The rate of linear convergence of the Douglas–Rachford algorithm for subspaces is the cosine of the Friedrichs angle

Heinz H. Bauschke,* J.Y. Bello Cruz,† Tran T.A. Nghia,‡ Hung M. Phan§ and Xianfu Wang¶

April 23, 2014

Abstract

The Douglas–Rachford splitting algorithm is a classical optimization method that has found many applications. When specialized to two normal cone operators, it yields an algorithm for finding a point in the intersection of two convex sets. This method for solving feasibility problems has attracted a lot of attention due to its good performance even in nonconvex settings.

In this paper, we consider the Douglas–Rachford algorithm for finding a point in the intersection of two subspaces. We prove that the method converges strongly to the projection of the starting point onto the intersection. Moreover, if the sum of the two subspaces is closed, then the convergence is linear with the rate being the cosine of the Friedrichs angle between the subspaces. Our results improve upon existing results in three ways: First, we identify the location of the limit and thus reveal the method as a best approximation algorithm; second, we quantify the rate of convergence, and third, we carry out our analysis in general (possibly infinite-dimensional) Hilbert space. We also provide various examples as well as a comparison with the classical method of alternating projections.

1 Introduction

Throughout this paper, we assume that

$$(1) \qquad \boxed{X \text{ is a real Hilbert space}}$$

with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. Let U and V be closed convex subsets of X such that $U \cap V \neq \varnothing$. The *convex feasibility problem* is to find a point in $U \cap V$. This is a basic

---

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

†IME, Federal University of Goias, Goiania, G.O. 74001-970, Brazil. E-mail: yunier.bello@gmail.com.

‡Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: nghia.tran@ubc.ca.

§Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: hung.phan@ubc.ca.

¶Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: shawn.wang@ubc.ca.

problem in the natural sciences and engineering (see, e.g., [5], [15], and [16]) — as such, a plethora of algorithms based on the nearest point mappings (projectors) P_U and P_V have been proposed to solve it.

One particularly popular method is the *Douglas–Rachford splitting algorithm* [22] which utilizes the *Douglas–Rachford splitting operator*[1]

$$(2) \qquad\qquad\qquad T := P_V(2P_U - \mathrm{Id}) + \mathrm{Id} - P_U$$

and $x_0 \in X$ to generate the sequence $(x_n)_{n \in \mathbb{N}}$ by

$$(3) \qquad\qquad\qquad (\forall n \in \mathbb{N}) \quad x_{n+1} := T x_n.$$

While the sequence $(x_n)_{n \in \mathbb{N}}$ may or may not converge to a point in $U \cap V$, the projected ("shadow") sequence

$$(4) \qquad\qquad\qquad (P_U x_n)_{n \in \mathbb{N}}$$

always converges (weakly) to a point in $U \cap V$ (see [30], [17], [32], [4], [7]). The Douglas–Rachford algorithm has been applied very successfully to various problems where U and V are not necessarily convex, even though the supporting formal theory is far from being complete (see, e.g., [1], [8], and [24]). Very recently, Hesse, Luke and Neumann [27] (see also [26]) considered projection methods for the (nonconvex) sparse affine feasibility problem. Their paper highlights the importance of understanding the Douglas–Rachford algorithm for the case when U and V are closed *subspaces*[2] of X; their basic convergence result is the following.

Fact 1.1 (Hesse–Luke–Neumann) (See [27, Theorem 4.6].) *Suppose that X is finite-dimensional and U and V are subspaces of X. Then the sequence $(x_n)_{n \in \mathbb{N}}$ generated by (3) converges to a point in* $\mathrm{Fix}\, T := \{x \in X \mid x = Tx\}$ *with a linear rate*[3].

The aim of this paper is three-fold. We complement Fact 1.1 by providing the following:

- We identify the limit of the shadow sequence $(P_U x_n)_{n \in \mathbb{N}}$ as $P_{U \cap V} x_0$; consequently and somewhat surprisingly, the Douglas–Rachford method in this setting not only solves a feasibility problem but actually a *best approximation problem*.

- We quantify the rate of convergence — it turns out to be the cosine of the Friedrichs angle between U and V; moreover, our estimate is sharp.

- Our analysis is carried out in general (possibly infinite-dimensional) Hilbert space.

The paper is organized as follows. In Sections 2 and 3, we collect various auxiliary results to facilitate the proof of the main results (Theorem 4.1 and Theorem 4.3) in Section 4. In Section 5, we analyze the Douglas–Rachford algorithm for two lines in the Euclidean plane. The results obtained are used in Section 6 for an infinite-dimensional construction illustrating the lack of linear convergence. In Section 7, we compare the Douglas–Rachford algorithm to the method of alternating projections. We report on numerical experiments in Section 8, and conclude the paper in Section 9.

Notation is standard and follows largely [7]. We write $U \oplus V$ to indicate that the terms of the Minkowski sum $U + V = \{u + v \mid u \in U, v \in V\}$ satisfy $U \perp V$.

---

[1] Throughout this paper, we write $\alpha := \beta$ to indicate that α is defined to be equal to β.

[2] Throughout this paper, we use the term "subspace" to denote a (not necessarily closed) vector subspace in the sense of Linear Algebra.

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

2 Auxiliary results

In this section, we collect various results to ease the derivation of the main results.

2.1 Firmly nonexpansive mappings

It is well known (see [30], [23], or [8]) that the Douglas–Rachford operator T (see (2)) is *firmly nonexpansive*, i.e.,

$$(5) \qquad (\forall x \in X)(\forall y \in X) \quad \|Tx - Ty\|^2 + \|(\operatorname{Id} - T)x - (\operatorname{Id} - T)y\|^2 \leq \|x - y\|^2.$$

The following result will be useful in our analysis.

Fact 2.1 (See [7, Corollary 5.16 and Proposition 5.27], or [11, Theorem 2.2], [3] and [14].) *Let $T\colon X \to X$ be linear and firmly nonexpansive, and let $x \in X$. Then $T^n x \to P_{\operatorname{Fix} T} x$.*

2.2 Products of projections and the Friedrichs angle

Unless otherwise stated, we assume from now on that

$$(6) \qquad \boxed{U \text{ and } V \text{ are closed subspaces of } X.}$$

The proof of the following useful fact can be found in [21, Lemma 9.2]:

$$(7) \qquad U \subseteq V \;\Rightarrow\; P_U(V) \subseteq V \;\Leftrightarrow\; P_V P_U = P_U P_V = P_{U \cap V}.$$

Our main results are formulated using the notion of the Friedrichs angle between U and V. Let us review the definition and provide the key results which are needed in the sequel.

Definition 2.2 *The* cosine of the Friedrichs angle *between U and V is*

$$(8) \qquad c_F := \sup \left\{ |\langle u, v \rangle| \mid u \in U \cap (U \cap V)^\perp, \; v \in V \cap (U \cap V)^\perp, \; \|u\| \leq 1, \; \|v\| \leq 1 \right\}.$$

We write $c_F(U, V)$ for c_F if we emphasize the subspaces utilized.

Fact 2.3 (fundamental properties of the Friedrichs angle) *Let $n \in \{1, 2, 3, \ldots\}$. Then the following hold:*

(i) *$U + V$ is closed $\Leftrightarrow c_F < 1$.*

(ii) *$c_F(U, V) = c_F(V, U) = c_F(U^\perp, V^\perp)$.*

(iii) *$c_F = \|P_V P_U - P_{U \cap V}\| = \|P_{V^\perp} P_{U^\perp} - P_{U^\perp \cap V^\perp}\|$.*

(iv) **(Aronszajn–Kayalar–Weinert)** *$\|(P_V P_U)^n - P_{U \cap V}\| = c_F^{2n-1}$.*

3

Proof. (i): See [20, Theorem 13] (ii): See [20, Theorem 16] (iii): See [21, Lemma 9.5(7)] and (ii) above. (iv): See [21, Theorem 9.31] (or the original works [2] and [29]). ∎

Note that Fact 2.3(ii)&(iii) yield

$$(9) \qquad\qquad c_F = 0 \quad \Leftrightarrow \quad P_V P_U = P_U P_V = P_{U \cap V}.$$

The classical Fact 2.3(iv) deals with *even* powers of alternating projectors. We complement this result by deriving the counterpart for *odd* powers.

Lemma 2.4 (odd powers of alternating projections) *Let* $n \in \{1, 2, 3, \dots\}$. *Then*

$$(10) \qquad\qquad \|P_U(P_V P_U)^n - P_{U \cap V}\| = c_F^{2n}.$$

Proof. If $c_F = 0$, then the conclusion is clear from (9). We thus assume that $c_F > 0$. By (7),

$$(11a) \quad ((P_U P_V)^n - P_{U \cap V})(P_V P_U - P_{U \cap V}) = (P_U P_V)^n P_V P_U - (P_U P_V)^n P_{U \cap V} - P_{U \cap V} P_V P_U + P_{U \cap V}^2$$

$$(11b) \qquad\qquad\qquad = (P_U P_V)^n P_U - P_{U \cap V} - P_{U \cap V} + P_{U \cap V}$$

$$(11c) \qquad\qquad\qquad = P_U (P_V P_U)^n - P_{U \cap V}.$$

It thus follows from Fact 2.3(iv)&(iii) that

$$(12) \qquad \|P_U(P_V P_U)^n - P_{U \cap V}\| \leq \|(P_U P_V)^n - P_{U \cap V}\| \cdot \|P_V P_U - P_{U \cap V}\| = c_F^{2n}.$$

Since $(P_U(P_V P_U)^n - P_{U \cap V})(P_U P_V - P_{U \cap V}) = (P_U P_V)^{n+1} - P_{U \cap V}$, we obtain from Fact 2.3(iv)&(iii) that

$$(13a) \quad c_F^{2n+1} = \|(P_U P_V)^{n+1} - P_{U \cap V}\| \leq \|P_U(P_V P_U)^n - P_{U \cap V}\| \cdot \|P_U P_V - P_{U \cap V}\|$$

$$(13b) \qquad\quad = \|P_U(P_V P_U)^n - P_{U \cap V}\| c_F.$$

Since $c_F > 0$, we obtain $c_F^{2n} \leq \|P_U(P_V P_U)^n - P_{U \cap V}\|$. Combining with (12), we deduce (10). ∎

3   Basic properties of the Douglas–Rachford splitting operator

We recall that the Douglas–Rachford operator is defined by

$$(14) \qquad\qquad T = T_{V,U} := P_V(2P_U - \mathrm{Id}) + \mathrm{Id} - P_U.$$

For future reference, we record the following consequence of Fact 2.1.

Corollary 3.1 *Let* $x \in X$. *Then* $T^n x \to P_{\mathrm{Fix}\, T} x$.

It will be very useful to work with reflectors, which we define next.

Definition 3.2 (reflector) *The* reflector *associated with U is*

$$(15) \qquad\qquad R_U := 2P_U - \mathrm{Id} = P_U - P_{U^\perp}.$$

The following simple yet useful result is easily verified.

4

Proposition 3.3 *The reflector R_U is a surjective isometry with*

$$(16) \qquad\qquad R_U^* = R_U^{-1} = -R_{U^\perp} = R_U.$$

We now record several reformulations of T and T^* which we will use repeatedly in the paper without explicitly mentioning it. Item (i) also appears in [19, Section 2.3]; it is interesting to observe that the operator T also occurs in the convergence analysis of the Douglas–Rachford algorithm applied to ℓ_1 minimization with linear constraints.

Proposition 3.4 *The following hold:*

    (i) $T = \frac{1}{2}\operatorname{Id} + \frac{1}{2}R_V R_U = P_V R_U + \operatorname{Id} - P_U = P_V P_U + P_{V^\perp} P_{U^\perp}$.

    (ii) $T^* = P_U P_V + P_{U^\perp} P_{V^\perp} = T_{U,V} = T^*_{V,U}$.

    (iii) $T_{V,U} = T_{V^\perp,U^\perp}$.

Proof. (i): Expand and use the linearity of P_U and P_V. (ii)&(iii): This follows from (i).   ■

The next result highlights the importance of the reflectors when traversing between T and T^*. Item (iv) was observed previously in [13, Remark 4.1].

Proposition 3.5 *The following hold:*

    (i) $R_U T^* = T R_U = T^* R_V = R_V T = P_U + P_V - \operatorname{Id}$.

    (ii) $T^*(R_V R_U) = (R_V R_U)T^* = T$ and $T(R_U R_V) = (R_U R_V)T = T^*$.

    (iii) $TT^* = T^*T$, *i.e., T is normal.*

    (iv) $2TT^* = T + T^*$.

    (v) TT^* *is firmly nonexpansive and self-adjoint.*

    (vi) $TT^* = P_V P_U P_V + P_{V^\perp} P_{U^\perp} P_{V^\perp} = P_V P_U + P_U P_V - P_U - P_V + \operatorname{Id} = P_U P_V P_U + P_{U^\perp} P_{V^\perp} P_{U^\perp}$.

Proof. (i): Indeed, using Proposition 3.4(i)&(ii), we see that

$$(17a) \qquad TR_U = (P_V P_U + P_{V^\perp} P_{U^\perp})(P_U - P_{U^\perp}) = P_V P_U - P_{V^\perp} P_{U^\perp} = P_V P_U - (\operatorname{Id} - P_V)(\operatorname{Id} - P_U)$$

$$(17b) \qquad\qquad = P_U + P_V - \operatorname{Id}$$

$$(17c) \qquad\qquad = P_U P_V - (\operatorname{Id} - P_U)(\operatorname{Id} - P_V) = P_U P_V - P_{U^\perp} P_{V^\perp} = (P_U P_V + P_{U^\perp} P_{V^\perp})(P_V - P_{V^\perp})$$

$$(17d) \qquad\qquad = T^* R_V.$$

Because $P_U + P_V - \operatorname{Id}$ is self-adjoint, so are TR_U and $T^* R_V$. Hence $R_U T^* = (TR_U)^* = TR_U = T^* R_V = (R_V T)^* = R_V T$ by Proposition 3.3.

    (ii): Clear from (i).

    (iii): Using (ii), we obtain $TT^* = T^*(R_V R_U)(R_U R_V)T = T^*T$.

    (iv): $4TT^* = (\operatorname{Id} + R_V R_U)(\operatorname{Id} + R_U R_V) = 2\operatorname{Id} + R_V R_U + R_U R_V = 2((\operatorname{Id} + R_V R_U)/2 + (\operatorname{Id} + R_U R_V)/2) = 2(T + T^*)$.

    (v): Since T and T^* are firmly nonexpansive, so is their convex combination $(T + T^*)/2$, which equals TT^* by (iv). It is clear that TT^* is self-adjoint.

5

(vi): It follows from Proposition 3.4(i)&(ii) that $TT^* = (P_V P_U + P_{V^\perp} P_{U^\perp})(P_U P_V + P_{U^\perp} P_{V^\perp}) = P_V P_U P_V + P_{V^\perp} P_{U^\perp} P_{V^\perp}$, which yields the first equality. Replacing $P_{U^\perp}$ and $P_{V^\perp}$ by $\mathrm{Id} - P_U$ and $\mathrm{Id} - P_V$, respectively, followed by expanding and simplifying results in the second equality. The last equality is proved analogously. ∎

Parts of our next result were also obtained in [26] and [27] when X is finite-dimensional. Concerning item (ii), we note in passing that if $U + V = X$, then $(U^\perp, V^\perp)$ satisfies the *inverse best approximation property*; see [18, Corollary 2.12] for details.

Proposition 3.6 *Let $n \in \mathbb{N}$. Then the following hold:*

(i) $\operatorname{Fix} T = \operatorname{Fix} T^* = \operatorname{Fix} T^* T = (U \cap V) \oplus (U^\perp \cap V^\perp)$.

(ii) $\operatorname{Fix} T = U \cap V \Leftrightarrow \overline{U + V} = X$.

(iii) $P_{\operatorname{Fix} T} = P_{U \cap V} + P_{U^\perp \cap V^\perp}$.

(iv) $P_{\operatorname{Fix} T} T = T P_{\operatorname{Fix} T} = P_{\operatorname{Fix} T}$.

(v) $P_U P_{\operatorname{Fix} T} = P_V P_{\operatorname{Fix} T} = P_{U \cap V} P_{\operatorname{Fix} T} = P_{U \cap V} = P_{U \cap V} P_{\operatorname{Fix} T} T^n = P_{U \cap V} T^n$.

Proof. (i): Set $A = N_U$ and $B = N_V$, where N_U denotes the normal cone operator defined by $N_U(x) = \{x^* \in X \mid \sup_{u \in U} \langle x^*, u - x \rangle \leq 0\}$ if $x \in U$ and $N_U(x) = \varnothing$ otherwise. Then $(A + B)^{-1}(0) = U \cap V$. Combining [6, Example 2.7 and Corollary 5.5(iii)] yields $\operatorname{Fix} T = (U \cap V) \oplus (U^\perp \cap V^\perp)$. By [11, Lemma 2.1], we have $\operatorname{Fix} T = \operatorname{Fix} T^*$. Since T and T^* are firmly nonexpansive, and $0 \in \operatorname{Fix} T \cap \operatorname{Fix} T^*$, we apply [7, Corollary 4.3 and Corollary 4.37] to deduce that $\operatorname{Fix} T^* T = \operatorname{Fix} T \cap \operatorname{Fix} T^*$.

(ii): Using (i), we obtain $\operatorname{Fix} T = U \cap V \Leftrightarrow U^\perp \cap V^\perp = \{0\} \Leftrightarrow \overline{U + V} = \overline{U^{\perp\perp} + V^{\perp\perp}} = (U^\perp \cap V^\perp)^\perp = \{0\}^\perp = X$.

(iii): This follows from (i).

(iv): Clearly, $T P_{\operatorname{Fix} T} = P_{\operatorname{Fix} T}$. Furthermore, $P_{\operatorname{Fix} T} T = T P_{\operatorname{Fix} T}$ by [11, Lemma 3.12].

(v): First, (iii) and (7) imply $P_U P_{\operatorname{Fix} T} = P_V P_{\operatorname{Fix} T} = P_{U \cap V} P_{\operatorname{Fix} T} = P_{U \cap V}$. This and (iv) give the remaining equalities. ∎

4 Main result

We now are ready for our main results concerning the dynamical behaviour of the Douglas–Rachford iteration.

Theorem 4.1 (powers of T) *Let $n \in \{1, 2, 3, \ldots\}$, and let $x \in X$. Then*

(18a) $$\|T^n - P_{\operatorname{Fix} T}\| = c_F^n,$$

(18b) $$\|(TT^*)^n - P_{\operatorname{Fix} T}\| = c_F^{2n},$$

and

(19) $$\|T^n x - P_{\operatorname{Fix} T} x\| \leq c_F^n \|x - P_{\operatorname{Fix} T} x\| \leq c_F^n \|x\|.$$

Proof. Set

(20) $$c := \|T - P_{\text{Fix } T}\| = \|TP_{(\text{Fix } T)^\perp}\|,$$

and observe that the second equality is justified since $P_{\text{Fix } T} = TP_{\text{Fix } T}$. Since T is (firmly) nonexpansive and normal (see Proposition 3.5(iii)), it follows from [11, Lemma 3.15(i)] that

(21) $$\|T^n - P_{\text{Fix } T}\| = c^n.$$

Proposition 3.4(i), Proposition 3.6(iii), and (7) imply

(22a) $$(T - P_{\text{Fix } T})x = \left(P_V P_U + P_{V^\perp} P_{U^\perp}\right)x - \left(P_{U \cap V} + P_{U^\perp \cap V^\perp}\right)x$$

(22b) $$= \underbrace{\left(P_V P_U - P_{U \cap V}\right)x}_{\in V} + \underbrace{\left(P_{V^\perp} P_{U^\perp} - P_{U^\perp \cap V^\perp}\right)x}_{\in V^\perp}$$

(22c) $$= \left(P_V P_U P_U x - P_{U \cap V} P_U x\right) + \left(P_{V^\perp} P_{U^\perp} P_{U^\perp} x - P_{U^\perp \cap V^\perp} P_{U^\perp} x\right)$$

(22d) $$= \underbrace{\left(P_V P_U - P_{U \cap V}\right) P_U x}_{\in V} + \underbrace{\left(P_{V^\perp} P_{U^\perp} - P_{U^\perp \cap V^\perp}\right) P_{U^\perp} x}_{\in V^\perp}.$$

Hence, using (22d) and Fact 2.3(iii), we obtain

(23a) $$\|(T - P_{\text{Fix } T})x\|^2 = \left\|\left(P_V P_U - P_{U \cap V}\right) P_U x\right\|^2 + \left\|\left(P_{V^\perp} P_{U^\perp} - P_{U^\perp \cap V^\perp}\right) P_{U^\perp} x\right\|^2$$

(23b) $$\leq \left\|P_V P_U - P_{U \cap V}\right\|^2 \|P_U x\|^2 + \left\|P_{V^\perp} P_{U^\perp} - P_{U^\perp \cap V^\perp}\right\|^2 \|P_{U^\perp} x\|^2$$

(23c) $$= c_F^2 \|P_U x\|^2 + c_F^2 \|P_{U^\perp} x\|^2$$

(23d) $$= c_F^2 \|x\|^2.$$

We deduce that

(24) $$c = \|T - P_{\text{Fix } T}\| \leq c_F.$$

Furthermore, (22b) implies

(25a) $$\|(T - P_{\text{Fix } T})x\|^2 = \left\|\left(P_V P_U - P_{U \cap V}\right)x\right\|^2 + \left\|\left(P_{V^\perp} P_{U^\perp} - P_{U^\perp \cap V^\perp}\right)x\right\|^2$$

(25b) $$\geq \left\|\left(P_V P_U - P_{U \cap V}\right)x\right\|^2.$$

This and Fact 2.3(iii) yield

(26) $$c = \|T - P_{\text{Fix } T}\| \geq \|P_V P_U - P_{U \cap V}\| = c_F.$$

Combining (24) and (26), we obtain $c = c_F$. Consequently, (21) yields (18a) while (18b) follows from [11, Lemma 3.15.(2)].

Finally, [11, Lemma 3.14(1)&(3)] results in $\|T^n x - P_{\text{Fix } T} x\| \leq c^n \|x - P_{\text{Fix } T} x\| = c_F^n \|x - P_{\text{Fix } T} x\| = c_F^n \|P_{\text{Fix } T^\perp} x\| \leq c_F^n \|x\|$. ∎

The following result yields further insights in various powers of T and T^*.

Proposition 4.2 *Let* $n \in \mathbb{N}$. *Then the following hold:*

(i) $(TT^*)^n = (T^*T)^n = (P_U P_V P_U)^n + (P_{U^\perp} P_{V^\perp} P_{U^\perp})^n = (P_V P_U P_V)^n + (P_{V^\perp} P_{U^\perp} P_{V^\perp})^n$ *if* $n \geq 1$.

(ii) $P_U(TT^*)^n = (P_U P_V)^n P_U = (TT^*)^n P_U$ and $P_V(TT^*)^n = (P_V P_U)^n P_V = (P_V P_U)^n P_V$.

(iii) $T^{2n} = (TT^*)^n (R_V R_U)^n$.

(iv) $T^{2n+1} = (TT^*)^n T (R_V R_U)^n = (TT^*)^n T^* (R_V R_U)^{n+1}$.

(v) $T^{2n} = \big((P_U P_V)^n P_U + (P_{U^\perp} P_{V^\perp})^n P_{U^\perp} \big)(R_V R_U)^n = \big((P_V P_U)^n P_V + (P_{V^\perp} P_{U^\perp})^n P_{V^\perp} \big)(R_V R_U)^n$.

(vi) $T^{2n+1} = \big((P_U P_V)^{n+1} + (P_{U^\perp} P_{V^\perp})^{n+1} \big)(R_V R_U)^{n+1} = \big((P_V P_U)^{n+1} + (P_{V^\perp} P_{U^\perp})^{n+1} \big)(R_V R_U)^n$.

Proof. (i): Proposition 3.5(iii)&(vi) yield

$$(27) \qquad (TT^*)^n = (P_U P_V P_U + P_{U^\perp} P_{V^\perp} P_{U^\perp})^n = (P_U P_V P_U)^n + (P_{U^\perp} P_{V^\perp} P_{U^\perp})^n;$$

the last equality follows similarly.

(ii): By (i) we have that $P_U(TT^*)^n = (P_U P_V P_U)^n = (P_U P_V)^n P_U$ and $(TT^*)^n P_U = (P_U P_V P_U)^n = (P_U P_V)^n P_U$. The proof of the remaining equalities is similar.

(iii): Since $T = T^* R_V R_U = R_V R_U T^*$ (see Proposition 3.5(ii)), we have $T^n = (T^*)^n (R_V R_U)^n$. Thus $T^{2n} = T^n (T^*)^n (R_V R_U)^n$, and therefore $T^{2n} = (TT^*)^n (R_V R_U)^n$ using Proposition 3.5(iii).

(iv): Using (iii) and Proposition 3.5(iii)&(ii), we have $T^{2n+1} = T T^{2n} = T(TT^*)^n (R_V R_U)^n = (TT^*)^n T (R_V R_U)^n = (TT^*)^n (T^* R_V R_U)(R_V R_U)^n = (TT^*)^n T^* (R_V R_U)^{n+1}$.

(v)&(vi): Using (iii), (i), (iv), and Proposition 3.4(ii), we have

$$(28) \quad T^{2n} = \big((P_U P_V P_U)^n + (P_{U^\perp} P_{V^\perp} P_{U^\perp})^n \big)(R_V R_U)^n = \big((P_U P_V)^n P_U + (P_{U^\perp} P_{V^\perp})^n P_{U^\perp} \big)(R_V R_U)^n$$

and

$$(29a) \qquad T^{2n+1} = \big((P_U P_V P_U)^n + (P_{U^\perp} P_{V^\perp} P_{U^\perp})^n \big)(P_U P_V + P_{U^\perp} P_{V^\perp})(R_V R_U)^{n+1}$$
$$(29b) \qquad = \big((P_U P_V)^{n+1} + (P_{U^\perp} P_{V^\perp})^{n+1} \big)(R_V R_U)^{n+1}.$$

Similarly, using (iii), (i), (iv), and Proposition 3.4(ii), we have

$$(30a) \qquad T^{2n} = \big((P_V P_U P_V)^n + (P_{V^\perp} P_{U^\perp} P_{V^\perp})^n \big)(R_V R_U)^n$$
$$(30b) \qquad = \big((P_V P_U)^n P_V + (P_{V^\perp} P_{U^\perp})^n P_{V^\perp} \big)(R_V R_U)^n$$

and

$$(31a) \qquad T^{2n+1} = \big((P_V P_U P_V)^n + (P_{V^\perp} P_{U^\perp} P_{V^\perp})^n \big)(P_V P_U + P_{V^\perp} P_{U^\perp})(R_V R_U)^n$$
$$(31b) \qquad = \big((P_V P_U)^{n+1} + (P_{V^\perp} P_{U^\perp})^{n+1} \big)(R_V R_U)^n.$$

The proof is complete. ∎

We are now ready for our main result. Note that item (i) is the counterpart of Fact 2.3(iv) for the Douglas–Rachford algorithm.

Theorem 4.3 (shadow powers of T) *Let $n \in \mathbb{N}$, and let $x \in X$. Then the following hold:*

(i) $\|P_U T^n - P_{U \cap V}\| = \|P_V T^n - P_{U \cap V}\| = c_F^n$.

(ii) $\max \big\{ \|P_U T^n x - P_{U \cap V} x\|, \|P_V T^n x - P_{U \cap V} x\| \big\} \le c_F^n \|x\|$.

(iii) $\|P_V T^{n+1} x - P_{U \cap V} x\| \leq c_F \|P_U T^n x - P_{U \cap V} x\| \leq c_F \|T^n x - P_{\text{Fix } T} x\| \leq c_F^{n+1} \|x\|.$

Proof. (i): Note that $P_{U \cap V} R_V R_U = P_{U \cap V} (P_V - P_{V^\perp}) R_U = P_{U \cap V} R_U = P_{U \cap V} (P_U - P_{U^\perp}) = P_{U \cap V}$. It follows that $P_{U \cap V} = P_{U \cap V} (R_V R_U)^n = P_{U \cap V} (R_V R_U)^{n+1}$. Hence, using Proposition 4.2(v), we have $P_U T^{2n} - P_{U \cap V} = (P_U P_V)^n P_U (R_V R_U)^n - P_{U \cap V} (R_V R_U)^n = ((P_U P_V)^n P_U - P_{U \cap V})(R_V R_U)^n$ and thus $\|P_U T^{2n} - P_{U \cap V}\| = c_F^{2n}$ by Lemma 2.4. It follows likewise from Proposition 4.2(vi) and Fact 2.3(iv) that $\|P_U T^{2n+1} - P_{U \cap V}\| = \|((P_U P_V)^{n+1} - P_{U \cap V})(R_V R_U)^{n+1}\| = \|(P_U P_V)^{n+1} - P_{U \cap V}\| = c_F^{2n+1}$. Thus, we have $\|P_U T^n - P_{U \cap V}\| = c_F^n$. The proof of $\|P_V T^n - P_{U \cap V}\| = c_F^n$ is analogous.

(ii): Clear from (i).

(iii): Using Proposition 3.4(i) and Proposition 3.6(v), we have

(32a) $\quad P_V T^{n+1} - P_{U \cap V} = P_V (P_V P_U + P_{V^\perp} P_{U^\perp}) T^n - P_{U \cap V} P_{\text{Fix } T} = P_V P_U T^n - P_{U \cap V} P_{\text{Fix } T} T^n$

(32b) $\quad\quad\quad = P_V P_U T^n - P_{U \cap V} + P_{U \cap V} - P_{U \cap V} T^n = (P_V P_U - P_{U \cap V})(P_U T^n - P_{U \cap V}).$

Combining this with Fact 2.3(iii) and (18a), we get $\|P_V T^{n+1} x - P_{U \cap V} x\| \leq c_F \|P_U T^n x - P_{U \cap V} x\| = c_F \|P_U (T^n x - P_{\text{Fix } T} x)\| \leq c_F \|T^n x - P_{\text{Fix } T} x\| \leq c_F^{n+1} \|x\|.$ ∎

Corollary 4.4 (linear convergence) *Let $x \in X$. Then $T^n x \to P_{(U \cap V) + (U^\perp \cap V^\perp)} x$, $P_U T^n x \to P_{U \cap V} x$, and $P_V T^n x \to P_{U \cap V} x$. If $U + V$ is closed, then convergence of these sequences is linear with rate $c_F < 1$.*

Proof. Corollary 3.1, Proposition 3.6(iii)&(v) and (7) imply $P_U T^n x \to P_U P_{\text{Fix } T} x = P_{U \cap V} x$ and analogously $P_V T^n x \to P_{U \cap V} x$. Recall from Fact 2.3(i) that $U + V$ is closed if and only if $c_F < 1$. The conclusion is thus clear from Theorem 4.3(iii). ∎

A translation argument gives the following result (see also [9, Theorem 3.17] for an earlier related result).

Corollary 4.5 (affine subspaces) *Suppose that U and V are closed affine subspaces of X such that $U \cap V \neq \varnothing$, and let $x \in X$. Then*

(33) $\quad\quad\quad T^n x \to P_{\text{Fix } T} x, \quad P_U T^n x \to P_{U \cap V} x, \quad \text{and} \quad P_V T^n x \to P_{U \cap V} x.$

If $(U - U) + (V - V)$ is closed, then the convergence is linear with rate $c_F(U - U, V - V) < 1$.

5 Two lines in the Euclidean plane

We present some geometric results concerning the lines in the plane which will not only be useful later but which also illustrate the results of the previous sections. In this section, we assume that $X = \mathbb{R}^2$, and we set

(34) $\quad e_0 := (1, 0), \quad e_{\pi/2} := (0, 1), \quad \text{and} \quad (\forall \theta \in [0, \pi/2]) \ e_\theta := \cos(\theta) e_0 + \sin(\theta) e_{\pi/2}.$

Define the (counter-clockwise) rotator by

(35) $\quad\quad\quad (\forall \theta \in \mathbb{R}_+) \quad R_\theta := \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$

and note that $R_\theta^{-1} = R_\theta^*$. Now let $\theta \in {]0, \pi/2]}$, and suppose that

(36) $\quad\quad\quad U = \mathbb{R} \cdot e_0 \quad \text{and} \quad V = \mathbb{R} \cdot e_\theta = R_\theta(U).$

9

Then

(37)
$$U \cap V = \{0\} \quad \text{and} \quad c_F(U, V) = \cos(\theta).$$

By, e.g., [7, Proposition 28.2(ii)], $P_V = R_\theta P_U R_\theta^*$. In terms of matrices, we thus have

(38)
$$P_U = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad P_V = \begin{pmatrix} \cos^2(\theta) & \sin(\theta)\cos(\theta) \\ \sin(\theta)\cos(\theta) & \sin^2(\theta) \end{pmatrix}.$$

Consequently, the corresponding Douglas–Rachford splitting operator is

(39a)
$$T = P_V(2P_U - \text{Id}) + \text{Id} - P_U = \begin{pmatrix} \cos^2(\theta) & -\sin(\theta)\cos(\theta) \\ \sin(\theta)\cos(\theta) & 1 - \sin^2(\theta) \end{pmatrix}$$

(39b)
$$= \cos(\theta) \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \cos(\theta) R_\theta.$$

Thus, $\text{Fix } T = \{0\} = P_U(\text{Fix } T)$,

(40)
$$(\forall n \in \mathbb{N}) \quad T^n = \cos^n(\theta) R_{n\theta}, \quad P_U T^n = \cos^n(\theta) \begin{pmatrix} \cos(n\theta) & -\sin(n\theta) \\ 0 & 0 \end{pmatrix},$$

and

(41)
$$(\forall x \in X) \quad \|T^n x\| = \cos^n(\theta) \|x\|, \quad \|P_U T^n x\| = \cos^n(\theta) |\cos(n\theta) x_1 - \sin(n\theta) x_2|.$$

Furthermore,

(42)
$$(\forall n \geq 1) \quad (P_V P_U)^n = \cos^{2n-1}(\theta) \begin{pmatrix} \cos(\theta) & 0 \\ \sin(\theta) & 0 \end{pmatrix}$$

and thus

(43)
$$(\forall x \in X) \quad \|(P_V P_U)^n x\| = \cos^{2n-1}(\theta) |x_1| \leq \cos^{2n-1}(\theta) \|x\|.$$

6 An example without linear rate of convergence

In this section, let us assume that our underlying Hilbert space is $\ell^2(\mathbb{N}) = \mathbb{R}^2 \oplus \mathbb{R}^2 \oplus \cdots$. It will be more suggestive to use boldface letters for vectors lying in, and operators acting on, $\ell^2(\mathbb{N})$. Thus,

(44)
$$\mathbf{X} = \ell^2(\mathbb{N}),$$

and we write is $\mathbf{x} = (x_n)_{n \in \mathbb{N}} = ((x_0, x_1), (x_2, x_3), \ldots)$ for a generic vector in $\mathbf{X}$. Suppose that $(\theta_n)_{n \in \mathbb{N}}$ is a sequence of angles in $]0, \pi/2[$ with $\theta_n \to 0^+$. We set $(\forall n \in \mathbb{N}) \ c_n := \cos(\theta_n) \to 1^-$. We will use notation and results from Section 5. We assume that

(45)
$$\mathbf{U} = \mathbb{R} \cdot e_0 \times \mathbb{R} \cdot e_0 \times \cdots \subseteq \mathbf{X}$$

and that

(46)
$$\mathbf{V} = \mathbb{R} \cdot e_{\theta_0} \times \mathbb{R} \cdot e_{\theta_1} \times \cdots \subseteq \mathbf{X}.$$

Then

(47) $$\mathbf{U} \cap \mathbf{V} = \{\mathbf{0}\} \quad \text{and} \quad c_F(\mathbf{U}, \mathbf{V}) = \sup_{n \in \mathbb{N}} c_F(\mathbb{R} \cdot e_0, \mathbb{R} \cdot e_{\theta_n}) = \sup_{n \in \mathbb{N}} c_n = 1.$$

The Douglas–Rachford splitting operator is

(48) $$\mathbf{T} = c_0 R_{\theta_0} \oplus c_1 R_{\theta_1} \oplus \cdots .$$

Now let $\mathbf{x} = (x_n)_{n \in \mathbb{N}} \in \mathbf{X}$ and $\gamma \in {]}0,1{[}$. Assume further that $\{n \in \mathbb{N} \mid x_n \neq 0\}$ is infinite. Then there exists $N \in \mathbb{N}$ such that $(x_{2N}, x_{2N+1}) \in \mathbb{R}^2 \smallsetminus \{(0,0)\}$ and $c_N > \gamma$. Hence

(49) $$\gamma^{-n} \|\mathbf{T}^n \mathbf{x}\| \geq \gamma^{-n} c_N^n \|(x_{2N}, x_{2N+1})\| \to +\infty;$$

consequently,

(50) $$\mathbf{T}^n \mathbf{x} \to \mathbf{0}, \text{ but not linearly with rate } \gamma.$$

Let us now assume in addition that $\theta_0 = \pi/3$ and $(\forall n \geq 1)\ \theta_n = \pi/(4n)$ and $x_n = 1/(n+1)$. Then there exists $\delta > 1$ and $N \geq 1$ such that $(\forall n \geq N)\ \gamma^{-1} c_n \geq \delta$. Hence, for every $n \geq N$, we have

(51) $$\gamma^{-n} \|P_{\mathbf{U}} \mathbf{T}^n \mathbf{x}\| \geq \gamma^{-n} c_n^n |\cos(n\theta_n) x_{2n} - \sin(n\theta_n) x_{2n+1}| \geq \frac{\delta^n}{2^{3/2}(n+1)(2n+1)} \to +\infty;$$

thus,

(52) $$P_{\mathbf{U}} \mathbf{T}^n \mathbf{x} \to \mathbf{0}, \text{ but not linearly with rate } \gamma.$$

In summary, these constructions illustrate that when the Friedrichs angle is zero, then one cannot expect linear convergence of the (projected) iterates of the Douglas–Rachford splitting operator.

7 Comparison with the method of alternating projections

Let us now compare our main results (Theorems 4.1 and 4.3) with the *method of alternating projections*, for which the following fundamental result is well known.

Fact 7.1 (Aronszajn) (See [2] or [21, Theorem 9.8].) *Let* $x \in X$. *Then*

(53) $$(\forall n \geq 1) \quad \|(P_V P_U)^n x - P_{U \cap V} x\| \leq c_F^{2n-1} \|x - P_{U \cap V} x\|.$$

In Fact 7.1, the rate c_F is best possible (see Fact 2.3(iv) and the results and comments in [21, Chapter 9]), and if the Friedrichs angle is 0, then slow convergence may occur (see, e.g., [10]).

From Theorem 4.1, Corollary 4.4, and Fact 7.1, we see that the rate of convergence c_F of $(T^n x)_{n \in \mathbb{N}}$ to $P_{\text{Fix } T} x$ and, *a fortiori*, of $(P_U T^n x)_{n \in \mathbb{N}}$ to $P_{U \cap V} x$, is clearly *slower* than the rate of convergence c_F^2 of $((P_V P_U)^n x)_{n \in \mathbb{N}}$ to $P_{U \cap V} x$. In other words, the Douglas–Rachford splitting method appears to be twice as slow as the method of alternating projections. While this is certainly the case for the iterates $(T^n x)_{n \in \mathbb{N}}$, the actual iterates of interest, namely $(P_U T^n x)_{n \in \mathbb{N}}$, in practice often (somewhat paradoxically) make striking non-monotone progress.

Let us illustrate this using the set up of Section 5 the notation and results of which we will utilize.

Consider first (41) and (43) with $x = e_0$ and $\theta = \pi/17$. In Figure 1, we show the first 100 iterates of the sequences $(\|T^n x\|)_{n \in \mathbb{N}}$ (red line), $(\|P_U T^n x\|)_{n \in \mathbb{N}}$ (blue line), and $(\|(P_V P_U)^n x\|)_{n \in \mathbb{N}}$ (green line). The sequences $(\|T^n x\|)_{n \in \mathbb{N}}$ and $(\|(P_V P_U)^n x\|)_{n \in \mathbb{N}}$, which are decreasing, represent the distance of the iterates to 0, the unique solution of the problem. While $(\|(P_V P_U)^n x\|)_{n \in \mathbb{N}}$ decreases faster than $(\|T^n x\|)_{n \in \mathbb{N}}$, the sequence of "shadows" $(\|P_U T^n x\|)_{n \in \mathbb{N}}$ exhibits a curious non-monotone "rippling" behaviour — it may be quite close to the solution soon after the iteration starts!
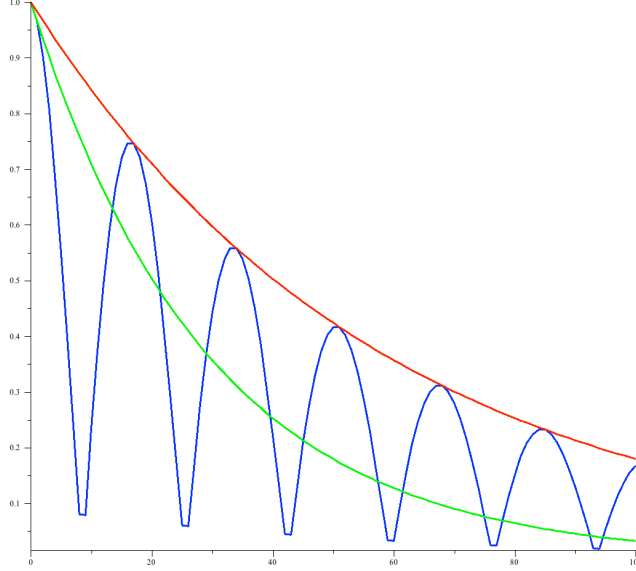
Figure 1: The distance of the first 100 terms, connected piecewise linearly, of the sequences $(T^n x)_{n \in \mathbb{N}}$ (red), $(P_U T^n x)_{n \in \mathbb{N}}$ (blue), and $((P_V P_U)^n x)_{n \in \mathbb{N}}$ (green) to the unique solution.

We next show in Figure 2 and Figure 3 the first 100 terms of $(\|T^n x\|)_{n \in \mathbb{N}}$ and $(\|(P_V P_U)^n x\|)_{n \in \mathbb{N}}$, where $\theta \colon [0,1] \to [0, \pi/2] \colon t \mapsto (\pi/2)t^3$ is parametrized to exhibit more clearly the behaviour for small angles. Clearly, and as predicted, smaller angles correspond to slower rates of convergence.
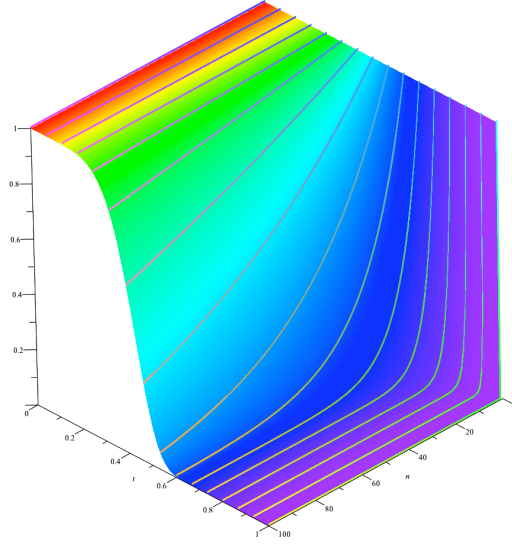
Figure 2: The distance of the first 100 terms of the sequence $(T^n x)_{n \in \mathbb{N}}$ to the unique solution when the angle ranges between 0 and $\pi/2$.
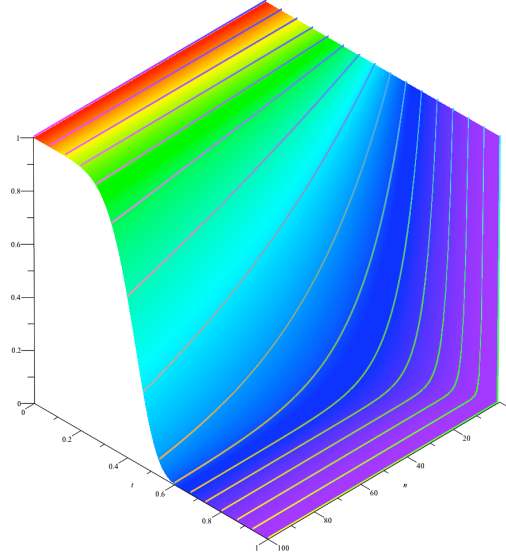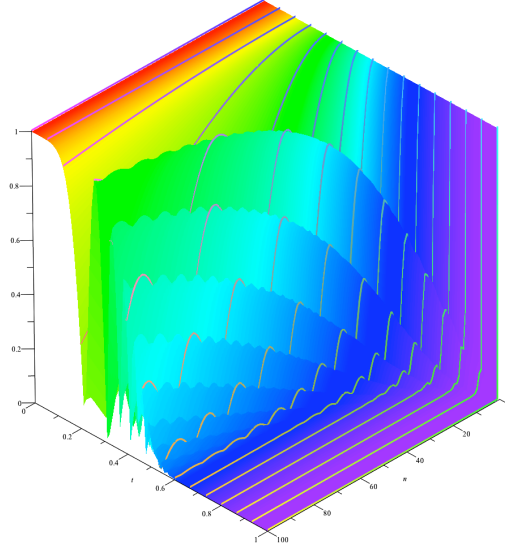

Figure 3: The distance of the first 100 terms of the sequence $((P_V P_U)^n x)_{n \in \mathbb{N}}$ to the unique solution when the angle ranges between 0 and $\pi/2$.

In Figure 4, we depict the "shadow sequence" $(\|P_U T^n x\|)_{n \in \mathbb{N}}$. Observe again the "rippling" phenomenon. While the situation of two lines appears at first to be quite special, it turns out that the same "rippling" also arises in quite different settings; see [12, Figures 4 and 6] and [19].

Figure 4: The distance of the first 100 terms of the "shadow" sequence $(P_U T^n x)_{n \in \mathbb{N}}$ to the unique solution when the angle ranges between 0 and $\pi/2$.

The figures in this section were prepared in $\texttt{Maple}^{\text{TM}}$ (see [31]).

8 Numerical experiments

In this section, we compare the Douglas–Rachford method (DRM) to the method of alternating projections (MAP) for finding $P_{U \cap V} x_0$. Our numerical set up is as follows. We assume that $X = \mathbb{R}^{50}$, and we randomly generated 100 pairs of subspaces U and V of X such that $U \cap V \neq \{0\}$. We then chose 10 random starting points, each with Euclidean norm 10. This resulted in a total of 1,000 instances for each algorithm. Note that the sequences to monitor are

(54) $$\left(P_U T^n x_0\right)_{n \in \mathbb{N}} \quad \text{and} \quad \left((P_V P_U)^n x_0\right)_{n \in \mathbb{N}}$$

for DRM and for MAP, respectively. Our stopping criterion tolerance was set to

(55) $$\varepsilon := 10^{-3}.$$

We investigated two different stopping criteria, which we detail and discuss in the following two sections.

8.1 Stopping criterion based on the true error

We terminated the algorithm when the current iterate of the monitored sequence $(z_n)_{n \in \mathbb{N}}$ satisfies

(56) $$d_{U \cap V}(z_n) < \varepsilon$$

for the first time. Note that in applications, we typically would not have access to this information but here we use it to see the true performance of the two methods.
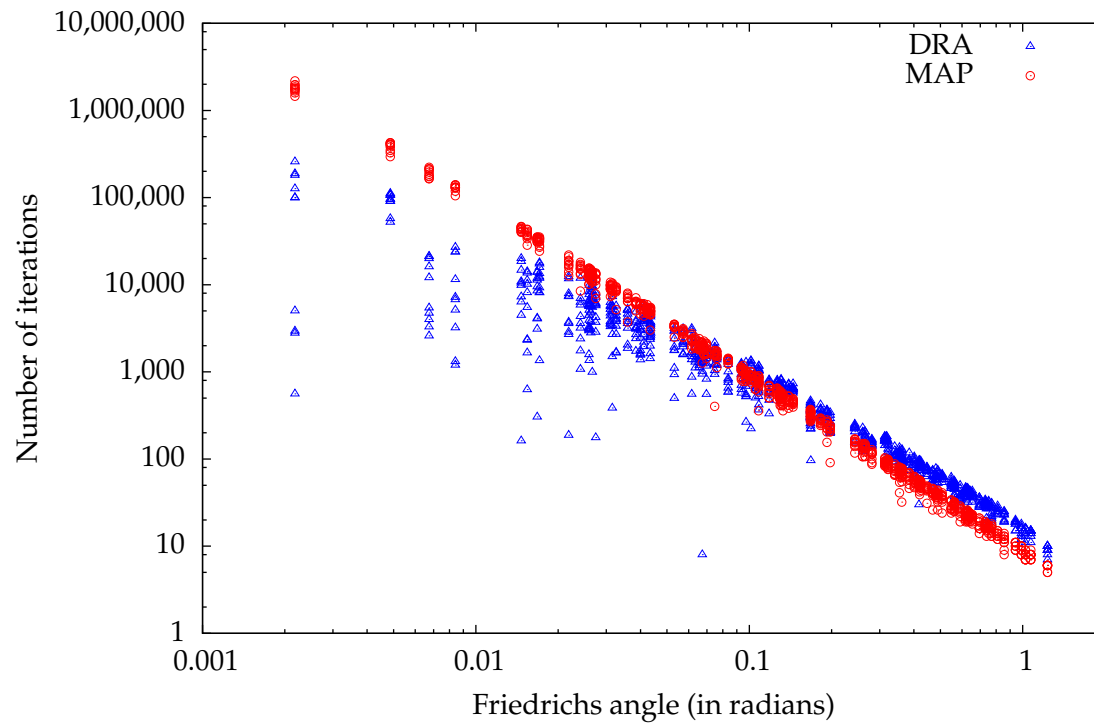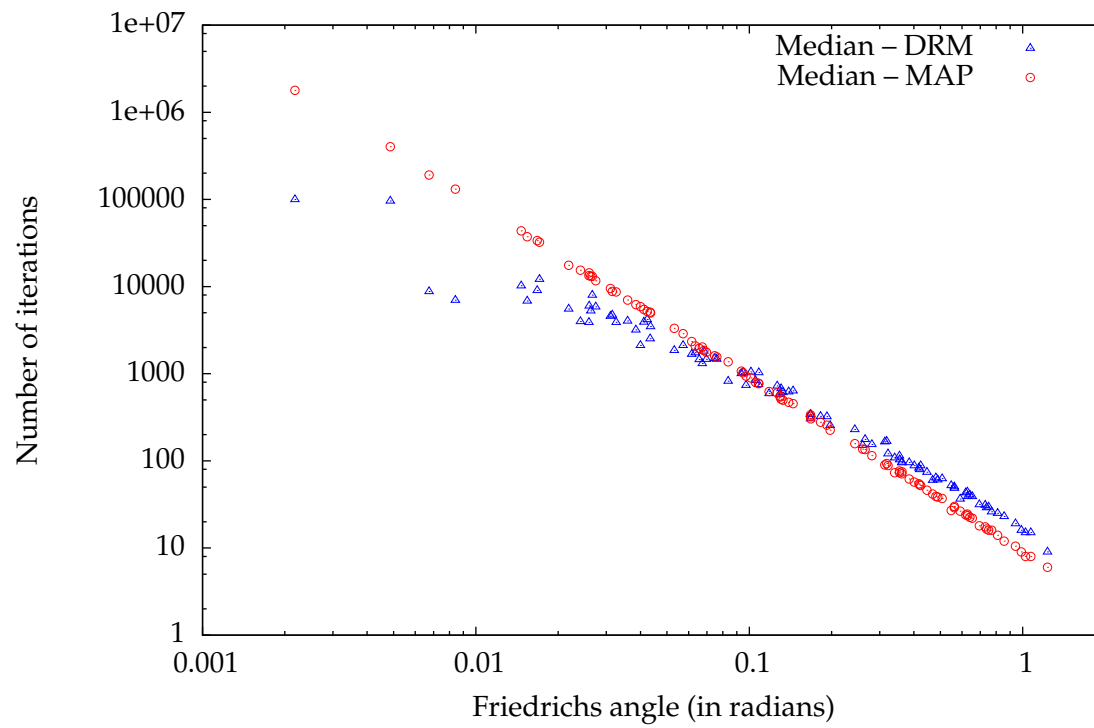
14

Figure 5: True error criterion

Figure 6: True error criterion

In Figure 5 and Figure 6, the horizontal axis represents the Friedrichs angle between the subspaces and the vertical axis represents the number of iterations. Results for all 1,000 runs are presented in Figure 5, while we show the *median* in Figure 6.

From the figures, we see that DRM is generally faster than MAP when the Friedrichs angle $\theta < 0.1$. In the opposite case, MAP is faster. This can be interpreted as follows. Since DRM converges with linear rate $c_F = \cos(\theta)$ while MAP does with rate c_F^2, we expect that MAP performs better when c_F is small, i.e., θ is large. But when the Friedrichs angle is small, the "rippling" behaviour of DRM appears to manifest itself (see also Figure 1). Note that MAP is not much faster than DRM, which suggests DRM as the better overall choice.

8.2 Stopping criterion based on individual distances

In practice, it is not always possible to obtain the true error. Thus, we utilized a reasonable alternative stopping criterion, namely when the monitored sequence $(z_n)_{n \in \mathbb{N}}$ satisfies

(57) $$\max \left\{ d_U(z_n), d_V(z_n) \right\} < \varepsilon$$
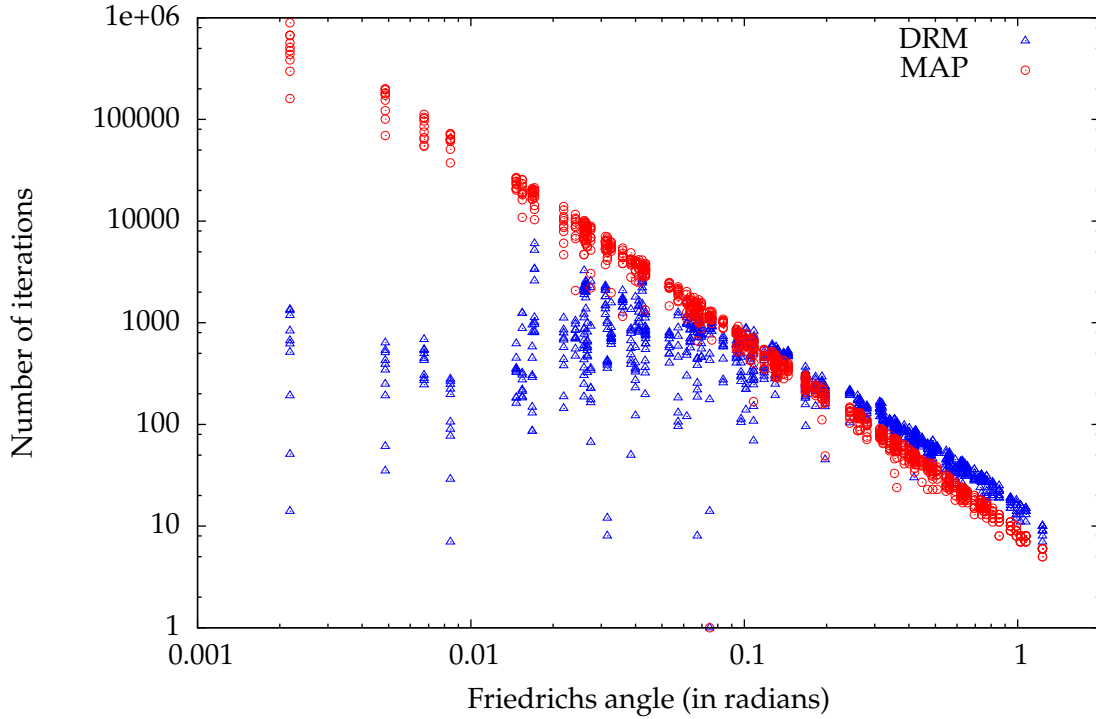
for the first time.

Figure 7: Max distance criterion

Figures 7 and 8 show the results when we use the max distance criterion with the same data. The behaviour is similar to the experiments with the true error criterion.

The figures in this section were computed with the help of `Julia` (see [28]) and `Gnuplot` (see [25]).
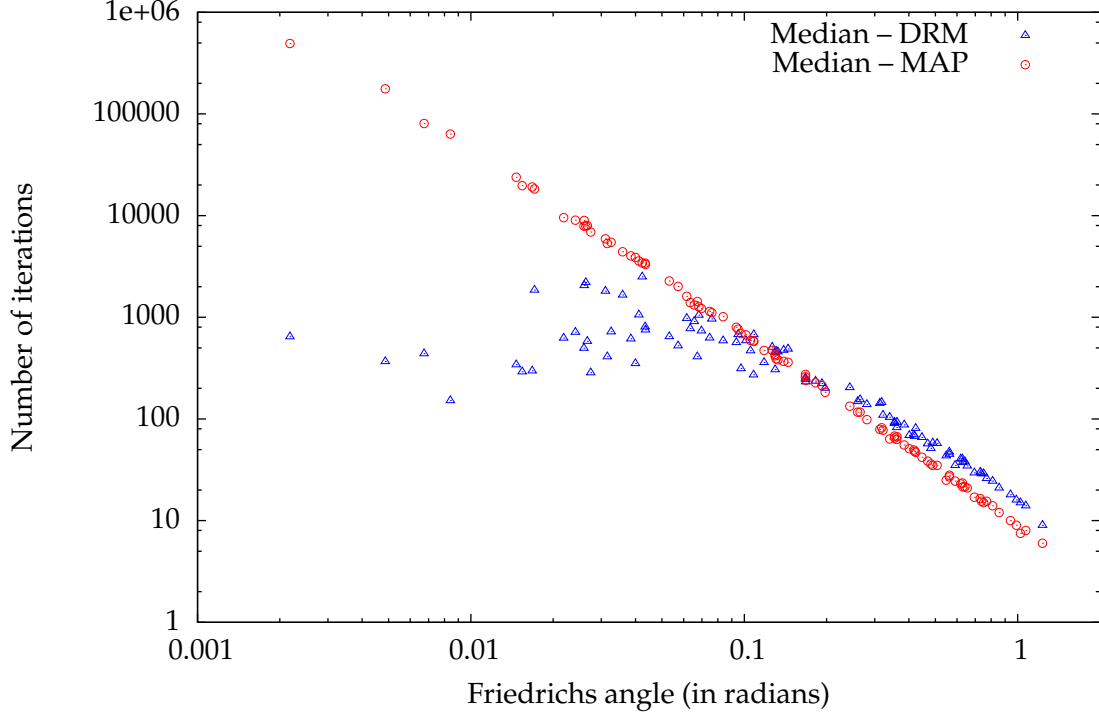
Figure 8: Max distance criterion

9 Conclusion

We completely analyzed the Douglas–Rachford splitting method for the important case of two subspaces. We determined the limit and the sharp rate of convergence. Lack of linear convergence was illustrated by an example in ℓ_2. Finally, we compared this method to the method of alternating projections and found the Douglas–Rachford method to be faster when the Friedrichs angle between the subspaces is small.

Acknowledgments

References

[1] F.J. Aragón Artacho, J.M. Borwein, and M.K. Tam, Recent results on Douglas-Rachford methods for combinatorial optimization problems, *Journal of Optimization Theory and Applications*, in press. DOI 10.1007/s10957-013-0488-0

[2] N. Aronszajn, Theory of reproducing kernels, *Transactions of the AMS* 68 (1950), 337–404.

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

[4] H.H. Bauschke, New demiclosedness principles for (firmly) nonexpansive operators, *Computational and Analytical Mathematics*, Chapter 2, Springer, 2013.

[5] H.H. Bauschke and J.M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Review* 38(3) (1996), 367–426.

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

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

[8] H.H. Bauschke, P.L. Combettes, and D.R. Luke, Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization, *Journal of the Optical Society of America* 19(7) (2002), 1334–1345.

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

[10] H.H. Bauschke, F. Deutsch, and H. Hundal, Characterizing arbitrarily slow convergence in the method of alternating projections, *International Transactions in Operational Research* 16 (2009), 413–425.

[11] H.H. Bauschke, F. Deutsch, H. Hundal, and S.-H. Park, Accelerating the convergence of the method of alternating projections, *Transactions of the AMS* 355(9) (2003), 3433–3461.

[12] H.H. Bauschke and V.R. Koch, Projection Methods: Swiss Army Knives for Solving Feasibility and Best Approximation Problems with Halfspaces, in *Infinite Products and Their Applications*, in press.

[13] J.M. Borwein and M.K. Tam, A cyclic Douglas–Rachford iteration scheme, *Journal of Optimization Theory and Applications* 160 (2014), 1–29.

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

[15] Y. Censor and S.A. Zenios, *Parallel Optimization*, Oxford University Press, 1997.

[16] P.L. Combettes, The foundations of set theoretic estimation, *Proceedings of the IEEE* 81(2) (1993), 182–208.

[17] P.L. Combettes, Iterative construction of the resolvent of a sum of maximal monotone operators, *Journal of Convex Analysis* 16(4) (2009), 727–748.

[18] P.L. Combettes and N.N. Reyes, Functions with prescribed best linear approximations, *Journal of Approximation Theory* 162(5) (2010), 1095–1116.

[19] L. Demanet and X. Zhang, Eventual linear convergence of the Douglas-Rachford iteration for basis pursuit, `http://arxiv.org/abs/1301.0542v3`, May 2013.

[20] F. Deutsch, The angle between subspaces of a Hilbert space, in *Approximation Theory, Wavelets and Applications*, S.P. Singh (editor), Kluwer, 1995, pp. 107–130.

[21] F. Deutsch, *Best Approximation in Inner Product Spaces*, Springer, 2001.

[22] J. Douglas and H.H. Rachford, On the numerical solution of heat conduction problems in two and three space variables, *Transactions of the AMS* 82 (1956), 421–439.

[23] 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.

[24] V. Elser, I. Rankenburg, and P. Thibault, Searching with iterated maps, *Proceedings of the National Academy of Sciences* 104(2) (2007), 418–423.

[25] GNU Plot, `http://sourceforge.net/projects/gnuplot`

[26] 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.

[27] R. Hesse, D.R. Luke, and P. Neumann, Alternating projections and Douglas-Rachford for sparse affine feasibility, `http://arxiv.org/abs/1307.2009v2`, March 2014.

[28] The Julia language, `http://julialang.org/`

[29] S. Kayalar and H. Weinert, Error bounds for the method of alternating projections, *Mathematics of Control, Signals, and Systems* 1 (1996), 43–59.

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

[31] Maple, `http://www.maplesoft.com/products/maple/`

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