The Douglas–Rachford algorithm for two (not necessarily intersecting) affine subspaces

Heinz H. Bauschke* and Walaa M. Moursi†

November 9, 2015

Abstract

The Douglas–Rachford algorithm is a classical and very successful splitting method for finding the zeros of the sums of monotone operators. When the underlying operators are normal cone operators, the algorithm solves a convex feasibility problem. In this paper, we provide a detailed study of the Douglas–Rachford iterates and the corresponding shadow sequence when the sets are affine subspaces that do not necessarily intersect. We prove strong convergence of the shadows to the nearest generalized solution. Our results extend recent work from the consistent to the inconsistent case. Various examples are provided to illustrates the results.

2010 Mathematics Subject Classification: Primary 47H09, 49M27, 65K05, 65K10, Secondary 47H05, 47H14, 49M29, 49N15.

Keywords: Affine subspace, Attouch–Théra duality, Douglas–Rachford splitting operator, firmly nonexpansive mapping, fixed point, generalized solution, linear convergence, maximally monotone operator, normal cone operator, normal problem, projection operator.

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

†Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada, and Mansoura University, Faculty of Science, Mathematics Department, Mansoura 35516, Egypt. E-mail: `walaa.moursi@ubc.ca`.

1 Introduction

Throughout this paper

$$X \text{ is a real Hilbert space,}$$

with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$. A (possibly) set-valued operator $A : X \rightrightarrows X$ is *monotone* if any two pairs (x,u) and (y,v) in the graph of A satisfy $\langle x-y, u-v\rangle \geq 0$, and is *maximally monotone* if it is monotone and any proper enlargement of the graph of A (in terms of set inclusion) destroys the monotonicity of A. Monotone operators play an important role in modern optimization and nonlinear analysis; see, e.g., the books [5], [11], [12], [14], [30], [31], [33], [34], and [35].

Let $A : X \rightrightarrows X$ be maximally monotone and let $\operatorname{Id} : X \to X$ be the identity operator. The *resolvent* of A is $J_A := (\operatorname{Id} + A)^{-1}$ and the *reflected resolvent* is $R_A := 2J_A - \operatorname{Id}$. It is well-known that J_A is single-valued, maximally monotone and firmly nonexpansive.

The sum problem for two maximally monotone operators A and B is to find $x \in X$ such that $0 \in Ax + Bx$. When $(A+B)^{-1}(0) \neq \varnothing$ one approach to solve the problem is the Douglas–Rachford splitting technique. Recall that the Douglas–Rachford splitting operator [25] for the ordered pair of operators (A,B) is defined by

$$T_{(A,B)} := \tfrac{1}{2}(\operatorname{Id} + R_B R_A) = \operatorname{Id} - J_A + J_B R_A. \tag{1}$$

Let $x_0 \in X$. When $(A+B)^{-1}(0) \neq \varnothing$ the *"governing sequence"* $(T^n_{(A,B)}x_0)_{n\in\mathbb{N}}$ produced by the Douglas–Rachford operator converges weakly to a point in $\operatorname{Fix} T_{(A,B)}$ [1] (see [25]) and the *"shadow sequence"* $(J_A T^n_{(A,B)}x_0)_{n\in\mathbb{N}}$ converges weakly to a point in $(A+B)^{-1}0$. For further information on the Douglas–Rachford algorithm, which is a fundamental algorithm in optimization[2], we refer the reader to [20], [25], [32], and also [5].

When $A := N_U$ and $B := N_V$ [3], where U and V are nonempty closed convex subsets of X, the sum problem is equivalent to the convex feasibility problem: Find $x \in U \cap V$. In this case, using [5, Example 23.4],

$$T := T_{(N_U,N_V)} = \operatorname{Id} - P_U + P_V R_U, \tag{2}$$

where $R_U = 2P_U - \operatorname{Id}$. In the inconsistent case, when $U \cap V = \varnothing$, the governing sequence is known to satisfy $\|T^n x_0\| \to +\infty$ while the shadow sequence $(P_U T^n x_0)_{n\in\mathbb{N}}$ remains

[1] $\operatorname{Fix} T = \{x \in X \mid x = Tx\}$ is the set of fixed points of T.

[2] An AMS MathSciNet search for the seminal paper [25] by Lions and Mercier reveals nearly 200 citations.

[3] Throughout the paper we use N_C and P_C to denote the *normal cone* and *projector* associated with a nonempty closed convex subset C of X, respectively.

bounded with weak cluster points being the best approximation pairs[4] relative to U and V provided they exist (see [6]). Unlike the method of alternating projections, which employs the operator $P_V P_U$, the Douglas–Rachford method is not fully understood in the inconsistent case. In fact, it is a fundamental problem in optimization to understand the problem in the case when there is no solution and the corresponding behaviour of associated optimization algorithms — perhaps the most famous occurrence is the birth of the least-squares method for solving inconsistent linear systems due to Gauss. The inconsistent case may often result in considerable technical difficulties: For instance, Newton's method for finding zeros applied to a quadratic function without zeros may lead to chaotic behaviour! In this light, our work can be seen as a step toward understanding the behaviour of the Douglas–Rachford algorithm in the inconsistent case. While the general case remains wide open (especially because of Example 4.9 and Remark 4.10), we are able to make significant progress in the case when U and V are affine subspaces. The Douglas–Rachford operator is used to define the "*normal problem*" (see [9]) when the original problem is possibly inconsistent. In this case the set of best approximation solutions relative to U (which are also known as the *normal solutions*, see [9]) is $U \cap (\mathrm{v} + V)$, where[5] $\mathrm{v} = P_{\overline{\mathrm{ran}}(\mathrm{Id} - T)}0$. It is natural to ask what can we learn about the algorithm in the highlight of the new concept of the normal problem.

In this paper, we study the case when U and V are closed affine subspaces that do not necessarily intersect in detail. This case has applications e.g., in image restoration of a spatially limited image from partial knowledge of its Fourier transform (see, e.g., [16]). The problem reduces to a convex feasibility problem for two affine subspaces, and one approach to solve the problem is the Gerchberg-Papoulis algorithm (see, [21] and [27]), which reduces to the method of alternating projections in this case, to find a solution. Furthermore, the Douglas–Rachford method for two closed affine subspaces has recently shown to be useful for solving the nonconvex sparse affine feasibility problem (see [22] and [23]) and basis pursuit problem (see [17]). Other possible applications arise via the parallel splitting method (see, e.g., [5, Proposition 25.7]), which adapts the Douglas–Rachford method to handle a finite sum of monotone operators. When specializing the operators to be normal cones of affine subspaces, the problem reduces to convex feasibility problem (possibly inconsistent) of an affine subspace and the diagonal subspace in the product space which corresponds to considering possibly more than two affine subspaces in the original space.

Our results show that the shadow sequence will always converge strongly to a best approximation solution in $U \cap (\mathrm{v} + V)$ and therefore we generalize the main results in [3], which can only find a solution when the problem is consistent. This is remarkable because we do not have to have prior knowledge about the *gap vector* v; the shadow

---

[4] Suppose that U and V are nonempty closed convex subsets of X and let $(\bar{u}, \bar{v}) \in U \times V$. Then $(\bar{u}, \bar{v})$ is a *best approximation pair* relative to U and V if $\|\bar{u} - \bar{v}\| = \inf\{\|u - v\| \mid u \in U, v \in V\}$.

[5] Throughout this paper we use ran A to denote *the range* of the operator A.

sequence is simply $(P_U T^n x_0)_{n\in\mathbb{N}}$. Our proofs critically rely on the new results developed in Theorem 3.2 and Proposition 4.3, which serve as a bridge between the well-developed results in the consistent case in [3] and those of the normal problem studied in [9].

Let us summarize the main contributions of this paper:

R1 We compare the sequences $(({}_{-\mathrm{v}}T)^n x)_{n\in\mathbb{N}}$, $((T_{-\mathrm{v}})^n x)_{n\in\mathbb{N}}$ [6] and $(T^n x + n\,\mathrm{v})_{n\in\mathbb{N}}$ when T is an affine nonexpansive operator [7] and $\mathrm{v} := P_{\overline{\mathrm{ran}}(\mathrm{Id}-T)}0 \in \mathrm{ran}(\mathrm{Id}-T)$[8]. We prove that the three sequences coincide (see Theorem 3.2). Surprisingly, when we drop the assumption of T being affine, the sequences can be dramatically different (see Example 3.3).

R2 We prove the strong convergence of the shadow sequence $(P_U T^n x_0)_{n\in\mathbb{N}}$ when U and V are affine subspaces that do not have to intersect (see Theorem 4.4). We identify the limit to be the best approximation solution; moreover, the rate of convergence is linear when $U+V$ is closed.

R3 In view of **R2** it is tempting to conjecture that the shadow sequence $(J_A T^n x_0)_{n\in\mathbb{N}}$ in the inconsistent case (i.e., when $(A+B)^{-1}0 = \varnothing$) converges in a more general setting. We illustrate the somewhat surprising fact that if A and B are affine — but not normal cone — operators (see Example 4.8), the sequence $(J_A T^n x_0)_{n\in\mathbb{N}}$ can be unbounded. In fact, we can have $\|J_A T^n x_0\| \to +\infty$ even though the sum problem has *normal solutions*[9]. This illustrates that normal cone operators have additional structure that makes **R2** possible.

Organization

The remainder of this paper is organized as follows. Section 2 contains a collection of new results concerning nonexpansive and firmly nonexpansive operators whose fixed point sets could possibly be empty. Section 3 focuses on affine nonexpansive operators and their corresponding inner and outer "normal" shifts. Various examples that illustrate our theory are provided. Section 4 is devoted to present the main results. We prove strong convergence of the shadows of the Douglas–Rachford iterates of two (not necessarily intersecting) affine subspaces.

[6] Let $w \in X$. We define the *inner shift* and *outer shift* of an operator T by w at $x \in X$ by $T_w x := T(x-w)$ and ${}_wTx := -w + Tx$, respectively.

[7] Recall that T is *nonexpansive* if $(\forall x \in X)(\forall y \in X)$ $\|Tx - Ty\| \le \|x - y\|$.

[8] In highlight of Fact 2.2 the vector v is unique and well-defined.

[9] The normal solutions are the counterpart of the best approximation solutions in the context of the normal problem [9] when the operators are not normal cone operators (see Section 4 for details).

Notation

Let C be a nonempty closed convex subset of X. The *recession cone* of C is $\operatorname{rec} C := \{x \in X \mid x + C \subseteq C\}$, the *polar cone* of C is $C^{\ominus} := \{u \in X \mid \sup_{c\in C}\langle c, u\rangle \leq 0\}$, and the *dual cone* of C is $C^{\oplus} = -C^{\ominus}$. When C is an affine subspace the *linear space parallel to* C is $\operatorname{par} C = C - C$. Otherwise, the notation we utilize is standard and follows, e.g., [5] and [29].

2 Nonexpansive and firmly nonexpansive operators

In this section, we collect various results on (firmly) nonexpansive operators that will be useful later. Let $w \in X$. Recall that for a single-valued or set-valued operator T we define the *inner shift* and *outer shift* by w at $x \in X$ by

$$T_w x := T(x - w) \quad \text{and} \quad {}_wTx := -w + Tx, \tag{3}$$

respectively.

Lemma 2.1. *Let $T : X \to X$ and let $w \in X$. Then the following hold:*

(i) $\operatorname{Fix}(T_{-w}) = -w + \operatorname{Fix}(w + T) = -w + \operatorname{Fix}({}_{-w}T)$.

(ii) $w \in \operatorname{ran}(\operatorname{Id} - T) \iff \operatorname{Fix}(w + T) \neq \varnothing \iff \operatorname{Fix}(T_{-w}) \neq \varnothing$.

(iii) $(\forall x \in X)(\forall n \in \mathbb{N}) \quad (T_{-w})^n x = -w + (w + T)^n(x + w) = -w + ({}_{-w}T)^n(x + w)$.

Proof. (i): Let $x \in X$. Then $x \in \operatorname{Fix}(T_{-w}) \iff x = T(x + w) \iff x + w = w + T(x + w) \iff x + w \in \operatorname{Fix}(w + T) \iff x \in -w + \operatorname{Fix}(w + T)$.

(ii): $w \in \operatorname{ran}(\operatorname{Id} - T) \iff (\exists x \in X)$ such that $w = x - Tx \iff (\exists x \in X)$ such that $x = w + Tx \iff \operatorname{Fix}(w + T) \neq \varnothing$. Now combine with (i).

(iii): We proceed by induction. The conclusion is clear when $n = 0$. Now assume that for some $n \in \mathbb{N}$ it holds that $(T_{-w})^n x = -w + (w + T)^n(x + w)$. Then $(T_{-w})^{n+1}x = T((T_{-w})^n x + w) = T(-w + (w + T)^n(x + w) + w) = -w + w + T((w + T)^n(x + w)) = -w + (w + T)^{n+1}(x + w)$, as claimed. ■

We recall the following important fact.

Fact 2.2 (Infimal displacement vector). (See, e.g., [2],[13] and [26].) *Let $T : X \to X$ be nonexpansive. Then $\overline{\operatorname{ran}}(\operatorname{Id} - T)$ is convex; consequently, the infimal displacement vector*

$$\mathrm{v} := P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)} 0 \tag{4}$$

is the unique and well-defined element in $\overline{\mathrm{ran}}(\mathrm{Id} - T)$ *such that* $\|\mathrm{v}\| = \inf_{x \in X} \|x - Tx\|$.

Unless stated otherwise, throughout this paper we assume that

$$T \text{ is a nonexpansive operator on X,} \tag{5}$$

and that

$$\mathrm{v} := P_{\overline{\mathrm{ran}}(\mathrm{Id} - T)} 0 \in \mathrm{ran}\,(\mathrm{Id} - T). \tag{6}$$

Example 2.3. *Suppose that one of the following holds:*

(i) *X is finite-dimensional and $T : X \to X$ is an affine operator.*

(ii) *U and V are nonempty closed convex subsets of X, $A = N_U$, $B = N_V$, T is defined as in* (2)*, and one of the following conditions holds:*

  (a) *V is bounded.*

  (b) *U and V are polyhedral subsets*[10] *of X.*

Then $\mathrm{v} \in \mathrm{ran}\,(\mathrm{Id} - T)$*, i.e.,* (6) *holds.*

Proof. (i): $\mathrm{ran}(\mathrm{Id} - T)$ is a finite-dimensional affine subspace, hence closed. Therefore, $\mathrm{ran}(\mathrm{Id} - T) = \overline{\mathrm{ran}}(\mathrm{Id} - T)$ and (6) holds. (ii): It follows from the Browder-Göhde-Kirk fixed point theorem (see, e.g., [5, Theorem 4.19]) applied to the case (ii)(a), and [10, Theorem 5.6.1] applied to the case (ii)(b), that $\mathrm{Fix}\, P_V P_U \neq \varnothing$. By [9, Proposition 3.18 and Proposition 3.3], $\mathrm{v} \in \mathrm{ran}\,(\mathrm{Id} - T)$ as claimed. ■

In view of (6) and Lemma 2.1(ii) we have

$$\mathrm{Fix}(T_{-\mathrm{v}}) \neq \varnothing \quad \text{and} \quad \mathrm{Fix}(\mathrm{v} + T) \neq \varnothing. \tag{7}$$

We start with the following useful result.

Lemma 2.4. *Let C be a nonempty closed convex subset of X and let $c \in C$ satisfies that $\|c\| = \|P_C 0\|$. Then $c = P_C 0$.*

Proof. See Appendix A. ■

[10] A subset of X is polyhedral if it is a finite intersection of closed half-spaces.

Proposition 2.5. *Let $y_0 \in \operatorname{Fix}(\mathrm{v} + T)$. Then the following hold*[11]*:*

(i) $y_0 - \mathbb{R}_+ \mathrm{v} \subseteq \operatorname{Fix}(\mathrm{v} + T)$.

(ii) $\operatorname{Fix}(\mathrm{v} + T) - \mathbb{R}_+ \mathrm{v} = \operatorname{Fix}(\mathrm{v} + T)$.

(iii) $-\mathbb{R}_+ \mathrm{v} \subseteq \operatorname{rec}(\operatorname{Fix}(\mathrm{v} + T))$.

(iv) $(\forall n \in \mathbb{N})\ T^n y_0 = y_0 - n\,\mathrm{v}$.

(v) $]-\infty, 1] \cdot \mathrm{v} + \operatorname{Fix} T_{-\mathrm{v}} \subseteq \operatorname{Fix}(\mathrm{v} + T)$. *In particular it holds that* $\operatorname{Fix}(T_{-\mathrm{v}}) \subseteq \operatorname{Fix}(\mathrm{v} + T)$.

(vi) *For every $x \in X$, the sequence $(T^n x + n\,\mathrm{v})_{n \in \mathbb{N}}$ is Féjer monotone with respect to both* $\operatorname{Fix}(\mathrm{v} + T)$ *and* $\operatorname{Fix}(T_{-\mathrm{v}})$.

(vii) *Suppose that $x_0 \in \operatorname{Fix}(T_{-\mathrm{v}})$ and set $(\forall n \in \mathbb{N})\ x_n = T^n x_0$. Then $x_n = x_0 - n\,\mathrm{v}$ and $(x_n)_{n \in \mathbb{N}}$ lies in* $\operatorname{Fix}(T_{-\mathrm{v}})$.

Proof. (i): First we use induction to show that

$$(\forall n \in \mathbb{N}) \quad y_0 - n\,\mathrm{v} \in \operatorname{Fix}(\mathrm{v} + T). \tag{8}$$

Clearly when $n = 0$ the base case holds true. Now suppose that for some $n \in \mathbb{N}$ it holds that $y_0 - n\,\mathrm{v} \in \operatorname{Fix}(\mathrm{v} + T)$, i.e.,

$$y_0 - n\,\mathrm{v} = \mathrm{v} + T(y_0 - n\,\mathrm{v}). \tag{9}$$

Using (6) and (9) we have

$$\begin{aligned} \|\mathrm{v}\| &\le \|(\mathrm{Id} - T)(y_0 - (n+1)\,\mathrm{v})\| = \|y_0 - (n+1)\,\mathrm{v} - T(y_0 - (n+1)\,\mathrm{v})\| \\ &= \|y_0 - n\,\mathrm{v} - \mathrm{v} - T(y_0 - (n+1)\,\mathrm{v})\| = \|T(y_0 - n\,\mathrm{v}) - T(y_0 - (n+1)\,\mathrm{v})\| \le \|\mathrm{v}\|. \end{aligned}$$

Consequently all the inequalities above are equalities and we conclude that $\|\mathrm{v}\| = \|y_0 - (n+1)\,\mathrm{v} - T(y_0 - (n+1)\,\mathrm{v})\|$. It follows from (6) and Lemma 2.4 that

$$y_0 - (n+1)\,\mathrm{v} - T(y_0 - (n+1)\,\mathrm{v}) = \mathrm{v}\,. \tag{10}$$

That is, $y_0 - (n+1)\,\mathrm{v} = \mathrm{v} + T(y_0 - (n+1)\,\mathrm{v})$, which proves (8). Now using [5, Corollary 4.15] we learn that $\operatorname{Fix}(\mathrm{v} + T)$ is convex, which when combined with (8) yields (i).

(ii): On the one hand it follows from (i) that $\operatorname{Fix}(\mathrm{v} + T) - \mathbb{R}_+ \mathrm{v} \subseteq \operatorname{Fix}(\mathrm{v} + T)$. On the other hand $\operatorname{Fix}(\mathrm{v} + T) = \operatorname{Fix}(\mathrm{v} + T) - 0 \cdot \mathrm{v} \subseteq \operatorname{Fix}(\mathrm{v} + T) - \mathbb{R}_+ \mathrm{v}$.

[11] Let $x \in X$. Then $\mathbb{R}_+ x = \{rx \mid r \in [0, +\infty[\}$.

(iii): This follows directly from (ii).

(iv): We use induction. Clearly $y_0 - 0\,\mathrm{v} = y_0 = T^0 y_0$. Now suppose that for some $n \in \mathbb{N}$ it holds $T^n y_0 = y_0 - n\,\mathrm{v}$. Using (i) we have $T^{n+1} y_0 = T(y_0 - n\,\mathrm{v}) = -\mathrm{v} + y_0 - n\,\mathrm{v} = y_0 - (n+1)\,\mathrm{v}$.

(v): Using Lemma 2.1(i) and Proposition 2.5(ii) we have $\left]-\infty, 1\right] \cdot \mathrm{v} + \operatorname{Fix}(T_{-\mathrm{v}}) = \mathrm{v} - \mathbb{R}_+\,\mathrm{v} + \operatorname{Fix}(T_{-\mathrm{v}}) = \operatorname{Fix}(\mathrm{v} + T) - \mathbb{R}_+\,\mathrm{v} = \operatorname{Fix}(\mathrm{v} + T)$. In particular we have $\operatorname{Fix}(T_{-\mathrm{v}}) = 0 \cdot \mathrm{v} + \operatorname{Fix}(T_{-\mathrm{v}}) \subseteq \operatorname{Fix}(\mathrm{v} + T)$.

(vi): Let $x \in X$ and let $y \in \operatorname{Fix}(\mathrm{v} + T)$. Then using (iv) we have for every $n \in \mathbb{N}$,

$$\begin{aligned} \|T^{n+1}x + (n+1)\,\mathrm{v} - y\| &= \|T^{n+1}x - (y - (n+1)\,\mathrm{v})\| = \|T^{n+1}x - T^{n+1}y\| \\ &\leq \|T^n x - T^n y\| = \|T^n x - (y - n\,\mathrm{v})\| = \|T^n x + n\,\mathrm{v} - y\|. \end{aligned}$$

The statement for $\operatorname{Fix}(T_{-\mathrm{v}})$ follows from (v).

(vii): Combine (v) and (iv) to get that $x_n = x_0 - n\,\mathrm{v}$. Now by Lemma 2.1(i) $x_0 + \mathrm{v} \in \operatorname{Fix}(\mathrm{v} + T)$. Using (i) we have $(\forall n \in \mathbb{N})\ x_0 + \mathrm{v} - n\,\mathrm{v} \in \operatorname{Fix}(\mathrm{v} + T)$ or equivalently by Lemma 2.1(i) $x_0 - n\,\mathrm{v} \in -\mathrm{v} + \operatorname{Fix}(\mathrm{v} + T) = \operatorname{Fix}(T_{-\mathrm{v}})$. ■

The next example is readily verified.

Example 2.6. *Let C be a nonempty closed convex subset of X and suppose that $T = \operatorname{Id} - P_C$. Then T is firmly nonexpansive*[12] *and $\mathrm{v} = P_C 0$. Let $x \in X$. Then $x \in \operatorname{Fix}(\mathrm{v} + T) \iff P_C x = \mathrm{v}$, while $x \in \operatorname{Fix} T_{-\mathrm{v}} \iff P_C(x + \mathrm{v}) = \mathrm{v}$.*

Proposition 2.7. *Suppose that $X = \mathbb{R}$, and that $\operatorname{Fix} T = \varnothing$. Let $x \in \mathbb{R}$ and set $(\forall n \in \mathbb{N})\ y_n = T^n x + n\,\mathrm{v}$. Then the following hold:*

(i) *$(y_n)_{n \in \mathbb{N}}$ converges.*

(ii) *$\mathbb{R} \to \mathbb{R} : x \mapsto \lim_{n \to \infty} (T^n x + n\,\mathrm{v})$ is nonexpansive.*

(iii) *Suppose that T is firmly nonexpansive. Then $\mathbb{R} \to \mathbb{R} : x \mapsto \lim_{n \to \infty} (T^n x + n\,\mathrm{v})$ is firmly nonexpansive.*

Proof. (i): In view of Proposition 2.5(vi) the sequence $(y_n)_{n \in \mathbb{N}}$ is Féjer monotone with respect to $\operatorname{Fix}(\mathrm{v} + T)$. Now by Proposition 2.5(i) we know that $\operatorname{Fix}(\mathrm{v} + T)$ contains an unbounded interval. Since $X = \mathbb{R}$ we conclude that $\operatorname{int} \operatorname{Fix}(\mathrm{v} + T) \neq \varnothing$. It follows from [5, Proposition 5.10] that $(y_n)_{n \in \mathbb{N}}$ converges.

[12] Recall that $T\colon X \to X$ is *firmly nonexpansive* if $(\forall x \in X)(\forall y \in X)\quad \|Tx - Ty\|^2 + \|(\operatorname{Id} - T)x - (\operatorname{Id} - T)y\|^2 \leq \|x - y\|^2$.

(ii): Let $y \in \mathbb{R}$. Then

$$\begin{aligned}\Big| \lim_{n\to\infty} (T^n x + n\,\mathrm{v}) - \lim_{n\to\infty} (T^n y + n\,\mathrm{v})\Big| &= \Big| \lim_{n\to\infty} (T^n x + n\,\mathrm{v} - T^n y - n\,\mathrm{v})\Big| \\ &= \lim_{n\to\infty} |T^n x - T^n y| \le \lim_{n\to\infty} |x - y| = |x - y|.\end{aligned} \tag{11}$$

(iii): It follows from [5, Proposition 4.2(iv)] that an operator is firmly nonexpansive if and only if it is nonexpansive and monotone. Therefore, in view of (ii), we need to check monotonicity. Without loss of generality let $y \in \mathbb{R}$ such that $x \le y$. Since T is firmly nonexpansive, hence monotone, one can verify that $(\forall n \in \mathbb{N})\ T^n x \le T^n y$ and therefore $(\forall n \in \mathbb{N})\ T^n x + n\,\mathrm{v} \le T^n y + n\,\mathrm{v}$. Now take the limit as $n \to \infty$. ■

When $X = \mathbb{R}$, it follows from Proposition 2.7(i) that the sequence $(T^n x + n\,\mathrm{v})_{n\in\mathbb{N}}$ converges. In view of Proposition 2.5(vi) the sequence $(T^n x + n\,\mathrm{v})_{n\in\mathbb{N}}$ is Féjer monotone with respect to $\operatorname{Fix}(\mathrm{v} + T)$ which might suggest that the limit lies in $\operatorname{Fix}(\mathrm{v} + T)$. We show in the following example that this is not true in general.

Example 2.8. *Suppose that $X = \mathbb{R}$ and that*

$$T : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} x - \alpha, & \text{if } x \le \alpha; \\ 0, & \text{if } \alpha < x \le \beta; \\ x - \beta, & \text{if } x > \beta, \end{cases} \tag{12}$$

where $0 < \alpha < \beta$. Then T is firmly nonexpansive but not affine, $\mathrm{v} = \alpha$, $\operatorname{Fix}(\mathrm{v} + T) = \left]-\infty, \alpha\right]$, $\operatorname{Fix} T_{-\mathrm{v}} = \left]-\infty, 0\right]$, *and*

$$T^n + n\,\mathrm{v} : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} x, & \text{if } x \le \alpha; \\ \alpha, & \text{if } \alpha < x \le \beta; \\ x - n(\beta - \alpha), & \text{if } x > \beta \text{ and } n \le \lfloor x/\beta \rfloor; \\ \min\left\{\alpha, x - \left\lfloor \frac{x}{\beta} \right\rfloor \beta\right\} + \left\lfloor \frac{x}{\beta} \right\rfloor \alpha, & \text{if } x > \beta \text{ and } n > \lfloor x/\beta \rfloor. \end{cases} \tag{13}$$

Consequently,

$$\lim_{n\to\infty} (T^n + n\,\mathrm{v}) : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} x, & \text{if } x \le \alpha; \\ \alpha, & \text{if } \alpha < x \le \beta; \\ \min\left\{\alpha, x - \left\lfloor \frac{x}{\beta} \right\rfloor \beta\right\} + \alpha \left\lfloor \frac{x}{\beta} \right\rfloor, & \text{if } x > \beta. \end{cases} \tag{14}$$

Therefore for every $x_0 \in \mathbb{R}$ the sequence $(T^n x_0 + n\,\mathrm{v})_{n\in\mathbb{N}}$ is eventually constant. However, if the starting point x_0 lies in the interval $\left]\beta, \infty\right[$, then $\lim_{n\to\infty} T^n x_0 + n\,\mathrm{v} = \min\{\alpha, x - \lfloor \frac{x}{\beta} \rfloor \beta\} + \alpha \lfloor \frac{x}{\beta} \rfloor \notin \operatorname{Fix}(\mathrm{v} + T)$.

Proof. See Appendix B. ■

3 Affine nonexpansive operators

In this section, we investigate properties of *affine* nonexpansive operators. This additional assumption allows for stronger results than those obtained in the previous section. We recall the following fact.

Fact 3.1. (See [5, Proposition 3.17].) *Let S be a nonempty subset of X, and let $y \in X$. Then*

$$(\forall x \in X) \;\; P_{y+S}x = y + P_S(x-y). \tag{15}$$

Theorem 3.2. *Let $L\colon X \to X$ be linear and nonexpansive, let $b \in X$, and suppose that $T\colon X \to X\colon x \mapsto Lx + b$. Suppose also that $\mathrm{v} \in \mathrm{ran}(\mathrm{Id} - T)$, and let $x \in X$. Then the following hold:*

(i) $\mathrm{v} = P_{\mathrm{Fix}\, L}(-b) \in \mathrm{Fix}\, L = (\mathrm{ran}(\mathrm{Id} - L))^{\perp}$, *and* $\mathrm{v} \neq 0 \iff b \notin \overline{\mathrm{ran}}(\mathrm{Id} - L)$.

(ii) $(\forall n \in \mathbb{N})\; T^n x = L^n x + \sum_{k=0}^{n-1} L^k b$.

(iii) $(\forall n \in \mathbb{N})\; T^n x + n\,\mathrm{v} = L^n x + \sum_{k=0}^{n-1} L^k P_{\overline{\mathrm{ran}}(\mathrm{Id} - L)} b$.

(iv) $(\forall n \in \mathbb{N})\; (T_{-\mathrm{v}})^n x = T^n x + n\,\mathrm{v}$.

(v) $(\forall n \in \mathbb{N})\; (T_{-\mathrm{v}})^n x = (\mathrm{v} + T)^n x$.

(vi) $\mathrm{Fix}(T_{-\mathrm{v}}) = -\mathrm{v} + \mathrm{Fix}(T_{-\mathrm{v}}) = -\mathrm{v} + \mathrm{Fix}({}_{-\mathrm{v}}T) = -\mathrm{v} + \mathrm{Fix}(\mathrm{v} + T) = \mathrm{Fix}(\mathrm{v} + T)$.

(vii) $\mathrm{Fix}(T_{-\mathrm{v}}) = \mathrm{Fix}(\mathrm{v} + T) = \mathbb{R}\,\mathrm{v} + \mathrm{Fix}(\mathrm{v} + T) = \mathbb{R}\,\mathrm{v} + \mathrm{Fix}(T_{-\mathrm{v}})$. *Consequently* v *lies in the* lineality space[13] *of* $\mathrm{Fix}(T_{-\mathrm{v}}) = \mathrm{Fix}(\mathrm{v} + T)$.

Proof. (i): Note that $\mathrm{ran}(\mathrm{Id} - T) = \mathrm{ran}(\mathrm{Id} - L) - b$ and hence $\overline{\mathrm{ran}}(\mathrm{Id} - T) = \overline{\mathrm{ran}}(\mathrm{Id} - L) - b$. Therefore, using Fact 3.1 we have $\mathrm{v} = P_{\overline{\mathrm{ran}}(\mathrm{Id} - T)}0 = P_{-b+\overline{\mathrm{ran}}(\mathrm{Id} - L)}0 = -b + P_{\overline{\mathrm{ran}}(\mathrm{Id} - L)}(0 - (-b)) = -b + P_{\overline{\mathrm{ran}}(\mathrm{Id} - L)}b$. Using [5, Fact 2.18(iv)] and [7, Lemma 2.1], we learn that $\overline{\mathrm{ran}}(\mathrm{Id} - L)^{\perp} = \ker(\mathrm{Id} - L^*) = \mathrm{Fix}\, L^* = \mathrm{Fix}\, L$, and hence

$$\mathrm{v} = (\mathrm{Id} - P_{\overline{\mathrm{ran}}(\mathrm{Id} - L)})(-b) = P_{(\overline{\mathrm{ran}}(\mathrm{Id} - L))^{\perp}}(-b) = P_{\mathrm{Fix}\, L}(-b). \tag{16}$$

Note that $\mathrm{v} \neq 0 \iff b \notin \overline{\mathrm{ran}}(\mathrm{Id} - L)$.

(ii): We prove this by induction. When $n = 0$ the conclusion is obviously true. Now suppose that for some $n \in \mathbb{N}$ it holds that

$$T^n x = L^n x + \sum_{k=0}^{n-1} L^k b. \tag{17}$$

[13] For the definition and a detailed discussion of the lineality space, we refer the reader to [28, page 65].

Then $T^{n+1}x = T(T^n x) = T(L^n x + \sum_{k=0}^{n-1} L^k b) = L(L^n x + \sum_{k=0}^{n-1} L^k b) + b = L^{n+1}x + \sum_{k=0}^{n} L^k b$, as claimed.

(iii): Note that $b = P_{\overline{\operatorname{ran}}(\operatorname{Id} - L)} b + P_{\operatorname{Fix} L} b$. Using (i) and (ii) yields

$$\begin{aligned} T^n x + n\,\mathrm{v} &= L^n x + \sum_{k=0}^{n-1} (L^k b + \mathrm{v}) = L^n x + \sum_{k=0}^{n-1} \Big(L^k b + L^k\,\mathrm{v}\Big) \\ &= L^n x + \sum_{k=0}^{n-1} \Big(L^k b - L^k P_{\operatorname{Fix} L} b\Big) = L^n x + \sum_{k=0}^{n-1} L^k (\operatorname{Id} - P_{\operatorname{Fix} L}) b \\ &= L^n x + \sum_{k=0}^{n-1} L^k P_{\overline{\operatorname{ran}}(\operatorname{Id} - L)} b. \end{aligned}$$

(iv): We prove this by induction. Note that by (i) $\mathrm{v} \in \operatorname{Fix} L$, hence $L\mathrm{v} = \mathrm{v}$. When $n = 0$ we have $(T_{-\mathrm{v}})^0 x = x = T^0 x + 0 \cdot \mathrm{v}$. Now suppose that for some $n \in \mathbb{N}$ it holds that $(T_{-\mathrm{v}})^n x = T^n x + n\,\mathrm{v}$. Then $(T_{-\mathrm{v}})^{n+1} x = T_{-\mathrm{v}}(T^n x + n\,\mathrm{v}) = T(T^n x + n\,\mathrm{v} + \mathrm{v}) = L(T^n x) + L((n+1)\,\mathrm{v}) + b = T^{n+1}x + (n+1)\,\mathrm{v}$.

(v) We use induction again. The base case is obviously true. Now suppose that for some $n \in \mathbb{N}$ it holds that $(\mathrm{v} + T)^n x = T^n x + n\,\mathrm{v}$. Then $(\mathrm{v} + T)^{n+1} x = \mathrm{v} + T(\mathrm{v} + T)^n x = \mathrm{v} + T(T^n x + n\,\mathrm{v}) = \mathrm{v} + L(T^n x + n\,\mathrm{v}) + b = \mathrm{v} + LT^n x + n\,\mathrm{v} + b = LT^n x + b + (n+1)\,\mathrm{v} = T^{n+1}x + (n+1)\,\mathrm{v}$. Now combine with (iv).

(vi): Using (v) with $n = 1$ we have $T_{-\mathrm{v}} = \mathrm{v} + T$. Now apply Lemma 2.1(i).

(vii): Using (vi) and the assumption that T is an affine operator, we have $\operatorname{Fix}(T_{-\mathrm{v}}) = \operatorname{Fix}(\mathrm{v} + T)$ is an affine subspace. Now let $y_0 \in \operatorname{Fix}(T_{-\mathrm{v}}) = \operatorname{Fix}(\mathrm{v} + T)$. Using Proposition 2.5(i) we have $-\mathbb{R}_+\,\mathrm{v} \subseteq \operatorname{Fix}(\mathrm{v} + T) - y_0 = \operatorname{par}\operatorname{Fix}(\mathrm{v} + T)$ and therefore $\mathbb{R}\,\mathrm{v} \subseteq \operatorname{par}\operatorname{Fix}(\mathrm{v} + T)$. Hence $y_0 + \mathbb{R}\,\mathrm{v} \subseteq \operatorname{Fix}(\mathrm{v} + T)$ which yields $\operatorname{Fix}(\mathrm{v} + T) + \mathbb{R}\,\mathrm{v} \subseteq \operatorname{Fix}(\mathrm{v} + T)$. Since the opposite inclusion is obviously true we conclude that (vii) holds. ■

Suppose T is nonexpansive but not affine. Theorem 3.2 might suggest that, for every $x \in X$, the sequences $(T^n x + n\,\mathrm{v})_{n\in\mathbb{N}}$, $((T_{-\mathrm{v}})^n x)_{n\in\mathbb{N}}$ and $((\mathrm{v} + T)^n x)_{n\in\mathbb{N}}$ coincide, and consequently $(T^n x + n\,\mathrm{v})_{n\in\mathbb{N}}$ is a sequence of iterates of a nonexpansive operator. Interestingly, this is not the case as we illustrate now.

Example 3.3. *Suppose that $X = \mathbb{R}$ and let $\beta > 0$. Suppose that*

$$T\colon \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} x - \beta, & x \le \beta; \\ \alpha\,(x - \beta)\,, & x > \beta, \end{cases} \tag{18}$$

where $0 < \alpha < 1$. Then $\operatorname{Fix} T = \varnothing$, $\mathrm{v} = \beta$, for every $n \in \mathbb{N}$

$$(T_{-\mathrm{v}})^n : \mathbb{R} \to \mathbb{R} : x \mapsto \alpha^n \max\{x, 0\} + \min\{x, 0\}\,, \tag{19}$$

$$(20)\qquad (\mathrm{v}+T)^n : \mathbb{R} \to \mathbb{R} : x \mapsto \alpha^n \max\{x-\beta, 0\} + \min\{x, \beta\},$$

and

$$(21)\qquad T^n + n\,\mathrm{v} : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} x, & \text{if } x \le \beta; \\ \alpha^n x - \left(\frac{\alpha(1-\alpha^n)}{1-\alpha}\right)\beta + n\beta, & \text{if } x > \beta, n < q(x); \\ \alpha^{q(x)} x - \left(\frac{\alpha(1-\alpha^{q(x)})}{1-\alpha}\right)\beta + q(x)\beta, & \text{if } x > \beta, n \ge q(x), \end{cases}$$

where $q(x) : \mathbb{R} \to \mathbb{N} : x \mapsto \left\lceil \log_\alpha \frac{\beta}{\alpha\beta + (1-\alpha)x} \right\rceil$. *Consequently,*

$$(22)\qquad (\forall x \in \mathbb{R}) \quad \lim_{n\to\infty} (T_{-\mathrm{v}})^n x = \min\{x, 0\},$$

$$(23)\qquad (\forall x \in \mathbb{R}) \quad \lim_{n\to\infty} (\mathrm{v}+T)^n x = \min\{x, \beta\},$$

and

$$(24)\qquad (\forall x \in \mathbb{R}) \quad \lim_{n\to\infty} (T^n x + n\,\mathrm{v}) = \begin{cases} x, & \text{if } x \le \beta; \\ \alpha^{q(x)} x - \left(\frac{\alpha(1-\alpha^{q(x)})}{1-\alpha}\right)\beta + q(x)\beta, & \text{if } x > \beta. \end{cases}$$

Moreover, there is no operator $S : \mathbb{R} \to \mathbb{R}$ *such that for every* $x \in \mathbb{R}$ *and for every* $n \in \mathbb{N}$ *we have* $S^n x = T^n x + n\,\mathrm{v}$.

Proof. See Appendix C. ■

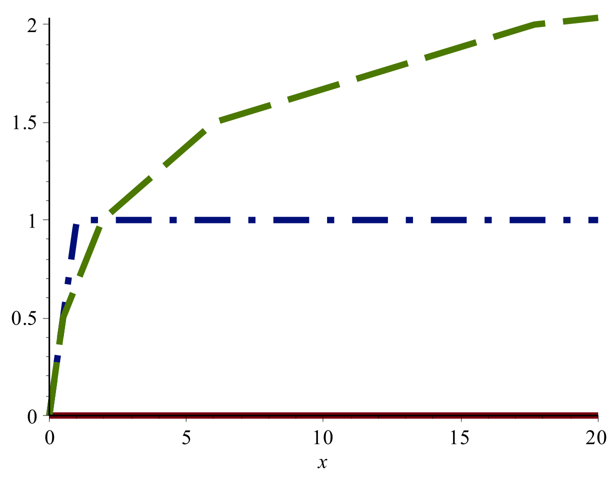

Figure 1: The solid curve represents $\lim_{n\to\infty}(T_{-\mathrm{v}})^n x$, the dashed dotted curve represents $\lim_{n\to\infty}(\mathrm{v}+T)^n x$, and the dashed curve represents $\lim_{n\to\infty} T^n x + n\,\mathrm{v}$, when $\alpha = 0.5$ and $\beta = 1$.

Figure 1 provides a plot of the functions defined by (22), (23) and (24) that illustrates that they are pairwise distinct.

4 The Douglas–Rachford operator for two affine subspaces

Unless otherwise stated we assume from now on that

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

The Attouch–Théra dual pair of (A, B) (see [1]) is the pair $(A, B)^* := (A^{-1}, B^{-\circledcirc})$, where

$$A^{\circledcirc} := (-\operatorname{Id}) \circ A \circ (-\operatorname{Id}) \quad \text{and} \quad A^{-\circledcirc} := (A^{-1})^{\circledcirc} = (A^{\circledcirc})^{-1}. \tag{25}$$

We shall use

$$Z := Z_{(A,B)} = (A + B)^{-1}(0) \qquad \text{and} \qquad K := K_{(A,B)} = (A^{-1} + B^{-\circledcirc})^{-1}(0), \tag{26}$$

to denote the primal and dual solutions respectively (see, e.g., [4]).

The *normal problem* associated with the ordered pair (A, B) (see [9]) is to find $x \in X$ such that

$$0 \in {}_{\mathrm{v}}Ax + B_{\mathrm{v}}x = Ax - \mathrm{v} + B(x - \mathrm{v}), \tag{27}$$

where

$$\mathrm{v} = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}, \tag{28}$$

and $T = T_{(A,B)}$ is defined by (1). We recall (see [15, Lemma 2.6(iii)] and [4, Corollary 4.9]) that

$$Z = J_A(\operatorname{Fix} T) \quad \text{and} \quad K = (\operatorname{Id} - J_A)(\operatorname{Fix} T), \tag{29}$$

and that T is self-dual (see [19, Lemma 3.6 on page 133] and [4, Corollary 4.3]), i.e.,

$$T_{(A,B)} = T_{(A,B)^*} = T_{(A^{-1},B^{-\circledcirc})}. \tag{30}$$

The *normal pair* associated with the ordered pair (A, B) is the pair $({}_{\mathrm{v}}A, B_{\mathrm{v}})$ and the *normal Douglas–Rachford operator* is $T_{({}_{\mathrm{v}}A,B_{\mathrm{v}})}$. Using [9, Proposition 2.24] we have

$$T_{({}_{\mathrm{v}}A,B_{\mathrm{v}})} = T_{-\mathrm{v}}. \tag{31}$$

The set of *normal solutions* is $Z_{\mathrm{v}} := Z_{({}_{\mathrm{v}}A,B_{\mathrm{v}})}$ and the set of *dual normal solutions* is $K_{\mathrm{v}} := K_{({}_{\mathrm{v}}A,B_{\mathrm{v}})}$.

Lemma 4.1. *The following hold:*

(i) $Z_{\mathrm{v}} = J_{\mathrm{v}A}(\operatorname{Fix}(T_{-\mathrm{v}})) = J_{(-\mathrm{v}+A)}(\operatorname{Fix}(T_{-\mathrm{v}})) = J_A(\operatorname{Fix}(T_{-\mathrm{v}}) + \mathrm{v}) = J_A(\operatorname{Fix}(\mathrm{v} + T))$.

(ii) $K_{\mathrm{v}} = (\operatorname{Id} - J_{\mathrm{v}A})(\operatorname{Fix}(T_{-\mathrm{v}})) = (\operatorname{Id} - J_{(-\mathrm{v}+A)})(\operatorname{Fix}(T_{-\mathrm{v}}))$.

(iii) $K_{\mathrm{v}} \neq \varnothing \iff Z_{\mathrm{v}} \neq \varnothing \iff \mathrm{v} \in \operatorname{ran}(\operatorname{Id} - T)$.

Proof. (i): Apply (29) to the normal pair $({}_{\mathrm{v}}A, B_{\mathrm{v}})$ and use (31) and (3). Now apply [5, Proposition 23.15(ii)]. The last equality follows from Lemma 2.1(i). (ii): Apply (29) to the normal pair $({}_{\mathrm{v}}A, B_{\mathrm{v}})$ then use (31) and [5, Proposition 23.15(iii)]. (iii): The first equivalence follows from applying [4, Proposition 2.4(v)] to the normal pair $({}_{\mathrm{v}}A, B_{\mathrm{v}})$. Now combine Lemma 2.1(ii) and (i). ■

In the following we assume that

$$\mathrm{v} = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)}0 \in \operatorname{ran}(\operatorname{Id} - T), \tag{32}$$

that

$$U \text{ and } V \text{ are nonempty closed convex subsets of } X \tag{33}$$

and that

$$A = N_U \text{ and } B = N_V. \tag{34}$$

Using [5, Example 23.4], (1) becomes

$$T_{U,V} := T_{(N_U, N_V)} = \operatorname{Id} - P_U + P_V R_U, \tag{35}$$

where $R_U = 2P_U - \operatorname{Id}$. In this case (see [9, Proposition 3.16])

$$\mathrm{v} = P_{\overline{U-V}}0, \tag{36}$$

or equivalently

$$-\mathrm{v} \in N_{\overline{U-V}}(\mathrm{v}). \tag{37}$$

The normal problem now is to find $x \in X$ such that

$$0 \in N_U x - \mathrm{v} + N_V(x - \mathrm{v}). \tag{38}$$

Lemma 4.2. *Let $w \in X$. Then the following hold:*

(i) $J_{-w+N_U} = J_{N_U}(\cdot + w) = P_U(\cdot + w)$.

(ii) $J_{N_U(\cdot - w)} = w + J_{N_U}(\cdot - w) = w + P_U(\cdot - w)$.

(iii) $N_V(\cdot - w) = N_{w+V}$.

(iv) *Suppose that U is an affine subspace and and that $w \in (\operatorname{par} U)^\perp$. Then $(\forall \alpha \in \mathbb{R})(\forall x \in X)$ $P_U(x + \alpha w) = P_U x$.*

Proof. (i) and(ii): See [5, Proposition 23.15(ii) and (iii) and Example 23.4]. (iii): One can easily verify that $(\forall w \in X)$ we have $\iota_V(\cdot - w) = \iota_{w+V}$. Therefore $N_V(\cdot - w) = \partial\iota_V(\cdot - w) = \partial\iota_{w+V} = N_{w+V}$. (iv): Let $a \in U$. Then $U = a + \operatorname{par} U$. Using Fact 3.1, we have $(\forall \alpha \in \mathbb{R})(\forall x \in X)$ $P_U(x + \alpha w) = P_{a+\operatorname{par} U}(x + \alpha w) = a + P_{\operatorname{par} U}(x + \alpha w - a) = a + P_{\operatorname{par} U}(x - a) + \alpha P_{\operatorname{par} U} w = a + P_{\operatorname{par} U}(x - a) = P_{a+\operatorname{par} U} x = P_U x$. ■

Proposition 4.3. *Suppose that U and V are closed affine subspaces of X and that $T = T_{U,V}$. Then the following hold:*

(i) *T is affine and $T = \operatorname{Id} - P_U - P_V + 2P_V P_U$.*

(ii) $\mathrm{v} \in (\operatorname{par} U)^\perp \cap (\operatorname{par} V)^\perp$.

(iii) $(\forall x \in X)$ $(\forall \alpha \in \mathbb{R})$ $P_U x = P_U(x + \alpha\,\mathrm{v})$.

(iv) $(\forall x \in X)$ $(\forall \alpha \in \mathbb{R})$ $P_V x = P_V(x + \alpha\,\mathrm{v})$.

(v) $T_{-\mathrm{v}} = \mathrm{v} + T = T_{N_U, N_V(\cdot - \mathrm{v})} = T_{U, \mathrm{v} + V}$.

(vi) $Z_{\mathrm{v}} = U \cap (\mathrm{v} + V)$.

(vii) $K_{\mathrm{v}} = (\operatorname{par} U)^\perp \cap (\operatorname{par} V)^\perp$.

(viii) $\operatorname{Fix}(T_{-\mathrm{v}}) = \operatorname{Fix}(\mathrm{v} + T) = Z_{\mathrm{v}} + K_{\mathrm{v}} = (U \cap (\mathrm{v} + V)) + ((\operatorname{par} U)^\perp \cap (\operatorname{par} V)^\perp)$.

Proof. (i): Note that $J_A = P_U$ and $J_B = P_V$ are affine (see e.g. [5, Corollary 3.20(i)]). Using (35) we have $T = \operatorname{Id} - P_U + P_V(2P_U - \operatorname{Id}) = \operatorname{Id} - P_U + 2P_V P_U - P_V$. Since the class of affine operators is closed under addition, subtraction and composition we deduce that T is affine.

(ii): It follows from [6, Proposition 2.7 & Remark 2.8(ii)] that $\mathrm{v} \in (\operatorname{rec} U)^\oplus \cap (\operatorname{rec} V)^\ominus = (\operatorname{par} U)^\perp \cap (\operatorname{par} V)^\perp$, where the last equality follows from [5, Proposition 6.22 and Proposition 6.23(v)].

(iii) and (iv): Combine (ii) with Lemma 4.2(iv).

(v) It follows from (iv) with α replaced by -1, and Lemma 4.2(ii) that $J_{N_V(\cdot - \mathrm{v})} = \mathrm{v} + P_V(\cdot - \mathrm{v}) = \mathrm{v} + P_V$. Consequently, using Theorem 3.2(v) and (1) we have $T_{-\mathrm{v}} = \mathrm{v} + T = \mathrm{Id} - P_U + \mathrm{v} + P_V R_U = T_{N_U, N_V(\cdot - \mathrm{v})}$. Finally, Lemma 4.2(iii) implies that $T_{N_U, N_V(\cdot - \mathrm{v})} = T_{N_U, N_{\mathrm{v}+V}} = T_{U, \mathrm{v}+V}$.

(vi): See [9, Proposition 3.16].

(vii): Let $z \in U \cap (\mathrm{v} + V) = Z_{\mathrm{v}}$ and note that, as subdifferential operators, N_U and N_V are paramonotone (see, e.g., [24]) and so are the translated operators $-\mathrm{v} + N_U$ and $N_V(\cdot - \mathrm{v})$. Therefore, in view of [4, Remark 5.4] and (ii) we have

$$\begin{aligned} K_{\mathrm{v}} &= (-\mathrm{v} + N_U z) \cap (-N_V(z - \mathrm{v})) = (-\mathrm{v} + (\mathrm{par}\, U)^{\perp}) \cap (\mathrm{par}\, V)^{\perp} \\ &= (\mathrm{par}\, U)^{\perp} \cap (\mathrm{par}\, V)^{\perp}. \end{aligned} \tag{39}$$

(viii): Since $-\mathrm{v} + N_U$ and $N_V(\cdot - \mathrm{v})$ are paramonotone, it follows from (v), (vii) and [4, Corollary 5.5] applied to the normal pair $({}_{\mathrm{v}}A, B_{\mathrm{v}})$ that $\mathrm{Fix}(T_{-\mathrm{v}}) = \mathrm{Fix}(\mathrm{v} + T) = Z_{\mathrm{v}} + K_{\mathrm{v}}$. Now combine with (vi) and (vii). ∎

We are now ready for our main result. It illustrates that, even in the inconsistent case, the "shadow sequence" $(P_U T^n x)_{n \in \mathbb{N}}$ behaves extremely well because it converges to a normal solution *without prior knowledge* of the infimal displacement vector. The proof of Theorem 4.4 relies on the work leading up to this point as well as the convergence analysis of the consistent case in [3].

Theorem 4.4 (Douglas–Rachford algorithm for two affine subspaces). *Let $x \in X$. Then $(\forall n \in \mathbb{N})$ we have*

$$P_U T^n x = P_U(T^n x + n\,\mathrm{v}) = P_U((T_{-\mathrm{v}})^n x) = P_U T^n_{U, \mathrm{v}+V} x = J_{-\mathrm{v} + N_U}((T_{-\mathrm{v}})^n x), \tag{40}$$

and

$$P_U T^n x \to P_{Z_{\mathrm{v}}} x = P_{U \cap (\mathrm{v}+V)} x. \tag{41}$$

Moreover, if $\mathrm{par}\, U + \mathrm{par}\, V$ *is closed (as is always the case when X is finite-dimensional) then the convergence is linear* [14] *with rate being the cosine of the Friedrichs angle*

$$c_F(\mathrm{par}\, U, \mathrm{par}\, V) := \sup_{\substack{u \in \mathrm{par}\, U \cap W^{\perp} \cap \mathrm{ball}(0;1) \\ v \in \mathrm{par}\, V \cap W^{\perp} \cap \mathrm{ball}(0;1)}} |\langle u, v \rangle| < 1, \tag{42}$$

where $W = \mathrm{par}\, U \cap \mathrm{par}\, V$ *and* $\mathrm{ball}(0;1)$ *is the closed unit ball.*

[14] Recall that $x_n \to x$ *linearly* with rate $\gamma \in \,]0,1[$ if $(\gamma^{-n}\|x_n - x\|)_{n \in \mathbb{N}}$ is bounded.

Proof. Let $n \in \mathbb{N}$. Using Proposition 4.3(iii) with (x, α) replaced by $(T^n x, n)$ we learn that $P_U T^n x = P_U(T^n x + n\,\mathrm{v})$. Now combine with Theorem 3.2(iv) to get the second identity. The third identity follows from applying Proposition 4.3(v). Finally note that using the first identity, Proposition 4.3(iii) with (x, α) replaced by $((T_{-\mathrm{v}})^n x, 1)$ and Lemma 4.2(i) we learn that $P_U T^n x = P_U((T_{-\mathrm{v}})^n x + \mathrm{v}) = J_{-\mathrm{v}+N_U}((T_{-\mathrm{v}})^n x)$. Now we prove (41). It follows from (32), Lemma 4.1(iii) and Proposition 4.3(vi) that $Z_{\mathrm{v}} = U \cap (\mathrm{v} + V) \neq \varnothing$. Now apply [3, Corollary 4.5]. ■

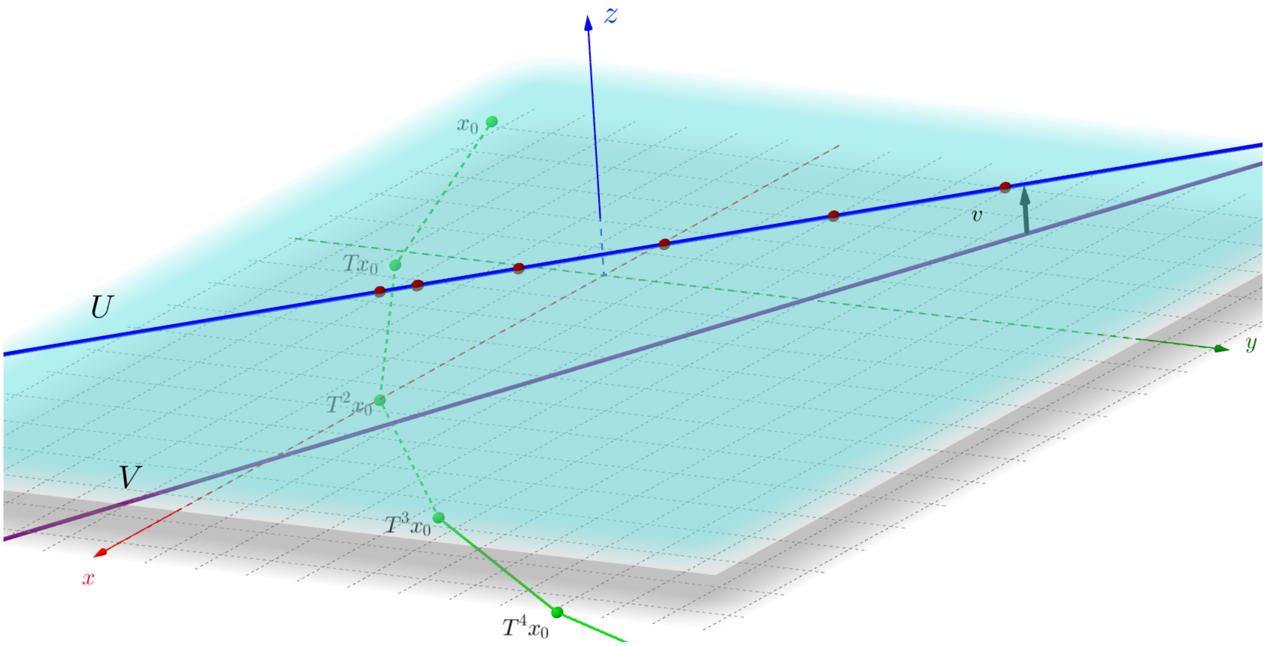

Figure 2: Two nonintersecting affine subspaces U (blue line) and V (purple line) in $\mathbb{R}^3$. Shown are also the first few iterates of $(T^n x_0)_{n\in\mathbb{N}}$ (green points) and $(P_U T^n x_0)_{n\in\mathbb{N}}$ (red points).

Figure 2 shows a Geogebra snapshot [18] of the Douglas–Rachford iterates and its shadows for two nonintersecting nonparallel lines U and V in $\mathbb{R}^3$.

Let us now comment on the comparison of the Douglas–Rachford algorithm to the method of alternating projections.

Remark 4.5. *Let $x \in X$ and $n \in \{1, 2, \ldots\}$. A straight-forward induction yields*

(43) $$(P_V P_U)^n x = -\mathrm{v} + (P_{\mathrm{v}+V} P_U)^n x$$

while Theorem 4.4 results in

(44) $$P_U T^n x = P_U T^n_{U,\mathrm{v}+V} x.$$

We conclude that executing the Douglas–Rachford algorithm or the Method of Alternating Projection (MAP), to solve the best approximation problem for U and V, is essentially the same (up to a shift with v in the case of MAP) as executing the two algorithms on the (consistent) normal problem. Consequently, we can use the numerical results established in [3, Sections 7 and 8] *to illustrate the efficiency of our algorithm. In passing, we recall that the numerical experiments presented in* [3, Section 7] *show that Douglas–Rachford algorithm reveal a faster rate of convergence than MAP when the angle between the subspaces is small.*

The following result is known (see e.g., [13, Corollary 1.5] and [2, Corollary 2.3]). We include a simple proof for completeness in Appendix D.

Proposition 4.6. *Suppose that $T : X \to X$ be firmly nonexpansive, and that $\mathrm{v} = P_{\overline{\mathrm{ran}}(\mathrm{Id}-T)}0 \in \mathrm{ran}(\mathrm{Id}-T)$. Then*

$$(\forall x \in X) \quad T^n x - T^{n+1}x \to \mathrm{v}\,. \tag{45}$$

Proposition 4.7 (When only one set is an affine subspace). *Suppose that U is an affine subspace of X, and that $T = T_{U,V}$. Then for every $x \in X$ the sequence $(P_U T^n x)_{n\in\mathbb{N}}$ is asymptotically regular, i.e., $P_U T^n x - P_U T^{n+1}x \to 0$.*

Proof. Using [6, Remark 2.8(ii)] we have $\mathrm{v} \in (\mathrm{par}\, U)^{\perp}$. It follows from Lemma 4.2(iv) applied with (x, α) replaced by $(T^{n+1}x, 1)$ and Proposition 4.6 that

$$\|P_U T^n x - P_U T^{n+1}x\| = \|P_U T^n x - P_U(T^{n+1}x + \mathrm{v})\| \le \|T^n x - T^{n+1}x - \mathrm{v}\| \to 0, \tag{46}$$

as claimed. ■

The following simple example shows that for two affine (but not normal cone) operators, the shadow sequence may fail to converge.

Example 4.8. *Suppose that $X = \mathbb{R}^2$ and let $S : \mathbb{R}^2 \to \mathbb{R}^2 : (x_1, x_2) \mapsto (-x_2, x_1)$, be the counter-clockwise rotator by $\pi/2$. Let $b \in \mathbb{R}^2 \smallsetminus \{(0,0)\}$. Suppose that $A := S$ and set $B := -S + b$. Then $\mathrm{zer}\, A \neq \varnothing$, $\mathrm{zer}\, B \neq \varnothing$ yet $\mathrm{zer}(A + B) = \varnothing$. Moreover, $\mathrm{v} = (\mathrm{Id} + S)(b)$, the set of normal solutions $Z_{\mathrm{v}} = \mathbb{R}^2$ and for every $x \in \mathbb{R}^2$ we have $\|J_A T^n x\| \to \infty$.*

Proof. Let $x \in \mathbb{R}^2$ and note that S and $-S$ are both linear, continuous, single-valued, monotone and $S^2 = (-S)^2 = -\mathrm{Id}$. It follows from [9, Proposition 2.10] that $J_A x = J_S x = \frac{1}{2}(\mathrm{Id} - S)x = \frac{1}{2}(x - Sx)$. Similarly using [5, Proposition 23.15(ii)] we can see that $J_B x = \frac{1}{2}(x - b + Sx - Sb)$. Therefore we have $R_A x = -Sx$ and $R_B x = -b + Sx - Sb$. Hence $R_B R_A x = S(-Sx) - Sb - b = -S^2 x - Sb - b = x - Sb - b = x - (\mathrm{Id} + S)b$. Consequently we have

$$(\forall x \in \mathbb{R}^2) \quad Tx = \tfrac{1}{2}(\mathrm{Id} + R_B R_A)x = x - \tfrac{1}{2}(\mathrm{Id} + S)b. \tag{47}$$

It follows from (47) that $\operatorname{ran}(\operatorname{Id} - T) = \{\frac{1}{2}(\operatorname{Id} + S)b\}$, hence $\mathrm{v} = \frac{1}{2}(\operatorname{Id} + S)b$ and $Tx = x - \mathrm{v}$. Therefore, using Theorem 3.2(v) $\operatorname{Fix}(\mathrm{v} + T) = \operatorname{Fix}(T_{-\mathrm{v}}) = \mathbb{R}^2$. Moreover, using (29) and (31) applied to the normal pair $({}_{\mathrm{v}}A, B_{\mathrm{v}})$ we learn that $Z_{\mathrm{v}} = J_{{}_{\mathrm{v}}A}(\operatorname{Fix}(T_{-\mathrm{v}})) = \mathbb{R}^2$. In view of Proposition 2.5(i) we have $T^n x = x - n\,\mathrm{v}$. Hence, using that J_A is linear, we get $J_A T^n x = J_A(x - n\,\mathrm{v}) = J_A x - nJ_A\,\mathrm{v}$. Now $J_A\,\mathrm{v} = \frac{1}{2}(\operatorname{Id} - S)(\frac{1}{2}(\operatorname{Id} + S)b) = \frac{1}{2}b \neq (0,0)$, which completes the proof. ■

With some more work, we can construct an example of the same type where the involved operators are even subdifferential operators:

Example 4.9 (The shadows may fail to converge to a normal solution). *Suppose that U and V are affine subspaces of X such that $U \cap V = \varnothing$, and set $\widetilde{A} := N_U^{-1}$ and $\widetilde{B} := N_V^{-\circledcirc}$. Then $Z = K = \varnothing$, $Z_{\mathrm{v}} = \mathrm{v} + \big((\operatorname{par} U)^{\perp} \cap (\operatorname{par} V)^{\perp}\big)$, $K_{\mathrm{v}} = -\mathrm{v} + (U \cap (\mathrm{v} + V))$, $\|J_{\widetilde{A}} T^n x\| \to \infty$, and $J_{\widetilde{A}^{-1}} T^n x = P_U T^n x \to P_{U \cap (\mathrm{v}+V)} x$. Consequently, even though $Z_{\mathrm{v}} \neq \varnothing$, the sequence of primal shadows has no convergent subsequence.*

Proof. Note that $\widetilde{A}^{-1} = N_U$, $\widetilde{B}^{-\circledcirc} = N_V$, therefore using (30) we have $T_{(\widetilde{A},\widetilde{B})} = T_{(\widetilde{A}^{-1},\widetilde{B}^{-\circledcirc})} = T_{U,V}$. Clearly, $K = (N_U + N_V)^{-1}(0) = U \cap V = \varnothing$, hence by (29) and (30), $Z = \varnothing$ and $\operatorname{Fix} T_{(\widetilde{A},\widetilde{B})} = \varnothing$. Using [9, Remark 3.5] and Proposition 4.3(vii)&(vi), we conclude that $Z_{\mathrm{v}} = \mathrm{v} + \big((\operatorname{par} U)^{\perp} \cap (\operatorname{par} V)^{\perp}\big)$ and $K_{\mathrm{v}} = -\mathrm{v} + (U \cap (\mathrm{v} + V))$. Let $x \in X$. It follows from Theorem 4.4 with A and B replaced with $\widetilde{A}^{-1}$ and $\widetilde{B}^{-\circledcirc}$ that $J_{\widetilde{A}^{-1}} T^n x = J_{N_U} T^n x = P_U T^n x \to P_{U \cap (\mathrm{v}+V)} x$. By [26, Corollary 6(a)], $\|T^n x\| \to \infty$. Moreover the inverse resolvent identity (see, e.g., [29, Lemma 12.14]) implies $J_{\widetilde{A}} T^n x = (\operatorname{Id} - P_U) T^n x = T^n x - P_U T^n x$. Finally, $\|J_{\widetilde{A}} T^n x\| = \|T^n x - P_U T^n x\| \geq \|T^n x\| - \|P_U T^n x\| \to \infty$. ■

Remark 4.10. *Example 4.8 and Example 4.9 give rise to the following open problem: What conditions imposed on the operators A and B guarantee that the shadow sequence of the Douglas–Rachford algorithm converges to a normal solution provided there is one?*

Acknowledgments

We thank two anonymous referees for their careful reading and pertinent comments. HHB was partially supported by the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chair Program.

References

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

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

[3] H.H. Bauschke, J.Y. Bello Cruz, T.T.A. Nghia, H.M. Phan and X. Wang: The rate of linear convergence of the Douglas-Rachford algorithm for subspaces is the cosine of the Friedrichs angle, *Journal of Approximation Theory* 185 (2014), 63–79.

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

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

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

[7] H.H. Bauschke, F. Deutsch, H. Hundal and S.-H. Park: Accelerating the convergence of the method of alternating projections, *Transactions of the American Mathematical Society* 355 (2003), 3433–3461.

[8] H.H. Bauschke, W.L. Hare and W.M. Moursi, On the range of the Douglas–Rachford operator, *Mathematics of Operations Research*, in press.

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

[10] H.H. Bauschke, J.M. Borwein, and A.S. Lewis, The method of cyclic projections for closed convex sets in Hilbert space, in *Recent Developments in Optimization Theory and Nonlinear Analysis (Jerusalem 1995), Contemporary Mathematics 204*, 1–38, 1997.

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

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

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

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

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

[16] P.L. Combettes, The convex feasibility problem in image recovery, *Advances in Imaging and Electron Physics* 25 (1995), 155–270.

[17] L. Demanet and X. Zhang, Eventual linear convergence of the Douglas–Rachford iteration for basis pursuit, *to appear in Mathematics of Computation, AMS.*

[18] Geogebra, `http://www.geogebra.org`

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

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

[21] R.W. Gerchberg, Super-resolution through error energy reduction, *Optica Acta* 21 (1974), 709–720.

[22] R. Hesse and D.R. Luke, Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems, *SIAM Journal on Optimization* 23 (2013), 2397–2419.

[23] R. Hesse, D.R. Luke, and P. Neumann, Alternating projections and Douglas-Rachford for sparse affine feasibility, *IEEE Transactions on Signal Processing* 62 (2014). No. 18.

[24] A.N. Iusem, On some properties of paramonotone operators, *Journal of Convex Analysis* 5 (1998), 269–278.

[25] P.L. Lions and B. Mercier, Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis* 16(6) (1979), 964–979.

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

[27] A. Papoulis, A new algorithm in spectral analysis and band-limited extrapolation, *IEEE Transactions on Circuits and Systems* 22 (1975), 735–742.

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

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

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

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

[32] B.F. Svaiter, On weak convergence of the Douglas–Rachford method, *SIAM Journal on Control and Optimization* 49 (2011), 280–287.

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

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

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

Appendix A

Proof of Lemma 2.4. Using that P_C is firmly nonexpansive we have $\|c - P_C 0\|^2 = \|P_C c - P_C 0\|^2 \leq \|c - 0\|^2 - \|(\operatorname{Id} - P_C)c - (\operatorname{Id} - P_C)0\|^2 = \|c\|^2 - \|c - P_C c + P_C 0\|^2 = \|c\|^2 - \|P_C 0\|^2 = 0$. ■

Appendix B

Proof of Example 2.8. Clearly

$$\operatorname{Id} - T = P_{[\alpha,\beta]} : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} \alpha, & \text{if } x \leq \alpha; \\ x, & \text{if } \alpha < x \leq \beta; \\ \beta, & \text{if } x > \beta. \end{cases} \tag{48}$$

Therefore, $\operatorname{ran}(\operatorname{Id} - T) = [\alpha, \beta]$, and consequently $\mathrm{v} = \alpha$. Moreover

$$(\forall x \in \mathbb{R}) \quad x \geq Tx + \alpha \geq T^2 x + 2\alpha \geq \cdots \geq T^n x + n\alpha \geq \cdots . \tag{49}$$

It is clear from Example 2.6 that

$$\operatorname{Fix}(\mathrm{v} + T) = \left]-\infty, \alpha\right]. \tag{50}$$

The statement for $\operatorname{Fix} T_{-\mathrm{v}}$ then follows from combining (50) and Lemma 2.1(i). The convergence of the sequence follows from Example 2.6 or Proposition 2.7(i). Now we prove (13). We claim that

$$T^n : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} x - n\alpha, & \text{if } x \leq \alpha; \\ (1-n)\alpha, & \text{if } \alpha < x \leq \beta; \\ x - n\beta, & \text{if } x > \beta \text{ and } n \leq \lfloor x/\beta \rfloor; \\ \min\left\{\alpha, x - \left\lfloor \frac{x}{\beta} \right\rfloor \beta\right\} + \left(\left\lfloor \frac{x}{\beta} \right\rfloor - n\right)\alpha, & \text{if } x > \beta \text{ and } n > \lfloor x/\beta \rfloor. \end{cases} \tag{51}$$

Using induction it is easy to verify the cases when $x \le \alpha$ and when $\alpha < x \le \beta$. Now we focus on the case when $x > \beta$. Set

$$K := \lfloor x/\beta \rfloor \ \text{ and } \ r := x - K\beta, \tag{52}$$

and note that $x = K\beta + r$, $K \in \{1, 2, 3, \ldots\}$ and $0 \le r < \beta$. In view of (49), if $n \in \{0, 1, 2, 3, \ldots, K\}$ we get $T^n x = x - n\beta = (K-n)\beta + r$. In particular,

$$T^K x = x - \lfloor x/\beta \rfloor \beta = r. \tag{53}$$

If $n > K$ we examine two cases. Case 1: $0 \le r \le \alpha$. It follows from (53) and (12) that $(\forall n \ge K)$ $T^n x = r + (K-n)\alpha$. Case 2: $\alpha < r < \beta$. Note that $T^{K+1}x = 0$, therefore using (53) and (12) we have $(\forall n > K)$ $T^n x = (K+1-n)\alpha = \alpha + (K-n)\alpha$, which proves (51). Now (13) follows from (51) because $\mathrm{v} = \alpha$. Letting $n \to \infty$ in (13) yields (14). Note that $\min\{\alpha, x - \lfloor \frac{x}{\beta} \rfloor \beta\} \ge 0$ and $\lfloor \frac{x}{\beta} \rfloor \ge 1$. By considering cases ($K = 1$ and $K \ge 1$), (14) implies that $\lim_{n\to\infty}(T^n x_0 + n\,\mathrm{v}) = \min\{\alpha, x - \lfloor \frac{x}{\beta} \rfloor \beta\} + \lfloor \frac{x}{\beta} \rfloor \alpha > \alpha \notin \left]-\infty, \alpha\right] = \operatorname{Fix}(\mathrm{v} + T)$.

■

Appendix C

Proof of Example 3.3. Considering cases, we easily check that

$$\operatorname{Id} - T \colon \mathbb{R} \to \mathbb{R} : x \mapsto (1-\alpha)\max\{x, \beta\} + \alpha\beta \ge \beta > 0. \tag{54}$$

Hence $\operatorname{Fix} T = \varnothing$ and $\mathrm{v} = \beta$ as claimed. Moreover, using (54) one can verify that

$$(\forall x \in X) \quad x \ge Tx + \beta > \cdots \ge T^n x + n\beta \ge T^{n+1}x + (n+1)\beta \ge \cdots. \tag{55}$$

We also verify that

$$(\forall x \in \mathbb{R}) \quad T_{-\mathrm{v}} : \mathbb{R} \to \mathbb{R} : x \mapsto \max\{x, 0\}\alpha + \min\{x, 0\}. \tag{56}$$

We now prove (19) by induction. Let $x \in \mathbb{R}$.

Clearly when $n = 0$ the base case holds true. Now suppose that for some $n \in \mathbb{N}$ (19) holds. If $x \le 0$ then $(T_{-\mathrm{v}})^n x = x \le 0$, and therefore, (56) implies that $(T_{-\mathrm{v}})^{n+1} x = T_{-\mathrm{v}}((T_{-\mathrm{v}})^n x) = T_{-\mathrm{v}} x = x$. Similarly we have $x > 0 \Rightarrow \alpha^n x = (T_{-\mathrm{v}})^n x > 0$, and consequently (56) implies that $(T_{-\mathrm{v}})^{n+1} x = T_{-\mathrm{v}}((T_{-\mathrm{v}})^n x) = T_{-\mathrm{v}}(\alpha^n x) = \alpha^{n+1}x$. The proof of (20) follows from combining (19) and Lemma 2.1(iii). Now we turn to (21). We consider two cases.

Case 1: $x \le \beta$. It is obvious using the definition of T that $(\forall n \in \mathbb{N})\ T^n x = x - n\beta$.
Case 2: $x > \beta$. Let $n \in \mathbb{N}$ be such that $T^n x > \beta$. By (55) and (18) we have

$$\begin{aligned} T^{n+1}x &= \alpha^{n+1}x - (\alpha^{n+1} + \alpha^n + \cdots + \alpha)\beta = \alpha^{n+1}x - \frac{\alpha(1-\alpha^{n+1})}{1-\alpha}\beta \\ &= \alpha^{n+1}\left(\frac{(1-\alpha)x + \alpha\beta}{1-\alpha}\right) - \frac{\alpha}{1-\alpha}\beta. \end{aligned} \tag{57}$$

In view of (55) there exists a unique integer, say, $q(x) \in \{1, 2, \ldots\}$ that satisfies $T^{q(x)-1}x > \beta$ and $T^{q(x)}x \le \beta$. Since $0 < \alpha < 1$, using (57) we have

$$\begin{aligned} T^{q(x)}x \le \beta \iff& \alpha^{q(x)}\left(\frac{(1-\alpha)x+\alpha\beta}{1-\alpha}\right) - \frac{\alpha}{1-\alpha}\beta \le \beta \\ \iff& \alpha^{q(x)}\left(\frac{(1-\alpha)x+\alpha\beta}{1-\alpha}\right) \le \frac{\beta}{1-\alpha} \iff \alpha^{q(x)}\left((1-\alpha)x+\alpha\beta\right) \le \beta \\ \iff& \alpha^{q(x)} \le \frac{\beta}{(1-\alpha)x+\alpha\beta} \iff q(x) \ge \log_\alpha \frac{\beta}{\alpha\beta+(1-\alpha)x}. \end{aligned}$$

Consequently, $q(x) = \left\lceil \log_\alpha \frac{\beta}{\alpha\beta+(1-\alpha)x} \right\rceil$. At this point, since $T^{q(x)}x \le \beta$, we must have $(\forall n \ge q(x))\ T^n x = T^{q(x)}x - (n - q(x))\beta$, which proves (21). The formulae (22), (23) and (24) are direct consequences of (19), (20) and (21), respectively. To prove the last claim note that if $S : \mathbb{R} \to \mathbb{R}$ is such that for every $n \in \mathbb{N}$ we have $S^n = T^n + n\,\mathrm{v}$, then setting $n = 1$ must yield

$$S = \mathrm{v} + T \colon \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} x, & x \le \beta; \\ \alpha(x-\beta)+\beta, & x > \beta. \end{cases} \tag{58}$$

Now compare (20) and (21). ■

Appendix D

Proof of Proposition 4.6. Let $y_0 \in \operatorname{Fix}(\mathrm{v} + T)$ and note that Proposition 2.5(iv) implies that $(\forall n \in \mathbb{N})\ (\mathrm{Id} - T)T^n y_0 = \mathrm{v}$. Since T is firmly nonexpansive, it follows from Proposition 2.5(vi) and [5, Proposition 5.4(ii)] that

$$\begin{aligned} \|T^n x - T^{n+1}x - \mathrm{v}\|^2 &= \|(\mathrm{Id}-T)T^n x - (\mathrm{Id}-T)T^n y_0\|^2 \\ &\le \|T^n x - T^n y_0\|^2 - \|T^{n+1}x - T^{n+1}y_0\|^2 \\ &= \|T^n x + n\,\mathrm{v} - y_0\|^2 - \|T^{n+1}x + (n+1)\,\mathrm{v} - y_0\|^2 \to 0. \end{aligned}$$

■