

Characterizing Maximal Monotone Operators with Unique Representation

Sotiris ARMENIAKOS & Aris DANIILIDIS

Abstract

We study maximal monotone operators $A : X \rightrightarrows X^*$ whose Fitzpatrick family reduces to a singleton; such operators will be called *uniquely representable*. We show that every such operator is cyclically monotone (hence, $A = \partial f$ for some convex function f) if and only if it is 3-monotone. In Radon–Nikodým spaces, under mild conditions (which become superfluous in finite dimensions), we prove that a subdifferential operator $A = \partial f$ is uniquely representable if and only if f is the sum of a support and an indicator function of suitable convex sets.

Keywords: Maximal monotone operator, Fitzpatrick function, subdifferential, Radon–Nikodým property.

AMS Classification: *Primary:* 46B20, 46N10, 47H05 ; *Secondary:* 49J53, 90C25.

1 Introduction

In this paper we denote by $(X, \|\cdot\|)$ a real Banach space, by X^* its dual space and by $\langle \cdot, \cdot \rangle$ the duality map between X and X^* . The graph of a set-valued operator $A : X \rightrightarrows X^*$ is denoted by

$$\text{Gr}(A) := \{(x, x^*) \in X \times X^* : x^* \in Ax\}.$$

An operator A is called monotone if for every $(x, x^*), (y, y^*) \in \text{Gr}(A)$, one has:

$$\langle x, x^* \rangle + \langle y, y^* \rangle \geq \langle x, y^* \rangle + \langle y, x^* \rangle.$$

A function $h : X \times X^* \rightarrow \mathbb{R} \cup \{+\infty\}$ is called a *representative function* of A if

(R_1) h is proper, convex and lower semicontinuous (in short, lsc);

(R_2) $h(x, x^*) \geq \langle x, x^* \rangle$, for all $(x, x^*) \in X \times X^*$;

(R_3) $h(x, x^*) = \langle x, x^* \rangle$, for all $(x, x^*) \in \text{Gr}(A)$.

We denote by $\mathcal{F}_A$ the family of all representative functions of the operator A and we will refer to it as the *Fitzpatrick* or the *representative* family of A . A classical example arises when $A = \partial f$ is the subdifferential of a proper convex lsc function f . In this case, the function

$$(x, x^*) \mapsto \Phi_f(x, x^*) = f(x) + f^*(x^*) \tag{1.1}$$

belongs to the Fitzpatrick family $\mathcal{F}_{\partial f}$ (this is a consequence of the Fenchel–Young inequality), where $f^* : X^* \rightarrow \mathbb{R} \cup \{+\infty\}$ denotes the convex conjugate of f . Recall the classical fact, due to

Rockafellar [24], that a maximal monotone operator is the subdifferential of a proper convex lsc function if and only if A is cyclically monotone (see Section 2 for the required preliminaries). In [13], Burachik and Svaiter established a new characterization of subdifferential operators in terms of the Fitzpatrick family: $A = \partial f$ for some proper convex lsc f if and only if $\mathcal{F}_A$ contains a separable representative function (that is, a function $h : X \times X^* \rightarrow \mathbb{R}$ of the form $h(x, x^*) = \phi(x) + \psi(x^*)$).

For maximal monotone operators, the so-called Fitzpatrick function

$$F_A(x, x^*) = \langle x, x^* \rangle - \inf_{(y, y^*) \in \text{Gr } A} \langle x - y, x^* - y^* \rangle \quad (1.2)$$

belongs to the Fitzpatrick family $\mathcal{F}_A$ and is in fact the pointwise minimum member of the family, see [18]. Moreover, if two maximal monotone operators A, B have a common representative function, then they are equal (that is, $\mathcal{F}_A \cap \mathcal{F}_B \neq \emptyset \implies A = B$).

For merely monotone operators, a different natural representative function can be constructed as follows. We first define the function

$$(x, x^*) \mapsto \phi(x, x^*) := \begin{cases} \langle x, x^* \rangle, & \text{if } (x, x^*) \in \text{Gr}(A) \\ +\infty, & \text{otherwise.} \end{cases}$$

Then the restriction on $X \times X^*$ of the function:

$$P_A(x, x^*) := \phi^{**}(x, x^*) \quad (1.3)$$

is a representative function, that is, $P_A \in \mathcal{F}_A$. In addition, P_A turns out to be the pointwise maximum member of the family $\mathcal{F}_A$ (see, e.g., [20]).

In particular, for a cyclically monotone operator (that is $A = \partial f$ for some proper convex, lsc function f) we have

$$\langle x, x^* \rangle \leq F_{\partial f}(x, x^*) \leq f(x) + f^*(x^*) \leq P_{\partial f}(x, x^*).$$

The above inequality reveals that the Fitzpatrick function $F_{\partial f}$ provides a more refined lower bound for the Fenchel-Young inequality. This interesting fact has been further explored in recent works initiated by Carlier, see [15, 5, 12].

Representation functions have been used extensively in the literature, mainly because they allow the use of convex analysis techniques in the study of monotone operators. Several applications can be found in optimization [10, 25, 14], machine learning [22], optimal transport [15], optimal control [19], stochastic differential equations [1] and Banach space theory [23].

Given a maximal monotone operator $A : X \rightrightarrows X^*$, one can generally construct more than one representative function and actually infinite, as the pointwise convex combination of any two representative functions is still a representative function. The operators for which the Fitzpatrick family collapses to a singleton will be called *uniquely representable*. These operators are in some sense singular. In this work, we address the following natural question:

Which maximal monotone operators are uniquely representable ?

This question has already been considered in the literature. For example in [2], it was shown that subdifferentials of proper lower semicontinuous sublinear functions, as well as indicator

functions of closed convex sets, are uniquely representable. We shall hereby extend this class of functions and obtain, under mild conditions, a complete characterization.

In [6], the case of linear monotone operators was studied and it was shown that such an operator is uniquely representable if and only if it is skew-symmetric. Since a linear monotone operator is a subdifferential if and only if it is symmetric, we deduce that the class of merely monotone operators contains operators which are far from being subdifferentials, yet enjoy a unique representative function. In striking contrast to the above, we shall see that imposing 3-monotonicity forces the operator to be cyclically monotone.

As a by-product of our analysis, we also suggest a new way to compute the Fitzpatrick function for subdifferential operators; see Lemma 4.9 and the subsequent remark.

Contributions.

In Section 3 we prove that a maximal monotone operator which is 3-monotone and has a unique representative function must be cyclically monotone. In other words:

Theorem A. *Let $A : X \rightrightarrows X^*$ be a maximal monotone operator such that $\mathcal{F}_A = \{F_A\}$. Then the following are equivalent:*

- (i). *A is 3-monotone ;*
- (ii). *A is cyclically monotone.*

More precisely, for every $v^ \in \text{Im} A$, the function $x \mapsto F_A(x, v^*)$ is proper, convex, lsc and*

$$A = \partial F_A(\cdot, v^*)$$

In Section 4 we focus our attention to the class of subdifferential operators (that is, $A = \partial f$) and we prove the following result:

Theorem B. *Let X be a space with the Radon–Nikodým property and $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex, lsc with $\text{int}(\text{dom } f) \neq \emptyset$ and $\text{int}(\text{dom } f^*) \neq \emptyset$. Then the Fitzpatrick family $\mathcal{F}_{\partial f}$ consists of a single element if and only if there exist a constant $c \in \mathbb{R}$, a functional $\bar{x}^* \in X^*$ and closed convex sets $K, C \subseteq X$ and $\mathcal{V} \subseteq X^*$ where*

- *C is a cone, $\mathcal{V}$ is w^* -closed convex*
- *$0 \in \mathcal{V} \perp K - K$.*

such that for every $\hat{x} \in K$ and $x \in X$

$$f(x) = \sigma_{\mathcal{V}}(x - \hat{x}) + i_{\overline{K+C}}(x) + \langle x, \bar{x}^* \rangle + c. \quad (1.4)$$

We prove that in finite dimensions the above technical assumptions on the domains of f and f^* can be dropped (see Theorem 4.18).

A classical result of Collier [16] guarantees that if X has the Radon–Nikodým property (in short, X is an RNP space), then the dual space X^* is w^* -Asplund, ensuring good differentiability properties for the conjugate functions f^* . This plays a central role in our approach. It is quite interesting to investigate whether the class of admissible spaces can be extended to classes of spaces where convex functions satisfy weaker (if any) differentiability properties.

Finally, without assumptions on X we generalize [2, Theorem 5.3]; see Theorem 4.12 and the subsequent remarks.

Organization of the paper. The paper is structured as follows. In Section 2, we recall fundamental results from convex analysis and differentiability theory in Banach spaces. Section 3 is devoted to the proof of our first main result (Theorem A) concerning 3-monotone operators and in Section 4 we focus on the particular case of subdifferential operators and prove several results leading eventually to Theorem B. We also discuss the finite dimensional case.

2 Preliminaries

In this section we fix our notation and recall notions and preliminary results that will be used throughout the paper.

Convex functions and subdifferentials. Let $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper convex function. A functional $x^* \in X^*$ is called an ε -subgradient of f at x if

$$f(y) - f(x) \geq \langle y - x, x^* \rangle - \varepsilon, \quad \text{for every } y \in \text{dom } f.$$

For $\varepsilon \geq 0$, the set of all ε -subgradients of f at x is denoted by $\partial_\varepsilon f(x)$ and called the ε -subdifferential of f at x . In the particular case that $\varepsilon = 0$, $\partial_0 f(x)$ is denoted simply by $\partial f(x)$ and called the subdifferential of f at x . The Fenchel–Legendre conjugate of f is the proper convex w^* -lsc function $f^*: X^* \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by

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

For every $(x, x^*) \in X \times X^*$, the Fenchel–Young inequality holds:

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

and ε -subgradients of f are characterized via f^* as follows

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

It is well known that if $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ is proper, convex, lsc then for any $x \in \text{int}(\text{dom } f)$ and $\varepsilon \geq 0$, the set of ε -subdifferentials is bounded. More precisely, there exists $M = M(x, f, \varepsilon)$ such that

$$\|x^*\| \leq M, \quad \text{for all } x^* \in \partial_\varepsilon f(x).$$

For a subset $A \subseteq X$, the indicator function i_A is defined as

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

For a set $B \subseteq X^*$, the support function σ_B is given by

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

Given a set $C \subseteq X^*$, the polar set $C^\circ \subseteq X$ is defined by

$$C^\circ := \{x \in X : \langle x, x^* \rangle \leq 1, \forall x^* \in C\}.$$

In the special case where C is a cone, one has

$$\sigma_C(x) = i_{C^\circ}(x).$$

Inf-convolution. Given two functions $f, g: E \rightarrow (-\infty, +\infty]$, where E is a real vector space, their *infimal convolution* is defined by

$$(f \square g)(x) := \inf_{y \in E} \{f(y) + g(x - y)\}, \quad x \in E.$$

For a function $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$, we define its lower semicontinuous envelope $\bar{f}$, as follows

$$\bar{f} := \sup\{g : g \text{ is lsc and } g \leq f\}, \quad (2.1)$$

that is, the greatest lower semicontinuous function dominated by f . We also recall the following identity:

$$(\bar{f \square g})^* = f^* + g^*, \quad (2.2)$$

for any two proper functions f, g .

Differentiability and RN spaces. A function $\phi: X \rightarrow Y$ is said to be *Fréchet differentiable* at x if there exists a linear bounded operator $A: X \rightarrow Y$ such that

$$\lim_{\|u\|_X \rightarrow 0} \frac{\|\phi(x + u) - \phi(x) - A(u)\|_Y}{\|u\|_X} = 0.$$

In this case, A is unique and denoted $D\phi(x)$ (the Fréchet derivative of f at x). In particular, for a proper convex $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$, a point x is called a point of Fréchet differentiability (or a *Fréchet point*) if the subdifferential mapping ∂f is single-valued and norm-to-norm upper semicontinuous at x . In this case, the Fréchet derivative agrees with the unique subgradient.

Convex functions exhibit strong regularity properties. In the finite-dimensional case, the Rademacher theorem ensures that every locally Lipschitz function (in particular every convex function) is differentiable almost everywhere. In the infinite-dimensional setting, convex functions retain partial differentiability properties on a large set if the underlying space is an *Asplund space*. A Banach space X is called Asplund if every convex continuous function defined on an open convex subset of X is Fréchet differentiable on a dense G_δ subset of its domain. Notably, reflexive spaces and Banach spaces with separable duals are Asplund.

Let us further recall that a *slice* $S(x^*, A, a)$ in a Banach space X is defined for $A \subset X$, $a > 0$ and $x^* \in X^*$ by

$$S(x^*, A, a) = \{x \in A : \langle x, x^* \rangle > \sigma_A(x^*) - a\}.$$

A nonempty set $A \subseteq X$ is called *dentable* if it admits slices of arbitrarily small diameter. A Banach space X is said to have the RNP (or called an RN space) if every nonempty bounded subset of A is dentable. One of the central results in the differentiability theory of convex functions is the fact that

$$X \text{ is Asplund} \iff X^* \text{ has the RNP.}$$

In this work we are going to use a dual version of the above due to Collier [16]

$$X \text{ has the RNP} \iff X^* \text{ is } w^*\text{-Asplund.}$$

A dual space X^* is called w^* -Asplund if every convex, continuous and w^* -lsc function on X^* is Fréchet differentiable on a G_δ dense subset of its domain. We remark further that if $\hat{a} = Df^*(v^*)$ for a Fréchet point v^* of f^* , then $\hat{a} \in X$. For further background on convex analysis and these functional-analytic preliminaries, see [17, 21, 26].

Monotone operators and representative functions. For an operator $A: X \rightrightarrows X^*$, we denote the domain by $\text{dom}(A) = \{x \in X : Ax \neq \emptyset\}$ and the range by

$$\text{Im}(A) = \{x^* \in X^* : \exists x \in X \text{ s.t. } x^* \in Ax\}.$$

An operator $A: X \rightrightarrows X^*$ is said to be n -cyclically monotone (or simply n -monotone) if for any collection $\{(a_i, a_i^*)\}_{i=1}^n \subseteq \text{Gr}(A)$ we have

$$\sum_{i=1}^n \langle a_i, a_i^* \rangle \geq \sum_{i=1}^n \langle a_{i+1}, a_i^* \rangle,$$

with the convention $a_{n+1} = a_1$. The operator A is called cyclically monotone if it is n -monotone for every $n \in \mathbb{N}$. It is called *monotone* if it is 2-monotone, which is equivalent to the condition

$$\langle x - y, x^* - y^* \rangle \geq 0, \quad \text{for every } (x, x^*), (y, y^*) \in \text{Gr}(A).$$

A monotone (resp. cyclically monotone) operator is called maximal monotone (resp. maximal cyclically monotone) if there is no strict extension that preserves monotonicity (resp. cyclic monotonicity). It is well-known that subdifferentials of proper convex lsc functions are maximal cyclically monotone operators.

The *Fitzpatrick function* associated to an operator A is the function $F_A: X \times X^* \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by (1.2), or equivalently by

$$F_A(x, x^*) := \sup_{(y, y^*) \in \text{Gr}(A)} \{\langle y, x^* \rangle + \langle x, y^* \rangle - \langle y, y^* \rangle\}. \quad (2.3)$$

Recall also the function P_A , defined by (1.3). The following properties are folklore, see e.g. [18, 20]. If $A: X \rightrightarrows X^*$ is maximal monotone, then $F_A \in \mathcal{F}_A$, and F_A is the pointwise minimum of $\mathcal{F}_A$, satisfying

$$F_A(x, x^*) = \langle x, x^* \rangle \iff (x, x^*) \in \text{Gr}(A).$$

For any monotone operator A , we have $P_A \in \mathcal{F}_A$. Moreover, P_A is the pointwise maximal member of $\mathcal{F}_A$. In addition, for $x \in X$ and $x^* \in X^*$ we have:

$$P_A(x, x^*) = F_A^*(x^*, x) = \sup_{(y, y^*) \in X \times X^*} \{\langle y, x^* \rangle + \langle x, y^* \rangle - F_A(y, y^*)\}. \quad (2.4)$$

In addition, P_A satisfies the following inequality for all $(x, x^*), (y, y^*) \in X \times X^*$ (see [8, Proposition 2.7])

$$P_A(x, x^*) + P_A(y, y^*) \geq \langle x, y^* \rangle + \langle y, x^* \rangle. \quad (2.5)$$

Two useful facts. We finish this section by mentioning two auxiliary results that will be used in the sequel. The first one follows from [7, Proposition 3.6] and [7, Corollary 3.7].

Fact 2.1. Let $f, g : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex and lsc such that for all $x \in X$:

$$f(x) \geq g(x)$$

Then:

(i). If X is finite dimensional and $f(x) = g(x)$ for a dense subset of $\text{rint}(\text{dom } g)$ then:

$$f(x) = g(x) \quad \text{for all } x \in X.$$

(ii). If $\text{int}(\text{dom } g) \neq \emptyset$ and $f(x) = g(x)$ for a dense subset of $\text{int}(\text{dom } g)$, then

$$f(x) = g(x) \quad \text{for all } x \in X.$$

We mention for completeness that the assumption $\text{int}(\text{dom } g) \neq \emptyset$ is essential in (ii) above, see [7, Proposition 3.4].

The second fact is a structural result for subgradients of a convex function. The result can be deduced from [9, Corollary 4.10].

Fact 2.2. Let $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex, lsc and $\text{int}(\text{dom } f) \neq \emptyset$. Then for every $x \in \text{dom } \partial f$ and $x^* \in \partial f(x)$, there exists $x_{\text{int}}^* \in \overline{\text{conv}}^{w*}(\bigcup_{y \in \text{int}(\text{dom } f)} \partial f(y))$ and $x_N^* \in N_{\text{dom } f}(x)$ such that:

$$x^* = x_{\text{int}}^* + x_N^*.$$

3 Case of Maximal Monotone Operators

The aim of this section is to establish Theorem A, asserting that every maximal monotone, 3-monotone operator with $\mathcal{F}_A$ singleton is in fact a subdifferential. To this end, let us first present some general properties that a maximal monotone operator A should necessarily satisfy if the Fitzpatrick family $\mathcal{F}_A$ is a singleton. We begin by the following observation.

Lemma 3.1. For a maximal monotone operator $A : X \rightrightarrows X^*$ the following are equivalent:

- (i). $\mathcal{F}_A$ is a singleton, i.e. $\mathcal{F}_A = \{F_A\}$.
- (ii). $F_A \equiv P_A$.
- (iii). For every $(x, x^*), (y, y^*) \in X \times X^*$

$$F_A(x, x^*) + F_A(y, y^*) \geq \langle x, y^* \rangle + \langle y, x^* \rangle. \quad (3.1)$$

If $A = \partial f$ for some proper convex lsc function $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$, then the above assertions are also equivalent to the following property, evoking the function Φ_f from (1.1):

- (iv). For every $(x, x^*) \in X \times X^*$

$$F_{\partial f}(x, x^*) = \Phi_f(x, x^*).$$

Proof. Since F_A is the minimal and P_A is the maximal member of $\mathcal{F}_A$, the equivalence between (i) and (ii) is straightforward. The implication (ii) $\implies$ (iii) follows from (2.5). Furthermore (3.1) together with (2.4) implies

$$F_A(x, x^*) \geq \sup_{(y, y^*) \in X \times X^*} \{\langle y, x^* \rangle + \langle x, y^* \rangle - F_A(y, y^*)\} = P_A(x, x^*)$$

in virtue of (2.4). Since we always have $F_A \leq P_A$, we deduce that (ii) holds.

Let us now assume that $A = \partial f$ for a proper convex lsc function $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$. Since both functions $\Phi_f, F_{\partial f}$ belong to the family $\mathcal{F}_{\partial f}$ we clearly have that (i) $\implies$ (iv). We shall show that (iv) $\implies$ (iii). Indeed, if (iv) holds, then in virtue of the Fenchel–Young inequality we obtain:

$$F_{\partial f}(x, x^*) + F_{\partial f}(y, y^*) = f(x) + f^*(y^*) + f(y) + f^*(x^*) \geq \langle x, y^* \rangle + \langle y, x^* \rangle$$

concluding the proof. $\square$

We proceed by observing that the class of operators for which $\mathcal{F}_A = \{F_A\}$ is invariant under both translations and dilations of the graph.

Proposition 3.2 (Translation and Dilation Invariance). *Let $A : X \rightrightarrows X^*$ be a maximal monotone operator such that $\mathcal{F}_A = \{F_A\}$. Then for every $\lambda_1, \lambda_2 > 0$ and $(w, w^*) \in X \times X^*$, the operator $\hat{A}$ with graph*

$$\text{Gr}(\hat{A}) = \{ (\lambda_1 x - w, \lambda_2 x^* - w^*) : (x, x^*) \in \text{Gr}(A) \}$$

is also maximal monotone and satisfies $\mathcal{F}_{\hat{A}} = \{F_{\hat{A}}\}$.

Proof. By the previous lemma, we have

$$F_A(x, x^*) + F_A(y, y^*) \geq \langle x, y^* \rangle + \langle y, x^* \rangle, \quad \text{for every } (x, x^*), (y, y^*) \in X \times X^*.$$

Let us fix $(w, w^*) \in X \times X^*$ and define A_1 by

$$\text{Gr}(A_1) := \text{Gr}(A) - \{(w, w^*)\}.$$

A direct calculation shows that

$$\begin{aligned} & F_{A_1}(x - w, x^* - w^*) + F_{A_1}(y - w, y^* - w^*) \\ &= F_A(x, x^*) + F_A(y, y^*) - \langle x + y, w^* \rangle - \langle w, x^* + y^* \rangle + 2\langle w, w^* \rangle. \end{aligned}$$

Using (3.1), the above expression is bounded from below by

$$\langle x - w, y^* - w^* \rangle + \langle y - w, x^* - w^* \rangle,$$

so in view of Lemma 3.1, in particular from (iii) $\implies$ (i), we have that $\mathcal{F}_{A_1} = \{F_{A_1}\}$. Fix now $\lambda_1, \lambda_2 > 0$ and define

$$\text{Gr}(A_2) := \{(\lambda_1 a, \lambda_2 a^*) : (a, a^*) \in \text{Gr}(A)\}.$$

Then

$$F_{A_2}(\lambda_1 x, \lambda_2 x^*) = \lambda_1 \lambda_2 F_A(x, x^*).$$

Therefore, if the operator A satisfies the inequality (3.1), so does the operator A_2 .

The assertion follows by combining the above cases. $\square$

We now introduce marginal functions associated to representative functions.

Lemma 3.3. *Let $A : X \rightrightarrows X^*$ be a monotone operator. For every $w \in \text{dom}(A)$ and $v^* \in \text{Im}(A)$ the marginal functions $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ and $g : X^* \rightarrow \mathbb{R} \cup \{+\infty\}$ defined through:*

$$f(x) = f_{A,w}(x) := \inf_{a^* \in X^*} \{P_A(x, a^*) - \langle w, a^* \rangle\},$$

and

$$g(x^*) = g_{A,v^*}(x^*) := \inf_{a \in X} \{P_A(a, x^*) - \langle a, v^* \rangle\}.$$

are proper convex and lsc.

Proof. We first prove that f is proper. Indeed, since $A(w) \neq \emptyset$, we can consider $w^* \in A(w)$, then since P_A is a representative function we have

$$P_A(w, w^*) = \langle w, w^* \rangle$$

and we deduce from (2.5) that

$$\begin{aligned} P_A(x, a^*) - \langle w, a^* \rangle &= P_A(x, a^*) + \underbrace{P_A(w, w^*) - \langle w, w^* \rangle}_{=0} - \langle w, a^* \rangle \\ &\geq \langle x, w^* \rangle - \langle w, w^* \rangle. \end{aligned}$$

Consequently $f(x) > -\infty$ for all $x \in X$. Moreover, f is convex and lsc, as a marginal function of the (jointly) convex function:

$$(x, a^*) \mapsto P_A(x, a^*) - \langle w, a^* \rangle$$

The assertion for g follows similarly. □

The above functions are quite useful in the study of maximal monotone operators. A variant appears in [8] to prove a central case of the Debrunner–Flor theorem. More importantly in our work, the duals of these functions interpolate between $F_A(\cdot, v^*)$ and $P_A(\cdot, v^*)$:

Lemma 3.4. *Let A be a maximal monotone operator. For every $w \in \text{dom}(A)$, $v^* \in \text{Im}(A)$ and $(x, x^*) \in X \times X^*$ we have:*

- (i). $F_A(x, v^*) \leq g_{A,v^*}^*(x) \leq P_A(x, v^*)$;
- (ii). $F_A(w, x^*) \leq f_{A,w}^*(x^*) \leq P_A(w, x^*)$.

Proof. (i). From (2.5), for all $a \in X$ and $x^* \in X^*$,

$$\begin{aligned} P_A(x, v^*) + P_A(a, x^*) &\geq \langle x, x^* \rangle + \langle a, v^* \rangle \implies \\ P_A(x, v^*) + \underbrace{\inf_{a \in X} \{P_A(a, x^*) - \langle a, v^* \rangle\}}_{g_{A,v^*}} &\geq \langle x, x^* \rangle \implies \\ P_A(x, v^*) &\geq \sup_{x^* \in X^*} (\langle x, x^* \rangle - g_{A,v^*}(x^*)) = g_{A,v^*}^*(x). \end{aligned}$$

On the other hand,

$$\begin{aligned}
F_A(x, v^*) &= \sup_{(a, a^*) \in \text{Gr}(A)} \{ \langle x, a^* \rangle + \langle a, v^* \rangle - \langle a, a^* \rangle \} \\
&= \sup_{(a, a^*) \in \text{Gr}(A)} \{ \langle x, a^* \rangle + \langle a, v^* \rangle - P_A(a, a^*) \} \\
&\leq \sup_{(a, a^*) \in \text{Gr}(A)} \left\{ \langle x, a^* \rangle + \sup_{a \in X} (\langle a, v^* \rangle - P_A(a, a^*)) \right\} \\
&= \sup_{(a, a^*) \in \text{Gr}(A)} \left\{ \langle x, a^* \rangle - \inf_{a \in X} (P_A(a, a^*) - \langle a, v^* \rangle) \right\} \\
&= \sup_{(a, a^*) \in \text{Gr}(A)} \{ \langle x, a^* \rangle - g_{A, v^*}(a^*) \} \leq \sup_{a^* \in X^*} \{ \langle x, a^* \rangle - g_{A, v^*}(a^*) \} = g_{A, v^*}^*(x).
\end{aligned}$$

(ii). It follows analogously. $\square$

As an immediate corollary we deduce the following.

Corollary 3.5. *Let $A : X \rightrightarrows X^*$ be a maximal monotone operator such that $\mathcal{F}_A = \{F_A\}$. Then for every $w \in \text{dom}(A)$ and $v^* \in \text{Im}(A)$ we have:*

- (i). $F_A(x, v^*) = g_{A, v^*}^*(x) = P_A(x, v^*)$;
- (ii). $F_A(w, x^*) = f_{A, w}^*(x^*) = P_A(w, x^*)$.

Proof. It follows directly from Lemma 3.4 and Lemma 3.1 (ii). $\square$

With this in hand, we are ready to prove one of the two main results of this work.

Proof of Theorem A.

(i) $\implies$ (ii). Let $(x, x^*) \in \text{Gr}(A)$ and $v^* \in \text{Im}(A)$ be arbitrary. By the Fenchel-Young inequality and the previous corollary we get

$$F_A(x, v^*) + g_{A, v^*}(x^*) = g_{A, v^*}^*(x) + g_{A, v^*}(x^*) \geq \langle x, x^* \rangle. \quad (3.2)$$

We need to show that equality holds, which directly yields that $x^* \in \partial g_{A, v^*}^*(x) = \partial F_A(\cdot, v^*)$. Arguing by contradiction, assume that the inequality is strict, that is,

$$F_A(x, v^*) + g_{A, v^*}(x^*) > \langle x, x^* \rangle.$$

Then, there would exist $(y, y^*) \in \text{Gr}(A)$ such that

$$\langle x, y^* \rangle + \langle y, v^* \rangle + g_{A, v^*}(x^*) > \langle x, x^* \rangle + \langle y, y^* \rangle.$$

But since $g_{A, v^*}(x^*) \leq P_A(y, x^*) - \langle y, v^* \rangle = F_A(y, x^*) - \langle y, v^* \rangle$, it would follow that

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

Thus there would exist $(z, z^*) \in \text{Gr}(A)$ such that

$$\langle x, y^* \rangle + \langle y, z^* \rangle + \langle z, x^* \rangle > \langle x, x^* \rangle + \langle y, y^* \rangle + \langle z, z^* \rangle,$$

contradicting the 3-monotonicity of A . Therefore in (3.2) equality should hold, namely,

$$F_A(x, v^*) + g_{A, v^*}(x^*) = \langle x, x^* \rangle,$$

which implies $x^* \in \partial F_A(\cdot, v^*)(x)$. Since the pair (x, x^*) was arbitrary in $\text{Gr}(A)$ and both operators A and $\partial F_A(\cdot, v^*)$ are maximal monotone, we conclude

$$A = \partial F_A(\cdot, v^*).$$

(ii) $\implies$ (i). If A is cyclically monotone, then clearly A is 3-monotone. For the "more precisely part", in virtue of the Rockafellar characterization theorem, $A = \partial f$ for a proper convex lsc function f . Fix now $v^* \in \text{Im}(A)$, so that $f^*(v^*) < +\infty$. Since $\mathcal{F}_A$ is a singleton, by Lemma 3.1(iv) we have

$$\Phi_f(x, v^*) = f(x) + f^*(v^*) = F_A(x, v^*), \quad \text{for all } x \in \text{dom } f.$$

Since v^* is fixed

$$A(x) = \partial f(x) = \partial(f(x) + f^*(v^*)) = \partial F_A(\cdot, v^*)(x),$$

that is,

$$A = \partial F_A(\cdot, v^*)$$

for all $v^* \in \text{Im}(A)$, which concludes the proof. $\square$

Remark 3.6. (i). It was proven in [6] that a linear monotone operator A is skew-symmetric if and only if $\mathcal{F}_A$ is a singleton. As a consequence, 3-monotonicity is a crucial assumption in the statement of Theorem A.

(ii). The notion of *paramonotone* (respectively, *3*-monotone*) operator interpolates between between the classes of monotone and 3-monotone operators, see [3, Chapter 22] and [11] for relevant definitions and an exposition. It is reasonable to conjecture that paramonotone (resp. 3*-monotone) maximal monotone operators with $\mathcal{F}_A$ singleton are necessarily subdifferentials. Unfortunately, our current techniques do not provide an answer to this question.

4 Case of Subdifferential Operators

4.1 Structural Properties of Subdifferentials with Unique Representative Function

By virtue of the previous section, we now focus on subdifferential operators $A = \partial f$ where function $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ is proper convex lsc. As mentioned in the introduction, the function $\Phi_f : X \times X^* \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by:

$$(x, x^*) \mapsto \Phi_f(x, x^*) = f(x) + f^*(x^*)$$

belongs to $\mathcal{F}_{\partial f}$ and consequently $\Phi_f \geq F_{\partial f}$ on $X \times X^*$. For any monotone operator, we define the following set-valued mapping $\mathcal{M}_A : X \times X^* \rightrightarrows \text{Gr}(A)$,

$$\mathcal{M}_A(x, x^*) = \{(a, a^*) \in \text{Gr}(A) : F_A(x, x^*) = \langle x, a^* \rangle + \langle a, x^* \rangle - \langle a, a^* \rangle\} \quad (4.1)$$

that is, the set of all points of $\text{Gr}(A)$ realizing the supremum in (2.3). In our case of interest we have the following result.

Proposition 4.1. *Let f be proper convex lsc and assume that $F_{\partial f}(x, x^*) = \Phi_f(x, x^*)$ for some $(x, x^*) \in \text{dom } \Phi_f$. Then for every $(a, a^*) \in \mathcal{M}_{\partial f}(x, x^*)$ we have:*

$$a^* \in \partial f(x) \cap \partial f(a) \quad \text{and} \quad a \in \partial f^*(x^*) \cap \partial f^*(a^*)$$

Proof. in view of (4.1) for every $(a, a^*) \in \mathcal{M}_{\partial f}(x, x^*)$ we have

$$F_{\partial f}(x, x^*) + \langle a, a^* \rangle = \langle a, x^* \rangle + \langle x, a^* \rangle.$$

Since $F_{\partial f}(x, x^*) = \Phi_f(x, x^*) = f(x) + f^*(x^*)$ and $f(a) + f^*(a^*) = \langle a, a^* \rangle$, we deduce:

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

Using the Fenchel–Young inequality we infer from the above that

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

and

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

Since $a^* \in \partial f(a)$, the conclusion follows. $\square$

The required technical tool to proceed is the following lemma.

Lemma 4.2. *Let $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex and lsc. Let $x \in \text{int}(\text{dom } f)$, $v^* \in \text{dom}(Df^*)$ and define $\hat{a} := Df^*(v^*)$. Assume that $F_{\partial f}(x, v^*) = \Phi_f(x, v^*)$. Then:*

- (i). $\mathcal{M}_{\partial f}(x, v^*) \neq \emptyset$
- (ii). *For every $(a, a^*) \in \mathcal{M}_{\partial f}(x, v^*)$ we have:*

$$a = \hat{a} = Df^*(v^*) \quad \text{and} \quad a^* \in \partial f(x) \cap \partial f(\hat{a}).$$

Proof. (i). Since $(x, v^*) \in \text{dom } f \times \text{dom } f^* = \text{dom } \Phi_f$ we have

$$F_{\partial f}(x, v^*) = \Phi_f(x, v^*) < +\infty.$$

Consider a positive sequence $\{\varepsilon_n\}_{n \in \mathbb{N}}$ with $\varepsilon_n \rightarrow 0$ and pick $\{a_n, a_n^*\}_{n \in \mathbb{N}} \subseteq \text{Gr}(\partial f)$ such that

$$F_{\partial f}(x, v^*) \leq \langle x, a_n^* \rangle + \langle a_n, v^* \rangle - \langle a_n, a_n^* \rangle + \varepsilon_n.$$

Using the assumption that $F_{\partial f}(x, v^*) = \Phi_f(x, v^*)$ and that $(a_n, a_n^*) \in \text{Gr}(\partial f)$, we may rewrite the above as

$$f(x) + f^*(v^*) + f(a_n) + f^*(a_n^*) \leq \langle x, a_n^* \rangle + \langle a_n, v^* \rangle + \varepsilon_n. \quad (4.2)$$

The above by the Fenchel–Young inequality (4.2) yields

$$f(a_n) + f^*(v^*) \leq \langle a_n, v^* \rangle + \varepsilon_n$$

and thus $a_n \in \partial_{\varepsilon_n} f^*(v^*)$. By Brønsted–Rockafellar ([21, Theorem 3.17]) we may find a sequence $\{b_n, b_n^*\}_{n \in \mathbb{N}} \subseteq \text{Gr}(\partial f)$ with

$$\max \{\|a_n - b_n\|, \|v^* - b_n^*\|\} \leq \sqrt{\varepsilon_n}. \quad (4.3)$$

Therefore $v^* = \lim_{n \rightarrow \infty} b_n^*$, in norm. Since ∂f^* is norm-to-norm upper semicontinuous at $(v^*, Df^*(v^*)) = (v^*, \hat{a})$ we get $b_n \xrightarrow{\|\cdot\|} \hat{a}$ and consequently by (4.3)

$$a_n \xrightarrow{\|\cdot\|} \hat{a}.$$

Furthermore, inequality (4.2) also yields

$$f(x) + f^*(a_n^*) \leq \langle x, a_n^* \rangle + \varepsilon_n \quad (4.4)$$

so $a_n^* \in \partial_{\varepsilon_n} f(x)$. As $x \in \text{int}(\text{dom } f)$, there exists $M > 0$ such that $\{a_n\}_{n \in \mathbb{N}} \in B(0, M)$. Let $a^* \in B(0, M)$ be a w^* -cluster point of $\{a_n^*\}_{n \in \mathbb{N}}$. Then

$$(\hat{a}, a^*) \in \overline{\{(a_n, a_n^*)\}_{n \geq 1}}^{(\|\cdot\|, w^*)} \subseteq \text{Gr}(\partial f),$$

because as $\text{int}(\text{dom } f) \neq \emptyset$, the graph of ∂f is $(\|\cdot\|, w^*)$ -closed and consequently $a^* \in \partial f(\hat{a})$. Moreover, since for every $n \in \mathbb{N}$

$$F_{\partial f}(x, v^*) \leq \langle x, a_n^* \rangle + \langle a_n, v^* \rangle - \langle a_n, a_n^* \rangle + \varepsilon_n$$

we deduce that $(\hat{a}, a^*) \in \mathcal{M}_{\partial f}(x, v^*)$ and clearly $\mathcal{M}_{\partial f}(x, v^*) \neq \emptyset$, which proves the first assertion of the lemma.

(ii). It follows from the previous proposition and the fact that $\partial f^*(v^*) = \{Df^*(v^*)\}$. $\square$

Before proceeding, let us recall that if v^* is a Fréchet point of f^* then

$$Df^*(v^*) \in \text{dom } f \subseteq X$$

Let us also denote by $\pi_{X^*} : X \times X^* \rightarrow X^*$ the canonical projection onto X^* . For convenience, given a proper convex lsc function $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ we define the sets

$$\mathcal{Z} := \text{Im}(Df^*) \subseteq \text{dom } f \quad \text{and} \quad K := \overline{\text{conv}}(\mathcal{Z}). \quad (4.5)$$

The previous results yield the following corollary in case ∂f has a unique representative, (i.e. $\mathcal{F}_{\partial f}$ is a singleton and Lemma 3.1 applies).

Corollary 4.3. *Let $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex and lsc. Assume that $\mathcal{F}_{\partial f}$ is a singleton and $\text{dom}(Df^*) \neq \emptyset$. Then for every $x \in \text{int}(\text{dom } f)$ and for every $\hat{a} \in \mathcal{Z} := \text{Im}(Df^*)$:*

$$\partial f(x) \cap \partial f(\hat{a}) \neq \emptyset.$$

In particular:

$$\mathcal{V} := \overline{\text{conv}}^{w^*} \left(\bigcup_{x \in \text{int}(\text{dom } f)} \partial f(x) \right) \subseteq \bigcap_{\hat{x} \in \text{conv}(\mathcal{Z})} \partial f(\hat{x}) \quad (4.6)$$

Proof. Let $x \in \text{int}(\text{dom } f)$ and fix $\hat{a} \in \mathcal{Z} \subseteq \text{dom } f$. Take $\tilde{x} \in \text{int}(\text{dom } f)$ such that

$$x \in [\hat{a}, \tilde{x}] \quad (4.7)$$

By Lemma 4.2, $\mathcal{M}_{\partial f}(x, v^*) \neq \emptyset$, where v^* is such that $Df^*(v^*) = \hat{a}$ and thus

$$\partial f(\tilde{x}) \cap \partial f(\hat{a}) \supseteq \pi_{X^*}(\mathcal{M}_{\partial f}(x, v^*)) \neq \emptyset. \quad (4.8)$$

It follows from (4.7) and (4.8) that

$$\partial f(x) = \partial f(\tilde{x}) \cap \partial f(\hat{a}) \subseteq \partial f(\hat{a}).$$

Therefore

$$\left(\bigcup_{x \in \text{int}(\text{dom } f)} \partial f(x) \right) \subseteq \partial f(\hat{a}), \text{ for every } \hat{a} \in \mathcal{Z}.$$

Finally, since $\partial f(\hat{a})$ is w^* -closed and convex, it follows that $\mathcal{V} \subseteq \partial f(\hat{a})$ for every $\hat{a} \in \mathcal{Z}$. Furthermore, as $\mathcal{Z} \subseteq \text{dom } f$ and $\text{dom } f$ is convex we deduce $\text{conv}(\mathcal{Z}) \subseteq \text{dom } f$. Now, consider $\hat{x} \in \text{conv}(\mathcal{Z})$, that is

$$\hat{x} = \sum_{i=1}^N \lambda_i \hat{a}_i, \quad \lambda_i \geq 0, \quad \sum_{i=1}^N \lambda_i = 1, \quad \hat{a}_i \in \mathcal{Z}.$$

Since $\bigcap_{i=1}^n \partial f(\hat{a}_i) \supseteq \mathcal{V} \neq \emptyset$ by [3, Proposition 22.10] it follows that:

$$f(\hat{x}) = \bigcap_{i=1}^n \partial f(\hat{a}_i).$$

Since $\hat{x}$ was arbitrarily chosen we conclude that

$$\mathcal{V} \subseteq \bigcap_{\hat{x} \in \text{conv}(\mathcal{Z})} \partial f(\hat{x}).$$

This proves the assertion. □

Recalling the notion from (4.5) we have the following formula describing f at interior points.

Proposition 4.4. *Let $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper convex and lsc function such that $\mathcal{F}_{\partial f}$ is a singleton. Assume that $0 \in \mathcal{V}$ and $\text{dom}(Df^*) \neq \emptyset$. Then*

$$K \subseteq \text{argmin } f \subseteq \text{dom } f \quad \text{and} \quad \mathcal{V} \perp K - K.$$

Moreover, for every $x \in \text{int}(\text{dom } f)$, $x^* \in \partial f(x)$ and $\hat{x} \in K$

$$f(x) = \langle x - \hat{x}, x^* \rangle + \min f.$$

Proof. By Corollary 4.3 we have $0 \in \partial f(\hat{x})$ for all $\hat{x} \in \mathcal{Z}$, thus $\mathcal{Z} \subseteq \text{argmin } f$. Since $\text{argmin } f$ is closed and convex, we deduce

$$\text{conv}(\mathcal{Z}) \subseteq \text{argmin } f$$

From (4.6) and the fact that $\text{Gr}(\partial f)$ is $(\|\cdot\|, w^*)$ -closed,

$$\mathcal{V} \subseteq \bigcap_{\hat{x} \in K} \partial f(\hat{x}).$$

Since $K \subseteq \text{argmin } f$, we have $f(\hat{x}) = \min f$, for all $\hat{x} \in K$. Let us now fix an arbitrary $\hat{x} \in K$ and pick any $x \in \text{int}(\text{dom } f)$ and $x^* \in \partial f(x) \neq \emptyset$. Then by the above inclusion we also have $x^* \in \partial f(\hat{x})$ and consequently

$$f(\hat{x}) + f^*(x^*) = \langle \hat{x}, x^* \rangle \quad \text{and} \quad f^*(x) + f^*(x^*) = \langle x, x^* \rangle. \quad (4.9)$$

Combining the above we deduce:

$$f(x) = \langle x, x^* \rangle - f^*(x^*) = \langle x, x^* \rangle - \langle \hat{x}, x^* \rangle + f(\hat{x}) = \langle x - \hat{x}, x^* \rangle + \min f$$

and the formula for f follows. Now, let $\hat{x}, \hat{y} \in K$. We deduce from the above that for every $x \in \text{int}(\text{dom } f)$ and $x^* \in \partial f(x)$ we have

$$f(x) = \langle x - \hat{x}, x^* \rangle + \min f = \langle x - \hat{y}, x^* \rangle + \min f$$

deducing

$$\langle \hat{x} - \hat{y}, x^* \rangle = 0.$$

Hence $\langle \hat{x} - \hat{y}, x^* \rangle = 0$ for all $\hat{x}, \hat{y} \in K$ and $x^* \in \bigcup_{x \in \text{int}(\text{dom } f)} \partial f(x)$, which yields $\mathcal{V} \perp K - K$. $\square$

We shall also need the following lemma.

Lemma 4.5. *Let $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper lsc convex function. Assume $\text{int}(\text{dom } f) \neq \emptyset$, $\text{dom}(Df^*) \neq \emptyset$ and that $F_{\partial f} = \Phi_f$ on $\text{dom } f \times \text{dom } f^* = \text{dom } \Phi_f$. Then:*

$$\text{dom}(Df^*) \subseteq \partial f(K) = \mathcal{V} + \hat{N} \subseteq \text{dom } f^*$$

where:

$$\hat{N} = \bigcup_{\hat{x} \in K} N_{\text{dom } f}(\hat{x}). \quad (4.10)$$

and $K, \mathcal{V}$ are defined by (4.5) and (4.6) respectively.

Proof. It follows readily from (4.5) that $\text{dom}(Df^*) \subseteq \partial f(K)$. By Corollary 4.3 for every $\hat{x} \in K$, we have $\mathcal{V} \subseteq \partial f(\hat{x})$. Since $\partial f(\hat{x}) = \partial f(\hat{x}) + N_{\text{dom } f}(\hat{x})$ we deduce that $\partial f(\hat{x}) \supseteq \mathcal{V} + N_{\text{dom } f}(\hat{x})$. By Fact 2.2 we obtain that $\partial f(\hat{x}) \subseteq \mathcal{V} + N_{\text{dom } f}(\hat{x})$ and consequently

$$\partial f(\hat{x}) = \mathcal{V} + N_{\text{dom } f}(\hat{x})$$

As $\hat{x} \in K$ is arbitrary, we conclude that:

$$\partial f(K) = \bigcup_{\hat{x} \in K} \partial f(\hat{x}) = \mathcal{V} + \bigcup_{\hat{x} \in K} N_{\text{dom } f}(\hat{x}) = \mathcal{V} + \hat{N}$$

This completes the proof. $\square$

With this in hand, we are ready to provide a formula for f^* in terms of $K, \mathcal{V}, \hat{N}$.

Proposition 4.6. *Let $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper convex and lsc. Assume that $\text{dom}(Df^*)$ is densely contained in $\text{dom } f^*$, $\text{int}(\text{dom } f) \cap \text{argmin } f \neq \emptyset$ and $\mathcal{F}_A$ is a singleton. Then for every $x^* \in X^*$ we have :*

$$f^*(x^*) = \sigma_K(x^*) + i_{\overline{\mathcal{V} + \hat{N}}}(x^*) - \min f \quad (4.11)$$

where $N = \overline{\text{conv}} \hat{N}$ and $K, \mathcal{V}, \hat{N}$ are defined by (4.5), (4.6) and (4.10) respectively.

Proof. Define $g : X^* \rightarrow \mathbb{R} \cup \{+\infty\}$ by

$$g(x^*) = \sigma_K(x^*) + i_{\overline{\mathcal{V} + \hat{N}}}(x^*) - \min f$$

Clearly g is proper, convex and lsc. Since $\text{dom } f^*$ is convex Lemma 4.5 gives

$$\text{dom}(Df^*) \subseteq \mathcal{V} + \text{conv} \hat{N} \subseteq \text{dom } f^*$$

Therefore

$$\text{dom } f^* \subseteq \overline{\text{dom } (Df^*)} = \overline{\mathcal{V} + N}$$

which yields

$$i_{\text{dom } f^*}(x^*) \geq i_{\overline{\mathcal{V} + N}}(x^*), \text{ for all } x^* \in X^* \quad (4.12)$$

Let $\bar{x} \in \text{int}(\text{dom } f)$ with $0 \in \partial f(\bar{x}) \subseteq \mathcal{V}$, i.e. $f(\bar{x}) = \min f$. Then since $0 \in \mathcal{V} \subseteq \partial f(\hat{x})$ for all $\hat{x} \in K$, (cf. Corollary 4.3) given $x^* \in X^*$ we have

$$\begin{aligned} f(\bar{x}) + f^*(x^*) &= \Phi_f(\bar{x}, x^*) = F_{\partial f}(\bar{x}, x^*) = \sup_{(a, a^*) \in \text{Gr}(\partial f)} \{ \langle a, x^* \rangle + \langle \bar{x}, a^* \rangle - \langle a, a^* \rangle \} \\ &\geq \sup_{\hat{x} \in K} \{ \langle \hat{x}, x^* \rangle + \langle \bar{x}, 0 \rangle - \langle \hat{x}, 0 \rangle \} = \sigma_K(x^*). \end{aligned}$$

Therefore

$$f^*(x^*) \geq \sigma_K(x^*) - \min f.$$

This together with (4.12) implies that for every $x^* \in X^*$

$$f^*(x^*) = f^*(x^*) + i_{\text{dom } f^*}(x^*) \geq \sigma_K(x^*) + i_{\overline{\mathcal{V} + N}}(x^*) - \min f = g(x^*)$$

Furthermore, for every $v^* \in \text{dom } (Df^*)$, by Proposition 4.4 we have

$$\hat{a} = Df^*(v^*) \in \text{argmin } f.$$

In particular

$$f^*(x^*) + \min f = f^*(v^*) + f(\hat{a}) = \langle \hat{a}, v^* \rangle \leq \sigma_K(v^*).$$

The above yields

$$f^*(v^*) = g(v^*), \text{ for every } v^* \in \text{dom } (Df^*).$$

Since $\text{dom } (Df^*)$ is dense in $\text{dom } g$, we apply Fact 2.1 to conclude $f^*(x^*) = g(x^*)$ for all $x^* \in X^*$, completing the proof. $\square$

Having obtained an explicit formula for f^* , we now compute f .

Proposition 4.7. *Let $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper convex and lsc. Assume that $\text{dom } (Df^*)$ is densely contained in $\text{dom } f^*$, $\text{int}(\text{dom } f) \cap \text{argmin } f \neq \emptyset$ and $\mathcal{F}_A$ is a singleton. Let $C = N^\circ = (\overline{\text{conv}} \hat{N})^\circ$ and $K, \mathcal{V}, \hat{N}$ be defined as in (4.5), (4.6), (4.10) respectively. Then for every $x \in X$ and $\hat{x} \in K$ we have:*

$$f(x) = \sigma_{\mathcal{V}}(x - \hat{x}) + i_{\overline{K+C}}(x) + \min f \quad (4.13)$$

Proof. Fix $\hat{x}_K \in K$ and define a function g as follows:

$$g(x) = \sigma_{\mathcal{V}}(x - \hat{x}_K) + i_{\overline{K+N^\circ}}(x) + \min f.$$

Recall that $\mathcal{V} \perp K - K$ by Proposition 4.4, thus for all $v^* \in \mathcal{V}$ and $\hat{x} \in K$

$$\langle \hat{x}, v^* \rangle = \langle \hat{x}_K, v^* \rangle, \quad \forall \hat{x} \in K.$$

This implies that

$$\sigma_{\mathcal{V}}(x - \hat{x}) = \sigma_{\mathcal{V}}(x - \hat{x}_K),$$

therefore the function g is well defined and is independent of the choice of $\hat{x}$. Note that

$$(i_{\overline{K+N^\circ}})^*(x^*) = \sigma_{\overline{K+N^\circ}}(x^*) = \sigma_{K+N^\circ}(x^*) = \sigma_K(x^*) + \sigma_{N^\circ}(x^*) = \sigma_K(x^*) + i_N(x^*),$$

where we used the fact that $\sigma_{N^\circ} \equiv i_N$ for the closed convex cone $N = \overline{\text{conv}} N$. Recalling (2.2), we infer

$$g^*(x^*) + \min f = \overline{(\sigma_{\mathcal{V}}^* + \langle \hat{x}_K, \cdot \rangle) \square (i_{\overline{K+N^\circ}})^*(x^*)} = \overline{(i_{\mathcal{V}} + \langle \hat{x}_K, \cdot \rangle) \square (\sigma_K + i_N)(x^*)}.$$

Furthermore:

$$\begin{aligned} (i_{\mathcal{V}} + \langle \hat{x}_K, \cdot \rangle) \square (\sigma_K + i_N)(x^*) &= \inf_{u^* + v^* = x^*} \{i_{\mathcal{V}}(u^*) + \langle \hat{x}_K, u^* \rangle + \sigma_K(v^*) + i_N(v^*)\} \\ &= \inf_{u^* \in \mathcal{V}} \{ \sigma_K(x^* - u^*) + \langle \hat{x}_K, u^* \rangle + i_N(x^* - u^*) \}. \end{aligned} \quad (4.14)$$

Since $\mathcal{V} \perp K - K$, we have again that for all $u^* \in \mathcal{V}$ and $\hat{x} \in K$ we have:

$$\langle \hat{x}, u^* \rangle = \langle \hat{x}_K, u^* \rangle.$$

Therefore,

$$\sigma_K(x^* - u^*) = \sup_{\hat{x} \in K} \{ \langle \hat{x}, x^* - u^* \rangle \} = \sup_{\hat{x} \in K} \langle \hat{x}, x^* \rangle - \langle \hat{x}_K, u^* \rangle = \sigma_K(x^*) - \langle \hat{x}_K, u^* \rangle.$$

Therefore the infimum in (4.14) disappears:

$$\inf_{u^* \in \mathcal{V}} \{ \sigma_K(x^* - u^*) + \langle \hat{x}_K, u^* \rangle + i_N(x^* - u^*) \} = \sigma_K(x^*) + \inf_{u^* \in \mathcal{V}} i_N(x^* - u^*) = \sigma_K(x^*) + i_{N+K}(x^*)$$

and we obtain

$$g^*(x^*) + \min f = \overline{\sigma_K + i_{\mathcal{V}+N}}(x^*). \quad (4.15)$$

By (4.11) $\sigma_K(x^*) + i_{\overline{\mathcal{V}+N}} := f^*(x^*) + \min f$. Since this function is proper convex and lsc, it follows that $\overline{\sigma_K + i_{\mathcal{V}+N}}(x^*) \geq \sigma_K(x^*) + i_{\overline{\mathcal{V}+N}}(x^*)$ and consequently (4.15) yields

$$g^*(x^*) + \min f \geq \sigma_K(x^*) + i_{\overline{\mathcal{V}+N}}(x^*) = f^*(x^*) + \min f$$

that is,

$$g^*(x^*) \geq f^*(x^*), \quad \text{for all } x^* \in X^*$$

Moreover

$$g^*(x^*) = f^*(x^*), \quad \text{for all } x^* \in \mathcal{V} + N$$

Since $\mathcal{V} + N$ is dense in $\text{dom } f^*$ and $\text{int}(\text{dom } f^*) \neq \emptyset$ we deduce by Fact 2.1 that

$$g^*(x^*) = f^*(x^*), \quad \forall x^* \in X^*.$$

As f, g are proper convex and lsc with $f^* = g^*$, it follows

$$f(x) = g(x) = \sigma_{\mathcal{V}}(x - \hat{x}_K) + i_{\overline{K+N}}(x) + \min f.$$

The proof is complete. $\square$

As an immediate corollary, we deduce the following result, which corresponds to the *only if* part of the statement of Theorem B.

Corollary 4.8. *Let X be an RN space, $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex, lsc with $\text{int}(\text{dom } f) \neq \emptyset$ and $\text{int}(\text{dom } f^*) \neq \emptyset$. Then if $\mathcal{F}_{\partial f}$ is a singleton, there exist a constant $c \in \mathbb{R}$, a functional $\bar{x}^* \in X^*$ and closed convex sets $K, C \subseteq X$ and $\mathcal{V} \subseteq X^*$ where*

- C is a cone, $\mathcal{V}$ is w^* -closed
- $0 \in \mathcal{V} \perp K - K$.

such that for every $\hat{x} \in K$

$$f(x) = \sigma_{\mathcal{V}}(x - \hat{x}) + i_{\overline{K+C}}(x) + \langle x, \bar{x}^* \rangle + c.$$

Proof. Let us first notice that f^* is a proper convex and w^* -lsc function. Since X is an RNP-space, the dual space X^* is w^* -Asplund and consequently the assumption $\text{int}(\text{dom } f^*) \neq \emptyset$ yields $\overline{\text{dom } (Df^*)} = \overline{\text{dom } f^*}$.

Let further $\bar{x} \in \text{int}(\text{dom } f)$, $\bar{x}^* \in \partial f(\bar{x})$ and define the function

$$\tilde{f}(x) := f(x) - \langle x - \bar{x}, \bar{x}^* \rangle, \quad x \in X.$$

Since $\mathcal{F}_{\partial f} = \{F_{\partial f}\}$, we obtain by translation invariance (Proposition 3.2) that $\mathcal{F}_{\partial \tilde{f}} = \{F_{\partial \tilde{f}}\}$. In addition,

$$\bar{x} \in \text{int}(\text{dom } \tilde{f}) \cap \text{argmin } \tilde{f} \quad \text{and} \quad \tilde{f}(\bar{x}) = f(\bar{x}) = \min \tilde{f}.$$

The assumptions of Proposition 4.7 are thus satisfied for $\tilde{f}$, allowing us to conclude that for every $x \in X$

$$f(x) = \sigma_{\mathcal{V}}(x) + i_{\overline{K+C}}(x) + \langle x, \bar{x}^* \rangle - \langle \bar{x}, \bar{x}^* \rangle + f(\bar{x})$$

where the sets $\mathcal{V}, C$ and K are defined as before. Setting $c = f(\bar{x}) - \langle \bar{x}, \bar{x}^* \rangle$ the assertion follows. $\square$

4.2 A class of subdifferentials with unique representative

The aim of this section is to prove that the class of functions introduced in Section 4.1, satisfy $\mathcal{F}_{\partial f} = \{F_{\partial f}\}$. Before proceeding, we introduce a notation. For a proper convex lsc function $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$, we denote

$$(f^* \oplus f)(x, x^*) = (f^*(x^*), f(x)) \in \mathbb{R}^2$$

and define

$$\partial_{\varepsilon}(f^* \oplus f)(x, x^*) = \{(y, y^*) \in X \times X^* : y^* \in \partial_a f(x), y \in \partial_b f^*(x^*), a + b \leq \varepsilon\}.$$

Consider further the function $K_{\partial f} : X \times X^* \rightarrow \mathbb{R} \cup \{+\infty\}$ defined via

$$K_{\partial f}(x, x^*) := \inf\{\varepsilon \geq 0 : \text{Gr}(\partial f) \cap \partial_{\varepsilon}(f^* \oplus f)(x, x^*) \neq \emptyset\}. \quad (4.16)$$

We are now ready to proceed with the main lemma of this section.

Lemma 4.9. *Let $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex and lsc. For $(w, v^*) \in \text{dom } f \times \text{dom } f^*$*

$$\Phi_f(w, v^*) = f(w) + f^*(v^*) = F_{\partial f}(w, v^*) + K_{\partial f}(w, v^*).$$

Proof. Fix $(w, v^*) \in \text{dom } f \times \text{dom } f^*$ and notice that since $F_{\partial f}$ and Φ_f belong to $\mathcal{F}_{\partial f}$ and the former is minimal, we have

$$F_{\partial f}(w, v^*) \leq f(w) + f^*(v^*) < +\infty.$$

For any arbitrary $\eta > 0$, there exist $a, b \geq 0$ and $(y, y^*) \in \text{Gr}(\partial f)$ with $y \in \partial_a f^*(v^*)$ and $y^* \in \partial_b f(w)$ such that $K_{\partial f}(w, v^*) + \eta \geq a + b$. This yields

$$f(w) + f^*(y^*) \leq \langle w, y^* \rangle + a \quad \text{and} \quad f(y) + f^*(v^*) \leq \langle y, v^* \rangle + b.$$

By (2.3) and the fact that $f(y) + f^*(y^*) = \langle y, y^* \rangle$ we have

$$\begin{aligned} F_{\partial f}(w, v^*) &\geq \langle w, y^* \rangle + \langle y, v^* \rangle - \langle y, y^* \rangle \\ &\geq f(w) + f^*(y^*) + f(y) + f^*(v^*) - f(y) - f^*(y^*) - a - b \\ &= f(w) + f^*(v^*) - a - b \geq f(w) + f^*(v^*) - K_{\partial f}(w, v^*) - \eta. \end{aligned}$$

Thus

$$K_{\partial f}(w, v^*) + F_{\partial f}(w, v^*) \geq f(w) + f^*(v^*). \quad (4.17)$$

For the reverse inequality, fix $\varepsilon > 0$ and let $(z, z^*) \in \text{Gr}(\partial f)$ be such that

$$F_{\partial f}(w, v^*) \leq \langle w, z^* \rangle + \langle z, v^* \rangle - \langle z, z^* \rangle + \varepsilon.$$

Then

$$\begin{aligned} f(w) + f^*(v^*) - F_{\partial f}(w, v^*) &\geq f(w) + f^*(v^*) - \langle w, z^* \rangle - \langle z, v^* \rangle + \langle z, z^* \rangle - \varepsilon \\ &= \underbrace{f(w) + f^*(z^*) - \langle w, z^* \rangle}_a + \underbrace{f(z) + f^*(v^*) - \langle z, v^* \rangle}_b - \varepsilon \\ &\geq K_{\partial f}(w, v^*) - \varepsilon \end{aligned}$$

which together with (4.17) concludes the proof. $\square$

Remark 4.10. Let us extract from the proof of Lemma 4.9 the following result: given a point $(x, x^*) \in \text{dom } f \times \text{dom } f^*$, the infimum in (4.16) is actually attained at some $\varepsilon \geq 0$ if and only if the supremum in the definition of the Fitzpatrick function (2.3) is attained.

An immediate consequence is the following result

Corollary 4.11. *Let $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex, and lsc. Then $\mathcal{F}_{\partial f} = \{F_{\partial f}\}$ if and only if*

$$\text{dom } F_{\partial f} = \text{dom } f \times \text{dom } f^*$$

and for every $(x, x^) \in \text{dom } f \times \text{dom } f^*$ and $\varepsilon > 0$*

$$\partial_\varepsilon (f^* \oplus f)(x, x^*) \cap \text{Gr}(\partial f) \neq \emptyset.$$

Proof. The two conditions imply that for every $(x, x^*) \in X \times X^*$

$$F_{\partial f}(x, x^*) = f(x) + f^*(x^*) = \Phi_f(x, x^*)$$

and the result follows from Lemma 3.1(iv). $\square$

With this criterion in hand, we now generalize [2, Theorem 5.3] to a broader class of functions.

Theorem 4.12. *Consider closed convex sets $K, C \subseteq X$ and $\mathcal{V} \subseteq X^*$ where*

- *C is a cone, $\mathcal{V}$ is w^* -closed*
- *$0 \in \mathcal{V} \perp K - K$*

Assume further the following compatibility condition:

$$\overline{(\sigma_K + i_{\mathcal{V}+C^\circ})}(x^*) = \sigma_K(x^*) + i_{\overline{\mathcal{V}+C^\circ}}(x^*). \quad (4.18)$$

Then for every $\hat{x} \in K$ define $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ by

$$f(x) = \sigma_{\mathcal{V}}(x - \hat{x}) + i_{\overline{K+C}}(x), \quad x \in X.$$

Then the Fitzpatrick family of its subdifferential reduces to a singleton: $\mathcal{F}_{\partial f} = \{F_{\partial f}\}$.

Proof. Reasoning as in Proposition 4.7, we deduce that f is well defined, independently of the choice of $\hat{x} \in K$. Fix $\hat{x}_K \in K$. Notice that since $\mathcal{V} \perp K - K$:

$$\min f = f(\hat{x}) = 0 \quad \text{for all } \hat{x} \in K$$

Step 1. We compute f^* and show that $\mathcal{V} \subseteq \partial f(\hat{x})$ for all $\hat{x} \in K$. Similarly to the computation of Proposition 4.7 we have:

$$f^*(x^*) = \overline{\sigma_K + i_{\mathcal{V}+C^\circ}}(x^*)$$

and using (4.18) we conclude that in fact

$$f^*(x^*) = \sigma_K(x^*) + i_{\overline{\mathcal{V}+C^\circ}}(x^*). \quad (4.19)$$

Furthermore, for any $v^* \in \mathcal{V}$ we have

$$\langle \hat{x}, v^* \rangle = \langle \hat{y}, v^* \rangle, \quad \text{for every } \hat{x}, \hat{y} \in K,$$

which yields

$$f^*(v^*) = \sigma_K(v^*) = \langle \hat{x}, v^* \rangle.$$

Since $f(\hat{x}) = 0$, for all $\hat{x} \in K$, it follows that

$$f(\hat{x}) + f^*(v^*) = \langle \hat{x}, v^* \rangle, \quad \text{for every } \hat{x} \in K$$

therefore in particular $\mathcal{V} \subseteq \bigcap_{\hat{x} \in K} \partial f(\hat{x}_K)$.

Step 2. We show that

$$\text{dom } F_{\partial f} = \text{dom } f \times \text{dom } f^* \quad (4.20)$$

To this end, let us recall, see e.g. [4, Theorem 2.6], that

$$\text{dom } f \times \text{dom } f^* \subseteq \text{dom } F_{\partial f} \subseteq \overline{\text{dom } f} \times \overline{\text{dom } f^*}. \quad (4.21)$$

If $w \in \overline{\text{dom } f} \setminus \text{dom } f$, then since

$$\overline{\text{dom } f} \subseteq \overline{K + C}$$

we infer that

$$\sigma_{\mathcal{V}}(w - \hat{x}_K) = +\infty.$$

Consequently there exists a sequence $\{v_n^*\}_{n \in \mathbb{N}} \in \mathcal{V} \subseteq \partial f(\hat{x}_K)$ such that $\langle w - \hat{x}_K, v_n^* \rangle \rightarrow +\infty$. It follows that for every $u^* \in X^*$,

$$F_{\partial f}(w, u^*) \geq \langle w, x_n^* \rangle + \langle 0, u^* \rangle - \langle \hat{x}_K, x_n^* \rangle = \langle w - \hat{x}_K, x_n^* \rangle \rightarrow +\infty.$$

As u^* is arbitrary, we obtain

$$\text{dom } F_{\partial f} \subseteq \text{dom } f \times \overline{\text{dom } f^*}.$$

On the other hand, by (4.19), $\overline{\text{dom } f^*} \subseteq \overline{\mathcal{V} + C^\circ}$. Thus, given $y^* \in \overline{\text{dom } f^*} \setminus \text{dom } f^*$ we have

$$\sigma_K(y^*) = +\infty,$$

and there exists a sequence $\{\hat{x}_n\}_{n \in \mathbb{N}} \subseteq K$ such that $\langle \hat{x}_n, y^* \rangle \rightarrow +\infty$. Since $0 \in \mathcal{V} \subseteq \partial f(\hat{x}_n)$ for all $n \in \mathbb{N}$, given $z \in X$ we have

$$F_{\partial f}(z, y^*) \geq \langle z, 0 \rangle + \langle \hat{x}_n, y^* \rangle - \langle \hat{x}_n, 0 \rangle \rightarrow +\infty.$$

Therefore, $\text{dom } F_{\partial f} \subseteq \text{dom } f \times \text{dom } f^*$ and by (4.21) we deduce that (4.20) holds.

Step 3. We prove that for every $(x, x^*) \in \text{dom } f \times \text{dom } f^*$ and $\varepsilon > 0$

$$\partial_\varepsilon(f^* \oplus f)(x, x^*) \cap \text{Gr}(\partial f) \neq \emptyset$$

To this end, fix $\varepsilon > 0$. From (4.19), we may find $\hat{y} \in K \subseteq \text{argmin } f$, such that $\sigma_K(x^*) \leq \langle \hat{y}, x^* \rangle + \varepsilon/2$. Since $f(\hat{y}) = 0$, we obtain

$$\hat{y} \in \partial_{\varepsilon/2} f^*(x^*).$$

We also have that $\partial_{\varepsilon/2} f(x) \cap \mathcal{V} \neq \emptyset$ for any $\varepsilon > 0$. To see this, as $x \in \text{dom } f$, take $v^* \in \mathcal{V}$ with

$$f(x) = \sigma_{\mathcal{V}}(x - \hat{x}_K) \leq \langle x - \hat{x}_K, v^* \rangle + \frac{\varepsilon}{2} = \langle x, v^* \rangle - \langle \hat{x}_K, v^* \rangle + \frac{\varepsilon}{2} = \langle x, v^* \rangle - f^*(v^*) + \frac{\varepsilon}{2}$$

where we used Step 1 and the fact that $f(\hat{x}_K) = 0$. Now, since $\mathcal{V} \subseteq \partial f(\hat{y})$ clearly

$$\partial_{\varepsilon/2} f(x) \cap \partial f(\hat{y}) \neq \emptyset$$

For $y^* \in \partial_{\varepsilon/2} f(x) \cap \partial f(\hat{y})$ we have:

$$(\hat{y}, y^*) \in \partial_\varepsilon(f^* \oplus f)(x, x^*) \cap \text{Gr}(\partial f) \neq \emptyset$$

Thus, both conditions of Corollary 4.11 are satisfied, concluding the proof. $\square$

Remark 4.13. (i). The compatibility condition (4.18) is satisfied in several natural situations. In particular, any of the following assumptions (a)-(b) ensures (4.18):

- (a) K is compact (b) $\mathcal{V} + C^\circ$ is closed (c) $\text{int}(\text{dom } f^*) \neq \emptyset$ (d) X is finite dimensional
(The later case which will be discussed in the next subsection.)

(ii). In the degenerate case where $K = \{0\}$ and $C = X$, we recover [2, Theorem 5.3], namely the result for sublinear functions. In a dual manner, if $\mathcal{V} = \{0\}$ and $C = \{0\}$, we recover the case of indicator functions of closed convex sets.

(iii). Computing conjugates we see that both f and f^* belong to the same family of functions (up to natural adjustments). This is consistent with [2, Theorem 5.7], which states that if the subdifferential of the function f^* satisfies $\mathcal{F}_{\partial f^*} = \{F_{\partial f^*}\}$, then so does the subdifferential of f .

Combining Theorem 4.12 with Corollary 4.8 we obtain Theorem B.

Proof of Theorem B.

It is sufficient to verify the compatibility condition (4.18). A close inspection of the proof of Proposition 4.7, shows that (4.18) holds whenever $\text{int}(\text{dom } f^*) \neq \emptyset$. $\square$

Combining the above result with the main result of Section 3 yields the following.

Corollary 4.14. *Let X be a space with the RNP, and $A : X \rightrightarrows X^*$ a 3-monotone and maximal monotone operator admitting only one representative function. If $\text{int}(\text{dom } (A))$ and $\text{int}(\text{Im } (A))$ are nonempty, then $A = \partial f$ for some proper convex lsc function f of the form (1.4).*

Remark 4.15. It remains open to determine whether the conditions $\text{int}(\text{dom } f) \neq \emptyset$ and $\text{int}(\text{dom } f^*) \neq \emptyset$ as well as the assumption that the ambient space is an RNP space are necessary for Theorem B (and consequently for Corollary 4.14).

On the other hand, the only non-subdifferential operators known to admit a unique representative function are the skew-symmetric linear operators. This motivates the following conjecture.

Conjecture 4.16. *A maximal monotone operator admitting a unique representation must either be the subdifferential of a function of the form (1.4) or a linear skew-symmetric operator.*

4.3 The Finite-Dimensional Case

We now focus on the finite-dimensional setting, where the technical assumptions required in Theorem 4.12 can be entirely removed. Indeed, in the aforementioned statements conditions on the Banach space X and on the domains of f and f^* were needed solely to guarantee the existence of a maximizer in (2.3), i.e., to establish Lemma 4.2. In finite dimensions, however, such a maximizer always exists under minimal assumptions, even without requiring the Fitzpatrick family being singleton. Throughout this section, E denotes a finite-dimensional Euclidean space.

Before proving the main result, we first recall some standard properties of convex functions in finite dimensions. Let $f : E \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper convex function. Its domain lies within the affine hull $\text{Aff}(\text{dom } f)$. The restriction of f to this space has nonempty interior, denoted by $\text{rint}(\text{dom } f)$. Fix $u \in \text{Aff}(\text{dom } f)$ and define the associated linear subspace of E given by

$$U := \text{Aff}(\text{dom } f) - \{u\}.$$

For each $x^* \in E^*$, we decompose

$$x^* = x_{\text{par}}^* + (x^*)^\perp,$$

where x_{par}^* is the projection onto U and $(x^*)^\perp \in U^\perp$. Consequently, for all $x, y \in \text{Aff}(\text{dom } f)$, we have

$$\langle x - y, (x^*)^\perp \rangle = 0. \quad (4.22)$$

For $\varepsilon \geq 0$ we define

$$\partial_\varepsilon^{\text{par}} f(x) := \{x_{\text{par}}^* : x^* \in \partial_\varepsilon f(x)\},$$

the projection onto U of the set of ε -subgradients of f at x . Restricting f to its affine hull,

$$\tilde{f} := f|_{\text{Aff}(\text{dom } f)} : \text{Aff}(\text{dom } f) \rightarrow \mathbb{R} \cup \{+\infty\},$$

we obtain $\text{rint}(\text{dom } f) = \text{int}(\text{dom } \tilde{f}) \neq \emptyset$. Hence for every $\varepsilon \geq 0$ and $x \in \text{rint}(\text{dom } f)$, the set $\partial_\varepsilon^{\text{par}} f(x)$ is bounded. More precisely:

$$\|y_{\text{par}}^*\| \leq M, \quad \text{for every } y^* \in \partial_\varepsilon f(x) \quad (4.23)$$

Analogously, let $v \in \text{Aff}(\text{dom } f^*)$ and define

$$V = \text{Aff}(\text{dom } f^*) - \{v\}.$$

Any $y \in E$ admits the decomposition

$$y = y_{\text{par}} + y^\perp,$$

with $y_{\text{par}} \in V$ and $y^\perp \in V^\perp$. If $y \in \partial_\varepsilon f^*(x^*)$ for some $\varepsilon \geq 0$ and $x^* \in \text{rint}(\text{dom } f^*)$, then y_{par} is also bounded, similar to (4.23).

Fixing $u \in \text{Aff}(\text{dom } f)$ and $v \in \text{Aff}(\text{dom } f^*)$, one checks that for any $(y, y^*) \in \text{Gr}(\partial f)$, the projection $(y_{\text{par}}, y_{\text{par}}^*)$ also lies in $\text{Gr}(\partial f)$. We therefore define the parallel projection of the graph with respect to U and V :

$$\text{Gr}^{\text{par}}(\partial f) := \{(y_{\text{par}}, y_{\text{par}}^*) : (y, y^*) \in \text{Gr}(\partial f)\} \subseteq (U \times V) \cap \text{Gr}(\partial f).$$

We are now in position to show that the supremum in (2.3) is always attained in finite dimensions whenever $(x, x^*) \in \text{rint}(\text{dom } f) \times \text{rint}(\text{dom } f^*)$:

Proposition 4.17. *Let $f: E \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper, convex and lsc on a finite-dimensional Euclidean space E . Suppose*

$$(x, x^*) \in \text{rint}(\text{dom } f) \times \text{rint}(\text{dom } f^*).$$

Then

$$\mathcal{M}_{\partial f}(x, x^*) \neq \emptyset.$$

where $\mathcal{M}_{\partial f}(x, x^)$ is defined as in (4.1).*

Proof. Since (x, x^*) lies in the product of the relative interiors of f and f^* , $F_{\partial f}, \Phi_f \in \mathcal{F}_{\partial f}$ and $F_{\partial f}$ is minimal

$$F_{\partial f}(x, x^*) \leq \Phi_f(x, x^*) = f(x) + f^*(x^*) < +\infty,$$

so the supremum is finite. Fix $(u, v) \in \text{Aff}(\text{dom } f) \times \text{Aff}(\text{dom } f^*)$ and define U, V as above. We first restrict the supremum to parallel components. Define

$$F_{\partial f}^{\text{par}}(x, x^*) := \sup_{(y_{\text{par}}, y_{\text{par}}^*) \in \text{Gr}^{\text{par}}(\partial f)} \{\langle x, y_{\text{par}}^* \rangle + \langle y_{\text{par}}, x^* \rangle - \langle y_{\text{par}}, y_{\text{par}}^* \rangle\}.$$

Clearly $F_{\partial f}(x, x^*) \geq F_{\partial f}^{\text{par}}(x, x^*)$. Take $\varepsilon > 0$ and $(y, y^*) \in \text{Gr}(\partial f)$ such that

$$F_{\partial f}(x, x^*) \leq \langle x, y^* \rangle + \langle y, x^* \rangle - \langle y, y^* \rangle + \varepsilon. \quad (4.24)$$

Writing $y = y_{\text{par}} + y^\perp$ and $y^* = y_{\text{par}}^* + (y^*)^\perp$, we infer from the orthogonality relation (4.22) that

$$\langle x - y, (y^*)^\perp \rangle = 0 \quad \text{and} \quad \langle y^\perp, x^* - y_{\text{par}}^* \rangle = 0.$$

Substituting into (4.24), we obtain

$$F_{\partial f}(x, x^*) \leq \langle x, y_{\text{par}}^* \rangle + \langle y_{\text{par}}, x^* \rangle - \langle y_{\text{par}}, y_{\text{par}}^* \rangle + \varepsilon \leq F_{\partial f}^{\text{par}}(x, x^*) + \varepsilon.$$

As $\varepsilon > 0$ is arbitrary, it follows that

$$F_{\partial f}(x, x^*) = F_{\partial f}^{\text{par}}(x, x^*).$$

Finally, by Lemma 4.9, the supremum may be restricted to points $(y_{\text{par}}, y_{\text{par}}^*)$ lying in

$$\partial_{\rho}^{\text{par}}(f^* \oplus f)(x, x^*) \cap \text{Gr}(\partial f), \quad (4.25)$$

for some $\rho > f(x) + f^*(x^*) - F_{\partial f}(x, x^*)$. Since (x, x^*) lies in the respective relative interiors of f, f^* , by (4.23) we have that $\partial_k^{\text{par}}(f^* \oplus f)(x, x^*)$ is bounded. As both $\partial_k^{\text{par}}(f^* \oplus f)(x, x^*)$ and $\text{Gr}(\partial f)$ are closed, the intersection in (4.25) is compact. This completes the proof, as the supremum of a continuous function over a compact set is attained. $\square$

In view of Proposition 4.17, all results from Section 4.1 extend to the finite-dimensional case, with no extra assumptions on $\text{dom } f$ and $\text{dom } f^*$, up to relative adjustments. Yielding the following full characterization:

Theorem 4.18. *Let $f: E \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex, lsc and E be finite dimensional Euclidean space. Then the Fitzpatrick family, $\mathcal{F}_{\partial f}$, consists of a single element if and only if there exist a constant $c \in \mathbb{R}$, a functional $\bar{x}^* \in E^*$ and closed convex sets $K, C, \mathcal{V} \subseteq E$ with C a cone with*

$$0 \in \mathcal{V} \perp K - K$$

such that, for every $\hat{x} \in K$

$$f(x) = \sigma_{\mathcal{V}}(x - \hat{x}) + i_{\overline{K+C}}(x) + \langle x, \bar{x}^* \rangle + c, \text{ for } x \in E.$$

Proof. The assertion follows by applying the arguments from Section 4.3 in the finite-dimensional context. In particular since E is finite dimensional, Rademacher's theorem ensures that f^* is differentiable almost everywhere on $\text{rint}(\text{dom } f^*)$ and Fitzpatrick supremum is attained in virtue of Proposition 4.17, allowing us to reproduce the arguments of Lemma 4.3 and Corollary 4.8 verbatim within E . We need only check the compatibility condition of Theorem 4.12:

$$f^*(x^*) = \overline{(\sigma_K + i_{\mathcal{V}+N})}(x^*) = \sigma_K(x^*) + i_{\overline{\mathcal{V}+N}}(x^*) := g(x^*)$$

The two convex lsc functions above, agree on $\mathcal{V} + N$ which is dense in $\text{rint}(\text{dom } g)$ and clearly as the lsc envelope of the sum is greater than the sum of the lsc envelopes, $f^* \geq g$ always. In virtue of Fact 2.1 the conclusion follows. $\square$

Acknowledgments The authors would like to thank Pierre-Cyril Aubin-Frankowski, Jérôme Bolte, Heinz Bauschke, Guillaume Carlier, Juan Enrique Martínez-Legaz, Nicolas Hadjisavvas and Sebastián Tapia-García for useful discussion and suggestions.

This research was funded in whole by the Austrian Science Fund (FWF) [DOI 10.55776/P36344N]. For open access purposes, the second author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

References

- [1] S. N. ARMSTRONG AND J.-C. MOURRAT, *Lipschitz regularity for elliptic equations with random coefficients*, Arch. Ration. Mech. Anal., 219 (2016), pp. 255–348.
- [2] S. BARTZ, H. H. BAUSCHKE, J. M. BORWEIN, S. REICH, AND X. WANG, *Fitzpatrick functions, cyclic monotonicity and Rockafellar’s antiderivative*, Nonlinear Anal., 66 (2007), pp. 1198–1223.
- [3] H. H. BAUSCHKE AND P. L. COMBETTES, *Convex analysis and monotone operator theory in Hilbert spaces*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer, Cham, second ed., 2017. With a foreword by Hedy Attouch.
- [4] H. H. BAUSCHKE, D. A. McLAREN, AND H. S. SENDOV, *Fitzpatrick functions: inequalities, examples, and remarks on a problem by S. Fitzpatrick*, Journal of Convex Analysis, 13 (2006), p. 499.
- [5] H. H. BAUSCHKE, S. SINGH, AND X. WANG, *On Carlier’s inequality*, arXiv preprint arXiv:2206.14872, (2022).
- [6] H. H. BAUSCHKE, X. WANG, AND L. YAO, *Monotone linear relations: maximality and Fitzpatrick functions*, J. Convex Anal., 16 (2009), pp. 673–686.
- [7] J. BENOIST AND A. DANILIDIS, *Coincidence theorems for convex functions*, J. Convex Anal., 9 (2002), pp. 259–268.
- [8] J. M. BORWEIN, *Maximal monotonicity via convex analysis*, J. Convex Anal., 13 (2006), pp. 561–586.
- [9] J. M. BORWEIN AND L. YAO, *Structure theory for maximally monotone operators with points of continuity*, J. Optim. Theory Appl., 157 (2013), pp. 1–24.
- [10] R. I. BOŢ, S.-M. GRAD, AND G. WANKA, *Maximal monotonicity for the precomposition with a linear operator*, SIAM J. Optim., 17 (2006), pp. 1239–1252.
- [11] H. BRÉZIS, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, vol. No. 5 of North-Holland Mathematics Studies, North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973. Notas de Matemática, No. 50. [Mathematical Notes].
- [12] R. S. BURACHIK AND J. E. MARTÍNEZ-LEGAZ, *A note on Carlier inequality*, (2025).
- [13] R. S. BURACHIK AND B. F. SVAITER, *Maximal monotone operators, convex functions and a special family of enlargements*, Set-Valued Analysis, 10 (2002), pp. 297–316.
- [14] R. S. BURACHIK AND B. F. SVAITER, *Maximal monotonicity, conjugation and the duality product*, Proc. Amer. Math. Soc., 131 (2003), pp. 2379–2383.
- [15] G. CARLIER, *Fenchel–Young inequality with a remainder and applications to convex duality and optimal transport*, SIAM Journal on Optimization, 33 (2023), pp. 1463–1472.
- [16] J. B. COLLIER, *The dual of a space with the Radon–Nikodým property*, Pacific J. Math., 64 (1976), pp. 103–106.
- [17] M. FABIAN, P. HABALA, P. HÁJEK, V. MONTESINOS, AND V. ZIZLER, *Banach space theory*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer, New York, 2011. The basis for linear and nonlinear analysis.
- [18] S. FITZPATRICK, *Representing monotone operators by convex functions*, in Workshop/Miniconference on Functional Analysis and Optimization, vol. 20, Australian National University, Mathematical Sciences Institute, 1988, pp. 59–66.
- [19] N. GHOUSOUB, *A variational theory for monotone vector fields*, J. Fixed Point Theory Appl., 4 (2008), pp. 107–135.
- [20] J.-E. MARTÍNEZ-LEGAZ AND B. F. SVAITER, *Monotone operators representable by lsc convex functions*, Set-Valued Analysis, 13 (2005), pp. 21–46.
- [21] R. R. PHELPS, *Convex functions, monotone operators and differentiability*, vol. 1364 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, second ed., 1993.
- [22] S. RAKOTOMANDIMBY, J.-P. CHANCELIER, M. DE LARA, AND M. BLONDEL, *Learning with Fitzpatrick losses*, 2024.
- [23] S. REICH AND S. SIMONS, *Fenchel duality, Fitzpatrick functions and the Kirszbraun–Valentine extension theorem*, Proc. Amer. Math. Soc., 133 (2005), pp. 2657–2660.

- [24] R. T. ROCKAFELLAR, *Characterization of the subdifferentials of convex functions*, Pacific J. Math., 17 (1966), pp. 497–510.
- [25] B. F. SVAITER, *On weak convergence of the Douglas-Rachford method*, SIAM J. Control Optim., 49 (2011), pp. 280–287.
- [26] C. ZĂLINESCU, *Convex analysis in general vector spaces*, World Scientific Publishing Co., Inc., River Edge, NJ, 2002.

Sotiris Armeniakos, Aris Daniilidis

Institut für Stochastik und Wirtschaftsmathematik, VADOR E105-04
 TU Wien, Wiedner Hauptstraße 8, 1040 Vienna, Austria

Email: sotirios.armeniakos@tuwien.ac.at, aris.daniilidis@tuwien.ac.at

Webpage: <https://www.arisdaniilidis.at>

Research supported by the Austrian Science Fund (FWF), grant DOI 10.55776/P-36344N.