

# THE CYCLIC DOUGLAS–RACHFORD METHOD FOR INCONSISTENT FEASIBILITY PROBLEMS

JONATHAN M. BORWEIN AND MATTHEW K. TAM

ABSTRACT. We analyse the behaviour of the newly introduced cyclic Douglas–Rachford algorithm for finding a point in the intersection of a finite number of closed convex sets. This work considers the case in which the target intersection set is possibly empty.

## 1. PRELIMINARIES AND NOTATION

Throughout we assume  $\mathcal{H}$  is a (real) *Hilbert space* with inner product  $\langle \cdot, \cdot \rangle$  and induced norm  $\| \cdot \|$ . We use  $\rightarrow$  (resp.  $\xrightarrow{w}$ ) to denote norm (resp. weak) convergence.

We consider the *convex feasibility problem*

$$(1.1) \quad \text{Find } x \in \bigcap_{i=1}^N C_i,$$

where  $C_i$  are closed convex subsets of  $\mathcal{H}$ . For convenience, we define  $C_0 := C_N$  and  $C_{N+1} := C_1$ .

When the intersection is empty, (1.1) is ill-posed. Instead, we seek an appropriate substitute for a point in the intersection. For example, if  $N = 2$  it is natural to consider the following variational problem

$$(1.2) \quad \inf_{(c_1, c_2) \in C_1 \times C_2} \|c_1 - c_2\|.$$

If the infimum is realised, we call the solution  $(c_1, c_2)$  a *best approximation pair* with respect to  $(C_1, C_2)$ . For  $N > 2$ , an appropriate generalization of (1.2) for characterizing the limit cycles of projection methods remains elusive and subtle. For details, see [2].

**1.1. General Theory.** In this section we recall some general theory regarding general nonexpansive mappings — applied later to projections and reflections — in Hilbert space. We give some definitions.

**Definition 1.1.** Let  $T : \mathcal{H} \rightarrow \mathcal{H}$ .

(a)  $T$  is *nonexpansive* if  $\|Tx - Ty\| \leq \|x - y\|$  for all  $x, y \in \mathcal{H}$ .

---

*Date:* March 7, 2014.

*2010 Mathematics Subject Classification.* 47H09, 47H10, 47N10.

*Key words and phrases.* cyclic Douglas–Rachford method, convex feasibility problem, projection, reflection.

(b)  $T$  is *firmly nonexpansive* if

$$\|Tx - Ty\|^2 + \|(I - T)x - (I - T)y\|^2 \leq \|x - y\|^2 \text{ for all } x, y \in \mathcal{H}.$$

(c)  $T$  is  $\lambda$ -*averaged* if  $T = (1 - \lambda)I + \lambda R$ , for some nonexpansive mapping  $R : \mathcal{H} \rightarrow \mathcal{H}$ .

(d)  $T$  is *demiclosed* if  $x_n \xrightarrow{w} x, Tx_n \rightarrow y \implies Tx = y$ .

(e)  $T$  is *asymptotically regular* at  $x \in \mathcal{H}$  if  $(I - T)T^n x \rightarrow 0$ .

(f) The set of *fixed points* of  $T$  is  $\text{Fix } T := \{x \in \mathcal{H} : Tx = x\}$ .

We collect facts concerning the interplay between these properties.

**Fact 1.2** (Nonexpansive properties). *The following hold.*

- (a) *If  $T$  is firmly nonexpansive then  $T$  is nonexpansive.*
- (b) *If  $\lambda \in [0, 1]$  and  $T$  is  $\lambda$ -averaged then  $T$  is nonexpansive.*
- (c)  *$T$  is firmly nonexpansive if and only if  $T$  is 1/2-averaged.*
- (d) *If  $T$  is nonexpansive then  $I - T$  is demiclosed.*
- (e) *If  $\lambda \in ]0, 1[$  and  $T$  is  $\lambda$ -averaged then, for any  $z \in \mathcal{H}$ ,*

$$\lim_{n \rightarrow \infty} \|T^{n+1}z - T^n z\| = \frac{1}{k} \lim_{n \rightarrow \infty} \|T^{n+k}z - T^n z\| = \lim_{n \rightarrow \infty} \frac{1}{n} \|T^n z\|,$$

for all  $k \geq 1$ . In particular, if  $(T^n z)_{n=1}^\infty$  is bounded, then  $T$  is asymptotically regular at  $x$ .

*Proof.* For (a)–(d) see, for example, [5, Ch. 4]. For (e), see [1, Th. 2.1].  $\square$

The following theorem will be useful in establishing convergence of our algorithms.

**Theorem 1.3** (Weak convergence of iterates). *Let  $\lambda \in ]0, 1[$ . Suppose  $(T_i)_{i=1}^m$  is a family of  $\lambda$ -averaged mappings from  $\mathcal{H}$  to  $\mathcal{H}$  such that  $\text{Fix}(T_m \dots T_1) \neq \emptyset$ . For any  $x_0 \in \mathcal{H}$  define*

$$x_{n+1} := (T_m \dots T_1)x_n.$$

*Then  $x_n - (T_m \dots T_1)x_n \rightarrow 0$  and there exist points  $y_1 \in \text{Fix}(T_m \dots T_1)$ ,  $y_2 \in \text{Fix}(T_1 T_m \dots T_2)$ ,  $\dots$ ,  $y_m \in \text{Fix}(T_{m-1} \dots T_1 T_m)$  such that*

$$\begin{aligned} x_n &\xrightarrow{w} y_1 = T_m y_m, \\ T_1 x_n &\xrightarrow{w} y_2 = T_1 y_1, \\ T_2 T_1 x_n &\xrightarrow{w} y_3 = T_2 y_2, \\ &\vdots \\ T_{m-2} \dots T_1 x_n &\xrightarrow{w} y_{m-1} = T_{m-2} y_{m-2}, \\ T_{m-1} T_{m-2} \dots T_1 x_n &\xrightarrow{w} y_m = T_{m-1} y_{m-1}. \end{aligned}$$

*Proof.* This is a special case of [5, Th. 5.22].  $\square$

**1.2. The Method of Cyclic Projections.** The (*nearest point*) projection onto a set  $C$  (if it exists) is the mapping  $P_C : \mathcal{H} \rightarrow C$  defined by

$$P_C x := \arg \min_{c \in C} \|c - x\|.$$

It is well known that if  $C$  is closed and convex,  $P_C$  is well defined (i.e., nearest points exist uniquely for all  $x \in \mathcal{H}$ ) (see, for example, [4, Prop. 2.1.2]). It has the variational characterization

$$P_C x = c \iff c \in C \text{ and } \langle x - P_C x, c - P_C x \rangle \leq 0 \text{ for all } c \in C.$$

For any  $y_0 \in \mathcal{H}$ , the *method of cyclic projections* can be described in terms of the iteration scheme

$$y_1^1 := P_{C_1} y_0, \quad y_n^{i+1} := P_{C_{i+1}} y_n^i, \quad y_{n+1}^1 := y_n^{N+1}.$$

We refer to the sequences  $(y_n^1)_{n=1}^\infty, (y_n^2)_{n=1}^\infty, \dots, (y_n^N)_{n=1}^\infty$  as the *cyclic projection sequences*.

Define

$$Q_i := P_{C_i} P_{C_{i-1}} \dots P_{C_1} P_{C_N} \dots P_{C_{i+1}}.$$

Note that, for each  $i$ , the sequence  $(y_n^i)_{n=1}^\infty$  is given by

$$y_{n+1}^i = Q_i y_n^i.$$

Suppose that each  $\text{Fix } Q_i$  is nonempty and let  $q^1 \in \text{Fix } Q_1$ . Define the sequence  $(q^i)_{i=1}^N$  by

$$q^{i+1} := P_{C_{i+1}} q^i \in \text{Fix } Q_{i+1}.$$

Define  $(d^i)_{i=1}^N$ , the sequence of *difference vectors*, by  $d^i := q^{i+1} - q^i$ . It can be shown that the difference vectors are well-defined (i.e., they are independent of the choice of  $q^1$ ). For further details see [4].

Recall the following dichotomy theorem.

**Theorem 1.4** (Cyclic projections dichotomy). *Exactly one of the following alternatives hold.*

- (a) *Each  $\text{Fix } Q_i$  is empty. Then  $\|y_n^i\| \rightarrow +\infty$ , for all  $i$ .*
- (b) *Each  $\text{Fix } Q_i$  is nonempty. Then, for each  $i$ ,  $(y_n^i)_{n=1}^\infty$  weakly converges to a point  $y^i$  such that  $y^{i+1} = P_{C_{i+1}} y^i$ , and the sequence  $(y_n^{i+1} - y_n^i)_{n=1}^\infty$  converges in norm to  $d^i$ .*

*Proof.* See [4, Th. 5.2.1]. □

**1.3. The Cyclic Douglas–Rachford Method.** The (*metric*) reflection with respect to a set  $C$  is the mapping  $R_C : \mathcal{H} \rightarrow \mathcal{H}$  given by

$$R_C := 2P_C - I,$$

where  $I$  denotes the *identity mapping*. If  $C$  is closed and convex,  $R_C$  is well defined. It has the variational characterization (see, for example, [8, Fac. 2.1])

$$R_C x = r \iff \frac{1}{2}(r + x) \in C \text{ and } \langle x - r, c - r \rangle \leq \frac{1}{2}\|x - r\|^2 \text{ for all } c \in C.$$

The *Douglas–Rachford operator* is the mapping  $T_{C_1, C_2} : \mathcal{H} \rightarrow \mathcal{H}$  given by

$$T_{C_1, C_2} := \frac{I + R_{C_2} R_{C_1}}{2}.$$

The *cyclic Douglas–Rachford operator* is the mapping  $T_{[C_1, C_2, \dots, C_N]} : \mathcal{H} \rightarrow \mathcal{H}$  given by

$$T_{[C_1, C_2, \dots, C_N]} := \prod_{i=1}^N T_{C_i, C_{i+1}},$$

where  $T_{N, N+1} := T_{N, 1}$ .

Where there is no ambiguity, we write  $T_{i, i+1}$  to mean  $T_{C_i, C_{i+1}}$ , and  $\sigma_i$  to mean the cyclic permutation of  $C_1, C_2, \dots, C_N$  beginning with  $C_i$ . Under this notation

$$T_{[\sigma_1]} := T_{[C_1, C_2, \dots, C_{N-1}, C_N]}, \quad T_{[\sigma_2]} := T_{[C_2, C_3, \dots, C_N, C_1]}, \quad \text{etc.}$$

For convenience, we define  $\sigma_0 := \sigma_N$  and  $\sigma_{N+1} := \sigma_1$ .

For any  $x_0 \in \mathcal{H}$ , the *cyclic Douglas–Rachford method* can be described in terms of the iteration scheme

$$(1.3) \quad x_1^1 := x_0, \quad x_n^{i+1} := T_{i, i+1} x_n^i, \quad x_{n+1}^1 := x_n^{N+1}.$$

We refer to the sequences  $(x_n^1)_{n=1}^\infty, (x_n^2)_{n=1}^\infty, \dots, (x_n^N)_{n=1}^\infty$  as the *cyclic Douglas–Rachford sequences*.

Note that, for each  $i$ , the sequence  $(x_n^i)_{n=1}^\infty$  is given by

$$(1.4) \quad x_{n+1}^i := T_{[\sigma_i]} x_n^i.$$

**Remark 1.5.** If  $z \in C_i$  then  $T_{i, i+1} z = P_{C_{i+1}} z$ . Hence, if  $x_0 = y_0 \in C_1$ , the cyclic projection and cyclic Douglas–Rachford sequences coincide. That is, for each  $i$ ,

$$y_n^i = x_n^i, \text{ for } n = 1, 2, 3, \dots$$

If  $x_0 \neq y_0$  and  $x_0 \notin C_1$ , it is entirely possible for the cyclic projection and cyclic Douglas–Rachford sequences to be distinct. For an example, see [8, Rem. 3.3].  $\diamond$

**Remark 1.6** (Alternating reflections). The classical Douglas–Rachford method, which applies to two sets problems, performs iterations by repeated application of a Douglas–Rachford operator, i.e.  $x_{n+1} := T_{1, 2}(x_n)$  for all  $n$  or  $x_{n+1} := T_{2, 1}(x_n)$  for all  $n$ . Thus, in the two sets case, the cyclic Douglas–Rachford method may be thought of as a traditional Douglas–Rachford algorithm in which the set chosen to be reflected on first is alternated.  $\diamond$

## 2. A DICHOTOMY THEOREM

We require a suite of seven preparatory lemmas.

**Lemma 2.1.** *For each index  $i$ ,*

- (a)  $T_{i, i+1}$  is  $1/2$ -averaged, and hence firmly nonexpansive.
- (b)  $T_{[\sigma_i]}$  is  $(1 - 2^{-N})$ -averaged.

*Proof.* (a) Since convex reflections are nonexpansive, it immediately follows that  $T_{i,i+1}$  is  $1/2$ -averaged. (b) Suppose that  $T$  is  $1/2$ -averaged and  $Q$  is  $(1 - 2^{-k})$ -averaged for some nonnegative integer  $k$ . We may write

$$T = \frac{1}{2}I + \frac{1}{2}R, \text{ and } Q = \frac{1}{2^k}I + \left(1 - \frac{1}{2^k}\right)S,$$

for nonexpansive mappings  $R$  and  $S$ . Observe

$$\begin{aligned} TQ &= \frac{1}{2}Q + \frac{1}{2}RQ = \frac{1}{2^{k+1}}I + \frac{2^k - 1}{2^{k+1}}S + \frac{2^k}{2^{k+1}}RQ \\ &= \frac{1}{2^{k+1}}I + \frac{2^{k+1} - 1}{2^{k+1}} \left( \frac{2^k - 1}{2^{k+1} - 1}S + \frac{2^k}{2^{k+1} - 1}RQ \right) \\ &= \frac{1}{2^{k+1}}I + \left(1 - \frac{1}{2^{k+1}}\right) \left( \frac{2^k - 1}{2^{k+1} - 1}S + \frac{2^k}{2^{k+1} - 1}RQ \right). \end{aligned}$$

Since  $S$  and  $RQ$  are nonexpansive, so is their convex combination, and hence  $TQ$  is  $(1 - 2^{k+1})$ -averaged. The equivalence now follows.  $\square$

The follow lemma shows that the cyclic Douglas–Rachford method has similar asymptotic behaviour to the method of cyclic projections. To exploit the nonexpansive properties of  $T_{[\sigma_i]}$  and  $T_{i,i+1}$ , we will sometimes choose  $y_0 := P_{C_1}x_0$ .

**Lemma 2.2.** *For any  $x_0 \in \mathcal{H}$ , choose  $y_0 := P_{C_1}x_0$ . As  $n \rightarrow \infty$ ,*

$$(x_n^i - x_n^{i+1}) - (y_n^i - y_n^{i+1}) \rightarrow 0,$$

for any index  $i$ .

*Proof.* By Remark 1.5, the method cyclic projection sequence starting at  $y_0$  can be consider as cyclic Douglas–Rachford sequence. Since  $T_{[\sigma_1]}$  is nonexpansive,

$$\|x_{n+1}^1 - y_{n+1}^1\| \leq \|x_n^1 - y_n^1\| \implies \lim_{n \rightarrow \infty} \|x_n^1 - y_n^1\| \text{ exists.}$$

Since  $T_{i,i+1}$  is firmly nonexpansive, for each  $i$ ,

$$\begin{aligned} \|x_n^1 - y_n^1\|^2 - \|x_{n+1}^1 - y_{n+1}^1\|^2 &= \sum_{i=1}^N (\|x_n^i - y_n^i\|^2 - \|x_n^{i+1} - y_n^{i+1}\|^2) \\ &\geq \sum_{i=1}^N \|(x_n^i - x_n^{i+1}) - (y_n^i - y_n^{i+1})\|^2. \end{aligned}$$

The result follows by taking the limit as  $n \rightarrow \infty$ .  $\square$

**Lemma 2.3.** *The sequence  $(x_n^j)_{n=1}^\infty$  is bounded if and only if  $\text{Fix } T_{[\sigma_j]}$  is nonempty.*

*Proof.* Suppose  $(x_n^j)_{n=1}^\infty$  is bounded. Then there exists a subsequence  $(x_{n_k}^j)_{k=1}^\infty$  weakly convergent to some point  $z$ . By Fact 1.2(e),  $(I - T_{[\sigma_j]})x_{n_k} \rightarrow 0$ . Since  $(I - T_{[\sigma_j]})$  is demiclosed,  $(I - T_{[\sigma_j]})z = 0 \implies z \in \text{Fix } T_{[\sigma_j]}$ .

Conversely, if  $z \in \text{Fix } T_{[\sigma_j]}$  nonexpansivity implies

$$\|z - x_n^j\| \leq \|z - x_1^j\| \implies \|x_n^j\| \leq \|z\| + \|z - x_1^j\|.$$

This completes the proof.  $\square$

**Lemma 2.4.** *The following four properties are equivalent.*

- (a) *The sequence  $(x_n^j)_{n=1}^\infty$  is bounded, for some index  $j$ .*
- (b) *The sequences  $(x_n^1)_{n=1}^\infty, (x_n^2)_{n=1}^\infty, \dots, (x_n^N)_{n=1}^\infty$  are bounded.*
- (c) *The sequence  $(y_n^j)_{n=1}^\infty$  is bounded, for some index  $j$ .*
- (d) *The sequences  $(y_n^1)_{n=1}^\infty, (y_n^2)_{n=1}^\infty, \dots, (y_n^N)_{n=1}^\infty$  are bounded.*

Furthermore, if  $(x_n^j)_{n=1}^\infty$  is unbounded then  $\|x_n^j\| \rightarrow +\infty$ .

*Proof.* Fix an index  $j$ . For any  $x_0 \in \mathcal{H}$ , choose  $y_0 := P_{C_1}x_0$ . Since  $T_{[\sigma_j]}$  is nonexpansive,

$$\|x_n^j - y_n^j\| \leq \|x_{n-1}^j - y_{n-1}^j\| \leq \dots \leq \|x_1^j - y_1^j\|.$$

By the triangle inequality

$$\|x_n^j\| \leq \|x_1^j - y_1^j\| + \|y_n^j\| \text{ and } \|y_n^j\| \leq \|x_n^j\| + \|x_1^j - y_1^j\|.$$

Thus  $(x_n^j)_{n=1}^\infty$  is bounded if and only if  $(y_n^j)_{n=1}^\infty$  is bounded. By Theorem 1.4,  $(y_n^j)_{n=1}^\infty$  is bounded if and only if  $(y_n^1)_{n=1}^\infty, (y_n^2)_{n=1}^\infty, \dots, (y_n^N)_{n=1}^\infty$ , and if  $(y_n^j)_{n=1}^\infty$  is unbounded then  $\|x_n^j\| \rightarrow +\infty$ . The result follows by combining these two statements.  $\square$

We observe

$$\begin{aligned} x_n^{i+1} &:= T_{i,i+1}x_n^i = \frac{x_n^i + R_{C_{i+1}}R_{C_i}x_n^i}{2} \\ &= \frac{x_n^i + 2P_{C_{i+1}}R_{C_i}x_n^i - R_{C_i}x_n^i}{2} \\ &= \frac{2x_n^i + 2P_{C_{i+1}}R_{C_i}x_n^i - 2P_{C_i}x_n^i}{2} \\ &= x_n^i + P_{C_{i+1}}R_{C_i}x_n^i - P_{C_i}x_n^i. \end{aligned} \tag{2.1}$$

$$\implies P_{C_{i+1}}R_{C_i}x_n^i = x_n^{i+1} - x_n^i + P_{C_i}x_n^i$$

**Lemma 2.5.** *For all  $i$  and for all  $n$ ,*

$$\|x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}\|^2 \leq \langle x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}, x_n^i - P_{C_i}x_n^i \rangle. \tag{2.2}$$

*In particular, for all  $i$  and all  $n$ ,*

$$\|x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}\| \leq \|x_n^i - P_{C_i}x_n^i\|. \tag{2.3}$$

*Proof.* By (2.1) and the variational characterization of convex projections,

$$\begin{aligned} & \|x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}\|^2 - \langle x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}, x_n^i - P_{C_i}x_n^i \rangle \\ &= \langle x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}, (x_n^{i+1} - x_n^i + P_{C_i}x_n^i) - P_{C_{i+1}}x_n^{i+1} \rangle \\ &= \langle x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}, P_{C_{i+1}}R_{C_i}x_n^i - P_{C_{i+1}}x_n^{i+1} \rangle \\ &\leq 0. \end{aligned}$$

This proves (2.2). Equation (2.3) now follows by an application of the Cauchy–Schwarz inequality.  $\square$

**Lemma 2.6.** *For all  $i$  and for all  $m$ ,*

$$(2.4) \quad \sum_{n=2}^m \sum_{i=1}^N \|(x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}) - (x_n^i - P_{C_i}x_n^i)\|^2 \\ \leq \langle x_2^1 - P_{C_1}x_2^1, x_1^N - P_{C_N}x_1^N \rangle - \langle x_{m+1}^1 - P_{C_1}x_{m+1}^1, x_m^N - P_{C_N}x_m^N \rangle,$$

where  $\langle x_2^1 - P_{C_1}x_2^1, x_1^N - P_{C_N}x_1^N \rangle$  and  $\langle x_{m+1}^1 - P_{C_1}x_{m+1}^1, x_m^N - P_{C_N}x_m^N \rangle$  are nonnegative. In particular, the double-sum in (2.4) is bounded, and hence, as  $n \rightarrow \infty$ ,

$$(x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}) - (x_n^i - P_{C_i}x_n^i) \rightarrow 0.$$

*Proof.* Applying Lemma 2.5,

$$\begin{aligned} & \sum_{n=2}^m \sum_{i=1}^N \|(x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}) - (x_n^i - P_{C_i}x_n^i)\|^2 \\ &= \sum_{n=2}^m \sum_{i=1}^N (\|x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}\|^2 - 2\langle x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}, x_n^i - P_{C_i}x_n^i \rangle \\ & \quad + \|x_n^i - P_{C_i}x_n^i\|^2) \\ &\leq \sum_{n=2}^m \sum_{i=1}^N (\langle x_n^i - P_{C_i}x_n^i, x_n^{i-1} - P_{C_{i-1}}x_n^{i-1} \rangle - \langle x_n^{i+1} - P_{C_{i+1}}x_n^{i+1}, x_n^i - P_{C_i}x_n^i \rangle). \\ &= \langle x_2^1 - P_{C_1}x_2^1, x_2^0 - P_{C_0}x_2^0 \rangle - \langle x_m^{N+1} - P_{C_{N+1}}x_m^{N+1}, x_m^N - P_{C_N}x_m^N \rangle \\ &= \langle x_2^1 - P_{C_1}x_2^1, x_1^N - P_{C_N}x_1^N \rangle - \langle x_{m+1}^1 - P_{C_1}x_{m+1}^1, x_m^N - P_{C_N}x_m^N \rangle. \end{aligned}$$

The nonnegativity of  $\langle x_2^1 - P_{C_1}x_2^1, x_1^N - P_{C_N}x_1^N \rangle$  and  $\langle x_{m+1}^1 - P_{C_1}x_{m+1}^1, x_m^N - P_{C_N}x_m^N \rangle$  is a consequence of (2.2).  $\square$

We now prove the analogue of Lemma 2.5 which will be applied to the limits (if they exist) of the cyclic Douglas–Rachford sequences. As before, we may deduce

$$(2.5) \quad P_{C_{i+1}}R_{C_i}x^i = x^{i+1} - x^i + P_{C_i}x^i.$$

**Lemma 2.7.** *Suppose  $(x^i)_{i=1}^N$  are points such that  $x^{i+1} = T_{i,i+1}x^i$ . Then, for all  $i$ ,*

$$P_{C_{i+1}}R_{C_i}x^i - P_{C_{i+1}}x^{i+1} = (x^{i+1} - x^i) - (P_{C_{i+1}}x^{i+1} - P_{C_i}x^i) = 0.$$

*Proof.* Consider the cyclic Douglas–Rachford sequences for initial point  $x_0 := x^1$ . Since  $x^i \in \text{Fix } T_{[\sigma_i]}$ , for each  $i$ , the result follows from (2.5) and Lemma 2.6.  $\square$

We are now ready to prove a dichotomy theorem which is the cyclic Douglas–Rachford method analogue of Theorem 1.4.

**Theorem 2.8** (Cyclic Douglas–Rachford dichotomy). *The following holds.*

(a) For each  $i$ ,

$$P_{C_{i+1}} R_{C_i} x_n^i - P_{C_{i+1}} x_n^i = (x_n^{i+1} - x_n^i) - (P_{C_{i+1}} x_n^{i+1} - P_{C_i} x_n^i) \rightarrow 0.$$

(b) Exactly one of the following alternatives hold.

(i) Each  $\text{Fix } T_{[\sigma_i]}$  is empty. Then  $\|x_n^i\| \rightarrow +\infty$ , for all  $i$ .

(ii) Each  $\text{Fix } T_{[\sigma_i]}$  is nonempty. Then, for each  $i$ ,

$$x_n^i \xrightarrow{w} x^i \in \text{Fix } T_{[\sigma_i]} \text{ with } x^{i+1} = T_{i,i+1} x^i.$$

Furthermore, for each  $i$ ,

$$\begin{aligned} x_n^{i+1} - x_n^i &= P_{C_{i+1}} R_{C_i} x_n^i - P_{C_i} x_n^i \rightarrow d^i, & P_{C_{i+1}} x_n^{i+1} - P_{C_i} x_n^i &\rightarrow d^i, \\ x^{i+1} - x^i &= P_{C_{i+1}} x^{i+1} - P_{C_i} x^i = d^i, & P_{C_{i+1}} R_{C_i} x^i &= P_{C_{i+1}} x^{i+1}. \end{aligned}$$

*Proof.* (a) follows by Lemma 2.6. (b) By appealing to Lemmas 2.3 and 2.4, we establish the two possible alternatives: either  $\text{Fix } T_{[\sigma_i]} = \emptyset$  and  $\|x_n^i\| \rightarrow +\infty$  for all  $i$ , or  $\text{Fix } T_{[\sigma_i]} \neq \emptyset$  for all  $i$ .

If each  $\text{Fix } T_{[\sigma_i]} \neq \emptyset$ , Lemma 2.1 together with Theorem 1.3 imply that the sequence  $(x_n^i)_{n=1}^\infty$  converges weakly to a point  $x^i \in \text{Fix } T_{[\sigma_i]}$  with  $x^{i+1} = T_{i,i+1} x^i$ . Lemma 2.2 with Theorem 1.4 implies  $x_n^{i+1} - x_n^i \rightarrow d^i$ , which together with (a) implies  $P_{C_{i+1}} x_n^{i+1} - P_{C_i} x_n^i \rightarrow d^i$ .

Lemma 2.7 together with Theorem 1.4 applied to the cyclic Douglas–Rachford sequences having initial point  $x^1$  yields  $x^{i+1} - x^i = P_{C_{i+1}} x^{i+1} - P_{C_i} x^i = d^i$  and  $P_{C_{i+1}} R_{C_i} x^i = P_{C_{i+1}} x^{i+1}$ .  $\square$

If  $\bigcap_{i=1}^N C_i \neq \emptyset$ , it can be shown that the limits  $(x^i)_{i=1}^N$  coincide (see, for example, [8, Lem. 2.3]). In this case, we obtain [8, Th. 3.1] as a special case of Theorem 2.8. That is, we have the following corollary.

**Corollary 2.9** (Consistent cyclic Douglas–Rachford iterations). *Suppose  $\bigcap_{i=1}^N C_i \neq \emptyset$ . Then the cyclic Douglas–Rachford sequences weakly converge to a common point  $x$  such that  $P_{C_i} x = P_{C_j} x$  for all indices  $i, j$ . In particular,  $P_{C_j} \in \bigcap_{i=1}^N C_i$  for any index  $j$ .*

**Remark 2.10.** The proof of Corollary 2.9 given in [8] for the consistent case is dependent on the fact that  $\text{Fix } T_{[\sigma_1]} = \bigcap_{i=1}^N \text{Fix } T_{i,i+1} \neq \emptyset$ . Since  $\text{Fix } T_{i,i+1} \neq \emptyset$  if and only if  $C_i \cap C_{i+1} \neq \emptyset$ , in the inconsistent case one can only guaranteed that  $\text{Fix } T_{[\sigma_1]} \supseteq \bigcap_{i=1}^N \text{Fix } T_{i,i+1} = \emptyset$ .  $\diamond$

**Remark 2.11** (Approximating the difference vectors). In Theorem 2.8, it was shown that the sequences

$$(x_n^{i+1} - x_n^i)_{n=1}^\infty, \quad (P_{C_{i+1}}R_{C_i}x_n^i - P_{C_i}x_n^i)_{n=1}^\infty, \quad (P_{C_{i+1}}x_n^{i+1} - P_{C_i}x_n^i)_{n=1}^\infty,$$

converge (in norm) to  $d^i$ . The latter two are suitable if one is interested in approximating  $d^i$  using a pair of points from  $C_i$  and  $C_{i+1}$ .  $\diamond$

**Remark 2.12** (Cyclic Douglas–Rachford as a favourable compromise). The behaviour of the cyclic Douglas–Rachford scheme is somewhere between that of the method of alternating projections and the classical Douglas–Rachford scheme. In this sense, it can be considered a compromise between the two schemes having some of the desirable properties of both. We elaborate.

Firstly, the cyclic Douglas–Rachford and classical Douglas–Rachford scheme perform the reflections with respect to the constraint sets, rather than using just a projection, as is the case of the method of cyclic projections. This can be seen as an advantage (at least heuristically). If a point is not contained in a particular constraint set, the reflection can potentially yield a strictly feasible point, whereas projections produce a point on the boundary (see Figure 1).

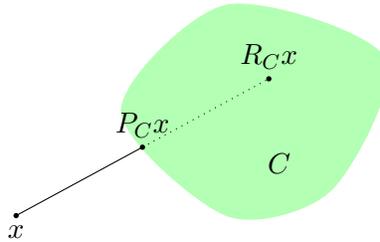


FIGURE 1.  $R_C x$  is strictly feasible, while  $P_C x$  is on the boundary of  $C$ .

On one hand, the cyclic Douglas–Rachford and classical Douglas–Rachford iterations both proceed by applying a two set Douglas–Rachford mapping (i.e. one of the form  $T_{i,i+1}$ ). In the consistent case, the limit obtained by both these schemes, once projected onto an appropriate constraint set, produces a solution to a feasibility problem. On the other, in the inconsistent case, the Douglas–Rachford scheme iterates are always unbounded. The behaviour described in Theorem 2.8 is much closer to that of the method of cyclic projections, described in Theorem 1.4.

Thus, if one wishes to diagnose infeasibility one might prefer Douglas–Rachford to the cyclic variant, but if one desires an estimate even in the infeasible case one would likely opt for the cyclic variant. The behaviour of the three methods is illustrated in the two possible two sets cases in Figure 2.  $\diamond$

### 3. THE TWO SET CASE

We now specialize the results of Section 2 for the case of problems having only two sets. Here the geometry of the problem is both better understood and more tractable. We introduce the following two sets

$$E := \{e \in C_1 : d(e, C_1) = d(C_1, C_2)\}, \quad F := \{f \in C_2 : d(f, C_1) = d(C_1, C_2)\}.$$

Further, the *displacement vector*,  $v$ , is defined by

$$v := P_{\overline{C_2 - C_1}}(0) = -P_{\overline{C_1 - C_2}}(0).$$

We recall some useful facts.

**Fact 3.1** (Properties of  $E, F$  and  $v$ ). *The following hold.*

- (i) *If  $C_1 \cap C_2 \neq \emptyset$  then  $C_1 \cap C_2 = E = F$  and  $v = 0$ .*
- (ii)  *$E + v = F$ , and  $\|v\| = d(C_1, C_2)$  if and only if  $C_2 - C_1$  is closed.*
- (iii)  *$E = \text{Fix } Q_1 = \text{Fix}(P_{C_1}P_{C_2})$  and  $F = \text{Fix } Q_2 = \text{Fix}(P_{C_2}P_{C_1})$ .*
- (iv)  *$v = d^1 = -d^2$ .*

*Proof.* See, for example, [3, Sec. 1] and [4, Fac. 5.2.2]. □

We are now ready to specialize the conclusions of Theorem 2.8. In particular, we show that the cyclic Douglas–Rachford scheme can be used to find best approximation pairs, provided they exist.

**Theorem 3.2** (Alternating Douglas–Rachford dichotomy). *The following holds.*

(a) *We have*

$$\begin{aligned} P_{C_2}R_{C_1}x_n^1 - P_{C_2}x_n^1 &= (x_n^2 - x_n^1) - (P_{C_2}x_n^2 - P_{C_1}x_n^1) \rightarrow 0, \\ P_{C_1}R_{C_2}x_n^2 - P_{C_1}x_n^2 &= (x_n^1 - x_n^2) - (P_{C_1}x_n^1 - P_{C_2}x_n^2) \rightarrow 0. \end{aligned}$$

(b) *Exactly one of the following alternatives hold.*

- (i)  *$E, F, \text{Fix } T_{[C_1, C_2]}, \text{Fix } T_{[C_2, C_1]} = \emptyset$ . Then  $\|x_n^1\|, \|x_n^2\| \rightarrow +\infty$ .*
- (ii)  *$E, F, \text{Fix } T_{[C_1, C_2]}, \text{Fix } T_{[C_2, C_1]} \neq \emptyset$ . Then*

$$x_n^1 \xrightarrow{w_i} x^1 \in \text{Fix } T_{[C_1, C_2]}, \quad x_n^2 \xrightarrow{w_i} x^2 \in \text{Fix } T_{[C_2, C_1]},$$

*with  $x^2 = T_{C_1, C_2}x^1$  and  $x^1 = T_{C_2, C_1}x^2$ . Furthermore,*

$$\begin{aligned} x_n^2 - x_n^1 &= P_{C_2}R_{C_1}x_n^1 - P_{C_1}x_n^1 \rightarrow v, & P_{C_2}x_n^2 - P_{C_1}x_n^1 &\rightarrow v, \\ x_{n+1}^1 - x_n^2 &= P_{C_1}R_{C_2}x_n^2 - P_{C_2}x_n^2 \rightarrow -v, & P_{C_1}x_{n+1}^1 - P_{C_2}x_n^2 &\rightarrow -v, \end{aligned}$$

*and  $x^2 - x^1 = P_{C_2}x^2 - P_{C_1}x^1 = v$ . In particular,*

$$P_{C_1}R_{C_2}x^2 = P_{C_1}x^1 \in E, \quad P_{C_2}R_{C_1}x^1 = P_{C_2}x^2 \in F.$$

*Proof.* Follows from Theorem 2.8 and Fact 3.1. □

Contrast Theorem 3.2 with its analogues for cyclic projections (Theorem 1.4) and for the classical Douglas–Rachford scheme (Theorem 3.3), which we state below for completeness.

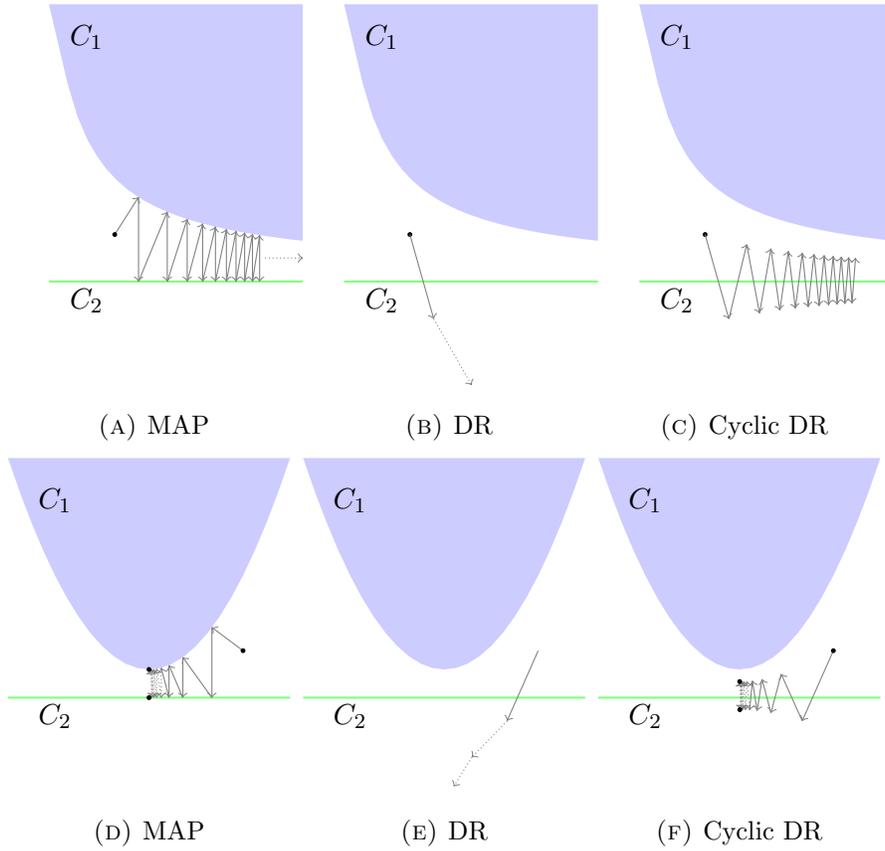


FIGURE 2. Behaviour of the three methods starting with the same initial point. In (A)–(F),  $C_2 := \mathbb{R} \times \{0\}$ . In (A)–(C),  $C_1 := \text{epi}(1 + 1/\cdot) \cap (\mathbb{R}_+ \times \mathbb{R}_+)$  and  $E, F$  are empty. In (D)–(F),  $C_1 := \text{epi}(1 + (\cdot)^2)$  and  $E, F$  are nonempty.

**Theorem 3.3** (Douglas–Rachford method dichotomy). *Let  $C_1, C_2 \subseteq \mathcal{H}$  be closed and convex. Let  $z_0 \in \mathcal{H}$  and set  $z_{n+1} := T_{C_1, C_2} z_n$ . Then*

- (a)  $z_{n+1} - z_n = P_{C_2} R_{C_1} z_n - P_{C_1} z_n \rightarrow v$  and  $P_{C_2} P_{C_1} z_n - P_{C_1} z_n \rightarrow v$ .  
 (b) Exactly one of the following alternatives holds.

(i)  $C_1 \cap C_2 \neq \emptyset$  and  $(z_n)_{n=1}^\infty$  converges weakly to a point in

$$\text{Fix } T_{C_1, C_2} = (C_1 \cap C_2) + N_{\overline{C_1 - C_2}}(0).$$

(ii)  $C_1 \cap C_2 = \emptyset$  and  $\|z_n\| \rightarrow +\infty$ .

- (c) Exactly one of the following two alternatives holds.

(i)  $E, F = \emptyset$ ,  $\|P_{C_1} z_n\| \rightarrow +\infty$  and  $\|P_{C_2} P_{C_1} z_n\| \rightarrow +\infty$ ;

(ii)  $E, F \neq \emptyset$ ,  $(P_{C_1} z_n)_{n=1}^\infty$  and  $(P_{C_2} P_{C_1} z_n)_{n=1}^\infty$  are bounded with weak cluster points in  $E$  and  $F$ , respectively. Furthermore, the weak cluster points of

$$((P_{C_1} z_n, P_{C_2} R_{C_1} z_n))_{n=1}^\infty \text{ and } ((P_{C_1} z_n, P_{C_2} P_{C_1} z_n))_{n=1}^\infty$$

are best approximation pairs relative to  $(C_1, C_2)$ .

*Proof.* See [6, Th. 3.13]. □

#### 4. FINAL REMARKS

We have analysed the behaviour of the cyclic Douglas–Rachford algorithm for finding a point in the intersection of a finite number of closed convex sets. Whilst each iteration of the the method is similar to that of the classical Douglas–Rachford scheme, its behaviour, particularly in the inconsistent case, is closer the that of the method of cyclic projections. With this in mind, one might consider the cyclic Douglas–Rachford scheme as a useful comprise between these methods.

Applied to two-set feasibility problems for which best approximation pairs exist, the cyclic Douglas–Rachford method produces a pair of points which when projected onto the appropriate sets yields a best approximation pair. This is important for applications in which consistency of the feasibility problem is not known *a priori*.

Finally we finish with two open questions:

- (1) Can one prove a version of the main result in [2], showing that for cyclic Douglas-Rachford applied to three sets there is no variational characterization of the fixed point sets  $\text{Fix } T_{[\sigma_i]}$  being nonempty?
- (2) What can be said about convergence rates for the Douglas-Rachford methods? Much less seems known than in the case of alternating projections. Recent linear convergence results for the classical Douglas–Rachford method applied to affine subspaces can be found in [7].

**Acknowledgements.** The authors would like to thank Liangjin Yao for his helpful suggestions.

#### REFERENCES

- [1] J.B. Baillon, R.E. Bruck and S. Reich. On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces. *Houst. J. Math.*, 4(1):1–9 (1978).
- [2] J.B. Baillon, P.L. Combettes and R. Cominetti. There is no variational characterization of the cycles in the method of periodic projections. *J. Func. Anal.*, 262(1):400–408 (2012).
- [3] H.H. Bauschke and J.M. Borwein. On the convergence of von Neumann’s alternating projection algorithm for two sets. *Set-Valued Anal.*, 1:185–212 (1993).
- [4] H.H. Bauschke, J.M. Borwein and A.S. Lewis. The method of cyclic projections for closed convex sets in Hilbert spaces. *Contemp. Math.*, 204:1–38 (1997).
- [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. *J. Approx. Theory*, 127:178–192 (2004).
- [7] H.H. Bauschke, J.Y. Cruz, 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. arXiv preprint [arXiv:1309.4709](https://arxiv.org/abs/1309.4709) (2013).

- [8] J.M. Borwein and M.K. Tam. A cyclic Douglas–Rachford iteration scheme. *J. Optim. Theory Appl.*, 160:1–29 (2014). doi: [10.1007/s10957-013-0381-x](https://doi.org/10.1007/s10957-013-0381-x).

(J.M. Borwein) CARMA CENTRE, UNIVERSITY OF NEWCASTLE, CALLAGHAN, NSW 2308, AUSTRALIA.

*E-mail address:* `jon.borwein@gmail.com`

(M.K. Tam) CARMA CENTRE, UNIVERSITY OF NEWCASTLE, CALLAGHAN, NSW 2308, AUSTRALIA.

*E-mail address:* `matthew.k.tam@gmail.com`