Generalized solutions for the sum of two maximally monotone operators

Heinz H. Bauschke[*], Warren L. Hare[†], and Walaa M. Moursi[‡]

December 19, 2013 (revised version)

Abstract

A common theme in mathematics is to define generalized solutions to deal with problems that potentially do not have solutions. A classical example is the introduction of least squares solutions via the normal equations associated with a possibly infeasible system of linear equations.

In this paper, we introduce a "normal problem" associated with finding a zero of the sum of two maximally monotone operators. If the original problem admits solutions, then the normal problem returns this same set of solutions. The normal problem may yield solutions when the original problem does not admit any; furthermore, it has attractive variational and duality properties. Several examples illustrate our theory.

2010 Mathematics Subject Classification: Primary 47H05, 47H09, 90C46; Secondary 49M27, 49N15, 90C25.

Keywords: Attouch–Théra duality, Douglas–Rachford splitting operator, firmly nonexpansive mapping, generalized solution, maximally monotone operator, resolvent.

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

[†]Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: warren.hare@ubc.ca.

[‡]Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: walaa.moursi@ubc.ca.

1 Motivation and Introduction

1.1 A motivation from Linear Algebra

A classical problem rooted in Linear Algebra and of central importance in the natural sciences is to solve a system of linear equations, say

$$Ax = b. \tag{1}$$

However, it may occur (due to noisy data, for instance) that (1) does not have a solution. An ingenious approach to cope with this situation, dating back to Carl Friedrich Gauss and his famous prediction of the asteroid Ceres (see, e.g., [12, Subsection 1.1.1] and [24, Epilogue in Section 4.6]) in 1801, is to consider the *normal equation* associated with (1), namely

$$A^*Ax = A^*b, \tag{2}$$

where A^* denotes the transpose of A. The normal equation (2) has extremely useful properties:

- If the original system (1) has a solution, then so does the associated system (2); furthermore, the sets of solutions of these two systems coincide in this case.
- The associated system (2) always has a solution.
- The solutions of the normal equations have a *variational interpretation* as *least squares solutions*: they are the minimizers of the function $x \mapsto \|Ax - b\|^2$.

Our goal in this paper is to introduce a "normal problem" associated with the problem of finding a zero of the sum of two monotone operators. The solutions of this normal problem will agree with the solutions of the original problem provided the latter set is nonempty. The normal problem will also have a variational interpretation as well as attractive duality properties. We start developing the framework required to explain this in the following subsection.

1.2 The sum problem and Attouch–Théra duality

Throughout this paper,

$$X \text{ is a real Hilbert space with inner product } \langle\cdot,\cdot\rangle \tag{3}$$

and induced norm $\|\cdot\|$. Recall that a set-valued operator $A\colon X \rightrightarrows X$ (i.e., $(\forall x \in X)$ $Ax \subseteq X$) is *monotone* if $(\forall(x,x^*) \in \operatorname{gra} A)(\forall(y,y^*) \in \operatorname{gra} A)$ $\langle x-y, x^*-y^*\rangle \geq 0$; A is *maximally monotone* if A is monotone and it is impossible to extend A while keeping monotonicity. Since subdifferential operators of proper lower semicontinuous convex functions are maximally monotone, as are continuous linear operators with a monotone symmetric part, it is not surprising that maximally monotone operators play an important role in modern optimization and variational analysis. For relevant books on monotone operator theory and convex analysis we refer the reader to, e.g., [8], [13], [15], [18], [28], [29], [30], [31], [34], [35], [36], and [37]. From now on, we assume that

$$A \text{ and } B \text{ are maximally monotone operators on } X. \tag{4}$$

Because it encompasses the problem of finding solutions to constrained convex optimization problems, a key problem in monotone operator theory is to find a zero of the sum $A+B$. Let us formalize this now.

Definition 1.1 (primal problem) *The* primal problem *associated with the (ordered) pair (A,B) is to determine the set of zeros of the sum,*

$$Z_{(A,B)} := (A+B)^{-1}(0), \tag{5}$$

also referred to as the set of primal solutions. *When there is no cause for confusion, we will write Z instead of $Z_{(A,B)}$.*

Since addition is commutative, it is clear that the order of the operators A and B is irrelevant and thus $Z_{(A,B)} = Z_{(B,A)}$. In contrast, the order for the dual problem matters. Before we formally define the dual problem, we must introduce some notation. First,

$$A^{\circledcirc} := (-\operatorname{Id}) \circ A \circ (-\operatorname{Id}). \tag{6}$$

Note that $A^{\circledcirc}$ is also maximally monotone as is $(A^{-1})^{\circledcirc} = (A^{\circledcirc})^{-1}$, which motivates the definition[1]

$$A^{-\circledcirc} := \big(A^{-1}\big)^{\circledcirc} = \big(A^{\circledcirc}\big)^{-1}. \tag{7}$$

Definition 1.2 (dual pair and (Attouch–Théra) dual problem) *The* dual pair *of (A,B) is $(A,B)^* := (A^{-\circledcirc}, B^{-1})$. The* (Attouch–Théra) dual problem *associated with the pair (A,B) is to determine the set of zeros of the sum,*

$$K_{(A,B)} := \big(A^{-\circledcirc} + B^{-1}\big)^{-1}(0), \tag{8}$$

also referred to as the set of dual solutions. *When there is no cause for confusion, we will write K instead of $K_{(A,B)}$.*

[1] This is similar to the notation A^{-T} for the transpose of the inverse of an invertible matrix in Linear Algebra.

This duality, pioneered by Attouch and Théra [4] (see also [23, page 40]), has very attractive properties, including the following:

- $(A, B)^{**} = (A, B)$.
- The dual problem of (A, B) is precisely the primal problem of $(A, B)^*$.
- The set of primal solutions is nonempty if and only if the set of dual solutions is nonempty.

1.3 Aim of this paper

Not every sum problem admits a solution: suppose that $A = N_U$ and $B = N_V$, where U and V are nonempty closed convex subsets of X. It is clear that Z, the set of primal solutions associated with (A, B), is equal to $U \cap V$ — *however, this intersection may be empty* in which case the primal problem does not have any solution.

Our aim in this paper is to define a normal problem *associated with the original sum problem with attractive and useful properties.* Similarly to the complete extension of classical linear equations via normal equations (see Section 1.1), our proposed approach achieves the following:

- If the original problem has a solution, then so does the normal problem and the sets of solutions to these problems coincide.
- The normal problem may have a solution even if the original problem does not have any.
- The solutions of the normal problem have a variational interpretation as *infimal displacement* solutions related to the Douglas–Rachford splitting operator.
- The normal problem interacts well with Attouch–Théra duality.

Due to some technical results that need to be reviewed and developed, we postpone the actual derivation and definition of the normal problem until Section 3.2. We conclude this introductory section with some comments on the organization and notation of this paper.

1.4 Organization of the paper

The remainder of the paper is organized as follows. In Section 2, we review Attouch–Théra duality (Section 2.1), firmly nonexpansive operators and resolvents (Section 2.2), the

Douglas–Rachford splitting operator (Section 2.3), and we also provide some auxiliary results on perturbations (Section 2.4). Our main results are in Section 3. The normal problem is introduced in Section 3.2, after presenting results on perturbation duality (Section 3.1). Examples and directions for future research are discussed in Section 3.3 and 3.4, respectively.

Throughout, we utilize standard notation from convex analysis and monotone operator theory (see, e.g., [8], [28], [29], or [34]).

2 Auxiliary results

2.1 Solution mappings for Attouch–Théra duality

Definition 2.1 (solution mappings) *The* dual and primal solution mappings *associated with* (A, B) *are*

(9) $$\mathbf{K}\colon X \rightrightarrows X\colon x \mapsto (-Ax) \cap (Bx)$$

and

(10) $$\mathbf{Z}\colon X \rightrightarrows X\colon x \mapsto (-A^{-\circledcirc}x) \cap (B^{-1}x),$$

respectively.

Note that the primal solution mapping $\mathbf{Z}$ of (A, B) is the dual solution mapping of $(A, B)^*$ and analogously for $\mathbf{K}$. The importance of these mappings stems from the following result, which shows that the solutions mappings relate the sets of solutions Z and K to each other:

Fact 2.2 (See [4] or [6, Proposition 3.1].) $\operatorname{dom}\mathbf{K} = Z$*,* $\operatorname{ran}\mathbf{K} = K$*,* $\operatorname{dom}\mathbf{Z} = K$*,* $\operatorname{ran}\mathbf{Z} = Z$*, and* $\mathbf{Z} = \mathbf{K}^{-1}$*.*

2.2 Firmly nonexpansive operators and resolvents

Most of the material in this section is standard. Facts without explicit references may be found in, e.g., [8], [21], or [22].

Definition 2.3 *Let* $T\colon X \to X$*. Then* T *is* nonexpansive*, if*

(11) $$(\forall x \in X)(\forall y \in X) \quad \|Tx - Ty\| \le \|x - y\|.$$

Furthermore, T *is* firmly nonexpansive *if*

(12) $$(\forall x \in X)(\forall y \in X) \quad \|Tx - Ty\|^2 + \|(\operatorname{Id} - T)x - (\operatorname{Id} - T)y\|^2 \le \|x - y\|^2.$$

Clearly, every firmly nonexpansive mapping is nonexpansive.

Fact 2.4 *Let $T\colon X \to X$. Then T is firmly nonexpansive if and only if $2T - \operatorname{Id}$ is nonexpansive.*

Fact 2.5 (infimal displacement vector) (See, e.g., [5], [17], and [25].) *Let $T\colon X \to X$ be nonexpansive. Then $\overline{\operatorname{ran}}(\operatorname{Id} - T)$ is convex; consequently, the* infimal displacement vector *$v := P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}0$ is the unique element in $\overline{\operatorname{ran}}(\operatorname{Id} - T)$ such that $(\forall x \in X)\ \|v\| \leq \|x - Tx\|$.*

Lemma 2.6 *Let $T_1\colon X \to X$ and $T_2\colon X \to X$ be nonexpansive. Set $v_1 := P_{\overline{\operatorname{ran}}(\operatorname{Id} - T_1T_2)}0$ and $v_2 := P_{\overline{\operatorname{ran}}(\operatorname{Id} - T_2T_1)}0$. Then $\|v_1\| = \|v_2\|$.*

Proof. By definition of v_1, there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in X such that $\|x_n - T_1T_2x_n\| \to \|v_1\|$. Hence $(\forall n \in \mathbb{N})\ \|v_2\| \leq \|(T_2x_n) - T_2T_1(T_2x_n)\| \leq \|x_n - T_1T_2x_n\|$ and thus $\|v_2\| \leq \|v_1\|$. We see analogously that $\|v_1\| \leq \|v_2\|$. ■

Definition 2.7 (resolvent and reflected resolvent) *The* resolvent *of A is the operator*

$$J_A := (\operatorname{Id} + A)^{-1}, \tag{13}$$

and the reflected resolvent *is*

$$R_A := 2J_A - \operatorname{Id}. \tag{14}$$

Fact 2.8 *J_A is firmly nonexpansive and R_A is nonexpansive. Furthermore,*

$$J_A + J_{A^{-1}} = \operatorname{Id}. \tag{15}$$

Example 2.9 Let U be a nonempty closed convex subset of X, and suppose that $A = N_U$ is the corresponding normal cone operator. Then $J_A = P_U$ is the projection operator onto U and $R_A = 2P_U - \operatorname{Id}$ is the corresponding reflector.

Proposition 2.10 *Suppose that $A\colon X \to X$ is continuous, linear, and single-valued such that A and $-A$ are monotone, and $A^2 = -\alpha \operatorname{Id}$, where $\alpha \in \mathbb{R}_+$. Then*

$$J_A = \frac{1}{1+\alpha}\big(\operatorname{Id} - A\big) \quad \text{and} \quad R_A = \frac{1-\alpha}{1+\alpha}\operatorname{Id} - \frac{2}{1+\alpha}A. \tag{16}$$

Proof. We have

$$J_AJ_{-A} = (\operatorname{Id} + A)^{-1}(\operatorname{Id} - A)^{-1} = \big((\operatorname{Id} - A)(\operatorname{Id} + A)\big)^{-1} = (\operatorname{Id} - A^2)^{-1} = (\operatorname{Id} + \alpha \operatorname{Id})^{-1} \tag{17a}$$

$$= \frac{1}{1+\alpha}\operatorname{Id}. \tag{17b}$$

It follows that $J_A = (1+\alpha)^{-1} J_{-A}^{-1} = (1+\alpha)^{-1}(\operatorname{Id} - A)$ and hence that

$$R_A = 2J_A - \operatorname{Id} = \frac{2}{1+\alpha}(\operatorname{Id} - A) - \operatorname{Id} = \frac{1-\alpha}{1+\alpha}\operatorname{Id} - \frac{2}{1+\alpha}A, \tag{18}$$

as claimed. ■

Example 2.11 Suppose that $X = \mathbb{R}^2$ and that $A\colon \mathbb{R}^2 \to \mathbb{R}^2\colon (x,y) \mapsto (-y,x)$ is the rotator by $\pi/2$. Then $A^2 = -\operatorname{Id}$; consequently, by Proposition 2.10, $J_A = (1/2)(\operatorname{Id} - A)$ and $R_A = -A$.

2.3 The Douglas–Rachford splitting operator

Definition 2.12 *The* Douglas–Rachford splitting operator *associated with* (A,B) *is*

$$T := T_{(A,B)} := J_A R_B + \operatorname{Id} - J_B. \tag{19}$$

We will simply use T *instead of* $T_{A,B}$ *provided there is no cause for confusion.*

Fact 2.13 *The following hold:*

(i) $T_{(A,B)} = \frac{1}{2}(\operatorname{Id} + R_A R_B)$; *consequently,* $T_{(A,B)}$ *is firmly nonexpansive.*

(ii) **(Eckstein)** (See [20, Lemma 3.6].) $T_{(A,B)} = T_{(A,B)^*} = T_{(A^{-\circledast}, B^{-1})}$.

(iii) **(Eckstein)** (See [20, Proposition 4.1].)

$$\operatorname{gra}(T) = \big\{(b+b^*, a+b^*) \;\big|\; (a,a^*) \in \operatorname{gra} A, (b,b^*) \in \operatorname{gra} B, b-a = b^*+a^*\big\}. \tag{20}$$

Corollary 2.14 *We have*

$$\operatorname{gra}(\operatorname{Id} - T) = \big\{(b+b^*, b-a) \;\big|\; (a,a^*) \in \operatorname{gra} A, (b,b^*) \in \operatorname{gra} B, b-a = b^*+a^*\big\}; \tag{21}$$

consequently,

$$\operatorname{ran}(\operatorname{Id} - T) = \big\{b-a \;\big|\; (a,a^*) \in \operatorname{gra} A,\ (b,b^*) \in \operatorname{gra} B,\ b-a = b^*+a^*\big\} \tag{22a}$$

$$\subseteq (\operatorname{dom} B - \operatorname{dom} A) \cap (\operatorname{ran} A + \operatorname{ran} B). \tag{22b}$$

Remark 2.15 Suppose that A is a linear operator (or, even more generally, a linear relation). Then J_A is linear (see [11, Theorem 2.1(xviii)]) and $\operatorname{Id} - T$ simplifies to $J_A + J_B - 2J_AJ_B$. This setting occurs sometimes in applications; see, e.g., [3, Section 4.3].

It is clear from the definition and Fact 2.13(i) that $\operatorname{Id} - T_{A,B}$ is also firmly nonexpansive. In fact, we note in passing that $\operatorname{Id} - T_{A,B}$ is itself a Douglas–Rachford splitting operator:

Proposition 2.16 $\operatorname{Id} - T_{(A,B)} = T_{(A^{-1},B)}$.

Proof. Using (15), we obtain

$$T_{A,B} + T_{A^{-1},B} = \operatorname{Id} - J_B + J_A R_B + \operatorname{Id} - J_B + J_{A^{-1}} R_B \tag{23a}$$
$$= 2\operatorname{Id} - 2J_B + (J_A + J_{A^{-1}})R_B \tag{23b}$$
$$= 2\operatorname{Id} - 2J_B + R_B \tag{23c}$$
$$= \operatorname{Id}, \tag{23d}$$

and the conclusion follows. ■

Fact 2.17 (See [6, Theorem 4.5].) *The mapping*

$$\Psi\colon \operatorname{gra} \mathbf{K} \to \operatorname{Fix} T\colon (z,k) \mapsto z + k \tag{24}$$

is a well defined bijection that is continuous in both directions, with $\Psi^{-1}\colon x \mapsto (J_B x, x - J_B x)$.

Corollary 2.18 (Combettes) (See [19, Lemma 20.6(iii)].) $J_B(\operatorname{Fix} T) = Z$.

2.4 Perturbation calculus

Definition 2.19 (shift operator and corresponding inner/outer perturbations) *Let* $w \in X$. *We define the associated* shift operator

$$S_w\colon X \to X\colon x \mapsto x - w, \tag{25}$$

and we extend S_w *to deal with subsets of* X *by setting* $(\forall C \subseteq X)\ S_w(C) := \bigcup_{c\in C}\{S_w(c)\}$. *We define the corresponding* inner and outer perturbations *of* A *by*

$$[A; w] := A \circ S_w\colon X \rightrightarrows X\colon x \to A(x - w), \tag{26}$$

and

$$[w; A] := S_w \circ A\colon X \rightrightarrows X\colon x \to Ax - w. \tag{27}$$

Observe that if $w \in X$, then the operators $[A; w]$ and $[w; A]$ are maximally monotone, with domains $S_{-w}(\operatorname{dom} A) = w + \operatorname{dom} A$ and $\operatorname{dom} A$, respectively.

Lemma 2.20 (perturbation calculus) *Let* $w \in X$. *Then the following hold:*

(i) $[A; w]^{-1} = [-w; A^{-1}]$.

(ii) $[w;A]^{-1} = [A^{-1}; -w]$.

(iii) $[A;w]^{\ovee} = [A^{\ovee}; -w]$.

(iv) $[w;A]^{\ovee} = [-w; A^{\ovee}]$.

(v) $[A;w]^{-\ovee} = [w; A^{-\ovee}]$.

(vi) $[w;A]^{-\ovee} = [A^{-\ovee}; w]$.

Proof. Let $(x,y) \in X \times X$. (i): $y \in [A;w]^{-1}x \Leftrightarrow x \in [A;w]y = A(y-w) \Leftrightarrow y - w \in A^{-1}x \Leftrightarrow y \in A^{-1}x + w = [-w; A^{-1}]x$. (ii): $y \in [w;A]^{-1}x \Leftrightarrow x \in [w;A]y \Leftrightarrow x \in Ay - w \Leftrightarrow x + w \in Ay \Leftrightarrow y \in A^{-1}(x+w) = [A^{-1}; -w]x$. (iii): $[A;w]^{\ovee} = -[A;w](-x) = -A(-x-w) = A^{\ovee}(x+w) = [A^{\ovee}; -w]$. (iv): $[w;A]^{\ovee}x = -[w;A](-x) = -(A(-x) - w) = A^{\ovee}x - (-w) = [-w; A^{\ovee}]x$. (v): Using (i) and (iv), we see that $[A;w]^{-\ovee} = ([A;w]^{-1})^{\ovee} = [-w; A^{-1}]^{\ovee} = [w; A^{-\ovee}]$. (vi): Using (ii) and (iii), we see that $[w;A]^{-\ovee} = ([w;A]^{-1})^{\ovee} = [A^{-1}; -w]^{\ovee} = [A^{-\ovee}; w]$. ■

As an application, we record the following result which will be useful later.

Corollary 2.21 (dual of inner-outer perturbation) *Let $w \in X$. Then*

$$\big([A;w], [w;B]\big)^* = \big([A;w]^{-\ovee}, [w;B]^{-1}\big) = \big([w; A^{-\ovee}], [B^{-1}; -w]\big). \tag{28}$$

Proof. Combine Definition 1.2 with Lemma 2.20(v)&(ii). ■

2.5 Perturbations of the Douglas–Rachford operator

We now turn to the Douglas–Rachford operator.

Proposition 2.22 *Let $w \in X$. Then the following hold:*

(i) *If $x \in \operatorname{Fix}[-w;T]$, then $x - w - J_Bx \in [w;B]J_Bx \cap (-[A;w]J_Bx)$.*

(ii) *If $y \in [w;B]z \cap (-[A;w]z)$, then $x := w + y + z \in \operatorname{Fix}[-w;T]$ and $z = J_Bx$.*

Proof. If $x \in X$, then $x - w - J_Bx \in [w;B]J_Bx$.

(i): Since $x \in \operatorname{Fix}([-w;T])$, we have $x - Tx = w$; equivalently, $J_Bx - w = J_AR_Bx$. Hence $2J_Bx - x = R_Bx \in (A + \operatorname{Id})(J_Bx - w) = [A;w]J_Bx + J_Bx - w$ and thus $-(x - w - J_Bx) \in [A;w]J_Bx$.

(ii): Since $y \in [w;B]z \cap (-[A;w]z) = (Bz - w) \cap (-A(z-w))$, we have $z = J_Bx$ and $z - w = J_A(-y + z - w)$. Hence $R_Bx = 2J_Bx - x = 2z - (w + y + z) = z - w - y$ and so $J_AR_Bx = J_A(z - w - y) = z - w$. Thus, $x - Tx = J_Bx - J_AR_Bx = z - (z - w) = w$. ■

Corollary 2.23 *Let $w \in X$. Then* $\operatorname{Fix}[-w;T] = w + \bigcup_{z\in X} \big(z + [w;B]z \cap (-[A;w]z)\big)$.

Proposition 2.24 *Let $w \in X$. Then*

$$T_{([A;w],[w;B])} = [T;-w] \tag{29}$$

and

$$\operatorname{Fix}[T;-w] = -w + \operatorname{Fix}[-w;T] = \bigcup_{z\in X} \Big(z + \big((Bz - w) \cap (-A(z-w))\big)\Big). \tag{30}$$

Proof. Let $x \in X$. Using, e.g., [8, Proposition 23.15], we obtain $J_{[A;w]}x = J_A(x-w) + w$ and $J_{[w;B]}x = J_B(x+w)$. Consequently, $R_{[A;w]}x = 2J_A(x-w) + 2w - x$ and $R_{[w;B]}x = 2J_B(x+w) - x$. It thus follows with Definition 2.12 that

$$\begin{aligned}
T_{([A;w],\ [w;B])}x &= x - J_{[w;B]}x + J_{[A;w]}R_{[w;B]}x && \text{(31a)}\\
&= x - J_B(x+w) + J_A\big(2J_B(x+w) - x - w\big) + w && \text{(31b)}\\
&= (x+w) - J_B(x+w) + J_A\big(R_B(x+w)\big) && \text{(31c)}\\
&= T(x+w) = [T;-w]x, && \text{(31d)}
\end{aligned}$$

and so (29) holds. Next, $x \in \operatorname{Fix}[T;-w] \Leftrightarrow x = T(x+w) \Leftrightarrow x + w = w + T(x+w) \Leftrightarrow x + w \in \operatorname{Fix}[-w;T]$, and have thus verified the left identity in (30). To see the right identity in (30), use Corollary 2.23. ■

We now obtain a generalization of Fact 2.17, which corresponds to the case when $w = 0$.

Proposition 2.25 *Let $w \in X$ and define*

$$\mathbf{K}_w \colon X \rightrightarrows X \colon x \mapsto (-A(x-w)) \cap (Bx - w). \tag{32}$$

Then

$$\Psi_w \colon \operatorname{gra}\mathbf{K}_w \to \operatorname{Fix}[-w;T] \colon (z,k) \mapsto z + k + w \tag{33}$$

is a well defined bijection that is continuous in both directions, with $\Psi_w^{-1} \colon x \mapsto (J_Bx, x - J_Bx - w)$.

Proof. For the pair $([A;w],[w;B])$, the dual solution mapping is $\mathbf{K}_w$ and the Douglas–Rachford operator is $[T;-w]$ by (29). Applying Fact 2.17 in this context, we obtain

$$\Phi \colon \operatorname{gra}\mathbf{K}_w \to \operatorname{Fix}[T;-w] \colon (z,k) \mapsto z + k \tag{34}$$

is continuous in both directions with $\Phi^{-1} \colon x \mapsto (J_{[w;B]}x, x - J_{[w;B]}x) = (J_B(x+w), x - J_B(x+w))$. Furthermore, S_{-w} is a bijection from $\operatorname{Fix}[T;-w]$ to $\operatorname{Fix}[-w;T]$ by (30). This shows that $\Psi_w = S_{-w} \circ \Phi$ and the result follows. ■

3 The normal problem

3.1 The w-perturbed problem

Definition 3.1 (w-perturbed problem) *Let $w \in X$. The* w-perturbation *of (A, B) is $([A; w], [w; B])$. The* w-perturbed problem *associated with the pair (A, B) is to determine the set of zeros*

$$(35) \qquad Z_w := Z_{([A;w],[w;B])} = \big([A; w] + [w; B]\big)^{-1}(0) = \big\{x \in X \bigm| w \in A(x - w) + Bx\big\}.$$

Note that the w-perturbed problem of (A, B) is precisely the primal problem of $([A; w], [w; B])$, i.e., of the w-perturbation of (A, B).

Before discussing the Douglas–Rachford operator of the w-perturbation, we point out that there are other seemingly unrelated notions of generalizing the sum; e.g., the variational sum introduced by Attouch, Baillon, and Théra [2] and the extended sum introduced by Revalski and Théra. (For a recent nice survey, see [26].) We also note that our perturbation is based on the inner perturbation $[A; w]$ and the outer perturbation $[w; B]$; this appears to make it different from the usual perturbation approaches based solely on inner perturbations; see, e.g., [14].

Proposition 3.2 (Douglas–Rachford operator of the w-perturbation) *Let $w \in X$. Then the Douglas–Rachford operator of the w-perturbation $([A; w], [w; B])$ of (A, B) is*

$$(36) \qquad T_{([A;w],[w;B])} = [T; -w].$$

Proof. This follows from (29) of Proposition 2.24. ■

Proposition 3.3 *Let $w \in X$. Then*

$$(37) \qquad Z_w = J_{[w;B]}\big(\operatorname{Fix}[T; -w]\big) = J_B\big(w + \big\{x \in X \bigm| x = T(x + w)\big\}\big).$$

Furthermore, the following are equivalent:

(i) $Z_w \neq \varnothing$.

(ii) $\operatorname{Fix}[T; -w] \neq \varnothing$.

(iii) $w \in \operatorname{ran}(\operatorname{Id} - T)$.

(iv) $w \in \operatorname{ran}([A; w] + B)$.

Proof. The identity (37) follows by combining Corollary 2.18 with Proposition 3.2. This also yields the equivalence of (i) and (ii). Let $x \in X$. Then $x \in Z_w \Leftrightarrow 0 \in [A;w]x + [w;B]x \Leftrightarrow w \in [A;w]x + Bx$, and we deduce the equivalence of (i) and (iv). Finally, $x \in \operatorname{Fix}[T;-w] \Leftrightarrow x = T(x+w) \Leftrightarrow w \in (\operatorname{Id} - T)(x+w)$, which yields the equivalence of (ii) and (iii). ■

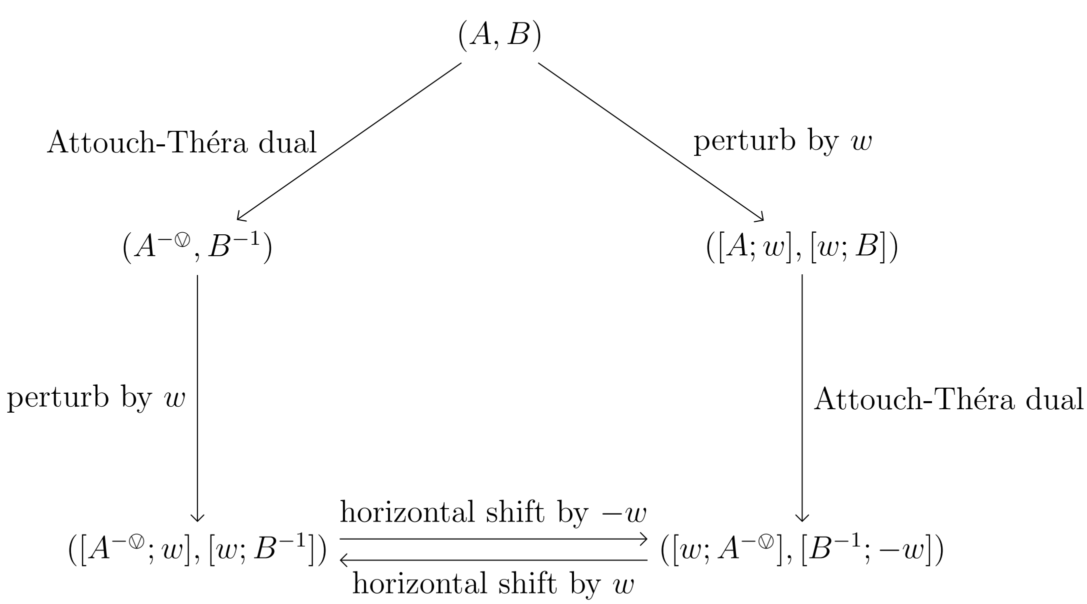
The equivalence of (i) and (iii) yields the following key result on which w-perturbations have nonempty solution sets.

Corollary 3.4 $\big\{w \in X \;\big|\; Z_w \neq \varnothing\big\} = \operatorname{ran}(\operatorname{Id} - T)$.

Remark 3.5 (Attouch–Théra dual of the perturbed problem) Consider the given pair of monotone operators (A, B). We could either first perturb and then take the Attouch–Théra dual or start with the Attouch–Théra dual and then perturb. It turns out that the order of these operations does not matter — up to a horizontal shift of the graphs. Indeed, for every $x \in X$, we have

$$\begin{aligned}
\big([A^{-\varovee};w] + [w;B^{-1}]\big)x &= A^{-\varovee}(x-w) + B^{-1}x - w && \text{(38a)}\\
&= A^{-\varovee}(x-w) - w + B^{-1}((x-w)+w) && \text{(38b)}\\
&= [w;A^{-\varovee}](x-w) + [B^{-1};-w](x-w) && \text{(38c)}\\
&= \big([w;A^{-\varovee}] + [B^{-1};-w]\big)(x-w). && \text{(38d)}
\end{aligned}$$

Hence $\operatorname{gra}([A^{-\varovee};w] + [w;B^{-1}]) = (w,0) + \operatorname{gra}([w;A^{-\varovee}] + [B^{-1};-w])$, which gives rise to the following diagram:



The following remark was prompted by a comment of a referee.

Remark 3.6 (forward-backward opeator) Suppose that B is at most single-valued. Then it is well known (see, e.g., [8, Proposition 25.1(iv)]) that

$$Z = \operatorname{Fix} F, \quad \text{where } F := J_A \circ (\operatorname{Id} - B) \tag{39}$$

is the *forward-backward operator*. Now let $w \in X$ and $x \in X$. We then have the equivalences $w = x - Fx \Leftrightarrow x - w = J_A(x - Bx) \Leftrightarrow x - Bx \in (\operatorname{Id} + A)(x - w) \Leftrightarrow w \in [A; w]x + Bx$. Combining with Proposition 3.3, we obtain

$$\big\{w \in X \;\big|\; Z_w \neq \varnothing\big\} = \operatorname{ran}(\operatorname{Id} - F), \tag{40}$$

which can be interpreted as the forward-backward counterpart of Corollary 3.4. Note, however, that no attractive duality theory is available in general since B^{-1} may fail to be at most single-valued.

3.2 The normal problem

We are now in a position to define the normal problem.

Definition 3.7 (infimal displacement vector and the normal problem) *The vector*

$$v(A, B) = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)} 0 \tag{41}$$

is the infimal displacement vector of (A, B). *The* normal problem *associated with* (A, B) *is the* $v(A, B)$*-perturbed problem of* (A, B)*, and the set of* normal solutions *is* $Z_{v(A,B)}$.

Remark 3.8 (new notions are well defined) The notions presented in Definition 3.7 are *well defined*: indeed, since T is firmly nonexpansive (Fact 2.13(i)), it is also nonexpansive and the existence and uniqueness of $v(A, B)$ follows from Fact 2.5.

Remark 3.9 (new notions extend original notions) Suppose that for the original problem (A, B), we have $Z = Z_0 = (A + B)^{-1}(0) \neq \varnothing$. By Corollary 3.4, $0 \in \operatorname{ran}(\operatorname{Id} - T)$ and so $v(A, B) = 0$. Hence the normal problem coincides with the original problem, as do the associated sets of solutions. Furthermore, as Wang [32] recently showed, Z is generically a singleton.

Remark 3.10 (normal problem may or may not have solutions) If the set of original solutions Z is empty, then the set of normal solutions may be either nonempty (see Example 3.18) or empty (see Example 3.19).

The original problem of finding a zero of $A + B$ is clearly symmetric in A and B. We now present a statement about the *magnitude* of the corresponding infimal displacement vectors:

Proposition 3.11 $\|v(A,B)\| = \|v(B,A)\|$.

Proof. It follows from Fact 2.13(i) that

$$\mathrm{Id} - T_{(A,B)} = \tfrac{1}{2}(\mathrm{Id} - R_A R_B) \quad \text{and} \quad \mathrm{Id} - T_{(B,A)} = \tfrac{1}{2}(\mathrm{Id} - R_B R_A). \tag{42}$$

Thus, using Lemma 2.6, we see that $\|v(A,B)\| = 2\|P_{\overline{\mathrm{ran}}(\mathrm{Id} - R_A R_B)}0\| = 2\|P_{\overline{\mathrm{ran}}(\mathrm{Id} - R_A R_B)}0\| = \|v(B,A)\|$. ■

Remark 3.12 ($v(A,B) \neq v(B,A)$ may occur) We will see in the sequel examples where $v(A,B) \neq 0$ but (i) $v(B,A) = -v(A,B)$ (see Remark 3.17); (ii) $v(B,A) \perp v(A,B)$ (see Example 3.20); or (iii) $v(A,B) = v(B,A)$ (see Example 3.22).

Remark 3.13 (self-duality: $v(A,B) = v(A^{-\circledcirc}, B^{-1})$) Since, by Fact 2.13(ii), $T_{(A,B)} = T_{(A^{\circledcirc}, B^{-1})}$, it is clear that $v(A,B) = v(A^{-\circledcirc}, B^{-1})$. It follows from Remark 3.5 that the operations of perturbing by $v(A,B)$ and taking the Attouch–Théra dual commute, up to a shift.

Remark 3.14 (computing $v(A,B)$) Note that the infimal displacement vector can be found as

$$(\forall x \in X) \quad v(A,B) = -\lim_{n\to\infty} \frac{T^n x}{n} = \lim_{n\to\infty} T^n x - T^{n+1}x; \tag{43}$$

see [5], [17], and [25]. Conceptionally, we can thus first find $v(A,B)$ via either iteration in (43), and followed by iterating the operator $x \mapsto T(x + v(A,B))$ (which is justified by Proposition 3.3) to find a normal solution.

3.3 Examples

Proposition 3.15 $(\mathrm{Id} - J_A)B^{-1}0 \subseteq \mathrm{ran}(\mathrm{Id} - T) \subseteq \mathrm{dom}\, B - \mathrm{dom}\, A$.

Proof. The right inclusion follows from (22). To tackle the left inclusion, suppose that $z \in B^{-1}0$ and set $w := z - J_A z$. Then $w = z - J_A z \in A(J_A z) = A(z-w) + 0 \subseteq [A; w](z) + Bz$. Hence, by Proposition 3.3, $w \in \mathrm{ran}(\mathrm{Id} - T)$. ■

Proposition 3.16 (normal cone operators) *Suppose that $A = N_U$ and $B = N_V$, where U and V are nonempty closed convex subsets of X. Then*

$$v(A,B) = P_{\overline{V-U}}0 \tag{44}$$

and the set of normal solutions is

$$V \cap (v(A,B) + U) = \mathrm{Fix}(P_V P_U). \tag{45}$$

Proof. Since $B^{-1}0 = V$ and $J_A = P_U$, Proposition 3.15 yields $C := \big\{v - P_U v \;\big|\; v \in V\big\} \subseteq \operatorname{ran}(\operatorname{Id} - T) \subseteq V - U$; hence,

$$\overline{C} \subseteq \overline{\operatorname{ran}}(\operatorname{Id} - T) \subseteq \overline{V - U}. \tag{46}$$

Set $g := P_{\overline{V-U}}0$. By [7, Theorem 4.1], there exists a sequence $(v_n)_{n\in\mathbb{N}}$ in V such that $v_n - P_U v_n \to g$. It follows that $(v_n - P_U v_n)_{n\in\mathbb{N}}$ lies in C and hence that $g \in \overline{C}$. Therefore $P_{\overline{C}}0 = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}0 = P_{\overline{V-U}}0$ and we obtain (44). (For an alternative proof, see [9, Theorem 3.5].)

Let $x \in X$. Then x is a normal solution if and only if

$$g \in N_U(x - g) + N_V(x). \tag{47}$$

Assume first that (47) holds. Then $x \in V$ and $x - g \in U$. Hence $x \in V \cap (g + U) = \operatorname{Fix}(P_V P_U)$ by [7, Lemma 2.2]. Conversely, assume $x \in V \cap (g + U) = \operatorname{Fix}(P_V P_U)$. Then $x \in V$, $x - g \in U$, $P_U x = x - g$ and $P_V(x - g) = x$. Hence $N_U(x - g) \supseteq \mathbb{R}_+ g$ and $N_V(x) \supseteq \mathbb{R}_- g$; consequently, $N_U(x - g) + N_V(x) \supseteq \mathbb{R}_+ g + \mathbb{R}_- g = \mathbb{R}g \ni g$ and therefore (47) holds. ■

Remark 3.17 Proposition 3.16 is consistent with the theory dealing with inconsistent feasibility problems (see, e.g., [7]). Note that it also yields the formula

$$v(A, B) = -v(B, A) \tag{48}$$

in this particular context.

Example 3.18 (no original solutions but normal solutions exist) Suppose that A and B are as in Proposition 3.16, that $U \cap V = \varnothing$, and V is also bounded. Then $\operatorname{Fix}(P_V P_U) \neq \varnothing$ by the Browder–Göhde–Kirk fixed point theorem (see, e.g., [8, Theorem 4.19]). So, the original problem has no solution but there exist normal solutions.

Example 3.19 (neither original nor normal solutions exist) Suppose that $X = \mathbb{R}^2$, that A and B are as in Proposition 3.16, that $U = \mathbb{R} \times \{0\}$, and that $V = \big\{(x, y) \in \mathbb{R}^2 \;\big|\; \beta + \exp(x) \leq y\big\}$, where $\beta \in \mathbb{R}_+$. Then $v(A, B) = (0, \beta)$ yet $\operatorname{Fix}(P_V P_U) = \varnothing$.

Example 3.20 Suppose that $X = \mathbb{R}^2$, let $L\colon \mathbb{R}^2 \to \mathbb{R}^2\colon (\xi, \eta) \mapsto (-\eta, \xi)$ be the rotator by $\pi/2$, let $a^* \in \mathbb{R}^2$ and $b^* \in \mathbb{R}^2$. Suppose that $(\forall x \in \mathbb{R}^2)$ $Ax = Lx + a^*$ and $Bx = -Lx - b^*$. Now let $(x, w) \in X \times X$. Then $x \in Z_w$ if and only if $0 = A(x - w) + Bx - w = L(x - w) + a^* - Lx - b^* - w$ and so $(\operatorname{Id} + L)w = a^* - b^*$, i.e., $w = J_L(a^* - b^*) = (1/2)(\operatorname{Id} - L)(a^* - b^*)$ by Example 2.11. It follows that

$$v(A, B) = \tfrac{1}{2}(\operatorname{Id} - L)(a^* - b^*). \tag{49}$$

An analogous argument yields

$$v(B, A) = \tfrac{1}{2}(\operatorname{Id} + L)(b^* - a^*). \tag{50}$$

Setting $d^* = b^* - a^*$, we have $4\langle v(A,B), v(B,A)\rangle = \langle Ld^* - d^*, Ld^* + d^*\rangle = \|Ld^*\|^2 - \|d^*\|^2 = 0$. and $v(A,B) + v(B,A) = Ld^*$. Thus if $d^* \neq 0$, i.e., $a^* \neq b^*$, then

$$v(A,B) \neq 0 \ \text{ and } \ v(A,B) \perp v(B,A). \tag{51}$$

The following remark was prompted by a comment of a referee.

Remark 3.21 (backward-backward operator) Consider Example 3.20. Using Proposition 2.10 (see also Example 2.11), we have $J_L = (\operatorname{Id} - L)/2$ and $J_{-L} = (\operatorname{Id} + L)/2$. Hence, by e.g. [8, Proposition 23.15(ii)], $J_A\colon x \mapsto J_L(x - a^*)$ and $J_B\colon x \mapsto J_{-L}(x + b^*)$. Now write $x = (\xi_1, \xi_2)$, $a^* = (\alpha_1, \alpha_2)$, and $b^* = (\beta_1, \beta_2)$. Then $\operatorname{Id} - J_BJ_A\colon (\xi_1, \xi_2) \mapsto (1/2)(\xi_1 - \beta_1 - \beta_2, \xi_2 + \alpha_1 + \alpha_2 - \beta_1 - \beta_2)$ and $\operatorname{Id} - J_AJ_B\colon (\xi_1, \xi_2) \mapsto (1/2)(\xi_1 + \alpha_1 + \alpha_2 - \beta_1 - \beta_2, \xi_2 + \alpha_1 + \alpha_2)$. Consequently, $\operatorname{ran}(\operatorname{Id} - J_BJ_A) = \operatorname{ran}(\operatorname{Id} - J_AJ_B) = X$ and the vectors $P_{\overline{\operatorname{ran}}(\operatorname{Id} - J_BJ_A)}(0)$ and $P_{\overline{\operatorname{ran}}(\operatorname{Id} - J_AJ_B)}(0)$, which both coincide with 0, do *not* indicate the lack of zeros of $A + B$. Hence, in our analysis, we cannot replace the Douglas–Rachford operator by the backward-backward operator (J_BJ_A or J_AJ_B), which occurs in the study of semigroups and the celebrated Trotter–Lie–Kato formula (see, e.g., [16, Section 3]), and expect similar results. (For systematic studies of the asymptotic behaviour of the two resolvents, see [10] and [33].)

Example 3.22 Suppose that there exists a^* and b^* in X such that $\operatorname{gra} A = X \times \{a^*\}$ and $\operatorname{gra} B = X \times \{b^*\}$. By (22), $\varnothing \neq \operatorname{ran}(\operatorname{Id} - T) \subseteq \{a^* + b^*\}$. Hence $v(A,B) = a^* + b^*$ and analogously $v(B,A) = a^* + b^*$. Thus, if $a^* + b^* \neq 0$, we have

$$v(A,B) \neq 0 \ \text{ and } \ v(A,B) = v(B,A). \tag{52}$$

Proposition 3.23 *Suppose that there exists continuous linear monotone operators L and M on X, and vectors a^* and b^* in X such that $(\forall x \in X)$ $Ax = Lx + a^*$ and $Bx = Mx + b^*$. Consider the problem*

$$\text{minimize } \|w\|^2 \quad \text{subject to} \quad (w,x) \in X \times X \text{ and } (\operatorname{Id} + L)w - (L + M)x = a^* + b^*. \tag{53}$$

Let $(w,x) \in X \times X$. Then (w,x) solves (53) $\Leftrightarrow$ $w = v(A,B)$ and x is a normal solution $\Leftrightarrow$ $w = P_{J_L(\operatorname{ran}(A+B))}0$ and $x \in (A+B)^{-1}(\operatorname{Id} + L)w$.

Proof. Then $w = [A;w]x + Bx \Leftrightarrow (\operatorname{Id} + L)w - (L + M)x = a^* + b^* \Leftrightarrow (\operatorname{Id} + L)w = (L + M)x + a^* + b^* \Leftrightarrow w = J_L\big((L + M)x + a^* + b^*\big) = J_L(A + B)x$. The conclusion thus follows from Proposition 3.3. ■

It is nice to recover a special case of our original motivation given in Section 1.1:

Example 3.24 (classical least squares solutions) Suppose that $X = \mathbb{R}^n$, let $M \in \mathbb{R}^{n\times n}$ be such that $M + M^*$ is positive semidefinite, and let $b \in \mathbb{R}^n$. Suppose that $(\forall x \in \mathbb{R}^n)$ $Ax = -b$ and $B = M$ so that the original problem is to find $x \in \mathbb{R}^n$ such that $Mx = b$. Then $v(A,B) = P_{\operatorname{ran} M}(b) - b$ and the normal solutions are precisely the least squares solutions.

Proof. We will use Proposition 3.23. The constraint in (53) turns into $(\mathrm{Id}+0)w-(0+M)x = 0+(-b)$, i.e., $w = Mx - b$ so that the optimization problem in (53) is

$$\text{minimize } \|Mx - b\|^2. \tag{54}$$

Hence the normal solutions in our sense are precisely the classical least squares solutions. Furthermore, $v(A,B) = P_{\mathrm{ran}(A+B)}0 = P_{-b+\mathrm{ran}\,M}(0) = P_{\mathrm{ran}\,M}(b) - b$. ■

3.4 Future research

We conclude by outlining some research directions:

- We noted in Remark 3.14 that $v := v(A,B)$ can be found by iterating (43). Subsequently iterating $x \mapsto T(x+v)$ will then lead to a normal solution. It would be desirable to devise an algorithm that approximates $v(A,B)$ and a corresponding normal solution (should it exist) *simultaneously*. Proposition 3.23, which leads us to solving a quadratic optimization problem, suggests that this may indeed be possible in general.
- Another avenue for future research is to consider more general sums of the form $A + L^*BL$, where L is a linear operator.
- It would be useful to obtain sufficient conditions for stability: Assuming that $A_n \to A$ and $B_n \to B$ in an appropriate sense (see, e.g., [1] and [29]), when can we conclude that the infimial displacement vectors $v(A_n, B_n)$ and the correponding normal solutions of (A_n, B_n) converge to those of (A,B)?
- Finally, it would be interesting to investigate our approach further in other areas including optimal control (see, e.g., [3, Section 4.3]) or PDEs (see, e.g., the study of the Schrödinger equation in [2, Section 9]).

Acknowledgments

We thank the referees for their careful reading and constructive comments. We also thank Patrick Combettes and Shawn Wang for helpful discussions. HHB was supported by the Natural Sciences and Engineering Research Council of Canada and the Canada Research Chair Program. WLH was supported by the Natural Sciences and Engineering Research Council of Canada. WMM was partially supported by NSERC grants of HHB and WLH.

References

[1] H. Attouch, *Variational Convergence for Functions and Operators*, Pitman, Boston, MA, 1984.

[2] H. Attouch, J.-B. Baillon, and M. Théra, Variational sum of monotone operators, *Journal of Convex Analysis* 1 (1994), 1–29.

[3] H. Attouch, M.-O. Czarnecki, and J. Peypouquet, Prox-penalization and splitting methods for constrained variational problems, *SIAM Journal on Optimization* 21 (2011), 149–173.

[4] H. Attouch and M. Théra, A general duality principle for the sum of two operators, *Journal of Convex Analysis* 3 (1996), 1–24.

[5] 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(1) (1978), 1–9.

[6] H.H. Bauschke, R.I. Boţ, W.L. Hare, and W.M. Moursi, Attouch-Théra duality revisited: paramonotonicity and operator splitting, *Journal of Approximation Theory* 164 (2012), 1065–1084.

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

[8] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, 2011.

[9] H.H. Bauschke, P.L. Combettes, and D.R. Luke, Finding best approximation pairs relative to two closed convex sets in Hilbert spaces, *Journal of Approximation Theory* 127 (2004), 178–192.

[10] H.H. Bauschke, P.L. Combettes, and S. Reich, The asymptotic behavior of the composition of two resolvents, *Nonlinear Analysis* 60 (2005), 283–301.

[11] H.H. Bauschke, S.M. Moffat, and X. Wang, Firmly nonexpansive mappings and maximally monotone operators: correspondence and duality, *Set-Valued and Variational Analysis* 20 (2012), 131–153.

[12] Å. Björck, *Numerical Methods for Least Squares Problems*, SIAM, 1996.

[13] J.M. Borwein and J.D. Vanderwerff, *Convex Functions*, Cambridge University Press, 2010.

[14] R.I. Boţ, *Conjugate Duality in Convex Optimization*, Springer, 2010.

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

[16] H. Brezis and A. Pazy, Semigroups of nonlinear contractions on convex sets, *Journal of Functional Analysis* 6 (1970), 237–281.

[17] R.E. Bruck and S. Reich, Nonexpansive projections and resolvents of accretive operators in Banach spaces, *Houston Journal of Mathematics* 3(4) (1977), 459–470.

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

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

[20] J. Eckstein, *Splitting Methods for Monotone Operators with Applications to Parallel Optimization*, Ph.D. thesis, MIT, 1989.

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

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

[23] B. Mercier, *Inéquations Variationnelles de la Mécanique* (Publications Mathématiques d'Orsay, no. 80.01), Orsay, France: Université de Paris-XI, 1980. `http://portail.mathdoc.fr/PMO/PDF/M_MERCIER-87.pdf`

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

[25] A. Pazy, Asymptotic behavior of contractions in Hilbert space, *Israel Journal of Mathematics* 9, 235–240 (1971).

[26] J.P. Revalski, Regularization procedures for monotone operators: recent advances, in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H.H. Bauschke et al. (eds), Springer, 2011.

[27] J.P. Revalski and M. Théra, Generalized sums of monotone operators, *Comptes Rendus de l'Académie des Sciences* 329 (1999), 979–984.

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

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

[30] S. Simons, *Minimax and Monotonicity*, Springer, 1998.

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

[32] X. Wang, Most maximally monotone operators have a unique zero and super-regular resolvent, *Nonlinear Analysis* 87 (2013), 69–82.

[33] X. Wang and H.H. Bauschke, Compositions and averages of two resolvents: relative geometry of fixed point sets and a partial answer to a question by C. Byrne, *Nonlinear Analysis* 74 (2011), 4550–4572.

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

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

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

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