Journal of Convex Analysis
Volume 13 (2006), No. 3+4, 499–523

Fitzpatrick Functions: Inequalities, Examples, and Remarks on a Problem by S. Fitzpatrick

Heinz H. Bauschke
Department of Mathematics, UBC Okanagan, Kelowna, British Columbia V1V 1V7, Canada
heinz.bauschke@ubc.ca

D. Alexander McLaren
Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario N1G 2W1, Canada
amclaren@uoguelph.ca

Hristo S. Sendov
Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario N1G 2W1, Canada
hssendov@uoguelph.ca

Dedicated to the memory of Simon Fitzpatrick.

Received: February 4, 2005
Revised manuscript received: July 12, 2005

In 1988, Simon Fitzpatrick defined a new convex function F_A — nowadays called the Fitzpatrick function — associated with a monotone operator A, and similarly a monotone operator G_f associated with a convex function f.

This paper deals with two different aspects of Fitzpatrick functions. In the first half, we consider the Fitzpatrick function of the subdifferential of a proper, lower semicontinuous, and convex function. A refinement of the classical Fenchel-Young inequality is derived and conditions for equality are investigated. The results are illustrated by several examples.

In the second half, we study the problem, originally posed by Fitzpatrick, of determining when $A = G_{F_A}$. Fitzpatrick proved that this identity is satisfied whenever A is a maximal monotone; however, he also observed that it can hold even in the absence of maximal monotonicity. We propose a new condition sufficient for this identity, formulated in terms of the polarity notions introduced recently by Martínez-Legaz and Svaiter. Moreover, on the real line, this condition is also necessary and it corresponds to the connectedness of A.

Keywords: Convex function, Fenchel conjugate, Fenchel-Young inequality, Fitzpatrick function, monotone operator, monotone set

2000 Mathematics Subject Classification: Primary 26B25, 47H05; Secondary 47H04, 52A41, 90C25

1. The subdifferential case: Introduction

Throughout the first half of this paper, we assume that X is a real reflexive Banach space, with norm $\|\cdot\|$, with dual space X^*, and with duality product $p = \langle\cdot|\cdot\rangle$. Recall that an operator $A\colon X \to 2^{X^*}$, which we identify entirely with its graph in $X \times X^*$ in the second

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

half of this paper, is *monotone*, if

$$(\forall (x,x^*) \in A)(\forall (y,y^*) \in A) \quad \langle x-y|x^*-y^*\rangle \geq 0. \tag{1}$$

If a monotone operator A possesses no proper extension that is still monotone, then A is said to be *maximal monotone*. The prime example of a maximal monotone operator is the *subdifferential operator* ∂f of a proper, lower semicontinuous, and convex function f (see [12]). For background material on monotone operators and convex analysis, we refer the reader to [13], [14], [16], and [20]. The notation we employ is standard. The projector (resp. normal cone operator, indicator function, distance function) for a given nonempty closed convex set C in X is denoted by P_C (resp. by N_C, by ι_C, by d_C).

In 1988, Simon Fitzpatrick (see [6, Definition 3.1]) investigated the following function which has been utilized recently in [3], [8], [9], [19], and [18].

Definition 1.1 (Fitzpatrick function). The *Fitzpatrick function* associated with an operator $A\colon X \to 2^{X^*}$ is defined by

$$\begin{aligned} F_A\colon X \times X^* \to \left]-\infty,+\infty\right] : (y,y^*) \ &\mapsto \sup_{(x,x^*)\in A} \big(\langle y|x^*\rangle + \langle x|y^*\rangle - \langle x|x^*\rangle \big) \quad (2) \\ &= \langle y|y^*\rangle - \inf_{(x,x^*)\in A} \langle y-x|y^*-x^*\rangle \,. \quad (3) \end{aligned}$$

Observe that $(X \times X^*)^* = X^* \times X$ and define $R\colon X \times X^* \to X^* \times X : (x,x^*) \mapsto (x^*,x)$. This operator is useful in the formulation of some basic properties of the Fitzpatrick function.

Fact 1.2. *Let $A\colon X \to 2^{X^*}$ be monotone, and let $(x,x^*) \in X \times X^*$. Then the following is true.*

(*i*) F_A *is convex, lower semicontinuous, and proper.*
(*ii*) $F_A = (\iota_A + p)^* \circ R$.
(*iii*) $F_A \leq F_A^* \circ R \leq p + \iota_A$, *with equality throughout at points in A.*
(*iv*) $F_{A^{-1}} \circ R = F_A$.

Proof. See [6, Proposition 4.1 and Proposition 4.2]. □

The first objective of this paper is to study the Fitzpatrick function $F_{\partial f}$ of the subdifferential operator of a convex, lower semicontinuous, and proper function f. The function $F_{\partial f}$ is particularly interesting because it allows the following refinement of the classical Fenchel-Young inequality:

$$(\forall x \in X)(\forall x^* \in X^*) \quad \langle x|x^*\rangle \leq F_{\partial f}(x,x^*) \leq f(x) + f^*(x^*). \tag{4}$$

This part of the paper is organized as follows. In Section 2, the inequalities (4) are derived, the domain of $F_{\partial f}$ is located as precisely as possible, and the question when equalities occur in (4) is investigated. These results are illustrated in Section 3, where numerous examples are presented. Section 4 provides a natural upper bound for the Fitzpatrick function of a sum of two monotone operators. This section concludes our work on Fitzpatrick functions of subdifferentials.

The remaining sections 5–8 deal with a problem by Simon Fitzpatrick; we refer the reader to Section 5 for a more detailed introduction to the second half of this paper.

2. The subdifferential case: Refined Fenchel-Young inequality

Proposition 2.1. *Let $f\colon X \to \left]-\infty,+\infty\right]$ be convex, lower semicontinuous, and proper. Then for all $(y,y^*) \in X \times X^*$, we have*

$$\langle y|y^*\rangle \le F_{\partial f}(y,y^*) \le f(y) + f^*(y^*) \le F^*_{\partial f}(y^*,y) \le \langle y|y^*\rangle + \iota_{\partial f}(y,y^*). \tag{5}$$

In particular, $\operatorname{dom} f \times \operatorname{dom} f^* \subset \operatorname{dom} F_{\partial f}$*. If* $(y,y^*) \in (\operatorname{dom} \partial f) \times (\operatorname{dom} \partial f^*)$ *satisfies* $F_{\partial f}(y,y^*) = f(y) + f^*(y^*)$ *and* $\big((x_n,x_n^*)\big)_{n\in\mathbb{N}}$ *is a sequence in* ∂f *such that* $\langle x_n|y^*\rangle + \langle y|x_n^*\rangle - \langle x_n|x_n^*\rangle \to F_{\partial f}(y,y^*)$*, then*

$$f(x_n) - \langle x_n|y^*\rangle \to \min_{x\in X}\big(f(x) - \langle x|y^*\rangle\big) = -f^*(y^*) \tag{6}$$

and

$$f^*(x_n^*) - \langle y|x_n^*\rangle \to \min_{x^*\in X^*}\big(f^*(x^*) - \langle y|x^*\rangle\big) = -f(y); \tag{7}$$

furthermore, the minimizers in (6) *and* (7) *are* $\partial f^*(y^*)$ *and* $\partial f(y)$*, respectively.*

Proof. The first inequality of (5) is equivalent to $\inf_{(x,x^*)\in\partial f}\langle y-x|y^*-x^*\rangle \le 0$, which is true because ∂f is maximal monotone (see [12] or [20, Theorem 3.1.11]). Let $\big((x_n,x_n^*)\big)_{n\in\mathbb{N}}$ be a sequence in ∂f such that $\langle y|x_n^*\rangle + \langle x_n|y^*\rangle - \langle x_n|x^*\rangle \to F_{\partial f}(y,y^*)$. Then

$$\begin{aligned} F_{\partial f}(y,y^*) &\leftarrow \langle y|x_n^*\rangle + \langle x_n|y^*\rangle - \langle x_n|x_n^*\rangle = \langle y-x_n|x_n^*\rangle + \langle x_n|y^*\rangle && (8)\\ &\le f(y) - f(x_n) + \langle x_n|y^*\rangle \le f(y) + f^*(y^*). && (9)\end{aligned}$$

This verifies the second inequality of (5) and it implies that $\operatorname{dom} f \times \operatorname{dom} f^* \subset \operatorname{dom} F_{\partial f}$. Taking the conjugate of this second inequality yields the third inequality of (5). The fourth inequality of (5) is clear since $(p+\iota_{\partial f})^*(y^*,y) = F_{\partial f}(y,y^*)$ (see Fact 1.2(*ii*)).

Now let $(y,y^*) \in (\operatorname{dom}\partial f) \times (\operatorname{dom}\partial f^*)$ be such that $F_{\partial f}(y,y^*) = f(y) + f^*(y^*)$. Then (8)–(9) imply that $\langle x_n|y^*\rangle - f(x_n) \to f^*(y^*)$. Hence (6) holds and an analogous argument verifies (7). The result concerning the minimizers follows from convex calculus. $\square$

Remark 2.2. Let $f\colon X \to \left]-\infty,+\infty\right]$ be convex, lower semicontinuous, and proper, and let $(y,y^*) \in \operatorname{dom} F_{\partial f}$.

(i) Parts of Proposition 2.1 were already proved in Fitzpatrick's original paper: [6, Corollaries 3.9 and 3.13] imply that $p \le F_{\partial f}$ and that $\operatorname{conv}\operatorname{dom}\partial f \times \operatorname{conv}\operatorname{dom}\partial f^* \subset \operatorname{dom} F_{\partial f}$. This last inclusion is a consequence of Proposition 2.1 since $\operatorname{dom}\partial f \subset \operatorname{dom} f$, $\operatorname{dom}\partial f^* \subset \operatorname{dom} f^*$, and both $\operatorname{dom} f$ and $\operatorname{dom} f^*$ are convex.

(ii) It is well known that equality in the Fenchel-Young inequality

$$\langle y|y^*\rangle \le f(y) + f^*(y^*) \tag{10}$$

holds precisely on the graph of ∂f, i.e.,

$$\langle y|y^*\rangle = f(y) + f^*(y^*) \quad \text{if and only if } (y,y^*) \in \partial f. \tag{11}$$

Proposition 2.1 yields

$$\langle y|y^*\rangle \le F_{\partial f}(y,y^*) \le f(y) + f^*(y^*). \tag{12}$$

These inequalities provide a refinement of the Fenchel-Young inequality (10).

(iii) It is natural to inquire when equality occurs in either of the inequalities (12). Fitzpatrick's [6, Corollary 3.9] provides a complete solution for the left-hand inequality: indeed, since ∂f is maximal monotone, his result states that

$$\langle y|y^*\rangle = F_{\partial f}(y,y^*) \quad \text{if and only if } (y,y^*)\in\partial f. \tag{13}$$

(iv) The problem of characterizing which points in $\operatorname{dom} F_{\partial f}$ satisfy

$$F_{\partial f}(y,y^*) \stackrel{?}{=} f(y)+f^*(y^*) \tag{14}$$

is interesting as it does not appear to have a simple solution (see Theorem 2.3 and Example 3.1). However, we note the implication

$$(y,y^*)\in\partial f \quad\Rightarrow\quad F_{\partial f}(y,y^*)=f(y)+f^*(y^*); \tag{15}$$

indeed, if $(y,y^*)\in\partial f$, then $\langle y|y^*\rangle = f(y)+f^*(y^*)$ (see (11)) and the identity now follows from (12). Theorem 2.3 below shows that if f is "sufficiently nice", then it is precisely the points in ∂f that satisfy (14).

Theorem 2.3. *Suppose that $f\colon X\to\mathbb{R}$ and $f^*\colon X^*\to\mathbb{R}$ are both Fréchet differentiable and convex, and that $(y,y^*)\in X\times X^*$. Then $F_{\nabla f}(y,y^*) = f(y)+f^*(y^*)$ if and only if $y^*=\nabla f(y)$.*

Proof. Observe that $\operatorname{dom} F_{\nabla f} = X\times X^*$ by Proposition 2.1. If $y^*=\nabla f(y)$, then $F_{\nabla f}(y,y^*)=f(y)+f^*(y^*)$ by (15). Conversely, let us assume that $(x_n)_{n\in\mathbb{N}}$ is a sequence in X such that

$$\langle y|\nabla f(x_n)\rangle + \langle x_n|y^*\rangle - \langle x_n|\nabla f(x_n)\rangle \to F_{\nabla f}(y,y^*) = f(y)+f^*(y^*). \tag{16}$$

The limit statements (6) and (7) of Proposition 2.1 combined with [20, Theorem 3.9.1] yield

$$x_n\to\nabla f^*(y^*) \quad\text{and}\quad \nabla f(x_n)\to y^*. \tag{17}$$

Now (16) and (17) result in $\langle y|y^*\rangle = f(y)+f^*(y^*)$. Therefore, $y^*=\nabla f(y)$. ☐

Theorem 2.4. *Let X be finite-dimensional, let both $f\colon X\to\mathbb{R}$ and $f^*\colon X^*\to\mathbb{R}$ be convex, and suppose that $(y,y^*)\in X\times X^*$. Then $F_{\partial f}(y,y^*)=f(y)+f^*(y^*)$ if and only if there exists $(z,z^*)\in\partial f$ such that $F_{\partial f}(y,y^*) = \langle y|z^*\rangle+\langle z|y^*\rangle-\langle z|z^*\rangle$, $(y,z^*)\in\partial f$, and $(z,y^*)\in\partial f$.*

Proof. Suppose first that there exists $(z,z^*)\in\partial f$ such that $F_{\partial f}(y,y^*) = \langle y|z^*\rangle+\langle z|y^*\rangle-\langle z|z^*\rangle$, $(y,z^*)\in\partial f$, and $(z,y^*)\in\partial f$. Using (5), we obtain

$$\begin{aligned}
F_{\partial f}(y,y^*) &= \langle y|z^*\rangle+\langle z|y^*\rangle-\langle z|z^*\rangle && (18)\\
&= f(y)+f^*(z^*)+f(z)+f^*(y^*)-\langle z|z^*\rangle && (19)\\
&\geq f(y)+f^*(y^*) && (20)\\
&\geq F_{\partial f}(y,y^*). && (21)
\end{aligned}$$

Conversely, let $\big((x_n,x_n^*)\big)_{n\in\mathbb{N}}$ be a sequence in ∂f such that

$$\langle x_n|y^*\rangle + \langle y|x_n^*\rangle - \langle x_n|x_n^*\rangle \to F_{\partial f}(y,y^*) = f(y)+f^*(y^*). \tag{22}$$

By Proposition 2.1, the sequences $(x_n)_{n\in\mathbb{N}}$ and $(x_n^*)_{n\in\mathbb{N}}$ are minimizing for the objective functions $f - \langle\cdot|y^*\rangle$ and $f^* - \langle y|\cdot\rangle$, which have minimizers $\partial f^*(y^*)$ and $\partial f(y)$, respectively. Both sets of minimizers are nonempty and compact, because X is finite-dimensional and the functions f and f^* have full domains. Therefore, by [5, Proposition I.37], the two corresponding optimization problems are *well posed in the generalized sense* [5, Section I.6], which implies that both minimizing sequences $(x_n)_{n\in\mathbb{N}}$ and $(x_n^*)_{n\in\mathbb{N}}$ possess subsequences converging to corresponding minimizers. After relabeling, if necessary, we assume that $x_n \to z \in \partial f^*(y^*)$ and that $x_n^* \to z^* \in \partial f(y)$. We deduce that $(z, y^*) \in \partial f$ and that $(y, z^*) \in \partial f$. Since $\big((x_n, x_n^*)\big)_{n\in\mathbb{N}}$ lies in ∂f, taking the limit shows that $(z, z^*) \in \partial f$. Furthermore, taking limits in (22) yields $\langle z|y^*\rangle + \langle y|z^*\rangle - \langle z|z^*\rangle = F_{\partial f}(y, y^*)$. $\square$

Remark 2.5. Suppose that X is finite-dimensional, that both $f\colon X \to \mathbb{R}$ and $f^*\colon X \to \mathbb{R}$ are convex and differentiable, and that $(y, y^*) \in X \times X^*$ satisfies $F_{\nabla f}(y, y^*) = f(y) + f^*(y^*)$. Then Theorem 2.4 guarantees the existence of $z \in X$ such that $F_{\nabla f}(y, y^*) = \langle z|y^*\rangle + \langle y|\nabla f(z)\rangle - \langle z|\nabla f(z)\rangle$ and $\nabla f(y) = \nabla f(z)$. Note that f is strictly convex since f^* is differentiable. Thus, $y = z$ and $f(y) + f^*(y^*) = F_{\nabla f}(y, y^*) = \langle y|y^*\rangle$. Therefore, $y^* = \nabla f(y)$, which is also a consequence of Theorem 2.3.

We now turn to the following result, which is useful for computing the Fitzpatrick function of a subdifferential as it almost precisely locates the domain.

Theorem 2.6. *Let $f\colon X \to\ ]-\infty, +\infty]$ be convex, lower semicontinuous, and proper. Then*

$$\operatorname{dom} f \times \operatorname{dom} f^* \subset \operatorname{dom} F_{\partial f} \subset \overline{\operatorname{dom}}\, f \times \overline{\operatorname{dom}}\, f^*. \tag{23}$$

Proof. The first inclusion of (23) follows from Proposition 2.1. Now set $C = \overline{\operatorname{dom}}\, f$ and suppose that $y \in X \smallsetminus C$. A result due to Simons (see [15, Lemma 2.(c)] or [20, Theorem 3.1.9(iii)]) implies that for every $\eta \in\ ]\inf f(X), +\infty[\ =\ ]\inf f(X), f(y)[$, there exists $(x_\eta, x_\eta^*) \in \partial f$ such that

$$L_\eta := \sup_{x\in X\smallsetminus\{y\}} \frac{\eta - f(x)}{\|y - x\|} \le 2\frac{\langle y - x_\eta|x_\eta^*\rangle}{\|y - x_\eta\|}. \tag{24}$$

Let us fix momentarily $x_0 \in \operatorname{dom} f$ and let us agree upon that limits taken in this proof correspond to letting η tend to $+\infty$. Then $x_0 \neq y$ and $L_\eta \ge (\eta - f(x_0))/\|x_0 - y\| \to +\infty$. Thus,

$$\langle y - x_\eta|x_\eta^*\rangle \ge \tfrac{1}{2}L_\eta\|y - x_\eta\| \ge \tfrac{1}{2}L_\eta d_C(y) \to +\infty. \tag{25}$$

It follows that, for every $y^* \in X^*$,

$$\begin{aligned}
F_{\partial f}(y, y^*) &\ge \langle y|x_\eta^*\rangle + \langle x_\eta|y^*\rangle - \langle x_\eta|x_\eta^*\rangle && (26)\\
&= \langle y - x_\eta|x_\eta^*\rangle + \langle x_\eta - y|y^*\rangle + \langle y|y^*\rangle && (27)\\
&\ge \tfrac{1}{2}L_\eta\|y - x_\eta\| - \|y^*\|\|x_\eta - y\| + \langle y|y^*\rangle && (28)\\
&\ge \big(\tfrac{1}{2}L_\eta - \|y^*\|\big)d_C(y) + \langle y|y^*\rangle && (29)\\
&\to +\infty. && (30)
\end{aligned}$$

Thus, $(y, y^*) \in \big((X \smallsetminus C) \times X^*\big) \smallsetminus \operatorname{dom} F_{\partial f}$ and hence

$$\operatorname{dom} F_{\partial f} \subset \overline{\operatorname{dom}}\, f \times X^*. \tag{31}$$

Applying this line of thought to f^* rather than f, and recalling that $F_{\partial f^*} \circ R = F_{(\partial f)^{-1}} \circ R = F_{\partial f}$ (see Fact 1.2(iv)), we deduce that

$$\operatorname{dom} F_{\partial f} \subset X \times \overline{\operatorname{dom}}\, f^*. \tag{32}$$

Altogether, $\operatorname{dom} F_{\partial f} \subset \overline{\operatorname{dom}}\, f \times \overline{\operatorname{dom}}\, f^*$. $\square$

Remark 2.7. Theorem 2.6 can be proved differently by utilizing new results of S. Simons [17]. Remark 3.2(ii) and Remark 3.5(ii) below show that it is impossible to improve the lower and upper bound of Theorem 2.6, respectively.

3. The subdifferential case: Examples

Recall that $N_C = \partial \iota_C$ denotes the normal cone operator.

Example 3.1 (indicator function). Suppose that X is a real Hilbert space and that C is a nonempty closed convex set in X. Then $F_{N_C}\colon X \times X \to \,]-\infty, +\infty] : (y, y^*) \mapsto \iota_C(y) + \iota_C^*(y^*)$.

Proof. Fix $(y, y^*) \in X \times X$. Then

$$F_{N_C}(y, y^*) = \sup_{x \in C,\, x^* \in N_C(x)} \big(\langle x|y^*\rangle + \langle y|x^*\rangle - \langle x|x^*\rangle \big) \tag{33}$$

$$= \sup_{x \in C} \big(\langle x|y^*\rangle + \sup \langle y - x|N_C(x)\rangle \big). \tag{34}$$

If $y \in C = \operatorname{dom} \iota_C$, then $(\forall x \in C)$ $\sup \langle y - x|N_C(x)\rangle = 0$ and hence $F_{N_C}(y, y^*) = \sup_{x \in C} \langle x|y^*\rangle = \iota_C^*(y^*) = \iota_C(y) + \iota_C^*(y^*)$. Now assume that $y \notin C$. Then $P_C y \in C$ and $z = y - P_C y \in N_C(P_C y) \smallsetminus \{0\}$. Hence $\sup \langle y - P_C y|N_C(P_C y)\rangle \geq \sup_{\rho \in [0,+\infty[} \langle z|\rho z\rangle = +\infty = \iota_C(y)$. Therefore, $F_{N_C}(y, y^*) = \iota_C(y) + \iota_C^*(y^*)$ in all cases. $\square$

Remark 3.2.

(i) Example 3.1 provides a function for which (14) holds *everywhere* on $X \times X^*$. This is in stark contrast to Theorem 2.3, which states that a “sufficiently nice” function f satisfies this identity only where it has to, namely on the graph of ∇f.

(ii) Now suppose that in $X = \mathbb{R}^2$, the set C is the epigraph of the function $\mathbb{R} \to \mathbb{R}\colon \rho \mapsto \frac{1}{2}|\rho|^2$. Then $\operatorname{dom} \iota_C^* = \operatorname{ran} N_C = \big\{(\xi_1, \xi_2) \in X \mid \xi_2 < 0\big\} \cup \{(0,0)\}$, which is a set that is not closed. Hence $\operatorname{dom} F_{N_C} = \operatorname{dom} \iota_C \times \operatorname{dom} \iota_C^*$ is also not closed. Therefore,

$$\operatorname{dom} f \times \operatorname{dom} f^* = \operatorname{dom} F_{\partial f} \subsetneqq \overline{\operatorname{dom}}\, f \times \overline{\operatorname{dom}}\, f^* \tag{35}$$

can occur in Theorem 2.6.

Example 3.3 (norm). Suppose X is a real Hilbert space and let $f = \|\cdot\|$. Then

$$F_{\partial f}\colon X \times X^* \to \,]-\infty, +\infty] : (y, y^*) \mapsto f(y) + f^*(y^*) = \begin{cases} \|y\|, & \text{if } \|y^*\| \leq 1; \\ +\infty, & \text{otherwise.} \end{cases} \tag{36}$$

Proof. Let $g = f^*$, i.e., g is the indicator function of the unit ball, and let $(y, y^*) \in X \times X$. Then Fact 1.2(iv) and Example 3.1 yield $F_{\partial f}(y, y^*) = \big(F_{(\partial f)^{-1}} \circ R\big)(y, y^*) = \big(F_{\partial g} \circ R\big)(y, y^*) = F_{\partial g}(y^*, y) = g(y^*) + g^*(y) = f^*(y^*) + f(y)$. $\square$

The logarithmic barrier, $-\ln$, is a classical function in convex analysis and optimization. It admits simple formulae for the corresponding Fitzpatrick function and its conjugate.

Example 3.4 (negative logarithm). Suppose $X = \mathbb{R}$ and let

$$f\colon X \to \left]-\infty,+\infty\right] : \rho \mapsto \begin{cases} +\infty, & \text{if } \rho \le 0; \\ -\ln(\rho), & \text{if } \rho > 0. \end{cases} \tag{37}$$

Then

$$F_{\partial f}\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}\colon (\rho,\rho^*) \mapsto \begin{cases} 1-2\sqrt{\rho(-\rho^*)}, & \text{if } \rho \ge 0 \text{ and } \rho^* \le 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{38}$$

Moreover,

$$F^*_{\partial f}\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}\colon (\rho^*,\rho) \mapsto \begin{cases} -1, & \text{if } \rho^* \le -1/\rho < 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{39}$$

Proof. It is well known that $f^*\colon \rho^* \mapsto -1 + f(-\rho^*)$. Consequently, Theorem 2.6 implies that $\left]0,+\infty\right[\times \left]-\infty,0\right[\subset \operatorname{dom} F_{\partial f} \subset \left[0,+\infty\right[\times \left]-\infty,0\right]$. Thus, let us fix $(\rho,\rho^*) \in \left[0,+\infty\right[\times \left]-\infty,0\right]$. Then

$$F_{\partial f}(\rho,\rho^*) = \sup_{\xi>0} \big(\xi\rho^* + \rho(-1/\xi) - \xi(-1/\xi)\big) = 1 + \sup_{\xi>0}\big(\xi\rho^* - \rho/\xi\big). \tag{40}$$

Considering three cases (namely, $\rho = 0$, $\rho^* = 0$, and $\rho\rho^* \neq 0$) and some calculus yield (38). After some further calculus, one arrives at (39). $\square$

Remark 3.5.

(i) In the setting of Example 3.4, the refined Fenchel-Young inequality (12) is equivalent to

$$(\forall \rho \ge 0)(\forall \sigma \ge 0) \quad -\rho\sigma \le 1 - 2\sqrt{\rho\sigma} \le -\ln(\rho) - 1 - \ln(\sigma), \tag{41}$$

which is, in turn, equivalent to $(\forall \tau \ge 0)\ -\tau \le 1 - 2\sqrt{\tau} \le -1 - \ln(\tau)$.

(ii) Furthermore, Example 3.4 shows that

$$\operatorname{dom} f \times \operatorname{dom} f^* \subsetneqq \operatorname{dom} F_{\partial f} = \overline{\operatorname{dom}}\, f \times \overline{\operatorname{dom}}\, f^* \tag{42}$$

can occur in Theorem 2.6.

The negative entropy is an important function in various branches of mathematical sciences; however, its corresponding Fitzpatrick appears to be relatively involved.

Example 3.6 (negative entropy). Suppose that $X = \mathbb{R}$ and let

$$f\colon X \to \left]-\infty,+\infty\right] : \rho \mapsto \begin{cases} +\infty, & \text{if } \rho < 0; \\ 0, & \text{if } \rho = 0; \\ \rho\ln(\rho) - \rho, & \text{if } \rho > 0. \end{cases} \tag{43}$$

Denote the inverse of the function $\left[0,+\infty\right[\to \left[0,+\infty\right[: \rho \to \rho e^{\rho}$ by W. The function W is known as the *Lambert W function*; see [4] for further information. Then

$$F_{\partial f}\colon (\rho,\rho^*) \mapsto \begin{cases} +\infty, & \text{if } \rho < 0; \\ \exp(\rho^* - 1), & \text{if } \rho = 0; \\ \rho\rho^* + \rho\big(W(\kappa) + \frac{1}{W(\kappa)} - 2\big), & \text{if } \rho > 0 \text{ and } \kappa = \rho e^{1-\rho^*}. \end{cases} \tag{44}$$

Proof. Note that $\operatorname{dom} f = [0, +\infty[$ and that $\operatorname{dom} f^* = \mathbb{R}$, since $f^* = \exp$. Since both domains are closed, Theorem 2.6 implies that

$$\operatorname{dom} F_{\partial f} = [0, +\infty[\times \mathbb{R}. \tag{45}$$

This establishes the first case in (44). Recall that $\partial f(\xi) = \varnothing$, if $\xi \leq 0$; $\partial f(\xi) = \{\ln(\xi)\}$, if $\xi > 0$. Thus, letting $(\rho, \rho^*) \in [0, +\infty[\times \mathbb{R}$, we obtain

$$F_{\partial f}(\rho, \rho^*) = \sup_{\xi > 0} \big(\rho \ln(\xi) + \rho^* \xi - \xi \ln(\xi)\big). \tag{46}$$

If $\rho = 0$, then $F_{\partial f}(0, \rho^*) = f^*(\rho^* - 1) = \exp(\rho^* - 1)$ and thus the second case in (44) is verified. Suppose that $\rho > 0$. Then the (strictly concave) function we supremize over in (46) has a unique maximizer which must be a critical point, i.e., it satisfies

$$\frac{\rho}{\xi} + \rho^* - 1 - \ln(\xi) = 0. \tag{47}$$

We now show that

$$\zeta = e^{\rho^* - 1 + W(\rho e^{1-\rho^*})} \tag{48}$$

satisfies (47). Using the definition of ζ for (49) and (51) and the definition of W for (50), we indeed obtain

$$\begin{aligned}
\ln(\zeta) &= \rho^* - 1 + W(\rho e^{1-\rho^*}) && (49)\\
&= \rho^* - 1 + \frac{\rho e^{1-\rho^*}}{e^{W(\rho e^{1-\rho^*})}} && (50)\\
&= \rho^* - 1 + \frac{\rho e^{1-\rho^*}}{\zeta e^{1-\rho^*}} && (51)\\
&= \rho^* - 1 + \frac{\rho}{\zeta}. && (52)
\end{aligned}$$

Letting $\kappa = \rho e^{1-\rho^*}$, we observe that

$$e^{\rho^* - 1 + W(\kappa)} = \frac{e^{\rho^* - 1}}{W(\kappa)} W(\kappa) e^{W(\kappa)} = \frac{e^{\rho^* - 1}}{W(\kappa)} \kappa = \frac{e^{\rho^* - 1}}{W(\kappa)} \rho e^{1-\rho^*} = \frac{\rho}{W(\kappa)}. \tag{53}$$

Therefore,

$$\begin{aligned}
F_{\partial f}(\rho, \rho^*) &= \rho\big(\rho^* - 1 + W(\kappa)\big) + \rho^* e^{\rho^* - 1 + W(\kappa)} - e^{\rho^* - 1 + W(\kappa)}\big(\rho^* - 1 + W(\kappa)\big) && (54)\\
&= \rho\big(\rho^* - 1 + W(\kappa)\big) + e^{\rho^* - 1 + W(\kappa)}\big(1 - W(\kappa)\big) && (55)\\
&= \rho\big(\rho^* - 1 + W(\kappa)\big) + \frac{\rho}{W(\kappa)}\big(1 - W(\kappa)\big) && (56)\\
&= \rho\rho^* + \rho\Big(W(\kappa) + \frac{1}{W(\kappa)} - 2\Big), && (57)
\end{aligned}$$

which completes the proof. $\square$

The discussion of a quadratic function requires some preliminary work.

Proposition 3.7. *Let X be a real Hilbert space and let $A\colon X \to X$ be continuous, positive semidefinite, linear, and symmetric. Define $q_A\colon X \to \mathbb{R}\colon x \mapsto \frac{1}{2}\langle x | Ax\rangle$. Then*

(*i*) $\nabla q_A = A$.
(*ii*) $q_A^* \circ A = q_A$.
(*iii*) $\operatorname{ran} A \subset \operatorname{dom} q_A^* \subset \overline{\operatorname{ran}}\, A$.
(*iv*) *If* $\operatorname{ran} A$ *is closed and* $A^\dagger$ *denotes the Moore-Penrose inverse of* A *(see* [7] *for further information), then* $\operatorname{dom} q_A^* = \operatorname{ran} A$ *and* $q_A^*|_{\operatorname{ran} A} = q_{A^\dagger}|_{\operatorname{ran} A}$.

Proof. (*i*) and (*ii*): See, e.g., [2, Theorem 3.6.(i)]. (*iii*): (See also [1, Proposition 12.3.6(iii)].) Since $\operatorname{dom} q_A = X$, it is clear from (*ii*) that $\operatorname{ran} A \subset \operatorname{dom} q_A^*$. Now let $x^* \in \operatorname{dom} q_A^*$, define $g\colon X \to \mathbb{R}\colon x \mapsto q_A(x) - \langle x|x^*\rangle$, and observe that $q_A^*(x^*) = -\inf_{x\in X} g(x)$. Hence g is bounded below. Using (*i*) and a well known consequence of Ekeland's variational principle (see, e.g., [20, Corollary 1.4.3]), we deduce that there exists a sequence $(x_n)_{n\in\mathbb{N}}$ such that $\nabla g(x_n) = \nabla q_A(x_n) - x^* = Ax_n - x^* \to 0$. Therefore, $x^* \in \overline{\operatorname{ran}}\, A$. (*iv*) Suppose that $\operatorname{ran} A$ is closed. Then (*iii*) implies that $\operatorname{dom} q_A^* = \operatorname{ran} A$. Now let $y \in \operatorname{ran} A$, say $y = Ax$. On the one hand, (*ii*) yields $q_A^*(y) = q_A^*(Ax) = q_A(x) = \frac{1}{2}\langle x|Ax\rangle$. On the other hand, using standard properties of the Moore-Penrose inverse [7], we deduce that $q_{A^\dagger}(y) = \frac{1}{2}\langle y|A^\dagger y\rangle = \frac{1}{2}\langle Ax|A^\dagger Ax\rangle = \frac{1}{2}\langle x|AA^\dagger Ax\rangle = \frac{1}{2}\langle x|Ax\rangle$. Altogether, $q_A^*(y) = q_{A^\dagger}(y)$. $\square$

Example 3.8 (quadratic function). Suppose X is a real Hilbert space, let $A\colon X \to X$ be a continuous positive semidefinite linear symmetric operator with closed range, and define $q_A\colon X \to \mathbb{R}\colon x \mapsto \frac{1}{2}\langle x|Ax\rangle$. Then

$$F_A\colon X \times X \to \left]-\infty, +\infty\right] : (y, y^*) \mapsto 2q_A^*\big(\tfrac{1}{2}y^* + \tfrac{1}{2}Ay\big) \tag{58}$$

and hence $\operatorname{dom} F_A = X \times \operatorname{ran} A$. Furthermore:

(i) If A is a bijection, then $F_A\colon X\times X \to \left]-\infty, +\infty\right] : (y, y^*) \mapsto \frac{1}{4}\langle y + A^{-1}y^*|y^* + Ay\rangle$.
(ii) If A is positive definite and $(y, y^*) \in \operatorname{dom} F_A$, then $F_A(y, y^*) = q_A(y) + q_A^*(y^*)$ precisely when $y^* = Ay$.

Proof. Fix $(y, y^*) \in X \times X$. Then (58) is verified by

$$F_A(y, y^*) = \sup_{x\in X}\big(\langle x|y^*\rangle + \langle y|Ax\rangle - \langle x|Ax\rangle\big) \tag{59}$$

$$= \sup_{x\in X}\big(\langle x|y^* + Ay\rangle - 2q_A(x)\big) \tag{60}$$

$$= 2\sup_{x\in X}\big(\langle x|\tfrac{1}{2}y^* + \tfrac{1}{2}Ay\rangle - q_A(x)\big) \tag{61}$$

$$= 2q_A^*\big(\tfrac{1}{2}y^* + \tfrac{1}{2}Ay\big). \tag{62}$$

As $\operatorname{dom} q_A^* = \operatorname{ran} A$ according to Proposition 3.7(*iii*), we deduce that $\operatorname{dom} F_A = X \times \operatorname{ran} A$. (i): Now assume that A is also a bijection. Then (59)–(62) and Proposition 3.7(*iv*) result in

$$F_A(y, y^*) = 2q_A^*\big(\tfrac{1}{2}y^* + \tfrac{1}{2}Ay\big) = 2q_{A^{-1}}\big(\tfrac{1}{2}y^* + \tfrac{1}{2}Ay\big) = \tfrac{1}{4}\langle y^* + Ay|A^{-1}y^* + y\rangle. \tag{63}$$

(ii): Finally, assume that A is positive definite, that $(y, y^*) \in \operatorname{dom} F_A = X \times \operatorname{ran} A$, say $y^* = Az$, and that $F_A(y, y^*) = q_A(y) + q_A^*(y^*)$. Using Proposition 3.7(*ii*) and (58), we deduce that

$$q_A(y) + q_A(z) = q_A(y) + q_A^*(Az) = q_A(y) + q_A^*(y^*) = F_A(y, y^*) = F_A(y, Az) \tag{64}$$

$$= 2q_A^*\big(\tfrac{1}{2}Az + \tfrac{1}{2}Ay\big) = 2q_A^*\big(A(\tfrac{1}{2}z + \tfrac{1}{2}y)\big) = 2q_A\big(\tfrac{1}{2}z + \tfrac{1}{2}y\big) \tag{65}$$

$$= \langle \tfrac{1}{2}z + \tfrac{1}{2}y|\tfrac{1}{2}Az + \tfrac{1}{2}Ay\rangle. \tag{66}$$

Thus

$$2\langle y|Ay\rangle + 2\langle z|Az\rangle = \langle z+y|Az+Ay\rangle = \langle z|Az\rangle + \langle z|Ay\rangle + \langle y|Az\rangle + \langle y|Ay\rangle\,; \tag{67}$$

equivalently,

$$0 = \langle y-z|Ay-Az\rangle = \langle y-z|A(y-z)\rangle\,. \tag{68}$$

Since A is positive definite, we conclude that $y = z$. Therefore, $y^* = Az = Ay$. □

Remark 3.9. Example 3.8(ii) illustrates that positive definite quadratic functions are similarly "nice" in the sense that among the points $(y, y^*) \in \operatorname{dom} F_A = X \times \operatorname{ran} A$, only the ones on the graph of A have the property that $F_A(y, y^*) = q_A(y) + q_A^*(y^*)$. See also Remark 2.2(iv) and Theorem 2.3.

Example 3.10 (energy). Suppose that X is a real Hilbert space and that $j\colon X \to \mathbb{R}\colon x \mapsto \frac{1}{2}\|x\|^2$ is the energy. Then $\nabla j = \mathrm{Id}$ and

$$F_{\mathrm{Id}}\colon X \times X \to \mathbb{R}\colon (y, y^*) \mapsto \tfrac{1}{4}\|y + y^*\|^2. \tag{69}$$

Proof. This is a direct consequence of Example 3.8(i). □

Remark 3.11. In the setting of Example 3.10, let $(y, y^*) \in X \times X$. We further compute that $F^*_{\mathrm{Id}}(y^*, y) = \|y\|^2 = \|y^*\|^2$, if $y = y^*$; $F_{\mathrm{Id}}(y^*, y) = +\infty$, otherwise. Thus, $F_{\mathrm{Id}}(y, y^*) = \frac{1}{2}\langle y|y^*\rangle + \frac{1}{2}\big(j(y) + j^*(y^*)\big)$ and $F^*_{\mathrm{Id}}(y^*, y) = \frac{1}{2}\big(j(y) + j^*(y^*)\big) + \frac{1}{2}\big(\langle y|y^*\rangle + \iota_{\mathrm{Id}}(y, y^*)\big)$. In this case, the Fitzpatrick function and its Fenchel conjugate are exactly the averages of the neighboring functions in the chain of inequalities (5). (This is false in general; see, e.g., Example 3.4.)

Our next goal is to compute the Fitzpatrick function for the subdifferential of the sum of the energy and the indicator function. This requires some preparation.

Proposition 3.12. *Suppose that X is a real Hilbert space, that $j = \frac{1}{2}\|\cdot\|^2$ is the energy function, and that C is a nonempty closed convex set in X. Let $z \in X$. Then $(j+\iota_C)^*(z) = \langle z|P_C z\rangle - j(P_C z)$. In particular, the following is true.*

(i) If C is a nonempty closed convex cone, then $(j + \iota_C)^*(z) = \frac{1}{2}\|P_C z\|^2$.

(ii) If C is the closed unit ball, then $(j + \iota_C)^*(z) = \begin{cases} \frac{1}{2}\|z\|^2, & \text{if } \|z\| \le 1; \\ \|z\| - \frac{1}{2}, & \text{otherwise.} \end{cases}$

Proof. By definition of the Fenchel conjugate, we have

$$(j + \iota_C)^*(z) = \sup_{x\in X}\big(\langle x|z\rangle - j(x) - \iota_C(x)\big) = -\inf_{x\in X}\big(\langle x|-z\rangle + j(x) + \iota_C(x)\big). \tag{70}$$

A point $x \in X$ attains the supremum if and only if it is a critical point, i.e., $0 \in -z + x + N_C(x)$, which in turn shows that $x = P_C z$. We deduce that $(j+\iota_C)^*(z) = \langle z|P_C z\rangle - j(P_C z)$. (i): If C is a nonempty closed convex cone, then $\langle P_C z|z - P_C z\rangle = 0$ and the formula follows. (ii): If C is the closed unit ball, then use $P_C z = z$, if $z \in C$; $P_C z = z/\|z\|$, otherwise, to obtain the result. □

Example 3.13 (energy plus indicator). Suppose that X is a real Hilbert space, that $j = \frac{1}{2}\|\cdot\|^2$ is the energy function, and that C is a nonempty closed convex set in X. Let $f = j + \iota_C$, and let $(y, y^*) \in X \times X$. Then

$$F_{\partial f}(y,y^*) = \iota_C(y) + 2(j+\iota_C)^*\big(\tfrac{1}{2}y + \tfrac{1}{2}y^*\big) \tag{71}$$

$$= \iota_C(y) + \tfrac{1}{4}\|y+y^*\|^2 - \big(\tfrac{1}{4}\|\cdot\|^2 + \iota_C^*\big)^*\big(\tfrac{1}{2}y + \tfrac{1}{2}y^*\big) \tag{72}$$

$$= \iota_C(y) + \langle P_C(\tfrac{1}{2}y + \tfrac{1}{2}y^*) | y + y^* - P_C(\tfrac{1}{2}y + \tfrac{1}{2}y^*)\rangle\,. \tag{73}$$

This simplifies in the following particular cases.

(i) If C is a nonempty closed convex cone, then $F_{\partial f}(y,y^*) = \iota_C(y) + \|P_C(\frac{1}{2}y + \frac{1}{2}y^*)\|^2$.

(ii) If C is the closed unit ball, then $F_{\partial f}(y,y^*) = \iota_C(y) + \begin{cases} \frac{1}{4}\|y+y^*\|^2, & \text{if } \|y+y^*\| \le 2; \\ \|y+y^*\| - 1, & \text{otherwise.} \end{cases}$

Moreover, in the general case, we have $F_{\partial f}(y,y^*) = f(y) + f^*(y^*) \in \mathbb{R}$ if and only if $(y,y^*) \in \partial f$.

Proof. Since $\partial f = \mathrm{Id} + N_C$, we note that

$$F_{\partial f}(y,y^*) = \sup_{x\in C}\big(\langle x|y^*\rangle + \langle x|y-x\rangle + \sup\langle y-x|N_C(x)\rangle\big). \tag{74}$$

If $y \in C$, then $(\forall x \in C)$ $\sup\langle y-x|N_C(x)\rangle = 0$. Otherwise, if $y \notin C$, then $P_Cy \in C$ and $\sup\langle y - P_Cy|N_C(P_Cy)\rangle = +\infty$ (because $y - P_Cy \in N_C(P_Cy) \smallsetminus \{0\}$ and $N_C(P_Cy)$ is a cone), which implies that $F_{\partial f}(y,y^*) = +\infty$. Thus, for the remainder of this proof, we assume that

$$y \in C. \tag{75}$$

Then (74) simplifies to

$$F_{\partial f}(y,y^*) = \sup_{x\in C}\big(\langle x|y^*+y\rangle - \|x\|^2\big) = 2\sup_{x\in C}\big(\langle x|\tfrac{1}{2}y^* + \tfrac{1}{2}y\rangle - j(x)\big) \tag{76}$$

$$= 2\sup_{x\in X}\big(\langle x|\tfrac{1}{2}y^* + \tfrac{1}{2}y\rangle - \big(j+\iota_C)(x)\big) = 2(j+\iota_C)^*\big(\tfrac{1}{2}y^* + \tfrac{1}{2}y\big). \tag{77}$$

This verifies (71). Observe that $\big(\frac{1}{2}(\frac{1}{2}j + \iota_C^*)^*\big)^* = \frac{1}{2}\big(\frac{1}{2}j + \iota_C^*\big)(2\cdot) = j + \iota_C^*$, which, after taking the Fenchel conjugate, results in $\frac{1}{2}(\frac{1}{2}j + \iota_C^*)^* = (j+\iota_C^*)^*$. This and the Moreau decomposition [11, Section 7] yield $j = (j\Box\iota_C) + (j\Box\iota_C^*) = (j+\iota_C^*)^* + (j+\iota_C)^* = \frac{1}{2}(\frac{1}{2}j + \iota_C^*)^* + (j+\iota_C)^*$, where here and elsewhere "$\Box$" denotes *infimal convolution.* Hence

$$2(j+\iota_C)^* = 2j - (\tfrac{1}{2}j + \iota_C^*)^*. \tag{78}$$

Now (77) and (78) imply (72). Furthermore, (73) is obtained from (71) and Proposition 3.12. Items (i) and (ii) follow from their counterparts in Proposition 3.12 and from (71).

Set $z = \frac{1}{2}y + \frac{1}{2}y^*$. Using (75), (76)–(77), Proposition 3.12, and some re-arranging, we deduce that

$$f(y) + f^*(y^*) - F_{\partial f}(y,y^*) = j(y) + \langle y^*|P_Cy^*\rangle - j(P_Cy^*) - \langle P_Cz|2z - P_Cz\rangle \tag{79}$$

$$= j(y - P_Cz) + \big(j(y^* - P_Cz) - j(y^* - P_Cy^*)\big) \tag{80}$$

$$\ge 0. \tag{81}$$

Furthermore, equality holds in (81) if and only if

$$y = P_C z \quad \text{and} \quad P_C z = P_C y^*. \tag{82}$$

Altogether, we see that

$$\begin{aligned} f(y) + f^*(y^*) = F_{\partial f}(y, y^*) &\Leftrightarrow \text{equality holds in (81)} \Leftrightarrow \text{(82) is true} && (83)\\ &\Leftrightarrow P_C y^* = y \Leftrightarrow y^* \in P_C^{-1} y \Leftrightarrow y^* \in (\mathrm{Id} + N_C) y && (84)\\ &\Leftrightarrow y^* \in \partial f(y), && (85) \end{aligned}$$

which completes the proof. □

Remark 3.14.

(i) The "Moreover" part in Example 3.13 shows that $j + \iota_C$ is as "nice" as quadratic functions (see Remark 3.9) or the smooth functions of Theorem 2.3 (see also Remark 2.2(iv)).

(ii) Refinements (12) of the Fenchel-Young inequality in the context of Example 3.13 arise for various choices of C; e.g., if C is a nonempty closed convex cone in X, then Example 3.13(i) yields

$$(\forall c \in C)(\forall z \in X) \quad \langle c|z\rangle \leq \tfrac{1}{4}\|P_C(c+z)\|^2 \leq \tfrac{1}{2}\|c\|^2 + \tfrac{1}{2}\|P_C z\|^2. \tag{86}$$

Example 3.15. Suppose that X is a real Hilbert space, that $j = \frac{1}{2}\|\cdot\|^2$ is the energy function, and that C is a nonempty closed convex set in X. Let $f = j - \frac{1}{2}d_C^2$, and let $(y, y^*) \in X \times X$. Then $\nabla f = P_C$ and

$$F_{P_C}(y, y^*) = \iota_C(y^*) + \left\langle P_C(\tfrac{1}{2}y + \tfrac{1}{2}y^*) | y + y^* - P_C(\tfrac{1}{2}y + \tfrac{1}{2}y^*)\right\rangle, \tag{87}$$

which simplifies in the following particular cases.

(i) If C is a nonempty closed convex cone, then $F_{P_C}(y, y^*) = \iota_C(y^*) + \|P_C(\frac{1}{2}y + \frac{1}{2}y^*)\|^2$.

(ii) If C is the closed unit ball, then $F_{P_C}(y, y^*) = \iota_C(y^*) + \begin{cases} \frac{1}{4}\|y+y^*\|^2, & \text{if } \|y+y^*\| \leq 2;\\ \|y+y^*\| - 1, & \text{otherwise.}\end{cases}$

Proof. This is an immediate consequence of Example 3.13 because $P_C = (\mathrm{Id} + N_C)^{-1}$. □

Remark 3.16. Suppose that X is a real Hilbert space and that K is a nonempty closed convex cone in X, with polar cone $K^\ominus$. Let us show that

$$F_{P_K}^* : X \times X \to \left]-\infty, +\infty\right] : (y^*, y) \mapsto \iota_K(y^*) + \|y^*\|^2 + \iota_{K^\ominus}(y - y^*). \tag{88}$$

Fix $(y^*, y) \in X \times X$. If $y^* \notin K$, then $F_{P_K}^*(y^*, y) = +\infty$ since $\operatorname{dom} F_{P_K} = X \times K$ by Example 3.15 and since $F_{P_K}^*(y^*, y) \geq F_{P_K}(y, y^*)$ by Fact 1.2(*iii*). So assume that $y^* \in K$. Then $x \in X$ is a critical point for the optimization problem

$$\sup_{x \in X} \left(\langle x|y^*\rangle - \tfrac{1}{4}\|P_K(x + x^*)\|^2 \right) = \sup_{x \in X} \left(\langle x|y^*\rangle - \tfrac{1}{2}\tfrac{1}{2}d_{K^\ominus}^2(x + x^*) \right) \tag{89}$$

if and only if $y^* = P_K\left(\frac{1}{2}x + \frac{1}{2}x^*\right)$ because $\nabla \frac{1}{2}d_{K^\ominus}^2 = \mathrm{Id} - P_{K^\ominus} = P_K$. It follows that if $x \in X$ is such a critical point, then $\sup_{x \in X} \left(\langle x|y^*\rangle - \frac{1}{4}\|P_K(x + x^*)\|^2 \right) = \langle x|y^*\rangle - \|y^*\|^2 =$

$2\langle\frac{1}{2}x+\frac{1}{2}x^*|y^*\rangle-\langle x^*|y^*\rangle-\|y^*\|^2=2\langle y^*|y^*\rangle-\langle x^*|y^*\rangle-\|y^*\|^2=\|y^*\|^2-\langle x^*|y^*\rangle$. Using this, (89), and Example 3.15(i), we deduce that

$$F^*_{P_K}(y^*,y)=\sup_{x^*\in X}\big(\langle y|x^*\rangle-\iota_K(x^*)+\sup_{x\in X}\big(\langle x|y^*\rangle-\tfrac{1}{4}\|P_K(x+x^*)\|^2\big)\big) \tag{90}$$

$$=\sup_{x^*\in X}\big(\langle y|x^*\rangle-\iota_K(x^*)+\|y^*\|^2-\langle y^*|x^*\rangle\big) \tag{91}$$

$$=\|y^*\|^2+\sup_{x^*\in X}\big(\langle y-y^*|x^*\rangle-\iota_K(x^*)\big) \tag{92}$$

$$=\|y^*\|^2+\iota_{K^\ominus}(y-y^*). \tag{93}$$

This verifies (88). As an illustration, suppose that $K=X$. Then $P_K=\mathrm{Id}$, $K^\ominus=\{0\}$ and (88) becomes $F^*_{\mathrm{Id}}(y^*,y)=\|y^*\|^2+\iota_{\{0\}}(y-y^*)$, a formula consistent with Remark 3.11.

We conclude this section with an example of a Fitzpatrick function on the real line that allows an explicit description.

Example 3.17. Suppose $X=\mathbb{R}$ and let $f=\frac{1}{3}|\cdot|^3$. Then $f'\colon\mathbb{R}\to\mathbb{R}\colon\rho\to\rho|\rho|$ and

$$F_{f'}\colon(\rho,\rho^*)\mapsto\max\big\{\alpha(\rho,\rho^*),\alpha(-\rho,-\rho^*)\big\}, \tag{94}$$

where $\alpha(\rho,\rho^*)=\max_{\xi\geq 0}\big(-\xi^3+\rho\xi^2+\rho^*\xi\big)$.

Remark 3.18. Let α be as in Example 3.17, let $(\rho,\rho^*)\in\mathbb{R}^2$, and set $\xi=\frac{1}{3}\big(\rho+\sqrt{\rho^2+3\rho^*}\big)$. A careful discussion shows that

$$\alpha(\rho,\rho^*)=\begin{cases}0, & \text{if } \rho\leq 0 \text{ and } \rho^*\leq 0;\\ 0, & \text{if } \rho>0 \text{ and } \rho^*\leq -\frac{1}{4}\rho^2;\\ -\xi^3+\rho\xi^2+\rho^*\xi>0, & \text{otherwise.}\end{cases} \tag{95}$$

4. The Fitzpatrick function of a sum

Let A and B be two monotone operators from X to 2^{X^*}. Fitzpatrick's [6, Problem 5.4] asks to characterize F_{A+B}. This problem does not appear to have a simple solution. Nonetheless, an upper bound is available and we illustrate it by utilizing some of the examples of Section 3.

Definition 4.1. For two given monotone operators A and B from X to 2^{X^*}, we define $\Phi_{\{A,B\}}$ by

$$\Phi_{\{A,B\}}\colon X\times X^*\to\,]-\infty,+\infty]:(y,y^*)\mapsto\big(F_A(y,\cdot)\Box F_B(y,\cdot)\big)(y^*). \tag{96}$$

Proposition 4.2. *Suppose that A and B are two monotone operators from X to 2^{X^*}. Then*

$$F_{A+B}\leq\Phi_{\{A,B\}}. \tag{97}$$

Proof. Let $(y,y^*)\in X\times X^*$ and $(x,x^*)\in(A+B)$, say $x^*=a^*+b^*$, where $a^*\in Ax$ and $b^*\in Bx$. Also, write $y^*=u^*+v^*$, where $\{u^*,v^*\}\subset X^*$. Then

$$\langle x|y^*\rangle+\langle y|x^*\rangle-\langle x|x^*\rangle \tag{98}$$

$$=\langle x|u^*+v^*\rangle+\langle y|a^*+b^*\rangle-\langle x|a^*+b^*\rangle \tag{99}$$

$$=\big(\langle x|u^*\rangle+\langle y|a^*\rangle-\langle x|a^*\rangle\big)+\big(\langle x|v^*\rangle+\langle y|b^*\rangle-\langle x|b^*\rangle\big) \tag{100}$$

$$\leq F_A(y,u^*)+F_B(y,v^*). \tag{101}$$

Supremizing over $(x, x^*) \in A + B$ results in $F_{A+B}(y, y^*) \le F_A(y, u^*) + F_B(y, v^*)$. In turn, infimizing over $u^* + v^* = y^*$ verifies (97). □

Remark 4.3. The upper bound $\Phi_{\{A,B\}}$ provided in Proposition 4.2 is sometimes — but not always — tight as the remainder of this section shows. It would be interesting to characterize the pairs of monotone operators (A, B) that satisfy the identity $F_{A+B} = \Phi_{\{A,B\}}$.

Example 4.4. Suppose that X is a real Hilbert space and that C and D are two closed convex sets in X such that the constraint qualification $C \cap \operatorname{int} D \neq \varnothing$ holds. Then $F_{N_C+N_D} = \Phi_{\{N_C,N_D\}}$.

Proof. Let $(y, y^*) \in X \times X$. Using Example 3.1 for (104) and (108) and the constraint qualification for (106) and (109) (see, e.g., [20, Theorem 2.8.7]), we obtain

$$\begin{aligned}
\Phi_{\{N_C,N_D\}}(y, y^*) &= \big(F_{N_C}(y,\cdot)\Box F_{N_D}(y,\cdot)\big)(y^*) && (102)\\
&= \inf_{u^*+v^*=y^*} \big(F_{N_C}(y, u^*) + F_{N_D}(y, v^*)\big) && (103)\\
&= \inf_{u^*+v^*=y^*} \big(\iota_C(y) + \iota_C^*(u^*) + \iota_D(y) + \iota_D^*(v^*)\big) && (104)\\
&= \iota_C(y) + \iota_D(y) + \big(\iota_C^*\Box\iota_D^*\big)(y^*) && (105)\\
&= \iota_{C\cap D}(y) + \big(\iota_D + \iota_D\big)^*(y^*) && (106)\\
&= \iota_{C\cap D}(y) + \iota_{C\cap D}^*(y^*) && (107)\\
&= F_{N_{C\cap D}}(y, y^*) && (108)\\
&= F_{N_C+N_D}(y, y^*). && (109)
\end{aligned}$$

Therefore, $F_{N_C+N_D} = \Phi_{\{N_C,N_D\}}$. □

Example 4.5. Suppose that X is a real Hilbert space and that C is a nonempty closed convex set in X. Then $F_{\mathrm{Id}+N_C} = \Phi_{\{\mathrm{Id},N_C\}}$

Proof. This is a continuation of Example 3.13, the notation of which we borrow here. Let $(y, y^*) \in X \times X$. Example 3.10 and Example 3.1 imply

$$\begin{aligned}
\Phi_{\{\mathrm{Id},N_C\}}(y, y^*) &= \big(F_{\mathrm{Id}}(y,\,\cdot\,)\Box F_{N_C}(y,\,\cdot\,)\big)(y^*) && (110)\\
&= \inf_{u^*+v^*=y^*} \big(F_{\mathrm{Id}}(y, u^*) + F_{N_C}(y, v^*)\big) && (111)\\
&= \inf_{u^*+v^*=y^*} \big(\tfrac14\|u^* + y\|^2 + \iota_C(y) + \iota_C^*(v^*)\big). && (112)
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&\inf_{u^*+v^*=y^*} \big(\tfrac14\|u^* + y\|^2 + \iota_C^*(v^*)\big) && (113)\\
&= \inf_{z^*} \big(\tfrac14\|(y + y^*) - z^*\|^2 + \iota_C^*(z^*)\big) && (114)\\
&= \inf_{z^*} \big(\tfrac14\big(\|y + y^*\|^2 - 2\,\langle y + y^*|z^*\rangle + \|z^*\|^2\big) + \iota_C^*(z^*)\big) && (115)\\
&= \tfrac14\|y + y^*\|^2 + \inf_{z^*}\big(- \langle\tfrac12(y + y^*)|z^*\rangle + \tfrac14\|z^*\|^2 + \iota_C^*(z^*)\big) && (116)\\
&= \tfrac14\|y + y^*\|^2 - \sup_{z^*}\big(\langle\tfrac12(y + y^*)|z^*\rangle - \big(\tfrac14\|\cdot\|^2 + \iota_C^*\big)(z^*)\big) && (117)\\
&= \tfrac14\|y + y^*\|^2 - \big(\tfrac14\|\cdot\|^2 + \iota_C^*\big)^*\big(\tfrac12 y + \tfrac12 y^*\big). && (118)
\end{aligned}$$

Combining (110)–(112), (113)–(118), and (71)–(72) results in

$$\begin{aligned}\Phi_{\{\mathrm{Id},N_C\}}(y,y^*) &= \inf_{u^*+v^*=y^*}\big(\tfrac14\|u^*+y\|^2+\iota_C(y)+\iota_C^*(v^*)\big) &(119)\\ &= \iota_C(y)+\tfrac14\|y+y^*\|^2-\big(\tfrac14\|\cdot\|^2+\iota_C^*\big)^*\big(\tfrac12 y+\tfrac12 y^*\big) &(120)\\ &= F_{\mathrm{Id}+N_C}(y,y^*). &(121)\end{aligned}$$

Therefore, $F_{\mathrm{Id}+N_C}=\Phi_{\{\mathrm{Id},N_C\}}$. □

Example 4.6. Suppose that X is a real Hilbert space and that K is a closed subspace of X. Then $F_{\mathrm{Id}}=F_{P_K+P_{K^\perp}}=\Phi_{\{P_K,P_{K^\perp}\}}$.

Proof. Let $(y,y^*)\in X\times X$. Example 3.15(i) and Example 3.10 imply

$$\begin{aligned}&\Phi_{\{P_K,P_{K^\perp}\}}(y,y^*) &(122)\\ &= \big(F_{P_K}(y,\cdot)\Box F_{P_{K^\perp}}(y,\cdot)\big)(y^*) = \inf_{z^*\in X}\big(F_{P_K}(y,z^*)+F_{P_{K^\perp}}(y,y^*-z^*)\big) &(123)\\ &= \inf_{z^*\in X}\big(\iota_K(z^*)+\tfrac14\|P_K(y+z^*)\|^2+\iota_{K^\perp}(y^*-z^*)+\tfrac14\|P_{K^\perp}(y+y^*-z^*)\|^2\big) &(124)\\ &= \tfrac14\|P_K(y+P_Ky^*)\|^2+\tfrac14\|P_{K^\perp}(y+P_{K^\perp}y^*)\|^2=\tfrac14\big(\|P_K(y+y^*)\|^2+\|P_{K^\perp}(y+y^*)\|^2\big) &(125)\\ &= \tfrac14\|y+y^*\|^2=F_{\mathrm{Id}}(y,y^*)=F_{P_K+P_{K^\perp}}(y,y^*). &(126)\end{aligned}$$

Therefore, the proof is complete. □

Our last result shows that the conclusion of Example 4.6 may fail if we work with cones rather than subspaces.

Example 4.7. Suppose that $X=\mathbb{R}$ and that $K=[0,+\infty[$. Then $F_{P_K+P_{-K}}=F_{\mathrm{Id}}\neq\Phi_{\{P_K,P_{-K}\}}$.

Proof. It is clear that $F_{P_K+P_{-K}}=F_{\mathrm{Id}}$. Now consider the point $(-1,1)\in\mathbb{R}^2$. By Example 3.10,

$$F_{\mathrm{Id}}(-1,1)=\tfrac14|-1+1|^2=0. \tag{127}$$

Utilizing Example 3.15, we obtain on the other hand

$$\begin{aligned}\Phi_{\{P_K,P_{-K}\}}(-1,1) &= \inf_{\rho\in\mathbb{R}}\big(\iota_K(\rho)+\tfrac14|P_K(-1+\rho)|^2+\iota_{-K}(1-\rho)+\tfrac14|P_{-K}(-1+1-\rho)|^2\big) &(128)\\ &= \tfrac14\inf_{\rho\geq1}\big(|P_K(\rho-1)|^2+|P_{-K}(-\rho)|^2\big) &(129)\\ &= \tfrac14\inf_{\rho\geq1}\big(|\rho-1|^2+|\rho|^2\big) &(130)\\ &= \tfrac14. &(131)\end{aligned}$$

Altogether, (127) and (128)–(131) imply that $F_{\mathrm{Id}}\neq\Phi_{\{P_K,P_{-K}\}}$. □

5. Fitzpatrick's problem: Introduction

From now on, we assume that X is a Euclidean space. In contrast to the previous sections, we shall emphasize the view point that monotone operators can be identified with their graphs, which we refer to as *monotone sets*. Definition 1.1 introduces the Fitzpatrick function associated with some underlying (usually) monotone set. In a somewhat "dual" spirit, Fitzpatrick also introduced a set, defined in terms of an underlying function.

Definition 5.1. Let $f\colon X \times X^* \to \left]-\infty, +\infty\right]$ be convex. Then the *Fitzpatrick set* associated with f is

$$G_f := \big\{(x, x^*) \in X \times X^* \mid (x^*, x) \in \partial f(x, x^*)\big\}. \tag{132}$$

Since $G_f \subset X \times X^*$, we may identify G_f with an operator from X to 2^{X^*}, as Fitzpatrick did originally (see [6, Definition 2.1]). In [6, Section 3], Fitzpatrick studied properties of the "composition" G_F and derived the following fundamental result.

Fact 5.2. *Let A be a nonempty monotone subset of $X \times X^*$. Then*

$$A \subset G_{F_A}. \tag{133}$$

Moreover,

$$A = G_{F_A}, \quad \textit{whenever } A \textit{ is maximal monotone.} \tag{134}$$

Proof. This is a restatement of [6, Corollary 3.5]. □

Fact 5.2 shows that every maximal monotone A is a "fixed point" of the composition G_F. However, Fitzpatrick also observed that this composition may have fixed points that are very far from being maximal monotone. Indeed, $A = \{(0, 0)\}$ has these properties; see [6, Example following Corollary 3.5]. Fitzpatrick's [6, Problem 5.2] specifically asks to characterize monotone sets A for which $A = G_{F_A}$.

The second objective of this paper is to provide a sufficient condition for this fixed point problem. This condition, which is formulated in terms of the polar of a monotone set introduced recently by Martínez-Legaz and Svaiter [8], is also necessary when $X = \mathbb{R}$. In fact, when $X = \mathbb{R}$, the condition requires that A be a nonempty *connected* monotone set. Maximal monotone operators and singletons are the two extreme cases of this condition in the sense that they are as large and as small as possible.

The remainder of the paper is organized as follows. In Section 6, we review and derive results that shall make the proofs in later sections more transparent. Section 7 contains the new sufficient condition for Fitzpatrick's fixed point problem. We then assume that $X = \mathbb{R}$. This particular setting allows us to show in Section 8 that the new sufficient condition is not only equivalent to connectedness but also necessary.

6. Fitzpatrick's problem: Auxiliary results

The following three results will be useful later.

Fact 6.1. *Let A be a maximal monotone subset of $X \times X^*$ and set $R_A := (\mathrm{Id} + A)^{-1}$. Then R_A is single-valued and the* Minty parameterization

$$M\colon X \to A\colon x \mapsto \big(R_A x, (\mathrm{Id} - R_A)x\big) \tag{135}$$

of A is bijective and continuous in both directions.

Proof. See, e.g., [14, Theorem 12.15]. $\square$

Proposition 6.2. *Let A be a nonempty monotone subset of $X \times X^*$ and define*

$$\mathcal{A}\colon X \times X^* \to 2^A\colon (x, x^*) \mapsto \big\{(a, a^*) \in A \mid F_A(x, x^*) = \langle x|a^*\rangle + \langle a|x^*\rangle - \langle a|a^*\rangle\big\}. \tag{136}$$

Let $(x, x^*) \in X \times X^*$. Then

$$\overline{\operatorname{conv}}\big\{(a^*, a) \mid (a, a^*) \in \mathcal{A}(x, x^*)\big\} \subset \partial F_A(x, x^*). \tag{137}$$

Moreover,

$$\overline{\operatorname{conv}}\big\{(a^*, a) \mid (a, a^*) \in \mathcal{A}(x, x^*)\big\} = \partial F_A(x, x^*), \quad \textit{whenever } A \textit{ is compact.} \tag{138}$$

Proof. Let us define, for every $(a, a^*) \in A$,

$$f_{(a,a^*)} : X \times X^* \to \mathbb{R} : (y, y^*) \mapsto \langle y|a^*\rangle + \langle a|y^*\rangle - \langle a|a^*\rangle \,. \tag{139}$$

The mapping $A \to \mathbb{R}\colon (a, a^*) \mapsto f_{(a,a^*)}(y, y^*)$ is continuous, for every $(y, y^*) \in X \times X^*$. Moreover, each $f_{(a,a^*)}$ is linear, hence convex and differentiable on $X \times X^*$, with gradient

$$\nabla f_{(a,a^*)}(x, x^*) = (a^*, a). \tag{140}$$

The inclusion (137) is well known (and easy to verify). When A is compact, the desired identity (138) follows from the Ioffe-Tikhomirov theorem; see, e.g., [20, Theorem 2.4.18]. $\square$

Proposition 6.3. *Let $f\colon X \times X^* \to \,]-\infty, +\infty]$ be convex and $(x, x^*) \in X \times X^*$. Then $(x, x^*) \in G_f$ if and only if $\frac{1}{2}f(x, x^*) + \frac{1}{2}f^*(x^*, x) = \langle x|x^*\rangle$.*

Proof. Using the definition of the duality product on the product space $X \times X^*$, we have

$$(x, x^*) \in G_f \Leftrightarrow (x^*, x) \in \partial f(x, x^*) \tag{141}$$
$$\Leftrightarrow f(x, x^*) + f^*(x^*, x) = \langle (x, x^*)|(x^*, x)\rangle \tag{142}$$
$$\Leftrightarrow \tfrac{1}{2}f(x, x^*) + \tfrac{1}{2}f^*(x^*, x) = \langle x|x^*\rangle \,, \tag{143}$$

which completes the proof. $\square$

We now turn to recent notions and results by Martínez-Legaz and Svaiter [8].

Definition 6.4. Let A be a nonempty monotone subset of $X \times X^*$. Then A is said to be *representable*, if there exists a lower semicontinuous convex function $h_A\colon X \times X^* \to \,]-\infty, +\infty]$ such that

$$h_A \geq p \quad \text{and} \quad A = \big\{(x, x^*) \in X \times X^* \mid h_A(x, x^*) = \langle x|x^*\rangle\big\}. \tag{144}$$

The *representable closure* of A is the intersection of all monotone extensions of A which are representable.

Definition 6.5. Let A be a nonempty subset of $X \times X^*$. Then the *polar (in the sense of Martínez-Legaz and Svaiter)* is given by

$$A^\mu := \big\{(x,x^*) \in X \times X^* \mid \inf_{(a,a^*)\in A} \langle x-a|x^*-a^*\rangle \geq 0\big\}. \tag{145}$$

Moreover, A is said to be *μ-closed* whenever A coincides with its *bipolar* $A^{\mu\mu} := (A^\mu)^\mu$.

It follows from Definition 6.5 that

$$A^\mu \text{ is closed;} \tag{146}$$

in particular, every μ-closed set is closed. Furthermore, Definition 1.1 (see also [8, equation (22)]) implies that

$$A^\mu = \big\{(x,x^*) \in X \times X^* \mid F_A(x,x^*) \leq p(x,x^*)\big\}. \tag{147}$$

An application of Zorn's Lemma (see also [8, Proposition 22]) shows that

$$A^\mu = \bigcup \big\{B \subset X \times X^* \mid A \subset B \text{ and } B \text{ is maximal monotone}\big\}, \tag{148}$$

which yields (see also [8, Proposition 21])

$$A^\mu = A, \;\text{ whenever } A \text{ is maximal monotone.} \tag{149}$$

See [8] for further properties of these new notions. We shall utilize the main result of Martínez-Legaz and Svaiter in the following form.

Fact 6.6. *Let A be a nonempty monotone subset of $X \times X^*$. Then A is representable if and only if A is μ-closed, in which case*

$$A = \bigcap \big\{B \subset X \times X^* \mid A \subset B \textit{ and } B \textit{ is maximal monotone}\big\}. \tag{150}$$

Proof. This follows from [8, Theorem 31, Corollary 32, and Corollary 33]. ☐

Corollary 6.7. *Let $f\colon X \times X^* \to\,]-\infty,+\infty]$ be convex, lower semicontinuous, and proper. Then G_f is representable, monotone, μ-closed, and hence closed.*

Proof. Letting $h_{G_f}\colon X \times X^* \to\,]-\infty,+\infty] : (x,x^*) \mapsto \frac{1}{2}f(x,x^*) + \frac{1}{2}f^*(x^*,x)$, we see that the conclusion follows from Proposition 6.3, Definition 6.4, Fact 6.6, and (146). ☐

Remark 6.8. Corollary 6.7 and Fact 5.2 imply that every maximal monotone set is representable. The representation is achieved by the corresponding Fitzpatrick function, which is in fact the (pointwise) infimum of all such representations; see [6, Theorem 3.10].

The next result states that the Fitzpatrick function is blind to taking the μ-closure.

Proposition 6.9. *Let A be a nonempty monotone subset of $X \times X^*$. Then $F_A = F_{A^{\mu\mu}}$.*

Proof. By [8, Corollary 14], we have $(\iota_A + p)^{**} = (\iota_{A^{\mu\mu}} + p)^{**}$. Taking the Fenchel conjugate yields $(\iota_A + p)^* = (\iota_{A^{\mu\mu}} + p)^*$. In turn, using Fact 1.2(*ii*), we conclude that $F_A = F_{A^{\mu\mu}}$. ☐

Proposition 6.10. *Let A be a nonempty monotone subset of $X \times X^*$ such that $A = G_{F_A}$. Then A is μ-closed and hence closed.*

Proof. Corollary 6.7 implies that G_{F_A} is μ-closed. Hence A is μ-closed and thus closed (by [8, Proposition 8] or by (146)). □

Remark 6.11. Closed monotone subsets of $X \times X^*$ may fail to be μ-closed. Indeed,

$$B := \big( ]-\infty, 0] \times \{-1\} \big) \cup \big([0, +\infty[\times \{+1\} \big) \tag{151}$$

is a closed monotone subset of $\mathbb{R}^2$, yet $B^{\mu\mu} \setminus B = \{0\} \times]-1, +1[\neq \varnothing$.

7. Fitzpatrick's problem: A sufficient condition

The goal of this section is to derive a new condition sufficient for the identity $A = G_{F_A}$. We shall build on the next two results.

Proposition 7.1. *Let A be a nonempty monotone subset of $X \times X^*$. Then $G_{F_A} \subset A^\mu$.*

Proof. By Fact 5.2, $A \subset G_{F_A}$. Corollary 6.7 implies that G_{F_A} is monotone. Altogether, using (148), we conclude that $G_{F_A} \subset A^\mu$. □

Proposition 7.2. *Let A be a nonempty monotone subset of $X \times X^*$. Then $G_{F_A} \subset \overline{\operatorname{conv}} A$.*

Proof. We argue by contradiction and thus assume the existence of some $(x, x^*) \in G_{F_A} \setminus \overline{\operatorname{conv}} A$. Then on the one hand

$$(x^*, x) \in \partial F_A(x, x^*), \tag{152}$$

and on the other hand the separation theorem yields $(y, y^*) \in X \times X^*$ and $\epsilon > 0$ such that

$$\langle (x, x^*) | (y^*, y) \rangle \geq 2\epsilon + \sup_{(a,a^*) \in A} \langle (a, a^*) | (y^*, y) \rangle . \tag{153}$$

Using (152), we see that

$$F_A(x + y, x^* + y^*) \geq F_A(x, x^*) + \langle (y, y^*) | (x^*, x) \rangle . \tag{154}$$

Thus, upon recalling the definition of $F_A(x + y, x^* + y^*)$, there must exist $(b, b^*) \in A$ such that

$$\langle x + y | b^* \rangle + \langle b | x^* + y^* \rangle - \langle b | b^* \rangle \geq -\epsilon + F_A(x, x^*) + \langle (x, x^*) | (y^*, y) \rangle . \tag{155}$$

This and (153) show that

$$\begin{aligned}
F_A(x, x^*) &\leq \langle x + y | b^* \rangle + \langle b | x^* + y^* \rangle - \langle b | b^* \rangle + \epsilon - \langle (x, x^*) | (y^*, y) \rangle && (156)\\
&\leq \langle x + y | b^* \rangle + \langle b | x^* + y^* \rangle - \langle b | b^* \rangle - \epsilon - \sup_{(a,a^*) \in A} \langle (a, a^*) | (y^*, y) \rangle && (157)\\
&\leq \langle x + y | b^* \rangle + \langle b | x^* + y^* \rangle - \langle b | b^* \rangle - \epsilon - \langle (b, b^*) | (y^*, y) \rangle && (158)\\
&= \langle x | b^* \rangle + \langle b | x^* \rangle - \langle b | b^* \rangle - \epsilon && (159)\\
&\leq F_A(x, x^*) - \epsilon, && (160)
\end{aligned}$$

which is absurd since $(x, x^*) \in \operatorname{dom} \partial F_A \subset \operatorname{dom} F_A$. □

Corollary 7.3. *Let A be a nonempty monotone subset of $X \times X^*$. Then*

$$G_{F_A} \subset A^\mu \cap \overline{\operatorname{conv}} A. \tag{161}$$

Proof. Combine Proposition 7.1 with Proposition 7.2. □

We are now ready for the main result of this section.

Theorem 7.4. *Let A be a nonempty monotone subset of $X \times X^*$ such that $A = A^\mu \cap \overline{\operatorname{conv}} A$. Then $G_{F_A} = A$.*

Proof. Fact 5.2 and Corollary 7.3 imply

$$A \subset G_{F_A} \subset A^\mu \cap \overline{\operatorname{conv}} A. \tag{162}$$

Since $A = A^\mu \cap \overline{\operatorname{conv}} A$, it now follows from (162) that $A = G_{F_A}$. □

Remark 7.5. Let us provide some examples of monotone operators A for which $A = G_{F_A}$, by means of satisfying the assumption in Theorem 7.4.

(i) A is maximal monotone (by (149)).
(ii) A is a nonempty subset of $X \times X^*$ that is closed, convex, and monotone.
(iii) A is a singleton. (This is a special case of (ii).)

Conditions (i) and (iii) were already pointed out by Fitzpatrick; see [6, Section 3]. Theorem 7.4 remains true when X is a reflexive Banach space, as a second glance at its proof reveals.

8. Fitzpatrick's problem and connectedness when $X = \mathbb{R}$

Throughout this section, we assume that $X = \mathbb{R}$ and that $X \times X^* = \mathbb{R} \times \mathbb{R}$ is partially ordered by the nonnegative orthant $\mathbb{R}^2_+ := [0, +\infty[\times [0, +\infty[$. Given (x, x^*) and (y, y^*) in $\mathbb{R} \times \mathbb{R}$, we thus write $(x, x^*) \preceq (y, y^*)$ if $(y - x, y^* - x^*) \in \mathbb{R}^2_+$, and $(x, x^*) \precneqq (y, y^*)$ if $(x, x^*) \preceq (y, y^*)$ and $(x, x^*) \neq (y, y^*)$.

Proposition 8.1. *Let A be a nonempty monotone subset of $\mathbb{R} \times \mathbb{R}$. Then there exists a subset S of $\mathbb{R}$ and an order-preserving parameterization $M \colon S \to A$ that is bijective and continuous in both directions.*

Proof. In view of Fact 6.1, we let S be the range of $\operatorname{Id} + A$ and M be the Minty parameterization of any maximal monotone extension of A. Resolvents on the real line are increasing, and A is (totally) ordered as a subset of $\mathbb{R} \times \mathbb{R}$; see [14, Exercise 12.9]. □

Definition 8.2. Let (x, x^*) and (y, y^*) belong to $\mathbb{R} \times \mathbb{R}$ such that $(x, x^*) \precneqq (y, y^*)$. Then the *box* with diagonal delimited by these points is defined by

$$\operatorname{box}\big((x, x^*), (y, y^*)\big) = \operatorname{conv}\big\{(x, x^*), (x, y^*), (y, x^*), (y, y^*)\big\}. \tag{163}$$

We allow the box to degenerate to a line segment which happens precisely when the line segment $\big[(x, x^*), (y, y^*)\big]$ is parallel to any of the coordinate axis.

Proposition 8.3. *Let A be a nonempty closed monotone subset of $\mathbb{R} \times \mathbb{R}$ that is not connected. Then there are two points (a, a^*) and (b, b^*) in A such that $(a, a^*) \precneqq (b, b^*)$,*

$$A \cap \operatorname{box}\big((a, a^*), (b, b^*)\big) = \big\{(a, a^*), (b, b^*)\big\}, \tag{164}$$

and hence

$$\operatorname{box}\big((a, a^*), (b, b^*)\big) \subset A^\mu. \tag{165}$$

Proof. Since A is not connected, it contains two points (c, c^*) and (d, d^*) belonging to distinct connected components of A. By monotonicity of A, we assume (without loss of generality) that $(c, c^*) \precneqq (d, d^*)$. Now let S and M be as in Proposition 8.1 so that there exist γ and δ in S such that $\gamma < \delta$ and $(c, c^*) = M(\gamma)$ and $(d, d^*) = M(\delta)$. Moreover, S is closed but not connected, and γ and δ belong to distinct connected components of S. Since $[\gamma, \delta] \not\subset S$, there exists $\rho \in]\gamma, \delta[\cap (\mathbb{R} \setminus S)$. Now let $]\alpha, \beta[$ be the largest open interval in $\mathbb{R} \setminus S$ that contains ρ. Then $\gamma \leq \alpha < \rho < \beta \leq \delta$ and $S \cap [\alpha, \beta] = \{\alpha, \beta\}$. Therefore, by Proposition 8.1, the points defined by $(a, a^*) := M(\alpha)$ and $(b, b^*) := M(\beta)$ satisfy (164), and this in turn implies (165). ☐

Theorem 8.4. *Let A be a nonempty closed monotone subset of $\mathbb{R} \times \mathbb{R}$. Then*

$$A = A^\mu \cap \overline{\operatorname{conv}} A \Leftrightarrow A \text{ is connected.} \tag{166}$$

Proof. We shall argue geometrically, exploiting the correspondence described in Proposition 8.1.

"$\Rightarrow$": Suppose to the contrary that A is not connected. With the help of Proposition 8.3, we see that there exist two points (a, a^*) and (b, b^*) in A such that

$$(a, a^*) \precneqq (b, b^*), \tag{167}$$

$$A \cap \operatorname{box}\big((a, a^*), (b, b^*)\big) = \big\{(a, a^*), (b, b^*)\big\}, \tag{168}$$

and

$$\operatorname{conv}\big\{(a, a^*), (b, b^*)\big\} \subset \operatorname{box}\big((a, a^*), (b, b^*)\big) \subset A^\mu. \tag{169}$$

But this implies

$$\operatorname{conv}\big\{(a, a^*), (b, b^*)\big\} \subset \big(A^\mu \cap \overline{\operatorname{conv}} A\big) \cap \operatorname{box}\big((a, a^*), (b, b^*)\big) = A \cap \operatorname{box}\big((a, a^*), (b, b^*)\big), \tag{170}$$

which contradicts (167)–(168).

"$\Leftarrow$": The inclusion

$$A \subset A^\mu \cap \overline{\operatorname{conv}} A \tag{171}$$

is always true, since A is monotone. Now let M and S be as in Proposition 8.1. By assumption, S is connected and closed, i.e., a closed (possibly unbounded) interval. Using Proposition 8.1, we define in $[-\infty, +\infty] \times [-\infty, +\infty]$ the two pairs

$$(a, a^*) := \lim_{S \ni s \to \inf S} M(s) \quad \text{and} \quad (b, b^*) := \lim_{S \ni s \to \sup S} M(s). \tag{172}$$

We now consider cases.

Case 1: $(a, a^*) \in \mathbb{R}^2$ and $(b, b^*) \in \mathbb{R}^2$. Then both (a, a^*) and (b, b^*) belong to A, and

$$A^\mu = A \cup \big((a, a^*) - \mathbb{R}^2_+\big) \cup \big((b, b^*) + \mathbb{R}^2_+\big). \tag{173}$$

Moreover, $\overline{\operatorname{conv}} A \subset \big((a, a^*) + \mathbb{R}^2_+\big) \cap \big((b, b^*) - \mathbb{R}^2_+\big)$. Altogether, $A^\mu \cap \overline{\operatorname{conv}} A \subset A$.

Case 2: $(a, a^*) \notin \mathbb{R}^2$ and $(b, b^*) \in \mathbb{R}^2$. Then $(b, b^*) \in A$, $A^\mu = A \cup \big((b, b^*) + \mathbb{R}^2_+\big)$, and $\overline{\operatorname{conv}} A \subset (b, b^*) - \mathbb{R}^2_+$. Thus, $A^\mu \cap \overline{\operatorname{conv}} A \subset A$.

Case 3: $(a, a^*) \in \mathbb{R}^2$ and $(b, b^*) \notin \mathbb{R}^2$. This is analogous to Case 2.

Case 4: $(a, a^*) \notin \mathbb{R}^2$ and $(b, b^*) \notin \mathbb{R}^2$. Then $A^\mu = A$ and thus $A^\mu \cap \overline{\operatorname{conv}} A = A$. □

Remark 8.5. Theorem 8.4 is false in higher-dimensional spaces. Indeed, Example 8.9 below provides a nonempty closed monotone subset A of $X \times X^*$ that is connected but for which $A \neq G_{F_A}$ and hence (by Theorem 7.4) $A \neq A^\mu \cap \overline{\operatorname{conv}} A$.

Theorem 8.6. *Let A be a nonempty monotone subset of $\mathbb{R} \times \mathbb{R}$ such that $A = G_{F_A}$. Then A is closed and connected.*

Proof. By Proposition 6.10, A is closed. Assume to the contrary that A is not connected. Now Proposition 8.3 yields two points (a, a^*) and (b, b^*) in A such that

$$(a, a^*) \lneqq (b, b^*) \tag{174}$$

and

$$A \cap \operatorname{box}\big((a, a^*), (b, b^*)\big) = \big\{(a, a^*), (b, b^*)\big\}. \tag{175}$$

Since A is monotone, we have

$$A \subset \big((a, a^*) - \mathbb{R}^2_+\big) \cup \big((b, b^*) + \mathbb{R}^2_+\big). \tag{176}$$

Now let (m, m^*) be the midpoint of the box, i.e.,

$$(m, m^*) := \tfrac{1}{2}(a, a^*) + \tfrac{1}{2}(b, b^*), \tag{177}$$

and let $(r, r^*) \in \mathbb{R}^2_+$. Then $0 \le 2rr^* + (b - a)r^* + r(b^* - a^*)$, which is equivalent to

$$m(b^* + r^*) + (b + r)m^* - (b + r)(b^* + r^*) \le \tfrac{1}{2}ab^* + \tfrac{1}{2}ba^* \tag{178}$$

and also to

$$m(a^* - r^*) + (a - r)m^* - (a - r)(a^* - r^*) \le \tfrac{1}{2}ab^* + \tfrac{1}{2}ba^*. \tag{179}$$

Since $(r, r^*) \in \mathbb{R}^2_+$ was chosen arbitrarily, we see that Definition 1.1, (176), (178), and (179) imply

$$F_A(m, m^*) = \sup_{(c,c^*) \in A} \big(mc^* + cm^* - cc^*\big) \le \tfrac{1}{2}ab^* + \tfrac{1}{2}ba^*. \tag{180}$$

Moreover, (177) results in

$$ma^* + am^* - aa^* = \tfrac{1}{2}ab^* + \tfrac{1}{2}ba^* = mb^* + bm^* - bb^*. \tag{181}$$

Utilizing Proposition 6.2 and its notation, we obtain from (180) and (181) that

$$\big\{(a, a^*), (b, b^*)\big\} \subset \mathcal{A}(m, m^*); \tag{182}$$

consequently,

$$(m^*, m) \in \partial F_A(m, m^*). \tag{183}$$

By Definition 5.1,

$$(m, m^*) \in G_{F_A}. \tag{184}$$

On the other hand, (174), (175), and (177) imply

$$(m, m^*) \notin A. \tag{185}$$

Altogether, (184) and (185) result in $(m, m^*) \in G_{F_A} \backslash A$, which contradicts our assumption that $A = G_{F_A}$. ☐

We are now in a position to completely settle Fitzpatrick's problem when $X = \mathbb{R}$.

Corollary 8.7. *Let A be a nonempty monotone subset of $\mathbb{R} \times \mathbb{R}$. Then the following are equivalent:*

(i) $A = G_{F_A}$.
(ii) *A is closed and connected.*
(iii) $A = A^\mu \cap \overline{\operatorname{conv}} A$.

Proof. "$(i) \Rightarrow (ii)$": Theorem 8.6. "$(ii) \Rightarrow (iii)$": Theorem 8.4. "$(iii) \Rightarrow (i)$": Theorem 7.4. ☐

Remark 8.8. The identity $A = A^\mu \cap \overline{\operatorname{conv}} A$ is a sufficient condition for the fixed point equation $A = G_{F_A}$. However, it is not necessary. Indeed, let $X := \mathbb{R}^2$ and identify $X \times X^*$ with $\mathbb{R}^4$. Set

$$A := [0, +\infty[\cdot (1, 0, 1, 0) \cup [0, +\infty[\cdot (0, 1, 0, 1). \tag{186}$$

Then $G_{F_A} = A$ yet $]0, +\infty[\cdot (1, 1, 1, 1) \subseteq (A^\mu \cap \overline{\operatorname{conv}} A) \smallsetminus A$. This example was obtained by the second author after acceptance of the manuscript. For details, see [10, Example 4.3.9].

We conclude with an example that illustrates the failure of Corollary 8.7 in $\mathbb{R}^2 \times \mathbb{R}^2$.

Example 8.9. Let $X := \mathbb{R}^2$ and identify $X \times X^*$ with $\mathbb{R}^4$ via $(x, x^*) = \big((x_1, x_2), (x_1^*, x_2^*)\big) \mapsto (x_1, x_2, x_1^*, x_2^*)$. Denote the standard unit vectors in $\mathbb{R}^4$ by e_1, e_2, e_3, e_4, and then set $a_1 := e_1 + e_3$, $a_2 := e_1 + e_3 + e_4$, $a_3 := e_1 + e_2 + e_4$, $a_4 := -e_1 + e_2 + e_4$, $a_5 := -e_1 - e_3 + e_4$, and $a_6 := -e_1 - e_3$. Now define a piecewise linear path from a_1 to a_6 by

$$A := \bigcup_{i \in \{1,2,3,4,5\}} [a_i, a_{i+1}]. \tag{187}$$

Clearly, A is closed and connected. By discussing cases, it is readily checked that A is monotone. Therefore,

$$A \text{ is a nonempty closed connected monotone subset of } X \times X^*. \tag{188}$$

Direct computation from the definition shows that $F_A(0,0) = -1$. Proposition 6.2 and its notation now imply that $\{a_1, a_6\} \subset \mathcal{A}(0,0)$. Since $(0,0) \in [a_1, a_6]$, we conclude $(0,0) \in \partial F_A(0,0)$ and thus

$$(0, 0) \in G_{F_A}. \tag{189}$$

On the other hand, $(0,0) \notin A$ by definition of A. Altogether,

$$A \neq G_{F_A}. \tag{190}$$

Combining (188) and (190) now shows that Corollary 8.7 cannot hold in this setting.

Acknowledgements. We wish to thank Stephen Simons and Benar Svaiter for sending us [17] and [8], respectively. We also thank two anonymous referees for their careful reading and insightful comments. H. H. Bauschke's work and H. S. Sendov's work was partially supported by the Natural Sciences and Engineering Research Council of Canada.

References

[1] H. H. Bauschke: Projection Algorithms and Monotone Operators, Ph.D. Thesis, Simon Fraser University, Burnaby, Canada, August 1996; available at http://www.cecm.sfu.ca/preprints/1996pp.html.

[2] H. H. Bauschke, J. M. Borwein: Maximal monotonicity of dense type, local maximal monotonicity, and monotonicity of the conjugate are all the same for continuous linear operators, Pac. J. Math. 189 (1999) 1–20.

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

[4] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, D. E. Knuth: On the Lambert W function, Adv. Comput. Math. 5 (1996) 329–359.

[5] A. L. Dontchev, T. Zolezzi: Well-Posed Optimization Problems, Lecture Notes in Mathematics 1543, Springer, Berlin (1993).

[6] S. Fitzpatrick: Representing monotone operators by convex functions, in: Functional Analysis and Optimization, Workshop / Miniconference (Canberra 1988), Proc. Cent. Math. Anal. Aust. Natl. Univ. 20, Australian National University, Canberra (1988) 59–65.

[7] C. W. Groetsch: Generalized Inverses of Linear Operators, Monographs and Textbooks in Pure and Applied Mathematics 37, Marcel Dekker, New York (1977).

[8] J.-E. Martínez-Legaz, B. F. Svaiter: Monotone operators representable by l.s.c. convex functions, Set-Valued Anal. 13 (2005) 21–46.

[9] J.-E. Martínez-Legaz, M. Théra: A convex representation of maximal monotone operators, J. Nonlinear Convex Anal. 2 (2001) 243–247.

[10] D. A. McLaren: Notes on the Fitzpatrick Function, M. Sc. Thesis, University of Guelph, Ontario, Canada, August 2005.

[11] J.-J. Moreau: Proximité et dualité dans un espace hilbertien, Bull. Soc. Math. Fr. 93 (1965) 273–299.

[12] R. T. Rockafellar: On the maximal monotonicity of subdifferential mappings, Pac. J. Math. 33 (1970) 209–216.

[13] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton (1970).

[14] R. T. Rockafellar, R. J.-B. Wets: Variational Analysis, Springer, Berlin (1998).

[15] S. Simons: Subdifferentials are locally maximal monotone, Bull. Aust. Math. Soc. 47 (1993) 465–471.

[16] S. Simons: Minimax and Monotonicity, Lecture Notes in Mathematics 1693, Springer, Berlin (1998).

[17] S. Simons: The Fitzpatrick function and the range of a sum, preprint (2005).

[18] S. Simons, C. Zălinescu: A new proof for Rockafellar's characterization of maximal monotone operators, Proc. Amer. Math. Soc. 132 (2004) 2969–2972.

[19] S. Simons, C. Zălinescu: Fenchel duality, Fitzpatrick functions and maximal monotonicity, J. Nonlinear Convex Anal. 6 (2005) 1–22.

[20] C. Zălinescu: Convex Analysis in General Vector Spaces, World Scientific, Singapore (2002).