Journal of Convex Analysis
Volume 3 (1996), No. 2, 255–268

More Than First-Order Developments of Convex Functions: Primal-Dual Relations

C. Lemaréchal, C. Sagastizábal
INRIA, BP 105, 78153 Le Chesnay, France.
e-mail: Claude.Lemarechal@inria.fr, Claudia.Sagastizabal@inria.fr

Received May 2, 1995
Revised manuscript received November 3, 1995

The subject of this paper concerns the remainder term in the first-order development of a (finite-valued) convex function. We study functions for which this term is comparable to a squared norm and we relate it to the corresponding remainder term of the conjugate function. We show that a convex function satisfies a quadratic growth condition if and only if its subdifferential satisfies a linear growth condition. Finally, we define a new concept of "tangential regularization", involving a local decomposition of $\mathbb{R}^N$, along the subspace where the function is "smooth" and the subspace parallel to the subdifferential.

Keywords : Conjugate functions, first-order developments, regularization, growth conditions.

1991 Mathematics Subject Classification: 26B05, 52A41.

1. Introduction

The very first motivation of this paper is to study second-order properties of convex functions. More precisely, we want to relate the second-order elements of a convex function φ with those of its Moreau-Yosida regularization

$$\Phi(x) := \min_{y \in \mathbb{R}^n} \left\{ \varphi(y) + \tfrac{1}{2}\|y - x\|^2 \right\} ;$$

this is done in the companion paper [5]. A key property of such a regularization is the dual relation $\Phi^* = \varphi^* + 1/2\|\cdot\|^2$, which shows that Φ^* and φ^* share identical smoothness properties. This observation strongly motivates the following question: how are related the second-order elements of a convex function with those of its conjugate? It is precisely this last question that we investigate here.
The answer is easy in the quadratic case: the Hessians of two mutually conjugate quadratic functions are inverse to each other, a result which has been generalized in [1] to the C^2 case. What we need now is a more general study, in which no first-order differentiability is assumed. This does make sense: for example, the univariate function $\varphi(\xi) := |\xi| + 1/2\xi^2$ is not C^1 but its conjugate is C^2. A close look at this example confirms the intuition given by the C^2 case: nonsmoothness of a function comes together with an affine behaviour of its conjugate.

ISSN 0944-6532 / \$ 2.50

The present paper gives a number of results related with the above question. They have an interest in their own right, and are exploited in [5], entirely devoted to a study of the Moreau-Yosida regularization and its Hessian. In Section 2, we start with some elementary properties of convex sets and normal cones, which will be instrumental for the sequel. The main content of the paper is then Section 3, which relates the remainder terms in the first-order developments of two mutually conjugate functions. Finally, Section 4 introduces a useful operation, which consists in "strong-convexifying" affine parts of a function, thereby regularizing its conjugate.

We use standard notation; the working space is $\mathbb{R}^N$, in which the scalar product is $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ is the associated norm. Our study considers finite-valued convex functions φ only. A function denoted by an upper-case letter (Φ) will be some regularized form of the corresponding lower-cased function (φ). Throughout this paper we will consider a particular subgradient $g_0 \in \partial\varphi(z_0)$ at a particular $z_0 \in \mathbb{R}^N$; in general, $\mathcal{N}$ and $\mathcal{T}$ will denote the normal and tangent cones $\mathrm{N}_{\partial\varphi(z_0)}(g_0)$ and $\mathrm{T}_{\partial\varphi(z_0)}(g_0)$. We will use $\mathrm{I}, \sigma, \overset{+}{\triangledown}$ to denote the indicator and support functions and the infimal convolution respectively.

2. On the geometry of convex sets

The material in this section is quite elementary and departs from the general flow of the paper. It can therefore be skipped during a first reading.

In the following results we characterize those convex cones that are subspaces. For the subspace $\mathcal{M}$ in Proposition 2.1, see Theorem 2.7 in [7].

Proposition 2.1. *Let $\mathcal{N}$ be a closed convex cone and call $\mathcal{M} := \mathcal{N} \cap \{-\mathcal{N}\}$ the largest subspace contained in $\mathcal{N}$. Then*

$$\mathcal{N} \textit{ is a subspace if and only if } \mathcal{N} \cap \mathcal{M}^\perp = \{0\}\,.$$

Moreover, for any $\nu_0 \in \mathcal{M}^\perp$ and $\nu \in \mathcal{N}$,

$$\langle \nu_0, \nu \rangle \neq 0 \quad \Longrightarrow \quad -\nu \notin \mathcal{N}\,. \tag{1}$$

Proof. When the convex cone $\mathcal{N}$ is a subspace, it is symmetric: $\mathcal{N} = -\mathcal{N} = \mathcal{M}$. In this case, $\mathcal{N} \cap \mathcal{M}^\perp = \mathcal{M} \cap \mathcal{M}^\perp = \{0\}$.

Conversely, when $\mathcal{N} \cap \mathcal{M}^\perp = \{0\}$, suppose for contradiction that $\mathcal{N}$ is not a subspace; we can take $\nu \in \mathcal{N} \backslash \mathcal{M}$, which can be expressed as a direct sum:

$$\nu = \nu_m + \nu_0 \quad \text{with } \nu_m \in \mathcal{M} \text{ and } 0 \neq \nu_0 \in \mathcal{M}^\perp\,.$$

Since $\mathcal{M}$ is a symmetric set, $-\nu_m \in \mathcal{M} \subset \mathcal{N}$ and, since $\mathcal{N}$ is a convex cone,

$$\nu_0 = \nu + (-\nu_m) \in \mathcal{N}\,;$$

thus we have exhibited a nonzero $\nu_0 \in \mathcal{N} \cap \mathcal{M}^\perp$. This is the required contradiction.

Finally, we have to prove (1). Take $\nu_0 \in \mathcal{M}^\perp$ and $\nu \in \mathcal{N}$. If $-\nu$ were in $\mathcal{N}$, it would be also in $\mathcal{M}$ and this would contradict $\langle \nu_0, \nu \rangle \neq 0$. $\square$

The next result deals with $\mathrm{N}_S(g_0)$ and $\mathrm{T}_S(g_0)$, the normal and tangent cones to a closed convex set S at $g_0 \in S$.

Proposition 2.2. *Assume $S \subset \mathbb{R}^N$ is a closed convex set and let $g_0 \in S$. Then*

$$\mathcal{N} := \mathrm{N}_S(g_0) \text{ is a subspace} \quad \text{if and only if} \quad g_0 \in \operatorname{ri} S\,,$$

where ri *denotes the relative interior.*

Proof. Call σ_S the support function of S. By definition of normal cones, $s \in \mathcal{N}$ exactly when $\langle s, g_0\rangle = \sigma_S(s)$. The convex cone $\mathcal{N}$ is a subspace if and only if it is symmetric. Thus, $\mathcal{N}$ is a subspace if and only if the following property holds:

$$s \in \mathcal{N} \implies \sigma_S(s) + \sigma_S(-s) = 0\,, \qquad [= \langle s, g_0\rangle + \langle -s, g_0\rangle]$$

or equivalently

$$\sigma_S(d) + \sigma_S(-d) > 0 \quad \implies \quad d \notin \mathcal{N}\,,$$

i.e., applying the definition of $\mathcal{N}$:

$$\sigma_S(d) + \sigma_S(-d) > 0 \quad \implies \quad \exists g_d \in S \,:\, \langle d, g_d - g_0\rangle > 0\,,$$

which in turn can be written

$$\sigma_S(d) + \sigma_S(-d) > 0 \quad \implies \quad \sigma_S(d) > \langle d, g_0\rangle\ .$$

By Theorem 13.1 in [7] or Theorem V.2.2.3(ii) in [4], this last property just expresses $g_0 \in \operatorname{ri} S$. □

We finish with an easy result.

Proposition 2.3. *The only nonempty compact convex sets in $\mathbb{R}^N$ that are relatively open are the singletons.*

Proof. The relative interior of a convex set S is the interior for the topology relative to $\operatorname{aff} S$, the affine hull of S; this last topological space being connected, it contains only two open and closed subsets: $\emptyset$ and itself. Thus, if the nonempty closed set S is relatively open, $S = \operatorname{aff} S$. The result follows since the only bounded affine sets are the singletons. □

A link between all these results, which will be useful for our purposes, is the following: if φ is a finite-valued convex function and $z \in \mathbb{R}^N$ is such that $\mathrm{N}_{\partial\varphi(z)}(g)$ is a subspace for all $g \in \partial\varphi(z)$, then φ is differentiable at z.

3. On first-order developments of convex functions

We now present a theory analogous to that of §X.4.2(b) in [4], and of [6], relating upper bounds on a convex function with lower bounds on its conjugate. However, instead of finding bounds depending on a particular subgradient at a given point, we consider here the first-order expansion of a convex function φ:

$$\varphi(z_0 + h) = \varphi(z_0) + \varphi'(z_0; h) + o(\|h\|)\,.$$

We want to estimate the remainder term (which is clearly nonnegative). To bound it from above means to find $\varepsilon > 0$ and some function r_u such that

$$\varphi(z_0 + h) \le \varphi(z_0) + \varphi'(z_0; h) + r_u(h) + \mathrm{I}_{B_\varepsilon}(h) \quad \text{for all } h\,. \tag{2}$$

For a lower bound we need likewise ε and r_l such that

$$\varphi(z_0 + h) + \mathrm{I}_{B_\varepsilon}(h) \ge \varphi(z_0) + \varphi'(z_0; h) + r_l(h) \quad \text{for all } h\,. \tag{3}$$

Here and throughout, $\mathrm{I}_{B_\varepsilon}$ denotes the indicator function of $B(0, \varepsilon)$. For both remainder functions $r = r_u$ and $r = r_l$, the following properties will always be assumed without further mentioning:

$$r \text{ is convex, nonnegative, differentiable at } 0,\ r(0) = 0,\ \nabla r(0) = 0.$$

Besides, the following property will be crucial for our development:

$$\text{there exists } c > 0 \text{ such that} \quad r(h) \ge \frac{1}{2} c \|h\|^2 \quad \text{for all } h\,. \tag{4}$$

Note that (3) and (4) with $r = r_l$ hold whenever φ is strongly convex, and they imply that the conjugate φ^* is finite-valued.
Our aim is to study the duality between (2) and (3): if φ satisfies one of these inequalities, does φ^* satisfy the other? The answer is given in Theorems 3.2 and 3.6 below, which play the role of Theorems X.4.2.7 and X.4.2.6 in [4] respectively.

3.1. On the Lipschitzian regularization

To derive dual relations, we will conjugate both sides of (2) and (3). For this, we will frequently compute conjugates of the type $(\psi + \mathrm{I}_{B_\varepsilon})^*$, i.e., infimal convolutions

$$\Psi^*_\varepsilon(g) := (\psi + \mathrm{I}_{B_\varepsilon})^*(g) = \min_{s \in \mathbb{R}^N} \{\psi^*(g - s) + \varepsilon \|s\|\}\,. \tag{5}$$

The function Ψ^*_ε is the so-called Lipschitzian regularization of ψ^*, introduced in [2], and illustrated by Fig. 1. An important property of this function is that Ψ^*_ε coincides with ψ^* on a certain set.

Figure 1: Lipschitzian Regularization

Lemma 3.1. *Let ψ be a finite-valued convex function and denote by Ψ^*_ε the Lipschitzian regularization of its conjugate ψ^*.*

(i) *If there are $c > 0$ and $\delta > 0$ such that*

$$\psi(h) \geq \psi(0) + \psi'(0; h) + \frac{1}{2}c\|h\|^2 \quad \text{for } \|h\| \leq \delta\,, \tag{6}$$

then $\Psi^*_\varepsilon(s) = \psi^*(s)$ for all $\varepsilon \leq \delta$ and $s \in \partial\psi(0) + B(0, \varepsilon c/2)$.

(ii) *If ψ has the form $\psi(h) = \frac{1}{2}C\|h\|^2$ (i.e. $\psi^* = \frac{1}{2C}\|\cdot\|^2$) for some $C > 0$, then*

$$\Psi^*_\varepsilon(z) = \begin{cases} \frac{1}{2C}\|z\|^2 & \text{if } \|z\| \leq \varepsilon C \\ -\frac{\varepsilon^2 C}{2} + \varepsilon\|z\| & \text{if not.} \end{cases}$$

(iii) *With ψ as in* (ii) *and $\eta \in]0, \varepsilon]$, there holds*

$$\Psi^*_\varepsilon(z) \geq -\frac{\eta^2 C}{2} + \eta\|z\| \qquad \text{for all } z.$$

Proof. [(i)] Apply Proposition XI.3.4.5 of [4]: ψ^* and its regularization Ψ^*_ε coincide on the set

$$\{s \in \mathbb{R}^N : \psi^*(s) = \Psi^*_\varepsilon(s)\} = \{s \in \mathbb{R}^N : \partial\psi^*(s) \cap B(0, \varepsilon) \neq \emptyset\}\,.$$

Since $s \in \partial\psi(0)$ means $0 \in \partial\psi^*(s)$, this coincidence set contains in particular $\partial\psi(0)$. To say that $\partial\psi^*(s) \cap B(0, \varepsilon) \neq \emptyset$ is to say that $\langle s, \cdot\rangle - \psi(\cdot)$ attains its maximum $\psi^*(s)$ on $B(0, \varepsilon)$. Accordingly, let us find an upper bound for

$$A := \sup_{\|h\| > \varepsilon} \{\langle s, h\rangle - \psi(h)\}\,.$$

For this, use the convexity of ψ on $[0, h]$: if $h \notin B(0, \varepsilon)$,

$$\psi(\varepsilon h/\|h\|) \leq \psi(0) + \frac{\varepsilon}{\|h\|}[\psi(h) - \psi(0)]\,,$$

which, after some algebraic manipulations using (6), gives when $\varepsilon \leq \delta$:

$$\psi(h) \geq \psi(0) + \psi'(0; h) + \frac{1}{2}c\varepsilon\|h\|\,.$$

Therefore

$$\begin{aligned} A &\leq \sup_{\|h\| > \varepsilon} \{\langle s, h\rangle - \psi(0) - \psi'(0; h) - \tfrac{1}{2}c\varepsilon\|h\|\} \\ &\leq -\psi(0) + \sup_{\|h\| \in \mathbb{R}^N} \{\langle s, h\rangle - \psi'(0; h) - \tfrac{1}{2}c\varepsilon\|h\|\} \\ &= -\psi(0) + \left(\psi'(0; \cdot) + \tfrac{1}{2}c\varepsilon\|\cdot\|\right)^* (s) \\ &= -\psi(0) + \left(\mathrm{I}_{\partial\psi(0)} \overset{+}{\vee} \mathrm{I}_{B(0, \varepsilon c/2)}\right)(s)\,. \end{aligned}$$

The last infimal convolution is zero when $s \in \partial\psi(0) + B(0, \varepsilon c/2)$. For every such s which is not in $\partial\psi(0)$, we thus have

$$A \leq -\psi(0) = \inf \psi^* < \psi^*(s)\,.$$

Then $\langle s, h\rangle - \psi(h)$ attains its maximum on $B(0, \varepsilon)$; (i) is proved.
[(ii)] For our quadratic function ψ, apply the dual definition (5):

$$\begin{aligned}\Psi^*_\varepsilon(z) = \left(\tfrac{1}{2}C\|\cdot\|^2 + \mathrm{I}_{B_\varepsilon}\right)^*(z) &= \sup_{x\in B(0,\varepsilon)} \{\langle z, x\rangle - \tfrac{1}{2}C\|x\|^2\} \\ &= C \sup_{x\in B(0,\varepsilon)} \{\langle z/C, x\rangle - \tfrac{1}{2}\|x\|^2\} \\ &= C \sup_{x\in B(0,\varepsilon)} \{\tfrac{1}{2}\|z/C\|^2 - \frac{1}{2}\|x - z/C\|^2\} \\ &= \tfrac{1}{2C}\|z\|^2 - \frac{C}{2} \inf_{x\in B(0,\varepsilon)} \|x - z/C\|^2 \\ &= \tfrac{1}{2C}\|z\|^2 - \frac{C}{2}\mathrm{d}^2_{B(0,\varepsilon)}(z/C)\,.\end{aligned}$$

To finish the proof observe that the distance function $\mathrm{d}_{B(0,\varepsilon)}(z/C)$ is equal to $\varepsilon - \|z\|/C$ whenever $z/C \notin B(0, \varepsilon)$.
[(iii)] We consider two cases. If $\|z\| \leq \varepsilon C$, the result follows from

$$\Psi^*_\varepsilon(z) - \left(-\eta^2 C/2 + \eta\|z\|\right) = \frac{1}{2C}\|z\|^2 + \frac{1}{2}\eta^2 C - \eta\|z\| = \frac{1}{2C}(\|z\| - \eta C)^2\,.$$

In the other case, set $q(\eta) := -\eta^2 C/2 + \eta\|z\|$ (so $q = \Psi^*_\varepsilon$ when $\eta = \varepsilon$). Compute $q'(\eta) = -\eta C + \|z\|$. We see that q is an increasing function when $-\eta C + \|z\| \geq 0$, in particular when $\eta \leq \varepsilon < \|z\|/C$. □

We are now in a position to establish our duality relations between (2) and (3).

3.2. Bounding the conjugate from below

Let us return to our developments (2), (3). When the remainder term r_u does not increase too fast near the origin, (2) becomes a rather natural growth condition: it expresses that φ does not differ too much from its first order approximation. In this case, a lower bound on φ^* does exist:

Theorem 3.2. *Let φ be a finite-valued convex function satisfying (2) at a given z_0. Assume also that $r = r_u$ satisfies (4) and*

$$r_u(h) \leq \frac{1}{2}C\|h\|^2 \quad \text{for all } h\,. \tag{7}$$

Then, for all $g_0 \in \partial\varphi(z_0)$ and $s \in B(0, \dfrac{\varepsilon c^2}{2C})$, we have

$$\varphi^*(g_0 + s) \geq \varphi^*(g_0) + \langle s, z_0\rangle + \min_{\gamma\in\partial\varphi(z_0)} r^*_u(g_0 + s - \gamma)\,. \tag{8}$$

Proof. We conjugate both sides in (2): for all $g \in \mathbb{R}^N$

$$\begin{aligned}\varphi^*(g) - \langle g, z_0\rangle &\geq \left[\varphi(z_0) + \varphi'(z_0;\cdot) + r_u + \mathrm{I}_{B_\varepsilon}\right]^*(g)\\ &= -\varphi(z_0) + \left(\mathrm{I}_{\partial\varphi(z_0)} \,\overset{+}{\vee}\, (r_u + \mathrm{I}_{B_\varepsilon})^*\right)(g)\\ &= -\varphi(z_0) + \left(\mathrm{I}_{\partial\varphi(z_0)} \,\overset{+}{\vee}\, R^*_\varepsilon\right)(g)\\ &= -\varphi(z_0) + \min_{\gamma\in\partial\varphi(z_0)} R^*_\varepsilon(g-\gamma)\,,\end{aligned} \tag{9}$$

where R^*_ε denotes the Lipschitzian regularization of r^*_u, see (5). We will prove that, for g close enough to $\partial\varphi(z_0)$, R^*_ε can be replaced by r^*_u in (9).
First, observe that (4) and (7) are transformed in the dual space to

$$\frac{1}{2C}\|\cdot\|^2 \leq r^*_u \leq \frac{1}{2c}\|\cdot\|^2\,. \tag{10}$$

These inequalities are transmitted to the Lipschitzian regularizations; by Lemma 3.1 (iii) we therefore have, for all $\eta \in]0,\varepsilon]$,

$$R^*_\varepsilon(z) \geq \frac{1}{2}\left(-\eta^2 C + 2\eta\|z\|\right) \qquad \text{for all } z.$$

Apply (10) to obtain by division

$$\frac{r^*_u(z)}{R^*_\varepsilon(z)} \leq \frac{\|z\|^2}{-\eta^2 cC + 2\eta c\|z\|} \qquad \text{for all } z. \tag{11}$$

Let γ_g denote an optimal γ in (9). From (10) and (11),

$$\frac{1}{2C}\|g-\gamma_g\|^2 \leq r^*_u(g-\gamma_g) \leq \frac{\|g-\gamma_g\|^2}{-\eta^2 cC + 2\eta c\|g-\gamma_g\|} R^*_\varepsilon(g-\gamma_g)\,.$$

To this chain of inequalities, we append additional upper bounds, using the optimality of γ_g, $R^*_\varepsilon \leq r^*_u$, and (10): for all $\gamma \in \partial\varphi(z_0)$,

$$\frac{1}{2C} \leq \frac{1}{-\eta^2 cC + 2\eta c\|g-\gamma_g\|}\,\frac{\|g-\gamma\|^2}{2c}\,.$$

After some algebra this results in

$$\|g-\gamma_g\| \leq \frac{C}{2\eta c^2}\|g-\gamma\|^2 + \frac{\eta C}{2} \qquad \begin{array}{l}\text{for } \gamma_g \text{ optimal in (9)}\,, \eta \leq \varepsilon\\ \text{and } \gamma \text{ arbitrary in } \partial\varphi(z_0).\end{array} \tag{12}$$

To obtain (8), take $\eta = \frac{\varepsilon c}{2C}$, $g_0 \in \partial\varphi(z_0)$ and $s \in B(0, \frac{\varepsilon c^2}{2C})$. For $g := g_0 + s$, let γ in (12) be the projection of g onto $\partial\varphi(z_0)$: we do have $\|g-\gamma_g\| \leq \varepsilon c/2$. Then Lemma 3.1 (i) allows us to replace R^*_ε by r^*_u in (9). The result follows, since $\varphi(z_0) + \varphi^*(g_0) = \langle g_0, z_0\rangle$. □

The remainder term in (8) may be deemed abstract. However it can be relaxed to a more explicit form.

Corollary 3.3. *Let φ be a finite-valued convex function satisfying* (2)*, with $r_u = \frac{1}{2}C\|\cdot\|^2$. Then, for all $g_0 \in \partial\varphi(z_0)$ and $s \in B(0, \varepsilon C/2)$, we have*

$$\varphi^*(g_0 + s) \geq \varphi^*(g_0) + \langle s, z_0\rangle + \frac{1}{2C}\|g_0 + s - \mathcal{P}(g_0 + s)\|^2 , \tag{13}$$

where $\mathcal{P}$ is the projection onto $\partial\varphi(z_0)$.
As a result

$$\langle s, x - z_0\rangle \geq \frac{\|s\|^2}{2C} \quad \text{for all } s \in \mathrm{N}_{\partial\varphi(z_0)}(g_0) \cap B(0, \varepsilon C/2) \text{ and } x \in \partial\varphi^*(g_0 + s) . \tag{14}$$

Proof. Use Theorem 3.2 with $r_u = \frac{1}{2}C\|\cdot\|^2$, so that $r_u^* = \frac{1}{2C}\|\cdot\|^2$. To prove (14), use the subgradient inequality $\varphi^*(g_0) \geq \varphi^*(g_0 + s) - \langle s, x\rangle$ for $x \in \partial\varphi^*(g_0 + s)$, and observe that $g_0 + s$ is projected onto g_0. $\square$

This result becomes even more suggestive in terms of the Moreau decomposition of the space $\mathbb{R}^N$. Recall that, if $\mathcal{N}$ and $\mathcal{T}$ are two mutually polar cones, then $\mathbb{R}^N = \mathcal{N} \oplus \mathcal{T}$. More precisely, any $s \in \mathbb{R}^N$ can be written

$$s = s_{\mathcal{N}} + s_{\mathcal{T}} \qquad \text{where } s_{\mathcal{N}} := \mathrm{Proj}_{\mathcal{N}}(s) \text{ and } s_{\mathcal{T}} := \mathrm{Proj}_{\mathcal{T}}(s)$$

(see Theorem III.3.2.5 of [4], for example). We will use this decomposition with $\mathcal{N} = \mathrm{N}_{\partial\varphi(z_0)}(g_0)$ and $\mathcal{T} = \mathrm{T}_{\partial\varphi(z_0)}(g_0)$.

Proposition 3.4. *Let φ^* be a closed convex function satisfying* (13) *for all $g_0 \in \partial\varphi(z_0)$ and $s \in B(0, \varepsilon C/2)$. Then, for all such g_0 and s, we have*

$$\varphi^*(g_0 + s) \geq \varphi^*(g_0) + \langle s, z_0\rangle + \frac{1}{2C}\|s_{\mathcal{N}}\|^2 . \tag{15}$$

Proof. By definition of a tangent cone, $\mathcal{T}$ contains $\cup_{t>0}[\partial\varphi(z_0) - g_0]/t$, hence $\partial\varphi(z_0) \subset g_0 + \mathcal{T}$ and $\|g_0 + \cdot - \mathcal{P}(g_0 + \cdot)\| \geq \|g_0 + \cdot - \mathrm{Proj}_{g_0+\mathcal{T}}(g_0 + \cdot)\| = \|\cdot - \mathrm{Proj}_{\mathcal{T}}(\cdot)\| = \|\mathrm{Proj}_{\mathcal{N}}(\cdot)\|$. Then (13) proves the result. $\square$

The Cauchy-Schwarz inequality in (14) suggests a sort of "tangential radial Lipschitz behaviour" of particular subgradients. Indeed, for x close to z_0, consider those $g = g_0 + s \in \partial\varphi(x)$ that are projected onto $g_0 \in \partial\varphi(z_0)$: they satisfy $\|g - g_0\| \leq 2C\|x - z_0\|$. Instead of proving this informal observation, we rather state a more global result:

Corollary 3.5. *For a finite-valued convex function φ and $g_0 \in \partial\varphi(z_0)$, the following statements are equivalent:*

$$\exists\varepsilon, C > 0 : \|h\| \leq \varepsilon \Rightarrow \varphi(z_0 + h) \leq \varphi(z_0) + \varphi'(z_0; h) + \frac{C}{2}\|h\|^2 , \tag{16}$$

$$\exists\delta, D > 0 : \|h\| \leq \delta \Rightarrow \partial\varphi(z_0 + h) \subset \partial\varphi(z_0) + B(0, D\|h\|) . \tag{17}$$

Proof. Let ε and C be as in (16). Because $\partial\varphi(z_0)$ is compact and $\partial\varphi$ has a closed graph, we can find $\delta > 0$ such that

$$\|h\| \leq \delta \Rightarrow \partial\varphi(z_0 + h) \subset \partial\varphi(z_0) + B(0, \varepsilon C/2)\,.$$

Take now arbitrary $h \in B(0,\delta)$, $g \in \partial\varphi(z_0 + h)$ and set $g_0 := \mathcal{P}(g)$. Then $s := g - g_0 \in \mathrm{N}_{\partial\varphi(z_0)}(g_0) \cap B(0, \varepsilon C/2)$ and $z_0 + h \in \partial\varphi^*(g)$. When (16) holds, Corollary 3.3 applies with $r_u = \frac{1}{2}C\|\cdot\|^2$ and (17) follows from (14).

Conversely, write the Mean-Value Theorem: for some $\theta \in]0,1[$ and some $g_\theta \in \partial\varphi(z_0 + \theta h)$,

$$\begin{aligned}\varphi(z_0 + h) - \varphi(z_0) - \varphi'(z_0; h) &= \langle g_\theta, h\rangle - \max_{g \in \partial\varphi(z_0)} \langle g, h\rangle \\ &= \min_{g \in \partial\varphi(z_0)} \langle g_\theta - g, h\rangle \\ &\leq \min_{g \in \partial\varphi(z_0)} \|g_\theta - g\|\,\|h\|\,.\end{aligned}$$

Assume (17) and take $h \in B(0,\delta)$, hence $\theta h \in B(0,\delta)$. Then the last term is smaller than $\theta D\|h\|^2 \leq D\|h\|^2$ and the growth property (16) holds. □

When φ has a gradient at z_0, this result becomes a coarse form of Theorem 2.12 in [3]; the difference is that the latter deals with approximations, rather than inequalities or inclusions.

3.3. Bounding the conjugate from above

Now, we proceed to study the effect on φ^* of the growth property (3), (4). We will see that this property is quite strong.

It is not clear whether our assumptions (4) and (7) are really essential for Theorem 3.2. At least, (7) is natural: useful lower bounds for φ^* need nontrivial upper bounds for φ; but the role of (4) is more obscure. By contrast, the assumptions in the following dual counterpart to Theorem 3.2 seem rather minimal.

Theorem 3.6. *Let φ be a finite-valued convex function satisfying* (3) *at a given z_0, with $r = r_l$ satisfying* (4)*. Then, for all $g_0 \in \partial\varphi(z_0)$, φ^* is differentiable at g_0 ($\nabla\varphi^*(g_0) = z_0$) and, for all $s \in B(0, \varepsilon c/2)$, we have*

$$\varphi^*(g_0 + s) \leq \varphi^*(g_0) + \langle s, z_0\rangle + \min_{\gamma \in \partial\varphi(z_0)} r_l^*(g_0 + s - \gamma)\,. \tag{18}$$

Proof. Proceed as in the proof of Theorem 3.2. Clearly (6) holds with $\psi = \varphi$ and $\delta = \varepsilon$, so Lemma 3.1 (i) can be applied to conjugate the lefthand side of (3):

$$\varphi^*(g) - \langle g, z_0\rangle \leq -\varphi(z_0) + \min_{\gamma \in \partial\varphi(z_0)} r_l^*(g - \gamma) \quad \text{for } g \in \partial\varphi(z_0) + B(0, \varepsilon c/2)\,.$$

Set $g = g_0 + s$, with s as stated, and observe that $\varphi(z_0) + \varphi^*(g_0) = \langle g_0, z_0\rangle$ to obtain (18). Now, in view of (4), $r_l^* \leq \frac{1}{2c}\|\cdot\|^2$, so (18) gives

$$\varphi^*(g_0 + s) - \varphi^*(g_0) - \langle s, z_0\rangle \leq \frac{1}{2c} \min_{\gamma \in \partial\varphi(z_0)} \|g_0 + s - \gamma\|^2 \leq \frac{1}{2c}\|s\|^2\,.$$

Because the lefthand side is nonnegative, this clearly implies $\nabla\varphi^*(g_0) = z_0$. $\square$

It is interesting to compare this result with Proposition X.4.2.6 in [4]. Our upper bound (18) is stronger than (X.4.2.5) in [4], just because our starting assumption (3) is stronger. In §3.2, (8) was relaxed to (13) and also to (15). Here, (18) could likewise be relaxed to

$$\varphi^*(g_0 + s) \leq \varphi^*(g_0) + \langle s, z_0\rangle + \frac{1}{2c}\|s\|^2 .$$

Actually, we prefer another relaxed form of (18), which gives a counterpart to (15):

Corollary 3.7. *Let φ be a finite-valued convex function satisfying* (3)*, with $r = r_l$ satisfying* (4)*, and let $g_0 \in \partial\varphi(z_0)$. Take the Moreau decomposition $s = s_{\mathcal{T}} + s_{\mathcal{N}}$ of an arbitrary s along $\mathcal{T}$ and $\mathcal{N}$, the tangent and normal cones to $\partial\varphi(z_0)$ at g_0. Then*

$$\limsup_{t\downarrow 0} \frac{\varphi^*(g_0 + ts) - \varphi^*(g_0) - t\,\langle s, z_0\rangle}{t^2} \leq \frac{1}{2c}\|s_{\mathcal{N}}\|^2 . \tag{19}$$

Proof. When $t \downarrow 0$, by definition of the tangent cone, there is some $g_t = g_0 + ts_{\mathcal{T}} + o(t) \in \partial\varphi(z_0)$. From here, the result is an easy consequence of Theorem 3.6. However, we prefer a more direct proof, which does not use the preceding machinery. From (3) and knowing that $\varphi'(z_0; h) \geq \langle g_0, h\rangle$,

$$\varphi(z_0 + h) \geq \varphi(z_0) + \langle g_t, h\rangle + \frac{1}{2}c\|h\|^2 + o(\|h\|^2)$$

for all h. Apply Proposition X.4.2.6 in [4]: for all ν,

$$\varphi^*(g_t + \nu) \leq \varphi^*(g_t) + \langle \nu, z_0\rangle + \frac{1}{2c}\|\nu\|^2 + o(\|\nu\|^2).$$

Combining the subgradient inequality at g_t, this gives

$$\varphi^*(g_t + \nu) \leq \varphi^*(g_0) + \langle g_t - g_0 + \nu, z_0\rangle + \frac{1}{2c}\|\nu\|^2 + o(\|\nu\|^2) .$$

Take $\nu := ts - g_t + g_0 = ts_{\mathcal{N}} + o(t)$ and divide by t^2. The result follows because $o(\|\nu\|^2) = o(t^2)$. $\square$

Beware that (19) is much stronger than (15) (written in the dual space and mutatis mutandis). In particular, it implies the existence of $\nabla\varphi^*(g_0) = z_0$. In fact Theorem 3.6 explains why full duality cannot hold between the pairs (2), (7) and (3), (4). Indeed, suppose that the following statement were true: if (2), (7) hold, then (3), (4) hold with φ replaced by φ^* and $r_l = r_u^*$. Then Theorem 3.6, used with φ and φ^* interchanged, would imply the existence of $\nabla\varphi(z_0)$.

4. A tangential regularization

We just mentioned that the growth properties (2), (7) and (3), (4) cannot be dual to each other, unless $\nabla\varphi(z_0)$ exists. In fact, Proposition 3.4 shows that, when φ satisfies an appropriate form of (2), then φ^* satisfies (15). Now (15) is a weakened form of (3)-(4) (for φ^*) and the following informal calculation shows that it cannot be improved because its last quadratic term cannot contain the $s_{\mathcal{T}}$-component. Given a point z_0 and $g_0 \in \partial\varphi(z_0)$, assume there is another $g \in \partial\varphi(z_0)$ and set $s_{\mathcal{T}} := g - g_0 \in \mathcal{T}$. Then, for $t \in]0,1[$,

$$\varphi^*(g_0 + ts_{\mathcal{T}}) = -\varphi(z_0) + \langle g_0 + ts_{\mathcal{T}}, z_0\rangle = \varphi^*(g_0) + t\,\langle s_{\mathcal{T}}, z_0\rangle \,.$$

We conclude that φ^* is locally affine in $\mathcal{T}$.

Knowing that (2), (7) imply (15) and nothing more, a natural question therefore arises: conversely, when (15) holds, what can be expected from φ in terms of the growth property (2)? The knack is to introduce an artificial $s_{\mathcal{T}}$-term in φ^*: indeed, the function

$$\phi^*_{\mathcal{T}}(s) := \varphi^*(g_0 + s) + \frac{1}{2}\|s_{\mathcal{T}}\|^2 \tag{20}$$

satisfies the appropriate growth condition (3) (for φ^*) and triggers the dualization mechanism of section 3.3.

Proposition 4.1. *Let φ be a finite-valued convex function and $g_0 \in \partial\varphi(z_0)$. The conjugate of $\phi^*_{\mathcal{T}}$ in* (20) *is*

$$\phi_{\mathcal{T}}(x) = \min_{y\in x+\mathcal{T}}\{\varphi(y) - \langle g_0, y\rangle + \tfrac{1}{2}\|x-y\|^2\}\,, \tag{21}$$

whose subdifferential is

$$\partial\phi_{\mathcal{T}}(x) = -g_0 + \{g \in \partial\varphi(p_{\mathcal{T}}(x)) :\ \operatorname{Proj}_{\mathcal{T}}(g_0 - g) = p_{\mathcal{T}}(x) - x\}\,. \tag{22}$$

Here $p_{\mathcal{T}}(x)$ is the unique minimizer in (21)*, characterized by*

$$\exists g \in \partial\varphi(p_{\mathcal{T}}(x))\ : p_{\mathcal{T}}(x) = x + \operatorname{Proj}_{\mathcal{T}}(g_0 - g)\,. \tag{23}$$

In particular, $p_{\mathcal{T}}(z_0) = z_0$, $\phi_{\mathcal{T}}(z_0) = \varphi(z_0) - \langle g_0, z_0\rangle$ and $\nabla\phi_{\mathcal{T}}(z_0) = 0$ exists.

Proof. Take the conjugates of the functions making up the sum in (20): $(\varphi^*(g_0 + \cdot))^* = \varphi(\cdot) - \langle g_0, \cdot\rangle$, and

$$\begin{aligned}(\tfrac{1}{2}\|\operatorname{Proj}_{\mathcal{T}}(\cdot)\|^2)^* &= (\tfrac{1}{2}\|\cdot - \operatorname{Proj}_{\mathcal{N}}(\cdot)\|^2)^* = (\min_{u\in\mathcal{N}}\{\tfrac{1}{2}\|\cdot - u\|^2\})^* \\ &= (\tfrac{1}{2}\|\cdot\|^2 \mathbin{\overset{+}{\vee}} \mathrm{I}_{\mathcal{N}})^* = \tfrac{1}{2}\|\cdot\|^2 + \mathrm{I}_{\mathcal{T}}\,.\end{aligned}$$

Their infimal convolution is the righthand side in (21); this is a finite-valued convex function of x, which is therefore the conjugate of $\phi^*_{\mathcal{T}}$.

To obtain the expression of $\partial\phi_{\mathcal{T}}$, write

$$\phi_{\mathcal{T}}(x) = \min_{y+z=x}\{\varphi(y) - \langle g_0, y\rangle + \tfrac{1}{2}\|z\|^2 + \mathrm{I}_{\mathcal{T}}(-z)\}$$

and apply Theorem XI.3.4.1 in [4]:

$$\partial\phi_{\mathcal{T}}(x) = [\partial\varphi(p_{\mathcal{T}}(x)) - g_0] \cap [x - p_{\mathcal{T}}(x) - \mathrm{N}_{\mathcal{T}}(p_{\mathcal{T}}(x) - x)] \,.$$

Thus the subgradients of $\phi_{\mathcal{T}}$ at x are the vectors $g - g_0$, with $g \in \partial\varphi(p_{\mathcal{T}}(x))$, such that $x - p_{\mathcal{T}}(x) - (g - g_0) \in \mathrm{N}_{\mathcal{T}}(p_{\mathcal{T}}(x) - x)$, i.e., $x - p_{\mathcal{T}}(x) - (g - g_0) \in \mathrm{N}_{\mathcal{T}}(p_{\mathcal{T}}(x) - x)$. Use

$$\gamma \in \mathrm{N}_{\mathcal{T}}(p_{\mathcal{T}}(x) - x) \iff p_{\mathcal{T}}(x) - x = \mathrm{Proj}_{\mathcal{T}}(\gamma + p_{\mathcal{T}}(x) - x) \,, \tag{24}$$

with $\gamma = x - p_{\mathcal{T}}(x) - (g - g_0)$: altogether, the subgradients of $\phi_{\mathcal{T}}$ at x are the vectors $g - g_0$, with $g \in \partial\varphi(p_{\mathcal{T}}(x))$ such that $p_{\mathcal{T}}(x) - x = \mathrm{Proj}_{\mathcal{T}}(g_0 - g)$. This is just (22).
The optimality condition characterizing the unique minimum in (21) is $0 \in \partial\varphi(p_{\mathcal{T}}(x)) - g_0 + p_{\mathcal{T}}(x) - x + \mathrm{N}_{x+\mathcal{T}}(p_{\mathcal{T}}(x))$, i.e.,

$$\exists g \in \partial\varphi(p_{\mathcal{T}}(x)) \, : g_0 - g + x - p_{\mathcal{T}}(x) \in \mathrm{N}_{x+\mathcal{T}}(p_{\mathcal{T}}(x)) \,.$$

Observe that $\mathrm{N}_{x+\mathcal{T}}(p_{\mathcal{T}}(x)) = \mathrm{N}_{\mathcal{T}}(p_{\mathcal{T}}(x) - x)$, and use again the equivalence (24): the characterization above can be written $p_{\mathcal{T}}(x) - x = \mathrm{Proj}_{\mathcal{T}}(g_0 - g)$, which is just (23).
In particular, when $x = z_0$, we can take $g = g_0 \in \partial\varphi(z_0)$, and this shows that z_0 satisfies the characterization (23) of $p_{\mathcal{T}}(z_0)$. In other words, $p_{\mathcal{T}}(z_0) = z_0$, and the value $\phi_{\mathcal{T}}(0)$ follows easily. Finally, we have

$$\partial\phi_{\mathcal{T}}(0) = [\partial\varphi(z_0) - g_0] \cap \mathrm{N}_{\mathcal{T}}(0) \subset \mathcal{T} \cap \mathrm{N}_{\mathcal{T}}(0) = \mathcal{T} \cap \mathcal{N} \,,$$

which is the singleton $\{0\}$. $\square$

From its definition (21), $\phi_{\mathcal{T}}$ appears as a "tangential" Moreau-Yosida regularization in the cone $\mathcal{T}$; and $p_{\mathcal{T}}$ appears likewise as a "tangential" proximal operator. This explains our notation.
The above result is true for a rather arbitrary closed convex function φ^*. When additional properties are assumed, $\phi_{\mathcal{T}}$ enjoys the properties involved in Corollary 3.5.

Proposition 4.2. *Let the convex functions φ and φ^* be both finite-valued, with φ^* satisfying* (13) *for $g_0 \in \partial\varphi(z_0)$ and $\|s\|$ small enough. Then*

$$\exists \varepsilon, C' : \|h\| \le \varepsilon \Rightarrow \phi_{\mathcal{T}}(z_0 + h) \le \phi_{\mathcal{T}}(z_0) + \frac{C'}{2}\|h\|^2 \,,$$

or equivalently

$$\exists \delta, D > 0 : \|h\| \le \delta \Rightarrow \partial\phi_{\mathcal{T}}(z_0 + h) \subset B(0, D\|h\|) \,. \tag{25}$$

Proof. From its definition (20), $\phi^*_{\mathcal{T}}$ is finite-valued. Using Proposition 3.4, $\phi^*_{\mathcal{T}}(s) \ge \phi^*_{\mathcal{T}}(0) + \frac{1}{2C'}|s\|^2$, with $C' := \max\{1, C\}$; so (3) holds, with φ replaced by $\phi^*_{\mathcal{T}}$ and $r_l = \frac{1}{2C'}\|\cdot\|^2$. Then Theorem 3.6 can be applied. Knowing that $\nabla\phi_{\mathcal{T}}(z_0) = 0$ (Proposition 4.1), we obtain the growth property on $\phi_{\mathcal{T}}$. The radial Lipschitz property (25) then follows from Corollary 3.5. $\square$

Another closely related question concerns the regularity of the mapping $p_{\mathcal{T}}(\cdot)$. More precisely, is this tangential proximal operator Lipschitzian? Our next result gives a partial answer: under mild assumptions, $p_{\mathcal{T}}(\cdot)$ is radially Lipschitzian at z_0.

Corollary 4.3. *Let the convex functions φ and φ^* be both finite-valued, with φ^* satisfying* (13) *for $g_0 \in \partial\varphi(z_0)$ and $\|s\|$ small enough. Then*

$$\exists \delta, L > 0 : \ \|x - z_0\| \leq \delta \Rightarrow \|p_{\mathcal{T}}(x) - z_0\| \leq L\|x - z_0\| \,,$$

or equivalently

$$\exists \delta, D > 0 : \ \|x - z_0\| \leq \delta \Rightarrow \|p_{\mathcal{T}}(x) - x\| \leq D\|x - z_0\| \,. \tag{26}$$

Proof. The equivalence between both statements is straighforward from the triangular inequality, let us prove (26). Combining (22) from Proposition 4.1 with (25) from Proposition 4.2, there exist $\delta, D > 0$ such that $\|x - z_0\| \leq \delta$ implies

$$-g_0 + \{g \in \partial\varphi(p_{\mathcal{T}}(x)) : \ \operatorname{Proj}_{\mathcal{T}}(g_0 - g) = p_{\mathcal{T}}(x) - x\} \subset B(0, D\|x - z_0\|) \,.$$

Take a vector $-g_0 + g$ as described in the above lefthand side. Since $\operatorname{Proj}_{\mathcal{T}}(\cdot)$ is nonexpansive and $\operatorname{Proj}_{\mathcal{T}}(0) = 0$, we can write

$$\|p_{\mathcal{T}}(x) - x\| = \|\operatorname{Proj}_{\mathcal{T}}(g_0 - g)\| \leq \|g_0 - g\| \leq D\|x - z_0\|$$

and (26) follows. □

Let us give a simple example showing how these results can be interpreted.
Take $\varphi := \max(\varphi_1, \varphi_2)$, with each φ_i smooth. Then (2), (7) always hold. Take z_0 such that $\varphi(z_0) = \varphi_1(z_0) = \varphi_2(z_0)$ but $\nabla\varphi_1(z_0) \neq \nabla\varphi_2(z_0)$. Then take $\alpha \in]0, 1[$ and $g_0 = \alpha\nabla\varphi_1(z_0) + (1-\alpha)\nabla\varphi_2(z_0)$, so that $g_0 \in \operatorname{ri}\partial\varphi(z_0)$. We claim that, for x close to z_0, $p_{\mathcal{T}}(x)$ satisfies $\varphi_1(p_{\mathcal{T}}(x)) = \varphi_2(p_{\mathcal{T}}(x))$. In fact, if $\varphi_1(p_{\mathcal{T}}(x)) > \varphi_2(p_{\mathcal{T}}(x))$, then $\nabla\varphi(p_{\mathcal{T}}(x)) = \nabla\varphi_1(p_{\mathcal{T}}(x))$ is far from g_0 and, from Proposition 4.1, $\nabla\phi_{\mathcal{T}}(x) = \nabla\varphi(p_{\mathcal{T}}(x)) - g_0$ is far from $0 = \nabla\phi_{\mathcal{T}}(z_0)$. This contradicts the continuity of $\nabla\phi_{\mathcal{T}}$.
Roughly speaking, when x describes a ball around z_0, $p_{\mathcal{T}}(x)$ describes a portion of the surface $\mathcal{S} := \{z : \varphi_1(z) = \varphi_2(z)\}$ where φ is not differentiable. We show now that, when restricted to this surface, φ is pretty smooth at z_0.

Theorem 4.4. *Let φ be a (finite-valued) strongly convex function satisfying* (2)*, with r_u satisfying* (7)*, and take $g_0 \in \operatorname{ri}\partial\varphi(z_0)$. Then there exists $D > 0$ such that, with $p_{\mathcal{T}}(x)$ defined by* (23)*,*

$$\varphi(p_{\mathcal{T}}(x)) \leq \varphi(z_0) + \langle g_0, p_{\mathcal{T}}(x) - z_0\rangle + \frac{1}{2}D\|p_{\mathcal{T}}(x) - z_0\|^2 \,,$$

for x close enough to z_0 such that $\|(x - z_0)_{\mathcal{T}}\| = O(\|(x - z_0)_{\mathcal{N}}\|)$.

Proof. We are in the conditions of Corollary 3.3 and Proposition 4.2 (recall in particular that (4) holds and implies $\varphi^* < +\infty$); so we can write $\phi_{\mathcal{T}}(x) \leq \phi_{\mathcal{T}}(z_0) + \frac{1}{2}\|x - z_0\|^2$. Using (21) we get

$$\varphi(p_{\mathcal{T}}(x)) - \langle g_0, p_{\mathcal{T}}(x)\rangle + \frac{1}{2}\|p_{\mathcal{T}}(x) - x\|^2 \leq \varphi(z_0) - \langle g_0, z_0\rangle + \frac{1}{2}\|x - z_0\|^2 \,,$$

from which we obtain

$$\varphi(p_{\mathcal{T}}(x)) \leq \varphi(z_0) + \langle g_0, p_{\mathcal{T}}(x) - z_0\rangle + \frac{1}{2}\|x - z_0\|^2 \,.$$

By assumption, $\|x - z_0\|^2 = O(\|(x - z_0)_{\mathcal{N}}\|^2)$. Since $g_0 \in \operatorname{ri}\partial\varphi(z_0)$, $\mathcal{T}$ is a subspace (Proposition 2.2); then $(x - z_0)_{\mathcal{N}} = (x - p_{\mathcal{T}}(x))_{\mathcal{N}} + (p_{\mathcal{T}}(x) - z_0)_{\mathcal{N}} = (p_{\mathcal{T}}(x) - z_0)_{\mathcal{N}}$ (remember $x - p_{\mathcal{T}}(x) \in \mathcal{T}$). The result follows because $\|(p_{\mathcal{T}}(x) - z_0)_{\mathcal{N}}\| \leq \|p_{\mathcal{T}}(x) - z_0\|$. $\square$

When restricted to points of the form $p_{\mathcal{T}}(\cdot)$ (and z_0 is such a point), the function φ is differentiable at z_0 and locally comparable to a quadratic function. From there, the way is open to obtaining a real second-order development. This is done in [5].

References

[1] J.-P. Crouzeix: A relationship between the second derivative of a convex function and of its conjugate, Mathematical Programming, 13 (1977) 364–365.

[2] J.-B. Hiriart-Urruty: Lipschitz r-continuity of the approximate subdifferential of a convex function, Mathematica Scandinavica, 47 (1980) 123–34.

[3] J.-B. Hiriart-Urruty: The approximate first-order and second-order directional derivatives for a convex function, In J.-P. Cecconi and T. Zolezzi, editors, Mathematical Theories of Optimization, number 979 in Lecture Notes in Mathematics (1983), Springer-Verlag, 154–166.

[4] J.-B. Hiriart-Urruty and C. Lemaréchal: Convex Analysis and Minimization Algorithms, Springer-Verlag, 1993. (two volumes).

[5] C. Lemaréchal and C. Sagastizábal: Practical aspects of the Moreau-Yosida regularization: theoretical preliminaries, SIAM Journal on Optimization, 1995. To appear.

[6] J.-P. Penot: Subhessians, superhessians and conjugation, Nonlinear Analysis: Theory, Methods and Applications, 23(6) (1994) 689–702.

[7] R.T. Rockafellar: Convex Analysis, Princeton University Press, Princeton, New Jersey, 1970.