Global Behavior of the Douglas–Rachford Method for a Nonconvex Feasibility Problem

Francisco J. Aragón Artacho* Jonathan M. Borwein† Matthew K. Tam‡

June 22, 2015

Abstract

In recent times the Douglas–Rachford algorithm has been observed empirically to solve a variety of nonconvex feasibility problems including those of a combinatorial nature. For many of these problems current theory is not sufficient to explain this observed success and is mainly concerned with questions of local convergence. In this paper we analyze global behavior of the method for finding a point in the intersection of a half-space and a potentially non-convex set which is assumed to satisfy a property weaker than compactness. The special case in which the second set is finite provides a prototypical setting for combinatorial optimization problems.

Keywords Douglas–Rachford algorithm, global convergence, feasibility problem, half-space, non-convex

MSC2010: 90C26, 65K05

1 Introduction

Recent computational experiments have demonstrated the ability of *Douglas–Rachford methods* to successfully solve a variety of significant non-convex optimization problems. Examples include combinatorial optimization [3, 1, 13], low-rank matrix reconstruction [4], boolean satisfiability [16], sphere packing [16, 13], matrix completion [1], image reconstruction [10], and road design [8]. Since this success in non-convex settings is not uniform, the underlying theory is in need of significant enhancement.

Douglas–Rachford methods belong to the family of so called *projection algorithms* and are traditionally analyzed using nonexpansivity properties when the problem is convex. For non-convex problems theory is still developing. One approach to proving convergence is to replace assumptions of convexity with regularity properties not tantamount to convexity [14, 18, 9, 15]. Examples of the properties required include: *superregularity*, *strong regularity of sets*, (ϵ, δ)*-regularity*, and exploiting convex structure (*i.e.,* realizing a non-convex constraint as the union of certain convex sets). Being local in nature, these assumptions lead to local results; that is, they hold in a sufficiently small neighbourhood of a solution. An alternate, albeit less general, approach for proving convergence is to

*Department of Statistics and Operations Research, University of Alicante, Spain. Email: francisco.aragon@ua.es

†CARMA Centre, University of Newcastle, Australia. Email: jon.borwein@gmail.com

‡CARMA Centre, University of Newcastle, Australia. Email: matthew.tam@uon.edu.au

instead focus on specific non-convex feasibility problems. In this direction, global convergence of the method has been established for the two set feasibility problem involving a line (or hyperplane) and sphere [12, 2, 11].

In the non-convex feasibility problems for which Douglas–Rachford methods have been successful, it is the methods' apparent global convergence properties which deserve greater attention, and hence on which we shall focus. This is the case, for instance, in problems having discrete or combinatorial constraints where local convergence can be deduced from the general theory of convex sets as a consequence of the *local convexity* of the constraints. Furthermore, an implementation of a Douglas–Rachford method would always *round* the current iterate, and check if this rounded iterate is a solution to the problem. That is, in practice, the algorithm is never *run locally*.

The purpose of this paper is to analyze global properties of the basic Douglas–Rachford method. Of the two approaches already discussed, ours bears more resemblance to the latter. We focus on a two set feasibility problem with the first set allowed to be quite general — it must satisfy an assumption which encompasses all compact sets. The trade-off for this assumption is that a more constrained structure is required of the second set — it must be a closed half-space. The fact that this analysis is far from immediate, illustrates the subtles of the Douglas–Rachford method's global behavior, even for typical problems.

The remainder of this paper is organized as follows. In Section 2 we introduce the problem to be solved and the basic Douglas–Rachford algorithm. In Section 3 we study properties of the Douglas–Rachford operator. In Section 4 we analyze the global behavior of the method and give our main results. In Section 5 we give various examples and counter-examples. Finally, in Section 6, we make our concluding remarks.

2 Preliminaries

Through this paper our setting is the real finite dimensional Hilbert space $\mathbb{R}^m$ equipped with inner-product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. We consider the *feasibility problem*

$$\text{Find } x \in H \cap Q, \tag{1}$$

where $Q \subset \mathbb{R}^m$ is a closed set, and $H \subset \mathbb{R}^m$ is a (closed) half-space. We will be concerned with the case in which Q has additional properties (see Assumptions 4.1 & 4.2), but is intended to be as general as possible – these assumptions are explicitly stated where needed. It is convenient to represent H, and its *dividing hyperplane* L, in the form

$$H := \{x \in \mathbb{R}^m \mid \langle a, x \rangle \leq b\}, \qquad L := \{x \in \mathbb{R}^m \mid \langle a, x \rangle = b\},$$

where $b \in \mathbb{R}$, and $a \in \mathbb{R}^m$ with $\|a\| = 1$.

To solve (1) we employ the specific *Douglas–Rachford operator*, denoted $\mathrm{DR} : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$, given by

$$\mathrm{DR}(x) := \frac{x + R_H(R_Q(x))}{2}. \tag{2}$$

Here R_A denotes the *reflector* of a point x with respect to the set $A \subset \mathbb{R}^m$ given by $R_A(x) := 2P_A(x) - x$, and P_A stands for the *projector* of x onto A given by

$$P_A(x) := \left\{ z \in A \;\middle|\; \|x - z\| = \inf_{a \in A} \|x - a\| \right\}.$$

When x is a *fixed point* of the Douglas–Rachford operator (*i.e.,* $x \in \mathrm{DR}(x)$) there is an element of $P_Q(x)$ which solves (1) as is easy to confirm. This suggests that iterating the Douglas–Rachford operator to find a fixed point is a potential method for solving (1).

Given an initial point $x_0 \in \mathbb{R}^m$, we say the sequence $\{x_k\}_{k\in\mathbb{N}}$ is a *Douglas–Rachford iteration* if

$$x_{k+1} \in \mathrm{DR}(x_k) \quad \text{for} \quad k \in \mathbb{N}.$$

We now make explicit the precise form of the Douglas–Rachford operator for our feasibility problem in (1). The projections onto L and H are given by

$$P_L(x) = x - (\langle a, x\rangle - b)a, \quad P_H(x) = \begin{cases} x & \text{if } \langle a, x\rangle \le b, \\ x - (\langle a, x\rangle - b)a & \text{if } \langle a, x\rangle > b. \end{cases}$$

We assume P_Q is effectively computable. The corresponding reflections are given by

$$R_L(x) = x - 2(\langle a, x\rangle - b)a, \quad R_H(x) = \begin{cases} x & \text{if } \langle a, x\rangle \le b, \\ x - 2(\langle a, x\rangle - b)a & \text{if } \langle a, x\rangle > b, \end{cases}$$

and $R_Q(x) = 2P_Q(x) - x$. The Douglas–Rachford operator (2) may therefore be expressed as

$$\mathrm{DR}(x) = \bigcup_{q\in P_Q(x)} \mathrm{DR}(x, q),$$

where

$$\mathrm{DR}(x, q) := \begin{cases} q & \text{if } \langle a, 2q - x\rangle \le b, \\ q + (\langle a, x\rangle + b - 2\langle a, q\rangle)a & \text{if } \langle a, 2q - x\rangle > b. \end{cases} \tag{3}$$

Under this notation, a sequence $\{x_k\}_{k\in\mathbb{N}}$ is a Douglas–Rachford iteration if

$$x_{k+1} = \mathrm{DR}(x_k, q_k) \quad \text{for some} \quad q_k \in P_Q(x_k).$$

We shall refer to the sequence $\{q_k\}_{k\in\mathbb{N}}$ as an *auxiliary sequence* for $\{x_k\}$.

An example of a Douglas–Rachford iteration in the plane when Q is a set of four points is given in Figure 1. The iteration converges to a solution of (1) in eight steps. This behavior is explained by Theorem 4.2.

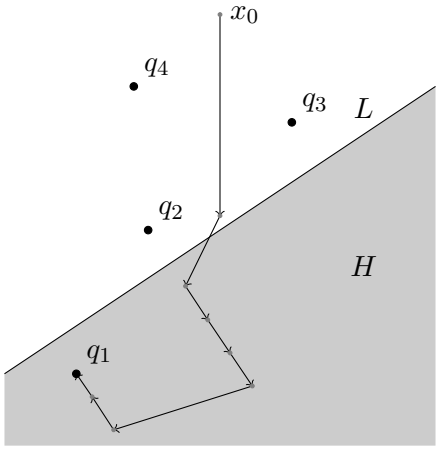


Figure 1: A Douglas–Rachford iteration in $\mathbb{R}^2$ with $Q = \{q_1, q_2, q_3, q_4\}$.

Remark 2.1. In practice, an implementation of the Douglas–Rachford iteration would hope to terminate as soon as $q_{k_0} \in Q \cap H$ for some $k_0 \in \mathbb{N}$. Observe, however, that (3)

does not necessarily ensure that the Douglas–Rachford sequence or its auxiliary sequence remain constant for $k \geq k_0$. It is therefore important to distinguish the Douglas–Rachford iteration from an algorithm arising from its implementation (see Algorithm 1). ◊

It is worth emphasising that, unlike the Douglas–Rachford algorithm, other projection algorithms can fail when applied to (1). An example is given in Example 2.1. The simplest method from this family is the *alternating projection algorithm*. It iterates by alternatively applying projectors onto Q and H. Precisely, given an initial point $x_0 \in \mathbb{R}^m$, it generates a sequence $\{x_k\}$ given by $x_{k+1} \in P_H(P_Q(x_k))$.

Example 2.1 (Failure of alternating projections). In general, von Neumann's alternating projection is unable to find a point in the intersection of H and Q (and hence the same is true for the *cyclic Douglas–Rachford algorithm* [5, 6]). Figure 2 shows a simple example with a doubleton $Q = \{q_1, q_2\} \subset \mathbb{R}^2$. In this example, $P_Q(P_H(q_1)) = q_1$ and the algorithm cycles between q_1 and $P_H(q_1)$ for any starting point $x_0 \in P_Q^{-1}(q_1)$. ◊

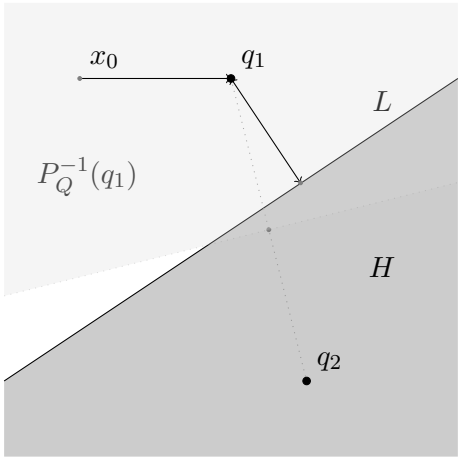


Figure 2: Failure of the alternating projection algorithm for initial points in $P_Q^{-1}(q_1)$.

We next give some examples of feasibility problems of the form in (1) to which our results apply.

Example 2.2 (Finite union of compact convex sets). Suppose Q is a finite union of compact convex sets in $\mathbb{R}^m$. Applied to this problem, [9, Theorem 1] yields a local result that applies near *strong fixed-points* of $\mathrm{DR}(\cdot)$ (these are points for which $\mathrm{DR}(x) = \{x\}$). Namely, that the Douglas–Rachford sequence locally converges to a *fixed-point* and any cluster of its auxiliary sequence solves (1). Note that the existence of a strong-ixed points, x, impiles $P_Q(x) \cap H \neq \emptyset$ (see [9]).

Theorem 4.3 complements this result by showing that, whenever $Q \cap H \neq \emptyset$, the auxiliary sequence always has at least one cluster point, regardless of the behavior of the Douglas–Rachford sequence. ◊

Example 2.3 (Knapsack lower bound feasibility). The classical *0-1 knapsack problem* is the binary program

$$\min \{\langle c, x\rangle \mid x \in \{0,1\}^n,\ \langle a, x\rangle \leq b\},$$

for non-negative vectors $a, c \in \mathbb{R}^n_+$ and a non-negative real number $b \in \mathbb{R}_+$.

The *0-1 knapsack lower-bound feasibility problem* is a case of (1) with constraints

$$H := \{x \in \mathbb{R}^n \mid \langle a, x\rangle \leq b\}, \qquad Q := \{x \in \{0,1\}^n \mid \langle c, x\rangle \geq \lambda\},$$

where $\lambda \in \mathbb{R}_+$. As a decision problem it is NP-complete [17, I.5 Corollary 6.11].

Applied to this problem, Theorem 4.1 shows that the Douglas–Rachford method either finds a solution in finitely many iterations, or none exists and the norm of the Douglas–Rachford sequence diverges to infinity. Note that, in general, P_Q usually cannot be computed efficiently. ◊

Example 2.4. Consider the problem of minimizing a linear function $f = \langle a, \cdot\rangle$ over a compact set Q for which P_Q is effectively computable or approximable. Problem (1) is a useful relaxation of this minimization when an upper-bound b on the optimal value is known. Theorem 4.3 then applies. ◊

Example 2.5 (A sphere and a half-space). Application of the Douglas–Rachford method to the problem of finding a point in the intersection of a sphere and an affine line (more generally an affine subspace) was originally investigated in [12, 2], with global convergence later proven using a Lyapunov stability argument in [11]. Here we consider the case in which the affine line is replaced with a half-space.

Let Q be the unit sphere and H a half-space in $\mathbb{R}^2$. By symmetry, we may assume $a = (0, 1)$. Let $x_0 \neq 0$ with $x_0(1) > 0$. Then $x_k(1) > 0$ and $q_k = \frac{x_k}{\|x_k\|}$ for all $k \in \mathbb{N}$, and the iteration given by (3) becomes

$$x_{k+1}(1) = \frac{x_k(1)}{\|x_k\|}, \quad x_{k+1}(2) = \begin{cases} \frac{x_k(2)}{\|x_k\|} & \text{if } \left(\frac{2}{\|x_k\|} - 1\right) x_k(2) \leq b, \\ \left(1 - \frac{1}{\|x_k\|}\right) x_k(2) + b & \text{if } \left(\frac{2}{\|x_k\|} - 1\right) x_k(2) > b. \end{cases}$$

If $Q \cap H \neq \emptyset$ (or equivalently $b \geq -1$) then Theorem 4.3 below ensures $d(q_k, H) \to 0$. It then follows that either $q_{k_0} \in H \cap Q$ for some $k_0 \in \mathbb{N}$ (*i.e.,* a solution can be found in finitely many iterations), or $q_k(2) \to b$ and hence $q_k \to (\sqrt{1-b^2}, b) \in Q \cap H$. ◊

3 Properties of the Douglas–Rachford Operator

In this section we investigate the behavior of the Douglas–Rachford operator (2) and the corresponding iteration without imposing any additional assumptions on the closed set Q.

We begin by distinguishing two cases depending on whether the point is contained in the half-space. Our first proposition shows the image of H under the Douglas–Rachford operator to be a subset of H.

The following remains true (in the convex case) when H is any obtuse cone. see Remark 3.20 from Bauschke's DR paper.

Proposition 3.1. *If $x \in H$ then* $\mathrm{DR}(x) \subset H$.

Proof. Choose any $q \in P_Q(x)$. We will distinguish two cases, depending on whether q belongs to H. If $q \notin H$, then

$$\langle a, 2q - x\rangle = 2(\langle a, q\rangle - b) + (b - \langle a, x\rangle) + b > b,$$

whence, $\mathrm{DR}(x, q) = q + (\langle a, x\rangle + b - 2\langle a, q\rangle)a$. Thus,

$$\langle a, \mathrm{DR}(x, q)\rangle = \langle a, q\rangle + \langle a, x\rangle + b - 2\langle a, q\rangle \leq 2b - \langle a, q\rangle < b,$$

and we have $\mathrm{DR}(x,q) \in H$.

Suppose now that $q \in H$. If $\langle a, 2q - x\rangle \le b$ we have $\mathrm{DR}(x,q) = q \in H$. Otherwise, if $\langle a, x\rangle + b < 2\langle a, q\rangle$, we have $\mathrm{DR}(x,q) = q + (\langle a, x\rangle + b - 2\langle a, q\rangle)a$ and

$$\langle a, \mathrm{DR}(x,q)\rangle = \langle a, x\rangle + b - \langle a, q\rangle < \langle a, q\rangle \le b.$$

Thus, $\mathrm{DR}(x,q) \in H$. □

Our second proposition characterizes behavior of the Douglas–Rachford operator for points which lie outside the half-space.

Proposition 3.2. *Suppose that $x \notin H$ and $q \in P_Q(x)$. The following holds:*

(i) if $q \in H$, then $\mathrm{DR}(x,q) = q$;
(ii) if $q \notin H$, then
  (a) if $d(x,H) \ge 2d(q,H)$, then $\mathrm{DR}(x,q) = q$ and $\mathrm{DR}(\mathrm{DR}(x,q)) = P_L(q)$;
  (b) if $d(x,H) < 2d(q,H)$, then $\mathrm{DR}(x,q) = q + (\langle a, x\rangle + b - 2\langle a, q\rangle)a$, and
    (I) if $d(x,H) \le d(q,H)$, then $\mathrm{DR}(x,q) \in H$;
    (II) if $d(x,H) > d(q,H)$, then $d(\mathrm{DR}(x,q),H) = d(x,H) - d(q,H)$. Furthermore, if $q \in P_Q(\mathrm{DR}(x,q))$, then $\mathrm{DR}(\mathrm{DR}(x,q),q) \in H$.

Proof. (i) If $q \in H$, then

$$\langle a, 2q - x\rangle = 2(\langle a, q\rangle - b) + b + (b - \langle a, x\rangle) < b;$$

whence, $\mathrm{DR}(x,q) = q$, as claimed.

(ii)(a) We have $d(x,H) \ge 2d(q,H) \iff \langle a, 2q - x\rangle \le b$. Thus, $\mathrm{DR}(x,q) = q$ and

$$\mathrm{DR}(\mathrm{DR}(x,q)) = \mathrm{DR}(q) = q - (\langle a, q\rangle - b)a = P_L(q),$$

as required.

(ii)(b)(I) If $\langle a, x\rangle - b = d(x,H) \le d(q,H) = \langle a, q\rangle - b$, then

$$\langle a, \mathrm{DR}(x,q)\rangle = \langle a, q\rangle + \langle a, x\rangle + b - 2\langle a, q\rangle = \langle a, x - q\rangle + b \le b,$$

whence, $\mathrm{DR}(x,q) \in H$.

(ii)(b)(II) If $d(x,H) > d(q,H)$, then

$$\langle a, \mathrm{DR}(x,q)\rangle = \langle a, x - q\rangle + b > b,$$

and

$$d(\mathrm{DR}(x,q),H) = \langle a, x - q\rangle = (\langle a, x\rangle - b) - (\langle a, q\rangle - b) = d(x,H) - d(q,H).$$

Further, suppose that $q \in P_Q(\mathrm{DR}(x,q))$. Then,

$$\begin{aligned}\langle a, 2q - \mathrm{DR}(x,q)\rangle &= \langle a, 2q - q - (\langle a, x\rangle + b - 2\langle a, q\rangle)a\rangle \\ &= \langle a, q\rangle - (\langle a, x\rangle - b) + 2(\langle a, q\rangle - b) \\ &= \langle a, q\rangle - d(x,H) + 2d(q,H) > b.\end{aligned}$$

Hence,

$$\begin{aligned}\mathrm{DR}(\mathrm{DR}(x,q),q) &= q + (\langle a, q + (\langle a, x\rangle + b - 2\langle a, q\rangle)\, a\rangle + b - 2\langle a, q\rangle)\, a \\ &= q + (\langle a, x\rangle - 3\langle a, q\rangle + 2b)a.\end{aligned}$$

Finally, we have

$$\begin{aligned}\langle a, \mathrm{DR}(\mathrm{DR}(x,q),q)\rangle &= (\langle a,x\rangle - b) - 2(\langle a,q\rangle - b) + b \\ &= d(x,H) - 2d(q,H) + b < b;\end{aligned}$$

whence, $\mathrm{DR}(\mathrm{DR}(x,q),q) \in H$. □

By combining Propositions 3.1 & 3.2, we shall deduce the following lemma concerning the behavior of Douglas–Rachford iterations which never enter the half-space.

Lemma 3.1. *Let $\{x_k\}$ be a Douglas–Rachford sequence with auxiliary sequence $\{q_k\}$. If $x_k \notin H$ for each $k \in \mathbb{N}$, then $q_k \notin H$ for all $k \in \mathbb{N}$, the sequence $\{d(x_k, L)\}$ is strictly monotone decreasing, and*

$$\lim_{k\to\infty} d(x_k, L) = \lim_{k\to\infty} d(q_k, L) = 0.$$

Proof. If $x_k \notin H$ for all k, by Propositions 3.1 & 3.2, it must be that $d(q_k,H) < d(x_k,H) < 2d(q_k,H)$ and $d(x_{k+1},H) = d(x_k,H) - d(q_k,H)$. Thus the sequence $\{d(x_k,H)\}$ is strictly decreasing and bounded below by zero. Since $\lim_{k\to\infty} d(x_k,H)$ exists, we deduce

$$\lim_{k\to\infty} d(q_k,H) = \lim_{k\to\infty} [d(x_k,H) - d(x_{k+1},H)] = 0.$$

Thus $\lim_{k\to\infty} d(q_k,L) = \lim_{k\to\infty} d(q_k,H) = 0$, and, as $d(x_k,H) < 2d(q_k,H)$, we have $\lim_{k\to\infty} d(x_k,L) = \lim_{k\to\infty} d(x_k,H) = 0$. □

We now turn our attention to the precise structure of the Douglas–Rachford operator at points which lie within the half-space. The following proposition gives a relationship between consecutive terms in a Douglas–Rachford sequence.

Proposition 3.3. *For any $x \in H$ and any $q \in P_Q(x)$ such that $q \notin H$ one has*

$$\begin{aligned}\mathrm{DR}(x,q) &= q - (d(x,L) + 2d(q,L))a, \\ d(\mathrm{DR}(x,q),L) &= d(q,L) + d(x,L).\end{aligned}$$

Proof. Since $x \in H$ and $q \notin H$ we have that $\langle a,q\rangle > b \geq \langle a,x\rangle$. Thus,

$$\langle a, 2q - x\rangle = \langle a,q\rangle + \langle a, q-x\rangle > b,$$

and we have

$$\begin{aligned}\mathrm{DR}(x,q) &= q - (-\langle a,x\rangle - b + 2\langle a,q\rangle)a = q - (b - \langle a,x\rangle + 2(\langle a,q\rangle - b)) \\ &= q - (d(x,L) + 2d(q,L))a.\end{aligned}$$

Then,

$$d(\mathrm{DR}(x,q),L) = |\langle a,q\rangle - b - d(x,L) - 2d(q,L)| = d(q,L) + d(x,L),$$

which completes the proof. □

When the Douglas–Rachford point x lies in H, our next proposition relates x to an auxiliary points which lie in $Q \setminus H$.

Proposition 3.4. *Let $x \in H$ and $q \in P_Q(x) \setminus H$. Then, if $p \in P_Q(\mathrm{DR}(x,q)) \setminus H$, with $p \neq q$, one has*

$$d(p,H) + \|\mathrm{DR}(x,q) - q\| \leq d(q,H) + d(\mathrm{DR}(x,q),Q). \tag{4}$$

Furthermore, we have $d(p,H) < d(q,H)$.

Proof. Let $z := \mathrm{DR}(x, q) \in H$. Since $p \in P_Q(z)$, we have $\|z - q\| \geq \|z - p\|$. Observe that

$$0 < d(p, H) + d(z, L) = \langle a, p\rangle - b + b - \langle a, z\rangle = \langle a, p - z\rangle \leq \|p - z\|. \tag{5}$$

If the inequality in (5) is not strict, there is some $\lambda \in \mathbb{R}$ such that $\lambda a = p - z$. Then,

$$\langle a, p - z\rangle = \|p - z\| \implies \lambda = \lambda\|a\|^2 = \|p - z\| > 0.$$

If $\|z - q\| = \|z - p\| = \lambda$, using Proposition 3.3, we obtain a contradiction with the assumption that $p \neq q$:

$$\|p - q\| = \|\lambda a + z - q\| = \|\lambda a - (d(x, L) + 2d(q, L))a\| = |\lambda - \|z - q\|| = 0.$$

Thus, $\|z - q\| > \|z - p\| = \lambda$, and by Proposition 3.3 we have

$$\begin{aligned} d(p, H) - d(q, H) &= \langle a, p - q\rangle = \langle a, \lambda a + z - q\rangle \\ &= \lambda + \langle a, z - q\rangle = \lambda - (d(x, L) + 2d(q, L)) \\ &= \lambda - \|z - q\| < 0, \end{aligned} \tag{6}$$

whence,

$$d(p, H) + \|z - q\| = d(q, H) + \|z - p\| = d(q, H) + d(z, Q),$$

and we are done.

Otherwise, suppose that the inequality in (5) is strict, i.e.,

$$d(p, H) + d(z, L) < \|p - z\|. \tag{7}$$

By Proposition 3.3 we know that

$$\|z - q\| = d(x, L) + 2d(q, L) \quad \text{and} \quad d(z, L) = d(q, L) + d(x, L);$$

whence,

$$\|z - q\| = d(q, L) + d(z, L).$$

Thus, by (7),

$$d(q, H) = d(q, L) = \|z - q\| - d(z, L) > \|z - q\| - \|z - p\| + d(p, H) \geq d(p, H),$$

which proves (4). The last assertion in the statement follows from (6) and the inequality above. □

Remark 3.1. In particular, Proposition 3.4 shows that once a Douglas–Rachford sequence enters the half-space, it is not possible for its auxiliary sequence to cycle within points from $Q \setminus H$. Indeed, once an element $q_{k_0} \in Q \setminus H$ appears as the k_0-th term auxiliary sequence either the sequence remains constant with $q_k = q_{k_0}$ for all $k \geq k_0$, or there exists $k_1 \geq k_0$ such that $q_{k_0} \notin P_Q(x_k)$ for all $k \geq k_1$. ◊

4 Analysis of the Algorithm

In this section we establish our main results which analyze the global convergence properties of the Douglas–Rachford method assuming some additional structure on the closed set Q. These assumption encompass, but are not limited to, the setting in which Q is a finite set, or a compact set.

We first consider the case in which the following assumption holds. This assumption might be of interest, for instance, when Q is an integer lattice.

Assumption 4.1. For every $q \in Q \setminus H$ there is at most a finite number of points in $Q \setminus H$ whose distance to L is smaller than $d(q, L)$. ◊

The following lemma shows that, under this assumption, any Douglas–Rachford sequence eventually enters the half-space.

Lemma 4.1. *Suppose Assumption 4.1 holds. Then any Douglas–Rachford sequence $\{x_k\}$ enters and remains in H after a finite number of steps. That is, there exists $k_0 \in \mathbb{N}$ such that $x_k \in H$ for all $k \geq k_0$.*

Proof. Suppose there exists a Douglas–Rachford sequence $\{x_k\}$ such that $x_k \notin H$ for all $k \in \mathbb{N}$, with auxiliary sequence $\{q_k\}$. By Proposition 3.1 & Proposition 3.2, $q_k \notin H$ for all $k \in \mathbb{N}$, and

$$d(x_{k+1}, H) = d(x_k, H) - d(q_k, H) \quad \text{for all } k \in \mathbb{N}.$$

By telescoping we obtain

$$0 < d(x_{k+1}, H) = d(x_1, H) - \sum_{j=1}^{k} d(q_j, H) \leq d(x_1, H) - k\, d(Q \setminus H, H).$$

By Assumption 4.1, we have $d(Q \setminus H, H) > 0$. A contradiction is obtained for sufficiently large k. □

Remark 4.1. Evidently, Lemma 4.1 remains valid on replacing Assumption 4.1 with the assumption $d(Q \setminus H, H) > 0$. ◊

The following example shows that the conclusions of Lemma 4.1 need not hold without something akin to Assumption 4.1.

Example 4.1. Consider the following subsets of the real line:

$$Q := \left\{ \frac{2}{3^k} \;\middle|\; k = 0, 1, 2, \dots \right\} \cup \{0\}, \quad H := \{x \in \mathbb{R} \mid x \leq 0\}.$$

For initial point $x_0 = 1$, the Douglas–Rachford iteration and auxiliary sequence are

$$x_k = \frac{1}{3^k}, \qquad q_k := P_Q(x_k) = \frac{2}{3^{k+1}}.$$

Both $\{x_n\}$ and $\{q_n\}$ are positive real numbers, and hence never enter H. ◊

We formulate now our first main result regarding convergence of the method under Assumption 4.1.

Theorem 4.1. *Suppose Assumption 4.1 holds. Let $\{x_k\}$ be a Douglas–Rachford sequence with auxiliary sequence $\{q_k\}$. Then either: (i) there exists $k_0 \in \mathbb{N}$ such that $q_{k_0} \in Q \cap H \neq \emptyset$, or (ii) $H \cap Q = \emptyset$. Moreover, in the latter case, $\|x_k\| \to +\infty$.*

Proof. Suppose $q_k \notin H$ for all $k \in \mathbb{N}$. By Lemma 4.1, there is some $k_0 \in \mathbb{N}$ such that $x_k \in H$ for all $k \geq k_0$. By Assumption 4.1 and Proposition 3.4, there is some $k_1 \in \mathbb{N}$ with $k_1 \geq k_0$ and some $q \in Q \setminus H$ such that $P_Q(x_k) = q$ for all $k \geq k_1$. Then, by Proposition 3.3,

$$\begin{aligned} x_{k+1} &= q - (d(x_k, L) + 2d(q, L))a = q - (d(x_{k-1}, L) + 3d(q, L))a \\ &= \dots = q - (d(x_{k_1}, L) + (2 + k - k_1)d(q, L))a, \end{aligned} \tag{8}$$

for all $k \geq k_1$. Hence,

$$\|x_{k+1}\| \geq d(x_{k_1}, L) + (2 + k - k_1)d(q, L) - \|q\|,$$

whence, $\|x_k\| \to \infty$. Further, suppose by contradiction that there is $p \in H \cap Q$. For any $k \geq k_1$, by (8), one has

$$\begin{aligned}\|x_{k+1} - p\|^2 &= \|(x_{k+1} - q) + (q - p)\|^2 \\ &= \|x_{k+1} - q\|^2 + \|q - p\|^2 + 2\langle x_{k+1} - q, q - p\rangle \\ &= \|x_{k+1} - q\|^2 + \|q - p\|^2 - 2\langle a, q - p\rangle(d(x_{k_1}, L) + (2 + k - k_1)d(q, L)).\end{aligned}$$

Observe that, since $q \notin H$ and $p \in H$, we have

$$\langle a, q - p\rangle = (\langle a, q\rangle - b) + (b - \langle a, p\rangle) > 0.$$

Thus, there is some $k_2 \in \mathbb{N}$, with $k_2 \geq k_1$, such that

$$\|q - p\|^2 - 2\langle a, q - p\rangle(d(x_{k_1}, L) + (2 + k - k_1)d(q, L)) < 0, \quad \forall k \geq k_2.$$

Hence,

$$\|x_{k+1} - p\|^2 < \|x_{k+1} - q\|^2, \quad \forall k \geq k_2,$$

that is, we obtain a contradiction with the fact that $P_Q(x_{k+1}) = q$. □

Theorem 4.1 suggests the following algorithm for finding a point $Q \cap H$.

Algorithm 1: The Douglas–Rachford algorithm for solving (1) with Assumption 4.1.

Input: $x_0 \in \mathbb{R}^m$;
Choose any $q_0 \in P_Q(x_0)$;
Set $k := 0$;
while $q_k \notin H$ **do**
  $x_{k+1} := \mathrm{DR}(x_k, q_k)$;
  Choose any $q_{k+1} \in P_Q(x_{k+1})$;
  $k := k + 1$;
end
Output: $q_k \in Q \cap H$;

The following corollary provides justification for Algorithm 4.1.

Corollary 4.1. *Suppose Assumption 4.1 holds. Then Algorithm 1 either: (i) terminates finitely to a point in $H \cap Q$, or (ii) $Q \cap H = \emptyset$ and $\|x_k\| \to +\infty$.*

Proof. Follows directly from Theorem 4.1. □

It is easily seen that any finite set satisfies Assumption 4.1, hence Theorem 4.1 applies. However, as we now show, a stronger result holds in this special case. We first require the following two lemmata.

Lemma 4.2. *Let $\{x_k\}$ be a Douglas–Rachford sequence with auxiliary sequence $\{q_k\}$. Suppose there exists $k_0 \in \mathbb{N}$ such that $x_{k_0} \in H$ and $q_{k_0} \in H$. Then $q_k \in H$ and $d(q_{k+1}, L) \geq d(q_k, L)$ for all $k \geq k_0$. Furthermore, for any $k \geq k_0$, one has $d(q_{k+1}, L) = d(q_k, L)$ if and only if $q_{k+1} = q_k$.*

Proof. We distinguish two cases. First suppose $\langle a, 2q_{k_0} - x_{k_0}\rangle \leq b$. By (3) we have $x_{k_0+1} = q_{k_0}$ and hence $q_{k_0+1} = q_{k_0}$.

Now suppose $\langle a, 2q_{k_0} - x_{k_0}\rangle > b$. By (3),

$$x_{k_0+1} = q_{k_0} - (\langle a, 2q_{k_0} - x_{k_0}\rangle - b)a = q_{k_0} - d(2q_{k_0} - x_{k_0}, H)a. \tag{9}$$

Then, as $q_{k_0+1} \in P_Q(x_{k_0+1})$, we have

$$\begin{aligned}
\langle a, q_{k_0+1}\rangle &= \langle a, q_{k_0+1} - x_{k_0+1}\rangle + \langle a, x_{k_0+1}\rangle \\
&\leq \|q_{k_0+1} - x_{k_0+1}\| + \langle a, q_{k_0}\rangle - d(2q_{k_0} - x_{k_0}, H) \qquad (10)\\
&\leq \|q_{k_0} - x_{k_0+1}\| + \langle a, q_{k_0}\rangle - d(2q_{k_0} - x_{k_0}, H) \qquad (11)\\
&= \langle a, q_{k_0}\rangle \leq b,
\end{aligned}$$

whence $q_{k_0+1} \in H$ and $d(q_{k_0+1}, L) \geq d(q_{k_0}, L)$.

Further, let us assume that $d(q_{k_0+1}, L) = d(q_{k_0}, L)$. Then, we have $\langle a, q_{k_0+1}\rangle = \langle a, q_{k_0}\rangle$. Therefore the inequalities in (10) and (11) must be equalities, from where we deduce

$$\langle a, q_{k_0+1} - x_{k_0+1}\rangle = \|q_{k_0+1} - x_{k_0+1}\| = \|q_{k_0} - x_{k_0+1}\|. \tag{12}$$

Hence, by (12) and (9), we have

$$\begin{aligned}
\langle q_{k_0+1} - x_{k_0+1}, x_{k_0+1} - q_{k_0}\rangle &= \langle q_{k_0+1} - x_{k_0+1}, -d(2q_{k_0} - x_{k_0}, H)a\rangle \\
&= -\|x_{k_0+1} - q_{k_0}\|\langle q_{k_0+1} - x_{k_0+1}, a\rangle \\
&= -\|x_{k_0+1} - q_{k_0}\|^2.
\end{aligned}$$

Thus, using again (12), we obtain

$$\begin{aligned}
\|q_{k_0+1} - q_{k_0}\|^2 &= \|q_{k_0+1} - x_{k_0+1}\|^2 + \|x_{k_0+1} - q_{k_0}\|^2 + 2\langle q_{k_0+1} - x_{k_0+1}, x_{k_0+1} - q_{k_0}\rangle \\
&= \|q_{k_0+1} - x_{k_0+1}\|^2 - \|x_{k_0+1} - q_{k_0}\|^2 = 0;
\end{aligned}$$

that is, $q_{k_0+1} = q_{k_0}$. The result now follows by induction. □

Lemma 4.3. *Let $\{x_k\}$ be a Douglas–Rachford sequence with auxiliary sequence $\{q_k\}$. Suppose there exists $k_0 \in \mathbb{N}$ such that $x_k \in H$ and $q_k = q \in H$ for all $k \geq k_0$. Then there exists $k_1 \in \mathbb{N}$ with $k_1 \geq k_0$ such that $x_k = x_{k_1}$ for all $k \geq k_1$.*

Proof. If $\langle a, 2q - x_j\rangle \leq b$ for some $j \geq k_0$, then by (3) we deduce $x_k = q$ for all $k \geq j+1$ and we are done.

Otherwise, by (3), we have $x_{k+1} = q + (\langle a, x_k\rangle + b - 2\langle a, q\rangle)a$ for all $k \geq k_0$. Thus,

$$\begin{aligned}
\langle a, 2q - x_{k+1}\rangle &= \langle a, 2q - (q + (\langle a, x_k\rangle + b - 2\langle a, q\rangle)a)\rangle \\
&= \langle a, 2q - x_k\rangle - (b - \langle a, q\rangle) \\
&= \langle a, 2q - x_k\rangle - d(q, L),
\end{aligned}$$

whence,

$$b < \langle a, 2q - x_k\rangle = \langle a, 2q - x_{k_0}\rangle - (k - k_0)d(q, L).$$

This implies $d(q, L) = 0$ and then,

$$x_{k+1} = q + (\langle a, x_k\rangle + b - 2\langle a, q\rangle)a = q + (\langle a, x_k\rangle - b)a,$$

from which we deduce $x_k = q + (\langle a, x_k\rangle - b)a$ for all $k \geq k_0 + 1$. This completes the proof. □

The following theorem is a refinement of Theorem 4.1 when Q is assumed to be finite.

Theorem 4.2. *Suppose Q is finite. Let $\{x_k\}$ be a Douglas–Rachford sequence with auxiliary sequence $\{q_k\}$. Then either: (i) $\{x_k\}$ and $\{q_k\}$ are eventually constant and the limit of $\{q_k\}$ is contained in $H \cap Q \neq \emptyset$, or (ii) $H \cap Q = \emptyset$ and $\|x_k\| \to +\infty$.*

Proof. By Theorem 4.1, it suffices to show that the sequences $\{x_k\}$ and $\{q_k\}$ are eventually constant and the limit of $\{q_k\}$ is contained in $H \cap Q \neq \emptyset$, assuming there exists a $k_0 \in \mathbb{N}$ such that $x_{k_0} \in H$ and $q_{k_0} \in H$.

To this end, suppose there is some $k_0 \in \mathbb{N}$ such that $x_{k_0} \in H$ and $q_{k_0} \in H$. By Proposition 3.1 and Lemma 4.2 we have $x_k \in H, q_k \in H$ and $d(q_{k+1}, L) \geq d(q_k, L)$ with equality if and only if $q_{k+1} = q_k$, for all $k \geq k_0$. Since Q is finite, the latter implies that there exists $k_1 \geq k_0$ such that $q_k = q_{k_1}$ for all $k \geq k_1$. By Lemma 4.3, there exists a $k_2 \geq k_1$ such $x_k = x_{k_2}$ for all $k \geq k_2$. This completes the proof. □

We now consider the following assumption which holds, in particular, for any compact set Q.

Assumption 4.2. The function $\iota_Q + d(\cdot, H)$ has compact lower-level sets. In particular, for every $q \in Q$, the set

$$\mathcal{Q}(q) := \{p \in Q \mid d(p, H) \leq d(q, H)\},$$

is compact. ◊

Assumption 4.2 may also be written in the form $\lim_{x \in Q, \|x\| \to \infty} d(x, H) = +\infty$.

The following is our final main result. It characterizes behavior of the algorithm for Q satisfying Assumption 4.2. In particular, it shows that the Douglas–Rachford algorithm can be used to determine consistency of the feasibility problem (1) and to find a solution when such exist.

Theorem 4.3. *Suppose Assumption 4.2 holds. Let $\{x_k\}$ be a Douglas–Rachford sequence with auxiliary sequence $\{q_k\}$. Then either: (i) $d(q_k, H) \to 0$ and the set of cluster points of the auxiliary sequence is non-empty and contained in $Q \cap H$, or (ii) $d(q_k, H) \to \beta$ for some $\beta > 0$ and $H \cap Q = \emptyset$. Moreover, in the latter case, $\|x_k\| \to +\infty$.*

Proof. Let $\{x_k\}$ be a Douglas–Rachford sequence with auxiliary sequence $\{q_k\}$. We distinguish two cases.

First suppose that $x_k \notin H$ for all k. By Proposition 3.2, $q_k \notin H$ for all k and, by Lemma 3.1, $d(q_k, H) \to 0$. Hence there exists $k_0 \in \mathbb{N}$ such $d(q_k, H) \leq d(q_0, H)$ for all $k \geq k_0$, whence $q_k \in \mathcal{Q}(q_0)$ for all $k \geq k_0$. Since $\mathcal{Q}(q_0)$ is compact and $d(\cdot, H)$ is continuous, the set of cluster points of the auxiliary sequence $\{q_k\}$ is non-empty and contained in $Q \cap H$.

Next suppose there is some $k_0 \in \mathbb{N}$ such that $x_{k_0} \in H$. Then, by Proposition 3.1, we have $x_k \in H$ for all $k \geq k_0$. On one hand, if there exists $k_1 \geq k_0$ such that $q_{k_1} \in H$ then, by Lemma 4.2, $q_k \in H$ for all $k \geq k_1$. Since $\mathcal{Q}(q_{k_1}) = Q \cap H$ is compact, it follows that the set of cluster points of the auxiliary sequence $\{q_k\}$ is non-empty and contained in $Q \cap H$, and we are done.

On the other hand, suppose $q_k \notin H$ for all $k \geq k_0$. Then by Proposition 3.4, $q_k \in \mathcal{Q}(q_{k_0})$ for all $k \geq k_0$, and $\{d(q_k, H)\}_{k=k_0}^{\infty}$ is monotone decreasing and bounded below by zero, hence

$$\beta := \inf_{k \geq k_0} d(q_k, H) = \lim_{k \to \infty} d(q_k, H) \geq 0.$$

By Proposition 3.3, for all $k \geq k_0$,

$$\begin{aligned} x_{k+1} &= q_k - (d(x_k, L) + 2d(q_k, L))a \\ &= q_k - (d(x_{k-1}, L) + d(q_{k-1}, L) + 2d(q_k, L))a \\ &= \ldots = q_k - \left(d(x_{k_0}, L) + d(q_k, L) + \sum_{j=k_0}^{k} d(q_j, L) \right) a. \end{aligned}$$

For all $k \geq k_0$, we may therefore express $x_{k+1} = q_k - \lambda_k a$ where

$$\lambda_k := d(x_{k_0}, L) + d(q_k, L) + \sum_{j=k_0}^{k} d(q_j, L) \geq (k - k_0)\beta \geq 0. \tag{13}$$

If $\beta = 0$ then, by the compactness of $\mathcal{Q}(q_{k_0})$ and continuity of $d(\cdot, H)$, the set of cluster points of the auxiliary sequence $\{q_k\}$ is non-empty and contained in $Q \cap H$.

Conversely, assume $\beta > 0$. By the compactness of $\mathcal{Q}(q_{k_0})$ there exists $K > 0$ such that $\|q\| \leq K$ for all $q \in \mathcal{Q}(q_{k_0})$, and thus

$$\|x_{k+1}\| \geq \lambda_k - \|q_k\| \geq \lambda_k - K \geq (k - k_0)\beta - K \to +\infty. \tag{14}$$

To complete the proof, we must show $Q \cap H = \emptyset$. To this end, suppose there is a $p \in Q \cap H$. We claim that, for all $k \geq k_0$,

$$\|x_k - q_k\| > \left\| x_k - P_L(q_k) - \frac{\beta}{2} a \right\|.$$

To prove this claim, first observe

$$\begin{aligned} \left(\langle a, q_k \rangle - b - \frac{\beta}{2} \right) + 2\langle a, x_k - q_k \rangle &= -\left(\langle a, q_k \rangle - b \right) - \frac{\beta}{2} - 2\left(b - \langle a, x_k \rangle \right) \\ &= -d(q_k, L) - \frac{\beta}{2} - 2d(x_k, L) < 0. \end{aligned}$$

Since $\left(\langle a, q_k \rangle - b - \frac{\beta}{2} \right) = d(q_k, L) - \frac{\beta}{2} \geq \beta - \frac{\beta}{2} = \frac{\beta}{2} > 0$, we deduce the claimed result

$$\begin{aligned} \left\| x_k - P_L(q_k) - \frac{\beta}{2} a \right\|^2 &= \left\| x_k - q_k + \left(\langle a, q_k \rangle - b - \frac{\beta}{2} \right) a \right\|^2 \\ &= \|x_k - q_k\|^2 + \left(\langle a, q_k \rangle - b - \frac{\beta}{2} \right)^2 + 2\left(\langle a, q_k \rangle - b - \frac{\beta}{2} \right) \langle a, x_k - q_k \rangle \\ &< \|x_k - q_k\|^2. \end{aligned}$$

Further, for all $k \geq k_0$, we have

$$\begin{aligned} \|x_{k+1} - q_{k+1}\|^2 &> \left\| x_{k+1} - P_L(q_{k+1}) - \frac{\beta}{2} a \right\|^2 \\ &= \|x_{k+1} - p\|^2 + \left\| p - P_L(q_{k+1}) - \frac{\beta}{2} a \right\|^2 + 2\left\langle x_{k+1} - p, p - P_L(q_{k+1}) - \frac{\beta}{2} a \right\rangle. \end{aligned}$$

Since

$$\begin{aligned} \left\langle a, P_L(q_{k+1}) + \frac{\beta}{2} a - p \right\rangle &= \left\langle a, q_{k+1} - \left(\langle a, q_{k+1} \rangle - b \right) a + \frac{\beta}{2} a - p \right\rangle \\ &= b - \langle a, p \rangle + \frac{\beta}{2} = d(p, L) + \frac{\beta}{2}, \end{aligned}$$

we obtain

$$\begin{aligned}\left\langle x_{k+1}-p, p-P_L(q_{k+1})-\frac{\beta}{2}a\right\rangle &= \left\langle q_k-\lambda_k a-p, p-P_L(q_{k+1})-\frac{\beta}{2}a\right\rangle \\ &= \left\langle q_k-p, p-P_L(q_{k+1})-\frac{\beta}{2}a\right\rangle \\ &\quad +\lambda_k\left\langle a, P_L(q_{k+1})+\frac{\beta}{2}a-p\right\rangle \\ &= \left\langle q_k-p, p-P_L(q_{k+1})-\frac{\beta}{2}a\right\rangle+\lambda_k\left(d(p,L)+\frac{\beta}{2}\right).\end{aligned}$$

Then,

$$\begin{aligned}\|x_{k+1}-q_{k+1}\|^2 > \|x_{k+1}-p\|^2 &+ \left\|p-P_L(q_{k+1})-\frac{\beta}{2}a\right\|^2 \\ &+2\left\langle q_k-p, p-P_L(q_{k+1})-\frac{\beta}{2}a\right\rangle+2\lambda_k\left(d(p,L)+\frac{\beta}{2}\right).\end{aligned}$$

By the compactness of $\mathcal{Q}(q_{k_0})$ and since $H\cap Q\subseteq\mathcal{Q}(q_{k_0})$, we have

$$\eta := \min_{w,z\in\mathcal{Q}(q_{k_0})}\left\{\left\|p-P_L(z)-\frac{\beta}{2}a\right\|^2+2\left\langle w-p, p-P_L(z)-\frac{\beta}{2}a\right\rangle\right\}\in\mathbb{R}.$$

Therefore,

$$\|x_{k+1}-q_{k+1}\|^2 > \|x_{k+1}-p\|^2+\eta+2\lambda_k\left(d(p,L)+\frac{\beta}{2}\right).$$

Finally, since $\lambda_k\to+\infty$ (by (13)) and $d(p,L)+\frac{\beta}{2}>0$, there is some $k_1\geq k_0$ such that $\eta+2\lambda_k\left(d(p,L)+\frac{\beta}{2}\right)>0$ for all $k\geq k_1$. Then

$$\|x_{k+1}-q_{k+1}\|^2 > \|x_{k+1}-p\|^2,\quad \forall k\geq k_1,$$

which contradicts the fact that $q_{k+1}\in P_Q(x_{k+1})$. We therefore have that $H\cap Q=\emptyset$ and the proof is complete. □

Every compact set satisfies Assumption 4.2, and thus we deduce the following important corollary.

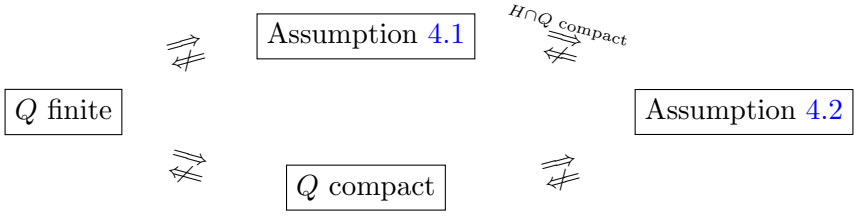
Corollary 4.2. *Suppose Q is a compact set. Let $\{x_k\}$ be a Douglas–Rachford sequence with auxiliary sequence $\{q_k\}$. Then either: (i) $d(q_k,H)\to 0$ and the set of cluster points of the auxiliary sequence is non-empty and contained in $Q\cap H$, or (ii) $d(q_k,H)\to\beta$ for some $\beta>0$ and $H\cap Q=\emptyset$. Moreover, in the latter case, $\|x_k\|\to+\infty$.*

Proof. Follows immediately from Theorem 4.3. □

Remark 4.2. A closer look at the proof of Theorem 4.3 in the case of $H\cap Q=\emptyset$, shows that (14) gives information regarding the rate of divergence of the Douglas–Rachford sequence. More precisely, it shows that the sequence diverges with *at least linear rate* in the sense that $\|x_{k+1}\|\geq(k-k_0)\beta-K$ for some $\beta,K>0$. ♢

Remark 4.3. In general, the Douglas–Rachford sequence need not converge finitely for compact Q (see Theorem 4.2). Example 4.1 serves as a counter-example. ♢

Remark 4.4. The following diagram summarizes the relationships amongst our assumptions.



◊

5 Examples and Counter-Examples

In this section, we give a number of examples which highlight the role of the hyperplane in Theorems 4.3 and 4.1. These examples are interesting in light of results such as [7, 14] which exploit linear structure to analyze the Douglas–Rachford method.

Example 5.1 (Failure for two half-spaces). The algorithm no longer remains globally convergent on replacing the half-space by a cone resulting from the intersection of two half-spaces, as is shown in Figure 3. ◊

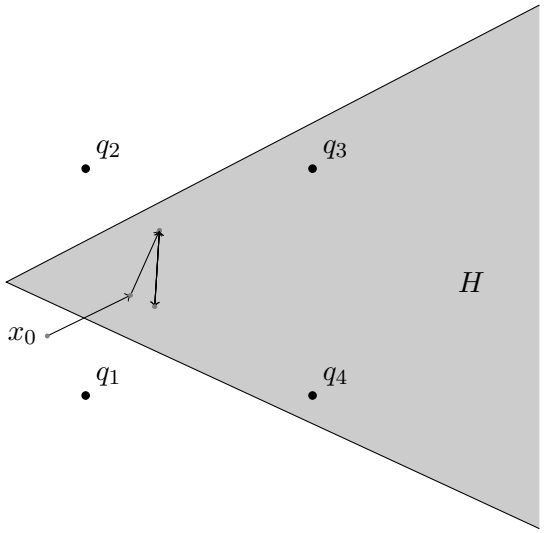


Figure 3: A 2-cycle of the Douglas–Rachford algorithm when H is a cone.

Example 5.2 (Failure for hyperplane). The Douglas–Rachford operator for Q and L (rather than H) is given by

$$\mathrm{DR}(x) = \bigcup_{q \in P_Q(x)} \mathrm{DR}(x,q) = \bigcup_{q \in P_Q(x)} q + (\langle a, x\rangle + b - 2\langle a, q\rangle)a.$$

In this case, Proposition 3.4 no longer holds, hence the algorithm need not converge. An example with cycling behavior is given in Figure 4. Moreover, cycles are still possible,

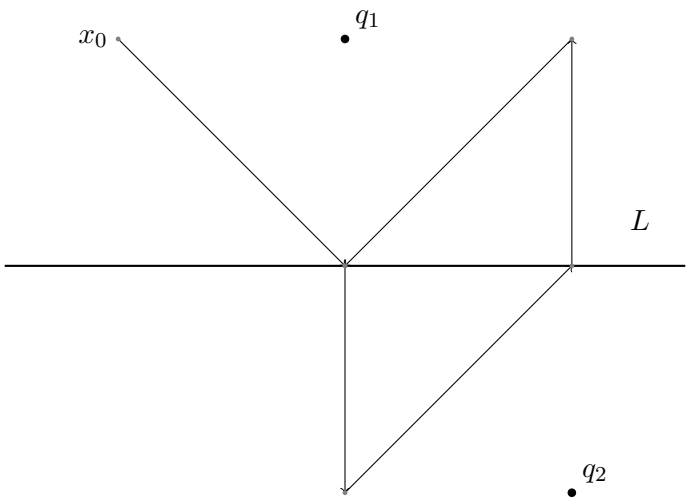


Figure 4: A 4-cycle of the Douglas–Rachford algorithm with hyperplane constraint.

even in the product formulation in terms of the diagonal space, see Example 5.4 , as the following example shows. Let $C_1 = C_2 := \{0, 1\}$. Consider

$$C := C_1 \times C_2 = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

and

$$D := \{(x, y) \in \mathbb{R}^2 \mid x = y\}.$$

If $x_0 = (-1/2, 1)^T$, then $x_1 = (1/4, 3/4)^T$, $x_2 = (3/4, 1/4)^T$ and $x_3 = x_1$. ◊

Example 5.3 (Failure with slab constraints). The algorithm no longer remains globally convergent on replacing the half-space by a *slab* constraint. That is,

$$H := \{x \mid c \geq \langle a, x \rangle \geq d\},$$

with $d < c$, both finite. An example is given in Figure 5.

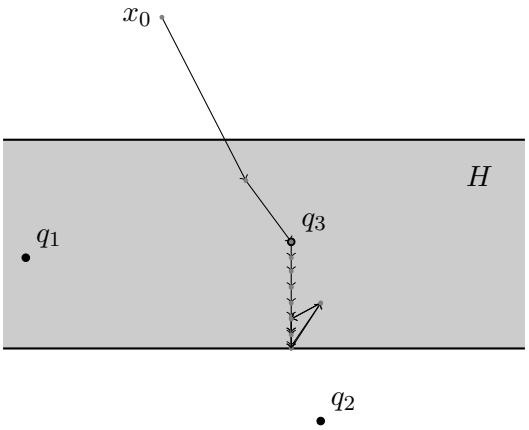


Figure 5: A 4-cycle of the Douglas–Rachford algorithm with a slab constraint.

◊

Example 5.4 (Failure of the product reformulation). *Pierra's product space reformulation* casts any feasibility problem with finitely many constraints as an equivalent two set problem in a larger product space. As the Douglas–Rachford method can only be directly applied to two set feasibility problem, this reformulation is crucial in many applications.

Consider the constraint sets

$$H := \{x \in \mathbb{R}^2 \mid x_2 \leq 1\}, \qquad Q := \{0,1\} \times \{0,1\}.$$

Applied to this problem, the global convergence of the Douglas–Rachford method is covered by Theorem 4.2. However, as we will show the product space reformulation destroys convergence.

Consider the product space $\mathbb{R}^2 \times \mathbb{R}^2$ with the reformulation with constraints

$$C := H \times Q = \{(x,y) \mid x \in H,\, y \in Q\}, \qquad D = \{(x,y) \mid x = y\}.$$

Observe that $x \in Q \cap H$ if and only if $(x,x) \in C \cap D$. Furthermore, note that the neither of the sets C or D are half-spaces, hence our results no longer apply.

Consider the Douglas–Rachford iteration with the reflection first performed with respect to the *diagonal space* D for initial point $(x_0, y_0) = ((0,2/5),(0,4/5))$. Then $(x_1, y_1) = ((0,3/5),(0,1/5))$ and $(x_2, y_2) = (x_0, y_0)$. That is, 2-cycle is obtained.

If the reflection was instead performed first with respect to $H \times Q$ for initial point $(x_0, y_0) = ((0,4/5),(0,2/5))$, a 2-cycle is still obtained. In this case, we have $(x_1, y_1) = ((0,1/5),(0,3/5))$ and $(x_2, y_2) = (x_0, y_0)$. ◊

6 Conclusion

We have established global convergence and described global behavior of the Douglas–Rachford method applied to the two-set feasibility problem of finding a point in the intersection of a half-space and a second potentially non-convex set, assumed to have minimal additional structure. Indeed, the difference in behavior of the method applied to a half-space instead of an affine constraint is quite striking.

Acknowledgements F.J. Aragón Artacho was supported by MINECO of Spain, as part of the Ramón y Cajal program (RYC-2013-13327) and the I+D grant MTM2014-59179-C2-1-P. J.M. Borwein was supported, in part, by the Australian Research Council. M.K. Tam was supported by an Australian Post-Graduate Award.

References

[1] F.J. Aragón Artacho, J.M. Borwein & M.K. Tam.: Douglas-Rachford feasibility methods for matrix completion problems. *ANZIAM J.*, 55(4):299–326, 2014.

[2] F.J. Aragón Artacho & J.M. Borwein.: Global convergence of a non-convex Douglas–Rachford iteration. *J. Glob. Optim.*, 57(3):753–769, 2013.

[3] F.J. Aragón Artacho, J.M. Borwein & M.K. Tam.: Recent results on Douglas–Rachford methods for combinatorial optimization problem. *J. Optim. Theory. Appl.*, 163(1):1–30, 2014.

[4] J.M. Borwein & M.K. Tam.: Reflection methods for inverse problems with application to protein conformation determination. *Generalized Nash Equilibrium Problems, Bilevel programming and MPEC*, New Delhi, 2012.

[5] J.M. Borwein & M.K. Tam.: A cyclic Douglas–Rachford Iteration Scheme. *J. Optim. Theory Appl.,* 160(1):1–29, 2014.

[6] J.M. Borwein & M.K. Tam.: The cyclic Douglas–Rachford method for inconsistent feasibility problems. *J. Nonlinear Convex Anal.,* 16(4):537-584, 2015.

[7] H.H. Bauschke, M.N. Dao, and W.M. Moursi.: The Douglas–Rachford algorithm in the affine-convex case. `arXiv:1505.06408v1`

[8] H.H. Bauschke & V.R. Koch.: Projection Methods: Swiss Army Knives for solving feasibility and best approximation problems with halfspaces. *Infinite Products and Their Applications* (Haifa), 2012.

[9] H.H. Bauschke and D. Noll.: On the local convergence of the Douglas–Rachford algorithm. *Arch. Math.* 102:589–600, 2014.

[10] H.H. Bauschke, P.L. Combettes, & D.R. Luke.: Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization. *JOSA A* 19(7):1334–1345, 2002.

[11] J. Benoist.: The Douglas–Rachford algorithm for the case of the sphere and the line. *J. Global Optim.* in press, 2015.

[12] J.M. Borwein & B. Sims.: The Douglas–Rachford algorithm in the absence of convexity. In *Fixed-Point Algorithms for Inverse Problems in Science and Engineering.* Springer New York, p. 93–109, 2011.

[13] V. Elser, I. Rankenburg & P. Thibault.: Searching with iterated maps. Proc. Nat. Acad. Sc., 104(2):418–423, 2007.

[14] R. Hesse & D.R. Luke.: Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems. *SIAM J. Optim.*, 23(4):2397–2419, 2013.

[15] R. Hesse, D.R. Luke & P. Neumann.: Alternating projections and Douglas–Rachford for sparse affine feasibility. *IEEE. Trans. Sign. Proc.*, 62(18):4868–4881, 2014.

[16] S. Gravel & V. Elser.: Divide and concur: A general approach to constraint satisfaction. *Phys. Rev. E*, 78(3):036706, 2008.

[17] G.L. Nemhauser & L.A. Wolsey.: *Integer and combinatorial optimization.* John Wiley & Sons, 1999.

[18] H.M. Phan.: Linear convergence of the Douglas–Rachford method for two closed sets. *Optim.* in press, 2015.

This paper deals discusses finite convergences results for a modification of the Douglas–Rachford iteration applies to a convex set, and a convex cone with non-empty interior:
S.-Y. Matsushita & L. Xu.: On the finite convergence of the Douglas–Rachford method.