This is the submitted version of the article published by Springer Verlag: Burachik, R.S., Martínez-Legaz, J.E., Rezaie, M. et al: An additive subfamily of enlargements of a maximally monotone operator, published in Set-Valued and Variational Analysis, December 2015, Volume 23, Issue 4, pp 643 -665. The final version is available at: https://link.springer.com/article/10.1007/s11228-015-0340-9

An additive subfamily of enlargements of a maximally monotone operator

Regina S. Burachik · Juan Enrique Martínez-Legaz · Mahboubeh Rezaie · Michel Théra

Dedicated to professor L. Thibault

Received: date / Accepted: date

Abstract We introduce a subfamily of additive enlargements of a maximally monotone operator. Our definition is inspired by the early work of Simon Fitzpatrick. These enlargements constitute a subfamily of the family of enlargements introduced by Svaiter. When the operator under consideration is the subdifferential of a convex lower semicontinuous proper function, we prove that some members of the subfamily are *smaller* than the classical ε-subdifferential enlargement widely used in convex analysis. We also recover the epsilon-subdifferential within the subfamily. Since they are all additive, the enlargements in our subfamily can be seen as structurally closer to the ε-subdifferential enlargement.

Keywords Maximally monotone operator · ε-subdifferential mapping · subdifferential operator · convex lower semicontinuous function · Fitzpatrick function ·

The research of Juan-Enrique Martínez-Legaz was supported by the MINECO of Spain, Grant MTM2011-29064-C03-01, and the Australian Research Council, project DP140103213. He is affiliated to MOVE (Markets, Organizations and Votes in Economics).

The research of Michel Théra was partially supported by the MINECO of Spain, Grant MTM2011-29064-C03-03, and the Australian Research Council, project DP110102011.

First author
School of Information Technology and Mathematical Sciences, University of South Australia, Mawson Lakes, SA 5095, Australia
E-mail: regina.burachik@unisa.edu.au

Second author
Departament d'Economia i d'Historia Economica, Universitat Autonoma de Barcelona, Spain
E-mail: Juan-Enrique.Martinez-Legaz@uab.cat

Third author
University of Isfahan, Iran
E-mail: mrezaie@sci.ui.ac.ir

Fourth author
Université de Limoges, France and Centre for Informatics and Applied Optimisation, Federation University Australia
E-mail: michel.thera@unilim.fr

enlargement of an operator · Brøndsted- Rockafellar enlargements · additive enlargements · Brøndsted- Rockafellar property · Fenchel-Young function.

Mathematics Subject Classification (2000) 49J52 · 48N15 · 90C25 · 90C30 · 90C46

1 Introduction

Let X be a real Banach space with continuous dual X^*. By a generalized equation governed by a maximally monotone operator $T : X \rightrightarrows X^*$, we mean the problem of finding $x \in X$ such that

$$0 \in T(x). \tag{1}$$

This model has been extensively used as a mathematical formulation of fundamental problems in optimization and fixed point theory. Main illustrations follow.

- If X is a Hilbert space, I is the identity map, $F : X \longrightarrow X$ is a nonexpansive mapping, and $T = I - F$, then solving (1) is equivalent to finding a fixed point of F.
- If T is a maximally monotone operator from a Hilbert space into itself, then the set of solutions of (1) is the set of fixed points of the so-called resolvent map $R := (I + \lambda T)^{-1}$, with $\lambda > 0$, or the set of fixed points of the Cayley operator $C := 2R - I$.
- As observed by Rockafellar [30, Theorem 37.4], when $L : X \times X \to \mathbb{R}$ is a concave-convex function (for instance the Lagrangian of a convex program), finding a saddle point of L is equivalent to solving $(0, 0) \in \partial L(x, y)$, where $\partial L(x, y) = \partial_x(-L)(x, y) \times \partial_y L(x, y)$, and ∂_x and ∂_y are the convex subdifferentials operators with respect to the first and the second variable, respectively.
- If f is a lower semicontinuous proper convex function and $T = \partial f$, the subdifferential of f, then the set of solutions of (1) is the set of minimizers of f.

Solving inclusion (1) is tantamount to finding a point of the form $(x, 0)$ in the graph of T. If T is not point-to-point, then it lacks semicontinuity properties. Namely, if Tx is not a singleton, then T cannot be inner-semicontinuous at x (see [6, Theorem 4.6.3]). This fact makes the problem ill-behaved, making the required computations hard. Enlargements of T are point-to-set mappings (the terms set-valued mapping and multifunction are also used) which have a graph larger than the graph of T. These mappings, however, have better continuity properties than T itself. Moreover, they stay "close" to T, so they allow to define perturbations of problem (1), without losing information on T. In this way, we can define well-behaved approximations of problem (1), which (i) are numerically more robust, and (ii) whose solutions approximate accurately the solutions of (1). The use of enlargements in the study of problem (1) has been a fruitful approach, from both practical and theoretical reasons. A typical example of the usefulness of enlargements in the analysis of (1) arises when considering a convex optimization problem, i.e., the case in which $T = \partial f$, where $f : X \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex, and lower semicontinuous function. It is a well-known fact that $T = \partial f$ is maximally monotone. This has been proved by

Moreau for Hilbert spaces [24] and by Rockafellar [31] for Banach spaces. The ϵ-*subdifferential of* f, introduced by Brøndsted and Rockafellar in [5] (see Definition 4), is an enlargement of $T = \partial f$ which had a crucial role in the development of algorithms for solving (1), as well as in allowing a better understanding of the properties of the mapping ∂f itself (see, e.g., [31]). This is why the ϵ-subdifferential has been intensively studied since its introduction in 1965, not only from an abstract point of view, but also for constructing specific numerical methods for convex nonsmooth optimization (see, e.g., [1,17,18,20,32]). Using the optimization problem as a benchmark, but having the general problem (1) in mind, it is relevant to study enlargements of an arbitrary maximally monotone operator T. To be useful, the enlargements of T must share with the ϵ-subdifferential most of its good properties. By good properties we mean local boundedness, demi-closedness of the graph, Lipschitz continuity, and Brøndsted-Rockafellar property. Indeed, given an arbitrary maximally monotone operator T defined on a reflexive Banach space, Svaiter introduced in [35] a family of enlargements, denoted by $\mathbb{E}(T)$, which share with the ϵ-subdifferential all these good properties. There are, however, properties of the ϵ-subdifferential which are not shared by every element of $\mathbb{E}(T)$. To make this statement precise, we recall the largest member of the family $\mathbb{E}(T)$, denoted by T^{BE}. The enlargement $T^{\mathrm{BE}} : \mathbb{R}_+ \times X \rightrightarrows X^*$ has been the intense focus of research (see, e.g. [6–9,11,22,28,35]), and is defined as follows. We say that

$$x^* \in T^{\mathrm{BE}}(\epsilon, x) \iff \forall (y, y^*) \in \mathrm{gph}\ T \text{ we have } \langle y - x, y^* - x^* \rangle \geq -\epsilon. \quad (2)$$

The discrepancy between some elements of $\mathbb{E}(T)$ and the ϵ-subdifferential arises from the fact that, when $T = \partial f$, the biggest enlargement T^{BE} is larger than the ϵ-subdifferential. Namely, $\partial_\epsilon f(\cdot) \subset (\partial f)^{BE}(\epsilon, \cdot)$, and the inclusion can be strict, as noticed by Martínez-Legaz and Théra, see [22]. Hence, it is natural to expect that some properties of the ϵ-subdifferential *will not* be shared by every element of $\mathbb{E}(T)$, and in particular, they will not be shared by T^{BE}. Such a property is *additivity*. In the context of enlargements of arbitrary maximally monotone operators, this property was introduced in [9] and further studied in [35,36]. It is stated as follows. An enlargement $E : \mathbb{R}_+ \times X \rightrightarrows X^*$ is *additive* if for every $x_1^* \in E(\epsilon_1, x_1)$ and every $x_2^* \in E(\epsilon_2, x_2)$, it holds that

$$\langle x_1 - x_2, x_1^* - x_2^* \rangle \geq -(\epsilon_1 + \epsilon_2).$$

The ϵ-subdifferential is additive. Moreover, it is *maximal* among all those enlargements of ∂f with this property. In other words, if another enlargement of ∂f is additive and contains the graph of the ϵ-subdifferential, then it must coincide with the ϵ-subdifferential enlargement. We describe the latter property as being *maximally additive* (or *max-add*, for short). Namely, an enlargement E is max-add when it is additive and, if the graph of another additive enlargement E' contains the graph of E, then we must have $E = E'$. Since the ϵ-subdifferential is max-add, the members of the family $\mathbb{E}(T)$ that are max-add do share an extra property with the ϵ-subdifferential, and in this sense, they can be seen as structurally "closer" to the ϵ-subdifferential. As hinted above, not all enlargements $E \in \mathbb{E}(T)$ are additive. However, it was proved in [35] that the smallest enlargement, denoted by T^{SE}, is additive. The existence of a max-add element in $\mathbb{E}(T)$ is then obtained in [35] as a consequence of Zorn's lemma. In the present paper, we define a whole family of additive elements of $\mathbb{E}(T)$, denoted by $\mathbb{E}_{\mathcal{H}}(T)$. The family $\mathbb{E}_{\mathcal{H}}(T)$ has max-add elements, and the existence of these elements is deduced through a constructive proof. For the case in which $T = \partial f$, we show that some specific elements of $\mathbb{E}_{\mathcal{H}}(T)$ are contained in the ϵ-subdifferential enlargement. Additionally, a specific element of our family coincides with the ϵ-subdifferential when $T = \partial f$.

The layout of the paper is as follows. First, we define our family of enlargements of a maximally monotone operator T. Our definition is inspired by the early work of Fitzpatrick presented in [13], but can as well be seen as a subfamily of $E(T)$. Second, we prove that all members of our subfamily are additive. We also introduce a new definition related to additivity, which helps us in the proofs. We deduce, in a constructive way, the existence of max-add elements in $\mathbb{E}(T)$. Finally, we consider the case $T = \partial f$. For this case we prove that some members of the subfamily are *smaller* than the ϵ-subdifferential enlargement, and we recover the ϵ-subdifferential as a member of our subfamily.

2 Basic Definitions

Throughout this paper, we assume that X is a real **reflexive** Banach space with continuous dual X^*, and pairing between them denoted by $\langle \cdot, \cdot \rangle$. We will use the same symbol $\| \cdot \|$ for the norms in X and X^*, and w will stand for the weak topologies on X and X^*. We consider the Cartesian product $X \times X^*$ equipped with the product topology determined by the norm topology in X and the weak topology in X^*. In this case the dual of $X \times X^*$ can be identified with $X^* \times X$ and hence, the dual product is defined as $\langle (x, x^*), (y^*, y) \rangle = \langle x, y^* \rangle + \langle y, x^* \rangle$.

For a given (in general, multivalued) operator $T : X \rightrightarrows X^*$, its graph is denoted by

$$\mathrm{gph}\,(T) := \{(x, x^*) \in X \times X^* : x^* \in T(x)\}.$$

Recall that $T : X \rightrightarrows X^*$ is said to be *monotone* if and only if

$$\langle y - x, y^* - x^* \rangle \geq 0 \qquad \forall (x, x^*), (y, y^*) \in \mathrm{gph}\,(T).$$

A monotone operator T is called *maximally monotone* if and only if the condition $\langle y - x, y^* - x^* \rangle \geq 0$ for every $(y, y^*) \in \mathrm{gph}\,(T)$, implies $(x, x^*) \in \mathrm{gph}\,(T)$. Equivalently, it amounts to saying that T has no monotone extension (in the sense of graph inclusion).

In what follows, $f : X \to \mathbb{R} \cup \{+\infty\}$ will be a convex function. Recall that f is *proper* if the set $\mathrm{dom}\,(f) := \{x \in X : f(x) < +\infty\}$ is nonempty. The *subdifferential* of f is the multivalued mapping $\partial f : X \rightrightarrows X^*$ defined by

$$\partial f(x) := \{x^* \in X^* : f(y) - f(x) \geq \langle y - x, x^* \rangle, \, \forall y \in X\}, \tag{3}$$

if $x \in \mathrm{dom}\,(f)$, and $\partial f(x) := \emptyset$, otherwise. Given $\epsilon \geq 0$, the *ϵ-subdifferential of f* is the multivalued mapping $\partial_\epsilon f : X \rightrightarrows X^*$ defined by

$$\partial_\epsilon f(x) := \{x^* \in X^* : f(y) - f(x) \geq \langle y - x, x^* \rangle - \epsilon, \, \forall y \in X\}, \tag{4}$$

if $x \in \mathrm{dom}\,(f)$, and $\partial_\epsilon f(x) := \emptyset$, otherwise. The case $\epsilon = 0$ gives the subdifferential of f at x. The set $\partial_\epsilon f(x)$ is nonempty for every $\epsilon > 0$ if and only if f is lower semicontinuous at x. Note that the ϵ-subdifferential can be viewed as an approximation of the subdifferential. Indeed, in [22, Theorem 1] a formula expressing, for a lower semicontinuous convex extended-real-valued function, its $\epsilon-$subdifferential in terms of its subdifferential was established.

As we will see later in Subsection 2.1, enlargements are multifunctions defined on $\mathbb{R}_+ \times X$. Consequently, we need a different notation for the epsilon-subdifferential (4). This enlargement will be denoted as follows:

$$\breve{\partial} f(\epsilon, x) := \partial_\epsilon f(x).$$

We call the enlargement $\breve{\partial} f$ the *Brøndsted-Rockafellar enlargement of* ∂f. The Fenchel-Moreau conjugate of f is denoted by $f^* : X^* \to \mathbb{R} \cup \{+\infty\}$ and is defined by

$$f^*(x^*) := \sup\{\langle x, x^*\rangle - f(x) : x \in X\}. \tag{5}$$

Observe that f^* is lower semicontinuous with respect to the weak topology on X^*. In what follows, we shall denote by f^{FY} the Fenchel-Young function associated to f:

$$f^{FY}(x, x^*) := f(x) + f^*(x^*) \text{ for all } (x, x^*) \in X \times X^*.$$

Then f^{FY} is a convex, proper and $(\|\cdot\| \times w)$-lower semicontinuous function on $X \times X^*$ and it is well known that f^{FY} completely characterizes the graph of the subdifferential of f:

$$\partial f(x) = \{x^* \in X^* : f^{FY}(x, x^*) = \langle x, x^*\rangle\}. \tag{6}$$

Moreover, f^{FY} also completely characterizes the graph of the Brøndsted-Rockafellar enlargement of ∂f. Namely,

$$x^* \in \breve{\partial} f(\epsilon, x) \text{ if and only if } f^{FY}(x, x^*) \le \langle x, x^*\rangle + \epsilon. \tag{7}$$

If Z is a general Banach space and $f, g : Z \to \mathbb{R} \cup \{+\infty\}$, the *infimal convolution* of f with g is denoted by $f \oplus g$ and defined by

$$(f \oplus g)(z) := \inf_{z_1 + z_2 = z}\{f(z_1) + g(z_2)\}.$$

If $q : Z \to \mathbb{R} \cup \{+\infty\}$, the *closure* of q is denoted by $\mathrm{cl}\,(q)$ and defined by:

$$\mathrm{epi}(\mathrm{cl}\,(q)) = \mathrm{cl}\,(\mathrm{epi}(q)).$$

We will use the following well-known property:

$$(f + g)^* = \mathrm{cl}\,(f^* \oplus g^*) \le (f^* \oplus g^*). \tag{8}$$

2.1 The family $\mathbb{E}(T)$

We mentioned above two examples of enlargements, the enlargement $\breve{\partial} f$ of $T = \partial f$, and the enlargement T^{BE} of an arbitrary maximally monotone operator. Each of these is a member of a family of enlargements of ∂f and T, respectively. For a maximally monotone operator T, denote by $\mathbb{E}(T)$ the following family of enlargements defined in [35] and [11].

Definition 2.1 *Let $T : X \rightrightarrows X^*$. We say that a point-to-set mapping $E : \mathbb{R}_+ \times X \rightrightarrows X^*$ belongs to the family $\mathbb{E}(T)$ when*

(E_1) *$T(x) \subset E(\epsilon, x)$ for all $\epsilon \geq 0, x \in X$;*
(E_2) *If $0 \leq \epsilon_1 \leq \epsilon_2$, then $E(\epsilon_1, x) \subset E(\epsilon_2, x)$ for all $x \in X$;*
(E_3) *The transportation formula holds for E. More precisely, let $x_1^* \in E(\epsilon_1, x_1)$, $x_2^* \in E(\epsilon_2, x_2)$, and $\alpha \in [0, 1]$. Define*

$$\hat{x} := \alpha x_1 + (1 - \alpha) x_2,$$
$$\hat{x}^* := \alpha x_1^* + (1 - \alpha) x_2^*,$$

$$\begin{aligned} \epsilon &:= \alpha \epsilon_1 + (1 - \alpha)\epsilon_2 + \alpha \langle x_1 - \hat{x}, x_1^* - \hat{x}^* \rangle + (1 - \alpha) \langle x_2 - \hat{x}, x_2^* - \hat{x}^* \rangle \\ &= \alpha \epsilon_1 + (1 - \alpha)\epsilon_2 + \alpha(1 - \alpha) \langle x_1 - x_2, x_1^* - x_2^* \rangle. \end{aligned}$$

Then $\epsilon \geq 0$ and $\hat{x}^* \in E(\epsilon, \hat{x})$.

The following lemma, which is well-known but hard to track down, states that the transportation formula holds for the Brøndsted-Rockafellar enlargement. We include its simple proof here for convenience of the reader.

Lemma 2.1 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be convex. Then the transportation formula holds for $\breve{\partial} f$.*

Proof Assume that $x_1^* \in \breve{\partial} f(\epsilon_1, x_1)$, $x_2^* \in \breve{\partial} f(\epsilon_2, x_2)$ and $\alpha \in [0, 1]$, and let $\hat{x}$, $\hat{x}^*$ and ϵ be as in condition (E_3) of Definition 2.1. Let us first show that $\epsilon \geq 0$. By assumption, we have

$$\begin{aligned} f(x_2) - f(x_1) &\geq \langle x_2 - x_1, x_1^* \rangle - \epsilon_1, \\ f(x_1) - f(x_2) &\geq \langle x_1 - x_2, x_2^* \rangle - \epsilon_2. \end{aligned}$$

Summing up these inequalities and re-arranging the resulting expression gives

$$\langle x_1 - x_2, x_1^* - x_2^* \rangle \geq -\epsilon_1 - \epsilon_2.$$

We can now write

$$\alpha(1 - \alpha)\langle x_1 - x_2, x_1^* - x_2^* \rangle \geq \alpha(1 - \alpha)(-\epsilon_1 - \epsilon_2) \geq -\alpha\epsilon_1 - (1 - \alpha)\epsilon_2.$$

Using the definition of ϵ in (E_3) , we deduce that $\epsilon \geq 0$. In order to finish the proof, we use the assumption on x_1^*, x_2^* to write

$$\begin{aligned} \alpha \left(f(z) - f(x_1) \right) &\geq \alpha \left(\langle z - x_1, x_1^* \rangle - \epsilon_1 \right) \\ (1 - \alpha) \left(f(z) - f(x_2) \right) &\geq (1 - \alpha) \left(\langle z - x_2, x_2^* \rangle - \epsilon_2 \right). \end{aligned}$$

Summing up these inequalities, and using the convexity of f, we obtain, after some simple algebra,

$$f(z) - f(\hat{x}) \geq \langle z - \hat{x}, \hat{x^*} \rangle - \epsilon,$$

and hence $\hat{x^*} \in \breve{\partial} f(\epsilon, \tilde{x})$, as wanted. ■

Remark 2.1 *If* $\mathrm{gph}\,(T)$ *is nonempty, the family* $\mathbb{E}(T)$ *is nonempty and its biggest enlargement is* T^{BE}*, defined in* (2)*. Using this fact, one can easily prove that for every* $E \in \mathbb{E}(T)$ *and* $x \in X$*, one has* $E(0, x) = T(x)$*. Moreover, from Lemma 2.1 and the definitions, it follows that the enlargement* $\breve{\partial} f \in \mathbb{E}(\partial f)$ *(see also [11]).*

2.2 Convex representations of T

As a consequence of (6) and (7), the function f^{FY} is an example of a convex function that completely characterizes the graph of the operator ∂f, as well as the graph of the Brøndsted-Rockafellar enlargement. For an arbitrary maximally monotone operator T, Fitzpatrick defined in [13, Definition 3.1] an ingenious proper convex $(\|\cdot\| \times w)$-lower semicontinuous function, here denoted by $\mathcal{F}_T$, that has the same properties:

$$\mathcal{F}_T(x, x^*) := \sup\{\langle y, x^* \rangle + \langle x - y, y^* \rangle : (y, y^*) \in \mathrm{gph}\,(T)\}. \tag{9}$$

It satisfies:

- $\mathcal{F}_T(x, x^*) \geq \langle x, x^* \rangle$.
- In analogy to (6), we have that (see [13]):

$$\mathrm{gph}\,(T) := \{(x, x^*) \in X \times X^* : \ \mathcal{F}_T(x, x^*) = \mathcal{F}_T^*(x^*, x) = \langle x, x^* \rangle\}.$$

- In analogy to (7), we have that (see [11]):

$$x^* \in T^{\mathrm{BE}}(\epsilon, x) \text{ if and only if } \mathcal{F}_T(x, x^*) \leq \langle x, x^* \rangle + \epsilon.$$

  Therefore $\mathcal{F}_T$ completely characterizes the graph of the operator T, as well as the graph of its enlargement T^{BE}. When $T = \partial f$, we can relate $\mathcal{F}_T$ and f^{FY} as follows.

$$\forall (x, x^*) \in X \times X^*, \quad \langle x, x^* \rangle \leq \mathcal{F}_{\partial f}(x, x^*) \leq f(x) + f^*(x^*) = f^{FY}(x, x^*).$$

Remark 2.2 *Note that the Fitzpatrick function associated to a subdifferential operator could be different from the Fenchel-Young function. Indeed, if* X *is a Hilbert space and* $f : X \longrightarrow \mathbb{R}$ *is given by* $f(x) := \frac{1}{2}\|x\|^2$*, then* $f^{FY}(x, y) := \frac{1}{2}\left(\|x\|^2 + \|y\|^2\right)$ *and* $\mathcal{F}_{\partial f}(x, y) = \frac{1}{4}\|x + y\|^2$.

The Fitzpatrick function was unnoticed for several years until it was rediscovered by Martínez-Legaz and Théra [23]. However, we recently discovered, by reading a paper by Flåm [14], that this function had already been used by Krylov [19] before Fitzpatrick. According to the fact that it bridges monotone operators to convex functions, it has been the subject of an intense research with applications in different areas such as the variational representation of (nonlinear) evolutionary PDEs, and the development of variational techniques for the analysis of their structural stability; see e.g., [29, 16, 37, 38]; more surprisingly, Flåm [14] gave an economic interpretation of the Fitzpatrick function.

Moreover, in [13] Fitzpatrick also defined a family of convex functions associated to T. We recall this definition next.

Definition 2.2 *Let $T : X \rightrightarrows X^*$ be maximally monotone. Define $\mathcal{H}(T)$ as the family of lower semicontinuous convex functions $h : X \times X^* \longrightarrow \mathbb{R} \cup \{+\infty\}$ such that*

$$h(x,x^*) \geq \langle x^*, x\rangle, \forall x \in X, x^* \in X^*, \tag{10}$$

$$x^* \in T(x) \Rightarrow h(x,x^*) = \langle x^*, x\rangle. \tag{11}$$

The family $\mathcal{H}(T)$ was studied in [13] in connection with the operator T itself. It was proved in [13] that the smallest element of this family is precisely $\mathcal{F}_T$.

Clearly, relations (10) and (11) imply that one can express a monotone relation as a minimization problem: setting $\Theta(x,x^*) := \mathcal{F}_T(x,x^*) - \langle x, x^*\rangle$, we have that

$$x^* \in T(x) \iff \Theta(x,x^*) = \inf_{(y,y^*)\in X\times X^*} \Theta(y,y^*) = 0.$$

Moreover, it can be observed that, for a prescribed x^* in the range ot T, i.e. a point $x^* \in T(x)$ for some x, one can solve the inclusion $x^* \in T(x)$ just by minimizing the functional $\Theta(\cdot, x^*)$.

In the paper [39], Visintin presents an interesting application of Fitzpatrick functions to the Calculus of Variations. As pointed out above, one can express a monotone relation as a minimization problem in which the minimum value is prescribed as zero. In [39] it is shown that, by generalizing the Fitzpatrick approach, one can express a monotone relation as a minimization problem, without the need of prescribing the minimum value as zero. This is convenient in many practical problems in which the minimum value is not known, including problems from the Calculus of Variations.

Given a maximally monotone operator, [11, Corollary 3.7] shows that the converse of (11) also holds. Namely, for all $h \in \mathcal{H}(T)$ one has

$$h(x,x^*) = \langle x, x^*\rangle \iff (x,x^*) \in \mathrm{gph}\,(T).$$

The use of Fitzpatrick functions has led to considerable simplifications in the proofs of some classical properties of maximally monotone operators; see, for instance, the work by Burachik and Svaiter [11], Simons and Zălinescu [34], Penot and Zălinescu [27], Boţ et al. [4], Simons [33], and Marques Alves and Svaiter [21]. It was proved by Burachik and Svaiter [11] that the family $\mathcal{H}(T)$ is in a one-to-one relationship with the family $\mathbb{E}(T)$ of enlargements of T, introduced and studied by Svaiter in [35]. More connections between $\mathbb{E}(T)$ and $\mathcal{H}(T)$ were studied in [9, 11, 10]. The correspondence from T to $\mathcal{H}(T)$ associates to a given maximally monotone operator, functions defined in $X \times X^*$. In the paper [13], Fitzpatrick also defined a correspondence which goes in the opposite direction. Namely, given a proper convex function $h : X \times X^* \to \mathbb{R} \cup \{+\infty\}$, Fitzpatrick defined the operator $T_h : X \rightrightarrows X^*$, given by

$$T_h(x) := \{x^* : (x^*, x) \in \partial h(x,x^*)\} \tag{12}$$

Remark 2.3 *Let $f : X \times X^* \longrightarrow \mathbb{R} \cup \{+\infty\}$ be convex and lower semicontinuous, and consider again $f^{FY}(x, x^*) = f(x) + f^*(x^*)$. Then Example 2.3 in [13] proves that $T_{f^{FY}} = \partial f$. We extend this result in Theorem 4.1(ii). Namely, we will extend this equality between two maximally monotone operators to an equality between two enlargements of $T = \partial f$.*

The following theorem summarizes those results in [13] which will be relevant to our study.

Theorem 2.1 *[13] Let $T : X \rightrightarrows X^*$ be monotone and $f : X \times X^* \longrightarrow \mathbb{R} \cup \{+\infty\}$ be convex. Let $T_{\mathcal{F}_T}$ be defined as in* (12) *for $h := \mathcal{F}_T$. The following facts hold.*

(a) For any $x \in X$ one has $T(x) \subset T_{\mathcal{F}_T}(x)$. If T is maximally monotone then $T = T_{\mathcal{F}_T}$;
(b) If T is maximally monotone, then $\mathcal{F}_T \in \mathcal{H}(T)$. Moreover, $\mathcal{F}_T$ is the smallest convex function in $\mathcal{H}(T)$;
(c) The operator T_f is monotone.

We end this subsection by extending Theorem 2.1(a) to every $h \in \mathcal{H}(T)$; the result is an easy consequence of [13, Theorem 2.4 and Proposition 2.2].

Proposition 2.1 *Let T be maximally monotone, and fix $h \in \mathcal{H}(T)$. Then $T = T_h$.*

Proof Since $h \in \mathcal{H}(T)$, we have $\text{gph } T \subset \{(x, x^*) \in X \times X^* : h(x, x^*) = \langle x^*, x\rangle\}$; hence, by [13, Theorem 2.4], the inclusion $\text{gph } T \subset \text{gph } T_h$ holds. On the other hand, by [13, Proposition 2.2], T_h is monotone. Using this fact together with the preceding inclusion and the maximal monotonicity of T, we get $T = T_h$. ■

2.3 Autoconjugate convex representations of T

Every element $h \in \mathcal{H}(T)$ is defined on $X \times X^*$, while h^* is defined on $X^* \times X$. Recall that the dual of $X \times X^*$ can be identified with $X^* \times X$ through the product

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

In order to work with functions defined on $X \times X^*$, we will use the permutation function $i : X \times X^* \longrightarrow X^* \times X$ defined by $i\,(x, x^*) := (x^*, x)$. The composition $h^* \circ i$ will thus be defined on $X \times X^*$. Notice that the mapping $h \longmapsto h^* \circ i$ is precisely the operator $\mathcal{J} : \mathcal{H}(T) \longrightarrow \mathcal{H}(T)$ defined in [11], which, as shown in [11, Remark 5.4], is an involution.

Definition 2.3 *Let $T : X \rightrightarrows X^*$ be maximally monotone. Every $h \in \mathcal{H}(T)$ is called a* convex representation of T. *When $h \in \mathcal{H}(T)$ satisfies*

$$h^* \circ i = h,$$

we say that h is an autoconjugate convex representation of T.

The function f^{FY}, which characterizes the epsilon-subdifferential enlargement of ∂f, is an autoconjugate convex representation of ∂f, as can be easily checked. Hence, it is natural to look for autoconjugate convex representations when searching for an enlargement structurally closer to the epsilon-subdifferential. This observation generates a great interest in constructing autoconjugate convex representations of an arbitrary operator T. Outside the subdifferential case, the operator $T := \partial f + S$, where S is a skew-adjoint linear operator ($S^* = -S$), admits the autoconjugate convex representation given by $f(x) + f^*(-S(x) + x^*)$ (see for instance Example 2.6 in [3], and Ghoussoub [15]). The interest of having autoconjugate convex representations is also given by the next theorem:

Theorem 2.2 *An operator $T : X \rightrightarrows X^*$ is maximally monotone if and only if it admits an autoconjugate convex representation.*

Proof Svaiter proved in [36, Proposition 2.2 and Theorem 2.4] that for every maximally monotone operator T, there exists $h \in \mathcal{H}(T)$ such that h is auto conjugate. See also Bauschke and Wang [2, Theorem 5.7]. The converse follows from a result by Burachik and Svaiter [12, Theorem 3.1]. ■

Remark 2.4 *The "only if" part of Theorem 2.2 proved in [36, Proposition 2.2 and Theorem 2.4] is valid in any real Banach space. The "if" part proved in [12, Theorem 3.1] assumes X is reflexive.*

Remark 2.5 *The papers [25, 26, 36] present non-constructive examples of autoconjugate convex representations of T. Constructive examples of autoconjugate convex representations of T can be found in [3, 2, 27]. The one found in [27] requires a mild constraint qualification, namely, that the affine hull of the domain of T is closed. The other ones do not require any constraint qualification. We will show later other constructive examples of autoconjugate representations of T.*

We introduce now another map, defined on the set

$$\mathcal{H} := \bigcup \{\mathcal{H}(T) : T : X \rightrightarrows X^* \text{ is maximally monotone}\},$$

which will have an important role in the definition of our enlargements and in obtaining autoconjugate convex representations of T.

Remark 2.6 *For every lower semicontinuous proper convex function $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$, one has $f^{FY} \in \mathcal{H}(\partial f) \subset \mathcal{H}$. Furthermore, it is easy to see that f^{FY} is an autoconjugate convex representation of ∂f.*

Definition 2.4 *The map $\mathcal{A}$: $\mathcal{H} \to \mathcal{H}$ is defined by*

$$\mathcal{A}h := \frac{1}{2}(h + h^* \circ i). \tag{13}$$

Remark 2.7 *It follows from [11, Theorem 5.1 and Proposition 5.3] that $\mathcal{A}h$ and $h^* \circ i$ belong to $\mathcal{H}(T)$, for every $h \in \mathcal{H}(T)$; therefore, the map $\mathcal{A}$ is well defined.*

For the next theorem we need to define the following sets:

$$\mathcal{H}_{*\leq} := \{h \in \mathcal{H} : h^* \circ i \leq h\}, \quad \mathcal{H}_{*=} := \{h \in \mathcal{H} : h^* \circ i = h\},$$
$$\mathcal{H}_{*\geq} := \{h \in \mathcal{H} : h^* \circ i \geq h\}. \tag{14}$$

Theorem 2.3 *Consider the operator $\mathcal{A}$ given in Definition 2.4, and the sets defined in* (14). *The following statements hold.*

(i) The operator $\mathcal{A}$ maps $\mathcal{H}$ into $\mathcal{H}_{*\leq}$. The operator $h \longmapsto (\mathcal{A}h)^* \circ i$ maps $\mathcal{H}$ into $\mathcal{H}_{*\geq}$;
(ii) The set of fixed points of $\mathcal{A}$ is $\{h \in \mathcal{H}_{*\leq} : h^* \circ i = h \text{ on } \operatorname{dom}(h)\}$;
(iii) Let $h \in \mathcal{H}$. For every $n \geq 1$, one has $\operatorname{dom}(\mathcal{A}^n h) = \operatorname{dom}(h) \cap \operatorname{dom}(h^* \circ i)$;
(iv) Let $h \in \mathcal{H}$. The sequences $\{\mathcal{A}^n h\}_{n\geq 1} \subset \mathcal{H}_{*\leq}$ and $\{(\mathcal{A}^n h)^* \circ i\}_{n\geq 1} \subset \mathcal{H}_{*\geq}$ are pointwise non-increasing and non-decreasing, respectively. The pointwise limit $\mathcal{A}^\infty h$ of the first one satisfies $\operatorname{dom}(\mathcal{A}^\infty h) = \operatorname{dom}(h) \cap \operatorname{dom}(h^* \circ i)$. *If $\mathcal{A}^\infty h$ is lower semicontinuous, it is a fixed point of $\mathcal{A}$*;
(v) Let $h \in \mathcal{H}$. For every $n \geq 1$, one has

$$(\mathcal{A}^n h)^* \circ i \leq \mathcal{A}^\infty h \leq \mathcal{A}^n h; \tag{15}$$

(vi) Let $h \in \mathcal{H}$. The sequence $\{(\mathcal{A}^n h)^* \circ i\}_{n\geq 1}$ converges pointwise to $\mathcal{A}^\infty h$ on $\operatorname{dom}(h) \cap \operatorname{dom}(h^* \circ i)$;
(vii) Let $T : X \rightrightarrows X^*$ be maximally monotone. If $h \in \mathcal{H}(T)$ and $\mathcal{A}^\infty h$ is lower semicontinuous, then $\mathcal{A}^\infty h \in \mathcal{H}(T)$.

Proof (i) We need to show that $(\mathcal{A}h)^* \circ i \leq \mathcal{A}h$. Indeed, using the properties of the conjugation operator, we can write

$$\begin{aligned}
\left((\mathcal{A}h)^* \circ i\right)(x, x^*) &= (\mathcal{A}h)^*(x^*, x) \\
&= \left(\frac{h + h^* \circ i}{2}\right)^*(x^*, x) = \frac{1}{2}(h + h^* \circ i)^*(2\,x^*, 2\,x) \\
&\leq \frac{1}{2}(h^* \oplus (h^* \circ i)^*)(2\,x^*, 2\,x) \\
&\leq \frac{1}{2}\left(h^*(x^*, x) + (h^* \circ i)^*(x^*, x)\right) \\
&= \frac{1}{2}\left((h^* \circ i)(x, x^*) + h(x, x^*)\right) = \mathcal{A}h(x, x^*),
\end{aligned}$$

where we have used (8) in the first inequality, the definition of infimal convolution in the second inequality, and the equality $h^{**} = h$ in the last step (recall that h is lower semicontinuous, convex and proper). The fact that the operator $h \longmapsto (\mathcal{A}h)^* \circ i$ maps $\mathcal{H}$ into $\mathcal{H}_{*\geq}$ follows from (i) and the fact that $\mathcal{A}h^{**} = \mathcal{A}h \geq (\mathcal{A}h)^* \circ i$.

(ii) If the equality $h^* \circ i = h$ holds on $\operatorname{dom}(h)$, then we clearly have $\mathcal{A}h = h$ at points where h is finite. At points where h is infinite, $\mathcal{A}h$ must also be infinite (because h^* is proper), and hence $\mathcal{A}h = h$ everywhere. Conversely, assume that $\mathcal{A}h = h$. If $(x, x^*) \in \operatorname{dom}(h)$, then the equality $\mathcal{A}h = h$ yields $(h^* \circ i)(x, x^*) < +\infty$ and hence $(h^* \circ i)(x, x^*) = h(x, x^*)$. This implies that $h^* \circ i = h$ on $\operatorname{dom}(h)$.

(iii) We prove the claim by induction. It is clear from the definition that $\operatorname{dom}(\mathcal{A}h) = \operatorname{dom}(h) \cap \operatorname{dom}(h^* \circ i)$, so the claim is true for $n = 1$. Assume that $\operatorname{dom}(\mathcal{A}^n h) = \operatorname{dom}(h) \cap \operatorname{dom}(h^* \circ i)$. Using (i) yields

$$(\mathcal{A}^n h)^* \circ i \leq \mathcal{A}^n h, \tag{16}$$

for every $n \geq 2$. This implies that

$$\operatorname{dom}(\mathcal{A}^n h) \subset \operatorname{dom}\left((\mathcal{A}^n h)^* \circ i\right).$$

Using the definition of $\mathcal{A}$, the inclusion above, and the induction hypothesis, we can write

$$\begin{aligned} &\operatorname{dom}\left(\mathcal{A}^{n+1} h\right) \\ &= \operatorname{dom}(\mathcal{A}^n h) \cap \operatorname{dom}\left((\mathcal{A}^n h)^* \circ i\right) \\ &= \operatorname{dom}(\mathcal{A}^n h) = \operatorname{dom}(h) \cap \operatorname{dom}(h^* \circ i), \end{aligned}$$

which proves the claim by induction.

(iv) By (16) we can write

$$\mathcal{A}^{n+1} h = \frac{1}{2}\left(\mathcal{A}^n h + (\mathcal{A}^n h)^* \circ i\right) \leq \mathcal{A}^n h, \tag{17}$$

showing that the sequence $\{\mathcal{A}^n h\}$ is pointwise non-increasing. By the order reversing property of the conjugation operator, the sequence $\left\{(\mathcal{A}^n h)^* \circ i\right\}$ is non-decreasing.

Denote by D_0 the set $\operatorname{dom}(h) \cap \operatorname{dom}(h^* \circ i)$. We claim that, for $(x, x^*) \in D_0$, the two sequences $\{\mathcal{A}^n h(x, x^*)\}$ and $\left\{\left((\mathcal{A}^n h)^* \circ i\right)(x, x^*)\right\}$ are adjacent (that is, $\{\mathcal{A}^n h(x, x^*)\}$ is non-increasing, $\left\{\left((\mathcal{A}^n h)^* \circ i\right)(x, x^*)\right\}$ is non-decreasing, and $\lim_{n\to\infty}\left(\mathcal{A}^n h(x, x^*) - \left((\mathcal{A}^n h)^* \circ i\right)(x, x^*)\right) = 0$). We can write

$$\begin{aligned} 0 &\leq (\mathcal{A}^{n+1} h - \left(\mathcal{A}^{n+1} h\right)^* \circ i)(x, x^*) \leq (\mathcal{A}^{n+1} h - (\mathcal{A}^n h)^* \circ i)(x, x^*) \\ &= \frac{1}{2}(\mathcal{A}^n h - (\mathcal{A}^n h)^* \circ i)(x, x^*) < +\infty, \end{aligned}$$

where we have used (i) and (iii) in the left-most inequality, (17) in the second one, the definition of $\mathcal{A}$ in the equality, and the fact that $(x, x^*) \in D_0$ together with (iii) in the last inequality. Hence, we obtain

$$0 \leq (\mathcal{A}^n h - (\mathcal{A}^n h)^* \circ i)(x, x^*) \leq \frac{1}{2^{n-1}}(\mathcal{A}h - \left((\mathcal{A}h)^* \circ i\right))(x, x^*),$$

the second inequality following by induction from the above inequality $(\mathcal{A}^{n+1} h - \left(\mathcal{A}^{n+1} h\right)^* \circ i)(x, x^*) \leq \frac{1}{2}(\mathcal{A}^n h - (\mathcal{A}^n h)^* \circ i)(x, x^*)$, and the claim is established. By (i), the sequence $\{\mathcal{A}^n h(x, x^*)\}$ is bounded below by the function $\pi := \langle \cdot, \cdot \rangle$. Therefore, for every fixed $(x, x^*) \in D_0$, the sequence $\{\mathcal{A}^n h(x, x^*)\} \subset \mathbb{R}$ is non-increasing and bounded below by $\langle x, x^* \rangle \in \mathbb{R}$. The completeness axiom thus yields

$$\mathbb{R} \ni \lim_{n\to\infty}(\mathcal{A}^n h)(x, x^*) = \inf_n (\mathcal{A}^n h)(x, x^*) \geq \langle x, x^* \rangle.$$

Note that, if $(x, x^*) \notin D_0$, then $(\mathcal{A}^n h)(x, x^*) = +\infty$ for all n, so in this case we have $(\mathcal{A}^\infty h)(x, x^*) = +\infty$. From its definition, we have that $\mathcal{A}^\infty h$ is proper and convex, and

$$\operatorname{dom}(\mathcal{A}^\infty h) = \operatorname{dom}(h) \cap \operatorname{dom}(h^* \circ i).$$

To prove that $\mathcal{A}^\infty h$ is a fixed point of $\mathcal{A}$ provided that it is lower semicontinuous we will use (ii). We have just shown that $D_0 = \operatorname{dom}(\mathcal{A}^\infty h)$. We need to prove that $(\mathcal{A}^\infty h)^* \circ i = \mathcal{A}^\infty h$ on D_0. Indeed, take $(x, x^*) \in D_0$. By (iii) and (16), the sequences $\{(\mathcal{A}^n h)(x, x^*)\}$ and $\{((\mathcal{A}^n h)^* \circ i)(x, x^*)\}$ are contained in $\mathbb{R}$. We have shown that the sequence $\{(\mathcal{A}^n h)(x, x^*)\}$ converges monotonically. Since, as noted earlier, the sequences $\{(\mathcal{A}^n h)(x, x^*)\}$ and $\{((\mathcal{A}^n h)^* \circ i)(x, x^*)\}$ are adjacent, they have the same limit $(\mathcal{A}^\infty h)(x, x^*)$. Using this fact, for every $(x, x^*) \in D_0$ we can write

$$\begin{aligned}
\left((\mathcal{A}^\infty h)^* \circ i\right)(x, x^*) &= (\mathcal{A}^\infty h)^*(x^*, x) \\
&= \left(\inf_n \mathcal{A}^n h\right)^*(x^*, x) = \sup_n (\mathcal{A}^n h)^*(x^*, x) \\
&= \sup_n \left((\mathcal{A}^n h)^* \circ i\right)(x, x^*) = \lim_{n\to\infty} \left((\mathcal{A}^n h)^* \circ i\right)(x, x^*) \\
&= \lim_{n\to\infty} (\mathcal{A}^n h)(x, x^*) = (\mathcal{A}^\infty h)(x, x^*),
\end{aligned}$$

showing that $(\mathcal{A}^\infty h)^* \circ i = \mathcal{A}^\infty h$ on D_0. For $(x, x^*) \notin D_0 = \operatorname{dom}(\mathcal{A}^\infty h)$, we trivially have

$$\left((\mathcal{A}^\infty h)^* \circ i\right)(x, x^*) \le (\mathcal{A}^\infty h)(x, x^*) = +\infty = \lim_{n\to\infty} (\mathcal{A}^n h)(x, x^*),$$

the latter equality following from (iii). Hence $\mathcal{A}^\infty h \in \mathcal{H}_{*\le}$. This and (ii) prove that $\mathcal{A}^\infty h$ is a fixed point of $\mathcal{A}$.

(v) The second inequality in (15) follows from the definition of $\mathcal{A}^\infty h$. The first inequality follows from the monotonicity of the sequences $\{\mathcal{A}^n h\}_{n\ge1}$ and $\{(\mathcal{A}^n h)^* \circ i\}_{n\ge1}$ combined with (16). Indeed, for every $n \ge 1$ we have

$$(\mathcal{A}^n h)^* \circ i \le \sup_{m\ge1} (\mathcal{A}^m h)^* \circ i \le \inf_{m\ge1} \mathcal{A}^m h = \mathcal{A}^\infty h.$$

(vi) From (v), for every $(x, x^*) \in \operatorname{dom}(h) \cap \operatorname{dom}(h^* \circ i)$ we have

$$(\mathcal{A}^\infty h)(x, x^*) = \lim_n \left((\mathcal{A}^n h)^* \circ i\right)(x, x^*).$$

Since, according to the proof of (iv), for $(x, x^*) \in \operatorname{dom}(h) \cap \operatorname{dom}(h^* \circ i)$ the sequences $\{(\mathcal{A}^n h)(x, x^*)\}$ and $\{((\mathcal{A}^n h)^* \circ i)(x, x^*)\}$ are adjacent, we immediately obtain that $\{((\mathcal{A}^n h)^* \circ i)(x, x^*)\}$ converges to $(\mathcal{A}^\infty h)(x, x^*)$.

(vii) By (v) and (i) we have

$$(\mathcal{A}^\infty h)(x, x^*) \ge \left((\mathcal{A}h)^* \circ i\right)(x, x^*) \ge \langle x, x^* \rangle$$

for every $(x, x^*) \in X \times X^*$. Hence, if $x^* \in T(x)$, by (v) and Remark 2.7, we have

$$\langle x, x^* \rangle \le (\mathcal{A}^\infty h)(x, x^*) \le (\mathcal{A}h)(x, x^*) = \langle x, x^* \rangle,$$

which proves that $\mathcal{A}^\infty h \in \mathcal{H}(T)$ provided that $\mathcal{A}^\infty h$ is lower semicontinuous. ∎

Remark 2.8 *Since, according to the proof of (iv), one has* $\left((\mathcal{A}^{\infty}h)^{*}\circ i\right)(x,x^{*}) = (\mathcal{A}^{\infty}h)(x,x^{*})$ *for every* $(x,x^{*}) \in \operatorname{dom}(\mathcal{A}^{\infty}h)$ *, the function* $\mathcal{A}^{\infty}h$ *is lower semicontinuous on its domain. Therefore, for the lower semicontinuity assumption of (iv) and (vii) to hold, it is sufficient that* $\mathcal{A}^{\infty}h$ *be lower semicontinuous on the boundary of its domain. In particular, this condition automatically holds if the set* $\operatorname{dom}(h) \cap \operatorname{dom}(h^{*}\circ i)$ *is closed.*

Remark 2.9 *Since, by Remark 2.6, the Fenchel-Young function* f^{FY} *associated with a lower semicontinuous proper convex function* $f : X \longrightarrow \mathbb{R}\cup\{+\infty\}$ *is an autoconjugate representation, it is a fixed point of* $\mathcal{A}$.

We note that the function $\mathcal{A}^{\infty}h$ may fail to be an autoconjugate convex representation of T, as the following example shows.

Example 2.1 *Let* T *be the identity in a Hilbert space. In this case* $\mathcal{F}_T(x,x^{*}) = \frac{\|x+x^{*}\|^2}{4}$ *and*

$$(\mathcal{F}_T^{*}\circ i)(x,x^{*}) = \begin{cases} \|x\|^2, & \text{if } x = x^{*}, \\ +\infty, & \text{if } x \neq x^{*}. \end{cases}$$

If we take $h := \mathcal{F}_T^{*}\circ i$ *then it is easy to check that* $\mathcal{A}h = h$*, and hence* $\mathcal{A}^{\infty}h = h$*. On the other hand,* $\mathcal{A}^{\infty}h = h$ *is not an autoconjugate , since* $h^{*}\circ i = \mathcal{F}_T$*. We have, however,* $h^{*}\circ i = h$ *on the diagonal, that is, on* $\operatorname{dom}(h)$.

The preceding example shows that, in general, $\mathcal{A}^{\infty}h$ may fail to be an autoconjugate of T. The next result establishes an assumption on h under which $\mathcal{A}^{\infty}h$ is an autoconjugate of T.

Corollary 2.1 *With the notation of Theorem 2.3, let* $h \in \mathcal{H}_{*\leq}$*. Assume that the following qualification condition holds:*

$$(\mathcal{QC}) \qquad \operatorname{dom}(h) = \operatorname{dom}(h^{*}\circ i).$$

Then

$$(\mathcal{A}^{\infty}h)^{*}\circ i = \mathcal{A}^{\infty}h.$$

Proof By Theorem 2.3 parts (iv) and (ii), we have $(\mathcal{A}^{\infty}h)^{*}\circ i = \mathcal{A}^{\infty}h$ on $\operatorname{dom}(h)$. On the other hand, by (v) and (i) of Theorem 2.3, we have $\mathcal{A}^{\infty}h \leq h$, and hence $(\mathcal{A}^{\infty}h)^{*} \geq h^{*}$. Therefore

$$\operatorname{dom}((\mathcal{A}^{\infty}h)^{*}\circ i) \subset \operatorname{dom}(h^{*}\circ i) = \operatorname{dom}(h).$$

Since, by Theorem 2.3(iv), we have $\operatorname{dom}(\mathcal{A}^{\infty}h) = \operatorname{dom}(h)$, it follows that $\mathcal{A}^{\infty}h = +\infty = (\mathcal{A}^{\infty}h)^{*}\circ i$ on $(X\times X^{*})\setminus\operatorname{dom}(h)$. ■

The limit $\mathcal{A}^{\infty}h$ found in the previous result provides a constructive example of autoconjugate convex representation, in the following sense.

Theorem 2.4 *Consider the operator $\mathcal{A}$ given in Definition 2.4. Let* $\mathcal{A}^{\infty}h$ *be as in Theorem 2.3 and* $(x,x^*)\in X\times X^*$. *If* $\max\left\{(\mathcal{A}h)(x,x^*),\left((\mathcal{A}h)^*\circ i\right)(x,x^*)\right\}=+\infty$, *then* $(\mathcal{A}^{\infty}h)(x,x^*)=+\infty$. *If* $\max\left\{(\mathcal{A}h)(x,x^*),(\mathcal{A}h)^*(x^*,x)\right\}<+\infty$ *and* $\epsilon>0$, *setting* $n>1+\log_2\left((\mathcal{A}h)(x,x^*)-\left((\mathcal{A}h)^*\circ i\right)(x,x^*)\right)-\log_2\epsilon$, *one has* $(\mathcal{A}^{n}h)(x,x^*)-\epsilon\leq(\mathcal{A}^{\infty}h)(x,x^*)$; *this provides a convenient stopping criterion for effectively computing* $(\mathcal{A}^{\infty}h)(x,x^*)$ *with an error smaller than* ϵ *by means of the iteration* $(\mathcal{A}^{n}h)(x,x^*)$.

Proof If $(\mathcal{A}h)^*(x^*,x)=+\infty$, then $(\mathcal{A}^{\infty}h)(x,x^*)=+\infty$. Let us then assume that $(\mathcal{A}h)^*(x^*,x)<+\infty$. If $(\mathcal{A}h)(x,x^*)=+\infty$, then, by (iv) and (iii) of Theorem 2.3, we also have $(\mathcal{A}^{\infty}h)(x,x^*)=+\infty$. Consider now the case when both $(\mathcal{A}h)(x,x^*)$ and $(\mathcal{A}h)^*(x^*,x)$ are finite. For $n\geq 1$, from the inequalities

$$(\mathcal{A}h)^*\circ i\leq(\mathcal{A}^{n}h)^*\circ i\leq\mathcal{A}^{\infty}h\leq\mathcal{A}^{n}h\leq\mathcal{A}h$$

and the fact that the sequence $\left\{(\mathcal{A}^{n}h)^*\right\}$ is increasing, it follows that

$$\begin{aligned}0&\leq\left(\mathcal{A}^{n+1}h\right)(x,x^*)-(\mathcal{A}^{\infty}h)(x,x^*)\\&\leq\left(\mathcal{A}^{n+1}h\right)(x,x^*)-\left(\left(\mathcal{A}^{n+1}h\right)^*\circ i\right)(x,x^*)\\&\leq\left(\mathcal{A}^{n+1}h\right)(x,x^*)-\left((\mathcal{A}^{n}h)^*\circ i\right)(x,x^*)\\&=\frac{1}{2}\left((\mathcal{A}^{n}h)(x,x^*)+\left((\mathcal{A}^{n}h)^*\circ i\right)(x,x^*)\right)-\left((\mathcal{A}^{n}h)^*\circ i\right)(x,x^*)\\&=\frac{1}{2}\left((\mathcal{A}^{n}h)(x,x^*)-\left((\mathcal{A}^{n}h)^*\circ i\right)(x,x^*)\right);\end{aligned}$$

hence

$$\left(\mathcal{A}^{n+1}h\right)(x,x^*)-(\mathcal{A}^{\infty}h)(x,x^*)\leq\frac{1}{2^n}\left((\mathcal{A}h)(x,x^*)-\left((\mathcal{A}h)^*\circ i\right)(x,x^*)\right);$$

therefore, if $(\mathcal{A}h)(x,x^*)\neq\left((\mathcal{A}h)^*\circ i\right)(x,x^*)$ (otherwise, $(\mathcal{A}^{\infty}h)(x,x^*)=(\mathcal{A}h)(x,x^*)$), then for $\epsilon>0$, taking $n>1+\log_2\left((\mathcal{A}h)(x,x^*)-\left((\mathcal{A}h)^*\circ i\right)(x,x^*)\right)-\log_2\epsilon$, we have

$$\frac{1}{2^{n-1}}\left((\mathcal{A}h)(x,x^*)-\left((\mathcal{A}h)^*\circ i\right)(x,x^*)\right)<\epsilon,$$

and hence

$$(\mathcal{A}^{n}h)(x,x^*)-\epsilon\leq(\mathcal{A}^{\infty}h)(x,x^*)\leq(\mathcal{A}^{n}h)(x,x^*).$$

■

3 A family of enlargements

We now introduce and investigate a family of enlargements, denoted $\mathbb{E}_{\mathcal{H}}(T)$, which is inspired by Fitzpatrick's paper [13]. As we will see next, $\mathbb{E}_{\mathcal{H}}(T) \subset \mathbb{E}(T)$ and therefore this new family inherits all the good properties of the elements of $\mathbb{E}(T)$. In its definition we will use the ϵ-subdifferential of a function $h \in \mathcal{H}(T)$. Equation (12) gives an operator associated to a convex function h defined on $X \times X^*$. We now extend this operator so that it results in a point-to-set mapping defined on $\mathbb{R}_+ \times X$.

Definition 3.1 *Let $T : X \rightrightarrows X^*$ be maximally monotone. For $h \in \mathcal{H}(T)$ and $\epsilon \geq 0$, we define $\breve{T}_h : \mathbb{R}_+ \times X \rightrightarrows X^*$ by*

$$\breve{T}_h(\epsilon, x) := \{x^* \in X^* : (x^*, x) \in \breve{\partial} h(2\epsilon, x, x^*)\}. \tag{18}$$

Remark 3.1 *By Definition 3.1 and* (12)*, if $h \in \mathcal{H}(T)$ then*

$$T_h(x) = \breve{T}_h(0, x).$$

Remark 3.2 *Following Burachik and Svaiter [11], we can associate with $h \in \mathcal{H}(T)$ an enlargement $L^h \in \mathbb{E}(T)$ as follows. For $\epsilon \geq 0$ and $x \in X$ we set*

$$L^h(\epsilon, x) := \{x^* \in X^* : h\,(x, x^*) \leq \langle x, x^* \rangle + \epsilon\}.$$

Conversely, it was also shown in [11] that with every $E \in \mathbb{E}(T)$ we can associate a unique $h \in \mathcal{H}(T)$ such that $E = L^h$.

Remark 3.3 *For every lower semicontinuous proper convex function $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$, one has $L^{f^{FY}} = \breve{\partial} f$.*

We next give a specific notation to the unique h associated with an enlargement $E \in \mathbb{E}(T)$.

Definition 3.2 *Given $E \in \mathbb{E}(T)$, denote by h_E the unique $h \in \mathcal{H}(T)$ such that $E = L^h$.*

Proposition 3.1 *Let $T : X \rightrightarrows X^*$ be maximally monotone, and fix $h \in \mathcal{H}(T)$. Then $\breve{T}_h = L^{\mathcal{A}h} \in \mathbb{E}(T)$, where $\mathcal{A}$ is as in Definition 2.4.*

Proof Our proof will follow from Remark 3.2. Indeed, from Fenchel-Young inequality, we see that $x^* \in \breve{T}_h(\epsilon, x)$ if and only if

$$h(x, x^*) + h^*(x^*, x) \leq \langle (x, x^*), (x^*, x) \rangle + 2\epsilon = 2\left(\langle x, x^* \rangle + \epsilon\right).$$

Equivalently, $x^* \in \breve{T}_h(\epsilon, x)$ if and only if $x^* \in L^{\mathcal{A}h}(\epsilon, x)$. Hence $\breve{T}_h = L^{\mathcal{A}h} \in \mathbb{E}(T)$. ∎

Corollary 3.1 *Let $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuos proper convex function. Then $\breve{\partial} f = \breve{T}_{f^{FY}}$.*

Proof By Proposition 3.1 and Remarks 2.9 and 3.3, we have

$$\breve{T}_{f^{FY}} = L^{\mathcal{A}f^{FY}} = L^{f^{FY}} = \breve{\partial} f.$$

■

Motivated by the preceding result, for a maximally monotone $T : X \rightrightarrows X^*$ we define the following family of enlargements.

$$\mathbb{E}_{\mathcal{H}}(T) := \{E \in \mathbb{E}(T) : \text{there exists } h \in \mathcal{H}(T) \text{ s.t. } E = \breve{T}_h\}. \tag{19}$$

Proposition 3.1 yields the following result.

Corollary 3.2 *Let T and h be as in Proposition 3.1.*

(i) For every $x \in \operatorname{dom}(T)$ and every $\epsilon \geq 0$, the set $\breve{T}_h(\epsilon, x)$ is convex;
(ii) The graph of the mapping $\breve{T}_h$ is demi-closed. *Namely, if $\{x_n\} \subset X$ converges strongly (weakly) to x, $\{x_n^*\} \subset \breve{T}_h(\epsilon_n, x_n)$ converges weakly (strongly, respectively) to x^*, and $\{\epsilon_n\}$ converges to $\epsilon \geq 0$, then $x^* \in \breve{T}_h(\epsilon, x)$. In particular, $\breve{T}_h(\epsilon, x)$ is weakly closed;*
(iii) $\breve{T}_h(0, x) = T(x)$ for every $x \in X$.

Proof For part (i), we use Proposition 3.1. Indeed, the function $\mathcal{A}h \in \mathcal{H}(T)$ is convex, and hence it is direct to check that the set $L^{\mathcal{A}h}(\epsilon, x) = \breve{T}_h(\epsilon, x)$ is convex. For part (ii), we use again Proposition 3.1 and [35, Proposition 4.3]. The latter states that every $E \in \mathcal{H}(T)$ has a demi-closed graph. Part (iii) also follows directly from Proposition 2.1:

$$\breve{T}_h(0, x) = T_h(x) = T(x).$$

■

The following result is another straightforward consequence of Proposition 3.1.

Corollary 3.3 *Let $T : X \rightrightarrows X^*$ be maximally monotone. Then, for every $h \in \mathcal{H}(T)$, $\epsilon \geq 0$ and $x \in X$, we have $\breve{T}_h(\epsilon, x) \subset T^{\mathrm{BE}}(\epsilon, x)$; in particular, $\breve{T}_{\mathcal{F}_T}(\epsilon, x) \subset T^{\mathrm{BE}}(\epsilon, x)$.*

Proof The proof is a direct consequence of Proposition 3.1 and the fact that T^{BE} is the biggest element in $\mathcal{H}(T)$. ■

3.1 **Additivity**

Following [35], an enlargement $E \in \mathbb{E}(T)$ is said to be *additive* when for every $x^* \in E(\epsilon_1, x)$ and every $y^* \in E(\epsilon_2, y)$ we have

$$\langle x - y, x^* - y^* \rangle \geq -(\epsilon_1 + \epsilon_2). \tag{20}$$

Given $E \in \mathbb{E}(T)$, and h_E as in Definition 3.2, we say that h_E is *additive* whenever E is additive. In other words, $h \in \mathcal{H}(T)$ is additive if and only if L^h is additive. We define the following sets

$$\mathcal{H}_a(T) := \{h \in \mathcal{H}(T) : L^h \text{ is additive}\},$$

$$\mathbb{E}_a(T) := \{E \in \mathbb{E}(T) : E \text{ is additive}\}.$$

In the definition below, we use some sets and notation introduced in [36].

Definition 3.3 *For a maximally monotone $T : X \rightrightarrows X^*$ and $h \in \mathcal{H}(T)$, define*

$$S(h) := \{g \in \mathcal{H}(T) : h \geq g \geq g^* \circ i\}. \tag{21}$$

We say that $g \in S(h)$ is minimal (on $S(h)$) if, whenever there is $g' \in S(h)$ such that $g' \leq g$, we must have $g = g'$.

Remark 3.4 *Given $E \in \mathbb{E}(T)$, let $h_E \in \mathcal{H}(T)$ be as in Definition 3.2. In other words, we have $E = L^{h_E}$. Proposition 5.5 in [11] states that*

$$E = L^{h_E} \in \mathbb{E}_a(T) \iff h_E^* \circ i \leq h_E.$$

Remark 3.5 *Definition 3.3 and Remark 3.4 imply that*

$$E \in \mathbb{E}_a(T) \iff S(h_E) \neq \emptyset \iff h_E \in S(h_E).$$

Indeed, if $S(h_E) \neq \emptyset$ then there exists $g \in \mathcal{H}(T)$ such that $h_E \geq g \geq g^* \circ i$. This implies that

$$h_E{}^* \circ i_E \leq g^* \circ i \leq g \leq h_E,$$

so $h_E \in \mathcal{H}_a(T)$ by Remark 3.4. In this situation, $h_E \in S(h_E)$.

It was observed in [35] that additivity, as a property of the graph, can be maximal with respect to inclusion. We recall next this maximality property, an introduce the relation of *mutual additivity* between enlargements as well.

Definition 3.4 *Let $T : X \rightrightarrows X^*$ be maximally monotone.*

(a) *We say that $E \in \mathbb{E}_a(T)$ is* maximally additive *(or* max-add, *for short), if, whenever there exists $E' \in \mathbb{E}_a(T)$ such that*

$$E(\epsilon, x) \subset E'(\epsilon, x), \forall\, \epsilon \geq 0,\, x \in X,$$

then we must have $E = E'$.

(b) *Let $E, E' \in \mathbb{E}(T)$. We say that E and E' are* mutually additive*, if for all $\epsilon, \eta \geq 0$, $x, y \in X$, $x^* \in E(\epsilon, x)$ and $y^* \in E'(\eta, y)$ we have*

$$\langle x - y, x^* - y^* \rangle \geq -(\epsilon + \eta). \tag{22}$$

We denote this situation as $E \sim_a E'$.

Remark 3.6 *Note that E is additive if and only if $E \sim_a E$. Note also that the relation $\sim_a$ is symmetric.*

Remark 3.7 *Take $h \in \mathcal{H}(T)$ such that $h \geq h^* \circ i$. Theorem 2.4 in [36] proves that*

$$h_0 \in S(h) \text{ is minimal in } S(h) \text{ if and only if } h_0^* \circ i = h_0. \tag{23}$$

In other words, minimal elements of $S(h)$ are autoconjugate convex representations of T. We will see that the latter property characterizes max-add enlargements.

Remark 3.8 *For $T = \partial f$, it was proved by Svaiter [35] that the ϵ-subdifferential is max-add. For an arbitrary T, it was proved in [35] that the smallest enlargement of T is always additive, and the existence of a max-add enlargement was deduced in [35] using Zorn's lemma. On the other hand, additivity does not necessarily hold for T^{BE}, the biggest enlargement of T. More precisely, the following weaker inequality was established by Burachik and Svaiter, see [9].*

Theorem 3.1 *Let H be a Hilbert space, $T : H \rightrightarrows H$ be maximally monotone, and $\epsilon, \eta \geq 0$. Then,*

$$\langle x - y, x^* - y^* \rangle \geq -\left(\sqrt{\epsilon} + \sqrt{\eta}\right)^2 \quad \forall x^* \in T^{\mathrm{BE}}(\epsilon, x), y^* \in T^{\mathrm{BE}}(\eta, y).$$

The following result is independent from the enlargement $\breve{T}_h$ and is interesting in its own right.

Proposition 3.2 *Let $T : X \rightrightarrows X^*$ be maximally monotone and $E, E' \in \mathbb{E}(T)$. Assume that $h, h' \in \mathcal{H}(T)$ are such that $E = L^h$ and $E' = L^{h'}$. The following hold.*

(i) *$E \sim_a E'$ if and only if $h^* \circ i \leq h'$. In particular, E is additive if and only if $h^* \circ i \leq h$;*
(ii) *$h = h^* \circ i$ if and only if L^h is max-add.*

In particular, E is mutually additive with $L^{h^* \circ i}$. Inasmuch $L^{h^* \circ i}$ is the largest enlargement which is mutually additive with L^h, it can thus be seen as the "additive complement" of L^h. Moreover, max-add enlargements are characterized by the fact that they coincide with their additive complement.

Proof (i) Assume that (22) holds for all $x^* \in E(\epsilon, x) = L^h(\epsilon, x)$ and all $y^* \in E'(\eta, y) = L^{h'}(\eta, y)$. For every $(x, x^*) \in \mathrm{dom}\,(h)$, set $\epsilon := h(x, x^*) - \langle x, x^* \rangle \geq 0$. Similarly, for $(y, y^*) \in \mathrm{dom}\,(h')$, set $\eta := h'(y, y^*) - \langle y, y^* \rangle \geq 0$. Using (22), we obtain

$$\begin{aligned}\langle (y, y^*), (x^*, x) \rangle - h'(y, y^*) &= \langle y, x^* \rangle + \langle x, y^* \rangle - \langle y, y^* \rangle - \eta \\ &= \langle x, x^* \rangle - \langle x - y, x^* - y^* \rangle - \eta \leq \langle x, x^* \rangle + \epsilon \\ &= h(x, x^*).\end{aligned}$$

Since $(y, y^*) \in \mathrm{dom}\,(h')$ is arbitrary, we can take supremum in the left hand side to obtain $h'^*(x^*, x) \leq h(x, x^*)$, which, taking conjugates, yields $h^* \circ i \leq h'$. Conversely, assume that $h^* \circ i \leq h'$. Take $x^* \in E(\epsilon, x) = L^h(\epsilon, x)$ and $y^* \in E'(\eta, y) =$

$L^{h'}(\eta, y)$. Using the assumption, together with these inclusions and Fenchel-Young inequality, we get

$$\begin{aligned}&\langle (x,x^*),(y^*,y)\rangle\\ \le\ & h(x,x^*)+(h^*\circ i)\,(y,y^*)\\ \le\ & h(x,x^*)+h'^*(y,y^*)\le \langle x,x^*\rangle+\langle y,y^*\rangle+\epsilon+\eta.\end{aligned}$$

Re-arranging the left-most and right-most expressions we obtain (22). The last statement follows by taking $E' = E$ in (i).
(ii) The proof is based on Remark 3.7. Indeed, consider the set $S(h)$ given in Definition 3.3. We claim that L^h is max-add if and only if $h \in S(h)$ and h is minimal in $S(h)$. If the claim is true, then Remark 3.7 readily gives $h^* \circ i = h$. Let us proceed to prove the claim. Indeed, assume first that L^h is max-add. By Remark 3.5, we have $h \in S(h)$. It remains to show that h is minimal in $S(h)$. Let $h' \in S(h)$ be such that $h' \le h$. We must show that $h' = h$. Since $h' \in S(h)$, we have $h'^* \circ i \le h'$, and hence $L^{h'}$ is additive by Remark 3.5. Since $h' \le h$, we have that $L^h(\epsilon, x) \subset L^{h'}(\epsilon, x)$ for all $\epsilon \ge 0$ and all $x \in X$. Using now the fact that L^h is max-add and $L^{h'}$ is additive, we conclude that $L^h = L^{h'}$. Given any enlargement $E \in \mathbb{E}(T)$, the map from E to h_E is a bijection. This fact, together with the equality $L^h = L^{h'}$, allows us to conclude that $h = h'$. Hence h is minimal in $S(h)$. Conversely, assume that $h \in S(h)$ and that h is a minimal element of $S(h)$. Let $h' \in \mathcal{H}_a(T)$ be such that $L^h(\epsilon, x) \subset L^{h'}(\epsilon, x)$ for all $\epsilon \ge 0$ and all $x \in X$. This implies that $h' \le h$; indeed, if $(x, x^*) \in \operatorname{dom}(h)$ then, setting $\epsilon := h(x, x^*) - \langle x, x^*\rangle \ge 0$, we have $(x, x^*) \in L^h(\epsilon, x)$ and hence $(x, x^*) \in L^{h'}(\epsilon, x)$, that is,

$$h'\,(x,x^*) \le \langle x,x^*\rangle + \epsilon = h(x,x^*).$$

Moreover, since $h' \in \mathcal{H}_a(T)$, by Remark 3.4 we have that $h'^* \circ i \le h'$, and hence $h' \in S(h)$. The minimality of h now implies that $h = h'$. In other words, $L^h = L^{h'}$ and therefore L^h is max-add. This completes the proof of the claim. As mentioned above, now (ii) follows directly from the claim and Remark 3.7. The fact that E is mutually additive with $L^{h^*\circ i}$ follows by taking $h' := h^* \circ i$ in part (i). By (i), $L^{h^*\circ i}$ is the largest of all enlargements mutually additive with E. By (ii), E is max add if and only if $h = h^* \circ i$. Equivalently, $L^h = L^{h^*\circ i}$ and hence E coincides with its additive complement. This completes the proof. ■

Remark 3.9 *Using Zorn's lemma, it was proved in [36] that there exists $h \in \mathcal{H}(T)$ such that $h^* \circ i = h$, and hence there are max-add elements in the family $\mathbb{E}(T)$. Other non-constructive examples of autoconjugate convex representations of T can be found in [25, 26]. It is then natural to ask for a constructive example. Indeed, in the case when we are provided with a convex representation h of T whose domain coincides with the domain of $h^* \circ i$, we can constructively obtain both a max-add enlargement by means of Corollary 3.5 below and an autoconjugate convex representation of T (however the coincidence of those domains is an essential condition for having such a possibility, as Example 2.1 shows). Such convex representations can be found in [27] and [3]. The one found in [27] requires a mild constraint qualification, namely, that the affine hull of the domain of T is closed. The other ones* *do not require any constraint qualification. Corollary 2.1 gives an alternative non-constructive proof of the existence of autoconjugate convex representations in the case of operators for which a suitable convex representation is available.*

It was shown in [11] that there is a largest element in the family $\mathcal{H}(T)$, which we denote here by σ_T. It is shown in [11] that $\sigma_T = \operatorname{cl}\operatorname{conv}(\pi + \delta_{G(T)})$, where the notation $\delta_{G(T)}$ is used for the indicator function of the graph of T. Moreover, according to [11, eq. (9)] we have $\sigma_T = (\mathcal{F}_T)^* \circ i$, and this function characterizes the smallest enlargement, i.e., $T^{SE} = L^{\sigma_T}$.

We recover a result from [35] as a corollary of Proposition 3.2.

Corollary 3.4 *The biggest enlargement T^{BE} and the smallest enlargement T^{SE} are mutually additive.*

Proof This follows from the fact that $T^{\mathrm{BE}} = L^{\mathcal{F}_T}$ and $T^{\mathrm{SE}} = L^{\sigma_T}$, together with the equality $\sigma_T = (\mathcal{F}_T)^* \circ i$. ∎

We next show that all members of our family are additive.

Corollary 3.5 *For every $h \in \mathcal{H}(T)$ we have that $\breve{T}_h$ is additive. The enlargement $\breve{T}_h$ is max-add if and only if $(\mathcal{A}h)^* \circ i = \mathcal{A}h$. Consequently, if $h^* \circ i = h$ then $\breve{T}_h$ is max-add.*

Proof Recall (Proposition 3.1) that

$$\breve{T}_h = L^{\mathcal{A}h}. \tag{24}$$

By Proposition 3.2(i), it is enough to prove that $(\mathcal{A}h)^* \circ i \le \mathcal{A}h$, which is precisely the conclusion of the first assertion in Theorem 2.3(i). This proves the first statement. The second statement follows from (24) and Proposition 3.2(ii). If $h = h^* \circ i$ then we have $\mathcal{A}h = (\mathcal{A}h)^* \circ i$. Indeed, if $h = h^* \circ i$ it is direct to check that $\mathcal{A}h = h$. So $(\mathcal{A}h)^* \circ i = h^* \circ i = h = \mathcal{A}h$. By (24) and Proposition 3.2(ii), we conclude that $\breve{T}_h$ is max-add. ∎

A consequence of the above results is that, besides the smallest enlargement, a whole subfamily of enlargements happens to be additive. If they derive from an autoconjugate h, then they are max-add and hence they can be regarded as "structurally closer" to the epsilon subdifferential. We will see in the next section that a particular member of this subfamily is precisely the ϵ-subdifferential when $T = \partial f$.

4 The case $T := \partial f$

We want now to establish the relation between our new enlargement and the ϵ-subdifferential in the case $T := \partial f$.

Lemma 4.1 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous proper convex function, and let $h \in \mathcal{H}(\partial f)$ be such that $h \le f^{FY}$. Then, for every $(x, x^*) \in X \times X^*$, one has*

$$h^{FY}\left((x, x^*), (x^*, x)\right) \ge \langle x, x^* \rangle + f^{FY}(x, x^*).$$

Proof From the inequality $h \leq f^{FY}$ it follows that $h^* \geq \left(f^{FY}\right)^* = f^{FY} \circ i$; hence, using that $h(x,x^*) \geq \langle x, x^*\rangle$, we obtain

$$h^{FY}\left((x,x^*),(x^*,x)\right) = h(x,x^*) + h^*(x^*,x) \geq \langle x,x^*\rangle + f^{FY}(x,x^*).$$

■

Theorem 4.1 *Let f and h be as in Lemma 4.1. Then, for every $\epsilon > 0$ and $x \in X$, one has*

$$\breve{T}_h(\frac{\epsilon}{2},x) \subset \breve{\partial} f(\epsilon,x)$$

Proof Let $x^* \in \breve{T}_h(\frac{\epsilon}{2},x) = L^{\mathcal{A}h}(\frac{\epsilon}{2},x)$. Then, by Lemma 4.1, we have

$$\begin{aligned} f^{FY}(x,x^*) &\leq h^{FY}\left((x,x^*),(x^*,x)\right) - \langle x,x^*\rangle \\ &= 2\mathcal{A}h(x,x^*) - \langle x,x^*\rangle \\ &\leq 2\left(\langle x,x^*\rangle + \tfrac{\epsilon}{2}\right) - \langle x,x^*\rangle = \langle x,x^*\rangle + \epsilon, \end{aligned}$$

where we used the definition of $\mathcal{A}$ in the first equality, and the assumption on x^* in the second inequality. This shows that $x^* \in \breve{\partial} f(\epsilon,x)$. ■

Remark 4.1 *As observed in Remark 2.3, when $\epsilon = 0$ in Theorem 4.1(ii) we recover the equality $T_h = \partial f$, proved in [13, Example 2.3].*

When $h := \mathcal{F}_{\partial f}$, we can strengthen the inclusion in Theorem 4.1:

Proposition 4.1 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous proper convex function. Then, for every $\epsilon \geq 0$ and $x \in X$, we have*

$$\breve{T}_{\mathcal{F}_{\partial f}}(\frac{\epsilon}{2},x) \subset T^{\mathrm{SE}}(\epsilon,x).$$

Proof Suppose that $x^* \in \breve{T}_{\mathcal{F}_{\partial f}}(\frac{\epsilon}{2},x)$. Then we can write

$$\frac{1}{2}(\mathcal{F}_{\partial f} + \mathcal{F}^*_{\partial f} \circ i)(x,x^*) \leq \langle x,x^*\rangle + \frac{\epsilon}{2}.$$

Using the fact that $\mathcal{F}_{\partial f} \geq \langle \cdot,\cdot\rangle$, the last inequality yields

$$\left(\mathcal{F}^*_{\partial f} \circ i\right)(x,x^*) \leq \langle x,x^*\rangle + \epsilon. \tag{25}$$

Equivalently, $x^* \in L^{\mathcal{F}^*_{\partial f}\circ i}(\epsilon,x) = T^{\mathrm{SE}}(\epsilon,x)$ (see the proof of Corollary 3.4). ■

Remark 4.2 *Since h^{FY} is autoconjugate , we see that $\breve{T}_{h^{FY}}$ is max-add. Indeed, this fact follows from Theorem 4.1 and [35, Theorem 6.4] (see also Corollary 3.5). Is this the only max-add enlargement of $T := \partial f$? The answer is no, since an example in [3] shows three different autoconjugate convex representations of a subdifferential operator, which result in three different max-add enlargements.*

Acknowledgements The authors would like to thank the two anonymous referees for their valuable comments and suggestions, which have led to an improved paper. The authors are very grateful to Heinz Bauschke and Benar Fux Svaiter for their comments and corrections on an earlier version of this manuscript. Bauschke kindly indicated to us an additional reference for Remark 2.5, while Svaiter kindly indicated that the results in [36] constitute the earliest proof of the "only if" part of Theorem 2.2.

References

1. Ya. I. Alber, A. N. Iusem, and M. V. Solodov. On the projected subgradient method for nonsmooth convex optimization in a Hilbert space. *Math. Programming*, 81(1, Ser. A):23–35, 1998.
2. H. H. Bauschke and X. Wang. The kernel average for two convex functions and its application to the extension and representation of monotone operators. *Trans. Amer. Math. Soc.*, 361(11):5947–5965, 2009.
3. H.H. Bauschke, X. Wang, and L. Yao. Autoconjugate representers for linear monotone operators. *Math. Program.*, 123(1, Ser. B):5–24, 2010.
4. R.I. Boţ, E.R. Csetnek, and G. Wanka. A new condition for maximal monotonicity via representative functions. *Nonlinear Anal.*, 67(8):2390–2402, 2007.
5. A. Brøndsted and R. T. Rockafellar. On the subdifferentiability of convex functions. *Proc. Amer. Math. Soc.*, 16:605–611, 1965.
6. R. S. Burachik and A. N. Iusem. *Set-valued mappings and enlargements of monotone operators*, volume 8 of *Springer Optimization and Its Applications*. Springer, New York, 2008.
7. R. S. Burachik, A. N. Iusem, and B. F. Svaiter. Enlargement of monotone operators with applications to variational inequalities. *Set-Valued Anal.*, 5(2):159–180, 1997.
8. R. S. Burachik, C. A. Sagastizábal, and B. F. Svaiter. ϵ-enlargements of maximal monotone operators: theory and applications. In *Reformulation: nonsmooth, piecewise smooth, semismooth and smoothing methods (Lausanne, 1997)*, volume 22 of *Appl. Optim.*, pages 25–43. Kluwer Acad. Publ., Dordrecht, 1999.
9. R. S. Burachik and B.F. Svaiter. ϵ-enlargements of maximal monotone operators in Banach spaces. *Set-Valued Anal.*, 7(2):117–132, 1999.
10. R. S. Burachik and B.F. Svaiter. Operating enlargements of monotone operators: new connections with convex functions. *Pac. J. Optim.*, 2(3):425–445, 2006.
11. R.S. Burachik and B. F. Svaiter. Maximal monotone operators, convex functions and a special family of enlargements. *Set-Valued Anal.*, 10(4):297–316, 2002.
12. R.S. Burachik and B. F. Svaiter. Maximal monotonicity, conjugation and the duality product. *Proc. Amer. Math. Soc.*, 131(8):2379–2383, 2003.
13. S. Fitzpatrick. Representing monotone operators by convex functions. In *Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988)*, volume 20 of *Proc. Centre Math. Anal. Austral. Nat. Univ.*, pages 59–65. Austral. Nat. Univ., Canberra, 1988.
14. S.D. Flåm. Monotonicity and market equilibrium. *Preprint, 2014*.
15. N. Ghoussoub. Maximal monotone operators are selfdual vector fields and vice-versa. *http://arxiv.org/abs/math/0610494v1*, 2006.
16. N. Ghoussoub. A variational theory for monotone vector fields. *J. Fixed Point Theory Appl.*, 4(1):107–135, 2008.
17. K. C. Kiwiel. *Methods of descent for nondifferentiable optimization*, volume 1133 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1985.
18. K. C. Kiwiel. Proximity control in bundle methods for convex nondifferentiable minimization. *Math. Programming*, 46(1, (Ser. A)):105–122, 1990.
19. N. Krylov. Properties of monotone mappings. *Lith. Math. J.*, 22:140–145, 1982.
20. C. Lemaréchal and C. Sagastizábal. Variable metric bundle methods: from conceptual to implementable forms. *Math. Programming*, 76(3, Ser. B):393–410, 1997.
21. M. Marques Alves and B. F. Svaiter. A new proof for maximal monotonicity of subdifferential operators. *J. Convex Anal.*, 15(2):345–348, 2008.
22. J.-E. Martínez-Legaz and M. Théra. ϵ-subdifferentials in terms of subdifferentials. *Set-Valued Anal.*, 4(4):327–332, 1996.
23. J.-E. Martínez-Legaz and M. Théra. A convex representation of maximal monotone operators. *J. Nonlinear Convex Anal.*, 2(2):243–247, 2001. Special issue for Professor Ky Fan.
24. J.J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, 93:273–299, 1965.
25. J.-P. Penot. Autoconjugate functions and representations of monotone operators. *Bull. Austral. Math. Soc.*, 67(2):277–284, 2003.
26. J.-P. Penot. The relevance of convex analysis for the study of monotonicity. *Nonlinear Anal.*, 58(7-8):855–871, 2004.
27. J.-P. Penot and C. Zălinescu. Some problems about the representation of monotone operators by convex functions. *ANZIAM J.*, 47(1):1–20, 2005.

28. J. P. Revalski and M. Théra. Enlargements and sums of monotone operators. *Nonlinear Anal.*, 48(4, Ser. A: Theory Methods):505–519, 2002.
29. T. Roche, R. Rossi, and U. Stefanelli. Stability results for doubly nonlinear differential inclusions by variational convergence. *SIAM J. on Control and Optim.*, 52(2):1071–1107, 2014.
30. R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.
31. R. T. Rockafellar. On the maximal monotonicity of subdifferential mappings. *Pacific J. Math.*, 33:209–216, 1970.
32. H. Schramm and J. Zowe. A version of the bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results. *SIAM J. Optim.*, 2(1):121–152, 1992.
33. S. Simons. *From Hahn-Banach to monotonicity*, volume 1693 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2008.
34. S. Simons and C. Zălinescu. A new proof for Rockafellar's characterization of maximal monotone operators. *Proc. Amer. Math. Soc.*, 132(10):2969–2972, 2004.
35. B. F. Svaiter. A family of enlargements of maximal monotone operators. *Set-Valued Anal.*, 8(4):311–328, 2000.
36. B. F. Svaiter. Fixed points in the family of convex representations of a maximal monotone operator. *Proc. Amer. Math. Soc.*, 131(12):3851–3859, 2003.
37. A. Visintin. Scale-transformations of maximal monotone relations in view of homogenization. *Boll. Unione Mat. Ital. (9)*, 3(3):591–601, 2010.
38. A. Visintin. Variational formulation and structural stability of monotone equations. *Calc. Var. Partial Differential Equations*, 47(1-2):273–317, 2013.
39. A. Visintin. An extension of the Fitzpatrick theory. *Comm. on Pure and Appl. Anal.*, 13(5):2039–2058, 2014.