# DECOMPOSABLE SUMS AND THEIR IMPLICATIONS ON NATURALLY QUASICONVEX RISK MEASURES

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

Barış Bilir

September 2020

DECOMPOSABLE SUMS AND THEIR IMPLICATIONS ON NATURALLY QUASICONVEX RISK MEASURES
By Barış Bilir
September 2020

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)

Elisa Mastrogiacomo

Firdevs Ulus

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

# ABSTRACT

# DECOMPOSABLE SUMS AND THEIR IMPLICATIONS ON NATURALLY QUASICONVEX RISK MEASURES

Barış Bilir
M.S. in Industrial Engineering
Advisor: Çağın Ararat
September 2020

When measuring risk in finance, it is natural to expect that risk decreases with diversification. For risk measures, convexity and quasiconvexity are the two properties which capture the concept of diversification. In between these two properties, there is natural quasiconvexity. Natural quasiconvexity is an old but not so well-known property which is weaker than convexity but stronger than quasiconvexity. In the literature, a lot of effort is put on the analysis of the convexity and the quasiconvexity properties of risk measures. However, a detailed discussion on naturally quasiconvex risk measures is still missing and this thesis aims to fill this gap. Natural quasiconvexity is equivalent to a property called $\star$-quasiconvexity. By making use of this equivalence, we relate naturally quasiconvex risk measures to additively decomposable sums. A notion called convexity index, which is defined in in the literature in 1980s, plays a crucial role in the discussion of additively decomposable sums. Next, we turn our attention to naturally quasiconvex risk measures. By making use of the results on additively decomposable sums, we prove that natural quasiconvexity and convexity are exactly the same properties for conditional risk measures defined on $L^p$, for $p \geq 1$, under some mild conditions. Lastly, we study naturally quasiconvex risk measures on $L^2$ as a special case.

*Keywords:* Risk measures, natural quasiconvexity, additively decomposable sums, convexity index.

# ÖZET

# AYRIŞTIRILABİLİR TOPLAMLAR VE DOĞAL YARI-DIŞBÜKEY RİSK ÖÇERLER ÜZERİNDEKİ SONUÇLARI

Barış Bilir
Endüstri Mühendisliği, Yüksek Lisans
Tez Danışmanı: Çağın Ararat
Eylül 2020

Finansta, risk ölçümünde çeşitlendirme ile riskin düşmesi beklenen bir durumdur. Risk ölçerlerde bu durum dışbükeylik ve yarı-dışbükeylik özellikleri ile sağlanır. Bu iki özelliğin arasında doğal yarı-dışbükeylik özelliği vardır. Doğal yarı-dışbükeylik, eski bir özellik olmakla birlikte pek bilinmemektedir, yarı-dışbükeylikten güçlü fakat dışbükeylikten daha zayıf bir özelliktir. Akademik kaynaklarda, risk ölçerlerin dışbükeylik ve yarı-dışbükeylik özelliklerini irdelemek için yapılmış birçok çalışma bulunmaktadır. Ancak, doğal yarı-dışbükey risk ölçerler üzerine yapılmış kapsamlı bir çalışma mevcut değildir. Bu tezde, literatürdeki bu boşluğu giderme amacı güdülmüştür. Doğal yarı-dışbükeylik, $\star$-yarı-dışbükeylik özelliğine denktir. Bu iki özelliğin denkliğinden faydalanarak, doğal yarı-dışbükey risk ölçerler ayrıştırılabilir toplamlar olarak yazılabilir. Bu sebepten, doğal yarı-dışbükey risk ölçerleri incelemek için öncelikle ayrıştırılabilir toplamların çalışılması gerekmektedir. Bu tezin ilk bölümünde ayrışılabilir toplamlar üzerinde durulmuştur. 1980'li yıllarda tanımlanan dışbükeylik indisi, ayrıştırılabilir toplamları analiz etmemizde önemli bir araç olmuştur. Tezin ikinci kısmında, doğal yarı-dışbükey risk ölçerler anlatılmıştır. İlk kısımda elde ettiğimiz sonuçlardan faydalanarak, makul koşullar altında ve her $p \geq 1$ için, $L^p$ uzaylarında tanımlı risk ölçerler için dışbükeylik ve doğal yarı-dışbükeylik özelliklerinin denk olduğu gösterilmiştir. Son olarak $L^2$ uzayında tanımlı risk ölçerler, özel bir durum olarak incelenmiştir.

*Anahtar sözcükler*: Risk ölçer, doğal yarı-dışbükeylik, ayrıştırılabilir toplam fonksiyonları, dışbükeylik indisi.

# Acknowledgement

First and foremost, I am deeply indebted to my thesis supervisor Çağın Ararat for his guidance, patience, and expertise. Without him, this thesis would not be possible. I learned everything I know about probability, doing research and teaching from him. He is the person who changed my life entirely and I will be in debt to him forever.

I would also like to extend my deepest gratitude to Elisa Mastrogiacomo for her invaluable contributions to this thesis, being on my committe and her hospitality during our time in Italy.

I would also like to thank Fridevs Ulus, who graciously agreed to be on my committe, for her valuable time to read and review this thesis.

I am grateful to my parents Mehmet, Şansel and my brother Kemal for their endless love and constant support, which worth more than I can express on paper.

I would like to extend my sincere thanks to my true friend Furkan Saygın Şener for his support and guidance. I couldn't have asked for a better friend than him.

Thanks also to Betül Akbal, Furkan Altunyaldız, Özlem Eda Başalma,Taner Keskin, Erkin Oto and Güzey Şıkman for their invaluable friendship.

Special thanks to Seyit Emre Düzoylum for his help while submitting this thesis and his accompany on probability journey.

Lastly, I would like to thank the one who always cheers me up and makes my life beautiful.

# Contents

# List of Figures

# Chapter 1

# Introduction

A major issue in finance is the measurement of risk of a financial position and the tools that are used to measure it are called risk measures. A risk measure is said to be conditional if it measures the risk of a financial position not today but at an intermediate time. In the study of risk measures, a lot of emphases is given on the analysis of properties of risk measures, however, a property called natural quasiconvexity is still not very well-understood. In this thesis, we are concerned with investigating natural quasiconvexity property for conditional risk measures.

An axiomatic approach to (monetary) risk measures is given in the seminal paper Artzner et al. [1], in which the so-called coherent risk measures are defined. Within the framework of the aforementioned paper, financial positions are modelled as random variables and a risk measure is a real-valued functional defined on a suitable linear space of financial positions. Under the axioms of a coherent risk measure, greater returns for all possible scenarios means lower risk (monotonicity), a deterministic amount added to the position reduces the risk by the same amount (translativity or cash additivity) and risk increases in a sublinear way (subadditivity and positive homogeneity). Coherent risk measures are then generalized to convex risk measures in Föllmer and Schied [2], Frittelli and Gianin [3] and Heath [4]. The motivation behind this generalization is that the risk of a financial position may increase in a nonlinear way with the size of a position. Therefore, subadditivity and positive homogenity axioms are replaced with the weaker convexity axiom. It is noteworthy that convexity explicitly captures the idea that diversification does not increase risk. Later, it is argued in El Karoui and Ravanelli [5] that cash additivity should be replaced with cash sub-additivity since the former ignores uncertainty about interest rates. Moreover, it is a well-known result that quasiconvexity and convexity are equivalent under cash additivity. (see, e.g., Marinacci and Montrucchio [6, Corollary 4.2]). Hence, the replacement of cash additivity with cash sub-additivity pave the way for drawing the distinction between quasiconvexity and convexity properties of risk measures. Important works along these lines are in Cerreia-Vioglio et al. [7], Drapeau and Kupper [8], Frittelli and Maggis [9]. Furthermore, it is argued in Cerreia-Vioglio et al. [10] that quasiconvexity is indeed the right mathematical formulation of diversification under cash sub-additivity.

In the conditional setting, a risk measure gives the risk of a financial position not necessarily in the present time. In other words, it measures the risk of a financial position possibly at an intermediate time. For discussions on conditional risk measures we refer the reader to Bion-Nadal [11], Detlefsen and Scandolo [12], Frittelli and Maggis [9], Riedel [13], Frittelli and Gianin [14] and Ruszczyński and Shapiro [15].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G} \subseteq \mathcal{F}$ and $p \geq 1$. A conditional risk measure $\rho\colon L^p(\Omega, \mathcal{F}, \mathbb{P}) \to L^p(\Omega, \mathcal{G}, \mathbb{P})$ is called naturally quasiconvex if for every $X, Y \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ and $\lambda \in [0,1]$ there exists $\mu \in [0,1]$ such that

$$\rho(\lambda X + (1-\lambda)Y) \leq \mu\rho(X) + (1-\mu)\rho(Y) \quad \text{for every } X, Y \in \mathcal{X},$$

Clearly, natural quasiconvexity is a property that is stronger than quasiconvexity but weaker than convexity. Furthermore, natural quasiconvexity is equivalent to a property called $\star$-quasiconvexity. It turns out that, in view of the equivalance of natural quasiconvexity and $\star$-quasiconvexity, naturally quasiconvex risk measures are highly related with additively decomposable sums, which we discuss next.

In Debreu and Koopmans [16] and Crouzeix and Lindberg [17], an indepth discussion on additively decomposable sums can be found. In their setting, a real-valued function $s$ defined on a product space $X_1 \times \ldots \times X_n$ is called additively decomposable if

$$s(x_1, \ldots x_n) = f_1(x_1) + \ldots + f_n(x_n), \tag{1.0.1}$$

where $X_i$ is an open convex subset of $\mathbb{R}^n$ and $f_i \colon X_i \mapsto \mathbb{R}$ for each $i \in \{1, \ldots, n\}$ for each $i \in \{1, \ldots, n\}$. They define for every real-valued function defined on an open set of $\mathbb{R}$ an index, namely convexity index, which can be used to determine whether it is convex or not. They showed that in the setting of (1.0.1), quasiconvexity of $s$ have implications on convexity of each $f_i$, via convexity indices. To study natuarlly quasiconvex functions, we generalize their results to more general settings, in particular to functions definded on general vector spaces and to additively decomposable infinite sums.

The remainder of this thesis is organized as follows. In Chapter 2, additively decomposable sums on general vector spaces are studied. We introduce the convexity index for extended real-valued functions defined on general vector spaces and present some important properties of it in Section 2.1. Then, in Section 2.2, we show that an additively decomposable finite sum is quasiconvex if and only if either all functions that appear in the sum is convex or all except one is convex together with a condition on the sum of the convexity indices of them. It is noteworthy that Sections 2.1 and 2.2 have strong links with Debreu and Koopmans [16] and Crouzeix and Lindberg [17] since most of the results in these sections are generalizations of their results to real-valued functions defined on general topological vector spaces. To make that generalization, it appears that we only need lower semi-continuity assumption on each function that appears in the additively decomposable sum. In Section 2.3, we showed that almost the same result with the one in the previous section holds for additively infinite sums. In Chapter 3, we studied naturally quasiconvex conditional risk measures. Section 3.1 is a brief introduction to risk measures. In Section 3.2, we discuss natural quasiconvexity and give an equivalent characterization of it. In Section 3.3, we show that convexity and natural quasiconvexity are exactly the same properties for conditional risk measures, under some mild conditions. Finally, in Section 3.4, we defin a new property, namely locality with respect to a basis, and show that, under that property, naturally quasiconvexity and convexity are equivalent with respect to the preorder defined by the cone generated by the elements of the basis that the risk measure local with respect to.

# Chapter 2

# Decomposable sums in general topological vector spaces

## 2.1 Convexity index

Convexity index for real-valued functions defined on $\mathbb{R}^n$ is first introduced in Debreu and Koopmans [16] and redefined in Crouzeix and Lindberg [17]. The main concern of this section is to extend the definition of convexity index to extended real-valued functions defined on general topological vector spaces. This will be a building block in the study of additively decoposable sums in general vector spaces.

As we work with extended real-valued functions, the following conventions for the arithmetic on $\bar{\mathbb{R}} = [-\infty, +\infty]$ are used throughout the paper. We have $0 \cdot (+\infty) = 0 \cdot (-\infty) = 0$, $\frac{1}{0} = +\infty$. We also set $e^z = +\infty$ if $z = +\infty$ and $e^z = 0$ if $z = -\infty$.

Let $X$ be a topological vector space and $f \colon X \to \bar{\mathbb{R}}$ a function which we keep fixed except stated otherwise. We define the *effective domain* of $f$ as the set

$$\operatorname{dom} f \coloneqq \{x \in X \colon f(x) < +\infty\}.$$

We assume that $f$ is proper in the sense that $\operatorname{dom} f \neq \emptyset$ and $f(x) > -\infty$ for every $x \in X$. For each $\lambda \in \mathbb{R}$, we associate to $f$ the function $r_\lambda \colon X \to \mathbb{R}$ defined by

$$r_\lambda(x) := e^{-\lambda f(x)}, \quad x \in X. \tag{2.1.1}$$

First, we present an auxiliary result related to the convexity properties of $r_\lambda$, $\lambda \in \mathbb{R}$, which is helpful to have a better grasp of the definition of convexity index.

**Lemma 2.1.1.** *The following results hold for $r_\lambda$, $\lambda \in \mathbb{R}$, associated to $f \colon X \to \bar{\mathbb{R}}$.*

*(i) Let $\lambda < 0$. Then, $r_\lambda$ is convex if and only if $r_\mu$ is convex for each $\mu < \lambda$.*

*(ii) Let $\lambda > 0$. Then, $r_\lambda$ is concave if and only if $r_\mu$ is concave for each $\mu \in [0, \lambda)$.*

*(iii) If $r_\lambda$ is concave for some $\lambda > 0$, then $r_\mu$ is convex for each $\mu < 0$.*

**Proof.** Let $\lambda \neq 0$ and $\mu < \lambda$. Then, we may write $r_\mu = k_{\mu,\lambda} \circ r_\lambda$, where $k_{\mu,\lambda}(t) := t^{\frac{\mu}{\lambda}}$ for each $t \in [0, +\infty]$ with the conventions $0^0 = (+\infty)^0 = 1$, $(+\infty)^a = +\infty$ for $a > 0$, and $(+\infty)^a = 0$ for $a < 0$.

(a) Let $\lambda < 0$. Suppose that $r_\lambda$ is convex and let $\mu < \lambda$. Let $x_1, x_2 \in X$ and $\eta \in [0, 1]$. For every $t \in [0, +\infty]$, we have

$$k'_{\mu,\lambda}(t) = \frac{\mu}{\lambda} t^{\frac{\mu}{\lambda}-1} \geq 0, \quad k''_{\mu,\lambda}(t) = \Big(\frac{\mu}{\lambda}\Big)\Big(\frac{\mu}{\lambda} - 1\Big) t^{\frac{\mu}{\lambda}-2} \geq 0$$

since $\frac{\mu}{\lambda} > 1$. Hence $k_{\mu,\lambda}$ is convex and increasing. Observe that

$$\begin{aligned}
r_\mu(\eta x_1 + (1-\eta)x_2) &= k_{\mu,\lambda} \circ r_\lambda(\eta x_1 + (1-\eta)x_2) \\
&\leq k_{\mu,\lambda}(\eta r_\lambda(x_1) + (1-\eta) r_\lambda(x_2)) \\
&\leq \eta k_{\mu,\lambda} \circ r_\lambda(x_1) + (1-\eta) k_{\mu,\lambda} \circ r_\lambda(x_2) \\
&= \eta r_\mu(x_1) + (1-\eta) r_\mu(x_2),
\end{aligned}$$

where the first inequality holds since $r_\lambda$ is convex, $k_{\mu,\lambda}$ is increasing and the second inequality follows from the convexity of $k_{\mu,\lambda}$. Therefore, $r_\mu$ is convex.

Conversely, assume that $r_\mu$ is convex for each $\mu < \lambda$. Let $x_1, x_2 \in X$ and $\eta \in [0,1]$. For every $\mu < \lambda$, we have

$$r_\mu(\eta x_1 + (1-\eta)x_2) \leq \eta r_\mu(x_1) + (1-\eta)r_\mu(x_2).$$

Thanks to the continuity of the power function $\mu \mapsto e^{-\mu a}$ on $(-\infty, 0)$ for each fixed $a \in \bar{\mathbb{R}}$, we may let $\mu \to \lambda$ and get

$$r_\lambda(\eta x_1 + (1-\eta)x_2) \leq \eta r_\lambda(x_1) + (1-\eta)r_\lambda(x_2).$$

Hence $r_\lambda$ is convex.

(b) Let $\lambda > 0$. Suppose that $r_\lambda$ is concave. Clearly, $r_0 \equiv 1$ is convex. Let $\mu \in (0, \lambda)$. Let $x_1, x_2 \in X$ and $\eta \in [0,1]$. Similar to (i), it is easy to check that $k_{\mu,\lambda}$ is concave and increasing since $0 \leq \frac{\mu}{\lambda} < 1$. Observe that

$$\begin{aligned} r_\mu(\eta x_1 + (1-\eta)x_2) &= k_{\mu,\lambda} \circ r_\lambda(\eta x_1 + (1-\eta)x_2) \\ &\geq k_{\mu,\lambda}(\eta r_\lambda(x_1) + (1-\eta)r_\lambda(x_2)) \\ &\geq \eta k_{\mu,\lambda} \circ r_\lambda(x_1) + (1-\eta)k_{\mu,\lambda} \circ r_\lambda(x_2) \\ &= \eta r_\mu(x_1) + (1-\eta)r_\mu(x_2), \end{aligned}$$

where the first inequality holds since $r_\lambda$ is concave, $k_{\mu,\lambda}$ is increasing and the second inequality follows from concavity of $k$. Therefore, $r_\mu$ is concave.

Conversely, assume that $r_\mu$ is concave for each $\mu \in [0, \lambda)$. Let $x_1, x_2 \in X$ and $\eta \in [0,1]$. For every $\mu \in [0, \lambda)$, we have

$$r_\mu(\eta x_1 + (1-\eta)x_2) \geq \eta r_\mu(x_1) + (1-\eta)r_\mu(x_2).$$

By letting $\mu \to \lambda$ similar to (i), we get

$$r_\lambda(\eta x_1 + (1-\eta)x_2) \geq \eta r_\lambda(x_1) + (1-\eta)r_\lambda(x_2).$$

Hence $r_\lambda$ is concave.

(c) Assume that there exists $\lambda > 0$ such that $r_\lambda$ is concave. Let $\mu < 0$, $x_1, x_2 \in X$

and $\eta \in [0,1]$. Similar to (i) and (ii), we may conclude that $k_{\mu,\lambda}$ is convex and decreasing since $\frac{\mu}{\lambda} < 0$. Observe that

$$\begin{aligned} r_\mu(\eta x_1 + (1-\eta)x_2) &= k_{\mu,\lambda} \circ r_\lambda(\eta x_1 + (1-\eta)x_2) \\ &\leq k_{\mu,\lambda}(\eta r_\lambda(x_1) + (1-\eta) r_\lambda(x_2)) \\ &\leq \eta k_{\mu,\lambda} \circ r_\lambda(x_1) + (1-\eta) k_{\mu,\lambda} \circ r_\lambda(x_2) \\ &= \eta r_\mu(x_1) + (1-\eta) r_\mu(x_2), \end{aligned}$$

where the first inequality holds since $r_\lambda$ is concave, $k_{\mu,\lambda}$ is decreasing and the second inequality follows from convexity of $k$. Therefore, $r_\mu$ is convex.

$\square$

Let us consider the following two cases. First, assume that there exists $\lambda < 0$ such that $r_\lambda$ is not convex. Then, according to Lemma 2.1.1(i), $r_\gamma$ is not convex for every $\gamma \in [\lambda, 0)$. Moreover, if there exists $\mu < \lambda$ such that $r_\mu$ is convex, then $r_\gamma$ is convex for every $\gamma \in (-\infty, \mu]$. Second, assume otherwise that $r_\lambda$ is convex for every $\lambda < 0$. In view of Lemma 2.1.1(iii), this is a necessary condition for $r_\mu$ to be concave for some $\mu > 0$. Moreover, if $r_\mu$ is concave for some $\mu > 0$, then $r_\gamma$ is concave for every $\gamma \in [0, \mu)$ by Lemma 2.1.1(ii). Figure 3.2 and Figure 2.2 below depict these two cases.

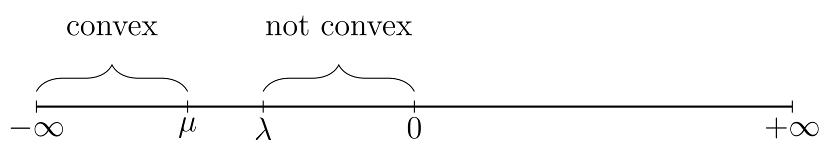


Figure 2.1: $r_\lambda$ is not convex for some $\lambda < 0$.

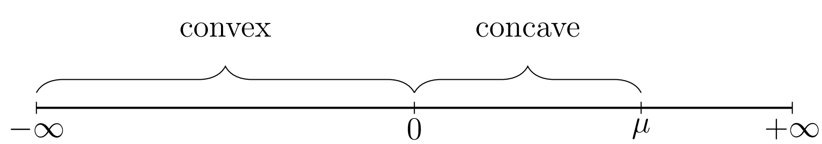


Figure 2.2: $r_\lambda$ is convex for every $\lambda < 0$.

Considering the cases described above, an intriguing query arises to determine the largest $\lambda < 0$ for which $r_\lambda$ is convex (first case) and the largest $\lambda \geq 0$ for which $r_\lambda$ is concave (second case). The following definition is motivated by these ideas.

**Definition 2.1.2.** *The convexity index $c(f) \in \bar{\mathbb{R}}$ of $f$ is defined as follows:*

*(i) if there exists $\bar{\lambda} < 0$ such that $r_{\bar{\lambda}}$ is not convex, then*

$$c(f) := \sup\{\lambda < 0 \colon r_\lambda \text{ is convex}\}.$$

*(ii) if $r_{\bar{\lambda}}$ is convex for every $\bar{\lambda} < 0$, then*

$$c(f) := \sup\{\lambda \geq 0 \colon r_\lambda \text{ is concave}\}.$$

**Remark 2.1.3.** By the discussion preceding Definition 2.1.2, it is clear that $c(f) \in [-\infty, 0)$ in case (i) and $c(f) \in [0, +\infty]$ in case (ii).

As its name suggests, convexity index of $f$ tells whether it is convex or not as stated in the next theorem.

**Theorem 2.1.4.** *The function $f$ is convex if and only if $c(f) \geq 0$.*

**Proof.** Assume that $f$ is a convex function. Let $\bar{\lambda} < 0$. For every $\eta \in [0,1]$, $x_1, x_2 \in X$,

$$\begin{aligned} r_{\bar{\lambda}}(\eta x_1 + (1-\eta)x_2) &= e^{-\bar{\lambda} f(\eta x_1 + (1-\eta)x_2)} \\ &\leq e^{-\bar{\lambda}(\eta f(x_1) + (1-\eta) f(x_2))} \\ &\leq \eta r_{\bar{\lambda}}(x_1) + (1-\eta) r_{\bar{\lambda}}(x_2), \end{aligned}$$

where the first inequality follows from the convexity of $f$ and monotonicity of $t \mapsto e^{-\bar{\lambda} t}$ , the second inequality holds since $t \mapsto e^{-\bar{\lambda} t}$ defined on $[0, +\infty]$ is convex. Therefore, $r_{\bar{\lambda}}$ is convex. Hence, by Definition 2.1.2,

$$c(f) = \sup\{\lambda \geq 0 \colon r_\lambda \text{ is concave}\}$$

so that $c(f) \geq 0$.

Conversely, assume that $c(f) \geq 0$. Let $x_1, x_2 \in X$ and $\eta \in [0,1]$. We claim that

$$f(\eta x_1 + (1-\eta)x_2) \leq \eta f(x_1) + (1-\eta) f(x_2). \tag{2.1.2}$$

Note that (2.1.2) holds trivially if $f(x_1) = +\infty$ or $f(x_2) = +\infty$. Hence, we consider the following cases:

(a) $f(x_1) < +\infty$, $f(x_2) < +\infty$, $f(\eta x_1 + (1-\eta)x_2) < +\infty$,

(b) $f(x_1) < +\infty$, $f(x_2) < +\infty$, $f(\eta x_1 + (1-\eta)x_2) = +\infty$.

For each $\lambda \in \mathbb{R}$, define

$$\begin{aligned} k(\lambda) &\coloneqq r_\lambda(\eta x_1 + (1-\eta)x_2) - \eta r_\lambda(x_1) - (1-\eta) r_\lambda(x_2) \\ &= e^{-\lambda f(\eta x_1 + (1-\eta)x_2)} - \eta e^{-\lambda f(x_1)} - (1-\eta) e^{-\lambda f(x_2)}. \end{aligned}$$

By Remark 2.1.3, $r_\lambda$ is convex for every $\lambda < 0$, which implies that $k(\lambda) \leq 0$ for every $\lambda < 0$. Observe that $k(\lambda) = +\infty$ for every $\lambda < 0$ if case (b) is true, which is in contradiction with the previous statement. Therefore, we are only left with case (a). Noting that $k(0) = 0$, we get

$$k'_-(0) \coloneqq \lim_{\lambda \to 0^-} \frac{k(\lambda) - k(0)}{\lambda} \geq 0.$$

Furthermore, $k$ is differentiable everywhere on $\mathbb{R}$, in particular, at $\lambda = 0$. Hence, the derivative and the left derivative of $k$ at 0 are equal, that is,

$$0 \leq k'_-(0) = k'(0) = -f(\eta x_1 + (1-\eta)x_2) + \eta f(x_1) + (1-\eta) f(x_2).$$

Therefore, (2.1.2) holds. Since $\eta, x_1, x_2$ are arbitrary, $f$ is convex. $\square$

Before proceeding further, we give a basic result which will be useful in the following sections.

**Lemma 2.1.5.** *Let $w \in \mathbb{R}_+$. $c(wf) = \frac{1}{w} c(f)$.*

**Proof.** The result follows trivially when $w = 0$. Suppose $w > 0$. We need to consider the following two cases:

(a) $f$ is convex. Then, $wf$ is convex. Theorem 2.1.4 together with Remark 2.1.3 implies that $c(wf) = \sup\{\lambda \in \mathbb{R} \colon \lambda \geq 0, e^{-\lambda w f} \text{ is concave}\}$. Observe that

$$\begin{aligned} c(wf) &= \sup\{\lambda \in \mathbb{R} \colon \lambda \geq 0, e^{-\lambda w f} \text{ is concave}\} \\ &= \frac{1}{w}\sup\{\lambda w \in \mathbb{R} \colon \lambda \geq 0, e^{-\lambda w f} \text{ is concave}\} \\ &= \frac{1}{w}\sup\{\mu \in \mathbb{R} \colon \lambda \geq 0, e^{-\mu f} \text{ is concave}\} \\ &= \frac{1}{w}c(f). \end{aligned}$$

(b) $f$ is non-convex. Then, $wf$ is non-convex. Theorem 2.1.4 together with Remark 2.1.3 implies that $c(wf) = \sup\{\lambda \in \mathbb{R} \colon \lambda < 0, e^{-\lambda w f} \text{ is convex}\}$. Similar to the previous case, we have

$$\begin{aligned} c(wf) &= \sup\{\lambda \in \mathbb{R} \colon \lambda < 0, e^{-\lambda w f} \text{ is convex}\} \\ &= \frac{1}{w}\sup\{\lambda w \in \mathbb{R} \colon \lambda < 0, e^{-\lambda w f} \text{ is convex}\} \\ &= \frac{1}{w}\sup\{\mu \in \mathbb{R} \colon \lambda < 0, e^{-\mu f} \text{ is convex}\} \\ &= \frac{1}{w}c(f). \end{aligned}$$

Hence, the result holds. □

As discussed in Crouzeix and Lindberg [17, Proposition 3(b)], constant real-valued functions defined on $\mathbb{R}^n$ can be characterized in terms of convexity index. In the proof of this result, they use the fact that convexity of a function with finite values implies continuity. Such a result does not hold for extended real-valued functions defined on general topological vector spaces. Fortunately, the characterization is still valid under a mild condition, as shown in the next theorem.

**Theorem 2.1.6.** *Assume that $f$ is lower semicontinuous. Then, $f$ is a constant function if and only if $c(f) = +\infty$.*

**Proof.** Assume that $f$ is a constant function. Then, for every $\lambda \in \mathbb{R}$, the function $r_\lambda$ is constant, hence convex. By Definition 2.1.2, we have

$$c(f) = \sup\{\lambda \geq 0 \colon r_\lambda \text{ is concave}\}. \tag{2.1.3}$$

Moreover, for each $\lambda \geq 0$, the constant function $r_\lambda$ is also concave. Therefore, $c(f) = +\infty$.

Conversely, assume that $c(f) = +\infty$. Then, by Theorem 2.1.4, $f$ is convex. To get a contradiction, suppose that $f$ is not constant. We need to consider the following two cases:

(a) Suppose that $f$ takes only one finite value, that is, $f(X) = \{c, +\infty\}$ for some $c \in \mathbb{R}$. Then, for every $\lambda < 0$, we have $r_\lambda(x) = e^{-\lambda c}$ for every $x \in \operatorname{dom} f$ and $r_\lambda(x) = +\infty$ for every $x \in X \setminus \operatorname{dom} f$; hence $r_\lambda$ is convex. So (2.1.3) is valid. Moreover, since $f$ is proper and lower semicontinuous, $\operatorname{dom} f = \{x \in X \colon f(x) \leq c\}$ is a closed set. Let $\lambda > 0$. Then, $r_\lambda(x) = e^{-\lambda c} > 0$ for every $x \in \operatorname{dom} f$ and $r_\lambda(x) = 0$ for every $x \in X \setminus \operatorname{dom} f$. Let $x_1 \in \operatorname{dom} f$ and $x_2 \in X \setminus \operatorname{dom} f$. Since $X$ is a topological vector space, the function $[0,1] \ni \eta \mapsto x^\eta := \eta x_1 + (1-\eta)x_2 \in X$ is continuous. Hence, $\lim_{\eta \to 0} x^\eta = x_2$. Since $x_2 \in X \setminus \operatorname{dom} f$ and $X \setminus \operatorname{dom} f$ is an open set, continuity at $\eta = 0$ implies that there exists $\bar{\eta} \in (0,1)$ such that $x^{\bar{\eta}} \in X \setminus \operatorname{dom} f$. In particular, $f(x^{\bar{\eta}}) = +\infty$ and $r_\lambda(x^{\bar{\eta}}) = 0$. It follows that

$$r_\lambda(x^{\bar{\eta}}) = 0 < \bar{\eta} e^{-\lambda c} = \bar{\eta} r_\lambda(x_1) + (1-\bar{\eta}) r_\lambda(x_2)$$

so that $r_\lambda$ is not concave. Therefore, $c(f) = 0$ by (2.1.3), which is a contradiction to $c(f) = +\infty$. So this case is eliminated.

(b) Suppose that $f$ takes at least two finite values, that is, there exist $x_1, x_2 \in X$ such that $f(x_1) < f(x_2) < +\infty$. Since $f$ is convex and $f(x_1) < f(x_2)$, for every $\eta \in (0,1)$, we have

$$f(\eta x_1 + (1-\eta)x_2) \leq \eta f(x_1) + (1-\eta) f(x_2) < f(x_2). \tag{2.1.4}$$

Moreover, since $f$ is lower semi-continuous at $x_2$,

$$f(x_2) \leq \liminf_{x \to x_2} f(x),$$

where

$$\liminf_{x \to x_2} f(x) = \sup \left\{ \inf \left\{ f(x) \colon x \in U \setminus \{x_2\} \right\} : U \subseteq X \text{ is open, } x_2 \in U, U \setminus \{x_2\} \neq \emptyset \right\}.$$

Since $f(x_1) < f(x_2) \leq \liminf_{x \to x_2} f(x)$, there exists an open neighborhood $\bar{U}$ of $x_2$ such that

$$f(x_1) < \inf\{f(x) \colon x \in \bar{U} \setminus \{x_2\}\}.$$

Observe that

$$f(x_1) < f(x) \text{ for every } x \in \bar{U}. \tag{2.1.5}$$

Similar to case (a), the function $[0,1] \ni \eta \mapsto x^\eta := \eta x_1 + (1-\eta)x_2 \in X$ is continuous with $\lim_{\eta \to 0} x^\eta = x_2$, which implies that there exists $\bar{\eta} \in (0,1)$ such that $x^{\bar{\eta}} \in \bar{U}$. By (2.1.5), $f(x_1) < f(x^{\bar{\eta}})$. Together with (2.1.4), we have

$$f(x_1) < f(x^{\bar{\eta}}) < f(x_2). \tag{2.1.6}$$

By Remark 2.1.3, $r_\lambda$ is concave for every $\lambda \geq 0$ since $c(f) = +\infty$. Therefore,

$$r_\lambda(x^{\bar{\eta}}) - \bar{\eta} r_\lambda(x_1) - (1-\bar{\eta}) r_\lambda(x_2) \geq 0 \text{ for every } \lambda > 0. \tag{2.1.7}$$

Let $\lambda > 0$. With some algebraic operations, (2.1.7) can be rewritten as

$$\frac{r_\lambda(x_1) - r_\lambda(x_2)}{r_\lambda(x^{\bar{\eta}}) - r_\lambda(x_2)} \leq \frac{1}{\bar{\eta}}. \tag{2.1.8}$$

Note that

$$\begin{aligned}
\frac{r_\lambda(x_1) - r_\lambda(x_2)}{r_\lambda(x^{\bar\eta}) - r_\lambda(x_2)} &= \frac{r_\lambda(x_1) - r_\lambda(x^{\bar\eta}) + r_\lambda(x^{\bar\eta}) - r_\lambda(x_2)}{r_\lambda(x^{\bar\eta}) - r_\lambda(x_2)} \\
&= \frac{r_\lambda(x_1) - r_\lambda(x^{\bar\eta})}{r_\lambda(x^{\bar\eta}) - r_\lambda(x_2)} + 1 \\
&\geq \frac{r_\lambda(x_1) - r_\lambda(x^{\bar\eta})}{r_\lambda(x^{\bar\eta}) - r_\lambda(x_2)} \\
&\geq \frac{r_\lambda(x_1) - r_\lambda(x^{\bar\eta})}{r_\lambda(x^{\bar\eta})} \\
&\geq \frac{r_\lambda(x_1)}{r_\lambda(x^{\bar\eta})} - 1 \\
&= e^{-\lambda(f(x_1) - f(x^{\bar\eta}))} - 1
\end{aligned}$$

By (2.1.6), $f(x_1) - f(x^{\bar\eta}) < 0$. Therefore,

$$\begin{aligned}
\limsup_{\lambda\to\infty} \frac{r_\lambda(x_1) - r_\lambda(x_2)}{r_\lambda(x^{\bar\eta}) - r_\lambda(x_2)} &\geq \limsup_{\lambda\to\infty} e^{-\lambda(f(x_1) - f(x^{\bar\eta}))} - 1 \\
&= \lim_{\lambda\to\infty} e^{-\lambda(f(x_1) - f(x^{\bar\eta}))} - 1 \\
&= +\infty.
\end{aligned}$$

This implies that there exists $\bar\lambda > 0$ such that

$$\frac{r_{\bar\lambda}(x_1) - r_{\bar\lambda}(x_2)}{r_{\bar\lambda}(x^{\bar\eta}) - r_{\bar\lambda}(x_2)} > \frac{1}{\bar\eta}.$$

By (2.1.8), $r_{\bar\lambda}$ is not concave, which is in contradiction with (2.1.7). Hence, $f$ is constant.

□

## 2.2 Finite Decomposable Sums

In this section, additively decomposable quasiconvex functions defined on general topological vector spaces is discussed. The main results of this section are as the followings: (i) Theorem 2.2.2, which characterizes a quasiconvex decomposable sum in terms of sum of the convexity indices of its coordinate functions, and (ii) Theorem 2.2.5 in which it is presented that given a quasiconvex decomposable sum, either all coordinate functions are convex, or all except one are convex and convexity indices of coordinate functions satisfies an additive formula, and vice versa.

First we prove a technical result that we need in the following part.

**Lemma 2.2.1.** *If $f$ is quasiconvex and not convex, then there exist $x_1, x_2 \in$ $\operatorname{dom} f$, $\bar{t} \in [0,1)$ and $\alpha \in \mathbb{R}$ such that*

$$
\begin{aligned}
&f(x_2) < \alpha \leq f(x_1), \\
&f(tx_1 + (1-t)x_2) \leq \alpha + (f(x_1) - f(x_2))(t - \bar{t}) \qquad && \textit{for every } t \in [0, \bar{t}], \\
&\alpha \leq f(tx_1 + (1-t)x_2) < \alpha + (f(x_1) - f(x_2))(t - \bar{t}) && \textit{for every } t \in (1, \bar{t}, 1].
\end{aligned}
$$

**Proof.** Since $f$ is not convex, there exists $x_1, x_2 \in X$ such that

$$\Delta := \sup_{t \in [0,1]} g(t) > 0, \tag{2.2.1}$$

where $g(t) := f(tx_1 + (1-t)x_2) - tf(x_1) - (1-t)f(x_2)$. Without loss of generality, assume that $f(x_1) \geq f(x_2)$. For every $t \in [0,1]$,

$$g(t) \leq f(tx_1 + (1-t)x_2) - f(x_2) \leq f(x_1) - f(x_2),$$

where the first inequality holds since $f(x_1) \geq f(x_2)$ and the second inequality follows from quasiconvexity of $f$. Hence,

$$0 < \Delta \leq f(x_1) - f(x_2). \tag{2.2.2}$$

For every $\gamma \in [0, \Delta)$, define $t_\gamma := \sup\{t \in [0,1] \colon g(t) \geq \gamma\}$. Obviously $t_\gamma \geq 0$. We claim that

$$t_\gamma \leq \frac{f(x_1) - f(x_2) - \gamma}{f(x_1) - f(x_2)}.$$

Suppose not. Then, there exists $t^* \in \left(\frac{f(x_1)-f(x_2)-\gamma}{f(x_1)-f(x_2)}, t_\gamma\right]$ such that

$$g(t^*) \geq \gamma. \tag{2.2.3}$$

Observe that

$$t^* > \frac{f(x_1) - f(x_2) - \gamma}{f(x_1) - f(x_2)} \quad \iff \quad t^* f(x_1) + (1 - t^*) f(x_2) > f(x_1) - \gamma.$$

Hence,

$$\begin{aligned} g(t^*) &= f(t^* x_1 + (1 - t^*) x_2) - t^* f(x_1) - (1 - t^*) f(x_2) \\ &< f(t^* x_1 + (1 - t^*) x_2) - f(x_1) + \gamma \\ &\leq \max\{f(x_1), f(x_2)\} - f(x_1) + \gamma \\ &= \gamma, \end{aligned}$$

where the inequality follows from quasiconvexity of $f$ while the last equality follows from the assumption that $f(x_1) \geq f(x_2)$. This is in contradiction with (2.2.3). Therefore,

$$0 < t_\gamma \leq \frac{f(x_1) - f(x_2) - \gamma}{f(x_1) - f(x_2)}, \quad \text{for every } \gamma \in [0, \Delta). \tag{2.2.4}$$

Moreover, if $\gamma_1 \leq \gamma_2$, then $\{t \in [0,1] \colon g(t) \geq \gamma_2\} \subseteq \{t \in [0,1] \colon g(t) \geq \gamma_1\}$. Hence,

$$t_{\gamma_1} \geq t_{\gamma_2} \quad \text{for every } \gamma_1 \leq \gamma_2. \tag{2.2.5}$$

Let us define

$$\bar{t} \coloneqq \lim_{\gamma \to \Delta} t_\gamma \quad \text{and} \quad \alpha \coloneqq \Delta + \bar{t} f(x_1) + (1 - \bar{t}) f(x_2).$$

It follows from (2.2.2) and letting $\gamma \to \Delta$ in (2.2.4) that

$$0 \leq \bar{t} \leq \frac{f(x_1) - f(x_2) - \Delta}{f(x_1) - f(x_2)} < 1. \tag{2.2.6}$$

Therefore,

$$\begin{aligned} f(x_2) &< f(x_2) + \Delta + \bar{t}(f(x_1) - f(x_2)) \\ &= \alpha \\ &\le f(x_2) + \Delta + \frac{f(x_1) - f(x_2) - \Delta}{f(x_1) - f(x_2)}(f(x_1) - f(x_2)) = f(x_1), \end{aligned}$$

where the strict inequality holds since $\Delta + \bar{t}(f(x_1) - f(x_2))$ is positive by (2.2.2), (2.2.6), and the inequality follows from (2.2.6). Hence, we have

$$f(x_2) < \alpha \le f(x_1). \tag{2.2.7}$$

By the definition of $\Delta$, for every $t \in [0, 1]$,

$$f(tx_1+(1-t)x_2) \le \Delta+tf(x_1)+(1-t)f(x_2) = \alpha+(t-\bar{t})(f(x_1)-f(x_2)). \tag{2.2.8}$$

Now, assume that the inequality above holds with equality for some $\hat{t} \in [0, 1]$. Then, $g(\hat{t}) = \Delta$. Observe that, $g(\hat{t}) > \gamma$ for every $\gamma \in [0, \Delta)$. By (2.2.5), $\hat{t} \le t_\gamma$ for every $\gamma \in [0, \Delta)$. Letting $\gamma \to \Delta$, we have $\hat{t} \le \bar{t}$. Thus, if $t > \bar{t}$ we have

$$f(tx_1 + (1 - t)x_2) < \alpha + (t - \bar{t})(f(x_1) - f(x_2)). \tag{2.2.9}$$

Summing up, from (2.2.8) and (2.2.9) we have

$$\begin{aligned} f(tx_1 + (1 - t)x_2) &\le \alpha + (t - \bar{t})(f(x_1) - f(x_2)) \quad \text{for every } t \in [0, \bar{t}], \\ f(tx_1 + (1 - t)x_2) &< \alpha + (t - \bar{t})(f(x_1) - f(x_2)) \quad \text{for every } t \in (\bar{t}, 1]. \end{aligned} \tag{2.2.10}$$

Let $t \in (\bar{t}, 1]$. It remains to show that $\alpha \le f(tx_1 + (1-t)x_2)$. By the definition of $\bar{t}$ and (2.2.5), we have $t_\gamma \ge \bar{t}$ for every $\gamma \in [0, \Delta)$. Hence, there exists $\overline{\gamma} \in (0, \Delta)$ such that

$$\bar{t} \le t_\gamma < t \quad \text{for every } \gamma \in [\overline{\gamma}, \Delta). \tag{2.2.11}$$

For all $\gamma \in [\overline{\gamma}, \Delta)$ and $\epsilon \in \left(0, \frac{t_\gamma}{2}\right)$, there exists $t_{\gamma,\epsilon} \in [t_\gamma - \epsilon, t_\gamma]$ such that $g(t_{\gamma,\epsilon}) \ge \gamma$. By the definition of $g$, this is equivalent to

$$f(t_{\gamma,\epsilon}x_1 + (1 - t_{\gamma,\epsilon})x_2) \ge \gamma + f(x_2) + t_{\gamma,\epsilon}(f(x_1) - f(x_2)) > f(x_2). \tag{2.2.12}$$

Note that $t \mapsto f(tx_1+(1-t)x_2)$ is quasiconvex on $[0,1]$ as it is the composition of an affine function with a quasiconvex function. Observe that $t_{\gamma,\epsilon}$ can be written as a convex combination of 0 and $t$ since $0 < t_{\gamma,\epsilon} \leq t_\gamma < t$. Therefore,

$$f(t_{\gamma,\epsilon}x_1 + (1-t_{\gamma,\epsilon})x_2) \leq \max\{f(x_2), f(tx_1+(1-t)x_2)\}.$$

By (2.2.12), $f(t_{\gamma,\epsilon}x_1 + (1-t_{\gamma,\epsilon})x_2) > f(x_2)$. Hence,

$$f(tx_1+(1-t)x_2) \geq f(t_{\gamma,\epsilon}x_1 + (1-t_{\gamma,\epsilon})x_2). \tag{2.2.13}$$

It follows from (2.2.11), (2.2.12) and (2.2.13) that

$$\begin{aligned} f(tx_1+(1-t)x_2) &\geq \gamma + f(x_2) + t_{\gamma,\epsilon}(f(x_1)-f(x_2)) \\ &\geq \gamma + f(x_2) + (t_\gamma - \epsilon)(f(x_1)-f(x_2)) \\ &\geq \gamma + f(x_2) + (\bar{t} - \epsilon)(f(x_1)-f(x_2)) \end{aligned}$$

Letting $\epsilon \to 0$ and $\gamma \to \Delta$ gives

$$f(tx_1+(1-t)x_2) \geq \Delta + f(x_2) + \bar{t}(f(x_1)-f(x_2)) = \alpha. \tag{2.2.14}$$

The result follows from (2.2.7), (2.2.10) and (2.2.14). $\square$

For the rest of this section, we fix $n \in \mathbb{N}$ and let $X_i$ be a topological vector space and $f_i$ be a proper extended real-valued function defined on $X_i$ for each $i \in \{1, 2, \ldots, n\}$. Furthermore, we define the function $s\colon X_1 \times \ldots \times X_n \mapsto \mathbb{R} \cup \{+\infty\}$ by

$$s(x_1, \ldots, x_n) \coloneqq f_1(x_1) + \ldots + f_n(x_n). \tag{2.2.15}$$

In Debreu and Koopmans [16] and Crouzeix and Lindberg [17] implications of quasiconvexity of $s$ on $f_i$ when $X_i$ is an open subset of $\mathbb{R}^n$ is studied. In this section, we extend their results to general topological vector spaces.

**Theorem 2.2.2.** *Assume that $f_1, f_2, \ldots, f_n$ are non-constant. $s$ is quasiconvex* *if and only if*

$$c(f_1) + \ldots + c(f_n) \geq 0$$

**Proof.** It is enough to consider the case $n = 2$ and we write $f = f_1$, $g = f_2$, $X_1 = X$ and $X_2 = Y$. Assume that $s$ is quasiconvex and assume to the contrary that $c(f) + c(g) < 0$. Given $\overline{y} \in Y$, for every $x_1, x_2 \in X$, and $t \in [0, 1]$

$$\begin{aligned} f(tx_1 + (1-t)x_2) + g(\overline{y}) &= s(tx_1 + (1-t)x_2, \overline{y}) \\ &\leq \max\{s(x_1, \overline{y}), s(x_2, \overline{y})\} \\ &= \max\{f(x_1) + g(\overline{y}), f(x_2) + g(\overline{y})\} \\ &= \max\{f(x_1), f(x_2)\} + g(\overline{y}). \end{aligned}$$

Subtracting $g(\overline{y})$ from both sides yields

$$f(tx_1 + (1-t)x_2) \leq \max\{f(x_1), f(x_2)\}. \tag{2.2.16}$$

Therefore, $f$ is quasiconvex. By a similar argument, $g$ is also quasiconvex. Without loss of generality we can assume that $c(f) \leq c(g)$. Then, there exists $\lambda < 0$ such that

$$c(f) < \lambda < -c(g). \tag{2.2.17}$$

For every $(x, y) \in X \times Y$, define

$$\overline{f}(x) := e^{-\lambda f(x)}, \quad \overline{g}(y) := e^{\lambda g(y)} \quad \text{and} \quad \overline{s}(x, y) := e^{-\lambda s(x,y)}.$$

$\overline{s}$ is a quasiconvex function since it is the composition of a quasiconvex function with a non-decreasing function. It follows from Definition 2.1.2 and Remark 2.1.3 that $\overline{f}$ is not convex. Moreover, $-\overline{g}$ is not convex. To see this, assume to the contrary that $\overline{g}$ is concave. Then, $r_\mu$ associated with $g$ is convex for every $\mu < 0$ by Lemma 2.1.1-(iii). Hence, $c(g) = \sup\{\gamma \in \mathbb{R} \colon \gamma \geq 0, r_\gamma \text{ is concave}\}$ by Definition 2.1.2. Since $\overline{g}$ is assumed to be concave, $-\lambda \in \{\gamma \in \mathbb{R} \colon \gamma \geq 0, r_\gamma \text{ is concave}\}$. By Lemma 2.1.1, $c(g) \geq -\lambda$. This is in contradiction with (2.2.17).

Observe that $\overline{f}$ and $-\overline{g}$ are composition of $f$ and $g$ respectively with a non-decreasing function. Therefore $\overline{f}$ and $-\overline{g}$ are quasiconvex.

Next, we will apply Lemma 2.2.1 to $\overline{f}$ and $-\overline{g}$ since they are quasiconvex but not convex functions. There exist $x_1, x_2 \in \operatorname{dom}\overline{f}$, $y_1, y_2 \in \operatorname{dom}\overline{g}$, $\overline{t}, \overline{u} \in [0,1)$ and $\alpha, \beta > 0$ such that

$$\begin{aligned}
&0 < \theta(0) < \alpha \leq \theta(1), && && (2.2.18)\\
&\theta(t) \leq \alpha + m(t-\overline{t}) && \text{if } 0 \leq t \leq \overline{t}, && (2.2.19)\\
&\alpha \leq \theta(t) < \alpha + m(t-\overline{t}) && \text{if } \overline{t} < t \leq 1, && (2.2.20)\\
&\mu(0) < -\beta \leq \mu(1) < 0, && && (2.2.21)\\
&\mu(u) \leq -\beta + n(u-\overline{u}) && \text{if } 0 \leq u \leq \overline{u}, && (2.2.22)\\
&-\beta \leq \mu(u) < -\beta + n(u-\overline{u}) && \text{if } \overline{u} < u \leq 1,, && (2.2.23)
\end{aligned}$$

where

$$\begin{aligned}
\theta(t) &:= \overline{f}(tx_1 + (1-t)x_2) && \text{for every } t \in [0,1],\\
\mu(u) &:= -\overline{g}(uy_1 + (1-u)y_2) && \text{for every } u \in [0,1],\\
m &:= \theta(1) - \theta(0) > 0,\\
n &:= \mu(1) - \mu(0) > 0.
\end{aligned}$$

Define

$$\begin{aligned}
&\xi(t,u) := -\frac{\theta(t)}{\mu(u)} \quad \text{for all } (t,u) \in [0,1]\times[0,1],\\
&S := \left\{(t,u) \in [0,1]\times[0,1] \colon \xi(t,u) < \frac{\alpha}{\beta}\right\},\\
&T := \{(t,u) \in [0,1]\times[0,1] \colon m\beta(t-\overline{t}) + n\alpha(u-\overline{u}) \leq 0\}.
\end{aligned}$$

Notice that

$$\xi(t,u) = e^{-\lambda s(tx_1+(1-t)x_2, uy_1+(1-u)y_2)}.$$

$\xi$ is quasiconvex in $(t, u)$ since $s$ is quasiconvex. As a strict lower level set of a quasiconvex function, $S$ is a convex set. Moreover,

$$S = \Big\{(t, u) \in [0, 1] \times [0, 1]\colon \beta\theta(t) + \alpha\mu(u) < 0\Big\}.$$

From (2.2.19), (2.2.20), (2.2.22) and (2.2.23), we immediately derive

$$S \cap \Big((\bar{t}, 1] \times (\overline{u}, 1]\Big) = \emptyset, \tag{2.2.24}$$

$$S \supseteq T \cap \Big([0, \bar{t}] \times (\overline{u}, 1]\Big), \tag{2.2.25}$$

$$S \supseteq T \cap \Big((\bar{t}, 1] \times [0, \overline{u}]\Big), \tag{2.2.26}$$

There are four possible cases:

(a) Suppose that $\bar{t} > 0$ and $\overline{u} > 0$. Let

$$k := \frac{m\beta}{n\alpha},$$
$$\epsilon := \min\Big\{\bar{t}, \frac{1-\overline{u}}{k}, 1-\bar{t}, \frac{\overline{u}}{k}\Big\}.$$

It is not hard to see that $(\bar{t}-\epsilon, \overline{u}+k\epsilon) \in T\cap\Big([0,\bar{t}]\times(\overline{u},1]\Big)$ and $(\bar{t}+\epsilon, \overline{u}-k\epsilon) \in T\cap\Big((\bar{t},1]\times[0,\overline{u}]\Big)$. In view of (2.2.25) and (2.2.26), $(\bar{t}-\epsilon, \overline{u}+k\epsilon), (\bar{t}+\epsilon, \overline{u}-k\epsilon) \in S$. It follows from convexity of $S$ that

$$(\bar{t}, \overline{u}) = \frac{1}{2}(\bar{t}-\epsilon, \overline{u}+k\epsilon) + (1-\frac{1}{2})(\bar{t}+\epsilon, \overline{u}-k\epsilon) \in S.$$

Hence, $\beta\theta(\bar{t}) + \alpha\mu(\overline{u}) < 0$. On the other hand, by (2.2.19) and (2.2.22), $\beta\theta(\bar{t}) \le \alpha\beta$ and $\alpha\mu(\overline{u}) \le -\alpha\beta$. Then, we have either $\alpha\mu(\overline{u}) < -\alpha\beta$ or $\beta\theta(\bar{t}) < \alpha\beta$.

Assume that the latter holds. Choose $u^* \in (\overline{u}, 1]$ such that

$$u^* - \overline{u} < \frac{\alpha\beta - \beta}{n\alpha},$$

i.e. such that

$$n\alpha(u^* - \overline{u}) < \alpha\beta - \beta\theta(\bar{t}).$$

For instance, one can choose $u^* = \overline{u} + \frac{1}{2}\frac{\alpha\beta - \beta\theta(\bar{t})}{n\alpha}$. It follows from (2.2.23) that

$$\beta\theta(\bar{t}) + \alpha\mu(u^*) < \beta\theta(\bar{t}) - \alpha\beta + n\alpha(u^* - \overline{u}) < 0.$$

Therefore, $(\bar{t}, u^*) \in S$.

Next, note that one can choose $(t^{**}, u^{**}) \in (\bar{t}, 1] \times [0, \overline{u}]$ such that

$$m\beta(t^{**} - \bar{t}) + n\alpha(u^{**} - \overline{u}) \leq 0.$$

Hence, $(t^{**}, u^{**}) \in T \cap \Big((\bar{t}, 1] \times [0, \overline{u}]\Big)$, and consequently $(t^{**}, u^{**}) \in S$ by (2.2.26). Notice that, since $\bar{t} < t^{**}$ and $u^{**} < \overline{u} < u^*$, there exists $\overline{\lambda} \in (0, 1)$ such that

$$(\overline{\lambda}t^{**} + (1 - \overline{\lambda})\bar{t}, \overline{\lambda}u^{**} + (1 - \overline{\lambda})u^*) \in (\bar{t}, 1] \times (\overline{u}, 1].$$

Moreover, $S$ is convex. Therefore,

$$(\overline{\lambda}t^{**} + (1 - \overline{\lambda})\bar{t}, \overline{\lambda}u^{**} + (1 - \overline{\lambda})u^*) \in S.$$

This is in contradiction with (2.2.24).

Lastly, assume $\alpha\mu(\overline{u}) < -\alpha\beta$. Observe that there exists $t^* \in (\bar{t}, 1]$ such that $m\beta(t^* - \bar{t}) < -\alpha\beta - \alpha\mu(\overline{u})$. Hence,

$$\beta\theta(t^*) + \alpha\mu(\overline{u}) < \alpha\beta + m\beta(t^* - \bar{t}) + \alpha\mu(\overline{u}) < 0,$$

where the first strict inequality follows from (2.2.20). Therefore $(t^*, \overline{u}) \in S$. Next, we can choose $(t^{**}, u^{**}) \in [0, \bar{t}] \times (\overline{u}, 1]$ such that

$$m\beta(t^{**} - \bar{t}) + n\alpha(u^{**} - \overline{u}) \leq 0.$$

Therefore, $(t^{**}, u^{**}) \in T \cap [0, \bar{t}] \times (\overline{u}, 1]$. By (2.2.25), $(t^{**}, u^{**}) \in S$. Observe that, since $t^{**} < \bar{t} < t^*$ and $\overline{u} < u^{**}$. Therefore, there exists $\overline{\lambda} \in (0, 1)$ such that

$$\big(\overline{\lambda}t^{**} + (1 - \overline{\lambda})t^*, \overline{\lambda}u^{**} + (1 - \overline{\lambda})\overline{u}\big) \in (\bar{t}, 1] \times (\overline{u}, 1].$$

Moreover $S$ is convex. Hence,

$$(\overline{\lambda}t^{**} + (1-\overline{\lambda})\bar{t}, \overline{\lambda}u^{**} + (1-\overline{\lambda})u^{*}) \in S.$$

This is a contradiction by (2.2.24).

(b) Suppose that $\bar{t} = 0$ and $\overline{u} > 0$. By (2.2.18), we have $\alpha - \theta(0) > 0$. Choose $u^* \in (\overline{u}, 1]$ such that

$$n\alpha(u^* - \overline{u}) < \beta(\alpha - \theta(0)).$$

For instance, one can choose $u^* = \overline{u} + \frac{1}{2}\frac{\beta(\alpha-\theta(0))}{n\alpha}$. Observe that

$$\beta\theta(0) + \alpha\mu(u^*) < \beta\theta(0) - \beta\alpha + n\alpha(u^* - \overline{u}) < 0.$$

where the first strict inequality follows from (2.2.23). By the definition of $S$, $(0, u^*) \in S$. Choose $(t^{**}, u^{**}) \in (\bar{t}, 1] \times [0, \overline{u}]$ such that

$$m\beta(t^{**} - \bar{t}) + n\alpha(u^{**} - \overline{u}) \leq 0.$$

In other words, $(t^{**}, u^{**}) \in T$. By (2.2.26), $(t^{**}, u^{**}) \in S$. Since $S$ is convex and $(0, u^*), (t^{**}, u^{**}) \in S$, for any $\lambda \in [0, 1]$

$$(\lambda t^{**} + (1-\lambda)0, \lambda u^{**} + (1-\lambda)u^*) \in S.$$

On the other hand, $0 = \bar{t} < t^{**}$ and $u^{**} \leq \overline{u} < u^*$. Therefore, there exists $\overline{\lambda} \in (0, 1)$ such that

$$(\overline{\lambda}t^{**} + (1-\overline{\lambda})0, \overline{\lambda}u^{**} + (1-\overline{\lambda})u^*) \in (\bar{t}, 1] \times (\overline{u}, 1].$$

This is a contradiction by (2.2.24).

(c) Suppose that $\bar{t} > 0$ and $\overline{u} = 0$. By (2.2.21), we have $\mu(0) + \beta < 0$. Choose $t^* \in (\bar{t}, 1]$ such that

$$m\beta(t^* - \bar{t}) < -\alpha(\beta + \mu(0)).$$

For instance, one can choose $t^* = \bar{t} - \frac{1}{2}\frac{\alpha(\beta-\mu(0))}{m\beta}$. Observe that

$$\beta\theta(t^*) + \alpha\mu(0) < \alpha\beta + m\beta(t^* - \bar{t}) + \alpha\mu(0) < 0.$$

where the first strict inequality follows from (2.2.20). By the definition of $S$, $(t^*, 0) \in S$. Choose $(t^{**}, u^{**}) \in [0, \bar{t}] \times (\overline{u}, 1]$ such that

$$m\beta(t^{**} - \bar{t}) + n\alpha(u^{**} - \overline{u}) \leq 0.$$

In other words, $(t^{**}, u^{**}) \in T$. By (2.2.25), $(t^{**}, u^{**}) \in S$. Since $S$ is convex and $(t^*, 0), (t^{**}, u^{**}) \in S$, for any $\lambda \in [0, 1]$

$$(\lambda t^{**} + (1-\lambda)t^*, \lambda u^{**} + (1-\lambda)0) \in S.$$

On the other hand, $t^{**} \leq \bar{t} < t^*$ and $0 = \overline{u} < u^{**}$. Therefore, there exists $\overline{\lambda} \in (0, 1)$ such that

$$(\overline{\lambda} t^{**} + (1-\overline{\lambda})t^*, \overline{\lambda} u^{**} + (1-\overline{\lambda})0) \in (\bar{t}, 1] \times (\overline{u}, 1].$$

This is a contradiction by (2.2.24).

(d) Suppose that $\bar{t} = \overline{u} = 0$. By (2.2.18) and (2.2.21), we have $\mu(0) + \beta < 0$ and $\alpha - \theta(0) > 0$. Choose $t^*, u^* \in (0, 1]$ such that

$$\begin{aligned} \beta m t^* &< -\alpha(\beta + \mu(0)), \\ \alpha n u^* &< \beta(\alpha - \theta(0)). \end{aligned}$$

For instance, choose $t^* = \frac{1}{2}\frac{\beta(\alpha-\mu(0))}{\alpha n}$ and $u^* = -\frac{1}{2}\frac{\alpha(\beta+\mu(0))}{\beta m}$. Observe that,

$$\begin{aligned} \beta\theta(t^*) + \alpha\mu(0) &< \alpha\beta + \beta m(t^* - \bar{t}) + \alpha\mu(0) < 0, \\ \beta\theta(0) + \alpha\mu(u^*) &< -\beta\alpha + \alpha n(u - \overline{u}) + \beta\theta(0) < 0, \end{aligned}$$

where the first strict inequalities on the two lines follow from (2.2.20) and (2.2.23), respectively. By the definition of $S$, $(t^*, 0) \in S$ and $(0, u^*) \in S$. Since $S$ is convex, $(\lambda t^*, (1-\lambda)u^*) \in S$ for any $\lambda \in (0, 1)$. On the other hand,

$(\lambda t^*, (1-\lambda)u^*) \in (\bar{t}, 1] \times (\overline{u}, 1]$ which is a contradiction by (2.2.24).

Since (a), (b), (c) and (d) all lead to a contradiction, $c(f) + c(g) \geq 0$.

Conversely, assume that $c(f) + c(g) \geq 0$. Since $f$ and $g$ are non-constant, by Theorem 2.1.6 $c(f)$ and $c(g)$ are finite. If $c(f) \geq 0$ and $c(g) \geq 0$, by Theorem 2.1.4, $f, g$ are convex. As the sum of two convex functions $s$ is convex, hence it is also quasiconvex. If $c(f)$ and $c(g)$ are not both positive, then by symmetry it is enough to consider the case $c(f) < 0 < c(g)$. Define,

$$\psi(x) := e^{c(g)f(x)} \qquad \text{for every } x \in X,$$
$$\zeta(y) := e^{-c(g)g(y)} \qquad \text{for every } y \in Y.$$

It follows from Definition 2.1.2 and Lemma 2.1.1-(i) that $x \mapsto e^{-c(f)f(x)}$ is convex. Since $-c(g) \leq c(f)$, by Lemma 2.1.1-(i), $\psi$ is convex. Moreover, $\zeta$ is concave directly from the Definition 2.1.2 and Lemma 2.1.1-(ii). Let $\rho$ on $X \times Y$ defined by

$$\rho(x, y) := e^{c(g)[f(x)+g(y)]} = \frac{\psi(x)}{\zeta(y)}.$$

Notice that for every $(x_1, y_1), (x_2, y_2) \in X \times Y$ and $\lambda \in [0, 1]$

$$\rho\big(\lambda(x_1, y_1) + (1-\lambda)(x_2, y_2)\big) = \frac{\psi(\lambda x_1 + (1-\lambda)x_2)}{\zeta(\lambda y_1 + (1-\lambda)y_2)} \leq \frac{\lambda\psi(x_1) + (1-\lambda)\psi(x_2)}{\lambda\zeta(y_1) + (1-\lambda)\zeta(y_2)} \tag{2.2.27}$$

where the inequality follows from convexity of $\psi$ and concavity of $\zeta$. Now, consider the following two cases:

Suppose that $\frac{\psi(x_2)}{\zeta(y_2)} \leq \frac{\psi(x_1)}{\zeta(y_1)}$. Then,

$$\begin{aligned}
&\psi(x_2)\zeta(y_1) \leq \psi(x_1)\zeta(y_2)\\
\Longrightarrow\ &(1-\lambda)\psi(x_2)\zeta(y_1) \leq (1-\lambda)\psi(x_1)\zeta(y_2)\\
\Longrightarrow\ &\lambda\psi(x_1)\zeta(y_1) + (1-\lambda)\psi(x_2)\zeta(y_1) \leq \lambda\psi(x_1)\zeta(y_1) + (1-\lambda)\psi(x_1)\zeta(y_2)\\
\Longrightarrow\ &\zeta(y_1)[\lambda\psi(x_1) + (1-\lambda)\psi(x_2)] \leq \psi(x_1)[\lambda\zeta(y_1) + (1-\lambda)\zeta(y_2)]\\
\Longrightarrow\ &\frac{\lambda\psi(x_1) + (1-\lambda)\psi(x_2)}{\lambda\zeta(y_1) + (1-\lambda)\zeta(y_2)} \leq \frac{\psi(x_1)}{\zeta(y_1)}.
\end{aligned}$$

Therefore,

$$\frac{\lambda\psi(x_1) + (1-\lambda)\psi(x_2)}{\lambda\zeta(Y_1) + (1-\lambda)\zeta(y_2)} \leq \rho(x_1, y_1). \tag{2.2.28}$$

Otherwise, $\frac{\psi(x_2)}{\zeta(y_2)} > \frac{\psi(x_1)}{\zeta(y_1)}$ holds. A similar argument with the previous case gives

$$\frac{\lambda\psi(x_1) + (1-\lambda)\psi(x_2)}{\lambda\zeta(y_1) + (1-\lambda)\zeta(y_2)} < \rho(x_2, y_2). \tag{2.2.29}$$

By (2.2.27), (2.2.28) and (2.2.29)

$$\rho(\lambda(x_1, y_1) + (1-\lambda)(x_2, y_2)) \leq \max\{\rho(x_1, y_1), \rho(x_2, y_2)\}.$$

Therefore, $\rho$ is quasiconvex. Furthermore,

$$\log\big(\rho(x, y)\big) = c(g)[f(x) + g(y)] = c(g)s(x, y)$$

is quasiconvex since quasiconvexity is stable under composition with a non-decreasing function. Hence, $s$ is quasiconvex. $\square$

The following lemma is needed in the proof of Theorem 2.2.4.

**Lemma 2.2.3.** *Let $\rho_1, \rho_2$ be positive real-valued lower semi-continuous and convex functions defined on the topological vector spaces $X$ and $Y$, respectively. Define $\rho$ on $X \times Y$ by*

$$\rho(x, y) := [\rho_1(x)]^{\alpha}[\rho_2(y)]^{\beta},$$

*where $\alpha, \beta > 0$ and $\alpha + \beta = 1$. Assume that $\rho$ is concave. Then, at least one of the functions $\rho_1$ and $\rho_2$ is concave.*

**Proof.** Assume to the contrary that both $\rho_1$ and $\rho_2$ are not concave. Choose $\overline{y} \in Y$ such that $\rho_2(\overline{y}) > 0$. It follows from concavity of $\rho$ that for every $x_1, x_2 \in X$ and $\lambda \in [0, 1]$,

$$\rho(\lambda x_1 + (1-\lambda)x_2, \overline{y}) \geq \lambda\rho(x_1, \overline{y}) + (1-\lambda)\rho(x_2, \overline{y}).$$

Dividing both sides by $[\rho_2(\overline{y})]^\beta$ yields

$$[\rho_1(\lambda x + (1-\lambda)x_2)]^\alpha \geq \lambda[\rho_1(x_1)]^\alpha + (1-\lambda)[\rho_1(x_2)]^\alpha. \tag{2.2.30}$$

Note that $\lambda[\rho_1(x_1)]^\alpha + (1-\lambda)[\rho_1(x_2)]^\alpha \geq \min\{[\rho_1(x_1)]^\alpha, [\rho_1(x_2)]^\alpha\} = \min\{\rho_1(x_1), \rho_1(x_2)\}^\alpha$. Therefore,

$$[\rho_1(\lambda x_1 + (1-\lambda)x_2)]^\alpha \geq \min\{\rho_1(x_1), \rho_1(x_2)\}^\alpha,$$

which implies

$$\rho_1(\lambda x_1 + (1-\lambda)x_2) \geq \min\{\rho_1(x_1), \rho_1(x_2)\}. \tag{2.2.31}$$

The supposition together with (2.2.31) implies that $-\rho_1$ is quasiconvex but not convex. According to Lemma 2.2.1, there exist $x_1, x_2 \in X$, $\bar{t} \in [0,1)$ and $\gamma \in [0,1]$ such that

$$\theta(0) > \gamma \geq \theta(1) > 0, \tag{2.2.32}$$

$$\theta(t) \geq \gamma - m(t-\bar{t}) \qquad \text{if } 0 \leq t \leq \bar{t}, \tag{2.2.33}$$

$$\gamma \geq \theta(t) > \gamma - m(t-\bar{t}) \qquad \text{if } \bar{t} < t \leq 1, \tag{2.2.34}$$

where $\theta(t) := \rho_1(tx_1 + (1-t)x_2)$ for all $t \in [0,1]$ and $m := \theta(0) - \theta(1) > 0$.

The function $\overline{\theta}(t)\colon t \mapsto [\theta(t)]^\alpha$ is concave on $[0,1]$ by (2.2.30) and the fact that the composition of a concave function with an affine function is concave. As a real-valued concave function defined on [0,1], $\overline{\theta}$ is continuous on $(0,1)$ and have finite left, right derivatives. Consequently, $\theta = \overline{\theta}^{\frac{1}{\alpha}}$ is continuous as it is the composition of two continuous. It is not hard to see that

$$\begin{aligned} \overline{\theta}'_+(t) &= \alpha(\theta(t))^{\alpha-1}\theta'_+(t), \\ \overline{\theta}'_-(t) &= \alpha(\theta(t))^{\alpha-1}\theta'_-(t). \end{aligned} \tag{2.2.35}$$

Hence, $\theta$ also have finite left and right derivatives. Moreover,by (2.2.32) we have

$$\overline{\theta}(0) > \gamma^\alpha. \tag{2.2.36}$$

We claim that $\bar{t} \neq 0$. Suppose not; hence by (2.2.33)

$$\begin{aligned}
&\gamma \geq \theta(t) \quad \text{for every } t \in (0,1] \\
\Longrightarrow\ &\gamma^{\alpha} \geq \overline{\theta}(t) \quad \text{for every } t \in (0,1] \\
\Longrightarrow\ &\gamma^{\alpha} \geq \overline{\theta}(0),
\end{aligned}$$

where the last implication holds since $\overline{\theta}$ is right-continuous at 0. This is in contradiction with (2.2.36); therefore, $\bar{t} \neq 0$, and consequently $\overline{\theta}$ is continuous at $\bar{t}$. It results from (2.2.33) and (2.2.34) that $\theta(\bar{t}) = \gamma$ and $\theta'_{-}(\bar{t}) \leq -m \leq \theta'_{+}(\bar{t})$. Furthermore, $\overline{\theta}'_{+}(\bar{t}) \leq \overline{\theta}'_{-}(\bar{t})$ since $\overline{\theta}$ is concave. In view of (2.2.35), $\theta'_{+}(\bar{t}) = \theta'_{-}(\bar{t}) = -m$. Thus, $\theta$ has derivative at $\bar{t}$ and $\theta'(\bar{t}) < 0$. It follows from (2.2.34) that

$$\theta(t) > \theta(\bar{t}) + \theta'(\bar{t})(t - \bar{t}) \quad \text{if } \bar{t} < t \leq 1. \tag{2.2.37}$$

Reasoning as above, we can prove that there exists $y_1, y_2 \in Y$, $\overline{u} \in (0,1)$ such that $\mu'(\overline{u}) < 0$ and

$$\mu(u) > \mu(\overline{u}) + \mu'(\overline{u})(u - \overline{u}) \quad \text{if } \overline{u} < u \leq 1. \tag{2.2.38}$$

where

$$\mu(u) := \rho_2(uy_1 + (1-u)y_2).$$

Set

$$\zeta(t,u) := [\theta(t)]^{\alpha}[\mu(u)]^{\beta} \quad \text{for every } t, u \in [0,1].$$

For every $\delta > 0$, define

$$\begin{aligned}
h &:= -\delta \frac{\theta(\bar{t})}{\theta'(\bar{t})}, \\
k &:= -\delta \frac{\mu(\overline{u})}{\mu'(\overline{u})}.
\end{aligned}$$

Next, choose $\delta$ so that $0 < h < 1 - \bar{t}$ and $0 < k < 1 - \overline{u}$. For instance, one can choose $\delta = \frac{1}{2}\min\big\{-\frac{\theta'(\bar{t})}{\theta(\bar{t})}(1-\bar{t}), -\frac{\mu'(\overline{u})}{\mu(\overline{u})}(1-\overline{u})\big\}$. By (2.2.34) and (2.2.38), we have

$$\theta(\bar{t}+h) > \theta(\bar{t}) + \theta'(\bar{t})h = \theta(\bar{t})[1-\delta]$$
$$\mu(\overline{u}+k) > \mu(\overline{u}) + \mu'(\overline{u})k = \mu(\overline{u})[1-\delta]$$

Hence,

$$\zeta(\bar{t}+h, \overline{u}+k) > [\theta(\bar{t})]^{\alpha}[\mu(\overline{u})]^{\beta}[1-\delta]. \tag{2.2.39}$$

On the other hand, for every $\lambda \in [0,1]$,

$$\begin{aligned}
\zeta(\bar{t}+\lambda h, \overline{u}+\lambda k) &= \zeta\big((\bar{t},\overline{u}) + \lambda(h,k)\big) \\
&= \zeta\big(\lambda(\bar{t}+h,\overline{u}+k) + (1-\lambda)(\bar{t},\overline{u})\big) \\
&\geq \lambda\zeta(\bar{t}+h,\overline{u}+k) + (1-\lambda)\zeta(\bar{t},\overline{u}) \\
&= \zeta(\bar{t},\overline{u}) + \lambda[\zeta(\bar{t}+h,\overline{u}+k) - \zeta(\bar{t},\overline{u})],
\end{aligned}$$

where the inequality follows from concavity of $\zeta$ on $[0,1]\times[0,1]$. With some algebra, we get

$$\zeta(\bar{t},\overline{u}) + \frac{\zeta\big((\bar{t},\overline{u}) + \lambda(h,k)\big) - \zeta(\bar{t},\overline{u})}{\lambda} \geq \zeta(\bar{t}+h,\overline{u}+k)$$

Moreover, directional derivative of $\zeta$ along the direction $(h,k)$ and $(\bar{t},\overline{u})$ exists. By letting $\lambda \to 0$, we get

$$\zeta(\bar{t},\overline{u}) + \frac{\partial\zeta}{\partial t}(\bar{t},\overline{u})h + \frac{\partial\zeta}{\partial u}(\bar{t},\overline{u})k \geq \zeta(\bar{t}+h,\overline{u}+k).$$

Plugging in the values of the partial derivatives yields

$$\zeta(\bar{t},\overline{u})\Big[1 + \alpha h\frac{\theta'(\bar{t})}{\theta(\bar{t})} + \beta k\frac{\mu'(\overline{u})}{\mu(\overline{u})}\Big] \geq \zeta(\bar{t}+h,\overline{u}+k),$$

It follows from the defintion of $h$, $k$ and $\zeta$ that

$$[\theta(\bar{t})]^{\alpha}[\mu(\overline{u})]^{\beta}[1-\delta] \geq \zeta(\bar{t}+h,\overline{u}+k),$$

which is in contradiction with (2.2.39). Therefore, at least one of the functions $\rho_1$ and $\rho_2$ is concave. $\square$

The next proposition gives an additivity formula for the convexity index.

**Proposition 2.2.4.** *Assume that $f_i$ is lower semi-continuous and convex for each $i \in \{1, 2, \dots n\}$. Then,*

$$\frac{1}{c(s)} = \sum_{i=1}^{n} \frac{1}{c(f_i)}, \tag{2.2.40}$$

*where $s$ is defined as in* (2.2.15).

**Proof.** It is enough to consider the following three cases with $n = 2$. Note that, since $s$ is convex, by Theorem 2.1.4 implies that $c(s) \geq 0$.

(a) Suppose that $c(f_1) = 0$ and $c(f_2) \in [0, +\infty)$. By Definition 2.1.2, $r_\lambda$ associated with $f_1$ is not concave for any $\lambda > 0$. Suppose for a contradiction that there exists $\overline{\lambda} > 0$ such that $r_{\overline{\lambda}}$ associated with $s$ is concave. Then, for every $(x_1, x_2), (x_1', x_2) \in X_1 \times \operatorname{dom} f_2$, and $t \in [0, 1]$

$$e^{-\overline{\lambda}[f_1(tx_1+(1-t)x_1')+f_2(x_2)]} \geq te^{-\overline{\lambda}[f_1(x_1)+f_2(x_2)]} + (1-t)e^{-\overline{\lambda}[f_1(x_1')+f_2(x_2)]}.$$

Dividing both sides by $e^{-\overline{\lambda} f_2(x_2)}$ yields

$$e^{-\overline{\lambda} f_1(\eta x_1+(1-\eta)x_1')} \geq \eta e^{-\overline{\lambda} f_1(x_1)} + (1-\eta)e^{-\overline{\lambda} f_1(x_1')},$$

where $\eta := \frac{t}{e^{-\overline{\lambda} f_2(x_2}}$. In other words, $r_{\overline{\lambda}}$ associated with $f_1$ is concave. This is in contradiction with the assumption $c(f_1) = 0$. Therefore, $r_\lambda$ associated with $s$ is not concave for every $\lambda > 0$. By Definition 2.1.2, $c(s) = 0$ and (2.2.40) holds.

(b) Suppose that $c(f_1) = +\infty$ and $c(f_2) \in [0, +\infty]$. By Theorem 2.1.6, $f_1 \equiv K$, where $K \in \mathbb{R} \cup \{+\infty\}$ is a constant. If $K = +\infty$, then $s(x_1, x_2) = +\infty$ for every $(x_1, x_2) \in X_1 \times X_2$. Therefore, (2.2.40) holds trivially. If $c(s) = +\infty$, then $f_1, f_2$ are constant. Hence, (2.2.40) holds trivially. So assume $K > +\infty$ and $c(f_2) < +\infty$. Observe that

$$e^{-c(f_2)s(x_1,x_2)} = e^{-c(f_2)K} e^{-c(f_2)f_2(x_2)}, \quad (x_1, x_2) \in X_1 \times X_2.$$

$e^{-c(f_2)K} > 0$ is a constant and $x_2 \mapsto e^{-c(f_2)f_2(x_2)}$ is concave on $X_2$. Therefore, $r_{c(f_2)}$ associated with $s$ is concave. Definition 2.1.2 together with Lemma 2.1.1 implies that $c(f_2) \le c(s)$.

On the other hand, $e^{-c(s)s}$ is concave by Definition 2.1.2. Moreover,

$$e^{-c(s)s} = e^{-c(s)K} e^{-c(s)f_2}$$

where $e^{-c(s)K} > 0$ is a constant. Therefore, $e^{-c(s)f_2}$ is concave. It follows from Definition 2.1.2 and Lemma 2.1.1 that $c(s) \le c(f_2)$. Hence, $c(s) = c(f_2)$ and (2.2.40) holds.

(c) Suppose that $0 < c(f_i) < +\infty$ for $i = 1, 2$. Set $\alpha_i = \frac{1}{c(f_i)}$ for $i = 1, 2$ and $\alpha = \alpha_1 + \alpha_2$. Then,

$$e^{-\frac{1}{\alpha}s(x_1,x_2)} = \left(e^{-c(f_1)f_1(x_1)}\right)^{\frac{\alpha_1}{\alpha}} \left(e^{-c(f_2)f_2(x_2)}\right)^{\frac{\alpha_2}{\alpha}}.$$

Note that $x_i \mapsto e^{-c(f_i)f_i(x_i)}$ is concave on $X_i$ for $i = 1, 2$ and $g(u, v) = u^{\frac{\alpha_1}{\alpha}} v^{\frac{\alpha_2}{\alpha}}$ is non-decreasing in each argument, concave on $\mathbb{R}_+ \times \mathbb{R}_+$. Therefore, $(x_1, x_2) \mapsto e^{-\frac{1}{\alpha}s(x_1,x_2)}$ is concave on $X_1 \times X_2$ as it is the composition of a concave functions with a function which is non-decreasing in each argument. In view of Definition 2.1.2, we have

$$\frac{1}{\alpha} \le c(s) \iff \frac{1}{c(f_1)} + \frac{1}{c(f_2)} \ge \frac{1}{c(s)} \tag{2.2.41}$$

Next, assume that $\mu \in [0, \alpha)$. We have,

$$e^{-\frac{1}{\mu}f_i(x_i)} = \left(e^{-\frac{\alpha}{\mu}f_i(x_i)}\right)^{\frac{1}{\alpha}} = \left(e^{-\frac{\alpha}{\mu}c(f_i)f_i(x_i)}\right)^{\frac{\alpha_i}{\alpha}}, \quad x_i \in X_i$$

for each $i \in \{1, 2\}$. Since $\frac{\alpha}{\mu} > 1$, it follows from Definition 2.1.2 that $x_i \mapsto e^{-\frac{\alpha}{\mu}c(f_i)f_i}$ is not concave for each $i \in \{1, 2\}$. Moreover,

$$e^{-\frac{1}{\mu}s(x_1,x_2)} = \left(e^{-\frac{\alpha}{\mu}c(f_1)f_1(x_1)}\right)^{\frac{\alpha_1}{\alpha}} \left(e^{-\frac{\alpha}{\mu}c(f_2)f_2(x_2)}\right)^{\frac{\alpha_2}{\alpha}}, \quad (x_1, x_2) \in X_1 \times X_2.$$

In view of Lemma 2.2.3, $e^{-\frac{1}{\mu}s}$ is not concave and consequently $c(s) < \frac{1}{\mu}$.

Letting $\mu \to \alpha$, we have $c(s) \leq \frac{1}{\alpha}$. Thus,

$$\frac{1}{c(f_1)} + \frac{1}{c(f_2)} \leq \frac{1}{c(s)} \tag{2.2.42}$$

Summing up, (2.2.41) and (2.2.42) imply (2.2.40).

□

Now, we are ready to give the main result of this section.

**Theorem 2.2.5.** *Assume that $f_i$ is a non-constant lower semi-continuous function for each $i \in \{1, 2, \ldots, n\}$. Then, $s$ is quasiconvex if and only if one of the following conditions holds:*

*(i)* $f_1, f_2, \ldots, f_n$ *are convex.*

*(ii)* *All of* $f_1, f_2, \ldots f_n$ *except one are convex and*

$$\frac{1}{c(f_1)} + \frac{1}{c(f_2)} + \ldots + \frac{1}{c(f_n)} \leq 0. \tag{2.2.43}$$

**Proof.** Assume that $s$ is quasiconvex. Then, $f_1, \ldots, f_n$ are quasiconvex. We claim that at most one of the functions $f_1, \ldots f_n$ is not convex. Suppose for a contradiction that $f_{i_1}, \ldots, f_{i_k}$ are not convex for some $k \geq 2$ so that from Theorem 2.1.4 we deduce $c(f_{i_j}) < 0$ for each $j \in \{1, \ldots k\}$. Thus

$$c(f_{i_1}) + \ldots + c(f_{i_k}) < 0. \tag{2.2.44}$$

Next, define $\overline{s}$ on $X_{i_1} \times \ldots \times X_{i_k}$ by

$$\overline{s}(x_{i_1}, \ldots, x_{i_k}) := f_{i_1}(x_{i_1}) + \ldots + f_{i_k}(x_{i_k}).$$

Clearly, $\overline{s}$ is quasiconvex. Then, by Theorem 2.2.2,

$$c(f_{i_1}) + \ldots + c(f_{i_k}) \geq 0,$$

which is in contradiction with (2.2.44). Hence, the claim holds. If $f_1, f_2, \ldots, f_n$ are convex, then we are done. Suppose that $f_m$ is not convex for some $m \in \{1, \ldots, n\}$ and define $\hat{s}$ on $\bigtimes_{i\in\{1,\ldots,n\}\setminus\{m\}} X_i$ by

$$\hat{s}(x_1, \ldots, x_{m-1}, x_{m+1}, \ldots, x_n) = \sum_{i\in\{1,\ldots,n\}\setminus\{m\}} f_i(x_i).$$

By Theorem 2.2.2,

$$c(\hat{s}) + c(f_m) \geq 0.$$

Since $c(f_m) < 0$, we can write it as

$$\begin{aligned}
&\frac{1}{c(\hat{s})} \leq -\frac{1}{c(f_m)} \\
\Longrightarrow &\frac{1}{c(\hat{s})} + \frac{1}{c(f_m)} \leq 0 \\
\Longrightarrow &\frac{1}{c(f_1)} + \frac{1}{c(f_2)} + \ldots + \frac{1}{c(f_n)} \leq 0,
\end{aligned}$$

where the last results follows from Proposition 2.2.4.

Conversely, assume that either $f_1, f_2, \ldots f_n$ are convex or all of $f_1, f_2, \ldots f_n$ except one are convex and (2.2.43) holds. If the former holds, then $s$ is quasiconvex trivially. Suppose that the latter holds together with $f_m$ is not convex for some $m \in \{1, \ldots, n\}$ and $\hat{s}$ is defined as above. Proposition 2.2.4 together with (2.2.43) implies that

$$\frac{1}{c(\hat{s})} + \frac{1}{c(f_m)} \leq 0.$$

Since $c(f_m) < 0$, we can write it as

$$\begin{aligned}
&\frac{1}{c(\hat{s})} \leq -\frac{1}{c(f_m)} \\
\Longrightarrow &\, c(\hat{s}) \geq -c(f_m) \\
\Longrightarrow &\, c(\hat{s}) + c(f_m) \geq 0.
\end{aligned}$$

In view of Theorem 2.2.2, $s$ is quasiconvex. $\square$

**Remark 2.2.6.** The only if part of Theorem 2.2.5 still holds if we drop the assumption of $f_i$ being non-constant for each $i \in \{1, \dots, n\}$.

## 2.3 Infinite Decomposable Sums

The aim of this section is to extend Theorem 2.2.5 to infinite decomposable sums. Similar to the notation of the previous sections, $X_i$ is a topological vector spaces and $f_i$ is an extended real-valued function defined on $X_i$ for each $i \in \mathbb{N}$. For every $x \in X$, let

$$s_n(x) \coloneqq \sum_{i=1}^{n} f_i(x_i),$$
$$s(x) \coloneqq \lim_{n\to\infty} s_n(x),$$

where $X = \bigtimes_{i=1}^{\infty} X_i$, provided that the limit exists. It can be easily noticed that the lower semi-continuity of $f_i$ on $X_i$ for $i = 1, \dots, n$ implies lower semi-continuity of $s_n$ on $X$. This is due to the fact that $f_i$ is also lower semi-continuous on $X$ and lower semi-continuity is preserved under finite sums. Before presenting the main result of this section, we give two separate conditions under which lower semi-continuity is preserved under infinite sums.

**Lemma 2.3.1.** *Assume that $f_i$ is a lower semi-continuous function for each $i \in \mathbb{N}$ and either one of the following holds:*

*(i) $f_i$ is finite for each $i \in \mathbb{N}$ and $(s_n)_{n\in\mathbb{N}}$ is uniformly convergent.*

*(ii) $f_i \geq 0$ for each $i \in \mathbb{N}$.*

*Then, $s$ is lower semi-continuous.*

**Proof.** Assume that $(i)$ holds. Let $x_0 \in X$. Note that $s_n(x_0) \in \mathbb{R}$ since each $f_i$

is finite. Thus, we have

$$\begin{aligned} s(x_0) &= s(x_0) - s_n(x_0) + s_n(x_0) \\ &\leq s_n(x_0) + |s(x_0) - s_n(x_0)| \\ &\leq s_n(x_0) + \sup_{x \in X} |s(x) - s_n(x)|. \end{aligned} \tag{2.3.1}$$

It follows from lower semi-continuity of $s_n$ at $x_0$ together with (2.3.1) that

$$s(x_0) \leq \sup_{x \in X} |s(x) - s_n(x)| + \liminf_{x \to x_0} s_n(x). \tag{2.3.2}$$

Moreover, similar to (2.3.1), for every $x' \in X$ and $n \in \mathbb{N}$,

$$s_n(x') \leq \sup_{x \in X} |s(x) - s_n(x)| + s(x').$$

By taking limit infimums of both sides as $x' \to x_0$, we get

$$\liminf_{x' \to x_0} s_n(x') \leq \sup_{x \in X} |s(x) - s_n(x)| + \liminf_{x' \to x_0} s(x').$$

Together with (2.3.2), we have

$$s(x_0) \leq 2 \sup_{x \in X} |s(x) - s_n(x)| + \liminf_{x \to x_0} s(x).$$

It follows from uniform convergence of $(s_n)_{n \in \mathbb{N}}$ that

$$s(x_0) \leq \liminf_{x \to x_0} s(x).$$

Now, assume that $(ii)$ holds. For every $r \geq 0$ and $n \in \mathbb{N}$, define

$$\begin{aligned} U_r &= \{x \in X \colon s(x) > r\}, \\ U_r^n &= \{x \in X \colon s_n(x) > r\}. \end{aligned}$$

Fix $r \in \mathbb{R}$. We claim that

$$U_r = \bigcup_{n=1}^{\infty} U_r^n. \tag{2.3.3}$$

Let $\overline{x} \in U_r$. Observe that, $(s_n(\overline{x}))_{n \in \mathbb{N}}$ is a non-decreasing sequence in $\mathbb{R}$ since $f_i$

is a non-negative function for each $i \in \mathbb{N}$. Hence, there exists $N \in \mathbb{N}$ such that $s_N(\overline{x}) > r$. Then, $\overline{x} \in U_r^N$, and consequently, $U_r \subseteq \bigcup_{n=1}^{\infty} U_r^n$.

Next, for any given $n \in \mathbb{N}$, let $x' \in \bigcup_{n=1}^{\infty} U_r^n$. Then, there exists $N' \in \mathbb{N}$ such that $x' \in U_r^{N'} \subseteq U_r$ since $(s_n(x'))_{n\in\mathbb{N}}$ is a non-decreasing sequence in $\mathbb{R}$. Hence, $\bigcup_{n=1}^{\infty} U_r^n \subseteq U_r$, and consequently, the claim holds.

Moreover, $U_r^n$ is open since $s_n$ is lower semi-continuous on $X$ for each $n \in \mathbb{N}$. It follows from (2.3.3) that $s$ is lower semi-continuous on $X$. □

**Remark 2.3.2.** The converse implication of Lemma 2.3.1 is also valid for a more general setting. To be more specific, assume $s$ is lower semi-continuous and let $r \in \mathbb{R}$. Then, $\{(x_1, x_2, \ldots) \in X \colon s(x) < r\}$ is an open set. Since projection on general vector spaces is an open mapping, $\{x_i \in X_1 \colon f(x_1) < r\}$ is an open set in $X_i$ for each $i \in \mathbb{N}$. In other words, lower semi-continuity of $s$ implies that $f_i$ is lower semi-continuous for each $i \in \mathbb{N}$.

Now, we are ready to give the main result of this section.

**Theorem 2.3.3.** *Assume that $f_i$ are non-constant lower semi-continuous functions for each $i \in \mathbb{N}$ such that either $f_i \geq 0$ for each $i \in \mathbb{N}$ or $f_i$ is finite for each $i \in \mathbb{N}$ together with $(s_n)_{n\in\mathbb{N}}$ is uniformly absolutely convergent. Then, $s$ is quasiconvex if and only if one of the following conditions holds:*

*(i) $f_i$ is convex for each $i \in \mathbb{N}$.*

*(ii) All of $f_1, f_2, \ldots$ except one are convex and*

$$\sum_{i=1}^{\infty} \frac{1}{c(f_i)} \leq 0. \tag{2.3.4}$$

**Proof.** Assume that $s$ is quasiconvex. Then, $f_i$ is quasiconvex for each $i \in \mathbb{N}$. We claim that at most one of $f_i$ is not convex for $i \in \mathbb{N}$. Suppose not. Fix $n_0 \in \mathbb{N}$

such that $n_0 \geq 2$ and define $\overline{s}$ on $\bigtimes_{i=n_0+1}^{\infty} X_i$ by

$$\overline{s}(x_{n_0+1}, x_{n_0+1}, \ldots) \coloneqq \sum_{j=n_0+1}^{\infty} f_j(x_j).$$

Then, it follows from Theorem 2.2.5 that at most one of $f_1, \ldots f_{n_0}, \overline{s}$ is not convex. We need to consider the following two cases:

(a) $\overline{s}$ is convex. Then, at most one of $f_1, \ldots, f_{n_0}$ is not convex.

(b) $\overline{s}$ is not convex. Let us choose $i_0, i_1 > n_0$ such that $f_{i_0}, f_{i_1}$ are not convex. It follows from the quasiconvexity of $s$ on $X$ that the function $\hat{s}$ defined on $X_1 \times \ldots \times X_{n_0} \times X_{i_0} \times X_{i_1}$ by

$$\hat{s}(x_1, \ldots x_{n_0}, x_{i_0}, x_{i_1}) \coloneqq f_1(x_1) + \ldots + f_{n_0}(x_{n_0}) + f_{i_0}(x_{i_0}) + f_{i_1}(x_{i_1})$$

is quasiconvex. But then, by Theorem 2.2.5, at least one of $f_{i_0}, f_{i_1}$ needs to be convex.

Therefore, either $f_i$ is convex for each $i \in \mathbb{N}$ or all of $f_1, f_2, \ldots$ except one are convex. If the former holds, then we are done. Suppose that the latter holds. Without loss of generality, since $s_n$ is absolutely convergent, we can assume $f_1$ be the function which is quasiconvex but not convex. For every $n \in \mathbb{N}$, define

$$a_n \coloneqq \sum_{i=1}^{n} \frac{1}{c(f_{i+1})}.$$

Observe that $(a_n)_{n \in N}$ is a non-decreasing sequence of real numbers since $\frac{1}{c(f_i)} \geq 0$ for every $i \in \mathbb{N} \setminus \{1\}$ by Theorem 2.1.4. Moreover, quasiconvexity of $s$ implies that $s_n$ is quasiconvex. In view of Theorem 2.2.5, for every $n \in \mathbb{N}$

$$\frac{1}{c(f_1)} + a_n \leq 0,$$
$$a_n \leq -\frac{1}{c(f_1)}.$$

Observe that, $\lim_{n\to\infty} a_n$ exists and it is finite since the sequence $(a_n)_{n\in\mathbb{N}}$ is non-decreasing and bounded above. Then,

$$\begin{aligned} &\lim_{n\to\infty} a_n \leq -\frac{1}{c(f_1)} \\ \Longrightarrow\ &\lim_{n\to\infty} a_n + \frac{1}{c(f_1)} \leq 0 \\ \Longrightarrow\ &\sum_{i=1}^{\infty} \frac{1}{c(f_i)} \leq 0. \end{aligned}$$

Conversely, assume that either $f_i$ is convex for each $i \in \mathbb{N}$ or all of $f_1, f_2, \ldots$ except one are convex and (2.3.4) holds. If the former holds, then $s$ is quasiconvex trivially. Suppose that the latter holds. Without loss of generality, since $s_n$ is absolutely convergent, we can assume that $f_1$ is not convex. Note that

$$\frac{1}{c(f_1)} + \frac{1}{c(f_2)} \leq \sum_{i=1}^{\infty} \frac{1}{c(f_i)}$$

since $f_i$ is convex, and consequently, $c(f_i) \geq 0$ for each $i \in \mathbb{N} \setminus \{1\}$ by Theorem 2.1.4. It follows from (2.3.4) that

$$\frac{1}{c(f_1)} + \frac{1}{c(f_2)} \leq 0. \tag{2.3.5}$$

In view of Theorem 2.2.5, convexity of $f_2$ together with (2.3.5) implies that $f_1 + f_2$ is quasiconvex. Then, by Theorem 2.2.2,

$$c(f_1) + c(f_2) \geq 0. \tag{2.3.6}$$

Next, define $\overline{\overline{s}}$ on $\bigtimes_{i=3}^{\infty} X_i$ by

$$\overline{\overline{s}}(x_3, x_4, \ldots) = \sum_{i=3}^{\infty} f_i(x_i).$$

Clearly, $\overline{\overline{s}}$ is convex, and consequently, $c(\overline{\overline{s}}) \geq 0$ by Theorem 2.1.4. By (2.3.6), we have

$$c(f_1) + c(f_2) + c(\overline{\overline{s}}) \geq 0.$$

It follows from Theorem 2.2.2 that $s$ is quasiconvex. $\square$

# Chapter 3

# Naturally quasiconvex conditional risk measures

Throughout this section, we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{G} \subseteq \mathcal{F}$ and $L^p(\mathcal{F}) := L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ for each $p \in [1, \infty]$. We assume that $L^p(\mathcal{F})$ space is equipped with the topology induced by the $L^p$-norm for each $p \in [1, \infty)$ and with the weak-$*$ topology $\sigma(L^\infty(\mathcal{F}), L^1(\mathcal{F}))$ for $p = \infty$. We denote the dual space of $L^p(\mathcal{G})$ with $L^q(\mathcal{G})$ , where $\frac{1}{p} + \frac{1}{q} = 1$ and $L^p_+(\mathcal{G}) := \{Y \in L^p(\mathcal{G}) \colon Y \geq 0\}$ . The dual cone of $L^p_+(\mathcal{G})$ is denoted as $L^p_+(\mathcal{G})^+$ and defined by

$$L^p_+(\mathcal{G})^+ := \{Y \in L^q(\mathcal{G}) \colon \mathbb{E}[YZ] \geq 0 \text{ for every } Z \in L^p(\mathcal{G})\} \cong L^q_+(\mathcal{G})$$

All of the equalities and inequalities among random variables hold $\mathbb{P}$-almost surely. For every $A \in \mathcal{F}$, the stochastic indicator function is denoted by $\mathbb{1}_A(\omega)$ being 1 if $\omega \in A$ and 0 if $\omega \notin A$.

## 3.1 Risk measures

In this section, we briefly discuss risk measures and review their basic properties. Let $p \in [1, \infty]$. In the following, we consider the risk measures which are defined on $L^p(\mathcal{F})$ and take values in the closed linear subspace $L^p(\mathcal{G})$ of its domain. In this setting, a risk measure gives the risk of a financial position possibly at an intermediate time. In other words, we are in the conditional setting. Notice that the case where the risk measure is measured at the present time, namely the static case, can be covered by taking $\mathcal{G} = \{\emptyset, \Omega\}$ (or a trivial $\sigma$-algebra) as a special case.

Below, we define a conditional risk measure as a functional with the minimal set of properties.

**Definition 3.1.1.** *A mapping $\rho\colon L^p(\mathcal{F}) \to L^p(\mathcal{G})$ is a conditional risk measure if it satisfies the following properties:*

*(i) Monotonicity : $X \leq Y$ implies $\rho(X) \geq \rho(Y)$ for every $X, Y \in L^p(\mathcal{F})$.*

*(ii) Quasiconvexity : $\rho(\lambda X + (1-\lambda)Y) \leq \max\{\rho(X), \rho(Y)\}$ for every $X, Y \in L^p(\mathcal{F})$ and $\lambda \in [0,1]$.*

The next definition provides some further properties of conditional risk measures which are of importance.

**Definition 3.1.2.** *A conditional risk measure $\rho\colon L^p(\mathcal{F}) \to L^p(\mathcal{G})$ is called*

*(i) translative if for every $Z \in L^p(\mathcal{G})$,*

$$\rho(X + Z) = \rho(X) - Z.$$

*(ii) local if for every $X \in L^p(\mathcal{F})$ and $A \in \mathcal{G}$,*

$$\rho(X\mathbb{1}_A)\mathbb{1}_A = \rho(X)\mathbb{1}_A.$$

*(iii) convex if for every $X, Y \in L^p(\mathcal{F})$ and $\lambda \in [0,1]$,*

$$\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y).$$

**Remark 3.1.3.** Locality property holds if and only if for every $X, U \in L^p(\mathcal{F})$ and $A \in \mathcal{G}$

$$\rho(X\mathbb{1}_A + U\mathbb{1}_{A^c}) = \rho(X)\mathbb{1}_A + \rho(U)\mathbb{1}_{A^c}.$$

Translativity captures the following idea. Suppose that $X, Z$ are financial positions such that worth of $X$ and $Z$ are fully known at the times $T_1$ and $T_2$, respectively, where $T_1 < T_2$. Suppose further that $\rho$ is a risk measure which measures risk of a financial position at time $T_1$ whose value is fully known at time $T_2$. Then, one can expect that at time $T_1$ the risk of holding $X + Z$ is equal to the risk of holding $X$ and liquidating $Z$.

Locality means that the events that will not happen in the future have no contribution to the value of risk. Note that this is exactly having a tree-like probabilistic setting as in Figure 3.1 when $\mathcal{F}$ is finitely generated.

Suppose $X$ and $Y$ are two financial positions. One can diversify by investing some fraction $\lambda \in [0, 1]$ of the resources on $X$ and the remaining on $Y$. In other words, one can diversify by choosing to invest on the portfolio $\lambda X + (1 - \lambda)Y$ instead of taking position in only one of the assets. The convexity property exactly means that the risk of the diversified position is less than or equal to the weighted sum of individual risks with weights equal to the fractions of the assets in the diversified portfolio.

As discussed earlier, quasiconvexity is also considered as a formulation of diversification. Under quasiconvexity, for each possible scenario, risk of the diversified portfolio is less than or equal to the risk of the riskier asset. Although both convexity and quasiconvexity refers to the same concept, it is clear that convexity is a more conservative property to capture the concept of diversification.

## 3.2 Natural quasiconvexity

The two properties which are of the utmost importance for the remaining part are natural quasiconvexity and $\star$-quasiconvexity. Natural quasiconvexity and $\star$-quasiconvexity are introduced by Tanaka [18] and Jeyakumar et al. [19], respectively. First, let us start with definition of naturally quasiconvex conditional risk measure.

**Definition 3.2.1.** *A conditional risk measure $\rho : L^p(\mathcal{F}) \to L^p(\mathcal{G})$ is naturally quasiconvex if for every $X, Y \in L^p(\mathcal{F})$ and $\lambda \in [0,1]$ there exists $\mu \in [0,1]$ such that*

$$\rho(\lambda X + (1-\lambda)Y) \leq \mu\rho(X) + (1-\mu)\rho(Y).$$

Clearly, natural quasiconvexity is a property which is stonger than quasiconvexity but weaker than convexity. Next, we give definition of $\star$-quasiconvex risk measure.

**Definition 3.2.2.** *A conditional risk measure $\rho : L^p(\mathcal{F}) \to L^p(\mathcal{G})$ is $\star$-quasiconvex if for every $Z^* \in L^q_+(\mathcal{G})$*

$$X \mapsto \mathbb{E}[\rho(X)Z^*]$$

*is quasiconvex on $L^p(\mathcal{F})$.*

Next, we show equivalence of natural quasiconvexity and $\star$-quasiconvexity. Set-valued case of the following theorem can be found in Kuroiwa [20].

**Theorem 3.2.3.** *A conditional risk measure $\rho : L^p(\mathcal{F}) \to L^p(\mathcal{G})$ is naturally quasiconvex if and only if it is $\star$-quasiconvex.*

**Proof.** Assume that $\rho$ is naturally quasiconvex and let $Z^* \in L^q_+(\mathcal{G})$. Then, for every $X, Y \in L^p(\mathcal{F})$ and $\lambda \in [0,1]$ there exists $\mu \in [0,1]$ such that

$$\mathbb{E}[\rho(\lambda X + (1-\lambda)Y)Z^*] \leq \mu\mathbb{E}[\rho(X)Z^*] + (1-\mu)\mathbb{E}[\rho(Y)Z^*].$$

Clearly,

$$\mu\mathbb{E}[\rho(X)Z^*] + (1-\mu)\mathbb{E}[\rho(Y)Z^*] \leq \max\{\mathbb{E}[\rho(X)Z^*], \mathbb{E}[\rho(Y)Z^*]\}.$$

Hence,

$$\mathbb{E}[\rho(\lambda X + (1-\lambda)Y)Z^*] \leq \max\{\mathbb{E}[\rho(X)Z^*], \mathbb{E}[\rho(Y)Z^*]\}$$

Conversely, assume that $\rho$ is $\star$-quasiconvex and assume to the contrary that $\rho$ is not naturally quasiconvex. Then, there exists $X, Y \in L^p(\mathcal{F})$, $\lambda \in [0,1]$ and $A \in \mathcal{G}$ with $\mathbb{P}(A) > 0$ such that for every $\mu \in [0,1]$

$$\rho(\lambda X(\omega) + (1-\lambda)Y(\omega)) > \mu\rho(X(\omega)) + (1-\mu)\rho(Y(\omega)),$$

for every $\omega \in A$. Since singletons are closed convex bounded sets, by seperation theorem, there exists $Z^* \in L^q_+(\mathcal{G})$ and $\alpha \in \mathbb{R}$ such that

$$\mathbb{E}[\rho(\lambda X + (1-\lambda)Y)Z^*\mathbb{1}_A] > \alpha > \mathbb{E}[(\mu\rho(X) + (1-\mu)\rho(Y))Z^*\mathbb{1}_A].$$

Observe that, by setting $\mu = 0$ and $\mu = 1$, we get

$$\begin{aligned} \mathbb{E}[\rho(\lambda X + (1-\lambda)Y)Z^*\mathbb{1}_A] &> \mathbb{E}[\rho(Y)Z^*\mathbb{1}_A], \\ \mathbb{E}[\rho(\lambda X + (1-\lambda)Y)Z^*\mathbb{1}_A] &> \mathbb{E}[\rho(X)Z^*\mathbb{1}_A]. \end{aligned}$$

Hence,

$$\mathbb{E}[\rho(\lambda X + (1-\lambda)Y)Z^*\mathbb{1}_A] > \max\{\mathbb{E}[\rho(X)Z^*], \mathbb{E}[\rho(Y)Z^*\mathbb{1}_A]\}.$$

Note that $Z^*\mathbb{1}_A \in L^q_+(\mathcal{G})$. Then, by Definition 3.2.2, $\rho$ is not $\star$-quasiconvex which is a contradiction. Therefore, $\rho$ is naturally quasiconvex. $\square$

## 3.3 Relationship between convexity and natural qausiconvexity

In light of the discussion of the previous section, natural quasiconvexity and convexity are closely related. Indeed, we will prove they are equivalent properties for risk measures under some mild assumptions. To be more precise, we need the risk measure to be in some sense non-constant lower semi-continuous and satisfy the locality property. In the example below, the role of being non-constant lower semi-continuous can be clearly seen since we make use of Theorem 2.2.5 which also requires these properties. On the other hand, locality is not explicitly emphasized but can be understood from the tree-like like structure we work on as can be seen in the Figure 3.3.1. The next example illustrates these ideas when $\mathcal{F}$ is finitely generated.

**Example 3.3.1.** Let $\Omega = \{\omega_1, \dots \omega_{10}\}$, $X : \Omega \to \mathbb{R}$ be an integrable, $\mathcal{F}$-measurable random variable and $\rho : L^p(\mathcal{F}) \to L^p(\mathcal{G})$ be a conditional risk measure. Take $\mathcal{F}$ and $\mathcal{G}$ as

$$\begin{aligned}
\mathcal{F} &= \sigma\big(\{\{w_1\}, \dots, \{w_{10}\}\}\big), \\
\mathcal{G} &= \sigma\big(\{\{w_1, w_2, w_3, w_4\}, \{w_5, w_6, w_7\}, \{w_8, w_9, w_{10}\}\}\big).
\end{aligned}$$

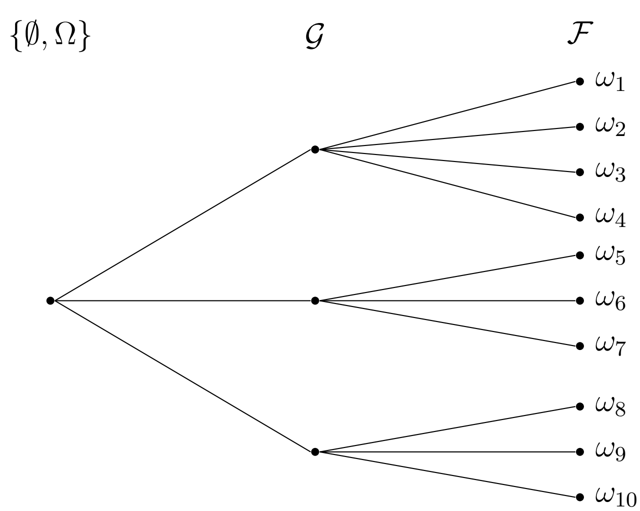


Figure 3.1: Relationship between convex risk measures and natural quasiconvex risk measures when $\mathcal{F}$ is finitely generated.

In this setting, $L^p(\mathcal{F}) \cong \mathbb{R}^4 \times \mathbb{R}^3 \times \mathbb{R}^3 \cong \mathbb{R}^{10}$ and $L^p(\mathcal{G}) \cong \mathbb{R}^3$. Since $X \in L^p(\mathcal{F})$, it can be seen as a vector in $\mathbb{R}^{10}$.

$$x = \begin{bmatrix} X(\omega_1) \\ \vdots \\ X(\omega_{10}) \end{bmatrix} \in \mathbb{R}^{10}$$

Furthermore, $\rho$ can be seen as a vector valued function. For every $x \in \mathbb{R}^{10}$,

$$\rho(x) = \begin{bmatrix} \rho_1(x_{1:4}) \\ \rho_2(x_{5:7}) \\ \rho_3(x_{8:10}) \end{bmatrix} \in \mathbb{R}^3$$

where $\rho_1 : \mathbb{R}^4 \to \mathbb{R}$, $\rho_2 : \mathbb{R}^3 \to \mathbb{R}$ and $\rho_3 : \mathbb{R}^3 \to \mathbb{R}$.

Assume that $\rho$ is naturally qusiconvex and $\rho_i$ is lower semi-continuous, non-constant for each $i \in \{1, 2, 3\}$. Note that, in view of Remark 2.3.2, lower semi-continuity of $\rho$ is sufficient for $\rho_i$ being lower semi-continuous for each $i \in \{1, 2, 3\}$. It follows from Theorem 3.2.3 that $\rho$ is $\star$-quasiconvex. Then, by Definition 3.2.1, for every $w = (w_1, w_2, w_3) \in \mathbb{R}^3_+$,

$$w_1\rho_1(x_{1:4}) + w_2\rho_2(x_{5:7}) + w_3\rho_3(x_{8:10}) \tag{3.3.1}$$

is quasiconvex on $\mathbb{R}^4 \times \mathbb{R}^3 \times \mathbb{R}^3$. Consider the following four cases:

(a) $c(\rho_1) \geq 0$, $c(\rho_2) \geq 0$ and $c(\rho_3) > 0$.

(b) $c(\rho_1) \geq 0$, $c(\rho_2) \geq 0$ and $c(\rho_3) < 0$. Then, it follows from Theorem 2.1.4 that $\rho_1, \rho_2$ are convex and $\rho_3$ is non-convex. In view of (3.3.1) and Theorem 2.2.5, we have

$$\frac{1}{c(w_1\rho_1)} + \frac{1}{c(w_2\rho_2)} + \frac{1}{c(w_3\rho_3)} \leq 0. \tag{3.3.2}$$

It follows from Lemma 2.1.5 together with (3.3.2) that

$$\frac{1}{c(w_1\rho_1)} + \frac{1}{c(w_2\rho_2)} + \frac{1}{c(w_3\rho_3)} = \frac{w_1}{c(\rho_1)} + \frac{w_2}{c(\rho_2)} + \frac{w_3}{c(\rho_3)} \tag{3.3.3}$$

Notice that, since $\rho_i$ is non-constant for each $i \in \{1,2,3\}$, $c(\rho_i) < +\infty$ for each $i \in \{1,2,3\}$. Let us set $(w_1, w_2, w_3) = (c(\rho_1), c(\rho_2), -c(\rho_3))$ in (3.3.3). Then,

$$\frac{1}{c(w_1\rho_1)} + \frac{1}{c(w_2\rho_2)} + \frac{1}{c(w_3\rho_3)} = 1,$$

which is in contradiction with (3.3.2). Hence, this case is eliminated.

(c) $c(\rho_1) \geq 0$, $c(\rho_2) < 0$ and $c(\rho_3) < 0$. Then, $\rho_2$ and $\rho_3$ are not convex by Theorem 2.1.4. On the other hand, since (3.3.1) holds, it follows from Theorem 2.2.5 that at most one of $\rho_1$, $\rho_2$ or $\rho_3$ is non-convex. We reached a contradiction, hence this case is eliminated.

(d) $c(\rho_1) < 0$, $c(\rho_2) < 0$ and $c(\rho_3) < 0$. By a similar argument with the case (c), we can eliminate this case.

Observe that the remaining four cases are symmetric either with the case (b) or (c). Therefore, we are left with the case (a) and $\rho_1, \rho_2, \rho_3$ are convex. Hence, $\rho$ is convex.

In Example 3.3.1, we need non-constancy of $\rho_i$ for each $i \in \{1,2,3\}$ to draw the connection between natrually quasiconvexity and convexity. We need a similar assumption to generalize Example 3.3.1, which is stated in the following assumption.

**Assumption 3.3.2.** *For every $X \in L^p(\mathcal{F})$ and $A \in \mathcal{G}$ with $\mathbb{P}(A) > 0$*

$$X \mapsto \mathbb{E}[\rho(X)1_A]$$

*is non-constant.*

To motivate this assumption, let us give an example of a quasiconvex risk measure which satisfies Assumption 3.3.2.

**Example 3.3.3.** Consider the certainty equivalent risk measure

$$\rho(X) = \ell^{-1}\mathbb{E}[\ell(-X)|\mathcal{G}]$$

where $\ell\colon \mathbb{R} \to \mathbb{R}$ is a continuous increasing loss function. Let $x_1, x_2 \in \mathbb{R}$ such that $x_1 \neq x_2$. Observe that

$$\begin{aligned}\rho(x_1) &= \ell^{-1}\mathbb{E}[\ell(-x_1)|\mathcal{G}] = \ell\big(\ell^{-1}(-x_1)\big) = -x_1,\\ \rho(x_2) &= \ell^{-1}\mathbb{E}[\ell(-x_2)|\mathcal{G}] = \ell\big(\ell^{-1}(-x_2)\big) = -x_2.\end{aligned}$$

Hence,

$$\begin{aligned}\mathbb{E}[\rho(x_1)\mathbb{1}_A] - \mathbb{E}[\rho(x_2)\mathbb{1}_A] &= \mathbb{E}\big[\big(\rho(x_1) - \rho(x_2)\big)\mathbb{1}_A\big]\\ &= \mathbb{E}[(x_2 - x_1)\mathbb{1}_A]\\ &= (x_2 - x_1)\mathbb{P}(A)\\ &> 0,\end{aligned}$$

where the strict inequality holds since $x_1 \neq x_2$ and $\mathbb{P}(A) > 0$. Hence,

$$X \mapsto \mathbb{E}[\rho(X)\mathbb{1}_A]$$

is non-constant.

The next theorem is a generalization of Example 3.3.1.

**Theorem 3.3.4.** *Assume $\rho\colon L^p(\mathcal{F}) \to L^p(\mathcal{G})$ is a lower semi-continuous local conditional risk measure which satisfies Assumption 3.3.2. Then, $\rho$ is convex if and only if it is naturally quasiconvex.*

**Proof.** Assume $\rho$ is convex. Then, $\rho$ is naturally quasiconvex trivially.

Conversely, assume that $\rho$ is naturally quasiconvex. Let us fix $A \in \mathcal{G}$ such that $\mathbb{P}(A) > 0$. Then, as a special case of Theorem 3.2.3, for every $Y \in L^q_+(\mathcal{G})$ such that $Y = y_1 \mathbb{1}_A + y_2 \mathbb{1}_{A^c}$,

$$X \mapsto \mathbb{E}[\rho(X)Y]$$

is quasiconvex on $L^p(\mathcal{F})$, where $y_i > 0$ for each $i \in \{1, 2\}$. Moreover,

$$\begin{aligned}\mathbb{E}[\rho(X)Y] &= y_1 \mathbb{E}[\rho(X)\mathbb{1}_A] + y_2 \mathbb{E}[\rho(X)\mathbb{1}_{A^c}] \\ &= y_1 \mathbb{E}[\rho(X\mathbb{1}_A)\mathbb{1}_A] + y_2 \mathbb{E}[\rho(X\mathbb{1}_{A^c})\mathbb{1}_{A^c}],\end{aligned}$$

where the last equality follows from the locality of $\rho$. Hence,

$$X \mapsto y_1 \mathbb{E}[\rho(X\mathbb{1}_A)\mathbb{1}_A] + y_2 \mathbb{E}[\rho(X\mathbb{1}_{A^c})\mathbb{1}_{A^c}] \tag{3.3.4}$$

is quasiconvex on $L^p(\mathcal{F})$. Clearly, $X = X\mathbb{1}_A + X\mathbb{1}_{A^c}$. Let us denote $X_1 := X\mathbb{1}_A$ and $X_2 := X\mathbb{1}_{A^c}$, $L^p(A) := L^p\Big(A, \mathcal{F} \cap A, \frac{\mathbb{P}|_{\mathcal{F}\cap A}}{\mathbb{P}(A)}\Big)$ and $L^p(A^c) := L^p\Big(A^c, \mathcal{F} \cap A^c, \frac{\mathbb{P}|_{\mathcal{F}\cap A^c}}{\mathbb{P}(A^c)}\Big)$. Observe that $X_1 \in L^p(A)$ and $X_2 \in L^p(A^c)$. Thus,

$$L^p(\mathcal{F}) \cong L^p(A) \times L^p(A^c). \tag{3.3.5}$$

Next, define

$$\begin{aligned}&\rho_1 := \rho|_{L^p(A)}, \\ &\rho_2 := \rho|_{L^p(A^c)}, \\ &f_1 \colon L^p(A) \to \mathbb{R} :: X_1 \mapsto \mathbb{E}[\rho_1(X_1)\mathbb{1}_A], \\ &f_2 \colon L^p(A^c) \to \mathbb{R} :: X_2 \mapsto \mathbb{E}[\rho_2(X_2)\mathbb{1}_{A^c}].\end{aligned}$$

In view of (3.3.4) and (3.3.5),

$$(X_1, X_2) \mapsto y_1 f_1(X_1) + y_2 f_2(X_2) \tag{3.3.6}$$

is quasiconvex on $L^p(A) \times L^p(A^c)$. Note that $f_1, f_2$ are lower semi-continuous by Remark 2.3.2. We need to consider the following four cases:

(a) $c(f_1) \geq 0$, $c(f_2) \geq 0$.

(b) $c(f_1) \geq 0$, $c(f_2) < 0$. Then, it follows from Theorem 2.1.4 that $f_1$ is convex and $f_2$ is not convex. In view of (3.3.6) and Theorem 2.2.5, we have

$$\frac{1}{c(y_1 f_1)} + \frac{1}{c(y_2 f_2)} \leq 0. \tag{3.3.7}$$

Note that $f_1$ and $f_2$ are non-constant by Assumption 3.3.2, therefore $c(f_i) < +\infty$ for each $i \in \{1, 2\}$ . By Lemma 2.1.5, we have

$$\frac{1}{c(y_1 f_1)} + \frac{1}{c(y_2 f_2)} = \frac{y_1}{c(f_1)} + \frac{y_2}{c(f_2)} \tag{3.3.8}$$

Let us set $(y_1, y_2) = \big(c(f_1), -\frac{1}{2}c(f_2)\big)$ in (3.3.8). Then,

$$\frac{1}{c(y_1 f_1)} + \frac{1}{c(y_2 f_2)} = \frac{1}{2},$$

which is in contradiction with (3.3.7). Hence, this case is eliminated.

(c) $c(f_1) < 0$, $c(f_2) \geq 0$. This case is symmetric with case (b), hence eliminated.

(d) $c(f_1) < 0$, $c(f_2) < 0$. Then, $f_1$ and $f_2$ are not convex by Theorem 2.1.4. On the other hand, since (3.3.6) holds, it follows from Theorem 2.2.5 that at least one of $f_1$ or $f_2$ is convex. We reached a contradiction, hence this case is eliminated.

Summing up, we are left with the case (a). In particular, by making use of the locality property of $\rho_1$, we have $X_1 \mapsto \mathbb{E}[\rho_1(X_1)\mathbb{1}_A]$ is convex on $L^p(A)$. But then, for every $X, Y \in L^p(\mathcal{F})$ and $\lambda \in [0, 1]$

$$\begin{aligned}
\mathbb{E}[\rho((\lambda X + (1-\lambda)Y))\mathbb{1}_A] &= \mathbb{E}[\rho((\lambda X + (1-\lambda)Y)\mathbb{1}_A)\mathbb{1}_A] \\
&= \mathbb{E}[\rho_1(\lambda X\mathbb{1}_A + (1-\lambda)Y\mathbb{1}_A)\mathbb{1}_A] \\
&\leq \mathbb{E}[(\lambda\rho_1(X\mathbb{1}_A) + (1-\lambda)\rho_1(Y\mathbb{1}_A))\mathbb{1}_A] \\
&= \mathbb{E}[\lambda\rho(X\mathbb{1}_A)\mathbb{1}_A + (1-\lambda)\rho(Y\mathbb{1}_A)\mathbb{1}_A] \\
&= \mathbb{E}[(\lambda\rho(X) + (1-\lambda)\rho(Y))\mathbb{1}_A],
\end{aligned}$$

where the first and the last equalities holds since $\rho$ is local and the inequality follows from the convexity of $\rho_1$. But then, for every $A \in \mathcal{F}$ with $\mathbb{P}(A) > 0$

$$\mathbb{E}[\rho((\lambda X + (1-\lambda)Y))\mathbb{1}_A] \leq \mathbb{E}[(\lambda\rho(X) + (1-\lambda)\rho(Y))\mathbb{1}_A] \tag{3.3.9}$$

It is a well known result that (3.3.9) holds if and only if $\rho((\lambda X + (1-\lambda)Y)) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$ (see, e.g., Çınlar [21, Remark 2.3(d)]). Hence, $\rho$ is convex on $L^p(\mathcal{F})$. □

## 3.4 Relationship between convexity and natural quasiconvexity on $L^2$

Now, we turn our attention to the case $\rho\colon L^2(\mathcal{F}) \to L^2(\mathcal{G})$. Before we get into the discussion, let us review some basic properties of $L^2(\mathcal{F})$ and $L^2(\mathcal{G})$.

$L^2(\mathcal{F})$ and $L^2(\mathcal{G})$ are seperable Hilbert spaces with respect to the inner product

$$\langle X, Y\rangle = \mathbb{E}[XY]. \tag{3.4.1}$$

It is a well-known result that every seperable Hilbert space has at least one countable orthonormal basis. Thus, one can find countable orthonormal bases for $L^2(\mathcal{F})$ and $L^2(\mathcal{G})$. Moreover, $L^2(\mathcal{G})$ is a closed subspace of $L^2(\mathcal{F})$ and we have

$$L^2(\mathcal{F}) = L^2(\mathcal{G}) \oplus L^2(\mathcal{G})^{\perp},$$

where $L^2(\mathcal{G})^{\perp} = \{X \in L^2(\mathcal{F})\colon \langle X, Y\rangle = 0 \text{ for every } Y \in L^2(\mathcal{G})\}$. We also define, for every $X \in L^2(\mathcal{F})$, the orthogonal complement of $X$ by

$$X^{\perp} := X - \mathbb{E}[X|\mathcal{G}].$$

Let $\{e_i\colon i \in I\}$ be a countable orthonormal basis of $L^2(\mathcal{G})$. It is orthonormal in the sense that

(i) $\|e_i\| = 1$ for each $i \in I$,

(ii) $\langle e_i, e_j \rangle = 0$ for each $i, j \in I$ such that $i \neq j$,

where $\|\cdot\|$ is the $L^2$-norm induced by the inner product in (3.4.1). Furthermore, for every $Y \in L^2(\mathcal{G})$, we have

$$Y = \sum_{i \in I} \langle Y, e_i \rangle e_i.$$

A natural question is whether or not there is an extension of $\{e_i \colon i \in I\}$ to a countable orthonormal basis of $L^2(\mathcal{F})$. Luckily, one can find such an extension and a proof of this result can be found in Hunter and Nachtergaele [22, Theorem 6.29]. In light of these arguments, let $I' \supseteq I$ such that $\{e_i \colon i \in I'\}$ is a countable orthonormal basis of $L^2(\mathcal{F})$.

Let us come back to the discussion of this section. We start with defining a new property, called locality with respect to a given basis. This property precisely says that when one calculates the inner product of $\rho(X)$ with an element of the basis, one can replace $X$ with the sum of its projection on that basis and its orthogonal complement.

**Definition 3.4.1.** *$\rho$ is called local with respect to the basis $\{e_i \colon i \in I\}$ if*

$$\langle \rho(X), e_i \rangle = \langle \rho(\langle X, e_i \rangle e_i + X^{\perp}), e_i \rangle$$

*for each $i \in I$.*

Before proceeding further, let us give example of a risk measure which is local with respect to a basis.

**Example 3.4.2.** Let $\Omega = [0,1]$, $\mathcal{F}$ be the Borel $\sigma$-algebra on $[0,1]$, $\mathcal{G} = \mathcal{B}\big([0, \frac{1}{2}]\big) \vee \big\{(\frac{1}{2}, 1]\big\}$ and $\mathbb{P}$ be the Lebesgue measure . Then,

$$L^2(\mathcal{G}) = \Big\{X \in L^2([0,1]) \colon X \mathbb{1}_{\left(\frac{1}{2},1\right]} = 0,\ X \mathbb{1}_{\left(\frac{1}{2},1\right]} = c \text{ for some constant } c \in \mathbb{R}\Big\},$$
$$L^2(\mathcal{G})^{\perp} = \Big\{Z \in L^2([0,1]) \colon Z \mathbb{1}_{\left[0,\frac{1}{2}\right)} = 0\Big\}$$

For every $i \in \mathbb{N}$, let us define

$$
\begin{aligned}
e_i &:= \Big\{\frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}, \frac{1}{2}\cos(4\pi i\omega)\mathbb{1}_{\left(0,\frac{1}{2}\right]}, 2\mathbb{1}_{\left[\frac{1}{2},1\right]}\Big\},\\
\beta_i &:= \Big\{\frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{\left(\frac{1}{2},1\right]}, \frac{1}{2}\cos(4\pi i\omega)\mathbb{1}_{\left(\frac{1}{2},1\right]}\Big\}.
\end{aligned}
$$

$\{e_i\}_{i=1}^{\infty}$ and $\{\beta_i\}_{i=1}^{\infty}$ are orthonormal bases for $L^2(\mathcal{G})$ and $L^2(\mathcal{G})^{\perp}$, respectively. Furthermore, $\{e_i, \beta_i\}_{i=1}^{\infty}$ is a basis for $L^2(\mathcal{F})$. Note that

$$
\mathbb{E}[X|\mathcal{G}] = X\mathbb{1}_{\left[0,\frac{1}{2}\right]} + \mathbb{E}[2X\mathbb{1}_{\left(\frac{1}{2},1\right]}].
$$

Let us consider the conditional entropic risk measure $\rho\colon L^2(\mathcal{F}) \to L^2(\mathcal{G})$ defined by

$$
\rho(X) = \log\big(\mathbb{E}[e^{-X}|\mathcal{G}]\big).
$$

For every $X \in L^2(\mathcal{F})$, define

$$
X_s := \big\langle X, \frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}\big\rangle.
$$

Note that

$$
\begin{aligned}
\rho\Big(X_s + X^{\perp}\Big) &= \log\big(\mathbb{E}\big[e^{X_s\frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{[0,\frac{1}{2}]}+X^{\perp}}\big|\mathcal{G}\big]\big)\\
&= \log\big(e^{X_s\frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{[0,\frac{1}{2}]}+X^{\perp}}\mathbb{1}_{[0,\frac{1}{2})} + \mathbb{E}[2e^{X_s\frac{1}{2}\sin(4\pi i\omega)+X^{\perp}}\mathbb{1}_{(\frac{1}{2},1]}]\big).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\Big\langle \rho\Big(X_s\frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}+X^{\perp}\Big),\frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}\Big\rangle\\
&=\int_0^{\frac{1}{2}}\log\Big(e^{-X_s\frac{1}{2}\sin(4\pi n\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}-X^{\perp}}\mathbb{1}_{\left[0,\frac{1}{2}\right)}+\mathbb{E}\big[2e^{X_s\sin(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}+X^{\perp}}\mathbb{1}_{\left(\frac{1}{2},1\right]}\big]\Big)\frac{1}{2}\sin(4\pi i\omega)d\omega\\
&=\int_0^{\frac{1}{2}}(-X_s\frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}-X^{\perp})\frac{1}{2}\sin(4\pi i\omega)d\omega\\
&=\int_0^{\frac{1}{2}}-X_s\frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}\frac{1}{2}\sin(4\pi i\omega)d\omega\\
&=-X_s\frac{1}{4}\int_0^{\frac{1}{2}}\sin^2(4\pi i\omega)d\omega\\
&=-X_s.
\end{aligned}
\tag{3.4.2}
$$

On the other hand,

$$
\begin{aligned}
\Big\langle \rho(X),\frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}\Big\rangle &=\int_0^{\frac{1}{2}}\log\big(e^{-X(\omega)}+\mathbb{E}[2e^{X_s}\mathbb{1}_{\left(\frac{1}{2},1\right]}]).\big)\frac{1}{2}\sin(4\pi i\omega)d\omega\\
&=\int_0^{\frac{1}{2}}\log\big(e^{-X(\omega)}\big)\frac{1}{2}\sin(4\pi i\omega)d\omega\\
&=\int_0^{\frac{1}{2}}-X(\omega)\frac{1}{2}\sin(4\pi i\omega)d\omega.\\
&=-X_s.
\end{aligned}
\tag{3.4.3}
$$

In view of (3.4.2) and (3.4.3), we have

$$\Big\langle \rho(X),\frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}\Big\rangle=\Big\langle \rho\Big(X_s\frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}+X^{\perp}\Big),\frac{1}{2}\sin(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}\Big\rangle.$$

Similar calculations yields that

$$\Big\langle \rho(X),\frac{1}{2}\cos(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}\Big\rangle=\Big\langle \rho\Big(X_s\frac{1}{2}\cos(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}+X^{\perp}\Big),\frac{1}{2}\cos(4\pi i\omega)\mathbb{1}_{\left[0,\frac{1}{2}\right]}\Big\rangle.$$

Hence, conditional entropic risk measure is local with respect to $\{e_i,\beta_i\}_{i=1}^{\infty}$.

In the following, we are concerned with connections between locality with respect to a given basis and the two of the properties given in Definition 3.1.2, namely locality and translativity. We start with the connection between locality with respect to a given basis and locality, but first we need some structural properties on the measure space that we work on.

Let $\mathcal{G} = \sigma(\{B_i \colon i \in I\})$, where $\{B_i \colon i \in I\}$ is a set of disjoint sets. Then, for any given $Y \in L^2(\mathcal{G})$, one can find $y_i \in \mathbb{R}$ such that

$$Y = \sum_{i \in I} y_i \mathbb{1}_{B_i}.$$

In other words, $\{\mathbb{1}_{B_i} \colon i \in I\}$ spans $L^2(\mathcal{G})$. Moreover,

$$\langle \mathbb{1}_{B_i}, \mathbb{1}_{B_j} \rangle = 0$$

for every $i, j \in I$ such that $i \neq j$. Therefore, $\left\{ \frac{\mathbb{1}_{B_i}}{\sqrt{\mathbb{P}(B_i)}} \colon i \in I \right\}$ is a countable orthonormal basis for $L^2(\mathcal{G})$. Through out this section, we set

$$e_i^{\text{ind}} = \frac{\mathbb{1}_{B_i}}{\sqrt{\mathbb{P}(B_i)}}$$

for each $i \in I$ . We refer the countable orthonormal basis $\left\{ \frac{\mathbb{1}_{B_i}}{\sqrt{\mathbb{P}(B_i)}} \right\}_{i \in I}$ as the $e^{\text{ind}}$ basis.

We begin with the connection between the old locality (Definition 3.1.2) and the new locality (Definition 3.4.1).

**Theorem 3.4.3.** *If $\rho$ is local then it is local with respect to $e^{ind}$.*

**Proof.** Let $X \in L^2(\mathcal{F})$ and $i \in I$. It follows from the definition of $e^{\text{ind}}$ that

$$\rho\big(\langle X, e_i^{\text{ind}} \rangle e_i^{\text{ind}} + (X - \mathbb{E}[X|\mathcal{G}])\big) = \rho\Big(\frac{\mathbb{E}[X\mathbb{1}_{B_i}]\mathbb{1}_{B_i}}{\mathbb{P}(B_i)} + \sum_{j \in I} (X - \mathbb{E}[X|\mathcal{G}])\mathbb{1}_{B_j}\Big). \quad (3.4.4)$$

Note that

$$\frac{\mathbb{E}[X\mathbb{1}_{B_i}]\mathbb{1}_{B_i}}{\mathbb{P}(B_i)} = \mathbb{E}[X|B_i]\mathbb{1}_{B_i} = \mathbb{E}[X\mathbb{1}_{B_i}|\mathcal{G}]\mathbb{1}_{B_i} \quad (3.4.5)$$

Moreover,

$$\sum_{j\in I}(X-\mathbb{E}[X|\mathcal{G}])\mathbb{1}_{B_j}=\sum_{j\in I}(X-\mathbb{E}[X\mathbb{1}_{B_j}|\mathcal{G}])\mathbb{1}_{B_j} \tag{3.4.6}$$

It follows from (3.4.4), (3.4.5) and (3.4.6) that

$$\rho\big(\langle X,e_i^{\text{ind}}\rangle e_i^{\text{ind}}+(X-\mathbb{E}[X|\mathcal{G}])\big)=\rho\Big(X\mathbb{1}_{B_i}+\sum_{\substack{j\in I\\ j\neq i}}(X-\mathbb{E}[X\mathbb{1}_{B_j}|\mathcal{G}])\mathbb{1}_{B_j}\Big) \tag{3.4.7}$$

Since $\rho$ is local, by using the alternative definition of locality (Remark 3.1.3), we can re-write (3.4.7) as

$$\rho\Big(X\mathbb{1}_{B_i}+\sum_{\substack{j\in I\\ j\neq i}}(X-\mathbb{E}[X\mathbb{1}_{B_j}|\mathcal{G}])\mathbb{1}_{B_j}\Big)=\rho(X)\mathbb{1}_{B_i}+\rho\Big(\sum_{\substack{j\in I\\ j\neq i}}(X-\mathbb{E}[X\mathbb{1}_{B_j}|\mathcal{G}])\mathbb{1}_{B_j}\Big)\mathbb{1}_{B_i^c} \tag{3.4.8}$$

In view of (3.4.4), (3.4.7) and (3.4.8), we have

$$\rho\big(\langle X,e_i^{\text{ind}}\rangle e_i^{\text{ind}}+(X-\mathbb{E}[X|\mathcal{G}])\big)=\rho(X)\mathbb{1}_{B_i}+\rho\Big(\sum_{\substack{j\in I\\ j\neq i}}(X-\mathbb{E}[X\mathbb{1}_{B_j}|\mathcal{G}])\mathbb{1}_{B_j}\Big)\mathbb{1}_{B_i^c}$$

By taking inner product of both sides with $e_i^{\text{ind}}$, we get

$$\begin{aligned}\langle\rho\big(\langle X,e_i^{\text{ind}}\rangle e_i^{\text{ind}}+(X-\mathbb{E}[X|\mathcal{G}])\big),e_i^{\text{ind}}\rangle &= \langle\rho(X)\mathbb{1}_{B_i},e_i^{\text{ind}}\rangle\\ &\quad+\Big\langle\rho\Big(\sum_{\substack{j\in I\\ j\neq i}}(X-\mathbb{E}[X\mathbb{1}_{B_j}|\mathcal{G}])\mathbb{1}_{B_j}\Big)\mathbb{1}_{B_i^c},e_i^{\text{ind}}\Big\rangle\\ &=\langle\rho(X)\mathbb{1}_{B_i},e_i\rangle\\ &=\mathbb{E}[\rho(X)\frac{\mathbb{1}_{B_i}}{\mathbb{P}(B_i)}]\\ &=\langle\rho(X),e_i^{\text{ind}}\rangle,\end{aligned}$$

where the second equality holds since $e_i^{\text{ind}}$ is orthogonal to $\mathbb{1}_{B_i^c}$. $\square$

Next, we examine the relation between translativity and locality with respect to a given basis. In the usual sense, one can't draw any connection between translativity and locality (Definition 3.1.2). An important result regarding the newly defined locality (Definition 3.4.1) is that, given a countable orthonormal basis, translativity implies locality with respect to that basis, as shown in the next theorem.

**Theorem 3.4.4.** *Assume that $\rho$ is translative and let $\{e_i\colon i \in I\}$ be any countable orthonormal basis of $L^2(\mathcal{G})$. Then, $\rho$ is local with respect to $\{e_i\colon i \in I\}$ .*

**Proof.** Let $X \in L^2(\mathcal{F})$. Observe that, since $\rho$ is translative, we have

$$\langle \rho(\langle X, e_i\rangle e_i + X^{\perp}), e_i\rangle = \langle \rho(X) - \langle X, e_i\rangle e_i + \mathbb{E}[X|\mathcal{G}], e_i\rangle. \tag{3.4.9}$$

Note that

$$\langle \rho(X) - \langle X, e_i\rangle e_i + \mathbb{E}[X|\mathcal{G}], e_i\rangle = \langle \rho(X), e_i\rangle - \langle X, e_i\rangle\langle e_i, e_i\rangle + \langle \mathbb{E}[X|\mathcal{G}], e_i\rangle.$$

Since $\{e_i\colon i \in I\}$ is orthonormal, $\langle e_i, e_i\rangle = 1$. Hence, we have

$$\langle \rho(X) - \langle X, e_i\rangle e_i + \mathbb{E}[X|\mathcal{G}], e_i\rangle = \langle \rho(X), e_i\rangle - \langle X, e_i\rangle + \mathbb{E}[\mathbb{E}[X|\mathcal{G}]e_i].$$

It follows from the self-adjointness of conditional expectation that $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]e_i] = \langle X, e_i\rangle$. Thus,

$$\langle \rho(X) - \langle X, e_i\rangle e_i + \mathbb{E}[X|\mathcal{G}], e_i\rangle = \langle \rho(X), e_i\rangle. \tag{3.4.10}$$

By (3.4.9) and (3.4.10), we have

$$\langle \rho(\langle X, e_i\rangle e_i + X^{\perp}), e_i\rangle = \langle \rho(X), e_i\rangle,$$

that is , $\rho$ is local with respect to $\{e_i\colon i \in I\}$. □

An important result of this section is Conjecture 3.4.7, given below. It gives the relation between natural quasiconvexity and convexity for risk measures on $L^2$ that are local with respect to a seperable basis. In the proof of this result, we make use of Theorem 2.3.3, which is concerned with countable sums. Before, we get into the result, we need to introduce some notations and an assumption.

Let us define for every $i \in I$

$$\mathcal{Y}_i \coloneqq \{\alpha e_i \colon \alpha \in \mathbb{R}\},$$
$$\mathcal{X}_i \coloneqq \mathcal{Y}_i \times L^2(\mathcal{G})^{\perp}.$$

We also define $\mathcal{S} \subseteq \bigtimes_{i\in I} \mathcal{X}_i$ by

$$\mathcal{S} \coloneqq \Big\{\big((Y_1, Z_1), \dots\big) \in \bigtimes \mathcal{X}_i \colon Y_i = \langle X, e_i\rangle e_i,\ Z_i = X^{\perp} \text{ for some } X \in L^2(\mathcal{F}),\ \text{for each } i \in I\Big\}.$$

Note that for every $X \in L^2(\mathcal{F})$, there is a unique representation of $X$ such that

$$X = \sum_{i\in I} \big(\langle X, e_i\rangle e_i\big) + X^{\perp}$$

Hence, $\phi\colon \mathcal{S} \to L^2(\mathcal{F})$ defined by

$$\phi\Big(\big((Y_1, Z), (Y_2, Z), \dots)\big)\Big) \coloneqq \Big(\sum_{i=1}^{\infty} Y_i\Big) + Z$$

is a bijection. One can easliy veriy that, $\phi$ is also linear, and consequently its an isomporhim between $L^2(\mathcal{F})$ and $S$. Furhtermore, let $f_i \colon \mathcal{X}_i \mapsto \mathbb{R}$ is defined as

$$f_i(X_i) \coloneqq \langle \rho(\langle X, e_i\rangle e_i + X^{\perp}), e_i\rangle$$

Next, we define the preorder $\preceq$ on $L^2(\mathcal{G})$ by

$$Y \preceq V \iff \langle Y, e_i\rangle \le \langle V, e_i\rangle \text{ for each } i \in I.$$

The ordering cone which corresponds to $\preceq$ is given as

$$C = \{Y \in L^2(\mathcal{G}) \colon Y \succeq 0\}.$$

Furthermore, the dual cone of $C$ can be written explicitly as

$$C^{+} = \{Y \in L^2(\mathcal{G}) \colon \langle Y, V\rangle \ge 0 \text{ for every } V \in C\}.$$

We will also need the following assumption in the sketch of proof of Conjecture

3.4.7.

**Assumption 3.4.5.** *Let $\{e_i \colon i \in I\}$ be a countable orthonormal basis for $L^2(\mathcal{G})$. Then,*

$$X \mapsto \langle \rho(X), e_i \rangle$$

*is non-constant.*

The following is an example of a risk measure for which Assumption 3.4.5 holds.

**Example 3.4.6.** Let $\mathcal{G} = \sigma(\{B_i \colon i \in I\})$, where $\{B_i \colon i \in I\}$ is a set of disjoint sets and $\mathbb{P}(B_i) > 0$. Recall that $e^{\text{ind}}$ basis is defined as $e_i^{\text{ind}} = \frac{\mathbb{1}_{B_i}}{\sqrt{\mathbb{P}(B_i)}}$ for each $i \in I$. Consider the certainty equivalent risk measure

$$\rho(X) = \ell^{-1}\mathbb{E}[\ell(-X)|\mathcal{G}]$$

where $\ell \colon \mathbb{R} \to \mathbb{R}$ is a continuous increasing loss function. Let $x_1, x_2 \in \mathbb{R}$ such that $x_1 \neq x_2$. We claim that Assumption 3.4.5 holds for $\rho$ and $e^{\text{ind}}$ basis. Observe that, for every $i \in I$,

$$\begin{aligned}
\rho(\langle x_1, e_i^{\text{ind}} \rangle e_i^{\text{ind}} + x_1^{\perp}) &= \rho(\mathbb{E}[x_1 e_i^{\text{ind}}] e_i^{\text{ind}}) \\
&= \ell^{-1}\mathbb{E}[\ell(-\mathbb{E}[x_1 e_i^{\text{ind}}] e_i^{\text{ind}})|\mathcal{G}] \\
&= \ell^{-1}\ell(-\mathbb{E}[x_1 e_i^{\text{ind}}] e_i) \\
&= -\mathbb{E}[x_1 e_i^{\text{ind}}] e_i^{\text{ind}}.
\end{aligned}$$

A similar calculation yields that

$$\rho(\langle x_2, e_i^{\text{ind}} \rangle e_i^{\text{ind}} + x_2^{\perp}) = -\mathbb{E}[x_2 e_i^{\text{ind}}] e_i^{\text{ind}}.$$

Then,

$$
\begin{aligned}
&\langle \rho(\langle x_1, e_i^{\text{ind}} \rangle e_i^{\text{ind}} + x_1^{\perp}), e_i^{\text{ind}} \rangle - \langle \rho(\langle x_2, e_i^{\text{ind}} \rangle e_i^{\text{ind}} + x_2^{\perp}), e_i^{\text{ind}} \rangle \\
&= \mathbb{E}[x_2 e_i^{\text{ind}}] e_i^{\text{ind}} - \mathbb{E}[x_1 e_i^{\text{ind}}] e_i^{\text{ind}} \\
&= (x_2 - x_1)\mathbb{E}[e_i^{\text{ind}}] \\
&= (x_2 - x_1)\mathbb{E}\left[\frac{\mathbb{1}_{B_i}}{\sqrt{\mathbb{P}(B_i)}}\right] \\
&= (x_2 - x_1)\sqrt{\mathbb{P}(B_i)} \\
&> 0,
\end{aligned}
$$

where the strict inequality holds since $x_1 \neq x_2$ and $\mathbb{P}(B_i) > 0$. Hence,

$$
X \mapsto \langle \rho(\langle X, e_i^{\text{ind}} \rangle e_i^{\text{ind}} + X^{\perp}), e_i^{\text{ind}} \rangle
$$

is non-constant.

Now, we are ready to give skectch of the result.

**Conjecture 3.4.7.** Assume that $\rho$ is naturally quasiconvex with respect to $\preceq$, continuous, local with respect to a basis $\{e_i \colon i \in I\}$ and Assumption 3.4.5 holds. Then, $\rho$ is convex with respect to $\preceq$.

**Sketch of Proof.** Since $\rho$ is naturally quasiconvex, in view of Theorem 3.2.3 $\rho$ is $\star$-quasiconvex. In other words, for every $Y \in C^+$

$$
X \mapsto \mathbb{E}[\rho(X)Y] \tag{3.4.11}
$$

is quasiconvex on $L^2(\mathcal{F})$. Observe that

$$
\begin{aligned}
\mathbb{E}[\rho(X)Y] &= \left\langle \rho(X), \sum_{i\in I} \langle Y, e_i \rangle e_i \right\rangle \\
&= \sum_{i\in I} \langle Y, e_i \rangle \langle \rho(X), e_i \rangle \\
&= \sum_{i\in I} \langle Y, e_i \rangle \langle \rho(\langle X, e_i \rangle e_i + X^{\perp}), e_i \rangle
\end{aligned} \tag{3.4.12}
$$

where the third equality holds since $\rho$ is local with respect to the basis $\{e_i \colon i \in I\}$. Furthermore, we define for every $Y \in C^+$, $s_Y \colon \bigtimes \mathcal{X}_i \to \overline{\mathbb{R}}$ by

$$s_Y(X_1, X_2, \ldots) := \begin{cases} \sum_{i\in I} \langle Y, e_i\rangle f_i(X_i) & \text{if } (X_1, X_2, \ldots) \in S \\ +\infty & \text{otherwise} \end{cases}. \tag{3.4.13}$$

In view of (3.4.11) and (3.4.12), $s_Y$ is quasiconvex on $\bigtimes_{i\in I} \mathcal{X}_i$ for every $Y \in C^+$. Note that $f_i$ is continuous since it is composition of two continuous function for each $i \in I$. In particular, $f_i$ is lower semi-continuous. By applying Theorem 2.3.3 to (3.4.13) (see Remark 3.4.8), we can conclude that either one of the following two cases holds.

(a) $X_i \mapsto \langle Y, e_i\rangle f_i(X_i)$ is convex for each $i \in I$.

(b) $X_i \mapsto \langle Y, e_i\rangle f_i(X_i)$ is convex for all $i \in I$ except one and

$$\sum_{i\in I} \frac{1}{c(\langle Y, e_i\rangle f_i)} \leq 0. \tag{3.4.14}$$

Supoose that (b) holds. It follows from Theorem 2.1.4 that $c(f_i) \geq 0$ for all except for some $i_0 \in I$. It follows from Lemma 2.1.5 and (3.4.14), we have

$$\sum_{i\in I} \frac{\langle Y, e_i\rangle}{c(f_i)} \leq 0. \tag{3.4.15}$$

Clearly, $e_i \in C$ for every $i \in I$. Moreover, since $\{e_i \colon i \in I\}$ is an orthonormal basis, $e_i \in C^+$ for each $i \in I$. Let us fix some $i_1 \in I \setminus \{i_0\}$ and set $Y = e_{i_1}$. Then,

$$\sum_{i\in I} \frac{\langle Y, e_i\rangle}{c(f_i)} = \frac{\langle e_{i_1}, e_{i_1}\rangle}{c(f_{i_1})} = \frac{1}{c(f_{i_1})} > 0, \tag{3.4.16}$$

where the first equality follows from orthonormality of $\{e_i \colon i \in I\}$. Since (3.4.16) is in contradiction with (3.4.15), we can eliminate this case.

Therefore, case (a) holds. Since one can choose $Y$ such that $\langle Y, e_i\rangle > 0$ and division of a real-valued convex function with a positive real number is convex ,

$f_i$ is convex on $\mathcal{X}_i$. Then, in view of locality of $\rho$, for every $X, V \in L^2(\mathcal{F})$ and $\lambda \in [0,1]$

$$\langle \rho(\lambda X + (1-\lambda)V), e_i \rangle \leq \lambda \langle \rho(X), e_i \rangle + (1-\lambda)\langle \rho(V), e_i \rangle$$

for each $i \in I$. In other words, $f_i$ is convex with respect to $\preceq$ for each $i \in I$. Consequently, $\rho$ is convex with respect to $\preceq$ on $L^2(\mathcal{F})$. $\square$

**Remark 3.4.8.** In the proof of Theorem 3.4.7, $s_Y$ is an additively decomposable sum if and only if $\sum_{i \in I} \langle Y, e_i \rangle f_i(X_i) = +\infty$ for every $(X_1, X_2 \ldots) \in \bigtimes_{i \in I} \mathcal{X}_i \setminus S$. However, we don't have such a result, hence $s_Y$ is not an additively decomposable sum. Moreover, if $L^2$ can be represented as an infinite product of topological vector spaces, then we can have an additively decomposable infinite sum by defining $s_Y$ on that representation. However, there is no such representation of $L^2(\mathcal{F})$. We propose the following two ways, which needs to be worked on, to solve that problem:

(i) If Theorem 2.3.3 generalized to a case where $s$ converges only on a convex subset $S$ of $\bigtimes_{i=1}^{+\infty} X_i$, then this generalization can be applied to $s_Y$ to get the results of Theorem 3.4.7.

(ii) If $\rho$ can be approximated as a limit of decomposable sums defined on finite product spaces and natural quasiconvexity of $\rho$ implies natural quasiconvexity of each additively decomposable sum, then one can show that $\rho$ is convex by using results of Section 2.2 and a limit argument.

Lastly, to apply theorem Theorem 2.3.3 to $s_Y$, $s_Y$ must be uniformly absolutely convergent. Either the uniformly absolutely convergence assumption in Theorem 2.3.3 needs to relaxed or uniformly absolutely convergence of $s_Y$ should be proved in order to complete the proof of Theorem 3.4.7.

The following figure is a summary of Chapter 3.

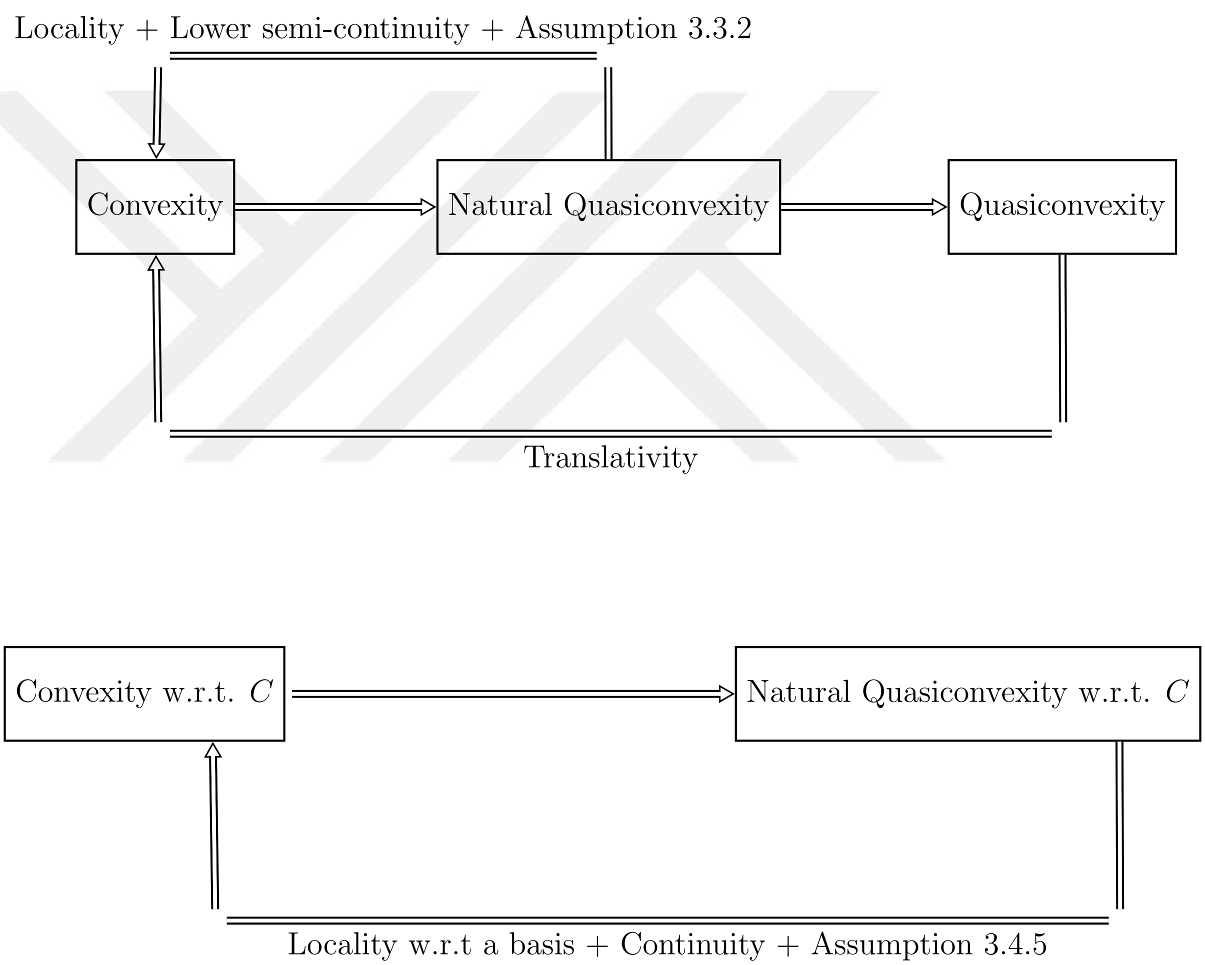


Figure 3.2: Summary of Chapter 3

# Chapter 4

# Conclusion

In this thesis, we investigate naturally quasiconvex risk measures. In particular, we discuss connections between natural quasiconvexity and convexity, locality and translativity properties for risk measures.

In Debreu and Koopmans [16] and Crouzeix and Lindberg [17], additively decomposable sums defined on open subsets of $\mathbb{R}^n$ are studied. We generalize their results to additively decomposable functions defined on general topological vector spaces. To achieve this generalization, in Section 2.1 we define convexity index for real-valued functions that are defined on a topological vector space. We show that all of the results regarding convexity indices for functions defined on $\mathbb{R}^n$ also holds for real-valued functions defined on a general topological vector space. In section 2.2, Theorem 2.2.5, generalizes the main result in Debreu and Koopmans [16] and Crouzeix and Lindberg [17] to general topological vector spaces. In particular, we show that quasiconvexity of an additively decomposable sum have either all coordinate functions convex or at most one is convex together with a condition on the sum of reciprocals of convexity indices of each coordinate function. To make the generalizations in Section 2.1 and 2.2, all we need to assume is $f$ is lower semicontinuous. In Theorem 2.3.3 of Section 2.3, we extend the main result of the previous section to additively decomposable infinite sums, which was not available in the aforementioned papers, even for the functions defined on $\mathbb{R}^n$. Study of additively decomposable infinite sums pave the way for the analysis we have in Section 3.4.

In Chapter 3, Section 3.1 and 3.2 are brief introductions to risk measures, natural quasiconvexity and $\star$-quasiconvexity. An important result in these sections is Theorem 3.2.3 which states that natural quasiconvexity and $\star$-quasiconvexity are equivalent properties. The most remarkable result in this thesis is Theorem 3.3.4 in Section 3.3, where we show that convexity and natural quasiconvexity is equivalent for local risk measures under some mild conditions. We also present a numeric example in Example 3.3.1 to give the intuition behind this theorem. In the last section, we are concerned with risk measures on $L^2$ and we define a new property called locality with respect to a basis. We showed in Example 3.4.2 that conditional entropic risk measure, which is a well-known example of a convex risk measure, is local with respect to a basis. Relationships between locality with respect to a basis with locality and translativity are given in Theorem 3.4.3 and Theorem 3.4.4, respectively. Lastly, we give a sketch of the proof of Theorem 3.4.7, which is concerned with the connection between convexity and natural convexity for risk measures that are local with respect to a basis.

# Bibliography


[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical finance*, vol. 9, no. 3, pp. 203–228, 1999.

[2] H. Föllmer and A. Schied, "Convex measures of risk and trading constraints," *Finance and stochastics*, vol. 6, no. 4, pp. 429–447, 2002.

[3] M. Frittelli and E. R. Gianin, "Putting order in risk measures," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1473–1486, 2002.

[4] D. Heath, "Back to the future. plenary lecture," in *First World Congress of the Bachelier Finance Society, Paris*, 2000.

[5] N. El Karoui and C. Ravanelli, "Cash subadditive risk measures and interest rate ambiguity," *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, vol. 19, no. 4, pp. 561–590, 2009.

[6] M. Marinacci and L. Montrucchio, "Introduction to the mathematics of ambiguity," 2003.

[7] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, "Complete monotone quasiconcave duality," *Mathematics of Operations Research*, vol. 36, no. 2, pp. 321–339, 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, "Dual representation of quasi-convex conditional maps," *SIAM Journal on Financial Mathematics*, vol. 2, no. 1, pp. 357–382, 2011.

[10] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, "Risk measures: rationality and diversification," *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, vol. 21, no. 4, pp. 743–774, 2011.

[11] J. Bion-Nadal, "Conditional risk measures and robust representation of convex risk measures," *Ecole Polytechnique, CMAP*, no. 557, 2004.

[12] K. Detlefsen and G. Scandolo, "Conditional and dynamic convex risk measures," *Finance and Stochastics*, vol. 9, no. 4, pp. 539–561, 2005.

[13] F. Riedel, "Dynamic coherent risk measures," *Stochastic processes and their applications*, vol. 112, no. 2, pp. 185–200, 2004.

[14] M. Frittelli and E. R. Gianin, "Dynamic convex risk measures," *Risk measures for the 21st century*, pp. 227–248, 2004.

[15] A. Ruszczyński and A. Shapiro, "Conditional risk mappings," *Mathematics of operations research*, vol. 31, no. 3, pp. 544–561, 2006.

[16] G. Debreu and T. C. Koopmans, "Additively decomposed quasiconvex functions," *Mathematical Programming*, vol. 24, no. 1, pp. 1–38, 1982.

[17] J.-P. Crouzeix and P. O. Lindberg, "Additively decomposed quasiconvex functions," *Mathematical Programming*, vol. 35, no. 1, pp. 42–57, 1986.

[18] T. Tanaka, "Generalized quasiconvexities, cone saddle points, and minimax theorem for vector-valued functions," *Journal of Optimization Theory and Applications*, vol. 81, no. 2, pp. 355–377, 1994.

[19] V. Jeyakumar, W. Oettli, and M. Natividad, "A solvability theorem for a class of quasiconvex mappings with applications to optimization," *Journal of Mathematical Analysis and Applications*, vol. 179, pp. 537–537, 1993.

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

[21] E. Çınlar, *Probability and stochastics*, vol. 261. Springer Science & Business Media, 2011.

[22] J. K. Hunter and B. Nachtergaele, *Applied analysis.* World Scientific Publishing Company, 2001.