ENLARGEMENTS: A BRIDGE BETWEEN MAXIMAL MONOTONICITY AND CONVEXITY

REGINA S. BURACHIK

ABSTRACT

Perhaps the most important connection between maximally monotone operators and convex functions is the fact that the subdifferential of a convex function is maximally monotone. This connects convex functions with a proper subset of maximally monotone operators (i.e., the cyclically monotone operators). Our focus is to explore maps going in the opposite direction, namely those connecting an arbitrary maximally monotone map with convex functions. In this survey, we present results showing how enlargements of a maximally monotone operator T provide this connection. Namely, we recall how the family of enlargements is in fact in a bijective correspondence with a whole family of convex functions. Moreover, each element in either of these families univocally defines T. We also show that enlargements are not merely theoretical artifacts, but have concrete advantages and applications, since they are, in some sense, better behaved than T itself. Enlargements provide insights into existing tools linked to convex functions. A recent example is the use of enlargements for defining a distance between two point-to-set maps, one of them being maximally monotone. We recall this new distance here, and briefly illustrate its applications in characterizing solutions of variational problems.

MATHEMATICS SUBJECT CLASSIFICATION 2020

Primary 47H04; Secondary 47H05, 49J53, 90C25, 58E35

KEYWORDS

Maximally monotone operators, set valued maps, enlargements, convex analysis, convex functions, inclusion problem, variational inequalities, Bregman distances

INTERNATIONAL CONGRESS OF MATHEMATICIANS
2022 JULY 6–14

Proc. Int. Cong. Math. 2022, Vol. 7, pp. 5212–5233
DOI 10.4171/ICM2022/107

Published by EMS Press

1. INTRODUCTION

In the framework of real Banach spaces, convex analysis and the theory of monotone operators have shown a fascinating interplay: the theory of one can be used for developing further the theory of the other, and the advances in one of them generates advances of the other. This cross-fertilization between both theories is a key topic in modern Functional Analysis, and has been captivating mathematicians since the middle of the 20th century. The natural way of going from convex analysis to maximally monotone maps is via the subdifferential of a convex function, which is maximally monotone [29,35]. More precisely, subdifferentials are the (proper) subset of *cyclically monotone* maps [33].

Given a maximally monotone operator T, we approximate T by another set-valued map, called from now on an *enlargement* of T. Our goal is to show that enlargements provide a fruitful way of going from maximally monotone operators to convex functions. As an example, enlargements are used to prove a formula involving infimal convolution of convex functions [20], or to show new equivalences for the situation in which two convex functions differ by a constant [13], or for characterizing solutions of difference of convex (DC) problems [7]. More examples of this interplay can be found, for instance, in [9, CHAPTER 5], as well as in [31,37]. We will see in what follows the crucial rôle of convex analysis in establishing outer-semicontinuity of the enlargements.

The concept of enlargement was first hinted in 1996 by Martínez-Legaz and Théra in [28]. Independently, the enlargement was formally defined and studied for the first time in the 1997 paper [10]. The results we quote in this survey span a time period from the late 1990s until today. We also quote crucial results obtained by Svaiter in [38], where the formal definition of the family of enlargements is introduced, and a fundamental link with convexity is established. Another key result mentioned here is the introduction of a family of convex functions associated with T, suggested by Fitzpatrick in 1988 in [23]. A beautiful fact is that these two families, seemingly independent from each other, are actually in a bijective correspondence, as we will see in Theorem 30.

A main motivation for studying maximally monotone maps and their approximations is the inclusion problem, stated as

$$\text{Find } x^* \in X \text{ such that } z \in T(x^*), \tag{1.1}$$

where $T : X \rightrightarrows X^*$ is a maximally monotone operator between a Banach space X and its dual X^*. Model (1.1) is used for solving fundamental problems, such as optimality conditions for (smooth and nonsmooth) optimization problems, fixed point problems, variational inequalities, and solutions of nonlinear equations. If T is point-to-point, the inclusion above becomes an equality.

This survey is organized as follows. In Section 2 we give the theoretical setting and the main definitions and basic results that we will need in later sections. In this section we recall the Fenchel–Young function, and also the Fitzpatrick function. In Section 3 we define the family of enlargements and give prototypical examples. In this section we describe the structure of the family (in terms of smaller and larger elements), and recall some of its continuity properties and the Brøndsted–Rockafellar property. We end this section recalling a bijective correspondence between the family of enlargements of T and a family of convex functions. In Section 4 we define a family of convex functions associated with the family of enlargements, and illustrate this definition with examples. In this section we also describe the structure of this family, with smallest and largest elements, in analogy with the situation for the family of enlargements. In Section 5 we recall a new distance between point-to-set maps induced by the family of convex functions associated with T. We illustrate this new concept with applications to variational problems, and we include some open questions on these distances. Section 6 contains a few final words.

To facilitate reading, most technical proofs are kept to a minimum, and the focus is set on the main ideas and key points of the results. Interested readers can consult the references given regarding each result. Some results combine several existing facts, and their proofs are given to illustrate the type of analysis used in this topic.

2. PRELIMINARIES

Throughout this paper, X is a real *reflexive* Banach space with topological dual X^* and duality pairing between them denoted by $\langle \cdot, \cdot \rangle$. The norm in any space is denoted by $\| \cdot \|$. We use w to represent the weak topologies both on X and X^*. When using the weak topology, we will mention it explicitly, otherwise the strong topology is assumed. Let Z be a topological space and consider a subset $A \subset Z$, we denote by $\overline{A}$ its closure with respect to the strong topology, by $\operatorname{int}(A)$ the *interior* of A and by $\operatorname{co}(A)$ the *convex hull* of A.

2.1. Basic facts and tools

Recall the following definitions concerning extended real valued functions.

Definition 1. Let Z be a topological space and consider a function $f : Z \to \mathbb{R} \cup \{+\infty\}$.

(i) The *epigraph* of f is the set

$$\operatorname{epi}(f) := \big\{(z,t) \in Z \times \mathbb{R} : f(z) \leq t\big\}.$$

(ii) The *domain* of f is the set

$$\operatorname{dom} f := \big\{x \in Z : f(x) < +\infty\big\}.$$

(iii) The function f is said to be *proper* if $\operatorname{dom} f \neq \emptyset$.

(iv) The function f is said to be *lower-semicontinuous (lsc)* if $\operatorname{epi}(f)$ is closed.

The following definitions are relevant to convex functions.

Definition 2. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ will be a convex function.

(i) The *subdifferential* of f is the point-to-set map $\partial f : X \rightrightarrows X^*$ defined by

$$\partial f(x) := \big\{x^* \in X^* : f(y) - f(x) \geq \langle y - x, x^* \rangle, \forall y \in X\big\},$$
$$\text{if } x \in \operatorname{dom} f, \tag{2.1}$$

and $\partial f(x) := \emptyset$, otherwise.

(ii) Given $\epsilon \geq 0$, the ϵ-*subdifferential of* f is the point-to-set map $\partial_\epsilon f : X \rightrightarrows X^*$ defined by

$$\partial_\epsilon f(x) := \{x^* \in X^* : f(y) - f(x) \geq \langle y - x, x^* \rangle - \epsilon, \forall y \in X\}, \\ \text{if } x \in \operatorname{dom} f, \tag{2.2}$$

and $\partial_\epsilon f(x) := \emptyset$, otherwise. Note that $\partial_0 f = \partial f$. To keep ϵ as a variable, we restate the ϵ-subdifferential of f in a way that does not involve ϵ. Namely, we consider the point-to-set map $\breve{\partial} f : \mathbb{R}_+ \times X \rightrightarrows X^*$ defined by

$$\breve{\partial} f(\varepsilon, x) := \partial_\epsilon f(x), \tag{2.3}$$

and call the ensuing enlargement the *Brøndsted–Rockafellar* enlargement.

(iii) The *Fenchel–Moreau conjugate* of f, denoted as $f^* : X^* \to \mathbb{R} \cup \{+\infty\}$, is defined by

$$f^*(x^*) := \sup\{\langle x, x^* \rangle - f(x) : x \in X\}. \tag{2.4}$$

(iv) The *Fenchel–Young function* associated to f is the function $f^{\mathrm{FY}} : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f^{\mathrm{FY}}(x, x^*) := f(x) + f^*(x^*) \quad \text{for all } (x, x^*) \in X \times X^*. \tag{2.5}$$

Remark 1. Note that f^* is always convex and (weakly) lsc. Hence, it is also strongly lsc by convexity. Therefore, f^{FY} is a convex, proper, and $(\|\cdot\| \times w)$-lsc function on $X \times X^*$. A remarkable and well-known fact is that f^{FY} completely characterizes the operator ∂f, in the following sense:

$$\partial f(x) = \{x^* \in X^* : f^{\mathrm{FY}}(x, x^*) = \langle x, x^* \rangle\}. \tag{2.6}$$

More precisely, the definitions yield

$$\begin{aligned} &f(x) + f^*(v) \geq \langle x, v \rangle, \quad \forall (x, v) \in X \times X^*, \\ &f(x) + f^*(v) = \langle x, v \rangle \iff x \in X,\ v \in \partial f(x). \end{aligned} \tag{2.7}$$

Moreover, f^{FY} completely characterizes the map $\breve{\partial} f$ (see (2.3)). Indeed, the definitions yield

$$\breve{\partial} f(\varepsilon, x) = \{v \in X^* : f^{\mathrm{FY}}(x, v) = f(x) + f^*(v) \leq \langle x, v \rangle + \varepsilon\}. \tag{2.8}$$

Since we always have $f^{\mathrm{FY}}(x, v) \geq \langle x, v \rangle$, when $\epsilon = 0$, (2.8) collapses into (2.6). System (2.7) constitutes the main inspiration for defining a family of convex functions induced by a maximally monotone operator T.

The map $\breve{\partial} f$ has a fundamental rôle in variational analysis. It is used for (i) developing algorithms for nonsmooth optimization (e.g., the ε-subgradient method, bundle methods, perturbed proximal methods), (ii) characterizing minimizing/stationary sequences, and (iii) characterizing approximate solutions of optimization problems. A crucial theoretical impact is the fact that it was used by Rockafellar to show maximality of ∂f in [35].

All information on a point-to-set map is encapsulated in its graph.

Definition 3. Let Z, Y be topological spaces and $F : Z \rightrightarrows Y$ a point-to-set map. The *graph* of F is the set

$$G(F) := \{(z, y) \in Z \times Y : y \in F(z)\}.$$

Given a point-to-set map F, we define $\overline{F}$ as the point-to-set map by the following equality:

$$G(\overline{F}) := \overline{G(F)},$$

where the closure in the right-hand side is taken with respect to the product topology in $Z \times Y$.

We next recall well known concepts related with point-to-set maps from X to X^*.

Definition 4. Let $T : X \rightrightarrows X^*$ be a point-to-set map.

(i) The *domain* of T is denoted by $D(T)$ and defined by

$$D(T) := \{x \in X : T(x) \neq \emptyset\},$$

and the *range* of T, denoted by $R(T)$, is defined by

$$R(T) := \{v \in X^* : \text{exists } x \in D(T) \text{ such that } v \in T(x)\},$$

(ii) T is said to be *monotone* when

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

(iii) A monotone operator T is called *maximally monotone* provided

$$\langle y - x, y^* - x^* \rangle \geq 0 \quad \forall (y, y^*) \in G(T) \text{ implies } (x, x^*) \in G(T).$$

Equivalently, T is maximally monotone when $G(T)$ cannot be properly extended (in the sense of the inclusion), without violating the monotonicity condition given in (ii).

Continuity properties are associated with closedness of the graph. Hence, the topology we use determines the continuous maps. Besides from closedness w.r.t. the strong and weak topologies, we will consider sequential closedness with respect to the strong topology in X and the weak topology in X^*.

Recall the standard notation for strong and weak convergence in a reflexive Banach space: Given a sequence $(z_n) \subset X$, and an element $z \in X$, we denote by $z_n \to z$ the strong convergence of (z_n) to z. Given a sequence $(w_n) \subset X^*$ and an element $w \in X^*$,we denote by $w_n \rightharpoonup w$ the weak convergence of (w_n) to w. In this situation, we say that the sequence (z_n, w_n) *converges (sw)* (for strong-weak) to (z, w).

Definition 5. Let X be a reflexive Banach space and fix $S \subset X \times X^*$. We say that S is *sequentially strong-weak-closed*, denoted as $(sw)_s$-closed, if for every sequence $(x_n, v_n) \subset S$ the following condition holds:

$$\text{If } x_n \to x, \quad v_n \rightharpoonup v, \text{ then } (x, v) \in S.$$

We say that S is *sequentially weakly closed*, if S contains all weak limits of its weakly convergent sequences.

Remark 2. If $S \subset X \times X^*$ is weakly closed, then it is sequentially weakly closed (this is true for every topological space). The converse is in general not true (see, e.g, [2, EXAMPLE 3.33]), unless S is convex.

Fact 6. (i) *If S is weakly closed then S is $(sw)_s$-closed.*

(ii) *Assume that S is convex and S is strongly closed. Then, it is weakly closed (and hence $(sw)_s$-closed).*

Proof. (i) If S is weakly closed then, by the previous remark, S is sequentially weakly closed. Since every strongly convergent sequence is weakly convergent, this implies that S is $(sw)_s$-closed.

(ii) If S is strongly closed and convex, by [9, COROLLARY 3.4.16], S is weakly closed. By (i), it is $(sw)_s$-closed. ∎

2.2. The Fitzpatrick function

From Remark 1, we see that the function f^{FY} completely characterizes the operator $T := \partial f$ and its enlargement $\breve{\partial} f$. A fundamental step in extending this type of link to an arbitrary maximally monotone operator T was performed by Fitzpatrick in 1988 in [23], who defined the following function, now called the *Fitzpatrick function associated with T*:

$$\begin{aligned} F_T(x,v) &:= \sup\bigl\{\langle y,v\rangle + \langle x-y,u\rangle \;:\; (y,u) \in G(T)\bigr\} \\ &= \sup\bigl\{\langle y-x,v-u\rangle + \langle x,v\rangle \;:\; (y,u) \in G(T)\bigr\}. \end{aligned}$$

By [23, THEOREMS 3.4 AND 3.8], we have that

$$\begin{aligned} F_T(x,v) &= \langle x,v\rangle \quad \text{if and only if } (x,v) \in G(T), \\ F_T(x,v) &\geq \langle x,v\rangle \quad \text{or every } x \in X, v \in X^*, \end{aligned} \tag{2.9}$$

Note the similarity with system (2.7). In other words, F_T characterizes $G(T)$, in a similar way as f^{FY} characterizes $G(\partial f)$. The Fitzpatrick function remained unnoticed for several years until it was rediscovered in [28]. In [24] Flåm gave an economic interpretation of the Fitzpatrick function, and also mentioned that this function was already used in 1982 by Krylov in [27]. The Fitzpatrick function allows for tractable reformulations of hard problems, including variational representation of (nonlinear) evolutionary PDEs, and the development of variational techniques for the analysis of their structural stability; see, e.g., [25, 32, 40, 41]. In [23, THEOREM 3.10], Fitzpatrick proved that F_T is the smallest function among all those that verify system (2.9). We will further explore this fact in later sections. Namely, we will revisit system (2.9) when defining the family of convex functions associated with the enlargements of T.

3. ENLARGEMENTS OF MAXIMALLY MONOTONE MAPS

We start this section by recalling the definition of enlargement, introduced by Svaiter in [38]. Then, we will focus on some well-known properties of set valued maps: local boundedness, outersemicontinuity, Lipschitz continuity, and the Brøndsted–Rockafellar property.

To explore continuity properties, we will often consider the closure of a point-to-set map $E : \mathbb{R} \times X \rightrightarrows X^*$. By Definition 3, for $F := E$, $\overline{E}$ is the point-to-set map such that $G(\overline{E}) = \overline{G(E)}$, where the closure is taken with respect to the strong topology in all spaces. In what follows, T is a fixed maximally monotone operator. From now on, most of the proofs will be omitted to alleviate the reading. All these proofs can be found in the references given for every result. The proofs which I do provide are meant to illustrate the type of mathematical tools used in the analysis, without making the text too technical.

3.1. The family of enlargements $\mathbb{E}(T)$

The theoretical framework that follows is based on the groundbreaking definition of a family of enlargements of T, introduced by Svaiter in [38].

Definition 7. Let $T : X \rightrightarrows X^*$ be a maximally monotone map. We say that a point-to-set mapping $E : \mathbb{R}_+ \times X \rightrightarrows X^*$ is an *enlargement* of T if the following hold:

(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

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

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

The set of all maps verifying (E_1)–(E_3) is denoted by $\mathbb{E}(T)$. We say that E is *closed* if $G(E)$ is (strongly) closed. The set of all closed enlargements is denoted by $\mathbb{E}_c(T)$.

Remark 3. Condition (E_1) ensures that E is an enlargement of T, while (E_2) indicates that the enlargement is increasing with respect to ϵ. Condition (E_3) allows constructing new elements in $G(E)$ by using convex combinations of known elements in $G(E)$. As we will see below, this condition is essential for establishing the link between enlargements of maximally monotone operators and convex functions. Note that if conditions (E_1)–(E_3) hold for E, then they also hold for $\overline{E}$, hence if $E \in \mathbb{E}(T)$, then $\overline{E} \in \mathbb{E}(T)$. By taking $\epsilon_1 = \epsilon_2 = 0$ and $\alpha_1, \alpha_2 \in (0, 1)$ in (E_3), we deduce that $E(0, \cdot)$ is a monotone map. By (E_1), we also have $E(0, \cdot) \supset T$. By maximality, we must have $E(0, \cdot) = T$.

Example 8. The set $\breve{\partial} f(\epsilon, x)$ is nonempty for every $\epsilon > 0$ if and only if f is lower semicontinuous at x. The map $\breve{\partial} f$ is an enlargement of $T = \partial f$. The fact that it verifies (E_1)–(E_2)

follows directly from the definitions. The proof of condition (E_3) for $\breve{\partial} f$ can be found, e.g., in [14, LEMMA 2.1]. It follows from the definitions that $\breve{\partial} f$ is a closed enlargement of ∂f, namely, $\breve{\partial} f \in \mathbb{E}_c(\partial f)$.

We recall next an example of an enlargement of an arbitrary maximally monotone operator T.

Example 9. Define the point-to-set map $T^e : \mathbb{R}_+ \times X \rightrightarrows X^*$ as follows:

$$T^e(\varepsilon, x) := \begin{cases} \{v \in X^* : \langle u - v, y - x\rangle \geq -\varepsilon, \forall (y, u) \in G(T)\}, & \forall x \in D(T), \\ \emptyset, & \text{if } x \notin D(T). \end{cases}$$

As mentioned in the Introduction, this enlargement of T was explicitly defined for the first time in [10]. The fact that it verifies conditions (E_1)–(E_2) follows directly from the definition. The transportation formula, i.e., condition (E_3) for T^e, is established in [17, 21]. It follows from the definitions that it is a closed enlargement of T, namely, $T^e \in \mathbb{E}_c(T)$. The enlargement T^e has been used for developing (i) inexact prox-like methods for variational inequalities [8, 10, 11, 15, 18], (ii) bundle-type methods for finding zeroes of maximally monotone operators [21, 30], and (iii) a unifying convergence analysis for algorithms for variational inequalities [16]. More recently, T^e has been used for developing inexact versions of the Douglas–Rachford algorithm for finding zeroes of sums of maximally monotone operators [1, 22, 39]. This list is by no means complete, but serves as evidence of the impact this concept has had on the development of inexact methods for variational inequalities and related problems. One of the reasons for this enlargement to have so many applications is the fact that, as we will see in Section 3.2, it has better continuity properties than the original T.

Remark 4. We mentioned above the fact that F_T characterizes T (see (2.9)). Moreover, F_T also characterizes T^e. Indeed, it follows directly from the definitions that

$$F_T(x, v) \leq \langle x, v\rangle + \epsilon \quad \text{if and only if } v \in T^e(\epsilon, x). \tag{3.1}$$

Remark 5. When $T = \partial f$, we always have from the definitions that $\breve{\partial} f(\epsilon, x) \subset (\partial f)^e(\epsilon, x)$. The opposite inclusion can be strict, as observed in [28], see also [9, EXAMPLE 5.2.5(IV)].

We mentioned above that $T^e \in \mathbb{E}_c(T)$, we can say more about its "location" within this family. The following result was established in [38].

Theorem 10. *The family $\mathbb{E}(T)$ has a largest and a smallest element (with respect to the inclusion of their graphs). The largest element is T^e, and the smallest element is*

$$T^s(\epsilon, x) = \bigcap_{E \in \mathbb{E}(T)} E(\epsilon, x).$$

Moreover, T^e is the largest element in $\mathbb{E}_c(T)$, and $\overline{T^s}$ is the smallest element in $\mathbb{E}_c(T)$. In other words, for every $E \in \mathbb{E}_c(T)$, we have

$$G(\overline{T^s}) \subset G(E) \subset G(T^e).$$

3.2. Local boundedness

Consider an iterative method for solving problem (1.1) that generates a sequence $(x^k, v^k) \subset G(T)$. Assume that the sequence $(x^k) \subset X$ is convergent. In this situation, can we say something about the behavior of the sequence (v^k)? The answer to this question requires more knowledge about $G(T)$. In fact, if T is maximally monotone, and (x^k) has its limit in the interior of $D(T)$, then $(v^k) \subset X^*$ is bounded and hence has a weakly convergent subsequence by Bourbaki–Alaoglu's theorem. The fundamental property needed here is the *local boundedness* of T in the interior of its domain.

Definition 11. Let X be a topological space and Y a metric space. A point-to-set map $F : X \rightrightarrows Y$ is said to be *locally bounded at* $x \in D(F)$ if there exists an open neighborhood U of x such that $F(U) := \bigcup_{z\in U} F(z)$ is bounded, and it is said to be *locally bounded* when it is locally bounded at every $x \in D(F)$.

Maps that are monotone are locally bounded at every point of the interior of their domains. When they are also maximal, they are not locally bounded at any point of the boundary of their domains. The latter fact means that we cannot expect enlargements to be locally bounded at any point of the boundary of their domains. Hence, we concentrate on points in the interior of their domains. Since $G(E) \supset \{0\} \times G(T)$ and we use enlargements to approximate T, we need to ensure that the local boundedness property is not lost when replacing T by E. In fact, we will see that $G(E)$ is not "too large," in the sense that the local boundedness property in the interior of the domains is still preserved.

Local boundedness of maximally monotone maps was established by Rockafellar in [34], and later extended to more general cases by Borwein and Fitzpatrick [3]. To make our study specific for enlargements, we will use a refined notion of local boundedness for point to set maps defined on $\mathbb{R}_+ \times X$.

Definition 12. Let $E : \mathbb{R}_+ \times X \rightrightarrows X^*$ be a point-to-set mapping. We say that E is *affine locally bounded* at $x \in X$ when there exists an open neighborhood V of x and positive constants L, M such that

$$\sup_{\substack{y\in V,\\ v\in E(\varepsilon,y)}} \|v\| \leq L\varepsilon + M.$$

Remark 6. By Theorem 10, $G(T^e) \supset G(E)$ for every $E \in \mathbb{E}(T)$. Therefore, all local boundedness properties enjoyed by T^e are inherited by $E \in \mathbb{E}(T)$. This means that it is enough to study the local boundedness property for T^e. This result is [17, COROLLARY 3.10], which we recall next.

Theorem 13 (Affine local boundedness). *If $T : X \rightrightarrows X^*$ is monotone, then T^e is affine locally bounded in* $\operatorname{int}D(T)$. *In other words, for all $x \in \operatorname{int}D(T)$ there exist a neighborhood V of x and positive constants L, M such that*

$$\sup\left\{\|v\| : v \in T^e(\varepsilon, y), y \in V\right\} \leq L\varepsilon + M$$

for all $\varepsilon \geq 0$.

Remark 7. The local boundedness property allows us to give an answer to the question posed at the start of Section 3.2. We mentioned there that local boundedness of T implies that, when the limit of (x^k) is in the interior of $D(T)$, the sequence (v^k) has a subsequence which is a weakly convergent. Theorem 13 and Remark 6 show that this fact is still true for any enlargement $E \in \mathbb{E}(T)$.

3.3. Lipschitz continuity

Consider again a method that generates a sequence $(x^k) \subset D(T)$ and assume that $X = X^* = \mathbb{R}^n$. Assume again that your sequence (x^k) converges to some $x \in D(T)$. Given any fixed $v \in Tx$, can you find a sequence $v^k \in Tx^k$ such that (v^k) converges to v? Enlargements of T do verify this property. They actually verify a much stronger property: Lipschitz continuity. On the other hand, the above mentioned property is not true for maximally monotone operators. Indeed, if $f(t) = |t|$ then $T = \partial f$ does not verify it at $t = 0$. Indeed, maximally monotone operators satisfy this property at a point x if and only if Tx is a single point. The latter fact is shown in [34] (see also [9, THEOREM 4.6.3]).

Definition 14. Let Z and Y be Banach spaces and $F : Z \rightrightarrows Y$ a point-to-set map. Let U be a subset of $D(F)$ such that F is closed-valued on U. The mapping F is said to be *Lipschitz continuous* on U if there exists a *Lipschitz constant* $\kappa > 0$ such that for all $x, x' \in U$ it holds that

$$F(x) \subset F(x') + \kappa \|x - x'\| B(0,1),$$

where $B(0,1) := \{y \in Y : \|y\| \le 1\}$. In other words, for every $x \in U$, $v \in F(x)$, and $x' \in U$, there exists $v' \in F(x')$ such that

$$\|v - v'\| \le \kappa \|x - x'\|.$$

The fact that enlargements of T are Lipschitz continuous at every point in the interior of their domains was proved in [17, THEOREM 3.14]. For more details on these properties, see [9, CHAPTER 5].

3.4. On graphs, closedness, and convexity

Assume now that your method generates a sequence $(x^k, v^k) \subset G(T)$ which has a limit point $(x, v) \in X \times X^*$. When can you ensure that $(x, v) \in G(T)$? In other words, is $G(T)$ closed? Closedness of the graph of a point-to-set map is a type of continuity called *outer-semicontinuity* [9, DEFINITION 2.5.1(A) AND THEOREM 2.5.4]. When T is maximally monotone, $G(T)$ is strongly closed and also $(sw)_s$-closed, see [9, PROPOSITION 4.2.1]. We will show in this section that enlargements enjoy the same kind of outer-semicontinuity. We will see that convexity has a crucial rôle in establishing this. Let $E : \mathbb{R}_+ \times X \rightrightarrows X^*$ be any point-to-set map. Definition 3 gives

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

To see $G(E)$ as the epigraph of a (possibly not convex) function, rearrange the order of its variables and consider the set:

$$\tilde{G}(E) := \big\{(x, v, t) \in X \times X^* \times \mathbb{R}_+ : v \in E(t, x)\big\}.$$

We see next how this set determines a convex function that characterizes $G(E)$. The next result from [38, LEMMA 3.2] is the key in linking enlargements with convexity.

Lemma 15. *Let $\Phi : X \times X^* \times \mathbb{R} \to X \times X^* \times \mathbb{R}$ be defined as*

$$\Phi(x, v, \epsilon) = \big(x, v, \epsilon + \langle v, x\rangle\big),$$

and let $E : \mathbb{R}_+ \times X \rightrightarrows X^*$ be any point-to-set map. The following statements are equivalent:

(i) *E verifies (E_3);*

(ii) *$\Phi(\tilde{G}(E)) \subset X \times X^* \times \mathbb{R}$ is a convex set.*

Fact 16. *Let $E \in \mathbb{E}(T)$.*

(i) *$\tilde{G}(E)$ is $(sw)_s$-closed if and only if $\Phi(\tilde{G}(E))$ is $(sw)_s$-closed.*

(ii) *$\tilde{G}(E)$ is (strongly) closed if and only if $\Phi(\tilde{G}(E))$ is (strongly) closed.*

Proof. (i) Assume that $\tilde{G}(E)$ is $(sw)_s$-closed and take a sequence $(x_n, v_n, s_n) \in \Phi(\tilde{G}(E))$ such that $x_n \to x$, $v_n \rightharpoonup v$, and $s_n \to s$. By definition of Φ, this means that $(x_n, v_n, s_n - \langle x_n, v_n\rangle) \in \tilde{G}(E)$ and

$$x_n \to x, \quad v_n \rightharpoonup v, \quad \text{and} \quad s_n - \langle x_n, v_n\rangle \to s - \langle x, v\rangle.$$

Since $\tilde{G}(E)$ is $(sw)_s$-closed, we deduce that $(x, v, s - \langle x, v\rangle) \in \tilde{G}(E)$. Equivalently, $(x, v, s) \in \Phi(\tilde{G}(E))$ and therefore $\Phi(\tilde{G}(E))$ is sequentially $(sw)_s$-closed. The converse implication follows identical steps, mutatis mutandis. The proof of (ii) follows from the fact that Φ and Φ^{-1} are continuous with respect to the strong topology. ∎

Convexity is crucial for ensuring outer-semicontinuity of the enlargements of T.

Corollary 17. *Let $E \in \mathbb{E}_c(T)$. Then E is $(sw)_s$-outer-semicontinuous.*

Proof. Since $E \in \mathbb{E}_c(T)$, we know that $G(E)$ (and equivalently, $\tilde{G}(E)$) is (strongly) closed. By Fact 16(ii) and Lemma 15, $\Phi(\tilde{G}(E))$ is (strongly) closed and convex. By Fact 6(ii), it is $(sw)_s$-closed. Finally, Fact 16(i) yields that $\tilde{G}(E)$ $(sw)_s$-closed, showing the outer-semicontinuity. ∎

3.5. Brøndsted and Rockafellar property

Since $\{0\} \times G(T) \subset G(E)$, the transportation formula allows producing elements in $G(E)$ by using elements $G(T)$. Can we use elements in $G(E)$ to approach those in $G(T)$?

For $T = \breve{\partial} f$, Brøndsted and Rockafellar [5] showed that any $(\varepsilon, x, v) \in G(\breve{\partial} f)$ can be approximated by an element $(x', v') \in G(\partial f)$ in the following way:

$$\text{For all } \eta > 0, \text{ there exists } v' \in \partial f(x') \text{ such that } \|x - x'\| \leq \frac{\varepsilon}{\eta} \text{ and } \|v - v'\| \leq \eta.$$

This result is known as *Brøndsted–Rockafellar's lemma*. The remarkable fact is that enlargements enjoy this property, too. While this property holds for $T = \breve{\partial} f$ in any Banach space, more general enlargements require reflexivity of the space. A consequence of the Brøndsted and Rockafellar property is the fact that the domain and range of an enlargement is dense in the domain and range of T, respectively. We sketch the proof of the Brøndsted and Rockafellar property here because it is beautiful and is based on crucial results on maximally monotone operators. For that, we recall that the duality mapping in a Banach space is defined as $J := \partial g$ for $g(x) := \frac{1}{2}\|x\|^2$. Using the definition of the subdifferential, it can be shown [9, PROPOSITION 4.4.4(I)] that the duality mapping verifies

$$J(x) := \{v \in X^* : \langle x, v \rangle = \|x\|^2, \|x\| = \|v\|\}. \tag{3.2}$$

The proof uses a key property of T in reflexive Banach spaces. Namely, the surjectivity of $T + \alpha J$ for $\alpha > 0$. This surjectivity property, which is, moreover, a characterization of maximally monotone operators in reflexive spaces, was established by Rockafellar in [36]. Since $G(T^e) \supset G(E)$ for every $E \in \mathbb{E}_c(T)$, it is enough to show that the Brøndsted–Rockafellar property holds for T^e.

Theorem 18. *Let $(\varepsilon, x_\varepsilon, v_\varepsilon) \in G(T^e)$ be given. For all $\eta > 0$, there exists $(x, v) \in G(T)$ such that*

$$\|v - v_\varepsilon\| \leq \frac{\varepsilon}{\eta} \quad \textit{and} \quad \|x_\varepsilon - x\| \leq \eta.$$

Proof. The claim trivially holds if $\varepsilon = 0$ because in this case by (E_1) we can take $(x, v) = (x_\varepsilon, v_\varepsilon) \in G(T)$. Assume that $\varepsilon > 0$. For any fixed $\beta > 0$, define $T_\beta(\cdot) := \beta T(\cdot) + J(\cdot - x_\varepsilon)$. Since T is maximally monotone, the surjectivity property mentioned above implies that there exist $x \in X$ and $v \in Tx$ such that $\beta v_\varepsilon \in \beta v + J(x - x_\varepsilon)$, which rearranges as $\beta(v_\varepsilon - v) \in J(x - x_\varepsilon)$. Using the fact that $(\varepsilon, x_\varepsilon, v_\varepsilon) \in G(T^e)$ and the definition of J in (3.2), we have

$$-\varepsilon \leq \langle v_\varepsilon - v, x_\varepsilon - x \rangle = -\frac{1}{\beta}\|x - x_\varepsilon\|^2 = -\beta\|v - v_\varepsilon\|^2,$$

and the result follows by taking $\beta := \eta^2/\varepsilon$ and rearranging the expression above. ■

Since $G(E) \supset \{0\} \times G(T)$, we may wonder whether the range and domain of E might be much larger that those of T. The precise situation is a consequence of Theorem 18, and is stated next. Again, it is enough to establish this result for T^e.

Corollary 19. *The following hold:*

(i) $R(T) \subset R(T^e) \subset \overline{R(T)}$;

(ii) $D(T) \subset D(T^e) \subset \overline{D(T)}$.

Proof. The rightmost inclusions in (i) and (ii) follow from the previous theorem by taking $\eta \to +\infty$ for (i) and $\eta \to 0$ for (ii), respectively. The leftmost inclusions follow from (E_1). ■

4. A FAMILY OF CONVEX FUNCTIONS ASSOCIATED WITH $\mathbb{E}(T)$

We have seen that the convexity emanating from condition (E_3) ensures that enlargements are outer-semicontinuous. How can this fact be used to associate enlargements with convex functions? All information on E is encapsulated in the set $\tilde{G}(E)$. Hence, we start by using this set to define the epigraph of a function defined in $X \times X^*$. The results in this section either directly use those in [19], or combine these for ease of presentation.

Definition 20. Let $S \subset X \times X^* \times \mathbb{R}$. The *lower envelope* of S is the function $\gamma : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\gamma(x, v) := \inf\big\{t \in \mathbb{R} : (x, v, t) \in S\big\},$$

with the convention that $\inf \emptyset = +\infty$.

Fact 21. *Let $S \subset X \times X^* \times \mathbb{R}$ be a nonempty set and let γ be its lower envelope as in Definition 20. The following properties hold:*

(i) $S \subset \operatorname{epi}(\gamma)$.

(ii) *If S is closed and has epigraphical structure (i.e., if $(x, v, t) \in S$ then $(x, v, s) \in S$ for every $s > t$) then $S = \operatorname{epi}(\gamma)$.*

(iii) *S is closed if and only if γ is lower semicontinuous.*

(iv) *S is closed and convex if and only if γ is convex, proper, and lower semicontinuous.*

Proof. The proofs of (i), (iii), and (iv) follow directly from the definitions. For (ii), use the definition of infimum and the closedness of S to deduce that $(x, v, \gamma(x, v)) \in S$. Now the epigraphical structure yields $\operatorname{epi}(\gamma) \subset S$. ■

The following simple lemma shows when $\tilde{G}(E) = \operatorname{epi}(\lambda)$ for some function λ.

Lemma 22. *Let $E : \mathbb{R}_+ \times X \rightrightarrows X^*$ and let $\lambda : X \times X^* \to \mathbb{R} \cup \{+\infty\}$. The following statements are equivalent:*

(i) $\tilde{G}(E) = \operatorname{epi}(\lambda)$.

(ii) $\lambda \geq 0$ *and* $E(t, x) = \{v \in X^* : \lambda(x, v) \leq t\}$ *for all* $t \geq 0$, $x \in X$.

Proof. [(i) → (ii)] By definition of E, $\tilde{G}(E) \subset X \times X^* \times \mathbb{R}_+$, so $D(E) \subset \mathbb{R}_+ \times X$. Hence, $\lambda \geq 0$. Now use (i) to write

$$v \in E(b, x) \leftrightarrow (x, v, b) \in \tilde{G}(E) \overset{\text{(i)}}{=} \operatorname{epi}(\lambda) \leftrightarrow \lambda(x, v) \leq b.$$

[(ii) → (i)] We have

$$(x, v, b) \in \tilde{G}(E) \leftrightarrow b \geq 0,\ v \in E(b, x) \overset{\text{(ii)}}{\leftrightarrow} \lambda(x, v) \leq b \leftrightarrow (x, v, b) \in \operatorname{epi}(\lambda). \qquad \blacksquare$$

Imposing closedness assumptions, we obtain lsc of λ, as we see in the next result, which is [19, PROPOSITION 3.1].

Theorem 23. *Let $E : \mathbb{R}_+ \times X \rightrightarrows X^*$ be a point-to-set map such that $G(E)$ is closed and verifies (E_2). Define $\lambda_E : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ as*

$$\lambda_E(x, v) := \inf\{t \geq 0 : (x, v, t) \in \tilde{G}(E)\} = \inf\{t \geq 0\ :\ v \in E(t, x)\}.$$

In other words, λ_E is the lower envelope of $\tilde{G}(E)$. The following hold:

(i) $\tilde{G}(E) = \operatorname{epi}(\lambda_E)$.

(ii) *λ_E is strongly lsc.*

(iii) $\lambda_E \geq 0$.

(iv) $E(t, x) = \{v \in X^* : \lambda_E(x, v) \leq t\}$ *for all $t \geq 0$, $x \in X$.*

Furthermore, λ_E is the unique function that verifies (iii) and (iv). In particular, the mapping $E \mapsto \lambda_E$ is one-to-one in $\mathbb{E}_c(T)$.

Proof. Items (i) and (ii) follow from Fact 21(ii)–(iii). Item (iii) follows from the definition, and item (iv) follows from (iii) and (i). For the uniqueness of λ_E, use Lemma 22. The injectivity of the mapping $E \mapsto \lambda_E$ follows from (i). ■

The set $\tilde{G}(E)$ is (in general) not convex, and the same holds for λ_E. For generating a convex function, we use again the map Φ.

Definition 24. Let $E \in \mathbb{E}_c(T)$ and let λ_E be as in Theorem 23. Define $\Lambda_E : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ by the equality

$$\operatorname{epi}(\Lambda_E) := \Phi(\operatorname{epi}(\lambda_E)).$$

Namely,

$$\Lambda_E(x, v) = \lambda_E(x, v) + \langle x, v \rangle, \tag{4.1}$$

for every $(x, v) \in X \times X^*$.

The following result constitutes the link between enlargements and convex functions.

Theorem 25. *Let $E \in \mathbb{E}_c(T)$ and let Λ_E as in Definition 24. The following hold:*

(i) *Λ_E is convex and (strongly) lsc.*

(ii) *Λ_E verifies*

$$\Lambda_E(x, v) \geq \langle x, v \rangle, \quad \textit{for every } (x, v) \in X \times X^*,$$
$$\Lambda_E(x, v) = \langle x, v \rangle, \quad \textit{if } v \in Tx.$$

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

Proof. To prove (i), use Definition 24 to write

$$\mathrm{epi}(\Lambda_E) = \Phi\big(\mathrm{epi}(\lambda_E)\big) = \Phi\big(\tilde{G}(E)\big),$$

where we used Theorem 23(i) in the second equality. The convexity now follows from Lemma 15. To prove that the function Λ_E is lsc, use the fact that $E \in \mathbb{E}_c(T)$ and Fact 16(ii). The inequality in (ii) follows from (4.1) and Theorem 23(iii). As for the equality in (ii), assume that $v \in Tx$. By (E_1), this implies that $v \in E(0, x)$ and, by definition, $\lambda_E(x, v) \leq 0$. Since $\lambda_E \geq 0$, we must have $\lambda_E(x, v) = 0$. The latter, combined with (4.1), yields $\Lambda_E(x, v) = \langle x, v\rangle$. The last assertion holds by (4.1) and the fact that, given E, the function λ_E is uniquely defined. ■

The following result is [19, Corollary 3.2] and characterizes E in terms of the convexity of Λ_E.

Corollary 26. *Let E be a point-to-set map with closed graph which verifies (E_1)–(E_2). Let Λ_E be as in Definition 24. The following statements are equivalent:*

(i) $E \in \mathbb{E}_c(T)$.

(ii) Λ_E *is convex.*

(iii) E *verifies* (E_3).

Proof. The equivalence between (i) and (iii) follows directly from the assumptions and the definitions. The equivalence between (ii) and (iii) follows from the definition of Λ_E and Lemma 15. Indeed, the definition of Λ_E and Theorem 23(i) gives

$$\mathrm{epi}(\Lambda_E) = \Phi\big(\mathrm{epi}(\lambda_E)\big) = \Phi\big(\tilde{G}(E)\big).$$

By Lemma 15, E verifies (E_3) if and only if $\Phi(\tilde{G}(E))$ is convex, and by the above expression the latter is equivalent to the convexity of Λ_E. ■

Remark 8. Let $E = T^e \in \mathbb{E}_c(T)$. In this case, we have that $\Lambda_{T^e} = F_T$, the Fitzpatrick function associated with T. This is a consequence of (2.9), and Remark 4. Indeed, (2.9) and the latter remark imply that conditions (iii) and (iv) in Theorem 23 hold for $\lambda(x, v) := F_T(x, v) - \langle x, v\rangle$. By the uniqueness property stated in the same theorem, we must have $\lambda_{T^e}(x, v) = F_T(x, v) - \langle x, v\rangle$. Since we know that $T^e \in \mathbb{E}_c(T)$, (4.1) now yields $\Lambda_{T^e} = F_T$.

Theorem 25 helps us identify the relevant set of convex functions, which we define next.

Definition 27. Let $\mathcal{H}_0$ be the set of all convex and (strongly) lower semicontinuous functions defined on $X \times X^*$. The *Fitzpatrick family of* T is the set

$$\mathcal{H}(T) := \big\{h \in \mathcal{H}_0 : h(x, v) \geq \langle x, v\rangle \text{ for all } (x, v) \in X \times X^*, \\ \text{and } h(x, v) = \langle x, v\rangle \text{ whenever } v \in Tx\big\}.$$

Remark 9. By taking $S = \operatorname{epi}(h)$ in Fact 6, we see that h is (strongly) lower semicontinuous and convex if and only if it is weakly lower semicontinuous, and the latter is equivalent to h being $(sw)_s$-lower semicontinuous.

Example 28. Using Definition 27, Remark 1 states that $f^{\mathrm{FY}} \in \mathcal{H}(\partial f)$. Similarly, the system (2.9) implies that $F_T \in \mathcal{H}(T)$ for every T maximally monotone operator.

Theorem 25 provides a map from $\mathbb{E}(T)$ to $\mathcal{H}(T)$. Namely, the map $E \mapsto \Lambda_E$, with Λ_E fully characterizing $G(T)$ in the sense of condition (ii) in the statement of the theorem. We will define next the inverse of this map. Define $\pi : X \times X^* \to \mathbb{R}$ as $\pi(x, v) := \langle x, v\rangle$ for every $(x, v) \in X \times X^*$.

Theorem 29. *Let $h : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ be a convex and lsc function. Consider the point-to-set map $L^h : \mathbb{R} \times X \rightrightarrows X^*$ defined by*

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

for every $t \in \mathbb{R}$, $x \in X$. The following statements are equivalent:

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

(ii) $h \geq \pi$, $D(L^h) \subset \mathbb{R}_+ \times X$ *and* $L^h \in \mathbb{E}_c(T)$.

(iii) $\Lambda_{L^h} = h$ *and* $L^h(0, x) \supset Tx$.

Proof. [(i) $\to$ (ii)] The definition on $\mathcal{H}(T)$ directly gives $h \geq \pi$. The latter inequality also yields $D(L^h) \subset \mathbb{R}_+ \times X$. Indeed, $v \in L^h(t, x)$ if and only if $h(x, v) \leq t + \pi(x, v)$. Since $h \geq \pi$, this implies that $t \geq 0$, so $D(L^h) \subset \mathbb{R}_+ \times X$. We need to show that $E := L^h$ verifies (E_1)–(E_3). Condition (E_2) follows trivially from the definition of L^h. We next check (E_1). By (i), we know that $h(x, v) = \langle x, v\rangle$ for every $v \in Tx$. Therefore,

$$Tx \subset \big\{u \in X^* : h(x, u) \leq \langle x, v\rangle\big\} = L^h(0, x),$$

so (E_1) holds. To verify (E_3) we will show that $\Phi(\tilde{G}(L^h))$ is convex and use Lemma 15. Indeed, by definition of Φ and L^h,

$$\begin{aligned}\Phi\big(\tilde{G}(L^h)\big) &= \big\{\big(x, v, t + \langle x, v\rangle\big) : v \in L^h(t, x)\big\}\\ &= \big\{\big(x, v, t + \langle x, v\rangle\big) : h(x, v) \leq t + \langle x, v\rangle\big\} =: \Phi_h.\end{aligned}$$

We claim that $\Phi_h = \operatorname{epi}(h)$. Indeed, it is clear that $\Phi_h \subset \operatorname{epi}(h)$. To show that $\operatorname{epi}(h) \subset \Phi_h$, take any $(x, v, s) \in \operatorname{epi}(h)$. We need to show that we can write $s = t + \langle x, v\rangle$ for some $t \geq 0$. Indeed, by (i), we know that $h(x, v) \geq \langle x, v\rangle$. Hence, $\langle x, v\rangle \leq h(x, v) \leq s$ so $t = s - \langle x, v\rangle \geq 0$. This shows that the claim is true and $\Phi_h = \operatorname{epi}(h)$. Since h is convex, the above expression gives the convexity of $\Phi(\tilde{G}(L^h))$, and Lemma 15 furnishes (E_3).
[(ii) $\to$ (iii)] The inclusion $L^h(0, x) \supset Tx$ is (E_1), which holds because $L^h \in \mathbb{E}_c(T)$. Let us show that $\Lambda_{L^h} = h$. By (ii), we can apply Theorem 23(iv) to $E := L^h$ and obtain

$$L^h(t, x) = \big\{v \in X^* : h(x, v) - \langle x, v\rangle \leq t\big\} = \big\{v \in X^* : \lambda_{L^h}(x, v) \leq t\big\},$$
$$\forall t \geq 0,\ \forall x \in X,$$

where used the definition of L^h in the first equality. Use $t := \lambda_{L^h}(x, v)$ in the middle set to obtain $h(x, v) - \langle x, v\rangle \leq \lambda_{L^h}(x, v)$. By (ii), we have that $h \geq \pi$ so that $t := h(x,v) - \pi(x,v) \geq 0$ can be used in the right-most set to derive $\lambda_{L^h}(x,v) \leq h(x,v) - \langle x,v\rangle$. Therefore, $\lambda_{L^h}(x,v) + \langle x,v\rangle = h(x,v)$. Using this equality and the definition of Λ_{L^h}, we have that

$$\Lambda_{L^h}(x,v) = \lambda_{L^h}(x,v) + \langle x,v\rangle = h(x,v),$$

so (iii) is proved.

[(iii) $\to$ (i)] Since h is lsc, L^h is a point-to-set map with closed graph that verifies (E_2). By Theorem 23(iii), we have that $\lambda_{L^h} \geq 0$. The definitions and (iii) now yield

$$h = \Lambda_{L^h} = \lambda_{L^h} + \pi \geq \pi,$$

which gives the inequality in the definition of $\mathcal{H}(T)$. Now we need to prove that $h(x,v) = \pi(x,v)$ whenever $v \in Tx$. Since $Tx \subset L^h(0,x)$, we can use Theorem 23(iv) for $t = 0$ to deduce that $\lambda_{L^h}(x,v) \leq 0$. By Theorem 23(iii) again, we have that $\lambda_{L^h} \geq 0$. Altogether, $\lambda_{L^h}(x,v) = 0$ if $v \in Tx$. Using (iii), we can write

$$h(x,v) = \Lambda_{L^h}(x,v) = \lambda_{L^h}(x,v) + \pi(x,v) = \pi(x,v),$$

when $v \in Tx$. Therefore, $h \in \mathcal{H}(T)$. ■

Remark 10. For every $h \in \mathcal{H}(T)$, we have that $h = \pi$ if and only if $v \in Tx$. Indeed, the "if" part follows because $h \in \mathcal{H}(T)$. Conversely, let $h = \pi$ and take $E := L^h \in \mathbb{E}_c(T)$. By Theorem 23(iv), $v \in L^h(0,x)$ and by Theorem 29 and Remark 3 $T = L^h(0,\cdot)$ so $v \in Tx$.

To complete this section, we combine the results above to obtain a bijection between $\mathbb{E}(T)$ and $\mathcal{H}(T)$. In the following result, we consider the sets $\mathbb{E}(T)$ and $\mathcal{H}(T)$ as partially ordered. In $\mathbb{E}(T)$ we use the partial order of the inclusion of the graphs: $E_1 \preceq E_2$ if and only if $G(E_1) \subset G(E_2)$. In $\mathcal{H}(T)$ we use the natural partial order of pointwise comparison in $X \times X^*$: $h_1 \leq h_2$ if and only if $h_1(x,v) \leq h_2(x,v)$ for every $(x,v) \in X \times X^*$.

Theorem 30. *The map $\Theta : \mathbb{E}_c(T) \to \mathcal{H}(T)$ defined as $\Theta(E) := \Lambda_E$ is a bijection, with inverse $\Theta^{-1}(h) = L^h$. Considering the partially ordered spaces $(\mathbb{E}(T), \preceq)$ and $(\mathcal{H}(T), \leq)$, the bijection Θ is "order reversing", i.e., if $E_1 \preceq E_2$ then $h_2 := \Theta(E_2) \leq h_1 := \Theta(E_1)$. Therefore, $\Theta(T^e) = F_T$ and F_T is the smallest element of the family $\mathcal{H}(T)$.*

Proof. By Theorem 25, we know that $\Theta(E) = \Lambda_E \in \mathcal{H}(T)$ and the function Θ is injective. The function Θ is surjective because by Theorem 29, $L^h \in \mathbb{E}_c(T)$ and hence $\Theta(L^h) = \Lambda_{L^h} = h$ for every $h \in \mathcal{H}(T)$. The fact that Θ is order reversing is proved via the following chain of equivalent facts:

$$E_1 \preceq E_2 \overset{\text{definition}}{\longleftrightarrow} \tilde{G}(E_1) \subset \tilde{G}(E_2) \overset{\text{Theorem 23(i)}}{\longleftrightarrow} \operatorname{epi}(\lambda_{E_1}) \subset \operatorname{epi}(\lambda_{E_1})$$
$$\leftrightarrow \lambda_{E_2} \leq \lambda_{E_1} \overset{(4.1)}{\leftrightarrow} \Theta(E_2) = \Lambda_{E_2} \leq \Theta(E_1) = \Lambda_{E_1}.$$

Since $G(T^e) \supset G(E)$ for every $E \in \mathbb{E}_c(T)$, we can use the fact that Θ is order reversing and Remark 8 to obtain

$$\Theta(T^e) = \Lambda_{T^e} \overset{\text{Remark 8}}{=} F_T \leq \Theta(E),$$

for every $E \in \mathbb{E}_c(T)$. Since $\Theta(\mathbb{E}_c(T)) = \mathcal{H}(T)$, the claim is proved. ■

5. A DISTANCE INDUCED BY $\mathcal{H}(T)$

In this section we recall how the family $\mathcal{H}(T)$ is used for defining a new distance between set-valued maps. More results can be found in [6, 7, 12]. Let $h \in \mathcal{H}(T)$, and let $S : X \rightrightarrows X^*$ be any point-to-set operator. Following [12, DEFINITION 3.1], for each $(x, y) \in X \times X$, define

$$\mathcal{D}_S^{\flat,h}(x,y) := \begin{cases} \inf_{v \in Sy}(h(x,v) - \langle x, v\rangle) & \text{if } (x,y) \in \operatorname{dom} S \times \operatorname{dom} T, \\ +\infty & \text{otherwise} \end{cases} \tag{5.1a}$$

$$\text{and} \quad \mathcal{D}_S^{\sharp,h}(x,y) := \begin{cases} \sup_{v \in Sy}(h(x,v) - \langle x, v\rangle) & \text{if } (x,y) \in \operatorname{dom} S \times \operatorname{dom} T, \\ +\infty & \text{otherwise.} \end{cases} \tag{5.1b}$$

We call these distances *Generalized Bregman distances* (GBDs). When S is point to point, all three collapse into one, $\mathcal{D}_S^h := \mathcal{D}_S^{\flat,h} = \mathcal{D}_S^{\sharp,h}$. The GBDs specialize to the Bregman distance. Given a proper, convex, and differentiable function $f : X \to \mathbb{R}$, the (classical) associated Bregman distance [4] is defined as

$$\mathcal{D}_f(x,y) := f(x) - f(y) + \langle y - x, \nabla f(y)\rangle. \tag{5.2}$$

It is easy to check that GBDs reduce to (5.2) when $T = S = \nabla f$ and $h := f^{\mathrm{FY}}$ [12, PROPOSITION 3.5]. We show next that GBDs can be used to characterize approximate solutions of the sum problem:

$$\text{find } x \in X \text{ such that } 0 \in Sx + Tx, \tag{5.3}$$

where $S, T : X \rightrightarrows X^*$ are point-to-set maps, with T being maximally monotone. The proof of the next result, inspired by [12, PROPOSITION 3.7], is [7, PROPOSITION 2.2].

Proposition 31. *Suppose that $T : X \rightrightarrows X^*$ is a maximally monotone operator and $S : X \rightrightarrows X^*$ is a point-to-set operator. Fix any $h \in \mathcal{H}(T)$, $\varepsilon \in \mathbb{R}_+$, and $x \in X$. Consider the following statements:*

(a) $0 \in L^h(\varepsilon, x) + Tx$.

(b) $\mathcal{D}_{-S}^{\flat,h}(x,x) \leq \varepsilon$.

Then (a) $\Longrightarrow$ (b). Moreover, if $\operatorname{dom} S$ *is open and S is locally bounded with weakly closed images, then the two statements are equivalent.*

Next we illustrate how optimality conditions for minimizing a DC (difference of convex) function can be expressed by means of a particular type of GBD. Let $f : X \to$

$\mathbb{R} \cup \{+\infty\}$ and $g : X \to \mathbb{R}$ be proper and convex lsc functions. We consider the problem of finding $x \in X$ that globally minimizes $f - g$. In the proposition below, the equivalence between statements (a) and (b) are well known in finite-dimensional spaces (see, e.g., [26, THEOREM 3.1]). Its extension to reflexive Banach spaces can be found in [7, PROPOSITION 2.3].

Proposition 32. *The following statements are equivalent:*

(a) *x is a global minimum of $f - g$ on X.*

(b) *For all $\varepsilon \geq 0$, $\breve{\partial} g(\varepsilon, x) \subseteq \breve{\partial} f(\varepsilon, x)$.*

(c) *For all $\varepsilon \geq 0$, $\mathcal{D}_{\breve{\partial} g}^{\sharp, f^{\mathrm{FY}}}(x, x) \leq \varepsilon$.*

5.1. Some open problems related to GBDs

(a) When $T = S = \partial f$, will the GBD induced by $h = F_T$ have some advantages when compared with the classical Bregman distance (induced by $h = f^{\mathrm{FY}}$)? What do the resulting *generalized projections* look like, when compared with the classical Bregman projections?

(b) Can these distances be used to regularize/penalize proximal-like iterations for variational inequalities?

(c) In view of Proposition 31, can we use the GBDs to develop new algorithms for solving problem (5.3)?

(d) In view of Proposition 32, can we obtain an optimality condition for the inclusion problem $0 \in Tx - Sx$ with T, S maximally monotone?

(e) Can a similar result to Proposition 32 be obtained for enlargements of $T = \partial f$ different from $\breve{\partial} f$?

6. FINAL WORDS

Many crucial properties have been left uncovered, and it is my hope that those mentioned here will motivate researchers to explore the yet undiscovered paths that link maximally monotone operators with convex functions.

I conclude with a tribute to Asen Dontchev, who passed away on 16 September 2021. Asen was an outstanding mathematician with crucial contributions to set-valued analysis, and especially to the topic of inclusion problems and their approximations.

REFERENCES

[1] M. M. Alves and M. Geremia, Iteration complexity of an inexact Douglas–Rachford method and of a Douglas–Rachford–Tseng's F-B four-operator splitting method for solving monotone inclusions. *Numer. Algorithms* **82** (2019), 263–295.

[2] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*. Springer, New York, 2011.

[3] J. M. Borwein and S. Fitzpatrick, Local boundedness of monotone operators under minimal hypotheses. *Bull. Aust. Math. Soc.* **39** (1989), 439–441.

[4] L. M. Bregman, The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming. *USSR Comput. Math. Math. Phys.* **7** (1967), 200–217.

[5] A. Brøndsted and R. T. Rockafellar, On the subdifferentiability of convex functions. *Proc. Amer. Math. Soc.* **16** (1965), 605–611.

[6] R. S. Burachik, M. N. Dao, and S. B. Lindstrom, The generalized Bregman distance. *SIAM J. Optim.* **31** (2021), 404–424.

[7] R. S. Burachik, M. N. Dao, and S. B. Lindstrom, Generalized Bregman envelopes and proximity operators. *J. Optim. Theory Appl.* **190** (2021), 744–778.

[8] R. S. Burachik and J. Dutta, Inexact proximal point methods for variational inequality problems. *SIAM J. Optim.* **20** (2010), 2653–2678.

[9] R. S. Burachik and A. N. Iusem, *Set-valued mappings and enlargements of monotone operators*. Springer, 2008.

[10] R. S. Burachik, A. N. Iusem, and B. F. Svaiter, Enlargement of monotone operators with applications to variational inequalities. *Set-Valued Var. Anal.* **5** (1997), 159–180.

[11] R. S. Burachik, J. Lopes, and G. Da Silva, An inexact interior point proximal method for the variational inequality problem. *Comput. Appl. Math.* **28** (2009), 15–36.

[12] R. S. Burachik and J. E. Martínez-Legaz, On Bregman-type distances for convex functions and maximally monotone operators. *Set-Valued Var. Anal.* **26** (2018), 369–384.

[13] R. S. Burachik and J. E. Martínez-Legaz, On a sufficient condition for equality of two maximal monotone operators. *Set-Valued Var. Anal.* **18** (2020), 327–335.

[14] R. S. Burachik, J. E. Martínez-Legaz, M. Rezaie, and M. Théra, An additive subfamily of enlargements of a maximally monotone operator. *Set-Valued Var. Anal.* **23** (2015), 643–665.

[15] R. S. Burachik, C. A. Sagastizábal, and S. Sheimberg, An inexact method of partial inverses and a parallel bundle method. *Optim. Methods Softw.* **21** (2006), 385–400.

[16] R. S. Burachik, S. Sheimberg, and B. F. Svaiter, Robustness of the hybrid extragradient proximal-point algorithm. *J. Optim. Theory Appl.* **111** (2001), 117–136.

[17] R. S. Burachik and B. F. Svaiter, ε-enlargements of maximal monotone operators in Banach spaces. *Set-Valued Var. Anal.* **7** (1999), 117–132.

[18] R. S. Burachik and B. F. Svaiter, A relative error tolerance for a family of generalized proximal point methods. *Math. Oper. Res.* **26** (2001), 816–831.

[19] R. S. Burachik and B. F. Svaiter, Maximal monotone operators, convex functions and a special family of enlargements. *Set-Valued Var. Anal.* **10** (2002), 297–316.

[20] R. S. Burachik and B. F. Svaiter, Operating enlargements of monotone operators: new connection with convex functions. *Pac. J. Optim.* **2** (2006), 425–445.

[21] R. S. Burachik, C. A. Sagastizábal, and B. F. Svaiter, Bundle methods for maximal monotone operators. In *Ill-posed variational problems and regularization techniques*, edited by M. Théra and R. Tichatschke, pp. 49–64, Springer, Berlin, Heidelberg, 1999.

[22] J. Eckstein and W. Yao, Relative-error approximate versions of Douglas–Rachford splitting and special cases of the ADMM. *Math. Program.* **170** (2017), 417–444.

[23] S. Fitzpatrick, Representing monotone operators by convex functions. In *Workshop/miniconference on functional analysis and optimization*, pp. 59–65, Proc. Centre Math. Anal. Austral. Nat. Univ. 20, Austral. Nat. Univ., Canberra, 1988.

[24] S. D. Flåm, Monotonicity and market equilibrium. *Set-Valued Var. Anal.* **24** (2016), 403–421.

[25] N. Ghoussoub, A variational theory for monotone vector fields. *J. Fixed Point Theory Appl.* **4** (2008), no. 1, 107–135.

[26] J.-B. Hiriart-Urruty, Conditions for global optimality. In *Handbook of global optimization*, edited by R. Horst and P. M. Pardalos, pp. 1–26, Springer, 1995.

[27] N. Krylov, Properties of monotone mappings. *Lith. Math. J.* **22** (1982), 140–145.

[28] J. E. Martínez-Legaz and M. Théra, ε-subdifferentials in terms of subdifferentials. *Set-Valued Var. Anal.* **4** (1996), 327–332.

[29] J. Moreau, Proximité et dualité dans un espace Hilbertien. *Bull. Soc. Math. France* **93** (1965), 273–299.

[30] L. Nagesseur, A bundle method using two polyhedral approximations of the ϵ-enlargement of a maximal monotone operator. *Comput. Optim. Appl.* **64** (2016), 75–100.

[31] J.-P. Penot, The relevance of convex analysis for the study of monotonicity. *Nonlinear Anal.* **58** (2004), 855–871.

[32] T. Roche, R. Rossi, and U. Stefanelli, Stability results for doubly nonlinear differential inclusions by variational convergence. *SIAM J. Control Optim.* **52** (2014), no. 2, 1071–1107.

[33] R. T. Rockafellar, Characterization of the subdifferentials of convex functions. *Pacific J. Math.* **17** (1966), 497–510.

[34] R. T. Rockafellar, Local boundedness of nonlinear monotone operators. *Michigan Math. J.* **6** (1969), 397–407.

[35] R. T. Rockafellar, On the maximal monotonicity of subdifferential mappings. *Pacific J. Math.* **33** (1970), 209–216.

[36] R. T. Rockafellar, On the maximality of sums of nonlinear monotone operators. *Trans. Amer. Math. Soc.* **149** (1970), 75–88.

[37] S. Simons, *From Hahn–Banach to monotonicity. 2nd edn.*, Lecture Notes in Math., Springer, NY, 2008.

[38] B. F. Svaiter, A Family of Enlargements of Maximal Monotone Operators. *Set-Valued Var. Anal.* **8** (2000), 311–328.

[39] B. F. Svaiter, A weakly convergent fully inexact Douglas–Rachford method with relative error tolerance. *ESAIM Control Optim. Calc. Var.* **25** (2019), no. 57, 311–328.

[40] A. Visintin, Variational formulation and structural stability of monotone equations. *Calc. Var. Partial Differential Equations* **47** (2013), no. 1–2, 273–317.

[41] A. Visintin, An extension of the Fitzpatrick theory. *Commun. Pure Appl. Anal.* **13** (2014), no. 5, 2039–2058.

REGINA S. BURACHIK

University of South Australia, Adelaide, Australia, regina.burachik@unisa.edu.au