A simplified conjugation scheme for lower semi-continuous functions

Journal:	*Optimization*
Manuscript ID:	GOPT-2015-0013.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	22-May-2015
Complete List of Authors:	Elias, Leonardo; Universidade Federal do Paraná, Martinez-Legaz, Juan Enrique; Universitat Autonoma de Barcelona, ;
Keywords:	Lower semi-continuous function, Generalized convex conjugation, Optimization duality theory
2010 Mathematics Subject Classification:	90C26, 26B25

A simplified conjugation scheme for lower semi-continuous functions

Leonardo M. Elias*
Programa de Pós-Graduação em Matemática
Universidade Federal do Paraná
CP 19081, 81531-980 Curitiba, PR
Brazil

Juan E. Martínez-Legaz†
Departament d'Economia i d'Història Econòmica
Universitat Autònoma de Barcelona
08193 Bellaterra
Spain

Abstract

We present two generalized conjugation schemes for lower semi-continous functions defined on a real Banach space whose norm is Fréchet differentiable off the origin, and sketch their applications to optimization duality theory. Both approaches are based upon a new characterization of lower semi-continuous functions as pointwise suprema of a special class of continuous functions.

Keywords: Lower semi-continuous function; generalized convex conjugation; optimization duality theory

1 Introduction

This paper elaborates on a conjugation theory for lower semi-continuous functions introduced in a recent paper [1], in which, for functions defined on $\mathbb{R}^n$,

*This work was developed during a visit of this author to the Departament d'Economia i d'Història Econòmica of the Universitat Autònoma de Barcelona. He gratefully ackowledges the financial support received from CAPES and the warm hospitality of the second author as well as of the members from the host institution. This research was supported by CAPES, under grant BEX 3843/14-9. This author is affiliated to CAPES Foundation, Ministry of Education of Brazil, Brasília - DF 70040-020, Brazil.

†This author was supported by the MICINN of Spain, Grant MTM2011-29064-C03-01, and under Australian Research Council's Discovery Projects funding scheme (project number DP140103213). He is affiliated to MOVE (Markets, Organizations and Votes in Economics).

the authors present a conjugation scheme such that the second conjugate of any function coincides with the pointwise supremum of its minorants of the type $x \longmapsto \langle x, p(x) \rangle + r$, with $\langle \cdot, \cdot \rangle$, $p : \mathbb{R}^n \to \mathbb{R}^n$ and r denoting the Euclidean scalar product, a continuous mapping, and a constant, respectively [1, Proposition 3.2(iii)]. To simplify that conjugation scheme, we first characterize this special class of functions: In the case of a real Banach space X whose norm is Fréchet differentiable off the origin, a function admits such a representation if, and only if, it is continuous on X and Fréchet differentiable at the origin (Theorem 4). Thanks to this characterization, we can simplify the conjugation theory of [1] by considering a "dual" space consisting of real valued functions on X rather than continuous mappings $p : X \to X$, the latter being, in general, much more complex mathematical objects than the former. Notice that a continuous mapping $p : X \to X$ yields a unique function of the type $x \longmapsto \langle x, p(x) \rangle$, but there are in general infinitely many such continuous mappings p yielding the same function. Consequently, dealing with functions $X \to \mathbb{R}$ rather than mappings $X \to X$, we drastically reduce the dimension of the "dual" space. We present two alternative conjugation schemes. In the first one, the "dual" variables are pairs consisting of continuous linear functionals and general continuous real valued functions; the consideration of continuous linear functionals allows us to relate our conjugation operator to classical convex conjugation. The second scheme is simpler, in that the arguments of the conjugate functions are just continuous real valued functions; however, since the dual space X^* does not appear in the picture, this second scheme is not directly related to Fenchel conjugation. We briefly sketch the application of our conjugation schemes to optimization duality theorem.

The rest of the paper is organized as follows. Section 2 presents the fundamental results on which our conjugation approaches are based. One of its main results states that every lower semi-continuous function is the pointwise supremum of a collection of continuous functions which are Fréchet differentiable at the origin and whose gradients at that point are equal to 0. In Section 3 we develop our first conjugation scheme and show its relationship with classical convex conjugation as well as its application to optimization duality theory. Section 4 presents our second approach to generalized conjugation for lower semi-continuous functions and the duality theory based on it. In Section 5 we illustrate both duality schemes by means of simple examples.

2 Preliminaries

Throughout this paper, we will assume that $(X, \|\cdot\|)$ is a real Banach space whose norm is Fréchet differentiable off the origin. Hilbert spaces are examples of such spaces; moreover, if $(X, \|\cdot\|)$ is of that type, μ is a measure on X, and $p \in (1, \infty)$, then the norm of $L^p(X, \mu)$ has that differentiability property, too [3, Theorem 3.1]. We will denote by $(X^*, \|\cdot\|_*)$ the dual space of X, and by $\langle \cdot, \cdot \rangle : X \times X^* \to \mathbb{R}$ the duality pairing. We recall that the duality mapping

$J : X \rightrightarrows X^*$ is $J := \partial \left(\frac{1}{2} \|\cdot\|^2 \right)$ or, equivalently,

$$J(x) := \left\{ x^* \in X^* \mid \|x\|^2 = \langle x, x^* \rangle = \|x^*\|_*^2 \right\} \qquad (x \in X).$$

Since a convex function is Fréchet differentiable precisely when its subdifferential is single-valued and norm-to-norm continuous (see, e.g., [6, Introduction]), our differentiability assumption on $\|\cdot\|$ is equivalent to the single-valuedness and norm-to-norm continuity of J. For the sake of notational simplicity, we will identify the singleton $J(x)$ with its element; under this identification, $J(x)$ is characterized by the equalities

$$\|x\|^2 = \langle x, J(x) \rangle = \|J(x)\|_*^2. \tag{1}$$

The following classical theorem will play a key role in the generalized conjugation theory we will develop.

Theorem 1 *[10, p. 133] If E is a metric space, $f : E \to \mathbb{R} \cup \{+\infty\}$ is lsc, $g : E \to \mathbb{R} \cup \{-\infty\}$ is usc, and $g \leq f$, then there exists a continuous function $h : E \to \mathbb{R}$ such that $g \leq h \leq f$.*

The lsc hull (that is, the largest semi-continuous minorant) of a function f will be denoted $cl f$. As is well known, f is lsc at a point x if, and only if, $cl f(x) = f(x)$.

Corollary 2 *Let E be a metric space, $f : E \to \mathbb{R} \cup \{+\infty\}$, and $x \in E$. Then f is lsc at x if, and only if,*

$$f(x) = \sup \{ h(x) \mid h : E \to \mathbb{R} \text{ is continuous, } h \leq f \}, \tag{2}$$

and the supremum is attained if $x \in dom(f)$.

Proof. Let as assume that f is lsc at x. We will prove that the inequality

$$f(x) \geq \sup \{ h(x) \mid h : E \to \mathbb{R} \text{ is continuous, } h \leq f \}$$

holds with the equal sign and the supremum in the right hand side of this inequality is attained if $x \in dom(f)$, by showing that, for every real number $\lambda \leq f(x)$, there exists a continuous function $h_\lambda : E \to \mathbb{R}$ satisfying $h_\lambda \leq f$ and $h_\lambda(x) \geq \lambda$. Define $g_\lambda : E \to \mathbb{R} \cup \{-\infty\}$ by $g_\lambda(x) := \lambda$ and $g_\lambda(y) := -\infty$ for $y \neq x$. Note that g_λ is usc and satisfies $g_\lambda(y) \leq cl f(y)$ for all $y \in E$, since we have $cl f(x) = f(x)$ by the lower semi-continuity of f at x. Hence, by Theorem 1, there exists a continuous function $h : E \to \mathbb{R}$ such that

$$g_\lambda(y) \leq h_\lambda(y) \leq cl f(y) \text{ for all } y \in E.$$

Therefore $h_\lambda \leq f$ and $\lambda = g_\lambda(x) \leq h_\lambda(x)$.

Conversely, assume that (2) holds. Since the function

$$y \longmapsto \sup \{ h(y) \mid h : E \to \mathbb{R} \text{ is continuous, } h \leq f \}$$

is an lsc minorant of f, it is a minorant of $cl f(y)$, too. Hence, by (2), we have

$$cl f(x) = f(x),$$

that is, f is lsc at x. ■

The following theorem is a complement to the Fenchel-Moreau conjugation theory developed in [1, Section 4]. It can be regarded as a generalized subdifferentiability result.

Theorem 3 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ and $x \in dom(f) \setminus \{0\}$. Then f is lsc at x if, and only if, there exists a continuous mapping $p : X \to X^*$ satisfying $p(0) = 0$ and*

$$f(y) \geq f(x) + \langle y, p(y) \rangle - \langle x, p(x) \rangle \text{ for all } y \in X. \tag{3}$$

Proof. If f is lsc at x, then, by Corollary 2, there exists a continuous function $h : X \to \mathbb{R}$ such that $h \leq f$ and $h(x) = f(x)$. Let $\delta := \inf_{\|y\| \leq \|x\|} h(y)$. We define $p : X \to X^*$ by

$$p(y) := \frac{h(y) - \delta}{(\max\{\|y\|, \|x\|\})^2} J(y).$$

Clearly, p is continuous and satisfies

$$\langle y, p(y) \rangle = (h(y) - \delta) \left(\min \left\{ \frac{\|y\|}{\|x\|}, 1 \right\} \right)^2.$$

Therefore, $\langle y, p(y) \rangle \leq h(y) - \delta \leq f(y) - \delta$ for all $y \in X$, which implies $\langle y, p(y) \rangle - f(y) \leq -\delta$ for all $y \in X$. Furthermore, $\langle x, p(x) \rangle = h(x) - \delta = f(x) - \delta$. Consequently, (3) holds.

Conversely, assume that (3) holds, and define $h_0 : E \to \mathbb{R}$ by

$$h_0(y) := f(x) + \langle y, p(y) \rangle - \langle x, p(x) \rangle.$$

Since h_0 is continuous and, by (3), we have $h_0 \leq f$, it turns out that

$$\sup \{h(x) \mid h : E \to \mathbb{R} \text{ is continuous, } h \leq f\} \geq h_0(x) = f(x).$$

Hence, (2) holds, and therefore the lower semi-continuity of f at x follows from Corollary 2. ■

Theorem 3 shows the role of functions of the type $x \longmapsto \langle x, p(x) \rangle$, with p continuous. In fact, the generalized conjugation theory developed in [1] is built on such functions. Thanks to the following characterization, in the next section we will present a new, simplified version of that theory.

Theorem 4 *Let $f : X \to \mathbb{R}$ be a function. The following statements are equivalent:*

(i) *There exists a continuous mapping* $p : X \to X^*$ *such that* $f(x) = \langle x, p(x)\rangle$ *for every* $x \in X$*.*

(ii) *The function* f *is continuous on* X *and Fréchet differentiable at* 0*, and satisfies* $f(0) = 0$*.*

Proof. $(i) \Rightarrow (ii)$ Continuity and the equality $f(0) = 0$ are obvious, whereas Fréchet differentiability follows from the continuity of p, given that

$$0 \leq \frac{|f(x) - f(0) - \langle x, p(0)\rangle|}{\|x\|} = \frac{|\langle x, p(x) - p(0)\rangle|}{\|x\|} \leq \|p(x) - p(0)\|_* .$$

$(ii) \Rightarrow (i)$ Define $p : X \to X^*$ by

$$p(x) = \begin{cases} \nabla f(0) + \frac{f(x) - \langle x, \nabla f(0)\rangle}{\|x\|^2} J(x), & \text{if } x \neq 0 \\ \nabla f(0), & \text{otherwise.} \end{cases} \tag{4}$$

Since f is continuous, so is p on $X \backslash \{0\}$. Continuity at 0 also holds, in view of (1):

$$\begin{aligned} \lim_{x \to 0} \|p(x) - p(0)\|_* &= \lim_{x \to 0} \frac{|f(x) - \langle x, \nabla f(0)\rangle|}{\|x\|^2} \|J(x)\|_* \\ &= \lim_{x \to 0} \frac{|f(x) - \langle x, \nabla f(0)\rangle|}{\|x\|} = 0. \end{aligned}$$

For $x \in X \setminus \{0\}$, using again (1) we obtain

$$\begin{aligned} \langle x, p(x)\rangle &= \langle x, \nabla f(0) + \frac{f(x) - \langle x, \nabla f(0)\rangle}{\|x\|^2} J(x)\rangle \\ &= \langle x, \nabla f(0)\rangle + \frac{f(x) - \langle x, \nabla f(0)\rangle}{\|x\|^2} \langle x, J(x)\rangle = f(x), \end{aligned}$$

whereas the equality $\langle x, p(x)\rangle = f(x)$ also holds, obviously, for $x = 0$. ∎

While preparing a revised version of this paper, we have learned that a statement equivalent to Theorem 4 can be found in Theorem 3.11 of the very recent article [5].

Corollary 5 *Let* $f : X \to \mathbb{R} \cup \{+\infty\}$ *and* $x \in X$*. Then* f *is lsc at* x *if, and only if,*

$$f(x) = \sup\{h(x) \mid h : X \to \mathbb{R} \text{ is continuous on } X \text{ and Fréchet differentiable at } 0,\ \nabla h(0) = 0,\ h \leq f\}; \tag{5}$$

the supremum is attained if $x \in dom(f) \backslash \{0\}$*.*

Proof. The "if" statement follows from the corresponding statement of Corollary 2.

Conversely, assume that f is lsc at x. In view of Corollary 2, in order to prove (5) we may assume, without loss of generality, that f is continuous on X. Moreover, since $f = \sup_{\lambda \in \mathbb{R}} f_\lambda$, with

$$f_\lambda := \min \{f, \lambda\},$$

we can further assume that f is finite valued.

Let us consider first the case when $x = 0$. By the continuity of f at 0, if $\lambda < f(0)$, the function f_λ takes the constant value λ on a neigborhood of 0, and hence f_λ is Fréchet differentiable with $\nabla f_\lambda (0) = 0$. Therefore, (5) follows from the equality $f(0) = \sup \{f_\lambda (0) \mid \lambda < f(0)\}$.

If $x \neq 0$, then, taking p as in Theorem 3 and defining $h : X \to \mathbb{R}$ by $h(y) := \langle y, p(y) \rangle + f(x) - \langle x, p(x) \rangle$, from Theorem 4, implication (i) $\Longrightarrow$ (ii), it follows that h is continuous on X, Fréchet differentiable at 0, and satisfies $\nabla h(0) = 0$; moreover, by (3), we have $h \leq f$. Therefore, since $h(x) = f(x)$, this proves that the supremum in (5) is attained at this precise h and that equality (5) holds. ■

Remark 6 *Notice that, when $\nabla f(0) = 0$, the mapping p defined by (4) is pointwise norm minimizing, that is, for every $x \in X$ the least norm solution of the linear equation $\langle x, y^* \rangle = f(x)$ is precisely $p(x)$. This is obvious for $x = 0$, whereas for $x \neq 0$ the assertion follows by observing that, for any solution y^*, in view of (1) one has*

$$\|p(x)\|_* = \frac{|f(x)|}{\|x\|^2} \|J(x)\|_* = \frac{|\langle x, y^* \rangle|}{\|x\|} \leq \|y^*\|_* .$$

When $\nabla f(0) \neq 0$, the mapping p of (4) is obtained by adding the constant mapping $x \longmapsto \nabla f(0)$ to the pointwise norm minimizing mapping corresponding to $g := f - \nabla f(0)$.

In view of Theorem 4, Corollary 5 may be regarded as a refinement of [11, Theorem 6.1] (see also [5, Theorem 3.11]), which characterizes lsc functions on a Hilbert space as pointwise suprema of functions of the type $x \mapsto \langle x, p(x) \rangle + \beta$, with $p : X \to X^*$ continuous and β constant. Our set of minorants h is smaller, since their gradients vanish at the origin, a condition which is absent in the statement of [11, Theorem 6.1] (as well as in that of [5, Theorem 3.11]); indeed, in the formulation of [11, Theorem 6.1], it would read $p(0) = 0$.

It is worth pointing out that the special role that the origin plays in Theorem 4 and (5) can be equally played by any other point $y \in X$, as the next result state. However, for the sake of simplicity, in the next sections we will state all of our results only for the case when $y := 0$.

Corollary 7 *Let $f : X \to \mathbb{R}$ be a continuous function and $y \in X$. Then f is Fréchet differentiable at y and satisfies $f(y) = 0$ if, and only if, there exists a continuous mapping $p : X \to X^*$ such that $f(x) = \langle x - y, p(x) \rangle$ for every $x \in X$.*

Corollary 8 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ and $x, y \in X$. Then f is lsc at x if, and only if,*

$$f(x) = \sup\{h(x) \mid h : X \to \mathbb{R} \text{ is continuous on } X \text{ and Fréchet differentiable at } y,\ \nabla h(y) = 0,\ h \leq f\};$$

the supremum is attained if $x \in \operatorname{dom}(f) \setminus \{y\}$.

The proofs of the two latter corollaries easily follow from Theorem 4 and Corollary 5, replacing f by $g := f(\cdot + y)$ in their statements.

In the case when X is a Hilbert space, after identifying it with its dual by means of the Riesz-Fréchet Representation Theorem, it turns out that J is the identity mapping. Therefore, by a straightforward modification of the proof of Theorem 4, we get the following result, which provides two characterizations of everywhere differentiable functions on Hilbert spaces.

Theorem 9 *If H is a real Hilbert space and $y \in H$, for $f : H \to \mathbb{R}$ the following statements are equivalent:*

(i) There exists a continuous mapping $p : H \to H$, Fréchet differentiable on $H \setminus \{y\}$, such that $f(x) = \langle x - y, p(x) \rangle$ for every $x \in H$.

(ii) The function f is Fréchet differentiable and satisfies $f(y) = 0$.

3 The first conjugation scheme

In this section we will develop our first conjugation theory for lsc functions. We will follow the approach to generalized conjugation theory presented in [4, Section 2].

We introduce the following set of functions

$$\mathcal{H}_X := \{h : X \to \mathbb{R} \mid h \text{ is continuous on } X \text{ and Fréchet differentiable at } 0,\ h(0) = 0,\ \nabla h(0) = 0\}$$

and the coupling function $c : X \times (X^* \times \mathcal{H}_X) \to \mathbb{R}$ defined by

$$c(x, (x^*, h)) := \langle x, x^* \rangle + h(x).$$

The c-conjugate function of $f : X \to \overline{\mathbb{R}}$ is $f^c : X^* \times \mathcal{H}_X \to \overline{\mathbb{R}}$, given by

$$f^c(x^*, h) = \sup_{x \in X} \{\langle x, x^* \rangle + h(x) - f(x)\}. \tag{6}$$

Notice that the classical Fenchel conjugation operator * is related to c-conjugation by the following formulas:

$$f^c(x^*, h) = (f - h)^* (x^*), \qquad f^*(x^*) = f^c(x^*, 0). \tag{7}$$

The "inverse" conjugation operator corresponds to the coupling function c' : $(X^* \times \mathcal{H}_X) \times X \to \mathbb{R}$ defined by

$$c'((x^*, h), x) := c(x, (x^*, h)).$$

We say that $f : X \to \overline{\mathbb{R}}$ is c-subdifferentiable at $x \in dom(f)$ if there exists $(x^*, h) \in X^* \times \mathcal{H}_X$ such that

$$f(y) - f(x) \geq \langle y - x, x^* \rangle + h(y) - h(x) \text{ for all } y \in X. \tag{8}$$

Then (x^*, h) is said to be a c-subgradient of f at x. The set of all c-subgradients of f at x, denoted $\partial_c f(x)$, is called the c-subdifferential of f at x. If $f(x) \notin \mathbb{R}$, we set $\partial_c f(x) := \emptyset$. The relationship between the c-subdifferential and the classical Fenchel subdifferential ∂ is as follows:

$$(x^*, h) \in \partial_c f(x) \iff x^* \in \partial (f - h)(x), \qquad x^* \in \partial f(x) \iff (x^*, 0) \in \partial_c f(x). \tag{9}$$

The c'-conjugate function $\varphi^{c'} : X \to \overline{\mathbb{R}}$ of $\varphi : X^* \times \mathcal{H}_X \to \overline{\mathbb{R}}$ is defined by

$$\varphi^{c'}(x) = \sup_{(x^*, h) \in X^* \times \mathcal{H}_X} \{\langle x, x^* \rangle + h(x) - \varphi(x^*, h)\}. \tag{10}$$

This c'-conjugate function relates to Fenchel conjugation according to the following formula, which easily follows from (10):

$$\varphi^{c'}(x) = \sup_{h \in \mathcal{H}_X} \left\{h(x) + \varphi(\cdot, h)^*(x)\right\}. \tag{11}$$

The biconjugate of f is $f^{cc'} := (f^c)^{c'}$. Combining (11) with (7), one gets

$$f^{cc'}(x) = \sup_{h \in \mathcal{H}_X} \left\{h(x) + (f - h)^{**}(x)\right\}. \tag{12}$$

For a full description of the generalized conjugation framework we are using, we refer to [4]; more details on abstract convexity can be found in [8, 9].

The following result is an easy consequence of [4, Proposition 6.2].

Theorem 10 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$. Then*

$$f^{cc'} = \sup\{g : X \to \mathbb{R} \mid g \textit{ is continuous on } X \textit{ and Fréchet differentiable at } 0,\ g \leq f\}.$$

Combining Theorem 10 with Corollary 5, we obtain the following result.

Corollary 11 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ and $x \in X$. Then f is lsc at x if, and only if, $f^{cc'}(x) = f(x)$. Moreover, if $x \in dom(f) \setminus \{0\}$ and f is lsc at x, then $\partial_c f(x) \neq \emptyset$.*

Combining Corollary 11 with (12), and using the well known fact that every lsc proper convex function f coincides with its second Fenchel conjugate f^{**}, one immediately sees that the supremum in (12) is attained at $h = 0$.

The next proposition relates the c-subdifferential at the origin to the Fréchet subdifferential. We recall that $x^* \in X^*$ is a Fréchet (or regular, following the terminology of [7]) subgradient of f at $x \in dom\,(f)$ if

$$\liminf_{y \to x} \frac{f\,(y) - f\,(x) - \langle y - x, x^* \rangle}{\|y - x\|} \geq 0.$$

The set of all such Fréchet subgradients is called the Fréchet subdifferential of f at x; we will denote this set by $\partial_F\,(x)$.

Proposition 12 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be such that $0 \in dom\,(f)$. Then the projection of $\partial_c f(0)$ onto X^* coincides with $\partial_F\,(0)$.*

Proof. If x^* belongs to the projection of $\partial_c f(0)$ onto X^*, then there exists $h \in \mathcal{H}_X$ such that (8) holds for $x = 0$, which means that the function $g := x^* + h + f\,(0)$ is a minorant of f. Since $g\,(0) = f\,(0)$ and $\nabla g\,(0) = x^*$, by [7, Proposition 8.5] we have $x^* \in \partial_F f\,(0)$.

Conversely, let $x^* \in \partial_F f\,(0)$. Then, again by [7, Proposition 8.5], on some closed neighborhhood V of 0 there is a smooth function $g \leq f$ with $g\,(0) = f\,(0)$ and $\nabla g\,(0) = x^*$. We clearly have $g \leq cl f$. We extend g to the whole of X by setting $g\,(y) := -\infty$ for $y \in X \setminus V$. This preserves the upper semi-continuity of g as well as the inequality $g \leq cl f$. Hence, by Theorem 1, there exists a continuous function $h : X \to \mathbb{R}$ such that $g \leq h \leq cl f$. We now use the Tietze extension theorem to prove the existence of a nonnegative continuous extension $j : X \to \mathbb{R}$ of the restriction of $h - g$ to V. Thus $\widetilde{g} := h - j$ is a continuous extension of the restriction of g to V, and we have $\widetilde{g} \leq h \leq cl f \leq f$ as well as $\widetilde{g}\,(0) = g\,(0) = f\,(0)$ and $\nabla \widetilde{g}\,(0) = \nabla g\,(0) = x^*$. Therefore, the function $h := \widetilde{g} - x^* - f\,(0)$ belongs to $\mathcal{H}_X$. Since $f \geq \widetilde{g} = f\,(0) + x^* + h$, it follows that $(x^*, h) \in \partial_c f(0)$, which shows that x^* belongs to the projection of $\partial_c f(0)$ onto X^*. ■

Corollary 13 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be such that $0 \in dom\,(f)$. Then $\partial_c f(0) \neq \emptyset$ if, and only if, $\partial_F\,(0) \neq \emptyset$.*

In particular, from Corollary 13 it follows that if f is Fenchel subdifferentiable or Fréchet differentiable at the origin, then $\partial_c f(0) \neq \emptyset$.

We will now apply this generalized conjugation scheme to duality theory, following the approach of [4, Section 3].

Let S be an arbitrary set, and let $\phi : S \times X \to \overline{\mathbb{R}}$ be a perturbed objective function. The optimization problem under consideration is

$$(\mathcal{P}) \qquad \text{minimize } \phi\,(s, 0)\,; \tag{13}$$

the perturbed problem corresponding to the perturbation parameter $x \in X$ is

$$(\mathcal{P}_x) \qquad \text{minimize } \phi\,(s, x)\,.$$

The associated perturbation function is $p : X \to \overline{\mathbb{R}}$, defined by

$$p(x) := \inf_{s \in S} \phi(s, x).$$

The dual problem to $(\mathcal{P})$ is

$$(\mathcal{D}_c) \qquad \text{maximize } -p^c(x^*, h).$$

A straightforward computation shows that the dual objective function is given by

$$-p^c(x^*, h) = \inf_{x \in X,\ s \in S} \{\phi(s, x) - \langle x, x^* \rangle - h(x)\}.$$

From [4, Theorem 6.7] and Corollary 11, the following result follows.

Theorem 14 *The optimal values of $(\mathcal{P})$ and $(\mathcal{D}_c)$ coincide if, and only if, p is lsc at* 0.

The c-Lagrangian function $L_c : S \times X^* \times \mathcal{H}_X \to \overline{\mathbb{R}}$ is given by

$$L_c(s, x^*, h) = -\phi_s^c(x^*, h),$$

with $\phi_s : X \to \overline{\mathbb{R}}$ denoting the partial mapping $\phi_s(x) := \phi(s, x)$. We thus have

$$L_c(s, x^*, h) = \inf_{x \in X} \{\phi(s, x) - \langle x, x^* \rangle - h(x)\}.$$

For a full description of generalized convex duality, we refer to [4, Section 3].

4 The second conjugation scheme

In this section we will present a conjugation scheme different from the one developed in Section 3. It has the advantage of using a simpler “dual" space, namely, $\mathcal{H}_X$ instead of $X^* \times \mathcal{H}_X$, but its disadvantage is that, unlike in the case of (7), (9), (11) and (12), there is no direct relationship between this new conjugation operator and classical convex conjugation.

We set $d : X \times \mathcal{H}_X \to \mathbb{R}$

$$d(x, h) := h(x)$$

and $d' : \mathcal{H}_X \times X \to \mathbb{R}$

$$d'(h, x) := d(x, h).$$

The d-conjugate function of $f : X \to \overline{\mathbb{R}}$ is $f^d : \mathcal{H}_X \to \overline{\mathbb{R}}$, defined by

$$f^d(h) = \sup_{x \in X} \{h(x) - f(x)\}.$$

We say that $f : X \to \overline{\mathbb{R}}$ is d-subdifferentiable at $x \in dom(f)$ if there exists $h \in \mathcal{H}_X$ such that

$$f(y) - f(x) \geq h(y) - h(x) \text{ for all } y \in X.$$

Then h is said to be a d-subgradient of f at x. The set of all d-subgradients of f at x, denoted $\partial_d f(x)$, is called the d-subdifferential of f at x. If $f(x) \notin \mathbb{R}$, we set $\partial_d f(x) := \emptyset$.

The d'-conjugate function $\varphi^{d'} : X \to \overline{\mathbb{R}}$ of $\varphi : \mathcal{H}_X \to \overline{\mathbb{R}}$ is defined by

$$\varphi^{d'}(x) = \sup_{h \in \mathcal{H}_X} \{h(x) - \varphi(h)\}.$$

We say that $\varphi : \mathcal{H}_X \to \overline{\mathbb{R}}$ is d'-subdifferentiable at $h \in dom(\varphi)$ if there exists $x \in X$ such that

$$\varphi(j) - \varphi(h) \geq j(x) - h(x) \text{ for all } j \in \mathcal{H}_X.$$

Then x is said to be a d'-subgradient of φ at h. The set of all d'-subgradients of φ at h, denoted $\partial_{d'}\varphi(h)$, is called the d'-subdifferential of φ at h. If $\varphi(h) \notin \mathbb{R}$, we set $\partial_{c'}\varphi(h) := \emptyset$.

The biconjugate of f is $f^{dd'} := (f^d)^{d'}$. The following result is an immediate consequence of [4, Proposition 6.2].

Theorem 15 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$. Then*

$$f^{dd'} = \sup\{g : X \to \mathbb{R} \mid g \text{ is continuous on } X \text{ and Fréchet differentiable at } 0,\ \nabla g(0) = 0,\ g \leq f\};$$

Corollary 16 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ and $x \in X$. Then f is lsc at x if, and only if, $f^{dd'}(x) = f(x)$. Moreover, if $x \in dom(f) \setminus \{0\}$ and f is lsc at x, then $\partial_d f(x) \neq \emptyset$.*

We now apply this conjugation scheme to construct a duality theory. We will use the same approach as in Section 3, but with the coupling function d instead of c. So, the dual problem to (13) is now

$$(\mathcal{D}_d) \qquad \text{maximize } -p^d(h).$$

The dual objective function turns out to be

$$-p^d(h) = \inf_{x \in X,\ s \in S} \{\phi(s,x) - h(x)\}.$$

Similarly to Theorem 14, using Corollary 16 we now obtain the following duality theorem.

Theorem 17 *The optimal values of $(\mathcal{P})$ and $(\mathcal{D}_d)$ coincide if, and only if, p is lsc at* 0.

The d-Lagrangian function $L_d : S \times \mathcal{H}_X \to \overline{\mathbb{R}}$ is given by

$$L_d(s,h) = -\phi_s^d(h).$$

A straightforward computation yields

$$L_d(s,h) = \inf_{x \in X} \{\phi(s,x) - h(x)\}.$$

For more details on the relationship between (P) and (D_d) and the way the d-Lagrangian function links these two optimization problems, we again refer to [4, Section 3].

5 Examples

In this section we will use two optimization problems to illustrate the duality schemes proposed in the preceding sections. The first one will be an indefinite quadratic problem $(\mathcal{P})$ with linear constraints, in which the perturbation function is lower semi-continuous at 0 and hence the optimal values of $(\mathcal{P})$, $(\mathcal{D}_c)$ and $(\mathcal{D}_d)$ coincide. The second one will be an example given by Duffin [2] in order to show that a duality gap may occur in convex duality when the Slater condition does not hold; since in classical convex duality the absence of a duality gap is equivalent to the lower semi-continuity of the perturbation function at the origin, our dual problems will also exhibit a duality gap in this example.

Example 18 *Consider the problem*

$$\begin{array}{ll} \textit{minimize} & -s_1^2 + s_2^2 + s_1 s_2 \\ \textit{subject to} & s_1 + s_2 \leq 1 \\ & s_1 - s_2 \leq 0 \\ & s_1 \geq 0,\ s_2 \geq 0, \end{array}$$

and let $\phi : \mathbb{R}^4 \to \mathbb{R} \cup \{+\infty\}$ *be the function defined by*

$$\phi(s_1, s_2, x_1, x_2) := \begin{cases} -s_1^2 + s_2^2 + s_1 s_2, & \text{if } (s_1, s_2, x_1, x_2) \in \Omega, \\ +\infty, & \text{otherwise,} \end{cases}$$

where

$$\Omega := \left\{(s_1, s_2, x_1, x_2) \in \mathbb{R}_+^4 : s_1 + s_2 + x_1 \leq 1,\ s_1 - s_2 + x_2 \leq 0,\ x_1 + x_2 \leq 1\right\}.$$

Note that

$$(\mathcal{P}) \qquad \textit{minimize } \phi\,(s, 0)$$

is equivalent to the given problem. Since, for $(s_1, s_2, x_1, x_2) \in \Omega$, *one has*

$$\begin{aligned} \phi(s_1, s_2, x_1, x_2) &= -s_1^2 + s_2^2 + s_1 s_2 \geq -s_1^2 + (s_1 + x_2)^2 + s_1(s_1 + x_2) \\ &= s_1^2 + 3x_2 s_1 + x_2^2 \geq x_2^2, \end{aligned}$$

the associated perturbation function p *satisfies*

$$p(x_1, x_2) \geq x_2^2 = \phi(0, x_2, x_1, x_2) \geq p(x_1, x_2)$$

for every $(x_1, x_2) \in \mathbb{R}^2$ *such that* $x_1 \geq 0$, $x_2 \geq 0$ *and* $x_1 + x_2 \leq 1$; *hence*

$$p(x_1, x_2) = \begin{cases} x_2^2, & \text{if } x_1 \geq 0,\ x_2 \geq 0,\ x_1 + x_2 \leq 1, \\ +\infty, & \text{otherwise.} \end{cases}$$

Therefore, $p(0) = 0$ *and, since* p *is lsc, we conclude that the duality gap must be* 0 *for both dual problems* $(\mathcal{D}_c)$ *and* $(\mathcal{D}_d)$. *Indeed, in the case of* $(\mathcal{D}_c)$, *the pair* $(0, h_0)$, *with* $h_0(x_1, x_2) := x_2^2$, *is an optimal solution, since* $-p^c\,(0, h_0) = \inf\left\{\phi(s_1, s_2, x_1, x_2) - x_2^2\right\} = 0$ *(the infimum is attained, e.g., at* $(s_1, s_2, x_1, x_2) :=$

$(0, 0, 0, 0)$), *whereas in the case of* $(\mathcal{D}_d)$, *the same function* h_0 *is an optimal solution, since we also have* $-p^d(h_0) = -p^c(0, h_0) = 0$.

Note that every continuous function $r : \mathbb{R}^2 \to \mathbb{R}^2$ *satisfying* $\langle x, r(x)\rangle = h_0(x)$ *for all* $x \in \mathbb{R}^2$ *is an optimal solution to the dual problem corresponding to the coupling function considered in [1](defined following the standard duality scheme of [4, Section 3]). One can take, for instance,* $r(x_1, x_2) := (0, x_2)$, *but, clearly, for any continuous mapping* $w : \mathbb{R}^2 \to \mathbb{R}^2$ *satisfying* $\langle x, w(x)\rangle = 0$ *for all* $x \in \mathbb{R}^2$, *the function* $r + w$ *is an optimal solution, too. Thus, by means of this example we also illustrate the fact that each optimal solution of either* $(\mathcal{D}_c)$ *or* $(\mathcal{D}_d)$ *is associated, in a natural way, to infinitely many optimal solutions to the dual problem arising according to the conjugation scheme in [1]; therefore, the optimal solution sets (as well as the feasible sets) of our dual problems are substantially smaller than those of the equivalent (from the viewpoint of optimal values) dual problem arising from [1].*

Example 19 *[2, Section 3] Consider the problem*

$$\begin{array}{ll} \textit{minimize} & e^{s_2} \\ \textit{subject to} & \sqrt{s_1^2 + s_2^2} \le s_1, \end{array}$$

and let $\phi : \mathbb{R}^3 \to \mathbb{R} \cup \{+\infty\}$ *be the function defined by*

$$\phi(s_1, s_2, x) := \begin{cases} e^{s_2}, & \textit{if } \sqrt{s_1^2 + s_2^2} \le s_1 - x, \\ +\infty, & \textit{otherwise;} \end{cases}$$

clearly,

$$(\mathcal{P}) \qquad \textit{minimize } \phi(s, 0)$$

is equivalent to the given problem. One can easily check that the associated perturbation function p *is given by*

$$p(x) = \begin{cases} 1, & \textit{if } x = 0 \\ 0, & \textit{if } x < 0 \\ +\infty, & \textit{if } x > 0. \end{cases}$$

Therefore, $p(0) = 1$ *and, since* p *is not lsc at* 0, *there exists a strictly positive duality gap for both dual problems* $(\mathcal{D}_c)$ *and* $(\mathcal{D}_d)$. *We shall verify that the optimal values of* $(\mathcal{D}_c)$ *and* $(\mathcal{D}_d)$ *are both* 0, *thus showing that the duality gap is equal to* 1. *Indeed, for every* $(x^*, h) \in \mathbb{R} \times \mathcal{H}_{\mathbb{R}}$ *we have*

$$\begin{aligned} -p^c(x^*, h) &= \inf_{x \in \mathbb{R}} \{p(x) - xx^* - h(x)\} \le \inf_{x<0} \{p(x) - xx^* - h(x)\} \\ &= \inf_{x<0} \{-xx^* - h(x)\} \le 0, \end{aligned}$$

the last inequality being a consequence of the continuity of h *at* 0; *moreover,*

$$-p^d(h) = -p^c(0, h) \le 0.$$

Since, for $h_0 \equiv 0$*, we have*

$$-p^d(h_0) = -p^c(0, h_0) = \inf_{x \in \mathbb{R}} p(x) = 0,$$

we conclude that $(0, h_0)$ *and* h_0 *are optimal soultions to* $(\mathcal{D}_c)$ *and* $(\mathcal{D}_d)$*, respectively, and that the optimal values of these dual problems are both* 0.

Acknowledgments. We are grateful to Wilfredo Sosa and two anonymous referees for their valuable comments and suggestions.

References

[1] J. Cotrina, E. W. Karas, A. A. Ribeiro, W. Sosa, Y. J. Yun. *Moreau-Fenchel conjugation for lower semi-continuous functions*, Optimization 60 (2011), no. 8-9, 1045–1057.

[2] R. J. Duffin. *Convex analysis treated by linear programming*, Mathemathical Programming 4 (1973), 125–143.

[3] I. E. Leonard, K. Sundaresan. *Geometry of Lebesgue-Bochner function spaces—smoothness,* Trans. Amer. Math. Soc. 198 (1974), 229–251.

[4] J. E. Martínez-Legaz. *Generalized convex duality and its economic applications*, Handbook of generalized convexity and generalized monotonicity, 237–292, Springer, 2005.

[5] J.-P. Penot. *Conjugacies adapted to lower semicontinuous functions*, Optimization 64 (2015), no. 3, 473–494.

[6] R. R. Phelps. *Convex functions, monotone operators and differentiability.* Springer, 1993.

[7] R. T. Rockafellar, R. J.-B. Wets. *Variational analysis.* Springer, 1998.

[8] A. Rubinov, *Abstract convexity and global optimization.* Kluwer Academic Publishers, 2000.

[9] I. Singer, *Abstract convex analysis.* John Wiley & Sons, 1997.

[10] K. R. Stromberg. *Introduction to classical real analysis.* Wadsworth International, 1981.

[11] W. Sosa. *Representation of continuous functions and its applications*, J. Optim. Theory Appl. 159 (2013), no. 3, 795–804.