On the minimal displacement vector of compositions and convex combinations of nonexpansive mappings

Heinz H. Bauschke* and Walaa M. Moursi†

October 1, 2019

Communicated by Michael Overton

Abstract

Monotone operators and (firmly) nonexpansive mappings are fundamental objects in modern analysis and computational optimization. It was shown in 2012 that if finitely many firmly nonexpansive mappings have or "almost have" fixed points, then the same is true for compositions and convex combinations. More recently, sharp information about the minimal displacement vector of compositions and of convex combinations of firmly nonexpansive mappings was obtained in terms of the displacement vectors of the underlying operators.

Using a new proof technique based on the Brezis–Haraux theorem and reflected resolvents, we extend these results from firmly nonexpansive to general averaged nonexpansive mappings. Various examples illustrate the tightness of our results.

2010 Mathematics Subject Classification: Primary 47H05, 47H09; Secondary 47H10, 90C25.

Keywords: Averaged nonexpansive mapping, Brezis–Haraux theorem, displacement map, maximally monotone operator, minimal displacement vector, nonexpansive mapping, resolvent.

1 Introduction

Throughout, we assume that

$$X \text{ is a real Hilbert space with inner product } \langle\cdot,\cdot\rangle \tag{1}$$

and induced norm $\|\cdot\|$. Recall that $T\colon X \to X$ is *nonexpansive* (i.e., 1-*Lipschitz continuous)* if $(\forall(x,y) \in X\times X)$ $\|Tx - Ty\| \leq \|x-y\|$ and that it is *firmly nonexpansive* if $(\forall(x,y) \in X\times X)$

*Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: `heinz.bauschke@ubc.ca`.

†Department of Electrical Engineering, Stanford University, 350 Serra Mall, Stanford, CA 94305, USA. E-mail: `wmoursi@stanford.edu`.

$\|Tx - Ty\|^2 \leq \langle x - y, Tx - Ty\rangle$. Furthermore, recall that a set-valued operator $A\colon X \rightrightarrows X$ is *maximally monotone* if it is *monotone*, i.e., $\{(x,x^*),(y,y^*)\} \subseteq \operatorname{gra} A \Rightarrow \langle x-y, x^*-y^*\rangle \geq 0$ and if the graph of A cannot be properly enlarged without destroying monotonicity[1]. These notions are of central importance in modern optimization; see, e.g., [3], [16], [17], and the references therein. Maximally monotone operators and firmly nonexpansive mappings are closely related to each other (see [20], [23], and [14]) because if $A\colon X \rightrightarrows X$ is maximally monotone, then its *resolvent*

$$\mathrm{J}_A = (\mathrm{Id} + A)^{-1} \tag{2}$$

is firmly nonexpansive, and if $T\colon X \to X$ is firmly nonexpansive, then $T^{-1} - \mathrm{Id}$ is maximally monotone[2]. In a similar vein, the classes of firmly nonexpansive and simply nonexpansive mappings are bijectively linked because the *reflected resolvent*

$$\mathrm{R}_A = 2\mathrm{J}_A - \mathrm{Id} \tag{3}$$

is nonexpansive, and every nonexpansive map arises in this way.

The interest in these operators stems from the fact that minimizers of convex functions are zeros of maximally monotone operators which in turn are fixed points of (firmly) nonexpansive mappings. For basic background material in fixed point theory and monotone operator theory, we refer the reader to [3], [8], [10], [16], [17], [24], [25], [27], [28], [29], [30], [31], [32], [33], and [34].

However, not every problem has a solution; equivalently, not every resolvent has a fixed point. Let us make this concrete by assuming that $R\colon X \to X$ is nonexpansive. The deviation of R from possessing a fixed point is captured by the notion of the *minimal displacement vector* which is well defined by[3]

$$\mathrm{v}_R = P_{\overline{\mathrm{ran}}(\mathrm{Id} - R)}(0). \tag{4}$$

If $\mathrm{v}_R = 0$, then either R has a fixed point or R "almost" has a fixed point in the sense that there is a sequence $(x_n)_{n\in\mathbb{N}}$ in X such that $x_n - Rx_n \to 0$ (see also [26, Section 3 and 4]).

Let us now assume that $m \in \{2,3,4,\ldots\}$ and that we are given m nonexpansive operators $R_1,\ldots,R_m$ on X, with corresponding minimal displacement vectors v_{R_i}. A natural question is the following:

> *What can be said about the minimal displacement vector v_R of R, when R is either a composition or a convex combination of $R_1,\ldots,R_m$, in terms of the given minimal displacement vectors $\mathrm{v}_{R_1},\ldots,\mathrm{v}_{R_m}$?*

In 2012, the authors of [5] proved the that if each $\mathrm{v}_{R_i} = 0$, then so is $\mathrm{v}_R = 0$ (in the composition and the convex combination case) *provided that each R_i is firmly nonexpansive* (see also [2] for the

[1] We shall write $\operatorname{dom} A = \{x \in X \mid Ax \neq \varnothing\}$ for the *domain* of A, $\operatorname{ran} A = A(X) = \bigcup_{x\in X} Ax$ for the *range* of A, and $\operatorname{gra} A = \{(x,u) \in X \times X \mid u \in Ax\}$ for the *graph* of A.

[2] Here and elsewhere, Id denotes the *identity* operator on X.

[3] Given a nonempty closed convex subset C of X, we denote its *projection mapping* or projector by P_C.

earlier case when each R_i is a projector). It is noteworthy that these results have been studied fairly recently by Kohlenbach in [19] and [18] from the viewpoint of *proof mining*. In the past year, these results were extended in [6] to derive bounds on the displacement vector, *but still under the assumption of firm nonexpansiveness*.

> *In this paper we obtain precise information on the minimal displacement vector* v_R *under the much less restrictive assumption that each* R_i *is merely averaged (rather than firmly) nonexpansive.*

The important class of averaged nonexpansive mappings (see [1] and the comprehensive study [11] for more) is much larger than the class of firmly nonexpansive mappings. Indeed, the former class is closed under compositions (but not the latter) and every nonexpansive mapping can be approximated by a sequence of averaged nonexpansive mappings. The key tool to derive our results is the celebrated Brezis–Haraux theorem [9], which is applied in a completely novel way in this work.

Our new results, outlined next, significantly generalize the results in [5] and [6] in various directions:

R1 We obtain very powerful formulae for the ranges of the displacement mapping of compositions and convex combinations. These formulae precisely describe the closure of the range displacement mapping of compositions and convex combinations in terms of the closures of the ranges of the displacement mappings of the individual operators (see (25) and (40)).

R2 Regarding the minimal displacement vector of compositions, we relax the assumption that all mappings are firmly nonexpansive to all but one map are averaged nonexpansive (see Theorem 4).

R3 We show that the conclusion of **R1** is sharp, by providing a counterexample when more than one map fail to be averaged (see Example 4.3).

R4 Regarding the minimal displacement vector of convex combinations, we relax the assumption that all mappings are firmly nonexpansive to them being merely nonexpansive (see Theorem 5.1).

R5 We discuss the attainment of the gap vector of the compositions and the connection to cyclic and noncyclic shifts of the compositions (see Proposition 4.8 and Remark 4.9).

The remainder of this paper is organized as follows. In Section 2, we collect various auxiliary results which will make the proofs of the main results more structured and pleasant. We then turn to compositions of two mappings in Section 3. The new main results concerning compositions are presented in Section 4 while convex combinations are dealt with in Section 5.

Finally, our notation is standard and follows [3] to which we also refer for facts not explicitly mentioned here.

2 Auxiliary results

This section contains various results that will aid in the derivation of the main results in subsequent sections.

2.1 Resolvents and reflected resolvents

Let $C\colon X \rightrightarrows X$ be maximally monotone. The *inverse resolvent identity* (see, e.g., [3, (23.17)] or [28, Lemma 12.14])

$$\mathrm{J}_{C^{-1}} = \mathrm{Id} - \mathrm{J}_C \tag{5}$$

is fundamental, as is the *Minty parametrization* (see, e.g., [3, Remark 23.23(ii)])

$$\operatorname{gra} C = \big\{(\mathrm{J}_C x, \mathrm{Id} - \mathrm{J}_C x) \;\big|\; x \in X\big\} = \big\{(\mathrm{J}_C x, \mathrm{J}_{C^{-1}} x) \;\big|\; x \in X\big\}. \tag{6}$$

Because the reflected resolvent is $\mathrm{R}_C = 2\mathrm{J}_C - \mathrm{Id}$, we obtain $\mathrm{Id} - \mathrm{R}_C = \mathrm{Id} - (2\mathrm{J}_C - \mathrm{Id}) = 2(\mathrm{Id} - \mathrm{J}_C)$ and further not only

$$2\operatorname{ran}(C) = 2\operatorname{ran}(\mathrm{Id} - \mathrm{J}_C) = \operatorname{ran}(\mathrm{Id} - \mathrm{R}_C) \tag{7}$$

by using (6), but also

$$\mathrm{R}_C = \mathrm{J}_C - \mathrm{J}_{C^{-1}}. \tag{8}$$

Let x and y be in X. Then $\mathrm{J}_{-y+C}(x) = \mathrm{J}_C(x+y)$ and $\mathrm{J}_{C(\cdot-y)}(x) = y + \mathrm{J}_C(x-y)$. Hence $\mathrm{R}_{-y+C}(x) = 2\mathrm{J}_C(x+y) - x = y + \mathrm{R}_C(x+y)$ and $\mathrm{R}_{C(\cdot-y)}(x) = y + \mathrm{R}_C(x-y)$. It follows that $\mathrm{R}_{B(\cdot-y)}\mathrm{R}_{-y+A}x = y + \mathrm{R}_B\mathrm{R}_A(x+y)$ and that $-2y + (x+y) - \mathrm{R}_B\mathrm{R}_A(x+y) = x - \mathrm{R}_{B(\cdot-y)}\mathrm{R}_{-y+A}(x)$. This yields the useful translation formula

$$\operatorname{ran}(\mathrm{Id} - \mathrm{R}_B\mathrm{R}_A) = 2y + \operatorname{ran}(\mathrm{Id} - \mathrm{R}_{B(\cdot-y)}\mathrm{R}_{-y+A}). \tag{9}$$

2.2 Averaged nonexpansive mappings

Let $R\colon X \to X$ and let $\alpha \in [0,1[$. Recall that R is α-averaged if $R = (1-\alpha)\,\mathrm{Id} + \alpha N$, where N is nonexpansive; equivalently, by [11, Lemma 2.1] or [3, Proposition 4.35]:

$$(\forall x \in X)(\forall y \in X) \quad \|(\mathrm{Id} - R)x - (\mathrm{Id} - R)y\|^2 \le \tfrac{\alpha}{1-\alpha}\big(\|x-y\|^2 - \|Rx - Ry\|^2\big). \tag{10}$$

If we don't wish to stress the constant α we refer to R simply as averaged or averaged nonexpansive. We have the following useful result.

Fact 2.1. *Let $m \in \{2,3,\ldots\}$, and let $R_1,\ldots,R_m$ be averaged on X. Then $R_m \cdots R_1$ is also averaged.*

Proof. See [1] and [12, Proposition 2.5]. ∎

2.3 Cocoercive operators

Let $\mu > 0$ and let $A\colon X \to X$. Then A is *μ-cocoercive* (this is also known as "inverse strongly monotone") if μA is firmly nonexpansive, i.e.,

$$(\forall x \in X)(\forall y \in X) \quad \langle x - y, Ax - Ay\rangle \geq \mu\|Ax - Ay\|^2; \tag{11}$$

equivalently, A^{-1} is μ-strongly monotone, i.e., $A^{-1} - \mu\,\mathrm{Id}$ is monotone (see [3, Example 22.7]).

The following result is implicitly contained in Moursi and Vandenberghe's [22, Proposition 2.1(iii)], and it extends previous work by Giselsson [15, Proposition 5.3].

Proposition 2.2. *Let $A\colon X \rightrightarrows X$ be maximally monotone, and let $\mu > 0$. Then A is μ-cocoercive if and only if R_A is $(1+\mu)^{-1}$-averaged.*

Proof. It is straightforward to verify that

$$(\forall u \in X)(\forall v \in X) \quad 4\big(\langle v, u - v\rangle - \mu\|u - v\|^2\big) = \|u\|^2 - \|2v - u\|^2 - 4\mu\|u - v\|^2 \tag{12}$$

Using the Minty parametrization (see (6)), we see that

$$\begin{aligned}
&A \text{ is } \mu\text{-cocoercive} \Leftrightarrow\\
&\quad(\forall x \in X)(\forall y \in X) \quad \langle \mathrm{J}_A x - \mathrm{J}_A y, (x-y) - (\mathrm{J}_A x - \mathrm{J}_A y)\rangle \geq \mu\|(x-y) - (\mathrm{J}_A x - \mathrm{J}_A y)\|^2 \Leftrightarrow\\
&(\forall x \in X)(\forall y \in X) \;\; 4\big(\langle \mathrm{J}_A x - \mathrm{J}_A y, (x-y) - (\mathrm{J}_A x - \mathrm{J}_A y)\rangle - \mu\|(x-y) - (\mathrm{J}_A x - \mathrm{J}_A y)\|^2\big) \geq 0.
\end{aligned} \tag{13}$$

On the other hand, using [3, Proposition 4.35], we have

$$\begin{aligned}
&\mathrm{R}_A \text{ is } (1+\mu)^{-1}\text{-cocoercive} \Leftrightarrow\\
&(\forall x \in X)(\forall y \in X)\; \|\mathrm{R}_A x - \mathrm{R}_A y\|^2 \leq \|x - y\|^2 - \frac{1-(1+\mu)^{-1}}{(1+\mu)^{-1}}\|(x-y) - (\mathrm{R}_A x - \mathrm{R}_A y)\|^2 \Leftrightarrow\\
&\quad(\forall x \in X)(\forall y \in X)\; \|x-y\|^2 - \|2(\mathrm{J}_A x - \mathrm{J}_A y) - (x-y)\|^2 - 4\mu\|(x-y) - (\mathrm{J}_A x - \mathrm{J}_A y)\|^2 \geq 0.
\end{aligned} \tag{14}$$

Now combine (13), (14), and (12) with $u = x - y$ and $v = \mathrm{J}_A x - \mathrm{J}_A y$. ■

Lemma 2.3. *Let A and B be maximally monotone on X. Suppose that there exists C in $\{A, B\}$ such that $C\colon X \to X$ is cocoercive. Then $\overline{\mathrm{ran}}(A+B) = \overline{\mathrm{ran}\,A + \mathrm{ran}\,B}$ and $\mathrm{int}\,\mathrm{ran}(A+B) = \mathrm{int}(\mathrm{ran}\,A + \mathrm{ran}\,B)$.*

Proof. Because $\mathrm{dom}\,C = X$, the sum rule (see, e.g., [3, Corollary 25.5(i)]) yields the maximal monotonicity of $A + B$. Moreover, C is 3^* monotone by [3, Example 25.20(i)]. Altogether, the conclusion follows from the Brezis–Haraux theorem (see, e.g., [3, Theorem 25.24(ii)]). ■

2.4 On the range of a displacement map

Let A and B be maximally monotone operators on X. Because of

$$\operatorname{Id} - \mathrm{R}_B\mathrm{R}_A = 2\mathrm{J}_A - 2\mathrm{J}_B\mathrm{R}_A = 2\mathrm{J}_A - 2\mathrm{J}_B(\mathrm{J}_A - \mathrm{J}_{A^{-1}}) = 2\operatorname{Id} - 2\mathrm{J}_A + 2(\operatorname{Id} - \mathrm{J}_B)\mathrm{R}_A, \tag{15}$$

we have $\operatorname{ran}(\operatorname{Id} - \mathrm{R}_B\mathrm{R}_A) \subseteq 2\operatorname{ran}(\operatorname{Id} - \mathrm{J}_A + (\operatorname{Id} - \mathrm{J}_B)\mathrm{R}_A) \subseteq \operatorname{ran}(\operatorname{Id} - \mathrm{R}_A) + \operatorname{ran}(\operatorname{Id} - \mathrm{R}_B)$ by (7). It follows that

$$\overline{\operatorname{ran}}(\operatorname{Id} - \mathrm{R}_B\mathrm{R}_A) \subseteq \overline{\operatorname{ran}(\operatorname{Id} - \mathrm{R}_A) + \operatorname{ran}(\operatorname{Id} - \mathrm{R}_B)}. \tag{16}$$

3 Composition of two mappings

In this section, we study the composition of two mappings.

Lemma 3.1. *Let A and B be maximally monotone on X. Suppose that there exists C in $\{A, B\}$ such that $C\colon X \to X$ and C is cocoercive. Then the following hold:*

(i) $\overline{\operatorname{ran}(\operatorname{Id} - \mathrm{R}_A) + \operatorname{ran}(\operatorname{Id} - \mathrm{R}_B)} \subseteq \overline{\operatorname{ran}}(\operatorname{Id} - \mathrm{R}_B\mathrm{R}_A)$.
(ii) $\overline{\operatorname{ran}(\operatorname{Id} - \mathrm{R}_A) + \operatorname{ran}(\operatorname{Id} - \mathrm{R}_B)} \subseteq \overline{\operatorname{ran}}(\operatorname{Id} - \mathrm{R}_A\mathrm{R}_B)$.

Proof. We start by establishing the following

`Claim:` If $\widetilde{A}$ and $\widetilde{B}$ are maximally monotone on X, and $0 \in \overline{\operatorname{ran}}(\widetilde{A} + \widetilde{B})$, then $0 \in \overline{\operatorname{ran}}(\operatorname{Id} - \mathrm{R}_{\widetilde{B}}\mathrm{R}_{\widetilde{A}})$.

To this end, assume there exist sequences $(x_n, u_n)_{n\in\mathbb{N}}$ in $\operatorname{gra}\widetilde{A}$ and $(x_n, v_n)_{n\in\mathbb{N}}$ in $\operatorname{gra}\widetilde{B}$ such that

$$u_n + v_n \to 0. \tag{17}$$

The Minty parametrizations (see (6)) of $\operatorname{gra}\widetilde{A}$ and $\operatorname{gra}\widetilde{B}$ give

$$(\forall n \in \mathbb{N})\ \ x_n = \mathrm{J}_{\widetilde{A}}(x_n + u_n),\ u_n = \mathrm{J}_{\widetilde{A}^{-1}}(x_n + u_n) \text{ and } v_n = \mathrm{J}_{\widetilde{B}^{-1}}(x_n + v_n); \tag{18}$$

hence, $x_n - u_n = (\mathrm{J}_{\widetilde{A}} - \mathrm{J}_{\widetilde{A}^{-1}})(x_n + u_n) = \mathrm{R}_{\widetilde{A}}(x_n + u_n)$. Set

$$(\forall n \in \mathbb{N})\ \ z_n = 2\mathrm{J}_{\widetilde{A}}(x_n + u_n) - 2\mathrm{J}_{\widetilde{B}}\mathrm{R}_{\widetilde{A}}(x_n + u_n) \in 2\operatorname{ran}(\mathrm{J}_{\widetilde{A}} - \mathrm{J}_{\widetilde{B}}\mathrm{R}_{\widetilde{A}}) \overset{(15)}{=} \operatorname{ran}(\operatorname{Id} - \mathrm{R}_{\widetilde{B}}\mathrm{R}_{\widetilde{A}}). \tag{19}$$

Thus

$$(\forall n \in \mathbb{N})\ \ z_n = 2x_n - 2\mathrm{J}_{\widetilde{B}}(x_n - u_n). \tag{20}$$

Next, on the one hand,

$$\begin{aligned}(\forall n \in \mathbb{N})\ \ z_n - 2(u_n + v_n) &\overset{(20)}{=} 2(x_n - u_n) - 2\mathrm{J}_{\widetilde{B}}(x_n - u_n) - 2v_n \overset{(5)}{=} 2\mathrm{J}_{\widetilde{B}^{-1}}(x_n - u_n) - 2v_n \\ &\overset{(18)}{=} 2\mathrm{J}_{\widetilde{B}^{-1}}(x_n - u_n) - 2\mathrm{J}_{\widetilde{B}^{-1}}(x_n + v_n). \end{aligned} \tag{21}$$

On the other hand, as $\mathrm{J}_{\widetilde{B}^{-1}}$ is nonexpansive, we have $\|\mathrm{J}_{\widetilde{B}^{-1}}(x_n - u_n) - \mathrm{J}_{\widetilde{B}^{-1}}(x_n + v_n)\| \leq \|(x_n - u_n) - (x_n + v_n)\| = \|u_n + v_n\| \to 0$ by (17). Altogether, $z_n - 2(u_n + v_n) \to 0$ and thus $z_n \to 0$ by (17). Recalling that $(z_n)_{n\in\mathbb{N}}$ lies in $\operatorname{ran}(\mathrm{Id} - \mathrm{R}_{\widetilde{B}}\mathrm{R}_{\widetilde{A}})$ from (19), we finally deduce the `Claim` that $0 \in \overline{\operatorname{ran}}(\mathrm{Id} - \mathrm{R}_{\widetilde{B}}\mathrm{R}_{\widetilde{A}})$.

Having established the `Claim`, we now let $y \in X$.

Case 1: $C = B$.
(i): Indeed, the following implications

$$\begin{aligned}
y \in \overline{\operatorname{ran}(\mathrm{Id} - \mathrm{R}_A) + \operatorname{ran}(\mathrm{Id} - \mathrm{R}_B)} &\Leftrightarrow y/2 \in \overline{\operatorname{ran} A + \operatorname{ran} B} && \text{(by (7))}\\
&\Leftrightarrow y/2 \in \overline{\operatorname{ran} A + \operatorname{ran} B(\cdot - y/2)} && \text{(as } \operatorname{ran} B = \operatorname{ran} B(\cdot - y/2))\\
&\Leftrightarrow y/2 \in \overline{\operatorname{ran}}(A + B(\cdot - y/2)) && \text{(by Lemma 2.3)}\\
&\Leftrightarrow 0 \in \overline{\operatorname{ran}}((-y/2 + A) + B(\cdot - y/2)) &&\\
&\Rightarrow 0 \in \overline{\operatorname{ran}}(\mathrm{Id} - \mathrm{R}_{B(\cdot - y/2)}\mathrm{R}_{-y/2+A}) && \text{(by the Claim)}\\
&\Leftrightarrow 0 \in -y + \overline{\operatorname{ran}}(\mathrm{Id} - \mathrm{R}_B\mathrm{R}_A) && \text{(by (9))}\\
&\Leftrightarrow y \in \overline{\operatorname{ran}}(\mathrm{Id} - \mathrm{R}_B\mathrm{R}_A)
\end{aligned}$$

yield the conclusion.

(ii): Similarly to (i), the following implications

$$\begin{aligned}
y \in \overline{\operatorname{ran}(\mathrm{Id} - \mathrm{R}_A) + \operatorname{ran}(\mathrm{Id} - \mathrm{R}_B)} &\Leftrightarrow y/2 \in \overline{\operatorname{ran} A + \operatorname{ran} B} && \text{(by (7))}\\
&\Leftrightarrow y/2 \in \overline{\operatorname{ran} A(\cdot - y/2) + \operatorname{ran} B} && \text{(as } \operatorname{ran} A = \operatorname{ran} A(\cdot - y/2))\\
&\Leftrightarrow y/2 \in \overline{\operatorname{ran}}(A(\cdot - y/2) + B) && \text{(by Lemma 2.3)}\\
&\Leftrightarrow 0 \in \overline{\operatorname{ran}}(A(\cdot - y/2) + (-y/2 + B)) &&\\
&\Rightarrow 0 \in \overline{\operatorname{ran}}(\mathrm{Id} - \mathrm{R}_{A(\cdot - y/2)}\mathrm{R}_{-y/2+B}) && \text{(by the Claim)}\\
&\Leftrightarrow 0 \in -y + \overline{\operatorname{ran}}(\mathrm{Id} - \mathrm{R}_A\mathrm{R}_B) && \text{(by (9))}\\
&\Leftrightarrow y \in \overline{\operatorname{ran}}(\mathrm{Id} - \mathrm{R}_A\mathrm{R}_B)
\end{aligned}$$

yield the conclusion.

Case 2: $C = A$.
(i): Apply item (ii) of *Case 1*, with A, B replaced by B, A. (ii): Apply item (i) of *Case 1*, with A, B replaced by B, A. ■

Theorem 3.2. *Let A and B be maximally monotone. Suppose that there exists $C \in \{A, B\}$ such that $C\colon X \to X$ and C is cocoercive. Then*

$$\overline{\operatorname{ran}(\mathrm{Id} - \mathrm{R}_A) + \operatorname{ran}(\mathrm{Id} - \mathrm{R}_B)} = \overline{\operatorname{ran}}(\mathrm{Id} - \mathrm{R}_B\mathrm{R}_A) = \overline{\operatorname{ran}}(\mathrm{Id} - \mathrm{R}_A\mathrm{R}_B). \tag{22}$$

Proof. Indeed, we have

$$\overline{\operatorname{ran}(\mathrm{Id} - \mathrm{R}_A) + \operatorname{ran}(\mathrm{Id} - \mathrm{R}_B)} \subseteq \overline{\operatorname{ran}}(\mathrm{Id} - \mathrm{R}_B\mathrm{R}_A) \cap \overline{\operatorname{ran}}(\mathrm{Id} - \mathrm{R}_A\mathrm{R}_B) \qquad \text{(by Lemma 3.1)}$$

$$\subseteq \overline{\operatorname{ran}}(\operatorname{Id} - \mathrm{R}_B\mathrm{R}_A) \cup \overline{\operatorname{ran}}(\operatorname{Id} - \mathrm{R}_A\mathrm{R}_B)$$
$$\subseteq \overline{\operatorname{ran}(\operatorname{Id} - \mathrm{R}_A) + \operatorname{ran}(\operatorname{Id} - \mathrm{R}_B)}. \qquad \text{(by (16))}$$

Hence all inclusions are in fact equalities and we are done. ■

Theorem 3.3 (composition of two nonexpansive mappings). *Let R_1 and R_2 be nonexpansive on X, and suppose that R_1 or R_2 is actually averaged nonexpansive. Then*

$$\overline{\operatorname{ran}}(\operatorname{Id} - R_2R_1) = \overline{\operatorname{ran}}(\operatorname{Id} - R_1R_2) = \overline{\operatorname{ran}(\operatorname{Id} - R_1) + \operatorname{ran}(\operatorname{Id} - R_2)}. \tag{23}$$

Proof. Since R_1 and R_2 are nonexpansive, there exist maximally monotone operators A and B on X such that $R_1 = \mathrm{R}_A$ and $R_2 = \mathrm{R}_B$ by using [3, Corollary 23.9 and Proposition 4.4] applied to $\frac{1}{2}(\operatorname{Id} + R_1)$ and $\frac{1}{2}(\operatorname{Id} + R_2)$.

Case 1: R_1 is averaged nonexpansive.
By Proposition 2.2, $A\colon X \to X$ and A is cocoercive. The conclusion is now clear from Theorem 3.2.

Case 2: R_2 is averaged nonexpansive.
Argue similarly to *Case 1*, or apply *Case 1* with R_1 and R_2 interchanged. ■

We conclude this section with an m-operator version of Theorem 3.3 (which we will sharpen in Section 4).

Proposition 3.4. *Let $m \in \{1, 2, \ldots\}$, and let $R_1, \ldots, R_m$ be averaged nonexpansive operators on X. Then*

$$\overline{\operatorname{ran}}(\operatorname{Id} - R_m \cdots R_1) = \overline{\operatorname{ran}(\operatorname{Id} - R_1) + \cdots + \operatorname{ran}(\operatorname{Id} - R_m)}. \tag{24}$$

Proof. The proof is by induction on m. The base case, $m = 1$, is trivial. Now assume that the result is true for some integer $m \geq 1$, and that we are given $m + 1$ averaged nonexpansive operators $R_1, \ldots, R_{m+1}$ on X. By Fact 2.1, $R_m \cdots R_1$ is averaged nonexpansive. Applying Theorem 3.3 to $R_m \cdots R_1$ and R_{m+1} we obtain

$$\begin{aligned}
&\overline{\operatorname{ran}}(\operatorname{Id} - R_{m+1} \cdots R_1)\\
&\quad = \overline{\operatorname{ran}}\big(\operatorname{Id} - R_{m+1}(R_m \cdots R_1)\big)\\
&\quad = \overline{\operatorname{ran}(\operatorname{Id} - R_{m+1}) + \operatorname{ran}(\operatorname{Id} - R_m \cdots R_1)} && \text{(by Theorem 3.3)}\\
&\quad = \overline{\operatorname{ran}(\operatorname{Id} - R_{m+1}) + \overline{\operatorname{ran}(\operatorname{Id} - R_1) + \cdots + \operatorname{ran}(\operatorname{Id} - R_m)}} && \text{(use inductive hypothesis)}\\
&\quad = \overline{\operatorname{ran}(\operatorname{Id} - R_1) + \cdots + \operatorname{ran}(\operatorname{Id} - R_{m+1})},
\end{aligned}$$

and the proof is complete. ■

4 Compositions

By combining Theorem 3.3 with Proposition 3.4, we are now ready for the main result on compositions.

Theorem 4.1 (main result on compositions). *Let $m \in \{1,2,\ldots\}$, and let $R_1,\ldots,R_m$ be nonexpansive on X. Suppose there exists $j \in \{1,\ldots,m\}$ such that each R_i is averaged nonexpansive whenever $i \neq j$. Let σ be a permutation of $\{1,\ldots,m\}$. Then*

$$\overline{\operatorname{ran}}(\operatorname{Id} - R_m \cdots R_1) = \overline{\operatorname{ran}(\operatorname{Id} - R_1) + \cdots + \operatorname{ran}(\operatorname{Id} - R_m)} \tag{25a}$$

$$= \overline{\operatorname{ran}}\big(\operatorname{Id} - R_{\sigma(m)} \cdots R_{\sigma(1)}\big); \tag{25b}$$

consequently,

$$\mathrm{v}_{R_{\sigma(m)}\cdots R_{\sigma(1)}} = \mathrm{v}_{R_m \cdots R_1} = P_{\overline{\operatorname{ran}(\operatorname{Id} - R_1)+\cdots+\operatorname{ran}(\operatorname{Id} - R_m)}}(0) \tag{26a}$$

and

$$\|\mathrm{v}_{R_{\sigma(m)}\cdots R_{\sigma(1)}}\| = \|\mathrm{v}_{R_m\cdots R_1}\| \leq \|\mathrm{v}_{R_1}\| + \cdots + \|\mathrm{v}_{R_m}\|. \tag{26b}$$

Proof. The result is clear when $m \in \{1,2\}$. So suppose $m \geq 3$. The conclusion follows in the *boundary case*, i.e., $j = 1$ or $j = m$, by combining Proposition 3.4 with Theorem 3.3. So let us assume additionally that $2 \leq j \leq m-1$. Then $T_m \cdots T_1 = S_2 S_1$, where $S_1 = R_{j-1}\cdots R_1$ is averaged nonexpansive and $S_2 = R_m \cdots R_j$ is nonexpansive. On the one hand, $\overline{\operatorname{ran}}(\operatorname{Id} - S_1) = \overline{\operatorname{ran}(\operatorname{Id} - R_1) + \cdots + \operatorname{ran}(\operatorname{Id} - R_{j-1})}$ by Proposition 3.4. On the other hand, $\overline{\operatorname{ran}}(\operatorname{Id} - S_2) = \overline{\operatorname{ran}(\operatorname{Id} - R_j) + \cdots \operatorname{ran}(\operatorname{Id} - R_m)}$ by the boundary case. Altogether, (25a) follows by Theorem 3.3 applied to S_1 and S_2. Finally, (25b) is a direct consequence of (25a) while (26) is implied by (25). ■

Remark 4.2. *Some comments are in order.*

(i) *The inequality in* (26b) *significantly generalizes [6, Theorem 2.2], where each R_i was assumed to be firmly nonexpansive!*
(ii) *The inequality in* (26b) *is sharp even when each R_i is firmly nonexpansive; see [6, Example 2.4].*

The following example shows that at least one of the mappings in Theorem 4.1 must be averaged.

Example 4.3. *Let u_1, u_2 be in X, and let $i \in \{1,2\}$. Set $R_i \colon x \mapsto -x - u_i$, which is nonexpansive but not averaged. Now $(\forall x \in X)$ we have $x - R_i x = 2x + u_i$. Hence*

$$\operatorname{ran}(\operatorname{Id} - R_i) = X. \tag{27}$$

We also have $(\forall x \in X)$ $R_2R_1x = x + u_1 - u_2$ and $R_1R_2x = x + u_2 - u_1$. It follows that $(\forall x \in X)$ $x - R_2R_1x = u_2 - u_1$ and $x - R_1R_2x = u_1 - u_2$. Thus $\operatorname{ran}(\operatorname{Id} - R_2R_1) = \{u_2 - u_1\}$ *and* $\operatorname{ran}(\operatorname{Id} - R_1R_2) = \{u_1 - u_2\}$*; in turn,*

$$\mathrm{v}_{R_2R_1} = u_2 - u_1 \ \textit{ yet } \ \mathrm{v}_{R_1R_2} = u_1 - u_2. \tag{28}$$

Therefore, the conclusions of Theorem 4.1 fail for general nonexpansive mappings whenever $u_1 \neq u_2$.

When $m = 2$, the following positive result for cyclic permutations of the mappings can be found in [4, Lemma 2.6]:

Proposition 4.4. *Let $m \in \{2,3,\ldots\}$ and let $R_1,\ldots,R_m$ be nonexpansive on X. Then*

$$\|\mathrm{v}_{R_mR_{m-1}\cdots R_1}\| = \|\mathrm{v}_{R_1R_m\cdots R_2}\| = \cdots = \|\mathrm{v}_{R_{m-1}\cdots R_1R_m}\| \tag{29}$$

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence such that $\|x_n - R_mR_{m-1}\cdots R_1x_n\| \to \|\mathrm{v}_{R_mR_{m-1}\cdots R_1}\|$ Then $\|\mathrm{v}_{R_1R_mR_{m-1}\cdots R_2}\| \leq \|(R_1x_n) - (R_1R_m\cdots R_2)(R_1x_n)\| = \|R_1x_n - R_1(R_m\cdots R_1x_n)\| \leq \|x_n - R_m\cdots R_1x_n\| \to \|\mathrm{v}_{R_mR_{m-1}\cdots R_1}\|$. Hence

$$\|\mathrm{v}_{R_1R_mR_{m-1}\cdots R_2}\| \leq \|\mathrm{v}_{R_mR_{m-1}\cdots R_1}\| \tag{30}$$

and the result follows by continuing cyclically in this fashion. ■

Proposition 4.5. *Let $m = 3$ and $R_i = \delta_i \operatorname{Id} - a_i$, where $\delta_i \in \{-1,0,1\}$. Then $\operatorname{ran}(\operatorname{Id} - R_i) = \{a_i\}$, if $\delta_i = 1$; and $\operatorname{ran}(\operatorname{Id} - R_i) = X$, if $\delta_i \in \{-1,0\}$. Moreover, $R_3R_2R_1\colon X \to X\colon x \mapsto \delta_3\delta_2\delta_1 x - a_3 - \delta_3 a_2 - \delta_3\delta_2 a_1$ and so*

$$\operatorname{ran}(\operatorname{Id} - R_3R_2R_1) = \begin{cases} \{a_3 + \delta_3 a_2 + \delta_3\delta_2 a_1\}, & \textit{if } \delta_1\delta_2\delta_3 = 1; \\ X, & \textit{otherwise,} \end{cases} \tag{31}$$

which implies

$$\mathrm{v}_{R_3R_2R_1} = \begin{cases} a_3 + \delta_3 a_2 + \delta_3\delta_2 a_1, & \textit{if } \delta_1\delta_2\delta_3 = 1; \\ 0, & \textit{otherwise.} \end{cases} \tag{32}$$

Example 4.6. *Suppose that $m = 3$, $R_1x = -x$, $R_2x = -x + u$, and $R_3 = x - u$, where $u \in X \smallsetminus \{0\}$. Then R_1 and R_2 are nonexpansive but not averaged, while R_3 is firmly nonexpansive. Therefore*

$$\mathrm{v}_{R_3R_2R_1} = 0 \tag{33}$$

while

$$\mathrm{v}_{R_3R_1R_2} = 2u \neq 0. \tag{34}$$

Hence $\|\mathrm{v}_{R_3R_2R_1}\| = 0 < 2\|u\| = \|\mathrm{v}_{R_3R_1R_2}\|$.

Proof. Set $\delta_1 = -1$, $\delta_2 = -1$, $\delta_3 = 1$, and $a_1 = 0$, $a_2 = -u$, $a_3 = u$. Then $\delta_1\delta_2\delta_3 = 1$ and (32) yields

$$\mathrm{v}_{R_3R_2R_1} = u + 1\cdot(-u) + (1)(-1)\cdot 0 = 0. \tag{35}$$

Similarly,

$$\mathrm{v}_{R_3R_1R_2} = u + 1\cdot 0 + (1)(-1)\cdot(-u) = 2u \tag{36}$$

and the proof is complete. ■

Remark 4.7.

(i) *Example 4.6 illustrates that the assumption that $m-1$ — and not merely 1 — of the operators in Theorem 4.1 be averaged nonexpansive is critical.*

(ii) *Example 4.6 also shows that Proposition 4.4 fails for noncyclic permutations of the mappings.*

Proposition 4.8. *Let $m \in \{2, 3, \ldots, \}$, and let $R_1, \ldots, R_m$ be nonexpansive on X. Then we have the following equivalences:*

$$
\begin{aligned}
& \mathrm{v}_{R_m R_{m-1}\cdots R_1} \in \operatorname{ran}(\operatorname{Id} - R_m R_{m-1}\cdots R_1) \\
\Leftrightarrow\; & \mathrm{v}_{R_{m-1}\cdots R_1 R_m} \in \operatorname{ran}(\operatorname{Id} - R_{m-1}\cdots R_1 R_m) && \text{(37a)} \\
\Leftrightarrow\; & \cdots && \text{(37b)} \\
\Leftrightarrow\; & \mathrm{v}_{R_1 R_m\cdots R_2} \in \operatorname{ran}(\operatorname{Id} - R_1 R_m\cdots R_2). && \text{(37c)}
\end{aligned}
$$

Proof. By symmetry, it suffices to show that "⇒" in (37a) holds. To this end, assume that there is $y \in X$ such that $\mathrm{v}_{R_m R_{m-1}\cdots R_1} = y - R_m R_{m-1}\cdots R_1 y$. Either a direct argument or [7, Proposition 2.5(iv)] shows that $\mathrm{v}_{R_m R_{m-1}\cdots R_1} = (R_m R_{m-1}\cdots R_1)y - (R_m R_{m-1}\cdots R_1)^2 y$. Thus

$$
\begin{aligned}
\|\mathrm{v}_{R_{m-1}\cdots R_1 R_m}\| &\overset{(29)}{=} \|\mathrm{v}_{R_m\cdots R_2 R_1}\| \\
&= \|(R_m R_{m-1}\cdots R_1)y - (R_m R_{m-1}\cdots R_1)^2 y\| \\
&\le \|R_{m-1}\cdots R_1 y - (R_{m-1}\cdots R_1)R_m R_{m-1}\cdots R_1 y\| \\
&\le \|y - R_m R_{m-1}\cdots R_1 y\| \\
&= \|\mathrm{v}_{R_m R_{m-1}\cdots R_1}\| \\
&\overset{(29)}{=} \|\mathrm{v}_{R_{m-1}\cdots R_1 R_m}\|.
\end{aligned}
$$

Consequently, $\|\mathrm{v}_{R_{m-1}\cdots R_1 R_m}\| = \|R_{m-1}\cdots R_1 y - (R_{m-1}\cdots R_1 R_m)R_{m-1}\cdots R_1 y\|$. Therefore

$$
\begin{aligned}
\mathrm{v}_{R_{m-1}\cdots R_1 R_m} &= R_{m-1}\cdots R_1 y - (R_{m-1}\cdots R_1 R_m)R_{m-1}\cdots R_1 y \\
&\in \operatorname{ran}(\operatorname{Id} - R_{m-1}\cdots R_1 R_m),
\end{aligned}
$$

and the proof is complete. ■

Remark 4.9. *Proposition 4.8 shows that the minimal displacement vector is attained for all cyclic shifts of the composition. For noncyclic shifts, this result goes wrong as De Pierro observed in [13, Section 3 on page 193] (see also [7, Example 2.7]).*

Let us conclude this section with an application to the projected gradient descent method (see also [21] for an analysis of the forward-backward method in the possibly inconsistent case).

Corollary 4.10 (projected gradient descent). *Let $f\colon X \to \mathbb{R}$ be convex and differentiable on X, with ∇f being L-Lipschitz continuous, let C be a nonempty closed convex subset of X, let $\alpha \in \left]0, 2\right[$. Then the magnitude of the minimal dispacement vector of the projected gradient descent operator*

$$T\colon X \to X\colon x \mapsto \mathrm{P}_C \circ \left(\operatorname{Id} - \alpha \tfrac{1}{L}\nabla f\right) \tag{38}$$

satisfies $\|\mathrm{v}_T\| \le \alpha L^{-1} \inf \|\nabla f(X)\|$.

Proof. This follows from Theorem 4.1 with $m = 2$, $R_1 = \operatorname{Id} - \alpha L^{-1}\nabla f$ and $R_2 = P_C$, where $\mathrm{v}_{R_2} = 0$. Hence $\|\mathrm{v}_{R_2}\| = 0$ and $\|\mathrm{v}_{R_1}\| = \inf_{x\in X}\|x - R_1 x\| = \inf_{x\in X}\|\alpha L^{-1}\nabla f(x)\|$. Now use (26b). ■

5 Convex combinations

In this final section, we focus on convex combinations of nonexpansive mappings.

Theorem 5.1 (main result on convex combinations). *Let $m \in \{2,3,\ldots\}$, let $R_1,\ldots,R_m$ be nonexpansive on X, and let $\lambda_1,\ldots,\lambda_m$ be in $]0,1[$ such that $\sum_{i=1}^m \lambda_i = 1$. Set*

$$\overline{R} = \sum_{i=1}^m \lambda_i R_i. \tag{39}$$

Then

$$\overline{\operatorname{ran}}(\operatorname{Id} - \overline{R}) = \overline{\sum_{i=1}^m \lambda_i \operatorname{ran}(\operatorname{Id} - R_i)}. \tag{40}$$

Consequently,

$$\mathrm{v}_{\overline{R}} = \mathrm{P}_{\overline{\sum_{i=1}^m \lambda_i \operatorname{ran}(\operatorname{Id} - R_i)}}(0) \tag{41}$$

and

$$\|\mathrm{v}_{\overline{R}}\| \leq \|\sum_{i=1}^m \lambda_i \mathrm{v}_{R_i}\| \leq \sum_{i=1}^m \lambda_i \|\mathrm{v}_{R_i}\|. \tag{42}$$

Proof. Set $A_i := \operatorname{Id} - R_i$ for each $i \in \{1,\ldots,m\}$. By [3, Example 20.29 and Example 25.20], A_i is maximally and 3^* monotone. By [6, Lemma 3.1],

$$\begin{aligned}\overline{\operatorname{ran}}(\operatorname{Id} - \overline{R}) &= \overline{\operatorname{ran}}\big(\sum_{i=1}^m \lambda_i(\operatorname{Id} - R_i)\big) = \overline{\operatorname{ran}}\big(\sum_{i=1}^m \lambda_i A_i\big) = \overline{\sum_{i=1}^m \lambda_i \operatorname{ran} A_i} \\ &= \overline{\sum_{i=1}^m \lambda_i \operatorname{ran}(\operatorname{Id} - R_i)}.\end{aligned}$$

This yields (40) and thus (41). In view of (40), we have

$$\sum_{i=1}^m \lambda_i \mathrm{v}_{R_i} \in \sum_{i=1}^m \lambda_i \overline{\operatorname{ran}}(\operatorname{Id} - R_i) \subseteq \overline{\sum_{i=1}^m \lambda_i \operatorname{ran}(\operatorname{Id} - R_i)} = \overline{\operatorname{ran}}(\operatorname{Id} - \overline{R}). \tag{43}$$

Thus $\|\mathrm{v}_{\overline{R}}\| \leq \|\sum_{i=1}^m \|\lambda_i \mathrm{v}_{R_i}\|$ and (42) follows. ■

Remark 5.2. *[6, Example 3.4 and Example 3.5] illustrate that the inequalities in (42) are sharp and also that in general $\mathrm{v}_{\overline{R}} \neq \sum_{i=1}^m \lambda_i \mathrm{v}_{R_i}$.*

We conclude with an application that extends [5, Theorem 5.5] and [6, Corollary 3.3], where each v_{R_i} was equal to 0:

Corollary 5.3. *Let $m \in \{2,3,\ldots\}$, let $R_1,\ldots,R_m$ be nonexpansive on X, and let $\lambda_1,\ldots,\lambda_m$ be in $]0,1[$ such that $\sum_{i=1}^m \lambda_i = 1$. Set $\overline{R} = \sum_{i=1}^m \lambda_i R_i$ and assume that $\sum_{i=1}^m \lambda_i \mathrm{v}_{R_i} = 0$. Then $\mathrm{v}_{\bar{R}} = 0$.*

Proof. Clear from (42). ■

Acknowledgments

The authors thank two anonymous referees for constructive and insightful comments. The research of HHB was partially supported by a Discovery Grant of the Natural Sciences and Engineering Research Council of Canada. The research of WMM was partially supported by a Natural Sciences and Engineering Research Council of Canada Postdoctoral Fellowship.

References

[1] J.B. Baillon, R.E. Bruck, and S. Reich, On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces, *Houston Journal of Mathematics* 4 (1978), 1–9.

[2] H.H. Bauschke, The composition of finitely many projections onto closed convex sets in Hilbert space is asymptotically regular, *Proceedings of the American Mathematical Society* 131 (2003), 141–146.

[3] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Second Edition, Springer, 2017.

[4] H.H. Bauschke, W.L. Hare, and W.M. Moursi, Generalized solutions for the sum of two maximally monotone operators, *SIAM Journal on Control and Optimization* 52 (2014), 1034–1047.

[5] H.H. Bauschke, V. Martin-Marquez, S.M. Moffat, and X. Wang, Compositions and convex combinations of asymptotically regular firmly nonexpansive mappings are also asymptotically regular, *Fixed Point Theory and Applications* (2012), 2012:53.

[6] H.H. Bauschke and W.M. Moursi, The magnitude of the minimal displacement vector for compositions and convex combinations of firmly nonexpansive mappings, *Optimization Letters* 12 (2018), 1465–1474.

[7] H.H. Bauschke and W.M. Moursi, The Douglas–Rachford algorithm for two (not necessarily intersecting) affine subspaces, *SIAM Journal on Optimization* 26 (2016), 968–985.

[8] H. Brezis, *Operateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*, North-Holland/Elsevier, 1973.

[9] H. Brezis and A. Haraux, Image d'une Somme d'opérateurs Monotones et Applications, *Israel Journal of Mathematics* 23 (1976), 165–186.

[10] R.S. Burachik and A.N. Iusem, *Set-Valued Mappings and Enlargements of Monotone Operators*, Springer-Verlag, 2008.

[11] P.L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization* 53 (2004), 475–504.

[12] P.L. Combettes, I. Yamada, Compositions and convex combinations of averaged nonexpansive operators, *Journal of Mathematical Analysis and Applications* 425 (2014), 55–70.

[13] A. De Pierro, From parallel to sequential projection methods and vice versa in convex feasibility: results and conjectures. In: Inherently parallel algorithms in feasibility and optimization and their applications (Haifa, 2000), 187–201, *Studies in Computational Mathematics* 8 (2001), North-Holland, Amsterdam.

[14] J. Eckstein and D.P. Bertsekas, On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators, *Mathematical Programming* 55 (1992), 293–318.

[15] P. Giselsson, Tight global linear convergence rate bounds for Douglas–Rachford splitting, *Journal of Fixed Point Theory and Applications* 19 (2017), 2241–2270.

[16] K. Goebel and W.A. Kirk, *Topics in Metric Fixed Point Theory*, Cambridge University Press, 1990.

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

[18] U. Kohlenbach, A polynomial rate of asymptotic regularity for compositions of projections in Hilbert space, *Foundations of Computational Mathematics* 19 (2019), 83–99.

[19] U. Kohlenbach, G. López-Acedo and A. Nicolae, Quantitative asymptotic regularity results for the composition of two mappings, *Optimization* 66 (2017), 1291–1299.

[20] G.J. Minty, Monotone (nonlinear) operators in Hilbert spaces, *Duke Mathematical Journal* 29 (1962), 341–346.

[21] W.M. Moursi, The forward-backward algorithm and the normal problem, *Journal of optimization Theory and Applications* 176 (2018), 605–624.

[22] W.M. Moursi and L. Vandenberghe, Douglas–Rachford splitting for the sum of a Lipschitz continuous and a strongly monotone operator, *Journal of Optimization Theory and Applications* 183 (2019), 179–198.

[23] S. Reich, Extension problems for accretive sets in Banach spaces, *Journal of Functional Analysis* 26 (1977), 378–395.

[24] S. Reich, The fixed point property for non-expansive mappings, *American Mathematical Monthly* 83 (1976), 266–268.

[25] S. Reich, The fixed point property for non-expansive mappings, II, *American Mathematical Monthly* 87 (1980), 292–294.

[26] S. Reich, On the asymptotic behavior of nonlinear semigroups and the range of accretive operators, *Journal of Mathematical Analysis and Applications* 79 (1981), 113–126.

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

[28] R.T. Rockafellar and R.J-B. Wets, *Variational Analysis*, Springer-Verlag, corrected 3rd printing, 2009.

[29] S. Simons, *Minimax and Monotonicity*, Springer-Verlag, 1998.

[30] S. Simons, *From Hahn-Banach to Monotonicity*, Springer-Verlag, 2008.

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

[32] E. Zeidler, *Nonlinear Functional Analysis and Its Applications I: Fixed Point Theorems*, Springer-Verlag, 1993.

[33] E. Zeidler, *Nonlinear Functional Analysis and Its Applications II/A: Linear Monotone Operators*, Springer-Verlag, 1990.

[34] E. Zeidler, *Nonlinear Functional Analysis and Its Applications II/B: Nonlinear Monotone Operators*, Springer-Verlag, 1990.