

Analysis of the Three-operator Davis-Yin Splitting in the Inconsistent Case

by

Andrew Naguib

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2025

© Andrew Naguib 2025

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis analyzes the Davis–Yin three-operator splitting method in the inconsistent case, where the underlying monotone inclusion problem may fail to have a solution. The Davis–Yin algorithm extends the Douglas–Rachford and forward–backward splitting methods and is effective in reformulating optimization and inclusion problems as fixed-point iterations. Our study investigates its behavior when no fixed point exists. We prove, under mild assumptions, that the Davis–Yin shadow sequence converges to a solution of the normal problem, which represents a minimal perturbation of the original formulation.

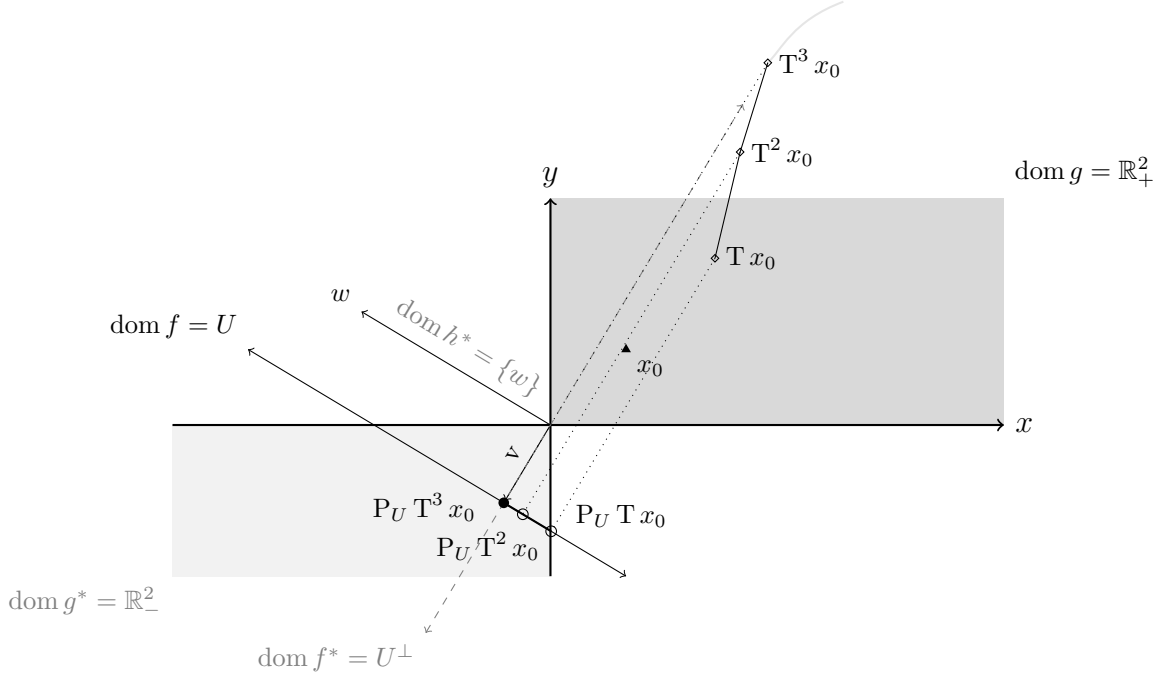


Figure 1: Illustration of our contribution, Theorem 4.6, as instantiated in Example 4.7.

Acknowledgements

First and foremost, I owe my deepest gratitude to my supervisor, Professor Walaa Moursi, for directing this work and for her guidance, which honed the rigour of my research. Specifically, as I navigated the transition from computer science to mathematics, she generously devoted countless hours to mentoring me to ensure the process was thorough and edifying.

The constructive feedback of my thesis readers, Professor Levent Tunçel and Professor Stephen A. Vavasis, was essential in refining and improving the presentation of the ideas; I am thankful for their meticulous review, valuable suggestions, and time.

At every turn, the unwavering support of my family, Amir, Thérèse, and Youhanna, has anchored me and propelled me forward.

Many thanks are due to my friends in the Combinatorics and Optimization department, whose stimulating philosophical conversations and camaraderie made my entire graduate experience both intellectually enriching and genuinely enjoyable. I would like to especially mention: Mahtab Alghasi, David Alemán Espinosa, Santiago Estupiñán-Salamanca, Matthew Hough, Leo Jung, Sina Kalantarzadeh, Matthew Kroeker, Erica Liu, Timothy C. Miller, Lukas Nabergall, Fernanda Rivera Omana, Martin Pei, Mathieu Rundström, Kartik Singh, Mariia Sobchuk, Jerónimo Valencia, Noah Weninger, and Yang Xiao.

Eventually, I extend my gratitude to the administrative team, whose timely responses and immediate accommodation of requests ensured a complete graduate experience. I would like to recognize: Genevieve Belford, Cody Miller-Sweeney, Melissa Quantrill, and Carol Seely-Morrison.

Dedication

With love, to Amir, Thérèse, and Youhanna.

Table of Contents

Author's Declaration	ii
Abstract	iii
Acknowledgements	iv
Dedication	v
List of Figures	viii
List of Symbols	ix
1 Introduction	1
2 Preliminaries	4
2.1 Sets and Operators	4
2.2 Overview of Convex Analysis	5
2.3 Notions of Nonexpansiveness	7
2.4 Monotone Operators	8
2.5 Fejér Monotonicity	9
2.6 Resolvents and Proximity Operators	10

3	On the Davis-Yin Operator	11
3.1	The Range of $\text{Id} - T$	12
3.2	A Necessary and Sufficient Condition for the Convergence of the Shadow Sequence	16
3.3	On the Sets $\text{Fix } T(\cdot + v)$ and $\text{Fix}(T + v)$	22
3.4	Towards Fejér Monotonicity	24
3.5	An Equivalent Description of Z	28
4	Convergence of the Shadow Sequence	31
5	Conclusion	46
	References	47

List of Figures

1	Illustration of our contribution, Theorem 4.6, as instantiated in Example 4.7.	iii
4.1	Two disjoint sets, $\text{dom } f = U$ and $\text{dom } g = \mathbb{R}_+^2$, for which the Davis–Yin iterates $T^n x$, starting at $x_0 = (1, 1)$, diverge and fail to minimize $h = \langle w, \cdot \rangle$, even though their projections, the shadow sequence, $P_U T^n x$ settle at the minimal displacement vector v .	40
4.2	Two disjoint sets, $\text{dom } f = U$ (the translated y -axis) and $\text{dom } g = \mathbb{R}_+^2$, for which both the Davis–Yin iterates $T^n x$ starting at $x_0 = (1, 1)$ and its associated shadow sequence $P_U T^n x$ diverge and fail to minimize $h = \langle w, \cdot \rangle$ as a consequence of $v_R \neq 0$.	41
4.3	Two disjoint sets, $\text{dom } f = U$ (the x -axis) and $\text{dom } g = W$ (the epigraph of the translated exponential function) for which the iterates of Davis–Yin, $T^n x$, starting at $x_0 = (1, 1)$ and its associated shadow sequence, $P_U T^n x$, diverge due to $v \in \overline{\text{ran}}(\text{Id} - T)$ but $v \notin \text{ran}(\text{Id} - T)$.	44

List of Symbols

Real Line

$\mathbb{R}$	The set of real numbers	p. 1
$\mathbb{R}_+$	The set of nonnegative real numbers $[0, +\infty[$	p. 7
$\mathbb{R}_{++}$	The set of strictly positive real numbers $]0, +\infty[$	p. 1
$\mathbb{R}_{--}$	The set of strictly negative real numbers $] -\infty, 0[$	p. 38
$\mathbb{N}$	The set of positive integers $\{0, 1, \dots\}$	p. 5
$\overline{\lim} x_n$	Limit superior of $(x_n)_{n \in \mathbb{N}}$	p. 5
$\underline{\lim} x_n$	Limit inferior of $(x_n)_{n \in \mathbb{N}}$	p. 5

Topology

$x_n \rightarrow x$	Strong convergence in Hilbert space	p. 5
$x_n \rightharpoonup x$	Weak convergence in Hilbert space	p. 5

Sets

$U + W$	The Minkowski sum of the sets U and W	p. 4
αU	Scaling of a set U by a real number α	p. 4
$z + U$	Translation of a set U by a vector z	p. 4
P_D	Projector onto a nonempty closed convex set D	p. 6
$D^\perp$	Orthogonal complement of a set D	p. 5

$D^\ominus$	Polar cone of a cone D	p. 5
N_D	Normal cone operator of a set D	p. 5
$\text{rec } D$	Recession cone of a set D	p. 20
ι_D	Indicator function of a set D	p. 5

Functions

$\Gamma_0(X)$	The set of all proper, convex, and lower semicontinuous functions defined from X to $] - \infty, +\infty]$	p. 6
$\text{dom } f$	Domain of a function f	p. 5
$\overline{\text{dom } f}$	Closure of the domain of a function f	p. 5
$\text{epi } f$	Epigraph of a function f	p. 5
$\text{argmin}_{x \in X} f$	The set of minimizers of a function f	p. 6
$f \square g$	Infimal convolution of the functions f and g	p. 7
$f \boxminus g$	Exact infimal convolution of the functions f and g	p. 7
P_f	Proximity operator of a function f	p. 10
f^*	The Fenchel conjugate of a function f	p. 6
∂f	The subdifferential operator of a function f	p. 6

Set-valued and Single-valued Operators

$A : X \rightrightarrows X$	A is a set-valued operator from X to X	p. 4
$\text{gra } A$	Graph of an operator A	p. 4
$\text{dom } A$	Domain of an operator A	p. 4
$\text{ran } A$	Range of an operator A	p. 4
$\text{zer } A$	Set of zeros of an operator A	p. 4
A^{-1}	Inverse of an operator A	p. 4
J_A	Resolvent of an operator A	p. 10
R_A	Reflected resolvent of an operator A	p. 10
$A(\cdot + w)$	Inner perturbation of operator A by a vector w	p. 4
$A + w$	Outer perturbation of operator A by a vector w	p. 4
$T : X \rightarrow X$	T is a single-valued operator from X to X	p. 4
$\text{Fix } T$	Set of fixed points of an operator $T : X \rightarrow X$	p. 8
Id	Identity operator	p. 1

Chapter 1

Introduction

Throughout this work, we assume

X is a (real) Hilbert space with inner product $\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{R}$,

and induced norm $\| \cdot \|$. The identity operator on X is $\text{Id} : X \rightarrow X : x \mapsto x$.

This thesis studies the three-operator splitting scheme introduced by Davis and Yin [1], known as “Davis-Yin Splitting”. The central question is whether Davis-Yin splitting can offer useful insights for the class of problems to which it is commonly applied, but in the *inconsistent* case. That is, when the set of optimal solutions is *empty*, which we answer in the affirmative (cf. Theorem 4.6) under mild assumptions. Analyzing the behavior of splitting methods to assess the consistency of the underlying problem has been an active area of research [2, 3, 4]. The *monotone inclusion problem* that Davis-Yin splitting considers is

$$\text{find } x \in X \text{ such that } 0 \in Ax + Bx + Cx, \quad (1.1)$$

where A, B , and C are maximally monotone operators (cf. Definition 2.40), and C is, additionally, β -cocoercive (cf. Definition 2.29) for some $\beta \in \mathbb{R}_{++}$. The Davis-Yin operator is defined as:

$$(\gamma \in]0, 2\beta[) \quad T_\gamma := T_{(\gamma A, \gamma B, \gamma C)} = J_{\gamma B}(R_{\gamma A} - \gamma C J_{\gamma A}) + \text{Id} - J_{\gamma A}, \quad (1.2)$$

where $J_{\gamma B}$ and $J_{\gamma A}$ are the resolvents of the corresponding operators and $R_{\gamma A}$ is the reflected resolvent (cf. Definition 2.53). Note that (1.2) is not invariant under swapping A and B

(see [5] for a detailed study of this aspect in the case $C = 0$). In this work, we consider $\gamma \in]0, \beta[$, which guarantees that γC remains firmly nonexpansive. Therefore, for the remainder of this thesis, we omit the γ in $\gamma A, \gamma B, \gamma C$, and, consequently, T_γ and instead write A, B, C , and consequently, T , i.e.,

$$T := T_{(A,B,C)} = J_B(R_A - C J_A) + \text{Id} - J_A. \quad (1.3)$$

Let $x_0 \in X$ be given. The splitting algorithm generates the governing sequence according to the following update rule:

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = T x_n \quad (1.4)$$

When $\text{zer}(A + B + C) = \emptyset$, it follows from [1, Lemma 2.2] that $\text{Fix } T = \emptyset$, i.e., the fixed point set of T is empty. Consequently, $\lim_{n \in \mathbb{N}} \|T^n x\| = \infty$ by [6, Corollary 1.4]. Therefore, we instead examine the problem

$$\text{find } x \in X \text{ such that } 0 \in Ax + B(x - v) + Cx - v,$$

where v (cf. Definition 2.36) is the element of least norm in $\overline{\text{ran}}(\text{Id} - T)$, the closure of the range of $\text{Id} - T$. The problem (3.39) will be referred to as the “normal problem”.

When (A, B, C) is replaced by $(\partial f, \partial g, \nabla h)$ where f, g , and h are proper (cf. Definition 2.13), lower semicontinuous (cf. Definition 2.11), convex (cf. Definition 2.10) functions defined on X , and ∇h additionally Lipschitz continuous (cf. Definition 2.27), we provide a necessary and sufficient condition for the weak convergence of the “shadow” sequence $(J_A x_n)_{n \in \mathbb{N}} = (J_{\partial f} x_n)_{n \in \mathbb{N}} = (P_f x_n)_{n \in \mathbb{N}}$ to a minimizer of (3.39), where P_f denotes the proximity operator of f (cf. Definition 2.58). Additionally, we provide other intriguing results of independent interest:

1. In Proposition 3.10, we derive a precise description for $\overline{\text{ran}}(\text{Id} - T)$.
2. Lemma 3.24 provides two core descriptions of the extended solution set (cf. Definition 3.23) and the set of fixed points of the perturbed operators: $\text{Fix } T(\cdot + v)$ and $\text{Fix}(T + v)$.
3. In Corollary 3.19, we establish a necessary and sufficient condition for the asymptotic regularity of $(J_A(x_n))_{n \in \mathbb{N}}$ and $(J_{A^{-1}}(x_n))_{n \in \mathbb{N}}$.
4. Our main result, which is stated in Theorem 4.6, establishes the weak convergence of $(P_f(x_n))_{n \in \mathbb{N}}$ to a minimizer of (3.39).

The authors of [7] introduced the averaged alternating reflections method for finding a best approximation pair $(u, w) \in U \times W$ such that $\|u - w\| = d(U, W) := \inf \|U - W\|$, where U and W are nonempty closed convex sets in X *that may not intersect*. More importantly, the authors established that the shadow sequence $P_W(x_n)_{n \in \mathbb{N}}$ is bounded and that its cluster points lie in $W \cap (U + v)$ (cf. [7, Theorem 3.13]), where P_W is the projector (the best approximation operator) onto W and that $v \in U - W$ if and only if $d(U, W)$ is attained. The same result motivated the work of [8] on the Douglas-Rachford operator, establishing connections between the generalized set of minimizers of the *normal problem* and the range of the displacement mapping, which were key to Proposition 3.8, among other findings. The authors of [8] provided a detailed characterization of the range of the displacement mapping of Douglas-Rachford in [9], which we utilized in Proposition 3.10 to establish a similar result for range of the displacement mapping for Davis-Yin. The previous contributions culminated in [10], proving the weak convergence of the shadow sequence in the inconsistent case of Douglas-Rachford Splitting in the convex setting (a specific nonconvex case was analyzed in [11]). There are other avenues through which one could respond to the inconsistent case, e.g., [12].

Chapter 2

Preliminaries

In this chapter, we review the relevant definitions, facts, and notation used in our work.

2.1 Sets and Operators

Let U, W be subsets of X , let $z \in X$. The Minkowski sum of U and W is given by $U + W = \{u + w \mid u \in U, w \in W\}$. The translation of U by z is defined as $z + U = \{z\} + U$. For any $\alpha \in \mathbb{R}$, the scaling of U by α is defined as $\alpha U = \{\alpha u \mid u \in U\}$. The notation $T : X \rightarrow X$ indicates that T maps each $x \in X$ to a point $Tx \in X$. Similarly, $A : X \rightrightarrows X$ denotes that A maps each $x \in X$ to a (possibly empty) set $Ax \subseteq X$. Furthermore, let w and x be in X . The inner perturbation of A by w at x is given by $A(\cdot + w)x = A(x + w)$, while the outer perturbation is given by $(A + w)x = Ax + w$.

Definition 2.1. Let $A : X \rightrightarrows X$. Then, A is characterized by its *graph*

$$\text{gra } A = \{(x, u) \in X \times X \mid u \in Ax\}. \quad (2.1)$$

The domain and the range of A are defined as:

$$\text{dom } A = \{x \in X \mid Ax \neq \emptyset\} \text{ and } \text{ran } A = A(X). \quad (2.2)$$

Definition 2.2. Let $A : X \rightrightarrows X$. The *inverse* of A , denoted by A^{-1} is defined through its graph

$$\text{gra } A^{-1} = \{(u, x) \in X \times X \mid (x, u) \in \text{gra } A\}. \quad (2.3)$$

Definition 2.3. Let $A : X \rightrightarrows X$. Then, the set of zeros of A is $\text{zer } A = A^{-1}(0) = \{x \in X \mid 0 \in Ax\}$.

2.2 Overview of Convex Analysis

Definition 2.4. The *limit superior* and *limit inferior* of a sequence $(x_n)_{n \in \mathbb{N}}$ in $[-\infty, +\infty]$ are denoted $\underline{\lim} x_n$ and $\overline{\lim} x_n$, respectively. Moreover, it follows that

$$\underline{\lim} x_n \leq \overline{\lim} x_n. \quad (2.4)$$

Definition 2.5. A sequence $(x_n)_{n \in \mathbb{N}}$ in X *converges strongly* to a point $x \in X$ if $\|x_n - x\| \rightarrow 0$, which will be denoted $x_n \rightarrow x$. A sequence $(x_n)_{n \in \mathbb{N}}$ in X *converges weakly* to a point $x \in X$ if for every $u \in X$, $\langle x_n, u \rangle \rightarrow \langle x, u \rangle$, which will be denoted $x_n \rightharpoonup x$.

Definition 2.6. A subset D of X is *convex* if for all $\alpha \in]0, 1[$, we have $\alpha D + (1 - \alpha)D = D$.

Definition 2.7. Let D be a subset of X , The *orthogonal complement* of D is denoted by $D^\perp$. That is,

$$D^\perp = \{u \in X \mid (\forall x \in D) \quad \langle x, u \rangle = 0\} \quad (2.5)$$

Definition 2.8. Let D be a cone of X . The *polar cone* of D is $D^\ominus = \{u \in X \mid \sup \langle D, u \rangle \leq 0\}$.

Definition 2.9. Let D be a nonempty convex subset of X and let $x \in X$. The *normal cone* to D at x is

$$N_D x = \begin{cases} (D - x)^\ominus = \{u \in X \mid \sup \langle D - x, u \rangle \leq 0\}, & \text{if } x \in D; \\ \emptyset, & \text{otherwise.} \end{cases} \quad (2.6)$$

Definition 2.10. Let $f : X \rightarrow]-\infty, +\infty]$. The *domain* of f is $\text{dom } f = \{x \in X \mid f(x) < +\infty\}$. The function f is *convex* if its *epigraph* $\text{epi } f = \{(x, \alpha) \in X \times \mathbb{R} \mid f(x) \leq \alpha\}$ is a convex set. In addition, the closure of $\text{dom } f$ is denoted by $\overline{\text{dom } f}$.

Definition 2.11. Let $f : X \rightarrow [-\infty, +\infty]$ and $x \in X$. Then, f is *lower semicontinuous* at x if, for every sequence $(x_n)_{n \in \mathbb{N}}$ in X , $x_n \rightarrow x$ implies

$$f(x) \leq \underline{\lim} f(x_n). \quad (2.7)$$

and *weakly lower semicontinuous* at x if $x_n \rightharpoonup x$ implies (2.7).

Example 2.12 ([13, Example 1.25]). The *indicator function* of a subset D of X , i.e., the function

$$\iota_D : X \rightarrow [-\infty, +\infty] : x \mapsto \begin{cases} 0, & \text{if } x \in D; \\ +\infty & \text{otherwise,} \end{cases} \quad (2.8)$$

is lower semicontinuous if and only if D is closed.

Definition 2.13. Let $f : X \rightarrow]-\infty, +\infty]$. Then, f is *proper* if $\text{dom } f := \{x \in X \mid f(x) < +\infty\} \neq \emptyset$.

Fact 2.14 ([13, Theorem 9.1]). *Let $f : X \rightarrow]-\infty, +\infty]$ be convex and lower semicontinuous. Then, f is weakly lower semicontinuous.*

Definition 2.15. The set of proper lower semicontinuous convex functions from X to $]-\infty, +\infty]$ is denoted by $\Gamma_0(X)$.

Fact 2.16 ([13, Proposition 8.17] and [13, Lemma 1.27]). *The set of proper, convex, and lower semicontinuous functions $\Gamma_0(X)$ is closed under addition.*

Definition 2.17. Let $f : X \rightarrow [-\infty, +\infty]$. The *Fenchel conjugate* of f is

$$f^* : X \rightarrow [-\infty, +\infty] : u \mapsto \sup_{x \in X} (\langle x, u \rangle - f(x)). \quad (2.9)$$

Definition 2.18. Let $f : X \rightarrow]-\infty, +\infty]$ be proper. The *subdifferential* of f is $\partial f : X \rightrightarrows X : x \mapsto \{u \in X \mid (\forall y \in X) \langle y - x, u \rangle + f(x) \leq f(y)\}$.

Fact 2.19 ([13, Theorem 16.29 and Corollary 16.39]). *Let $f \in \Gamma_0(X)$. Then,*

$$\text{ran } \partial f = \text{dom } \partial f^* = \text{dom } f^*. \quad (2.10)$$

Fact 2.20 ([13, Proposition 27.1]). *Let $f \in \Gamma_0(X)$. Then, $\text{argmin}_{x \in X} f = \text{zer } \partial f$.*

Fact 2.21 ([13, Corollary 16.48(iii)]). *Let f and g be in $\Gamma_0(X)$ such that $\text{dom } g = X$. Then, $\partial(f + g) = \partial f + \partial g$.*

Fact 2.22 ([10, Proposition 2.1]). *Let f and g be in $\Gamma_0(X)$ and let $\text{zer}(\partial f + \partial g) \neq \emptyset$. Then, $\text{argmin}_{x \in X} (f + g) = \text{zer}(\partial(f + g)) = \text{zer}(\partial f + \partial g)$.*

Definition 2.23. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X . Then, $(x_n)_{n \in \mathbb{N}}$ is *asymptotically regular* if $x_n - x_{n+1} \rightarrow 0$.

Fact 2.24 ([13, Theorem 3.16]). *Suppose that D is a nonempty closed convex subset in X . Then, for every point $x \in X$, there exists a unique point $P_D x \in D$ such that $\|x - P_D x\| = \inf \|x - D\|$. The point $P_D x$ is characterized by*

$$P_D x \in D \text{ and } (\forall u \in D) \quad \langle u - P_D x, x - P_D x \rangle \leq 0. \quad (2.11)$$

Definition 2.25. Let $f : X \rightarrow]-\infty, +\infty]$ and $g : X \rightarrow]-\infty, +\infty]$. The *infimal convolution* of f and g is

$$f \square g : X \rightarrow [-\infty, +\infty] : x \mapsto \inf_{y \in X} (f(y) + g(x - y)), \quad (2.12)$$

and it is *exact* at a point $x \in X$ if $(f \square g)(x) = \min_{y \in X} f(y) + g(x - y)$; $f \square g$ is *exact* if it is exact at every point of its domain, in which case it is denoted $f \boxdot g$.

Fact 2.26 ([13, Proposition 12.6]). *Let $f : X \rightarrow]-\infty, +\infty]$ and $g : X \rightarrow]-\infty, +\infty]$. Then,*

$$\text{dom}(f \square g) = \text{dom } f + \text{dom } g. \quad (2.13)$$

2.3 Notions of Nonexpansiveness

Definition 2.27. Let $T : X \rightarrow X$. Then, T is Lipschitz continuous with constant $\alpha \in \mathbb{R}_+$ if

$$(\forall x \in X)(\forall y \in X) \quad \|Tx - Ty\| \leq \alpha \|x - y\|. \quad (2.14)$$

Definition 2.28. Let $T : X \rightarrow X$. Then, T is *nonexpansive* if it is Lipschitz continuous with constant 1, i.e.,

$$(\forall x \in X)(\forall y \in X) \quad \|Tx - Ty\| \leq \|x - y\|, \quad (2.15)$$

and is *firmlly nonexpansive* if

$$(\forall x \in X)(\forall y \in X) \quad \|Tx - Ty\|^2 \leq \langle x - y, Tx - Ty \rangle. \quad (2.16)$$

Definition 2.29. Let $T : X \rightarrow X$. Then, T is β -cocoercive, where $\beta \in \mathbb{R}_{++}$, if βT is *firmlly nonexpansive*.

Remark 2.30. If T is β -cocoercive, then it follows from Cauchy-Schwarz inequality and (2.16) that T is Lipschitz continuous with constant $1/\beta$.

Fact 2.31 ([14, Corollaire 10] and [13, Corollary 18.17]). *Let $f \in \Gamma_0(X)$. Then, ∇f is nonexpansive if and only if ∇f is firmlly nonexpansive.*

Definition 2.32. Let $T : X \rightarrow X$ be nonexpansive, and let $\alpha \in]0, 1[$. Then, T is α -averaged if there exists a nonexpansive $N : X \rightarrow X$ such that $T = (1 - \alpha)\text{Id} + \alpha N$.

Fact 2.33 ([15, Lemma 4]). *Let $T : X \rightarrow X$ be a nonexpansive. Then, $\overline{\text{ran}}(\text{Id} - T)$ is convex.*

Definition 2.34. Let $T : X \rightarrow X$. Then, the set of fixed points of T is

$$\text{Fix } T := \{x \in X \mid Tx = x\}. \quad (2.17)$$

Fact 2.35 ([16, Lemma 2.1]). *Let $T : X \rightarrow X$ and $w \in X$. Then, $\text{Fix } T(\cdot + w) = -w + \text{Fix}(T + w)$.*

Definition 2.36. Let $T : X \rightarrow X$. Then, *the minimal displacement vector (or the element of least norm)* is well-defined (see Fact 2.33) as

$$v := P_{\overline{\text{ran}}(\text{Id} - T)}(0). \quad (2.18)$$

Fact 2.37 ([15, Lemma 2]). *Let $D \subseteq X$ be nonempty, closed, and convex and let v be the element of least norm in D . If $u_n \in D$ and $\|u_n\| \rightarrow \|v\|$ then $u_n \rightarrow v$ and $\|u_n - v\|^2 \leq \|u_n\|^2 - \|v\|^2$.*

Fact 2.38 ([6, Corollary 1.5; Proposition 1.3]). *Let $T : X \rightarrow X$ be averaged and let $x \in X$. Then, $\lim_{n \in \mathbb{N}} (T^n x - T^{n+1} x) = v$, where v is defined according to (2.18).*

2.4 Monotone Operators

Definition 2.39. Let $A : X \rightrightarrows X$. Then, A is *monotone* if

$$(\forall (x, u) \in \text{gra } A)(\forall (y, w) \in \text{gra } A) \quad \langle x - y, u - w \rangle \geq 0. \quad (2.19)$$

Definition 2.40. Let $A : X \rightrightarrows X$. Then, A is *maximally monotone* if it is monotone and there is no monotone operator $B : X \rightrightarrows X$ such that $\text{gra } B$ properly contains $\text{gra } A$. Equivalently, for every $(x, u) \in X \times X$,

$$(x, u) \in \text{gra } A \quad \Leftrightarrow \quad (\forall (y, w) \in \text{gra } A) \quad \langle x - y, u - w \rangle \geq 0. \quad (2.20)$$

Fact 2.41 ([13, Corollary 25.5(i)]). *Let $A, B : X \rightrightarrows X$ be maximally monotone operators such that $\text{dom } B = X$. Then, $A + B$ is maximally monotone.*

Definition 2.42. Let $A : X \rightrightarrows X$. Then, A is *3^* -monotone* if A is monotone and

$$(\forall x \in \text{dom } A)(\forall y^* \in \text{ran } A) \quad \inf_{(u, w) \in \text{gra } A} \langle x - u, y^* - w \rangle > -\infty. \quad (2.21)$$

Fact 2.43 ([13, Moreau; Theorem 20.25]). *Let $f \in \Gamma_0(X)$. Then, ∂f is maximally monotone.*

Fact 2.44 ([13, Example 25.13]). *Let $f : X \rightarrow]-\infty, +\infty]$ be proper. Then, ∂f is 3^* -monotone.*

Fact 2.45 (Brézis-Haraux; [13, Theorem 25.24]). *Let $A : X \rightrightarrows X$ and $B : X \rightrightarrows X$ be monotone operators such that $A+B$ is maximally monotone and A and B are 3^* monotone. Then,*

$$\overline{\text{ran}}(A + B) = \overline{\text{ran } A + \text{ran } B}. \quad (2.22)$$

Definition 2.46. Let $A : X \rightrightarrows X$. Then, A is *paramonotone* if it is monotone and

$$(\forall (x, u) \in \text{gra } A)(\forall (y, w) \in \text{gra } A) \quad \langle x - y, u - w \rangle = 0 \Rightarrow (x, w) \in \text{gra } A \text{ and } (y, u) \in \text{gra } A.$$

Fact 2.47 ([13, Example 22.4]). *Let $f : X \rightarrow]-\infty, +\infty]$ be proper and convex. Then, ∂f is paramonotone.*

2.5 Fejér Monotonicity

Definition 2.48. Let D be a nonempty subset of X and let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X . Then $(x_n)_{n \in \mathbb{N}}$ is *Fejér-monotone* with respect to D if

$$(\forall x \in D)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - x\| \leq \|x_n - x\|. \quad (2.23)$$

Fact 2.49 ([13, Proposition 5.4]). *Let D be a nonempty subset of X and $(x_n)_{n \in \mathbb{N}}$ be a sequence in X . Suppose that $(x_n)_{n \in \mathbb{N}}$ is Fejér monotone with respect to D . Then, $(x_n)_{n \in \mathbb{N}}$ is bounded.*

Lemma 2.50. *Let $u \in X$ and D be a nonempty subset of X . Suppose that $(x_n)_{n \in \mathbb{N}}$ is Fejér monotone with respect to D . Then, $(u + x_n)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $u + D$.*

Proof. Let $x \in D$. Note that $\|x_n - x\| \leq \|x_{n+1} - x\| \Leftrightarrow \|(u + x_n) - (u + x)\| \leq \|(u + x_{n+1}) - (u + x)\|$. Hence, $(u + x_n)$ is Fejér monotone with respect to $u + D$. $\square$

Fact 2.51 ([17, Lemma 2.2]). *Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X that is Fejér monotone with respect to a nonempty closed convex subset C of X . Let w_1 and w_2 be weak cluster points of $(x_n)_{n \in \mathbb{N}}$. Then $w_1 - w_2 \in (C - C)^\perp$, i.e., $\langle w_1 - w_2, C - C \rangle = 0$.*

Fact 2.52 ([16, Proposition 2.4(iv) and (vi)]). Let $T : X \rightarrow X$ be nonexpansive and suppose that $y_0 \in \text{Fix}(T + v)$, where v be defined according to (2.18). Then, for all $n \in \mathbb{N}$,

$$(i) \quad T^n y_0 = y_0 - nv.$$

(ii) For every $x \in X$, the sequence $(T^n x + nv)_{n \in \mathbb{N}}$ is Fejér monotone with respect to both $\text{Fix}(T + v)$ and $\text{Fix } T(\cdot + v)$.

2.6 Resolvents and Proximity Operators

Definition 2.53. Let $A : X \rightrightarrows X$. The *resolvent* of A is $J_A := (\text{Id} + A)^{-1}$. The *reflected resolvent* is

$$R_A : X \rightarrow X : x \mapsto 2J_A x - x. \quad (2.24)$$

Fact 2.54 ([13, Proposition 23.19]). Let $A : X \rightrightarrows X$ be maximally monotone. Then

$$J_{A^{-1}} = \text{Id} - J_A. \quad (2.25)$$

Fact 2.55 ([18, Theorem 4]). Let $A : X \rightrightarrows X$ be maximally monotone. Then,

$$\text{gra } A = \{(J_A x, x - J_A x) \mid x \in X\}. \quad (2.26)$$

Fact 2.56 ([13, Corollary 23.11]). Let $A : X \rightrightarrows X$ be maximally monotone. Then, the following hold:

(i) $J_A : X \rightarrow X$ and $J_{A^{-1}} = \text{Id} - J_A : X \rightarrow X$ are firmly nonexpansive and maximally monotone.

(ii) $R_A : X \rightarrow X$ is nonexpansive.

Fact 2.57 ([13, Proposition 23.2(i)]). Let $A : X \rightrightarrows X$. Then, $\text{ran } J_A = \text{dom } A$. Consequently, $\text{ran } J_{A^{-1}} = \text{dom } A^{-1} = \text{ran } A$.

Definition 2.58. The *proximity operator* of $f \in \Gamma_0(X)$ is

$$P_f : X \rightarrow X : x \mapsto \underset{y \in X}{\text{argmin}} \left(f(y) + \frac{1}{2} \|x - y\|^2 \right). \quad (2.27)$$

Example 2.59 ([13, Example 23.3]). Let $f \in \Gamma_0(X)$. Then, $J_{\partial f} = P_f$.

Fact 2.60 ([13, Proposition 16.44]). Let $f \in \Gamma_0(X)$. Then, $P_f = (\text{Id} + \partial f)^{-1}$.

Fact 2.61 ([13, Proposition 12.26]). Let $f \in \Gamma_0(X)$ and let x and p be in X . Then

$$p = P_f x \Leftrightarrow (\forall y \in X) \quad \langle y - p, x - p \rangle + f(p) \leq f(y). \quad (2.28)$$

Chapter 3

On the Davis-Yin Operator

Throughout the rest of this work, we assume that

$$A : X \rightrightarrows X \text{ and } B : X \rightrightarrows X \text{ are maximally and } 3^*\text{-monotone operators.} \quad (3.1)$$

Moreover, we assume that

$$C : X \rightarrow X \text{ is maximally monotone and firmly nonexpansive.} \quad (3.2)$$

Consequently, C is, additionally, 3^* -monotone by [13, Example 25.20] and paramonotone (cf. Definition 2.46) by [13, Example 22.8]. The Davis-Yin operator for the triplet (A, B, C) is given by:

$$T := T_{(A,B,C)} = J_B(R_A - C J_A) + \text{Id} - J_A. \quad (3.3)$$

The corresponding minimal displacement vector (cf. Definition 2.36) for T is defined as

$$v := P_{\overline{\text{ran}(\text{Id} - T)}}(0). \quad (3.4)$$

We additionally introduce the following vectors:

$$v_D := P_{\overline{\text{dom } A - \text{dom } B}}(0) \text{ and } v_R := P_{\overline{\text{ran } A + \text{ran } B + \text{ran } C}}(0). \quad (3.5)$$

By [13, Proposition 29.3],

$$(v_D, v_R) = P_{\overline{\text{dom } A - \text{dom } B} \times \overline{\text{ran } A + \text{ran } B + \text{ran } C}}((0, 0)). \quad (3.6)$$

Another closely related splitting algorithm is the Douglas-Rachford method [19], which is defined for two maximally monotone operators $\tilde{A} : X \rightrightarrows X$ and $\tilde{B} : X \rightrightarrows X$ as follows:

$$\tilde{T} := \tilde{T}_{(\tilde{A}, \tilde{B})} = J_{\tilde{B}} R_{\tilde{A}} + \text{Id} - J_{\tilde{A}}. \quad (3.7)$$

Note that (3.3) recovers (3.7) when $C = 0$.

Remark 3.1. As a consequence of Fact 2.41, we have that

$$B + C \text{ is maximally monotone.} \quad (3.8)$$

Moreover, by Fact 2.45, $\overline{\text{ran}}(B + C) = \overline{\text{ran}} B + \overline{\text{ran}} C$. In addition, $\text{dom}(B + C) = \text{dom } B \cap \text{dom } C = \text{dom } B \cap X = \text{dom } B$. Suppose that B is paramonotone. Then, $B + C$ is paramonotone by [20, Proposition 3.1].

3.1 The Range of $\text{Id} - T$

This section studies the behavior of the operator $\text{Id} - T$, as detailed in Proposition 3.2, and provides a characterization of the closure of its range in Proposition 3.10. To achieve this, we introduce two perturbed operators in Proposition 3.4, which play a crucial role in several results, including the intermediate result in Proposition 3.8: $\text{ran}(\text{Id} - T) = \text{ran}(\text{Id} - \tilde{T})$.

Proposition 3.2. *We have*

$$\text{Id} - T = J_A - J_B(R_A - C J_A) = J_{A^{-1}} + J_{B^{-1}}(R_A - C J_A) + C J_A. \quad (3.9)$$

Consequently,

$$(i) \quad J_A T^n x - J_B(J_A - C J_A) T^n x \rightarrow v.$$

$$(ii) \quad J_{A^{-1}} T^n x + J_{B^{-1}}(R_A - C J_A) T^n x + C J_A T^n x \rightarrow v.$$

Proof. The claim in (3.9) follows by applying Fact 2.54:

$$\text{Id} - T \stackrel{(3.3)}{=} J_A - J_B(R_A - C J_A) \quad (3.10a)$$

$$\begin{aligned} &\stackrel{(2.25)}{=} J_A - (R_A - C J_A - J_{B^{-1}}(R_A - C J_A)) \\ &= J_A - R_A + C J_A + J_{B^{-1}}(R_A - C J_A) \end{aligned}$$

$$\stackrel{(2.24)}{=} J_A - 2 J_A + \text{Id} + C J_A + J_{B^{-1}}(R_A - C J_A)$$

$$\stackrel{(2.25)}{=} J_{A^{-1}} + J_{B^{-1}}(R_A - C J_A) + C J_A. \quad (3.10b)$$

Clearly, $\text{dom } T = X$. Hence, by Fact 2.38, for all $x \in X$ and $n \in \mathbb{N}$, we have

$$(\text{Id} - T) T^n x \stackrel{(3.10a)}{=} J_A T^n x - J_B(J_A - C J_A) T^n x \quad (3.11a)$$

$$\stackrel{(3.10b)}{=} J_{A^{-1}} T^n x + J_{B^{-1}}(R_A - C J_A) T^n x + C J_A T^n x \quad (3.11b)$$

$$= T^n x - T^{n+1} x \rightarrow v, \quad (3.11c)$$

which proves (i) and (ii). $\square$

Corollary 3.3. *We have*

$$\text{ran}(\text{Id} - T) \subseteq (\text{dom } A - \text{dom } B) \cap (\text{ran } A + \text{ran } B + \text{ran } C) \quad (3.12)$$

Proof. By (3.9) and Fact 2.57,

$$\text{ran}(\text{Id} - T) = \text{ran}(J_A - J_B(R_A - C J_A)) \quad (3.13a)$$

$$\subseteq \text{ran } J_A - \text{ran } J_B(R_A - C J_A) \quad (3.13b)$$

$$\subseteq \text{dom } A - \text{dom } B. \quad (3.13c)$$

Moreover,

$$\text{ran}(\text{Id} - T) = \text{ran}(J_{A^{-1}} + J_{B^{-1}}(R_A - C J_A) + C J_A) \quad (3.14a)$$

$$\subseteq \text{ran } J_{A^{-1}} + J_{B^{-1}}(R_A - C J_A) + C J_A \quad (3.14b)$$

$$\subseteq \text{ran } A + \text{ran } B + \text{ran } C, \quad (3.14c)$$

which completes the proof. $\square$

Proposition 3.4. *Let w and x be in X . Then,*

$$(i) \quad T_{(A,B(\cdot-w)-w,C)} x = T_{(A,B,C)} x + w =: T + w.$$

$$(ii) \quad T_{(A-w,B(\cdot-w),C)} x = T_{(A,B,C)}(x + w) =: T(x + w).$$

Proof. Let $M : X \rightrightarrows X$ be a maximally monotone set-valued operator, we have, from [13, Proposition 23.17], that

$$J_{M(\cdot-w)} x = J_M(x + w) \quad \text{and} \quad J_{M-w} x = J_M(x - w) + w. \quad (3.15)$$

Hence,

$$J_{(B(\cdot-w)-w)} x = J_{B(\cdot-w)}(x - w) + w = J_B(x) + w. \quad (3.16)$$

Consequently,

$$T_{(A,B(\cdot-w)-w,C)} x = x - J_A x + J_{B(\cdot-w)-w}(R_A - C J_A) x \quad (3.17a)$$

$$\stackrel{(3.16)}{=} x - J_A x + J_{B(\cdot-w)}(2 J_A x - x - w - C J_A x) + w \quad (3.17b)$$

$$\stackrel{(3.16)}{=} x - J_A x + J_B(2 J_A x - x - w + w - C J_A x) + w \quad (3.17c)$$

$$\stackrel{(2.24)}{=} x - J_A x + J_B(R_A - C J_A) x + w \quad (3.17d)$$

$$\stackrel{(3.3)}{=} T_{(A,B,C)} x + w, \quad (3.17e)$$

which proves (i). Similarly,

$$T_{(A-w, B(\cdot-w), C)} x = x - J_{A-w} x + J_{B(\cdot-w)}(R_{A-w} - C J_{A-w})x \quad (3.18a)$$

$$\stackrel{(3.15)}{=} x - J_A(x+w) + J_{B(\cdot-w)}(2J_A(x+w) - x - C J_A(x+w)) \quad (3.18b)$$

$$\stackrel{(3.15)}{=} (x+w) - J_A(x+w) + J_B(2J_A(x+w) - x - w - C J_A(x+w)) \quad (3.18c)$$

$$\stackrel{(2.24)}{=} (x+w) - J_A(x+w) + J_B(R_A(x+w) - C J_A(x+w)) \quad (3.18d)$$

$$= (x+w) - J_A(x+w) + J_B(R_A - C J_A)(x+w) \quad (3.18e)$$

$$\stackrel{(3.3)}{=} T_{(A, B, C)}(x+w), \quad (3.18f)$$

which proves (ii). $\square$

Definition 3.5. Let $A : X \rightrightarrows X$ and $B : X \rightrightarrows X$. Then, the *primal problem* associated with the ordered pair (A, B) is

$$\text{find } x \in X \text{ such that } 0 \in Ax + Bx. \quad (3.19)$$

The corresponding Attouch-Théra *dual pair* [21, 22] is $(A, B)^* := (A^{-1}, B^{-\odot})$, where $B^{\odot} := (-\text{Id}) \circ B \circ (-\text{Id})$. As such, we let

$$Z_{(A, B)} := \text{zer}(A + B) = (A + B)^{-1}(0) = \{x \in X \mid 0 \in A + B\} \quad (3.20)$$

and

$$K_{(A, B)} := \text{zer}(A^{-1} + B^{-\odot}) = (A^{-1} + B^{-\odot})^{-1}(0) = \{x \in X \mid 0 \in A^{-1} + B^{-\odot}\}, \quad (3.21)$$

denote the primal and dual solutions of (3.19), respectively.

Fact 3.6 ([8, Proposition 3.3]). *Let $w \in X$. Then, the following are equivalent:*

$$(i) \quad Z_{(\tilde{A}-w, \tilde{B}(\cdot-w))} \neq \emptyset.$$

$$(ii) \quad \text{Fix } \tilde{T}(\cdot + w) \neq \emptyset.$$

$$(iii) \quad w \in \text{ran}(\text{Id} - \tilde{T}).$$

Lemma 3.7. *Let $w \in W$. Then, the following are equivalent:*

$$(i) \quad Z_{(A-w, B(\cdot-w)+C)} \neq \emptyset.$$

$$(ii) \quad Z_{(A,B(\cdot-w)+C-w)} \neq \emptyset.$$

$$(iii) \quad \text{Fix } T(\cdot + w) \neq \emptyset.$$

$$(iv) \quad \text{Fix}(T + w) \neq \emptyset.$$

$$(v) \quad w \in \text{ran}(\text{Id} - T).$$

Proof.

(i) $\Leftrightarrow$ (ii) This is immediate from (3.20).

(i) $\Leftrightarrow$ (iii) Indeed, this is [1, Lemma 3.1] applied to $T_{(A-w, B(\cdot-w), C)}$, which corresponds, by Proposition 3.4(i), to $T(\cdot + w)$.

(ii) $\Leftrightarrow$ (iv): Similarly, this is [1, Lemma 3.1] to $T_{(A, B(\cdot-w)-w, C)}$, which corresponds, by Proposition 3.4(ii), to $T + w$.

(iii) $\Leftrightarrow$ (v): Let $x \in X$. Then, $x \in \text{Fix } T(\cdot + w) \stackrel{(2.17)}{\Leftrightarrow} x = T(\cdot + w) \Leftrightarrow (x + w) - T(\cdot + w) = w \Leftrightarrow w \in \text{ran}(\text{Id} - T)(x + w)$. $\square$

Proposition 3.8. *We have*

$$\text{ran}(\text{Id} - T) = \text{ran}(\text{Id} - \tilde{T}). \quad (3.22)$$

Proof. Let $w \in X$. Then, replacing $\tilde{A}$ and $\tilde{B}$ by $A + w$ and $(B + C(\cdot + w) - w)$, respectively, in (3.7) yields, in view of Remark 3.1, Fact 3.6, Lemma 3.7, and Proposition 3.4, the following chain of implications:

$$\begin{aligned} w \in \text{ran}(\text{Id} - \tilde{T}) &\Leftrightarrow \text{Fix } \tilde{T}(\cdot + w) \neq \emptyset \Leftrightarrow Z_{(\tilde{A}-w, \tilde{B}(\cdot-w))} \neq \emptyset \\ &\stackrel{(3.20)}{\Leftrightarrow} (\exists x \in X) \quad 0 \in \tilde{A}x + \tilde{B}(x - w) - w \\ &\Leftrightarrow (\exists x \in X) \quad 0 \in Ax + B(x - w) + Cx - w \\ &\stackrel{(3.20)}{\Leftrightarrow} Z_{(A-w, B(\cdot-w)+C)} = Z_{(A, B(\cdot-w)+C-w)} \neq \emptyset \\ &\Leftrightarrow \text{Fix } T(\cdot + w) \neq \emptyset \Leftrightarrow \text{Fix}(T + w) \neq \emptyset \\ &\Leftrightarrow w \in \text{ran}(\text{Id} - T), \end{aligned} \quad (3.23)$$

which completes the proof. $\square$

Fact 3.9 ([9, Proposition 6.1]). *Suppose that $\tilde{A}$ and $\tilde{B}$ are 3^* -monotone. Then,*

$$\overline{\text{ran}}(\text{Id} - \tilde{T}) = \overline{(\text{dom } \tilde{A} - \text{dom } \tilde{B})} \cap \overline{(\text{ran } \tilde{A} + \text{ran } \tilde{B})}. \quad (3.24)$$

Proposition 3.10. *We have*

$$\overline{\text{ran}}(\text{Id} - T) = \overline{(\text{dom } A - \text{dom } B)} \cap \overline{(\text{ran } A + \text{ran } B + \text{ran } C)}. \quad (3.25)$$

Proof. Let $\tilde{A}$ and $\tilde{B}$ be replaced by A and $B + C$ in (3.7). As a consequence of (3.1), Remark 3.1, Fact 3.9, Proposition 3.8, and Fact 2.45, we have

$$\overline{\text{ran}}(\text{Id} - \tilde{T}) \stackrel{(3.24)}{=} \overline{(\text{dom } \tilde{A} - \text{dom } \tilde{B})} \cap \overline{(\text{ran } \tilde{A} + \text{ran } \tilde{B})} \quad (3.26a)$$

$$= \overline{(\text{dom } A - \text{dom}(B + C))} \cap \overline{(\text{ran } A + \text{ran}(B + C))} \quad (3.26b)$$

$$\stackrel{(2.22)}{=} \overline{(\text{dom } A - (\text{dom } B \cap \text{dom } C))} \cap \overline{(\text{ran } A + \overline{\text{ran}} B + \overline{\text{ran}} C)} \quad (3.26c)$$

$$= \overline{(\text{dom } A - \text{dom } B)} \cap \overline{(\text{ran } A + \text{ran } B + \text{ran } C)} \quad (3.26d)$$

$$\stackrel{(3.22)}{=} \overline{\text{ran}}(\text{Id} - T), \quad (3.26e)$$

which completes the proof. $\square$

3.2 A Necessary and Sufficient Condition for the Convergence of the Shadow Sequence

The main result of this section, Corollary 3.19, provides a necessary and sufficient condition for the asymptotic regularity (cf. Definition 2.23) of the shadow sequence. The identities in Lemma 3.11 are key components of our approach and will be used to establish important relationships in Theorem 3.12. Using the characterization of $\overline{\text{ran}}(\text{Id} - T)$ obtained in Section 3.1, we derive critical properties of v , v_D , and v_R in Proposition 3.16. These results enable us, among other outcomes, to analyze the behavior of the sequences $(J_A T^n x - J_A T^{n+1} x)_{n \in \mathbb{N}}$ and $(J_{A^{-1}} T^n - J_{A^{-1}} T^{n+1})_{n \in \mathbb{N}}$, leading to the main conclusion of this section.

Lemma 3.11. *Let a, b, c and z be in X . Then the following hold:*

$$(i) \quad \|z\|^2 = \|a + b\|^2 + \|z - a - b - c\|^2 - \|c\|^2 + 2\langle a, z - a \rangle + 2\langle b, z - 2a - c - b \rangle + 2\langle c, z - a \rangle.$$

$$(ii) \quad \|z - a\|^2 + \|a\|^2 = \|a + b\|^2 + \|z - a - b - c\|^2 - \|c\|^2 + 2\langle b, z - 2a - c - b \rangle + 2\langle c, z - a \rangle.$$

$$(iii) \quad \langle z, a + b \rangle = \|a + b\|^2 + \langle a, z - a \rangle + \langle b, z - 2a - b - c \rangle + \langle b, c \rangle.$$

Proof. All of (i), (ii), and (iii) are obtained with direct expansion. $\square$

Theorem 3.12. *Let x and y be in X . Then, the following hold:*

$$(i) \quad \begin{aligned} \|x - y\|^2 &= \|\mathbf{T}x - \mathbf{T}y\|^2 + \|(\text{Id} - \mathbf{T})x - (\text{Id} - \mathbf{T})y - (C\mathbf{J}_A x - C\mathbf{J}_A y)\|^2 \\ &\quad + 2\langle \mathbf{J}_{A^{-1}}x - \mathbf{J}_{A^{-1}}y, \mathbf{J}_A x - \mathbf{J}_A y \rangle \\ &\quad + 2\langle \mathbf{J}_{B^{-1}}(\mathbf{R}_A - C\mathbf{J}_A)x - \mathbf{J}_{B^{-1}}(\mathbf{R}_A - C\mathbf{J}_A)y, \\ &\quad \quad \mathbf{J}_B(\mathbf{R}_A - C\mathbf{J}_A)x - \mathbf{J}_B(\mathbf{R}_A - C\mathbf{J}_A)y \rangle \\ &\quad + \langle C\mathbf{J}_A x - C\mathbf{J}_A y, \mathbf{J}_A x - \mathbf{J}_A y \rangle \\ &\quad + \langle C\mathbf{J}_A x - C\mathbf{J}_A y, \mathbf{J}_A x - \mathbf{J}_A y \rangle - \|C\mathbf{J}_A x - C\mathbf{J}_A y\|^2. \end{aligned}$$

$$(ii) \quad \begin{aligned} \|\mathbf{J}_A x - \mathbf{J}_A y\|^2 + \|\mathbf{J}_{A^{-1}}x - \mathbf{J}_{A^{-1}}y\|^2 - \|\mathbf{J}_A \mathbf{T}x - \mathbf{J}_A \mathbf{T}y\|^2 + \|\mathbf{J}_{A^{-1}} \mathbf{T}x - \mathbf{J}_{A^{-1}} \mathbf{T}y\|^2 &= \\ \|(\text{Id} - \mathbf{T})x - (\text{Id} - \mathbf{T})y - (C\mathbf{J}_A x - C\mathbf{J}_A y)\|^2 & \\ + 2\langle \mathbf{J}_A \mathbf{T}x - \mathbf{J}_A \mathbf{T}y, \mathbf{J}_{A^{-1}} \mathbf{T}x - \mathbf{J}_{A^{-1}} \mathbf{T}y \rangle & \\ + 2\langle \mathbf{J}_{B^{-1}}(\mathbf{R}_A - C\mathbf{J}_A)x - \mathbf{J}_{B^{-1}}(\mathbf{R}_A - C\mathbf{J}_A)y, \mathbf{J}_B(\mathbf{R}_A - C\mathbf{J}_A)x - \mathbf{J}_B(\mathbf{R}_A - C\mathbf{J}_A)y \rangle & \\ + \langle C\mathbf{J}_A x - C\mathbf{J}_A y, \mathbf{J}_A x - \mathbf{J}_A y \rangle & \\ + \langle C\mathbf{J}_A x - C\mathbf{J}_A y, \mathbf{J}_A x - \mathbf{J}_A y \rangle - \|C\mathbf{J}_A x - C\mathbf{J}_A y\|^2. & \end{aligned}$$

$$(iii) \quad \|\mathbf{J}_A \mathbf{T}x - \mathbf{J}_A \mathbf{T}y\|^2 + \|\mathbf{J}_{A^{-1}} \mathbf{T}x - \mathbf{J}_{A^{-1}} \mathbf{T}y\|^2 \leq \|\mathbf{J}_A x - \mathbf{J}_A y\|^2 + \|\mathbf{J}_{A^{-1}}x - \mathbf{J}_{A^{-1}}y\|^2.$$

$$(iv) \quad \begin{aligned} 0 \leq \langle x - y, (\text{Id} - \mathbf{T})x - (\text{Id} - \mathbf{T})y \rangle - \|(\text{Id} - \mathbf{T})x - (\text{Id} - \mathbf{T})y\|^2 & \\ - \langle C\mathbf{J}_A x - C\mathbf{J}_A y, \mathbf{J}_B(\mathbf{R}_A - C\mathbf{J}_A)x - \mathbf{J}_B(\mathbf{R}_A - C\mathbf{J}_A)y \rangle. & \end{aligned}$$

Proof. Applying Lemma 3.11(i) with

$$\begin{cases} z = x - y \\ a = \mathbf{J}_{A^{-1}}x - \mathbf{J}_{A^{-1}}y \\ b = \mathbf{J}_{B^{-1}}(\mathbf{R}_A - C\mathbf{J}_A)x - \mathbf{J}_{B^{-1}}(\mathbf{R}_A - C\mathbf{J}_A)y \\ c = C\mathbf{J}_A x - C\mathbf{J}_A y \end{cases}$$

yields the following:

$$\begin{aligned}
\|x - y\|^2 &= \|(\text{Id} - T)x - (\text{Id} - T)y - (C J_A x - C J_A y)\|^2 + \|Tx - Ty\|^2 \\
&\quad + 2\langle J_{A^{-1}}x - J_{A^{-1}}y, x - y - (J_{A^{-1}}x - J_{A^{-1}}y) \rangle \\
&\quad + 2\langle J_{B^{-1}}(R_A - C J_A)x - J_{B^{-1}}(R_A - C J_A)y, \\
&\quad \quad x - y - 2(J_{A^{-1}}x - J_{A^{-1}}y) - (C J_A x - C J_A y) \\
&\quad \quad - J_{B^{-1}}(R_A - C J_A)x - J_{B^{-1}}(R_A - C J_A)y \rangle \\
&\quad + 2\langle C J_A x - C J_A y, x - y - J_{A^{-1}}x - J_{A^{-1}}y \rangle - \|C J_A x - C J_A y\|^2 \quad (3.27a) \\
&\stackrel{(2.25)}{=} \|(\text{Id} - T)x - (\text{Id} - T)y - (C J_A x - C J_A y)\|^2 + \|Tx - Ty\|^2 \\
&\quad + 2\langle J_{A^{-1}}x - J_{A^{-1}}y, J_A x - J_A y \rangle \\
&\quad + 2\langle J_{B^{-1}}(R_A - C J_A)x - J_{B^{-1}}(R_A - C J_A)y, \\
&\quad \quad J_B(R_A - C J_A)x - J_B(R_A - C J_A)y \rangle \\
&\quad + 2\langle C J_A x - C J_A y, J_A x - J_A y \rangle - \|C J_A x - C J_A y\|^2, \quad (3.27b)
\end{aligned}$$

which proves (i). Moreover,

$$\|x - y\|^2 \stackrel{(2.25)}{=} \|J_A x - J_A y + J_{A^{-1}}x - J_{A^{-1}}y\|^2 \quad (3.28a)$$

$$\begin{aligned}
&= \|J_A x - J_A y\|^2 + \|J_{A^{-1}}x - J_{A^{-1}}y\|^2 \\
&\quad + 2\langle J_A x - J_A y, J_{A^{-1}}x - J_{A^{-1}}y \rangle. \quad (3.28b)
\end{aligned}$$

Hence, replacing (x, y) by (Tx, Ty) in (3.28) yields

$$\|Tx - Ty\|^2 \stackrel{(2.25)}{=} \|J_A Tx - J_A Ty + J_{A^{-1}}Tx - J_{A^{-1}}Ty\|^2 \quad (3.29a)$$

$$\begin{aligned}
&= \|J_A Tx - J_A Ty\|^2 + \|J_{A^{-1}}Tx - J_{A^{-1}}Ty\|^2 \\
&\quad + 2\langle J_A Tx - J_A Ty, J_{A^{-1}}Tx - J_{A^{-1}}Ty \rangle, \quad (3.29b)
\end{aligned}$$

which, when combined with (i) in view of Lemma 3.11(ii), proves (ii). Since C is a firmly nonexpansive operator, it follows, by (2.16), that

$$\langle C J_A x - C J_A y, J_A x - J_A y \rangle - \|C J_A x - C J_A y\|^2 \geq 0. \quad (3.30)$$

Moreover, as a consequence (2.26) and (3.1), we have that

$$\langle J_A Tx - J_A Ty, J_{A^{-1}}Tx - J_{A^{-1}}Ty \rangle \stackrel{(2.26)}{\geq} 0, \quad (3.31)$$

$$\langle J_{B^{-1}}(R_A - C J_A)x - J_{B^{-1}}(R_A - C J_A)y, J_B(R_A - C J_A)x - J_B(R_A - C J_A)y \rangle \stackrel{(2.26)}{\geq} 0, \quad (3.32)$$

$$\langle C J_A x - C J_A y, J_A x - J_A y \rangle \stackrel{(2.19)}{\geq} 0. \quad (3.33)$$

Hence, each term on the right hand side of (ii) is nonnegative from which (iii) is immediately derived. Eventually, applying Lemma 3.11(iii) with

$$\begin{cases} z = x - y \\ a = J_A x - J_A y \\ b = -(J_B(R_A - C J_A)x - J_B(R_A - C J_A)y) \\ c = -(C J_A x - C J_A y) \end{cases}$$

proves (iv) as follows:

$$\begin{aligned} 0 &\stackrel{(2.26)}{\leq} \langle J_B(R_A - C J_A)x - J_B(R_A - C J_A)y, 2(J_A x - J_A y) - (x - y) - (C J_A x - C J_A y) \\ &\quad - (J_B(R_A - C J_A)x - J_B(R_A - C J_A)y) \rangle \quad (3.34a) \\ &\stackrel{(2.25)}{=} \langle J_B(R_A - C J_A)x - J_B(R_A - C J_A)y, J_{B^{-1}}(R_A - C J_A)x - J_{B^{-1}}(R_A - C J_A)y \rangle \quad (3.34b) \\ &\stackrel{(3.9)}{=} \langle x - y, (\text{Id} - T)x - (\text{Id} - T)y \rangle - \|(\text{Id} - T)x - (\text{Id} - T)y\|^2 \\ &\quad - \langle C J_A x - C J_A y, J_B(R_A - C J_A)x - J_B(R_A - C J_A)y \rangle. \quad \square \end{aligned}$$

Remark 3.13. Theorem 3.12(iv) recovers [23, Lemma 3.2], which is a precursor to proving the convergence of (1.4) in the *consistent* case.

Corollary 3.14. *Let $x \in X$ and let $y \in X$. Then, the following hold:*

- (i) $\sum_{k \in \mathbb{N}} \|(\text{Id} - T) T^k x - (\text{Id} - T) T^k y - (C J_A T^k x - C J_A T^k y)\|^2 < +\infty.$
- (ii) $\sum_{k \in \mathbb{N}} \langle J_{A^{-1}} T^k x - J_{A^{-1}} T^k y, J_A T^k x - J_A T^k y \rangle < +\infty.$
- (iii) $\sum_{k \in \mathbb{N}} \langle J_{B^{-1}}(R_A - C J_A) T^k x - J_{B^{-1}}(R_A - C J_A) T^k y, J_B(R_A - C J_A) T^k x - J_B(R_A - C J_A) T^k y \rangle < +\infty.$
- (iv) $\sum_{k \in \mathbb{N}} \langle C J_A T^k x - C J_A T^k y, J_A T^k x - J_A T^k y \rangle < +\infty.$
- (v) $\sum_{k \in \mathbb{N}} \|C J_A T^k x - C J_A T^k y\|^2 < +\infty.$

Proof. Let $(x, y) \in X \times X$. Anti-telescoping Theorem 3.12(i) when applied $n \in \mathbb{N}$ times

with $(T^n x, T^n y)$ yields the following:

$$\begin{aligned}
& \stackrel{(3.27)}{0} \leq \|x - y\|^2 - \|T^n x - T^n y\|^2 \\
& = \sum_{k \in \mathbb{N}} \|T^k x - T^k y\|^2 - \|T^{k+1} x - T^{k+1} y\|^2 \\
& = (\|x - y\|^2 - \|T^1 x - T^1 y\|^2) + (\|T^1 x - T^1 y\|^2 - \|T^2 x - T^2 y\|^2) + \dots \\
& \quad + (\|T^{k-1} x - T^{k-1} y\|^2 - \|T^k x - T^k y\|^2) \\
& \stackrel{(3.27)}{=} \sum_{k \in \mathbb{N}} \underbrace{\|(\text{Id} - T) T^k x - (\text{Id} - T) T^k y - (C J_A T^k x - C J_A T^k y)\|^2}_{\geq 0} \\
& \quad + 2 \sum_{k \in \mathbb{N}} \underbrace{\langle J_{A^{-1}} T^k x - J_{A^{-1}} y, J_A T^k x - J_A T^k y \rangle}_{\geq 0 \text{ by (2.26)}} \\
& \quad + 2 \sum_{n \in \mathbb{N}} \underbrace{\langle J_{B^{-1}}(R_A - C J_A) T^k x - J_{B^{-1}}(R_A - C J_A) T^k y, \\
& \quad \quad J_B(R_A - C J_A) T^k x - J_B(R_A - C J_A) T^k y \rangle}_{\geq 0 \text{ by (2.26)}} \\
& \quad + \sum_{k \in \mathbb{N}} \underbrace{\langle C J_A T^k x - C J_A T^k y, J_A T^k x - J_A T^k y \rangle}_{\geq 0 \text{ by (3.2)}} \\
& \quad + \sum_{k \in \mathbb{N}} \underbrace{\langle C J_A T^k x - C J_A T^k y, J_A x - J_A T^k y \rangle - \|C J_A T^k x - C J_A T^k y\|^2}_{\geq 0 \text{ by (3.2) and (2.16)}} \\
& \stackrel{(3.27)}{\leq} \|x - y\|^2 < \infty. \quad \square
\end{aligned}$$

Definition 3.15. Let D be a nonempty convex subset of X . The *recession cone* of D is $\text{rec } D = \{x \in X \mid x + D \subseteq D\}$.

Proposition 3.16. *We have*

- (i) $v_D \in (-\text{rec } \overline{\text{dom}} A)^\ominus \cap (\text{rec } \overline{\text{dom}} B)^\ominus = (-\text{rec } \overline{\text{dom}} A)^\ominus \cap (\text{rec } \overline{\text{dom}} B)$.
- (ii) $v_D \in (-\text{rec } \overline{\text{ran}} A) \cap (\text{rec}(\overline{\text{ran}} B + \overline{\text{ran}} C))$.
- (iii) $v_R \in (-\text{rec } \overline{\text{ran}} A)^\ominus \cap (-\text{rec}(\overline{\text{ran}} B + \overline{\text{ran}} C))^\ominus = -(\text{rec } \overline{\text{ran}} A)^\ominus \cap (\text{rec}(\overline{\text{ran}} B + \overline{\text{ran}} C))^\ominus$.
- (iv) $v_R \in (-\text{rec } \overline{\text{dom}} A) \cap (-\text{rec } \overline{\text{dom}} B) = -(\text{rec } \overline{\text{dom}} A \cap \text{rec } \overline{\text{dom}} B)$.
- (v) $\langle v_D, v_R \rangle = 0$.

$$(vi) \quad v_D + v_R \in (\overline{\text{dom } A - \text{dom } B}) \cap (\overline{\text{ran } A + \text{ran } B + \text{ran } C}).$$

$$(vii) \quad v = v_D + v_R.$$

$$(viii) \quad \|v\|^2 = \|v_D\|^2 + \|v_R\|^2 = \|(v_D, v_R)\|^2.$$

Proof. The result follows from [10, Proposition 3.1] when applied to A and $B + C$ in view of Remark 3.1, Proposition 3.10, and (3.4). $\square$

Lemma 3.17. *We have*

$$(i) \quad J_A - J_A T = J_{A^{-1}} T + J_{B^{-1}}(R_A - C J_A) + C J_A. \quad \text{Consequently, } \text{ran}(J_A - J_A T) \subseteq \text{ran } A + \text{ran } B + \text{ran } C.$$

$$(ii) \quad J_{A^{-1}} - J_{A^{-1}} T = J_A T - J_B(R_A - C J_A). \quad \text{Consequently, } \text{ran}(J_{A^{-1}} - J_{A^{-1}} T) \subseteq \text{dom } A - \text{dom } B.$$

Proof. Note that, as a consequence of (2.25), $J_A - J_{B^{-1}}(R_A - C J_A) = J_A - (R_A - C J_A) + J_B(R_A - C J_A) = J_A - 2J_A + \text{Id} + C J_A + J_B(R_A - C J_A) = \text{Id} - J_A + J_B(R_A - C J_A) + C J_A = T + C J_A = J_A T + J_{A^{-1}} T + C J_A \Leftrightarrow J_A - J_A T = J_{A^{-1}} + J_{B^{-1}}(R_A - C J_A) + C J_A$, which, when combined with Fact 2.57, yields (i). Similarly, $J_{A^{-1}} + J_B(R_A - C J_A) = \text{Id} - J_A + J_B(R_A - C J_A) = T = J_A T + J_{A^{-1}} T \Leftrightarrow J_{A^{-1}} - J_{A^{-1}} T = J_A T - J_B(R_A - C J_A)$, which, when combined with Fact 2.57, yields (ii). $\square$

Proposition 3.18. *Let $x \in X$. Then, the following hold:*

$$(i) \quad \|J_A T^n x - J_A T^{n+1} x\|^2 + \|J_{A^{-1}} T^n x - J_{A^{-1}} T^{n+1} x\|^2 \rightarrow \|v\|^2 = \|v_D\|^2 + \|v_R\|^2.$$

$$(ii) \quad J_A T^n x - J_A T^{n+1} x \rightarrow v_R.$$

$$(iii) \quad J_{A^{-1}} T^n x - J_{A^{-1}} T^{n+1} x \rightarrow v_D.$$

Proof. From Lemma 3.17(i) and Lemma 3.17(ii), we conclude that $(J_A T^n x - J_A T^{n+1} x, J_{A^{-1}} T^n x - J_{A^{-1}} T^{n+1} x)$ lies in $(\text{ran } A + \text{ran } B + \text{ran } C) \times (\text{dom } A - \text{dom } B)$. Moreover, note that

$$\|J_A T^n x - J_A T^{n+1} x\|^2 + \|J_{A^{-1}} T^n x - J_{A^{-1}} T^{n+1} x\|^2 \tag{3.35a}$$

$$= \|T^n x - T^{n+1} x\|^2 - 2\langle J_A T^n x - J_A T^{n+1} x, J_{A^{-1}} T^n x - J_{A^{-1}} T^{n+1} x \rangle \tag{3.35b}$$

$$\rightarrow \|v\|^2 - 2 \cdot 0 = \|v\|^2 = \|v_D\|^2 + \|v_R\|^2, \tag{3.35c}$$

where (3.35c) follows from Corollary 3.14(ii), Fact 2.38, and Proposition 3.16(viii). As a consequence of Fact 2.37 in view of (3.6),

$$(J_A T^n x - J_A T^{n+1} x, J_{A^{-1}} T^n x - J_{A^{-1}} T^{n+1} x) \rightarrow (v_R, v_D). \quad (3.36)$$

Equivalently,

$$J_A T^n x - J_A T^{n+1} x \rightarrow v_R \quad (3.37)$$

and

$$J_{A^{-1}} T^n x - J_{A^{-1}} T^{n+1} x \rightarrow v_D, \quad (3.38)$$

establishing (ii) and (iii), respectively. $\square$

Corollary 3.19. *Let $x \in X$. Then,*

(i) $(J_A T^n x)_{n \in \mathbb{N}}$ *is asymptotically regular if and only if $v_R = 0$.*

(ii) $(J_{A^{-1}} T^n x)_{n \in \mathbb{N}}$ *is asymptotically regular if and only if $v_D = 0$.*

Proof. The claims follow from Proposition 3.18(ii) and 3.18(iii), respectively. $\square$

3.3 On the Sets $\text{Fix } T(\cdot + v)$ and $\text{Fix}(T + v)$

Let T be as defined in (3.3). Then, the *normal problem* associated with the triplet (A, B, C) is

$$\text{find } x \in X \text{ such that } 0 \in Ax + B(x - v) + Cx - v. \quad (3.39)$$

The primal and dual pairs (cf. Definition 3.5) are $(A, B(\cdot - v) + C - v)$ and $(A^{-1}, (B(\cdot - v) + C - v)^{-\odot})$, respectively. Additionally, we let

$$Z := Z_{(A, B(\cdot - v) + C - v)} = \{x \in X \mid 0 \in Ax + B(x - v) + Cx - v\} \quad (3.40)$$

and

$$K := K_{(A, B(\cdot - v) + C - v)} = \{x \in X \mid 0 \in A^{-1} + (B(\cdot - v) + C - v)^{-\odot}\}, \quad (3.41)$$

denote the primal and dual solutions of (3.39), respectively. The main result of this section is the characterization of the fixed point sets $\text{Fix } T(\cdot + v)$ and $\text{Fix}(T + v)$ in terms of Z and K .

Definition 3.20. Let z be in X . We define $\mathbf{K} : X \rightrightarrows X : z \mapsto (Az - v) \cap (-(B(z - v) + Cz))$. Then, $\text{gra } \mathbf{K} = \{(z, k) \mid k \in (Az - v) \cap (-(B(z - v) + Cz))\}$.

Lemma 3.21. *The mapping $\phi : \text{gra } \mathbf{K} \rightarrow \text{Fix } T(\cdot + v) : (z, k) \rightarrow z + k$ is a well-defined bijection, with $\phi^{-1} : x \mapsto (J_A(x + v), x - J_A(x + v))$.*

Proof. The conclusion follows from [24, Theorem 4.5] when applied to $A - v$ and $(B + C(\cdot + v))(\cdot - v)$. $\square$

Lemma 3.22. *The mapping $\psi : \text{gra } \mathbf{K} \rightarrow \text{Fix}(T + v) : (z, k) \rightarrow z + k + v$ is a well-defined bijection, with $\psi^{-1} : x \mapsto (J_A x, x - J_A(x) - v)$.*

Proof. Note that, from Fact 2.35, $\text{Fix } T(\cdot + v) = -v + \text{Fix}(T + v)$. Hence, the mapping $\tau : \text{Fix } T(\cdot + v) \rightarrow \text{Fix}(T + v) : \bar{x} \mapsto \bar{x} + v$ defines a bijection. Moreover, $\text{gra } \psi = \text{gra}(\tau \circ \phi)$. Therefore, the result is a consequence of Lemma 3.21. $\square$

Definition 3.23 ([25, Equation (2.2)]). The extended solution set is defined as

$$S := S_e(A, B(\cdot - v) + C - v) = \{(z, k) \in X \times X \mid k \in Az \cap (-(B(z - v) + Cz - v))\}.$$

Lemma 3.24. *S has the following equivalent descriptions:*

$$(i) \ S = \{(J_{(A-v)} \times J_{(A-v)^{-1}})(\bar{x}, \bar{x}) \mid \bar{x} \in \text{Fix } T(\cdot + v)\}.$$

$$(ii) \ S = (0, -v) + \{(J_A \times J_{A^{-1}})(\bar{x}, \bar{x}) \mid \bar{x} \in \text{Fix}(T + v)\}.$$

Suppose that A and B are, in addition, paramonotone. Then, the following also holds:

$$(iii) \ \text{Fix } T(\cdot + v) = Z + K. \text{ Consequently, } \text{Fix}(T + v) = Z + K + v.$$

$$(iv) \ S = Z \times K.$$

Proof. Note that (i) and (ii) immediately follow from Lemma 3.21 and Lemma 3.22, respectively. Recall that $B + C$ is paramonotone (cf. Remark 3.1). Hence, applying [24, Corollary 5.5(ii)] to A and $B(\cdot - v) + C - v$ yields that $\text{gra } \mathbf{K}$ is the rectangle $Z \times K$, which, when combined with Lemma 3.21, Lemma 3.22, and Fact 2.35 proves (iii) and, when combined with (i), proves (iv). $\square$

3.4 Towards Fejér Monotonicity

In this section, we demonstrate the Fejér monotonicity of the sequence

$$(J_A T^n + n v_R, J_{A^{-1}} T^n x + n v_D)_{n \in \mathbb{N}}$$

with respect to S , the extended solution (cf. Definition 3.23), which directly leads to the boundedness of the sequence $(J_A T^n x)_{n \in \mathbb{N}}$ (resp. $(J_{A^{-1}} T^n x)_{n \in \mathbb{N}}$) when $v_R = 0$ (resp. $v_D = 0$) as established in Corollary 3.28 (resp. Corollary 3.29). To achieve this, we first prove the key identities presented in Lemma 3.25, along with their generalized forms in Proposition 3.26.

Lemma 3.25. *Suppose that $\bar{x} \in \text{Fix}(T + v)$. Then, the following hold:*

$$(i) \quad T \bar{x} = \bar{x} - v.$$

$$(ii) \quad J_A T \bar{x} = J_A \bar{x} - v_R.$$

$$(iii) \quad J_{A^{-1}} T \bar{x} = J_{A^{-1}} \bar{x} - v_D.$$

Proof. Note that (i) immediately follows by assumption. From Lemma 3.17(i) and 3.17(ii), we have

$$J_A \bar{x} - J_{A^{-1}} T \bar{x} \in \text{ran } A + \text{ran } B + \text{ran } C \quad (3.42)$$

and

$$J_{A^{-1}} \bar{x} - J_{A^{-1}} T \bar{x} \in \text{dom } A - \text{dom } B. \quad (3.43)$$

Then, from Proposition 3.16(viii), (3.5) and (i), we have

$$\|(v_R, v_D)\|^2 = \|v_R\|^2 + \|v_D\|^2 \stackrel{(3.5)}{\leq} \|(J_A \bar{x} - J_A T \bar{x}, J_{A^{-1}} \bar{x} - J_{A^{-1}} T \bar{x})\|^2 \quad (3.44a)$$

$$= \|J_A \bar{x} - J_A T \bar{x}\|^2 + \|J_{A^{-1}} \bar{x} - J_{A^{-1}} T \bar{x}\|^2 \quad (3.44b)$$

$$\begin{aligned} &\leq \|J_A \bar{x} - J_A T \bar{x}\|^2 + \|J_{A^{-1}} \bar{x} - J_{A^{-1}} T \bar{x}\|^2 \\ &\quad + 2 \underbrace{\langle J_A \bar{x} - J_A T \bar{x}, J_{A^{-1}} \bar{x} - J_{A^{-1}} T \bar{x} \rangle}_{\geq 0 \text{ by (2.26)}}, \end{aligned} \quad (3.44c)$$

$$= \|J_A \bar{x} - J_A T \bar{x} + J_{A^{-1}} \bar{x} - J_{A^{-1}} T \bar{x}\|^2 \quad (3.44d)$$

$$\stackrel{(2.25)}{=} \|\bar{x} - T \bar{x}\|^2 \quad (3.44e)$$

$$\stackrel{(i)}{=} \|\bar{x} - (\bar{x} - v)\|^2 \quad (3.44f)$$

$$= \|v\|^2 = \|v_R\|^2 + \|v_D\|^2 = \|(v_R, v_D)\|^2. \quad (3.44g)$$

Hence, the inequalities reduce to equalities. Consequently,

$$\langle J_A \bar{x} - J_A T \bar{x}, J_{A^{-1}} \bar{x} - J_{A^{-1}} T \bar{x} \rangle = 0 \quad (3.45)$$

and, by Fact 2.24, Fact 2.33, Proposition 3.10, and (3.6),

$$(J_A \bar{x} - J_A T \bar{x}, J_{A^{-1}} \bar{x} - J_{A^{-1}} T \bar{x}) = (v_R, v_D), \quad (3.46)$$

which proves (ii) and (iii). $\square$

Proposition 3.26. *Suppose that $\bar{x} \in \text{Fix}(T + v)$. Let $n \in \mathbb{N}$. Then, the following hold*

$$(i) \quad T^n \bar{x} = \bar{x} - n v.$$

$$(ii) \quad J_A T^n \bar{x} = J_A \bar{x} - n v_R.$$

$$(iii) \quad J_{A^{-1}} T^n \bar{x} = J_{A^{-1}} \bar{x} - n v_D.$$

Proof. Applying Fact 2.52(i) yields (i). We proceed to simultaneously proving (ii) and (iii) by induction. When $n = 0$, both statements trivially hold in view of (i). The cases when $n = 1$ follow from (ii) and (iii). Hence, suppose that for some $n > 1$, it holds that $(J_A T^n \bar{x}, J_{A^{-1}} T^n \bar{x}) = (J_A \bar{x} - n v_R, J_{A^{-1}} \bar{x} - n v_D)$. Then, Lemma 3.17(i) and 3.17(ii) yield, with x replaced by $\bar{x}$, that

$$J_A T^n \bar{x} - J_A T^{n+1} \bar{x} \in \text{ran } A + \text{ran } B + \text{ran } C \quad (3.47)$$

and

$$J_{A^{-1}} T^n \bar{x} - J_{A^{-1}} T^{n+1} \bar{x} \in \text{dom } A - \text{dom } B. \quad (3.48)$$

As a consequence,

$$\|(v_R, v_D)\|^2 \stackrel{(3.5)}{\leq} \|(J_A T^n \bar{x} - J_A T^{n+1} \bar{x}, J_{A^{-1}} T^n \bar{x} - J_{A^{-1}} T^{n+1} \bar{x})\|^2 \quad (3.49a)$$

$$= \|J_A T^n \bar{x} - J_A T^{n+1} \bar{x}\|^2 + \|J_{A^{-1}} T^n \bar{x} - J_{A^{-1}} T^{n+1} \bar{x}\|^2 \quad (3.49b)$$

$$\begin{aligned} &\leq \|J_A T^n \bar{x} - J_A T^{n+1} \bar{x}\|^2 + \|J_{A^{-1}} T^n \bar{x} - J_{A^{-1}} T^{n+1} \bar{x}\|^2 \\ &\quad + 2 \underbrace{\langle J_A T^n \bar{x} - J_A T^{n+1} \bar{x}, J_{A^{-1}} T^n \bar{x} - J_{A^{-1}} T^{n+1} \bar{x} \rangle}_{\geq 0 \text{ by (2.26)}} \end{aligned} \quad (3.49c)$$

$$= \|J_A T^n \bar{x} - J_A T^{n+1} \bar{x} + J_{A^{-1}} T^n \bar{x} - J_{A^{-1}} T^{n+1} \bar{x}\|^2 \quad (3.49d)$$

$$\stackrel{(2.25)}{=} \|T^n \bar{x} - T^{n+1} \bar{x}\|^2 \quad (3.49e)$$

$$\stackrel{(i)}{=} \|\bar{x} - n v - (\bar{x} - (n+1) v)\|^2 \quad (3.49f)$$

$$= \|v\|^2 = \|(v_D, v_R)\|^2. \quad (3.49g)$$

Hence, by Fact 2.24, Fact 2.33, and (3.6),

$$(\mathbf{v}_R, \mathbf{v}_D) = (J_A T^n \bar{x} - J_A T^{n+1} \bar{x}, J_{A^{-1}} T^n \bar{x} - J_{A^{-1}} T^{n+1} \bar{x}) \quad (3.50)$$

and

$$\langle J_A T^n \bar{x} - J_A T^{n+1} \bar{x}, J_{A^{-1}} T^n \bar{x} - J_{A^{-1}} T^{n+1} \bar{x} \rangle = 0, \quad (3.51)$$

which, when combined with the inductive hypothesis, complete the proof as follows:

$$\begin{aligned} (J_A T^{n+1} \bar{x}, J_{A^{-1}} T^{n+1} \bar{x}) &= (J_A T^n \bar{x}, J_{A^{-1}} T^n \bar{x}) - (\mathbf{v}_R, \mathbf{v}_D), \\ &= (J_A \bar{x} - n \mathbf{v}_R, J_{A^{-1}} \bar{x} - n \mathbf{v}_D) - (\mathbf{v}_R, \mathbf{v}_D) \\ &= (J_A \bar{x} - (n+1) \mathbf{v}_R, J_{A^{-1}} \bar{x} - (n+1) \mathbf{v}_D). \quad \square \end{aligned}$$

Theorem 3.27. *Suppose that $\bar{x} \in \text{Fix}(T + \mathbf{v})$. Let $x \in X$ and $n \in \mathbb{N}$. Then, the following hold:*

- (i) *The sequence $(J_A T^n x + n \mathbf{v}_R, J_{A^{-1}} T^n x + n \mathbf{v}_D)_{n \in \mathbb{N}}$ is Fejér monotone with respect to the set $\{(J_A \times J_{A^{-1}})(\bar{x}, \bar{x}) \mid \bar{x} \in \text{Fix}(T + \mathbf{v})\}$.*
- (ii) *The sequence $(J_A T^n x + n \mathbf{v}_R, J_{A^{-1}} T^n x + n \mathbf{v}_D)_{n \in \mathbb{N}}$ is bounded.*
- (iii) *The sequence $(J_B(R_A - C J_A) T^n x + n \mathbf{v}_R, J_{B^{-1}}(R_A - C J_A) T^n x + n \mathbf{v}_D)_{n \in \mathbb{N}}$ is bounded.*
- (iv) *The sequence $((0, -\mathbf{v}) + (J_A T^n x + n \mathbf{v}_R, J_{A^{-1}} T^n x + n \mathbf{v}_D))_{n \in \mathbb{N}}$ is Fejér monotone with respect to S .*

Suppose that A and B are, in addition, paramonotone. Then, the following also holds:

- (v) *The sequence $((0, -\mathbf{v}) + (J_A T^n x + n \mathbf{v}_R, J_{A^{-1}} T^n x + n \mathbf{v}_D))_{n \in \mathbb{N}}$ is Fejér monotone with respect to $Z \times K$.*

Proof. Note that, as a consequence of Theorem 3.12(iii), for all $(x, y) \in X \times X$,

$$\begin{aligned} &\|J_A T^{n+1} x - J_A T^{n+1} y\|^2 + \|J_{A^{-1}} T^{n+1} x - J_{A^{-1}} T^{n+1} y\|^2 \\ &\leq \|J_A T^n x - J_A T^n y\|^2 + \|J_{A^{-1}} T^n x - J_{A^{-1}} T^n y\|^2. \end{aligned} \quad (3.52)$$

As a consequence of Proposition 3.26(ii) and 3.26(iii), and replacing y in (3.52) with $\bar{x} \in \text{Fix}(\mathbf{v} + T)$, we obtain the following

$$\begin{aligned} &\|J_A T^{n+1} x - (J_A \bar{x} - (n+1) \mathbf{v}_R)\|^2 + \|J_{A^{-1}} T^{n+1} x - (J_{A^{-1}} \bar{x} - (n+1) \mathbf{v}_D)\|^2 \\ &\leq \|J_A T^n x - (J_A \bar{x} - n \mathbf{v}_R)\|^2 + \|J_{A^{-1}} T^n x - (J_{A^{-1}} \bar{x} - n \mathbf{v}_D)\|^2, \end{aligned} \quad (3.53)$$

which is equivalent to

$$\begin{aligned} & \|J_A T^{n+1} x + (n+1) v_R - J_A \bar{x}\|^2 + \|J_{A^{-1}} T^{n+1} x + (n+1) v_D - J_{A^{-1}} \bar{x}\|^2 \\ & \leq \|J_A T^n x + n v_R - J_A \bar{x}\|^2 + \|J_{A^{-1}} T^n x + n v_D - J_{A^{-1}} \bar{x}\|^2, \end{aligned} \quad (3.54)$$

which, by (2.23), proves (i) and, in view of Fact 2.49, proves (ii). By using (3.11), (iii) immediately follows from (ii) and Proposition 3.2(i). As for (iv), it is a consequence of (i), Lemma 2.50 when applied with $u = (0, -v)$ and $D = \{(J_A \times J_{A^{-1}})(\bar{x}, \bar{x}) \mid \bar{x} \in \text{Fix}(T+v)\}$, and Lemma 3.24(ii). When A and B are paramonotone, (v) follows from Remark 3.1, (iv) and Lemma 3.24(iv). $\square$

Corollary 3.28. *Suppose that $v_R = 0$ and $\bar{x} \in \text{Fix}(T+v)$. Let $x \in X$ and $n \in \mathbb{N}$. Then, the following hold*

- (i) $J_A T^n \bar{x} = J_A(\bar{x} - n v_D) = J_A \bar{x}$.
- (ii) $(J_A T^n x)_{n \in \mathbb{N}}$ is bounded.
- (iii) Consequently, $(C J_A T^n x)_{n \in \mathbb{N}}$ is bounded.
- (iv) $(J_B(R_A - C J_A) T^n x)_{n \in \mathbb{N}}$ is bounded.

Proof. Using Proposition 3.26(i) and Proposition 3.26(ii), we can rewrite (i) as $J_A T^n \bar{x} = J_A(\bar{x} - n v)$. From Proposition 3.16(vii), this is equivalent to $J_A T^n(\bar{x} - n(v_D + v_R))$. When $v_R = 0$, $v = v_D$ and the result follows from Proposition 3.26(ii) and the initial assumption on $\bar{x}$. The claim in (ii) is a consequence of Theorem 3.27(ii) and (i). Consequently, (iii) follows from (ii) and Remark 2.30. Eventually, from Proposition 3.2, we rewrite (iv) as $((T - \text{Id} - J_A) T^n x)_{n \in \mathbb{N}} = (T^{n+1} x - T^n x - J_A T^n x)_{n \in \mathbb{N}}$. Hence, the result follows from Fact 2.38 and (ii). $\square$

Corollary 3.29. *Suppose that $v_D = 0$ and $\bar{x} \in \text{Fix}(T+v)$. Let $x \in X$ and $n \in \mathbb{N}$. Then, the following hold*

- (i) $J_{A^{-1}} T^n \bar{x} = J_{A^{-1}}(\bar{x} - n v_R) = J_{A^{-1}} \bar{x}$.
- (ii) $(J_{A^{-1}} T^n x)_{n \in \mathbb{N}}$ is bounded.
- (iii) $(J_{B^{-1}}(R_A - C J_A) T^n x)_{n \in \mathbb{N}}$ is bounded.

Proof. In the same spirit as Corollary 3.28, we rewrite (i) as $J_A T^n(\bar{x} - n(v_D + v_R))$, by employing Proposition 3.26(i), Proposition 3.26(iii), and Proposition 3.16(vii). Then, the result follows when $v_D = 0$ and from the assumption on $\bar{x}$. Similarly, (ii) is a consequence of Theorem 3.27(ii) in view of (i). $\square$

3.5 An Equivalent Description of Z

When $v_R = 0$, we provide an equivalent description of Z , the set of minimizers to (3.39), however, eliminating the need for the outer perturbation, that is, one could solve the following monotone inclusion problem instead:

$$\text{find } x \in X \text{ such that } 0 \in Ax + B(x - v_D) + Cx. \quad (3.55)$$

The following proposition is crucial to this result.

Proposition 3.30. *Let $(\alpha, \beta) \in \mathbb{R}_+ \times \mathbb{R}_-$ and $(x, x^*) \in X \times X$. Suppose that $x^* \in -Ax \cap (B(x - v) + Cx - v)$. Then, the following hold:*

- (i) $x^* + \alpha v_D \in -Ax \cap (B(x - v) + Cx - v)$.
- (ii) $x + \beta v_R \in A^{-1}(-x^*) \cap (v + (B + C(\cdot + v) - v)^{-1}x^*)$.
- (iii) $x^* \in -\alpha v_D - Ax \cap (B(x + \beta v_R - v) + C(x + \beta v_R) - v)$.

Proof. Suppose that $(a, a^*) \in \text{gra } A$ and $(q, q^*) \in \text{gra}(B + C)$. Then, by the assumption on x^* , we have $-x^* \in Ax$. Furthermore,

$$\langle a - x, a^* - (-x^* - \alpha v_D) \rangle = \langle a - x, a^* - (-x^*) \rangle + \langle a - x, \alpha v_D \rangle \quad (3.56a)$$

$$\stackrel{(2.19)}{\geq} \langle a - x, \alpha v_D \rangle \quad (3.56b)$$

$$= \alpha \langle v - (a - (x - v)), 0 - v_D \rangle \quad (3.56c)$$

$$= \alpha \langle v_D - (a - (x - v)), 0 - v_D \rangle \quad (3.56d)$$

$$\geq 0, \quad (3.56e)$$

where (3.56b) follows by the monotonicity of A , (3.56d) by Proposition 3.16(vii) and Proposition 3.16(v). Eventually, (3.56e) is the negation of (2.11) which is a consequence of (3.5). Additionally, $x^* \in B(x - v) + Cx - v$ if and only if there exists $(b^*, c^*) \in X \times X$ such that $x^* = b^* + c^* - v$ where $b^* \in B(x - v)$ and $c^* = Cx$. Similarly, $(q, q^*) \in \text{gra}(B(\cdot - v) + C - v)$ if and only if there exists $(q_1^*, q_2^*) \in X \times X$ such that $q^* = q_1^* + q_2^* - v$ where $q_1^* \in B(q - v)$ and $q_2^* = Cq$. Then,

$$\langle q - (x - v), q_1^* - (b^* + \alpha v_D) \rangle = \langle q - (x - v), q_1^* - b^* \rangle + \alpha \langle q - (x - v), 0 - v_D \rangle \quad (3.57a)$$

$$\geq \alpha \langle q - (x - v), 0 - v_D \rangle \quad (3.57b)$$

$$= \alpha \langle v_D - (x - q), 0 - v_D \rangle \quad (3.57c)$$

$$\stackrel{(3.5)}{\geq} 0. \quad (3.57d)$$

As a consequence of (2.20), we have that $b^* + \alpha v_D \in B(x - v)$, which implies that $b^* + c^* + \alpha v_D - v = x^* + \alpha v_D \in B(x - v) + Cx - v$. Hence, $x^* + \alpha v_D \in -Ax \cap (B(x - v) + Cx - v)$, which is (i). Furthermore, by the assumption on x^* , $x \in A^{-1}(-x^*) \cap (v + (B + C(\cdot + v) - v)^{-1}x^*)$. As such,

$$\langle a^* - (-x^*), a - (x + \beta v_R) \rangle = \langle a^* - (-x^*), a - x \rangle + \beta \langle a^* - (-x^*), 0 - v_R \rangle \quad (3.58a)$$

$$\stackrel{(2.19)}{\geq} \beta \langle a^* - (-x^*), 0 - v_R \rangle \quad (3.58b)$$

$$= \beta \langle a^* + x^*, 0 - v_R \rangle \quad (3.58c)$$

$$= \beta \langle a^* + (x^* + v) - v, 0 - v_R \rangle \quad (3.58d)$$

$$= \beta \langle a^* + (x^* + v) - v_R, 0 - v_R \rangle \quad (3.58e)$$

$$\stackrel{(3.5)}{\geq} 0, \quad (3.58f)$$

where (3.59d) follows from Proposition 3.16(vii) and Proposition 3.16(v). Moreover,

$$\langle q^* - x^*, q - (x + \beta v_R - v) \rangle = \langle q^* - x^*, q - (x - v) \rangle + \beta \langle q^* - x^*, 0 - v_R \rangle \quad (3.59a)$$

$$\stackrel{(3.8)}{\geq} \beta \langle q^* - x^*, 0 - v_R \rangle \quad (3.59b)$$

$$= \beta \langle (q^* + v) - x^* - v, 0 - v_R \rangle \quad (3.59c)$$

$$= \beta \langle (q^* + v) + (-x^*) - v_R, 0 - v_R \rangle \quad (3.59d)$$

$$\stackrel{(3.5)}{\geq} 0, \quad (3.59e)$$

where (3.59d) follows from Proposition 3.16(vii) and Proposition 3.16(v). Hence, by (2.20), $x + \beta v_R \in A^{-1}(-x^*) \cap (v + (B + C(\cdot + v) - v)^{-1}x^*)$, establishing (ii). Then, from (i), $x^* + \alpha v_D \in -Ax$, if and only if $-x^* - \alpha v_D \in Ax$. Similarly, from (ii), $x + \beta v_R \in (v + (B + C(\cdot + v) - v)^{-1}x^*)$, if and only if, $x^* \in (B + C(\cdot + v) - v)(x + \beta v_R - v) = B(x + \beta v_R - v) + C(x + \beta v_R) - v$, from both of which we deduce (iii). $\square$

Proposition 3.31. *Let $\hat{Z} := \{x \in X \mid 0 \in Ax + B(x - v_D) + Cx - v_R\}$. Then, $\hat{Z} \subseteq Z$. Additionally, if $v_R = 0$, then $\hat{Z} = Z$.*

Proof. Suppose that $x \in \hat{Z}$ and $(q, q^*) \in \text{gra}(B + C)$, then, there exists $x^* \in -Ax \cap$

$(-v_R + B(x - v_D) + Cx) = -Ax \cap (-v_R + (B + C(\cdot + v_D))(x - v_D))$. Consequently,

$$\langle q - (x - v), q^* - (x^* + v) \rangle = \langle q - (x - v_D - v_R), q^* - (x^* + v_D + v_R) \rangle \quad (3.60a)$$

$$\begin{aligned} &= \langle q - (x - v_D), q^* - (x^* + v_R) \rangle \\ &\quad + \langle q - (x - v_D), 0 - v_D \rangle \\ &\quad + \langle v_R, q^* - (x^* + v_R) \rangle \end{aligned} \quad (3.60b)$$

$$\geq \langle v_D - (x - q), 0 - v_D \rangle + \langle v_R, q^* - x^* - v_R \rangle \quad (3.60c)$$

$$\geq 0. \quad (3.60d)$$

Proposition 3.16(vii) and Proposition 3.16(v) yield (3.60a) and (3.60b), respectively. As a consequence of Remark 3.1, $\langle q - (x - v_D), 0 - v_D \rangle$ is nonnegative, which justifies (3.60c). Eventually, note that $(x, q) \in \text{dom } A \times \text{dom}(B + C) = \text{dom } A \times \text{dom } B$ and $(-x^*, q) \in \text{ran } A \times \text{ran}(B + C)$. Hence, (3.60d) follows from (3.5) and (2.11). Again, Remark 3.1 yields $x^* + v \in (B + C(\cdot + v))(x - v)$, which when combined with $-x^* \in Ax$, shows that $x \in Z$. Suppose that $v_R = 0$, consequently, by Proposition 3.16(vii), $v = v_D$. Let $x \in Z$, then, there exists $x^* \in X$ such that $x^* \in -Ax \cap (-v + B(x - v) + Cx)$. Furthermore, as a consequence of Proposition 3.30(iii), $x^* \in -v_D - Ax \cap (B(x - v) + Cx - v_D)$ if and only if $0 \in Ax + B(x - v_D) + Cx$, i.e., $x \in \hat{Z}$. $\square$

Chapter 4

Convergence of the Shadow Sequence

This chapter builds on the results established in Chapter 3 to prove our main result, Theorem 4.6, the weak convergence of the shadow sequence. For this purpose, we replace (A, B, C) by $(\partial f, \partial g, \nabla h)$ where

$$f, g, h \in \Gamma_0(X) \quad \text{and} \quad \nabla h \text{ is, additionally, nonexpansive.} \quad (4.1)$$

Hence, $\partial f, \partial g$ and ∇h are maximally monotone by Fact 2.43 and 3*-monotone by Fact 2.44. Moreover, ∇h is firmly nonexpansive by Fact 2.31. Consequently, by Fact 2.14, we have that

$$f, g, \text{ and } h \text{ are weakly lower semicontinuous.} \quad (4.2)$$

Additionally, we assume that

$$P_{\overline{\text{ran}(\text{Id} - T)}}(0) =: v \in \text{ran}(\text{Id} - T) \text{ and } v_R = 0. \quad (4.3)$$

Then, it follows, from Proposition 3.16(vii), that

$$v = v_D := P_{\overline{\text{dom } f - \text{dom } g}}(0). \quad (4.4)$$

By (2.11) and (4.4), we have

$$(\forall (a, b) \in \overline{\text{dom } f} \times \overline{\text{dom } g}) \quad \langle 0 - v, (a - b) - v \rangle \leq 0. \quad (4.5)$$

Moreover, by using Example 2.59, we have

$$T = T_{(\partial f, \partial g, \nabla h)} = \text{Id} - P_f + P_g(R_f - \nabla h P_f). \quad (4.6)$$

Furthermore, by Example 2.59 and Proposition 3.2(i), we have

$$(\forall x \in X) \quad P_f T^n x - P_g(R_f - \nabla h P_f) T^n x = T^n x - T^{n+1} x \rightarrow v. \quad (4.7)$$

Examples 4.8 and 4.10 below illustrate why the assumptions in (4.3) are necessary.

Lemma 4.1. *Let $x \in X$ and $y \in \text{dom } f \cap (v + \text{dom } g)$. Then, the following hold:*

$$0 \geq \langle y - P_f x, v \rangle, \quad (4.8)$$

$$f(y) \geq f(P_f x) + \langle y - P_f x, v + P_{f^*} x \rangle, \quad (4.9)$$

$$\begin{aligned} g(y - v) &\geq g(P_g(R_f - \nabla h P_f)x) \\ &\quad + \langle y - P_g(R_f - \nabla h P_f)x - v, P_f x - \nabla h P_f x - P_g(R_f - \nabla h P_f)x - v \rangle \\ &\quad - \langle y - P_g(R_f - \nabla h P_f)x - v, v + P_{f^*} x \rangle, \end{aligned} \quad (4.10)$$

$$h(y) \geq h(P_f x) + \langle y - P_f x, \nabla h P_f x \rangle. \quad (4.11)$$

Proof. Let $(a, b) = (P_f x, y - v) \in \text{dom } \partial f \times \text{dom } g \subseteq \text{dom } f \times \text{dom } g$, where the latter inclusion follows from [13, Proposition 16.4(i)]. Then, replacing (a, b) by $(P_f x, y - v)$ in (4.5) yields (4.8) as follows:

$$\langle P_f x - (y - v) - v, 0 - v \rangle = \langle y - P_f x, v \rangle \leq 0. \quad (4.12)$$

Moreover, by (2.28), (4.8) and, in view of Example 2.59, (2.25), we prove (4.9) as follows:

$$f(y) \geq f(P_f x) + \langle y - P_f x, x - P_f x \rangle \quad (4.13)$$

$$\geq f(P_f x) + \langle y - P_f x, P_{f^*} x \rangle + \langle y - P_f x, v \rangle \quad (4.14)$$

$$\stackrel{(4.8)}{\geq} f(P_f x) + \langle y - P_f x, v + P_{f^*} x \rangle. \quad (4.15)$$

Similarly, by replacing (a, b) in (4.5) with $(y, P_g(R_f - \nabla h P_f)x)$, we have

$$\langle y - P_g(R_f - \nabla h P_f)x - v, v \rangle \leq 0. \quad (4.16)$$

Hence, by (2.28) and, in view of Example 2.59, (2.25), we have

$$\begin{aligned}
g(y - v) &\geq g(P_g(R_f - \nabla h P_f)x) \\
&\quad + \langle y - P_g(R_f - \nabla h P_f)x - v, (R_f - \nabla h P_f)x - P_g(R_f - \nabla h P_f)x \rangle \\
&= g(P_g(R_f - \nabla h P_f)x) + 2 \underbrace{\langle y - P_g(R_f - \nabla h P_f)x - v, v \rangle}_{\geq 0 \text{ by (4.16)}} \\
&\quad + \langle y - P_g(R_f - \nabla h P_f)x - v, P_f x - \nabla h P_f x - P_g(R_f - \nabla h P_f)x - v \rangle \\
&\quad - \langle y - P_g(R_f - \nabla h P_f)x - v, v + P_{f^*} x \rangle \tag{4.17} \\
&\geq g(P_g(R_f - \nabla h P_f)x) \\
&\quad + \langle y - P_g(R_f - \nabla h P_f)x - v, P_f x - \nabla h P_f x - P_g(R_f - \nabla h P_f)x - v \rangle \\
&\quad - \langle y - P_g(R_f - \nabla h P_f)x - v, v + P_{f^*} x \rangle, \tag{4.18}
\end{aligned}$$

where (4.17) follows from (2.24) together with Fact 2.54, however, in view of (2.25) and Example 2.59, which establishes 4.10. Eventually, (4.11) follows from Definition 2.18. $\square$

Proposition 4.2. *Suppose that $y \in \text{dom } f \cap (v + \text{dom } g)$ and $x \in X$. For all $n \in \mathbb{N}$, define*

$$\epsilon_n := \langle y - P_g(R_f - \nabla P_f) T^n x - v, P_f T^n x - P_g(R_f - \nabla P_f) T^n x - v \rangle, \tag{4.19}$$

$$\delta_n := \langle P_f T^n x - P_g(R_f - \nabla P_f) T^n x - v, P_f T^n x - (T^n x + n v) \rangle, \tag{4.20}$$

$$\gamma_n := \langle P_g(R_f - \nabla h P_f) T^n x - P_f T^n x + v, \nabla h P_f T^n x \rangle. \tag{4.21}$$

Then,

$$\lim_{n \in \mathbb{N}} \epsilon_n = 0, \quad \lim_{n \in \mathbb{N}} \delta_n = 0, \quad \text{and} \quad \lim_{n \in \mathbb{N}} \gamma_n = 0. \tag{4.22}$$

Moreover, the following holds for all $n \geq 1$:

$$\begin{aligned}
f(y) + g(y - v) + h(y) &\geq f(P_f T^n x) + g(P_g(R_f - \nabla h P_f) T^n x) + h(P_f T^n x) \\
&\quad + \epsilon_n + \delta_n + \gamma_n. \tag{4.23}
\end{aligned}$$

Proof. Let (A, B, C) be replaced by $(\partial f, \partial g, \nabla h)$ in Corollary 3.28(ii), Corollary 3.28(iii), and Corollary 3.28(iv). Then, in view of Example 2.59, we learn that

$$(P_f T^n x)_{n \in \mathbb{N}}, \quad (\nabla h P_f T^n x)_{n \in \mathbb{N}}, \quad \text{and} \quad (P_g(R_f - \nabla h P_f) T^n x)_{n \in \mathbb{N}} \text{ are bounded.} \tag{4.24}$$

Moreover, by Fact 2.52(ii), we have that

$$(T^n x + n v)_{n \in \mathbb{N}} \text{ is bounded.} \tag{4.25}$$

Hence, in view of (4.7), we prove (4.22) as follows:

$$\lim_{n \in \mathbb{N}} \epsilon_n = \lim_{n \in \mathbb{N}} \langle \underbrace{y - P_g(R_f - \nabla P_f) T^n x - v}_{\text{cf. (4.24)}}, \underbrace{P_f T^n x - P_g(R_f - \nabla P_f) T^n x - v}_{\rightarrow v \text{ by (4.7)}} \rangle = 0, \quad (4.26)$$

$$\lim_{n \in \mathbb{N}} \delta_n = \lim_{n \in \mathbb{N}} \langle \underbrace{P_f T^n x - P_g(R_f - \nabla P_f) T^n x - v}_{\rightarrow v \text{ by (4.7)}}, \underbrace{P_f T^n x}_{\text{cf. (4.24)}} - \underbrace{(T^n x + n v)}_{\text{cf. (4.25)}} \rangle = 0, \quad (4.27)$$

$$\lim_{n \in \mathbb{N}} \gamma_n = \lim_{n \in \mathbb{N}} \langle \underbrace{P_g(R_f - \nabla h P_f) T^n x - P_f T^n x + v}_{\rightarrow -v \text{ by (4.7)}}, \underbrace{\nabla h P_f T^n x}_{\text{cf. (4.24)}} \rangle = 0. \quad (4.28)$$

Let $n \geq 1$, summing (4.9), (4.10), and (4.11), with x replaced by $T^n x$, yields that

$$\begin{aligned} f(y) + g(y - v) + h(y) &\geq f(P_f T^n x) + g(P_g(R_f - \nabla h P_f) T^n x) + h(P_f T^n x) \\ &\quad + \langle y - P_f T^n x, v + P_{f^*} T^n x \rangle + \langle y - P_f T^n x, \nabla h P_f T^n x \rangle \\ &\quad + \langle y - P_g(R_f - \nabla h P_f) T^n x - v, \\ &\quad \quad P_f T^n x - \nabla h P_f T^n x - P_g(R_f - \nabla h P_f) T^n x - v \rangle \\ &\quad - \langle y - P_g(R_f - \nabla h P_f) T^n x - v, v + P_{f^*} T^n x \rangle \quad (4.29a) \\ &= f(P_f T^n x) + g(P_g(R_f - \nabla h P_f) T^n x) + h(P_f T^n x) \\ &\quad + \langle P_g(R_f - \nabla h P_f) T^n x - P_f T^n x + v, \nabla h P_f T^n x \rangle \\ &\quad + \langle y - P_g(R_f - \nabla P_f) T^n x - v, P_f T^n x - P_g(R_f - \nabla P_f) T^n x - v \rangle \\ &\quad + \langle P_g(R_f - \nabla h P_f) T^n x - P_f T^n x + v, v + P_{f^*} T^n x \rangle \quad (4.29b) \\ &= f(P_f T^n x) + g(P_g(R_f - \nabla h P_f) T^n x) + h(P_f T^n x) + \epsilon_n + \gamma_n \\ &\quad + \langle P_g(R_f - \nabla h P_f) T^n x - P_f T^n x + v, v + P_{f^*} T^n x \rangle \quad (4.29c) \\ &\stackrel{(2.25)}{=} f(P_f T^n x) + g(P_g(R_f - \nabla h P_f) T^n x) + h(P_f T^n x) + \epsilon_n + \gamma_n \\ &\quad + \langle P_g(R_f - \nabla h P_f) T^n x - P_f T^n x + v, \\ &\quad \quad v + T^n x + n v - P_f T^n x - n v \rangle \quad (4.29d) \\ &= f(P_f T^n x) + g(P_g(R_f - \nabla h P_f) T^n x) + h(P_f T^n x) + \epsilon_n + \gamma_n \\ &\quad + \langle P_f T^n x - P_g(R_f - \nabla h P_f) T^n x - v, P_f T^n x - (T^n x + n v) \rangle \\ &\quad + \langle P_g(R_f - \nabla h P_f) T^n x - P_f T^n x + v, (1 - n) v \rangle \quad (4.29e) \\ &= f(P_f T^n x) + g(P_g(R_f - \nabla h P_f) T^n x) + h(P_f T^n x) + \epsilon_n + \delta_n + \gamma_n \\ &\quad + \underbrace{(1 - n) \langle P_g(R_f - \nabla h P_f) T^n x - P_f T^n x + v, v - 0 \rangle}_{\leq 0} \quad (4.29f) \\ &\quad \quad \quad \geq 0 \text{ by (4.5)} \\ &\geq f(P_f T^n x) + g(P_g(R_f - \nabla h P_f) T^n x) + h(P_f T^n x) \\ &\quad + \epsilon_n + \delta_n + \gamma_n. \end{aligned}$$

□

Lemma 4.3. *Suppose that $Z = \{x \in X \mid 0 \in \partial f(x) + \partial g(x - v) + \nabla h(x) - v\} \neq \emptyset$. Then, the following hold:*

$$(i) \quad Z = \operatorname{argmin}_{x \in X} (-\langle x, v \rangle + f(x) + g(x - v) + h(x)).$$

$$(ii) \quad Z = \operatorname{argmin}_{x \in X} (f(x) + g(x - v) + h(x)).$$

Proof. Fact 2.21 yields that $\partial(g(\cdot - v) + h) = \partial g(\cdot - v) + \nabla h$. Moreover, applying Fact 2.22 to $(f, g(\cdot - v) + h - \langle v, \cdot \rangle)$ in view of Fact 2.16 proves (i) as follows:

$$\operatorname{argmin}_{x \in X} (f(x) + g(x - v) + h(x) - \langle v, x \rangle) = \operatorname{zer}(\partial(f + g(\cdot - v) + h - \langle v, \cdot \rangle)) \quad (4.30a)$$

$$= \operatorname{zer}(\partial f + \partial(g(\cdot - v) + h - \langle v, \cdot \rangle)) \quad (4.30b)$$

$$= \operatorname{zer}(\partial f + \partial g(\cdot - v) + \nabla h - v) \stackrel{(3.40)}{=} Z. \quad (4.30c)$$

Note that, in view of (4.3), by replacing (A, B, C) with $(\partial f, \partial g, \nabla h)$ in Proposition 3.31, we learn that

$$Z = \{x \in X \mid 0 \in \partial f(x) + \partial g(x - v) + \nabla h(x)\}. \quad (4.31)$$

Hence, reapplying Fact 2.22 to $(f, g(\cdot - v) + h)$ proves (ii) as follows:

$$\operatorname{argmin}_{x \in X} (f(x) + g(x - v) + h(x)) = \operatorname{zer}(\partial(f + g(\cdot - v) + h)) \quad (4.32a)$$

$$= \operatorname{zer}(\partial f + \partial(g(\cdot - v) + h)) \quad (4.32b)$$

$$= \operatorname{zer}(\partial f + \partial g(\cdot - v) + \nabla h) \stackrel{(4.31)}{=} Z. \quad \square$$

The strength of the aforementioned result will be elaborated in Proposition 4.4 to conclude that the shadow sequence is indeed a minimizing sequence.

Proposition 4.4. *Let $x \in X$ and set $\mu := \min_{x^* \in X} (f(x^*) + g(x^* - v) + h(x^*))$. Then, the following hold:*

$$(i) \quad (P_f T^n x)_{n \in \mathbb{N}} \text{ is bounded and its weak cluster points lie in } Z.$$

$$(ii) \quad (P_g(R_f - \nabla h P_f) T^n x)_{n \in \mathbb{N}} \text{ is bounded and its weak cluster points lie in } Z - v.$$

Suppose that $(P_f T^n x)_{n \in \mathbb{N}} \rightharpoonup \bar{z}$. Then,

$$(iii) \quad f(P_f T^n x) \rightarrow f(\bar{z}),$$

$$(iv) \quad g(P_g(R_f - \nabla h P_f) T^n x) \rightarrow g(\bar{z} - v),$$

$$(v) \quad h(P_f T^n x) \rightarrow h(\bar{z}).$$

$$(vi) \quad f(P_f T^n x) + g(P_g(R_f - \nabla h P_f) T^n x) + h(P_f T^n x) \rightarrow \mu.$$

Proof. For all $n \geq 1$, define $(u_n, w_n) = (P_f T^n x, P_g(R_f - \nabla h P_f) T^n x)$. From Corollary 3.28(ii) and Corollary 3.28(iii), applied with (A, B, C) replaced by $(\partial f, \partial g, \nabla h)$, we learn that $(u_n)_{n \in \mathbb{N}}$ and $(w_n)_{n \in \mathbb{N}}$ are bounded. Suppose that $\bar{z}$ is a weak cluster point of $(u_n)_{n \in \mathbb{N}}$. In view of (4.7), $\bar{z} - v$ is a weak cluster point for $(w_n)_{n \in \mathbb{N}}$. Let $y \in \text{dom } f \cap \text{dom } g(\cdot - v)$. By (4.2), we have

$$\infty > f(y) + g(y - v) + h(y) \tag{4.33a}$$

$$\stackrel{(4.23)}{\geq} \overline{\lim} (f(u_n) + g(w_n) + h(u_n)) \tag{4.33b}$$

$$\stackrel{(2.4)}{\geq} \underline{\lim} (f(u_n) + g(w_n) + h(u_n)) \tag{4.33c}$$

$$\geq \underline{\lim} f(u_n) + \underline{\lim} g(w_n) + \underline{\lim} h(u_n) \tag{4.33d}$$

$$\stackrel{(2.7)}{\geq} f(\bar{z}) + g(\bar{z} - v) + h(\bar{z}), \tag{4.33e}$$

which implies that $\bar{z} \in \text{dom } f \cap \text{dom } g(\cdot - v)$. Let u_{k_n} be such that $\lim_{n \in \mathbb{N}} f(u_{k_n}) = \overline{\lim} f(u_n)$, which yields the following:

$$f(y) + g(y - v) + h(y) \stackrel{(4.23)}{\geq} \lim_{n \in \mathbb{N}} f(u_{k_n}) + \overline{\lim} g(w_{k_n}) + \overline{\lim} h(u_{k_n}) \tag{4.34a}$$

$$= \overline{\lim} f(u_n) + \overline{\lim} g(w_{k_n}) + \overline{\lim} h(u_{k_n}) \tag{4.34b}$$

$$\stackrel{(2.4)}{\geq} \underline{\lim} f(u_n) + \overline{\lim} g(w_{k_n}) + \overline{\lim} h(u_{k_n}) \tag{4.34c}$$

$$\stackrel{(2.7)}{\geq} f(\bar{z}) + g(\bar{z} - v) + h(\bar{z}). \tag{4.34d}$$

Since y is arbitrary, we can choose $y = \bar{z}$ so that the inequalities in (4.34) reduce to equalities. Hence, we learn that $\overline{\lim} f(u_n) = \underline{\lim} f(u_n)$ which implies that $\lim f(u_{k_n})$ exists. Using the same argument for g with w_{k_n} and h with u_{k_n} yields also that $\lim_{n \in \mathbb{N}} g(w_{k_n})$ and

$\lim_{n \in \mathbb{N}} h(u_{k_n})$ exist. Hence, the following sequences converge: $(f(u_n))_{n \in \mathbb{N}}$, $(g(w_n))_{n \in \mathbb{N}}$, and $(h(u_n))_{n \in \mathbb{N}}$. In addition, by choosing $y = \bar{z}$ in (4.33), we have that

$$\infty > f(y) + g(y - v) + h(y) \geq \lim_{n \in \mathbb{N}} f(u_n) + \lim_{n \in \mathbb{N}} g(w_n) + \lim_{n \in \mathbb{N}} h(u_n) \quad (4.35a)$$

$$= f(\bar{z}) + g(\bar{z} - v) + h(\bar{z}) \quad (4.35b)$$

$$\geq \mu. \quad (4.35c)$$

Selecting y such that $f(y) + g(y - v) + h(y)$ is as close to μ as possible yields that

$$\lim_{n \in \mathbb{N}} f(u_n) + \lim_{n \in \mathbb{N}} g(w_n) + \lim_{n \in \mathbb{N}} h(u_n) = f(\bar{z}) + g(\bar{z} - v) + h(\bar{z}) = \mu, \quad (4.36)$$

which, when combined with Lemma 4.3(ii), proves (i) and, in view of (4.7), proves (ii). We now show that $\lim_{n \in \mathbb{N}} f(u_n) = \lim_{n \in \mathbb{N}} f(u_{k_n}) = f(\bar{z})$. By way of contradiction, suppose that $\lim_{n \in \mathbb{N}} f(u_n) = \lim_{n \in \mathbb{N}} f(u_{k_n}) > f(\bar{z})$. Then (4.35), in view of Definition 2.11, implies three possibilities:

1. $\lim_{n \in \mathbb{N}} g(w_{k_n}) < g(\bar{z} - v)$ and $\lim_{n \in \mathbb{N}} h(u_{k_n}) = h(\bar{z})$, which contradicts, by (2.7), the weak lower semicontinuity of g .
2. $\lim_{n \in \mathbb{N}} g(w_{k_n}) = g(\bar{z} - v)$ and $\lim_{n \in \mathbb{N}} h(u_{k_n}) < h(\bar{z})$, which contradicts, by (2.7), the weak lower semicontinuity of h .
3. $\lim_{n \in \mathbb{N}} g(w_{k_n}) < g(\bar{z} - v)$ and $\lim_{n \in \mathbb{N}} h(u_{k_n}) < h(\bar{z})$, which contradicts, by (2.7), the weak lower semicontinuity of g and h , respectively.

Hence,

$$\lim_{n \in \mathbb{N}} f(u_n) = f(\bar{z}), \quad \lim_{n \in \mathbb{N}} g(w_n) = g(\bar{z} - v), \quad \text{and} \quad \lim_{n \in \mathbb{N}} h(u_n) = h(\bar{z}), \quad (4.37)$$

which proves (iii), (iv), and (v). The claim in (vi) follows from (iii), (iv), and (v). $\square$

Corollary 4.5. *Let $x \in X$. Then,*

- (i) *The sequence $(P_f T^n x)_{n \in \mathbb{N}}$ has a unique cluster point, which is a minimizer of $f + g(\cdot - v) + h$.*
- (ii) *The sequence $(P_g(R_f - \nabla h P_f) T^n x)_{n \in \mathbb{N}}$ has a unique cluster point, which is a minimizer of $f(\cdot + v) + g + h(\cdot + v)$.*

Proof. Note that $\partial f, \partial g(\cdot - v)$, and ∇h are paramonotone by Fact 2.47. Consequently, applying Theorem 3.27(v) with (A, B, C) replaced by $(\partial f, \partial g, \nabla h)$, in view of Example 2.59, yields that the sequence $(P_f T^n x, -v + P_{f^*} T^n x + n v)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $Z \times K$. Let z_1 and z_2 be two cluster points of $(P_f T^n x)_{n \in \mathbb{N}}$. Then, $\{z_1, z_2\} \subseteq \operatorname{argmin}_{x \in X} (f(x) + g(x - v) + h(x)) = Z$. As such, $z_1 - z_2 \in Z - Z$. However, from Fact 2.51, $z_1 - z_2 \in (Z - Z)^\perp$. Altogether, $z_1 - z_2 \in (Z - Z) \cap (Z - Z)^\perp = \{0\}$. Hence, $z_1 = z_2$. The analogous conclusion for the sequence $(P_g(R_f - \nabla h P_f) T^n x)_{n \in \mathbb{N}}$ follows from (4.7). $\square$

Theorem 4.6. *Let f, g, h be in $\Gamma_0(X)$ and ∇h is, additionally, nonexpansive. Suppose that $0 \in \operatorname{dom} f^* + \operatorname{dom} g^* + \operatorname{dom} h^*$, $v \in \operatorname{ran}(\operatorname{Id} - T)$, and $Z \neq \emptyset$. Let $x \in X$ and define $\mu := \min_{x^* \in X} (f(x^*) + g(x^* - v) + h(x^*))$. Then, there exists $\bar{z} \in X$ such that the following hold:*

- (i) $P_f T^n x \rightharpoonup \bar{z}$, which is a minimizer of $f + g(\cdot - v) + h$.
- (ii) $P_g(R_f - \nabla h P_f) T^n x \rightharpoonup \bar{z} - v$, which is a minimizer of $f(\cdot + v) + g + h(\cdot + v)$.
- (iii) $f(P_f T^n x) \rightarrow f(\bar{z})$.
- (iv) $g(P_g(R_f - \nabla h P_f) T^n x) \rightarrow g(\bar{z} - v)$.
- (v) $h(P_f T^n x) \rightarrow h(\bar{z})$.
- (vi) $f(P_f T^n x) + g(P_g(R_f - \nabla h P_f) T^n x) + h(P_f T^n x) \rightarrow f(\bar{z}) + g(\bar{z} - v) + h(\bar{z}) = \mu$.

Proof. Note that $0 \in \operatorname{dom} f^* + \operatorname{dom} g^* + \operatorname{dom} h^* \subseteq \overline{\operatorname{dom} f^* + \operatorname{dom} g^* + \operatorname{dom} h^*} \stackrel{(2.10)}{=} \overline{\operatorname{ran} \partial f + \operatorname{ran} \partial g + \operatorname{ran} \partial h} = \overline{\operatorname{ran} \partial f + \operatorname{ran} \partial g + \operatorname{ran} \nabla h}$, which, by (3.5), implies that $v_R = 0$. Hence, in view of (4.3), (i)-(vi) follow from Proposition 4.4 and Corollary 4.5. $\square$

Example 4.7. Suppose that $X = \mathbb{R}^m$, where $m \geq 1$. Let $x_0 \in X$, $a \in \mathbb{R}_{++}^m$, $b \in \mathbb{R}_{--}$, $U = \{x \in X \mid \langle a, x \rangle = b\}$, $f = \iota_U$, $g = \iota_{\mathbb{R}_+^m}$, and $h = \langle w, \cdot \rangle$, where $w \in X$ is given. Then, for all $n \in \mathbb{N}$, the following hold:

- (i) $U \cap \mathbb{R}_+^m = \emptyset$.
- (ii) $A = \partial f = N_U$ and $J_A = J_{N_U} = P_U$.
- (iii) $B = \partial g = N_{\mathbb{R}_+^m}$ and $J_B = J_{N_{\mathbb{R}_+^m}} = P_{\mathbb{R}_+^m}$.
- (iv) $C = \nabla h \equiv w$.

- (v) $f^* = \iota_{U^\perp}$, $g^* = \iota_{\mathbb{R}_-^m}$, and $h^* = \iota_{\{w\}}$. Consequently, $\text{dom } f^* = U^\perp$, $\text{dom } g^* = \mathbb{R}_-^m$, and $\text{dom } h^* = \{w\}$.
- (vi) $\text{dom } f^* + \text{dom } g^* + \text{dom } h^* = \mathbb{R}^m$. Consequently, $v_R = 0$.
- (vii) $\text{ran}(\text{Id} - T) = U - \mathbb{R}_+^m$.
- (viii) $\text{ran}(\text{Id} - T)$ is polyhedral. Hence, $\text{ran}(\text{Id} - T)$ is closed and convex.
- (ix) $v \in \text{ran}(\text{Id} - T)$.
- (x) $Z = \underset{x^* \in X}{\text{argmin}} \left(\iota_U(x^*) + \iota_{\mathbb{R}_+^m}(x^* - v) + \langle w, x^* \rangle \right) \neq \emptyset$.
- (xi) $P_f T^n x_0 = P_U T^n x_0 = T^n x_0 - \frac{\langle T^n x_0, a \rangle - b}{\|a\|^2} a$.
- (xii) $P_g T^n x_0 = P_{\mathbb{R}_+^m} T^n x_0 = \max\{0, T^n x_0\}$.
- (xiii) $T = T_{(N_U, N_{\mathbb{R}_+^m}, \nabla h)} = \text{Id} - P_U + P_{\mathbb{R}_+^m}(2P_U - \text{Id} - w)$.
- (xiv) $P_U T^n x_0 \rightarrow \bar{z} \in Z$.

Proof.

- (i): This is immediate from the construction of $a \in \mathbb{R}_{++}^m$ and $b \in \mathbb{R}_{--}$.
- (ii) & (iii): These follow from [13, Example 16.13 and Example 23.4].
- (iv): Clear.
- (v): This follows from [13, Example 13.3(iii) and Example 13.3(ii)] and Definition 2.17.
- (vi): Note that $(U - U) \cap \mathbb{R}_+^m = \{0\}$. Then, by [13, Proposition 6.27], $\{0\} = (U - U) \cap \mathbb{R}_+^m = ((U - U)^\perp)^\ominus \cap (\mathbb{R}_-^m)^\ominus = ((U - U)^\perp + \mathbb{R}_-^m)^\ominus = (U^\perp + \mathbb{R}_-^m)^\ominus$. Hence, by [13, Corollary 6.34], $(U^\perp + \mathbb{R}_-^m)^{\ominus\ominus} = (U^\perp + \mathbb{R}_-^m) = \{0\}^\ominus = \mathbb{R}^m$.
- (vii): By Proposition 3.10 and [13, Example 6.43 and Example 6.40], $\text{ran}(\text{Id} - T) = (U - \mathbb{R}_+^m) \cap ((U - U)^\perp + \mathbb{R}_-^m + \{w\})$. Now, combine with (vi).
- (viii): Indeed, it follows from (vii) in view of [26, Corollary 19.3.2] that $\text{ran}(\text{Id} - T)$ is a polyhedral set. Consequently, it is closed and convex.
- (ix): This follows from (viii) and (3.4) in view of [6, Proposition 2.5].

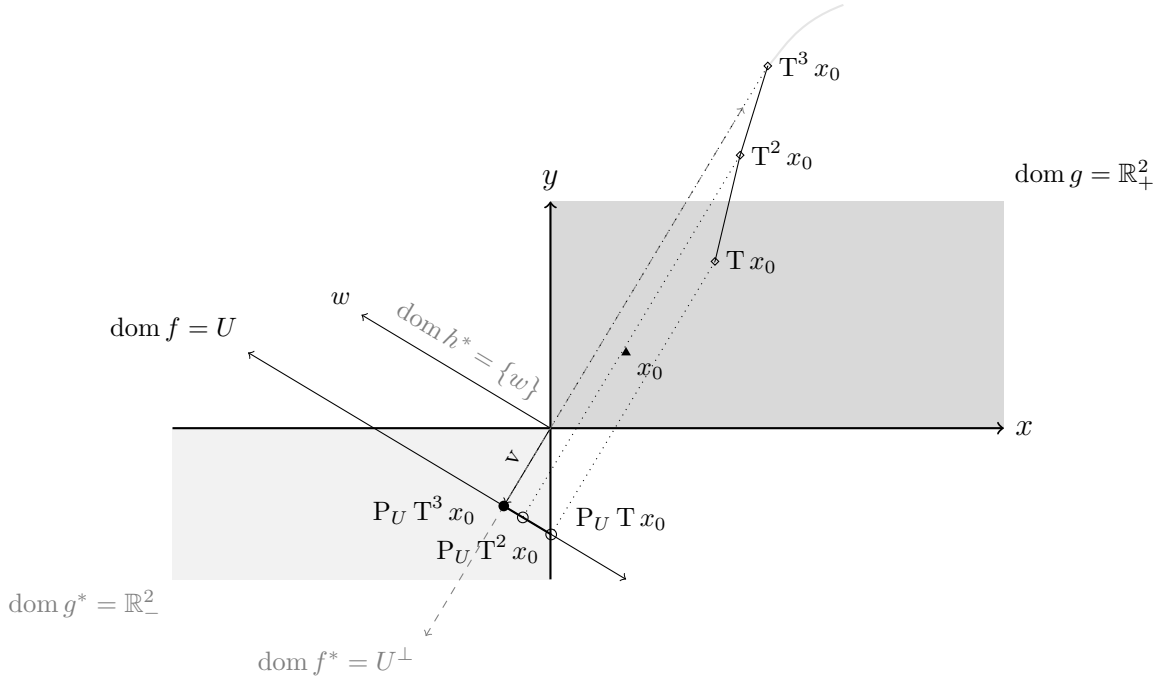


Figure 4.1: Two disjoint sets, $\text{dom } f = U$ and $\text{dom } g = \mathbb{R}^2_+$, for which the Davis–Yin iterates $T^n x$, starting at $x_0 = (1, 1)$, diverge and fail to minimize $h = \langle w, \cdot \rangle$, even though their projections, the shadow sequence, $P_U T^n x$ settle at the minimal displacement vector v .

(x): Combine Lemma 3.7 when applied with (A, B, C, w) replaced by $(\partial f, \partial g, \nabla h, v)$ and Lemma 4.3.

(xi): This is [27, Example 8.11].

(xii): This is [27, Example 8.9].

(xiii): This is (3.3) applied with (A, B, C) replaced by $(\partial f, \partial g, \nabla h)$ when combined with (ii) and (iii).

(xiv): Apply Theorem 4.6(i) and [13, Fact 2.33]. □

Concretely, suppose that $X = \mathbb{R}^2$ and let $x_0 = (1, 1)$, $a = (3, 5)$, $b = -7$, and $w = (-2.5, 1.5)$. Moreover, that $U = \{(x_1, x_2) \mid 3x_1 + 5x_2 = -7\}$, $f^*(y) = \iota_{U^\perp}$ where $U^\perp = \mathbb{R} \cdot (3, 5)$, $g^*(y) = \iota_{\mathbb{R}^2_-}(y)$, and $h^*(y) = \iota_{\{w\}}(y)$. These settings are depicted in Figure 4.1.

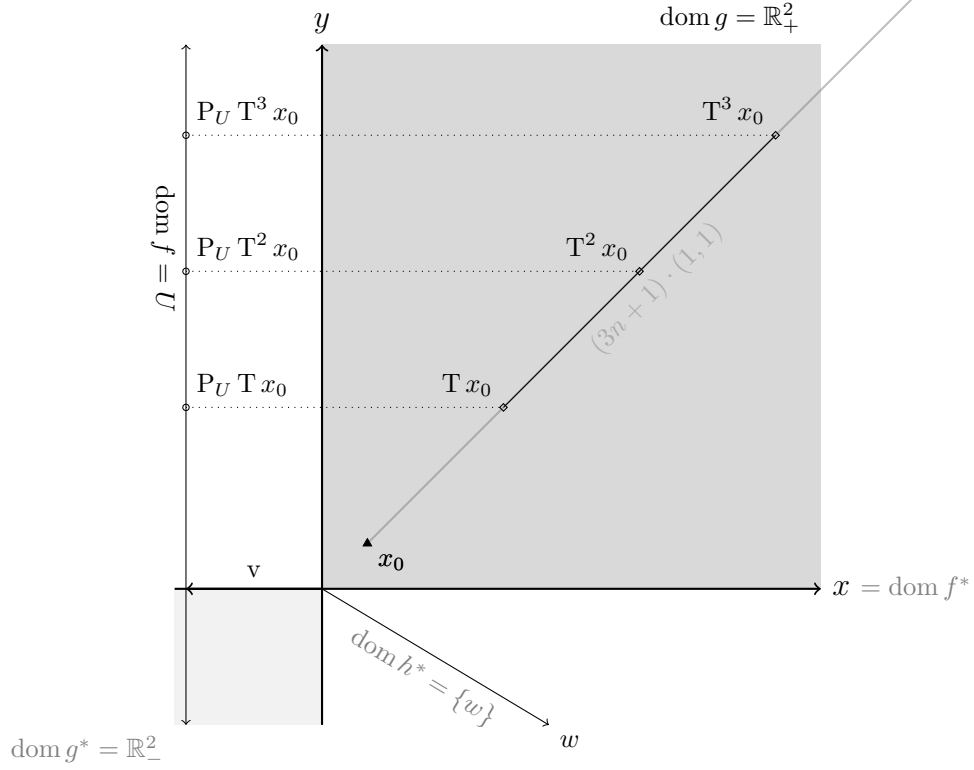


Figure 4.2: Two disjoint sets, $\text{dom } f = U$ (the translated y -axis) and $\text{dom } g = \mathbb{R}_+^2$, for which both the Davis-Yin iterates $T^n x$ starting at $x_0 = (1, 1)$ and its associated shadow sequence $P_U T^n x$ diverge and fail to minimize $h = \langle w, \cdot \rangle$ as a consequence of $v_R \neq 0$.

The following example elaborates the necessity of the constraint qualification $0 \in \text{dom } f^* + \text{dom } g^* + \text{dom } h^*$ for the convergence of the shadow sequence, $P_f T^n x$.

Example 4.8. Suppose that $X = \mathbb{R}^2$. Let $x_0 = (1, 1)$, $U = \mathbb{R} \cdot (0, 1) + (-3, 0)$, $f = \iota_U$, $g = \iota_{\mathbb{R}_+^2}$, and $h = \langle w, \cdot \rangle$, where $w = (5, -3)$. Then, for all $n \in \mathbb{N}$, the following hold:

- (i) $U \cap \mathbb{R}_+^2 = \emptyset$.
- (ii) $A = \partial f = N_U$ and $J_A = J_{N_U} = P_U$
- (iii) $B = \partial g = N_{\mathbb{R}_+^2}$ and $J_A = J_{N_{\mathbb{R}_+^2}} = P_{\mathbb{R}_+^2}$
- (iv) $C = \nabla h \equiv w$.

- (v) $f^* = \iota_{U^\perp}$, $g^* = \iota_{\mathbb{R}_-^2}$, and $h^* = \iota_{\{w\}}$. Consequently, $\text{dom } f^* = U^\perp$, $\text{dom } g^* = \mathbb{R}_-^2$, and $\text{dom } h^* = \{w\}$.
- (vi) $(0, 0) \notin \overline{\text{dom } f^* + \text{dom } g^* + \text{dom } h^*}$.
- (vii) $P_f T^n x_0 = P_U T^n x_0 = (-3, 0) + (0, 1) T^n x_0$.
- (viii) $P_g T^n x_0 = P_{\mathbb{R}_+^2} T^n x_0 = \max\{0, T^n x_0\}$.
- (ix) $T = T_{(N_U, N_{\mathbb{R}_+^2}, \nabla h)} = \text{Id} - P_U + P_{\mathbb{R}_+^2} (2P_U - \text{Id} - w)$.
- (x) $\|P_U T^n x_0\| \rightarrow +\infty$.

The above settings are depicted in Figure 4.2.

Proof.

- (i): This is immediate by construction.
- (ii) & (iii): These follow from [13, Example 16.13 and Example 23.4].
- (iv): Clear.
- (v): This follows from [13, Example 13.3(iii) and Example 13.3(ii)] and Definition 2.17.
- (vi): Indeed, note that $\text{dom } f^* + \text{dom } g^* + \text{dom } h^* \stackrel{(2.10)}{=} \text{ran } \partial f + \text{ran } \partial g + \text{ran } \partial h = \text{ran } \partial f + \text{ran } \partial g + \text{ran } \nabla h = \text{ran } N_{U^\perp} + \text{ran } N_{\mathbb{R}_+^2} + \{(5, -3)\} = \mathbb{R} \cdot (1, 0) + \mathbb{R}_-^2 + \{(5, -3)\} = \{(p + q + 5, q - 3) \mid p \in \mathbb{R}, q \in \mathbb{R}_-\}$. But, to obtain $(0, 0)$, q must lie in $\mathbb{R}_+$, in particular, $q = 3$.
- (vii): This is [13, Proposition 29.1(i)].
- (viii): This is [27, Example 8.9].
- (ix): Replace (A, B, C) by $(\partial f, \partial g, \nabla h)$ in (3.3) and combine with (ii) and (iii).
- (x): Combine (vi) and (3.5) to learn that $v_R \neq 0$. Then, by a reverse triangle inequality, $\|P_U T^n x\| \geq n\|v_R\| - \|P_U T^n x + n v_R\|$. Moreover, in this case, there is a closed-form expression of the iterates, that is, $T^n x_0 = (3n + 1) \cdot (1, 1)$ and $P_U T^n x_0 = (-3, 3n + 1)$,

from which one obtains the same conclusion. We proceed to proving so by induction. To this end, let $n = 1$. Then,

$$T x_0 = x_0 - P_U x_0 + P_{\mathbb{R}_+^2} (2 P_U - \text{Id} - w) x_0 \quad (4.38a)$$

$$= (1, 1) - P_U(1, 1) + P_{\mathbb{R}_+^2} (2 P_U(1, 1) - (1, 1) - (5, -3)) \quad (4.38b)$$

$$= (4, 0) + (0, 4) = (3(1) + 1) \cdot (1, 1). \quad (4.38c)$$

Hence, suppose that for some $n > 1$, it holds that $T^n x_0 = (3n + 1) \cdot (1, 1)$ and observe that $T^{n+1} x_0 = T T^n x_0 = T((3n + 1) \cdot (1, 1))$. Then,

$$\begin{aligned} T((3n + 1) \cdot (1, 1)) &= (3n + 1) \cdot (1, 1) - P_U((3n + 1) \cdot (1, 1)) \\ &\quad + P_{\mathbb{R}_+^2} (2 P_U((3n + 1) \cdot (1, 1)) - ((3n + 1) \cdot (1, 1)) - (5, -3)) \end{aligned} \quad (4.39a)$$

$$\begin{aligned} &= (3n + 1) \cdot (1, 1) - (-3, 3n + 1) \\ &\quad + P_{\mathbb{R}_+^2} ((-6, 2((3n + 1) \cdot (1, 1))) - ((3n + 1) \cdot (1, 1)) - (5, -3)) \end{aligned} \quad (4.39b)$$

$$= (3(n + 1) + 1, 0) + P_{\mathbb{R}_+^2} ((-6 - (3n + 1), 3n + 1) - (5, -3)) \quad (4.39c)$$

$$= (3(n + 1) + 1, 0) + P_{\mathbb{R}_+^2} ((-12 - 3n, 3(n + 1) + 1)) \quad (4.39d)$$

$$= (3(n + 1) + 1, 0) + (0, 3(n + 1) + 1) = (3(n + 1) + 1) \cdot (1, 1). \quad \square$$

We next recover the conclusion of [10, Theorem 6.1] for the Douglas–Rachford splitting $\tilde{T}$. This follows as a straightforward application of Theorem 4.6.

Corollary 4.9 ([10, Theorem 6.1]). *Let f and g be in $\Gamma_0(X)$. Suppose that $0 \in \text{dom } f^* + \text{dom } g^*$, $v \in \text{ran}(\text{Id} - \tilde{T})$, and $\text{zer}(\partial f + \partial g(\cdot - v)) \neq \emptyset$. Let $x \in X$ and define $\mu := \min_{x^* \in X} (f(x^* + v) + g(x^*))$. Then, there exists $\bar{z} \in X$ such that the following hold:*

(i) $P_f \tilde{T}^n x \rightharpoonup \bar{z}$, which is a minimizer of $f + g(\cdot - v)$.

(ii) $P_g R_f \tilde{T}^n x \rightharpoonup \bar{z} - v$, which is a minimizer of $f(\cdot + v) + g$.

(iii) $f(P_f \tilde{T}^n x) \rightarrow f(\bar{z})$.

(iv) $g(P_g R_f \tilde{T}^n x) \rightarrow g(\bar{z} - v)$.

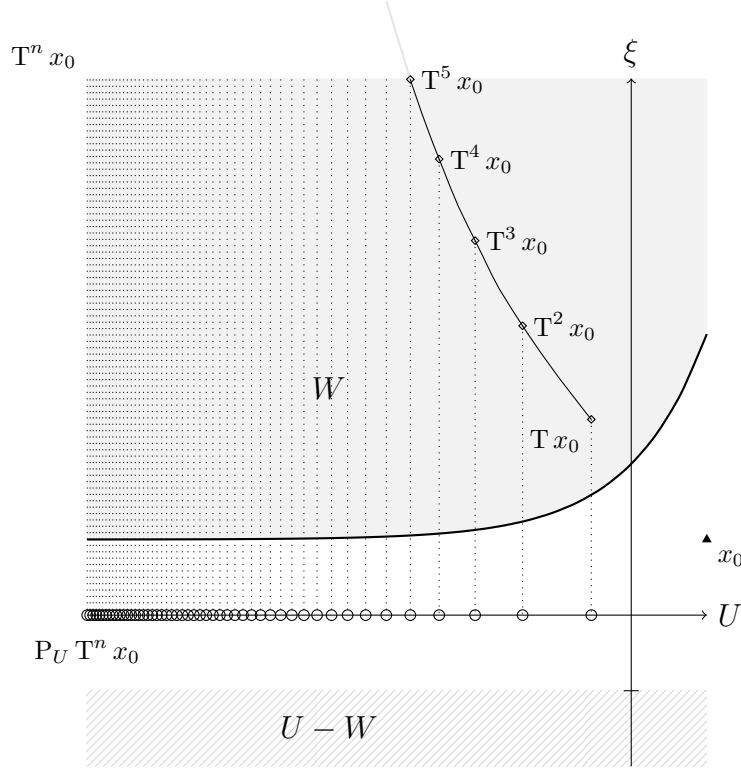


Figure 4.3: Two disjoint sets, $\text{dom } f = U$ (the x -axis) and $\text{dom } g = W$ (the epigraph of the translated exponential function) for which the iterates of Davis-Yin, $T^n x$, starting at $x_0 = (1, 1)$ and its associated shadow sequence, $P_U T^n x$, diverge due to $v \in \overline{\text{ran}}(\text{Id} - T)$ but $v \notin \text{ran}(\text{Id} - T)$.

$$(v) \quad f(P_f \tilde{T}^n) + g(P_g R_f \tilde{T}^n x) \rightarrow f(\bar{z}) + g(\bar{z} - v) = \mu.$$

Proof. Let $C = 0$ in (3.3). Then, the result immediately follows from Theorem 4.6. $\square$

The following example illustrates the behavior of the governing sequence, $T^n x$, and its associated shadow sequence, $P_f T^n x$, when $v \in \overline{\text{ran}}(\text{Id} - T) \setminus \text{ran}(\text{Id} - T)$.

Example 4.10. Suppose that $X = \mathbb{R}^2$. Let $x_0 = (1, 1)$, $(p, q) \in \mathbb{R} \times \mathbb{R}$, $U = \mathbb{R} \cdot (1, 0)$ and $W = \{(x, \xi) \mid e^x + 1 \leq \xi\}$ with $f = \iota_U$, $g = \iota_W$, and $h \equiv 0$. Then, for all $n \in \mathbb{N}$, the following hold:

- (i) $U \cap W = \emptyset$.

- (ii) $A = \partial f = N_U$ and $J_A = J_{N_U} = P_U$.
- (iii) $B = \partial g = N_W$ and $J_A = J_{N_W} = P_W$.
- (iv) $0 \in \overline{\text{dom } f^* + \text{dom } g^*}$.
- (v) $v \in \overline{\text{ran}}(\text{Id} - T) \setminus \text{ran}(\text{Id} - T)$.
- (vi) $P_f(p, q) = P_U(p, q) = (0, q)$.
- (vii) $P_g(p, q) = P_W(p, q) = (r, e^r + 1)$ if $(p, q) \notin W$, where r is the unique solution to $p = r + (e^r + 1 - q)e^r$.
- (viii) $T = T_{(\partial f, \partial g, 0)} = \tilde{T} = \text{Id} - P_U + P_W(2P_U - \text{Id})$.
- (ix) $\|P_U T^n x_0\| \rightarrow +\infty$.

These settings are depicted in Figure 4.3.

Proof.

- (i): This is immediate by construction.
- (ii) & (iii): These follow from [13, Example 16.13 and Example 23.4].
- (iv): Indeed, note that, by [28, Theorem 3.1], $\overline{\text{ran}}(\text{Id} - P_W) = (\text{rec } W)^\ominus$. Moreover, by Fact 2.57 and Fact 2.54, $\overline{\text{dom } f^* + \text{dom } g^*} \stackrel{(2.10)}{=} \overline{\text{ran } \partial f + \text{ran } \partial g} = \overline{\mathbb{R} \cdot (0, 1) + \overline{\text{ran } N_W}} = \overline{\mathbb{R} \cdot (0, 1) + \overline{\text{ran } J_{N_W}^{-1}}} = \overline{\mathbb{R} \cdot (0, 1) + \overline{\text{ran}(\text{Id} - J_{N_W})}} = \overline{\mathbb{R} \cdot (0, 1) + \overline{\text{ran}(\text{Id} - P_W)}} = \overline{\mathbb{R} \cdot (0, 1) + (\text{rec } W)^\ominus}$, from which the claim immediately follows.
- (v): By (iv) and Proposition 3.16(viii), $v = v_D$. Moreover, $U - W = \{(u - w, -\zeta) \mid u \in \mathbb{R}, w \in \mathbb{R}, \zeta \geq e^u + 1\} = \mathbb{R} \times (-\infty, -1)$, which is the *open* half-space below $\xi = -1$ (cf. Figure 4.3). Hence, $v \notin U - W$. Consequently, by Corollary 3.3, $v \notin \text{ran}(\text{Id} - T)$.
- (vi): This is immediate from Fact 2.24.
- (vii): This is [13, Proposition 23.35].
- (viii): Replace (A, B, C) by $(\partial f, \partial g, 0)$ in (3.3) and combine with (ii) and (iii).
- (ix): Note that $v \in U - W \Leftrightarrow 0 \in U - (v + W) \Leftrightarrow U \cap (v + W) \neq \emptyset$. Hence, by (v), $U \cap (v + W) = \emptyset$. Then, the claim follows as a consequence of [7, Theorem 3.3(iii)(a)]. $\square$

Chapter 5

Conclusion

This thesis studies the behavior of the Davis–Yin three-operator splitting method in the *inconsistent* setting. We show, under mild assumptions, that its “shadow” sequence converges weakly to a minimizer of the associated normal problem.

We further showed that the asymptotic regularity of the shadow sequence hinges precisely on the element of least norm in the closure of the range of the displacement operator, that is, when $v_R = 0$. This yields a necessary and sufficient condition. Moreover, we characterize the fixed-point sets of perturbed Davis–Yin operators and relate them to Eckstein’s extended solution set. Moreover, we offered an elegant reformulation of the normal problem solutions whenever asymptotic regularity holds.

Together, these results reinforce Davis–Yin splitting as a powerful tool for three-operator problems by extending its applicability to the inconsistent case.

References

- [1] Damek Davis and Wotao Yin. A Three-Operator Splitting Scheme and its Optimization Applications. *Set-Valued and Variational Analysis*, 25(4):829–858, 2017.
- [2] Goran Banjac, Paul Goulart, Bartolomeo Stellato, and Stephen Boyd. Infeasibility Detection in the Alternating Direction Method of Multipliers for Convex Optimization. *Journal of Optimization Theory and Applications*, 183(2):490–519, 2019.
- [3] Jisun Park and Ernest K. Ryu. Accelerated infeasibility detection of constrained optimization and fixed-point iterations. In *Proceedings of the 40th International Conference on Machine Learning*, ICML’23. JMLR.org, 2023.
- [4] David Applegate, Mateo Díaz, Haihao Lu, and Miles Lubin. Infeasibility Detection with Primal-Dual Hybrid Gradient for Large-Scale Linear Programming. *SIAM Journal on Optimization*, 34(1):459–484, 2024.
- [5] Heinz H. Bauschke and Walaa M. Moursi. On the order of the operators in the Douglas–Rachford algorithm. *Optimization Letters*, 10(3):447–455, 2016.
- [6] Ronald E. Bruck and Simeon Reich. Nonexpansive projections and resolvents of accretive operators in Banach spaces. 3(4):459–470, 1977.
- [7] Heinz H. Bauschke, Patrick L. Combettes, and D.Russell Luke. Finding best approximation pairs relative to two closed convex sets in Hilbert spaces. *Journal of Approximation Theory*, 127(2):178–192, 2004.
- [8] Heinz H. Bauschke, Warren L. Hare, and Walaa M. Moursi. Generalized Solutions for the Sum of Two Maximally Monotone Operators. *SIAM Journal on Control and Optimization*, 52(2):1034–1047, 2014.
- [9] Heinz H. Bauschke, Warren L. Hare, and Walaa M. Moursi. On the Range of the Douglas–Rachford Operator. *Math. Oper. Res.*, 41(3):884–897, 2016.

- [10] Heinz H. Bauschke and Walaa M. Moursi. On the Douglas–Rachford Algorithm for Solving Possibly Inconsistent Optimization Problems. *Mathematics of Operations Research*, 49(1):58–77, 2023.
- [11] Heinz H. Bauschke, Minh N. Dao, and Scott B. Lindstrom. The Douglas–Rachford algorithm for a hyperplane and a doubleton. *Journal of Global Optimization*, 74(1):79–93, 2019.
- [12] D. Russell Luke. Finding Best Approximation Pairs Relative to a Convex and Prox-Regular Set in a Hilbert Space. *SIAM Journal on Optimization*, 19(2):714–739, 2008.
- [13] Heinz H. Bauschke and Patrick L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics, Ouvrages de mathématiques de la SMC. Springer International Publishing, 2nd edition, 2017.
- [14] Jean-Bernard Baillon and Georges Haddad. Quelques propriétés des opérateurs angle-bornés et n -cycliquement monotones. *Israel Journal of Mathematics*, 26(2):137–150, 1977.
- [15] A. Pazy. Asymptotic behavior of contractions in Hilbert space. *Israel Journal of Mathematics*, 9(2):235–240, 1971.
- [16] Heinz H. Bauschke and Walaa M. Moursi. The Douglas–Rachford Algorithm for Two (Not Necessarily Intersecting) Affine Subspaces. *SIAM Journal on Optimization*, 26(2):968–985, 2016.
- [17] H. Bauschke, Minh Dao, and Walaa Moursi. On Fejér monotone sequences and non-expansive mappings. *Linear and Nonlinear Analysis*, 1:287–295, 2015.
- [18] George J. Minty. Monotone (nonlinear) operators in Hilbert space. *Duke Mathematical Journal*, 29(3):341 – 346, 1962.
- [19] Jim Douglas and H. H. Rachford. On the numerical solution of heat conduction problems in two and three space variables. *Transactions of the American Mathematical Society*, 82(2):421–439, 1956.
- [20] Heinz H. Bauschke, Xianfu Wang, and Liangjin Yao and. Rectangularity and para-monotonicity of maximally monotone operators. *Optimization*, 63(4):487–504, 2014.
- [21] Hedy Attouch and Michel Théra. A General Duality Principle for the Sum of Two Operators. *Journal of Convex Analysis*, 3:1–24, 1996.

- [22] B. Mercier. *Inéquations variationnelles de la mécanique*. Publications mathématiques d'Orsay. Université de Paris-Sud, Département de mathématique, 1980.
- [23] Francisco J. Aragón Artacho and David Torregrosa Belén. A Direct Proof of Convergence of Davis–Yin Splitting Algorithm Allowing Larger Stepsizes. *Set-Valued and Variational Analysis*, 30(3):1011–1029, 2022.
- [24] Heinz H. Bauschke, Radu I. Boţ, Warren L. Hare, and Walaa M. Moursi. At-touch–Théra duality revisited: Paramonotonicity and operator splitting. *Journal of Approximation Theory*, 164(8):1065–1084, 2012.
- [25] Jonathan Eckstein and B. F. Svaiter. General Projective Splitting Methods for Sums of Maximal Monotone Operators. *SIAM Journal on Control and Optimization*, 48(2):787–811, 2009.
- [26] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.
- [27] Heinz H. Bauschke and Walaa M. Moursi. *An Introduction to Convexity, Optimization, and Algorithms*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2023.
- [28] Eduardo H. Zarantonello. Projections on Convex Sets in Hilbert Space and Spectral Theory: Part I. Projections on Convex Sets: Part II. Spectral Theory. In Eduardo H. Zarantonello, editor, *Contributions to Nonlinear Functional Analysis*, pages 237–424. Academic Press, 1971.
- [29] Walaa M. Moursi. The range of the Douglas–Rachford operator in infinite-dimensional Hilbert spaces, 2022.
- [30] Heinz H. Bauschke and Walaa M. Moursi. On the Douglas–Rachford Algorithm for Solving Possibly Inconsistent Optimization Problems, 2024.
- [31] P. L. Lions and B. Mercier. Splitting Algorithms for the Sum of Two Nonlinear Operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979.
- [32] Gregory B Passty. Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *Journal of Mathematical Analysis and Applications*, 72(2):383–390, 1979.
- [33] Ernest K. Ryu, Yanli Liu, and Wotao Yin. Douglas–Rachford splitting and ADMM for pathological convex optimization. *Computational Optimization and Applications*, 74(3):747–778, 2019.