DUAL REPRESENTATIONS OF QUASICONVEX COMPOSITIONS WITH APPLICATIONS TO SYSTEMIC RISK

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

INDUSTRIAL ENGINEERING

By

Mücahit Aygün

JULY 2021

DUAL REPRESENTATIONS OF QUASICONVEX COMPOSITIONS WITH APPLICATIONS TO SYSTEMIC RISK
By Mücahit Aygün
JULY 2021

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Çağın Ararat(Advisor)

Özlem Çavuş İyigün

Elisa Mastrogiacomo

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

DUAL REPRESENTATIONS OF QUASICONVEX COMPOSITIONS WITH APPLICATIONS TO SYSTEMIC RISK

Mücahit Aygün
M.S. in Industrial Engineering
Advisor: Çağın Ararat
JULY 2021

The importance of measuring risk in an interconnected financial system has been appreciated recently, due in part to the global financial crisis. In the literature, systemic risk measures are generally represented by the composition of a univariate risk measure and an aggregation function, a function that encodes the structure of the financial network. Having dual representations for systemic risk measures is helpful in providing economic interpretations and offering duality-based computational methods. For a univariate risk measure, a key assumption is that diversification should not increase risk. The mathematical translation of this assumption was considered as convexity earlier in the history of risk measures. Recently, quasiconvexity has been considered as a more accurate translation of diversification. For a single quasiconvex risk measure, dual representations are available in the literature based on the so-called penalty functions. The use of a quasiconvex risk measure in composition with a concave aggregation function results in a quasiconvex systemic risk measure, a multivariate functional on a space of random vectors.

Motivated by the problem of finding dual representations for quasiconvex systemic risk measures, we study quasiconvex compositions in an abstract infinite-dimensional setting. We calculate an explicit formula for the penalty function of the composition in terms of the penalty functions of the ingredient functions. The proof makes use of a nonstandard minimax inequality (rather than equality as in the standard case) that is available in the literature. In the last part of the thesis, we apply our results in concrete probabilistic settings for systemic risk measures, in particular, in the context of the Eisenberg-Noe clearing model. We also provide novel economic interpretations of the dual representations in these settings.

Keywords: quasiconvex function, composition of functions, minimax inequality, risk measure, systemic risk, dual representation, penalty function.

ÖZET

YARIDIŞBÜKEY BİLEŞKELERİN ÇİFTEŞ TEMSİLLERİ VE SİSTEMİK RİSK ÜZERİNE UYGULAMALARI

Mücahit Aygün
Endüstri Mühendisliği, Yüksek Lisans
Tez Danışmanı: Çağın Ararat
TEMMUZ 2021

Birbirine bağlı bileşenleri olan bir sistemin riskini ölçmenin önemi, özellikle 2008 Ekonomik Krizi'nden sonra daha iyi anlaşılmıştır. Bilimsel yazında sistemik risk ölçüleri, genellikle bir risk ölçüsü ile finansal ağın yapısını özetleyen bir yığışma fonksiyonunun bileşkesi olarak tanımlanmaktadır. Sistemik risk ölçüleri için çifteş temsillerin var olması, onları hesaplamada ve ekonomik anlamda yorumlamada kritik bir öneme sahiptir. Bir risk ölçüsü için temel varsayımlardan birisi, çeşitlendirmenin riski azaltacağıdır. Bu özelliğin matematiksel karşılığı olarak, risk ölçülerinin tartışıldığı ilk yıllarda dışbükeylik kullanılmaktaydı. Ama son zamanlarda yarıdışbükeyliğin, çeşitliliğin daha doğru bir matematiksel karşılığı olduğu düşünülmeye başlandı. Tek bir yarıdışbükey risk ölçüsü için bilimsel yazında çifteş temsiller bulunmaktadır, bu temsillerde kullanılan çifteş fonksiyonlara ceza fonksiyonu denir. Yarıdışbükey bir risk ölçüsüyle içbükey bir yığışım fonksiyonunun bileşkesi ile yarıdışbükey sistemik risk ölçüleri elde edilir; bunlar, uygun bir rassal vektörler uzayında tanımlı, yani çokdeğişkenli, fonksiyonellerdir.

Bu tezde, dışbükey sistemik risk ölçülerinden hareketle, dışbükey bileşke fonksiyonları soyut ve sonsuz boyutlu bir çerçevede çalışacağız. Bileşke fonksiyonun ceza fonksiyonunu bileşkeyi oluşturan fonksiyonların ceza fonksiyonları cinsinden hesaplayan bir formül kanıtlayacağız. Kanıtın temelinde, bilimsel yazında sıkça kullanılan enküçük-enbüyük eşitliğinin aksine, bilimsel yazında bulunan ama kullanımı standart olmayan bir enküçük-enbüyük eşitsizliği yer alacak. Tezin son kısmındaysa sonuçlarımızı sistemik risk ölçüleri için somut ve olasılıksal çerçevelerde uygulayarak çifteş temsiller elde edeceğiz. İnceleyeceğimiz sistemler arasında Eisenberg-Noe modeline göre çalışan takas sistemleri de olacak. Ayrıca çifteş temsiller için ekonomik anlamda yorumlarda bulunacağız.

Anahtar sözcükler: yarıdışbükey fonksiyon, bileşke fonksiyon, enküçük-enbüyük eşitsizliği, risk ölçüsü, çifteş temsil, ceza fonksiyonu.

Acknowledgement

Firstly and foremost I am extremely grateful to my advisor Çağın Ararat for his patience, guidance and expertise. I have learned how to do research, new approaches during this process. I felt very motivated and had new perspectives after every meeting with him. I admired his work ethic, teaching style and respectful connection with students since first day I took a course from him. He will always be a role model throughout my academic career.

I would like to express my gratitude to Elisa Mastrogiacomo and Özlem Çavuş İyigün for being in my committee and their valuable time to read and review this thesis.

I feel very fortunate to have my wife Zeynep in my life. The process for writing a master thesis might be very stressful, especially during the Covid-19 time for most of the people. However, I was very peaceful since the first day we get married thanks to her understanding, endless love and beautiful smile.

I would like to thank my parents Mürüvvet and Yaşar, and my brothers Ahmet and Esad for their love and support for me to do what I want throughout my life. They always provide facilities for me to be successful and happy.

Thanks for Metin, Haliloğlu, Elbi, Burak, Fatih and Süleyman for their valuable friendship. They have always been with me whenever I need them.

Special thanks for Onur, Yasin, Deniz, Emin, Uğur and Keleş for discussing my questions related to this thesis whenever I ask in addition to their valuable friendship.

I want to thank Selman, Yusuf, Alperen, Merdan, Ahi, Said, Tosun and TDP for socializing me during my master.

I acknowledge the financial support provided by TÜBİTAK, The Scientific and Technological Research Council of Turkey, under project number 117F438.

Finally, I am grateful to all of my professors in Bilkent University. They have always been nice, understanding and helpful.

Contents

Chapter 1

Introduction

This thesis is concerned with extended real-valued functions of the form $f \circ g$, where f and g are functions defined on some general preordered topological vector spaces. We look for minimal assumptions on f and g to ensure that their composition $f \circ g$ is a monotone, quasiconvex, and lower semicontinuous function. In our main results, we provide novel duality formulae in which the dual function for $f \circ g$ is calculated in terms of the same type of functions for f and g.

In the literature, the study of $f \circ g$ from a duality point of view is not new in the convex case. For a single function, we have Fenchel-Moreau theorem which provides a dual representation for a convex lower semicontinuous function in terms of its Legendre-Fenchel conjugate (Theorem 12.2 in [1]). Then, it is natural to ask how and when we can have a dual representation for the composition of convex functions. This question has been answered in the literature, e.g., by Theorem 2.8.10 in [2], Theorem 3 in [3]; see also the more recent work [4].

As a natural extension of the convex case, we look for dual representations of $f \circ g$ when it is guaranteed to be quasiconvex, which seems to be an open problem to the best of our knowledge. For a single function, the quasiconvex duality theory of [5] provides a suitable replacement of conjugate functions in convex duality. This is further explored in [6] within an abstract framework and also in [7, 8, 9] within the context of risk measures. In line with [8], the dual functions for quasiconvex duality will be referred to as *penalty functions* in this thesis.

Our motivation for studying quasiconvex compositions also comes from financial mathematics, specifically, from the theory of systemic risk measures as we describe briefly next.

Initiated by the seminal work [10], risk measures have been studied extensively in the financial mathematics and operations research literature. In the original framework of [10], *coherent risk measures* are defined as monotone, convex, translative and positively homogeneous functionals defined on a space of real-valued random variables. These random variables could be used to model the uncertain future worths of investments, and a risk measure assigns to each random variable its minimum deterministic capital requirement. Among the properties of coherent risk measures, *monotonicity* is a natural requirement which asserts that the risk of an investment with consistently higher future values should be lower. *Convexity* is related to diversification; under this property, the risk of a mixture (i.e., convex combination) of two portfolios is not higher than the same type of mixture of the individual risks. The positive homogeneity property is a scaling property that is relaxed for defining *convex risk measures* in [11]. The *translativity* property asserts that a deterministic increase in the value of a portfolio decreases its risk by the same amount. This is indeed the property that justifies the interpretation of the value of a risk measure as capital requirement.

One might question whether convexity provides the correct encoding of the impact of diversification on risk. A weaker alternative is *quasiconvexity*, which bounds the risk of a mixture only by the maximum of the individual risks, hence the statement "Diversification does not increase risk." is reflected properly. Under translativity, convexity is equivalent to quasiconvexity, which is a weaker condition that is also related to diversification. Hence, the switch from convexity to quasiconvexity implies working with non-translative functionals in general. Indeed, the work [8] proposes a minimalist framework for risk measures in which only monotonicity and quasiconvexity are taken for granted, such functionals are called *quasiconvex risk measures*; see also [9]. For the use of quasiconvex risk measures in the context of financial optimization problems, see [12, 13, 14].

The theory of risk measures outlined above is for univariate (i.e., real-valued) random variables. In more complex settings such as markets with transaction costs (e.g., [15, 16]) and financial networks with interdependencies (e.g., [17, 18, 19, 20]), it becomes necessary to evaluate the risks of random vectors. For this thesis, we are particularly interested in the later situation where the participating financial institutions are subject to correlated sources of risk, typically affecting the future values of their assets. Hence, the resulting future values are naturally modeled as correlated random vectors, explaining the multivariate nature of the problem. At the same time, the institutions form a network through mutual obligations and the aforementioned uncertainty affects the ability of the institutions to meet these obligations. Hence, the aim of a *systemic risk measure* is to quantify the overall risk associated to the financial network.

In the pioneering work [17], a systemic risk measure R is defined as the composition of a univariate risk measure ρ with a so-called *aggregation function* Λ: $R = \rho \circ \Lambda$. The role of the aggregation function is to summarize the impact of the random shock vector X, on the economy (or society) as a scalar random quantity $\Lambda(X)$. The definition of Λ is made precise by the structure of the network and the accompanying clearing mechanism. For instance, one can consider a clearing system in the framework of [21] and define the aggregation function as the total payment made to society as in [20], in which case it is an increasing concave function. The output of Λ is further given as input to a convex risk measure ρ to calculate the value of $R(X)$. The resulting systemic risk measure R is a monotone convex functional that is not translative in general. In [20], dual representations for *convex* systemic risk measures are studied in detail. The mathematical machinery used in that work is the conjugation formulae Theorem 2.8.10 in [2] and Theorem 3 in [3] for convex compositions.

When ρ is only assumed to be a quasiconvex risk measure, the resulting systemic risk measure R is also quasiconvex. Providing dual representations for this case is the starting point of this thesis. However, we will first study the problem in greater generality. As stated at the beginning, we will explore the dual representation of a quasiconvex composition $f \circ g$, where the ingredients f, g are defined on general preordered topological vector spaces. To the best of our knowledge, the quasiconvex analogues of the conjugation results in [2] and [3] are not known in the literature. We provide a solution to this problem by proving a formula for the penalty function of $f \circ g$, roughly speaking, in terms of the penalty functions of f and g. More precisely, apart from the more technical continuity conditions, we will assume that f is an extended real-valued monotone, quasiconvex function. Since g is a vector-valued function (in a possibly infinite-dimensional space), choosing the right notion of quasiconvexity requires extra care. To this end, we will use the notion of *natural quasiconvexity*, which is introduced for vector-valued functions in [22] and for set-valued functions in [23]. When g is a monotone, naturally quasiconcave function, the resulting composition $f \circ g$ is a monotone, quasiconvex function.

For the proof of our main duality theorem (Theorem 4.1.6), we need a nonstandard minimax result since the assumptions of the standard minimax theorem in [24]. We are able to overcome this issue by using the minimax inequality in [25] (see also [26, 27]), which works under weaker conditions. With additional arguments that use the properties of the involved functions, we are able to turn the minimax inequality into an equality. Hence, the proof of the main theorem makes a novel use of minimax theory.

As a special case of Theorem 4.1.6, we consider convex compositions and recover the conjugation formula in [2] and [3].

After building the general theory, we go back to our motivating problem on systemic risk measures. Using a quasiconvex univariate risk measure ρ and a concave aggregation function Λ, we are able to provide a dual representation for the systemic risk measure $R = \rho \circ \Lambda$ in a probabilistic framework. We also discuss the economic interpretations of the dual variables and penalty functions in terms of the underlying financial network.

The rest of this thesis is organized as follows. In Chapter 2, we review some basic notions and results about convex and quasiconvex functions. Chapter 3 is dedicated to some more technical notions for vector-valued functions: natural quasiconvexity, regular monotonicity and lower demicontinuity. The main part is Chapter 4, where we prove the main theorem on quasiconvex compositions together with some important special cases. In Chapter 5, we apply the theory to obtain dual representations for systemic risk measures. Among the various examples that we study, Eisenberg-Noe model is discussed separately as it has a more sophisticated aggregation function. Some proofs of the results in Chapter 4 and Chapter 5 are collected in Chapter 6, the appendix.

Chapter 2

Convex and quasiconvex functions

2.1 Preliminaries

Let us begin with some basic notations and definitions that we use throughout the paper. We denote by $\overline{\mathbb{R}} \coloneqq \mathbb{R} \cup \{+\infty, -\infty\}$ the extended real line. For each $n \in \mathbb{N} \coloneqq \{1, 2, \ldots\}$, we denote by $\mathbb{R}^n$ the n-dimensional Euclidean space, by $\mathbb{R}^n_+$ the set of all $z = (z_1, \ldots, z_n)^\mathsf{T} \in \mathbb{R}^n$ with $z_i \geq 0$ for each $i \in \{1, \ldots, n\}$, and by $\mathbb{R}^n_{++}$ the set of all $z \in \mathbb{R}^n$ with $z_i > 0$ for each $i \in \{1, \ldots, n\}$. When $n = 1$, we write $\mathbb{R}_+ = \mathbb{R}^1_+$ and $\mathbb{R}_{++} = \mathbb{R}^1_{++}$. Let $\mathcal{X}$ be a Hausdorff locally convex topological vector space. We denote by $\mathcal{X}^*$ its topological dual space endowed with the weak* topology $\sigma(\mathcal{X}^*, \mathcal{X})$. The bilinear duality mapping on $\mathcal{X}^* \times \mathcal{X}$ is denoted by $\langle \cdot, \cdot \rangle$. For nonempty sets $A, B \subseteq \mathcal{X}$ and $\lambda \in \mathbb{R}$, we define the sum $A + B \coloneqq \{x + y \mid x \in A, y \in B\}$ and the product $\lambda A \coloneqq \{\lambda x \mid x \in A\}$ in the Minkowski sense. When $A = \{x\}$ for some $x \in \mathcal{X}$, we write $x + B \coloneqq \{x\} + B$.

Throughout this chapter, let $f \colon \mathcal{X} \to \overline{\mathbb{R}}$ be a function. Given $m \in \mathbb{R}$, the *m-sublevel set* of f is defined as

$$S^f_m \coloneqq \{x \in \mathcal{X} \mid f(x) \leq m\}.$$

The next lemma provides a representation of f via its sublevel sets. It is a known result and we give its proof for the convenience of the reader. This representation will be useful when obtaining a dual representation for f.

Lemma 2.1.1. *Let $x \in \mathcal{X}$. Then, it holds*

$$f(x) = \inf\{m \in \mathbb{R} \mid x \in S_m^f\}. \tag{2.1.1}$$

Proof. First, suppose that $f(x) = +\infty$. Then, there is no $m \in \mathbb{R}$ such that $m \geq f(x)$. Hence, $\inf\{m \in \mathbb{R} \mid x \in S_m^f\} = \inf \emptyset = +\infty$. Next, suppose that $f(x) = -\infty$. Then, we have $-\infty < m$ for every $m \in \mathbb{R}$ so that $\inf\{m \in \mathbb{R} \mid x \in S_m^f\} = \inf \mathbb{R} = -\infty$. Finally, suppose that $f(x) \in \mathbb{R}$. For every $m \in \mathbb{R}$ such that $x \in S_m^f$, we have $f(x) \leq m$ by definition. Therefore, $f(x) \leq \inf\{m \in \mathbb{R} \mid x \in S_m^f\}$. Since $f(x) \in \mathbb{R}$, there exists $n \in \mathbb{R}$ such that $n < \inf\{m \in \mathbb{R} \mid x \in S_m^f\}$. Hence, $x \notin S_n^f$, which implies that $f(x) > n$. Taking supremum over all such n gives

$$f(x) \geq \sup\left\{n \in \mathbb{R} \mid n < \inf\{m \in \mathbb{R} \mid x \in S_m^f\}\right\} = \inf\{m \in \mathbb{R} \mid x \in S_m^f\},$$

which completes the proof. $\square$

The function f is called *positively homogeneous* if $f(\lambda x) = \lambda f(x)$ for every $\lambda > 0$ and $x \in \mathcal{X}$. It is called *proper* if $f(x) > -\infty$ for every $x \in \mathcal{X}$ and $f(x) < +\infty$ for at least one $x \in \mathcal{X}$. The *conjugate function* or the *Legendre-Fenchel transform* $f^* \colon \mathcal{X}^* \to \overline{\mathbb{R}}$ of f is defined by

$$f^*(x^*) \coloneqq \sup_{x \in \mathcal{X}} \left(\langle x^*, x\rangle - f(x)\right), \quad x^* \in \mathcal{X}^*.$$

As an important special case, we may take $f = I_A$ for some $A \subseteq \mathcal{X}$, where I_A is the (convex analytic) *indicator function* of A defined by

$$I_A(x) \coloneqq \begin{cases} 0 & \text{if } x \in A, \\ +\infty & \text{if } x \in \mathcal{X} \setminus A. \end{cases}$$

Then, the conjugate function of I_A is the *support function* of A given by

$$I_A^*(x^*) = \sup_{x \in A} \langle x^*, x\rangle, \quad x^* \in \mathcal{X}^*. \tag{2.1.2}$$

Definition 2.1.2. *The function f is called* quasiconvex *if*

$$f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\} \tag{2.1.3}$$

for every $x, y \in \mathcal{X}$ and $\lambda \in [0,1]$. It is called quasiconcave *if $(-f)$ is quasiconvex.*

Remark 2.1.3. It is well-known that f is a quasiconvex function if and only if its sublevel set S_m^f is convex for every $m \in \mathbb{R}$ (Sect. 2.1, p. 41 in [2]).

Definition 2.1.4. *Let $x \in \mathcal{X}$. The function f is called* lower semicontinuous at x *if $f(x) \leq \liminf_{i \in I} f(x_i)$ whenever $(x_i)_{i \in I}$ is a net in $\mathcal{X}$ that converges to x. It is called* lower semicontinuous *if it is lower semicontinuous at each $x \in \mathcal{X}$. It is called* upper semicontinuous (at x) *if $(-f)$ is lower semicontinuous (at x).*

Remark 2.1.5. It is well-known that f is lower semicontinuous if and only if its sublevel set S_m^f is closed for every $m \in \mathbb{R}$ (Lemma 2.39 in [28]).

Remark 2.1.6. It is well-known that every proper closed convex subset of a locally convex topological vector space equals the intersection of all closed half spaces that contain it (Corollary 5.83 in [28]). In view of Remark 2.1.3 and Remark 2.1.5, when f is proper, lower semicontinuous and quasiconvex, for each $m \in \mathbb{R}$, the set S_m^f can be written as an intersection of closed halfspaces.

2.2 The order structure

To be able to handle monotone functions (e.g., in the risk measure applications in Section 5.1 and Section 5.2), we introduce an order structure on $\mathcal{X}$. To that end, let $C \subseteq \mathcal{X}$ be a convex cone and define a relation $\leq_C$ on $\mathcal{X}$ by

$$x \leq_C y \quad \Leftrightarrow \quad y - x \in C \tag{2.2.1}$$

for each $x, y \in \mathcal{X}$. It follows that $\leq_C$ is a *vector preorder*, that is, $x \leq_C y$ implies $x+z \leq_C y+z$ and $\lambda x \leq_C \lambda y$ for every $x, y, z \in \mathcal{X}$ and $\lambda > 0$.

Remark 2.2.1. It can be checked that every vector preorder $\preccurlyeq$ on $\mathcal{X}$ can be written as $\preccurlyeq = \leq_C$, where $C := \{x \in \mathcal{X} \mid 0 \preccurlyeq x\}$ is a convex cone. Hence, the assumption that C is a convex cone is not a restriction on the vector preorder of interest.

The elements of C are called *positive elements* of $\mathcal{X}$. Let us define the *(positive) dual cone* of C by

$$C^+ := \{x^* \in \mathcal{X}^* \mid \langle x^*, x\rangle \geq 0 \text{ for all } x \in C\},$$

which is a closed convex cone in $\mathcal{X}^*$. Then, the cone of *strictly positive elements* of $\mathcal{X}$ is defined by

$$C^{\#} = \left\{x \in C \mid \langle x^*, x\rangle > 0 \text{ for all } x^* \in C^+ \setminus \{0\}\right\}. \tag{2.2.2}$$

Given $\pi \in C^{\#}$, we may scale the elements of C^+ and obtain the closed convex set

$$C^+_\pi := \{x^* \in C^+ \mid \langle x^*, \pi\rangle = 1\}.$$

Remark 2.2.2. When $\mathcal{X}$ is finite-dimensional, $C^{\#}$ coincides with the interior of C. However, in the infinite-dimensional setting, we prefer working with $C^{\#}$ since the interior of C can be empty for many important examples including Lebesgue spaces; see Example 2.12 in [29], for instance.

The next lemma shows that C^+ can be recovered from the (much) smaller set C^+_π whenever $\pi \in C^{\#}$.

Lemma 2.2.3. *Assume that $C^{\#} \neq \emptyset$ and let $\pi \in C^{\#}$. Then, we have $C^+ \setminus \{0\} = \mathbb{R}_{++}C^+_\pi$.*

Proof. Let $\lambda > 0$ and $x^* \in C^+_\pi$. By definition, $C^+_\pi \subseteq C^+ \setminus \{0\}$ and $C^+ \setminus \{0\}$ is a cone so that $\lambda x^* \in C^+ \setminus \{0\}$. Hence, $\mathbb{R}_{++}C^+_\pi \subseteq C^+ \setminus \{0\}$. Conversely, let $x^* \in C^+ \setminus \{0\}$. We have $\langle x^*, \pi\rangle > 0$. Taking $z^* := \frac{x^*}{\langle x^*, \pi\rangle}$, we have $\langle z^*, \pi\rangle = 1$, which implies that $z^* \in C^+_\pi$. Moreover, taking $\lambda = \frac{1}{\langle x^*, \pi\rangle} > 0$, we have $x^* = \lambda z^* \in \mathbb{R}_{++}C^+_\pi$. Hence, $C^+ \setminus \{0\} \subseteq \mathbb{R}_{++}C^+_\pi$. $\square$

Thanks to the order structure provided by $\leq_C$, we may define the monotonicity of sets and functions. We say that a set $A \subseteq \mathcal{X}$ is *monotone* if $x \leq_C y$ and $x \in A$ imply $y \in A$, for every $x, y \in \mathcal{X}$. Similarly, we say that f is a *decreasing* function (with respect to C) if $x \leq_C y$ implies $f(x) \geq f(y)$ for every $x, y \in \mathcal{X}$; we say that f is an *increasing* function (with respect to C) if it is decreasing with respect to $-C$.

Remark 2.2.4. It is easy to see that f is decreasing if and only if its sublevel sets are monotone. Indeed, suppose that f is decreasing, and let $m \in \mathbb{R}$, $x \in S^f_m$. If $x \leq_C y$ for some $y \in \mathcal{X}$, then we have $m \geq f(x) \geq f(y)$ so that $y \in S^f_m$. Hence, S^f_m is monotone. Conversely, suppose that S^f_m is monotone for every $m \in \mathbb{R}$, and take $x, y \in \mathcal{X}$ such that $x \leq_C y$. If $f(x) = +\infty$, then $f(x) \geq f(y)$ holds trivially. Assume that $f(x) < +\infty$. Clearly, $x \in S^f_m$ for every $m \in \mathbb{R}$ with $m \geq f(x)$. Since $x \leq_C y$, $y \in S^f_m$ and S^f_m is monotone, we have $y \in S^f_m$, that is, $m \geq f(y)$ for every such m. Letting $m \to f(x)$ gives $f(x) \geq f(y)$. Hence, f is decreasing.

2.3 Dual representations

In convex analysis, Fenchel-Moreau theorem provides a dual representation for a proper lower semicontinous convex function f in terms of its conjugate function f^*:

$$f(x) = \sup_{x^* \in \mathcal{X}^*} \left(\langle x^*, x \rangle - f^*(x) \right), \quad x \in \mathcal{X}.$$

One immediate consequence of this theorem is the following lemma; we give its proof for the convenience of the reader. We will use this lemma in the proof of Proposition 6.1.3, which is a significant tool for proving Theorem 4.1.6, the main theorem of the paper.

Lemma 2.3.1. *Let $A \subseteq \mathcal{X}$ be a set and B its closed convex hull. Then,*

$$I_A^*(x^*) = \sup_{x \in B} \langle x^*, x \rangle, \quad x^* \in \mathcal{X}^*.$$

Proof. By Fenchel-Moreau theorem, we have $I_A^{**} = I_B$ since B is the closed convex hull of A. Then, taking the conjugate functions of both sizes and applying Fenchel-Moreau theorem once more, we get $I_A^* = I_A^{***} = I_B^*$. On the other hand, we have $I_B^*(x^*) = \sup_{x \in B} \langle x^*, x \rangle$ for every $x^* \in \mathcal{X}^*$. Hence, the result follows. □

For monotone functions, the following refinement of Fenchel-Moreau theorem is possible.

Proposition 2.3.2. *Suppose that f is proper, decreasing, convex and lower semicontinuous. Then, we have*

$$f(x) = \sup_{x^* \in C^+} \left(\langle -x^*, x \rangle - f^*(-x^*) \right), \quad x \in \mathcal{X}. \tag{2.3.1}$$

Proof. We first prove that $f^*(x^*) = +\infty$ when $x^* \notin -C^+$. Note that, in this case, there exists $c \in C$ such that $\langle x^*, c \rangle > 0$. Let $x_0 \in \operatorname{dom} f$ and $\lambda > 0$. Since f is decreasing, we have $f(x_0 + \lambda c) \leq f(x_0)$ so that $\langle x^*, x_0 + \lambda c \rangle - f(x_0 + \lambda c) \geq \lambda \langle x^*, c \rangle + \langle x^*, x_0 \rangle - f(x_0)$. Since $\langle x^*, c \rangle > 0$, letting $\lambda \to \infty$ implies that

$$f^*(x^*) \geq \sup_{\lambda > 0} \left(\langle x^*, x + \lambda c \rangle - f(x + \lambda c) \right) \geq +\infty.$$

Hence, $f^*(x^*) = +\infty$. Combining this with Fenchel-Moreau theorem yields (2.3.1). □

For a quasiconvex function, a suitable generalization of conjugation is possible by the so-called minimal penalty function, which is defined in terms of the support function of the negative of sublevel sets. The precise definition is given next.

Definition 2.3.3. *The minimal penalty function* $\alpha_f \colon \mathcal{X}^* \times \mathbb{R} \to \overline{\mathbb{R}}$ *associated with* f *is defined by*

$$\alpha_f(x^*, m) := \sup_{x \in S^f_m} \langle x^*, -x \rangle, \quad x^* \in \mathcal{X}^*, m \in \mathbb{R}.$$

Remark 2.3.4. We can extend this definition for $m = +\infty$ and $m = -\infty$ and $x^* \neq 0$ by letting $\alpha_f(x^*, +\infty) := +\infty$ and $\alpha_f(x^*, -\infty) := -\infty$. These values are consistent with the original definition. When we look at the case $m = +\infty$, we are considering the whole space $\{x \in \mathcal{X} \mid f(x) \leq \infty\} = \mathcal{X}$ in the supremum, which gives that the supremum is $+\infty$. Similarly, for the case $m = -\infty$, we are considering the supremum over the empty set, which is $-\infty$.

The next two remarks state some elementary properties of the minimal penalty function α_f.

Remark 2.3.5. It is clear that the minimal penalty function α_f is positively homogeneous in the first argument, that is, $\alpha_f(\lambda x^*, m) = \lambda \alpha_f(x^*, m)$ for every $x^* \in \mathcal{X}^*$, $m \in \mathbb{R}$. Moreover, α_f is increasing in the second argument. Indeed, taking $m_1, m_2 \in \mathbb{R}$ with $m_1 \leq m_2$, we have $S^f_{m_1} \subseteq S^f_{m_2}$ so that $\alpha_f(x^*, m_1) \leq \alpha_f(x^*, m_2)$ for every $x^* \in \mathcal{X}^*$.

Remark 2.3.6. By (2.1.2) and Definition 2.3.3, we have

$$\alpha_f(x^*, m) = \sup_{x \in S^f_m} \langle -x^*, x \rangle = I^*_{S^f_m}(-x^*), \quad x^* \in \mathcal{X}^*, m \in \mathbb{R}.$$

We continue with a lemma which serves as a basis for dual representations since it characterizes a set in the primal space $\mathcal{X}$ in terms of the elements of the dual space $\mathcal{X}^*$.

Lemma 2.3.7. *Let* $A \subseteq \mathcal{X}$ *be a nonempty, closed, convex and monotone set. Then, for every* $x \in \mathcal{X}$*, we have*

$$x \in A \quad \Leftrightarrow \quad \forall x^* \in C^+ \setminus \{0\} \colon \langle x^*, -x \rangle \leq \sup_{y \in A} \langle x^*, -y \rangle .$$

Proof. For $A = \mathcal{X}$, the result is clear. Suppose that $A \neq \mathcal{X}$ and let $x \in A$. Clearly, $\langle x^*, -x \rangle \leq \sup_{y \in A} \langle x^*, -y \rangle$ for each $x^* \in C^+ \setminus \{0\}$. We prove the converse by contrapositive.

To that end, let $x \in \mathcal{X} \setminus A$. Since A is closed and convex, by Hahn-Banach separation theorem, there exists $x^* \in \mathcal{X}^* \setminus \{0\}$ such that

$$\langle x^*, -x \rangle > \sup_{y \in A} \langle x^*, -y \rangle .$$

We claim that $x^* \in C^+$. To get a contradiction, suppose that $x^* \notin C^+$ so that $\langle x^*, -c \rangle > 0$ for some $c \in C$. Since C is a cone, for every $\lambda > 0$, we have $\lambda c \in C$. Now take $y \in A$. Since A is monotone, we have $y + \lambda c \in A$ and hence,

$$\langle x^*, -x \rangle > \langle x^*, -(y + \lambda c) \rangle = \langle x^*, -y \rangle + \lambda \langle x^*, -c \rangle \tag{2.3.2}$$

for every $\lambda > 0$. Since $\langle x^*, -c \rangle > 0$, letting $\lambda \to \infty$, the expression on the right of (2.3.2) diverges to $+\infty$ but the expression on the left is constant, which yields a contradiction. Therefore, $x^* \in C^+ \setminus \{0\}$ and the proof is complete. □

Remark 2.3.8. If f is a decreasing, lower semicontinuous and quasiconvex function, then the sublevel sets $S_m^f, m \in \mathbb{R}$, satisfy the properties in Lemma 2.3.7 by Remarks 2.1.3, 2.1.5, 2.2.4. Hence, by Definition 2.3.3, we have

$$x \in S_m^f \quad \Leftrightarrow \quad \forall x^* \in C^+ \setminus \{0\} \colon \ \langle x^*, -x \rangle \leq \alpha_f (x^*, m) .$$

Similarly, if f is an increasing, lower semicontinuous and quasiconvex function, then f is decreasing with respect to $-C$ so that

$$x \in S_m^f \quad \Leftrightarrow \quad \forall x^* \in C^- \setminus \{0\} \colon \ \langle x^*, -x \rangle \leq \alpha_f (x^*, m) ,$$

where $C^- := -C^+ = \{x^* \in \mathcal{X}^* \mid \langle x^*, x \rangle \leq 0 \text{ for all } x \in C\}$.

When f is lower semicontinuous and quasiconvex, its dual representation will be stated in terms of a special pseudoinverse of α_f, which we recall in the next definition.

Definition 2.3.9. *Let $\alpha \colon C^+ \times \mathbb{R} \to \overline{\mathbb{R}}$ be a function. We define its left inverse $\alpha^{-l} \colon C^+ \times \mathbb{R} \to \overline{\mathbb{R}}$ with respect to the second argument by*

$$\alpha^{-l}(x^*, s) := \sup \{m \in \mathbb{R} \mid \alpha(x^*, m) < s\} = \inf \{m \in \mathbb{R} \mid \alpha(x^*, m) \geq s\} , \tag{2.3.3}$$

for each $x^* \in C^+$ and $s \in \mathbb{R}$.

The following lemma provides simple strong duality results that will be useful in later calculations.

Lemma 2.3.10. *Let $\alpha\colon C^+ \times \mathbb{R} \to \overline{\mathbb{R}}$ be a function which is increasing in its second argument.*

(i) Let $r\colon \mathcal{X}^* \to \overline{\mathbb{R}}$ be a function and $A \subseteq \mathcal{X}^*$ a nonempty set. Then, we have

$$\inf\left\{m \in \mathbb{R} \mid \forall x^* \in A\colon r(x^*) \leq \alpha(x^*, m)\right\} = \sup_{x^* \in A} \alpha^{-l}(x^*, r(x^*)).$$

(ii) Let S be a nonempty set and $r\colon \mathcal{X}^* \times S \to \overline{\mathbb{R}}$. Then, for every $x^* \in \mathcal{X}^*$, we have

$$\inf\left\{m \in \mathbb{R} \mid \forall s \in S\colon r(x^*, s) \leq \alpha(x^*, m)\right\} = \sup_{s \in S} \alpha^{-l}(x^*, r(x^*, s)).$$

Proof. Let us prove (i) first. By the definition of left inverse, the claimed equality is equivalent to

$$\inf\left\{m \in \mathbb{R} \mid \forall x^* \in A\colon r(x^*) \leq \alpha(x^*, m)\right\} = \sup_{x^* \in A} \inf\left\{m \in \mathbb{R} \mid r(x^*) \leq \alpha(x^*, m)\right\}. \quad (2.3.4)$$

The $\geq$ part is true by weak duality. For the other side, to get a contradiction, assume that there exists $\tilde{m} \in \mathbb{R}$ such that

$$\inf\left\{m \in \mathbb{R} \mid \forall x^* \in A\colon r(x^*) \leq \alpha_f(x^*, m)\right\} > \tilde{m} > \sup_{x^* \in A} \inf\left\{m \in \mathbb{R} \mid r(x^*) \leq \alpha_f(x^*, m)\right\}.$$

The first inequality implies that there exists $\tilde{x}^* \in A$ such that

$$r(\tilde{x}^*) > \alpha_f(\tilde{x}^*, \tilde{m}). \quad (2.3.5)$$

The second inequality implies that $\tilde{m} > \inf\{m \in \mathbb{R} \mid r(\tilde{x}^*) \leq \alpha_f(\tilde{x}^*, m)\}$. Hence, by the monotonicity of α_f, we must have $r(\tilde{x}^*) \leq \alpha_f(\tilde{x}^*, \tilde{m})$, which is a contradiction to (2.3.5). Hence, (2.3.4) follows.

We have a similar proof for (ii). Note that the desired equality is equivalent to

$$\inf\left\{m \in \mathbb{R} \mid \forall s \in S\colon r(x^*, s) \leq \alpha(x^*, m)\right\} = \sup_{s \in S} \inf\left\{m \in \mathbb{R} \mid r(x^*, s) \leq \alpha(x^*, m)\right\}. \quad (2.3.6)$$

The $\geq$ part is true by weak duality. For the other side, to get a contradiction, assume that there exists $\tilde{m} \in \mathbb{R}$ such that

$$\inf \{m \in \mathbb{R} \mid \forall s \in S\colon r(x^*, s) \leq \alpha_f(x^*, m)\} > \tilde{m} > \sup_{s \in S} \inf \{m \in \mathbb{R} \mid r(x^*, s) \leq \alpha_f(x^*, m)\} .$$

The first inequality implies that there exists a $\tilde{s} \in S$ such that

$$r(x^*, \tilde{s}) > \alpha_f(x^*, \tilde{m}). \tag{2.3.7}$$

The second inequality implies that $\tilde{m} > \inf \{m \in \mathbb{R} \mid r(x^*, \tilde{s}) \leq \alpha_f(x^*, m)\}$. Hence, by the monotonicity of α_f, we have $r(x^*, \tilde{s}) \leq \alpha_f(x^*, \tilde{m})$, a contradiction to (2.3.7). Hence, (2.3.6) follows. $\square$

We state the dual representation theorem for lower semicontinuous quasiconvex functions, which is a part of Theorem 3 in [8]. It is formulated in terms of the left inverse of the minimal penalty function. We provide the proof for completeness.

Theorem 2.3.11. *Suppose that $f\colon \mathcal{X} \to \overline{\mathbb{R}}$ is a decreasing, lower semicontinuous and quasiconvex function. Then, f has the dual representation*

$$f(x) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha_f^{-l} \left(x^*, \langle x^*, -x\rangle\right), \quad x \in \mathcal{X}, \tag{2.3.8}$$

where α_f^{-l} is the left inverse of α_f.

Proof. Let $x \in \mathbb{R}$. By Lemma 2.1.1 and Remark 2.3.8 we have

$$\begin{aligned} f(x) &= \inf\{m \in \mathbb{R} \mid x \in S_m^f\} = \inf \left\{m \in \mathbb{R} \mid \forall x^* \in C^+ \setminus \{0\}\colon \langle x^*, -x\rangle \leq \alpha_f(x^*, m)\right\} \\ &= \sup_{x^* \in C^+ \setminus \{0\}} \alpha_f^{-l} \left(x^*, \langle x^*, -x\rangle\right), \end{aligned}$$

where the last equality is a direct result of Lemma 2.3.10(i) since α_f is increasing by Remark 2.3.5. $\square$

We will generalize Theorem 2.3.11 for the composition of quasiconvex functions in Chapter 4. In [8], a decreasing quasiconvex function on $\mathcal{X}$ is called a *risk measure* as a generalization of convex and coherent risk measures studied in the financial mathematics literature; see Chapter 4 of [11], for instance. Hence, Theorem 2.3.11 provides a dual representation for a lower semicontinuous (quasiconvex) risk measure. For the current discussion, we keep using the general terminology of convex analysis and do not use the term *risk measure*. In Sections 5.1 and 5.2, we will focus on applications in systemic risk measures, where we also introduce the financial background as necessary.

In applications, it might be necessary to consider a function that is defined on some subset of the vector space $\mathcal{X}$. The next corollary is for this purpose. To that end, let $\mathcal{K} \subseteq \mathcal{X}$ be a monotone convex set such that $C \subseteq \mathcal{K}$. In particular, we have $\mathcal{K} + C \subseteq \mathcal{K}$. Given a function $g\colon \mathcal{K} \to \overline{\mathbb{R}}$, we may extend g to $\mathcal{X}$ as a function $\bar{g}$ defined by $\bar{g}(x) = g(x)$ for $x \in \mathcal{K}$, and by $\bar{g}(x) = +\infty$ for $x \in \mathcal{X} \setminus \mathcal{K}$. Hence, the sublevel sets, minimal penalty function, and algebraic properties (quasiconvexity, monotonicity, etc.) of g are defined as those of $\bar{g}$.

Corollary 2.3.12. *Let $g\colon \mathcal{K} \to \overline{\mathbb{R}}$ be a quasiconvex, decreasing and lower semicontinuous (with respect to the relative topology) function. Then, we have*

$$g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha_g^{-l}\left(x^*, \langle x^*, -x \rangle\right), \quad x \in \mathcal{K}. \tag{2.3.9}$$

Proof. Let us define a function $\tilde{g}\colon \mathcal{X} \to \mathbb{R}$ by

$$\tilde{g}(x) := \inf\{m \in \mathbb{R} \mid x \in \operatorname{cl}(S_m^g)\}, \quad x \in \mathcal{X}.$$

Note that $S_m^{\tilde{g}} = \operatorname{cl}(S_m^g)$ for each $m \in \mathbb{R}$. Let $m \in \mathbb{R}$. Since g is quasiconvex, it follows that $S_m^{\tilde{g}}$ is closed and convex. To show that it is also monotone, let $x \in S_m^{\tilde{g}} = \operatorname{cl}(S_m^g)$, $c \in C$. Let $U \subseteq \mathcal{X}$ be a neighborhood of $x + c$. Since $\mathcal{X}$ is a topological vector space, $(U - c)$ is an open set; hence, it is a neighborhood of x. Therefore, $(U - c) \cap S_m^g \neq \emptyset$. Let $z \in (U - c) \cap S_m^g$ so that $z + c \in U$. On the other hand, since g is decreasing, S_m^g is monotone, which yields that $z + c \in S_m^g$. It follows that $U \cap S_m^g \neq \emptyset$. Since U is an arbitrary neighborhood of $x + c$, we conclude that $x + c \in \operatorname{cl}(S_m^g) = S_m^{\tilde{g}}$. Hence, $S_m^{\tilde{g}}$ is monotone. By Remarks 2.1.3, 2.1.5, 2.2.4, it follows that $\tilde{g}$ is decreasing, lower semicontinuous, and quasiconvex. Using Theorem 2.3.11, we get

$$\tilde{g}(x) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha_{\tilde{g}}^{-l}\left(x^*, \langle x^*, -x \rangle\right), \quad x \in \mathcal{X}. \tag{2.3.10}$$

By definition, $S_m^{\tilde{g}}$ is the closed convex hull of S_m^g for each $m \in \mathbb{R}$. Hence, Lemma 2.3.1 yields

$$\alpha_{\tilde{g}}(x^*, m) = \sup_{y \in S_m^{\tilde{g}}} \langle x^*, -y \rangle = \sup_{y \in S_m^g} \langle x^*, -y \rangle = \alpha_g(x^*, m), \quad x^* \in \mathcal{X}^*, m \in \mathbb{R}. \tag{2.3.11}$$

For $x \in \mathcal{K}$, by Lemma 2.1.1, we have

$$\tilde{g}(x) = \inf \left\{ m \in \mathbb{R} \mid x \in S_m^{\tilde{g}} \right\} = \inf \left\{ m \in \mathbb{R} \mid x \in S_m^{\tilde{g}} \cap \mathcal{K} \right\}. \tag{2.3.12}$$

We claim that $S_m^{\tilde{g}} \cap \mathcal{K} = S_m^g$. Indeed, it is clear that $S_m^{\tilde{g}} \cap \mathcal{K} = \mathrm{cl}(S_m^g) \cap \mathcal{K} \supseteq S_m^g$. On the other hand, since g is lower semicontinuous with respect to the relative topology, we have $S_m^g = A \cap \mathcal{K}$ for some closed set $A \subseteq \mathcal{X}$. Since $S_m^g \subseteq A$, we have $\mathrm{cl}(S_m^g) \subseteq A$. It follows that $\mathrm{cl}(S_m^g) \cap \mathcal{K} \subseteq A \cap \mathcal{K} = S_m^g$. Hence, the claim follows. Then, (2.3.12) yields

$$\tilde{g}(x) = \inf \{ m \in \mathbb{R} \mid x \in S_m^g \} = g(x).$$

After combining this result with (2.3.10) and (2.3.11), we obtain (2.3.9). □

When f is a proper lower semicontinuous convex function, two dual representations are possible: the one provided by Fenchel-Moreau theorem, and the one provided by Theorem 2.3.11 since f is also quasiconvex. To establish the link between the two representations, we calculate the left inverse of the minimal penalty function in terms of the conjugate function.

Proposition 2.3.13. *Suppose that $f \colon \mathcal{X} \to \mathbb{R}$ is convex and lower semicontinuous. If $m \in \mathbb{R}$ is such that the strict sublevel set $\{x \in \mathcal{X} \mid f(x) < m\}$ is nonempty, then*

$$\alpha_f(x^*, m) = \inf_{\lambda > 0} \left(\lambda m + \lambda f^* \left(-\frac{x^*}{\lambda} \right) \right), \quad x^* \in \mathcal{X}^*. \tag{2.3.13}$$

Moreover, for the left inverse, we have

$$\alpha_f^{-l}(x^*, s) = \sup_{\gamma \geq 0} \left(\gamma s - f^*(-\gamma x^*) \right), \quad x^* \in \mathcal{X}^*, s \in \mathbb{R}. \tag{2.3.14}$$

Proof. Let $x^* \in \mathcal{X}^*$ and $m \in \mathbb{R}$ such that $\{x \in \mathcal{X} \mid f(x) < m\} \neq \emptyset$. Note that $\alpha_f(x^*, m) = \sup_{x \in S_m^f} \langle x^*, -x \rangle$ can be seen as the optimal value of the following convex optimization problem:

$$\text{maximize } \langle x^*, -x \rangle \text{ subject to } f(x) \leq m, \; x \in \mathcal{X}.$$

By supposition, Slater's condition holds, that is, there exists $x_0 \in \mathcal{X}$ such that $f(x_0) < m$. Hence, we have strong duality for this problem, that is,

$$\alpha_f(x^*, m) = \inf_{\lambda \geq 0} \sup_{x \in \mathcal{X}} \left(\langle x^*, -x\rangle - \lambda(f(x) - m)\right).$$

When $\lambda = 0$, $\sup_{x\in\mathcal{X}} \left(\langle x^*, -x\rangle - \lambda(f(x) - m)\right) = \sup_{x\in\mathcal{X}} \langle x^*, -x\rangle = +\infty$. Hence, we can evaluate the infimum over $\lambda > 0$. Then,

$$\begin{aligned}\alpha_f(x^*, m) &= \inf_{\lambda > 0} \sup_{x \in \mathcal{X}} \left(\langle x^*, -x\rangle - \lambda(f(x) - m)\right) \\ &= \inf_{\lambda>0} \left(\lambda m + \sup_{x\in\mathcal{X}} \left(\langle x^*, -x\rangle - \lambda f(x)\right)\right) = \inf_{\lambda>0} \left(\lambda m + \lambda f^*\left(-\frac{x^*}{\lambda}\right)\right).\end{aligned}$$

We have proved (2.3.13). For $m > \inf_{x\in\mathcal{X}} f(x)$, the strict sublevel set $\{x \in \mathcal{X} \mid f(x) < m\}$ is nonempty. Let us define the set $F := (\inf_{x\in\mathcal{X}} f(x), +\infty)$. For $m < \inf_{x\in\mathcal{X}} f(x)$, we have $\alpha_f(x^*, m) = -\infty$. Also, note that $\inf_{x\in\mathcal{X}} f(x) = -\sup_{x\in\mathcal{X}}(0 - f(x)) = -f^*(0)$. Then, to prove (2.3.14), for each $s \in \mathbb{R}$, we have

$$\begin{aligned}\alpha_f^{-l}(x^*, s) &= \inf\left\{m \in \mathbb{R} \mid \alpha_f(x^*, m) \geq s\right\} \\ &= \inf_{m\in F} m \vee \inf\left\{m \in \mathbb{R} \mid \forall \lambda > 0\colon \lambda m + \lambda f^*\left(-\frac{x^*}{\lambda}\right) \geq s\right\} \\ &= \inf_{x\in\mathcal{X}} f(x) \vee \inf\left\{m \in \mathbb{R} \mid \forall\lambda > 0\colon m \geq \frac{s}{\lambda} - f^*\left(-\frac{x^*}{\lambda}\right)\right\} \\ &= -f^*(0) \vee \sup_{\lambda>0}\left(\frac{s}{\lambda} - f^*\left(-\frac{x^*}{\lambda}\right)\right) = -f^*(0) \vee \sup_{\gamma>0}\left(\gamma s - f^*\left(-\gamma x^*\right)\right) \\ &= \sup_{\gamma\geq 0}\left(\gamma s - f^*\left(-\gamma x^*\right)\right),\end{aligned}$$

which completes the proof. □

Remark 2.3.14. Under the assumptions of Proposition 2.3.13, rewriting the dual representation in Theorem 2.3.11 using Proposition 2.3.13 and the fact that C^+ is a cone gives

$$\begin{aligned}f(x) &= \sup_{x^*\in C^+\setminus\{0\}} \alpha_f^{-l}(x^*, \langle x^*, -x\rangle) = \sup_{x^*\in C^+\setminus\{0\}} \sup_{\gamma\geq 0} \left(\langle \gamma x^*, -x\rangle - f^*(-\gamma x^*)\right) \\ &= \sup_{x^*\in C^+} \left(\langle x^*, -x\rangle - f^*(-x^*)\right), \quad x \in \mathcal{X}.\end{aligned}$$

Hence, in the convex case, the representation in Theorem 2.3.11 reproduces the standard Fenchel-Moreau-type representation in Proposition 2.3.2.

Chapter 3

Naturally quasiconvex vector-valued functions

Throughout this chapter, let $\mathcal{X}, \mathcal{Y}$ be Hausdorff locally convex topological vector spaces with vector preorders $\leq_C, \leq_D$, where $C \subseteq \mathcal{X}$ and $D \subseteq \mathcal{Y}$ are closed convex cones. We denote by $2^{\mathcal{Y}}$ the power set of $\mathcal{Y}$. Let $f\colon \mathcal{Y} \to \overline{\mathbb{R}}$ and $g\colon \mathcal{X} \to \mathcal{Y}$ be functions. The main focus of this paper is to provide dual representations for a quasiconvex composite function of the form $f \circ g$. While Chapter 2 provides the background for the study of extended real-valued function, we dedicate this chapter to the study of vector-valued functions.

We start by recalling some generalized notions of convexity for vector-valued functions.

Definition 3.0.1. *Consider the following notions for $g\colon \mathcal{X} \to \mathcal{Y}$.*

1. *g is called D-*convex *if*

$$g(\lambda x_1 + (1-\lambda)x_2) \leq_D \lambda g(x_1) + (1-\lambda) g(x_2)$$

for every $x_1, x_2 \in \mathcal{X}$ and $\lambda \in (0,1)$.

2. *g is called D-*concave *if $-g$ is D-convex.*

3. *g is called D-*naturally quasiconvex *if, for every $x_1, x_2 \in \mathcal{X}$ and $\lambda \in [0,1]$, there exists*

$\mu \in [0,1]$ *such that*

$$g(\lambda x_1 + (1-\lambda)x_2) \leq_D \mu g(x_1) + (1-\mu)g(x_2).$$

*4. g is called D-*naturally quasiconcave *if $-g$ is naturally D-quasiconvex.*

From Definition 3.0.1, it is clear that D-convexity implies D-natural quasiconvexity. For real-valued functions with $D = \mathbb{R}_+$, D-natural quasiconvexity coincides with quasiconvexity; see the notes after Definition 2.1 in [23].

For the function $g\colon \mathcal{X} \to \mathcal{Y}$, let us consider the *scalarization* $h^g_{y^*}(x)\colon \mathcal{X} \to \mathbb{R}$ defined by

$$h^g_{y^*}(x) := \langle y^*, g(x) \rangle\,, \quad x \in \mathcal{X}, \tag{3.0.1}$$

for each $y^* \in D^+ \setminus \{0\}$. The next proposition provides useful characterizations of the convexity and D-natural quasiconvexity of g in terms of the analogous properties of the family of scalarizations.

Proposition 3.0.2. *We have the following equivalences for g and its scalarizations.*

(i) g is D-convex if and only if $h^g_{y^*}$ is quasiconvex for each $y^* \in D^+ \setminus \{0\}$.

(ii) The function g is D-naturally quasiconvex if and only if $h^g_{y^*}$ is quasiconvex for every $y^* \in D^+ \setminus \{0\}$.

Proof. We prove (i) first. Assume that g is D-convex and take $y^* \in D^+ \setminus \{0\}$. Let $x_1, x_2 \in \mathcal{X}$ and $\lambda \in (0,1)$. Then, the D-convexity of g implies that $g(\lambda x_1 + (1-\lambda)x_2) \leq_D \lambda g(x_1) + (1-\lambda)g(x_2)$, that is, $\lambda g(x_1) + (1-\lambda)g(x_2) - g(\lambda x_1 + (1-\lambda)x_2) \in D$. Hence,

$$\langle y^*, \lambda g(x_1) + (1-\lambda)g(x_2) - g(\lambda x_1 + (1-\lambda)x_2) \rangle \geq 0.$$

so that

$$\begin{aligned} h^g_{y^*}(\lambda x_1 + (1-\lambda)x_2) &= \langle y^*, g(\lambda x_1 + (1-\lambda)x_2) \rangle \\ &\leq \langle y^*, \lambda g(x_1) + (1-\lambda)g(x_2) \rangle = \lambda h^g_{y^*}(x_1) + (1-\lambda)h^g_{y^*}(x_2). \end{aligned}$$

Therefore, $h^g_{y^*}$ is convex.

Conversely, assume that $h^g_{y^*}$ is convex for each $y^* \in D^+ \setminus \{0\}$. Let $x_1, x_2 \in \mathcal{X}$ and $\lambda \in (0,1)$. For each $y^* \in D^+ \setminus \{0\}$, since $h^g_{y^*}$ is convex, we have

$$\begin{aligned}\langle y^*, g(\lambda x_1 + (1-\lambda)x_2)\rangle &= h^g_{y^*}(\lambda x_1 + (1-\lambda)x_2) \\ &\leq \lambda h^g_{y^*}(x_1) + (1-\lambda)h^g_{y^*}(x_2) = \langle y^*, \lambda g(x_1) + (1-\lambda)g(x_2)\rangle .\end{aligned}$$

Hence, $\langle y^*, \lambda g(x_1) + (1-\lambda)g(x_2) - g(\lambda x_1 + (1-\lambda)x_2)\rangle \geq 0$ for every $y^* \in D^+ \setminus \{0\}$, that is, $\lambda g(x_1) + (1-\lambda)g(x_2) - g(\lambda x_1 + (1-\lambda)x_2) \in D$, that is,

$$g(\lambda x_1 + (1-\lambda)x_2) \leq_D \lambda g(x_1) + (1-\lambda)g(x_2).$$

Therefore, g is D-convex, which completes the proof of (i).

Next, we prove (ii). Assume that g is D-naturally quasiconvex. Let $y^* \in D^+ \setminus \{0\}$ and consider $h^g_{y^*}$. Let $x_1, x_2 \in \mathcal{X}$ and $\lambda \in [0,1]$. Since g is D-naturally quasiconvex, there exists $\mu \in [0,1]$ such that $g(\lambda x_1 + (1-\lambda)x_2) \leq_D \mu g(x_1) + (1-\mu)g(x_2)$. Hence,

$$\begin{aligned}h^g_{y^*}(\lambda x_1 + (1-\lambda)x_2) &= \langle y^*, g(\lambda x_1 + (1-\lambda)x_2)\rangle \\ &\leq \langle y^*, \mu g(x_1) + (1-\mu)g(x_2)\rangle \\ &\leq \max\{\langle y^*, g(x_1)\rangle, \langle y^*, g(x_2)\rangle\} = \max\{h^g_{y^*}(x_1), h^g_{y^*}(x_2)\}.\end{aligned}$$

Therefore, $h^g_{y^*}$ is quasiconvex.

Conversely, assume that $h^g_{y^*}$ is quasiconvex for each $y^* \in D^+ \setminus \{0\}$. To get a contradiction, suppose that g is not D-naturally quasiconvex. Hence, there exist $x_1, x_2 \in \mathcal{X}$ and $\lambda \in [0,1]$ such that

$$(\mathrm{co}\,(\{g(x_1), g(x_2)\}) - g(\lambda x_1 + (1-\lambda)x_2)) \cap D = \emptyset.$$

Since D is closed and convex, and the (shifted) line segment $\mathrm{co}(\{g(x_1), g(x_2)\}) - g(\lambda x_1 + (1-\lambda)x_2)$ is compact and convex, by Hahn-Banach strong separation theorem, there exists $y^*_0 \in \mathcal{Y}^* \setminus \{0\}$ such that

$$\inf_{d \in D} \langle y^*_0, d\rangle > \sup_{y \in (\mathrm{co}(\{g(x_1), g(x_2)\}) - g(\lambda x_1 + (1-\lambda)x_2))} \langle y^*_0, y\rangle \tag{3.0.2}$$

Since D is a cone, $\inf_{d \in D} \langle y^*, d \rangle$ is either 0 or $-\infty$. However, the term on the right of (3.0.2) is finite. Hence, we must have $\inf_{d \in D} \langle y_0^*, d \rangle = 0$ so that $y_0^* \in D^+$ as well. Then, using this information in (3.0.2) implies

$$\langle y_0^*, \mu g(x_1) + (1-\mu) g(x_2) \rangle < \langle y_0^*, g(\lambda x_1 + (1-\lambda) x_2) \rangle$$

for every $\mu \in [0,1]$. It follows that

$$\max\{\langle y_0^*, g(x_1) \rangle, \langle y_0^*, g(x_2) \rangle\} < \langle y_0^*, g(\lambda x_1 + (1-\lambda) x_2) \rangle,$$

which contradicts the quasiconvexity of $h^g_{y_0^*}$. Hence, g is D-naturally quasiconvex, which completes the proof of (ii). $\square$

Remark 3.0.3. The equivalent condition in Proposition 3.0.2(ii) is sometimes called $*$-*quasiconvexity*; see, for instance, Definition 2.1 in [23].

If $f \colon \mathcal{Y} \to \overline{\mathbb{R}}$ is a decreasing convex function, and $g \colon \mathcal{X} \to \mathcal{Y}$ is a D-concave function, then it is easy to check that the composition $f \circ g \colon \mathcal{X} \to \overline{\mathbb{R}}$ is a convex function. The following proposition provides an analogue of this observation when the resulting composition is quasiconvex.

Proposition 3.0.4. *Suppose that f is quasiconvex and decreasing, and g is D-naturally quasiconcave. Then, $f \circ g \colon \mathcal{X} \to \overline{\mathbb{R}}$ is quasiconvex.*

Proof. Let $x_1, x_2 \in \mathcal{X}$ and $\lambda \in [0,1]$. Since g is naturally D-quasiconcave, there exists $\mu \in [0,1]$ such that $\mu g(x_1) + (1-\mu) g(x_2) \leq_D g(\lambda x_1 + (1-\lambda) x_2)$. Using the monotonicity and quasiconvexity of f, we obtain

$$f(g(\lambda x_1 + (1-\lambda) x_2)) \leq f(\mu g(x_1) + (1-\mu) g(x_2)) \leq \max\{f(g(x_1)), f(g(x_2))\}.$$

Hence $f \circ g$ is quasiconvex. $\square$

When $f \colon \mathcal{Y} \to \overline{\mathbb{R}}$ is quasiconvex and decreasing, note that Proposition 3.0.4 gives a sufficient condition on $g \colon \mathcal{X} \to \mathcal{Y}$ that is weaker than D-convexity so that the composition $f \circ g$ is quasiconvex. In the rest of the chapter, we investigate further properties of g that will help us in obtaining a dual representation for $f \circ g$ in Chapter 4. To that end, we start by studying monotonicity for the vector-valued case.

Definition 3.0.5. *The function* $g\colon \mathcal{X} \to \mathcal{Y}$ *is called* decreasing *if* $x_1 \leq_C x_2$ *implies* $g(x_2) \leq_D g(x_1)$ *for every* $x_1, x_2 \in \mathcal{X}$*; it is called* increasing *if* $x_1 \leq_C x_2$ *implies* $g(x_1) \leq_D g(x_2)$ *for every* $x_1, x_2 \in \mathcal{X}$*.*

The preservation of monotonicity under compositions is formulated next.

Proposition 3.0.6. *Suppose that* f *is decreasing and* g *is increasing. Then,* $f \circ g$ *is decreasing.*

Proof. Let $x_1, x_2 \in \mathcal{X}$ such that $x_1 \leq_C x_2$. Since g is increasing, we have $g(x_1) \leq_D g(x_2)$. Since f is decreasing, we obtain $f(g(x_1)) \geq f(g(x_2))$. Hence, $f \circ g$ is decreasing. □

The next proposition connects the monotonicity of g and that of its scalarizations.

Proposition 3.0.7. *Suppose that* g *is decreasing. Then,* $h^g_{y^*}$ *is decreasing for every* $y^* \in D^+ \setminus \{0\}$*.*

Proof. Let $y^* \in D^+ \setminus \{0\}$. Let $x_1, x_2 \in \mathcal{X}$ such that $x_1 \leq_C x_2$. Since g is decreasing, we have $g(x_2) \leq_D g(x_1)$, that is, $g(x_1) - g(x_2) \in D$. Hence, $\langle y^*, g(x_1) - g(x_2)\rangle \geq 0$, that is, $h^g_{y^*}(x_1) \geq h^g_{y^*}(x_2)$. Therefore, $h^g_{y^*}$ is decreasing. □

In Chapter 4, we will also need a notion of strict monotonicity for a vector-valued function. The next definition gives one that suits our purposes. Recall that $C^\#$ and $D^\#$ are the (convex) cones of strictly positive elements in $\mathcal{X}$ and $\mathcal{Y}$, respectively; see (2.2.2). Although these cones are not closed in general, we define their induced preorders $\leq_{C^\#}$ and $\leq_{D^\#}$ as in (2.2.1).

Definition 3.0.8. *The function* g *is called* regularly increasing *if it is increasing and* $x_1 \leq_{C^\#} x_2$ *implies* $g(x_1) \leq_{D^\#} g(x_2)$ *for every* $x_1, x_2 \in \mathcal{X}$*; it is called* regularly decreasing *if it is decreasing and* $x_1 \leq_{C^\#} x_2$ *implies* $g(x_2) \leq_{D^\#} g(x_1)$ *for every* $x_1, x_2 \in \mathcal{X}$*.*

To be able to employ Definition 3.0.8, we need to work under the following assumption.

Assumption 3.0.9. *The cones* $C^\#$ *and* $D^\#$ *are nonempty.*

We proceed with a continuity concept for g, which is defined through its set-valued extension $G\colon \mathcal{X} \to 2^{\mathcal{Y}}$ given by

$$G(x) \coloneqq g(x) + D, \quad x \in \mathcal{X}.$$

Given $M \subseteq \mathcal{Y}$, the sets

$$G^L(M) \coloneqq \{x \in \mathcal{X} \mid G(x) \cap M \neq \emptyset\}, \quad G^U(M) \coloneqq \{x \in \mathcal{X} \mid G(x) \subseteq M\}$$

are called the *lower inverse image* and *upper inverse image* of M under G, respectively. It is easy to check that $(G^U(M))^c = G^L(M^c)$ and $(G^L(M))^c = G^U(M^c)$.

Definition 3.0.10. *The function g is called D-lower demicontinuous if the lower inverse image $G^L(M)$ is open for every open halfspace $M \subseteq \mathcal{Y}$.*

When $\mathcal{Y} = \mathbb{R}$ and $D = \mathbb{R}_+$, note that Definition 3.0.10 coincides with the usual notion of lower semicontinuity, see Remark 2.1.5.

Remark 3.0.11. Note that g is D-lower demicontinuous if and only if the upper inverse image $G^U(M)$ is closed for every closed halfspace $M \subseteq \mathcal{Y}$. This follows from the observations that M is a closed halfspace if and only if M^c is an open halfspace, and that $G^U(M) = (G^L(M^c))^c$.

We conclude this chapter by relating the D-lower demicontinuity of g with the upper semicontinuity of its scalarizations.

Proposition 3.0.12. *The function g is D-lower demicontinuous if and only if $h^g_{y^*}$ is upper semicontinuous for every $y^* \in D^+ \setminus \{0\}$.*

Proof. Let $m \in \mathbb{R}$ and $y^* \in D^+ \setminus \{0\}$. Let us define the sets

$$A_{m,y^*} \coloneqq \{x \in \mathcal{X} \mid h^g_{y^*}(x) \geq m\}, \quad B_{m,y^*} \coloneqq G^U(M_{m,y^*}) = \{x \in \mathcal{X} \mid g(x) + D \subseteq M_{m,y^*}\},$$

where

$$M_{m,y^*} \coloneqq \{y \in \mathcal{Y} \mid \langle y^*, y\rangle \geq m\}.$$

We claim that $A_{m,y^*} = B_{m,y^*}$. First, let $x \in A_{m,y^*}$ and take $d \in D$. Hence, $\langle y^*, g(x)\rangle \geq m$ and $\langle y^*, d\rangle \geq 0$. Combining these two inequalities, we get $\langle y^*, g(x) + d\rangle \geq m$, that is,

$g(x) + d \in M_{m,y^*}$. Since $d \in D$ is arbitrary, we have $g(x) + D \subseteq M_{m,y^*}$. Hence, $x \in B_{m,y^*}$. Conversely, let $x \in B_{m,y^*}$. In particular, $g(x) \in M_{m,y^*}$, that is, $h^g_{y^*}(x) = \langle y^*, g(x)\rangle \geq m$. Hence, $x \in A_{m,y^*}$, which completes the proof of the claim. By this claim and Remark 3.0.11, the statement of the proposition follows immediately. $\square$

Let $y^* \in D^+ \setminus \{0\}$. In view of Proposition 3.0.2 and Proposition 3.0.7, when g is D-naturally quasiconcave increasing and D-lower demicontinuous, the function $-h^g_{y^*}$ is quasiconvex, decreasing and lower semicontinuous. In this case, we may apply Theorem 2.3.11 for $-h^g_{y^*}$ to get

$$-h^g_{y^*}(x) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha^{-l}_{-h^g_{y^*}}(x^*, \langle x^*, -x\rangle), \quad x \in \mathcal{X}. \tag{3.0.3}$$

The availability of (3.0.3) will be useful in Chapter 4 when obtaining the dual representation of a quasiconvex compositions.

Chapter 4

Quasiconvex compositions

The aim of this chapter is to establish dual representations for quasiconvex compositions. We continue working in the framework of Chapter 3, where we have locally convex topological vector spaces $\mathcal{X}, \mathcal{Y}$ with respective preorders $\leq_C, \leq_D$.

4.1 The main theorem

Let us fix two functions $f\colon \mathcal{Y} \to \overline{\mathbb{R}}$ and $g\colon \mathcal{X} \to \mathcal{Y}$. To motivate the discussion, we make the following simple observation: if f is decreasing and quasiconvex, and g is increasing and D-naturally quasiconcave, then $f \circ g$ is decreasing and quasiconvex by Proposition 3.0.4 and Proposition 3.0.6. Hence, in view of Theorem 2.3.11, a dual representation for $f \circ g$ is readily available once $f \circ g$ is guaranteed to be lower semicontinuous. This is achieved in the next proposition by suitable continuity assumptions on f and g.

Proposition 4.1.1. *Suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is increasing, D-lower demicontinuous, and D-naturally quasiconcave. Then, $f \circ g$ is a decreasing, lower semicontinuous, and quasiconvex function. Moreover, for every $x \in \mathcal{X}$, we have*

$$f \circ g(x) = \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l} \left(y^*, \sup_{x^* \in C^+ \setminus \{0\}} \alpha_{-h_{y^*}^g}^{-l} \left(x^*, \langle x^*, -x \rangle \right) \right) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha_{f \circ g}^{-l} \left(x^*, \langle x^*, -x \rangle \right). \tag{4.1.1}$$

Proof. By Proposition 3.0.4 and Proposition 3.0.6, $f \circ g$ is decreasing and quasiconvex. Let us show that it is also lower semicontinuous. To that end, let $m \in \mathbb{R}$. Note that

$$S_m^{f\circ g} = \{x \in \mathcal{X} \mid f \circ g(x) \leq m\} = \{x \in \mathcal{X} \mid g(x) \in S_m^f\} = \{x \in \mathcal{X} \mid G(x) \subseteq S_m^f\} = G^U(S_m^f). \tag{4.1.2}$$

Here, only the third equality needs a proof. Since f is decreasing, S_m^f is monotone. Let $x \in \mathcal{X}$ with $g(x) \in S_m^f$, and $d \in D$. Since S_m^f is monotone, we have $g(x) + d \in S_m^f$. As this is true for every $d \in D$, we have $G(x) = g(x) + D \subseteq S_m^f$. Conversely, let $x \in \mathcal{X}$ with $G(x) \subseteq S_m^f$. Since $0 \in D$, we have $g(x) \in g(x) + D = G(x) \subseteq S_m^f$. These observations verify the third equality in (4.1.2).

By Remark 2.1.6, we may write $S_m^f = \bigcap_{M\in\mathcal{M}} M$, where $\mathcal{M}$ is the collection of all closed half spaces M such that $S_m^f \subseteq M$. Therefore,

$$G^U(S_m^f) = G^U\left(\bigcap_{M\in\mathcal{M}} M\right) = \bigcap_{M\in\mathcal{M}} G^U(M).$$

Since g is D-lower demicontinuous, $G^U(M)$ is closed for each $M \in \mathcal{M}$. By (4.1.2), it follows that $S_m^{f\circ g} = G^U(S_m^f)$ is closed. Therefore, $f \circ g$ is lower semicontinuous by Remark 3.0.11.

By applying Theorem 2.3.11, we obtain the second dual representation in (4.1.1) immediately.

Finally, we show the first equality in (4.1.1). Let $x \in \mathcal{X}$. By applying Theorem 2.3.11 for f at the point $g(x)$, we get

$$f(g(x)) = \sup_{y^*\in D+\setminus\{0\}} \alpha_f^{-l}(y^*, \langle y^*, -g(x)\rangle).$$

On the other hand, by (3.0.3), we have

$$\langle y^*, -g(x)\rangle = -h_{y^*}^g(x) = \sup_{x^*\in C^+\setminus\{0\}} \alpha_{-h_{y^*}^g}^{-l}(x^*, \langle x^*, -x\rangle), \quad y^* \in D^+ \setminus \{0\}.$$

Combining the last two observations gives the first equality in (4.1.1). $\square$

Under the assumptions of Proposition 4.1.1, $f \circ g$ has a dual representation in the sense of Theorem 2.3.11. We have a a more explicit dual representation for $f \circ g$ in the next theorem.

Theorem 4.1.2. *Suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is increasing, D-lower demicontinuous, and D-naturally quasiconcave. We have*

$$f \circ g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}\Big(y^*, \alpha^{-l}_{-h^g_{y^*}}(x^*, \langle x^*, -x \rangle)\Big), \quad x \in \mathcal{X}.$$

Proof. By Lemma 2.1.1, Lemma 2.3.7 and Lemma 2.3.10(i), we have

$$\begin{aligned} f \circ g(x) &= \inf\{m \in \mathbb{R} \mid g(x) \in S^f_m\} && (4.1.3) \\ &= \inf\left\{m \in \mathbb{R} \mid \forall y^* \in D^+ \setminus \{0\} \colon \langle y^*, -g(x) \rangle \le \alpha_f(y^*, m)\right\} \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \inf\left\{m \in \mathbb{R} \mid \langle y^*, -g(x) \rangle \le \alpha_f(y^*, m)\right\}. && (4.1.4) \end{aligned}$$

By using (3.0.3) and then applying Lemma 2.3.10(i), we obtain

$$\begin{aligned} f \circ g(x) &= \sup_{y^* \in D^+ \setminus \{0\}} \inf\left\{m \in \mathbb{R} \mid \langle y^*, -g(x) \rangle \le \alpha_f(y^*, m)\right\} \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \inf\left\{m \in \mathbb{R} \mid \sup_{x^* \in C^+ \setminus \{0\}} \alpha^{-l}_{-h^g_{\tilde{y}^*}}(x^*, \langle -x^*, x \rangle) \le \alpha_f(y^*, m)\right\} \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \inf\left\{m \in \mathbb{R} \mid \forall x^* \in C^+ \setminus \{0\}, \alpha^{-l}_{-h^g_{\tilde{y}^*}}(x^*, \langle -x^*, x \rangle) \le \alpha_f(y^*, m)\right\} \\ &= \sup_{y^* \in D^+ \setminus \{0\}} \sup_{x^* \in C^+ \setminus \{0\}} \inf\left\{m \in \mathbb{R} \mid \alpha^{-l}_{-h^g_{\tilde{y}^*}}(x^*, \langle -x^*, x \rangle) \le \alpha_f(y^*, m)\right\} \\ &= \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}\Big(y^*, \alpha^{-l}_{-h^g_{y^*}}(x^*, \langle x^*, -x \rangle)\Big), \end{aligned}$$

which gives the conclusion of the theorem. □

The main problem is to calculate the minimal penalty function $\alpha_{f \circ g}$ as well as its left inverse $\alpha^{-l}_{f \circ g}$ in terms of the same type of functions for f and g (more precisely, the scalarizations of g). The solution of this problem will be provided by Theorem 4.1.6 and Corollary 4.1.8. It turns out that these results work under a mild compactness assumption on D^+ as we describe next.

Definition 4.1.3. *A set $\bar{D}^+ \subseteq D^+$ is called a* cone generator *for D^+ if every $y^* \in D^+ \setminus \{0\}$ can be written as $y^* = \lambda \bar{y}^*$ for some $\lambda > 0$ and $\bar{y}^* \in \bar{D}^*$.*

It is clear that if $\bar{D}^+$ is a cone generator for D^+, then D^+ is the conic hull of $\bar{D}^+$.

Remark 4.1.4. Suppose that $D^{\#} \neq \emptyset$ and let $\pi \in D^{\#}$. Then, D^+_π is a closed convex cone generator for D^+ thanks to Lemma 2.2.3.

In Section 4.4, we will discuss the existence and compactness of cone generators for several examples that show up frequently in applications. For the theoretical development of this chapter, we work under the following assumption.

Assumption 4.1.5. *There exists a convex and compact cone generator $\bar{D}^+$ for D^+.*

Next, we state the main theorem of the paper, which provides a formula for the minimal penalty function of $f \circ g$. Its proof is presented separately in Section 6.1. The proof consists of several auxiliary results together with the use of a minimax inequality in [25] for two functions. Assumption 4.1.5 will be crucial in applying this inequality.

Theorem 4.1.6. *Suppose that Assumption 3.0.9 and Assumption 4.1.5 hold. In addition, suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is regularly increasing, D-lower demicontinuous, and D-naturally quasiconcave. Then, for every $x^* \in C^+ \setminus \{0\}$ and $m \in \mathbb{R}$, we have*

$$\alpha_{f\circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \alpha_{-h^g_{y^*}}(x^*, \alpha_f(y^*, m)) = \inf_{y^* \in D^+_\pi \setminus \{0\}} \alpha_{-h^g_{y^*}}(x^*, \alpha_f(y^*, m)) . \tag{4.1.5}$$

Remark 4.1.7. It should be noted that $\bar{D}^+$ does not have to be the same as D^+_π but the second equality in (4.1.5) still holds.

The next corollary complements Theorem 4.1.6 by providing a formula for the left inverse of the minimal penalty funtion of $f \circ g$, which is the actual function that shows up in the dual representation of $f \circ g$ in Proposition 4.1.1. Its proof is given in Section 6.1.

Corollary 4.1.8. *In the setting of Theorem 4.1.6, for every $x^* \in C^+ \setminus \{0\}$ and $s \in \mathbb{R}$, we have*

$$\alpha^{-l}_{f\circ g}(x^*, s) = \sup_{y^* \in D^+ \setminus \{0\}} \alpha^{-l}_f\Big(y^*, \alpha^{-l}_{-h^g_{y^*}}(x^*, s)\Big) \tag{4.1.6}$$

and

$$f \circ g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha^{-l}_f\Big(y^*, \alpha^{-l}_{-h^g_{y^*}}(x^*, \langle x^*, -x\rangle)\Big), \quad x \in \mathcal{X}. \tag{4.1.7}$$

Remark 4.1.9. The second part Corollary 4.1.8 gives the same result as Theorem 4.1.2. In Corollary 4.1.8, we use the stronger Theorem 4.1.6.

4.2 Two important special cases

We consider special cases of the setting in Section 4.1 where at least one of the functions in the composition is convex/concave. In these cases, it is possible to obtain simplified formulae for the minimal penalty function of the composition. As before, we work with two functions $f\colon \mathcal{Y} \to \overline{\mathbb{R}}$ and $g\colon \mathcal{X} \to \mathcal{Y}$.

We first work on the case where both f and g satisfy a stronger convexity assumption so that $f \circ g$ becomes convex. As the next corollary shows, the reduced form of the dual representation is consistent with the ones available for convex compositions in the literature; see, for instance, Theorem 2.8.10 in [2] and Theorem 3 in [3].

Corollary 4.2.1. *Suppose that $f\colon \mathcal{Y} \to \overline{\mathbb{R}}$ is convex, decreasing and lower semicontinuous; and that g is increasing, D-lower demicontinuous, and D-concave. Then, we have*

$$f \circ g(x) = \sup_{x^* \in C^+} \sup_{y^* \in D^+} \left(\langle x^*, -x \rangle - (-h^g_{y^*})^*(-x^*) - f^*(-y^*) \right), \quad x \in \mathcal{X}.$$

Proof. Let $x \in \mathcal{X}$. First, let us prove the following scaling property for an arbitrary $\gamma \geq 0$ and $x^* \in C^+ \setminus \{0\}$.

$$\gamma \alpha^{-l}_{-h^g_{y^*}}(x^*, \langle x^*, -x \rangle) = \alpha^{-l}_{-h^g_{\gamma y^*}}(x^*, \langle x^*, -x \rangle). \tag{4.2.1}$$

Let us consider the case $\gamma > 0$. By Definition 2.3.9, we have

$$\begin{aligned} \alpha^{-l}_{-h^g_{\gamma y^*}}(x^*, \langle x^*, -x \rangle) &= \inf \left\{ m \in \mathbb{R} \mid \alpha_{-h^g_{\gamma y^*}}(x^*, m) \geq \langle x^*, -x \rangle \right\} \\ &= \inf \left\{ m \in \mathbb{R} \mid \sup_{z \in S^m_{-h^g_{\gamma y^*}}} \langle x^*, -z \rangle \geq \langle x^*, -x \rangle \right\}. \end{aligned}$$

We have the following relations:

$$z \in S^m_{-h^g_{\gamma y^*}} \iff \langle \gamma y^*, -g(z) \rangle \leq m \iff \langle y^*, -g(z) \rangle \leq \frac{m}{\gamma} \iff z \in S^{\frac{m}{\gamma}}_{-h^g_{y^*}}.$$

Therefore, we get

$$\begin{aligned}
\alpha^{-l}_{-h^g_{\gamma y^*}}(x^*, \langle x^*, -x\rangle) &= \inf\left\{ m \in \mathbb{R} \mid \sup_{z \in S^m_{-h^g_{\gamma y^*}}} \langle x^*, -z\rangle \geq \langle x^*, -x\rangle \right\} \\
&= \inf\left\{ m \in \mathbb{R} \mid \sup_{z \in S^{\frac{m}{\gamma}}_{-h^g_{y^*}}} \langle x^*, -z\rangle \geq \langle x^*, -x\rangle \right\} \\
&= \gamma \inf\left\{ n \in \mathbb{R} \mid \sup_{z \in S^n_{-h^g_{y^*}}} \langle x^*, -z\rangle \geq \langle x^*, -x\rangle \right\} \\
&= \gamma \alpha^{-l}_{-h^g_{y^*}}(x^*, \langle x^*, -x\rangle),
\end{aligned}$$

where the third equality comes from a change of variable which is $\frac{m}{\gamma} = n$ and the last equality is by Definition 2.3.9. For the case $\gamma = 0$, we will prove that $\alpha^{-l}_{-h^g_{\gamma y^*}}(x^*, \langle x^*, -x\rangle) = 0$. Observe that $S^m_{-h^g_{\gamma y^*}} = \{z \in \mathcal{X} \mid \langle 0, g(z)\rangle \leq m\}$. Therefore, if $m \geq 0$, $S^m_{-h^g_{\gamma y^*}} = \mathcal{X}$ and $S^m_{-h^g_{\gamma y^*}} = \emptyset$ if $m < 0$. Therefore, for $\gamma = 0$, we have

$$\begin{aligned}
\alpha^{-l}_{-h^g_{\gamma y^*}}(x^*, \langle x^*, -x\rangle) &= \inf\left\{ m \in \mathbb{R} \mid \sup_{z \in S^m_{-h^g_{\gamma y^*}}} \langle x^*, -z\rangle \geq \langle x^*, -x\rangle \right\} \\
&= \inf\left\{ m \geq 0 \mid \sup_{z \in \mathcal{X}} \langle x^*, -z\rangle \geq \langle x^*, -x\rangle \right\} = 0.
\end{aligned}$$

We have proved (4.2.1), now we continue with the main result. By Theorem 4.1.2, we have

$$f \circ g(x) = \sup_{x^* \in C^+ \backslash \{0\}} \sup_{y^* \in D^+ \backslash \{0\}} \alpha^{-l}_f\Big(y^*, \alpha^{-l}_{-h^g_{y^*}}(x^*, \langle x^*, -x\rangle)\Big).$$

Once we apply the second part of Proposition 2.3.13 to f and use (4.2.1), we get

$$\begin{aligned}
f \circ g(x) &= \sup_{x^* \in C^+ \backslash \{0\}} \sup_{y^* \in D^+ \backslash \{0\}} \sup_{\gamma \geq 0} \Big(\gamma \alpha^{-l}_{-h^g_{y^*}}(x^*, \langle x^*, -x\rangle) - f^*(-\gamma y^*)\Big) \\
&= \sup_{x^* \in C^+ \backslash \{0\}} \sup_{y^* \in D^+ \backslash \{0\}} \sup_{\gamma \geq 0} \Big(\alpha^{-l}_{-h^g_{\gamma y^*}}(x^*, \langle x^*, -x\rangle) - f^*(-\gamma y^*)\Big) \\
&= \sup_{x^* \in C^+ \backslash \{0\}} \sup_{\tilde{y}^* \in D^+} \Big(\alpha^{-l}_{-h^g_{\tilde{y}^*}}(x^*, \langle x^*, -x\rangle) - f^*(-\tilde{y}^*)\Big),
\end{aligned}$$

where the last equation comes from the change of variable $\gamma y^* = \tilde{y}^*$ since D^+ is a cone. Now let us apply Proposition 2.3.13 to $-h^g_{\tilde{y}^*}$ and obtain

$$\begin{aligned}
f\circ g(x) &= \sup_{x^*\in C^+\setminus\{0\}}\sup_{\tilde{y}^*\in D^+}\Big(\alpha^{-l}_{-h^g_{\tilde{y}^*}}(x^*,\langle x^*,-x\rangle)-f^*(-\tilde{y}^*)\Big)\\
&= \sup_{x^*\in C^+\setminus\{0\}}\sup_{\tilde{y}^*\in D^+}\left(\sup_{\beta\geq 0}\left(\beta\left\langle x^*,-x\right\rangle-(-h^g_{\tilde{y}^*})^*(-\beta x^*)\right)-f^*(-\tilde{y}^*)\right)\\
&= \sup_{\tilde{y}^*\in D^+}\sup_{x^*\in C^+\setminus\{0\}}\sup_{\beta\geq 0}\big(\langle\beta x^*,-x\rangle-(-h^g_{\tilde{y}^*})^*(-\beta x^*)-f^*(-\tilde{y}^*)\big)\\
&= \sup_{\tilde{y}^*\in D^+}\sup_{\tilde{x}^*\in C^+}\big(\langle\tilde{x}^*,-x\rangle-(-h^g_{\tilde{y}^*})^*(-\tilde{x}^*)-f^*(-\tilde{y}^*)\big),
\end{aligned}$$

where the last equation comes from the change of variable $\beta x^*=\tilde{x}^*$ since C^+ is a cone. □

Next, we work on the case where only one of the functions in the composition has a stronger convexity assumption. While Corollary 4.2.1 reproduces earlier results in the literature, the next result is novel to this work to the best of our knowledge. In Section 5.1, we will use this result to obtain new dual representations for quasiconvex systemic risk measures.

Proposition 4.2.2. *Suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is increasing, D-lower demicontinuous, and D-concave. Then, $f\circ g$ is an decreasing, lower semicontinuous, and quasiconvex function; and the following dual representation holds:*

$$f\circ g(x)=\sup_{x^*\in C^+\setminus\{0\}}\sup_{y^*\in D^+\setminus\{0\}}\alpha^{-l}_f\Big(y^*,\langle x^*,-x\rangle-(-h^g_{y^*})^*(-x^*)\Big),\quad x\in\mathcal{X}.\tag{4.2.2}$$

Assume further that g is also regularly increasing and Assumption 4.1.5 holds. The following results hold.

(i) Let $x^*\in C^+\setminus\{0\}$ and $m\in\mathbb{R}$ such that $\alpha_f(y^*,m)\in\mathbb{R}$ and $S^{\alpha_f(y^*,m)}_{-h^g_{y^*}}\neq\emptyset$ for all $y^*\in D^+\setminus\{0\}$. Then,

$$\alpha_{f\circ g}(x^*,m)=\inf_{y^*\in D^+\setminus\{0\}}\big((-h^g_{y^*})^*(-x^*)+\alpha_f(y^*,m)\big).$$

(ii) For every $x^*\in C^+\setminus\{0\}$ and $s\in\mathbb{R}$,

$$\alpha^{-l}_{f\circ g}(x^*,s)=\sup_{y^*\in D^+\setminus\{0\}}\alpha^{-l}_f\big(y^*,(-(-h^g_{y^*})^*(0))\vee(s-(-h^g_{y^*})^*(-x^*))\big).$$

Proof. Note that $x \mapsto \langle y^*, -g(x)\rangle$ is convex and lower semicontinuous by Proposition 3.0.2 and Proposition 3.0.12. Using (4.1.4) in the proof of Theorem 4.1.2 and the Fenchel-Moreau theorem, we have

$$\begin{aligned}
f \circ g(x) &= \sup_{y^* \in D^+ \setminus \{0\}} \inf \{m \in \mathbb{R} \mid \langle y^*, -g(x)\rangle \leq \alpha_f(y^*, m)\} \\
&= \sup_{y^* \in D^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \sup_{x^* \in C^+ \setminus \{0\}} \left(\langle -x^*, x\rangle - (-h^g_{y^*})^*(-x^*)\right) \leq \alpha_f(y^*, m) \right\} \\
&= \sup_{y^* \in D^+ \setminus \{0\}} \sup_{x^* \in C^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \langle -x^*, x\rangle - (-h^g_{y^*})^*(-x^*) \leq \alpha_f(y^*, m) \right\} \\
&= \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}\Big(y^*, \langle x^*, -x\rangle - (-h^g_{y^*})^*(-x^*)\Big),
\end{aligned}$$

where the third equality comes from Lemma 2.3.10(ii). Hence, (4.2.2) follows.

From now on, we assume that g is regularly increasing and Assumption 4.1.5 holds. To prove (i), let $x^* \in C^+ \setminus \{0\}$ and $m \in \mathbb{R}$ with $S^{\alpha_f(y^*,m)}_{-h^g_{y^*}} \neq \emptyset$. By Theorem 4.1.6, we have

$$\alpha_{f \circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \alpha_{-h^g_{y^*}}(x^*, \alpha_f(y^*, m)).$$

Also, take $x \in S^{\alpha_f(y^*,m)}_{-h^g_{y^*}}$ and let $c \in C^{\#}$. Then, there exists $d \in D^{\#}$ such that $g(x + c) = g(x) + d$ since g is regularly increasing. Therefore, by using the definition of $D^{\#}$, we get

$$\langle y^*, -g(x+c)\rangle = \langle -y^*, g(x) + d\rangle = \langle -y^*, g(x)\rangle + \langle -y^*, d\rangle < \langle -y^*, g(x)\rangle \leq \alpha_f(y^*, m),$$

which gives that the strict sublevel set $\{x \in \mathcal{X} \mid -h^g_{y^*}(x) < \alpha_f(y^*, m)\}$ is nonempty. Hence, by Proposition 2.3.13, we have

$$\alpha_{f \circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \inf_{\beta > 0} \left(\beta(-h^g_{y^*})^* \left(-\frac{x^*}{\beta}\right) + \beta \alpha_f(y^*, m) \right).$$

Then, by Theorem 2.3.1 in [2] on the elementary rules of conjugation, we have

$$\alpha_{f \circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \inf_{\beta > 0} \left((-\beta h^g_{y^*})^*(-x^*) + \beta \alpha_f(y^*, m) \right).$$

By the positive homogeneity of $y^* \mapsto \alpha_f(y^*, m)$ and also that of $y^* \mapsto h^g_{y^*}(x)$ for each $x \in \mathcal{X}$, we get

$$\alpha_{f \circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \inf_{\beta > 0} \left((-h^g_{\beta y^*})^*(-x^*) + \alpha_f(\beta y^*, m) \right).$$

Finally, since D^+ is cone, we can make a change of variables and obtain (i).

We prove (ii) next. By Corollary 4.1.8, the second part of Proposition 2.3.13, and the definition of left inverse, we have

$$\begin{aligned}
\alpha^{-l}_{f\circ g}(x^*, s) &= \sup_{y^*\in D^+\setminus\{0\}} \alpha^{-l}_f\Big(y^*, \alpha^{-l}_{-h^g_{y^*}}(x^*, s)\Big) \\
&= \sup_{y^*\in D^+\setminus\{0\}} \alpha^{-l}_f\Big(y^*, \sup_{\gamma\geq 0}\big(\gamma s - (-h^g_{y^*})^*(-\gamma x^*)\big)\Big) \\
&= \sup_{y^*\in D^+\setminus\{0\}} \inf\left\{m\in\mathbb{R} \mid \sup_{\gamma\geq 0}\big(\gamma s - (-h^g_{y^*})^*(-\gamma x^*)\big) \leq \alpha_f(y^*, m)\right\} \\
&= \sup_{y^*\in D^+\setminus\{0\}} \sup_{\gamma\geq 0}\inf\big\{m\in\mathbb{R} \mid \gamma s - (-h^g_{y^*})^*(-\gamma x^*) \leq \alpha_f(y^*, m)\big\},
\end{aligned}$$

where the last equality comes from Lemma 2.3.10(ii). By the conjugation formula, for $\gamma > 0$, we have

$$\begin{aligned}
(-h^g_{y^*})^*(-\gamma x^*) &= \sup_{x\in\mathcal{X}}\left(\langle -\gamma x^*, x\rangle + \langle y^*, g(x)\rangle\right) \\
&= \gamma \sup_{x\in\mathcal{X}}\left(\langle -x^*, x\rangle + \left\langle \frac{y^*}{\gamma}, g(x)\right\rangle\right) = \gamma(-h^g_{\frac{y^*}{\gamma}})^*(-x^*).
\end{aligned}$$

For $\gamma = 0$, we have

$$\begin{aligned}
\inf\big\{m\in\mathbb{R} \mid \gamma s - (-h^g_{y^*})^*(-\gamma x^*) \leq \alpha_f(y^*, m)\big\} &= \inf\big\{m\in\mathbb{R} \mid -(-h^g_{y^*})^*(0) \leq \alpha_f(y^*, m)\big\} \\
&= \alpha^{-l}_f(y^*, -(-h^g_{y^*})^*(0)).
\end{aligned}$$

Therefore, by using the previous two equations and the positive homogeneity of α_f, we get

$$
\begin{aligned}
&\alpha^{-l}_{f\circ g}(x^*,s)\\
&= \sup_{y^*\in D^+\setminus\{0\}} \sup_{\gamma\geq 0}\inf\left\{m\in\mathbb{R} \mid \gamma s - (-h^g_{y^*})^*(-\gamma x^*) \leq \alpha_f(y^*,m)\right\}\\
&= \sup_{y^*\in D^+\setminus\{0\}} \left(\alpha_f^{-l}(y^*, -(-h^g_{y^*})^*(0)) \vee \sup_{\gamma>0}\inf\left\{m\in\mathbb{R} \mid \gamma s - \gamma(-h^g_{\frac{y^*}{\gamma}})^*(-x^*) \leq \alpha_f(y^*,m)\right\}\right)\\
&= \sup_{y^*\in D^+\setminus\{0\}} \left(\alpha_f^{-l}(y^*, -(-h^g_{y^*})^*(0)) \vee \sup_{\gamma>0}\inf\left\{m\in\mathbb{R} \mid s - (-h^g_{\frac{y^*}{\gamma}})^*(-x^*) \leq \alpha_f\left(\frac{y^*}{\gamma},m\right)\right\}\right)\\
&= \sup_{y^*\in D^+\setminus\{0\}} \alpha_f^{-l}(y^*, -(-h^g_{y^*})^*(0)) \vee \sup_{\substack{y^*\in D^+\setminus\{0\},\\ \gamma>0}} \inf\left\{m\in\mathbb{R} \mid s - (-h^g_{\frac{y^*}{\gamma}})^*(-x^*) \leq \alpha_f\left(\frac{y^*}{\gamma},m\right)\right\}\\
&= \sup_{y^*\in D^+\setminus\{0\}} \alpha_f^{-l}(y^*, -(-h^g_{y^*})^*(0)) \vee \sup_{y^*\in D^+\setminus\{0\}} \inf\left\{m\in\mathbb{R} \mid (s - (-h^g_{y^*})^*(-x^*)) \leq \alpha_f(y^*,m)\right\}\\
&= \sup_{y^*\in D^+\setminus\{0\}} \alpha_f^{-l}(y^*, -(-h^g_{y^*})^*(0)) \vee \sup_{y^*\in D^+\setminus\{0\}} \alpha_f^{-l}\Big(y^*, s - (-h^g_{y^*})^*(-x^*)\Big)\\
&= \sup_{y^*\in D^+\setminus\{0\}} \Big(\alpha_f^{-l}\big(y^*, -(-h^g_{y^*})^*(0)\big) \vee \alpha_f^{-l}\big(y^*, s - (-h^g_{y^*})^*(-x^*)\big)\Big).
\end{aligned}
$$

By the monotonicity of α_f^{-l}, we can also write the last line as

$$
\sup_{y^*\in D^+\setminus\{0\}} \alpha_f^{-l}\big(y^*, (-(-h^g_{y^*})^*(0)) \vee (s - (-h^g_{y^*})^*(-x^*))\big),
$$

which completes the proof. □

4.3 Quasiconvex composition on a convex set

We turn our attention to the case where the composition is considered on a monotone convex set $\mathcal{K}\subseteq\mathcal{X}$ with $C\subseteq\mathcal{K}$, see Corollary 2.3.12, the analogous result for a single function. The treatment here will be relevant for some applications in Chapter 5.

We work with two functions $f\colon\mathcal{Y}\to\overline{\mathbb{R}}$ and $g\colon\mathcal{K}\to\mathcal{Y}$. The following results extend Theorem 4.1.6 and Theorem 4.1.2. Their proofs are given in Section 6.1.

Corollary 4.3.1. *Suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is regularly increasing, D-lower demicontinuous (with respect to the relative topology), and D-naturally quasiconcave. Then, $f\circ g$ is an decreasing, lower semicontinuous, and* *quasiconvex function. Moreover, for each $x^* \in C^+ \setminus \{0\}$ and $m \in \mathbb{R}$, we have*

$$\alpha_{f \circ g}(x^*, m) = \inf_{y^* \in D^+ \setminus \{0\}} \alpha_{-h^g_{y^*}} \left(x^*, \alpha_f(y^*, m)\right).$$

Proposition 4.3.2. *Suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is increasing, D-lower demicontinuous (with respect to the relative topology), and D-naturally quasiconcave. Then, we have*

$$f \circ g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \alpha^{-l}_{f \circ g} \left(x^*, \langle x^*, -x \rangle\right), \quad x \in \mathcal{K}, \tag{4.3.1}$$

and

$$f \circ g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha^{-l}_f \Big(y^*, \alpha^{-l}_{-h^g_{y^*}} \left(x^*, \langle x^*, -x \rangle\right)\Big), \quad x \in \mathcal{K}. \tag{4.3.2}$$

For a more specific case, we have the following proposition.

Proposition 4.3.3. *Suppose that f is decreasing, lower semicontinuous, and quasiconvex; and that g is increasing, D-lower demicontinuous (with respect to the relative topology), and concave. Then, we have*

$$f \circ g(x) = \sup_{x^* \in C^+ \setminus \{0\}} \sup_{y^* \in D^+ \setminus \{0\}} \alpha^{-l}_f \Big(y^*, \langle x^*, -x \rangle - (-h^g_{y^*})^*(-x^*)\Big), \quad x \in \mathcal{X}. \tag{4.3.3}$$

4.4 Compact cone generators

In this section, we will discuss the existence of compact convex cone generators in some specific spaces and show that Theorem 4.1.6 is applicable in these spaces. This will justify the use of our results in the context of systemic risk measures in Section 5.1.

As mentioned in Remark 4.1.4, D^+_π is a closed convex generator but it is not always compact. However, we do not have to restrict ourselves to this generator and can search for other generators because after guaranteeing the existence of a compact convex cone generator $\bar{D}^+$, we can still work on D^+_π thanks to Equation (4.1.5).

4.4.1 Finite-dimensional spaces

Let us take $\mathcal{Y} = \mathbb{R}^n$ with the Euclidean norm $\|\cdot\|$, as a natural consequence $\mathcal{Y}^* = \mathbb{R}^n$ with the same norm $\|\cdot\|$. Choose a convex cone D and denote the unit ball with $B = \{y \in \mathbb{R}^n : \|y\| \leq 1\}$. We will show the existence of a compact convex generator for D^+ so we are able to use our main theorem for the case $\mathcal{Y} = \mathbb{R}^n$.

Proposition 4.4.1. *The set $\bar{D}^+ := D^+ \cap B$ is compact and convex, and it is a cone generator for D^+.*

Proof. Since D^+ and B are closed and convex their intersection will also be closed and convex. Also, B is compact since it is closed and bounded. By using this fact and $\bar{D}^+$ is a closed subset of B we have $\bar{D}^+$ is also compact.

Now let us show that $\bar{D}^+$ generates D^+. Take an element $y^* \in D^+ \setminus \{0\}$. We have $\frac{y^*}{\|y^*\|} \in D^+$ since D^+ is a cone and $\left\|\frac{y^*}{\|y^*\|}\right\| = 1$ which implies that it is in B and hence in $\bar{D}^+$. We can write $y^* = \|y^*\| \frac{y^*}{\|y^*\|}$ where $\|y^*\| > 0$ and $\frac{y^*}{\|y^*\|} \in \bar{D}^+$; hence, $\bar{D}^+$ is a cone generator for D^+. □

4.4.2 Lebesgue spaces

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $p \in [1, +\infty]$, $n \in \mathbb{N}$. We denote by $L^0(\mathbb{R}^n)$ the space of n-dimensional random vectors that are identified up to $\mathbb{P}$-almost sure equality. We denote by $L^p(\mathbb{R}^n)$ the space of all $X \in L^0(\mathbb{R}^n)$ such that $\|X\|_p < +\infty$, where $\|X\|_p := (\mathbb{E}[\|X\|^p])^{1/p}$ for $p < +\infty$ and $\|X\|_p := \inf\{c > 0 \mid \mathbb{P}\{\|X\| \leq c\} = 1\}$ for $p = +\infty$. For $p \in \{0\} \cup [1, +\infty]$ and a set $A \subseteq \mathbb{R}^n$, we denote by $L^p(A)$ the set of all $X \in L^p(\mathbb{R}^n)$ such that $\mathbb{P}\{X \in A\} = 1$.

In this section, we fix $p \in [1, +\infty)$ and consider the case $\mathcal{Y} = L^p(\mathbb{R}^n)$, which is equipped with the norm $\|\cdot\|_p$ and the induced topology. As a consequence, $\mathcal{Y}^* = L^q(\mathbb{R}^n)$ with the norm $\|\cdot\|_q$ and we consider it with weak topology $\sigma(\mathcal{Y}^*, \mathcal{Y})$, where the conjugate exponent $q \in (1, +\infty]$ is defined by the relation $\frac{1}{p} + \frac{1}{q} = 1$. Let $D \subseteq \mathcal{Y}$ be a closed convex cone and denote the unit ball in $L^q(\mathbb{R}^n)$ by $B_q^n = \{Y^* \in L^q(\mathbb{R}^n) \mid \|Y^*\|_q \leq 1\}$. We will show the existence of a compact convex cone generator for D^+ so we are able to use Theorem 4.1.6

for the case $\mathcal{Y} = L^p(\mathbb{R}^n)$.

Proposition 4.4.2. *The set $\bar{D}^+ := D^+ \cap B_q^n$ is a $\sigma(\mathcal{Y}^*, \mathcal{Y})$-compact convex set and it is a cone generator for D^+.*

Proof. Since D^+ and B_q^n are closed convex sets, so is their intersection $\bar{D}^+$. Also, B_q^n is $\sigma(L^q(\mathbb{R}^n), L^p(\mathbb{R}^n))$-compact by Banach-Alaoglu Theorem (Theorem IV.21 in [30]). By using this fact and that $\bar{D}^+$ is a closed subset of B_q, we conclude that $\bar{D}^+$ is also compact.

Next, we show that $\bar{D}^+$ generates D^+. Let $Y^* \in D^+ \setminus \{0\}$. We have $\frac{Y^*}{\|Y^*\|_q} \in D^+$ since D^+ is a cone and $\left\| \frac{Y^*}{\|Y^*\|_q} \right\|_q = 1$ which implies it is in B_q and hence in $\bar{D}^+$. We can write $Y^* = \|Y^*\|_q \frac{Y^*}{\|Y^*\|_q}$, where $\|Y^*\|_q > 0$ and $\frac{Y^*}{\|Y^*\|_q} \in \bar{D}^+$. Hence, $\bar{D}^+$ is a cone generator for D^+. $\square$

Let us look at the special case $n = 1$ and take $D = L^p(\mathbb{R}_+)$ which is the set of all almost surely positive elements of $L^p(\mathbb{R})$, then $D^+ = L^q(\mathbb{R}_+)$. Also, we can take $\pi = 1$ and get $D_1^+ \cong \mathcal{M}_1^q(\mathbb{P})$, the set of all probability measures $\mathbb{Q}$ that are absolutely continuous with respect to $\mathbb{P}$ and with Radon-Nikodym derivatives $\frac{d\mathbb{Q}}{d\mathbb{P}}$ in $L^q(\mathbb{R}_+)$. Therefore the formula in Theorem 4.1.6 can be rewritten as

$$\alpha_{f\circ g}(X^*, m) = \inf_{\mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P})} \alpha_{-h^g_{\frac{d\mathbb{Q}}{d\mathbb{P}}}} \left(X^*, \alpha_f \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m \right) \right). \tag{4.4.1}$$

We can work with any closed convex cone generator after guaranteeing the existence of a compact convex cone generator since we do not need compactness for the second equation in Theorem 4.1.6. Therefore, the dual representation in Theorem 4.1.6 can be written in this way since $D_{\mathbf{1}}^+$ is a closed convex cone generator by Remark 4.1.4.

Chapter 5

Applications to systemic risk measures

In this chapter, we will explore the implications of the general theory developed in Chapter 4 for some quasiconvex risk measures for interconnected financial systems. Such risk measures are referred to as *systemic risk measures*, which is of recent interest in the financial mathematics literature. We refer the interested reader to [17], [19], [18], [20].

5.1 Results on general systemic risk measures

Risk measures are used in financial mathematics for calculating capital requirements of financial positions. There are different theoretical approaches for risk measures, each of them working with its own set of assumptions. However, the most common elements of a risk measure are monotonicity and diversification. Monotonicity corresponds to the fact that if value of an financial asset increases then its risk must decrease. Hence, a risk measure should be a decreasing function. Diversification means that when we diverse our investment it should not increase the risk. Some of the literature (e.g. [11]) uses convexity in order to satisfy this condition but quasiconvexity is more suitable for the concept of diversification. Also, cash additivity is sometimes used as an assumption for risk measures (e.g. in [10], [11]). When we are trying to weaken this assumption by cash subadditivity, it is suggested in [31] that convexity should be replaced with quasiconvexity. Therefore the assumptions in our main theorem (Theorem 4.1.6) will not restrict us for the applications in risk measures.

We want to calculate the risk of a financial system so there is also need to define a risk concept for systems, which are systemic risk measures. In the literature, most of the systemic risk measures (e.g.[19], [17]) are in the form $R(X) = \rho \circ \Lambda(X)$, where ρ is a risk measure and Λ is the aggregation function. which will be defined at Definition 5.1.1. We want to save the properties of monotonicity and diversification in this form; therefore, assuming the quasiconvexity and monotonicity of the aggregation functions is useful. Regularly increasing property for the aggregation function can be considered as if the system has a significant positive change then its value changes significantly too. Also the lower semicontinuity assumption is a mild regularity assumption. Hence, our assumptions in the theorems are not restrictive for the systemic risk measures.

We will work on the spaces $\mathcal{X} = L^p(\mathbb{R}^n)$ and $\mathcal{Y} = L^p(\mathbb{R})$, where $p \in [1, +\infty]$. These spaces are equipped with their norm topologies when $p < +\infty$ and with the weak* topologies when $p = +\infty$. In all cases, we have $\mathcal{X}^* = L^q(\mathbb{R}^n)$ and $\mathcal{Y}^* = L^q(\mathbb{R})$, with their weak topologies, where $q \in [1, +\infty]$ is determined by $\frac{1}{p} + \frac{1}{q} = 1$. We denote by $\mathcal{M}_n^q(\mathbb{P})$ the set of all vectors $\mathbb{S} = (\mathbb{S}_1, \ldots, \mathbb{S}_n)$, where $\mathbb{S}_i$ is a probability measure on $(\Omega, \mathcal{F})$ that is absolutely continuous with respect to $\mathbb{P}$ and $\frac{d\mathbb{S}_i}{d\mathbb{P}} \in L^q(\mathbb{R}_+)$ for each $i \in \{1, \ldots, n\}$. We take $C = L^p(\mathbb{R}_+^n)$ and $D = L^p(\mathbb{R}_+)$; hence the dual cones are given by $C^+ = L^q(\mathbb{R}_+^n)$ and $D^+ = L^q(\mathbb{R}_+)$. With this choice of D, for the sake of convenience, we will remove D from the terminology; for instance, we will simply call a function concave if it is D-concave.

Definition 5.1.1. *(i) A function $\rho \colon L^p(\mathbb{R}) \to \overline{\mathbb{R}}$ is called a* quasiconvex risk measure *if ρ is quasiconvex and decreasing.*

(ii) A function $\Lambda \colon L^p(\mathbb{R}^n) \to L^p(\mathbb{R})$ is called an aggregation function *if it is increasing.*

(iii) A function $R \colon L^p(\mathbb{R}^n) \to \overline{\mathbb{R}}$ is called a quasiconvex systemic risk measure *if*

$$R = \rho \circ \Lambda,$$

where ρ is a quasiconvex risk measure and Λ is an aggregation function.

When we consider the dual representation in Theorem 4.1.2 in the systemic risk measure context, as long as it satisfies our assumptions, it reads as

$$R(X) = \rho \circ \Lambda(X) = \sup_{X^* \in L^q(\mathbb{R}^n_+)\backslash\{0\}} \sup_{Y^* \in L^q(\mathbb{R}_+)\backslash\{0\}} \alpha_\rho^{-l}\left(Y^*, \alpha^{-l}_{-h^{\Lambda}_{Y^*}}\left(X^*, -\mathbb{E}\left[X^\mathsf{T} X^*\right]\right)\right).$$

As a generalization of the economic interpretations in [20] for the convex case, we may interpret the above formula as follows. The aggregation function Λ calculates the effect of financial institutions on society, and the risk measure ρ calculates the risk of society. In this dual representation, we first consider the lost of the system which is evaluated under X^* by $\mathbb{E}[-X^\mathsf{T} X^*]$ and control the plausibility of X^* when the condition of society is considered as Y^* via the function $\alpha^{-l}_{-h^{\Lambda}_{Y^*}}$. Then we calculate the plausibility of the Y^* by the function α_ρ^{-l} and take the supremum of them which can be considered as calculating worst case. Now we will focus on more specific conditions and look for more specific interpretations after passing to the probabilistic settings.

Note that when ρ is quasiconvex, the systemic risk measure in the form $R = \rho \circ \Lambda$ is a quasiconvex systemic risk measure whether aggregation function is concave or natural quasiconcave. Therefore, even when ρ is quasiconvex and the aggregation function is concave we will have dual representations for systemic risk measures which are new in the literature to the best of our knowledge. Now, we will give some examples for this purpose.

First, we start with quasiconvex risk measures of the form

$$\rho(Y) = \ell^{-1}\left(\mathbb{E}[\ell \circ (-Y)]\right), \quad Y \in L^p(\mathbb{R}),$$

where $p \in [1, +\infty]$, and $\ell \colon \mathbb{R} \to (-\infty, \infty]$ is a proper lower semicontinuous convex increasing function, called a *loss function*. For simplicity, we assume that ℓ is differentiable. Such ρ is called the *certainty equivalent* associated to loss function ℓ. It is found in [8] that

$$\alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m\right) = \mathbb{E}_\mathbb{Q}\left[h \circ \left(\beta \frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right],$$

where $\mathbb{Q} \in \mathcal{M}_1(\mathbb{P})$, h is the right inverse of the derivative ℓ', $\beta = \beta(\mathbb{Q}, m)$ is the solution of the equation $\mathbb{E}\left[\ell \circ h \circ (\beta \frac{d\mathbb{Q}}{d\mathbb{P}})\right] = \ell^+(m)$ under some integrability and positivity conditions.

Let us provide some specific examples for the loss function ℓ and recall the penalty functions for the corresponding certainty equivalents, already calculated in ([8], Example 8).

Later, we will combine these choices of ℓ to construct systemic risk measures.

Example 5.1.2. (i) (Quadratic loss function) Let us take $p = 2$, and $\ell(s) = s^2/2 + s$ for $s \geq -1$ and $\ell(s) = -\frac{1}{2}$ for $s < -1$. Then for $m \geq -1$ we have $\alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m\right) = -1$, and for each $\mathbb{Q} \in \mathcal{M}_1^2(\mathbb{P})$ and $m < -1$, we have

$$\alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m\right) = (1+m)\left\|\frac{d\mathbb{Q}}{d\mathbb{P}}\right\|_2 - 1.$$

Moreover, we have

$$\alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, s\right) = \frac{s+1}{\left\|\frac{d\mathbb{Q}}{d\mathbb{P}}\right\|_2} - 1$$

if $s > -1$ and $\alpha_\rho^{-l}(\frac{d\mathbb{Q}}{d\mathbb{P}}, s) = -\infty$ elsewhere.

(ii) (Logarithmic loss function) Let us take $p = 1$ or $p = +\infty$, and $\ell(s) = -\ln(-s)$ for $s < 0$ and $\ell(s) = +\infty$ for $s \geq 0$. Then, for each $\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})$, we have

$$\alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m\right) = me^{\mathbb{E}\left[\ln(\frac{d\mathbb{Q}}{d\mathbb{P}})\right]}, \quad m < 0,$$

and

$$\alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, s\right) = se^{-\mathbb{E}\left[\ln(\frac{d\mathbb{Q}}{d\mathbb{P}})\right]}, \quad s < 0.$$

(iii) (Power loss function) Let us take $p = 1$ or $p = +\infty$, and fix some $\gamma \in (0,1)$. Take $\ell(s) = -\frac{(-s)^{1-\gamma}}{1-\gamma}$ for $s \leq 0$ and $\ell(s) = \infty$ for $s > 0$. Then, for each $\mathbb{Q} \in \mathcal{M}_1^\infty(\mathbb{P})$, we have

$$\alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m\right) = \frac{m}{\left\|\frac{d\mathbb{Q}}{d\mathbb{P}}\right\|_{\frac{\gamma-1}{\gamma}}}, \quad m < 0,$$

and

$$\alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, s\right) = s\left\|\frac{d\mathbb{Q}}{d\mathbb{P}}\right\|_{\frac{\gamma-1}{\gamma}}, \quad s < 0.$$

Here, for $Y^* \in L^1(\mathbb{R})$, we use the notation $\|Y^*\|_a := (\mathbb{E}[|Y^*|^a])^{\frac{1}{a}}$ for $a < 1$ as well, although $\|\cdot\|_a$ is not a norm in general.

In addition to certainty equivalent, we also revisit the *economic index of riskiness* as another example of a quasiconvex risk measure. Based on the loss function ℓ, this risk measure is defined by

$$\rho(Y) = \frac{1}{\sup\{\lambda > 0 \mid \mathbb{E}[\ell \circ (-\lambda Y)] \leq c_0\}},$$

where $c_0 \in \mathbb{R}$ is fixed. To make this risk measure well-defined, ℓ is usually assumed to have the superlinear growth condition $\lim_{s\to\infty} \ell(s)/s = \infty$ and p is chosen in accordance with ℓ. Following the arguments in [8], it can be shown that

$$\alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m\right) = \mathbb{E}_{\mathbb{Q}}\left[mh \circ \left(m\beta\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right], \quad \mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}), m \in \mathbb{R},$$

where $\beta = \beta(\mathbb{Q}, m)$ is the solution of the equation $\mathbb{E}\left[\ell \circ h \circ (m\beta\frac{d\mathbb{Q}}{d\mathbb{P}})\right] = c_0$.

The following example is the analogue of Example 5.1.2(ii) for the economic index of riskiness; see Examples 3 and 9 in [8] for more details.

Example 5.1.3. Let us take $p = 1$ and $c_0 > 0$, and consider $\ell(s) = -\ln(1-s)$ for $s < 1$ and $\ell(s) = +\infty$ for $s \geq 1$. Then, for each $\mathbb{Q} \in \mathcal{M}_1^\infty(\mathbb{P})$, we have

$$\alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m\right) = m\left(1 - \exp\left(\mathbb{E}\left[\ln\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] - c_0\right)\right),$$

and

$$\alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, s\right) = \frac{s}{1 - \exp\left(\mathbb{E}\left[\ln\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] - c_0\right)},$$

where $\exp(x) = e^x$ for $x \in \mathbb{R}$.

These examples of risk measures are taken from [8] which works on lower semicontinuous, quasiconvex and monotone risk measures which are exactly same assumptions with Theorem 4.1.2 and Theorem 4.1.6 in our study. Therefore, we can directly use these results.

In many applications, the aggregation function Λ is defined in terms of a deterministic increasing function $\tilde{\Lambda}\colon \mathbb{R}^n \to \mathbb{R}$ via $\Lambda(X) := \tilde{\Lambda} \circ X$, that is,

$$\Lambda(X(\cdot))(\omega) := \tilde{\Lambda}(X(\omega)), \quad \omega \in \Omega, \tag{5.1.1}$$

for every $X \in L^p(\mathbb{R}^n)$, which we assume for the rest of this chapter. Here, the implicit assumption on $\tilde{\Lambda}$ is that the resulting function Λ is a true aggregation function, that is, $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R}^n)$. On the other hand, to ensure lower demicontinuity of

Λ, we need to impose sufficient regularity on $\tilde{\Lambda}$. This is done in the following lemma.

Lemma 5.1.4. *Let $\tilde{\Lambda}\colon \mathbb{R}^n \to \mathbb{R}$ be an increasing function and define Λ by* (5.1.1). *Suppose that $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R}^n)$, where $p \in [1, +\infty]$.*

(i) Suppose that $\tilde{\Lambda}$ is concave and bounded from above. Then, Λ is concave and lower demicontinuous.

(ii) Suppose that $\tilde{\Lambda}$ is linear. Then, Λ is linear and lower demicontinuous.

(iii) Suppose that $\tilde{\Lambda}$ is regularly increasing with respect to cones $\mathbb{R}^n_+$ and $\mathbb{R}_+$. Then, Λ is regularly increasing.

Proof. The proof is given in Section 6.2. □

We revisit some simple examples of the deterministic function $\tilde{\Lambda}$ from [20]. In each example, we also calculate the conjugate function Φ given by

$$\tilde{\Phi}(x^*) := (-\tilde{\Lambda})^*(-x^*) = \sup_{x\in\mathbb{R}^n} \left(\Lambda(x) - (x^*)^{\mathsf{T}} x\right), \quad x^* \in \mathbb{R}^n,$$

for future use. A more sophisticated aggregation function based on a clearing mechanism will be discussed separately in Section 5.2.

Example 5.1.5. (i) (Total profit-loss model) Take $\tilde{\Lambda}(x) = \sum_{i=1}^n x_i$, then by ([20],4.1) we have

$$\tilde{\Phi}(x^*) = \begin{cases} 0 & \text{if } x^* = \mathbf{1}, \\ \infty & \text{else.} \end{cases}$$

Note that the condition that $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R}^n)$ is satisfied for every choice of $p \in [1, +\infty]$.

(ii) (Total loss model) Take $\tilde{\Lambda}(x) = -\sum_{i=1}^n x_i^-$, then by ([20],4.2) we have

$$\tilde{\Phi}(x^*) = \begin{cases} 0 & \text{if } x_i^* \in [0,1] \text{ for every } i \in \{1, \dots, n\}, \\ \infty & \text{else.} \end{cases}$$

As in the previous example, for every choice of $p \in [1, +\infty]$, we have $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R}^n)$ is satisfied.

(iii) (Exponential aggregation model) Take $\tilde{\Lambda}(x) = -\sum_{i=1}^{n} e^{-x_i-1}$, then by ([20], 4.3) we have

$$\tilde{\Phi}(x^*) = \sum_{i=1}^{n} x_i^* \ln(x_i^*),$$

where $\ln(0) := -\infty$ and $0\ln(0) := 0$ for convention. Note that the condition that $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R}^n)$ is satisfied for $p = +\infty$. We cannot use Theorem 4.1.6 since we can not guarantee the existence of a compact cone generator in Proposition 4.4.2. However, we can still use the dual representation in Theorem 4.1.2.

Thanks to Lemma 5.1.4, in our applications, we can use each of the aggregation functions in Example 5.1.5.

Assuming the structure for the aggregation function in (5.1.1), we calculate the penalty function of a systemic risk measure when the underlying aggregation function is concave and regularly increasing. For convenience, for each $X^* \in L^q(\mathbb{R}^n)$, let us define the set

$$T_{X^*} := \{Y^* \in L^q(\mathbb{R}_+) \mid \mathbb{P}(X^* \neq 0, Y^* = 0) = 0\}. \tag{5.1.2}$$

Similar to (5.1.1), we also define

$$\Phi(X^*) := \tilde{\Phi} \circ X^*, \quad X^* \in L^q(\mathbb{R}^n). \tag{5.1.3}$$

Proposition 5.1.6. *Let $\tilde{\Lambda}\colon \mathbb{R}^n \to \mathbb{R}$ be a concave, regularly increasing function that is either bounded from above or linear. Define Λ by (5.1.1). Suppose that $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R}^n)$, where $p \in [1, +\infty)$. Let ρ be a lower semicontinuous quasiconvex risk measure. Let $X^* \in L^q(\mathbb{R}^n)$ and $m \in \mathbb{R}$ such that the strict sublevel set $\{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}[-Y^*\Lambda(X)] < m\}$ is nonempty for every $Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}$. Then,*

$$\alpha_{\rho\circ\Lambda}(X^*, m) = \inf_{Y^* \in T_{X^*}} \left(\mathbb{E}\left[Y^* \Phi\left(\frac{X^*}{Y^*}\right) 1_{\{Y^*>0\}} \right] + \alpha_\rho(Y^*, m) \right).$$

Proof. The proof is given in Section 6.2. □

Next, we aim to rewrite Proposition 5.1.6 in terms of probability measures. By doing this, we will be able to provide economic interpretations of the dual representation in view of model uncertainty. Since $D_1^+ = \{\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})\}$ is a closed convex cone generator for $D^+ = L^q(\mathbb{R}_+)$, we can write every $Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}$ as $Y^* = \lambda \frac{d\mathbb{Q}}{d\mathbb{P}}$ for some $\lambda > 0$ and $\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})$ by Remark 4.1.4. Similarly every $X^* \in C^+ = L^q(\mathbb{R}_+^n)$ can be written as $X^* = w \cdot \frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}}$, where $w \in \mathbb{R}_+^n \setminus \{0\}$, $\mathbb{S} = (\mathbb{S}_1, \dots, \mathbb{S}_n) \in \mathcal{M}_n^q(\mathbb{P})$, and $w \cdot \frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}} := (w_1 \frac{d\mathbb{S}_1}{d\mathbb{P}}, \dots, w_n \frac{d\mathbb{S}_n}{d\mathbb{P}})$. The interpretation of these dual variables is as follows. In the presence of model uncertainty, we consider $\mathbb{Q}$ as a probability measure that is assigned to an external entity (e.g., society) and, for each $i \in \{1, \dots, n\}$, $\mathbb{S}_i$ is a probability measure that is assigned to the internal entity i (e.g., a bank in the network) with corresponding weight w_i. Also, since we consider X^* and Y^* satisfying the condition $\mathbb{P}(X^* \neq 0, Y^* = 0) = 0$ in Proposition 5.1.6, it follows from Lemma 6.3 of [20] that $w_i \mathbb{S}_i$ is a finite measure that is absolutely continuous with respect to $\mathbb{Q}$, and we can write

$$\frac{\frac{w \cdot \mathbf{d}\mathbb{S}}{d\mathbb{P}}}{\frac{d\mathbb{Q}}{d\mathbb{P}}} = \frac{w \cdot \mathbf{d}\mathbb{S}}{d\mathbb{Q}},$$

where all Radon-Nikodym derivatives are well-defined. Therefore, the formula in Proposition 5.1.6 can be written as

$$\alpha_{\rho\circ\Lambda}\left(w \cdot \frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}}, m\right) = \inf_{\substack{\lambda>0, \mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P}), \\ w_i\mathbb{S}_i \ll \mathbb{Q}}} \left(\lambda \alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m\right) + \mathbb{E}_{\mathbb{Q}}\left[\lambda \Phi\left(\frac{w \cdot \mathbf{d}\mathbb{S}}{\lambda d\mathbb{Q}}\right)\right]\right). \tag{5.1.4}$$

We calculate the total penalty of choosing probability measure $\mathbb{S}$ and weight w for the financial institutions by considering all possibilities of society measure $\mathbb{Q}$. It is computed as summing the penalty of society being in the alternative model $\mathbb{Q}$ and the effect of the financial institutions on the society, so it has as an additive property when we are calculating the penalty function.

Proposition 5.1.7. *Let $\tilde{\Lambda} \colon \mathbb{R}^n \to \mathbb{R}$ be a concave, regularly increasing function that is either bounded from above or linear. Define Λ by (5.1.1). Suppose that $\Lambda(X) \in L^p(\mathbb{R}, \mathbb{P})$ for every $X \in L^p(\mathbb{R}^n)$, where $p \in [1, +\infty)$. Let ρ be a lower semicontinuous quasiconvex risk measure.*

(i) Suppose that $\tilde{\Lambda}$ is bounded from above, that is, $\Phi(0) < +\infty$. Then, we have

$$\alpha_{\rho\circ\Lambda}^{-l}(X^*, s) = \max\Bigg\{ \sup_{Y^* \in L_+^q(\mathbb{R}) \setminus \{0\}} \alpha_\rho^{-l}\left(Y^*, -\Phi(0)\mathbb{E}[Y^*]\right), \\ \sup_{Y^* \in T_{X^*}} \alpha_\rho^{-l}\left(Y^*, s - \mathbb{E}\left[Y^* \Phi\left(\frac{X^*}{Y^*}\right) 1_{\{Y^* > 0\}}\right]\right) \Bigg\},$$

where T_{X^*} is defined by (5.1.2)*. In particular, when we change variables to the probabilistic settings, we get*

$$\alpha^{-l}_{\rho\circ\Lambda}\left(w\cdot\frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}},s\right)=\max\left\{\sup_{\mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P})}\alpha^{-l}_{\rho}\left(\frac{d\mathbb{Q}}{d\mathbb{P}},-\Phi(0)\right),\right.$$
$$\left.\sup_{\substack{\mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P}),\lambda>0\\ w_i\mathbb{S}_i\ll\mathbb{Q}}}\alpha^{-l}_{\rho}\left(\frac{d\mathbb{Q}}{d\mathbb{P}},\frac{s}{\lambda}-\mathbb{E}_{\mathbb{Q}}\left[\Phi\left(\frac{w\cdot\mathbf{d}\mathbb{S}}{\lambda d\mathbb{Q}}\right)\right]\right)\right\}.$$

(ii) Suppose that $\tilde{\Lambda}$ is unbounded from above, that is, $\Phi(0)=+\infty$. Then, we have

$$\alpha^{-l}_{\rho\circ\Lambda}(X^*,s)=\sup_{Y^*\in T_{X^*}}\alpha^{-l}_{\rho}\left(Y^*,s-\mathbb{E}\left[Y^*\Phi\left(\frac{X^*}{Y^*}\right)1_{\{Y^*>0\}}\right]\right),$$

and

$$\alpha^{-l}_{\rho\circ\Lambda}\left(w\cdot\frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}},s\right)=\sup_{\substack{\mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P}),\lambda>0\\ w_i\mathbb{S}_i\ll\mathbb{Q}}}\alpha^{-l}_{\rho}\left(\frac{d\mathbb{Q}}{d\mathbb{P}},\frac{s}{\lambda}-\mathbb{E}_{\mathbb{Q}}\left[\Phi\left(\frac{w\cdot\mathbf{d}\mathbb{S}}{\lambda d\mathbb{Q}}\right)\right]\right).$$

Proof. The proof is given in Section 6.2. □

In the next proposition, we give a dual representation for quasiconvex systemic risk measures. Unlike the previous two propositions, we allow for $p=+\infty$ here as we do not rely on the expression for the penalty function.

Proposition 5.1.8. *Let $\tilde{\Lambda}\colon\mathbb{R}^n\to\mathbb{R}$ be a concave, increasing function that is either bounded from above or linear. Define Λ by* (5.1.1)*. Suppose that $\Lambda(X)\in L^p(\mathbb{R})$ for every $X\in L^p(\mathbb{R}^n)$, where $p\in[1,+\infty]$. Let ρ be a lower semicontinuous quasiconvex risk measure. Then, we have*

$$R(X)=\rho\circ\Lambda(X)=\sup_{\substack{w\in\mathbb{R}^n_+\setminus\{0\},\mathbb{S}\in\mathcal{M}_n^q(\mathbb{P})\\ \mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P}),w_i\mathbb{S}_i\ll\mathbb{Q}}}\alpha^{-l}_{\rho}\left(\frac{d\mathbb{Q}}{d\mathbb{P}},-\mathbb{E}_{\mathbb{Q}}\left[\Phi\left(\frac{w\cdot\mathbf{d}\mathbb{S}}{d\mathbb{Q}}\right)\right]-w^{\mathsf{T}}\mathbb{E}_{\mathbb{S}}\left[X\right]\right)\tag{5.1.5}$$

for every $X\in L^p(\mathbb{R}^n)$.

Proof. The proof is given in Section 6.2. □

Now, we are ready to give the implication of Theorem 4.1.6 and Theorem 4.1.2 by combining Examples 5.1.2 and 5.1.3 with Example 5.1.5. Therefore, we will give examples for the dual representation of the composition for a quasiconvex function and a concave function which is new in the literature to the best of our knowledge.

Example 5.1.9. (Total profit-loss model with economic index of riskiness) Take $\tilde{\Lambda}(x) = \sum_{i=1}^{n} x_i$ and $p \in [1, +\infty)$. By (5.1.4), we have the formula

$$\alpha_{\rho\circ\Lambda}\left(w \cdot \frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}}, m\right) = \inf_{\substack{\lambda>0, \mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P}),\\ w_i\mathbb{S}_i \ll \mathbb{Q}}} \left(\lambda\alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, m\right) + \mathbb{E}_\mathbb{Q}\left[\lambda\Phi\left(\frac{w\cdot\mathbf{d}\mathbb{S}}{\lambda d\mathbb{Q}}\right)\right]\right).$$

From the calculation in Example 5.1.5(i), we can see that it is enough to consider only the case where $\frac{w\cdot\mathbf{d}\mathbb{S}}{\lambda d\mathbb{P}} = 1$ almost surely. Therefore, we get

$$\alpha_{\rho\circ\Lambda}\left(w \cdot \frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}}, m\right) = \begin{cases} \lambda\alpha_\rho\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right) & \text{if } w\cdot\mathbf{d}\mathbb{S} = \lambda d\mathbb{Q}\mathbf{1} \text{ for some } \mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P}), \lambda>0, \\ \infty & \text{else.} \end{cases}$$

In order to give a more specific example, let us take ρ as the economic index of riskiness in Example 5.1.3 corresponding to the logarithmic loss function with $p = 1$. In this case, we obtain

$$\alpha_{\rho\circ\Lambda}\left(w \cdot \frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}}, m\right) = m\lambda\left(1 - \exp\left(\mathbb{E}\left[\ln\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] - c_0\right)\right)$$

if $w\cdot\mathbf{d}\mathbb{S} = \lambda d\mathbb{Q}\mathbf{1}$ for some $\mathbb{Q}\in\mathcal{M}_1^\infty(\mathbb{P})$ and $\lambda > 0$, and $\alpha_{\rho\circ\Lambda}(w\cdot\frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}}, m) = +\infty$ otherwise.

Example 5.1.10. (i) Let $\tilde{\Lambda}(x) = \sum_{i=1}^{n} x_i$ be the aggregation function in Example 5.1.5(i) and $p \in [1, +\infty]$. Then, by Proposition 5.1.8 and Example 5.1.5,

$$\rho\circ\Lambda(X) = \sup_{\mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P})} \alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\sum_{i=1}^{n}\mathbb{E}_\mathbb{Q}[X_i]\right).$$

In particular, if we take ρ as the certainty equivalent corresponding to the power loss function (Example 5.1.2(iii)) and $p = 1$, then by Proposition 5.1.8 and Example 5.1.2, we get

$$\rho\circ\Lambda(X) = \sup_{\mathbb{Q}\in\mathcal{M}_1^\infty(\mathbb{P})} -\left\|\frac{d\mathbb{Q}}{d\mathbb{P}}\right\|_{\frac{\gamma-1}{\gamma}} \sum_{i=1}^{n}\mathbb{E}_\mathbb{Q}[X_i].$$

(ii) Let us take the total loss model in Example 5.1.5 and $p \in [1, +\infty]$. Then we have the following dual representation by Proposition 5.1.8

$$R(X) = \rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}^n_+ \setminus \{0\}, \mathbb{S} \in \mathcal{M}^q_n(\mathbb{P}) \\ \frac{w_i d\mathbb{S}_i}{d\mathbb{P}} \leq 1, \\ \mathbb{Q} \in \mathcal{M}^q_1(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \alpha_\rho^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -w^\mathsf{T} \mathbb{E}_\mathbb{S} [X] \right). \tag{5.1.6}$$

As a more specific example, take $p = 2$ and consider the quadratic loss function in Example 5.1.2(i), which gives that

$$R(X) = \rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}^n_+ \setminus \{0\}, \mathbb{S} \in \mathcal{M}^2_n(\mathbb{P}) \\ \frac{w_i d\mathbb{S}_i}{d\mathbb{P}} \leq 1, w^\mathsf{T} \mathbb{E}_\mathbb{S}[X] < 1 \\ \mathbb{Q} \in \mathcal{M}^2_1(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \left(\frac{-w^\mathsf{T} \mathbb{E}_\mathbb{S} [X] + 1}{\left\| \frac{d\mathbb{Q}}{d\mathbb{P}} \right\|_2} - 1 \right) \tag{5.1.7}$$

(iii) Let us suppose that ρ is the certainty equivalent corresponding to the logarithmic loss function in Example 5.1.2(ii) with $p = +\infty$. Then, by Proposition 5.1.8 and Example 5.1.2, we have

$$\rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}^n_+ \setminus \{0\}, \mathbb{S} \in \mathcal{M}^1_n(\mathbb{P}) \\ \mathbb{Q} \in \mathcal{M}^1_1(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} - \frac{\mathbb{E}_\mathbb{Q} \left[\Phi \left(\frac{w \cdot \mathbf{d}\mathbb{S}}{d\mathbb{Q}} \right) \right] + w^\mathsf{T} \mathbb{E}_\mathbb{S} [X]}{e^{\mathbb{E}\left[\ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right]}}. \tag{5.1.8}$$

In particular, let us assume that $\tilde{\Lambda}$ is the exponential aggregation function in Example 5.1.3(iii). Then, (5.1.8) simplifies as

$$\rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}^n_+ \setminus \{0\}, \mathbb{S} \in \mathcal{M}^1_n(\mathbb{P}) \\ \mathbb{Q} \in \mathcal{M}^1_1(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} - \frac{\sum_{i=1}^n \left(\mathbb{E}_\mathbb{Q} \left[\frac{w_i d\mathbb{S}_i}{d\mathbb{Q}} \ln \left(\frac{w_i d\mathbb{S}_i}{d\mathbb{Q}} \right) \right] + w_i \mathbb{E}_{\mathbb{S}_i} [X_i] \right)}{e^{\mathbb{E}\left[\ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right]}}.$$

By looking at the Equation (5.1.8), we can make some interpretations. We are calculating the effect of X on the financial institutions by $\mathbb{E}_\mathbb{S}[X]$ and arrange their importance level by changing the weights w. We calculate some concept of divergence of the system model $\mathbb{S}$ to the alternative model of society $\mathbb{Q}$ by $\mathbb{E}_\mathbb{Q}[\Phi(\frac{w \cdot \mathbf{d}\mathbb{S}}{d\mathbb{Q}})]$. Hence if we also consider the minus sign, in the numerator we are calculating the risk of choosing the weights w and the condition of financial institutions $\mathbb{S}$ when we are in the alternative model $\mathbb{Q}$ for the society. Then, we should look for to what extend this calculation under this alternative model has impact on us. Therefore, if an alternative model is more realistic which means if it is closer to the actual probability measure $\mathbb{P}$, then it should affect us more. The denominator $e^{\mathbb{E}[\ln(\frac{d\mathbb{Q}}{d\mathbb{P}})]}$ calculates the distance of alternative model to our real probability measure $\mathbb{P}$ so the denominator handles the adjustment issues. Briefly, if an alternative model $\mathbb{Q}$ is more realistic than the denominator will be a small number, therefore the calculation under this model in the numerator will have more impact on our risk measure. By taking the supremum over all models, weights and society measures we are looking for the worst case in order to calculate our systemic risk measure. It should also be noted that our interpretation is valid for the other risk functions in Examples 5.1.2 and 5.1.3 since we can look from the same perspective, the denominator is always a method for calculating the distance between the alternative model $\mathbb{Q}$ and the real model $\mathbb{P}$ and the numerator is somehow risk under the this alternative model.

5.2 Eisenberg-Noe model

In the real-world applications, our aggregation function might not be defined on the whole space but rather a smaller subset. In this section, we will discuss the Eisenberg-Noe clearing model for which the aggregation function is of the form $\Lambda\colon L^p(\mathbb{R}^n_+) \to L^p(\mathbb{R})$ induced by a deterministic function $\tilde{\Lambda}\colon \mathbb{R}^n_+ \to \mathbb{R}$ via (5.1.1). Before describing this model in detail, as a preparation, we first prove a slightly different version of Proposition 5.1.6 using the cone $L^p(\mathbb{R}^n_+)$. In this section, we will define Φ via (5.1.3), where

$$\tilde{\Phi}(x^*) = \sup_{x\in\mathbb{R}^n_+} (\Lambda(x) - (x^*)^\mathsf{T} x), \quad x^* \in \mathbb{R}^n.$$

Proposition 5.2.1. *Let $\rho : L^p(\mathbb{R}) \to \overline{\mathbb{R}}$ be a lower semicontinuous, quasiconvex risk measure; and $\tilde{\Lambda}\colon \mathbb{R}^n_+ \to \mathbb{R}$ a concave, regularly increasing function that is bounded from above. Moreover, suppose that the strict sublevel set $\left\{X \in L^p(\mathbb{R}^n_+) \mid \mathbb{E}\left[-Y^*\Lambda(X)\right] < m\right\}$ is nonempty for every $Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}$. Then, for every $X^* \in L^q(\mathbb{R}^n_+)$ and $m \in \mathbb{R}$,*

$$\alpha_{\rho\circ\Lambda}(X^*, m) = 0 \wedge \inf_{Y^*\in L^q(\mathbb{R}_{++})} \left(\alpha_\rho\left(Y^*, m\right) + \mathbb{E}\left[Y^*\Phi\left(\frac{X^*}{Y^*}\right)\right]\right)$$

Proof. The proof is given in Section 6.2. □

Proposition 5.2.2. *Let $\rho : L^p(\mathbb{R}) \to \overline{\mathbb{R}}$ be a lower semicontinuous, quasiconvex risk measure, and $\tilde{\Lambda}\colon \mathbb{R}^n \to \mathbb{R}$ a concave, increasing function that is either bounded from above or linear.*

Suppose that $\Lambda(X) \in L^p(\mathbb{R})$ for every $X \in L^p(\mathbb{R}^n)$. Then for every $X \in L^p_+(\mathbb{R}^n)$, we have

$$\rho \circ \Lambda(X) = \sup_{X^* \in L^q(\mathbb{R}^n_+) \setminus \{0\}} \sup_{Y^* \in L^q(\mathbb{R}_{++})} \alpha_\rho^{-l} \left(Y^*, -\mathbb{E}\left[(X^*)^\mathsf{T} X + Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right]\right).$$

Proof. The proof is given in Section 6.2. □

As in Section 5.1, we may switch to probability measures by writing $X^* = w \cdot \frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}}$ and $Y^* = \lambda \frac{d\mathbb{Q}}{d\mathbb{P}}$, where $w \in \mathbb{R}^n_+ \setminus \{0\}$, $\lambda > 0$, $\mathbb{Q} \in \mathcal{M}^q_1(\mathbb{P})$, and $\mathbb{S} \in \mathcal{M}^q_n(\mathbb{P})$. Again, by [20, Lemma 6.3], we have $w_i \mathbb{S}_i \ll \mathbb{Q}$ if $Y^* \in L^q(\mathbb{R}_{++}, \mathbb{P})$. By using the same arguments in the proof of Proposition 5.1.7, we will have

$$\rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}^n_+ \setminus \{0\}, \mathbb{S} \in \mathcal{M}^q_n(\mathbb{P}) \\ \mathbb{Q} \in \mathcal{M}^q_1(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \alpha_\rho^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\mathbb{E}_\mathbb{Q}\left[\Phi\left(\frac{w \cdot \mathbf{d}\mathbb{S}}{d\mathbb{Q}}\right)\right] - w^\mathsf{T} \mathbb{E}_\mathbb{S}[X]\right). \tag{5.2.1}$$

Next, we review the clearing model in [21], which takes into account the liabilities between the members of the financial network, hence the structure of the network. In this model, financial institutions are considered as nodes and their liabilities are considered as arcs in a graph. More precisely, let $\mathcal{N} = \{0, 1, \dots, n\}$ denote the nodes, where nodes $1, \dots, n$ typically represent the banks and special node 0 represents society. For each $i, j \in \mathcal{N}$, let $\ell_{ij} \geq 0$ denote the nominal liability of member i to member j. Naturally, we assume no self-liabilities, that is, $\ell_{ii} = 0$ for each $i \in \mathcal{N}$; and society has no liabilities to banks, that is, $\ell_{0i} = 0$ for every $i \in \mathcal{N}$. We also assume that every bank has nonzero liability to society, that is, $\ell_{i0} > 0$ for every $i \in \mathcal{N} \setminus \{0\}$. Then, the relative liability of member i to member j is defined by

$$a_{ij} \coloneqq \frac{\ell_{ij}}{\bar{p}_i},$$

where $\bar{p}_i \coloneqq \sum_{j=0}^n \ell_{ij}$ is the total liability of member i. Finally, let $x \in \mathbb{R}^n_+$ denote a possible realization of the uncertain value of the assets of the banks. A clearing payment vector $p(x) \in \mathbb{R}^n$ is defined as a solution of the following fixed point problem:

$$p_i(x) = \min\left\{\bar{p}_i, \sum_{j=1}^n a_{ji} p_j(x)\right\} \text{ for } i \in \mathcal{N} \setminus \{0\}.$$

Clearly, every clearing payment vector $p = p(x)$ is a feasible solution for the following linear programming problem.

$$\begin{aligned}
\max \quad & \textstyle\sum_{i=1}^{n} a_{i0} p_i \\
\text{s.t.} \quad & p_i \leq x_i + \textstyle\sum_{j=1}^{n} a_{ji} p_j \quad \text{for } i = 1, \ldots, n, \\
& p_i \in [0, \bar{p}_i] \quad \text{for } i = 1, \ldots, n.
\end{aligned} \tag{5.2.2}$$

It is shown in Lemma 4 of [21] that every optimal solution of this problem is a clearing payment vector for our system. In addition to this, it is shown in [21] that for any $x \in \mathbb{R}^n_+$ the above linear problem is feasible, and hence it has an optimal solution. Denote the optimal value of this problem by $\tilde{\Lambda}(x)$. It should be noted that $\tilde{\Lambda}(x) \in \mathbb{R}_+$ since $a_{i0} > 0$ by definition and $p_i \in [0, \bar{p}_i]$. $\tilde{\Lambda}$ calculates the effect of the realized values of the assets on society. Therefore, $\tilde{\Lambda}$ can be considered as an aggregation function. Let us take $D = L^p(\mathbb{R}_+)$ and $D^+ = L^q(\mathbb{R}_+)$, then $\tilde{\Lambda}$ is concave and increasing as it is stated in Sect. 4.4 of [20]. It is bounded by $\sum_{i=1}^{n} a_{i0}\bar{p}_i$ so we can use Lemma 5.1.4.

Let us calculate the corresponding conjugate function: for every $x^* \in \mathbb{R}^n_+$, using (5.2.2), we have

$$\begin{aligned}
\tilde{\Phi}(x^*) &= \sup_{x \in \mathbb{R}^n_+} \left(-x^{\mathsf{T}} x^* + \tilde{\Lambda}(x) \right) \\
&= \sup_{0 \leq p \leq \bar{p}} \left(\sum_{i=1}^{n} a_{i0} p_i - \inf_{\substack{x \geq 0 \\ x \geq p - A^{\mathsf{T}} p}} \sum_{i=1}^{n} x_i^* x_i \right) \\
&= \sup_{0 \leq p \leq \bar{p}} \sum_{i=1}^{n} \left(a_{i0} p_i - (x_i^*) \left(p_i - \sum_{j=1}^{n} a_{ji} p_j \right)^+ \right).
\end{aligned}$$

Then, by applying Proposition 5.2.2, we have

$$\rho \circ \Lambda(X) = \sup_{\substack{X^* \in L^q(\mathbb{R}^n_+) \setminus \{0\} \\ Y^* \in L^q(\mathbb{R}_{++})}} \alpha_\rho^{-l} \left(Y^*, -\mathbb{E}\left[X^{\mathsf{T}} X^* + Y^* \Phi\left(\frac{X^*}{Y^*} \right) \right] \right).$$

We can pass to the probabilistic setting by using (5.2.1) as follows:

$$\rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}^n_+ \setminus \{0\}, \mathbb{S} \in \mathcal{M}^q_n(\mathbb{P}) \\ \mathbb{Q} \in \mathcal{M}^q_1(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \alpha_\rho^{-l} \left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\mathbb{E}_{\mathbb{Q}}\left[\Phi\left(w \cdot \frac{d\mathbb{S}}{d\mathbb{Q}} \right) \right] - w^{\mathsf{T}} \mathbb{E}_{\mathbb{S}}[X] \right).$$

As a more specific example, if we use the certainty equivalent that corresponds to the logarithmic loss function (see Example 5.1.2(ii)) for the case $p = 1$, then the dual representation simplifies as

$$\rho \circ \Lambda(X) = \sup_{\substack{w \in \mathbb{R}^n_+ \setminus \{0\}, \mathbb{S} \in \mathcal{M}^\infty_n(\mathbb{P}) \\ \mathbb{Q} \in \mathcal{M}^\infty_1(\mathbb{P}), w_i \mathbb{S}_i \ll \mathbb{Q}}} \frac{-\mathbb{E}_{\mathbb{Q}}\left[\Phi\left(w \cdot \frac{d\mathbb{S}}{d\mathbb{Q}}\right)\right] - w^{\mathsf{T}} \mathbb{E}_{\mathbb{S}}[X]}{e^{\mathbb{E}\left[\ln\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]}}. \tag{5.2.3}$$

In order to interpret (5.2.3), we first look at the explicit expression for Φ with the help of Theorem 14.60 in [32]:

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left[\Phi\left(w \cdot \frac{d\mathbb{S}}{d\mathbb{Q}}\right)\right] &= \mathbb{E}_{\mathbb{Q}}\left[\sup_{0 \le p \le \bar{p}} \sum_{i=1}^n \left(a_{i0} p_i - w_i \frac{d\mathbb{S}_i}{d\mathbb{Q}} \Big(p_i - \sum_{j=1}^n a_{ji} p_j\Big)^+\right)\right] \\
&= \sup_{P \in L^1([0,\bar{p}])} \sum_{i=1}^n \mathbb{E}_{\mathbb{Q}}\left[a_{i0} P_i - w_i \frac{d\mathbb{S}_i}{d\mathbb{Q}} \Big(P_i - \sum_{j=1}^n a_{ji} P_j\Big)^+\right].
\end{aligned}$$

Consider bank $i \in \{1, \ldots, n\}$. The term $a_{i0} P_i$ represents the gain of society that comes from bank i after clearing, and $(P_i - \sum_{j=1}^n a_{ji} P_j)^+$ corresponds to the net gain of the bank i, which is then multiplied by the weight $w_i \frac{d\mathbb{S}_i}{d\mathbb{Q}}$ of bank i relative to society. Thus, we obtain the "relative net gain" of society by calculating the difference between the gain of society and weighted net gain of banks after the clearing. Observe that if $w_i \frac{d\mathbb{S}_i}{d\mathbb{Q}}$ is small, then the net gain of bank i will have less negative impact on the relative net gain of society. We are calculating the supremum over the clearing vectors so we are trying to maximize the relative net gain of the society.

To interpret the dual representation in (5.2.3), we are calculating the effect of a random shock on the system with weights by $-w^{\mathsf{T}} \mathbb{E}_{\mathbb{S}}[X]$ and sum it with the "relative net loss" of society by $-\mathbb{E}_{\mathbb{Q}}[\Phi(w \cdot \frac{d\mathbb{S}}{d\mathbb{Q}})]$. In the denominator, as in the previous case (5.1.8), we are calculating the distance between the probability measure $\mathbb{Q}$ of the society and the actual probability measure $\mathbb{P}$. When the incompatibility between the society and system gets bigger which means the distance between the probability measure $\mathbb{S}$ of the system and the probability measure $\mathbb{Q}$ of the society increases, then the relative net loss of society will increase too. In the denominator, we are looking at the plausibility of the probability measure of the society. If the probability measure $\mathbb{Q}$ of the society is more realistic which means if it is closer to the real probability measure $\mathbb{P}$, the denominator will be small so the effect of the numerator will be much more in more realistic cases. Briefly, if the incompatibility between the system and society is high then relative loss will be higher in this case, and the model for the society is more realistic this will be considered more seriously in our dual representation and we will calculate the worst-case by taking the supremum.

Chapter 6

Appendix

In this chapter, we will give some definitions and propositions which will lead us to the proof our important results in Chapter 4 and Chapter 6.

6.1 Proof of some results in Chapter 4

The main purpose of this section is to prove Theorem 4.1.6. As a preparation for the proof, we will establish a sequence of technical results. In particular, these results will ensure that we may apply the minimax inequality in [25].

We work in the setting of Chapter 4: we consider two functions $f\colon \mathcal{Y} \to \overline{\mathbb{R}}$ and $g\colon \mathcal{X} \to \mathcal{Y}$. We also suppose that Assumption 4.1.5 holds, that is, $\bar{D}^+$ is a convex and compact cone generator for D^+. Given $m \in \mathbb{R}$ and $y^* \in D^+$, let us define the sets

$$A_{y^*}^m \coloneqq \{x \in \mathcal{X} \mid \langle y^*, -g(x)\rangle \leq \alpha_f(y^*, m)\},$$
$$\tilde{A}_{y^*}^m \coloneqq \{x \in \mathcal{X} \mid \langle y^*, -g(x)\rangle < \alpha_f(y^*, m)\}.$$

Clearly, $\tilde{A}_{y^*}^m \subseteq A_{y^*}^m$. Also, observe that $A_{y^*}^m$ is actually the sublevel set of $-h_{y^*}^g$; see (3.0.1). Therefore, when the function g is D-naturally quasiconcave, increasing and D-lower demicontinuous, the set $A_{y^*}^m$ is a closed, convex and monotone set by Propositions 3.0.2, 3.0.7, 3.0.12. We give the precise relationship between the sets $\tilde{A}_{y^*}^m$ and $A_{y^*}^m$ in the following proposition.

Proposition 6.1.1. *Suppose that Assumption 3.0.9 holds. In addition, suppose that* $g\colon \mathcal{X} \to \mathcal{Y}$ *is D-naturally quasiconcave, regularly increasing and D-lower demicontinuous; and let* $m \in \mathbb{R}$, $y^* \in D^+ \setminus \{0\}$. *Then,*

$$A^m_{y^*} = \operatorname{cl} \tilde{A}^m_{y^*} = \operatorname{cl}\operatorname{conv} \tilde{A}^m_{y^*}. \tag{6.1.1}$$

Proof. If $A^m_{y^*} = \emptyset$, then the result is obvious. Let us assume that $A^m_{y^*} \neq \emptyset$ and prove that $A^m_{y^*}$ is the closure of $\tilde{A}^m_{y^*}$. Since $\tilde{A}^m_{y^*} \subseteq A^m_{y^*}$ and $A^m_{y^*}$ is closed, we have $\operatorname{cl} \tilde{A}^m_{y^*} \subseteq A^m_{y^*}$.

Now let us take $x \in A^m_{y^*}$, and fix some $c \in C^\#$ and $\lambda > 0$. It is clear that $\lambda c \in C^\#$ since $C^\#$ is a cone. Moreover, since g is regularly increasing, we have $g(x + \lambda c) - g(x) \in D^\#$. In particular, since $y^* \in D^+ \setminus \{0\}$, we have $\langle y^*, g(x+\lambda c) - g(x)\rangle > 0$. Therefore,

$$\begin{aligned}\langle y^*, -g(x+\lambda c)\rangle &= \langle y^*, -g(x)\rangle - \langle y^*, g(x+\lambda c) - g(x)\rangle \\ &\leq \alpha_f(y^*, m) - \langle y^*, g(x+\lambda c) - g(x)\rangle < \alpha_f(y^*, m).\end{aligned}$$

Hence, $x + \lambda c \in \tilde{A}^m_{y^*}$. The net $(x+\lambda c)_{\lambda > 0} \subseteq \tilde{A}^m_{y^*}$ converges to x as $\lambda \to 0$, which implies that $x \in \operatorname{cl} \tilde{A}^m_{y^*}$. Hence, $A^m_{y^*} \subseteq \operatorname{cl} \tilde{A}^m_{y^*}$ and the first equality in (6.1.1).

Finally, since $A^m_{y^*}$ is convex, we have

$$A^m_{y^*} = \operatorname{conv}\left(\operatorname{cl} \tilde{A}^m_{y^*}\right) \subseteq \operatorname{cl}\left(\operatorname{conv} \tilde{A}^m_{y^*}\right) \subseteq A^m_{y^*}.$$

Hence, the second equality in (6.1.1) follows as well. □

Remark 6.1.2. Let $m \in \mathbb{R}$, $y^* \in D^+ \setminus \{0\}$. Hence, we may write $y^* = \lambda \bar{y}^*$ for some $\lambda > 0$ and $\bar{y}^* \in \bar{D}^+$. Then, it is easy to see that

$$x \in A^m_{y^*} \quad \Leftrightarrow \quad x \in A^m_{\bar{y}^*}$$

for each $x \in \mathcal{X}$. Hence, $A^m_{y^*} = A^m_{\bar{y}^*}$.

Next, for each $m \in \mathbb{R}$ and $x^* \in C^+$, we define two auxiliary functions $K^m_{x^*}, \tilde{K}^m_{x^*}\colon \mathcal{X} \times \bar{D}^+ \to \overline{\mathbb{R}}$ by

$$K^m_{x^*}(x, y^*) = \langle x^*, -x\rangle - I_{A^m_{y^*}}(x), \qquad \tilde{K}^m_{x^*}(x, y^*) = \langle x^*, -x\rangle - I_{\tilde{A}^m_{y^*}}(x),$$

for each $(x, y^*) \in \mathcal{X} \times \bar{D}^+$. The next proposition shows the relation between these two functions.

Proposition 6.1.3. *Let $m \in \mathbb{R}$ and $x^* \in C^+$. Suppose that g is D-naturally quasiconcave, regularly increasing and D-lower demicontinuous. Then, for each $y^* \in \bar{D}^+$, we have*

$$\sup_{x \in \mathcal{X}} \tilde{K}^m_{x^*}(x, y^*) = \sup_{x \in \mathcal{X}} K^m_{x^*}(x, y^*).$$

Proof. Let $y^* \in \bar{D}^+$. By definition, we have

$$\sup_{x \in \mathcal{X}} \tilde{K}^m_{x^*}(x, y^*) = \sup_{x \in \mathcal{X}} \big(\langle -x^*, x \rangle - I_{\tilde{A}^m_{y^*}}(x) \big) = I^*_{\tilde{A}^m_{y^*}}(-x^*). \tag{6.1.2}$$

Moreover, by Lemma 2.3.1 and Proposition 6.1.1, we have

$$I^*_{\tilde{A}^m_{y^*}}(-x^*) = \sup_{x \in A^m_{y^*}} \langle -x^*, x \rangle .$$

Similar to (6.1.2), we also have

$$\sup_{x \in \mathcal{X}} K^m_{x^*}(x, y^*) = \sup_{x \in A^m_{y^*}} \langle -x^*, x \rangle .$$

Combining these, we obtain the desired result. □

We will use a minimax theorem in the proof of Theorem 4.1.6 so we need to show the properties of these functions and the connection of these functions to our problem. Lets start with the properties of these functions.

Proposition 6.1.4. *Let $m \in \mathbb{R}$ and $x^* \in C^+$. Suppose that g is D-naturally quasiconcave. The following properties hold.*

(i) Suppose further that g is D-lower demicontinuous. Then, $K^m_{x^*}$ is concave and upper semicontinuous in its first argument, and quasiconvex in its second argument.

(ii) The function $\tilde{K}^m_{x^*}$ is concave in its first argument, and quasiconvex and lower semicontinuous in its second argument.

Proof. We prove (i) first. Let $y^* \in \bar{D}^+$. Since $A^m_{y^*}$ is a closed convex set, $I_{A^m_{y^*}}$ is a lower semicontinuous convex function. Hence, $x \mapsto K^m_{x^*}(x, y^*)$ is an upper semicontinuous concave function.

Next, let us fix $x \in \mathcal{X}$. We claim that $y^* \mapsto I_{A^m_{y^*}}(x)$ is a quasiconvex function. Indeed, let $y_1^*, y_2^* \in \bar{D}^+$ and $\lambda \in [0,1]$. Since $\bar{D}^+$ is convex, $\lambda y_1^* + (1-\lambda)y_2^* \in D^+_{cg}$. If $x \in A^m_{y_1^*}$ or $x \in A^m_{y_2^*}$, then we have

$$\min\left\{I_{A^m_{y_1^*}}(x), I_{A^m_{y_2^*}}(x)\right\} = 0 \leq I_{A^m_{\lambda y_1^* + (1-\lambda)y_2^*}}(x)$$

by the definiton of indicator function. On the other hand, suppose that $x \notin A^m_{y_1^*}$ and $x \notin A^m_{y_2^*}$. Then, $\langle y_1^*, -g(x)\rangle > \alpha_f(y_1^*, m)$ and $\langle y_2^*, -g(x)\rangle > \alpha_f(y_2^*, m)$. Hence,

$$\begin{aligned}
\langle \lambda y_1^* + (1-\lambda)y_2^*, -g(x)\rangle &> \lambda\alpha_f(y_1^*, m) + (1-\lambda)\alpha_f(y_2^*, m) \\
&= \lambda \sup_{y \in S^f_m} \langle y_1^*, -y\rangle + (1-\lambda)\sup_{y \in S^f_m}\langle y_2^*, -y\rangle \\
&\geq \sup_{y \in S^f_m} \langle \lambda y_1^* + (1-\lambda)y_2^*, -y\rangle = \alpha_f(\lambda y_1^* + (1-\lambda)y_2^*, m).
\end{aligned}$$

Therefore, $x \notin A^m_{\lambda y_1^* + (1-\lambda)y_2^*}$ so that

$$\min\left\{I_{A^m_{y_1^*}}(x), I_{A^m_{y_2^*}}(x)\right\} \leq +\infty = I_{A^m_{\lambda y_1^* + (1-\lambda)y_2^*}}(x).$$

It follows that $y^* \mapsto I_{A^m_{y^*}}(x)$ is quasiconvex, hence so is $y^* \mapsto K^m_{x^*}(x, y^*)$.

Next, we prove (ii). Let $y^* \in \bar{D}^+$. We claim that $\tilde{A}^m_{y^*}$ is a convex set. Indeed, let $x_1, x_2 \in \tilde{A}^m_{y^*}$ and $\lambda \in [0,1]$. Since $-h^g_{y^*}$ is quasiconvex, we have

$$-h^g_{y^*}(\lambda x_1 + (1-\lambda)x_2) \leq \max\left\{-h^g_{y^*}(x_1), -h^g_{y^*}(x_2)\right\} < \alpha_f(y^*, m),$$

which implies that $\lambda x_1 + (1-\lambda)x_2 \in \tilde{A}^m_{y^*}$. Hence, the claim follows. It follows that $I_{\tilde{A}^m_{y^*}}$ is a convex function and $x \mapsto \tilde{K}^m_{x^*}(x, y^*)$ is a concave function.

Let us fix $x \in \mathcal{X}$. We show that $y^* \mapsto I_{\tilde{A}^m_{y^*}}(x)$ is quasiconvex. Let $y_1^*, y_2^* \in \bar{D}^+$ and $\lambda \in [0,1]$. Since $\bar{D}^+$ is convex $\lambda y_1^* + (1-\lambda)y_2^* \in \bar{D}^+$. If $x \in \tilde{A}^m_{y_1^*}$ or $x \in \tilde{A}^m_{y_2^*}$, then we have

$$\min\left\{I_{\tilde{A}^m_{y_1^*}}(x), I_{\tilde{A}^m_{y_2^*}}(x)\right\} = 0 \leq I_{\tilde{A}^m_{\lambda y_1^* + (1-\lambda)y_2^*}}(x).$$

Suppose that $x \notin \tilde{A}^m_{y_1^*}$ and $x \notin \tilde{A}^m_{y_2^*}$. Hence, $\langle y_1^*, -g(x)\rangle \geq \alpha_f(y_1^*, m)$, $\langle y_2^*, -g(x)\rangle \geq \alpha_f(y_2^*, m)$, and

$$\begin{aligned}\langle \lambda y_1^* + (1-\lambda)y_2^*, -g(x)\rangle &\geq \lambda\alpha_f(y_1^*, m) + (1-\lambda)\alpha_f(y_2^*, m)\\ &= \lambda \sup_{y\in S_m^f} \langle y_1^*, -y\rangle + (1-\lambda) \sup_{y\in S_m^f} \langle y_2^*, -y\rangle\\ &\geq \sup_{y\in S_m^f} \langle \lambda y_1^* + (1-\lambda)y_2^*, -y\rangle = \alpha_f(\lambda y_1^* + (1-\lambda)y_2^*, m),\end{aligned}$$

which implies $x \notin \tilde{A}^m_{\lambda y_1^*+(1-\lambda)y_2^*}$. Hence,

$$\min\left\{I_{\tilde{A}^m_{y_1^*}}(x), I_{\tilde{A}^m_{y_2^*}}(x)\right\} \leq +\infty = I_{\tilde{A}^m_{\lambda y_1^*+(1-\lambda)y_2^*}}(x),$$

which completes the proof of quasiconvexity. It follows that $y^* \mapsto \tilde{K}^m_{x^*}(x, y^*)$ is quasiconvex.

Finally, to prove lower semicontinuity, let us define the set

$$E_x^m = \left\{y^* \in \bar{D}^+ \mid \langle y^*, -g(x)\rangle < \alpha_f(y^*, m)\right\}.$$

Note that

$$E_x^m = \left\{y^* \in \bar{D}^+ \mid \langle y^*, -g(x)\rangle < \sup_{y\in S_m^f} \langle y^*, -y\rangle\right\} = \left\{y^* \in \bar{D}^+ \mid 0 < \sup_{y\in S_m^f} \langle y^*, -y + g(x)\rangle\right\}.$$

Since the supremum of a family of affine functions is lower semicontinuous, it follows that E_x^m is open. On the other hand, for each $y^* \in \bar{D}^+$, it is clear that $y^* \in E_x^m$ if and only if $x \in \tilde{A}^m_{y^*}$, that is, $I_{\tilde{A}^m_{y^*}}(x) = I_{E_x^m}(y^*)$. So we actually have

$$\tilde{K}^m_{x^*}(x, y^*) = \langle x^*, -x\rangle - I_{\tilde{A}^m_{y^*}}(x) = \langle x^*, -x\rangle - I_{E_x^m}(y^*). \tag{6.1.3}$$

Since E_x^m is open, the function $I_{E_x^m}$ is upper semicontinuous. By (6.1.3), $y^* \mapsto \tilde{K}^m_{x^*}(x, y^*)$ is lower semicontinuous. □

Now, we will show the connection of these functions to the our problem.

Proposition 6.1.5. *Suppose that f is decreasing, lower semicontinuous and quasiconvex, and that g is D-naturally quasiconcave and D-lower demicontinuous. Then, for each*

$(x^*, m) \in C^+ \times \mathbb{R}$,

$$\alpha_{f\circ g}(x^*, m) = \sup_{x\in\mathcal{X}} \inf_{y^*\in\bar{D}^+} K^m_{x^*}(x, y^*).$$

Proof. Let $(x^*, m) \in C^+ \times \mathbb{R}$. Since f is decreasing, lower semicontinuous and quasiconvex, by Remarks 2.3.8 and 6.1.2, we have

$$\begin{aligned}
\alpha_{f\circ g}(x^*, m) &= \sup_{x\in S^{f\circ g}_m} \langle x^*, -x\rangle = \sup\{\langle x^*, -x\rangle \mid g(x) \in S^f_m,\ x \in \mathcal{X}\} \\
&= \sup_{x\in\mathcal{X}} \left\{\langle x^*, -x\rangle \mid \forall y^* \in D^+ \setminus \{0\}\colon\ \langle y^*, -g(x)\rangle \le \alpha_f(y^*, m)\right\} \\
&= \sup_{x\in\mathcal{X}} \left\{\langle x^*, -x\rangle \mid \forall y^* \in \bar{D}^+\colon\ \langle y^*, -g(x)\rangle \le \alpha_f(y^*, m)\right\} = \sup_{x\in B^m} \langle x^*, -x\rangle,
\end{aligned}$$

where

$$B^m := \bigcap_{y^*\in\bar{D}^+} A^m_{y^*}.$$

Hence,

$$\sup_{x\in B^m} \langle x^*, -x\rangle = \sup_{x\in\mathcal{X}} (\langle x^*, -x\rangle - I_{B^m}(x)) = \sup_{x\in\mathcal{X}} \inf_{y^*\in\bar{D}^+} (\langle x^*, -x\rangle - I_{A^m_{y^*}}(x)) = \sup_{x\in\mathcal{X}} \inf_{y^*\in\bar{D}^+} K^m_{x^*}(x, y^*).$$

Therefore, the result follows. $\square$

Proposition 6.1.6. *Let $(x^*, m) \in C^+ \times \mathbb{R}$. Then, we have*

$$\inf_{y^*\in D^+\setminus\{0\}} \alpha_{-h^g_{y^*}}(x^*, \alpha_f(y^*, m)) = \inf_{y^*\in\bar{D}^+} \alpha_{-h^g_{y^*}}(x^*, \alpha_f(y^*, m)) = \inf_{y^*\in\bar{D}^+} \sup_{x\in\mathcal{X}} K^m_{x^*}(x, y^*). \quad (6.1.4)$$

Proof. Let $\bar{y}^* \in \bar{D}^+$. Clearly, we have

$$\sup_{x\in\mathcal{X}} K^m_{x^*}(x, \bar{y}^*) = \sup_{x\in\mathcal{X}}(\langle x^*, -x\rangle - I_{A^m_{\bar{y}^*}}(x)) = \sup_{x\in A^m_{\bar{y}^*}} \langle x^*, -x\rangle.$$

Hence,

$$\begin{aligned}
\inf_{\bar{y}^*\in\bar{D}^+\setminus\{0\}} \alpha_{-h^g_{\bar{y}^*}}(x^*, \alpha_f(\bar{y}^*, m)) &= \inf_{\bar{y}^*\in D^+\setminus\{0\}} \sup_{x\in\mathcal{X}} \{\langle x^*, -x\rangle \mid -h^g_{\bar{y}^*}(x) \le \alpha_f(\bar{y}^*, m)\} \\
&= \inf_{\bar{y}^*\in D^+\setminus\{0\}} \sup_{x\in\mathcal{X}} \{\langle x^*, -x\rangle \mid \langle \bar{y}^*, -g(x)\rangle \le \alpha_f(\bar{y}^*, m)\} \\
&= \inf_{\bar{y}^*\in D^+\setminus\{0\}} \sup_{x\in A^m_{\bar{y}^*}} \langle x^*, -x\rangle = \inf_{\bar{y}^*\in D^+\setminus\{0\}} \sup_{x\in\mathcal{X}} K^m_{x^*}(x, \bar{y}^*),
\end{aligned}$$

which completes the proof of the second equality in (6.1.4). On the other hand, given $y^* \in D^+ \setminus \{0\}$, we may write $y^* = \lambda \bar{y}^*$ for some $\lambda > 0$ and $\bar{y}^* \in \bar{D}^+$, and we have

$$\alpha_{-h^g_{y^*}}(x^*, \alpha_f(y^*, m)) = \sup_{x \in A^m_{y^*}} \langle x^*, -x \rangle = \sup_{x \in A^m_{\bar{y}^*}} \langle x^*, -x \rangle = \alpha_{-h^g_{\bar{y}^*}}(x^*, \alpha_f(\bar{y}^*, m))$$

by Remark 6.1.2. Hence, the first equality in (6.1.4) follows. □

From this point on, we work under Assumption 4.1.5, that is, we assume that $\bar{D}^+$ is compact while D^+_π is not necessarily compact. In particular, Proposition 6.1.6 can be applied to both. With the tools developed above, we are ready to prove the main theorem of the paper. For the completeness of this thesis, we give the statements of the minimax inequality in [25] and the well-known minimax equality in [24].

Theorem 6.1.7 (Sion 1958)**.** *Let $\mathcal{U}, \mathcal{V}$ be nonempty convex sets of two topological vector spaces, and consider a function $f\colon \mathcal{U} \times \mathcal{V} \to \mathbb{R}$. Suppose that f is quasiconcave and upper semicontinuous in its first argument, and quasiconvex and lower semicontinuous in its second argument. Moreover, suppose that one of $\mathcal{U}, \mathcal{V}$ is a compact set. Then, we have*

$$\inf_{u \in \mathcal{U}} \sup_{v \in \mathcal{V}} f(u, v) = \sup_{v \in \mathcal{V}} \inf_{u \in \mathcal{U}} f(u, v).$$

Since the lower semicontinuity of the function $K^m_{x^*}$ is missing, it seems that we are not able to use Sion's minimax theorem in our setting. However, the following minimax inequality will be useful in our proof.

Theorem 6.1.8 (Liu 1978)**.** *In the setting of Theorem 6.1.7, consider two functions $f, \tilde{f}\colon \mathcal{U} \times \mathcal{V} \to \overline{\mathbb{R}}$ satisfying the following conditions:*

- *f is upper semicontinuous in its first argument and quasiconvex in its second argument,*
- *$\tilde{f}$ is quasiconcave in its first argument and lower semicontinuous in its second argument,*
- *$\tilde{f}(u, v) \le f(u, v)$ for all $u \in \mathcal{U}$ and $v \in \mathcal{V}$,*
- *$\mathcal{U}$ is compact.*

Then, we have

$$\inf_{u \in \mathcal{U}} \sup_{v \in \mathcal{V}} \tilde{f}(u,v) \leq \sup_{v \in \mathcal{V}} \inf_{u \in \mathcal{U}} f(u,v).$$

With the help of the above minimax inequality, we are ready to complete the proof of Theorem 4.1.6.

Proof of Theorem 4.1.6. Let $x^* \in C^+ \setminus \{0\}$ and $m \in \mathbb{R}$. For each $y^* \in \bar{D}^+$, since $\tilde{A}^m_{y^*} \subseteq A^m_{y^*}$, we have $I_{\tilde{A}^m_{y^*}}(x) \geq I_{A^m_{y^*}}(x)$ and hence

$$\tilde{K}^m_{x^*}(x,y^*) \leq K^m_{x^*}(x,y^*) \tag{6.1.5}$$

for every $x \in \mathcal{X}$. By Proposition 6.1.4, $K^m_{x^*}$ is upper semicontinuous and concave in its first variable, and quasiconvex in its second variable; $\tilde{K}^m_{x^*}$ is concave in its first variable, and quasiconvex and lower semicontinuous in its second variable. These properties, together with (6.1.5), and the convexity and compactness of $\bar{D}^+$, are sufficient to apply the minimax inequality of [25] (see also Theorem 3.1 in [27] and Corollary 11 in [26]) to the functions $K^m_{x^*}, \tilde{K}^m_{x^*}$. Consequently, we obtain

$$\inf_{y^* \in \bar{D}^+} \sup_{x \in \mathcal{X}} \tilde{K}^m_{x^*}(x,y^*) \leq \sup_{x \in \mathcal{X}} \inf_{y^* \in \bar{D}^+} K^m_{x^*}(x,y^*). \tag{6.1.6}$$

By Proposition 6.1.3, we have

$$\sup_{x \in \mathcal{X}} \tilde{K}^m_{x^*}(x,y^*) = \sup_{x \in \mathcal{X}} K^m_{x^*}(x,y^*).$$

Hence, (6.1.6) yields

$$\inf_{y^* \in \bar{D}^+} \sup_{x \in \mathcal{X}} K^m_{x^*}(x,y^*) \leq \sup_{x \in \mathcal{X}} \inf_{y^* \in \bar{D}^+} K^m_{x^*}(x,y^*).$$

However, the reverse inequality already holds by weak duality. Therefore, we get

$$\inf_{y^* \in \bar{D}^+} \sup_{x \in \mathcal{X}} K^m_{x^*}(x,y^*) = \sup_{x \in \mathcal{X}} \inf_{y^* \in \bar{D}^+} K^m_{x^*}(x,y^*).$$

Finally, by Propositions 6.1.5 and 6.1.6, we have

$$\begin{aligned}\alpha_{f\circ g}(x^*,m) &= \sup_{x\in\mathcal{X}}\inf_{y^*\in\bar{D}^+} K^m_{x^*}(x,y^*) = \inf_{y^*\in\bar{D}^+}\sup_{x\in\mathcal{X}} K^m_{x^*}(x,y^*)\\ &= \inf_{y^*\in D^+\setminus\{0\}} \alpha_{-h^g_{y^*}}(x^*,\alpha_f(y^*,m)) = \inf_{y^*\in\bar{D}^+} \alpha_{-h^g_{y^*}}(x^*,\alpha_f(y^*,m)).\end{aligned}$$

Finally, by Remark 4.1.4 and Proposition 6.1.6 applied to D^+_π, we have

$$\alpha_{f\circ g}(x^*,m) = \inf_{y^*_\pi\in D^+_\pi} \alpha_{-h^g_{y^*_\pi}}\left(x^*,\alpha_f\left(y^*_\pi,m\right)\right),$$

which completes the proof. □

Proof of Corollary 4.1.8. Let $x^*\in C^+\setminus\{0\}$ and $s\in\mathbb{R}$. Following the definition of left inverse and using Theorem 4.1.6, we have

$$\begin{aligned}\alpha^{-l}_{f\circ g}(x^*,s) &= \inf\{m\in\mathbb{R} \mid \alpha_{f\circ g}(x^*,m)\geq s\}\\ &= \inf\left\{m\in\mathbb{R} \mid \inf_{y^*\in D^+\setminus\{0\}} \alpha_{-h^g_{y^*}}(x^*,\alpha_f(y^*,m))\geq s\right\}\\ &= \inf\left\{m\in\mathbb{R} \mid \forall y^*\in D^+\setminus\{0\}\colon \alpha_{-h^g_{y^*}}(x^*,\alpha_f(y^*,m))\geq s\right\}.\end{aligned}$$

We claim that the following minimax equality holds:

$$\begin{aligned}&\inf\left\{m\in\mathbb{R} \mid \forall y^*\in D^+\setminus\{0\}\colon \alpha_{-h^g_{y^*}}(x^*,\alpha_f(y^*,m))\geq s\right\}\\ &= \sup_{y^*\in D^+\setminus\{0\}}\inf\left\{m\in\mathbb{R} \mid \alpha_{-h^g_{y^*}}(x^*,\alpha_f(y^*,m))\geq s\right\}.\end{aligned} \tag{6.1.7}$$

The $\geq$ part of this inequality holds as a weak duality property. Next, we show the $\leq$ part. To get a contradiction, suppose that there exists $\bar{m}\in\mathbb{R}$ such that

$$\begin{aligned}&\inf\left\{m\in\mathbb{R} \mid \forall y^*\in D^+\setminus\{0\}\colon \alpha_{-h^g_{y^*}}(x^*,\alpha_f(y^*,m))\geq s\right\}\\ &> \bar{m} > \sup_{y^*\in D^+\setminus\{0\}}\inf\left\{m\in\mathbb{R} \mid \alpha_{-h^g_{y^*}}(x^*,\alpha_f(y^*,m))\geq s\right\}.\end{aligned} \tag{6.1.8}$$

The first inequality in (6.1.8) implies the existence of $\bar{y}^*\in D^+\setminus\{0\}$ satisfying

$$\alpha_{-h^g_{\bar{y}^*}}(x^*,\alpha_f(\bar{y}^*,\bar{m})) < s. \tag{6.1.9}$$

On the other hand, the second inequality in (6.1.8) implies that

$$\bar{m} > \inf\left\{m \in \mathbb{R} \mid \alpha_{-h^g_{\bar{y}^*}}(x^*, \alpha_f(\bar{y}^*, m)) \geq s\right\}.$$

Hence, there exists $m_{\bar{y}^*} < \bar{m}$ such that

$$\alpha_{-h^g_{\bar{y}^*}}(x^*, \alpha_f(\bar{y}^*, m_{\bar{y}^*})) \geq s. \tag{6.1.10}$$

Since α_f is increasing in the second argument by Remark 2.3.5, we have $\alpha_f(\bar{y}^*, \bar{m}) \geq \alpha_f(\bar{y}^*, m_{\bar{y}^*})$. Hence, by (6.1.10), the monotonicity of $\alpha_{-h^g_{\bar{y}^*}}$, and (6.1.9), we obtain

$$s \leq \alpha_{-h^g_{\bar{y}^*}}(x^*, \alpha_f(\bar{y}^*, m_{\bar{y}^*})) \leq \alpha_{-h^g_{\bar{y}^*}}(x^*, \alpha_f(\bar{y}^*, \bar{m})) < s,$$

which is a contradiction. Hence, (6.1.7) follows so that

$$\alpha^{-l}_{f\circ g}(x^*, s) = \sup_{y^* \in D^+\setminus\{0\}} \inf\left\{m \in \mathbb{R} \mid \alpha_{-h^g_{y^*}}(x^*, \alpha_f(y^*, m)) \geq s\right\}. \tag{6.1.11}$$

Let $y^* \in D^+ \setminus \{0\}$. We claim that

$$\inf\left\{m \in \mathbb{R} \mid \alpha_{-h^g_{y^*}}(x^*, \alpha_f(y^*, m)) \geq s\right\} = \inf\left\{m \in \mathbb{R} \mid \alpha_f(y^*, m) \geq \alpha^{-l}_{-h^g_{y^*}}(x^*, s)\right\}. \tag{6.1.12}$$

For each $m \in \mathbb{R}$, by the definition of left inverse,

$$\alpha_{-h^g_{y^*}}(x^*, \alpha_f(y^*, m)) \geq s \quad \Rightarrow \quad \alpha_f(y^*, m) \geq \alpha^{-l}_{-h^g_{y^*}}(x^*, s).$$

Hence, the $\geq$ part of (6.1.12) follows. Next, we prove that $\leq$ part. To get a contradiction, suppose that

$$\inf\left\{m \in \mathbb{R} \mid \alpha_{-h^g_{y^*}}(x^*, \alpha_f(y^*, m)) \geq s\right\} > \tilde{m} > \inf\left\{m \in \mathbb{R} \mid \alpha_f(y^*, m) \geq \alpha^{-l}_{-h^g_{y^*}}(x^*, s)\right\}$$

for some $\tilde{m} \in \mathbb{R}$. By the first inequality, we have

$$\alpha_{-h^g_{y^*}}(x^*, \alpha_f(y^*, \tilde{m})) < s;$$

and by the second inequality together with the monotonicity of α_f, we have

$$\alpha_f(y^*, \tilde{m}) \geq \alpha^{-l}_{-h^g_{y^*}}(x^*, s).$$

Hence, by the monotonicity of $\alpha_{-h^g_{y^*}}$,

$$s \leq \alpha_{-h^g_{y^*}}(x^*, \alpha^{-l}_{-h^g_{y^*}}(x^*, s)) \leq \alpha_{-h^g_{y^*}}(x^*, \alpha_f(y^*, \tilde{m})) < s,$$

a contradiction. Therefore, (6.1.12) follows.

Combining (6.1.11) and (6.1.12) gives

$$\begin{aligned}
\alpha^{-l}_{f \circ g}(x^*, s) &= \sup_{y^* \in D^+ \setminus \{0\}} \inf \left\{ m \in \mathbb{R} \mid \alpha_f(y^*, m) \geq \alpha^{-l}_{-h^g_{y^*}}(x^*, s) \right\} \\
&= \sup_{y^* \in D^+ \setminus \{0\}} \alpha_f^{-l}(y^*, \alpha^{-l}_{-h^g_{y^*}}(x^*, s)),
\end{aligned}$$

which proves (4.1.6). By combining this with Proposition 4.1.1 and using the monotonicity of g, we obtain (4.1.7). □

Finally, we outline the proofs of the results in Section 4.3. Recall that we work with a monotone convex set $\mathcal{K} \subseteq \mathcal{X}$ with $C \subseteq \mathcal{K}$, and we consider two functions $f\colon \mathcal{Y} \to \overline{\mathbb{R}}$ and $g\colon \mathcal{K} \to \mathcal{Y}$. Let $x^* \in C^+$ and $m \in \mathbb{R}$. Similar to the constructions for the case $\mathcal{K} = \mathcal{Y}$ above, we define the sets

$$\mathbb{A}^m_{y^*} = \{x \in \mathcal{K} \mid \langle y^*, -g(x) \rangle \leq \alpha_f(y^*, m)\}, \qquad \tilde{\mathbb{A}}^m_{y^*} = \{x \in \mathcal{K} \mid \langle y^*, -g(x) \rangle < \alpha_f(y^*, m)\}$$

for each $y^* \in D^+$, and the functions $\mathbb{K}^m_{x^*}, \mathbb{K}^m_{x^*} \colon \mathcal{K} \times \bar{D}^+ \to \overline{\mathbb{R}}$ by

$$\mathbb{K}^m_{x^*}(x, y^*) = \langle x^*, -x \rangle - I_{\mathbb{A}^m_{y^*}}(x), \qquad \tilde{\mathbb{K}}_{x^*}(x, y^*) = \langle x^*, -x \rangle - I_{\tilde{\mathbb{A}}^m_{y^*}}(x).$$

After giving these definitions, by using same arguments we can adapt Propositions 6.1.1, 6.1.3, 6.1.4, 6.1.5 and 6.1.6, and Remark 6.1.2 for the following corollary.

Proof of Corollary 4.3.1. This result follows by the same arguments as in the proof of Theorem 4.1.6. □

Proof of Proposition 4.3.2. The proof of (4.3.1) follows the same arguments as the proof of Proposition 4.1.1. Here, we use Corollary 2.3.12 instead of Theorem 2.3.11. The proof of (4.3.2) follows by the same arguments as in Theorem 4.1.2. $\square$

Proof of Proposition 4.3.3. Proof of Proposition 4.2.2 is valid for this result. $\square$

6.2 Proofs for Chapter 5

Proof of Lemma 5.1.4. To prove that Λ is lower demicontinuous, by Remark 3.0.11, we need to prove that $\Lambda^U(M) = \{X \in L^p(\mathbb{R}^n) \mid \Lambda(X) + L^p(\mathbb{R}_+) \subseteq M\}$ is closed for every closed halfspace $M = \{Y \in L^p(\mathbb{R}) \mid \mathbb{E}\left[Y^*Y\right] \geq 0\}$, where $Y^* \in L^q(\mathbb{R})$.

We first claim that if $\Lambda(X) + L^p_+(\mathbb{R}) \subseteq M = \{Y \in L^p(\mathbb{R}) \mid \mathbb{E}\left[Y^*Y\right] \geq 0\}$ for some $X \in L^p(\mathbb{R}^n)$, then $Y^* \in L^q(\mathbb{R}_+)$. To see this, note that $\mathbb{E}\left[Y^*(\Lambda(X) + d)\right] \geq 0$ if and only if $\mathbb{E}[Y^*d] \geq -\mathbb{E}[Y^*\Lambda(X)]$ for every $d \in L^p_+(\mathbb{R})$. Assume that $\mathbb{E}[Y^*d] < 0$ for some $d \in L^p(\mathbb{R}_+)$. Since $L^p(\mathbb{R}_+)$ is a cone, for every $\lambda > 0$, we have $\lambda d \in L^p(\mathbb{R}_+)$. Also, $\lambda\mathbb{E}[Y^*d] \to -\infty$ as $\lambda \to 0$. However, $\lambda\mathbb{E}[Y^*d]$ is bounded by $-\mathbb{E}[Y^*\Lambda(X)]$, hence we get a contradiction. Therefore, $\mathbb{E}[Y^*d] \geq 0$ for all $d \in L^p(\mathbb{R}_+)$, which implies that $Y^* \in L^q(\mathbb{R}_+)$. This completes the proof of the claim.

In view of the claim, let us take $M = \{Y \in L^p(\mathbb{R}) \mid \mathbb{E}\left[Y^*Y\right] \geq 0\}$ for some $Y^* \in L^q(\mathbb{R}_+)$. We aim to show that $\{X \in L^p(\mathbb{R}^n) \mid \Lambda(X) + L^p(\mathbb{R}_+) \subseteq M\}$ is closed. Note that

$$\{X \in L^p(\mathbb{R}^n) \mid \Lambda(X) + L^p(\mathbb{R}_+) \subseteq M\} = \{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}\left[Y^*\Lambda(X)\right] \geq 0\}.$$

Let us first consider case (i), where $\tilde{\Lambda}$ is concave and bounded from above. Thanks to concavity, the set $\{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}[Y^*\tilde{\Lambda}(X)] \geq 0\}$ is convex.

Suppose that $p < +\infty$. Take a sequence $(X^k)_{k\in\mathbb{N}}$ in $\{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}\left[Y^*\Lambda(X)\right] \geq 0\}$ that converges to some $\tilde{X} \in L^p(\mathbb{R}^n)$ strongly. Hence, there exists a subsequence $(X^{k_\ell})_{\ell\in\mathbb{N}}$ that converges to $\tilde{X}$ almost surely. By the continuity of $\tilde{\Lambda}$, and then reverse Fatou's lemma, we get

$$\mathbb{E}[Y^*\Lambda(\tilde{X})] = \mathbb{E}[Y^*\tilde{\Lambda}\circ\tilde{X}] = \mathbb{E}\left[Y^*\lim_{\ell\to\infty}\tilde{\Lambda}\circ X^{k_\ell}\right]$$
$$\geq \limsup_{\ell\to\infty}\mathbb{E}[Y^*\tilde{\Lambda}\circ X^{k_\ell}] = \limsup_{\ell\to\infty}\mathbb{E}[Y^*\Lambda(X^{k_\ell})] \geq 0. \qquad (6.2.1)$$

Hence, $\tilde{X} \in \{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}[Y^*\Lambda(X)] \geq 0\}$ and this set is closed. Note that we can use reverse Fatou's lemma in the above calculation since $\tilde{\Lambda}$ is bounded from above so that $(Y^*\Lambda(X^{k_\ell}))_{\ell\in\mathbb{N}}$ is bounded from above.

Suppose that $p = +\infty$. To prove weak* closedness, let $r > 0$. By Krein-Šmulian theorem, it is enough to prove that $\{X \in L^\infty(\mathbb{R}^n) \mid \mathbb{E}[Y^*\Lambda(X)] \geq 0, \|X\|_\infty \leq r\}$ is closed in $L^1(\mathbb{R}^n)$. Let $(X^k)_{k\in\mathbb{N}}$ be a sequence in this set that converges to some $\tilde{X} \in L^1(\mathbb{R}^n)$ strongly in $L^1(\mathbb{R}^n)$. Hence, we may find a subsequence $(X^{k_\ell})_{\ell\in\mathbb{N}}$ that converges to $\tilde{X}$ almost surely. Repeating the argument in (6.2.1), we see that $\mathbb{E}[Y^*\Lambda(\tilde{X})] \geq 0$. On the other hand, we have $\|X^{k_\ell}\| \leq r$ for all $\ell \in \mathbb{N}$ with probability one. Hence, $\|\tilde{X}\| \leq r$ with probability one so that $\|\tilde{X}\|_\infty \leq r$. It follows that $\tilde{X} \in \{X \in L^\infty(\mathbb{R}^n) \mid \mathbb{E}[Y^*\Lambda(X)] \geq 0, \|X\|_\infty \leq r\}$, proving the closedness of this set in $L^1(\mathbb{R}^n)$.

Next we consider case (ii), where $\tilde{\Lambda}$ and hence Λ are linear. In particular, there exists $a \in \mathbb{R}^n$ such that $\tilde{\Lambda}(x) = a^\mathsf{T} x$ for every $x \in \mathbb{R}^n$. Suppose that $p < +\infty$. Let us take a net $(X^k)_{k\in I}$ in $\{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}[Y^*\Lambda(X)] \geq 0\}$ that converges to some $\tilde{X} \in L^p(\mathbb{R}^n)$ weakly, where I is an arbitrary index set. By linearity and weak convergence, we have

$$\mathbb{E}[Y^*\Lambda(\tilde{X})] = \mathbb{E}[Y^*\tilde{\Lambda}\circ\tilde{X}] = \mathbb{E}[(Y^*a)^\mathsf{T}\tilde{X}] = \lim_{k\in I}\mathbb{E}[(Y^*a)^\mathsf{T}X^k] \geq 0,$$

so that $\tilde{X} \in \{X \in L^p(\mathbb{R}^n) \mid \mathbb{E}[Y^*\Lambda(X)] \geq 0\}$, and this set is weakly closed, hence it is also strongly closed. The case $p = +\infty$ can be treated by Krein-Šmulian theorem as above.

For (iii), let us first observe that $(L^p(\mathbb{R}^n_+))^\# = L^p(\mathbb{R}^n_{++})$ and $(L^p(\mathbb{R}_+))^\# = L^p(\mathbb{R}_{++})$. Now take $X, \bar{X} \in L^p(\mathbb{R}^n)$ with $X \leq_{L^p(\mathbb{R}^n_{++})} \bar{X}$. Hence, for almost every $\omega \in \Omega$, we have $X(\omega) \leq_{\mathbb{R}^n_{++}} \bar{X}(\omega)$. Since $\tilde{\Lambda}$ is regularly increasing, we have

$$\Lambda(X)(\omega) = \tilde{\Lambda}(X(\omega)) < \tilde{\Lambda}(\bar{X}(\omega)) = \Lambda(\bar{X})(\omega)$$

for almost every $\omega \in \Omega$. Therefore, $\Lambda(X) \leq_{L^p(\mathbb{R}_{++})} \Lambda(\bar{X})$. So Λ is regularly increasing. □

Proof of Proposition 5.1.6 . Let $Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}$. Since we have D-concavity, finding the penalty function is a concave maximization problem. Moreover, since the strict sublevel set is nonempty, Slater's condition holds. Hence, we can use strong duality and obtain

$$\begin{aligned}
\alpha_{(-h^{\Lambda}_{Y^*})}(X^*, m) &= \sup_{X \in L^p(\mathbb{R}^n)} \left\{ \mathbb{E}\left[-(X^*)^{\mathsf{T}} X \right] \mid \mathbb{E}\left[-Y^* \Lambda(X) \right] \leq m \right\} \\
&= \inf_{\lambda > 0} \sup_{X \in L^p(\mathbb{R}^n)} \left(\mathbb{E}\left[-(X^*)^{\mathsf{T}} X \right] - \lambda \mathbb{E}\left[-Y^* \Lambda(X) \right] + \lambda m \right) \\
&= \inf_{\lambda > 0} \sup_{X \in L^p(\mathbb{R}^n)} \left(\mathbb{E}\left[-(X^*)^{\mathsf{T}} X + \lambda Y^* \Lambda(X) \right] + \lambda m \right) \\
&= \inf_{\lambda > 0} \left(\mathbb{E}\left[\sup_{x \in \mathbb{R}^n} \left(-(X^*)^{\mathsf{T}} x + \lambda Y^* \tilde{\Lambda}(x) \right) \right] + \lambda m \right),
\end{aligned}$$

where the second equality comes from strong duality (we can ignore the case $\lambda = 0$ as it produces an objective value of $+\infty$) and the fourth equality follows by Theorem 14.60 in [32].

Note that for every $x^* \in \mathbb{R}^n$ and $y^* \in \mathbb{R}_+$, we have

$$\sup_{x \in \mathbb{R}^n} (-x^{\mathsf{T}} x^* + \lambda y^* \tilde{\Lambda}(x)) = \begin{cases} 0 & \text{if } x^* = 0, y^* = 0, \\ \infty & \text{if } x^* \neq 0, y^* = 0, \\ \lambda y^* \tilde{\Phi}\left(\frac{x^*}{\lambda y^*} \right) & \text{if } y^* > 0. \end{cases} \tag{6.2.2}$$

Therefore,

$$\alpha_{(-h^{\Lambda}_{Y^*})}(X^*, m) = \begin{cases} \infty & \text{if } Y^* \notin T_{X^*}, \\ \inf_{\lambda > 0} \left(\mathbb{E}\left[\lambda Y^* \Phi\left(\frac{X^*}{\lambda Y^*} \right) 1_{\{Y^* > 0\}} \right] + \lambda m \right) & \text{if } Y^* \in T_{X^*}, \end{cases} \tag{6.2.3}$$

and by Theorem 4.1.6,

$$\alpha_{\rho \circ \Lambda}(X^*, m) = \inf_{Y^* \in L^q_+(\mathbb{R}) \setminus \{0\}} \alpha_{(-h^{\Lambda}_{Y^*})} \left(X^*, \alpha_\rho \left(Y^*, m \right) \right).$$

By combining this equality with (6.2.3), it follows that

$$\alpha_{\rho \circ \Lambda}(X^*, m) = \inf_{Y^* \in T_{X^*}} \inf_{\lambda > 0} \left(\mathbb{E}\left[\lambda Y^* \Phi\left(\frac{X^*}{\lambda Y^*} \right) 1_{\{Y^* > 0\}} \right] + \lambda \alpha_\rho(Y^*, m) \right).$$

Then, since T_{X^*} is a cone and α_ρ is positively homogeneous, we get

$$\alpha_{\rho\circ\Lambda}(X^*, m) = \inf_{Y^* \in T_{X^*}} \left(\mathbb{E}\left[Y^* \Phi\left(\frac{X^*}{Y^*}\right) 1_{\{Y^*>0\}} \right] + \alpha_\rho(Y^*, m) \right),$$

as desired. □

Proof of Proposition 5.1.7 . By Corollary 4.1.8 and Proposition 2.3.13, we have

$$\begin{aligned}
\alpha^{-l}_{\rho\circ\Lambda}(X^*, s) &= \sup_{Y^* \in L^q(\mathbb{R}_+)\setminus\{0\}} \alpha^{-l}_\rho \left(Y^*, \alpha^{-l}_{-h^\Lambda_{Y^*}}(X^*, s) \right) \\
&= \sup_{Y^* \in L^q(\mathbb{R}_+)\setminus\{0\}} \alpha^{-l}_\rho \left(Y^*, \sup_{\gamma\geq 0} \left(\gamma s - (-h^\Lambda_{Y^*})^*(-\gamma X^*)\right) \right) \\
&= \sup_{Y^* \in L^q(\mathbb{R}_+)\setminus\{0\}} \inf \left\{ m \in \mathbb{R} \mid \alpha_\rho(Y^*, m) \geq \sup_{\gamma \geq 0} \left(\gamma s - (-h^\Lambda_{Y^*})^*(-\gamma X^*)\right) \right\} \\
&= \sup_{Y^* \in L^q(\mathbb{R}_+)\setminus\{0\}} \sup_{\gamma\geq 0} \alpha^{-l}_\rho \left(Y^*, \gamma s - (-h^\Lambda_{Y^*})^*(-\gamma X^*) \right), \qquad (6.2.4)
\end{aligned}$$

where the last equality comes from Lemma 2.3.10. Let us calculate the second argument of α^{-l}_ρ for bounded case $\Phi(0) < +\infty$. For $\gamma = 0$, by using Theorem 14.60 in [32], we have

$$-(-h^\Lambda_{Y^*})^*(0) = -\sup_{Z \in L^p(\mathbb{R}^n)} \mathbb{E}\left[Y^*\Lambda(Z)\right] = -\mathbb{E}\left[\sup_{z\in\mathbb{R}^n} Y^*\Lambda(z)\right] = -\Phi(0)\mathbb{E}[Y^*].$$

Here, the last equality follows by the following simple observation: for every $y^* \in \mathbb{R}_+$,

$$\sup_{z\in\mathbb{R}^n} y^*\Lambda(z) = \begin{cases} 0 & \text{if } y^* = 0, \\ y^*\Phi(0) & \text{else.} \end{cases}$$

For $\gamma > 0$, by Theorem 14.60 in [32], we get

$$(-h^\Lambda_{Y^*})^*(-\gamma X^*) = \sup_{Z \in L^p(\mathbb{R}^n)} \left(-\mathbb{E}\left[\gamma Z^{\mathsf{T}} X^*\right] + \mathbb{E}\left[Y^*\Lambda(Z)\right] \right) = \mathbb{E}\left[\sup_{z\in\mathbb{R}^n} \left(-\gamma z^{\mathsf{T}} X^* + Y^*\Lambda(z) \right) \right].$$

Using the calculation in (6.2.2), it follows that

$$(-h^\Lambda_{Y^*})^*(-\gamma X^*) = \begin{cases} \infty & \text{if } Y^* \notin T_{X^*}, \\ \mathbb{E}\left[Y^*\Phi\left(\frac{\gamma X^*}{Y^*}\right) 1_{\{Y^*>0\}} \right] & \text{if } Y^* \in T_{X^*}. \end{cases}$$

Since α_ρ^{-l} is increasing in the second argument, we can ignore the case $Y^* \notin T_{X^*}$, since the second argument of α_ρ^{-l} will be $-\infty$ in (6.2.4). By the positive homogeneity of α_ρ, for $\gamma > 0$, we have

$$\alpha_\rho^{-l}\left(Y^*, \gamma s - \mathbb{E}\left[Y^*\Phi\left(\frac{\gamma X^*}{Y^*}\right)1_{\{Y^*>0\}}\right]\right) = \alpha_\rho^{-l}\left(\frac{Y^*}{\gamma}, s - \mathbb{E}\left[\frac{Y^*}{\gamma}\Phi\left(\frac{\gamma X^*}{Y^*}\right)1_{\{Y^*>0\}}\right]\right).$$

By combining all the findings, we get

$$\begin{aligned}
&\alpha_{\rho\circ\Lambda}^{-l}(X^*, s) \\
&= \sup_{Y^*\in L^q(\mathbb{R}_+)\backslash\{0\}} \sup_{\gamma\geq 0} \alpha_\rho^{-l}\left(Y^*, \gamma s - (-h_{Y^*}^\Lambda)^*(-\gamma X^*)\right) \\
&= \sup_{Y^*\in L^q(\mathbb{R}_+)\backslash\{0\}} \alpha_\rho^{-l}\left(Y^*, -\Phi(0)\mathbb{E}[Y^*]\right) \vee \sup_{\substack{Y^*\in T_{X^*},\\ \gamma>0}} \alpha_\rho^{-l}\left(\frac{Y^*}{\gamma}, s - \mathbb{E}\left[\frac{Y^*}{\gamma}\Phi\left(\frac{\gamma X^*}{Y^*}\right)1_{\{Y^*>0\}}\right]\right) \\
&= \sup_{Y^*\in L^q(\mathbb{R}_+)\backslash\{0\}} \alpha_\rho^{-l}\left(Y^*, -\Phi(0)\mathbb{E}[Y^*]\right) \vee \sup_{Y^*\in T_{X^*}} \alpha_\rho^{-l}\left(Y^*, s - \mathbb{E}\left[Y^*\Phi\left(\frac{X^*}{Y^*}\right)1_{\{Y^*>0\}}\right]\right),
\end{aligned}$$

where the last equation comes from the fact that T_{X^*} is a cone. Now we can pass to the probabilistic setting. For the left side, make the change of variable $Y^* = \lambda\frac{d\mathbb{Q}}{d\mathbb{P}}$ where $\lambda > 0$ and $\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P})$. By using the positive homogeneity of α_ρ, we have

$$\begin{aligned}
\alpha_\rho^{-l}\left(Y^*, -\Phi(0)\mathbb{E}[Y^*]\right) &= \alpha_\rho^{-l}\left(\lambda\frac{d\mathbb{Q}}{d\mathbb{P}}, -\Phi(0)\mathbb{E}\left[\lambda\frac{d\mathbb{Q}}{d\mathbb{P}}\right]\right) \\
&= \alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\Phi(0)\mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\right]\right) = \alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\Phi(0)\right),
\end{aligned}$$

which gives

$$\sup_{Y^*\in L^q(\mathbb{R}_+)\backslash\{0\}} \alpha_\rho^{-l}\left(Y^*, -\Phi(0)\mathbb{E}[Y^*]\right) = \sup_{\mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P})} \alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\Phi(0)\right).$$

For the other part, we can make the change of variables $X^* = w \cdot \frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}}$ and $Y^* = \lambda\frac{d\mathbb{Q}}{d\mathbb{P}}$ as before and get

$$\sup_{Y^*\in T_{X^*}} \alpha_\rho^{-l}\left(Y^*, s - \mathbb{E}\left[Y^*\Phi\left(\frac{X^*}{Y^*}\right)1_{\{Y^*>0\}}\right]\right) = \sup_{\substack{\mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P}),\lambda>0\\ w_i\mathbb{S}_i\ll\mathbb{Q}}} \alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, \frac{s}{\lambda} - \mathbb{E}_\mathbb{Q}\left[\Phi\left(\frac{w\cdot\mathbf{d}\mathbb{S}}{\lambda d\mathbb{Q}}\right)\right]\right).$$

Finally, we have

$$\alpha_{\rho\circ\Lambda}^{-l}\left(w\cdot\frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}}, s\right) =$$
$$\max\left\{\sup_{\mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P})}\alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\Phi(0)\right), \sup_{\substack{\mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P}),\lambda>0\\ w_i\mathbb{S}_i\ll\mathbb{Q}}}\alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, \frac{s}{\lambda} - \mathbb{E}_{\mathbb{Q}}\left[\Phi\left(\frac{w\cdot\mathbf{d}\mathbb{S}}{\lambda d\mathbb{Q}}\right)\right]\right)\right\}.$$

For the unbounded case $\Phi(0) = +\infty$, we can ignore the first term above by the monotonicity of α_ρ^{-l}. □

Proof of Proposition 5.1.8. By Proposition 4.2.2, we have the following

$$R(X) = \rho\circ\Lambda(X) = \sup_{X^*\in L^q(\mathbb{R}_+^n)\setminus\{0\}}\ \sup_{Y^*\in L^q(\mathbb{R}_+)\setminus\{0\}}\alpha_\rho^{-l}\left(Y^*, -\mathbb{E}\left[X^\mathsf{T}X^*\right] - (-h_{Y^*}^\Lambda)^*(-X^*)\right).$$

We will calculate the second argument of α_ρ^{-l}. By Theorem 14.60 in [32], we get

$$(-h_{Y^*}^\Lambda)^*(-X^*) = \sup_{Z\in L^p(\mathbb{R}^n)}\left(-\mathbb{E}\left[Z^\mathsf{T}X^*\right] + \mathbb{E}\left[Y^*\Lambda(Z)\right]\right) = \mathbb{E}\left[\sup_{z\in\mathbb{R}^n}\left(-z^\mathsf{T}X^* + Y^*\tilde{\Lambda}(z)\right)\right].$$

By the calculation in (6.2.2), we have

$$(-h_{Y^*}^\Lambda)^*(-X^*) = \begin{cases}\infty & \text{if } Y^*\notin T_{X^*},\\ \mathbb{E}\left[Y^*\Phi\left(\frac{X^*}{Y^*}\right)1_{\{Y^*>0\}}\right] & \text{if } Y^*\in T_{X^*}.\end{cases}$$

Since α_ρ^{-l} is increasing in the second argument, we can ignore the case $Y^*\notin T_{X^*}$ since the second argument will be $-\infty$. Therefore, we have

$$R(X) = \sup_{X^*\in L^q(\mathbb{R}_+^n)}\ \sup_{Y^*\in T_{X^*}}\alpha_\rho^{-l}\left(Y^*, -\mathbb{E}\left[X^\mathsf{T}X\right] - \mathbb{E}\left[Y^*\Phi\left(\frac{X^*}{Y^*}\right)1_{\{Y^*>0\}}\right]\right).$$

We can make the change of variables $X^* = w\cdot\frac{\mathbf{d}\mathbb{S}}{d\mathbb{P}}$ and $Y^* = \lambda\frac{d\mathbb{Q}}{d\mathbb{P}}$ as before and we get

$$\begin{aligned}R(X) &= \sup_{X^*\in L^q(\mathbb{R}_+^n)}\ \sup_{Y^*\in T_{X^*}}\alpha_\rho^{-l}\left(Y^*, -\mathbb{E}\left[X^\mathsf{T}X\right] - \mathbb{E}\left[Y^*\Phi\left(\frac{X^*}{Y^*}\right)1_{\{Y^*>0\}}\right]\right)\\ &= \sup_{\substack{w\in\mathbb{R}_+^n\setminus\{0\},\mathbb{S}\in\mathcal{M}_n^q(\mathbb{P})\\ \mathbb{Q}\in\mathcal{M}_1^q(\mathbb{P}),w_i\mathbb{S}_i\ll\mathbb{Q}}}\alpha_\rho^{-l}\left(\frac{d\mathbb{Q}}{d\mathbb{P}}, -\mathbb{E}_{\mathbb{Q}}\left[\Phi\left(\frac{w\cdot\mathbf{d}\mathbb{S}}{d\mathbb{Q}}\right)\right] - w^\mathsf{T}\mathbb{E}_{\mathbb{S}}\left[X\right]\right),\end{aligned}$$

after using the positive homogeneity of α_ρ and writing w instead of $\frac{w}{\lambda}$. $\square$

Proof of Proposition 5.2.1. Since we have concavity, finding the penalty function will be a concave maximization problem. Since Slater's condition holds, we can use the strong duality.

$$\begin{aligned}
\alpha_{(-h^{\Lambda}_{Y^*})}(X^*,m) &= \sup_{X\in L^p(\mathbb{R}^n_+)} \left\{\mathbb{E}\left[-X^\mathsf{T}X^*\right] \mid \mathbb{E}\left[-Y^*\Lambda(X)\right]\le m\right\}\\
&= \inf_{\lambda\ge 0}\sup_{X\in L^p(\mathbb{R}^n_+)} \left(\mathbb{E}\left[-X^\mathsf{T}X^*+\lambda Y^*\Lambda(X)\right]+\lambda m\right)\\
&= \inf_{\lambda\ge 0}\mathbb{E}\left[\sup_{x\in\mathbb{R}^n_+}\left(-x^\mathsf{T}X^*+\lambda Y^*\tilde{\Lambda}(x)+\lambda m\right)\right],
\end{aligned}$$

where last equality is by Theorem 14.60 in [32]. For $\lambda = 0$, by using the fact that $X^*\in L^q(\mathbb{R}^n_+)$, we reach

$$\sup_{X\in L^p(\mathbb{R}^n_+)} \left(\mathbb{E}\left[-X^\mathsf{T}X^*+\lambda Y^*\Lambda(X)\right]+\lambda m\right) = \sup_{X\in L^p(\mathbb{R}^n_+)}\mathbb{E}\left[-X^\mathsf{T}X^*\right]=0.$$

On the other hand, by the calculation in (6.2.2), we have

$$\alpha_{(-h^{\Lambda}_{Y^*})}(X^*,m) = 0\wedge\inf_{\lambda>0}\left(\lambda m+\mathbb{E}\left[1_{\{Y^*>0\}}\lambda Y^*\Phi\left(\frac{X^*}{\lambda Y^*}\right)\right]\right),$$

and by Corollary 4.3.1, we obtain

$$\begin{aligned}
\alpha_{\rho\circ\Lambda}(X^*,m) &= \inf_{Y^*\in L^q(\mathbb{R}_+)\setminus\{0\}}\alpha_{(-h^{\Lambda}_{Y^*})}(X^*,\alpha_\rho(Y^*,m))\\
&= \inf_{Y^*\in L^q(\mathbb{R}_+)\setminus\{0\}} 0\wedge\inf_{\lambda>0}\left(\lambda\alpha_\rho(Y^*,m)+\mathbb{E}\left[1_{\{Y^*>0\}}\lambda Y^*\Phi\left(\frac{X^*}{\lambda Y^*}\right)\right]\right)\\
&= 0\wedge\inf_{Y^*\in L^q(\mathbb{R}_+)\setminus\{0\}}\left(\alpha_\rho(Y^*,m)+\mathbb{E}\left[1_{\{Y^*>0\}}Y^*\Phi\left(\frac{X^*}{Y^*}\right)\right]\right),
\end{aligned}$$

where last line follows since α is positively homogeneous in the first component and $L^q(\mathbb{R}_+)$ is a cone.

Next, let us fix some arbitrary $n\in\mathbb{N}$ and take

$$Y^*_n := \left(1-\frac{1}{n}\right)Y^*1_{\{Y^*>0\}}+\frac{1}{n}1_{\{Y^*=0\}}\in L^q(\mathbb{R}_{++}).$$

Then, we have

$$
\begin{aligned}
&\inf_{\bar{Y}^* \in L^q(\mathbb{R}_{++})} \left(\alpha_\rho \left(\bar{Y}^*, m\right) + \mathbb{E}\left[1_{\{\bar{Y}^*>0\}} \bar{Y}^* \Phi\left(\frac{X^*}{\bar{Y}^*}\right)\right]\right) \\
&\leq \alpha_\rho \left(Y_n^*, m\right) + \mathbb{E}\left[1_{\{Y_n^*>0\}} Y_n^* \Phi\left(\frac{X^*}{Y_n^*}\right)\right] \\
&= \sup_{Y \in S_m^\rho} -\mathbb{E}\left[Y Y_n^*\right] + \mathbb{E}\left[1_{\{Y_n^*>0\}} \sup_{x \in \mathbb{R}_+^n} \left(-X^{*T} x + Y_n^* \tilde{\Lambda}(x)\right)\right] \\
&\leq \left(1 - \frac{1}{n}\right) \alpha_\rho \left(Y^* 1_{\{Y^*>0\}}, m\right) + \frac{1}{n} \alpha_\rho \left(1_{\{Y^*=0\}}, m\right) + \left(1 - \frac{1}{n}\right) \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right] \\
&\quad + \frac{1}{n} \mathbb{E}\left[1_{\{1_{\{Y^*=0\}}>0\}} 1^*_{\{Y^*=0\}} \Phi\left(\frac{X^*}{1^*_{\{Y^*=0\}}}\right)\right],
\end{aligned}
$$

where the last equality comes from the fact that supremum of affine functions is convex and indicator function of a convex set is a convex function. These inequalities are valid for every $n \in \mathbb{N}$, hence by sending n to ∞, we get

$$
\begin{aligned}
&\inf_{\bar{Y}^* \in L^q(\mathbb{R}_{++})} \left(\alpha_\rho \left(\bar{Y}^*, m\right) + \mathbb{E}\left[1_{\{\bar{Y}^*>0\}} \bar{Y}^* \Phi\left(\frac{X^*}{\bar{Y}^*}\right)\right]\right) \\
&\leq \alpha_\rho \left(Y^* 1_{\{Y^*>0\}}, m\right) + \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right] \\
&= \alpha_\rho \left(Y^*, m\right) + \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right],
\end{aligned}
$$

where last equality is trivial since it is the set where $Y^* = 0$ and does not affect the expectation. Since this inequality true for every $Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}$, by taking infimum we will have the following

$$
\begin{aligned}
&\inf_{Y^* \in L^q(\mathbb{R}_{++})} \left(\alpha_\rho \left(Y^*, m\right) + \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right]\right) \\
&\leq \inf_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \left(\alpha_\rho \left(Y^*, m\right) + \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right]\right).
\end{aligned}
$$

Also since $L^q(\mathbb{R}_{++}) \subseteq L^q(\mathbb{R}_+) \setminus \{0\}$, the reverse inequality holds as well, hence we obtain

$$
\begin{aligned}
&\inf_{Y^* \in L^q(\mathbb{R}_{++})} \left(\alpha_\rho \left(Y^*, m\right) + \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right]\right) \\
&= \inf_{Y^* \in L^q(\mathbb{R}_+) \setminus \{0\}} \left(\alpha_\rho \left(Y^*, m\right) + \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right]\right),
\end{aligned}
\tag{6.2.5}
$$

as desired. $\square$

Proof of Proposition 5.2.2. By Proposition 4.3.3 we have

$$R(X) = \rho \circ \Lambda(X) = \sup_{X^* \in L^q(\mathbb{R}^n_+)\backslash\{0\}} \sup_{Y^* \in L^q(\mathbb{R}_+)\backslash\{0\}} \alpha_\rho^{-l} \left(Y^*, -\mathbb{E}\left[X^\mathsf{T} X^*\right] - (-h^\Lambda_{Y^*})^*(-X^*)\right).$$

We will calculate the second argument. By using Theorem 14.60 in [32], we get

$$(-h^\Lambda_{Y^*})^*(-X^*) = \sup_{Z \in L^p(\mathbb{R}^n_+)} \left(-\mathbb{E}\left[Z^\mathsf{T} X^*\right] + \mathbb{E}\left[Y^* \Lambda(Z)\right]\right) = \mathbb{E}\left[\sup_{z \in \mathbb{R}^n_+} \left(-z^\mathsf{T} X^* + Y^* \tilde{\Lambda}(z)\right)\right].$$

By (6.2.2), we have

$$(-h^\Lambda_{Y^*})^*(-X^*) = \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right].$$

Now, let us look for the following term by using Lemma 2.3.10:

$$\begin{aligned}
&\sup_{Y^* \in L^q(\mathbb{R}_+)\backslash\{0\}} \alpha_\rho^{-l}\left(Y^*, -\mathbb{E}\left[X^\mathsf{T} X^*\right] - \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right]\right) \\
&= \sup_{Y^* \in L^q(\mathbb{R}_+)\backslash\{0\}} \inf\left\{m \in \mathbb{R} \mid \alpha_\rho(Y^*, m) \geq -\mathbb{E}\left[X^\mathsf{T} X^*\right] - \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right]\right\} \\
&= \sup_{Y^* \in L^q(\mathbb{R}_+)\backslash\{0\}} \inf\left\{m \in \mathbb{R} \mid \alpha_\rho(Y^*, m) + \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right] \geq -\mathbb{E}\left[X^\mathsf{T} X^*\right]\right\} \\
&= \inf\left\{m \in \mathbb{R} \mid \forall Y^* \in L^q(\mathbb{R}_+) \setminus \{0\} \colon \alpha_\rho(Y^*, m) + \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right] \geq -\mathbb{E}\left[X^\mathsf{T} X^*\right]\right\} \\
&= \inf\left\{m \in \mathbb{R} \mid \inf_{Y^* \in L^q(\mathbb{R}_+)\backslash\{0\}} \left(\alpha_\rho(Y^*, m) + \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right]\right) \geq -\mathbb{E}\left[X^\mathsf{T} X^*\right]\right\} \\
&= \inf\left\{m \in \mathbb{R} \mid \inf_{Y^* \in L^q(\mathbb{R}_{++})} \left(\alpha_\rho\left(Y^*, m\right) + \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right]\right) \geq -\mathbb{E}\left[X^\mathsf{T} X^*\right]\right\} \\
&= \sup_{Y^* \in L^q(\mathbb{R}_{++})} \inf\left\{m \in \mathbb{R} \mid \alpha_\rho(Y^*, m) \geq -\mathbb{E}\left[X^\mathsf{T} X^*\right] - \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right]\right\} \\
&= \sup_{Y^* \in L^q(\mathbb{R}_{++})} \alpha_\rho^{-l}\left(Y^*, -\mathbb{E}\left[X^\mathsf{T} X^*\right] - \mathbb{E}\left[1_{\{Y^*>0\}} Y^* \Phi\left(\frac{X^*}{Y^*}\right)\right]\right).
\end{aligned}$$

Here, we use (6.2.5) in the fifth equality and Lemma 2.3.10 in the sixth equality. $\square$

Chapter 7

Conclusion

In this thesis, we give a dual representation for the composition of quasiconvex functions and apply this representation to systemic risk measures. We have the well-known Fenchel-Moreau theorem for the dual representation of a convex function. For the composition of convex functions, a dual represenation is shown in Theorem 2.8.10 of [2] and in Theorem 3 of [3]. For a single quasiconvex function, a dual representation is provided by Theorem 3 of [8]. The next question is to figure out the dual representation of the composition of quasiconvex functions and we give answer for this question in this thesis.

In Chapter 2, we focus on the dual representation of an extended real-valued quasiconvex function. Firstly, we define the quasiconvexity, monotonicity and lower semicontinuity concepts for a function and mention the properties of its sublevel sets in Section 2.1, we briefly define the order concept for vectors in Section 2.2 since we work on topological vector spaces. In Section 2.3, we define the minimal penalty function which will be used in the dual representation. We have shown the relation between the sublevel sets and the penalty function by using a separation argument in Remark 2.3.8. We have shown the dual representation for quasiconvex functions in Theorem 2.3.11, which is also a part of Theorem 3 in [8], by using the left inverse of the penalty function and using the characterization in Remark 2.3.8. Then, we look at the case where the function is defined on a convex set but not on the whole space in Corollary 2.3.12. Finally, we have shown the relation between the minimal penalty function and the Fenchel conjugate in 2.3.13.

In Chapter 3, we generalize the concepts of monotonicity, quasiconvexity and semicontinuity for vector-valued functions. We show the relation between the vector-valued function and its scalarizations in terms of these concepts. We also define a notion of strictly monotonicity for vector-valued functions, we call a function with this property regularly increasing.

Section 4.1 is the main part of this thesis. We give the dual representation result in this section. First, in Proposition 4.1.1, we have shown that Theorem 2.3.11 can be applied to the composition of two quasiconvex functions. Then, in Theorem 4.1.2, we give a dual representation for the composition of two quasiconvex functions with a set of desirable properties, namely, the extended real-valued function is decreasing, quasiconvex and lower semicontinuous; the vector-valued function is increasing, naturally quasiconcave and lower demicontinuous. The real problem is whether we can get an explicit formula for the penalty function of the composition. We answer this question by Theorem 4.1.6, the main theorem of this thesis, which works under the additional assumptions that the vector-valued function is regularly increasing and a compact cone generator exists for the dual of the ordering cone. The compactness condition is a necessary for the application of the minimax result in [25], which is usual in minimax-type results. Also, Theorem 4.1.6 states that after we guarantee the existence of a compact cone generator, we can switch to an arbitrary cone generator which does not have to be compact. This result is useful in the applications in Chapter 5 since, for instance, we can work on $\mathcal{M}_1^q(\mathbb{P})$, which is not compact, thanks to this property. In Corollary 4.1.8, we show that the implication of Theorem 4.1.6 on the left inverse of the minimal penalty function of the composition and get the same dual representation with the Theorem 4.1.2. In Section 4.2, we first look at the case where two functions are convex and get a consistent result with the dual representations in the literature such as Theorem 2.8.10 in [2]. Then, we look at the case where only the vector-valued function is convex in Proposition 4.2.2. This result is important since when the real-valued function is quasiconvex and the vector-valued function is convex, the composition is still a quasiconvex composition so this representation is new in the literature to the best of our knowledge. We also use this type of composition in our applications. We give the same type of results for functions that are defined on convex sets in Section 4.3. Finally, we discuss the existence of compact cone generators in concrete settings in Section 4.4. Relevant to the applications on systemic risk, we pay attention to the case of L^p spaces.

In Chapter 5, we work on systemic risk measures for the application of Theorem 4.1.6. In the literature (e.g., in [17]), the basic construction of a systemic risk measures is the composition of an aggregation function and a risk measure. Hence, Theorem 4.1.6 is immediately relevant for obtaining dual representations for systemic risk measures. We review some well-known quasiconvex risk measures and their penalty functions in Example 5.1.2 and Example 5.1.3. Then, we adapt Proposition 4.2.2 to the probabilistic setting in Proposition 5.1.6, Proposition 5.1.7 and Proposition 5.1.8. Finally, in Section 5.2, we apply the main result to the Eisenberg-Noe model, a prominent example of a clearing system. Chapter 6 has the proofs of some results in Section 4.1 and Chapter 5.

To sum up, in this thesis, we give a dual representation result for the composition of quasiconvex functions. We give its applications on systemic risk measures. For the future work, these dual representation results can be applied to improve the modeling capacity in different fields. For instance, using dual representations, new computational methods can be developed for quasiconvex programming problems.

Bibliography

[1] R. T. Rockafellar, *Convex Analysis.* Princeton University Press, 1970.

[2] C. Zălinescu, *Convex Analysis in General Vector Spaces.* World Scientific, 2002.

[3] R. I. Boţ, S.-M. Grad, and G. Wanka, "Generalized Moreau-Rockafellar results for composed convex functions," *Optimization*, vol. 58, no. 7, pp. 917–933, 2009.

[4] J. V. Burke, T. Hoheisel, and Q. V. Nguyen, "A study of convex convex-composite functions via infimal convolution with applications," *Mathematics of Operations Research*, vol. to appear, 2021.

[5] J.-P. Penot and M. Volle, "On quasi-convex duality," *Mathematics of Operations Research*, vol. 15, no. 4, pp. 597–625, 1990.

[6] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, "Complete monotone quasiconcave duality," *Mathematics of Operations Research*, vol. 36, pp. 321–339, 2011.

[7] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, "Risk measures: rationality and diversification," *Mathematical Finance*, vol. 21, pp. 743–774, 2011.

[8] S. Drapeau and M. Kupper, "Risk preferences and their robust representation," *Mathematics of Operations Research*, vol. 38, no. 1, pp. 28–62, 2013.

[9] M. Frittelli and M. Maggis, "Complete duality for quasiconvex dynamic risk measures on modules of the l^p-type," *Statistics and Risk Modeling*, vol. 31, no. 1, pp. 103–128, 2014.

[10] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, 1999.

[11] H. Föllmer and A. Schied, *Stochastic Finance: An Introduction in Discrete Time.* De Gruyter, fourth edition ed., 2016.

[12] E. Mastrogiacomo and E. Rosazza Gianin, "Portfolio optimization with quasiconvex risk measures," *Mathematics of Operations Research*, vol. 40, pp. 1042–1059, 2015.

[13] S. Källblad, "Risk- and ambiguity-averse portfolio optimization with quasiconvex utility functionals," *Finance and Stochastics*, vol. 21, pp. 397–425, 2017.

[14] Ç. Ararat, "Portfolio optimization with two quasiconvex risk measures," *Turkish Journal of Mathematics*, vol. 45, pp. 695–717, 2021.

[15] A. H. Hamel and F. Heyde, "Duality for set-valued measures of risk," *SIAM Journal on Financial Mathematics*, vol. 1, pp. 66–95, 2010.

[16] A. H. Hamel, F. Heyde, and B. Rudloff, "Set-valued risk measures for conical market models," *Mathematics and Financial Economics*, vol. 5, pp. 1–28, 2011.

[17] C. Chen, G. Iyengar, and C. C. Moallemi, "An axiomatic approach to systemic risk," *Management Science*, vol. 59, 2013.

[18] Z. Feinstein, B. Rudloff, and S. Weber, "Measures of systemic risk," *SIAM Journal on Financial Mathematics*, vol. 8, 2017.

[19] F. Biagini, J.-P. Fouque, M. Frittelli, and T. Meyer-Brandis, "A unified approach to systemic risk measures via acceptance sets," *Mathematical Finance*, vol. 29, 2019.

[20] Ç. Ararat and B. Rudloff, "Dual representations for systemic risk measures," *Mathematics and Financial Economics*, vol. 14, 2020.

[21] L. Eisenberg and T. H. Noe, "Systemic risk in financial systems," *Management Science*, vol. 47, 2001.

[22] T. Tanaka, "Generalized quasiconvexities, cone saddle points, and minimax theorem for vector-valued functions," *Journal of Optimization Theory and Applications*, vol. 81, pp. 355–377, 1994.

[23] D. Kuroiwa, "Convexity for set-valued maps," *Applied Mathematics Letters*, vol. 9, no. 2, pp. 97–101, 1996.

[24] M. Sion, "On general minimax theorems," *Pacific Journal of Mathematics*, vol. 8, pp. 171–176, 1958.

[25] F.-C. Liu, "A note on the von Neumann-Sion minimax principle," *Bulletin of the Institute of Mathematics. Academia Sinica*, vol. 6, no. 2, part 2, pp. 517–523, 1978.

[26] G. H. Greco and M. P. Moschen, "A minimax inequality for marginally semicontinuous functions," in *Minimax Theory and Applications* (B. Ricceri and S. Simons, eds.), pp. 41–51, Kluwer Academic Publishers, 1998.

[27] C.-Z. Cheng and B.-L. Lin, "Nonlinear two functions minimax theorems," in *Minimax Theory and Applications* (B. Ricceri and S. Simons, eds.), pp. 1–20, Kluwer Academic Publishers, 1998.

[28] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, third edition ed., 2006.

[29] J. Glück and M. R. Weber, "Almost interior points in ordered banach spaces and the long-term behaviour of strongly positive operator semigroups," *Studia Mathematica*, vol. 254, pp. 237–263, 2020.

[30] M. Reed and B. Simon, *I: Functional Analysis*. Methods of Modern Mathematical Physics, Academic Press, revised and enlarged edition ed., 1980.

[31] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, "Risk measures: Rationality and diversification," *Mathematical Finance*, vol. 21, no. 4, pp. 743–774, 2011.

[32] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*. Springer, third printing ed., 2009.