PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 131, Number 12, Pages 3851–3859
S 0002-9939(03)07083-7
Article electronically published on May 5, 2003

# FIXED POINTS IN THE FAMILY OF CONVEX REPRESENTATIONS OF A MAXIMAL MONOTONE OPERATOR

B. F. SVAITER

(Communicated by Jonathan M. Borwein)

Abstract. Any maximal monotone operator can be characterized by a convex function. The family of such convex functions is invariant under a transformation connected with the Fenchel-Legendre conjugation. We prove that there exists a convex representation of the operator which is a fixed point of this conjugation.

## 1. Introduction

Let $X$ be a real Banach space and $X^*$ its dual. It is usual to identify a point to set operator $T : X \rightrightarrows X^*$ with its graph, $\{(x, x^*) \in X \times X^* \,|\, x^* \in T(x)\}$. We will use the notation $\langle x, x^*\rangle$ for the duality product $x^*(x)$ of $x \in X$, $x^* \in X^*$.

An operator $T : X \rightrightarrows X^*$ is *monotone* if

$$(x, x^*), (y, y^*) \in T \Rightarrow \langle x - y, x^* - y^*\rangle \geq 0$$

and is *maximal monotone* if it is monotone and

$$\forall (y, y^*) \in T, \langle x - y, x^* - y^*\rangle \geq 0 \Rightarrow (x, x^*) \in T.$$

Krauss [11] managed to represent maximal monotone operators by subdifferentials of saddle functions on $X \times X$. After that, Fitzpatrick [8] proved that maximal monotone operators can be represented by convex functions on $X \times X^*$. Later on, Simons [19] studied maximal monotone operators using a min-max approach. Recently, the convex representation of maximal monotone operators was rediscovered by Burachik and Svaiter [7] and Martinez-Legaz and Théra [13]. In [7], some results on enlargements are used to perform a systematic study of the family of convex functions which represents a given maximal monotone operator. Here we are concerned with this kind of representation.

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

$$f^*(x^*) := \sup_{x \in x} \langle x, x^*\rangle - f(x).$$

Received by the editors July 31, 2002.
2000 *Mathematics Subject Classification.* Primary 47H05.
*Key words and phrases.* Maximal monotone operators, conjugation, convex functions.
This work was partially supported by CNPq Grant 301200/93-9(RN) and by PRONEX–Optimization.

The *subdifferential* of $f$ is the operator $\partial f : X \rightrightarrows X^*$,

$$\partial f(x) := \{x^* \in X^* \mid f(y) \geq f(x) + \langle y - x, x^* \rangle, \, \forall y \in X\}.$$

If $f$ is convex, lower semicontinuous and proper, then $\partial f$ is maximal monotone [17]. From the previous definitions, we have the *Fenchel–Young inequality:* for all $x \in X$, $x^* \in X^*$,

$$f(x) + f^*(x^*) \geq \langle x, x^* \rangle \,, \;\; f(x) + f^*(x^*) = \langle x, x^* \rangle \iff x^* \in \partial f(x).$$

So, defining $h_{\mathrm{FY}} : X \times X^* \to \overline{\mathbb{R}}$,

$$h_{\mathrm{FY}}(x, x^*) := f(x) + f^*(x^*), \tag{1.1}$$

we observe that this function fully characterizes $\partial f$. Assume that $f$ is convex, lower semicontinuous and proper. In this case, $\partial f$ is maximal monotone. Moreover, if we use the canonical injection of $X$ in to $X^{**}$, then $f^{**}(x) = f(x)$ for all $x \in X$. Hence, for all $(x, x^*) \in X \times X^*$,

$$(h_{\mathrm{FY}})^*(x, x^*) = h_{\mathrm{FY}}(x, x^*).$$

Our aim is to prove that any maximal monotone operator has a convex representation with a similar property.

From now on, $T : X \rightrightarrows X^*$ is a maximal monotone operator. Define, as in [8], $\mathcal{H}(T)$ to be the family of convex lower semicontinuous functions $h : X \times X^* \to \overline{\mathbb{R}}$ such that

$$\begin{aligned} \forall (x, x^*) \in X \times X^*, \quad & h(x, x^*) \geq \langle x, x^* \rangle, \\ & (x, x^*) \in T \Rightarrow h(x, x^*) = \langle x, x^* \rangle. \end{aligned} \tag{1.2}$$

This family is nonempty [8]. Moreover, for any $h \in \mathcal{H}(T)$, $h(x, x^*) = \langle x, x^* \rangle$ if and only if $(x, x^*) \in T$ [7]. Hence, any element of $\mathcal{H}(T)$ fully characterizes, or represents, $T$. Since the sup of convex lower semicontinuous function is also convex and lower semicontinuous, also using (1.2) we conclude that the sup of any (nonempty) subfamily of $\mathcal{H}(T)$ is still in $\mathcal{H}(T)$.

The dual of $X \times X^*$ is $X^* \times X^{**}$. So, for $(x, x^*) \in X \times X^*$, $(y^*, y^{**}) \in X^* \times X^{**}$,

$$\langle (x, x^*) \, , \, (y^*, y^{**}) \rangle = \langle x, y^* \rangle + \langle x^*, y^{**} \rangle.$$

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

$$\mathcal{J}h(x, x^*) := h^*(x^*, x), \tag{1.3}$$

where $h^*$ stands for the Fenchel-Legendre conjugate of $h$ and the canonical inclusion of $X$ in $X^{**}$ is being used. Equivalently,

$$\mathcal{J}h(x, x^*) = \sup_{(y, y^*) \in X \times X^*} \langle x, y^* \rangle + \langle y, x^* \rangle - h(y, y^*). \tag{1.4}$$

Trivially, $\mathcal{J}$ inverts the natural order of functions, i.e., if $h \geq h'$, then $\mathcal{J}h' \geq \mathcal{J}h$. The family $\mathcal{H}(T)$ is invariant under the application $\mathcal{J}$ [7]. The aim of this paper is to prove that there exists an element $h \in \mathcal{H}(T)$ such that $\mathcal{J}h = h$.

The application $\mathcal{J}$ can be studied in the framework of *generalized conjugation* [18, Ch. 11, Sec. L]. With this aim, define

$$\begin{aligned} &\Phi : (X \times X^*) \times (X \times X^*) :\to \mathbb{R}, \\ &\Phi((x, x^*), (y, y^*)) := \langle x, y^* \rangle + \langle y, x^* \rangle. \end{aligned}$$

Given $h : X \times X^* \to \overline{\mathbb{R}}$, let $h^\Phi$ be the conjugate of $h$ with respect to the coupling function $\Phi$,

$$h^\Phi(x, x^*) := \sup_{(y,y^*)\in X\times X^*} \Phi((x, x^*), (y, y^*)) - h(y, y^*). \tag{1.5}$$

Now we have

$$\mathcal{J}h = h^\Phi,$$

and, in particular,

$$h \geq h^{\Phi\Phi} = \mathcal{J}^2 h. \tag{1.6}$$

## 2. Proof of the main theorem

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

$$\sigma_T := \sup_{h\in\mathcal{H}(T)} h.$$

Since $\mathcal{H}(T)$ is "closed" under the sup operation, we conclude that $\sigma_T$ is the biggest element of $\mathcal{H}(T)$. Combining this fact with the inclusion $\mathcal{J}\sigma_T \in \mathcal{H}(T)$ we conclude that

$$\sigma_T \geq \mathcal{J}\sigma_T.$$

For a more detailed discussion on $\sigma_T$, we refer the reader to [7, eq. (35)]. The above inequality will be, in some sense, our departure point. Now define

$$\mathcal{H}_a(T) := \{h \in \mathcal{H}(T) \,|\, h \geq \mathcal{J}h\}.$$

The family $\mathcal{H}_a(T)$ is connected with a family of enlargements of $T$ which shares with the $\varepsilon$-subdifferential a special property (see [7]). We already know that $\sigma_T \in \mathcal{H}_a(T)$. Later on, we will use the following construction of elements in this set.

**Proposition 2.1.** *Take $h \in \mathcal{H}(T)$ and define*

$$\hat{h} = \max\ h, \mathcal{J}h.$$

*Then $\hat{h} \in \mathcal{H}_a(T)$.*

*Proof.* Since $h$ and $\mathcal{J}h$ are in $\mathcal{H}(T)$, $\hat{h} \in \mathcal{H}(T)$. By definition,

$$\hat{h} \geq h,\ \hat{h} \geq \mathcal{J}h.$$

Applying $\mathcal{J}$ on these inequalities and using (1.6) for majorizing $\mathcal{J}^2 h$ we obtain

$$\mathcal{J}h \geq \mathcal{J}\hat{h},\ h \geq \mathcal{J}\hat{h}.$$

Hence, $\hat{h} \geq \mathcal{J}\hat{h}$. □

For $h \in \mathcal{H}(T)$ define

$$L(h) := \{g \in \mathcal{H}(T) \,|\, h \geq g \geq \mathcal{J}g\}.$$

The operator $\mathcal{J}$ inverts the order. Therefore, $L(h) \neq \emptyset$ if and only if $h \geq \mathcal{J}h$, i.e., $h \in \mathcal{H}_a(T)$. We already know that $L(\sigma_T) \neq \emptyset$.

**Proposition 2.2.** *For any $h \in \mathcal{H}_a(T)$, the family $L(h)$ has a minimal element.*

*Proof.* We shall use the Zorn Lemma. Let $\mathcal{C} \subseteq L(h)$ be a (nonempty) chain, that is, $\mathcal{C}$ is totally ordered. Take $h' \in \mathcal{C}$. For any $h'' \in \mathcal{C}$, $h' \geq h''$ or $h'' \geq h'$. In the first case we have $h' \geq h'' \geq \mathcal{J}h''$, and in the second case, $h' \geq \mathcal{J}h' \geq \mathcal{J}h''$. Therefore,

$$h' \geq \mathcal{J}h'',\ \forall h', h'' \in \mathcal{C}. \tag{2.1}$$

Now define

$$\hat{g} = \sup_{h' \in \mathcal{C}} \mathcal{J}h'. \tag{2.2}$$

Since $\mathcal{H}(T)$ is invariant under $\mathcal{J}$ and also closed with respect to the sup, we have $\hat{g} \in \mathcal{H}(T)$. From (2.1), (2.2) it follows that

$$h' \geq \hat{g} \geq \mathcal{J}h',\ \forall h' \in \mathcal{C}.$$

Applying $\mathcal{J}$ to the above inequalities, and also using (1.6), we conclude that

$$h' \geq \mathcal{J}\hat{g} \geq \mathcal{J}h',\ \forall h' \in \mathcal{C}. \tag{2.3}$$

Since $\hat{g} \in \mathcal{H}(T)$, $\mathcal{J}\hat{g} \in \mathcal{H}(T)$. Taking the sup on $h' \in \mathcal{C}$, in the right-hand side of the last inequality, we get

$$\mathcal{J}\hat{g} \geq \hat{g}.$$

Applying $\mathcal{J}$, again, we obtain

$$\mathcal{J}\hat{g} \geq \mathcal{J}(\mathcal{J}\hat{g}).$$

Take some $h' \in \mathcal{C}$. By the definition of $L(h)$ and (2.3), we conclude that $h \geq h' \geq \mathcal{J}\hat{g}$. Hence $\mathcal{J}\hat{g}$ belongs to $L(h)$ and is a lower bound for any element of $\mathcal{C}$. Now we apply the Zorn Lemma to conclude that $L(h)$ has a minimal element. □

The minimal elements of $L(h)$ (for $h \in \mathcal{H}_a(T)$) are the natural candidates for being fixed points of $\mathcal{J}$. First we will show that they are fixed points of $\mathcal{J}^2$. Observe that, since $\mathcal{J}$ inverts the order of functions, $\mathcal{J}^2$ preserves it, i.e., if $h \geq h'$, then $\mathcal{J}^2 h \geq \mathcal{J}^2 h'$. Moreover, $\mathcal{J}^2$ maps $\mathcal{H}(T)$ in itself.

**Proposition 2.3.** *Take $h \in \mathcal{H}_a(T)$ and let $h_0$ be a minimal element of $L(h)$. Then $\mathcal{J}^2 h_0 = h_0$.*

*Proof.* First observe that $\mathcal{J}^2 h_0 \in \mathcal{H}(T)$. By assumption, $h_0 \geq \mathcal{J}h_0$. Applying $\mathcal{J}^2$ in this inequality we get

$$\mathcal{J}^2 h_0 \geq \mathcal{J}^2(\mathcal{J}h_0) = \mathcal{J}(\mathcal{J}^2 h_0).$$

Since $h \geq h_0$ and, by (1.6), $h_0 \geq \mathcal{J}^2 h_0$, we conclude that $h \geq \mathcal{J}^2 h_0 \geq \mathcal{J}(\mathcal{J}^2 h_0)$. Hence $\mathcal{J}^2 h_0 \in L(h)$. Again using the inequality $h_0 \geq \mathcal{J}^2 h_0$ and the minimality of $h_0$, the conclusion follows. □

**Theorem 2.4.** *Take $h \in \mathcal{H}(T)$ such that $h \geq \mathcal{J}h$. Then $h_0 \in L(h)$ is minimal (on $L(h)$) if and only if $h_0 = \mathcal{J}h_0$.*

*Proof.* Assume first that $h_0 = \mathcal{J}h_0$. If $h' \in L(h)$ and

$$h_0 \geq h',$$

then, applying $\mathcal{J}$ on this inequality and using the definition of $L(h)$ we conclude that

$$h' \geq \mathcal{J}h' \geq \mathcal{J}h_0 = h_0.$$

Combining the above inequalities we obtain $h' = h_0$. Hence $h_0$ is minimal on $L(h)$.

Assume now that $h_0$ is minimal on $L(h)$. By the definition of $L(h)$, $h_0 \geq \mathcal{J}h_0$. Suppose that for some $(x_0, x_0^*)$,

$$h_0(x_0, x_0^*) > \mathcal{J}h_0(x_0, x_0^*). \tag{2.4}$$

We shall prove that this assumption is contradictory. By Proposition 2.3, $h_0 = \mathcal{J}(\mathcal{J}h_0)$. Hence, the above inequality can be expressed as

$$\mathcal{J}(\mathcal{J}h_0)(x_0, x_0^*) > \mathcal{J}h_0(x_0, x_0^*),$$

or equivalently

$$\sup_{(y,y^*)\in X\times X^*} \langle y, x_0^*\rangle + \langle x_0, y^*\rangle - \mathcal{J}h_0(y, y^*) > \mathcal{J}h_0(x_0, x_0^*).$$

Therefore, there exists some $(y_0, y_0^*) \in X \times X^*$ such that

$$\langle y_0, x_0^*\rangle + \langle x_0, y_0^*\rangle - \mathcal{J}h_0(y_0, y_0^*) > \mathcal{J}h_0(x_0, x_0^*). \tag{2.5}$$

In particular, $\mathcal{J}h_0(y_0, y_0^*), \mathcal{J}h_0(x_0, x_0^*) \in \mathbb{R}$. Interchanging $\mathcal{J}h_0(y_0, y_0^*)$ with $\mathcal{J}h_0(x_0, x_0^*)$ we get

$$\langle y_0, x_0^*\rangle + \langle x_0, y_0^*\rangle - \mathcal{J}h_0(x_0, x_0^*) > \mathcal{J}h_0(y_0, y_0^*).$$

Therefore, also using (1.4), we get $\mathcal{J}(\mathcal{J}h_0(y_0, y_0^*)) > \mathcal{J}h_0(y_0, y_0^*)$. Again using the equality $\mathcal{J}^2 h_0 = h_0$ we conclude that

$$h_0(y_0, y_0^*) > \mathcal{J}h_0(y_0, y_0^*). \tag{2.6}$$

Define $\gamma : X \times X^* \to \mathbb{R}$, $g : X \times X^* \to \overline{\mathbb{R}}$,

$$\gamma(x, x^*) \;:=\; \langle x, y_0^*\rangle + \langle y_0, x^*\rangle - \mathcal{J}h_0(y_0, y_0^*), \tag{2.7}$$

$$g \;:=\; \max\ \gamma, \mathcal{J}h_0. \tag{2.8}$$

By (1.4), $h_0 \geq \gamma$. Since $h_0 \in L(h)$, $h_0 \geq \mathcal{J}h_0$. Therefore,

$$h_0 \geq g \geq \mathcal{J}h_0.$$

We claim that $g \in \mathcal{H}(T)$. Indeed, $g$ is a lower semicontinuous convex function. Moreover, since $h_0, \mathcal{J}h_0 \in \mathcal{H}(T)$, it follows from (1.2) and the above inequalities that $g \in \mathcal{H}(T)$. Now apply $\mathcal{J}$ to the above inequality to conclude that

$$h_0 \geq \mathcal{J}g \geq \mathcal{J}h_0.$$

Therefore, defining

$$\hat{g} = \max\ g, \mathcal{J}g, \tag{2.9}$$

we have $h > h_0 \geq \hat{g}$. By Proposition 2.1, $\hat{g} \in \mathcal{H}(T)$ and $\hat{g} \geq \mathcal{J}\hat{g}$. Combining these results with the minimality of $h_0$, it follows that $\hat{g} = h_0$. In particular,

$$\hat{g}(y_0, y_0^*) = h_0(y_0, y_0^*). \tag{2.10}$$

To conclude the proof we shall evaluate $\hat{g}(y_0, y_0^*)$. Using (2.7) we obtain

$$\gamma(y_0, y_0^*) = 2\langle y_0, y_0^*\rangle - \mathcal{J}h_0(y_0, y_0^*).$$

Since $\mathcal{J}h_0 \in \mathcal{H}(T)$, $\mathcal{J}h_0(y_0, y_0^*) \geq \langle y_0, y_0^*\rangle$. Hence, $\gamma(y_0, y_0^*) \leq \langle y_0, y_0^*\rangle$ and by (2.8)

$$g(y_0, y_0^*) = \mathcal{J}h_0(y, y^*). \tag{2.11}$$

Again using the inequality $g \geq \gamma$, we have

$$\mathcal{J}\gamma(y_0, y_0^*) \geq \mathcal{J}g(y_0, y_0^*).$$

Direct calculation yields $\mathcal{J}\gamma(y_0, y_0^*) = \mathcal{J}h_0(y, y^*)$. Therefore

$$\mathcal{J}h_0(y_0, y_0^*) \geq \mathcal{J}g(y_0, y_0^*). \tag{2.12}$$

Combining (2.11), (2.12) and (2.9) we obtain

$$\hat{g}(y_0, y_0^*) = \mathcal{J}h_0(y_0, y_0^*).$$

This equality, together with (2.10), yields $h_0(y_0, y_0^*) = \mathcal{J}h_0(y_0, y_0^*)$, in contradiction with (2.6). Therefore, $h_0(x, x^*) = \mathcal{J}h_0(x, x^*)$ for all $(x, x^*)$. $\square$

Since $\sigma_T \in \mathcal{H}_a(T)$, $L(\sigma_T) \neq \emptyset$ and there exists some $h \in L(\sigma_T)$ such that $\mathcal{J}h = h$. (Indeed $L(\sigma_T) = \mathcal{H}_a(T)$.)

## 3. Application

Let $f : X \rightrightarrows X^*$ be a proper lower semicontinuous convex function. We already know that $\partial f$ is maximal monotone. Define, for $\varepsilon \geq 0$,

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

Note that $\partial_0 f = \partial f$. We also have

$$\partial f(x) \subseteq \partial_\varepsilon f(x), \, \forall x \in X, \varepsilon \geq 0, \tag{3.1}$$

$$0 \leq \varepsilon_1 \leq \varepsilon_2 \Rightarrow \partial_{\varepsilon_1} f(x) \subseteq \partial_{\varepsilon_2} f(x), \, \forall x \in X. \tag{3.2}$$

Property (3.1) tells that $\partial_\varepsilon f$ *enlarges* $\partial f$. Property (3.2) shows that $\partial_\varepsilon f$ is nondecreasing (or increasing) in $\varepsilon$. The operator $\partial_\varepsilon f$ has been introduced in [3], and since that, it has had may theoretical and algorithmic applications [1, 14, 9, 10, 22, 12, 2].

Since $\partial f$ is maximal monotone, the enlarged operator $\partial_\varepsilon f$ loses monotonicity in general. Even though, we have

$$x^* \in \partial_\varepsilon f(x) \Rightarrow \langle x - y, x^* - y^* \rangle \geq -\varepsilon, \, \forall (y, y^*) \in \partial f. \tag{3.3}$$

Now, take

$$\begin{aligned} &x_1^* \in \partial_{\varepsilon_1} f(x_1), x_2^* \in \partial_{\varepsilon_1} f(x_2), \\ &p, q \geq 0, p + q = 1, \end{aligned} \tag{3.4}$$

and define

$$\begin{aligned} &(\bar{x}, \bar{x}^*) := p(x_1, x_1^*) + q(x_2, x_2^*), \\ &\bar{\varepsilon} := p\varepsilon_1 + q\varepsilon_2 + pq\langle x_1 - x_2, x_1^* - x_2^* \rangle. \end{aligned} \tag{3.5}$$

Using the previous definitions, and the convexity of $f$, it is trivial to check that

$$\bar{\varepsilon} \geq 0, \; \bar{x}^* \in \partial_{\bar{\varepsilon}} f(\bar{x}). \tag{3.6}$$

Properties (3.4), (3.5), (3.6) will be called a *transportation formula*. If $\varepsilon_1 = \varepsilon_2 = 0$, then we are using elements in the graph of $\partial f$ to construct elements in the graph of $\partial_\varepsilon f$. In (3.5), the product of elements in $\partial_\varepsilon f$ appears. This product admits the following estimation:

$$x_1^* \in \partial_{\varepsilon_1} f(x_1), x_2^* \in \partial_{\varepsilon_1} f(x_2) \Rightarrow \langle x_1 - x_2, x_1^* - x_2^* \rangle \geq -(\varepsilon_1 + \varepsilon_2). \tag{3.7}$$

Moreover, $\partial_\varepsilon f$ is maximal with respect to property (3.7). We will call property (3.7) *additivity*. The enlargement $\partial_\varepsilon f$ can be characterized by the function $h_{\mathrm{FY}}$, defined in (1.1),

$$x^* \in \partial_\varepsilon f(x) \iff h_{\mathrm{FY}}(x, x^*) \leq \langle x, x^* \rangle + \varepsilon.$$

The transportation formula (3.4), (3.5), (3.6) now follows directly of the convexity of $h_{\mathrm{FY}}$. Additivity follows from the fact that $h_{\mathrm{FY}} \geq \mathcal{J}h_{\mathrm{FY}}$, and maximality of the additivity follows from the fact that

$$h_{\mathrm{FY}} = \mathcal{J}h_{\mathrm{FY}}.$$

Define the *graph* of $\partial_\varepsilon f$ as

$$G(\partial_{(\cdot)} f(\cdot)) := \{(x, x^*, \varepsilon) \mid x^* \in \partial_\varepsilon f(x)\}.$$

Note that $G(\partial_{(\cdot)} f(\cdot))$ is closed. So we say that $\partial_\varepsilon f$ is *closed*.

Given $T : X \rightrightarrows X^*$, maximal monotone, it would be desirable to have an enlargement of $T$, say $T^\varepsilon$, with similar properties to the $\partial_\varepsilon f$ enlargement of $\partial f$. With this aim, such an object was defined in [4, 5] (in finite-dimensional spaces and in Banach spaces, respectively), for $\varepsilon \geq 0$, as

$$T^\varepsilon(x) := \{x^* \in X^* \mid \langle x - y, x^* - y^*\rangle \geq -\varepsilon, \ \forall (y, y^*) \in T\}. \tag{3.8}$$

The $T^\varepsilon$ enlargement of $T$ shares many properties with the $\partial_\varepsilon f$ enlargement of $\partial f$: the transportation formula, Lipschitz continuity (in the interior of its domain), and even the Brøndsted-Rockafellar property (in Reflexive Banach spaces). Since its introduction, it has had both theoretical and algorithmic applications [4, 6, 20, 21, 15, 16]. Even though, $T^\varepsilon$ is *not* the extension of the construct $\partial_\varepsilon f$ to a generic maximal monotone operator. Indeed, taking $T = \partial f$, we obtain

$$\partial_\varepsilon f(x) \subseteq (\partial f)^\varepsilon(x),$$

with examples of strict inclusion even in finite-dimensional cases [4]. Therefore, in general, $T^\varepsilon$ lacks the "additive" property (3.7). The $T^\varepsilon$ enlargement satisfies a weaker property [5]

$$x_1^* \in T^{\varepsilon_1}(x_1), x_2^* \in T^{\varepsilon_2}(x_2) \Rightarrow \langle x_1 - x_2, x_1^* - x_2^*\rangle \geq -(\sqrt{\varepsilon_1} + \sqrt{\varepsilon_2})^2.$$

The enlargement $T^\varepsilon$ is also connected with a convex function. Indeed,

$$\begin{aligned} x^* \in T^\varepsilon(x) \quad &\Longleftrightarrow \quad \langle x - y, x^* - y^*\rangle \geq -\varepsilon, \forall (y, y^*) \in T \\ &\Longleftrightarrow \quad \sup_{(y,y^*)\in T} \langle x - y, y^* - x\rangle \leq \varepsilon. \end{aligned}$$

The Fitzpatrick function, $\varphi_T$, is the smallest element of $\mathcal{H}(T)$ [8], and is defined as

$$\varphi_T(x, x^*) := \sup_{(y,y^*)\in T} \langle x - y, y^* - x\rangle + \langle x, x^*\rangle. \tag{3.9}$$

Therefore,

$$x^* \in T^\varepsilon(x) \iff \varphi_T(x, x^*) \leq \langle x, x^*\rangle + \varepsilon.$$

Now, the transportation formula for $T^\varepsilon$ follows from convexity of $\varphi_T$. In [7] it is proven that each enlargement $\hat{T}^\varepsilon$ of $T$, which has a closed graph, is nondecreasing and satisfies the transportation formula, is characterized by a function $\hat{h} \in \mathcal{H}(T)$, by the formula

$$x^* \in \hat{T}^\varepsilon(x) \iff \hat{h}(x, x^*) \leq \langle x, x^*\rangle + \varepsilon.$$

So, if we want to retain "additivity",

$$x_1^* \in \hat{T}^{\varepsilon_1}(x_1), x_2^* \in \hat{T}^{\varepsilon_2}(x_2) \Rightarrow \langle x_1 - x_2, x_1^* - x_2^*\rangle \geq -(\varepsilon_1 + \varepsilon_2).$$

We shall require $\hat{h} \geq \mathcal{J}\hat{h}$. The enlargements in this family, which are also maximal with respect to the additivity, are structurally closer to the $\partial_\varepsilon f$ enlargement, and are characterized by $\hat{h} \in \mathcal{H}(T)$,

$$\hat{h} = \mathcal{J}\hat{h}.$$

If there were only one element in $\mathcal{H}(T)$ as the fixed point of $\mathcal{J}$, then this element would be the "canonical" representation of $T$ by a convex function, and the associated enlargement would be the extension of the $\varepsilon$-subdifferential enlargement to $T$. Unfortunately, it is not clear whether we have uniqueness of such fixed points.

Existence of an additive enlargement of $T$, maximal with respect to "additivity", was proved in [23]. The convex representation of this enlargement turned out to be minimal in the family $\mathcal{H}_a(T)$, but the characterization of these minimal elements of $\mathcal{H}_a(T)$ as fixed point of $\mathcal{J}$ was lacking.

Since the function $\sigma_T$ has played a fundamental role in our proof, we redescribe it here. Let $\delta_T$ be the indicator function of $T$, i.e., in $T$ its value is 0 and elsewhere in $(X \times X^* \setminus T)$ its value is $+\infty$. Denote the duality product by $\pi : X \times X^* \to \mathbb{R}$, $\pi(x, x^*) = \langle x, x^* \rangle$. Then

$$\sigma_T(x, x^*) = \mathrm{cl} - \mathrm{conv}(\pi + \delta_T),$$

where $\mathrm{cl} - \mathrm{conv} f$ stands for the biggest lower semicontinuous convex function majorized by $f$. We refer the reader to [7] for a detailed analysis of this function.

## Acknowledgement

We thank the anonymous referee for the suggestions which helped to improve this paper.

## References


1. D.P. Bertsekas and S.K. Mitter. A descent numerical method for optimization problems with nondifferentiable cost functionals. *SIAM Journal on Control*, 11(4):637–652, 1973. MR **48:**7592
2. F. Bonnans, J.Ch. Gilbert, C.L. Lemaréchal and C.A. Sagastizábal. *Optimisation Numérique, aspects théoriques et pratiques.* Collection "Mathématiques et applications", SMAI-Springer-Verlag, Berlin, 1997. MR **99f:**90001
3. A. Brøndsted and R. T. Rockafellar. On the subdifferentiability of convex functions. *Proceedings of the American Mathematical Society*, 16:605–611, 1965. MR **31:**2361
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. MR **98h:**49010
5. Regina Sandra Burachik and B. F. Svaiter. $\varepsilon$-enlargements of maximal monotone operators in Banach spaces. *Set-Valued Analysis*, 7(2):117–132, 1999. MR **2000i:**47099
6. Regina S. Burachik, Claudia A. Sagastizábal, and B. F. Svaiter. $\epsilon$-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. MR **2000a:**49030
7. Burachik, R.S. and Svaiter, B.F.: Maximal monotone operators, convex functions and a special family of enlargements, *Set Valued Analysis,* (to appear).
8. Fitzpatrick, S.: Representing monotone operators by convex functions, *Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988)* 59–65, *Proc. Centre Math. Anal. Austral. Nat. Univ.*, 20 *Austral. Nat. Univ., Canberra,* 1988. MR **90i:**47054
9. J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms.* Number 305-306 in Grund. der math. Wiss. Springer-Verlag, 1993. (two volumes). MR **95m:**90001; MR **95m:**90002
10. K.C. Kiwiel. Proximity control in bundle methods for convex nondifferentiable minimization. *Mathematical Programming*, 46:105–122, 1990. MR **91c:**90084
11. Krauss, Eckehard: A representation of maximal monotone operators by saddle functions, *Rev. Roumaine Math. Pures Appl.* **30** (1985), 823–837. MR **87e:**47074
12. C. Lemaréchal, A. Nemirovskii, and Yu. Nesterov. New variants of bundle methods. *Mathematical Programming*, 69:111–148, 1995. MR **96g:**90075
13. Martinez-Legaz, J.-E. and Théra, M.: A convex representation of maximal monotone operators, *Journal of Nonlinear and Convex Analysis* **2** (2001), 243–247. MR **2002e:**49035

14. E.A. Nurminski. $\varepsilon$-subgradient mapping and the problem of convex optimization. *Cybernetics*, 21(6):796–800, 1986.
15. Revalski, J.P. and M. Théra, Generalized sums of monotone operators, *C. R. Acad. Sci. Paris Sér. I Math.* 329(11):979–984, 1999. MR **2000m:**47068
16. Revalski, J.P. and M. Théra, Variational and extended sums of monotone operators, *Ill-posed variational problems and regularization techniques (Trier, 1998)*, 229–246, Lecture Notes in Econom. and Math. Systems, 477, Springer, Berlin, 1999. MR **2000k:**47071
17. Rockafellar, R. T.: On the maximal monotonicity of subdifferential mappings, *Pacific Journal of Mathematics* **33** (1970), 209–216. MR **41:**7432
18. R. Tyrrell Rockafellar, Roger J-B. Wets. *Variational Analysis*. Springer Verlag, Berlin Heidelberg, 1998. MR **98m:**49001
19. Simons, S.: Minimax and monotonicity. Lecture Notes in Mathematics, 1693. Springer-Verlag, Berlin, 1998. MR **2001h:**49002
20. 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. MR **2001a:**90084
21. M. V. Solodov and B. F. Svaiter. Error bounds for proximal point subproblems and associated inexact proximal point algorithms. *Mathematical Programming*, 88(2):371–389, 2000. MR **2001f:**90052
22. H. Schramm and J. Zowe. A version of the bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results. *SIAM Journal on Optimization*, 2(1):121–152, 1992. MR **93b:**90089
23. B. F. Svaiter. A Family of Enlargements of Maximal Monotone Operators. *Set-Valued Analysis*, 8(4):311–328, December 2000. MR **2001i:**49037

IMPA Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Rio de Janeiro–RJ, CEP 22460-320 Brazil

*E-mail address*: benar@impa.br