

Primal-dual symmetric intrinsic methods for finding antiderivatives of cyclically monotone operators

Heinz H. Bauschke*, Yves Lucet† and Xianfu Wang‡

Second revision August 10, 2007
(first revision August 4, 2007; original version November 23, 2006)

Abstract

A fundamental result due to Rockafellar states that every cyclically monotone operator A admits an antiderivative f in the sense that the graph of A is contained in the graph of the subdifferential operator ∂f .

Given a method \mathfrak{m} that assigns every finite cyclically monotone operator A some antiderivative \mathfrak{m}_A , we say that the method is *primal-dual symmetric* if \mathfrak{m} applied to the inverse of A produces the Fenchel conjugate of \mathfrak{m}_A . Rockafellar's antiderivatives do not possess this property. Utilizing Fitzpatrick functions and the proximal average, we present novel primal-dual symmetric intrinsic methods. The antiderivatives produced by these methods provide a solution to a problem posed by Rockafellar in 2005. The results leading to this solution are illustrated by various examples.

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

Keywords: Antiderivative, convex function, cyclically monotone operator, Fenchel conjugate, Fitzpatrick function, maximal monotone operator, n -cyclically monotone operator, proximal average, Rockafellar's antiderivative, Rockafellar function, subdifferential operator.

1 Introduction

Suppose that X is a real Banach space with continuous dual X^* , dual pairing $\langle \cdot, \cdot \rangle$, and norm $\|\cdot\|$. We start by recalling some known notions and results concerning (cyclically) monotone operators.

*Mathematics, Irving K. Barber School, University of British Columbia Okanagan, Kelowna, British Columbia V1V 1V7, Canada. E-mail: heinz.bauschke@ubc.ca.

†Computer Science, Irving K. Barber School, University of British Columbia Okanagan, Kelowna, British Columbia V1V 1V7, Canada. E-mail: yves.lucet@ubc.ca.

‡Mathematics, Irving K. Barber School, University of British Columbia Okanagan, Kelowna, British Columbia V1V 1V7, Canada. E-mail: shawn.wang@ubc.ca.

These operators play a fundamental role in modern optimization as well as convex and variational analysis; see, e.g., [11, 12, 18, 20, 21, 22, 24] for further information and notation not explicitly defined here. Let A be a *set-valued operator* from X to X^* , i.e., $(\forall x \in X) Ax \subseteq X^*$; thus, A is a mapping from X to the power set of X^* . We use the notation $A: X \rightrightarrows X^*$ and remark that A can be identified with its *graph* $\text{gra } A := \{(x, x^*) \in X \times X^* \mid x^* \in Ax\}$. The *domain* of A is $\text{dom } A := \{x \in X \mid Ax \neq \emptyset\}$ and the *range* of A is $\text{ran } A := A(X) = \bigcup_{x \in X} Ax$. The *inverse* of A is the operator $A^{-1}: X^* \rightrightarrows X$, defined by $x \in A^{-1}x^* \Leftrightarrow x^* \in Ax$. Furthermore, let $n \in \{2, 3, \dots\}$. Then A is *n-cyclically monotone* [1, 2, 3, 7, 10, 23] if the implication

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

holds. Note that 2-monotonicity simplifies to

$$(\forall (x, x^*) \in \text{gra } A)(\forall (y, y^*) \in \text{gra } A) \quad \langle x - y, x^* - y^* \rangle \geq 0, \quad (2)$$

i.e., to ordinary *monotonicity*. *Cyclic monotonicity* describes the situation when A is *m-cyclically monotone* for every $m \in \{2, 3, \dots\}$. The operator A is *maximal n-cyclically monotone* if A is *n-cyclically monotone* and no proper extension (in the sense of inclusion of graphs) of A is *n-cyclically monotone*. Zorn's Lemma guarantees that every *n-cyclically monotone* operator admits a maximal *n-cyclically monotone* extension. At one end of the spectrum of maximal *n-cyclically monotone* operators are the maximal 2-monotone, i.e., the *maximal monotone* operators. At the other end are the *maximal cyclically monotone operators* which Rockafellar in a ground-breaking paper [19] (see Fact 3.4 below) revealed to be precisely the subdifferential operators of functions that are convex, lower semicontinuous, and proper.

This paper is motivated by the following question posed by R. T. Rockafellar in 2005 during open-problem sessions at conferences in Borovets (Bulgaria) and Banff (Canada).

Given a cyclically monotone operator A with finite graph, find a method that produces an antiderivative of A that preserves the natural symmetry induced by convex duality.

One motivation for the above question that we feel will become particularly relevant for applications as numerical convex analysis matures is the efficient storage and representation of convex functions. This is a surprisingly difficult problem. The perhaps most natural approach of storing grid points (x_i, y_i) causes significant problems because Lagrangian interpolation can fail to recover a convex function [13]. We now describe three other possible approaches. Firstly, one could solve for subgradients x_i^* at each point x_i , or store such data in the first place, and then recover the function via $f(x) = \max_i (\langle x - x_i, x_i^* \rangle + y_i)$. The resulting function is piecewise linear with full domain; thus, its conjugate has bounded domain. Secondly, one could restrict the model of the function to the convex hull of the points x_i and set the function equal to $+\infty$ outside. And thirdly, one could store the points and subgradients (x_i, x_i^*) along with a scalar y_0 and then recover a function

f that satisfies $x_i^* \in \partial f(x_i)$ and $f(x_0) = y_0$. However, the existing representations in the literature [9, 19] are based on piecewise linear functions; so, in the finite graph case, one has to unavoidably privilege either the primal or the dual space in the very model used to recover the function.

In this paper, we provide constructive answers to Rockafellar’s question. In fact, we shall exhibit methods for constructing antiderivatives that we call *primal-dual symmetric*. These methods have the property that when they are applied to A^{-1} instead of A , then the Fenchel conjugate of the antiderivative of A is obtained. The mere existence of such methods struck us initially as quite remarkable since antiderivatives are at best unique up to additive constants. These methods also allow for the design of models of convex functions that inherit the symmetry induced by convex duality in the given discrete data. Our constructions are based on Rockafellar’s classical construction of an antiderivative as well as on recent work on Fitzpatrick functions and the proximal average operator, which has close connection to fundamental objects of optimization such as Moreau envelopes and proximal mappings [4, 6]. Another pleasant consequence of primal-dual symmetric methods is their “slope 1” property — we believe that this will aid in efforts to represent convex functions in a numerically stable way (the “slope 1” property guarantees that the derivatives outside the domain of interest have slopes that are neither too small nor too large in magnitude). This is an area of active research that lies beyond the scope of this paper; see [15] for a one-dimensional framework that is capable to express such antiderivatives and that serves as a starting point for further research.

The remainder of this paper can be summarized as follows. In Section 2, we introduce the common ancestor and Fitzpatrick functions [2]. These functions have turned out to be immensely useful in the study of — and they are intimately tied to — n -cyclic monotonicity. We provide a recursion formula for the Fitzpatrick functions (Proposition 2.13) and show that they stabilize when applied to cyclically monotone operators with finite graph (Theorem 2.16). In Section 3, we revisit Rockafellar’s classical antiderivative result (Fact 3.4) in the context of Fitzpatrick functions. In fact, his antiderivative satisfies a certain minimality property (Theorem 3.5), it is related to the common ancestor function (Corollary 3.11), and a closed form can be found for some finite-graph operators on \mathbb{R} (Theorem 3.14) — parts of these results, of which we were unaware during the preparation of the originally submitted version of this paper, were previously obtained by Lambert et al. in their interesting work [14] in which they focus on finding upper and lower bounds for antiderivatives using a linear programming formulation. The supremum of all Rockafellar antiderivatives is expressible in terms of a Fitzpatrick function (Theorem 3.15 and Corollary 3.16). Section 4 introduces the notion of a *primal-dual symmetric* method for antiderivatives (Definition 4.6). Such methods provide antiderivatives that only depend on the graph — which makes them *intrinsic* — and that return the Fenchel conjugate of the antiderivative when applied to the inverse operator. Neither Rockafellar’s classical antiderivatives nor simple symmetrizations of them have this property (Proposition 4.7). Based on recent work on the proximal average operator, we proceed to present our main result which provides a general construction of primal-dual symmetric methods (Theorem 4.13). Concrete instances are proximal-average based symmetrizations of the maximum and of the average of Rockafellar’s antiderivatives (Example 4.19 and Example 4.20). We then present a result (Corollary 4.23) that leads to a resolution of Rockafellar’s problem (Corollary 4.26 and Remark 4.27). We conclude the paper with a numerical example (Example 4.28).

Our notation is standard. The subdifferential operator of a convex function f is denoted by ∂f , its Fenchel conjugate by f^* and its domain by $\text{dom } f$. For a set S , we use $\text{conv } S$, $\overline{\text{conv}} S$, $\text{int } S$, and \overline{S} to denote its convex hull, its closed convex hull, its interior, and its closure, respectively. For a nonempty convex subset C of X and a point $x \in C$, the tangent and the normal cone of C at x are denoted by $T_C(x)$ and by $N_C(x)$, respectively. Finally, the set of all functions from X to $]-\infty, +\infty]$ that are convex, lower semicontinuous, and proper is denoted by $\Gamma(X)$ or simply by Γ .

2 The common ancestor and Fitzpatrick functions

Definition 2.1 [2, Definition 2.1] *Let $A: X \rightrightarrows X^*$ and let $(a_1, a_1^*) \in \text{gra } A$. The common ancestor functions are defined by*

$$C_{A,2,(a_1,a_1^*)}: X \times X^* \rightarrow]-\infty, +\infty] : (x, x^*) \mapsto \langle x, a_1^* \rangle + \langle a_1, x^* \rangle - \langle a_1, a_1^* \rangle, \quad (3)$$

and, for every $n \in \{3, 4, \dots\}$, by

$$C_{A,n,(a_1,a_1^*)}: X \times X^* \rightarrow]-\infty, +\infty] \\ (x, x^*) \mapsto \sup_{\substack{(a_2,a_2^*) \in \text{gra } A, \\ \vdots \\ (a_{n-1},a_{n-1}^*) \in \text{gra } A}} \left(\sum_{i=1}^{n-2} \langle a_{i+1} - a_i, a_i^* \rangle \right) + \langle x - a_{n-1}, a_{n-1}^* \rangle + \langle a_1, x^* \rangle. \quad (4)$$

We also set

$$C_{A,\infty,(a_1,a_1^*)} = \sup_{n \in \{2,3,\dots\}} C_{A,n,(a_1,a_1^*)}. \quad (5)$$

It is clear that

$$(\forall n \in \{2, 3, \dots\}) \quad C_{A,n,(a_1,a_1^*)} \text{ is convex and lower semicontinuous,} \quad (6)$$

that the sequence

$$(C_{A,n,(a_1,a_1^*)})_{n \in \{2,3,\dots\}} \text{ is increasing and pointwise convergent to } C_{A,\infty,(a_1,a_1^*)}, \quad (7)$$

and that $C_{A,\infty,(a_1,a_1^*)}$ is convex and lower semicontinuous. Moreover,

$$\text{gra } A \text{ finite} \quad \Rightarrow \quad (\forall n \in \{2, 3, \dots\}) \quad C_{A,n,(a_1,a_1^*)} \text{ is polyhedral and continuous.} \quad (8)$$

The next result shows that common ancestor functions are closely related to n -cyclic monotonicity. The proof is straightforward and thus omitted.

Proposition 2.2 *Let $A: X \rightrightarrows X^*$ and let $n \in \{2, 3, \dots\}$. Then A is n -cyclically monotone if and only if*

$$(\forall (a, a^*) \in \text{gra } A) (\forall (b, b^*) \in \text{gra } A) \quad C_{A,n,(a,a^*)}(b, b^*) \leq \langle b, b^* \rangle. \quad (9)$$

Computationally convenient is the following recursive formula.

Proposition 2.3 (recursion) *Let $A: X \rightrightarrows X^*$, let $(a_1, a_1^*) \in \text{gra } A$, let $n \in \{2, 3, \dots\}$, and let $(x, x^*) \in X \times X^*$. Then*

$$C_{A, n+1, (a_1, a_1^*)}(x, x^*) = \sup_{(a, a^*) \in \text{gra } A} C_{A, n, (a_1, a_1^*)}(a, x^*) + \langle x - a, a^* \rangle. \quad (10)$$

Proof. By definition, $C_{A, n+1, (a_1, a_1^*)}(x, x^*)$ is the supremum of the terms

$$\begin{aligned} & \left(\sum_{i=1}^{n-1} \langle a_{i+1} - a_i, a_i^* \rangle \right) + \langle x - a_n, a_n^* \rangle + \langle a_1, x^* \rangle \\ &= \left(\sum_{i=1}^{n-2} \langle a_{i+1} - a_i, a_i^* \rangle \right) + \langle a_n - a_{n-1}, a_{n-1}^* \rangle + \langle a_1, x^* \rangle + \langle x - a_n, a_n^* \rangle, \end{aligned} \quad (11)$$

where $(a_2, a_2^*), \dots, (a_n, a_n^*)$ in $\text{gra } A$. Supremizing first over $(a_2, a_2^*), \dots, (a_{n-1}, a_{n-1}^*)$, followed by supremizing over (a_n, a_n^*) , we obtain the conclusion. \blacksquare

Due to their implementability, operators with finite graphs are of particular interest. The next result demonstrates that if a sufficiently high order of cyclic monotonicity is achieved, then the common ancestor functions stabilize.

Theorem 2.4 *Let $A: X \rightrightarrows X^*$, let $(a_1, a_1^*) \in \text{gra } A$, and let $n \in \{2, 3, \dots\}$. Suppose that A is n -cyclically monotone and that $\text{gra } A$ has at most n points. Then $C_{A, \infty, (a_1, a_1^*)} = C_{A, n+1, (a_1, a_1^*)}$.*

Proof. Take $(x, x^*) \in X \times X^*$ and take $m \in \{n+2, n+3, \dots\}$. It suffices to show that

$$C_{A, m, (a_1, a_1^*)}(x, x^*) \leq C_{A, m-1, (a_1, a_1^*)}(x, x^*), \quad (12)$$

since this and (7) then imply that $C_{A, n+1, (a_1, a_1^*)} = C_{A, n+2, (a_1, a_1^*)} = \dots = C_{A, \infty, (a_1, a_1^*)}$. Take $(a_2, a_2^*), \dots, (a_{m-1}, a_{m-1}^*)$ in $\text{gra } A$. Since $\text{gra } A$ contains at most n points and since $m-1 \geq n+1$, there exist integers k and l such that $1 \leq k < l \leq m-1$ and $a_k = a_l$. Hence

$$\begin{aligned} & \sum_{i=1}^{m-2} \langle a_{i+1} - a_i, a_i^* \rangle + \langle x - a_{m-1}, a_{m-1}^* \rangle + \langle a_1, x^* \rangle \\ &= \sum_{i=1}^{k-1} \langle a_{i+1} - a_i, a_i^* \rangle + \sigma + \sum_{i=l}^{m-2} \langle a_{i+1} - a_i, a_i^* \rangle + \langle x - a_{m-1}, a_{m-1}^* \rangle + \langle a_1, x^* \rangle, \end{aligned} \quad (13)$$

where

$$\sigma = \sum_{i=k}^{l-1} \langle a_{i+1} - a_i, a_i^* \rangle. \quad (14)$$

We claim that

$$\sigma \leq 0. \quad (15)$$

Note that σ contains $l - k$ terms. If $l - k \leq n$, then the n -cyclic monotonicity of A implies that $\sigma \leq 0$. Otherwise, $l - k > n$ and we may analogously and recursively split up σ until it is a finite sum of negative terms. This verifies (15). Now (13) implies (12). \blacksquare

Example 2.5 Suppose that X is a Hilbert space. Let $e \in X$ be such that $\|e\| = 1$ and define A via $\text{gra } A := \{(-e, -e), (e, e)\}$. Then A is (2-cyclically) monotone and for every $(x, x^*) \in X \times X$, we have

$$C_{A,2,(-e,-e)}(x, x^*) = -\langle x + x^*, e \rangle - 1, \quad (16)$$

$$C_{A,2,(e,e)}(x, x^*) = \langle x + x^*, e \rangle - 1, \quad (17)$$

$$C_{A,3,(-e,-e)}(x, x^*) = \max \{ -\langle x + x^*, e \rangle - 1, \langle x - x^*, e \rangle - 3 \}, \quad (18)$$

$$C_{A,3,(e,e)}(x, x^*) = \max \{ \langle x^* - x, e \rangle - 3, \langle x + x^*, e \rangle - 1 \}. \quad (19)$$

Theorem 2.6 Let $A: X \rightrightarrows X^*$ and let $(a_1, a_1^*) \in \text{gra } A$. Suppose that A is not cyclically monotone. Then $C_{A,\infty,(a_1,a_1^*)} \equiv +\infty$.

Proof. There exist n points $(a_2, a_2^*), \dots, (a_{n+1}, a_{n+1}^*)$ in $\text{gra } A$, where $n \in \{2, 3, \dots\}$, such that

$$\sigma := \sum_{i=2}^{n+1} \langle a_{i+1} - a_i, a_i^* \rangle > 0, \quad \text{where } a_{n+2} := a_2. \quad (20)$$

Take $(x, x^*) \in X \times X^*$. Take $k \in \{2, 3, \dots\}$, define

$$a_{kn+2} := a_{(k-1)n+2} := \dots := a_2, \quad (21)$$

$$a_{kn+1} := a_{(k-1)n+1} := \dots := a_{n+1}, \quad (22)$$

$$a_{kn} := a_{(k-1)n} := \dots := a_n, \quad (23)$$

$$\vdots \quad (24)$$

$$a_{(k-1)n+3} := a_{(k-2)n+3} := \dots := a_3, \quad (25)$$

and analogously for $a_{n+2}^*, \dots, a_{kn+2}^*$. Then

$$C_{A,kn+3,(a_1,a_1^*)}(x, x^*) \geq \sum_{i=1}^{kn+1} \langle a_{i+1} - a_i, a_i^* \rangle + \langle x - a_{kn+2}, a_{kn+2}^* \rangle + \langle a_1, x^* \rangle \quad (26)$$

$$= \langle a_2 - a_1, a_1^* \rangle + k\sigma + \langle x - a_2, a_2^* \rangle + \langle a_1, x^* \rangle \quad (27)$$

$$\rightarrow +\infty \quad \text{as } k \rightarrow +\infty. \quad (28)$$

Therefore, $\lim_{k \rightarrow +\infty} C_{A,kn+3,(a_1,a_1^*)}(x, x^*) = +\infty$ and the result now follows from (7). \blacksquare

Example 2.7 Suppose that X is a Hilbert space. Let $e \in X$ be such that $\|e\| = 1$ and define A via $\text{gra } A := \{(-e, e), (e, -e)\}$. Then A is not monotone, and for every $k \in \{2, 3, \dots\}$ and

$(x, x^*) \in X \times X$, we have

$$C_{A,2,(-e,e)}(x, x^*) = \langle x - x^*, e \rangle + 1, \quad (29)$$

$$C_{A,2,(e,-e)}(x, x^*) = \langle x^* - x, e \rangle + 1, \quad (30)$$

$$C_{A,2k-1,(-e,e)}(x, x^*) = 4(k-1) - 2 + \max \{ \langle x - x^*, e \rangle - 1, -\langle x + x^*, e \rangle + 1 \}, \quad (31)$$

$$C_{A,2k-1,(e,-e)}(x, x^*) = 4(k-1) - 2 + \max \{ \langle x + x^*, e \rangle + 1, \langle x^* - x, e \rangle - 1 \}, \quad (32)$$

$$C_{A,2k,(-e,e)}(x, x^*) = 4(k-1) + \max \{ \langle x - x^*, e \rangle + 1, -\langle x + x^*, e \rangle - 1 \}, \quad (33)$$

$$C_{A,2k,(e,-e)}(x, x^*) = 4(k-1) + \max \{ \langle x + x^*, e \rangle - 1, \langle x^* - x, e \rangle + 1 \}, \quad (34)$$

$$C_{A,\infty,(-e,e)}(x, x^*) = +\infty, \quad (35)$$

$$C_{A,\infty,(e,-e)}(x, x^*) = +\infty. \quad (36)$$

We now turn to Fitzpatrick functions.

Definition 2.8 (Fitzpatrick functions) [2, Definition 2.2] *Let $A: X \rightrightarrows X^*$. For every $n \in \{2, 3, \dots\}$, the Fitzpatrick function of A of order n is*

$$F_{A,n} := \sup_{(a,a^*) \in \text{gra } A} C_{A,n,(a,a^*)}. \quad (37)$$

The Fitzpatrick function of A of infinite order is

$$F_{A,\infty} := \sup_{n \in \{2, 3, \dots\}} F_{A,n} = \sup_{(a,a^*) \in \text{gra } A} C_{A,\infty,(a,a^*)}. \quad (38)$$

It is clear that each $F_{A,n}$ is convex and lower semicontinuous; moreover, if $\text{gra } A$ is finite, then each $F_{A,n}$ is polyhedral and continuous. The sequence $(F_{A,n})_{n \in \{2, 3, \dots\}}$ is increasing and pointwise convergent to $F_{A,\infty}$, which is convex and lower semicontinuous. An immediate consequence of Definition 2.8 is the following result.

Proposition 2.9 [2, Proposition 2.3] *Let $A: X \rightrightarrows X^*$ and let $n \in \{2, 3, \dots\}$. Then $F_{A,n}: X \times X^* \rightarrow [-\infty, +\infty]$ is convex and lower semicontinuous. At $(x, x^*) \in X \times X^*$, the value of $F_{A,n}$ is given by*

$$\sup_{\substack{(a_1, a_1^*) \in \text{gra } A, \\ \vdots \\ (a_{n-1}, a_{n-1}^*) \in \text{gra } A}} \left(\sum_{i=1}^{n-2} \langle a_{i+1} - a_i, a_i^* \rangle \right) + \langle x - a_{n-1}, a_{n-1}^* \rangle + \langle a_1, x^* \rangle. \quad (39)$$

Moreover,

$$F_{A,n} \geq \langle \cdot, \cdot \rangle \text{ on } \text{gra } A. \quad (40)$$

Proposition 2.10 *Let $A: X \rightrightarrows X^*$, let $n \in \{2, 3, \dots\}$, and let $(x, x^*) \in X \times X^*$. Then $F_{A^{-1},n}(x^*, x) = F_{A,n}(x, x^*)$ and $F_{A^{-1},\infty}(x^*, x) = F_{A,\infty}(x, x^*)$.*

Proof. Take $(b_1^*, b_1), \dots, (b_{n-1}^*, b_{n-1})$ in $\text{gra } A^{-1}$ and set

$$(\forall i \in \{1, \dots, n-1\}) \quad (a_i, a_i^*) := (b_{n-i}, b_{n-i}^*) \in \text{gra } A. \quad (41)$$

Then

$$\begin{aligned} & \sum_{i=1}^{n-2} \langle b_i, b_{i+1}^* - b_i^* \rangle + \langle b_{n-1}, x^* - b_{n-1}^* \rangle + \langle x, b_1^* \rangle \\ &= \sum_{i=1}^{n-2} \langle b_i, b_{i+1}^* \rangle - \sum_{i=1}^{n-2} \langle b_i, b_i^* \rangle + \langle b_{n-1}, x^* \rangle + \langle x, b_1^* \rangle \\ &= \sum_{i=1}^{n-2} \langle a_{i+1}, a_i^* \rangle - \sum_{i=1}^{n-2} \langle a_i, a_i^* \rangle + \langle a_1, x^* \rangle + \langle x, a_{n-1}^* \rangle \\ &= \sum_{i=1}^{n-2} \langle a_{i+1} - a_i, a_i^* \rangle + \langle x - a_{n-1}, a_{n-1}^* \rangle + \langle a_1, x^* \rangle. \end{aligned} \quad (42)$$

The result follows by supremizing. ■

Fact 2.11 [2, Proposition 2.4 and Corollary 2.5] *Let $A: X \rightrightarrows X^*$ and let $n \in \{2, 3, \dots\}$. Then*

$$A \text{ is } n\text{-cyclically monotone} \Leftrightarrow F_{A,n} \leq \langle \cdot, \cdot \rangle \text{ on } \text{gra } A \Leftrightarrow F_{A,n} = \langle \cdot, \cdot \rangle \text{ on } \text{gra } A, \quad (43)$$

and

$$A \text{ is cyclically monotone} \Leftrightarrow F_{A,\infty} \leq \langle \cdot, \cdot \rangle \text{ on } \text{gra } A \Leftrightarrow F_{A,\infty} = \langle \cdot, \cdot \rangle \text{ on } \text{gra } A. \quad (44)$$

Corollary 2.12 *Let $A: X \rightrightarrows X^*$ and let $n \in \{2, 3, \dots\}$. Then A is n -cyclically monotone if and only if A^{-1} is.*

The recursion formula for Fitzpatrick functions that we present next is an immediate consequence of Proposition 2.3. (A special case of it was utilized in [3].)

Proposition 2.13 (recursion) *Let $A: X \rightrightarrows X^*$, let $n \in \{2, 3, \dots\}$, and let $(x, x^*) \in X \times X^*$. Then*

$$F_{A,n+1}(x, x^*) = \sup_{(a, a^*) \in \text{gra } A} F_{A,n}(a, x^*) + \langle x, a^* \rangle - \langle a, a^* \rangle. \quad (45)$$

Combining Fact 2.11 and Proposition 2.13, we obtain the following result which underlines the importance of the values of the Fitzpatrick function on $\text{dom } A \times \text{ran } A$.

Corollary 2.14 *Let $A: X \rightrightarrows X^*$ and let $n \in \{3, 4, \dots\}$. Then A is n -cyclically monotone if and only if*

$$(\forall (a, a^*) \in \text{gra } A) (\forall (b, b^*) \in \text{gra } A) \quad F_{A,n-1}(a, b^*) - \langle a, b^* \rangle \leq \langle a - b, a^* - b^* \rangle. \quad (46)$$

Example 2.15 Let $A: X \rightrightarrows X^*$ be monotone such that its graph contains two points, and let $n \in \{2, 3, \dots\}$. Then $F_{A,n} = \langle \cdot, \cdot \rangle$ on $\text{dom } A \times \text{ran } A$; consequently, A is cyclically monotone.

Proof. The fact that $F_{A,n} = \langle \cdot, \cdot \rangle$ is proved readily by induction. The cyclic monotonicity of A now follows from Corollary 2.14 and from the monotonicity of A . (Alternatively, use Corollary 2.18 below.) \blacksquare

Theorem 2.16 Let $A: X \rightrightarrows X^*$ be such that $\text{gra } A$ contains at most n points, where $n \in \{2, 3, \dots\}$. Suppose that A is n -cyclically monotone. Then A is $(n+1)$ -cyclically monotone and

$$F_{A,n+1} = F_{A,n+2} = \dots = F_{A,\infty}. \quad (47)$$

Proof. Take

$$\{(b_1, b_1^*), \dots, (b_{n+1}, b_{n+1}^*)\} \subseteq \text{gra } A. \quad (48)$$

We must show that

$$\sigma := \sum_{i=1}^{n+1} \langle b_{i+1} - b_i, b_i^* \rangle \leq 0, \quad \text{where } b_{n+2} := b_1. \quad (49)$$

Since $\text{gra } A$ contains no more than n points, there exist integers k and l such that

$$b_k = b_l \quad \text{and} \quad 1 \leq k < l \leq n+1. \quad (50)$$

Then

$$\sigma = \sigma_1 + \sigma_2, \quad (51)$$

where

$$\sigma_1 := \sum_{i=k}^{l-1} \langle b_{i+1} - b_i, b_i^* \rangle \quad \text{and} \quad \sigma_2 := \sum_{i=l}^{n+1} \langle b_{i+1} - b_i, b_i^* \rangle + \sum_{i=1}^{k-1} \langle b_{i+1} - b_i, b_i^* \rangle \quad (52)$$

are two cyclic sums, each of which contains at least one term and hence at most n terms. Since A is n -cyclically monotone, we see that $\sigma_1 \leq 0$ and that $\sigma_2 \leq 0$. Therefore, $\sigma = \sigma_1 + \sigma_2 \leq 0$. The statement concerning the Fitzpatrick functions follows from Theorem 2.4 and Definition 2.8. \blacksquare

Example 2.17 Let $A: X \rightrightarrows X^*$ be such that $\text{gra } A = \{(a, a^*)\}$, for some $(a, a^*) \in X \times X^*$. Then A is cyclically monotone and for every $(x, x^*) \in X \times X^*$, we have

$$F_{A,2}(x, x^*) = F_{A,3}(x, x^*) = \dots = F_{A,\infty}(x, x^*) = \langle a, x^* \rangle + \langle x, a^* \rangle - \langle a, a^* \rangle. \quad (53)$$

Corollary 2.18 Let $A: X \rightrightarrows X^*$ be such that $\text{gra } A$ contains at most n points, where $n \in \{2, 3, \dots\}$. Suppose that A is n -cyclically monotone. Then A is cyclically monotone.

Proof. By Theorem 2.16, A is $(n+1)$ -cyclically monotone and

$$F_{A,n+1} = F_{A,\infty}. \quad (54)$$

On the other hand, Fact 2.11 yields

$$F_{A,n+1} = \langle \cdot, \cdot \rangle \text{ on } \text{gra } A. \quad (55)$$

The result follows by combining (54), (55), and Fact 2.11. \blacksquare

Corollary 2.19 *Let $A: X \rightrightarrows X^*$ be such that $\text{gra } A$ contains at most n points, where $n \in \{2, 3, \dots\}$. Then A is cyclically monotone if and only if*

$$(\forall (a, a^*) \in \text{gra } A) \quad F_{A,n}(a, a^*) = \langle a, a^* \rangle. \quad (56)$$

Proof. “ \Rightarrow ”: On $\text{gra } A$, we always have $\langle \cdot, \cdot \rangle \leq F_{A,n} \leq F_{A,\infty}$. Since A is cyclically monotone, $F_{A,\infty} = \langle \cdot, \cdot \rangle$ on $\text{gra } A$ and hence $F_{A,n} = \langle \cdot, \cdot \rangle$ on $\text{gra } A$. “ \Leftarrow ”: By Fact 2.11, A is n -cyclically monotone. The result now follows from Corollary 2.18. \blacksquare

Example 2.20 Suppose that X is a Hilbert space. Let $e \in X$ such that $\|e\| = 1$ and define A via $\text{gra } A = \{(-e, -e), (e, e)\}$. Then A is cyclically monotone but $F_{A,2} \neq F_{A,3}$; in fact, for every $(x, x^*) \in X \times X$, we have

$$F_{A,2}(x, x^*) = \max \{ -1 \pm \langle x + x^*, e \rangle \} \quad (57)$$

and

$$F_{A,3}(x, x^*) = \dots = F_{A,\infty}(x, x^*) = \max \{ -1 \pm \langle x + x^*, e \rangle, -3 \pm \langle x - x^*, e \rangle \}. \quad (58)$$

Proof. The operator A is cyclically monotone since $\text{gra } A \subset \text{gra } \text{Id} = \text{gra } \partial \frac{1}{2} \|\cdot\|^2$. The formulae for $F_{A,2}$ and $F_{A,3}$ follow from Example 2.5. Theorem 2.16 shows that $F_{A,3} = \dots = F_{A,\infty}$. Finally, we note that $F_{A,2}(2e, -2e) = -1$ whereas $F_{A,3}(2e, -2e) = 1$. \blacksquare

The next example, which is an immediate consequence of Example 2.7 and Definition 2.8, illustrates the nonmonotone case.

Example 2.21 Suppose that X is a Hilbert space. Let $e \in X$ be such that $\|e\| = 1$ and define A via $\text{gra } A := \{(-e, e), (e, -e)\}$. Then A is not monotone, and for every $k \in \{2, 3, \dots\}$ and $(x, x^*) \in X \times X$, we have

$$F_{A,2}(x, x^*) = 1 + \max \{ \pm \langle x - x^*, e \rangle \}, \quad (59)$$

$$F_{A,2k-1}(x, x^*) = 4(k-1) - 2 + \max \{ -1 \pm \langle x - x^*, e \rangle, 1 \pm \langle x + x^*, e \rangle \}, \quad (60)$$

$$F_{A,2k}(x, x^*) = 4(k-1) + \max \{ 1 \pm \langle x - x^*, e \rangle, -1 \pm \langle x + x^*, e \rangle \}, \quad (61)$$

$$F_{A,\infty}(x, x^*) = +\infty. \quad (62)$$

3 Rockafellar functions

Definition 3.1 *Let $A: X \rightrightarrows X^*$ and let $f \in \Gamma$. Then f is an antiderivative of A if*

$$\text{gra } A \subseteq \text{gra } \partial f. \quad (63)$$

The following result will turn out to be useful.

Proposition 3.2 *Suppose that X is reflexive. Let $A: X \rightrightarrows X^*$, let $f \in \Gamma$, and suppose that f is an antiderivative of A such that $\text{ran } \partial f \subseteq \text{conv ran } A$. Then $\text{dom } f^* = \text{conv ran } A$.*

Proof. On the one hand, since f is an antiderivative of A , we deduce that $\text{gra } A \subseteq \text{gra } \partial f \Leftrightarrow \text{gra } A^{-1} \subseteq \text{gra } (\partial f)^{-1} = \text{gra } \partial f^* \Rightarrow \text{ran } A = \text{dom } A^{-1} \subseteq \text{dom } \partial f^* \subseteq \text{dom } f^* \Rightarrow \text{conv ran } A \subseteq \text{conv dom } f^* = \text{dom } f^*$. Because $\overline{\text{ran } \partial f} \subseteq \text{conv ran } A$, we see that $\text{dom } f^* \subseteq \overline{\text{dom } f^*} = \overline{\text{dom } \partial f^*} = \overline{\text{dom } (\partial f)^{-1}} = \overline{\text{ran } \partial f} \subseteq \text{conv ran } A$. Altogether, $\text{dom } f^* = \text{conv ran } A$. \blacksquare

Definition 3.3 (Rockafellar function) *Let $A: X \rightrightarrows X^*$. Then the Rockafellar functions are defined by*

$$(\forall (a, a^*) \in \text{gra } A) \quad R_{A,(a,a^*)}: X \rightarrow]-\infty, +\infty]: x \mapsto \sup_{n \in \{2,3,\dots\}} C_{A,n,(a,a^*)}(x, 0). \quad (64)$$

The importance of the Rockafellar functions stems from a fundamental result due to Rockafellar (see [19] or [24, Proposition 2.4.3, Theorem 3.2.8, and Corollary 3.2.11]) which states that maximal cyclically monotone operators are precisely the subdifferential operators of convex, lower semicontinuous and proper functions. The following part of Rockafellar's result will be utilized later.

Fact 3.4 (Rockafellar) (See [19] or [24, Proposition 2.4.3 and Corollary 3.2.11].) *Let $A: X \rightrightarrows X^*$ be cyclically monotone and let $(a, a^*) \in \text{gra } A$. Then the following hold.*

- (i) $R_{A,(a,a^*)}$ is convex, lower semicontinuous, and proper, $R_{A,(a,a^*)}(a) = 0$, and $R_{A,(a,a^*)}$ is an antiderivative of A .
- (ii) If A is maximal cyclically monotone, then any two antiderivatives of A differ only by a constant.

Among all antiderivatives, Rockafellar functions have a special status due to the following minimality property, which was first observed in [14, Theorem 3.4] for cyclically monotone operators with finite graph.

Theorem 3.5 *Let $A: X \rightrightarrows X^*$ be cyclically monotone and let $a \in \text{dom } A$. Then*

$$(\forall a^* \in Aa) \quad R_{A,(a,a^*)} = \min \{ f \in \Gamma(X) \mid f \text{ is an antiderivative of } A \text{ with } f(a) \geq 0 \} \quad (65)$$

$$= \min \{ f \in \Gamma(X) \mid f \text{ is an antiderivative of } A \text{ with } f(a) = 0 \}. \quad (66)$$

Proof. Suppose that $f \in \Gamma$ is an antiderivative of A with $f(a) \geq 0$, take $a^* \in Aa$ and $x \in X$. Then for every $x \in X$, $n \in \{1, 2, \dots\}$ and $(a_1, a_1^*), \dots, (a_n, a_n^*)$ belonging to $\text{gra } A$, we have

$$\begin{aligned} f(x) &\geq f(x) - f(a_n) + \left(\sum_{i=1}^{n-1} f(a_{i+1}) - f(a_i) \right) + f(a_1) - f(a) \\ &\geq \langle x - a_n, a_n^* \rangle + \left(\sum_{i=1}^{n-1} \langle a_{i+1} - a_i, a_i^* \rangle \right) + \langle a_1 - a, a^* \rangle. \end{aligned} \quad (67)$$

This implies

$$f \geq R_{A,(a,a^*)}. \quad (68)$$

In view of Fact 3.4(i), the proof is complete. \blacksquare

Corollary 3.6 *Let $A: X \rightrightarrows X^*$ be cyclically monotone, let $a \in \text{dom } A$, let $a_1^* \in Aa$, and let $a_2^* \in Aa$. Then $R_{A,(a,a_1^*)} = R_{A,(a,a_2^*)}$.*

Corollary 3.6 and Theorem 2.6 make the following definition well defined.

Definition 3.7 *Let $A: X \rightrightarrows X^*$ and let $a \in \text{dom } A$. Then we set*

$$R_{A,a} := R_{A,(a,a^*)}, \quad (69)$$

where a^* is an arbitrary point in Aa .

Corollary 3.8 *Let $A: X \rightrightarrows X^*$ be cyclically monotone and let $a \in \text{dom } A$. Set $B: X \rightrightarrows X^*: x \mapsto \text{conv}(Ax)$. Then B is cyclically monotone and $R_{B,a} = R_{A,a}$.*

Proof. It is readily verified that B is cyclically monotone. Hence $R_{B,a}$ is an antiderivative of B and of A such that $R_{B,a}(a) = 0$. By Theorem 3.5, $R_{B,a} \geq R_{A,a}$. On the other hand, $R_{A,a}$ is also an antiderivative of B ; thus, again by Theorem 3.5, $R_{A,a} \geq R_{B,a}$. Altogether, $R_{B,a} = R_{A,a}$. \blacksquare

Corollary 3.9 *Suppose that X is reflexive. Let $A: X \rightrightarrows X^*$ be cyclically monotone and let $(a, a^*) \in \text{gra } A$. Then*

$$R_{A,a}^* = \max \{g \in \Gamma(X^*) \mid g \text{ is an antiderivative of } A^{-1} \text{ and } g(a^*) = \langle a, a^* \rangle\}. \quad (70)$$

Proof. Take $g \in \Gamma(X^*)$ such that g is an antiderivative of A^{-1} and $g(a^*) = \langle a, a^* \rangle$. Then $g^*(a) = 0$ and g^* is an antiderivative of A . By Theorem 3.5, $g^* \geq R_{A,a}$ and therefore $g^{**} = g \leq R_{A,a}^*$. \blacksquare

Corollary 3.9 results in the following interesting counterpart to Theorem 3.5; see also [14, Proposition 4.2].

Corollary 3.10 *Suppose that X is reflexive. Let $A: X \rightrightarrows X^*$ be cyclically monotone and let $(a, a^*) \in \text{gra } A$. Then*

$$R_{A^{-1},a^*}^* - \langle a, a^* \rangle = \max \{f \in \Gamma(X) \mid f \text{ is an antiderivative of } A \text{ and } f(a) = 0\}. \quad (71)$$

The next result will be used later.

Corollary 3.11 *Let $A: X \rightrightarrows X^*$, let $(a, a^*) \in \text{gra } A$, and let $n \in \{2, 3, \dots\}$. Suppose that $\text{gra } A$ contains at most n points and that A is n -cyclically monotone. Then A is cyclically monotone and*

for every $x \in X$, we have

$$R_{A,a}(x) = C_{A,n+1,(a,a^*)}(x, 0) \quad (72)$$

$$= \max_{\substack{(a_2, a_2^*) \in \text{gra } A, \\ \vdots \\ (a_n, a_n^*) \in \text{gra } A}} \langle x - a_n, a_n^* \rangle + \langle a_n - a_{n-1}, a_{n-1}^* \rangle + \cdots + \langle a_2 - a, a^* \rangle. \quad (73)$$

Consequently, $R_{A,a}$ is a polyhedral and continuous antiderivative of A with $\text{ran } \partial R_{A,a} \subset \text{conv ran } A$.

Proof. This follows from Corollary 2.18, Theorem 2.4, (8), Fact 3.4(i), and the Ioffe-Tikhomirov theorem (see, e.g., [24, Theorem 2.4.18]). \blacksquare

Fact 3.4(ii) implies that if A is maximal cyclically monotone, then the Rockafellar functions $\{R_{A,a}\}_{a \in \text{dom } A}$ differ only by constants. For finite-graph operators, this is no longer true as the following consequence of Example 2.5 and Definition 3.3 shows.

Example 3.12 Suppose that X is a Hilbert space. Let $e \in X$ be such that $\|e\| = 1$ and define A via $\text{gra } A := \{(-e, -e), (e, e)\}$. Then for every $x \in X$, we have

$$R_{A,-e}(x) = \max \{ -\langle x, e \rangle - 1, \langle x, e \rangle - 3 \} = -2 + |\langle x, e \rangle - 1| \quad (74)$$

and

$$R_{A,e}(x) = \max \{ \langle x, e \rangle - 1, -\langle x, e \rangle - 3 \} = -2 + |\langle x, e \rangle + 1|. \quad (75)$$

Consequently, $R_{A,e} \not\leq R_{A,-e}$ and $R_{A,-e} \not\leq R_{A,e}$.

Remark 3.13 Let $A: X \rightrightarrows X^*$ and suppose that A is not cyclically monotone. Then Theorem 2.6 and Definition 3.3 imply that the Rockafellar functions $\{R_{A,a}\}_{a \in \text{dom } A}$ are all identically equal to $+\infty$. For a concrete example, see Example 2.7.

Turning momentarily to the case when $X = \mathbb{R}$, we now present not only a considerable generalization of Example 3.12 but also an explicit formula for any Rockafellar function and its subdifferential operator of a cyclically monotone operator with finite graph. See also [14, Section 7].

Theorem 3.14 Let $A: \mathbb{R} \rightrightarrows \mathbb{R}$ have finite graph and suppose that the graph of $B: \mathbb{R} \rightrightarrows \mathbb{R}: x \mapsto \text{conv}(Ax)$ is

$$\bigcup_{i=1}^n (\{a_i\} \times [b_i^-, b_i^+]), \quad (76)$$

where $n \in \{1, 2, \dots\}$, $a_1 < a_2 < \cdots < a_n$, and $b_1^- \leq b_1^+ \leq b_2^- \leq \cdots \leq b_n^- \leq b_n^+$. Set $a_0 := -\infty$ and $a_{n+1} := +\infty$. Suppose that $k \in \{1, \dots, n\}$. Then R_{A,a_k} is given by

$$\mathbb{R} \rightarrow \mathbb{R}: x \mapsto \begin{cases} (x - a_i)b_i^- + \sum_{j=i+1}^k (a_{j-1} - a_j)b_j^-, & \text{if } a_{i-1} < x \leq a_i \leq a_k; \\ (x - a_i)b_i^+ + \sum_{j=k}^{i-1} (a_{j+1} - a_j)b_j^+, & \text{if } a_k \leq a_i \leq x < a_{i+1}, \end{cases} \quad (77)$$

and $\partial R_{A,a_k}$ is given by

$$\mathbb{R} \rightrightarrows \mathbb{R}: x \mapsto \begin{cases} \{b_i^-\}, & \text{if } a_{i-1} < x < a_i \leq a_k; \\ [b_i^-, b_{i+1}^-], & \text{if } x = a_i < a_k; \\ [b_k^-, b_k^+], & \text{if } x = a_k; \\ [b_{i-1}^+, b_i^+], & \text{if } a_k < x = a_i; \\ \{b_i^+\}, & \text{if } a_k \leq a_i < x < a_{i+1}. \end{cases} \quad (78)$$

Proof. Clearly, A and B are cyclically monotone. Denote the function described in (77) by R and observe that R is piecewise linear, continuous everywhere, and well defined at a_k with

$$R(a_k) = 0. \quad (79)$$

Moreover, (77) implies that ∂R is given by (78), which is clearly monotone. Thus

$$R \in \Gamma(\mathbb{R}). \quad (80)$$

Take $i \in \{1, 2, \dots, n\}$. If $i < k$, then $Aa_i \subseteq Ba_i = [b_i^-, b_i^+] \subseteq [b_i^-, b_{i+1}^-] = \partial R(a_i)$. If $i = k$, then $Aa_i = Aa_k \subseteq Ba_k = [b_k^-, b_k^+] = \partial R(a_k)$. If $k < i$, then $Aa_i \subseteq Ba_i = [b_i^-, b_i^+] \subseteq [b_{i-1}^+, b_i^+] = \partial R(a_i)$. Thus,

$$R \text{ is an antiderivative of } A. \quad (81)$$

Since $\bigcup_{i=1}^n \{(a_i, b_i^-), (a_i, b_i^+)\} \subseteq \text{gra } A$, we deduce from (73) and (77) that

$$R_{A,a_k} \geq R. \quad (82)$$

Hence (79), (80), (81), (82), and Theorem 3.5 imply that $R = R_{A,a_k}$. \blacksquare

The next result links Rockafellar functions to Fitzpatrick functions.

Theorem 3.15 *Let $A: X \rightrightarrows X^*$. Then*

$$(\forall (x, x^*) \in X \times X^*) \quad F_{A,\infty}(x, x^*) = \sup_{a \in \text{dom } A} \langle a, x^* \rangle + R_{A,a}(x). \quad (83)$$

Proof. (See also the proof of [2, Theorem 3.5] for a variant.) Take $(x, x^*) \in X \times X^*$. Using Definitions 2.8, 2.1, 3.3, and 3.7, we see that

$$F_{A,\infty}(x, x^*) = \sup_{n \in \{2, 3, \dots\}} F_{A,n}(x, x^*) = \sup_{n \in \{2, 3, \dots\}} \sup_{(a, a^*) \in \text{gra } A} C_{A,n,(a,a^*)}(x, x^*) \quad (84)$$

$$= \sup_{(a, a^*) \in \text{gra } A} \sup_{n \in \{2, 3, \dots\}} C_{A,n,(a,a^*)}(x, 0) + \langle a, x^* \rangle \quad (85)$$

$$= \sup_{(a, a^*) \in \text{gra } A} \langle a, x^* \rangle + R_{A,(a,a^*)}(x) \quad (86)$$

$$= \sup_{a \in \text{dom } A} \langle a, x^* \rangle + R_{A,a}(x), \quad (87)$$

as required. ■

We deduce that the Fitzpatrick function of infinite order with second variable set to zero is exactly the supremum of all Rockafellar functions.

Corollary 3.16 *Let $A: X \rightrightarrows X^*$. Then*

$$F_{A,\infty}(\cdot, 0) = \sup_{a \in \text{dom } A} R_{A,a}. \quad (88)$$

Remark 3.17 Let $A: X \rightrightarrows X^*$ be maximal cyclically monotone. Rockafellar [19] (see Fact 3.4) proved that $A = \partial f$, where $f \in \Gamma$ is uniquely determined up to additive constants. By [2, Theorem 3.5], $F_{A,\infty} = f \oplus f^*$. Thus Corollary 3.16 implies that

$$\sup_{a \in \text{dom } A} R_{A,a} \equiv +\infty \Leftrightarrow 0 \notin \text{dom } f^* \Leftrightarrow \inf f(X) = -\infty. \quad (89)$$

Corollary 3.18 *Let $A: X \rightrightarrows X^*$ be cyclically monotone with finite graph, let $x^* \in X^*$, and set $f := F_{A,\infty}(\cdot, x^*)$. Then f is polyhedral, continuous, with full domain, $\text{gra } A \subset \text{gra } \partial f$, and $\overline{\text{ran } \partial f} \subseteq \text{conv ran } A$.*

Proof. Since $\text{gra } A$ is finite, Theorem 3.15 yields

$$f = \max_{a \in \text{dom } A} \langle a, x^* \rangle + R_{A,a}. \quad (90)$$

The function f is continuous, polyhedral, with full domain, as it is the finite maximum of such functions (see Corollary 3.11). Fix $x \in X$ and set $D_x := \{a \in \text{dom } A \mid f(x) = \langle a, x^* \rangle + R_{A,a}(x)\}$. On the one hand, using the Ioffe-Tikhomirov theorem (see, e.g., [24, Theorem 2.4.18]) and Corollary 3.11, we have

$$\partial f(x) = \overline{\text{conv}}^* \bigcup_{a \in D_x} \partial R_{A,a}(x) \subseteq \overline{\text{conv}}^* \bigcup_{a \in D_x} \text{conv ran } A = \text{conv ran } A, \quad (91)$$

where $\overline{\text{conv}}^*$ denotes the weak* closed convex hull operator. Hence $\text{ran } \partial f \subseteq \text{conv ran } A$ and thus $\overline{\text{ran } \partial f} \subseteq \text{conv ran } A$, since $\text{conv ran } A$ is compact as a convex hull of finitely many points. On the other hand, Fact 3.4(i) implies that

$$(\forall a \in \text{dom } A) \quad \text{gra } A \subset \text{gra } \partial R_{A,a}. \quad (92)$$

Combining (91) and (92), we conclude altogether that $\text{gra } A \subset \text{gra } \partial f$. ■

Example 3.19 Suppose that X is a Hilbert space. Let $e \in X$ such that $\|e\| = 1$ and define A via $\text{gra } A = \{(-e, -e), (e, e)\}$. Then A is cyclically monotone and for every $x \in X$, we have

$$F_{A,\infty}(x, 0) = \max \{R_{A,-e}(x), R_{A,e}(x)\} = \max \{-1 \pm \langle x, e \rangle\} = -1 + |\langle x, e \rangle|. \quad (93)$$

Proof. Combine Example 3.12 and Corollary 3.16. ■

We conclude this section with a result which illustrates how Fitzpatrick functions give rise to the smallest nonnegative antiderivative.

Corollary 3.20 *Let $A : X \rightrightarrows X^*$ be cyclically monotone with finite graph. Then*

$$F_{A,\infty}(\cdot, 0) = \min \{f \in \Gamma(X) \mid f \text{ is an antiderivative of } A \text{ such that } f \geq 0 \text{ on } \text{dom } A\}. \quad (94)$$

Proof. Take $f \in \Gamma(X)$ such that f is an antiderivative of A and $f \geq 0$ on $\text{dom } A$. Then Theorem 3.5 implies that $(\forall a \in \text{dom } A) f \geq R_{A,a}$; hence, by Corollary 3.16,

$$f \geq \max_{a \in \text{dom } A} R_{A,a} = F_{A,\infty}(\cdot, 0). \quad (95)$$

In view of Corollary 3.18 and Fact 3.4(i), $F_{A,\infty}(\cdot, 0)$ is an antiderivative of A that is nonnegative on $\text{dom } A$. ■

4 Intrinsic and primal-dual symmetric methods

From now on,

\mathcal{A} is the set of all cyclically monotone operators on X with finite nonempty graphs. (96)

Definition 4.1 *An intrinsic method for finding antiderivatives — or simply an intrinsic method — is a mapping $m : \mathcal{A} \rightarrow \Gamma : A \mapsto m_A$ such that for every $A \in \mathcal{A}$, m_A is an antiderivative of A .*

Example 4.2 Let $A : X \rightrightarrows X^*$ be cyclically monotone and let $(a, a^*) \in \text{gra } A$. Then the Rockafellar function $R_{A,(a,a^*)} = R_{A,a}$ is an antiderivative (see Fact 3.4) but — due to the dependency on a and the resulting non-uniqueness of Rockafellar functions — there is no corresponding intrinsic method m that produces Rockafellar functions. See Example 3.12 for a concrete example.

Remark 4.3 Given $A \in \mathcal{A}$, an intrinsic method m provides an antiderivatives m_A as a mapping depending only on A , or equivalently, only on the (unordered) graph of A . This key property of intrinsic methods explains why the process of providing Rockafellar functions considered in Example 4.2 is not intrinsic. Similarly, if a method computes antiderivatives by using an enumeration of the graph of A and if a different enumeration may result in a different antiderivative, then such a method cannot be intrinsic.

We now provide two intrinsic methods.

Example 4.4 Let $m : \mathcal{A} \rightarrow \Gamma : A \mapsto F_{A,\infty}(\cdot, 0) = \max_{(a,a^*) \in \text{gra } A} R_{A,(a,a^*)}$. Corollary 3.16 and Corollary 3.18 imply that m is an intrinsic method. Moreover, for every $A \in \mathcal{A}$, the antiderivative m_A has full domain and $\overline{\text{ran } \partial m_A} \subseteq \text{conv ran } A$.

Example 4.5 Let $A: X \rightrightarrows X^*$ be cyclically monotone such that $\text{gra } A$ contains exactly n points, where $n \in \{1, 2, \dots\}$, and set

$$m_A := \sum_{(a, a^*) \in \text{gra } A} \frac{1}{n} R_{A, (a, a^*)}. \quad (97)$$

Then m_A is an antiderivative of A that is polyhedral and continuous with full domain and $\text{ran } \partial m_A \subseteq \text{conv ran } A$. Furthermore, the corresponding method $m: \mathcal{A} \rightarrow \Gamma: A \mapsto m_A$ is intrinsic.

Proof. Note that m_A is continuous and polyhedral with full domain, as a finite sum of such functions. The sum rule (see, e.g., [24, Theorem 2.8.7(iii)]) and Fact 3.4 imply that for every $(x, x^*) \in \text{gra } A$, we have $x^* \in Ax \subseteq \sum_{(a, a^*) \in \text{gra } A} \frac{1}{n} Ax \subseteq \sum_{(a, a^*) \in \text{gra } A} \frac{1}{n} \partial R_{A, (a, a^*)}(x) = \partial m_A(x)$. Corollary 3.11 shows that for every $x \in X$, we have $\partial m_A(x) = \sum_{(a, a^*) \in \text{gra } A} \frac{1}{n} \partial R_{A, (a, a^*)}(x) \subseteq \sum_{(a, a^*) \in \text{gra } A} \frac{1}{n} \text{conv ran } A = \text{conv ran } A$. Consequently, $\text{ran } \partial m_A \subseteq \text{conv ran } A$ and hence $\text{ran } \partial m_A \subseteq \text{conv ran } A = \text{conv ran } A$. It is clear that m is intrinsic. ■

We assume from now on that

$$X \text{ is a Hilbert space.} \quad (98)$$

Definition 4.6 An intrinsic method $\mathfrak{m}: \mathcal{A} \rightarrow \Gamma: A \mapsto \mathfrak{m}_A$ is primal-dual symmetric if

$$(\forall A \in \mathcal{A}) \quad \mathfrak{m}_{A^{-1}} = \mathfrak{m}_A^*. \quad (99)$$

Proposition 4.7 While intrinsic, neither

$$A \rightarrow \Gamma: A \mapsto F_{A, \infty}(\cdot, 0) = \max_{(a, a^*) \in \text{gra } A} R_{A, (a, a^*)} \quad (100)$$

nor

$$A \rightarrow \Gamma: A \mapsto \sum_{(a, a^*) \in \text{gra } A} \frac{1}{n_A} R_{A, (a, a^*)}, \quad \text{where } n_A \text{ is the number of points in } \text{gra } A, \quad (101)$$

is primal-dual symmetric.

Proof. On the one hand, both methods produce polyhedral continuous functions with *full domain*. On the other hand, the Fenchel conjugates of such functions have *bounded domain*. ■

Since antiderivatives are only (and at best, see Example 3.12) unique up to a constant, it is perhaps surprising that primal-dual symmetric methods even exist. The remainder of this section is devoted to the derivation of such methods. We shall require several known notions which we review now.

Let $A: X \rightrightarrows X$ be a monotone operator. The *resolvent* of A is (the single-valued, firmly nonexpansive operator) $J_A := (\text{Id} + A)^{-1}$, where Id denotes the identity operator. A classical result due to Minty [16] asserts that J_A has full domain if and only if A is maximal monotone. The proof of the following result is straightforward and hence omitted.

Proposition 4.8 *Let $A: X \rightrightarrows X$ and $B: X \rightrightarrows X$ be monotone operators. Then the following are equivalent.*

- (i) $\text{gra } A \subseteq \text{gra } B$.
- (ii) $\text{gra } A^{-1} \subseteq \text{gra } B^{-1}$.
- (iii) J_B is an extension of J_A , i.e., $J_B = J_A$ on $\text{dom } J_A = \text{ran}(\text{Id} + A)$.

We further recall that given $f \in \Gamma$, the *proximal mapping* [17] of f is $\text{Prox}(f) := J_{\partial f}$. It is clear from the definition that for two points x and x^* in X , one has

$$x^* \in \partial f(x) \Leftrightarrow x = \text{Prox}(f)(x + x^*). \quad (102)$$

Proposition 4.9 *Let $f \in \Gamma$, let $(a, a^*) \in \text{gra } \partial f$, and suppose that $y \in N_{\text{dom } f}(a)$. Then $a = \text{Prox}(f)(2y + a + a^*)$.*

Proof. Since $N_{\text{dom } f}(a)$ is a cone, we have $2y \in N_{\text{dom } f}(a)$. Hence $2y + a + a^* \in a + a^* + \partial \iota_{\text{dom } f}(a) \subseteq a + \partial f(a) + \partial \iota_{\text{dom } f}(a) \subseteq a + \partial(f + \iota_{\text{dom } f})(a) = a + \partial f(a) = (\text{Id} + \partial f)(a)$ and thus $a = \text{Prox}(f)(2y + a + a^*)$. \blacksquare

We need one more notion.

Definition 4.10 *Let $f_0 \in \Gamma$ and let $f_1 \in \Gamma$. The proximal midpoint average of f_0 and f_1 is the function*

$$\mathcal{P}(f_0, f_1) := \left(\frac{1}{2}(f_0 + \frac{1}{2}\|\cdot\|^2)^* + \frac{1}{2}(f_1 + \frac{1}{2}\|\cdot\|^2)^* \right)^* - \frac{1}{2}\|\cdot\|^2. \quad (103)$$

The proximal average, which is a generalization of the proximal midpoint average with a parameter $\lambda \in [0, 1]$ (the choice $\lambda = \frac{1}{2}$ yields the proximal midpoint average), was introduced in [6], and further studied in [4, 5, 8].

We require the following properties.

Fact 4.11 *Let $f_0 \in \Gamma$ and let $f_1 \in \Gamma$. Then the following hold.*

- (i) $\mathcal{P}(f_0, f_1) = \mathcal{P}(f_1, f_0)$.
- (ii) $(\mathcal{P}(f_0, f_1))^* = \mathcal{P}(f_0^*, f_1^*)$.
- (iii) $\mathcal{P}(f_0, f_1) \in \Gamma$.
- (iv) $\text{Prox}(\mathcal{P}(f_0, f_1)) = \frac{1}{2}\text{Prox}(f_0) + \frac{1}{2}\text{Prox}(f_1)$.
- (v) *If $f_0 \leq f_1$, then $f_0 \leq \mathcal{P}(f_0, f_1) \leq f_1$.*
- (vi) $(\forall \gamma \in \mathbb{R}) \mathcal{P}(f_0, f_0 + \gamma) = f_0 + \frac{1}{2}\gamma$.

Proof. (i): This is clear from the definition. (ii): See [6, Theorem 6.1]. (iii): This follows from (ii). (iv): See [6, Theorem 6.1]. (v) and (vi) follow readily from the definition. (See also [5, Remark 4.15 and Example 7.1].) ■

Corollary 4.12 *Let $A: X \rightrightarrows X$ be cyclically monotone, and let f_0 and f_1 be antiderivatives of A . Then $\mathcal{P}(f_0, f_1)$ is also an antiderivative of A .*

Proof. By assumption, $\text{gra } A \subseteq \text{gra } \partial f_0$ and $\text{gra } A \subseteq \text{gra } \partial f_1$. Using Proposition 4.8, we see that both $\text{Prox}(f_0)$ and $\text{Prox}(f_1)$ extend J_A . Thus, by Fact 4.11(iv), $\text{Prox}(\mathcal{P}(f_0, f_1))$ also extends J_A . Utilizing Proposition 4.8 once more, we deduce that $\mathcal{P}(f_0, f_1)$ is an antiderivative of A . ■

We are now ready for our main result.

Theorem 4.13 (symmetrization) *Let $m: \mathcal{A} \rightarrow \Gamma: A \mapsto m_A$ be an intrinsic method. Set*

$$\mathfrak{m}: \mathcal{A} \rightarrow \Gamma: A \mapsto \mathcal{P}(m_A, m_{A^{-1}}^*). \quad (104)$$

Then \mathfrak{m} is a primal-dual symmetric intrinsic method.

Proof. Fix $A \in \mathcal{A}$. Observe that Corollary 2.12 and Corollary 4.12 imply that \mathfrak{m}_A is an antiderivative of A ; thus, \mathfrak{m} is an intrinsic method. On the one hand, the definitions and Fact 4.11(i) yield

$$\mathfrak{m}_{A^{-1}} = \mathcal{P}(m_{A^{-1}}, m_{(A^{-1})^{-1}}^*) = \mathcal{P}(m_{A^{-1}}, m_A^*) = \mathcal{P}(m_A^*, m_{A^{-1}}). \quad (105)$$

On the other hand, Fact 4.11(ii) implies

$$\mathfrak{m}_A^* = \left(\mathcal{P}(m_A, m_{A^{-1}}^*) \right)^* = \mathcal{P}(m_A^*, m_{A^{-1}}^{**}) = \mathcal{P}(m_A^*, m_{A^{-1}}). \quad (106)$$

Altogether, we obtain that $\mathfrak{m}_{A^{-1}} = \mathfrak{m}_A^*$. Therefore, \mathfrak{m} is primal-dual symmetric. ■

Example 4.14 *Let $m: \mathcal{A} \rightarrow \Gamma: A \mapsto m_A$ be intrinsic, and set $\mathfrak{m}: \mathcal{A} \rightarrow \Gamma: A \mapsto \mathcal{P}(m_A, m_{A^{-1}}^*)$. Let $A \in \mathcal{A}$ be such that $\text{gra } A \subset \text{gra Id}$. Then $\mathfrak{m}_A = \frac{1}{2} \|\cdot\|^2$.*

Proof. Since $A = A^{-1}$, we have $\mathfrak{m}_A^* = \left(\mathcal{P}(m_A, m_{A^{-1}}^*) \right)^* = \left(\mathcal{P}(m_A, m_A^*) \right)^* = \mathcal{P}(m_A^*, m_A^{**}) = \mathcal{P}(m_A^*, m_A) = \mathcal{P}(m_A, m_A^*) = \mathcal{P}(m_A, m_{A^{-1}}^*) = \mathfrak{m}_A$ and the result follows. ■

Before we present further applications of Theorem 4.13, let us discuss a non-intrinsic variant based on the original Rockafellar function.

Theorem 4.15 *Let $A: X \rightrightarrows X^*$ be cyclically monotone and let $(a, a^*) \in \text{gra } A$. Set*

$$f_{A, (a, a^*)} := \mathcal{P}(R_{A, (a, a^*)}, R_{A^{-1}, (a^*, a)}^*). \quad (107)$$

Then $f_{A, (a, a^)}^* = f_{A^{-1}, (a^*, a)} := \mathcal{P}(R_{A^{-1}, (a^*, a)}, R_{A, (a, a^*)}^*)$. Moreover,*

$$A \text{ is maximal cyclically monotone} \Rightarrow f_{A, (a, a^*)} = R_{A, (a, a^*)} + \frac{1}{2} \langle a, a^* \rangle. \quad (108)$$

Proof. The proof of $f_{A,(a,a^*)}^* = f_{A^{-1},(a^*,a)}$ is analogous to the one of Theorem 4.13. Now assume that A is maximal cyclically monotone. Since $R_{A,(a,a^*)}$ is an antiderivative of A and $R_{A^{-1},(a^*,a)}$ is an antiderivative of A^{-1} , there exists $\gamma \in \mathbb{R}$ such that

$$R_{A^{-1},(a^*,a)}^* = R_{A,(a,a^*)} + \gamma. \quad (109)$$

Conjugating (109) followed by evaluating at a^* yields $0 = R_{A,(a,a^*)}^*(a^*) - \gamma = \langle a, a^* \rangle - R_{A,(a,a^*)}(a) - \gamma = \langle a, a^* \rangle - \gamma$. Hence

$$R_{A^{-1},(a^*,a)}^* = R_{A,(a,a^*)} + \langle a, a^* \rangle, \quad (110)$$

and this readily implies that $f_{A,(a,a^*)} = \mathcal{P}(R_{A,(a,a^*)}, R_{A,(a,a^*)} + \langle a, a^* \rangle) = R_{A,(a,a^*)} + \frac{1}{2}\langle a, a^* \rangle$. ■

If the intrinsic method m in Theorem 4.13 produces “nice” antiderivatives, then so does sometimes the symmetrized method \mathfrak{m} . Before we state the corresponding result more precisely, we recall the required properties of the proximal midpoint average. Since these properties were stated in finite-dimensional Hilbert spaces, we assume from now on that

$$X \text{ is a finite-dimensional Hilbert space.} \quad (111)$$

Recall that $f \in \Gamma$ is *piecewise linear-quadratic* if $\text{dom } f$ can be written as a finite union of polyhedral sets on which f is of the form $\langle x, Ax \rangle + \langle x, b \rangle + \gamma$, where $A: X \rightarrow X$ is linear, $b \in X$, and $\gamma \in \mathbb{R}$. The piecewise linear-quadratic functions on X have many nice properties, see [21, Sections 10.E and 11.D].

Fact 4.16 *Let $f_0 \in \Gamma$ and let $f_1 \in \Gamma$ be such that f_0 and f_1^* have full domain. Then the following hold.*

- (i) $\mathcal{P}(f_0, f_1)$ and $\mathcal{P}(f_0^*, f_1^*)$ have full domain.
- (ii) If f_0 and f_1 are piecewise linear-quadratic, then so is $\mathcal{P}(f_0, f_1)$.
- (iii) If f_0 is differentiable and f_1 is strictly convex, then both $\mathcal{P}(f_0, f_1)$ and its conjugate are differentiable and strictly convex.

Proof. (i): See [5, Theorem 6.2.(i)]. (ii): The functions f_0 , f_1 , and $\frac{1}{2}\|\cdot\|^2$ are piecewise linear-quadratic. The operations employed to create $\mathcal{P}(f_0, f_1)$ do not lead outside the class of piecewise linear-quadratic functions; consequently, $\mathcal{P}(f_0, f_1)$ is piecewise linear-quadratic as well. (See also [15, Corollary 5.3].) (iii): This follows from (i) and [5, Theorem 6.2.(ii) and Theorem 6.2.(iii)]. ■

Corollary 4.17 *Let $m: \mathcal{A} \rightarrow \Gamma: A \mapsto m_A$ be an intrinsic method that produces antiderivatives with full domain. Set*

$$\mathfrak{m}: \mathcal{A} \rightarrow \Gamma: A \mapsto \mathcal{P}(m_A, m_{A^{-1}}^*). \quad (112)$$

Then \mathfrak{m} is a primal-dual symmetric intrinsic method, and the following hold.

- (i) $(\forall A \in \mathcal{A}) \mathfrak{m}_A$ and \mathfrak{m}_A^* have full domain.

(ii) If $(\forall A \in \mathcal{A}) m_A$ is piecewise linear-quadratic, then $(\forall A \in \mathcal{A}) \mathfrak{m}_A$ and \mathfrak{m}_A^* are both piecewise linear-quadratic.

Proof. Theorem 4.13 states that \mathfrak{m} is primal-dual symmetric. Now fix $A \in \mathcal{A}$, set $f_0 := m_A$ and set $f_1 := m_{A^{-1}}^*$. Then $\mathfrak{m}_A = \mathcal{P}(f_0, f_1)$ and, by Fact 4.11(i)&(ii), $\mathfrak{m}_A^* = \mathcal{P}(f_0^*, f_1^*) = \mathcal{P}(f_1^*, f_0^*)$. (i): Since f_0 and f_1^* have full domain, Fact 4.16(i) (applied to f_0 and f_1 , and to f_1^* and f_0^*) implies that \mathfrak{m}_A and \mathfrak{m}_A^* have full domain. (ii): This is clear from Fact 4.16(ii). ■

Remark 4.18 Consider Corollary 4.17. In general, antiderivatives are neither differentiable nor strictly convex. However, if for a particularly nice instance $A \in \mathcal{A}$ both m_A and $m_{A^{-1}}$ are differentiable, then we deduce from Fact 4.16(iii) that \mathfrak{m}_A and \mathfrak{m}_A^* are both differentiable and strictly convex. Analogous comments can be made for other symmetrizations of (not necessarily intrinsic) methods based on the proximal midpoint average.

We are now able to provide two examples of primal-dual symmetric intrinsic methods with very nice properties. These are in striking contrast to Proposition 4.7.

Example 4.19 Set $m: \mathcal{A} \rightarrow \Gamma: A \mapsto \max_{(a,a^*) \in \text{gra } A} R_{A,(a,a^*)}$ and $\mathfrak{m}: \mathcal{A} \rightarrow \Gamma: A \mapsto \mathcal{P}(m_A, m_{A^{-1}}^*)$. Then \mathfrak{m} is a primal-dual symmetric intrinsic method and for every $A \in \mathcal{A}$, both \mathfrak{m}_A and \mathfrak{m}_A^* have full domain and are piecewise linear-quadratic antiderivatives of A and A^{-1} , respectively.

Proof. For every $A \in \mathcal{A}$, m_A is a convex polyhedral (hence piecewise linear-quadratic) antiderivative of A with full domain (Example 4.4). The conclusion is now a consequence of Corollary 4.17. ■

Example 4.20 Set $m: \mathcal{A} \rightarrow \Gamma: A \mapsto \frac{1}{n_A} \sum_{(a,a^*) \in \text{gra } A} R_{A,(a,a^*)}$, where n_A is the number of points in $\text{gra } A$, and $\mathfrak{m}: \mathcal{A} \rightarrow \Gamma: A \mapsto \mathcal{P}(m_A, m_{A^{-1}}^*)$. Then \mathfrak{m} is a primal-dual symmetric intrinsic method, and for every $A \in \mathfrak{m}$, both \mathfrak{m}_A and \mathfrak{m}_A^* have full domain and are piecewise linear-quadratic antiderivatives of A and A^{-1} , respectively.

Proof. For every $A \in \mathcal{A}$, m_A is a convex polyhedral (hence piecewise linear-quadratic) antiderivative of A with full domain (Example 4.5). The conclusion follows from Corollary 4.17. ■

In the remainder of this section, we aim to extract further nice properties enjoyed by these two methods. We require the following results on the proximal midpoint average.

Proposition 4.21 Let $f_0 \in \Gamma$, let $f_1 \in \Gamma$, and set $f := \mathcal{P}(f_0, f_1)$. Suppose that $a^* \in \partial f_0(a) \cap \partial f_1(a)$. Then

$$(\forall x \in N_{\text{dom } f_0}(a) \cap N_{\text{dom } f_1^*}(a^*)) \quad x + a^* \in \partial f(x + a). \quad (113)$$

Proof. Take $x \in N_{\text{dom } f_0}(a) \cap N_{\text{dom } f_1^*}(a^*)$. Proposition 4.9 yields

$$a = \text{Prox}(f_0)(2x + a + a^*). \quad (114)$$

The same result (applied to f_1^*) shows that $a^* = \text{Prox}(f_1^*)(2x + a^* + a)$, which is equivalent to $a^* = (\text{Id} - \text{Prox}(f_1^*))(2x + a + a^*)$, i.e., to

$$2x + a = \text{Prox}(f_1)(2x + a + a^*). \quad (115)$$

Add (114) and (115), divide the result by 2, and recall Fact 4.11(iv) to deduce that

$$x + a = \text{Prox}(f)((x + a) + (x + a^*)). \quad (116)$$

The conclusion now follows from (102). \blacksquare

Theorem 4.22 *Let $f_0 \in \Gamma$, let $f_1 \in \Gamma$, and set $f := \mathcal{P}(f_0, f_1)$. Suppose that $a^* \in \partial f_0(a) \cap \partial f_1(a)$ and set $N := (N_{\text{dom } f_0}(a) \cap N_{\text{dom } f_1^*}(a^*)) \cup (N_{\text{dom } f_1}(a) \cap N_{\text{dom } f_0^*}(a^*))$. Then the following hold.*

- (i) $(\forall y \in a + N) y + a^* - a \in \partial f(y)$.
- (ii) f is differentiable at every point $y \in a + \text{int } N$ with $\nabla f(y) = y + a^* - a$.

Proof. Proposition 4.21 implies (i). On $a + \text{int } N$, we note that $y \mapsto y + a^* - a$ is a continuous selection of ∂f ; therefore, $\nabla f(y) = y + a^* - a$ by [18, Proposition 2.8] and (ii) holds. \blacksquare

Corollary 4.23 *Let $m: \mathcal{A} \rightarrow \Gamma: A \mapsto m_A$ be an intrinsic method such that for every $A \in \mathcal{A}$, $\text{ran } \partial m_A \subseteq \text{conv ran } A$. Set*

$$\mathbf{m}: \mathcal{A} \rightarrow \Gamma: A \mapsto \mathcal{P}(m_A, m_{A^{-1}}^*), \quad (117)$$

take $A \in \mathcal{A}$, take $(a, a^) \in \text{gra } A$, and set $N := N_{\text{conv dom } A}(a) \cap N_{\text{conv ran } A}(a^*)$. Then the following hold.*

- (i) $(\forall y \in a + N) y + a^* - a \in \partial \mathbf{m}_A(y)$.
- (ii) $(\forall y \in a + \text{int } N) \mathbf{m}_A$ is differentiable at y with $\nabla \mathbf{m}_A(y) = y + a^* - a$.

Proof. Set $f_0 := m_A$ and set $f_1 := m_{A^{-1}}^*$ so that $\mathbf{m}_A = \mathcal{P}(f_0, f_1)$. Proposition 3.2 implies that $\text{dom } f_0^* = \text{dom } m_A^* = \text{conv ran } A$ and that $\text{dom } f_1 = \text{dom } m_{A^{-1}}^* = \text{conv ran } A^{-1} = \text{conv dom } A$. The conclusion is therefore a consequence of Theorem 4.22. \blacksquare

Remark 4.24 In view of Example 4.4 and Example 4.5, we observe that Corollary 4.23 is applicable when m is either as in Example 4.19 or as in Example 4.20.

Example 4.25 Let m and \mathbf{m} be as in Corollary 4.23, let $(a, a^*) \in X \times X$, suppose that $A: X \rightrightarrows X$ is given by $\text{gra } A = \{(a, a^*)\}$. Then there exists $\gamma \in \mathbb{R}$ such that $\mathbf{m}_A = \frac{1}{2}\|\cdot\|^2 + \langle \cdot, a^* - a \rangle + \gamma$.

Proof. Indeed, the set N in Corollary 4.23 is the entire space X and hence item (ii) of that result implies that $\nabla \mathbf{m}_A: X \rightarrow X: x \mapsto x + a^* - a$. \blacksquare

We observe next that on the real line, the subdifferential extending A is actually single-valued — i.e., it corresponds to a gradient — with slope one outside the box $(\text{conv dom } A) \times (\text{conv ran } A)$.

Corollary 4.26 *Suppose that $X = \mathbb{R}$, let m and \mathbf{m} be as in Corollary 4.23, and let $A: \mathbb{R} \rightrightarrows \mathbb{R}$ have finite graph $\{(a_1, a_1^*), \dots, (a_n, a_n^*)\}$, where $a_1 \leq a_2 \leq \dots \leq a_n$ and $a_1^* \leq a_2^* \leq \dots \leq a_n^*$. Then*

$$(\forall x < a_1) \quad \mathbf{m}'_A(x) = x - a_1 + a_1^* \quad \text{and} \quad (\forall x > a_n) \quad \mathbf{m}'_A(x) = x - a_n + a_n^*. \quad (118)$$

Proof. Since $\text{conv dom } A = [a_1, a_n]$ and $\text{conv ran } A = [a_1^*, a_n^*]$, the result follows from Corollary 4.23(ii) (applied at (a_1, a_1^*) and at (a_n, a_n^*)). ■

Remark 4.27 Primal-dual symmetry and the “slope one” property of the extension of the cyclically monotone operator A in Corollary 4.26 were properties deemed desirable by R. T. Rockafellar. In view of Remark 4.24, there exist two explicit methods that generate antiderivatives with these desirable properties. Although not the product of an intrinsic method, the function $f_{A,(a,a^*)}$ of Theorem 4.15 has the same properties.

We conclude this paper by numerically illustrating an antiderivative produced by the primal-dual symmetric intrinsic method of Example 4.19.

Example 4.28 Define $A: \mathbb{R} \rightrightarrows \mathbb{R}$ via $\text{gra } A := \{(a, \exp(a)) \mid a \in \{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}\}$. Because $\text{gra } A \subset \text{gra}(\exp)$, the operator A is cyclically monotone. We interpret A as a 5-point sample of the gradient of the exponential function. Let m and \mathbf{m} be as in Example 4.19. Figure 1 visualizes the exponential function, the antiderivative m_A , the antiderivative $m_{A^{-1}}^*$, and the antiderivative \mathbf{m}_A produced by the primal-dual symmetric intrinsic method \mathbf{m} . As predicted by Corollary 3.20, the function m_A is nonnegative on $\text{dom } A$. In Figure 2, we visualized the derivative of the exponential function, its 5-point sample corresponding to A , and the maximal cyclically monotone extension $\partial \mathbf{m}_A$. Note that by Theorem 3.14 the Rockafellar functions are piecewise linear and hence their subdifferential operators have a “stair case” graph. On the other hand, \mathbf{m}_A is piecewise linear-quadratic and its subdifferential operator displays the “slope one” property guaranteed by Corollary 4.26 outside the rectangle $\text{dom } A \times \text{ran } A$. Both plots were generated in Scilab utilizing software packages discussed in [15]; further details on the numerical implementation will appear elsewhere.

Acknowledgment

The authors thank Jean-Baptiste Hiriart-Urruty for making them aware of [14], and two referees for their pertinent comments. Heinz Bauschke was partially supported by the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chair Program. Yves Lucet was partially supported by the Natural Sciences and Engineering Research Council of Canada. Xianfu Wang was partially supported by the Natural Sciences and Engineering Research Council of Canada.

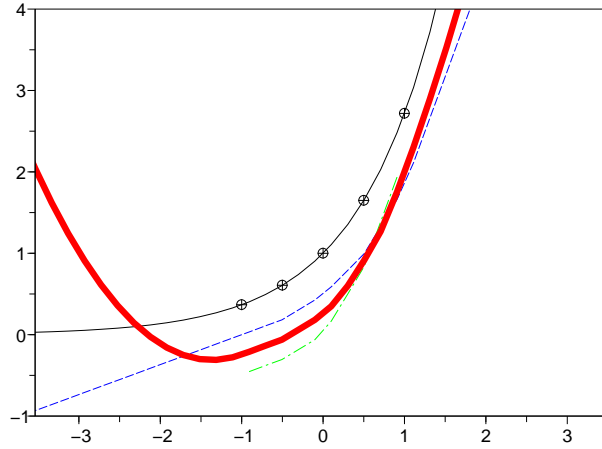


Figure 1: The graph of the exponential function (thin black curve) and the 5 points (circled) on its graph that led to the operator A , the antiderivative m_A (dashed blue curve), the antiderivative $m_{A^{-1}}^*$ (dashed-dotted green curve), and the proximal-average based antiderivative \mathbf{m}_A (thick red curve) are shown. Note that $m_A \geq 0$ on $\text{dom } A$, in accordance with Corollary 3.20.

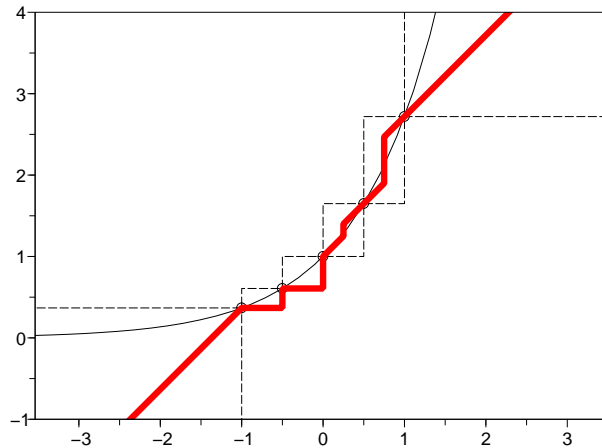


Figure 2: The finite graph operator A is shown as points (circled) on the graph of the exponential function (thin black curve), which is the same as its derivative. The reconstructed subdifferential operator $\partial \mathbf{m}_A$ (thick red curve) stays inside the rectangles (dashed) imposed on any monotone extension of A . Note the “slope one” property of $\partial \mathbf{m}_A$ outside the rectangle $\text{conv dom } A \times \text{conv ran } A$, as guaranteed by Corollary 4.26.

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] S. Bartz, H. H. Bauschke, J. M. Borwein, S. Reich, and X. Wang, “Fitzpatrick functions, cyclic monotonicity, and Rockafellar’s antiderivative”, *Nonlinear Analysis*, vol. 66, pp. 1198–1223, 2007.
- [3] H. H. Bauschke, J. M. Borwein, and X. Wang, “Fitzpatrick functions and continuous linear monotone operators”, to appear in *SIAM Journal on Optimization*.
- [4] H. H. Bauschke, R. Goebel, Y. Lucet, and X. Wang, “The proximal average: basic theory”, submitted.
- [5] H. H. Bauschke, Y. Lucet, and M. Trienis, “How to transform one convex function continuously into another”, to appear in *SIAM Review*.
- [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 and X. Wang, “A convex-analytical approach to extension results for n -cyclically monotone operators”, to appear in *Set-Valued Analysis*.
- [8] H. H. Bauschke and X. Wang, “The kernel average for two convex functions and its application to the extension and representation of monotone operators”, submitted. Preprint available at http://www.optimization-online.org/DB_HTML/2007/05/1658.html
- [9] J. Benoist and A. Daniilidis, “Subdifferential representation of convex functions: refinements and applications”, *Journal of Convex Analysis*, vol. 12, pp. 255–265, 2005.
- [10] J. M. Borwein, “Maximal monotonicity via convex analysis”, *Journal of Convex Analysis*, vol. 13, pp. 561–586, 2006.
- [11] J. M. Borwein and Q. J. Zhu, *Techniques of Variational Analysis*, Springer-Verlag, 2005.
- [12] H. Brézis, *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*, North-Holland, 1973.
- [13] P. Choné and H. V. J. Le Meur, “Non-convergence result for conformal approximation of variational problems subject to a convexity constraint”, *Numerical Functional Analysis and Optimization*, vol. 22, pp. 529–547, 2001.
- [14] D. Lambert, J.-P. Crouzeix, V. H. Nguyen, and J.-J. Strodiot, “Finite convex integration”, *Journal of Convex Analysis*, vol. 11, pp. 131–146, 2004.
- [15] Y. Lucet, H. H. Bauschke, and M. Trienis, “The piecewise linear-quadratic model for computational convex analysis”, to appear in *Computational Optimization and Applications*.

- [16] G. J. Minty, “On the maximal domain of a monotone function”, *Michigan Mathematical Journal*, vol. 8, pp. 135–137, 1961.
- [17] J.-J. Moreau, “Proximité et dualité dans un espace hilbertien”, *Bulletin de la Société Mathématique de France*, vol. 93, pp. 273–299, 1965.
- [18] R. R. Phelps, *Convex Functions, Monotone Operators, and Differentiability* (second edition), Lecture Notes in Mathematics, vol. 1364, Springer-Verlag, 1993.
- [19] R. T. Rockafellar, “On the maximal monotonicity of subdifferential mappings”, *Pacific Journal of Mathematics*, vol. 33, pp. 209–216, 1970.
- [20] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.
- [21] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, 1998.
- [22] S. Simons, *Minimax and Monotonicity*, Lecture Notes in Mathematics, vol. 1693, Springer-Verlag, 1998.
- [23] M. D. Voisei, “Extension theorems for k -monotone operators,” *Studii și Cercetări Științifice, Seria: Matematică, Universitatea din Bacău*, vol. 9, pp. 235–242, 1999.
- [24] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific Publishing, 2002.