The magnitude of the minimal displacement vector for compositions and convex combinations of firmly nonexpansive mappings

Heinz H. Bauschke* and Walaa M. Moursi†

March 29, 2018

Abstract

Maximally monotone operators and firmly nonexpansive mappings play key roles in modern optimization and nonlinear analysis. Five years ago, it was shown that if finitely many firmly nonexpansive operators are all asymptotically regular (i.e., they have or "almost have" fixed points), then the same is true for compositions and convex combinations.

In this paper, we derive bounds on the magnitude of the minimal displacement vectors of compositions and of convex combinations in terms of the displacement vectors of the underlying operators. Our results completely generalize earlier works. Moreover, we present various examples illustrating that our bounds are sharp.

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

Keywords: Asymptotically regular, firmly nonexpansive mapping, maximally monotone operator, minimal displacement vector, nonexpansive mapping, resolvent.

1 Introduction and Standing Assumptions

Throughout this paper,

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

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

†Simons Institute for the Theory of Computing, UC Berkeley, Melvin Calvin Laboratory, #2190, Berkeley, CA 94720, USA and Mansoura University, Faculty of Science, Mathematics Department, Mansoura 35516, Egypt. E-mail: `walaa.moursi@gmail.com`.

and induced norm $\|\cdot\|$. Recall that $T\colon X \to X$ is *firmly nonexpansive* (see, e.g., [3], [14], and [15] for further information) if $(\forall(x,y) \in X \times X)$ $\|Tx - Ty\|^2 \leq \langle x - y, Tx - Ty\rangle$ and 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 equivalent (see [18] and [12]) in the sense that if A is maximally monotone, then its *resolvent* $J_A := (\operatorname{Id} + A)^{-1}$ is firmly nonexpansive, and if T is firmly nonexpansive, then $T^{-1} - \operatorname{Id}$ is maximally monotone[2].

In optimization, one main problem is to find zeros of (sums of) maximally monotone operators — these zeros may correspond to critical points or solutions to optimization problems. In terms of resolvents, the corresponding problem is that of finding fixed points. For background material in fixed point theory and monotone operator theory, we refer the reader to [3], [7], [8], [10], [14], [15], [21], [22], [24], [23], [25], [26], [27], and [28]. However, not every problem has a solution; equivalently, not every resolvent has a fixed point. To make this concrete, let us assume that $T\colon X \to X$ is firmly nonexpansive. The deviation from T possessing a fixed point is captured by the notion of the *minimal (negative) displacement vector* which is well defined by[3]

$$v_T := P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}(0). \tag{2}$$

If T "almost" has a fixed point in the sense that $v_T = 0$, i.e., $0 \in \overline{\operatorname{ran}}(\operatorname{Id} - T)$, then we say that T is *asymptotically regular*. From now on, we assume that

$$I := \{1, 2, \ldots, m\}, \text{ where } m \in \{2, 3, 4, \ldots\}$$

and that we are given m firmly nonexpansive operators $T_1, \ldots, T_m$; equivalently, m resolvents of maximally monotone operators $A_1, \ldots, A_m$:

$$(\forall i \in I) \quad T_i = J_{A_i} = (\operatorname{Id} + A_i)^{-1} \text{ is firmly nonexpansive,}$$

and we abbreviate the corresponding minimal displacement vectors by

$$(\forall i \in I) \quad v_i := v_{T_i} = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T_i)}(0). \tag{3}$$

A natural question is the following: *What can be said about the minimal displacement vector of T when T is either a composition or a convex combination of $T_1, \ldots, T_m$?*

Five years ago, the authors of [5] proved the following:

> *If each T_i is asymptotically regular, then so are the corresponding compositions and convex combinations.*

[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.

This can be expressed equivalently as

$$(\forall i \in I)\ \ v_i = 0 \quad \Rightarrow \quad v_T = 0, \tag{4}$$

where T is either a composition or a convex combination of the family $(T_i)_{i\in I}$. It is noteworthy that these results have been studied recently by Kohlenbach [17] and [16] from the viewpoint of "proof mining".

In this work, we obtain *sharp bounds* on the magnitude of the minimal displacement vector of T that hold true *without any assumption of asymptotic regularity of the given operators.* The proofs rely on techniques that are new and that were introduced in [5] and [1] (where projectors were considered). The new results concerning compositions are presented in Section 2 while convex combinations are dealt with in Section 3. Finally, our notation is standard and follows [3] to which we also refer for standard facts not mentioned here.

2 Compositions

In this section, we explore compositions.

Proposition 2.1. $(\forall \varepsilon > 0)$ $(\exists x \in X)$ *such that* $\|x - T_m T_{m-1} \cdots T_1 x\| \leq \varepsilon + \sum_{k=1}^{m} \|v_k\|$.

Proof. The proof is broken up into several steps. Set

$$(\forall i \in I) \quad \widetilde{A}_i := -v_i + A_i(\cdot - v_i). \tag{5}$$

and observe that [3, Proposition 23.17(ii)&(iii)] yields

$$(\forall i \in I) \quad \widetilde{T}_i := J_{\widetilde{A}_i} = v_i + J_{A_i} = v_i + T_i. \tag{6}$$

We also work in

$$\mathbf{X} := X^m = \big\{\mathbf{x} = (x_i)_{i\in I} \;\big|\; (\forall i \in I)\ x_i \in X\big\}, \quad \text{with} \ \ \langle \mathbf{x}, \mathbf{y}\rangle = \sum_{i\in I} \langle x_i, y_i\rangle\,, \tag{7}$$

where we embed the original operators via

$$\mathbf{T}\colon X^m \to X^m\colon (x_i)_{i\in I} \mapsto (T_i x_i)_{i\in I} \ \text{ and } \ \mathbf{A}\colon X^m \rightrightarrows X^m\colon (x_i)_{i\in I} \mapsto \times(A_i x_i)_{i\in I}. \tag{8}$$

Denoting the identity on X^m by $\mathbf{Id}$, we observe that

$$J_{\mathbf{A}} = (\mathbf{Id} + \mathbf{A})^{-1} = T_1 \times \cdots \times T_m = \mathbf{T}. \tag{9}$$

Because $\operatorname{ran}(\mathbf{Id} - \mathbf{T}) = \operatorname{ran}(\operatorname{Id} - T_1) \times \cdots \times \operatorname{ran}(\operatorname{Id} - T_m)$ and hence $\overline{\operatorname{ran}}(\mathbf{Id} - \mathbf{T}) = \overline{\operatorname{ran}}(\operatorname{Id} - T_1) \times \cdots \times \overline{\operatorname{ran}}(\operatorname{Id} - T_m)$, we have (e.g., by using [3, Proposition 29.3])

$$\mathbf{v} := (v_i)_{i\in I} = P_{\overline{\operatorname{ran}}(\mathbf{Id} - \mathbf{T})}\mathbf{0}. \tag{10}$$

Finally, define the *cyclic right-shift operator*

$$\mathbf{R}\colon X^m \to X^m\colon (x_1, x_2, \ldots, x_m) \mapsto (x_m, x_1, \ldots, x_{m-1}) \quad\text{and}\quad \mathbf{M} := \mathbf{Id} - \mathbf{R}, \tag{11}$$

and the *diagonal* subspace

$$\boldsymbol{\Delta} := \big\{\mathbf{x} = (x)_{i\in I} \;\big|\; x \in X\big\}, \tag{12}$$

with orthogonal complement $\boldsymbol{\Delta}^{\perp}$.

CLAIM 1: $\mathbf{v} \in \overline{\operatorname{ran}}\,(\mathbf{A}(\cdot - \mathbf{v}) + \mathbf{M})$.
Indeed, (3) implies that $(\forall i \in I)\ v_i \in \overline{\operatorname{ran}}\,(\operatorname{Id} - T_i) = \overline{\operatorname{ran}}\,(\operatorname{Id} - J_{A_i}) = \overline{\operatorname{ran}}\, J_{A_i^{-1}} = \overline{\operatorname{dom}}\,(\operatorname{Id} + A_i^{-1}) = \overline{\operatorname{dom}}\ A_i^{-1} = \overline{\operatorname{ran}}\ A_i = \overline{\operatorname{ran}}\ A_i(\cdot - v_i)$. Hence, $\mathbf{v} \in \overline{\operatorname{ran}}\ \mathbf{A}(\cdot - \mathbf{v}) = \overline{\operatorname{ran} \mathbf{A}(\cdot - \mathbf{v}) + \mathbf{0}} \subseteq \overline{\operatorname{ran} \mathbf{A}(\cdot - \mathbf{v}) + \boldsymbol{\Delta}^{\perp}}$. On the other hand, we learn from [5, Corollary 2.6] (applied to $\mathbf{A}(\cdot - \mathbf{v})$) that $\overline{\operatorname{ran}}\,(\mathbf{A}(\cdot - \mathbf{v}) + \mathbf{M}) = \overline{\operatorname{ran} \mathbf{A}(\cdot - \mathbf{v}) + \boldsymbol{\Delta}^{\perp}}$. Altogether, we obtain that $\mathbf{v} \in \overline{\operatorname{ran}}\,(\mathbf{A}(\cdot - \mathbf{v}) + \mathbf{M})$ and CLAIM 1 is verified.

CLAIM 2: $(\forall \varepsilon > 0)\ (\exists (\mathbf{b}, \mathbf{x}) \in \mathbf{X} \times \mathbf{X})\ \|\mathbf{b}\| \leq \varepsilon$ and $\mathbf{x} = \mathbf{v} + \mathbf{T}(\mathbf{b} + \mathbf{R}\mathbf{x})$.
Fix $\varepsilon > 0$. In view of CLAIM 1, there exist $\mathbf{x} \in \mathbf{X}$ and $\mathbf{b} \in \mathbf{X}$ such that $\|\mathbf{b}\| \leq \varepsilon$ and $\mathbf{b} \in -\mathbf{v} + \mathbf{A}(\mathbf{x} - \mathbf{v}) + \mathbf{M}\mathbf{x}$. Hence, $\mathbf{b} + \mathbf{R}\mathbf{x} = \mathbf{b} + \mathbf{x} - \mathbf{M}\mathbf{x} \in \mathbf{x} + \mathbf{A}(\mathbf{x} - \mathbf{v}) - \mathbf{v} = (\mathbf{Id} + (-\mathbf{v} + \mathbf{A}(\cdot - \mathbf{v}))\mathbf{x}$. Thus, $\mathbf{x} = J_{-\mathbf{v} + \mathbf{A}(\cdot - \mathbf{v})}(\mathbf{b} + \mathbf{R}\mathbf{x}) = \mathbf{v} + \mathbf{T}(\mathbf{b} + \mathbf{R}\mathbf{x})$, where the last identity follows from (6), (9) and (10).

CLAIM 3: $(\forall \varepsilon > 0)\ (\exists (\mathbf{c}, \mathbf{x}) \in \mathbf{X} \times \mathbf{X})\ \|\mathbf{c}\| \leq \varepsilon$ and $\mathbf{x} = \mathbf{c} + \mathbf{v} + \mathbf{T}(\mathbf{R}\mathbf{x})$.
Fix $\varepsilon > 0$, let $\mathbf{b}$ and $\mathbf{x}$ be as in CLAIM 2, and set $\mathbf{c} := \mathbf{x} - \mathbf{v} - \mathbf{T}(\mathbf{R}\mathbf{x}) = \mathbf{T}(\mathbf{b} + \mathbf{R}\mathbf{x}) - \mathbf{T}(\mathbf{R}\mathbf{x})$. Then, since $\mathbf{T}$ is nonexpansive, $\|\mathbf{c}\| = \|\mathbf{T}(\mathbf{b} + \mathbf{R}\mathbf{x}) - \mathbf{T}(\mathbf{R}\mathbf{x})\| \leq \|\mathbf{b}\| \leq \varepsilon$, and CLAIM 3 thus holds.

CONCLUSION:
Let $\varepsilon > 0$. By CLAIM 3 (applied to $\varepsilon/\sqrt{m}$), there exists $(\mathbf{c}, \mathbf{x}) \in \mathbf{X} \times \mathbf{X}$ such that $\|\mathbf{c}\| \leq \varepsilon/\sqrt{m}$ and $\mathbf{x} = \mathbf{c} + \mathbf{v} + \mathbf{T}(\mathbf{R}\mathbf{x})$. Hence $\sum_{i\in I} \|c_i\| \leq \|\mathbf{c}\|\sqrt{m} \leq \varepsilon$ and $(\forall i \in I)\ x_i = c_i + v_i + T_i x_{i-1}$, where $x_0 := x_m$. The triangle inequality and the nonexpansiveness of each T_i thus yields

$$\begin{aligned}
\|T_m T_{m-1} \cdots T_1 x_0 - x_0\| &= \|T_m T_{m-1} \cdots T_1 x_0 - x_m\| \\
&= \big\|T_m T_{m-1} \cdots T_2 T_1 x_0 - T_m T_{m-1} \cdots T_2 x_1 \\
&\quad + T_m T_{m-1} \cdots T_3 T_2 x_1 - T_m T_{m-1} \cdots T_3 x_2 \\
&\quad + T_m T_{m-1} \cdots T_4 T_3 x_2 - T_m T_{m-1} \cdots T_4 x_3 \\
&\quad + \cdots \\
&\quad + T_m T_{m-1} x_{m-2} - T_m x_{m-1} \\
&\quad + T_m x_{m-1} - x_m\big\| \\
&\leq \|T_m T_{m-1} \cdots T_2 T_1 x_0 - T_m T_{m-1} \cdots T_2 x_1\| \\
&\quad + \|T_m T_{m-1} \cdots T_3 T_2 x_1 - T_m T_{m-1} \cdots T_3 x_2\| \\
&\quad + \|T_m T_{m-1} \cdots T_4 T_3 x_2 - T_m T_{m-1} \cdots T_4 x_3\| \\
&\quad + \cdots \\
&\quad + \|T_m T_{m-1} x_{m-2} - T_m x_{m-1}\| \\
&\quad + \|T_m x_{m-1} - x_m\|
\end{aligned}$$

$$
\begin{aligned}
&\le \|T_1x_0 - x_1\| + \|T_2x_1 - x_2\| + \|T_3x_2 - x_3\| \\
&\qquad + \cdots + \|T_{m-1}x_{m-2} - x_{m-1}\| + \|T_mx_{m-1} - x_m\| \\
&= \|c_1 + v_1\| + \|c_2 + v_2\| + \cdots + \|c_m + v_m\| \\
&\le \sum_{k=1}^{m} \|c_i\| + \sum_{k=1}^{m} \|v_i\| \\
&\le \varepsilon + \sum_{k=1}^{m} \|v_i\|, \qquad (13)
\end{aligned}
$$

as claimed. ■

We are now ready for our first main result.

Theorem 2.2. $\|v_{T_m \cdots T_2 T_1}\| \le \|v_{T_1}\| + \cdots + \|v_{T_m}\|$.

Proof. By Proposition 2.1, we have $(\forall \varepsilon > 0)$ $\|v_{T_m \cdots T_2 T_1}\| \le \varepsilon + \|v_{T_1}\| + \cdots + \|v_{T_m}\|$ and the result thus follows. ■

As an immediate consequence of Theorem 2.2, we obtain the first main result of [5]:

Corollary 2.3. [5, Corollary 3.2] *Suppose that $v_1 = \cdots = v_m = 0$. Then $v_{T_m \cdots T_2 T_1} = 0$.*

We now show that the bound on $\|v_{T_m \cdots T_2 T_1}\|$ given in Theorem 2.2 is sharp:

Example 2.4. *Suppose that $X = \mathbb{R}$, $T_1 \colon X \to X \colon x \mapsto x - a_1$, and $T_2 \colon X \to X \colon x \mapsto x - a_2$, where $(a_1, a_2) \in \mathbb{R} \times \mathbb{R}$. Then $(v_{T_1}, v_{T_2}, v_{T_2T_1}) = (a_1, a_2, a_1 + a_2)$ and $|a_1 + a_2| = |v_{T_2T_1}| \le |v_1| + |v_2| = |a_1| + |a_2|$; moreover, the inequality is an equality if and only if $a_1 a_2 \ge 0$.*

Proof. On the one hand, it is clear that $\operatorname{ran}(\operatorname{Id} - T_1) = \{a_1\}$ and likewise $\operatorname{ran}(\operatorname{Id} - T_2) = \{a_2\}$. Consequently, $(v_1, v_2) = (a_1, a_2)$. On the other hand, $T_2T_1 \colon X \to X \colon x \mapsto x - a_1 - a_2 = x - (a_1 + a_2)$, therefore $\operatorname{ran}(\operatorname{Id} - T_2T_1) = \{a_1 + a_2\}$. Hence, $v_{T_2T_1} = a_1 + a_2$, $|v_{T_2T_1}| = |a_1 + a_2|$ and $|v_1| + |v_2| = |a_1| + |a_2|$, and the conclusion follows. ■

The remaining results in this section concern the effect of cyclically permuting the operators in the composition.

Proposition 2.5. $v_{T_m T_{m-1} \cdots T_2 T_1} = v_{T_{m-1} T_{m-2} \cdots T_1 T_m} = \cdots = v_{T_1 T_m \cdots T_2}$.

Proof. We start by proving that if $S_1 \colon X \to X$ and $S_2 \colon X \to X$ are averaged[4], then

$$v_{S_2S_1} = v_{S_1S_2}. \qquad (14)$$

[4]Let $S \colon X \to X$. Then S is α-averaged if there exists $\alpha \in [0, 1[$ such that $S = (1 - \alpha)\operatorname{Id} + \alpha N$ and $N \colon X \to X$ is nonexpansive.

To this end, let $x \in X$ and note that S_2S_1 and S_1S_2 are α-averaged where $\alpha \in [0,1[$ by, e.g., [3, Remark 4.34(iii) and Proposition 4.44]. Using [19, Proposition 2.5(ii)] applied to S_2S_1 and S_1S_2 yields

$$\begin{aligned}
\|v_{S_2S_1} - v_{S_1S_2}\|^2 &\leftarrow \|(S_2S_1)^n x - (S_2S_1)^{n+1}x - ((S_1S_2)^n S_1 x - (S_1S_2)^{n+1}S_1x)\|^2 \\
&= \|(S_2S_1)^n x - (S_2S_1)^{n+1}x - (S_1(S_2S_1)^n x - S_1(S_2S_1)^{n+1}x)\|^2 \\
&= \|(\mathrm{Id} - S_1)(S_2S_1)^n x - (\mathrm{Id} - S_1)(S_2S_1)^{n+1}x\|^2 \\
&\le \tfrac{\alpha}{1-\alpha}(\|(S_2S_1)^n x - (S_2S_1)^{n+1}x\|^2 - \|S_1(S_2S_1)^n x - S_1(S_2S_1)^{n+1}x\|^2) \\
&\le \tfrac{\alpha}{1-\alpha}(\|(S_2S_1)^n x - (S_2S_1)^{n+1}x\|^2 - \|(S_1S_2)^n S_1 x - (S_1S_2)^{n+1}S_1x\|^2) \\
&\to \tfrac{\alpha}{1-\alpha}(\|v_{S_2S_1}\|^2 - \|v_{S_1S_2}\|^2) = 0,
\end{aligned} \tag{15}$$

where the last identity follows from [4, Lemma 2.6]. Because $T_{m-1}T_{m-2}\dots T_1$ is averaged by [3, Remark 4.34(iii) and Proposition 4.44], we can and do apply (14), with (S_1, S_2) replaced by $(T_{m-1}T_{m-2}\dots T_1, T_m)$, to deduce that $v_{T_mT_{m-1}\cdots T_2T_1} = v_{T_{m-1}T_{m-2}\cdots T_1T_m}$. The remaining identities follow similarly. ■

Proposition 2.6. *We have*

$$\begin{aligned}
v_{T_mT_{m-1}\cdots T_1} \in \operatorname{ran}(\mathrm{Id} - T_mT_{m-1}\cdots T_1) &\Leftrightarrow v_{T_{m-1}\cdots T_1T_m} \in \operatorname{ran}(\mathrm{Id} - T_{m-1}\cdots T_1T_m) &\quad (16a)\\
&\Leftrightarrow \cdots &\quad (16b)\\
&\Leftrightarrow v_{T_1T_m\cdots T_2} \in \operatorname{ran}(\mathrm{Id} - T_1T_m\dots T_2). &\quad (16c)
\end{aligned}$$

Proof. We prove the implication "⇒" of (16a): Suppose that $(\exists y \in X)\ v_{T_mT_{m-1}\cdots T_1} = y - T_mT_{m-1}\cdots T_1y$, i.e., $y \in \operatorname{Fix}(v_{T_m\cdots T_1} + T_m\cdots T_1)$. By [6, Proposition 2.5(iv)], we have $v_{T_mT_{m-1}\cdots T_1} = (T_m\cdots T_1)y - (T_m\cdots T_1)^2 y$. Using Proposition 2.5, we obtain

$$\begin{aligned}
\|v_{T_{m-1}\cdots T_1T_m}\| = \|v_{T_m\cdots T_2T_1}\| &= \|(T_mT_{m-1}\cdots T_1)y - (T_mT_{m-1}\cdots T_1)^2 y\| \\
&\le \|T_{m-1}\cdots T_1y - (T_{m-1}\cdots T_1T_m)T_{m-1}\cdots T_1y\| \\
&\le \|y - T_mT_{m-1}\cdots T_1y\| = \|v_{T_mT_{m-1}\cdots T_1}\| = \|v_{T_{m-1}\cdots T_1T_m}\|.
\end{aligned} \tag{17}$$

Consequently, $\|v_{T_{m-1}\cdots T_1T_m}\| = \|T_{m-1}\cdots T_1y - (T_{m-1}\cdots T_1T_m)T_{m-1}\cdots T_1y\|$ and hence

$$v_{T_{m-1}\dots T_1T_m} = T_{m-1}\cdots T_1y - (T_{m-1}\cdots T_1T_m)T_{m-1}\dots T_1y \in \operatorname{ran}(\mathrm{Id} - T_{m-1}\dots T_1T_m). \tag{18}$$

The opposite implication and the remaining $m-2$ equivalences are proved similarly. ■

The following example, taken from De Pierro's [11, Section 3 on page 193], illustrates that the conclusion of Proposition 2.6 does not necessarily hold if the operators are permuted noncyclically.

Example 2.7. *Suppose that* $X = \mathbb{R}^2$, $m = 3$, $C_1 = \mathbb{R} \times \{0\}$, $C_2 = \mathbb{R} \times \{1\}$, $C_3 = \{(x,y) \in \mathbb{R}^2 \mid y \ge 1/x > 0\}$, *and* $(T_1, T_2, T_3) = (P_{C_1}, P_{C_2}, P_{C_3})$. *Then,* $v_{T_3T_2T_1} = v_{T_3T_1T_2} = 0$, $v_{T_3T_2T_1} \in \operatorname{ran}(\mathrm{Id} - T_3T_2T_1)$ *but* $v_{T_3T_1T_2} \notin \operatorname{ran}(\mathrm{Id} - T_3T_1T_2)$.

Proof. Note that $T_2T_1 = P_{C_2}P_{C_1} = P_{C_2} = T_2$ and $T_1T_2 = P_{C_1}P_{C_2} = P_{C_1} = T_1$. Consequently, $(T_3T_2T_1, T_3T_1T_2) = (P_{C_3}P_{C_2}, P_{C_3}P_{C_1})$. The claim that $v_{T_3T_2T_1} = v_{T_3T_1T_2} = 0$ follows from [1, Theorem 3.1], or Theorem 2.2 applied with $m = 3$. This and [2, Lemma 2.2(i)] imply that $\operatorname{Fix} T_3T_2T_1 = \operatorname{Fix} P_{C_3}P_{C_2} = C_3 \cap C_2 \neq \varnothing$, whereas $\operatorname{Fix} T_3T_1T_2 = \operatorname{Fix} P_{C_3}P_{C_1} = C_3 \cap C_1 = \varnothing$. Hence, $v_{T_3T_2T_1} \in \operatorname{ran}(\operatorname{Id} - T_3T_2T_1)$ but $v_{T_3T_1T_2} \notin \operatorname{ran}(\operatorname{Id} - T_3T_1T_2)$. ■

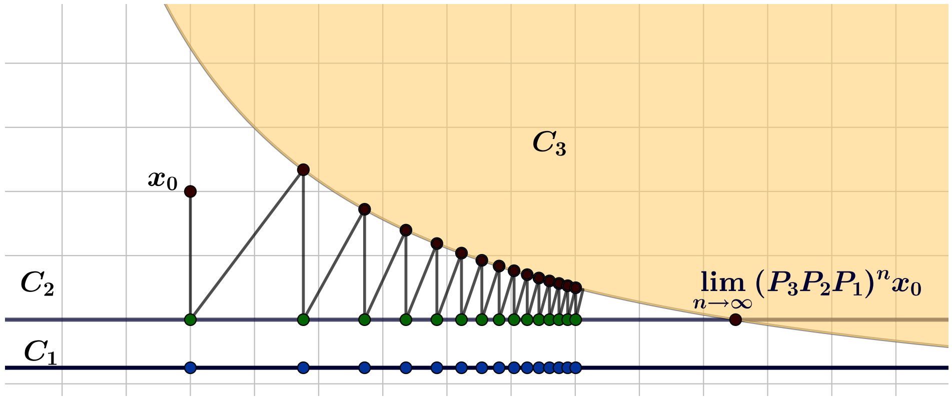

Figure 1: A GeoGebra [13] snapshot that illustrates the behaviour of the sequence $((P_3P_2P_1)^n x_0)_{n\in\mathbb{N}}$ in Proposition 2.6. The first few iterates of the sequences $(P_1(P_3P_2P_1)^n x_0)_{n\in\mathbb{N}}$ (blue points), $(P_2P_1(P_3P_2P_1)^n x_0)_{n\in\mathbb{N}}$ (green points), and $((P_3P_2P_1)^n x_0)_{n\in\mathbb{N}}$ (black points) are also depicted.

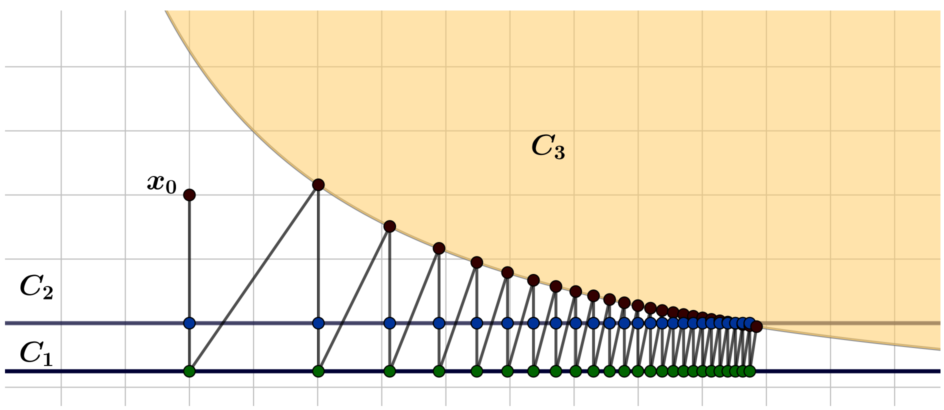

Figure 2: A GeoGebra [13] snapshot that illustrates the behaviour of the sequence $((P_3P_1P_2)^n x_0)_{n\in\mathbb{N}}$ in Proposition 2.6. The first few iterates of the sequences $(P_1(P_3P_1P_2)^n x_0)_{n\in\mathbb{N}}$ (green points), $(P_2P_1(P_3P_1P_2)^n x_0)_{n\in\mathbb{N}}$ (blue points), and $((P_3P_1P_2)^n x_0)_{n\in\mathbb{N}}$ (black points) are also depicted.

3 Convex Combinations

We start with the following useful lemma.

Lemma 3.1. *Suppose* $(\forall i \in I)$ A_i *is* 3^* *monotone*[5] *and* $\operatorname{dom} A_i = X$. *Let* $(\alpha_i)_{i\in I}$ *be a family of nonnegative real numbers. Then the following hold:*

(i) $\sum_{i\in I} \alpha_i A_i$ *is maximally monotone,* 3^* *monotone and* $\operatorname{dom}\left(\sum_{i\in I} \alpha_i A_i\right) = X$.
(ii) $\overline{\operatorname{ran}}\left(\sum_{i\in I} \alpha_i A_i\right) = \overline{\sum_{i\in I} \alpha_i \operatorname{ran} A_i}$.

Proof. Note that $(\forall i \in I)$, $\alpha_i A_i$ is maximally monotone, 3^* monotone and $\operatorname{dom} \alpha_i A_i = X$.

(i): The proof proceeds by induction. For $n = 2$, the 3^* monotonicity of $\alpha_1 A_1 + \alpha_2 A_2$ follows from [3, Proposition 25.22(ii)], whereas the maximal monotonicity of $\alpha_1 A_1 + \alpha_2 A_2$ follows from, e.g., [3, Proposition 25.5(i)]. Now suppose that for some $n \geq 2$ it holds that $\sum_{i=1}^{n} \alpha_i A_i$ is maximally monotone and 3^* monotone. Then $\sum_{i=1}^{n+1} \alpha_i A_i = \sum_{i=1}^{n} \alpha_i A_i + \alpha_{n+1} A_{n+1}$, which is maximally monotone and 3^* monotone, where the conclusion follows from applying the base case with $(\alpha_1, \alpha_2, A_1, A_2)$ replaced by $(1, \alpha_{n+1}, \sum_{i=1}^{n} \alpha_i A_i, A_{n+1})$.

(ii): Combine (i) and [20, Corollary 6]. ■

From this point onwards, let $(\lambda_i)_{i\in I}$ be in $]0, 1]$ with $\sum_{i\in I} \lambda_i = 1$, and set

$$\overline{T} := \sum_{i\in I} \lambda_i T_i. \tag{19}$$

We are now ready for our second main result.

Theorem 3.2. $\|v_{\overline{T}}\| \leq \|\sum_{i\in I} \lambda_i v_{T_i}\|$.

Proof. It follows from [3, Examples 20.7 and 25.20] that $(\forall i \in I)$ $\operatorname{Id} - T_i$ is maximally monotone, 3^* monotone and $\operatorname{dom}(\operatorname{Id} - T_i) = X$. This and Lemma 3.1(ii) (applied with (α_i, A_i) replaced by $(\lambda_i, \operatorname{Id} - T_i)$) imply that

$$\overline{\operatorname{ran}}\left(\operatorname{Id} - \overline{T}\right) = \overline{\operatorname{ran}} \sum_{i\in I} \lambda_i (\operatorname{Id} - T_i) = \overline{\sum_{i\in I} \lambda_i \operatorname{ran}(\operatorname{Id} - T_i)}. \tag{20}$$

Now, on the one hand, it follows from the definition of $v_{\overline{T}}$ that

$$\left(\forall y \in \overline{\operatorname{ran}}\left(\operatorname{Id} - \overline{T}\right)\right) \qquad \|v_{\overline{T}}\| \leq \|y\|. \tag{21}$$

On the other hand, the definition of v_i implies that $(\forall i \in I)$ $v_i \in \overline{\operatorname{ran}}(\operatorname{Id} - T_i)$. Hence, $\lambda_i v_i \in \lambda_i \overline{\operatorname{ran}}(\operatorname{Id} - T_i)$. Therefore, $\sum_{i\in I} \lambda_i v_i \in \sum_{i\in I} \lambda_i \overline{\operatorname{ran}}(\operatorname{Id} - T_i) \subseteq \overline{\sum_{i\in I} \lambda_i \operatorname{ran}(\operatorname{Id} - T_i)} = \overline{\operatorname{ran}}\left(\operatorname{Id} - \overline{T}\right)$, where the last identity follows from (20). Now apply (21) with y replaced by $\sum_{i\in I} \lambda_i v_i$. ■

As an easy consequence of Theorem 3.2, we obtain the second main result of [5]:

[5] We recall that a monotone operator $B \colon X \rightrightarrows X$ is 3^* *monotone* (see [9]) (this is also known as *rectangular*) if $(\forall (x, y^*) \in \operatorname{dom} B \times \operatorname{ran} B)$ $\sup_{(z,z^*)\in \operatorname{gra} B} \langle x - z, z^* - y^* \rangle < +\infty$.

Corollary 3.3. [5, Theorem 5.5] *Suppose that* $v_1 = \cdots = v_m = 0$. *Then* $v_{\overline{T}} = 0$.

The bound we provided in Theorem 3.2 is sharp as we illustrate now:

Example 3.4. *Let* $a \in X$ *and suppose that* $T\colon X \to X\colon x \mapsto x - a$. *Then* $v_T = a$ *and therefore* $\operatorname{Fix} T \neq \varnothing \Leftrightarrow a = 0$. *Set* $(\forall i \in I)\ T_i = T$. *Then* $\overline{T} = \sum_{i\in I} \lambda_i T_i = T$, $(\forall i \in I)\ v_i = v_{\overline{T}} = a$. *Consequently,* $\|v_{\overline{T}}\| = \|a\| = \|\sum_{i\in I} \lambda_i a\| = \|\sum_{i\in I} \lambda_i v_i\|$.

Example 3.4 suggests that the identity $v_{\overline{T}} = \sum_{i\in I} \lambda_i v_i$ holds true; however, the following example provides a negative answer to this conjecture.

Example 3.5. *Suppose that* $m = 2$, *that* $T_1\colon X \to X\colon x \mapsto x - a_1$, *and that* $T_2\colon X \to X\colon x \mapsto \frac{1}{2}x - a_2$, *where* $(a_1, a_2) \in (X \smallsetminus \{0\}) \times X$. *Then* $\operatorname{ran}(\operatorname{Id} - T_1) = \{a_1\}$, $\operatorname{ran}(\operatorname{Id} - T_2) = X$, $\operatorname{ran}(\operatorname{Id} - \overline{T}) = X$, *and* $0 = v_{\overline{T}} \neq \lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 a_1$.

Proof. On the one hand, one can easily verify that $(v_1, v_2) = (a_1, 0)$; hence, $\lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 a_1 \neq 0$. On the other hand, $\overline{T}\colon X \to X\colon x \mapsto \frac{\lambda_1+1}{2}x - (\lambda_1 a_1 + \lambda_2 a_2)$. Hence, $\overline{T}$ is a Banach contraction, and therefore, $\operatorname{Fix} \overline{T} \neq \varnothing$. Consequently, $v_{\overline{T}} = 0$. ■

Acknowledgments

The research of HHB was partially supported by a Discovery Grant of the Natural Sciences and Engineering Research Council of Canada. WMM was was partially supported by the DIMACS/Simons Collaboration on Bridging Continuous and Discrete Optimization through NSF grant # CCF-1740425.

References

[1] 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.

[2] H.H. Bauschke and J.M. Borwein, Dykstra's alternating projection algorithm for two sets, *Journal of Approximation Theory* 79 (1994), 418–443.

[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 Douglas–Rachford algorithm for two (not necessarily intersecting) affine subspace, *SIAM Journal on Optimization* 26, 968–985, 2016.
[7] J.M. Borwein and J.D. Vanderwerff, *Convex Functions*, Cambridge University Press, 2010.
[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] 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.
[12] 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.
[13] GeoGebra, http://www.geogebra.org.
[14] K. Goebel and W.A. Kirk, *Topics in Metric Fixed Point Theory*, Cambridge University Press, 1990.
[15] K. Goebel and S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, Marcel Dekker, 1984.
[16] U. Kohlenbach, A polynomial rate of asymptotic regularity for compositions of projections in Hilbert space, *Foundations of Computational Mathematics*, (2018), https://doi.org/10.1007/s10208-018-9377-0.
[17] U. Kohlenbach, G. López-Acedo and A. Nicolae, Quantitative asymptotic regularity results for the composition of two mappings, *Optimization* 66 (2017), 1291–1299.
[18] G.J. Minty, Monotone (nonlinear) operators in Hilbert spaces, *Duke Mathematical Journal* 29 (1962), 341–346.
[19] W.M. Moursi, The forward-backward algorithm and the normal problem, *Journal of optimization Theory and Applications* 176 (2018), 605–624.
[20] T. Pennanen, On the range of monotone composite mappings, *Journal of Nonlinear and Convex Analysis* 2 (2001), 193–202.
[21] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.
[22] R.T. Rockafellar and R.J-B. Wets, *Variational Analysis*, Springer-Verlag, corrected 3rd printing, 2009.
[23] S. Simons, *From Hahn-Banach to Monotonicity*, Springer-Verlag, 2008.
[24] S. Simons, *Minimax and Monotonicity*, Springer-Verlag, 1998.
[25] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific Publishing, 2002.
[26] E. Zeidler, *Nonlinear Functional Analysis and Its Applications I: Fixed Point Theorems*, Springer-Verlag, 1993.
[27] E. Zeidler, *Nonlinear Functional Analysis and Its Applications II/A: Linear Monotone Operators*, Springer-Verlag, 1990.
[28] E. Zeidler, *Nonlinear Functional Analysis and Its Applications II/B: Nonlinear Monotone Operators*, Springer-Verlag, 1990.