Maximal monotone operators, convex functions and a special family of enlargements.

Regina Sandra Burachik*
Engenharia de Sistemas e Computação,
COPPE–UFRJ, CP 68511,
Rio de Janeiro–RJ, 21945–970, Brazil.
regi@cos.ufrj.br

B.F. Svaiter†
IMPA, Instituto de Matemática Pura e Aplicada,
Estrada Dona Castorina, 110.
Rio de Janeiro, RJ, CEP 22460-320, Brazil.
benar@impa.br

June 28, 2000

Abstract

This work establishes new connections between maximal monotone operators and convex functions. Associated to each maximal monotone operator, there is a family of convex functions, each of which characterizes the operator. The basic tool in our analysis is a family of enlargements, recently introduced by Svaiter. This family of convex functions is in a one-to-one relation with a subfamily of these enlargements. We study the family of convex functions, and determine its extremal elements. An operator closely related to the Legendre-Fenchel

*Partially supported by PRONEX–Optimization.
†Partially supported by CNPq Grant 301200/93-9(RN) and by PRONEX–Optimization.

conjugacy is introduced and we prove that this family of convex functions is invariant under this operator. The particular case in which the operator is a subdifferential of a convex function is discussed. We propose a family of convex merit functions for the proximal point method, thus generalizing a recent finding of Solodov and Svaiter. Each element in the family of convex functions associated to the maximal monotone operator determines one of these merit functions.

keywords: maximal monotone operators, enlargements, convex functions, proximal point method, merit functions.

Mathematics Subject Classification (2000): 47H05, 46B99, 47H17.

1 Introduction and Motivation

Monotonicity and convexity are closely connected. Given a proper convex function f on a Banach space X, $f : X \to \mathbb{R} \cup \{+\infty\}$, the *subdifferential* of f at x is given by

$$\partial f(x) = \{u \in X^* | f(y) \geq f(x) + \langle u, y - x\rangle \text{ for all } y \in X\} .$$

The multifunction $\partial f : X \rightrightarrows X^*$, $x \mapsto \partial f(x)$ is monotone [9, 11]. If f is also lower semicontinuous (l.s.c. from now on), then ∂f is maximal monotone [15]. Nevertheless, maximal monotone operators in general are not subdifferentials.

The aim of this work is to further extend the connection between maximal monotone operators and convex functions. The basic tool we will use is a family of enlargements of maximal monotone operators introduced in [20]. We will show, by means of this family of enlargements, that associated to any maximal monotone operator T there exists a family of convex functions. Each of these functions characterizes the operator T and the whole family of convex functions is in a one-to-one relation with a subfamily of enlargements of the operator T. We also introduce an operator closely related to the Legendre-Fenchel conjugacy, which maps this family of functions in itself.

Besides the possible theoretical implications of this connection, we will give in the last section some application-oriented results. For finding a zero of a maximal monotone operator T in a Hilbert space, a most useful approach is the so-called proximal point method [16]. Recently, it was shown that each step of this method is equivalent to minimize a convex merit function [18].

We will show that each function in the family of convex functions associated with T determines a distinct merit function for the proximal point iteration, thus generalizing [18].

The study of enlargements will be central for connecting monotone operators with convex functions. Given A and B arbitrary sets and $F : A \rightrightarrows B$ a multifunction, an *enlargement* or *extension* of F is another multifunction $E : \mathbb{R}_+ \times A \to 2^B$ such that

$$F(x) \subseteq E(b, x) \quad \forall b \geq 0,\, x \in A.$$

A well-known and most important example of extension of a maximal monotone multifunction is the ε-subdifferential. Let again f be a proper l.s.c. convex function on X. The ε-subdifferential of f was introduced by Brøndsted and Rockafellar in [3]. It is defined as:

$$\partial_\varepsilon f(x) := \{v \in X^* \mid f(y) \geq f(x) + \langle v, y - x\rangle - \varepsilon \text{ for all } y \in X\},$$

for any $\varepsilon \geq 0$, $x \in X$. Note that $\partial f(x) \subseteq \partial_\varepsilon f(x)$ for any $\varepsilon \geq 0$, $x \in X$. As ∂f is maximal monotone, the ε-subdifferential is an example of enlargement of certain maximal monotone multifunctions (those which are subdifferentials). The ε-subdifferential enlargement (of ∂f) has many interesting properties, which are of theoretical and practical relevance. To mention some of them: a transportation formula, Lipschitz continuity in the interior of the domain [13, 8] and Brønsted and Rockafellar property [3].

For an arbitrary monotone multifunction $T : X \rightrightarrows X^*$, the following "$\varepsilon$-enlargement" was proposed and studied by Burachik and Svaiter [6]: Given $\varepsilon \geq 0$ and $x \in X$,

$$T^\varepsilon(x) := \{v \in X^* \mid \langle v - u, x - y\rangle \geq -\varepsilon \text{ for all } y \in X, u \in T(y)\}. \tag{1}$$

Again we have $T(x) \subseteq T^\varepsilon(x)$ for any $\varepsilon \geq 0$, $x \in X$ (follows from the monotonicity of T). The case in which X is a Hilbert space was considered in [5]. This enlargement has been originally introduced in finite dimensional spaces by Burachik, Iusem and Svaiter in [4]. When T is maximal monotone, the above "ε-enlargement" shares many properties with the ε-subdifferential enlargement of a closed proper convex function [5, 6]. In [20], a family of enlargements with similar properties was studied. This family of enlargements is characterized by three basic properties, which happen to be satisfied by the ε-subdifferential enlargement and the "ε-enlargement" above defined. So, Brøndsted-Rockafellar's ε-subdifferential enlargement and the "ε-enlargement" (defined above) are particular elements of this more general family of enlargements. Now we will use this family of enlargements to connect maximal monotone operators with convex functions.

This paper is organized as follows. In Section 2 we establish some notation and recall the definition of the family of enlargements introduced in [20], as well as some results presented there. In Section 3, we discuss the connection between convexity and enlargements on the family defined in the previous section. Then we introduce a family of convex functions, $\mathcal{H}$ associated to a maximal monotone operator. We prove that $\mathcal{H}$ is in one-to-one correspondence with the subfamily of closed enlargements. In Section 4 we discuss additional properties of $\mathcal{H}$ and study its extremal elements. In Section 5, an operator closely related with the conjugacy is introduced. We prove that $\mathcal{H}$ is invariant under this map. In Section 6 we study the particular features of the convex function associated to the Brønsted and Rockafellar ε-subdifferential enlargement. Moreover, we prove that one of these features is unique for $\mathcal{H}(\partial f)$. In Section 7 we use the theory of the previous sections for defining a family of convex merit functions for the proximal point method in a Hilbert space.

2 Basic definitions and Preliminary Results

From now on X is a real Banach space and X^* its dual. Given $x \in X$ and $v \in X^*$, $v(x)$ will be denoted indifferently by $\langle x, v\rangle$ and $\langle v, x\rangle$. In products of Banach spaces (e.g., $X \times X^*$, $\mathbb{R} \times X \times X^*$) we shall consider the canonical product topology.

Given a multifunction $T : X \rightrightarrows X^*$,

- The *graph* of T, $G(T)$ is

$$G(T) := \{(x, v) \,|\, x \in X \text{ and } v \in T(x)\}.$$

- T is *monotone* if $\langle u - v, x - y\rangle \geq 0$ for all $u \in T(x)$ and $v \in T(y)$, for all $x, y \in X$.

- T is *maximal monotone* if it is monotone and its graph is maximal with respect to this property, i.e., it is not properly contained in the graph of any other monotone operator.

The notation above is fairly standard. We will need some additional notation, to deal with enlargements.

Given $E : \mathbb{R}_+ \times X \rightrightarrows X^*$

- The *graph* of E, $G(E)$ is

$$G(E) := \{(b, x, v) \in \mathbb{R} \times X \times X^* \mid v \in E(b, x)\}.$$

  Define also

$$\tilde{G}(E) := \{(x, v, b) \in X \times X^* \times \mathbb{R} \mid v \in E(b, x)\}.$$

- The *closure* of E, $\overline{E} : \mathbb{R}_+ \times X \rightrightarrows X^*$ is defined as

$$\overline{E}(b, x) := \{v \in X^* \mid (b, x, v) \in \overline{G(E)}\}$$

  E is *closed* if $E = \overline{E}$.

- We say that E is an *enlargement* of $T : X \rightrightarrows X^*$ if for any $x \in X$, $b \geq 0$,

$$T(x) \subseteq E(b, x).$$

- We say that E is *non-decreasing* if for each $x \in X$, the application $\mathbb{R}_+ \ni \varepsilon \mapsto E(\varepsilon, x)$ is non-decreasing with respect to the inclusion, i.e.,

$$0 \leq \varepsilon \leq \varepsilon' \Rightarrow E(\varepsilon, x) \subseteq E(\varepsilon', x), \forall x \in X.$$

Remark 2.1 Let $E : \mathbb{R}_+ \times X \rightrightarrows X^*$.

As $G(E) \subseteq \mathbb{R}_+ \times X \times X^*$, $\overline{G(E)} \subseteq \mathbb{R}_+ \times X \times X^*$ and

$$G\left(\overline{E}\right) = \overline{G(E)}.$$

So, E is closed if and only if $G(E)$ is closed. In particular, $\overline{E}$ is closed.

It will be more convenient to work with $\tilde{G}(E)$ instead of $G(E)$. These two sets are basically the same, differing only in a permutation of coordinates. In particular,

$$\begin{aligned} \tilde{G}\left(\overline{E}\right) &= \overline{\tilde{G}(E)}, \\ \overline{E}(b, x) &= \{v \in X^* \mid (x, v, b) \in \overline{\tilde{G}(E)}\}, \ \forall b \geq 0, x \in X \end{aligned}$$

and E is closed if and only if $\tilde{G}(E)$ is closed.

Let f be a proper convex function on X. The multifunction ∂f is monotone and the ε-subdifferential of f is an enlargement of ∂f. As we work with enlargements as multifunctions on $\mathbb{R}_+ \times X$, we must have a notation for the multifunction $(\varepsilon, x) \mapsto \partial_\varepsilon f(x)$. It will be denoted by $\partial^{\text{BR}} f$. So,

$$\begin{aligned} &\partial^{\text{BR}} f : \mathbb{R}_+ \times X \rightrightarrows X^*, \\ &\partial^{\text{BR}} f(\varepsilon, x) = \{v \in X^* | f(y) \geq f(x) + \langle v, y - x\rangle - \varepsilon \text{ for all } y \in X\} \end{aligned} \tag{2}$$

and $\partial^{\text{BR}} f$ is an enlargement of ∂f. We will refer to $\partial^{\text{BR}} f$, after its authors, as *Brøndsted-Rockafellar enlargement of* ∂f. Trivially, $\partial^{\text{BR}} f$ is nondecreasing.

Given $T : X \rightrightarrows X^*$ *monotone*, the multifunction $(\varepsilon, x) \mapsto T^\varepsilon(x)$, defined in $\mathbb{R}_+ \times X$ according to (1), will be denoted by B^T.

Definition 2.2 *Let $T : X \rightrightarrows X^*$ be monotone. Then B^T is defined as*

$$\begin{aligned} &B^T : \mathbb{R}_+ \times X \rightrightarrows X^* \\ &B^T(b, x) := \{v \in X^* \mid \langle v - u, x - y\rangle \geq -b, \forall u \in T(y), y \in X\} \end{aligned} \tag{3}$$

Trivially, B^T is nondecreasing. Monotonicity of T implies in that B^T is an enlargement (of T).

The following "transportation formula" is satisfied by $\partial^{\text{BR}} f$ [10] and by B^T, when T is maximal monotone [5, 6].

Definition 2.3 *Denote by Δ_n the unit simplex of $\mathbb{R}^n$:*

$$\Delta_n := \left\{ \alpha \in \mathbb{R}^n \mid \alpha \geq 0, \sum_{i=1}^n \alpha_i = 1 \right\}. \tag{4}$$

We say that $E : \mathbb{R}_+ \times X \rightrightarrows X^*$ satisfies the transportation formula *if for any $n \in \mathbb{N}$, given n points*

$$(x^i, v^i, \varepsilon_i) \in \tilde{G}(E), \quad i = 1, \ldots, n$$

(or equivalently, $v^i \in E(\varepsilon_i, x^i)$) and some $\alpha \in \Delta_n$, defining

$$x = \sum_1^n \alpha_i x^i, \quad v = \sum_1^n \alpha_i v^i, \tag{5}$$

$$\varepsilon = \sum_1^n \alpha_i \varepsilon_i + \sum_1^n \alpha_i \langle x^i - x, v^i - v\rangle, \tag{6}$$

it holds that $(x, v, \varepsilon) \in \tilde{G}(E)$, or equivalently, $\varepsilon \geq 0$ and $v \in E(\varepsilon, x)$.

Observe that x, v, ε as in (5),(6) depend continuously on the n points (x^i, v^i, b_i). Hence, if E satisfies the transportation formula, then $\overline{E}$ likewise satisfies the transportation formula. Note also that if E is nondecreasing/an enlargement of T, then $\overline{E}$ is nondecreasing/an enlargement of T.

Now we recall the definition of a family of enlargements, introduced in [20], which will be central in our analysis.

Definition 2.4 *Let $T : X \rightrightarrows X^*$ be a monotone multifunction. Define $\mathbb{E}(T)$ as the family of multifunctions $E : \mathbb{R}_+ \times X \rightrightarrows X^*$ satisfying the following properties:*

(r1) *E is an enlargement of T, i.e.:*

$$E(\varepsilon, x) \supseteq T(x), \quad \forall \varepsilon \geq 0, x \in X.$$

(r2) *E is non-decreasing:*

$$E(\varepsilon', x) \supseteq E(\varepsilon, x), \quad \forall x \in X, \varepsilon' \geq \varepsilon \geq 0.$$

(r3) *E satisfies the transportation formula of Definition 2.3.*

The subfamily of those $E \in \mathbb{E}(T)$ which are closed will be denoted by $\mathbb{E}_c(T)$.

If f is a proper l.s.c. convex function, ∂f is maximal monotone and $\partial^{\text{BR}} f \in \mathbb{E}(\partial f)$. Furthermore, since the duality product is continuous on $X \times X^*$, f being l.s.c., from definition (2) it follows that the graph of $\partial^{\text{BR}} f$ is closed. So, $\partial^{\text{BR}} f \in \mathbb{E}_c(\partial f)$.

We are concerned with maximal monotone multifunctions. From now on, $T : X \rightrightarrows X^*$ is an *arbitrary* maximal monotone multifunction. Therefore (see [6, 20]) $B^T \in \mathbb{E}_c(T)$. In particular, $\mathbb{E}(T)$ is nonempty. The multifunction $S^T : \mathbb{R}_+ \times X \rightrightarrows X^*$,

$$S^T(\varepsilon, x) := \cap_{E \in \mathbb{E}(T)} E(\varepsilon, x). \tag{7}$$

is well defined. In [20] it was proved that $S^T \in \mathbb{E}(T)$ and that B^T, S^T are respectively the biggest and the smallest elements of $\mathbb{E}(T)$, in the (partial) order of inclusions of the graphs [20], i.e., $G(S^T) \subseteq G(E) \subseteq G(B^T)$ for any $E \in \mathbb{E}(T)$. Equivalently

$$E \in \mathbb{E}(T) \Rightarrow \tilde{G}(S^T) \subseteq \tilde{G}(E) \subseteq \tilde{G}(B^T).$$

Note that properties **(r1)**, **(r2)**, **(r3)** remain under the closure operation. Therefore, if $E \in \mathbb{E}(T)$, $\overline{E} \in \mathbb{E}_c(T)$. The family $\mathbb{E}_c(T)$ will play a central role in the next section.

Proposition 2.5 *$\overline{S^T}$, B^T are, respectively, the smallest and the biggest elements in $\mathbb{E}_c(T)$, with respect to the partial order of the inclusion of the graphs.*

Proof. As B^T is the biggest element of $\mathbb{E}(T)$ and is also closed, trivially B^T is the biggest element of $\mathbb{E}_c(T) \subseteq \mathbb{E}(T)$.

To prove that $\overline{S^T}$ is the smallest element of $\mathbb{E}_c(T)$, first observe that $\overline{S^T} \in \mathbb{E}_c(T)$ by the preceding discussion. Now take some $E \in \mathbb{E}_c(T) \subseteq \mathbb{E}(T)$. Then as S^T is the smallest element of $\mathbb{E}(T)$, $\tilde{G}(S^T) \subseteq \tilde{G}(E)$. Therefore,

$$\overline{\tilde{G}(S^T)} \subseteq \overline{\tilde{G}(E)}.$$

As E is closed and $\tilde{G}\left(\overline{S^T}\right) = \overline{\tilde{G}(S^T)}$, we conclude that $\tilde{G}\left(\overline{S^T}\right) \subseteq \tilde{G}(E)$. $\square$

A connection between transportation formula and convexity is given by [20, Lemma 3.2]. We restate this result in a more convenient formulation.

Lemma 2.6 *Let $\widehat{\Psi} : X \times X^* \times \mathbb{R} \to X \times X^* \times \mathbb{R}$,*

$$\widehat{\Psi}(x, v, b) := (x, v, b + \langle x, v\rangle).$$

Take $E : \mathbb{R}_+ \times X \rightrightarrows X^*$. Then, E satisfies the transportation formula if and only if $\widehat{\Psi}(\tilde{G}(E))$ is convex.

Proof. Let $R : \mathbb{R} \times X \times X^* \to X \times X^* \times \mathbb{R}$,

$$R(b, x, v) := (x, v, b).$$

Trivially $R^{-1} \circ \widehat{\Psi} \circ R(b, x, v) = (b + \langle x, v\rangle, x, v)$. According to [20, Lemma 3.2], E satisfies the transportation formula if and only if

$$R^{-1} \circ \widehat{\Psi} \circ R(G(E))$$

is convex. Since $R(G(E)) = \tilde{G}(E)$ and R is a linear isomorphism, Lemma 2.6 and [20, Lemma 3.2] are clearly equivalent. $\square$

3 Enlargements and convex functions

Any enlargement E can be studied via its graph, or equivalently, via the set $\tilde{G}(E) \subseteq X \times X^* \times \mathbb{R}_+$. Recall that an usual way to associate a scalar function to a set $A \subseteq X \times X^* \times \mathbb{R}$ is to consider the *lower envelope* of A [1], defined as $\alpha : X \times X^* \to \overline{\mathbb{R}}$,

$$\alpha(x, v) := \inf\{b \in \mathbb{R} \mid (x, v, b) \in A\}.$$

(In [14, Sec. 5] it is observed, in a finite dimensional setting, that the graph of α is the "lower boundary" of A). Trivially, $A \subseteq \mathrm{Epi}(\alpha)$. Additionally, $A = \mathrm{Epi}(\alpha)$ if A is closed and has an "epigraphical structure":

$$(x, v, b) \in A \Longrightarrow (x, v, b') \in A, \ \forall b' \geq b.$$

Given $E : \mathbb{R}_+ \times X \rightrightarrows X^*$, define $\lambda_E : X \times X^* \to \overline{\mathbb{R}}$,

$$\lambda_E(x, v) := \inf\{\varepsilon \geq 0 \mid v \in E(\varepsilon, x)\}. \tag{8}$$

Observe that

$$\lambda_E(x, v) = \inf\{b \mid (x, v, b) \in \tilde{G}(E)\}.$$

So, λ_E is the lower envelope of $\tilde{G}(E)$. For simplicity, we will call λ_E the lower envelope of E. As pointed out above, $\tilde{G}(E) \subseteq \mathrm{Epi}(\lambda_E)$. The conditions of $\tilde{G}(E)$ being closed and having an epigraphical structure are respectively equivalent to E being closed and nondecreasing. Enlargements satisfying these conditions are fully characterized by their lower envelope.

Proposition 3.1 *If $E : \mathbb{R}_+ \times X \rightrightarrows X^*$ is closed and nondecreasing then*

1. $\tilde{G}(E) = \mathrm{Epi}(\lambda_E)$,
2. λ_E *is (strongly) l.s.c.,*
3. $\lambda_E \geq 0$,
4. $E(b, x) = \{v \in X^* \mid \lambda_E(x, v) \leq b\}, \quad \forall b \in \mathbb{R}_+, x \in X.$

Furthermore, λ_E is the unique function from $X \times X^*$ to $\overline{\mathbb{R}}$ satisfying items 3, 4.

Proof. Item 1 follows from the facts mentioned above. Item 2 follows from closedness of E, item 1 and the fact that lower semicontinuity is equivalent to closedness of the epigraph. Item 3 holds trivially and item 4 follows from items 3 and 1.

For proving the last assertion of the proposition, observe that any function satisfying items 3,4 must satisfy also item 1. $\square$

Transformation/Application $\widehat{\Psi}$, defined in Lemma 2.6 correlates transportation formula and convexity. A similar transformation shall be performed on λ_E to get convexity. Given $E:\mathbb{R}_+\times X \rightrightarrows X^*$, define $\Lambda_E : X\times X^* \to \overline{\mathbb{R}}$,

$$\Lambda_E(x,v) := \lambda_E(x,v) + \langle x,v\rangle. \tag{9}$$

Note that Λ_E is the lower envelope of the set $\widehat{\Psi}(\tilde{G}(E))$ and

$$\mathrm{Epi}(\Lambda_E) = \widehat{\Psi}(\mathrm{Epi}(\lambda_E)). \tag{10}$$

Corollary 3.2 *Let E be a closed, non-decreasing enlargement of T. Then $E\in\mathbb{E}(T)$ if and only if Λ_E is convex.*

Proof. From Proposition 3.1 (item 1), $\tilde{G}(E) = \mathrm{Epi}(\lambda_E)$. Using also (10) we obtain

$$\mathrm{Epi}(\Lambda_E) = \widehat{\Psi}(\tilde{G}(E)).$$

Now, using Lemma 2.6 we conclude that E satisfies **(r3)** if and only if Λ_E is convex. As E satisfies **(r1)**, **(r2)** by assumption, the conclusion follows. □

Proposition 3.3 *Take $E\in\mathbb{E}_c(T)$. Then Λ_E is convex, l.s.c. and*

$$\begin{array}{l}\Lambda_E(x,v)\geq \langle x,v\rangle, \forall x\in X, v\in X^*,\\ v\in T(x)\Rightarrow \Lambda_E(x,v)=\langle x,v\rangle.\end{array} \tag{11}$$

Furthermore, the application $E\mapsto\Lambda_E$ is one-to-one in $\mathbb{E}_c(T)$.

Proof. Convexity of Λ_E follows from Corollary 3.2. Lower semicontinuity follows from Proposition 3.1 item 2. Item 3 of the same proposition implies the first condition in (11). To prove the second condition in (11), take $v\in T(x)$. As E is an enlargement of T, $v\in E(0,x)$, which readily implies $\lambda_E(x,v)=0$, and $\Lambda_E(x,v)=\langle x,v\rangle$.

By Proposition 3.1 item 1, the application $E\mapsto\lambda_E$ is one-to-one in $\mathbb{E}_c(T)$, which readily implies the last assertion of the proposition. □

Proposition 3.3 connects $\mathbb{E}_c(T)$ with a family of convex functions. Let us describe formally this family.

Definition 3.4 *Define $\mathcal{H}(T)$ as the family of l.s.c. convex functions $h: X\times X^*\to\overline{\mathbb{R}}$ such that*

$$h(x,v)\geq\langle x,v\rangle, \forall x\in X, v\in X^*, \tag{12}$$

$$v\in T(x)\Rightarrow h(x,v)=\langle x,v\rangle. \tag{13}$$

Note that, as T is maximal monotone, it has a nonempty graph and the elements of $\mathcal{H}(T)$ must be *proper* l.s.c. convex functions. Proposition 3.3 says that the application $E \mapsto \Lambda_E$ is one-to-one from $\mathbb{E}_c(T)$ into $\mathcal{H}(T)$. Our next aim is to prove that this application is a bijection between these families. Given $h : X \times X^* \to \overline{\mathbb{R}}$, define $L^h : \mathbb{R}_+ \times X \rightrightarrows X^*$,

$$L^h(\varepsilon, x) := \{v \in X^* \mid h(x, v) \leq \varepsilon + \langle x, v\rangle\}. \tag{14}$$

Proposition 3.5 *If $h \in \mathcal{H}(T)$, then $L^h \in \mathbb{E}_c(T)$ and $\Lambda_{L^h} = h$.*

Proof. Take $h \in \mathcal{H}(T)$. From (13) it follows that L^h is an enlargement of T. Trivially, L^h is nondecreasing. Furthermore, as h is l.s.c., L^h is closed.

Define $\ell : X \times X^* \to \overline{\mathbb{R}}$

$$\ell(x, v) := h(x, v) - \langle x, v\rangle.$$

Then, ℓ is nonnegative (condition (12)) and for any $\varepsilon \geq 0$, $x \in X$.

$$L^h(\varepsilon, x) = \{v \in X^* \mid \ell(x, v) \leq \varepsilon\}.$$

Hence, by Proposition 3.1, $\lambda_{L^h} = \ell$, which readily implies $\Lambda_{L^h} = h$. As h is convex, using Corollary 3.2, we conclude that $L^h \in \mathbb{E}_c(T)$. □

We summarize the above results in the following theorem.

Theorem 3.6 *The application*

$$\begin{array}{rcl} \mathbb{E}_c(T) & \to & \mathcal{H}(T) \\ E & \mapsto & \Lambda_E \end{array} \tag{15}$$

is a bijection, with inverse given by

$$\begin{array}{rcl} \mathcal{H}(T) & \to & \mathbb{E}_c(T) \\ h & \mapsto & L^h. \end{array} \tag{16}$$

Condition (13) on Definition 3.4 can be strengthened.

Corollary 3.7 *Take $h \in \mathcal{H}(T)$. Then for any $(x, v) \in X \times X^*$*

$$v \in T(x) \iff h(x, v) = \langle x, v\rangle.$$

Proof. If $v \in T(x)$, by (12) $h(x, v) = \langle x, v\rangle$. Conversely, suppose that $h(x, v) = \langle x, v\rangle$. Define $E = L^h$. Then

$$v \in E(0, x).$$

By Theorem 3.6 $E \in \mathbb{E}_c(T)$. In particular, $E \in \mathbb{E}(T)$. Using [20] we have $E(0, x) = T(x)$. So, $v \in T(x)$. □

4 Preliminary study of $\mathcal{H}(T)$

In this section we discuss some basic features of $\mathcal{H}(T)$. In the family of functions from $X \times X^*$ to $\overline{\mathbb{R}}$, we are considering the partial order:

$$h_1 \leq h_2 \Longleftrightarrow h_1(x,v) \leq h_2(x,v) \quad \forall (x,v) \in X \times X^*.$$

In the family of enlargements, we already have a partial order (inclusion of the graphs). With respect to these orders, the applications (15) and (16) are order inverting. Using Theorem 3.6 and Proposition 2.5, the following result follows.

Corollary 4.1 *The functions Λ_{B^T}, $\Lambda_{\overline{S^T}}$ belong to $\mathcal{H}(T)$ and are respectively the minimum and maximum of this family, i.e.,*

$$\Lambda_{B^T} \leq h \leq \Lambda_{\overline{S^T}}, \ \forall h \in \mathcal{H}(T).$$

The corollary above states that the elements of $\mathcal{H}(T)$ are located between Λ_{B^T} and $\Lambda_{\overline{S^T}}$. The converse is also true.

Corollary 4.2 *Let h be a l.s.c. convex function on X. If*

$$\Lambda_{B^T} \leq h \leq \Lambda_{\overline{S^T}}, \tag{17}$$

then $h \in \mathcal{H}(T)$.

Proof. Take h a l.s.c. convex function on X satisfying (17). As $\Lambda_{B^T} \in \mathcal{H}(T)$, it verifies (12) and combining this with (17), we get

$$\langle x, v\rangle \leq \Lambda_{B^T}(x,v) \leq h(x,v), \ \forall x \in X, \ v \in X^*.$$

So, (12) holds for h. Since $\Lambda_{\overline{S^T}} \in \mathcal{H}(T)$, it verifies (13) and combining this with (17), we get

$$v \in T(x) \Rightarrow h(x,v) \leq \Lambda_{\overline{S^T}}(x,v) = \langle x, v\rangle.$$

Combining the two above inequalities, we conclude that (13) also holds for h. ∎

The two preceding corollaries give another characterization of $\mathcal{H}(T)$:

$$\mathcal{H}(T) = \{h : X \times X^* \to \overline{\mathbb{R}} \ \text{convex, l.s.c.} \ | \Lambda_{B^T} \leq h \leq \Lambda_{\overline{S^T}}\}. \tag{18}$$

Since Λ_{B^T} and $\Lambda_{\overline{S^T}}$ hold special positions in the set $\mathcal{H}(T)$, these functions will be studied next. Before proceeding, we recall some well-known definitions of convex analysis.

Let $f : X \to \overline{\mathbb{R}}$, $A \subseteq X$

- The Legendre-Fenchel *conjugate* of f is the function $f^* : X^* \to \overline{\mathbb{R}}$,

$$f^*(v) := \sup\{\langle x, v\rangle - f(x) \mid x \in X\}.$$

- $\operatorname{conv}(f)$ is the biggest convex function majorized by f.
- The *indicator function* of A, is $\delta_A : X \to \overline{\mathbb{R}}$,

$$\delta_A(x) = \begin{cases} 0 & \text{if } x \in A, \\ +\infty & \text{otherwise.} \end{cases}$$

- The *convex hull* of A, denoted by $\operatorname{conv}(A)$ is the smallest convex set containing A.

First we investigate Λ_{B^T}. Brezis and Haraux introduced in [2] a scalar function connected with maximal monotone operators in Hilbert spaces. This concept extends trivially to the context of Banach spaces. Recall that $T : X \rightrightarrows X^*$ is an arbitrary maximal monotone operator. *Brezis-Haraux function* for T is $\beta_T : X \times X^* \to \overline{\mathbb{R}}$,

$$\beta_T(x, v) := \sup\{\langle y - x, v - u\rangle \mid (y, u) \in G(T)\}. \tag{19}$$

We claim that Brezis-Haraux function for T is the lower envelope of B^T, i.e.:

$$\beta_T = \lambda_{B^T}, \tag{20}$$

or, equivalently (from Proposition 3.1),

$$\tilde{G}(B^T) = \operatorname{Epi}(\beta_T).$$

First we will prove that

$$\beta_T \geq 0. \tag{21}$$

Suppose for contradictory purposes that $\beta_T(x^0, v^0) < 0$ for some $(x^0, v^0) \in X \times X^*$. Then,

$$\langle u - x^0, u - v^0\rangle > 0, \quad \forall (y, u) \in G(T).$$

Since T is maximal monotone, it follows that $(x^0, v^0) \in G(T)$. So, taking $(y, u) = (x^0, v^0) \in G(T)$ in (19) we conclude that $\beta_T(x^0, v^0) \geq 0$ in contradiction with the initial assumption.

Using (3) and (19), it follows that for any $\varepsilon \geq 0$ and $x \in X$,

$$B^T(\varepsilon, x) = \{v \in X^* \mid \beta_T(x, v) \leq \varepsilon\}. \tag{22}$$

The combination of (21) and (22) with the last statement of Proposition 3.1, establishes the claim (20).

Now we are ready to give a characterization of $\Lambda_{B^T}(x, v)$. Using (9), (20) and (19) we obtain

$$\begin{aligned} \Lambda_{B^T}(x, v) &= \beta_T(x, v) + \langle x, v\rangle \\ &= \sup\{\langle y - x, v - u\rangle + \langle x, v\rangle \mid (u, y) \in G(T)\} \\ &= \sup\{\langle y, v\rangle + \langle x, u\rangle - \langle y, u\rangle \mid (u, y) \in G(T)\}. \end{aligned} \tag{23}$$

Define the "ordered" duality product $\pi : X \times X^* \to \mathbb{R}$,

$$\pi(x, v) := \langle x, v\rangle.$$

With this notation, (23) reads

$$\Lambda_{B^T}(x, v) = \sup\{\langle y, v\rangle + \langle x, u\rangle - (\pi + \delta_{G(T)})(u, y) \mid (u, y) \in X \times X^*\}.$$

Therefore, for any $(x, v) \in X \times X^*$,

$$\Lambda_{B^T}(x, v) = \left(\pi + \delta_{G(T)}\right)^* (v, x), \tag{24}$$

where we are considering the natural injection $X \subseteq X^{**}$.

Now let us investigate $\Lambda_{\overline{S^T}}$. An auxiliary property of the closure of an enlargement will be needed. Recall that, given g an extended real function defined on a Banach space, the *lower closure* of g, $\operatorname{cl} g$ is defined as the greatest l.s.c. function majorized by g. Hence, (see [17, eq. 1(6), 1(7)])

$$\operatorname{cl} g(y) = \liminf_{y' \to y} g(y'), \quad \operatorname{Epi}(\operatorname{cl} g) = \overline{\operatorname{Epi}(g)}.$$

Proposition 4.3 *Let $E : \mathbb{R}_+ \times X \rightrightarrows X^*$. Then $\lambda_{\overline{E}} = \operatorname{cl} \lambda_E$.*

Proof. The proof, although trivial, will be given for the sake of completeness. By the definition, $\operatorname{cl} \lambda_E$ is l.s.c. and $\operatorname{cl} \lambda_E \leq \lambda_E$. Hence, $\tilde{G}(E) \subseteq \operatorname{Epi}(\operatorname{cl} \lambda_E)$ and $\operatorname{Epi}(\operatorname{cl} \lambda_E)$ is closed. Therefore, $\tilde{G}\left(\overline{E}\right) \subseteq \operatorname{Epi}(\operatorname{cl} \lambda_E)$ and

$$\operatorname{cl} \lambda_E \leq \lambda_{\overline{E}}. \tag{25}$$

Take $(x,v) \in X \times X^*$. If $\operatorname{cl} \lambda_E(x,v) = +\infty$, the above inequality implies trivially $\operatorname{cl} \lambda_E(x,v) = \lambda_{\overline{E}}(x,v)$. Suppose now

$$\operatorname{cl} \lambda_E(x,v) = b < +\infty.$$

As $\lambda_E \geq 0$, $b \geq 0$ and there exist a sequence $\{(x^k, v^k)\} \subseteq X \times X^*$ converging to (x,v) such that $\lambda_E(x^k, v^k)$ is finite and converges to b. Then, for each k, there exist a b_k, such that

$$\begin{array}{l} \lambda_E(x^k, v^k) \leq b^k \leq \lambda_E(x^k, v^k) + 1/k, \\ v^k \in E(b^k, x^k). \end{array}$$

Hence $\{(x^k, v^k, b_k)\} \subseteq \tilde{G}(E)$ converges to (x, v, b) and $v \in \overline{E}(b, x)$. This implies $\lambda_{\overline{E}}(x,v) \leq b = \operatorname{cl} \lambda_E(x,v)$. Using also (25) we conclude that equality holds also in this case. □

As a consequence,

$$\lambda_{\overline{S^T}} = \operatorname{cl} \lambda_{S^T}, \tag{26}$$

$$\Lambda_{\overline{S^T}} = \operatorname{cl} \Lambda_{S^T}, \tag{27}$$

where the first equality and continuity of the duality product were used to obtain the second equality. To proceed, we recall another result of [20].

Theorem 4.4 ([20, proof of Lemma 5.2 and Theorem 5.3]) *The set*

$$\{(b + \langle x, v\rangle, x, v) \mid v \in S^T(b, x)\}$$

is the convex hull of $\{(t, x, v) \mid t \geq \langle x, v\rangle,\ (x, v) \in G(T)\}$.

Let us restate this theorem in a more convenient notation.

Proposition 4.5 *Let* $\widehat{\Psi}$ *be as in Lemma 2.6. Then*

$$\widehat{\Psi}(\tilde{G}(S^T)) = \operatorname{conv} \operatorname{Epi}(\pi + \delta_{G(T)}).$$

Since Λ_E is the lower envelope of $\widehat{\Psi}(\tilde{G}(E))$, from the above proposition it follows that

$$\Lambda_{S^T} = \operatorname{conv}(\pi + \delta_{G(T)}), \tag{28}$$

$$\Lambda_{\overline{S^T}} = \operatorname{cl} \operatorname{conv}(\pi + \delta_{G(T)}), \tag{29}$$

where (27) was used in the last equality.

5 More on $\mathcal{H}(T)$

In this section we will study some additional features of $\mathcal{H}(T)$. Although $\mathcal{H}(T)$ is not contained in a vectorial space, it is, in some sense, convex. This fact is stated in the next proposition, whose proof follows directly from the Definition 3.4 or (18).

Proposition 5.1 *Take h_1, $h_2 \in \mathcal{H}(T)$ and $p, q \geq 0$, $p + q = 1$. Then*

$$ph_1 + qh_2 \in \mathcal{H}(T).$$

The set $\mathcal{H}(T)$ is also closed under the "sup operation":

Proposition 5.2 *Take $\{h_i\}_{i\in I}$ a nonempty family on $\mathcal{H}(T)$. Then*

$$\sup_{i\in I} h_i \in \mathcal{H}(T).$$

Given $\varphi : X \times X^* \to \overline{\mathbb{R}}$, define $\mathcal{J}\varphi : X \times X^* \to \overline{\mathbb{R}}$

$$\mathcal{J}\varphi(x, v) = \varphi^*(v, x),\ \forall x \in X, v \in X^*, \tag{30}$$

where the natural inclusion of X in X^{**} is being considered. Alternatively,

$$\mathcal{J}\varphi(x, v) = \sup\{\langle x, u\rangle + \langle v, y\rangle - \varphi(y, u) \mid (y, u) \in X \times X^*\}. \tag{31}$$

The operator $\mathcal{J}$ may also be seen in the framework of *generalized conjugacy.* Defining $A := X \times X^*$, $\Phi : A \times A \to \mathbb{R}$,

$$\Phi\Big((x, v), (y, u)\Big) := \langle x, u\rangle + \langle y, v\rangle,$$

and using the notation of [17, Ch. 11, Sec. L], we obtain

$$\mathcal{J}\varphi = \varphi^{\Phi}.$$

Our aim is to show that $\mathcal{H}(T)$ is invariant under $\mathcal{J}$. First observe that $\mathcal{J}\varphi$ is always l.s.c. and convex. Trivially, $\mathcal{J}$ is order reversing:

$$\varphi_1 \leq \varphi_2 \Rightarrow \mathcal{J}\varphi_2 \leq \mathcal{J}\varphi_1,$$

for any $\varphi_1, \varphi_2 : X \times X^* \to \overline{\mathbb{R}}$. Moreover,

$$\mathcal{J} \circ \mathcal{J}\varphi \leq \varphi. \tag{32}$$

From (30), $\mathcal{J}\varphi = \mathcal{J}\,\mathrm{cl}\,\varphi$ and $\mathcal{J}\varphi = \mathcal{J}\,\mathrm{conv}(\varphi)$. Therefore, using also (24) and Proposition 4.5 we obtain

$$\Lambda_{B^T} = \mathcal{J}\Lambda_{\overline{S^T}}. \tag{33}$$

Theorem 5.3 *The operator $\mathcal{J}$ maps $\mathcal{H}(T)$ into itself.*

Proof. Take $h \in \mathcal{H}(T)$. Then $\Lambda_{B^T} \leq h \leq \Lambda_{\overline{S^T}}$. Applying $\mathcal{J}$ on the members of this inequality, and taking into account that this operator is order reversing, we get

$$\mathcal{J}\Lambda_{\overline{S^T}} \leq \mathcal{J}h \leq \mathcal{J}\Lambda_{B^T}.$$

Using (33) in the first and the last terms of the above inequality, we have

$$\Lambda_{B^T} \leq \mathcal{J}h \leq \mathcal{J}\Big(\mathcal{J}\Lambda_{\overline{S^T}}\Big).$$

By (32), $\mathcal{J}\Big(\mathcal{J}\Lambda_{\overline{S^T}}\Big) \leq \Lambda_{\overline{S^T}}$. Now the result follows from (18) and the fact that $\mathcal{J}h$ is l.s.c. and convex. ∎

Remark 5.4 *In the general case, $\mathcal{J}^3 = \mathcal{J}$. If X is reflexive, then*

1. *$\mathcal{J}^2$ is the identity on $\mathcal{H}(T)$,*
2. *$\mathcal{J}$ is a bijection in $\mathcal{H}(T)$,*
3. *$\mathcal{J}(\Lambda_{B^T}) = \Lambda_{\overline{S^T}}$.*

Following the notation of [20], we say that $E : \mathbb{R}_+ \times X \rightrightarrows X^*$ is *additive* if

$$\langle v^1 - v^2, x^1 - x^2 \rangle \geq -(b_1 + b_2),\ \forall (x^1, v^1, b_1), (x^2, v^2, b_2) \in \tilde{G}(E). \tag{34}$$

An example of an additive enlargement is the ε-subdifferential. The enlargements in $\mathbb{E}(T)$ in general, are not additive. For this reason, additive elements of $\mathbb{E}(T)$ can be regarded as structurally closer to the ε-subdifferential enlargement than nonadditive elements. Denote by $\mathbb{E}_{ca}(T)$ the subfamily of closed additive elements of $\mathbb{E}(T)$. In [20], it was proved that $\overline{S^T} \in \mathbb{E}_{ca}(T)$. We will use $\mathcal{J}$ for characterizing $\mathbb{E}_{ca}(T)$.

Proposition 5.5 *Take $E \in \mathbb{E}_c(T)$. Then,*

$$E \in \mathbb{E}_{ca}(T) \iff \mathcal{J}\Lambda_E \leq \Lambda_E.$$

Proof. Take $E \in \mathbb{E}_{ca}(T)$. Then, by (34),

$$\langle v^1 - v^2, x^1 - x^2 \rangle \geq -(\lambda_E(x^1, v^1) + \lambda_E(x^2, v^2)),$$

for any $(x^1, v^1), (x^2, v^2) \in X \times X^*$. Algebraic manipulation yields,

$$\lambda_E(x^1, v^1) + \langle x^1, v^1 \rangle \geq \langle v^1, x^2 \rangle + \langle v^2, x^1 \rangle - (\lambda_E(x^2, v^2) + \langle x^2, v^2 \rangle),$$

or, equivalently,

$$\Lambda_E(x^1, v^1) \geq \langle v^1, x^2 \rangle + \langle v^2, x^1 \rangle - \Lambda_E(x^2, v^2).$$

Taking the sup in $(x^2, v^2) \in X \times X^*$, we conclude that $\Lambda_E(x^1, v^1) \geq \mathcal{J}\Lambda_E(x^1, v^1)$ for any $(x^1, v^1) \in X \times X^*$. So, $\mathcal{J}\Lambda_E \leq \Lambda_E$.

Conversely, suppose that $E \in \mathbb{E}_c(T)$ is such that $\mathcal{J}\Lambda_E \leq \Lambda_E$. Take

$$(x^1, v^1, b_1), (x^2, v^2, b_2) \in \tilde{G}(E).$$

Then,

$$\begin{array}{l} \lambda_E(x^1, v^1) \leq b_1, \\ \lambda_E(x^2, v^2) \leq b_2. \end{array} \tag{35}$$

Since $\mathcal{J}\Lambda_E \leq \Lambda_E$, in particular, $\mathcal{J}\Lambda_E(x^1, v^1) \leq \Lambda_E(x^1, v^1)$. Using also (31), we obtain

$$\langle x^2, v^1 \rangle + \langle x^1, v^2 \rangle - \Lambda_E(x^2, v^2) \leq \Lambda_E(x^1, v^1).$$

Substituting $\Lambda_E(x^i, v^i)$ by $\lambda_E(x^i, v^i) + \langle x^i, v^i \rangle$ for $i = 1, 2$, after some algebra, we get

$$\langle v^1 - v^2, x^1 - x^2 \rangle \geq -(\lambda_E(x^1, v^1) + \lambda_E(x^2, v^2)).$$

Using also (35), the desired inequality follows. $\square$

Define $\mathcal{H}_a(T)$ as

$$\mathcal{H}_a(T) := \{\Lambda_E \mid E \in \mathbb{E}_{ca}(T)\}.$$

From Proposition 5.5,

$$\mathcal{H}_a(T) := \{h \in \mathcal{H}(T) \mid \mathcal{J}h \leq h\}.$$

The subfamily $\mathcal{H}_a(T)$ has a similar structure to $\mathcal{H}(T)$.

Proposition 5.6 *If $h \in \mathcal{H}_a(T)$, then*

$$\{h' \in \mathcal{H}(T) \mid h' \geq h\} \subseteq \mathcal{H}_a(T).$$

Moreover,

1. *Take h_1, $h_2 \in \mathcal{H}_a(T)$ and $p, q \geq 0$, $p + q = 1$. Then*

$$ph_1 + qh_2 \in \mathcal{H}_a(T).$$

2. *Take $\{h_i\}_{i \in I}$ a nonempty family on $\mathcal{H}_a(T)$. Then*

$$\sup_{i \in I} h_i \in \mathcal{H}_a(T).$$

Observe that items 1 and 2 are formally identical to Propositions 5.1 and 5.2 respectively.

6 Features of the Brønsted-Rockafellar Enlargement

Let f be a proper l.s.c. convex function on X. It is known (f being a proper closed convex function) that f^* is a proper closed convex function on X^* and that (see [12, 15])

$$\begin{array}{c} f(x) + f^*(v) - \langle x, v \rangle \geq 0, \forall x \in X, v \in X^* \\ \text{with equality if and only if } v \in \partial f(x). \end{array}$$

Furthermore, (see [3])

$$\partial^{\mathrm{BR}} f(\varepsilon, x) = \{v \in X^* \mid f(x) + f^*(v) - \langle x, v \rangle \leq \varepsilon\}.$$

Therefore,

$$\lambda_{\partial^{\mathrm{BR}} f}(x, v) = f(x) + f^*(v) - \langle x, v \rangle,$$

and $\Lambda_{\partial^{\mathrm{BR}} f} : X \times X^* \to \overline{\mathbb{R}}$ is given by

$$\Lambda_{\partial^{\mathrm{BR}} f}(x, v) = f(x) + f^*(v). \tag{36}$$

Observe that $\Lambda_{\partial^{\mathrm{BR}} f}$ is a fixed point of $\mathcal{J}$:

$$\mathcal{J}\Lambda_{\partial^{\mathrm{BR}} f} = \Lambda_{\partial^{\mathrm{BR}} f}.$$

Indeed, $\mathcal{J}\Lambda_{\partial^{\mathrm{BR}} f}(x,v) = f^{**}(x) + f^*(v)$. As f is proper l.s.c. and convex, f^{**} coincides with f in X [12]. Another property of $\Lambda_{\partial^{\mathrm{BR}} f}$ is that it is a separable function on $X \times X^*$. In particular, $\mathcal{H}(\partial f)$ has a separable element.

Now we will prove that when $\mathcal{H}(T)$ has a separable element h, then T must be a subdifferential, and h must be the sum of f and f^*. In other words, this property is a distinctive feature of subdifferentials.

Theorem 6.1 *Suppose that some $h \in \mathcal{H}(T)$ is separable, i.e, there exist $f : X \to \mathbb{R} \cup \{+\infty\}$, $g : X^* \to \mathbb{R} \cup \{+\infty\}$ such that*

$$h(x,v) = f(x) + g(v), \quad \forall x \in X, v \in X^*. \tag{37}$$

Then, $T = \partial f$, $g = f^*$, and $h(x,v) = f(x) + f^*(v)$.

Proof. For establishing the assertion of the theorem, we will prove each of the following facts:

(i) $g(v) \geq f^*(v)\, \forall v \in X^*$ and $f(x) \geq g^*(x)\, \forall x \in X$

(ii) $g = f^*$ in $R(T)$ and $f = g^*$ in $D(T)$

(iii) $T = \partial f$,

(iv) $\partial g = \partial f^*$,(and hence $g \equiv f^*$).

(i): Using the fact that $h \in \mathcal{H}(T)$ and (37), we obtain

$$h(x,v) = f(x) + g(v) \geq \langle x, v \rangle.$$

Hence,

$$g(v) \geq \langle x, v \rangle - f(x),$$

for any $x \in X$. Taking supremum on $x \in X$, we get

$$g(v) \geq f^*(v),$$

for any $v \in X^*$. The proof of

$$f(x) \geq g^*(x),$$

for any $x \in X$ follows the same steps. This proves (i).

(ii): Now take $v \in Tx$. In this case we have

$$\langle x, v\rangle = h(x, v) = f(x) + g(v) \geq f(x) + f^*(v) \geq \langle x, v\rangle \qquad (38)$$

where we used (i) in the first inequality and Fenchel-Young inequality in the last one. Thus

$$g(v) = f^*(v).$$

Hence the first assertion of (ii) is established, and the same steps are used to establish the second assertion of (ii). As for (iii), observe that for any $v \in T(x)$, by (38)

$$\langle x, v\rangle = f(x) + f^*(v),$$

which implies that $v \in \partial f(x)$. As a consequence,

$$G(T) \subseteq G(\partial f),$$

and the maximality of T and yields (iii). In order to establish (iv), we consider the function $\tilde{h} : X^* \times X^{**} \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\tilde{h}(v, z) := g(v) + f^{**}(z).$$

We claim that $\tilde{h} \in \mathcal{H}(\partial f^*)$. Observe that, if this fact is true, then we can apply items (i), (ii), (iii) to $\tilde{h}$. More precisely, if $\tilde{h} \in \mathcal{H}(\partial f^*)$, applying (iii) to $h := \tilde{h}$ and $T := \partial f^*$, we get $\partial f^* = \partial g$. This equality, combined with (ii), yields $g = f^*$. So it is enough to prove the claim.

For establishing the claim, we must show that

(a)

$$\tilde{h}(v, z) \geq \langle z, v\rangle, \ \forall (v, z) \in X^* \times X^{**}.$$

and

(b)

$$\tilde{h}(v, z) = \langle z, v\rangle, \ \forall (v, z) \in G(f^*).$$

Using the first assertion of (i), we obtain

$$\tilde{h}(v, z) = f^{**}(z) + g(v) \geq f^{**}(z) + f^*(v) \geq \langle z, v\rangle,$$

and (a) is established. For proving (b), we quote Proposition 1 of [15], which characterizes the graph of ∂f^* in the following way:

$(z, v) \in G(\partial f^*)$ if and only if there exists a net $\{(x_i, v_i) \in G(\partial f)\}_{i \in I}$ such that $\{x_i\}_{i \in I}$ is a bounded net converging to z in the w^{**}-topology, and such that $\{v_i\}_{i \in I}$ is a net converging to v in the strong topology.

Consider now an element $(v, z) \in G(\partial f^*)$. Taking a net as in the above mentioned proposition:

$$\begin{aligned} \tilde{h}(v,z) &= f^{**}(z) + g(v) \leq \liminf_{i \in I}\{f^{**}(x_i) + g(v_i)\} \\ &= \liminf_{i \in I}\{f(x_i) + g(v_i)\} = \liminf_{i \in I} h(x_i, v_i) = \liminf_{i \in I}\langle x_i, v_i\rangle \\ &= \langle z, v\rangle \end{aligned}$$

We used above the lower semicontinuity of f^{**} and g, together with the fact that $f = f^{**}$ on X. The third equality holds by definition of h, the fourth one follows from (iii), the fact that $(x_i, v_i) \in G(\partial f)$ and definition of $\mathcal{H}(T)$. Thus, for $(v, z) \in G(\partial f^*)$:

$$\tilde{h}(v, z) \leq \langle z, v\rangle,$$

which combined with (a) yields (b). This completes the proof. $\square$

Remark 6.2 *Note that, under the assumptions of Theorem 6.1,*

$$h = \Lambda_{\partial^{\mathrm{BR}} f},$$

where we are using (36)*. As a consequence, only when T is the subdifferential, $\mathcal{H}(T)$ has a separable element. Moreover, this separable element is the image by $\Lambda_{(\cdot)}$ of the ε-subdifferential enlargement.*

7 Merit Functions for the Proximal Problem

From now on, X is a real Hilbert space. A relevant problem in this context is that of finding a zero of the maximal monotone operator $T : X \rightrightarrows X$, i.e.: find $x \in X$ such that

$$0 \in T(x). \tag{39}$$

The proximal point method[16] solves this problem, by generating a sequence $\{z^k\}$ such that

$$\begin{array}{l} z^k \text{ approximated solution of} \\ 0 \in c_k T(x) + x - z^{k-1}. \end{array} \tag{40}$$

Weak convergence of $\{z^k\}$ for a solution of (39) follows from the conditions

$$\begin{aligned} &\|z^k - (c_k T + I)^{-1} z^{k-1}\| \le e_k, \quad \sum e_k < +\infty, \\ &c_k \ge \bar{c} > 0. \end{aligned} \tag{41}$$

Note that the "k-th proximal subproblem": find x such that

$$0 \in c_k T(x) + x - z^{k-1}, \tag{42}$$

is clearly equivalent to the k-th "proximal system": find v, x such that

$$\left\{ \begin{array}{l} v \in T(x), \\ c_k v + x - z^{k-1} = 0. \end{array} \right. \tag{43}$$

As we will focus on the resolution of (43), we may simplify the notation taking $c_k = c$ and $z^{k-1} = z$. Hence, for fixed $c > 0$ and $z \in X$, we consider the proximal system: find v, x such that

$$\left\{ \begin{array}{l} v \in T(x), \\ cv + x - z = 0. \end{array} \right. \tag{44}$$

We will show that the proximal system (44) is equivalent to a convex minimization problem and that approximated solutions of this minimization problem yield approximated solutions of the proximal system.

Take $h \in \mathcal{H}(T)$ and $E \in \mathbb{E}_c(T)$ the associated enlargement, i.e., $h = \Lambda_E$. Define $\varphi^E_{c,z} : X \times X \to \overline{\mathbb{R}}$

$$\varphi^E_{c,z}(x, v) := \|cv + x - z\|^2 + 2c\lambda_E(x, v). \tag{45}$$

Then $\varphi^E_{c,z} \ge 0$. Furthermore

$$\varphi^E_{c,z}(x, v) = 0 \Leftrightarrow \begin{array}{l} v \in T(x), \\ cv + x - z = 0. \end{array}$$

Therefore, the minimization problem

$$\min \varphi^E_{c,z}(x, v), \quad (x, v) \in X \times X \tag{46}$$

and the proximal system (44) have the same (unique) solution. Note also that the optimal value of (46) is 0. Direct algebraic manipulation yields

$$\begin{aligned} &\varphi^E_{c,z}(x, v) = \|cv\|^2 - 2c\langle v, z\rangle + \|x - z\|^2 + 2c\left(\langle x, v\rangle + \lambda_E(x, v)\right) \\ &= \|x - z\|^2 + c^2\|v\|^2 - 2c\langle v, z\rangle + 2c\Lambda_E(x, v) \\ &= \|x - z\|^2 + c^2\|v\|^2 - 2c\langle v, z\rangle + 2ch(x, v). \end{aligned}$$

Hence, $\varphi^E_{c,z}$ is a proper closed convex function, strongly-$(1,c^2)$ convex in the following sense:

$$\text{if } (\xi,w) \in \partial\varphi^E_{c,z}(x,v) \text{ then, for any } y,u \in X$$
$$\varphi^E_{c,z}(y,u) \geq \varphi^E_{c,z}(x,v) + \langle \xi, y-x\rangle + \langle w, u-v\rangle + \|y-x\|^2 + c^2\|u-v\|^2 \tag{47}$$

We claim that $\varphi^E_{c,z}$ is a merit function for the proximal system (44) in the following sense:

1. $\varphi^E_{c,z}$ is nonnegative.
2. $\varphi^E_{c,z}$ assumes the value 0 only at the solution of (44).
3. If $\varphi^E_{c,z}(x,v)$ is "small" then (x,v) is "close" to the solution of the (44).

Only item 3 remains to be proved.

Proposition 7.1 *Let z^*, v^* be the exact solution of the proximal system* (44)*, i.e.,*

$$\begin{aligned} z^* &= (cT+I)^{-1}z \\ v^* &= (1/c)(z-z^*) \end{aligned}$$

(in particular, z^* is the solution of (42)*). Then, for any $x,v \in X$*

$$\|x-z^*\|^2 \leq \|x-z^*\|^2 + c^2\|v-v^*\|^2 \leq \varphi^E_{c,z}(x,v).$$

Proof. Only the last inequality must be proved. Recall that

$$\begin{aligned} \varphi^E_{c,z}(z^*,v^*) &= 0 \\ \varphi^E_{c,z} &\geq 0. \end{aligned}$$

Therefore, $(0,0) \in \partial\varphi^E_{c,z}(z^*,v^*)$. Using also (47) we get

$$\begin{aligned} \varphi^E_{c,z}(z^*,v^*) &\geq \varphi^E_{c,z}(x,v) + c^2\|v-v^*\|^2 + \|x-z^*\|^2 \\ &\geq c^2\|v-v^*\|^2 + \|x-z^*\|^2. \end{aligned}$$

□

The merit function $\varphi^{B^T}_{c,z}$ has been proposed and studied in [18]. Since Λ_{B^T} is the smallest element of $\mathcal{H}(T)$, it follows that

$$\varphi^{B^T}_{c,z} \leq \varphi^E_{c,z}$$

for any $E \in \mathbb{E}_c(T)$. An hybrid extragradient proximal point method for solving (39) has been introduced in [19, 18]. This method employs at each iteration an approximated minimizer of $\varphi^{BT}_{c_k,z^{k-1}}$, say $(\tilde{x}^k, \tilde{v}^k)$, to perform the extragradient step

$$z^k := z^{k-1} - c_k \tilde{v}^k.$$

The approximation criterion is:

$$\varphi^{BT}_{c_k,z^{k-1}}(\tilde{x}^k, \tilde{v}^k) \le \sigma \|\tilde{x}^k - z^{k-1}\|^2,$$

where $0 \le \sigma < 1$. An example of computable upper bounds for the merit function $\varphi^{BT}_{c,z}$ with algorithmic applications can be found in [18].

Now that we have merit functions for the generic proximal system (44), we discuss merit functions for the generic proximal problem: find x such that

$$0 \in cT(x) + x - z. \tag{48}$$

Define $\psi_1, \psi_2 : X \to \overline{\mathbb{R}}$ as

$$\begin{aligned} \psi_1(x) &:= \inf\{\varphi^E_{c,z}(x, v) \mid v \in X\}, \\ \psi_2(x) &:= \varphi^E_{c,z}(x, (1/c)(z - x)). \end{aligned}$$

The functions ψ_1 and ψ_2 are convex, nonnegative and have for minimizer the exact solution of (48). The optimal value of both is again zero. The merit function ψ_2 reduces to

$$\psi_2(x) = 2c\lambda_E(x, (1/c)(z - x)).$$

In the context of nonsmooth convex optimization, an approximated solution of the proximal problem

$$0 \in c\partial f(x) + (x - z),$$

is an x such that, for a small ε,

$$0 \in c\partial_\varepsilon f(x) + (x - z)$$

or, equivalently, $1/c(z - x) \in c\partial_\varepsilon f(x)$. This condition is equivalent to

$$\lambda_{\partial^{\mathrm{BR}} f}(x, 1/c(z - x)) \le \varepsilon,$$

or, equivalently,

$$\psi_2(x) \le 2c\varepsilon.$$

Thus, the merit function ψ_2 is a generalization, for maximal monotone operators, of a merit function which has been already applied implicitly in the context of nonsmooth convex optimization [7].

References

[1] Avriel, Mordecai. *Nonlinear programming, Analysis and methods.* Prentice-Hall Inc., Englewood Cliffs, N.J., 1976.

[2] Haïm Brezis and Alain Haraux. Image d'une somme d'opérateurs monotones et applications. *Israel J. Math.*, 23(2):165–186, 1976.

[3] A. Brøndsted and R. T. Rockafellar. On the subdifferentiability of convex functions. *Proceedings of the American Mathematical Society*, 16:605–611, 1965.

[4] Regina S. Burachik, Alfredo N. Iusem, and B. F. Svaiter. Enlargement of monotone operators with applications to variational inequalities. *Set-Valued Analysis*, 5(2):159–180, 1997.

[5] Regina S. Burachik, Claudia A. Sagastizábal, and B. F. Svaiter. ϵ-enlargements of maximal monotone operators: theory and applications. In M. Fukushima and L. Qi, editors, *Reformulation: nonsmooth, piecewise smooth, semismooth and smoothing methods (Lausanne, 1997)*, volume 22 of *Applied Optimization*, pages 25–43. Kluwer Acad. Publ., Dordrecht, 1999.

[6] Regina Sandra Burachik and B. F. Svaiter. ε-enlargements of maximal monotone operators in Banach spaces. *Set-Valued Analysis*, 7(2):117–132, 1999.

[7] Cominetti, R., Coupling the proximal point algorithm with approximation methods, *Journal of Optimization Theory and Applications*, **95**, (1997), 581–600.

[8] Hiriart-Urruty, J.-B.: Lipschitz r-continuity of the approximate subdifferential of a convex function, *Math. Scand.* **47** (1980), 123–134.

[9] Kachurovskii, R. I., On monotone operators and convex functionals, *Uspekhi Mathematicheskikh Nauk*, 15: 213-241, 1960.

[10] Lemarechal, C. *Extensions diverses des méthodes de gradient et applications.* Thèse d'Etát, Université de Paris IX, 1980.

[11] Minty, G. J., On the monotonicity of the gradient of a convex function, *Pacific Journal of Mathematics*, 14:243-247, 1964.

[12] J.-J. Moreau. Functionelles convexes. mimeographed lecture notes, Collège de France, Paris, 1967.

[13] Nurminskii, E. A., The continuity of the approximate subdifferential of a convex function. *Kibernetika* (Kiev) 5, 148–149, 1978.

[14] Rockafellar, R. Tyrrell *Convex analysis*, Princeton University Press, Princeton, N.J., 1970.

[15] R. T. Rockafellar. On the maximal monotonicity of subdifferential mappings. *Pacific Journal of Mathematics*, 33:209–216, 1970.

[16] R. Tyrrell Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control Optimization*, 14(5):877–898, 1976.

[17] R. Tyrrell Rockafellar, Roger J-B. Wets. *Variational Analysis.* Springer Verlag, Berlin Heidelberg, 1998.

[18] M. V. Solodov and B. F. Svaiter. Error bounds for proximal point subproblems and associated inexact proximal point algorithms. To appear *Mathematical Programming.*

[19] M. V. Solodov and B. F. Svaiter. An inexact hybrid extragradient-proximal point algorithm using the enlargement of a maximal monotone operator. *Set-Valued Analysis*, 7(4):323–345, December 1999.

[20] B. F. Svaiter. Regular enlargements of maximal monotone operators. To appear, *Set-Valued Analysis.*