Generalized monotone operators and their averaged resolvents

Heinz H. Bauschke¹ · Walaa M. Moursi² · Xianfu Wang¹

Received: 7 October 2019 / Accepted: 25 March 2020 / Published online: 25 April 2020
© Springer-Verlag GmbH Germany, part of Springer Nature and Mathematical Optimization Society 2020

Abstract
The correspondence between the monotonicity of a (possibly) set-valued operator and the firm nonexpansiveness of its resolvent is a key ingredient in the convergence analysis of many optimization algorithms. Firmly nonexpansive operators form a proper subclass of the more general—but still pleasant from an algorithmic perspective—class of averaged operators. In this paper, we introduce the new notion of conically nonexpansive operators which generalize nonexpansive mappings. We characterize averaged operators as being resolvents of comonotone operators under appropriate scaling. As a consequence, we characterize the proximal point mappings associated with hypoconvex functions as cocoercive operators, or equivalently; as displacement mappings of conically nonexpansive operators. Several examples illustrate our analysis and demonstrate tightness of our results.

Keywords Averaged operator · Cocoercive operator · Firmly nonexpansive mapping · Hypoconvex function · Maximally monotone operator · Nonexpansive mapping · Proximal operator

Mathematics Subject Classification Primary 47H05 · 47H09; Secondary 49N15 · 90C25

Dedicated to Marco López on the occasion of his 70th birthday.

¹ Mathematics, University of British Columbia, Kelowna, BC V1V 1V7, Canada
² Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON N2L 3G1, Canada
1 Introduction

In this paper, we assume that $X$ is a real Hilbert space, with inner product $\langle \cdot , \cdot \rangle$ and induced norm $\| \cdot \|$. Monotone operators form a beautiful class of operators that play a crucial role in modern optimization. This class includes subdifferential operators of proper lower semicontinuous convex functions as well as matrices with positive semidefinite symmetric part. (For detailed discussions on monotone operators and the connection to optimization problems, we refer the reader to [2,5–7,10,11,23,27–29,33,34], and the references therein.)

The correspondence between the maximal monotonicity of an operator and the firm nonexpansiveness of its resolvent is of central importance from an algorithmic perspective: to find a critical point of the former, iterate the latter!

Indeed, firmly nonexpansive operators belong to the more general and pleasant class of averaged operators (see Definition 2.1(ii)). Let $x_0 \in X$ and let $T : X \to X$ be averaged. Thanks to the Krasnosel’ski–Mann iteration (see [19,20] and also [2, Theorem 5.14]), the sequence $(T^n x_0)_{n \in \mathbb{N}}$ converges weakly to a fixed point of $T$. When $T$ is the proximal mapping associated with a proper lower semicontinuous convex function $f$, the set of fixed points of $T$ is the set of critical point of $f$; equivalently the set of minimizers of $f$. In fact, iterating $T$ in this case produces the famous proximal point algorithm, see [26]. Generalizations of the proximal point algorithm in the absence of monotonicity have been studied in, e.g., [12,17,25]. The main goal of this paper is to answer the question: Can we explore a new correspondence between a set-valued operator and its resolvent which generalizes the fundamental correspondence between monotone operators and firmly nonexpansive mappings (see Fact 2.2)? Our approach relies on the new notion of conically nonexpansive operators as well as the notions of $\rho$-monotonicity (respectively $\rho$-comonotonicity) for set-valued operators (see Definition 2.4 below) which, depending on the value of $\rho$, correspond to strong monotonicity, monotonicity or hypomonotonicity (respectively cocoercivity, monotonicity or cohypomonotonicity).

Although some correspondences between a monotone operator ($\rho \geq 0$) and its resolvent have been established in [3], our analysis here not only provides more quantifications but also goes beyond monotone operators. However, our proofs sometimes follow classical lines or rely on classical counterparts. We now summarize the three main results of this paper:

R1 We show that when $\rho > -1$, the resolvent of a $\rho$-monotone operator as well as the resolvent of its inverse are single-valued and both have full domain. This allows us to extend the classical theorem by Minty (see Fact 2.3) to this class of operators (see Theorem 2.17).

R2 We characterize conically nonexpansive operators (respectively averaged operators and nonexpansive operators) as resolvents of $\rho$-comonotone operators with $\rho > -1$ (respectively $\rho > -\frac{1}{2}$ and $\rho \geq -\frac{1}{2}$) (see Corollary 3.8 and also Table 1).
R3 As a consequence of R2, we obtain a novel characterization of the proximal point mapping associated with a hypoconvex function1 (under appropriate scaling of the function) as a cocoercive operator; or, equivalently, as the displacement mapping of a conically nonexpansive mapping (see Theorem 6.5).

The remainder of this paper is organized as follows. Section 2 is devoted to the study of the properties of ρ-monotone and ρ-comonotone operators. In Sect. 3, we provide a characterization of averaged operators as resolvents of ρ-comonotone operators. Section 4 provides useful correspondences between an operator and its resolvent as well as its reflected resolvent. In Sect. 5, we focus on ρ-monotone and ρ-comonotone linear operators. In the final Sect. 6, we establish the connection to hypoconvex functions.

The notation we use is standard and follows, e.g., [2] or [27].

2 ρ-monotone and ρ-comonotone operators

Let A : X ⇀ X. Recall that the resolvent of A is \( J_A = (\text{Id} + A)^{-1} \) and the reflected resolvent of A is \( R_A = 2J_A - \text{Id} \), where \( \text{Id} : X \to X \). The graph of A is \( \text{gra} A = \left\{ (x, u) \in X \times X \mid u \in Ax \right\} \). It is easy to check that

\[
(\forall (x, u) \in X \times X) \quad (x, u) \in \text{gra} J_A \iff (u, x - u) \in \text{gra} A
\]  

(1)

and that

\[
(\forall (x, u) \in X \times X) \quad (x, u) \in \text{gra} R_A \iff \left( \frac{1}{2}(x + u), \frac{1}{2}(x - u) \right) \in \text{gra} A.
\]  

(2)

We now recall the following well known definitions:

**Definition 2.1** Let D be a nonempty subset of X, let \( T : D \to X \), and let \( \alpha \in ]0, 1[ \).

(i) \( T \) is nonexpansive if \( (\forall (x, y) \in D \times D) \|Tx - Ty\| \leq \|x - y\| \).

(ii) \( T \) is \( \alpha \)-averaged if \( \alpha \in ]0, 1[ \) and there exists a nonexpansive operator \( N : D \to X \) such that \( T = (1 - \alpha) \text{Id} + \alpha N \); equivalently, if \( (\forall (x, y) \in D \times D) \) we have (see [2, Proposition 4.35])

\[
(1 - \alpha)\|\left( \text{Id} - T \right)x - \left( \text{Id} - T \right)y\|^2 \leq \alpha(\|x - y\|^2 - \|Tx - Ty\|^2).
\]  

(3)

(iii) \( T \) is firmly nonexpansive if \( T \) is \( \frac{1}{2} \)-averaged; equivalently, if \( (\forall (x, y) \in D \times D) \)

\[
\|Tx - Ty\|^2 + \|\left( \text{Id} - T \right)x - \left( \text{Id} - T \right)y\|^2 \leq \|x - y\|^2.
\]

We begin this section by stating the following two useful facts.

**Fact 2.2** (see, e.g., [14, Theorem 2]) Let D be a nonempty subset of X, let \( T : D \to X \), and set \( A = T^{-1} - \text{Id} \). Then \( T = J_A \). Moreover, the following hold:

(i) \( T \) is firmly nonexpansive if and only if A is monotone.

(ii) \( T \) is firmly nonexpansive and \( D = X \) if and only if A is maximally monotone.

---

1 This is also known as weakly convex function.
**Fact 2.3** (Minty’s Theorem) [21] (see also [2, Theorem 21.1]) Let \( A: X \rightrightarrows X \) be monotone. Then

\[
\text{gra } A = \left\{ (J_A x, (\text{Id} - J_A)x) \mid x \in \text{ran}(\text{Id} + A) \right\}.
\]

Moreover,

\( A \) is maximally monotone \( \iff \text{ran}(\text{Id} + A) = X. \)

**Definition 2.4** Let \( A: X \rightrightarrows X \) and let \( \rho \in \mathbb{R} \). We define the following notions:

(i) \( A \) is \( \rho \)-monotone if \( (\forall (x, u) \in \text{gra } A) (\forall (y, v) \in \text{gra } A) \) we have

\[
\langle x - y, u - v \rangle \geq \rho \|x - y\|^2.
\]

(ii) \( A \) is maximally \( \rho \)-monotone if \( A \) is \( \rho \)-monotone and there is no \( \rho \)-monotone operator \( B: X \rightrightarrows X \) such that \( \text{gra } B \) properly contains \( \text{gra } A \), i.e., for every \( (x, u) \in X \times X \),

\[
(x, u) \in \text{gra } A \iff (\forall (y, v) \in \text{gra } A) \langle x - y, u - v \rangle \geq \rho \|x - y\|^2.
\]

(iii) \( A \) is \( \rho \)-comonotone if \( (\forall (x, u) \in \text{gra } A) (\forall (y, v) \in \text{gra } A) \) we have

\[
\langle x - y, u - v \rangle \geq \rho \|u - v\|^2.
\]

(iv) \( A \) is maximally \( \rho \)-comonotone if \( A \) is \( \rho \)-comonotone and there is no \( \rho \)-comonotone operator \( B: X \rightrightarrows X \) such that \( \text{gra } B \) properly contains \( \text{gra } A \), i.e., for every \( (x, u) \in X \times X \),

\[
(x, u) \in \text{gra } A \iff (\forall (y, v) \in \text{gra } A) \langle x - y, u - v \rangle \geq \rho \|u - v\|^2.
\]

We shall say that \( A \) is \( \rho \)-monotone with optimal \( \rho \) if \( A \) is \( \rho \)-monotone and \( A \) is not \( \tilde{\rho} \)-monotone for every \( \tilde{\rho} > \rho \), and similar for the other notions.

**Remark 2.5**

(i) When \( \rho = 0 \), both \( \rho \)-monotonicity of \( A \) and \( \rho \)-comonotonicity of \( A \) reduce to the monotonicity of \( A \); equivalently to the monotonicity of \( A^{-1} \).

(ii) When \( \rho < 0 \), \( \rho \)-monotonicity is known as \( |\rho| \)-hypomonotonicity, see [27, Example 12.28] and [7, Definition 6.9.1]. In this case, the \( \rho \)-comonotonicity is also known as \( |\rho| \)-cohypomonotonicity (see [12, Definition 2.2]).

(iii) In passing, we point out that when \( \rho > 0 \), \( \rho \)-monotonicity of \( A \) reduces to \( \rho \)-strong monotonicity of \( A \), while \( \rho \)-comonotonicity of \( A \) reduces to \( \rho \)-cocoercivity\(^2\) of \( A \).

Unlike classical monotonicity, \( \rho \)-comonotonicity of \( A \) is not equivalent to \( \rho \)-comonotonicity of \( A^{-1} \). Instead, we have the following correspondences.

\(^2\) Let \( \beta > 0 \) and let \( T: X \rightarrow X \). Recall that \( T \) is \( \beta \)-cocoercive if \( \beta T \) is firmly nonexpansive, i.e.,

\[
(\forall (x, y) \in X \times X) \langle x - y, Tx - Ty \rangle \geq \beta \|Tx - Ty\|^2.
\]
Lemma 2.6 Let $A \colon X \rightharpoonup X$ and let $\rho \in \mathbb{R}$. Then the following are equivalent:

(i) $A$ is $\rho$-comonotone.
(ii) $A^{-1} - \rho \text{Id}$ is monotone.
(iii) $A^{-1}$ is $\rho$-monotone.

Proof “(i)$\Rightarrow$(ii)”: Let $\{(x, u), (y, v)\} \subseteq X \times X$. Then $\{(x, u), (y, v)\} \subseteq \text{gra}(A^{-1} - \rho \text{Id}) \iff [u \in A^{-1} x - \rho x \text{ and } v \in A^{-1} y - \rho y] \iff \{(x, u + \rho x), (y, v + \rho y)\} \subseteq \text{gra } A^{-1} \iff \{(u + \rho x, x), (v + \rho y, y)\} \subseteq \text{gra } A \Rightarrow \langle x - y, u - v + \rho(x - y) \rangle \geq \rho \|x - y\|^2 \Rightarrow \rho \|x - y\|^2 + \langle x - y, u - v \rangle \geq \rho \|x - y\|^2 \Rightarrow \langle u - v, x - y \rangle \geq 0$. “(ii)$\Rightarrow$(iii)”: Let $\{(x, u), (y, v)\} \subseteq \text{gra } A^{-1}$. Then $\{(x, u - \rho x), (y, v - \rho y)\} \subseteq \text{gra } (A^{-1} - \rho \text{Id})$. Hence $\langle x - y, u - v - \rho(x - y) \rangle \geq 0$; equivalently, $\langle x - y, u - v \rangle \geq \rho \|x - y\|^2$. “(iii)$\Rightarrow$(i)”: Let $\{(x, u), (y, v)\} \subseteq X \times X$. Then $\{(x, u), (y, v)\} \subseteq \text{gra } A \iff \{(u, x), (v, y)\} \subseteq \text{gra } A^{-1} \Rightarrow \langle x - y, u - v \rangle \geq \rho \|u - v\|^2$. \hfill \Box

Lemma 2.7 Let $A \colon X \rightharpoonup X$ and let $\rho \in \mathbb{R}$. Then the following hold:

(i) $\text{gra } A = \{(u + \rho x, x) \mid (x, u) \in \text{gra } (A^{-1} - \rho \text{Id})\}.
(ii) \text{gra } (A^{-1} - \rho \text{Id}) = \{(u, x - \rho u) \mid (x, u) \in \text{gra } A\}.

Proof (i): Let $(x, u) \in X \times X$. Then $(x, u) \in \text{gra } (A^{-1} - \rho \text{Id}) \iff u \in A^{-1} x - \rho x \iff u + \rho x \in A^{-1} x \iff x \in A(u + \rho x) \iff (u + \rho x, x) \in \text{gra } A$. This proves “$\supseteq$” in (i). The opposite inclusion is proved similarly. (ii): The proof is similar to the one of (i). \hfill \Box

Lemma 2.8 Let $A \colon X \rightharpoonup X$ and let $\rho \in \mathbb{R}$. Then the following are equivalent:

(i) $A$ is maximally $\rho$-comonotone.
(ii) $A^{-1} - \rho \text{Id}$ is maximally monotone.
(iii) $A^{-1}$ is maximally $\rho$-monotone.

Proof Note that Lemma 2.6 implies that $A$ is $\rho$-comonotone $\iff A^{-1} - \rho \text{Id}$ is monotone. Let $(v, y) \in X \times X$. “(i)$\Rightarrow$(ii)”: Then $(v, y)$ is $\rho$-monotonically related to $\text{gra } (A^{-1} - \rho \text{Id})$, i.e., $(\forall (u, x) \in \text{gra } (A^{-1} - \rho \text{Id})) \langle u - v, x - y \rangle \geq 0 \iff (\forall (u, x) \in \text{gra } (A^{-1} - \rho \text{Id})) \langle u - v, x - y \rangle + \rho \|u - v\|^2 \geq \rho \|u - v\|^2 \iff (\forall (u, x) \in \text{gra } (A^{-1} - \rho \text{Id})) \langle u - v, x + pu - (y + \rho v) \rangle \geq \rho \|u - v\|^2$. Because the last inequality holds for all $(u, x) \in \text{gra } (A^{-1} - \rho \text{Id})$, the parametrization of $\text{gra } A$ given in Lemma 2.7(i) and the maximal $\rho$-comonotonicity of $A$ imply that $\langle y + \rho v, v \rangle \in \text{gra } A$. Therefore, by Lemma 2.7(ii), $(v, y) \in \text{gra } (A^{-1} - \rho \text{Id})$.

“(ii)$\Leftarrow$(i)”: Then $(v, y)$ is $\rho$-comonotonoically related to $\text{gra } A$, i.e., $(\forall (x, u) \in \text{gra } A) \langle x - y, u - v \rangle \geq \rho \|u - v\|^2 \iff (\forall (x, u) \in \text{gra } A) \langle x - \rho u - (y - \rho v), u - v \rangle \geq 0$. It follows from Lemma 2.7(ii) and the maximal monotonicity of $A^{-1} - \rho \text{Id}$ that $(v, y - \rho v) \in \text{gra } (A^{-1} - \rho \text{Id})$; equivalently, using Lemma 2.7(ii), $(v, y) \in \text{gra } A$.

“(ii)$\Rightarrow$(iii)”: This follow from the equivalences $(\forall (u, x) \in \text{gra } A^{-1}) \langle v - u, x - y \rangle \geq \rho \|v - u\|^2 \iff (\forall (u, x) \in \text{gra } A^{-1}) \langle v - u, (y - \rho v) - (x - \rho u) \rangle \geq 0 \iff (\forall (u, x') \in \text{gra } (A^{-1} - \rho \text{Id})) \langle v - u, (\rho v - x') \rangle \geq 0$ and Lemma 2.7(ii). \hfill \Box

Remark 2.9 Note that when $\rho < 0$, the (maximal) monotonicity of $A^{-1} - \rho \text{Id}$ is equivalent to the (maximal) monotonicity of the Yosida approximation $(A^{-1} - \rho \text{Id})^{-1}$. Such a characterization is presented in [7, Proposition 6.9.3].

\footnote{Springer}
Proposition 2.10  Let \( A : X \rightrightarrows X \) be maximally \( \rho \)-comonotone for some \( \rho > -1 \). Then \( \text{ran}(\text{Id} + A^{-1}) = X \).

Proof By Lemma 2.8, \( A^{-1} - \rho \text{Id} \) is maximally monotone. Consequently, because \( 1 + \rho > 0 \), the operator \( \frac{1}{1+\rho}(A^{-1} - \rho \text{Id}) \) is maximally monotone. Applying (5) to \( \frac{1}{1+\rho}(A^{-1} - \rho \text{Id}) \), we have \( \text{ran}(\text{Id} + A^{-1}) = \text{ran}((1 + \rho) \text{Id} + (A^{-1} - \rho \text{Id})) = (1 + \rho) \text{ran}(\text{Id} + \frac{1}{1+\rho}(A^{-1} - \rho \text{Id})) = (1 + \rho)X = X \). \( \square \)

Proposition 2.11  Let \( A : X \rightrightarrows X \). Then the following hold:

(i) \( J_{A^{-1}} = \text{Id} - J_A \).
(ii) \( \text{ran}(\text{Id} + A^{-1}) = \text{dom}(J_A) = \text{ran}(\text{Id} + A) \).

Proof (i): This follows from [2, Proposition 23.7(ii) and Definition 23.1]. (ii): Using (i), we have \( \text{ran}(\text{Id} + A^{-1}) = \text{dom}(\text{Id} + A^{-1})^{-1} = \text{dom} J_{A^{-1}} = \text{dom}(\text{Id} - J_A) = (\text{dom} \text{Id}) \cap (\text{dom} J_A) = \text{dom} J_A = \text{ran}(\text{Id} + A) \). \( \square \)

Corollary 2.12  (surjectivity of \( \text{Id} + A \) and \( \text{Id} + A^{-1} \)) Let \( A : X \rightrightarrows X \) be maximally \( \rho \)-comonotone for some \( \rho > -1 \). Then

\[
\text{dom} J_A = \text{ran}(\text{Id} + A) = \text{dom}(\text{Id} - J_A) = \text{ran}(\text{Id} + A^{-1}) = X. \tag{10}
\]

Proof Combine Propositions 2.10 and 2.11. \( \square \)

Proposition 2.13  (resolvents that are at most single-valued) Let \( A : X \rightrightarrows X \) be \( \rho \)-comonotone for some \( \rho > -1 \). Then \( J_A = (\text{Id} + A)^{-1} \) and \( J_{A^{-1}} = \text{Id} - J_A \) are at most single-valued.

Proof Let \( x \in \text{dom} J_A = \text{ran}(\text{Id} + A) \) and let \( (u, v) \in X \times X \). Then \( \{u, v\} \subseteq J_A x \iff [x - u \in Au \text{ and } x - v \in Av] \iff \langle x - u, x - v \rangle \geq \rho \|u - v\|^2 \iff -\|u - v\|^2 \geq \rho \|u - v\|^2 \). Since \( \rho > -1 \), the last inequality implies that \( u = v \). Now combine with Proposition 2.11(i). \( \square \)

Corollary 2.14  (See also [13, Proposition 3.4]) Let \( A : X \rightrightarrows X \) be maximally \( \rho \)-comonotone for some \( \rho > -1 \). Then \( J_A = (\text{Id} + A)^{-1} \) and \( J_{A^{-1}} = \text{Id} - J_A \) are single-valued and \( \text{dom} J_A = \text{dom} J_{A^{-1}} = X \).

Proof Combine Corollary 2.12 with Proposition 2.13. \( \square \)

In Example 2.15 below, we illustrate that the assumption that \( \rho > -1 \) is critical in the conclusions of Corollary 2.12 and Proposition 2.13.

Example 2.15  Suppose that \( X \neq \{0\} \). Let \( C \) be a nonempty closed convex subset of \( X \) with associated projector \( P_C \), let \( r \in \mathbb{R}_+ \), set \( B = -\text{Id} - rP_C \), set \( A = B^{-1} \) and set \( \rho = -(1 + r) \leq -1 \). Then the following hold:

(i) \( B - \rho \text{Id} \) is maximally monotone.
(ii) \( A \) is maximally \( \rho \)-comonotone.
(iii) \( \text{ran}(\text{Id} + A) = \text{ran}(\text{Id} + A^{-1}) = (\rho + 1)C = -rC \).
(iv) \( \text{Id} + A \) is surjective \( \iff [C = X \text{ and } r > 0] \).
(v) \( J_A \) is at most single-valued \( \iff J_{A^{-1}} \) is at most single-valued \( \iff [C = X \text{ and } r > 0] \).

**Proof** (i): Indeed, \( B - \rho \text{Id} = -\text{Id} - r P_C + (1 + r) \text{Id} = r (\text{Id} - P_C) \). It follows from [2, Example 23.4 & Proposition 23.11(i)] that \( \text{Id} - P_C \) is maximally monotone. Moreover, \( A \) is maximally \( \rho \)-maximal if
\[
\langle u - v, - A y \rangle = - \langle u - v, \text{Id} + A\rangle 
\]
for some \( \rho > 0 \).

(ii): Combine (i) and Lemma 2.8.

(iii): The first identity follows from Proposition 2.11(ii). Now \( \text{ran(}\text{Id} + A^{-1}\rangle = \text{ran(}\text{Id} + B\rangle = \text{ran(}-r P_C\rangle = -r \text{ ran } P_C = -r C = (\rho + 1)C \).

(iv): This is a direct consequence of (iii).

(v): The first equivalence follows from Proposition 2.11(i). Note that \( [r = 0 \text{ or } C = [0]] \iff r C = [0] \iff r P_C \equiv 0 \iff B = - \text{Id} \iff \text{gra } J_{A^{-1}} = \text{gra } J_B = [0] \times X \).

Now suppose that \( r > 0 \). Then \( J_{A^{-1}} = J_B = (\text{Id} + B)^{-1} = (-r P_C)^{-1} = (\text{Id} + N_C) \circ (-r^{-1} \text{Id}) \). Therefore, \( J_{A^{-1}} \) is at most single-valued \( \iff C = X \) by, e.g., [2, Theorem 7.4].

**Proposition 2.16** Let \( A: X \rightrightarrows X \) be \( \rho \)-comonotone for some \( \rho > -1 \), and suppose that \( \text{ran(}\text{Id} + A\rangle = X \). Then \( A \) is maximally \( \rho \)-comonotone.

**Proof** Let \( (x, u) \in X \times X \) such that \( (\forall (y, v) \in \text{gra } A) \)
\[
\langle x - y, u - v \rangle \geq \rho \|u - v\|^2. \tag{11}
\]
It follows from the surjectivity of \( \text{Id} + A \) that there exists \( (y, v) \in X \times X \) such that \( v \in Ay \) and \( x + u = y + v \in (\text{Id} + A)y \). Consequently, (11) implies that \( \rho \|u - v\|^2 \leq \langle x - y, u - v \rangle = (-\langle u - v, u - v \rangle) = -\|u - v\|^2 \). Hence, because \( \rho > -1 \), we have \( u = v \) and thus \( x = y \) which proves the maximality of \( A \).

**Theorem 2.17** (Minty parametrization) Let \( A: X \rightrightarrows X \) be \( \rho \)-comonotone for some \( \rho > -1 \). Then
\[
\text{gra } A = \{ (J_A x, (\text{Id} - J_A)x) \mid x \in \text{ran(}\text{Id} + A\rangle \}. \tag{12}
\]
Moreover, \( A \) is maximally \( \rho \)-comonotone \( \iff \text{ran(}\text{Id} + A\rangle = X \), in which case
\[
\text{gra } A = \{ (J_A x, (\text{Id} - J_A)x) \mid x \in X \}. \tag{13}
\]

**Proof** Let \( (x, u) \in X \times X \). In view of Proposition 2.13, we have: \( (x, u) \in \text{gra } A \iff u \in Ax \iff x + u \in x + Ax = (\text{Id} + A)x \iff x = J_A (x + u) \iff [z := x + u \in \text{ran(}\text{Id} + A\rangle, x = J_A z \text{ and } u = x + u - x = x + u - J_A (x + u) = (\text{Id} - J_A)z] \). The equivalence of maximal \( \rho \)-comonotonicity of \( A \) and the surjectivity of \( \text{Id} + A \) follows from combining Corollary 2.12 and Proposition 2.16. □
3 $\rho$-comonotonicity and averagedness

We start this section with the following definition.

**Definition 3.1** Let $D$ be a nonempty subset of $X$, let $T : D \to X$, and let $\alpha \in [0, +\infty[$. Then $T$ is $\alpha$-conically nonexpansive (on $D$) if there exists a nonexpansive operator $N : D \to X$ such that $T = (1 - \alpha) \text{Id} + \alpha N$.

**Remark 3.2** In view of Definition 3.1, it is clear that $T$ is $\alpha$-averaged if and only if $T$ is both $\alpha$-conically nonexpansive and $\alpha \in ]0, 1[$. Similarly, $T$ is nonexpansive if and only if $T$ is $1$-conically nonexpansive.

One can directly verify the following result.

**Lemma 3.3** Let $(x, y) \in X \times X$ and let $\alpha \in \mathbb{R}$. Then

$$\alpha^2 \|x\|^2 - \|(\alpha - 1)x + y\|^2 = 2\alpha (x - y, y) - (1 - 2\alpha)\|x - y\|^2. \quad (14)$$

**Lemma 3.4** Let $D$ be a nonempty subset of $X$, let $N : D \to X$, let $\alpha \in \mathbb{R}$ and set $T = (1 - \alpha) \text{Id} + \alpha N$. Then $N$ is nonexpansive if and only if $(\forall (x, y) \in D \times D)$ we have

$$2\alpha (Tx - Ty, (\text{Id} - T)x - (\text{Id} - T)y) \geq (1 - 2\alpha)\|(\text{Id} - T)x - (\text{Id} - T)y\|^2. \quad (15)$$

**Proof** Let $(x, y) \in D \times D$. Applying Lemma 3.3 with $(x, y)$ replaced by $(x - y, Tx - Ty)$, we learn that

$$2\alpha (Tx - Ty, (\text{Id} - T)x - (\text{Id} - T)y) - (1 - 2\alpha)\|(\text{Id} - T)x - (\text{Id} - T)y\|^2 \quad (16a)$$

$$= \alpha^2 \|x - y\|^2 - \|(\alpha - 1)(x - y) + (1 - \alpha)(x - y) + \alpha(Nx - Ny)\|^2 \quad (16b)$$

$$= \alpha^2 \|x - y\|^2 - \|Nx - Ny\|^2. \quad (16c)$$

Now $N$ is nonexpansive $\iff \|x - y\|^2 - \|Nx - Ny\|^2 \geq 0$ holds for all $(x, y) \in D \times D$, and the conclusion follows. $\square$

We now provide various characterizations.

**Corollary 3.5** Let $D$ be a nonempty subset of $X$, let $T : D \to X$, and let $\alpha \in ]0, +\infty[$. Then the following hold:

(i) $T$ is nonexpansive $\iff (\forall (x, y) \in D \times D)$ \(2\alpha (Tx - Ty, (\text{Id} - T)x - (\text{Id} - T)y) \geq -\|(\text{Id} - T)x - (\text{Id} - T)y\|^2. \)

(ii) $T$ is $\alpha$-conically nonexpansive $\iff (\forall (x, y) \in D \times D)$ \(2\alpha (Tx - Ty, (\text{Id} - T)x - (\text{Id} - T)y) \geq (1 - 2\alpha)\|(\text{Id} - T)x - (\text{Id} - T)y\|^2. \)

(iii) $T$ is $\alpha$-conically nonexpansive $\iff \text{Id} - T$ is $\frac{1}{2\alpha}$-cocoercive.

**Proof** (i): Apply Lemma 3.4 with $\alpha = 1$. (ii): A direct consequence of Lemma 3.4. (iii): This follows easily from (ii). $\square$
Proposition 3.6 Let $D$ be a nonempty subset of $X$, let $T : D \to X$, let $\alpha \in ]0, +\infty[$, set $A = T^{-1} - \text{Id}$ and set $N = \frac{1}{\alpha} T - \frac{1-\alpha}{\alpha} \text{Id}$, i.e., $T = J_A = (1-\alpha) \text{Id} + \alpha N$. Then the following hold:

(i) $T$ is $\alpha$-conically nonexpansive $\iff N$ is nonexpansive $\iff A$ is $(\frac{1}{2\alpha} - 1)$-comonotone.

(ii) $[T$ is $\alpha$-conically nonexpansive and $D = X] \iff [N$ is nonexpansive and $D = X] \iff A$ is maximally $(\frac{1}{2\alpha} - 1)$-comonotone.

Proof (i): The first equivalence is Definition 3.1. We now turn to the second equivalence. “$\Rightarrow$”: Let $\{(x, u), (y, v)\} \subseteq \text{gra} A$. Then $(x, u) = (T(x+u), (\text{Id} - T)(x+u))$ and likewise $(y, v) = (T(y+v), (\text{Id} - T)(y+v))$. It follows from Lemma 3.4 applied with $(x, y)$ replaced by $(x+u, y+v)$ that $2\alpha \langle x - y, u - v \rangle \geq (1 - 2\alpha) \|u - v\|^2$. Since $\alpha > 0$, the conclusion follows by dividing both sides of the last inequality by $2\alpha$. “$\Leftarrow$”: Using Theorem 2.17, we learn that $(\forall (x, y) \in D \times D) \{(Tx, (\text{Id} - T)x), (Ty, (\text{Id} - T)y)\} \subseteq \text{gra} A$ and hence $\langle Tx - Ty, (\text{Id} - T)x - (\text{Id} - T)y \rangle \geq (\frac{1}{\alpha} - 1) \|\text{Id} - T\| x - (\text{Id} - T)y \| 2$, i.e., $2\alpha \langle Tx - Ty, (\text{Id} - T)x - (\text{Id} - T)y \rangle \geq (1 - 2\alpha) \|\text{Id} - T\| x - (\text{Id} - T)y \| 2$. Now use Lemma 3.4.

(ii): Note that dom $N = \text{dom} T = \text{ran} T^{-1} = \text{ran}(\text{Id} + A)$. Now combine (i) and Theorem 2.17. \hfill $\Box$

Proposition 3.7 Let $D$ be a nonempty subset of $X$, let $T : D \to X$, let $\alpha \in ]0, +\infty[$, set $A = T^{-1} - \text{Id}$, i.e., $T = J_A$, and set $\rho = \frac{1}{2\alpha} - 1 > -1$. Then the following equivalences hold:

(i) $T$ is $\alpha$-conically nonexpansive $\iff A$ is $\rho$-comonotone.

(ii) $[T$ is $\alpha$-conically nonexpansive and $D = X] \iff A$ is maximally $\rho$-comonotone.

(iii) $T$ is nonexpansive $\iff A$ is $(-\frac{1}{2})$-comonotone.

(iv) $[T$ is nonexpansive and $D = X] \iff A$ is maximally $(-\frac{1}{2})$-comonotone.

If we assume that $\alpha \in ]0, 1[$, equivalently, $\rho > -\frac{1}{2}$, then we additionally have:

(v) $T$ is $\alpha$-averaged $\iff A$ is $\rho$-comonotone.

(vi) $[T$ is $\alpha$-averaged and $D = X] \iff A$ is maximally $\rho$-comonotone.

Proof (i) & (ii): This follows from Proposition 3.6. (iii)–(iv): Combine (i) and (ii) with Remark 3.2. \hfill $\Box$

Corollary 3.8 (The characterization corollary).

Let $T : X \to X$. Then the following hold:

(i) $T$ is nonexpansive if and only if it is the resolvent of a maximally $(-\frac{1}{2})$-comonotone operator $A : X \rightrightarrows X$.

(ii) Let $\alpha \in ]0, +\infty[$. Then $T$ is $\alpha$-conically nonexpansive if and only if it is the resolvent of a maximally $\rho$-comonotone operator $A : X \rightrightarrows X$, where $\rho = \frac{1}{2\alpha} - 1 > -1$ (i.e., $\alpha = \frac{1}{2(\rho+1)}$).

(iii) Let $\alpha \in ]0, 1[$. Then $T$ is $\alpha$-averaged if and only if it is the resolvent of a maximally $\rho$-comonotone operator $A : X \rightrightarrows X$ where $\rho = \frac{1}{2\alpha} - 1 > -\frac{1}{2}$ (i.e., $\alpha = \frac{1}{2(\rho+1)}$).
Example 3.9 Suppose that $U$ is a closed linear subspace of $X$ and set $N = 2P_U - \text{Id}$. Let $\alpha \in ]0, +\infty[$, set $T_\alpha = (1 - \alpha) \text{Id} + \alpha N$, and set $A_\alpha = (T_\alpha)^{-1} - \text{Id}$. Then $T_\alpha$ is $\alpha$-conically nonexpansive and

$$A_\alpha = \begin{cases} N_U, & \text{if } \alpha = \frac{1}{2}; \\ \frac{2\alpha}{1 - 2\alpha} P_U, & \text{otherwise}. \end{cases}$$

Moreover, $A_\alpha$ is maximally \((\frac{1}{2\alpha} - 1)\)-comonotone.

**Proof** First note that $T_\alpha = (1 - \alpha) \text{Id} + \alpha (2P_U - \text{Id}) = (1 - 2\alpha) \text{Id} + 2\alpha P_U$. The case $\alpha = \frac{1}{2}$ is clear by, e.g., [2, Example 23.4]. Now suppose that $\alpha \in [0, +\infty[ \setminus \{\frac{1}{2}\}$, and let $(x, y) \in X \times X$. Then $y \in A_\alpha x \iff x + y \in (\text{Id} + A_\alpha)x \iff x = T_\alpha(x + y) = (1 - 2\alpha)(x + y) + 2\alpha P_U(x + y) \iff x = x + y - 2\alpha(\text{Id} - P_U)(x + y) \iff y = 2\alpha P_U(x + y) = 2\alpha P_U x + 2\alpha y \iff y = \frac{2\alpha}{1 - 2\alpha} P_U x$, and the conclusion follows in view of Corollary 3.8(ii). \hfill \Box

Proposition 3.10 Let $A : X \rightrightarrows X$ be such that $\text{dom } A \neq \emptyset$, let $\rho \in ]-1, +\infty[$, set $D = \text{ran}(\text{Id} + A)$, set $T = J_A$, i.e., $A = T^{-1} - \text{Id}$, and set $N = 2(\rho + 1)T - (2\rho + 1) \text{Id}$, i.e., $T = \frac{2\rho + 1}{2(\rho + 1)} \text{Id} + \frac{1}{2(\rho + 1)} N$. Then the following equivalences hold:

(i) $A$ is $\rho$-comonotone $\iff$ $N$ is nonexpansive.

(ii) $A$ is maximally $\rho$-comonotone $\iff$ $N$ is nonexpansive and $D = X$.

**Proof** (i): Set $\alpha = \frac{1}{2(\rho + 1)}$ and note that $\alpha > 0$. It follows from Proposition 2.13 that $T = J_A$ is single-valued. Now use Proposition 3.6(i). (ii): Combine (i) and Proposition 3.6(ii). \hfill \Box

Proposition 3.11 Let $A : X \rightrightarrows X$ be such that $\text{dom } A \neq \emptyset$, let $\rho \in ]-1, +\infty[$, set $D = \text{ran}(\text{Id} + A)$, set $T = J_A$, i.e., $A = T^{-1} - \text{Id}$, and set $\alpha = \frac{1}{2(\rho + 1)}$. Then we have the following equivalences:

(i) $A$ is $\rho$-comonotone $\iff$ $T$ is $\frac{1}{2(\rho + 1)}$-conically nonexpansive.

(ii) $A$ is maximally $\rho$-comonotone $\iff$ $T$ is $\alpha$-conically nonexpansive and $D = X$.

(iii) $A$ is $(-\frac{1}{2})$-comonotone $\iff$ $T$ is nonexpansive.

(iv) $A$ is maximally $(-\frac{1}{2})$-comonotone $\iff$ $T$ is nonexpansive and $D = X$.

(v) $[A$ is $\rho$-comonotone and $\rho > -\frac{1}{2}]$ $\iff$ $T$ is $\alpha$-averaged.

(vi) $[A$ is maximally $\rho$-monotone and $\rho > -\frac{1}{2}]$ $\iff$ $[T$ is $\alpha$-averaged and $D = X]$.

**Proof** This follows from Proposition 3.7. \hfill \Box

Corollary 3.12 Let $A : X \rightrightarrows X$ be maximally $\rho$-comonotone for some $\rho > -\frac{1}{2}$. Then $J_A : X \rightrightarrows X$ is $\frac{1}{2(\rho + 1)}$-averaged.

The following corollary can be found in [7, Proposition 6.9.6].

**Corollary 3.13** Let $A : X \rightrightarrows X$ be maximally $\rho$-comonotone and $\rho \geq -\frac{1}{2}$. Then $A^{-1}(0)$ is closed and convex.

**Proof** It is clear that $A^{-1}(0) = \text{Fix } J_A$. The conclusion now follows from combining [2, Corollary 4.24] and Proposition 3.11(iv). \hfill \Box

Table 1 below summarizes the main results of this section.
Table 1  Properties of an operator $A$ and its inverse $A^{-1}$ along with the corresponding resolvents $J_A$ and $J_{A^{-1}}$ respectively, for different values of $\rho \in \mathbb{R}$. Here, $A$ satisfies the implication: \[(x, u), (y, v) \in \text{gra } A \Rightarrow (x - y, u - v) \geq \rho \|u - v\|^2\]

| $\rho$ | $A$ | $A^{-1}$ | $J_A$ | $J_{A^{-1}}$ |
|-------|-----|---------|------|----------|
| -0    | $\rho$-cocoercive | $\rho$-strongly monotone | conically nonexpansive | $(\rho + 1)$-cocoercive |
| 0     | monotone | monotone | firmly nonexpansive | $\frac{1}{2(\rho + 1)}$-averaged |
| -0.5  | $\rho$-comonotone | $\rho$-monotone | nonexpansive | $(\rho + 1)$-cocoercive |
| -0.5  | $\rho$-comonotone | $\rho$-monotone | nonexpansive | $\frac{1}{2(\rho + 1)}$-averaged |
| -1    | $\rho$-comonotone | $\rho$-monotone | may fail to be at most single-valued | $\frac{1}{2(\rho + 1)}$-averaged |

### 4 Further properties of the resolvent $J_A$ and the reflected resolvent $R_A$

We start this section with the following useful lemma.

**Lemma 4.1** Let $D$ be a nonempty subset of $X$, let $T : D \to X$, and let $\alpha \in [0, 1]$. Then the following hold:

(i) $T$ is $\alpha$-averaged $\iff 2T - \text{Id} = (1 - 2\alpha) \text{Id} + 2\alpha N$ for some nonexpansive $N : D \to X$.

(ii) $[T = \frac{\alpha}{2} (\text{Id} + N)$ and $N : D \to X$ is nonexpansive] $\iff -(2T - \text{Id})$ is $\alpha$-averaged, in which case $T$ is a Banach contraction with Lipschitz constant $\alpha < 1$.

(iii) $T$ is $\frac{1}{2}$-strongly monotone $\iff 2T - \text{Id}$ is monotone.

**Proof** (i): We have: $T$ is $\alpha$-averaged $\iff [T = (1 - \alpha) \text{Id} + \alpha N$ and $N$ is nonexpansive] $\iff [2T - \text{Id} = (2 - 2\alpha) \text{Id} + 2\alpha N - \text{Id} = (1 - \alpha) \text{Id} + 2\alpha N$ and $N$ is nonexpansive].

(ii): We have: $[T = \frac{\alpha}{2} (\text{Id} + N)$ and $N$ is nonexpansive] $\iff [2T - \text{Id} = (\alpha - 1) \text{Id} + \alpha N = -(1 - \alpha) \text{Id} + \alpha (-N)$ and $N$ is nonexpansive].

(iii): We have: $T$ is $\frac{1}{2}$-strongly monotone $\iff T - \frac{1}{2} \text{Id}$ is monotone $\iff 2T - \text{Id}$ is monotone.

Before we proceed, we recall the following useful fact (see, e.g., [2, Proposition 4.35]).

**Fact 4.2** Let $D$ be a nonempty subset of $X$, let $T : D \to X$, and let $\alpha \in [0, 1]$. Then

$$T \text{ is } \alpha\text{-averaged } \iff \|Tx - Ty\|^2 + (1 - 2\alpha)\|x - y\|^2$$

---

$^3$ This is also known as $\alpha$-negatively averaged (see [15, Definition 3.7]).
\[ \leq 2(1-\alpha)(x - y, Tx - Ty) \]  
(18)

for all \((x, y) \in D \times D\).

**Proposition 4.3** Let \(A : X \rightrightarrows X\) and set \(D = \text{ran}(\text{Id} + A)\). Then the following hold:

(i) Let \(\beta > -\frac{1}{2}\). Then: \(A\) is \(\beta\)-comonotone \(\iff\) \(J_A : D \rightarrow X\) is \(\frac{1}{\sigma(1+\beta)}\)-averaged \(\iff\) \(R_A = (1 - \frac{1}{1+\beta}) \text{Id} + \frac{1}{1+\beta}N\) for some nonexpansive \(N : D \rightarrow X\).

(ii) Let \(\beta > 0\). Then: \(A\) is \(\beta\)-strongly monotone \(\iff\) \(J_A = \frac{1}{\sigma(\beta+1)}(\text{Id} + N) : D \rightarrow X\) and \(N : D \rightarrow X\) is nonexpansive \(\iff\) \(-R_A : D \rightarrow X\) is \(\frac{1}{\sigma+1}\)-averaged, in which case \(J_A\) is a Banach contraction with Lipschitz constant \(\frac{1}{\beta+1} < 1\).

(iii) \(A\) is nonexpansive on \(\text{dom} A \iff J_A\) is \(\frac{1}{2}\)-strongly monotone \(\iff\) \(R_A\) is monotone.

(iv) Let \(\alpha \in [0, 1]\). Then: \(A\) is \(\alpha\)-averaged on \(\text{dom} A \iff R_A\) is \(\frac{1-\alpha}{\alpha}\)-cocoercive.

(v) \(A\) is firmly nonexpansive on \(\text{dom} A \iff R_A\) is firmly nonexpansive.

**Proof** Let \(\{(x, u), (y, v)\} \subseteq X \times X\). Using (1), we have \(\{(x, u), (y, v)\} \subseteq \text{gra} J_A \iff \{(u, x - u), (v, y - v)\} \subseteq \text{gra} A\), which we shall use repeatedly.

(i): Let \(\{(x, u), (y, v)\} \subseteq \text{gra} J_A\) be arbitrarily chosen. We have

\[
A\text{ is } \beta\text{-comonotone}
\]

\[
\iff \beta \| (x - y) - (u - v) \|^2 \\
\leq \langle (x - y) - (u - v), u - v \rangle
\]

\[
\iff \beta \| x - y \|^2 + \beta \| u - v \|^2 - 2\beta \langle x - y, u - v \rangle \leq \langle x - y, u - v \rangle - \| u - v \|^2
\]

\[
\iff \beta \| x - y \|^2 + (\beta + 1) \| u - v \|^2 \leq (2\beta + 1) \langle x - y, u - v \rangle
\]

\[
\iff \| u - v \|^2 + \frac{\beta}{\beta+1} \| x - y \|^2 \leq \frac{2\beta+1}{\beta+1} \| x - y, u - v \|
\]

\[
\iff \| u - v \|^2 + (1 - \frac{1}{\beta+1}) \| x - y \|^2 \leq 2(1 - \frac{1}{\sigma(\beta+1)}) \| x - y, u - v \|
\]

\[
\iff J_A \text{ is } \frac{1}{\sigma(\beta+1)}\text{-averaged}
\]

\[
\iff R_A = (1 - \frac{1}{1+\beta}) \text{Id} + \frac{1}{1+\beta}N \text{for some nonexpansive} N : D \rightarrow X,
\]

where the last two equivalences follow from Fact 4.2 and Lemma 4.1(i), respectively.

(ii): We start by proving the equivalence of the first and third statements (see [15, Proposition 5.4] for “⇒” and also [24, Proposition 2.1(iii)]. Let \(\{(x, u), (y, v)\} \subseteq \text{gra}(-R_A)\) be chosen arbitrarily, i.e., \(\{(x, -u), (y, -v)\} \subseteq \text{gra} R_A\). In view of (2), this is equivalent to \(\{(\frac{1}{2}(x - u), \frac{1}{2}(x + u)), (\frac{1}{2}(y - v), \frac{1}{2}(y + v))\} \subseteq \text{gra} A\). We have

\[
A\text{ is } \beta\text{-strongly monotone}
\]

\[
\iff \langle (x - y) + (u - v), (x - y) - (u - v) \rangle \geq \beta \| (x - y) - (u - v) \|^2
\]

\[
\iff \| x - y \|^2 - \| u - v \|^2 \geq \beta \| x - y \|^2 + \beta \| u - v \|^2 - 2\beta \langle x - y, u - v \rangle
\]

\[
\iff 2\beta \langle x - y, u - v \rangle \geq (\beta - 1) \| x - y \|^2 + (\beta + 1) \| u - v \|^2
\]
where the last equivalence follows from Fact 4.2. Now apply Lemma 4.1(ii) to prove the equivalence of the second and third statements in (ii).

(iii): Let \( \{(x, u), (y, v)\} \subseteq \text{gra} \, J_A \) be chosen arbitrarily and note that (1) and (2) imply \( x - u \in Au, \, y - v \in Av, \, 2u - x \in (\text{Id} - A)u, \) and \( 2v - y \in (\text{Id} - A)v \). It follows from Corollary 3.5(i) applied with \((T, x, y)\) replaced by \((A, u, v)\) that

\[
A \text{ is nonexpansive } \iff \langle (x - y) - (u - v), 2(u - v) - (x - y) \rangle \\
\geq -\frac{1}{2} \|2(u - v) - (x - y)\|^2 \\
\iff -\|x - y\|^2 - 2\|u - v\|^2 + 3\langle x - y, u - v \rangle \\
\geq -2\|u - v\|^2 - \frac{1}{2}\|x - y\|^2 + 2\langle x - y, u - v \rangle \\
\iff \langle x - y, u - v \rangle \geq \frac{1}{2}\|x - y\|^2 \\
\iff J_A \text{ is } \frac{1}{2} \text{-strongly monotone} \\
\iff R_A \text{ is monotone},
\]

where the last equivalence follows from Lemma 4.1(iii).

(iv): Recall that for \( \{(x, u), (y, v)\} \subseteq X \times X \), we have, using (1) and (2), \( \{(x, u), (y, v)\} \subseteq \text{gra} \, R_A \iff \left\{ \left( \frac{1}{2}(x + u), \frac{1}{2}(x - u) \right), \left( \frac{1}{2}(y + v), \frac{1}{2}(y - v) \right) \right\} \subseteq \text{gra} \, A \). Now let \( \{(x, u), (y, v)\} \subseteq \text{gra} \, R_A \) be chosen arbitrarily. Applying Corollary 3.5(ii) with \((T, x, y)\) replaced by \( \left( A, \frac{1}{2}(x + u), \frac{1}{2}(y + v) \right) \) and Remark 3.2, we learn that

\[
A \text{ is } \alpha \text{-averaged } \iff 2\alpha \left( \frac{1}{2}((x - y) - (u - v)), u - v \right) \geq (1 - 2\alpha)\|u - v\|^2 \\
\iff \alpha \langle x - y, u - v \rangle - \alpha\|u - v\|^2 \geq (1 - 2\alpha)\|u - v\|^2 \\
\iff \frac{\alpha}{1 - \alpha} \langle x - y, u - v \rangle \geq \|u - v\|^2 \\
\iff R_A \text{ is } \frac{1 - \alpha}{\alpha} \text{-cocoercive.}
\]

(v): Apply (iv) with \( \alpha = \frac{1}{2} \). \( \square \)

**Remark 4.4** Proposition 4.3(i) generalizes the conclusion of [15, Proposition 5.3]. Indeed, if \( \beta > 0 \), then: \( A \) is \( \beta \)-cocoercive \( \iff R_A \) is \( \frac{1}{\beta + 1} \)-averaged. The proofs are similar.

### 5 \( \rho \)-monotone and \( \rho \)-comonotone linear operators

Given \( A \in \mathbb{R}^{n \times n} \), we denote the symmetric part of \( A \) by \( A_s = \frac{A + A^T}{2} \) and \( \lambda_{\text{min}}(A) \) stands for the smallest eigenvalue of \( A \) provided all eigenvalues of \( A \) are real.

**Proposition 5.1** Let \( A \in \mathbb{R}^{n \times n} \) and let \( \rho \in \mathbb{R} \). Then the following hold:
(i) $A$ is $\rho$-monotone $\iff \lambda_{\min}(A_\rho) \geq \rho$.
(ii) $A$ is $\rho$-comonotone $\iff \lambda_{\min}(A_\rho - \rho A^T A) \geq 0$.

**Proof** We will use the Lowner order notation “$A \succeq B$” only when $A$ and $B$ are both symmetric and $A - B$ is positive semidefinite. Let $x \in \mathbb{R}^n$ be chosen arbitrarily. (i): $A$ is $\rho$-monotone $\iff \langle x, Ax \rangle \geq \rho \|x\|^2 \iff \langle x, (A - \rho \text{Id})x \rangle \geq 0 \iff \langle x, (A - \rho \text{Id})x \rangle \geq 0 \iff A_\rho \succeq \rho \text{Id} \iff \lambda_{\min}(A_\rho) \geq \rho$. (ii): $A$ is $\rho$-comonotone $\iff \langle x, Ax \rangle \geq \rho \|Ax\|^2 \iff \langle x, (A_\rho - \rho A^T A)x \rangle \geq 0 \iff A_\rho - \rho A^T A \succeq 0$ $\iff \lambda_{\min}(A_\rho - \rho A^T A) \geq 0$.

**Example 5.2** Let $N : X \to X$ be continuous and linear such that $N^* = -N$ and $N^2 = -\text{Id}$. Then $N$ is nonexpansive. Moreover, let $\lambda \in \mathbb{R}$, set $T_\lambda = (1 - \lambda) \text{Id} + \lambda N$ and set $A_\lambda = (T_\lambda)^{-1} - \text{Id}$. Then the following hold:

(i) $A_\lambda = \frac{\lambda}{(1 - \lambda)\lambda + \lambda^2} ((1 - 2\lambda) \text{Id} - N)$.
(ii) $A_\lambda$ is $\rho$-monotone with optimal $\rho = \frac{\lambda(1 - 2\lambda)}{\lambda^2 + (1 - \lambda)^2}$.
(iii) $A_0 = 0$ is $\rho$-comonotone for every $\rho \in \mathbb{R}$, while $A_\lambda$ is $\rho$-comonotone with optimal $\rho = \frac{1 - 2\lambda}{2\lambda}$ provided that $\lambda \neq 0$.

**Proof** Let $x \in X$. Then $\|Nx\|^2 = \langle Nx, Nx \rangle = \langle x, N^*Nx \rangle = \langle x, -N^2x \rangle = \langle x, x \rangle = \|x\|^2$. Hence $N$ is nonexpansive; in fact, $N$ is an isometry. Now set

$$B_\lambda = \frac{\lambda}{(1 - \lambda)^2 + \lambda^2} ((1 - 2\lambda) \text{Id} - N).$$

(i): We have

$$\text{(Id} + B_\lambda)T_\lambda = \left(\text{Id} + \frac{\lambda}{(1 - \lambda)^2 + \lambda^2} ((1 - 2\lambda) \text{Id} - N)\right) ((1 - \lambda) \text{Id} + \lambda N)$$

$$= \frac{1}{(1 - \lambda)^2 + \lambda^2} ((1 - \lambda) \text{Id} - \lambda N)((1 - \lambda) \text{Id} + \lambda N)$$

$$= \frac{1}{(1 - \lambda)^2 + \lambda^2} ((1 - \lambda)^2 - \lambda^2 N^2) = \text{Id}.$$  

Similarly, one can show that $T_\lambda (\text{Id} + B_\lambda) = \text{Id}$ and the conclusion follows.

(ii): Using (i), we have

$$\langle x, A_\lambda x \rangle = \frac{\lambda}{(1 - \lambda)^2 + \lambda^2} ((1 - 2\lambda)\|x\|^2 - \langle Nx, x \rangle)$$

$$= \frac{\lambda(1 - 2\lambda)}{(1 - \lambda)^2 + \lambda^2} \|x\|^2.$$  

(iii): Using (i), we have

$$\|A_\lambda x\|^2 = \frac{\lambda^2}{((1 - \lambda)^2 + \lambda^2)^2} ((1 - 2\lambda)^2\|x\|^2 + \|Nx\|^2)$$

$$= \frac{\lambda^2}{((1 - \lambda)^2 + \lambda^2)^2} ((1 - 2\lambda)^2 + 1)\|x\|^2.$$
Therefore, combining with (25) we obtain

\[ \langle x, A_\lambda x \rangle = \frac{(1 - 2\lambda)((1 - \lambda)^2 + \lambda^2)}{\lambda((1 - 2\lambda)^2 + 1)} \cdot \frac{\lambda^2((1 - 2\lambda)^2 + 1)}{((1 - \lambda)^2 + \lambda^2)^2} \|x\|^2 \] (27a)
\[ = \frac{(1 - 2\lambda)((1 - \lambda)^2 + \lambda^2)}{\lambda((1 - 2\lambda)^2 + 1)} \|A_\lambda x\|^2 \] (27b)
\[ = \frac{1 - 2\lambda}{2\lambda} \|A_\lambda x\|^2, \] (27c)

and the conclusion follows.

Remark 5.3 Consider the setting of Example 5.2 and suppose in addition that \( X = \mathbb{R}^n \). Note that \( T_1 = N \) and hence \( \ker T_1 = \{0\} \) because \( N \) is an isometry. If \( \lambda \neq 1 \) and \( x \in \ker T_\lambda \), then \( \langle T_\lambda x, x \rangle = \langle (1 - \lambda) x, x \rangle = (1 - \lambda)\|x\|^2 = 0 \) and thus \( x = 0 \). In either case, \( \ker T_\lambda = \{0\} \) which yields \( \text{ran} T_\lambda = X \) and thus \( \text{dom} (T_\lambda)^{-1} = X \) as well. Thus, \( \text{dom} A_\lambda = X \). Now if \( \rho \in \mathbb{R} \) is such that \( A_\lambda - \rho \text{Id} \) is monotone, then \( A_\lambda - \rho \text{Id} \) is maximally monotone by [2, Example 20.34], i.e.,

\[ A_\lambda \text{ is maximally } \rho\text{-monotone} \] (28)
in Example 5.2(ii).

6 Hypococonvex functions

In this section, we apply results from the previous sections to characterize proximal mappings of hypoconvex functions. Throughout, we assume that \( f : X \to ]-\infty, +\infty] \) is a proper lower semicontinuous function, and that \( \mu, \lambda \) belong to \( ]0, +\infty[ \). The Moreau envelope function and proximal mapping of \( f \) are defined by

\[ e_\mu f(x) = \inf_{y \in X} \left( f(y) + \frac{1}{2\mu}\|x - y\|^2 \right), \]
\[ \text{Prox}_\mu f(x) = \arg\min_{y \in X} \left( f(y) + \frac{1}{2\mu}\|x - y\|^2 \right). \] (29)

Recall that \( f \) is \( \frac{1}{\lambda} \)-hypoconvex (see [27,32]) if \( \text{dom} f \) is convex and

\[ f((1 - \tau)x + \tau y) \leq (1 - \tau)f(x) + \tau f(y) + \frac{1}{2\lambda} \tau(1 - \tau)\|x - y\|^2, \] (30)

for all \( (x, y) \in X \times X \) and \( \tau \in ]0, 1[ \). Equivalently, \( f \) is \( \frac{1}{\lambda} \)-hypoconvex if and only if \( f + \frac{1}{2\lambda}\| \cdot \|^2 \) is convex.

Proposition 6.1 Suppose that \( f \) is \( \frac{1}{\lambda} \)-hypoconvex, and that \( 0 < \mu < \lambda \). Then the following hold:
(i) \( e_\mu f \) is locally Lipschitz on \( X \).

(ii) \( \text{Prox}_\mu f(x) \) is single-valued for every \( x \in X \).

**Proof** (i): See [18, Proposition 3.3(b)].

(ii): Let \( x \in X \). Then the function \( y \mapsto f(y) + \frac{1}{2\mu} \|x - y\|^2 \) is strongly convex because

\[
\begin{align*}
f(y) + \frac{1}{2\mu} \|x - y\|^2 &= \left( f(y) + \frac{1}{2\lambda} \|y\|^2 \right) + \frac{1}{2\mu} \|x - y\|^2 - \frac{1}{2\lambda} \|y\|^2 \\
&= \left( f(y) + \frac{1}{2\lambda} \|y\|^2 \right) + \left( \frac{1}{2\mu} - \frac{1}{2\lambda} \right) \|y\|^2 - \frac{1}{\mu} \langle x, y \rangle + \frac{1}{2\mu} \|y\|^2 \\
&= \left( f(y) + \frac{1}{2\lambda} \|y\|^2 \right) + \left( \frac{1}{2\mu} - \frac{1}{2\lambda} \right) \|y\|^2 - \frac{1}{\mu} \langle x, y \rangle + \frac{1}{2\mu} \|x\|^2
\end{align*}
\]

and \( f + \frac{1}{2\mu} \|\cdot\|^2 \) is convex. By [2, Proposition 11.15 and Corollary 11.9], the function \( y \mapsto f(y) + \frac{1}{2\mu} \|x - y\|^2 \) has exactly one minimizer over \( X \). \( \square \)

To characterize hypoconvex functions, let us introduce an abstract subdifferential. We shall use \( \partial f \) for the subdifferential mapping from convex analysis.

**Definition 6.2** An abstract subdifferential \( \partial_\# \) associates a subset \( \partial_\# f(x) \) of \( X \) to \( f \) at \( x \in X \), and it satisfies the following properties:

(i) \( \partial_\# f = \partial f \) if \( f \) is a proper lower semicontinuous convex function;

(ii) \( \partial_\# f = \nabla f \) if \( f \) is continuously differentiable;

(iii) \( 0 \in \partial_\# f(x) \) if \( f \) attains a local minimum at \( x \in \text{dom} f \);

(iv) for every \( x \in X \) and for every \( \beta \in \mathbb{R} \),

\[
\partial_\# \left( f + \beta \frac{\|\cdot - x\|^2}{2} \right) = \partial_\# f + \beta (\text{Id} - x).
\]

The Clarke–Rockafellar subdifferential, Mordukhovich subdifferential, and Fréchet subdifferential all satisfy Definition 6.2(i)–(iv), see, e.g., [8, 9, 22, 23, 27], so they are incarnations of \( \partial_\# \). Related but different abstract subdifferential operators have been used in [1, 16, 31]. The following result says that for a hypoconvex function the abstract, Clarke–Rockafellar, Mordukhovich, and Fréchet subdifferential operators all coincide.

**Proposition 6.3** If \( f : X \to ] - \infty, +\infty[ \) is a \( \frac{1}{\lambda} \)-hypoconvex function, then

\[
\partial_\# f = \partial \left( f + \frac{1}{2\lambda} \|\cdot\|^2 \right) - \frac{1}{\lambda} \text{Id}. \tag{34}
\]

**Proof** For the convex function \( f + \frac{1}{2\lambda} \|\cdot\|^2 \), apply Definition 6.2(i) and (iv) to obtain

\[
\partial \left( f + \frac{1}{2\lambda} \|\cdot\|^2 \right) = \partial_\# \left( f + \frac{1}{2\lambda} \|\cdot\|^2 \right) = \partial_\# f + \frac{1}{\lambda} \text{Id}
\]

from which (34) follows. \( \square \)
Let $f^*$ denote the Fenchel conjugate of $f$. The following result is well known in $\mathbb{R}^n$, see, e.g., [27, Exercise 12.61(b)(c), Example 11.26(d) and Proposition 12.19], and [32]. In fact, it also holds in a Hilbert space with a similar proof.

**Proposition 6.4** The following are equivalent:

(i) $f$ is $\frac{1}{\lambda}$-hypoconvex.

(ii) $\text{Id} + \lambda \partial \# f$ is maximally monotone.

(iii) $(\forall \mu \in ]0, \lambda[) \text{Prox}_{\mu f}$ is $\lambda/(\lambda - \mu)$-Lipschitz continuous with

$$\text{Prox}_{\mu f} = J_{\mu \partial \# f} = (\text{Id} + \mu \partial \# f)^{-1}.$$  \hspace{1cm} (35)

(iv) $(\forall \mu \in ]0, \lambda[) \text{Prox}_{\mu f}$ is single-valued and continuous.

**Proof** “(i)$\Rightarrow$(ii)”: By Proposition 6.3,

$$\partial \left( f + \frac{1}{2\lambda} \| \cdot \|^2 \right) = \partial \# f + \frac{1}{\lambda} \text{Id}.$$ 

Since $f + \frac{1}{2\lambda} \| \cdot \|^2$ is convex, $\text{Id} + \lambda \partial \# f$ is maximally monotone.

“(ii)$\Rightarrow$(iii)”: By Definition 6.2(iii) and (iv), $y \in \text{Prox}_{\mu f}(x)$ implies that

$$0 \in \partial \# \left( f(y) + \frac{1}{2\mu} \| y - x \|^2 \right) = \partial \# f(y) + \frac{1}{\mu}(y - x).$$

Thus, one has

$$(\forall x \in X) \text{Prox}_{\mu f}(x) \subseteq (\text{Id} + \mu \partial \# f)^{-1}(x).$$ \hspace{1cm} (36)

Using

$$\text{Id} + \mu \partial \# f = \frac{\lambda - \mu}{\lambda} \left( \text{Id} + \frac{\mu}{\lambda - \mu} (\text{Id} + \lambda \partial \# f) \right)$$

yields

$$(\text{Id} + \mu \partial \# f)^{-1} = J_A \circ \left( \frac{\lambda}{\lambda - \mu} \text{Id} \right),$$

where $A = \frac{\mu}{\lambda - \mu} (\text{Id} + \lambda \partial \# f)$ is maximally monotone by the assumption. Since $J_A$ is nonexpansive on $X$, $(\text{Id} + \mu \partial \# f)^{-1}$ is $\lambda/(\lambda - \mu)$-Lipschitz. Proposition 6.1(ii), together with (36), yield $\text{Prox}_{\mu f} = (\text{Id} + \mu \partial \# f)^{-1}$.

“(iii)$\Rightarrow$(iv)”: Clear.

“(iv)$\Rightarrow$(i)”: Let $x \in X$ and let $\mu \in ]0, \lambda[$. We have

$$e_{\mu f}(x) = \frac{1}{2\mu} \| x \|^2 - \left( f + \frac{1}{2\mu} \| \cdot \|^2 \right)^*(\frac{x}{\mu}),$$  \hspace{1cm} (37)
and $e_\mu f$ is locally Lipschitz by Proposition 6.1(i). In view of [4, Proposition 5.1], (iv)
implies that $e_\mu f$ is Fréchet differentiable with $\nabla e_\mu f = \mu^{-1}(\text{Id} - \text{Prox}_\mu f)$. Then $(f + \frac{1}{2\mu}\|\cdot\|^2)^*$ is Fréchet differentiable by (37). It follows from [30, Theorem 1] that $f + \frac{1}{2\mu}\|\cdot\|^2$ is convex. Since this holds for every $\mu \in ]0, \lambda[$, we obtain that $f + \frac{1}{2\lambda}\|\cdot\|^2$ is convex, so (i) follows.

We now provide a new refined characterization of hypoconvex functions in terms of the cocoercivity of their proximal operators; equivalently, of the conical nonexpansiveness of the displacement mapping of their proximal operators.

**Theorem 6.5** The following are equivalent:

(i) $f$ is $\frac{1}{\lambda}$-hypoconvex.

(ii) $(\forall \mu \in ]0, \lambda[) \text{Id} - \text{Prox}_\mu f$ is $\frac{\lambda}{2(\lambda - \mu)}$-conically nonexpansive.

(iii) $(\forall \mu \in ]0, \lambda[) \text{Prox}_\mu f$ is $\frac{\lambda - \mu}{\lambda}$-cocoercive.

**Proof** “(i)$\iff$(ii)”: Using Proposition 6.4 we have

\[ f \text{ is } \frac{1}{\lambda}\text{-hypoconvex} \]
\[ \iff \text{Id} + \lambda \partial f \text{ is maximally monotone} \]
\[ \iff (\forall \mu \in ]0, \lambda[) \frac{1}{\lambda} \text{Id} + \lambda \partial f \text{ is maximally monotone} \]
\[ \iff (\forall \mu \in ]0, \lambda[) \mu \partial f \text{ is maximally } (-\frac{1}{\lambda})\text{-monotone} \]
\[ \iff (\forall \mu \in ]0, \lambda[) (\mu \partial f)^{-1} \text{ is maximally } (-\frac{1}{\lambda})\text{-comonotone} \] (by Lemma 2.8)
\[ \iff (\forall \mu \in ]0, \lambda[) J_{(\mu \partial f)^{-1}} \text{ is } \frac{\lambda}{2(\lambda - \mu)} \text{-conically nonexpansive} \] (by Corollary 3.8(ii))
\[ \iff (\forall \mu \in ]0, \lambda[) J_{\mu \partial f} \text{ is } \frac{\lambda}{2(\lambda - \mu)} \text{-conically nonexpansive} \] (by Proposition 2.11(i))
\[ \iff (\forall \mu \in ]0, \lambda[) \text{Id} - \text{Prox}_\mu f \text{ is } \frac{\lambda}{2(\lambda - \mu)} \text{-conically nonexpansive} \] (by (35)).

“(ii)$\iff$(iii)”: Use Corollary 3.5(iii).

**Corollary 6.6** Suppose that $f : X \rightarrow \mathbb{R}$ is Fréchet differentiable such that $\nabla f$ is Lipschitz continuous with a constant $1/\lambda$. Then the following hold:

(i) $\text{Id} + \lambda \nabla f$ is maximally monotone.

(ii) $f$ is $\frac{1}{\lambda}$-hypoconvex.

(iii) $f + \frac{1}{2\lambda}\|\cdot\|^2$ is convex.

(iv) $(\forall \mu \in ]0, \lambda[) \text{Prox}_\mu f$ is single-valued.

(v) $(\forall \mu \in ]0, \lambda[) \text{Prox}_\mu f$ is $\frac{\lambda - \mu}{\lambda}$-cocoercive.

(vi) $(\forall \mu \in ]0, \lambda[) \text{Prox}_\mu f = J_{\mu \partial f} = (\text{Id} + \mu \nabla f)^{-1}$.

(vii) $(\forall \mu \in ]0, \lambda[) \text{Id} - \text{Prox}_\mu f$ is $\frac{\lambda}{2(\lambda - \mu)}$-conically nonexpansive.

**Proof** Definition 6.2(ii) implies that $(\forall x \in X) \partial f(x) = \{\nabla f(x)\}$. (i): Indeed, $\lambda \nabla f$ is nonexpansive. Now the conclusion follows from [2, Example 20.29]. (ii)–(vii): Combine (i) with Proposition 6.4 and Theorem 6.5.

Finally, we give two examples to illustrate our results.

© Springer
Example 6.7 Suppose that $X = \mathbb{R}$. Let $\lambda > 0$ and set $f_\lambda : x \mapsto \exp(x) - \frac{1}{2\lambda} x^2$. Then $f_\lambda$ is $\frac{1}{\lambda}$-hypoconvex by Proposition 6.4, $f'_\lambda : x \mapsto \exp(x) - \frac{x}{\lambda}$, and we have $\text{Id} + \lambda f'_\lambda = \lambda \exp$ is maximally monotone. Moreover, for every $\mu \in ]0, \lambda[$, we have

$$
\text{Prox}_{\mu f_\lambda} = \left((1 - \frac{\mu}{\lambda}) \text{Id} + \mu \exp\right)^{-1}
$$

(38a)

$$
= x \mapsto \begin{cases}
\ln\left(\frac{x}{\mu}\right), & \text{if } \mu = \lambda; \\
\frac{\lambda x}{\lambda - \mu} - \text{Lambert } W\left(\frac{\lambda \exp(\lambda x/(\lambda - \mu))}{\lambda - \mu}\right), & \text{if } \mu \in ]0, \lambda[.
\end{cases}
$$

(38b)

where the first identity in (38a) follows from Corollary 6.6(vi).

Example 6.8 Let $D$ be a nonempty closed convex subset of $X$, let $\lambda > 0$ and set $f_\lambda = \iota_D - \frac{1}{2\lambda} \|\cdot\|^2$. Then $f_\lambda$ is $\frac{1}{\lambda}$-hypoconvex by Proposition 6.4, and $\partial f_\lambda = N_D - \frac{1}{\lambda} \text{Id}$ by Proposition 6.3. Moreover, $\text{Id} + \lambda \partial f_\lambda = N_D$ is maximally monotone. Finally, using (35) and [2, Example 23.4] we have for every $\mu \in ]0, \lambda[$

$$
\text{Prox}_{\mu f_\lambda} = \left((1 - \frac{\mu}{\lambda}) \text{Id} + \mu N_D\right)^{-1}
$$

(39a)

$$
= \left((1 - \frac{\mu}{\lambda}) (\text{Id} + N_D)\right)^{-1} = (\text{Id} + N_D)^{-1} \circ \left(\frac{\lambda}{\lambda - \mu} \text{Id}\right)
$$

(39b)

$$
= P_D \circ \left(\frac{\lambda}{\lambda - \mu} \text{Id}\right)
$$

(39c)

In particular, if $D$ is a closed convex cone, then $\text{Prox}_{\mu f_\lambda} = \frac{\lambda}{\lambda - \mu} P_D$.

Acknowledgements The authors thank the editors and three anonymous referees for careful reading and constructive comments. HHB, WMM, and XW were partially supported by the Natural Sciences and Engineering Research Council of Canada.

References

1. Aussel, D., Corvellec, J.-N., Lassonde, M.: Mean value property and subdifferential criteria for lower semicontinuous functions. Trans. Am. Math. Soc. 347, 4147–4161 (1995)
2. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, 2nd edn. Springer, Berlin (2017)
3. Bauschke, H.H., Moffat, S.M., Wang, X.: Firmly nonexpansive mappings and maximally monotone operators: correspondence and duality. Set Valued Var Anal 20, 131–153 (2012)
4. Bernard, F., Thibault, L.: Prox-regular functions in Hilbert spaces. J. Math. Anal. Appl. 303, 1–14 (2005)
5. Borwein, J.M.: Fifty years of maximal monotonicity. Optim. Lett. 4, 473–490 (2010)
6. Brezis, H.: Operateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert. North-Holland/Elsevier, Amsterdam (1973)
7. Burachik, R.S., Iusem, A.N.: Set-Valued Mappings and Enlargements of Monotone Operators. Springer, Berlin (2008)
8. Clarke, F.H.: Generalized gradients and applications. Trans. Am. Math. Soc. 205, 247–262 (1975)
9. Clarke, F.H.: Optimization and Nonsmooth Analysis, Classics in Applied Mathematics, 2nd edn. SIAM, Philadelphia (1990)
10. Combettes, P.L.: The convex feasibility problem in image recovery. Adv. Imaging Electron Phys. 25, 155–270 (1995)
11. Combettes, P.L.: Solving monotone inclusions via compositions of nonexpansive averaged operators. Optimization 53, 475–504 (2004)
12. Combettes, P.L., Pennanen, T.: Proximal methods for cohyponomonotone operators. SIAM J. Control Optim. 43, 731–742 (2004)
13. Dao, M.N., Phan, H.M.: Adaptive Douglas–Rachford splitting algorithm for the sum of two operators. SIAM J. Optim. 29, 2697–2724 (2019)
14. Eckstein, J., Bertsekas, D.P.: On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. Math. Program. 55, 293–318 (1992)
15. Giselsson, P.: Tight global linear convergence rate bounds for Douglas–Rachford splitting. J. Fixed Point Theory Appl. 19, 2241–2270 (2017). https://doi.org/10.1007/s11784-017-0417-1
16. Ioffe, A.D.: Approximate subdifferentials and applications I: the finite-dimensional theory. Trans. Am. Math. Soc. 281, 389–416 (1984)
17. Iusem, A.N., Pennanen, T., Svaiter, P.F.: Inexact variants of the proximal point algorithm without monotonicity. SIAM J. Optim. 13, 1080–1097 (2007)
18. Jourani, A., Thibault, L., Zagrodny, D.: Differential properties of the Moreau envelope. J. Funct. Anal. 266, 1185–1237 (2014)
19. Krasnosel’ski˘ı, M.A.: Two remarks on the method of successive approximations. Uspekhi Matematicheskikh Nauk 10, 123–127 (1955)
20. Mann, W.R.: Mean value methods in iteration. Proc. Am. Math. Soc. 4, 506–510 (1953)
21. Minty, G.J.: Monotone (nonlinear) operators in Hilbert spaces. Duke Math. J. 29, 341–346 (1962)
22. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation I, Basic Theory. Springer, Berlin (2006)
23. Mordukhovich, B.S.: Variational Analysis and Applications, Springer Monographs in Mathematics. Springer, Berlin (2018)
24. Moursi, W.M., Vandenberghe, L.: Douglas–Rachford splitting for a Lipschitz continuous and a strongly monotone operator. J. Optim. Theory Appl. 183, 179–198 (2019). https://doi.org/10.1007/s10957-019-01517-8
25. Pennanen, T.: Local convergence of the proximal point algorithm and multiplier methods without monotonicity. Math. Oper. Res. 27, 170–191 (2002)
26. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. 14, 877–898 (1976)
27. Rockafellar, R.T., Wets, R.J.–B.: Variational Analysis. Springer, Berlin (2009). (corrected 3rd printing)
28. Simons, S.: Minimax and Monotonicity. Springer, Berlin (1998)
29. Simons, S.: From Hahn–Banach to Monotonicity. Springer, Berlin (2008)
30. Strömberg, T.: Duality between Fréchet differentiability and strong convexity. Positivity 15, 527–536 (2011)
31. Thibault, L., Zagrodny, D.: Integration of subdifferentials of lower semicontinuous functions on Banach spaces. J. Math. Anal. Appl. 189, 33–58 (1995)
32. Wang, X.: On Chebyshev functions and Klee functions. J. Math. Anal. Appl. 368, 293–310 (2010)
33. Zeidler, E.: Nonlinear Functional Analysis and Its Applications II/A: Linear Monotone Operators. Springer, Berlin (1990)
34. Zeidler, E.: Nonlinear Functional Analysis and Its Applications II/B: Nonlinear Monotone Operators. Springer, Berlin (1990)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.