Fitzpatrick functions, cyclic monotonicity and Rockafellar's antiderivative

Sedi Bartz[*], Heinz H. Bauschke[†], Jonathan M. Borwein[‡], Simeon Reich[§], and Xianfu Wang[¶]

December 31, 2005 — Version 1.23

Abstract

Several deeper results on maximal monotone operators have recently found simpler proofs using Fitzpatrick functions. In this paper, we study a sequence of Fitzpatrick functions associated with a monotone operator. The first term of this sequence coincides with the original Fitzpatrick function, and the other terms turn out to be useful for the identification and characterization of cyclic monotonicity properties. It is shown that for any maximal cyclically monotone operator, the pointwise supremum of the sequence of Fitzpatrick functions is closely related to Rockafellar's antiderivative. Several examples are explicitly computed for the purpose of illustration. In contrast to Rockafellar's result, a maximal 3-cyclically monotone operator need not be maximal monotone. A simplified proof of Asplund's observation that the rotation in the Euclidean plane by π/n is n-cyclically monotone but not $(n+1)$-cyclically monotone is provided. The Fitzpatrick family of the subdifferential operator of a sublinear and of an indicator function is studied in detail. We conclude with a new proof of Moreau's result concerning the convexity of the set of proximal mappings.

2000 Mathematics Subject Classification: Primary 47H05; Secondary 52A41, 90C25.

Keywords: Convex function, Fenchel conjugate, Fitzpatrick function, cyclically monotone operator, maximal monotone operator, subdifferential operator, sublinear function.

[*]Department of Mathematics, The Technion—Israel Institute of Technology, 32000 Haifa, Israel. E-mail: bartz@techunix.technion.ac.il.

[†]Mathematics, Irving K. Barber School, UBC Okanagan, Kelowna, British Columbia V1V 1V7, Canada. E-mail: heinz.bauschke@ubc.ca.

[‡]Faculty of Computer Science, Dalhousie University, 6050 University Avenue, Halifax, Nova Scotia B3H 1W5, Canada. E-mail: jborwein@cs.dal.ca.

[§]Department of Mathematics, The Technion—Israel Institute of Technology, 32000 Haifa, Israel. E-mail: sreich@techunix.technion.ac.il.

[¶]Mathematics, Irving K. Barber School, UBC Okanagan, Kelowna, British Columbia V1V 1V7, Canada. E-mail: shawn.wang@ubc.ca.

1 Introduction

Throughout this paper, we assume that

$$X \text{ is a real Banach space,} \tag{1}$$

with norm $\|\cdot\|$, with continuous dual space X^*, and with the pairing $p = \langle\cdot,\cdot\rangle$ between X and X^*. Our aim is to provide new results on monotone operators and their associated Fitzpatrick functions. We denote the *graph* of a set-valued operator $A\colon X \to 2^{X^*}$ by $\operatorname{gra} A = \big\{(x, x^*) \in X \times X^* \mid x^* \in Ax\big\}$ and recall the following notions.

Definition 1.1 *Let $A\colon X \to 2^{X^*}$. Then A is* n-cyclically monotone *if $n \in \{2, 3, \ldots\}$ and*

$$\left.\begin{array}{c}(a_1, a_1^*) \in \operatorname{gra} A\\ \vdots\\ (a_n, a_n^*) \in \operatorname{gra} A\\ a_{n+1} = a_1\end{array}\right\} \quad\Rightarrow\quad \sum_{i=1}^{n}\langle a_{i+1} - a_i, a_i^*\rangle \leq 0. \tag{2}$$

The operator A is monotone *if it is 2-cyclically monotone; equivalently, if*

$$\left.\begin{array}{c}(x, x^*) \in \operatorname{gra} A\\ (y, y^*) \in \operatorname{gra} A\end{array}\right\} \quad\Rightarrow\quad \langle x - y, x^* - y^*\rangle \geq 0. \tag{3}$$

The operator A is cyclically monotone *if for every $n \in \{2, 3, \ldots\}$, A is n-cyclically monotone.*

Definition 1.2 *Let $A\colon X \to 2^{X^*}$ and let $n \in \{2, 3, \ldots\}$. Then A is* maximal n-cyclically monotone *if A is monotone and no proper extension of A is n-cyclically monotone; A is* maximal cyclically monotone *if A is cyclically monotone and no proper extension of A is cyclically monotone; A is* maximal monotone *if A is maximal 2-cyclically monotone.*

Zorn's Lemma guarantees that every n-cyclically monotone operator admits an extension, in the sense of enlargement of the graph, to a maximal n-cyclically monotone extension. An analogous result holds for cyclically monotone operators. Monotone operators play a fundamental role in optimization; the reader is referred to [14, 42, 43, 48] for further information. Linear maximal n-cyclically monotone operators are discussed in [1]. One of the most remarkable results in monotone operator theory is due to Rockafellar [38] who proved that the maximally cyclically monotone operators are precisely the subdifferential operators of functions that are convex, lower semicontinuous, and proper. Hence every maximally cyclically monotone operator admits an antiderivative, which is unique up to an additive constant, see Fact 3.2 below. Other fundamental results concern characterizations of maximal monotonicity and the maximality of the sum of two maximal monotone operators under a constraint qualification. Through the use of the Fitzpatrick function [21], such results have recently found dramatically simpler proofs; see, e.g., [9] and [45]. Fitzpatrick functions play also a key role in various works on monotone operators [12, 16, 17, 18, 28, 29, 34, 44]. Let us now state the definition of a Fitzpatrick function.

Definition 1.3 *Let* $A\colon X \to 2^{X^*}$. *The* Fitzpatrick function *of* A *is*

$$F_A\colon X \times X^* \to \left]-\infty, +\infty\right] : (x, x^*) \mapsto \sup_{(a,a^*)\in \operatorname{gra} A} \big(\langle x, a^*\rangle + \langle a, x^*\rangle - \langle a, a^*\rangle\big). \tag{4}$$

Fitzpatrick [21] established the following fundamental properties.

Fact 1.4 *Let* $A\colon X \to 2^{X^*}$ *be maximal monotone. Then* F_A *is convex, lower semicontinuous, proper,* $p \le F_A$*, and* $p = F_A$ *on* $\operatorname{gra} A$*. Moreover,* F_A *is the smallest function with these properties.*

In this paper, we provide several novel results concerning Fitzpatrick functions and monotone operators. These results can be roughly classified into five categories, which also correspond to the contents of the remaining five sections.

① In Section 2, we introduce an increasing sequence of Fitzpatrick functions of order n, where $n \in \{2, 3, \ldots\}$. The first term of this sequence coincides with the original Fitzpatrick functions, which is a powerful tool in the study of maximal monotone operators. Analogously, we show that the Fitzpatrick functions of order n captures precisely (maximal) n-cyclic monotonicity properties. We also provide an example of a maximal 3-cyclically monotone operator that is *not* maximal monotone (Example 2.16). This is in striking contrast to Rockafellar's characterization of maximal cyclically monotone operators.

② The important case of subdifferential operators is considered in greater detail in Section 3. We show how the Fitzpatrick functions and Rockafellar's antiderivative arise from a common ancestor and prove, using deeper results on subgradients, that the sequence of Fitzpatrick functions converges to a function that has a close relationship to Rockafellar's antiderivative (Theorem 3.5).

③ Section 4 contains several examples of maximal monotone operators where the Fitzpatrick functions of all orders are computed in closed form. It is possible that outside the graph all Fitzpatrick functions are either identical (Example 4.2) or pairwise distinct (Example 4.4). We also provide a simple, more self-contained proof of Asplund's result [1] that rotations in the Euclidean plane by π/n are n-cyclically monotone but not $(n+1)$-cyclically monotone (Example 4.6).

④ The Fitzpatrick family of a given maximal monotone operator $A\colon X \to 2^{X^*}$ consists of all convex, lower semicontinuous, and proper functions that are greater than or equal to p everywhere and equal to p on the graph of A. By Fact 1.4, this family contains F_A. In Section 5, we extend results of Penot [34] and of Burachik and Fitzpatrick [16] by showing that the Fitzpatrick family is a singleton when A is the subdifferential operator of either a sublinear function (Theorem 5.3) or of an indicator function (Corollary 5.9).

⑤ The last Section 6 deals with convexity properties of the set of proximal mappings (resolvents of subdifferential operators). Moreau showed that in Hilbert space the set of proximal mappings is convex [33]. We provide a different proof of this result (Theorem 6.7) and also an example illustrating the nonconvexity of the set of firmly nonexpansive mappings outside Hilbert space (Example 6.4 and Remark 6.5).

Notation utilized is standard in Convex Analysis and Monotone Operator Theory; see, e.g., [12], [43], and [48].

2 Fitzpatrick functions and cyclic monotonicity

Let us introduce the following functions which may be interpreted as common ancestors of the Fitzpatrick function and Rockafellar's antiderivative.

Definition 2.1 *Let $A\colon X \to 2^{X^*}$, let $(a_1, a_1^*) \in \operatorname{gra} A$, and let $n \in \{2, 3, \ldots\}$. If $n = 2$, we set $C_{A,2,(a_1,a_1^*)}\colon X \times X^* \to\ ]-\infty, +\infty] : (x, x^*) \mapsto \langle x, a_1^*\rangle + \langle a_1, x^*\rangle - \langle a_1, a_1^*\rangle$. Now suppose that $n \in \{3, 4, \ldots\}$. Then the value of the function $C_{A,n,(a_1,a_1^*)}\colon X \times X^* \to\ ]-\infty, +\infty]$ at $(x, x^*) \in X \times X^*$ is defined by*

$$\sup_{\substack{(a_2,a_2^*)\in\operatorname{gra} A\\ \vdots \\ (a_{n-1},a_{n-1}^*)\in\operatorname{gra} A}} \left(\sum_{i=1}^{n-2}\langle a_{i+1} - a_i, a_i^*\rangle\right) + \langle x - a_{n-1}, a_{n-1}^*\rangle + \langle a_1, x^*\rangle; \tag{5}$$

equivalently, by

$$\sup_{\substack{(a_2,a_2^*)\in\operatorname{gra} A,\\ \vdots \\ (a_{n-1},a_{n-1}^*)\in\operatorname{gra} A}} \langle x, x^*\rangle + \left(\sum_{i=1}^{n-2}\langle a_{i+1} - a_i, a_i^*\rangle\right) + \langle x - a_{n-1}, a_{n-1}^*\rangle + \langle a_1 - x, x^*\rangle. \tag{6}$$

Definition 2.2 (Fitzpatrick functions) *Let $A\colon X \to 2^{X^*}$. For every $n \in \{2, 3, \ldots\}$, the* Fitzpatrick function of A of order n *is*

$$F_{A,n} = \sup_{(a,a^*)\in\operatorname{gra} A} C_{A,n,(a,a^*)}. \tag{7}$$

The Fitzpatrick function of A of infinite order *is $F_{A,\infty} = \sup_{n\in\{2,3,\ldots\}} F_{A,n}$.*

The first result is immediate from the definition.

Proposition 2.3 *Let $A\colon X \to 2^{X^*}$ and let $n \in \{2, 3, \ldots\}$. Then $F_{A,n}\colon X \times X^* \to [-\infty, +\infty]$ is convex and lower semicontinuous. At $(x, x^*) \in X \times X^*$, the value of $F_{A,n}$ is given by*

$$\sup_{\substack{(a_1,a_1^*)\in\operatorname{gra} A,\\ \vdots \\ (a_{n-1},a_{n-1}^*)\in\operatorname{gra} A}} \langle x, x^*\rangle + \left(\sum_{i=1}^{n-2}\langle a_{i+1} - a_i, a_i^*\rangle\right) + \langle x - a_{n-1}, a_{n-1}^*\rangle + \langle a_1 - x, x^*\rangle. \tag{8}$$

Moreover,

$$F_{A,n} \geq \langle \cdot, \cdot \rangle \ \text{on } \operatorname{gra} A. \tag{9}$$

If $n = 2$, then (8) simplifies to $\sup_{(a,a^*)\in\operatorname{gra} A} \langle x, a^*\rangle + \langle a, x^*\rangle - \langle a, a^*\rangle$, which is the original definition (see Definition 1.3) of the Fitzpatrick function [21] of A at $(x, x^*) \in X \times X^*$. Note that $(F_{A,n})_{n\in\{2,3,\ldots\}}$ is a sequence of increasing functions and that $F_{A,n} \to F_{A,\infty}$ pointwise. We now provide a characterization of n-cyclically monotone operators by Fitzpatrick functions of order n which directly generalizes [28, Proposition 2].

Proposition 2.4 *Let $A\colon X \to 2^{X^*}$ and let $n \in \{2, 3, \ldots\}$. Then the following are equivalent.*

(i) *A is n-cyclically monotone.*

(ii) *$F_{A,n} \leq \langle \cdot, \cdot\rangle$ on $\operatorname{gra} A$.*

(iii) *$F_{A,n} = \langle \cdot, \cdot\rangle$ on $\operatorname{gra} A$.*

Proof. Take $(x, x^*) \in \operatorname{gra} A$ and take $n - 1$ additional points $(a_1, a_1^*), \ldots, (a_{n-1}, a_{n-1}^*)$ in $\operatorname{gra} A$. "(i)⇒(ii)": In view of Definition 1.1, $\big(\sum_{i=1}^{n-2} \langle a_{i+1} - a_i, a_i^*\rangle\big) + \langle x - a_{n-1}, a_{n-1}^*\rangle + \langle a_1 - x, x^*\rangle \leq 0$ and thus $\langle x, x^*\rangle + \big(\sum_{i=1}^{n-2} \langle a_{i+1} - a_i, a_i^*\rangle\big) + \langle x - a_{n-1}, a_{n-1}^*\rangle + \langle a_1 - x, x^*\rangle \leq \langle x, x^*\rangle$. Recalling (8) and taking the supremum over $(a_1, a_1^*), \ldots, (a_{n-1}, a_{n-1}^*)$ in $\operatorname{gra} A$, we see that $F_{A,n}(x, x^*) \leq \langle x, x^*\rangle$. "(ii)⇒(iii)": This follows from (9). "(iii)⇒(i)": By (8), $\langle x, x^*\rangle + \big(\sum_{i=1}^{n-2} \langle a_{i+1} - a_i, a_i^*\rangle\big) + \langle x - a_{n-1}, a_{n-1}^*\rangle + \langle a_1 - x, x^*\rangle \leq \langle x, x^*\rangle$. Hence $\big(\sum_{i=1}^{n-2} \langle a_{i+1} - a_i, a_i^*\rangle\big) + \langle x - a_{n-1}, a_{n-1}^*\rangle + \langle a_1 - x, x^*\rangle \leq 0$, which implies that A is n-cyclically monotone. ■

Corollary 2.5 *Let $A\colon X \to 2^{X^*}$. Then A is cyclically monotone $\Leftrightarrow$ $F_{A,\infty} \leq \langle \cdot, \cdot\rangle$ on $\operatorname{gra} A$ $\Leftrightarrow$ $F_{A,\infty} = \langle \cdot, \cdot\rangle$ on $\operatorname{gra} A$.*

Proof. This is clear from Proposition 2.4 and the fact that $F_{A,\infty} = \sup_{n\in\{2,3,\ldots\}} F_{A,n}$. ■

Remark 2.6 For any function $f\colon X \times X^* \to\,]-\infty, +\infty]$, set, as in [21, Definition 2.1], $G_f\colon X \to 2^{X^*}\colon x \mapsto \big\{x^* \in X^* \mid (x^*, x) \in \partial f(x, x^*)\big\}$. Fitzpatrick proved that for any $A\colon X \to 2^{X^*}$ that is monotone, the operator $G_{F_{A,2}}$ is a monotone extension of A (and hence $A = G_{F_{A,2}}$ whenever A is maximal monotone); see [21, Corollary 3.5]. The following generalization holds. *Let $A : X \to 2^{X^*}$ be n-cyclically monotone for some $n \in \{2, 3, \ldots\}$. Then $G_{F_{A,n}}$ is a monotone extension of A. Consequently, if A is maximal monotone, then $A = G_{F_{A,n}}$.* Indeed, it follows from [21, Lemma 3.3] that $G_{F_{A,2}}$ is a monotone extension of A. Take $(x, x^*) \in \operatorname{gra} A$. Then $(x^*, x) \in \partial F_{A,2}(x, x^*)$ and thus, for any $(y, y^*) \in X \times X^*$,

$$F_{A,2}(x + y, x^* + y^*) - F_{A,2}(x, x^*) \geq \langle y, x^*\rangle + \langle x^*, x\rangle. \tag{10}$$

On the other hand, we have $F_{A,n} \geq F_{A,2}$ and also, by Proposition 2.4, $F_{A,n}(x, x^*) = \langle x, x^*\rangle = F_{A,2}(x, x^*)$. Altogether, $F_{A,n}(x + y, x^* + y^*) - F_{A,n}(x, x^*) \geq \langle y, x^*\rangle + \langle x^*, x\rangle$, and hence $(x^*, x) \in \partial F_{A,n}(x, x^*)$, i.e., $(x, x^*) \in \operatorname{gra} G_{F_{A,n}}$.

The next result deals with the extensibility of n-cyclic monotone operators.

Proposition 2.7 *Let $A\colon X \to 2^{X^*}$ be n-cyclically monotone for some $n \in \{2,3,\ldots\}$, let $(x,x^*) \in X \times X^*$, and define $B\colon X \to 2^{X^*}$ via $\operatorname{gra} B = \operatorname{gra} A \cup \{(x,x^*)\}$. Then B is n-cyclically monotone $\Leftrightarrow F_{A,n}(x,x^*) \leq \langle x,x^*\rangle$.*

Proof. "$\Rightarrow$": Proposition 2.4 (applied to B) shows that $F_{B,n}(x,x^*) \leq \langle x,x^*\rangle$. On the other hand, since $\operatorname{gra} A \subset \operatorname{gra} B$, (8) yields $F_{A,n}(x,x^*) \leq F_{B,n}(x,x^*)$. Altogether, $F_{A,n}(x,x^*) \leq \langle x,x^*\rangle$. "$\Leftarrow$": The result is clear if $(x,x^*) \in \operatorname{gra} A$ so we assume that $(x,x^*) \notin \operatorname{gra} A$. Take $(b_1,b_1^*),(b_2,b_2^*),\ldots,(b_n,b_n^*)$ in $\operatorname{gra} B$ and set $b_{n+1} = b_1$. If these n pairs all belong to $\operatorname{gra} A$, then $\sum_{i=1}^n \langle b_{i+1}-b_i, b_i^*\rangle \leq 0$ since A is n-cyclically monotone. Otherwise, the set of indices $I = \big\{i \in \{1,2,\ldots,n\} \mid (b_i,b_i^*) = (x,x^*)\big\}$ contains K elements, where $K \in \{1,2,\ldots,n\}$. Write $I = \{i_1,\ldots,i_k\}$, where $1 \leq i_1 < i_2 < \ldots < i_K \leq n$. After relabeling if necessary, we assume further that $i_1 = 1$ and we set $i_{K+1} = n+1$. Then

$$K\langle x,x^*\rangle + \sum_{i=1}^n \langle b_{i+1}-b_i, b_i^*\rangle = \sum_{k=1}^K \Big(\langle x,x^*\rangle + \sum_{i=i_k}^{i_{k+1}-1} \langle b_{i+1}-b_i, b_i^*\rangle\Big)$$
$$\leq K F_{A,n}(x,x^*)$$
$$\leq K\langle x,x^*\rangle. \tag{11}$$

We deduce that $\sum_{i=1}^n \langle b_{i+1}-b_i, b_i^*\rangle \leq 0$ and this implies the n-cyclic monotonicity of B. ■

The following result extends [21, Corollary 3.9].

Corollary 2.8 *Let $A\colon X \to 2^{X^*}$ be maximal n-cyclically monotone, for some $n \in \{2,3,\ldots\}$. Then $F_{A,n} > \langle\cdot,\cdot\rangle$ outside $\operatorname{gra} A$, and $F_{A,n} = \langle\cdot,\cdot\rangle$ on $\operatorname{gra} A$.*

Proof. Suppose to the contrary that there exists a point $(x,x^*) \in (X \times X^*) \smallsetminus (\operatorname{gra} A)$ such that $F_{A,n}(x,x^*) \leq \langle x,x^*\rangle$. Proposition 2.7 implies that the operator $B\colon X \to 2^{X^*}$, defined via $\operatorname{gra} B = \operatorname{gra} A \cup \{(x,x^*)\}$, is still n-cyclically monotone which contradicts the maximality assumption on A. Hence $F_{A,n} > \langle\cdot,\cdot\rangle$ outside $\operatorname{gra} A$. Finally, by Proposition 2.4, $F_{A,n} = \langle\cdot,\cdot\rangle$ on $\operatorname{gra} A$. ■

We now provide characterizations of maximal n-cyclic monotone operators within the classes of maximal monotone, n-cyclic monotone, and general set-valued operators.

Theorem 2.9 *Let $A\colon X \to 2^{X^*}$ be maximal monotone and let $n \in \{2,3,\ldots,\}$. Then the following are equivalent.*

(i) *A is n-cyclically monotone.*

(ii) *A is maximal n-cyclically monotone.*

(iii) $\operatorname{gra} A = \big\{(x,x^*) \in X \times X^* \mid F_{A,n}(x,x^*) = \langle x,x^*\rangle\big\}$.

Proof. "(i)⇒(ii)": This is clear since A is maximal monotone. "(ii)⇒(iii)": A direct consequence of Corollary 2.8. "(iii)⇐(i)": Take any $(x, x^*) \in \operatorname{gra} A$. Then $F_{A,n}(x, x^*) = \langle x^*, x\rangle$ by assumption. Hence Proposition 2.3 implies that

$$\sup_{\substack{(a_1,a_1^*)\in\operatorname{gra} A,\\ \vdots \\ (a_{n-1},a_{n-1}^*)\in\operatorname{gra} A}} \left(\sum_{i=1}^{n-2}\langle a_{i+1}-a_i, a_i^*\rangle\right) + \langle x-a_{n-1}, a_{n-1}^*\rangle + \langle a_1 - x, x^*\rangle = 0. \tag{12}$$

Therefore, for $n-1$ arbitrary pairs $(a_1, a_1^*), \cdots, (a_{n-1}, a_{n-1}^*)$ in $\operatorname{gra} A$, we have

$$\left(\sum_{i=1}^{n-2}\langle a_{i+1}-a_i, a_i^*\rangle\right) + \langle x-a_{n-1}, a_{n-1}^*\rangle + \langle a_1 - x, x^*\rangle \leq 0. \tag{13}$$

Since this holds for every $(x, x^*) \in \operatorname{gra} A$, we conclude that A is n-cyclically monotone. ■

Remark 2.10 Let $A\colon X \to 2^{X^*}$ be maximal monotone. For every $n \in \{2, 3, \ldots\}$, set $S_n = \big\{(x, x^*) \in X \times X^* \mid F_{A,n}(x, x^*) = \langle x, x^*\rangle\big\}$. It is clear from Corollary 2.8 and (8) that pointwise

$$\langle\cdot,\cdot\rangle \leq F_{A,2} \leq F_{A,3} \leq \cdots \leq F_{A,n} \to F_{A,\infty} \tag{14}$$

and hence that

$$\operatorname{gra} A = S_2 \supset S_3 \supset \cdots \supset S_n \supset S_{n+1} \supset \cdots. \tag{15}$$

We now have the following *dichotomy*. Either $S_n \equiv \operatorname{gra} A$, in which case A is a subdifferential operator (see Fact 3.2 below), or there is a minimal $n \in \{2, 3, \ldots\}$ such that $\operatorname{gra} A = S_n \supsetneqq S_{n+1}$. In the latter case, by Theorem 2.9, A is n-cyclically monotone but not $(n+1)$-cyclically monotone.

Theorem 2.11 *Let $A : X \to 2^{X^*}$ be n-cyclically monotone for some $n \in \{2, 3, \cdots\}$. Then A is maximal n-cyclically monotone* ⇔ $\big\{(x, x^*) \in X \times X^* \mid F_{A,n}(x, x^*) \leq \langle x, x^*\rangle\big\} \subset \operatorname{gra} A$.

Proof. "⇒": Take $(x, x^*) \in X \times X^*$ such that $F_{A,n}(x, x^*) \leq \langle x, x^*\rangle$. By Proposition 2.7, $\operatorname{gra} A \cup (x, x^*)$ is the graph of an n-cyclically monotone operator. Since A is maximal n-cyclically monotone, we deduce that $(x, x^*) \in \operatorname{gra} A$. "⇐": Assume to the contrary that A is not maximal n-cyclically monotone. Hence there exists $\widetilde{A}\colon X \to 2^{X^*}$ such that $\widetilde{A}$ is a n-cyclically monotone and $\operatorname{gra} A$ is a proper subset of $\operatorname{gra} \widetilde{A}$. Take $(x, x^*) \in \operatorname{gra} \widetilde{A} \smallsetminus \operatorname{gra} A$. Then $F_{\widetilde{A},n}(x, x^*) = \langle x, x^*\rangle$ by Proposition 2.4. On the other hand, $F_{A,n}(x, x^*) \leq F_{\widetilde{A},n}(x, x^*)$. Altogether, $F_{A,n}(x, x^*) \leq \langle x, x^*\rangle$. This contradicts the hypothesis since $(x, x^*) \notin \operatorname{gra} A$. ■

Corollary 2.12 *Let $A\colon X \to 2^{X^*}$ be cyclically monotone. Then A is maximal cyclically monotone* ⇔ $\big\{(x, x^*) \in X \times X^* \mid F_{A,\infty}(x, x^*) \leq \langle x, x^*\rangle\big\} \subset \operatorname{gra} A$.

Proof. "⇐": Assume A is not maximal cyclically monotone. Then there exists $(y, y^*) \in (X \times X^*) \smallsetminus \operatorname{gra} A$ such that $\operatorname{gra} A \cup \{(y, y^*)\}$ is the graph of some cyclically monotone operator, say B. By Corollary 2.5, $F_{A,\infty}(y, y^*) \leq F_{B,\infty}(y, y^*) \leq \langle y, y^*\rangle$. The hypothesis now implies that (y, y^*) belongs to $\operatorname{gra} A$, which is absurd. "$\Rightarrow$": Assume the inclusion is false, i.e., there exists $(y, y^*) \in X \times X^*$ such that $F_{A,\infty}(y, y^*) \leq \langle y, y^*\rangle$ yet $(y, y^*) \notin \operatorname{gra} A$. Thus, $(\forall n \in \{2, 3, \ldots\})$ $F_{A,n}(y, y^*) \leq F_{A,\infty}(y, y^*) \leq \langle y, y^*\rangle$. Let $B\colon X \to 2^{X^*}$ be given via $\operatorname{gra} B = \operatorname{gra} A \cup \{(y, y^*)\}$. Proposition 2.7 implies that $(\forall n \in \{2, 3, \ldots\})$ B is n-cyclically monotone. Hence B is cyclically monotone, which contradicts the maximality of A. ■

Theorem 2.13 *Let $A\colon X \to 2^{X^*}$ and let $n \in \{2, 3, \ldots, \}$. Then A is maximal n-cyclically monotone* $\Leftrightarrow$ $\operatorname{gra} A = \big\{(x, x^*) \in X \times X^* \mid F_{A,n}(x, x^*) \leq \langle x, x^*\rangle\big\}$.

Proof. Combine Theorem 2.11 and Proposition 2.4. ■

The next result, which is a refinement of [21, Proposition 4.2], provides an upper bound of $F_{A,2}$ in terms of the Fenchel conjugate of $F_{A,n}$.

Theorem 2.14 *Let X be reflexive and let $A\colon X \to 2^{X^*}$ be n-cyclically monotone for some $n \in \{2, 3, \ldots\}$. Then $\big(\forall (x, x^*) \in X \times X^*\big)$ $F_{A,2}(x, x^*) \leq F^*_{A,n}(x^*, x)$.*

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

$$\begin{aligned}
F^*_{A,n}(x^*, x) &= \sup_{(y,y^*)\in X\times X^*} \big(\langle (y, y^*), (x^*, x)\rangle - F_{A,n}(y, y^*)\big) \\
&\geq \sup_{(y,y^*)\in \operatorname{gra} A} \big(\langle y, x^*\rangle + \langle x, y^*\rangle - F_{A,n}(y, y^*)\big) \\
&= \sup_{(y,y^*)\in \operatorname{gra} A} \big(\langle y, x^*\rangle + \langle x, y^*\rangle - \langle y, y^*\rangle\big) \qquad (16) \\
&= F_{A,2}(x, x^*), \qquad (17)
\end{aligned}$$

where we used Proposition 2.4 in (16). ■

Corollary 2.15 *Let X be reflexive and let $A\colon X \to 2^{X^*}$ be maximal n-cyclically monotone for some $n \in \{2, 3, \ldots, \}$. Then A is maximal monotone* $\Leftrightarrow$ $(\forall (x, x^*) \in X \times X^*)$ $F^*_{A,n}(x^*, x) \geq \langle x, x^*\rangle$.

Proof. "$\Rightarrow$": Suppose that A is maximal (2-cyclically) monotone. Fix any $(x, x^*) \in X \times X^*$. On the one hand, by Corollary 2.8, $F_{A,2}(x, x^*) \geq \langle x, x^*\rangle$. On the other hand, by Theorem 2.14, $F^*_{A,n}(x^*, x) \geq F_{A,2}(x, x^*)$. Altogether, $F_{A,n}(x^*, x) \geq \langle x, x^*\rangle$, as claimed. "$\Leftarrow$": Suppose that $(\forall (x, x^*) \in X \times X^*)$ $F^*_{A,n}(x^*, x) \geq \langle x, x^*\rangle$. Since A is maximal n-cyclically monotone, Corollary 2.8 yields $(\forall (x, x^*) \in X \times X^*)$ $F_{A,n}(x, x^*) \geq \langle x, x^*\rangle$. Therefore, by [18, Theorem 3.1] (see also [45, Theorem 1.4] and [34, Theorem 6]), the set $\big\{(x, x^*) \in X \times X^* \mid F_{A,n}(x, x^*) = \langle x, x^*\rangle\big\}$ is the graph of a maximal monotone operator. Since $F_{A,n} \geq \langle \cdot, \cdot\rangle$, this set is equal to $\big\{(x, x^*) \in X \times X^* \mid F_{A,n}(x, x^*) \leq \langle x, x^*\rangle\big\}$, which, by Theorem 2.13, is $\operatorname{gra} A$. ■

The following example illustrates that maximal n-cyclic monotonicity does not imply maximal monotonicity.

Example 2.16 There exists a maximal 3-cyclically monotone operator on $\mathbb{R}^2$ that is *not* maximal monotone.

Proof. Let $X = \mathbb{R}^2$ and define an operator $A\colon X \to 2^X$ via $\operatorname{gra} A = \{(z_1, z_1^*), \ldots, (z_4, z_4^*)\} \subset \mathbb{R}^2 \times \mathbb{R}^2$, where

$$\begin{aligned} (z_1, z_1^*) &= \big((1,0),(0,1)\big), & (z_2, z_2^*) &= \big((0,1),(-1,0)\big),\\ (z_3, z_3^*) &= \big((-1,0),(-1,-2)\big), & (z_4, z_4^*) &= \big((0,-1),(0,-1)\big). \end{aligned} \tag{18}$$

It is elementary to verify that A is 3-cyclically monotone. Zorn's Lemma guarantees the existence of an operator $B\colon X \to 2^X$ such that $\operatorname{gra} A \subset \operatorname{gra} B$ and such that

$$B \text{ is maximal 3-cyclically monotone.} \tag{19}$$

We shall prove by contradiction that

$$(0,0) \notin \operatorname{dom} B. \tag{20}$$

Thus assume there exists $(\xi, \eta) \in X$ such that $\big((0,0),(\xi,\eta)\big) \in \operatorname{gra} B$. Define $\widetilde{A}$ via $\operatorname{gra} \widetilde{A} = \operatorname{gra} A \cup \big\{\big((0,0),(\xi,\eta)\big)\big\}$. Then $\operatorname{gra} \widetilde{A} \subset \operatorname{gra} B$. Hence $\widetilde{A}$ is 3-cyclically monotone, which leads to the following set of conditions on (ξ, η): $\big\{\xi \le 0,\ \eta \le 0,\ \xi \ge -1,\ \eta \ge -1,\ \eta \le 1,\ \xi \le -1,\ \xi \ge -2,\ \eta \ge -1,\ \xi \le 2,\ \xi \ge -3,\ \eta \le 0,\ \eta \ge -2,\ \xi \ge 0\big\}$. Since the subset $\{\xi \le -1, \xi \ge 0\}$ of these conditions is inconsistent, we have arrived at a contradiction. This verifies (20). We now claim that

$$B \text{ is not maximal monotone.} \tag{21}$$

Once more, we argue by contradiction and thus assume that B is maximal monotone. A result of Simons [43, Theorem 18.3] implies that

$$\operatorname{int} \operatorname{dom} B = \operatorname{int} \operatorname{conv} \operatorname{dom} B. \tag{22}$$

Now $\operatorname{gra} A \subset \operatorname{gra} B \Rightarrow \operatorname{dom} A \subset \operatorname{dom} B \Rightarrow \operatorname{int} \operatorname{conv} \operatorname{dom} A \subset \operatorname{int} \operatorname{conv} \operatorname{dom} B$. Hence, using the definition of A and (22), we deduce that

$$(0,0) \in \operatorname{int} \operatorname{conv}\{z_1, z_2, z_3, z_4\} = \operatorname{int} \operatorname{conv} \operatorname{dom} A \subset \operatorname{int} \operatorname{conv} \operatorname{dom} B = \operatorname{int} \operatorname{dom} B \subset \operatorname{dom} B. \tag{23}$$

But (20) and (23) are contradictory, and we therefore have proved that B is not maximal monotone. In view of (19) and (21), we see that the operator B does the job. ■

3 Rockafellar's antiderivative

Recall the definition of the function $C_{A,n,(a,a^*)}$ in Definition 2.1 for an operator $A\colon X \to 2^{X^*}$ and the two corresponding sets of parameters, namely points $(a, a^*) \in \operatorname{gra} A$ and integers $n \in \{2, 3, \ldots\}$. The Fitzpatrick function of order n arises by keeping n fixed while supremizing over (a, a^*). Analogously, we shall see that Rockafellar's antiderivative is (essentially) obtained by keeping (a, a^*) fixed while supremizing over n. Therefore, the function $C_{A,n,(a,a^*)}$ can be viewed as the "common ancestor" of the Fitzpatrick functions and Rockafellar's antiderivative.

Definition 3.1 (Rockafellar function) *Let $A\colon X \to 2^{X^*}$ and $(a, a^*) \in \operatorname{gra} A$. Then we define the* Rockafellar function *by*

$$R_{A,(a,a^*)}\colon X \to \left]-\infty, +\infty\right] : x \mapsto \sup_{n \in \{2,3,\ldots\}} C_{A,n,(a,a^*)}(x, 0) \tag{24}$$

The importance of the Rockafellar function stems from the following fundamental result due to Rockafellar (see [38] or [48, Proposition 2.4.3, Theorem 3.2.8, and Corollary 3.2.11]). It states that maximal cyclically monotone operators are precisely the subdifferential operators of convex, lower semicontinuous and proper functions. Moreover, the Rockafellar functions are antiderivatives, which are unique up to constants.

Fact 3.2 (Rockafellar) *Let $f\colon X \to \left]-\infty,+\infty\right]$ be convex, lower semicontinuous, and proper. Then ∂f is maximal monotone and cyclically monotone, hence maximal cyclically monotone. Conversely, let $A\colon X \to 2^{X^*}$ be maximal cyclically monotone and let $(a,a^*) \in \operatorname{gra} A$. Then $R_{A,(a,a^*)}$ is convex, lower semicontinuous, and proper, $R_{A,(a,a^*)}(a) = 0$, and*

$$A = \partial R_{A,(a,a^*)} \tag{25}$$

is maximal monotone. If $\partial f = A$, then $f(\cdot) = f(a) + R_{A,(a,a^*)}(\cdot)$.

We shall utilize the following result of Borwein (see [8, Theorem 1] or [48, Theorem 3.1.4(i)]).

Fact 3.3 *Let $f\colon X \to \left]-\infty,+\infty\right]$ be convex, lower semicontinuous, and proper. Then*

$$(\forall \varepsilon > 0)(\forall x \in \operatorname{dom} f)(\exists\, y_{\varepsilon,x} \in \operatorname{dom} \partial f) \quad \|x - y_{\varepsilon,x}\| < \varepsilon \ \text{ and } \ |f(x) - f(y_{\varepsilon,x})| < \varepsilon. \tag{26}$$

The proof of the following result is borrowed from [3, Corollary 1.3].

Proposition 3.4 *Let $f\colon X \to \left]-\infty,+\infty\right]$ be convex, lower semicontinuous, and proper. Then*

$$(\forall x^* \in X^*) \quad f^*(x^*) = \sup_{x \in \operatorname{dom} \partial f} \big(\langle x, x^*\rangle - f(x)\big). \tag{27}$$

Proof. Let $h = f + \iota_{\operatorname{dom} \partial f}$. Then Fact 3.3 states that $\operatorname{gra} h$ is dense in $\operatorname{gra} f$. Hence for every $x^* \in X^*$, $f^*(x^*) = \sup\langle \operatorname{gra} f, (x^*,-1)\rangle = \sup\langle \operatorname{gra} h, (x^*,-1)\rangle = \sup_{x\in\operatorname{dom}\partial f}\big(\langle x, x^*\rangle - f(x)\big)$. ■

We now compute the Fitzpatrick function of infinite order and its Fenchel conjugate for subdifferentials. Further information on Fitzpatrick functions of order 2 for subdifferentials can be found in [7].

Theorem 3.5 *Let $f\colon X \to \left]-\infty,+\infty\right]$ be convex, lower semicontinuous, and proper. Then for every $(x,x^*) \in X \times X^*$,*

$$\langle x, x^*\rangle \le F_{\partial f,2}(x,x^*) \le F_{\partial f,3}(x,x^*) \le \cdots \le F_{\partial f,n}(x,x^*) \to F_{\partial f,\infty}(x,x^*) = f(x) + f^*(x^*) \tag{28}$$

and

$$F^*_{\partial f,2}(x^*,x) \ge F^*_{\partial f,3}(x^*,x) \ge \cdots \ge F^*_{\partial f,n}(x^*,x) \to h(x^*,x) \ge F^*_{\partial f,\infty}(x^*,x) = f^*(x^*) + f(x), \tag{29}$$

where $h\colon X^* \times X\colon \left]-\infty,+\infty\right]$ is convex and $h(x^*,x) \ge h^{**}(x^*,x) = f^*(x^*) + f(x)$.

Proof. By Definition 2.2 and (8), it is clear that $(F_{\partial f,n})_{n\in\{2,3,\dots\}}$ is an increasing sequence converging pointwise to $F_{\partial f,\infty}$. By Fact 3.2, ∂f is maximal monotone so that Corollary 2.8 implies that $\langle\cdot,\cdot\rangle \le F_{\partial f,2}$. Take $(x,x^*) \in X \times X^*$. Using Definitions 2.1&3.1, Fact 3.2 and Proposition 3.4, we see that

$$\begin{aligned}
F_{\partial f,\infty}(x,x^*) &= \sup_{n\in\{2,3,\dots\}} F_{\partial f,n}(x,x^*) = \sup_{n\in\{2,3,\dots\}} \sup_{(a,a^*)\in\operatorname{gra}\partial f} C_{\partial f,n,(a,a^*)}(x,x^*) \\
&= \sup_{(a,a^*)\in\operatorname{gra}\partial f} \sup_{n\in\{2,3,\dots\}} C_{\partial f,n,(a,a^*)}(x,0) + \langle a,x^*\rangle \\
&= \sup_{(a,a^*)\in\operatorname{gra}\partial f} \langle a,x^*\rangle + R_{\partial f,(a,a^*)}(x) \\
&= \sup_{(a,a^*)\in\operatorname{gra}\partial f} \langle a,x^*\rangle + \big(f(x)-f(a)\big) = f(x) + \sup_{a\in\operatorname{dom}\partial f} \big(\langle a,x^*\rangle - f(a)\big) \\
&= f(x) + f^*(x^*).
\end{aligned} \tag{30}$$

This proves (28). By Fenchel conjugation, we learn that $\big(F^*_{\partial f,n}(x^*,x)\big)_{n\in\{2,3,\dots\}}$ is a decreasing sequence converging to $h(x^*,x)$, where $h\colon X^*\times X\colon\ ]-\infty,+\infty]$ is a convex function such that $h(x^*,x) \ge h^{**}(x^*,x) \ge f^*(x^*)+f(x)$. Conjugating this decreasing sequence yields $F_{\partial f,2}(x,x^*) \le F_{\partial f,3}(x,x^*) \le \cdots \le F_{\partial f,n}(x,x^*) \le \cdots \le h^*(x,x^*)$. Hence $F_{\partial,\infty}(x,x^*) = f(x)+f^*(x^*) \le h^*(x,x^*)$. Conjugating this inequality yields $h^{**}(x^*,x) \le f^*(x^*)+f(x)$, which completes the proof. ■

4 Examples

Example 4.1 (reciprocation) Let $X=\mathbb{R}$ and let

$$f\colon \mathbb{R}\to\ ]-\infty,+\infty] : \xi \mapsto \begin{cases} -\ln\xi, & \text{if } \xi>0;\\ +\infty, & \text{otherwise.}\end{cases} \tag{31}$$

Then $(\forall\xi\in\mathbb{R})$ $\partial f(\xi) = -1/\xi$, if $\xi>0$; $\partial f(\xi)=\varnothing$, otherwise, and $(\forall\eta\in X)$ $f^*(\eta) = f(-\eta)-1$. Furthermore,

$$(\forall n\in\{2,3,\dots\})(\forall(\xi,\eta)\in\mathbb{R}^2)\quad F_{\partial f,n}(\xi,\eta) = \begin{cases}(n-1)-n\sqrt[n]{-\eta\xi}, & \text{if } \xi\ge 0 \text{ and } \eta\le 0;\\ +\infty, & \text{otherwise}\end{cases} \tag{32}$$

and $F_{\partial f,\infty}(\xi,\eta) = \lim_n F_{\partial f,n}(\xi,\eta) = f(\xi)+f^*(\eta)$.

Proof. Take $(\xi,\eta)\in\mathbb{R}^2$. Using (8), we see that

$$\begin{aligned}
F_{\partial f,n}(\xi,\eta) &= \sup_{\substack{(\alpha_1,\beta_1)\in\operatorname{gra}\partial f,\\ \vdots\\ (\alpha_{n-1},\beta_{n-1})\in\operatorname{gra}\partial f}} \left(\sum_{i=1}^{n-2}(\alpha_{i+1}-\alpha_i)\beta_i\right)+(\xi-\alpha_{n-1})\beta_{n-1}+\alpha_1\eta \\
&= \sup_{\alpha_1>0,\ldots,\alpha_{n-1}>0} \alpha_1\eta+\left(\sum_{i=1}^{n-2}(\alpha_{i+1}-\alpha_i)\frac{-1}{\alpha_i}\right)+(\xi-\alpha_{n-1})\frac{-1}{\alpha_{n-1}} \\
&= (n-1)-\inf_{\alpha_1>0,\ldots,\alpha_{n-1}>0}\left[\alpha_1(-\eta)+\frac{\alpha_2}{\alpha_1}+\frac{\alpha_3}{\alpha_2}+\cdots+\frac{\alpha_{n-1}}{\alpha_{n-2}}+\frac{\xi}{\alpha_{n-1}}\right]. \qquad (33)
\end{aligned}$$

If $\xi<0$, then $F_{\partial f,n}(\xi,\eta)=+\infty$ (via $\alpha_1=\alpha_2=\cdots=\alpha_{n-2}=1$ and $\alpha_{n-1}\downarrow 0$ in (33)). Similarly, if $\eta>0$, then $F_{\partial f,n}(\xi,\eta)=+\infty$. We thus assume that $\xi\geq 0$ and $\eta\leq 0$. Then the bracketed term in (33) is clearly nonnegative. If $\xi=0$, then $F_{\partial f,n}(\xi,\eta)=n-1$ since the bracketed term can be made arbitrarily close to 0 (consider $\alpha_2=\alpha_1^2,\ldots,\alpha_{n-1}=\alpha_1^{n-1}$ and let $\alpha_1\downarrow 0$). An analogous argument shows that if $\eta=0$, then $F_{\partial f,n}(\xi,\eta)=n-1$ as well. Finally, we assume that $\xi>0$ and $\eta<0$. Then the Arithmetic-Mean-Geometric-Mean inequality shows that

$$\alpha_1(-\eta)+\frac{\alpha_2}{\alpha_1}+\frac{\alpha_3}{\alpha_2}+\cdots+\frac{\alpha_{n-1}}{\alpha_{n-2}}+\frac{\xi}{\alpha_{n-1}}\geq n\sqrt[n]{\alpha_1(-\eta)\frac{\alpha_2}{\alpha_1}\frac{\alpha_3}{\alpha_2}\cdots\frac{\alpha_{n-1}}{\alpha_{n-2}}\frac{\xi}{\alpha_{n-1}}}=n\sqrt[n]{-\eta\xi}, \qquad (34)$$

and equality in (34) holds exactly when

$$\alpha_1(-\eta)=\frac{\alpha_2}{\alpha_1}=\frac{\alpha_3}{\alpha_2}\cdots=\frac{\alpha_{n-1}}{\alpha_{n-2}}=\frac{\xi}{\alpha_{n-1}}, \qquad (35)$$

which is a system of equations that can be solved by forward or backward substitution. Therefore, $F_{\partial f,n}(\xi,\eta)=(n-1)-n\sqrt[n]{-\eta\xi}$. The formula for $F_{\partial f,\infty}$ follows from (28). ■

We now give an example where all Fitzpatrick functions coincide. This example will be considerably strengthened in Section 5.

Example 4.2 (normal cone operator) Let X be a real Hilbert space and let C be a nonempty closed convex subset of X. Set $N_C=\partial\iota_C$. Then

$$(\forall n\in\{2,3,\ldots\})(\forall(x,x^*)\in X\times X)\quad F_{N_C,n}(x,x^*)=F_{N_C,\infty}(x,x^*)=\iota_C(x)+\iota_C^*(x^*). \qquad (36)$$

Proof. Take $n\in\{2,3,\ldots\}$ and $(x,x^*)\in X\times X$. On the one hand, as shown in [7, Example 3.1], $F_{N_C,2}(x,x^*)=\iota_C(x)+\iota_C(x^*)$. On the other hand, (28) implies that $F_{N_C,2}(x,x^*)\leq F_{N_C,n}(x,x^*)\leq F_{N_C,\infty}(x,x^*)=\iota_C(x)+\iota_C^*(x^*)$. Altogether, we conclude that (36) holds. ■

Remark 4.3 Let $X=\mathbb{R}$ and let

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

Denote the inverse of the function $[0,+\infty[\ \to [0,+\infty[\ : \rho \to \rho e^\rho$ by W. (The function W is known as the *Lambert W function.*) Then by [7, Example 3.6]

$$F_{\partial f,2}\colon (\rho,\rho^*) \mapsto \begin{cases} +\infty, & \text{if } \rho < 0; \\ \exp(\rho^*-1), & \text{if } \rho = 0; \\ \rho\rho^* + \rho\big(W(\kappa) + \frac{1}{W(\kappa)} - 2\big), & \text{if } \rho > 0 \text{ and } \kappa = \rho e^{1-\rho^*}. \end{cases} \tag{38}$$

Unfortunately, we were unable to find a closed form for $F_{\partial f,n}$ for $n \in \{3,4,\ldots\}$.

Example 4.4 (identity) Let X be a real Hilbert space and $n \in \{2,3,\ldots\}$. Then

$$F_{\mathrm{Id},n}\colon X\times X \to \mathbb{R}\colon (x,x^*) \mapsto \tfrac{1}{2}\|x\|^2+\tfrac{1}{2}\|x^*\|^2-\tfrac{1}{2n}\|x-x^*\|^2 = \tfrac{n-2}{2n}\big(\|x\|^2+\|x^*\|^2\big)+\tfrac{1}{2n}\|x+x^*\|^2 \tag{39}$$

and $F_{\mathrm{Id},\infty}\colon X \times X \to \mathbb{R}\colon (x,x^*) \mapsto \frac{1}{2}\|x\|^2 + \frac{1}{2}\|x^*\|^2$.

Proof. Take $(x,x^*) \in X \times X$. Using (8), we see that

$$F_{\mathrm{Id},n}(x,x^*) = \sup_{x_1,\ldots,x_{n-1}} \Bigg[\sum_{i=1}^{n-2}\langle x_{i+1}-x_i, x_i\rangle + \langle x - x_{n-1}, x_{n-1}\rangle + \langle x^*, x_1\rangle\Bigg]. \tag{40}$$

The bracketed term can be rewritten as

$$\begin{aligned} &- \langle x_1,x_1\rangle + \langle x_2,x_1\rangle - \langle x_2,x_2\rangle + \langle x_2,x_3\rangle - \langle x_3,x_3\rangle + \cdots \\ &- \langle x_{n-2},x_{n-2}\rangle + \langle x_{n-2},x_{n-1}\rangle - \langle x_{n-1},x_{n-1}\rangle + \langle x,x_{n-1}\rangle + \langle x^*,x_1\rangle \\ =& -\tfrac{1}{2}\|x_1\|^2 - \tfrac{1}{2}\|x_1-x_2\|^2 - \tfrac{1}{2}\|x_2-x_3\|^2 - \cdots - \tfrac{1}{2}\|x_{n-2}-x_{n-1}\|^2 - \tfrac{1}{2}\|x_{n-1}\|^2 \\ &+ \langle x,x_{n-1}\rangle + \langle x^*,x_1\rangle \\ =& \tfrac{1}{2}\|x\|^2 + \tfrac{1}{2}\|x^*\|^2 - \tfrac{1}{2}\|x^*-x_1\|^2 - \tfrac{1}{2}\|x_1-x_2\|^2 - \cdots - \tfrac{1}{2}\|x_{n-2}-x_{n-1}\|^2 - \tfrac{1}{2}\|x_{n-1}-x\|^2, \end{aligned} \tag{41}$$

which is clearly a concave and differentiable function in $\mathbf{x} = (x_1,\ldots,x_{n-1})$, call it Ψ. Then $\nabla\Psi(\mathbf{x}) = \mathbf{0}$ is equivalent to the linear system

$$-2x_1 + x_2 + x^* = 0 \tag{42}$$
$$x_1 - 2x_2 + x_3 = 0 \tag{43}$$
$$x_2 - 2x_3 + x_4 = 0 \tag{44}$$
$$\vdots \tag{45}$$
$$x_{n-3} - 2x_{n-2} + x_{n-1} = 0 \tag{46}$$
$$x_{n-2} - 2x_{n-1} + x = 0. \tag{47}$$

This system has either no or a unique solution since $\Psi(\mathbf{x})$ is strictly concave in x_1 and x_{n-1} and since the other variables can be obtained by forward or backward substitution. In fact, the latter case is true: setting

$$(\forall k \in \{1,2,\ldots,n-1\}) \quad x_k = x^* + \tfrac{k}{n}(x-x^*), \tag{48}$$

one verifies that (42)–(47) holds and, by (41), that $F_{\mathrm{Id},n}(x,x^*) = \frac{1}{2}\|x\|^2 + \frac{1}{2}\|x^*\|^2 - \frac{1}{2n}\|x-x^*\|^2$. Finally, $\mathrm{Id} = \nabla\frac{1}{2}\|\cdot\|^2$ and $\big(\frac{1}{2}\|\cdot\|^2\big)^* = \frac{1}{2}\|\cdot\|^2$, so the formula for $F_{\mathrm{Id},\infty}$ follows from (28) (or by letting n tend to $+\infty$ in (39)). ■

Example 4.5 (skew operator) Let X be a Hilbert space and let $A\colon X \to X$ be a *skew-symmetric*, i.e., A is continuous, linear, and $A^* = -A$. Furthermore, suppose that $A \neq 0$ and that $n \in \{3, 4, \ldots\}$. Then $F_{A,2} = \iota_{\operatorname{gra} A}$ and $F_{A,n} \equiv +\infty$. Consequently, A is maximal monotone, but it is not n-cyclically monotone.

Proof. We shall use repeatedly that $(\forall y \in X)\ \langle y, Ay\rangle = 0$. Let $(x, x^*) \in X \times X$. Using (8), we see that

$$\begin{aligned} F_{A,2}(x,x^*) &= \sup_{y\in X}\big(\langle x, Ay\rangle + \langle y, x^*\rangle - \langle x, Ax\rangle\big) = \sup_{y\in X}\langle A^*x + x^*, y\rangle \\ &= \sup_{y\in X}\langle x^* - Ax, y\rangle = \iota_{\{0\}}(x^* - Ax) = \iota_{\operatorname{gra} A}(x, x^*). \end{aligned} \tag{49}$$

Since $A \neq 0$, we deduce that $A^* \neq 0$ and hence there exists $\widetilde{z} \in X$ such that $A^*\widetilde{z} + x^* \neq 0$. Then

$$\begin{aligned} F_{A,3}(x,x^*) &= \sup_{y,z}\big(\langle z - y, Ay\rangle + \langle x - z, Az\rangle + \langle y, x^*\rangle\big) = \sup_{y,z}\big(\langle z, Ay\rangle + \langle x, Az\rangle + \langle y, x^*\rangle\big) \\ &\geq \sup_{y}\langle A^*\widetilde{z} + x^*, y\rangle + \langle x, A\widetilde{z}\rangle = +\infty. \end{aligned} \tag{50}$$

Using (14), we see that $F_{A,n} \equiv +\infty$. The "Consequently" part follows from Theorem 2.9 and Theorem 2.13. ■

In [1], Asplund provides examples of matrices that are n-cyclically monotone but not $(n+1)$-cyclically monotone. However, his statement is not very explicit. For completeness, we give a simpler proof of his observation that the matrix corresponding to the rotation by π/n in the Euclidean plane is n-cyclically monotone yet not $(n+1)$-cyclically monotone.

Example 4.6 (rotations) Let $X = \mathbb{R}^2$ and let $n \in \{2, 3, \ldots\}$. Denote the matrix corresponding to counter-clockwise rotation by π/n by R_n, i.e.,

$$R_n = \begin{pmatrix} \cos(\pi/n) & -\sin(\pi/n) \\ \sin(\pi/n) & \cos(\pi/n) \end{pmatrix}. \tag{51}$$

Then R_n is maximal monotone and n-cyclically monotone, but R_n is not $(n+1)$-cyclically monotone.

Proof. It is clear that R_n is monotone, and that $R_n^* = R_n^{-1}$. Since $\operatorname{dom} R_n = X$, the maximal monotonicity of R_n is thus a consequence of [43, page 30]. Since R_2 is skew-symmetric, Example 4.5 implies that R_2 is not 3-cyclically monotone. We have verified the conclusion for $n = 2$ and thus assume for the remainder of the proof that $n \in \{3, 4, \ldots\}$. Let us show first that R_n is not $(n+1)$-cyclically monotone. Take $x \in X \smallsetminus \{0\}$. Since $R_n + R_n^*$ is invertible (in fact, a strictly positive multiple of the identity), there exists $a \in X$ such that $\frac{1}{2}R_n a + \frac{1}{2}R_n^* a = R_n^* x$. Note that $a \neq 0$ (since $x \neq 0$) and that $R_n a \neq R_n^* a$ (since $\pi/n < \pi$). The fact that R_n is an isometry and the parallelogram law thus yield $4\|a\|^2 = 2\|R_n a\|^2 + 2\|R_n^* a\|^2 = \|R_n a + R_n^* a\|^2 + \|R_n a - R_n^* a\|^2 > \|R_n a + R_n^* a\|^2 = \|2R_n^* x\|^2 = 4\|x\|^2$. Hence

$$\|a\| > \|x\|. \tag{52}$$

Furthermore, $R_n a + R_n^* a = 2R_n^* x$ implies that $2\langle a, R_n a\rangle = \langle a, R_n a + R_n^* a\rangle = 2\langle a, R_n^* x\rangle = 2\langle R_n a, x\rangle$. Using (52), we note that

$$-\langle a, R_n a\rangle + 2\langle x, R_n a\rangle = \langle a, R_n a\rangle = \|a\|^2 \cos(\pi/n) > \|x\|^2 \cos(\pi/n) = \langle x, R_n x\rangle. \tag{53}$$

We now take n points from $\operatorname{gra} R_n$ by setting

$$(\forall i \in \{1, 2, \ldots, n\}) \quad (a_i, a_i^*) = (R_n^{2i} a, R_n^{2i+1} a). \tag{54}$$

Then, since R_n is an isometry, we have for every $i \in \{1, 2, \ldots, n-1\}$,

$$\begin{aligned}\langle a_{i+1} - a_i, a_i^*\rangle &= \langle R_n^{2i+2} a - R_n^{2i} a, R_n^{2i+1} a\rangle = \langle R_n^{2i+2} a, R_n^{2i+1} a\rangle - \langle R_n^{2i} a, R_n^{2i+1} a\rangle \\ &= \langle R_n a, a\rangle - \langle a, R_n a\rangle = 0.\end{aligned} \tag{55}$$

Using (8), (54), (55), the fact that $R_n^{2n} = \operatorname{Id}$ and that R_n is an isometry, and (53), we deduce that

$$\begin{aligned}F_{R_n, n+1}(x, R_n x) &\geq \left(\sum_{i=1}^{n-1} \langle a_{i+1} - a_i, a_i^*\rangle\right) - \langle a_n, a_n^*\rangle + \langle x, a_n^*\rangle + \langle a_1, R_n x\rangle \\ &= -\langle R_n^{2n} a, R_n^{2n+1} a\rangle + \langle x, R_n^{2n+1} a\rangle + \langle R_n^2 a, R_n x\rangle \\ &= -\langle a, R_n a\rangle + \langle x, R_n a\rangle + \langle R_n a, x\rangle \\ &> \langle x, R_n x\rangle.\end{aligned} \tag{56}$$

Thus $F_{R_n, n+1} > p$ on $\operatorname{gra} R_n \smallsetminus \{(0,0)\}$, and therefore, by Proposition 2.4, R_n is not $(n+1)$-cyclically monotone.

It remains to show that R_n is n-cyclically monotone. Take $x_1 = (\xi_1, \eta_1), \ldots, x_n = (\xi_n, \eta_n)$ in X, and set $x_{n+1} = x_1$. We must show that

$$0 \geq \sum_{i=1}^{n} \langle x_{i+1} - x_i, R_n x_i\rangle. \tag{57}$$

We now identify $\mathbb{R}^2$ with $\mathbb{C}$ in the usual way: $x = (\xi, \eta)$ in $\mathbb{R}^2$ corresponds to $\xi + \mathrm{i}\eta$ in $\mathbb{C}$, where $\mathrm{i} = \sqrt{-1}$ and $\langle x, y\rangle = \operatorname{Re}(\overline{x} y)$ for x and y in $\mathbb{C}$. The operator R_n corresponds to complex multiplication by

$$\omega = \exp(\mathrm{i}\pi/n). \tag{58}$$

Thus our aim is to show that $0 \geq \operatorname{Re}\big(\sum_{i=1}^n \overline{(x_{i+1} - x_i)} \omega x_i\big) = \sum_{i=1}^n \operatorname{Re}\big(\overline{(x_{i+1} - x_i)} \omega x_i\big)$, an inequality which we now reformulate in $\mathbb{C}^n$. Denote the $n \times n$-identity matrix by $\mathbf{I}$ and set

$$\mathbf{B} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \vdots \\ \vdots & & \ddots & \ddots & \\ & & & & 0 \\ 0 & & & & 1 \\ 1 & 0 & \cdots & & 0 \end{pmatrix} \in \mathbb{C}^{n \times n} \quad \text{and} \quad \mathbf{R} = \omega \mathbf{I} \in \mathbb{C}^{n \times n}. \tag{59}$$

Identifying $\mathbf{x} \in \mathbb{C}^n$ with $(x_1, \ldots, x_n) \in X^n$, we note that (57) means $0 \geq \operatorname{Re}\big(((\mathbf{B}-\mathbf{I})\mathbf{x})^*\mathbf{R}\mathbf{x}\big)$; equivalently, $0 \geq \mathbf{x}^*(\mathbf{B}^*-\mathbf{I})\mathbf{R}\mathbf{x}+\mathbf{x}^*\mathbf{R}^*(\mathbf{B}-\mathbf{I})\mathbf{x}$. In other words, we need to show that the Hermitian matrix

$$\mathbf{C} = (\mathbf{I}-\mathbf{B}^*)\mathbf{R} + \mathbf{R}^*(\mathbf{I}-\mathbf{B}) = \begin{pmatrix} (\omega+\overline{\omega}) & -\overline{\omega} & 0 & \cdots & 0 & -\omega \\ -\omega & (\omega+\overline{\omega}) & \ddots & & & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & & & 0 \\ 0 & & & & (\omega+\overline{\omega}) & -\overline{\omega} \\ -\overline{\omega} & 0 & \cdots & 0 & -\omega & (\omega+\overline{\omega}) \end{pmatrix} \tag{60}$$

is positive semidefinite. The matrix $\mathbf{C}$ is a circulant (Toeplitz) matrix and thus belongs to a class of well-studied matrices that have close connections to Fourier Analysis. E.g., by [20, Chapter 3] or by [31, Exercise 5.8.12], the set of (n not necessarily distinct) eigenvalues of $\mathbf{C}$ is

$$\Lambda = \big\{q(1), q(\omega^2), \ldots, q(\omega^{2(n-1)})\big\}, \quad \text{where} \quad q\colon t \mapsto (\omega+\overline{\omega}) - \omega t - \overline{\omega} t^{n-1}. \tag{61}$$

Since $\omega^{2n} = 1$, we verify that $\Lambda = \big\{2\cos(\pi/n) - 2\cos((2k+1)\pi/n) \mid k \in \{0, 1, \ldots, n-1\}\big\}$ is a set of nonnegative real numbers, as required. ■

Remark 4.7 Let X and R_n be as in Example 4.6. We claim that

$$(\forall n \in \{2,3,4\}) \quad F_{R_n,n+1} \equiv +\infty. \tag{62}$$

Indeed, take $n \in \{2,3,4\}$, $(x,x^*) \in X \times X^*$ and $a \in X$, and set

$$(\forall i \in \{1,2,\ldots,n\}) \quad (a_i, a_i^*) = (R_n^{i-1}a, R_n^i a). \tag{63}$$

Then (8), the fact that R_n is an isometry, and the Cauchy-Schwarz inequality imply that

$$\begin{aligned} F_{R_n,n+1}(x,x^*) &\geq \bigg(\sum_{i=1}^{n-1} \langle R_n^i a - R_n^{i-1}a, R_n^i a\rangle\bigg) - \langle R_n^{n-1}a, R_n^n a\rangle + \langle x, R_n^n a\rangle + \langle a, x^*\rangle \\ &= \bigg(\sum_{i=1}^{n-1} \langle R_n a - a, R_n a\rangle\bigg) - \langle a, R_n a\rangle + \langle x, R_n^n a\rangle + \langle a, x^*\rangle \\ &\geq (n-1)\|a\|^2 - n\langle a, R_n a\rangle - \|x\|\|a\| - \|a\|\|x^*\| \\ &= \|a\|^2\big(n-1-n\cos(\pi/n)\big) - \|a\|\big(\|x\| + \|x^*\|\big). \end{aligned} \tag{64}$$

Since $n \in \{2,3,4\}$, we know that $n-1-n\cos(\pi/n) > 0$ and hence that (64) tends to $+\infty$ as $\|a\| \to +\infty$. Therefore, (62) is verified. However, our knowledge of $F_{R_n,n+1}$ for $n \in \{5,6,\ldots\}$ is rather limited.

5 The Fitzpatrick family for subdifferential operators of sublinear and of indicator functions

The following family of functions was first studied in [21] and then in [17].

Definition 5.1 (Fitzpatrick family) *Let $A\colon X \to 2^{X^*}$ be maximal monotone. The* Fitzpatrick family $\mathcal{F}_A$ *associated with A consists of all functions $F\colon X \times X^* \to \left]-\infty,+\infty\right]$ such that*

$$F \text{ is convex, lower semicontinuous, } p \leq F, \text{ and } p = F \text{ on } \operatorname{gra} A. \tag{65}$$

Let $A\colon X \to 2^{X^*}$ be maximal monotone. As mentioned in Fact 1.4, Fitzpatrick proved in [21] that $F_{A,2} \in \mathcal{F}_A$ and that it is the smallest member of $\mathcal{F}_A$:

$$\big(\forall (x,x^*) \in X \times X^*\big) \quad F_{A,2}(x,x^*) = \min\big\{F(x,x^*) \mid F \in \mathcal{F}_A\big\} = (\iota_{\operatorname{gra} A} + p)^*(x^*,x). \tag{66}$$

Fitzpatrick also proved that for every $(x,x^*) \in X \times X^*$, $F_{A,2}(x,x^*) \leq F^*_{A,2}(x^*,x) \leq \iota_{\operatorname{gra} A} + p$. This implies that $(\iota_{\operatorname{gra} A} + p)^{**}|_{X \times X^*}$ belongs to $\mathcal{F}_A$ and is, in fact, the largest member of $\mathcal{F}_A$:

$$\big(\forall (x,x^*) \in X \times X^*\big) \quad (\iota_{\operatorname{gra} A} + p)^{**}(x,x^*) = \max\big\{F(x,x^*) \mid F \in \mathcal{F}_A\big\}. \tag{67}$$

For any two functions $f\colon X \to \left]-\infty,+\infty\right]$ and $g\colon Y \to \left]-\infty,+\infty\right]$, where Y is a real Banach space, we denote the function $X \times Y \to \left]-\infty,+\infty\right] : (x,y) \mapsto f(x) + g(y)$ by $f \oplus g$. Note that if $f\colon X \to \left]-\infty,+\infty\right]$ is convex, lower semicontinuous, and proper, then ∂f is maximal monotone and $f \oplus f^* \in \mathcal{F}_{\partial f}$. In [17], Burachik and Svaiter established the following new characterization of the subdifferential operator: *If $\mathcal{F}_A$ contains a separable member $f \oplus g$, then necessarily $A = \partial f$ and $g = f^*$.* In this section, we shall present two examples (see Theorem 5.3 and Corollary 5.9) where the Fitzpatrick family $\mathcal{F}_{\partial f}$ reduces to the singleton $\{f \oplus f^*\}$. We shall require the following useful properties of sublinear functions (see, e.g., [3, page 26] or [48, Theorem 2.4.14]). Recall that $f\colon X \to \left]-\infty,+\infty\right]$ is *sublinear* if f is convex, $f(0) = 0$, and $(\forall x \in X)(\forall \lambda > 0)$ $f(\lambda x) = \lambda f(x)$.

Fact 5.2 *Let $f\colon X \to \left]-\infty,+\infty\right]$ be sublinear, lower semicontinuous, and proper. Then*

(i) $f^* = \iota_{\partial f(0)}$, *and*

(ii) $(\forall z \in X)(\forall \lambda > 0) \quad \partial f(\lambda z) = \partial f(z) = \big\{z^* \in \partial f(0) \mid \langle z, z^*\rangle = f(z)\big\}$.

(iii) $f = \iota^*_{\partial f(0)}|_X = \sup\langle \cdot, \partial f(0)\rangle$.

Theorem 5.3 *Let $f\colon X \to \left]-\infty,+\infty\right]$ be sublinear, lower semicontinuous, and proper. Then* $\mathcal{F}_{\partial f} = \{f \oplus f^*\}$.

Proof. We claim that

$$\iota_{\partial f(0)} \oplus f^{**} = f^* \oplus f^{**} = (\iota_{\operatorname{gra} \partial f} + p)^*. \tag{68}$$

The left equality in (68) is clear from Fact 5.2(i). Fix $(x^*, x^{**}) \in X^* \times X^{**}$ and assume first that $x^* \notin \partial f(0)$. Then

$$(f^* \oplus f^{**})(x^*, x^{**}) = \iota_{\partial f(0)}(x^*) + f^{**}(x^{**}) = +\infty \tag{69}$$

and there exists $w \in X$ such that $f(w) < \langle w, x^* \rangle$. Now Fact 3.3 guarantees the existence of a point $z \in \operatorname{dom} \partial f$ such that

$$f(z) < \langle z, x^* \rangle. \tag{70}$$

Using (70) and Fact 5.2(i), we estimate

$$\begin{aligned}
+\infty &\geq (\iota_{\operatorname{gra} \partial f} + p)^*(x^*, x^{**}) = \sup_{(y,y^*) \in \operatorname{gra} \partial f} \big(\langle y, x^* \rangle + \langle y^*, x^{**} \rangle - \langle y, y^* \rangle\big) && (71)\\
&\geq \sup_{\substack{\lambda > 0\\ z^* \in \partial f(\lambda z)}} \big(\langle \lambda z, x^* \rangle + \langle z^*, x^{**} \rangle - \langle \lambda z, z^* \rangle\big) \geq \sup_{\substack{\lambda > 0\\ z^* \in \partial f(0)\\ \langle z, z^* \rangle = f(z)}} \big(\lambda\big(\langle z, x^* \rangle - \langle z, z^* \rangle\big) + \langle z^*, x^{**} \rangle\big)\\
&= \sup_{\substack{\lambda > 0\\ z^* \in \partial f(0)\\ \langle z, z^* \rangle = f(z)}} \big(\lambda\big(\langle z, x^* \rangle - f(z)\big) + \langle z^*, x^{**} \rangle\big) = \sup_{\lambda > 0} \lambda\big(\langle z, x^* \rangle - f(z)\big) + \sup_{\substack{z^* \in \partial f(0)\\ \langle z, z^* \rangle = f(z)}} \langle z^*, x^{**} \rangle\\
&= +\infty. && (72)
\end{aligned}$$

Altogether, (69) and (71)–(72) imply that (68) holds when $x^* \notin \partial f(0)$. We now assume that $x^* \in \partial f(0)$. Using Fact 5.2(i), we then obtain

$$\begin{aligned}
(f^* \oplus f^{**})(x^*, x^{**}) &= \iota_{\partial f(0)}(x^*) + f^{**}(x^{**}) = f^{**}(x^{**}) = \iota^*_{\partial f(0)}(x^{**})\\
&= \sup_{y^* \in X^*} \big(\langle y^*, x^{**} \rangle - \iota_{\partial f(0)}(y^*)\big) = \sup_{y^* \in \partial f(0)} \big(\langle 0, x^* \rangle + \langle y^*, x^{**} \rangle - \langle 0, y^* \rangle\big)\\
&\leq \sup_{(y,y^*) \in \operatorname{gra} \partial f} \big(\langle y, x^* \rangle + \langle y^*, x^{**} \rangle - \langle y, y^* \rangle\big)\\
&= (\iota_{\operatorname{gra} \partial f} + p)^*(x^*, x^{**}). && (73)
\end{aligned}$$

On the other hand, [38] or [43, Lemma 34.6 and Theorem 34.8] guarantee the existence of a net (x_α) in X such that $x_\alpha \stackrel{*}{\rightharpoonup} x^{**}$ (i.e., weak* convergence) and $f(x_\alpha) \to f^{**}(x^{**})$. Using (66) and Theorem 3.5, the fact that any conjugate function is weak* lower semicontinuous and that $(x^*, x_\alpha) \stackrel{*}{\rightharpoonup} (x^*, x^{**})$, we deduce

$$\begin{aligned}
(\iota_{\operatorname{gra} \partial f} + p)^*(x^*, x^{**}) &\leq \underline{\lim} (\iota_{\operatorname{gra} \partial f} + p)^*(x^*, x_\alpha) = \underline{\lim}\, F_{\partial f, 2}(x_\alpha, x^*) \leq \underline{\lim}\, f(x_\alpha) + f^*(x^*)\\
&= f^{**}(x^{**}) + f^*(x^*) = (f^* \oplus f^{**})(x^*, x^{**}). && (74)
\end{aligned}$$

Combining (73) and (74), we conclude that $(\iota_{\operatorname{gra} \partial f} + p)^*(x^*, x^{**}) = (f^* \oplus f^{**})(x^*, x^{**})$ in this case as well. Therefore, we have verified (68) and hence

$$f^* \oplus f^{**} = (\iota_{\operatorname{gra} \partial f} + p)^*. \tag{75}$$

Now (75), the fact that $f^{**}|_X = f$ (see [43, Remark 34.4]), and (66) imply that $(\forall (x, x^*) \in X \times X^*)$ $(f \oplus f^*)(x, x^*) = f^*(x^*) + f(x) = (f^* \oplus f^{**})(x^*, x) = (\iota_{\operatorname{gra} \partial f} + p)^*(x^*, x) = F_{\partial f, 2}(x, x^*)$, i.e., that

$$F_{\partial f, 2} = f \oplus f^*. \tag{76}$$

Conjugating (75) we obtain $f^{**} \oplus f^{***} = (\iota_{\operatorname{gra}\partial f} + p)^{**}$; restricting this to $X \times X^*$ we arrive at

$$f \oplus f^* = (\iota_{\operatorname{gra}\partial f} + p)^{**}|_{X\times X^*}. \tag{77}$$

In view of (66)&(67), we deduce from (76)&(77) that $f \oplus f^*$ is both the smallest and the largest member of $\mathcal{F}_{\partial f}$, i.e., that $f \oplus f^*$ is the *only* member of $\mathcal{F}_{\partial f}$. ■

The proof of Theorem 5.3 relies critically on Fact 3.3. This deeper result can be avoided if f is everywhere subdifferentiable. However, as the following example shows, a proper sublinear lower semicontinuous function may fail to be subdifferentiable on its domain.

Example 5.4 Suppose that $X = \mathbb{R}^2$ and let $C = \big\{(\xi,\eta) \in X^* \mid 0 < 1/\xi \le \eta\big\}$. Then C is closed, convex, and nonempty, and the corresponding proper, lower semicontinuous, and sublinear support function $f = \iota_C^*$ is given by

$$f\colon X \to \left]-\infty,+\infty\right] : (\xi,\eta) \mapsto \begin{cases} -2\sqrt{\xi\eta}, & \text{if } \xi \le 0 \text{ and } \eta \le 0;\\ +\infty, & \text{otherwise.}\end{cases} \tag{78}$$

Moreover, the function f is not subdifferentiable at any point which belongs to the boundary of its domain, except for the origin.

Proof. Let x be a point which belongs to the boundary of the domain of f which is not the origin. Then $x = (0,\eta)$ where $\eta < 0$ or $x = (\xi,0)$ where $\xi < 0$. In any case, the supremum of x as a functional on C is not attained. If we show that $C = \partial f(0)$, it will then follow from the combination of Fact 5.2(ii)&(iii) that $\partial f(x) = \varnothing$. Indeed, from the definitions of $\partial f(0)$ and f it follows that $C \subset \partial f(0)$. If $y^* \in X^* \smallsetminus C$, then there is a functional $y \in X$ strictly separating C from y^*, that is, $\langle y, y^*\rangle > \sup_{x^*\in C}\langle y, x^*\rangle = f(y)$, which means that $y^* \notin \partial f(0)$, so $\partial f(0) \subset C$. Another point of view (due to S. Simons) is the following. For a fixed $\xi < 0$, the function $f(\xi,\cdot)$ has a vertical tangent at 0. It follows that f is not subdifferentiable at $(\xi, 0)$. ■

Remark 5.5 Several comments on Theorem 5.3 and Example 5.4 are in order.

(i) Theorem 5.3 was first observed by Penot [34] when X is reflexive. It was also established by Burachik and Fitzpatrick [16] when X is not necessarily reflexive, but f is everywhere finite; in fact, in the proof of Theorem 5.3 we generalize some of their calculations. (The statement of Theorem 5.3 appears in [16], but in a forthcoming corrigendum it is restated to hold when $\operatorname{dom} f = X$.)

(ii) Example 5.4 provides a proper, lower semicontinuous, and sublinear function that is not everywhere subdifferentiable on its domain, which is a cone in the Euclidean plane. We now provide two approaches to the construction of a function of this kind in higher-dimensional spaces.

  (a) Assume first that X is a real Hilbert space of dimension 3 or higher, and write $X = \mathbb{R}^2 \oplus Y$. Let C and f be as in Example 5.4, and set $D = C \oplus \{0\}$. Then $\iota_D^*\colon \mathbb{R}^2 \times Y \to \left]-\infty,+\infty\right] : \big((\xi,\eta),y\big) \mapsto f(\xi,\eta) + \iota_Y(y) = f(\xi,\eta)$ and hence $\operatorname{dom}\partial\iota_D^* = (\operatorname{dom}\partial f)\oplus Y \subsetneqq (\operatorname{dom} f)\oplus Y = \operatorname{dom}\iota_D^*$.

(b) Now assume that Y is a real Banach space, that $g\colon Y \to \left]-\infty,+\infty\right]$ is convex, lower semicontinuous, and proper, and that $y_0 \in (\operatorname{dom} g) \smallsetminus (\operatorname{dom} \partial g)$. Set $X = Y \times \mathbb{R}$ and $x_0 = (y_0, 1)$, and define the corresponding *perspective function* (see, e.g., [10, Exercise 24 on page 84], [13, Section 3.2.6], and [40, pages 35 and 67] for further information on this construction) by

$$h\colon X \to \left]-\infty,+\infty\right] : x = (y,\rho) \mapsto \begin{cases} \rho g(y/\rho), & \text{if } \rho > 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{79}$$

Since $\operatorname{epi} h$ is the convex cone generated by $(\operatorname{epi} g) \times \{1\}$, it is clear that $f = h^{**}|_X$ is sublinear, lower semicontinuous, and proper. The lower semicontinuity of g implies the lower semicontinuity of h on $(\operatorname{dom} g) \times \{1\}$; in particular, $(\forall y \in \operatorname{dom} g)\ f(y,1) = h(y,1) = g(y)$. We claim that

$$x_0 \in (\operatorname{dom} f) \smallsetminus (\operatorname{dom} \partial f). \tag{80}$$

It is clear that $x_0 \in \operatorname{dom} f$ since $f(x_0) = h(x_0) = g(y_0) \in \mathbb{R}$. Assume to the contrary that $\partial f(x_0) \neq \varnothing$, say $x^* = (y^*, \rho^*) \in \partial f(x_0)$. Then $(\forall y \in \operatorname{dom} g)\ \langle y - y_0, y^*\rangle = \langle (y,1) - (y_0,1), (y^*,\rho^*)\rangle \leq f(y,1) - f(y_0,1) = h(y,1) - h(y_0,1) = g(y) - g(y_0)$, which implies the absurdity $y^* \in \partial g(y_0)$. This verifies (80).

(iii) Let us provide some guidance on how to find Y, g, and y_0 with the properties required in (ii)(b). Assume that Y is a real Banach space and that S is a nonempty, bounded, closed, and convex subset of Y. Fix $d \in Y \smallsetminus \{0\}$. Following [25, Section 20.D] and [11, Example 5], we define

$$g\colon Y \to \left]-\infty,+\infty\right] : y \mapsto \min\big\{\rho \in \mathbb{R} \mid y + \rho d \in S\big\}. \tag{81}$$

Then g is convex, lower semicontinuous, proper, and $(\forall y \in Y)(\forall \rho \in \mathbb{R})\ g(y + \rho d) = g(y) - \rho$. Hence g is unbounded below — which implies that 0 is never a subgradient — and $(\forall y \in Y)$ $\partial g(y) = \partial g\big(y + g(y)d\big)$. Moreover,

$$(\forall y \in \operatorname{dom} g) \qquad y + g(y)d \in S \quad \text{and} \\ \partial g\big(y + g(y)d\big) \subset \big\{\text{nonzero support functionals for } S \text{ at } y + g(y)d\big\}. \tag{82}$$

Assume in addition that Y is separable, that $\operatorname{int} S = \varnothing$ (equivalently, that the core of S is empty [25, Lemma 17.E]), and that S is not contained in any closed hyperplane of Y. By [25, Exercise 2.18], there exists a point $y_0 \in S$ that is *not* a support point of S. Now take $d \in Y \smallsetminus \big(\operatorname{cone}(y_0 - S)\big)$ and let g be as in (81). Then $g(y_0) = 0$. In view of (82), this implies that $\partial g(y_0) = \varnothing$, as required.

(iv) Here is a concrete scenario of (iii). Let $\mathbb{N} = \{1, 2, \ldots\}$ and $Y = \ell_2(\mathbb{N})$. Set

$$S = \big\{(\eta_n)_{n\in\mathbb{N}} \in \ell_2(\mathbb{N}) \mid (\forall n \in \mathbb{N})\ \ |\eta_n| \leq 1/4^n\big\}, \ \ y_0 = (0)_{n\in\mathbb{N}}, \quad \text{and } d = (-1/2^n)_{n\in\mathbb{N}}, \tag{83}$$

and verify directly that the properties required in (iii) are satisfied. Fix $y = (\eta_n)_{n\in\mathbb{N}} \in Y$ and $\rho \in \mathbb{R}$. Then $y + \rho d \in S \Leftrightarrow (\forall n \in \mathbb{N})\ |\eta_n - \rho/2^n| \leq 1/4^n \Leftrightarrow (\forall n \in \mathbb{N})\ \rho \in [2^n\eta_n - 1/2^n, 2^n\eta_n + 1/2^n]$. Thus, letting g as in (81), we see that

$$g\colon Y \to \left]-\infty,+\infty\right] : y = (\eta_n)_{n\in\mathbb{N}} \mapsto \sup_{n\in\mathbb{N}} \big(2^n\eta_n - 1/2^n\big), \tag{84}$$

and that $\partial g(y_0) = \varnothing$ by (iii). Hence, the corresponding perspective function (see (79)) is

$$h\colon Y \times \mathbb{R} \to \left]-\infty, +\infty\right] : \big((\eta_n)_{n\in\mathbb{N}}, \rho\big) \mapsto \begin{cases} \sup_{n\in\mathbb{N}} \left(2^n \eta_n - \rho/2^n\right), & \text{if } \rho > 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{85}$$

In this particular setting, a little more care shows that $f = h^{**}$ is given explicitly by

$$f\colon Y \times \mathbb{R} \to \left]-\infty, +\infty\right] : \big((\eta_n)_{n\in\mathbb{N}}, \rho\big) \mapsto \begin{cases} \sup_{n\in\mathbb{N}} \left(2^n \eta_n - \rho/2^n\right), & \text{if } \rho \geq 0; \\ +\infty, & \text{otherwise.} \end{cases} \tag{86}$$

(v) Let us point out that (up to a minus sign) the function f of Example 5.4 is the perspective function of $y \mapsto -2\sqrt{y}$.

(vi) Let C be a weak* closed and convex subset of X^*, and let x_0 be a point in X such that the supremum of x_0 as a functional on C is finite but not attained. Set $f = \iota_C^*|_X = \sup\langle\cdot, C\rangle$. Then $C = \partial f(0)$ (see [48, Theorem 2.4.14(vi)] or the proof of Theorem 5.10 below). According to Fact 5.2(ii)&(iii), the subdifferential of the proper, lower semicontinuous, and sublinear function f is empty at $x_0 \in \operatorname{dom} f$. We note that it is precisely such a set (necessarily unbounded) which is presented in Example 5.4 and in Remark 5.5(ii)(a).

We aim to complement Theorem 5.3 with a result for indicator functions. The following fact goes back to Rockafellar [38, Proposition 1].

Fact 5.6 *Let $f\colon X \to \left]-\infty, +\infty\right]$ be convex, lower semicontinuous, and proper, and let $(x^*, x^{**}) \in X^* \times X^{**}$. Then $(x^*, x^{**}) \in \operatorname{gra} \partial f^*$ if and only if there exists a bounded net $(x_\alpha, x_\alpha^*) \in \operatorname{gra} \partial f$ such that $x_\alpha \overset{*}{\rightharpoonup} x^{**}$ and $x_\alpha^* \to x^*$.*

Theorem 5.7 *Let $f\colon X \to \left]-\infty, +\infty\right]$ be convex, lower semicontinuous, and proper. Then $F_{(\partial f)^{-1},2} = F_{\partial f^*,2}$. Moreover, if $F \in \mathcal{F}_{\partial f}$, then $F^* \in \mathcal{F}_{\partial f^*}$.*

Proof. Since $\operatorname{gra}(\partial f)^{-1} \subset \operatorname{gra} \partial f^*$, it is clear that $F_{(\partial f)^{-1},2} \leq F_{\partial f^*,2}$. Take $(x^*, x^{**}) \in X^* \times X^{**}$ and $(y^*, y^{**}) \in \operatorname{gra} \partial f^*$. By Fact 5.6, there exists a bounded net $(y_\alpha, y_\alpha^*) \in \operatorname{gra} \partial f$ such that $y_\alpha \overset{*}{\rightharpoonup} y^{**}$ and $y_\alpha^* \to y^*$. It follows that

$$F_{(\partial f)^{-1},2}(x^*, x^{**}) \geq \langle y_\alpha^*, x^{**}\rangle + \langle x^*, y_\alpha\rangle - \langle y_\alpha, y_\alpha^*\rangle \to \langle y^*, x^{**}\rangle + \langle x^*, y^{**}\rangle - \langle y^*, y^{**}\rangle. \tag{87}$$

Supremizing over $(y^*, y^{**}) \in \operatorname{gra} \partial f^*$ we get $F_{(\partial f)^{-1},2}(x^*, x^{**}) \geq F_{\partial f^*,2}(x^*, x^{**})$. Altogether, $F_{(\partial f)^{-1},2} = F_{\partial f^*,2}$. Take $F \in \mathcal{F}_{\partial f}$. Then

$$\begin{aligned} F^*(x^*, x^{**}) &= \sup_{(y,y^*)\in X\times X^*} \big(\langle y, x^*\rangle + \langle y^*, x^{**}\rangle - F(y, y^*)\big) \\ &\geq \sup_{(y,y^*)\in \operatorname{gra} \partial f} \big(\langle y, x^*\rangle + \langle y^*, x^{**}\rangle - \langle y, y^*\rangle\big) \\ &= F_{(\partial f)^{-1},2}(x^*, x^{**}) \\ &= F_{\partial f^*,2}(x^*, x^{**}) \\ &\geq \langle x^*, x^{**}\rangle. \end{aligned} \tag{88}$$

Since $F \geq F_{\partial f,2}$, we deduce $F^* \leq F^*_{\partial f,2}$ and hence

$$F^*|_{X^*\times X} \leq F^*_{\partial f,2}|_{X^*\times X}. \tag{89}$$

Assume that $(z,z^*) \in \operatorname{gra}\partial f$. Then $F^*_{\partial f,2}(z^*,z) = \sup_{(x,x^*)\in X\times X^*}\big(\langle x,z^*\rangle + \langle z,x^*\rangle - F_{\partial f,2}(x,x^*)\big) \leq \langle z,z^*\rangle$. In view of (88), we get

$$(\forall (z,z^*) \in \operatorname{gra}\partial f) \quad F^*_{\partial f,2}(z^*,z) = \langle z^*,z\rangle. \tag{90}$$

As before, take $(y^*,y^{**}) \in \operatorname{gra}\partial f^*$ and obtain a bounded net $(y_\alpha,y^*_\alpha) \in \operatorname{gra}\partial f$ such that $y_\alpha \overset{*}{\rightharpoonup} y^{**}$ and $y^*_\alpha \to y^*$. Using the weak* lower semicontinuity of F^*, (89), and (90), we deduce that

$$F^*(y^*,y^{**}) \leq \underline{\lim}\, F^*(y^*_\alpha,y_\alpha) \leq \underline{\lim}\, F^*_{\partial f,2}(y^*_\alpha,y_\alpha) = \underline{\lim}\langle y^*_\alpha,y_\alpha\rangle = \langle y^*,y^{**}\rangle. \tag{91}$$

Altogether, (88) and (91) imply that $F^* \in \mathcal{F}_{\partial f^*}$. ■

Let $A\colon X \to 2^{X^*}$ be maximal monotone and define $\widehat{A}\colon X^* \to 2^{X^{**}}$ via

$$\operatorname{gra}\widehat{A} = \big\{(y^*,y^{**}) \in X^* \times X^{**} \mid \inf_{(x,x^*)\in\operatorname{gra}A}\langle y^*-x^*,y^{**}-x\rangle \geq 0\big\}. \tag{92}$$

Assume that A is such that for every $(y^*,y^{**}) \in \operatorname{gra}\widehat{A}$, there exists a bounded net $(y_\alpha,y^*_\alpha) \in \operatorname{gra}A$ such that $y_\alpha \overset{*}{\rightharpoonup} y^{**}$ and $y^*_\alpha \to y^*$. This condition was originally proposed by Gossez [23] and it is now known in the literature as *dense type* or *type (D)*. Subdifferential operators are of type (D) [23] as are certain linear operators [4]; see [43] for further information. A second inspection of the proof of Theorem 5.7 reveals that the following, more general result is true.

Theorem 5.8 *Let $A\colon X \to 2^{X^*}$ be maximal monotone of type (D). Then $F_{A^{-1},2} = F_{\widehat{A},2}$. Moreover, if $F \in \mathcal{F}_A$, then $F^* \in \mathcal{F}_{\widehat{A}}$.*

The next result complements Theorem 5.3 and significantly sharpens Example 4.2.

Corollary 5.9 *Let $C \subset X$ be nonempty, convex, and closed. Then $\mathcal{F}_{\partial\iota_C} = \{\iota_C \oplus \iota_C^*\}$.*

Proof. It is well known that ι_C^* is sublinear, lower semicontinuous, and proper. Take $F \in \mathcal{F}_{\partial\iota_C}$. Theorem 5.7 and Theorem 5.3 yield $F^* \in \mathcal{F}_{\partial\iota_C^*} = \{\iota_C^* \oplus \iota_C^{**}\}$. Hence $F^* = \iota_C^* \oplus \iota_C^{**}$, which implies that $F^{**} = \iota_C^{**} \oplus \iota_C^{***}$ and further that $F = \iota_C \oplus \iota_C^*$. ■

We conclude this section with a new proof of a density theorem originally due to Phelps [35]. See also [8].

Theorem 5.10 *Let $C \subset X^*$ be nonempty, weak* closed, and convex. Then the set of all elements of X which attain their supremum on C is dense in the cone of elements of X which are bounded on C.*

Proof. (See also [3, Theorem 1.4].) Let $f = \iota_C^*|_X = \sup\langle \cdot, C\rangle$. We argue as in Example 5.4. From Fact 5.2(ii)&(iii), we know that $\operatorname{dom}\partial f$ consists of all the points in X (viewed as functionals belonging to X^{**}) that attain their supremum on the nonempty set $\partial f(0)$. We claim that

$$C = \partial f(0). \tag{93}$$

Indeed, the definition of f and the fact that $f(0) = 0$ imply that $C \subset \partial f(0)$. Suppose that $x^* \in X^* \smallsetminus C$. Then there exists an element $x \in X$ which strictly separates x^* and C, i.e., $\langle x^*, x\rangle > \sup\langle C, x\rangle = f(x) - f(0)$, which implies that $x^* \notin \partial f(0)$. This verifies (93). The result now follows since Fact 3.3 implies that $\operatorname{dom}\partial f$ is a dense subset of $\operatorname{dom} f$. ■

6 Resolvents of subdifferentials

From now on, J denotes the normalized *duality mapping* of the underlying Banach space, i.e.,

$$J = \partial\big(\tfrac{1}{2}\|\cdot\|^2\big). \tag{94}$$

The following result is due to Rockafellar [39, Corollary on page 78]; recently, Simons and Zălinescu [44] and Borwein [9] gave new, simpler and more analytic proofs of it based on the Fitzpatrick function.

Fact 6.1 *Suppose that X is reflexive and let $A\colon X \to 2^{X^*}$ be monotone. Then the following statements are true.*

(i) *If A is maximal monotone, then $J + A$ is surjective.*

(ii) *If both J and J^{-1} are single-valued (i.e., both X and X^* are "smooth") and $J+A$ is surjective, then A is maximal monotone.*

As pointed out by Fitzpatrick and by Bauschke, the assumption that both J and J^{-1} be single-valued in Fact 6.1(ii) is critical; see [43, page 39]. When X is a real Hilbert space, then $J = \operatorname{Id}$ and Fact 6.1 reduces to *Minty's theorem* [32] which states that a monotone operator A is maximal monotone if and only if $\operatorname{Id} + A$ is surjective, i.e., if and only if the (always single-valued) *resolvent* $(\operatorname{Id} + A)^{-1}$ has full domain.

The motivation for our discussion in this section is given in more detail in [6]. That paper is concerned with the convergence of iterations of certain mappings on a real Hilbert space. One relevant class of mappings is the class of projections and another is that of resolvents of subdifferentials. Algorithms using iterations of projections onto two (and more) closed convex subsets of a Hilbert space were proposed in order to find a point in the intersection of these sets. This was first done by von Neumann [46, 47] when he proved that the sequence $(P_A P_B)^n$ of products of projection operators converges pointwise to $P_{A\cap B}$ when A and B are closed subspaces. See [27] for a recent elementary geometric proof of von Neumann's result. Other iterations were considered as well. One of these algorithms uses the iterations of the midpoint average map of two nearest point projections, that is, $\frac{1}{2}P_A + \frac{1}{2}P_B$. The weak convergence of this iteration was established by Auslender [2]. Two decades ago Reich [37] raised the question if it is possible for this iteration to fail to converge in norm. Using an ingenious construction by Hundal [26] (see also [30]), the authors of [6] gave an affirmative answer to this question.

In 1976 Rockafellar [41] proposed and analyzed an algorithm based on iterations of resolvents of a subdifferential for minimizing the underlying function. Brézis and Lions [15] established weak convergence of these iterations under fairly general conditions, and they asked the question whether these iterations may fail to converge in norm. A counterexample was provided by Güler [24] (see also [5] for a recent variant). If one shows that the set of resolvents of subdifferentials is convex, then one can use the counterexample based on the iterations of averaged projections to obtain a counterexample for norm convergence of Rockafellar's algorithm [6, Corollary 7.1]. Indeed, the first person to prove this convexity result was Moreau [33]. Another proof of this fact using Fenchel conjugate calculus is presented in [6, Theorem 6.1]. In this section we prove that the set of resolvents of subdifferentials is convex using cyclic monotonicity. Recall [22, page 41] that an operator $T\colon C \to X$, where $C \subset X$, is *firmly nonexpansive* if

$$(\forall x \in C)(\forall y \in C)(\forall \rho \geq 0) \quad \|Tx - Ty\| \leq \|\rho(x-y) + (1-\rho)(Tx - Ty)\|. \tag{95}$$

The following two results are well known (see, e.g., [22]).

Fact 6.2 *Suppose that X is a real Hilbert space, let $C \subset X$ and let $T\colon C \to X$. Then the following are equivalent.*

(i) *T is firmly nonexpansive.*

(ii) $(\forall x \in C)(\forall y \in C)$ $\|Tx - Ty\|^2 \leq \langle x - y, Tx - Ty\rangle$.

(iii) $T = \frac{1}{2}\operatorname{Id} + \frac{1}{2}N$, *where N is nonexpansive (i.e., 1-Lipschitz continuous).*

(iv) $T = (\operatorname{Id} + A)^{-1}$ *is the resolvent of some monotone operator $A\colon X \to 2^X$.*

Corollary 6.3 *Suppose that X is a real Hilbert space and let $A\colon X \to 2^X$. Then A is maximal monotone if and only if its resolvent $(\operatorname{Id} + A)^{-1}$ is firmly nonexpansive with full domain.*

Proof. Combine Minty's theorem (see the paragraph following Fact 6.1) and Fact 6.2. ■

Our aim here is to prove the convexity of the set of resolvents of subdifferentials (also known as *proximal mappings*). This set is a convex subset of the set of firmly nonexpansive mappings, which is also easily seen to be convex by utilizing Fact 6.2(iii). As the following example illustrates, the set of firmly nonexpansive mappings need not be convex outside Hilbert spaces.

Example 6.4 Suppose that $X = \mathbb{R}^3$, where $\|x\| = \max_{i \in \{1,2,3\}} |\xi_i| = \|x\|_\infty$ for $x = (\xi_1, \xi_2, \xi_3) \in X$. Let $C = \{x, y\}$, where $x = (1, 0, 1)$ and $y = (0, 0, 0)$, and define $T_0\colon C \to X$ and

$T_1\colon C \to X$ via their graphs as $\operatorname{gra} T_0 = \big\{\big((1,0,1),(0,1,1)\big),\big((0,0,0),(0,0,0)\big)\big\}$ and $\operatorname{gra} T_1 = \big\{\big((1,0,1),(1,1,0)\big),\big((0,0,0),(0,0,0)\big)\big\}$. Then T_0 and T_1 are both firmly nonexpansive. Furthermore, for every $\lambda \in \,]0,1[$, the operator

$$T_\lambda = (1-\lambda)T_0 + \lambda T_1 \tag{96}$$

is *not* firmly nonexpansive.

Proof. For every $\rho \geq 0$, we have $\|T_0x - T_0y\| = 1 \leq \max\{1,\rho\} = \|(\rho,0,\rho) + (0,1-\rho,1-\rho)\| = \|\rho(x-y) + (1-\rho)(T_0x - T_0y)\|$ and similarly $\|T_1x - T_1y\| \leq \|\rho(x-y) + (1-\rho)(T_1x - T_1y)\|$, and this implies that T_0 and T_1 are firmly nonexpansive. Take $\lambda \in \,]0,1[$. Then

$$T_\lambda x = (1-\lambda)T_0x + \lambda T_1 x = (\lambda, 1, 1-\lambda) \quad \text{and} \quad T_\lambda y = (1-\lambda)T_0 y + \lambda T_1 y = (0,0,0). \tag{97}$$

Hence

$$\begin{aligned}\|T_\lambda x - T_\lambda y\| &= 1\\ &> \tfrac{1}{2}\|(1+\lambda, 1, 2-\lambda)\|\\ &= \|\tfrac{1}{2}(1,0,1) + (1-\tfrac{1}{2})(\lambda,1,1-\lambda)\|\\ &= \|\tfrac{1}{2}(x-y) + (1-\tfrac{1}{2})(T_\lambda x - T_\lambda y)\|,\end{aligned} \tag{98}$$

which implies that T_λ is not firmly nonexpansive. ■

Remark 6.5 Let $A\colon X \to 2^X$. Then A is *accretive* if $(\forall \lambda > 0)$ $(\operatorname{Id} + \lambda A)^{-1}$ is single-valued and nonexpansive on $\operatorname{dom}(\operatorname{Id} + \lambda A)^{-1} = \operatorname{ran}(\operatorname{Id} + \lambda A)$, and A is *maximal accretive* if A is accretive and no proper extension of A is accretive. Zorn's Lemma guarantees that every accretive operator admits a maximal accretive extension. Now let X be as in Example 6.4. Then a result of Crandall and Liggett [19, Theorem 2.5] implies that $A\colon X \to 2^X$ is maximal accretive if and only if it is accretive and $(\forall \lambda > 0)$ $\operatorname{ran}(\operatorname{Id} + \lambda A) = X$. Reich's [36, Lemma 7.1] shows that Crandall and Liggett's result is equivalent to the following: *Suppose that $D \subset X$ and that $T\colon D \to X$ is firmly nonexpansive. Then T admits a firmly nonexpansive extension $\widetilde{T}\colon X \to X$.* So let T_0 and T_1 be as in Example 6.4 and denote their firmly nonexpansive extensions to X by $\widetilde{T_0}$ and $\widetilde{T_1}$, respectively. Then Example 6.4 implies that for every $\lambda \in \,]0,1[$, the mapping $(1-\lambda)\widetilde{T_0} + \lambda \widetilde{T_1}$ is *not* firmly nonexpansive. In particular, the firmly nonexpansive mappings with full domain do not form a convex set.

Before proving the main result of this section, we need to translate the cyclic monotonicity characterization of a subdifferential to a characterization of its resolvent.

Theorem 6.6 *Suppose that X is a real Hilbert space and let $T\colon X \to X$. Then T is the resolvent of the maximal cyclically monotone operator $A\colon X \to 2^X$ $\Leftrightarrow$ T has full domain, T is firmly nonexpansive, and for every set of points $\{x_1, \ldots, x_n\}$, where $n \in \{2,3,\ldots\}$ and $x_{n+1} = x_1$, one has*

$$\sum_{i=1}^{n} \langle x_i - Tx_i, Tx_i - Tx_{i+1}\rangle \geq 0. \tag{99}$$

Proof. "$\Leftarrow$": In view of Corollary 6.3, it suffices to show that $A = T^{-1} - \operatorname{Id}$ is cyclically monotone. Fix $n \in \{2, 3, \ldots\}$, take n points $(y_1, y_1^*), \ldots, (y_n, y_n^*)$ in $\operatorname{gra} A$, and set $y_1 = y_{n+1}$. For every $i \in \{1, \ldots, n\}$, set $x_i = y_i + y_i^*$; hence $x_i \in (\operatorname{Id} + A)y_i$, $y_i = Tx_i$, $Tx_{n+1} = y_{n+1} = y_1 = Tx_1$, and $x_i - Tx_i = x_i - y_i = y_i^*$. Plugging all this in (99), we get

$$\sum_{i=1}^{n} \langle y_i - y_{i+1}, y_i^* \rangle \geq 0. \tag{100}$$

Thus A is cyclically monotone. "$\Rightarrow$": In view of Fact 3.2 and Corollary 6.3, we only need to verify (99). Take $n \in \{2, 3, \ldots\}$ and $x_1, \ldots, x_n$ in X, and set $x_{n+1} = x_1$. For every $i \in \{1, \ldots, n+1\}$, set $y_i = Tx_i$ and $y_i^* = x_i - Tx_i$; thus $Tx_i = (\operatorname{Id} + A)^{-1}x_i \Rightarrow x_i \in (\operatorname{Id} + A)Tx_i \Rightarrow Tx_i \in \operatorname{dom} A$ and $y_i^* = x_i - Tx_i \in ATx_i = Ay_i$. Therefore,

$$\sum_{i=1}^{n} \langle x_i - Tx_i, Tx_i - Tx_{i+1} \rangle = \sum_{i=1}^{n} \langle y_i - y_{i+1}, y_i^* \rangle \geq 0, \tag{101}$$

and this completes the proof. ■

Theorem 6.7 *Suppose that X is a real Hilbert space and let f and g be two functions from X to $]-\infty, +\infty]$ that are convex, lower semicontinuous, and proper. Further, let $\alpha \in]0, 1[$. Then there exists a proper, lower semicontinuous, and convex function $h \colon X \to]-\infty, +\infty]$ such that*

$$(\operatorname{Id} + \partial h)^{-1} = \alpha(\operatorname{Id} + \partial f)^{-1} + (1 - \alpha)(\operatorname{Id} + \partial g)^{-1}. \tag{102}$$

Proof. Set $S = (\operatorname{Id} + \partial f)^{-1}$ and $T = (\operatorname{Id} + \partial g)^{-1}$. Then S and T are firmly nonexpansive with full domain as they are the resolvents of the maximal monotone operators ∂f and ∂g, respectively. For convenience, set $\beta = 1 - \alpha$. Then $\alpha S + \beta T$ is firmly nonexpansive with full domain. In view of Fact 3.2 and Theorem 6.6, given $n \in \{2, 3, \ldots\}$ and $\{x_1, \ldots, x_n\} \subset X$ with $x_{n+1} = x_1$, it suffices to prove that

$$0 \leq \sum_{i=1}^{n} \langle x_i - (\alpha Sx_i + \beta Tx_i), (\alpha Sx_i + \beta Tx_i) - (\alpha Sx_{i+1} + \beta Tx_{i+1}) \rangle. \tag{103}$$

Now Theorem 6.6 implies that

$$0 \leq \alpha \sum_{i=1}^{n} \langle x_i - Sx_i, Sx_i - Sx_{i+1} \rangle \tag{104}$$

and that

$$0 \leq \beta \sum_{i=1}^{n} \langle x_i - Tx_i, Tx_i - Tx_{i+1} \rangle. \tag{105}$$

Furthermore, either by direct computation or by simply noticing that $\operatorname{Id} = \nabla\big(\tfrac{1}{2}\|\cdot\|^2\big)$ is cyclically monotone, we see that given any n points $y_1, \ldots, y_n$ satisfy $0 \leq \sum_{i=1}^{n} \langle y_i, y_i - y_{i+1} \rangle$, where $y_{n+1} = y_1$. In particular,

$$0 \leq \alpha\beta \sum_{i=1}^{n} \langle Tx_i - Sx_i, (Tx_i - Sx_i) - (Tx_{i+1} - Sx_{i+1}) \rangle. \tag{106}$$

Now for every $i \in \{1, \dots, n\}$,

$$\begin{aligned}
&\langle x_i - (\alpha S x_i + \beta T x_i), (\alpha S x_i + \beta T x_i) - (\alpha S x_{i+1} + \beta T x_{i+1}) \rangle \\
&= \alpha \langle x_i - S x_i, S x_i - S x_{i+1} \rangle + \beta \langle x_i - T x_i, T x_i - T x_{i+1} \rangle \\
&\quad + \alpha\beta \langle T x_i - S x_i, (T x_i - T x_{i+1}) - (S x_i - S x_{i+1}) \rangle .
\end{aligned} \tag{107}$$

Therefore, adding (104), (105), and (106) we arrive precisely at (103). ■

It is also possible to give — up to an additive constant — a formula for the function h of Theorem 6.7. To do this, we can use the formula given in Rockafellar's characterization of subdifferentials as maximal cyclically monotone mappings (Fact 3.2).

Acknowledgment

Sedi Bartz was partially supported by the Graduate School of the Technion. Heinz Bauschke was partially supported by the Natural Sciences and Engineering Research Council of Canada. Jonathan Borwein was partially supported by the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chair Program. Simeon Reich was partially supported by the Fund for the Promotion of Research at the Technion and by the Technion VPR Fund. Xianfu Wang was partially supported by the Natural Sciences and Engineering Research Council of Canada.

References

[1] E. Asplund, "A monotone convergence theorem for sequences of nonlinear mappings," *Nonlinear Functional Analysis*, Proceedings of Symposia in Pure Mathematics, American Mathematical Society, vol. XVIII Part 1, Chicago, pp. 1–9, 1970.

[2] A. Auslender, *Méthodes Numériques pour la Résolution des Problèmes d'Optimisation avec Constraintes*, Thèse, Faculté des Sciences, Grenoble, 1969.

[3] S. Bartz, *On the Subdifferential,* M. Sc. thesis, The Technion — Israel Institute of Technology, Haifa, October 2005.

[4] H. H. Bauschke and J. M. Borwein, "Maximal monotonicity of dense type, local maximal monotonicity, and monotonicity of the conjugate are all the same for continuous linear operators," *Pacific Journal of Mathematics,* vol. 189, pp. 1–20, 1999.

[5] H. H. Bauschke, J. V. Burke, F. R. Deutsch, H. S. Hundal, and J. D. Vanderwerff, "A new proximal point iteration that converges weakly but not in norm," *Proceedings of the American Mathematical Society,* vol. 133, pp. 1829–1835, 2005.

[6] H. H. Bauschke, E. Matoušková, and S. Reich, "Projection and proximal point methods: convergence results and counterexamples," *Nonlinear Analysis*, vol. 56, pp. 715–738, 2004.

[7] H. H. Bauschke, D. A. McLaren, and H. S. Sendov, "Fitzpatrick functions: inequalities, examples, and remarks on a problem by S. Fitzpatrick," to appear in *Journal of Convex Analysis.*

[8] J. M. Borwein, "A note on ε-subgradients and maximal monotonicity," *Pacific Journal of Mathematics*, vol. 103, pp. 307–314, 1982.

[9] J. M. Borwein, "Maximal monotonicity via convex analysis," to appear in *Journal of Convex Analysis.*

[10] J. M. Borwein and A. S. Lewis, *Convex Analysis and Nonlinear Optimization,* Springer-Verlag, 2000.

[11] J. M. Borwein and D. W. Tingley, "On supportless convex sets," *Proceedings of the American Mathematical Society,* vol. 94, pp. 471–476, 1985.

[12] J. M. Borwein and Q. J. Zhu, *Techniques of Variational Analysis,* Springer-Verlag, 2005.

[13] S. Boyd and L. Vandenberghe, *Convex Optimization,* Cambridge University Press, 2004.

[14] H. Brézis, *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert,* North-Holland, 1973.

[15] H. Brézis and P.-L. Lions, "Produits infinis de résolvantes," *Israel Journal of Mathematics,* vol. 29, pp. 329–345, 1978.

[16] R. S. Burachik and S. Fitzpatrick, "On the Fitzpatrick family associated to some subdifferentials," *Journal of Nonlinear and Convex Analysis*, vol. 6, pp. 165–171, 2005.

[17] R. S. Burachik and B. F. Svaiter, "Maximal monotone operators, convex functions and a special family of enlargements," *Set-Valued Analysis*, vol. 10, pp. 297–316, 2002.

[18] R. S. Burachik and B. F. Svaiter, "Maximal monotonicity, conjugation and the duality product," *Proceedings of the American Mathematical Society*, vol. 131, pp. 2379–2383, 2003.

[19] M. G. Crandall and T. M. Liggett, "A theorem and a counterexample in the theory of semigroups of nonlinear transformations," *Transactions of the American Mathematical Society,* vol. 160, pp. 263–278, 1971.

[20] P. J. Davis, *Circulant Matrices,* 2nd ed., Chelsea, 1994.

[21] S. Fitzpatrick, "Representing monotone operators by convex functions," *Workshop/Miniconference on Functional Analysis and Optimization (Canberra 1988)*, Proceedings of the Centre for Mathematical Analysis, Australian National University vol. 20, Canberra, Australia, pp. 59–65, 1988.

[22] K. Goebel and S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings,* Marcel Dekker, New York, 1984.

[23] J.-P. Gossez, "Opérateurs monotones non linéaires dans les espaces de Banach non réflexifs," *Journal of Mathematical Analysis and Applications,* vol. 34, pp. 371–395, 1971.

[24] O. Güler, "On the convergence of the proximal point algorithm for convex minimization," *SIAM Journal on Control and Optimization,* vol. 29, pp. 403–419, 1991.

[25] R. B. Holmes, *Geometric Functional Analysis and its Applications,* Springer-Verlag, 1975.

[26] H. S. Hundal, "An alternating projection that does not converge in norm," *Nonlinear Analysis,* vol. 57, pp. 35–61, 2004.

[27] E. Kopecká and S. Reich, "A note on the von Neumann alternating projections algorithm," *Journal of Nonlinear and Convex Analysis,* vol. 5, pp. 379–386, 2004.

[28] J.-E. Martínez-Legaz and B. F. Svaiter, "Monotone operators representable by l.s.c. convex functions," *Set-Valued Analysis,* vol. 13, pp. 21–46, 2005.

[29] J.-E. Martínez-Legaz and M. Théra, "A convex representation of maximal monotone operators," *Journal of Nonlinear and Convex Analysis,* vol. 2, pp. 243–247, 2001.

[30] E. Matoušková and S. Reich, "The Hundal example revisited," *Journal of Nonlinear and Convex Analysis*, vol. 4, pp. 411–427, 2003.

[31] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM Philadelphia, 2000.

[32] G. J. Minty, "Monotone (nonlinear) operators in Hilbert space," *Duke Mathematical Journal,* vol. 29, pp. 341–346, 1962.

[33] J.-J. Moreau, "Proximité et dualité dans un espace hilbertien," *Bulletin de la Société Mathématique de France*, vol. 93, pp. 273–299, 1965.

[34] J. P. Penot, "The relevance of convex analysis for the study of monotonicity," *Nonlinear Analysis,* vol. 58, pp. 855–871, 2004.

[35] R. R. Phelps, "Weak* support points of convex sets in E^*," *Israel Journal of Mathematics,* vol. 2, pp. 177–182, 1964.

[36] S. Reich, "Extension problems for accretive sets in Banach spaces," *Journal of Functional Analysis*, vol. 26, pp. 378–395, 1977.

[37] S. Reich, "A limit theorem for projections," *Linear and Multilinear Algebra*, vol. 13, pp. 281-290, 1983.

[38] R. T. Rockafellar, "On the maximal monotonicity of subdifferential mappings," *Pacific Journal of Mathematics,* vol. 33, pp. 209–216, 1970.

[39] R. T. Rockafellar, "On the maximality of sums of nonlinear monotone operators," *Transactions of the American Mathematical Society,* vol. 149, pp. 75–88, 1970.

[40] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[41] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM Journal on Control and Optimization,* vol. 14, pp. 877–898, 1976.

[42] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, Springer-Verlag, 1998.

[43] S. Simons, *Minimax and Monotonicity*, Lecture Notes in Mathematics, vol. 1693, Springer-Verlag, 1998.

[44] S. Simons and C. Zălinescu, "A new proof for Rockafellar's characterization of maximal monotone operators," *Proceedings of the American Mathematical Society*, vol. 132, pp. 2969–2972, 2004.

[45] S. Simons and C. Zălinescu, "Fenchel duality, Fitzpatrick functions and maximal monotonicity," *Journal of Nonlinear and Convex Analysis,* vol. 6, pp. 1–22, 2005.

[46] J. von Neumann, *Functional Operators II: The Geometry of Orthogonal Spaces*, Princeton University Press, Princeton, NJ, 1950. (This is a reprint of mimeographed lecture notes first distributed in 1933.)

[47] J. von Neumann, On rings of operators. Reduction theory, *Annals of Mathematics*, vol. 50, pp. 401–485, 1949.

[48] C. Zălinescu, *Convex Analysis in General Vector Spaces,* World Scientific Publishing, 2002.