Chapter 6
The Douglas-Rachford algorithm in the absence of convexity

Jonathan M. Borwein and Brailey Sims

Summary: The Douglas-Rachford iteration scheme, introduced half a century ago in connection with nonlinear heat flow problems, aims to find a point common to two or more closed constraint sets. Convergence of the scheme is ensured when the sets are convex subsets of a Hilbert space, however, despite the absence of satisfactory theoretical justification, the scheme has been routinely used to successfully solve a diversity of practical problems in which one or more of the constraints involved is non-convex. As a first step toward addressing this deficiency, we provide convergence results for a prototypical non-convex two-set scenario in which one of the sets is the Euclidean sphere.

Key words: non-convex feasibility problem, fixed point theory, dynamical system, iteration
AMS 2010 Subject Classification: 46B45, 47H10, 90C26

Introduction

In recent times variations of alternating projection algorithms have been applied in Hilbert space to various important applied problems—from optical aberration correction to three satisfiability, protein folding and construction of giant *Sudoku* puzzles [8]. While the theory of such methods is well understood in the convex case [3] and [4–6,12], there is little corresponding theory when some of the sets involved are non-convex—and that is the case for the examples mentioned above [8,9].

Jonathan M. Borwein
CARMA, School of Mathematical and Physical Sciences, University of Newcastle, NSW 2308, Australia, e-mail: jonathan.borwein@newcastle.edu.au

Brailey Sims
CARMA, School of Mathematical and Physical Sciences, University of Newcastle, NSW 2308, Australia, e-mail: brailey.sims@newcastle.edu.au

Our intention is to analyze the simplest non-convex prototype in Euclidean space: that of finding a point on the intersection of a sphere and a line or more generally a proper affine subset. The sphere provides an accessible model of many reconstruction problems in which the magnitude, but not the phase, of a signal is measured.

Preliminaries

For any closed subset A of a Hilbert space $(X, \langle\cdot,\cdot\rangle)$ we say that a mapping $P_A : D_A \subseteq X \longrightarrow A$ is a *closest point projection* of D_A onto A if $A \subseteq D_A$, $P_A^2 = P_A$ and

$$\|x - P_A(x)\| = \operatorname{dist}(x, A) := \inf\{\|x - a\| : a \in A\},$$

for all $x \in D_A$.

For a given closest point projection, P_A, onto A we take the *reflection* of x in A (relative to P_A) to be,

$$R_A := 2P_A - I.$$

In this note we will focus on the cases when the subset A is a sphere, which without loss of generality we take to be the unit sphere of the space; $S := \{x : \|x\| = 1\}$, or a line $L := \{x = \lambda a + \alpha b : \lambda \in \mathbf{R}\}$ where, without loss of generality, we take $\|a\| = \|b\| = 1$, $a \perp b$ and $\alpha > 0$.

The closest point projection of $x \neq 0$ onto the unit sphere S is,

$$P_S(x) := \frac{x}{\|x\|}$$

and so,

$$R_S(x) = \left(\frac{2}{\|x\|} - 1\right) x.$$

Excluding $x = 0$ from the domain of P_S, and hence also R_S, avoids the problem of non-unique closest points and hence the need to make a selection. The closest point projection of $x \in X$ onto L is the orthogonal projection,

$$P_L(x) := \langle x, a\rangle\, a + \alpha b$$

and so,

$$R_L(x) = 2\langle x, a\rangle\, a + 2\alpha b - x.$$

Given two closed sets A and B together with closest point projections P_A and P_B, starting from an arbitrary initial point $x_0 \in D_A$ the *Douglas-Rachford iteration scheme* (reflect-reflect-average), introduced in [7] for numerical solution of partial differential equations, is a method for finding a point in the intersection of the two sets. That is, it aims to find a feasible point for the possibly non-convex constraint $x \in A \cap B$. Explicitly it is the iterative scheme,

$$x_{n+1} := T_{A,B}(x_n),$$

where $T_{A,B}$ is the operator $T_{A,B} := \frac{1}{2}(R_B R_A + I)$. This method also goes under many other names, see [5].

When either of the sets is non-convex various compatibility restrictions between the domains and ranges of the mappings involved are required to ensure all iterates are defined. For instance; $R_A(D_A) \subseteq D_B$ and $\frac{1}{2}(R_B R_A + I)(D_A) \subseteq D_A$.

With our particular S and L we have for $x \neq 0$ that,

$$T_{S,L}(x) = \left(1 - \frac{1}{\|x\|}\right) x + \left(\frac{2}{\|x\|} - 1\right) \langle x, a\rangle\, a + \alpha b.$$

Thus, if X is N-dimensional and $(x(1), x(2), x(3), \cdots, x(N))$ denotes the coordinates of x relative to an orthonormal basis B whose first two elements are respectively a and b we have,

$$T_{S,L}(x) = \left(\frac{x(1)}{\rho}, \left(1 - \frac{1}{\rho}\right) x(2) + \alpha, \left(1 - \frac{1}{\rho}\right) x(3), \cdots, \left(1 - \frac{1}{\rho}\right) x(N)\right),$$

where $\rho := \|x\| = \sqrt{x(1)^2 + \cdots + x(N)^2}$.

Let us note that the only fixed points of $T_{S,L}$ are $\pm\sqrt{1-\alpha^2}a + \alpha b$, the two points of intersection of S with L.

In this case the *Douglas-Rachford* scheme becomes,

$$x_{n+1}(1) = x_n(1)/\rho_n, \tag{6.1}$$
$$x_{n+1}(2) = \alpha + (1 - 1/\rho_n)x_n(2), \quad \text{and} \tag{6.2}$$
$$x_{n+1}(k) = (1 - 1/\rho_n)x_n(k), \quad \text{for } k = 3, \cdots, N, \tag{6.3}$$

where $\rho_n := \|x_n\| = \sqrt{x_n(1)^2 + \cdots + x_n(N)^2}$.

From this it is clear that if the initial point x_0 lies in the hyperplane $\langle x, a\rangle = 0$; that is $x_0(1) = 0$, then all of the iterates remain in that hyperplane, which we will refer to as a *singular manifold* for the problem. We will analyze this case in greater detail in a subsequent section. Similarly, if the initial point lies in either of the two open half-spaces $\langle x, a\rangle > 0$ or $\langle x, a\rangle < 0$; that is, $x_0(1) > 0$ or $x_0(1) < 0$ respectively, then all subsequent iterates will remain in the same open half space. Further, by symmetry, it suffices to only consider initial points lying in the positive open half-space $x_0(1) > 0$.

Figure 6.1 shows two steps of the underlying geometric construction: the smaller (green) points are the intermediate reflections in the sphere. Most figures were constructed in *Cinderella*, a software geometry package, [www.cinderella.de]. A web applet version of the underlying Cinderella construction is available at
http://www.carma.newcastle.edu.au/~jb616/reflection.html.
Indeed, many of the insights for the proofs below came from examining the constructions. The number of iterations N, the height of the line (α), and the initial point are all dynamic—changing one changes the entire visible trajectory.

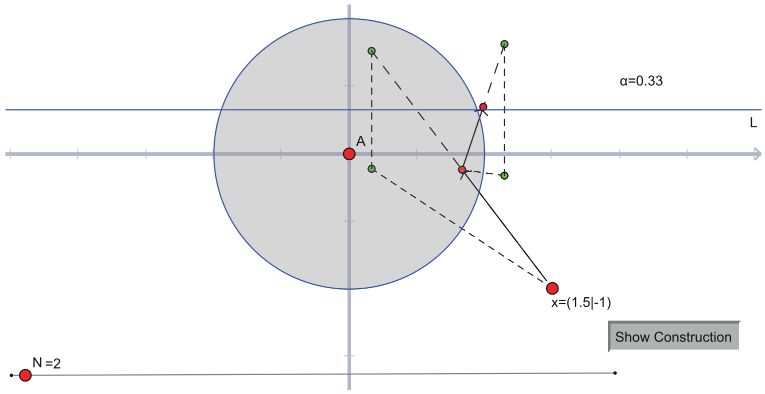

Fig. 6.1 Two steps showing the construction.

Success of the Douglas-Rachford scheme relies on convergence of the (*Picard*) iterates, $x_n = T^n_{A,B}(x_0)$, to a fixed point of the generally nonlinear operator $T_{A,B}$ in $A \cap B$, as $n \to \infty$. When both A and B are closed convex sets convergence of the scheme (in the weak topology) from any initial point in X to some point in $A \cap B$ was established by Lions and Mercier [12].

However, as noted, many practical situations yield feasibility problems in which one or more of the constraint sets is non-convex. That the Douglas-Rachford scheme works well in many of these situations has been observed and exploited for some years, despite the absence of any really satisfactory theoretical underpinning.

Remark 6.1 (divide-and-concur). If one wishes to find a point in the intersection of M sets $A_1, A_2, \dots A_k, \dots, A_M$ in X we can instead consider the subset $A := \prod_{k=1}^{M} A_k$ and the linear subset

$$B := \{x = (x_1, x_2, \dots, x_M) \colon x_1 = x_2 = \cdots = x_M\}$$

of the Hilbert space product $\prod_{k=1}^{M} X$. Then we observe that,

$$R_A(x) = \prod_{k=1}^{M} R_{A_k}(x_k),$$

so that the reflections may be 'divided' up and

$$P_B(x) = \left(\frac{x_1 + x_2 + \cdots + x_M}{M}, \dots, \frac{x_1 + x_2 + \cdots + x_M}{M} \right),$$

so that the projection and hence reflection on B are averaging ('concurrences'); thence comes the name. In this form the algorithm is particularly suited to parallelization [10].

We can also compose more reflections in serial as illustrated for reflect-reflect-reflect-average with spheres in Figure 6.2, where we observe iterates spiralling to a feasible point. ◇

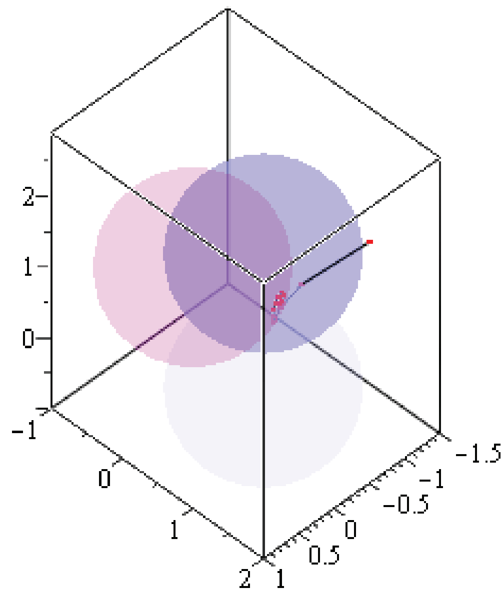

Fig. 6.2 Douglas-Rachford for three spheres in three-space.

Example 6.1 (linear equations). For the hyperplane $H := \{x\colon \langle b,x\rangle = \alpha\}$, where without loss of generality we take $\|b\| = 1$, the projection is

$$x \mapsto x + (\alpha - \langle b,x\rangle)\, b.$$

The consequent averaged-reflection version of the Douglas-Rachford recursion for a point in the intersection of M distinct hyperplanes is:

$$x \mapsto x + \frac{2}{M}\sum_{k=1}^{M}(\alpha_k - \langle b_k,x\rangle)\, b_k, \tag{6.4}$$

while the corresponding-averaged projection algorithm is:

$$x \mapsto x + \frac{1}{M}\sum_{k=1}^{M}(\alpha_k - \langle b_k, x\rangle)\, b_k \tag{6.5}$$

In more general situations, the difference between projection and reflection algorithms is even greater. ◇

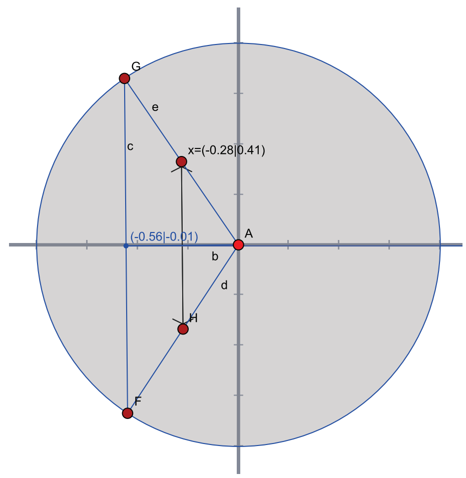

Fig. 6.3 Iterated reflection with a ray.

Remark 6.2 (The case of a half-line or segment). Note, even in two dimensions, alternating projections, alternating reflections, project-project and average, and reflect-reflect and average will all often converge to (locally nearest) infeasible points even when A is simply the ray $R := \{(x,0) : x \geqslant -1/2\}$ and B is the circle as before. They can also behave quite 'chaotically'. (See Figure 6.3 for a periodic illustration in *Cinderella* and Figure 6.4 for more complex behaviour.) So the affine nature of the convex set seems quite important. ◇

For any two closed sets A and B and feasible point $p \in A \cap B$ we say that the Douglas-Rachford scheme is *locally convergent* at p if there is a neighbourhood, N_p of p such that starting from any point x_0 in N_p the iterates $T^n_{A,B}(x_0)$ converge to p. The set comprising all initial points x_0 for which the iterates converge to p is the *basin of attraction* of p.

As a first step toward an understanding of the Douglas-Rachford scheme in the absence of convexity, we analyze its behaviour in the indicative situation when one of the sets is the non-convex sphere S and the other is the affine line L. We begin by establishing local convergence of the scheme when $0 \leq \alpha < 1$.

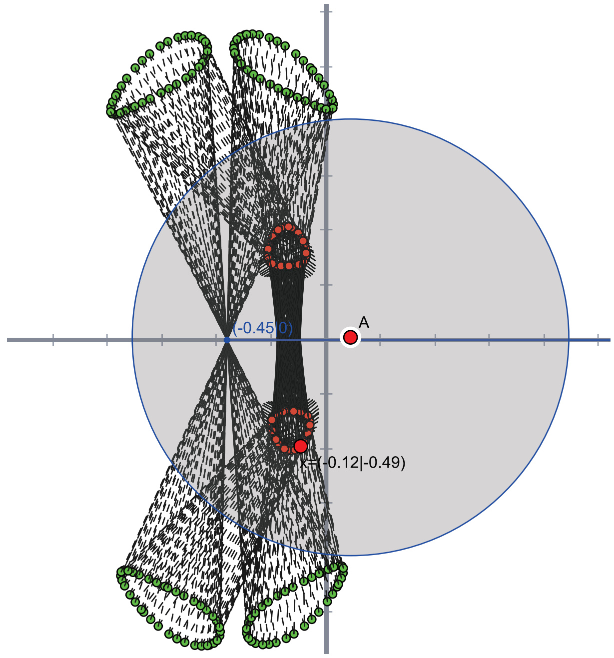

Fig. 6.4 More complex behaviour for a ray and circle.

Local convergence when $0 \leq \alpha < 1$

In this section we show, at least when X is finite dimensional, that for $0 \leq \alpha < 1$ local convergence at each of the feasible points is a consequence of the following theorem from the stability theory of difference equations.

Theorem 6.1 (Perron, [11], Corollary 4.7.2, page 104). *If $f : \boldsymbol{N} \times \boldsymbol{R}^m \longrightarrow \boldsymbol{R}^m$ satisfies,*

$$\lim_{x \to 0} \frac{\|f(n,x)\|}{\|x\|} = 0,$$

uniformly in n and M is a constant $m \times m$ matrix all of whose eigenvalues lie inside the unit disk, then the zero solution (provided it is an isolated solution; that is, there is a neighbourhood of 0 containing no other solution) of the difference equation,

$$x_{n+1} = Mx_n + f(n,x_n),$$

is exponentially asymptotically stable*; that is, there exists $\delta > 0$, $K > 0$ and $\zeta \in (0,1)$ such that if $\|x_0\| < \delta$ then $\|x_n\| \leq K\|x_0\|\zeta^n$.*

To apply this in our context, we begin by noting that the operator $T := T_{S,L}$ is differentiable at any non-zero point y with derivative the linear operator,

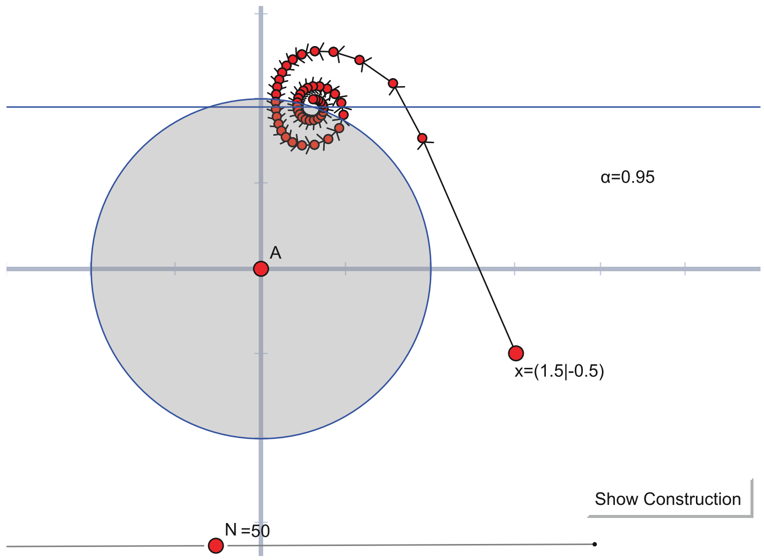

Fig. 6.5 Case with $\alpha = 0.95$**.**

$$T'_y(x) = \left\langle \left(\frac{2}{\|y\|} - 1\right) x - 2\frac{\langle x, y\rangle}{\|y\|^3} y,\ a \right\rangle a + \left(1 - \frac{1}{\|y\|}\right) x + \frac{\langle x, y\rangle}{\|y\|^3} y.$$

By symmetry it suffices to consider local convergence at the unique fixed point of $T_{S,L}$ lying in the positive open half-space $\langle x, a\rangle > 0$; namely, $p := \sqrt{1-\alpha^2}a + \alpha b$. Observing that, p is an isolated fixed point of $T_{S,L}$ (see the discussion before (6.1)) and, using $\|p\| = 1$ and $\langle p, a\rangle = \sqrt{1-\alpha^2}$, we obtain,

$$T'_p(x) = \left\langle x,\ \alpha^2 a - \alpha\sqrt{1-\alpha^2} b \right\rangle a \ + \ \left\langle x, \alpha\sqrt{1-\alpha^2} a + \alpha^2 b \right\rangle b.$$

Which, relative to the basis B corresponds to the $n \times n$ matrix,

$$\begin{pmatrix} \alpha^2 & -\alpha\sqrt{1-\alpha^2} & 0 & \cdots & 0 \\ \alpha\sqrt{1-\alpha^2} & \alpha^2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

From this we immediately deduce that the only points in the spectrum of T'_p are the eigenvalues 0, and $\alpha^2 \pm i\alpha\sqrt{1-\alpha^2}$.

Introducing the change of variable $\xi := x - p$ and defining f by,

$$f(\xi) := T_{S,L}(p+\xi) \; - \; T_{S,L}(p) \; - \; T_p'(\xi),$$

we see that the Douglas-Rachford scheme becomes,

$$\xi_{n+1} = T_{S,L}(p+\xi_n) - p = T_{S,L}(p+\xi_n) - T_{S,L}(p) = T_p'(\xi_n) + f(\xi_n).$$

Further, by the very definition of the derivative we have,

$$\lim_{\xi \to 0} \frac{\|f(\xi)\|}{\|\xi\|} = \lim_{\xi \to 0} \frac{\left\|T_{S,L}(p+\xi) - T_{S,L}(p) \; - \; T_p'(\xi)\right\|}{\|\xi\|} \; = \; 0.$$

Thus, all the conditions of Perron's theorem are satisfied, provided T_p' has its spectrum contained in the open unit disk. But, this follows immediately since both non-zero eigenvalues have modulus equal to $\alpha \; < \; 1$, establishing that locally the Douglas-Rachford scheme converges exponentially to $\xi = 0$; that is, to $x = p$. Thus, we have proved,

Theorem 6.2. *If $0 \leq \alpha < 1$ then the Douglas-Rachford scheme is locally convergent at each of the points $\pm\sqrt{1-\alpha^2}a + \alpha b$.*

Remark 6.3 (Explaining the spiral). It is also worthy of note that the non-zero eigenvalues both have arguments whose cosines have absolute value α, so 'spiraling', as illustrated in Figure 6.5, should be less rapid the larger the value of α, an observation born out by experiment. It should also be noted that when $\alpha = 1$; that is, the line L is tangential to the sphere S, Perron's theorem fails to apply, as in this case T_p' has eigenvalues lying on the unit circle. Indeed the conclusion of theorem 6.2 is false as we will shortly show. ◇

Convergence when $\alpha = 0$

We show that starting from any initial point with $x_0(1) > 0$ the Douglas-Rachford scheme converges to the feasible point $a = (1, 0, 0, \cdots, 0)$, as illustrated in Figure 6.6. In this case the scheme (6.1), (6.2), (6.3) reduces to,

$$\begin{aligned} &x_{n+1}(1) = x_n(1)/\rho_n, \quad \text{and} \\ &x_{n+1}(k) = (1 - 1/\rho_n)\, x_n(k), \quad \text{for } k = 2, \; \cdots, N, \end{aligned}$$

with $\rho_n = \|x_n\| = \sqrt{x_n(1)^2 + \cdots + x_n(N)^2} \geq x_n(1) > 0.$

Proposition 6.1. *If $\rho_n > 1$ then $\rho_{n+1}^2 < \rho_n^2$.*

Proof. We may estimate as follows.

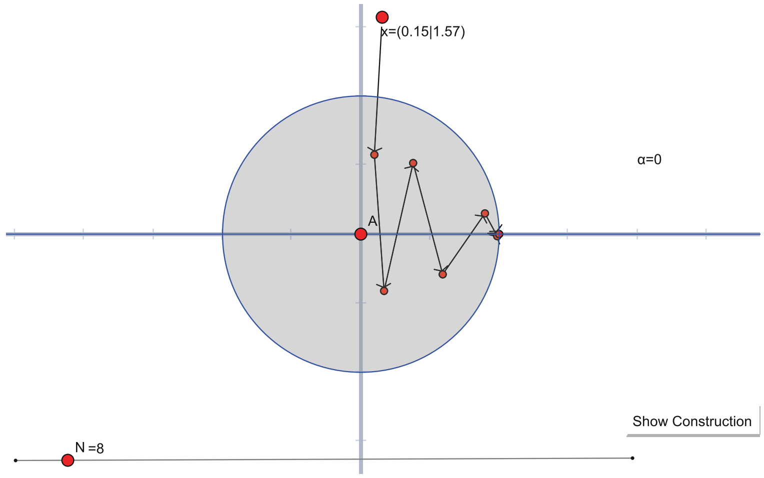

Fig. 6.6 Case with $\alpha = 0$.

$$
\begin{aligned}
\rho_{n+1}^2 &= \frac{x_n(1)^2}{\rho_n^2} + \left(1 - \frac{1}{\rho_n}\right)^2 \sum_{k=2}^{N} x_n(k)^2 \\
&= \frac{x_n(1)^2 + x_n(2)^2 + \cdots + x_n(N)^2}{\rho_n^2} + \left(1 - \frac{2}{\rho_n}\right) \sum_{k=2}^{N} x_n(k)^2 \\
&= 1 + \left(1 - \frac{2}{\rho_n}\right) \sum_{k=2}^{N} x_n(k)^2 \\
&\le 1 + \left(1 - \frac{2}{\rho_n} + \frac{1}{\rho_n^2}\right) \sum_{k=2}^{N} x_n(k)^2 \\
&= 1 + \left(1 - \frac{1}{\rho_n}\right)^2 \sum_{k=2}^{N} x_n(k)^2 \\
&\le 1 + \left(1 - \frac{1}{\rho_n}\right)^2 \rho_n^2 \\
&= 1 + (\rho_n - 1)^2 \\
&= \rho_n^2 + 2(1 - \rho_n) \\
&< \rho_n^2, \quad \text{as } \rho_n > 1.
\end{aligned}
$$

∎

Corollary 6.1. *If $\rho_n > 1$ for all n then $\rho_n \longrightarrow 1$.*

Proof. By the above proposition, the ρ_n are decreasing and so converge to some limit $\rho \geq 1$. But then, taking limits in $\rho_{n+1}^2 \leq \rho_n^2 + 2(1-\rho_n)$ leads to $\rho \leq 1$, so $\rho = 1$. ■

Proposition 6.2. *If $\rho_n \leq 1$ then so too is $\rho_{n+1} \leq 1$.*

Proof. From the first three lines in the proof of the above proposition we have

$$\begin{aligned}
\rho_{n+1}^2 &= 1 + \left(1 - \frac{2}{\rho_n}\right) \sum_{k=2}^{N} x_n(k)^2 \\
&\leq 1 - \sum_{k=2}^{N} x_n(k)^2, \quad \text{provided } \rho_n \leq 1 \\
&\leq 1.
\end{aligned}$$

■

Theorem 6.3. *If $\alpha = 0$ and the initial point has $x_0(1) > 0$ then the Douglas-Rachford scheme converges to the feasible point $(1, 0, 0, \cdots, 0)$.*

Proof. In case $\rho_n > 1$ for all n then, by the above corollary, $\rho_n \to 1$, so by the recurrence $x_n(k) \to 0$ for $k = 2, \cdots, N$ and $x_n \to (1, 0, 0, \cdots, 0)$.

On-the-other-hand, if this is not the case then there is a smallest n_0 with $\rho_{n_0} \leq 1$ and then either $\rho_{n'} = 1$ for some $n' \geq n_0$, in which case we have $x_{n'+1}(k) = 0$ for $k = 2, \cdots, N$, so $x_{n'+1} = (1, 0, \cdots, 0)$ and we have arrived at the feasible point after a finite number of steps, or alternatively from the last proposition $\rho_n < 1$ for all $n \geq n_0$. Consequently, the sequence $(x_n(1))_{n=n_0}^{\infty}$ is strictly increasing (hence convergent to some $x(1) \leq 1$) and so for $n \geq n_0$ we have $\rho_n \geq x_n(1) \geq x_{n_0} > 0$. But then, for each integer $k \geq 2$ and $n \geq n_0$ we see from the recurrence that,

$$\begin{aligned}
\left|\frac{x_{n+1}(k)}{x_{n+1}(1)}\right| &= (1-\rho_n)\left|\frac{x_n(k)}{x_n(1)}\right| \\
&\leq (1 - x_{n_0}(1))\left|\frac{x_n(k)}{x_n(1)}\right|.
\end{aligned}$$

Hence, $\dfrac{x_n(k)}{x_n(1)}$ converges to 0 and we conclude that $x_n \longrightarrow (1, 0, \cdots, 0)$. ■

The tangential case when $\alpha = 1$

When $\alpha = 1$ the only feasible point is $b = (0, 1, 0, \cdots, 0)$, however we show that starting from an initial point with $x_0(1) > 0$ the Douglas-Rachford scheme converges to a point $\hat{y}b := (0, \hat{y}, 0, \cdots, 0)$ with $\hat{y} > 1$, whose projection onto either S or L is the feasible point. The following result will be needed.

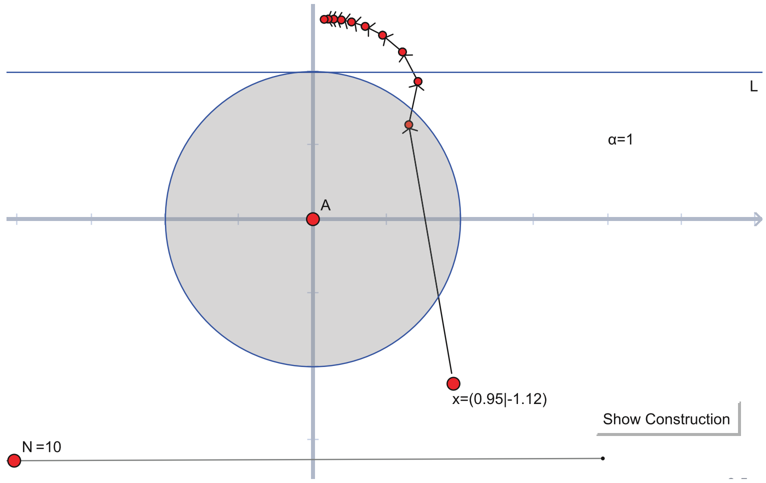

Fig. 6.7 Case with $\alpha = 1$.

Proposition 6.3. *If $\rho_n > 2$ then $\rho_{n+1} \le \rho_n$.*

Proof. The proof is similar to that of Proposition 6.1. We may estimate as follows.

$$
\begin{aligned}
\rho_{n+1}^2 &= \frac{x_n(1)^2}{\rho_n^2} + \left(\left(1-\frac{1}{\rho_n}\right)x_n(2)+1\right)^2 + \left(1-\frac{1}{\rho_n}\right)^2 \sum_{k=3}^{N} x_n(k)^2 \\
&= \frac{x_n(1)^2 + x_n(2)^2 + \cdots + x_n(N)^2}{\rho_n^2} \\
&\qquad + \left(1-\frac{2}{\rho_n}\right)\sum_{k=2}^{N} x_n(k)^2 + 2\left(1-\frac{1}{\rho_n}\right)x_n(2)+1 \\
&= 2 + \left(1-\frac{2}{\rho_n}\right)\sum_{k=2}^{N} x_n(k)^2 + 2\left(1-\frac{1}{\rho_n}\right)x_n(2) \\
&\le 2 + \left(1-\frac{2}{\rho_n}\right)\rho_n^2 + 2\left(1-\frac{1}{\rho_n}\right)\rho_n, \quad \text{as } \rho_n > 2 \\
&= \rho_n^2.
\end{aligned}
$$

■

To show the asserted behaviour, we begin by noting that from the recurrence,

$$
x_{n+1}(2) = x_n(2) \;+\; 1 - \frac{x_n(2)}{\rho_n} \ge x_n(2), \tag{6.6}
$$

since $\dfrac{x_n(2)}{\rho_n} \le 1$. Thus, the $x_n(2)$ are increasing and so either they converge to a finite limit, $\hat{y}$ say, or they diverge to $+\infty$.

In the first case, taking limits in the above equation (6.6) yields $\hat{y} = \lim_n x_n(2) = \lim_n \rho_n \ge 0$ and so $x_n \longrightarrow (0, \hat{y}, 0, \cdots, 0)$. To see that $\hat{y} > 1$ we argue as follows. We have $x_n(1) \to 0$. But (6.1) shows $x_{n+1}(1) = x_n(1)/\rho_n$ so we must have $\lim_n \rho_n > 1$.

To show that the second, divergent, case is impossible we appeal to Proposition 6.3. to deduce that if the $x_n(2)$ diverges to $+\infty$, we must have for all sufficiently large n that $2 < x_n(2) \le \rho_n$ and so eventually the ρ_n are decreasing and hence convergent to a finite limit which is necessarily greater than or equal to $\limsup_n x_n(2)$ which cannot therefore be infinite; a contradiction.

Consequently, we have proved,

Theorem 6.4. *When L is tangential to S at b (that is, when $\alpha = 1$), starting from any initial point with $x_0(1) \neq 0$, the Douglas-Rachford scheme converges to a point $\hat{y}b$ with $\hat{y} > 1$.*

This is consistent with the behaviour in the convex case [5,12].

Behaviour in the infeasible case when $\alpha > 1$

Satisfyingly, when there are no feasible solutions, starting from any point off the singular manifold, the Douglas-Rachford scheme diverges. More precisely,

Theorem 6.5. *If there are no feasible solutions (that is, when $\alpha > 1$) then starting from any initial point with $x_0(1) \neq 0$, we have that $x_n(2)$ and hence ρ_n diverge to $+\infty$ at a linear or faster rate in the sense that* $\liminf_n x_{n+1}(2) - x_n(2) \geqslant \alpha - 1$.

Proof. From the recursion we have,

$$\begin{aligned} x_{n+1}(2) - x_n(2) &= \alpha - \frac{x_n(2)}{\rho_n} \\ &> \alpha - 1, \quad \text{as } x_n(2) < \rho_n \\ &> 0, \end{aligned}$$

from which the result follows. ∎

It is also worth noting that, as a consequence of the above theorem and the recurrence, $x_n(1) \to 0$ and so asymptotically the iterates approach the hyperplane $\langle x, a \rangle = 0$.

Behaviour on the singular manifold, $\langle x, a \rangle = 0$

Here we consider the iterates of a non-zero initial point with $x_0(1) = 0$ and so $x_n(1) = 0$ for all n.

We again distinguish the cases; $\alpha = 0, 0 < \alpha < 1, \alpha = 1$. The case $\alpha > 1$ having already been dealt with in the previous section.

When $\alpha = 0$ it is readily seen that for any non-zero point x in the singular manifold we have $T_{S,L}(x) = \left(1 - \frac{1}{\|x\|}\right)x$. If $\|x\| = 1$ then the first iteration yields $x_1 = 0 \notin D_{T_{S,L}}$, so subsequent iterates are not defined. At points with $\|x\| < 1$ we see that $T_{S,L}$ has period two (that is, $T^2_{S,L}(x) = x$), while for $\|x\| > 1$ we have $T^2_{S,L}(x) = \left(1 - \frac{2}{\|x\|}\right)x$, so again the scheme breaks down as above, but after two iterations if $\|x\| = 2$.

We observe that the iterates of any non-zero point on the line $\{x : x = \lambda b,\ \lambda \in \mathbf{R}\}$ remain on this line and that when $\alpha = 1$ (that is, L is tangential to S at b) all points on the open half line corresponding to $\lambda > 0$ remain fixed under $T_{S,L}$.

In the other cases the scheme exhibits periodic behaviour when rational commensurability is present, while in the absence of such commensurability the behaviour may be quite chaotic. To make this precise we need to consider interval-valued mappings to deal with the jump at the origin. Luckily, the work in [1, 2] shows that various interval mapping analogues of Sharkovskii's theorem—"period three implies chaos"—are applicable. The interval mapping is needed to deal with the multivalued nature of the projection P_S at zero.

Remark 6.4 (Hilbert space analogues). It is not essential that X be finite dimensional for any of the arguments in sections 3, 4 and 5, since the iterates are tracked by a finite number of coordinates. However, since convergence (to zero) in the other dimensions is only coordinate wise, we can in general only guarantee weak convergence of the iterates. ◇

Some final remarks

A wealth of experimental evidence, using both *Maple* and the dynamic geometry package *Cinderella*, leads to the conclusion that the basin of attraction for $p = \sqrt{1-\alpha^2}a + hb$ is the open half space $\{x : \langle x, a\rangle > 0\}$ – the largest region possible. See also http://www.carma.newcastle.edu.au/~jb616/expansion.html.

Moreover, we found that for stable computation in Cinderella it was necessary to have access to precision beyond *Cinderella's* built-in double precision. This was achieved by taking input directly from *Maple*. We illustrate in Figure 6.8 which show various spurious red points on the left and accurate data on the right. The figures show the effect of roughly ten steps of the Douglas-Rachford iteration for 400 different starting points—where the points are coloured by their original distance from the vertical axis with red closest.

However, we are as yet unable to furnish a proof of this, leaving open the following conjecture:

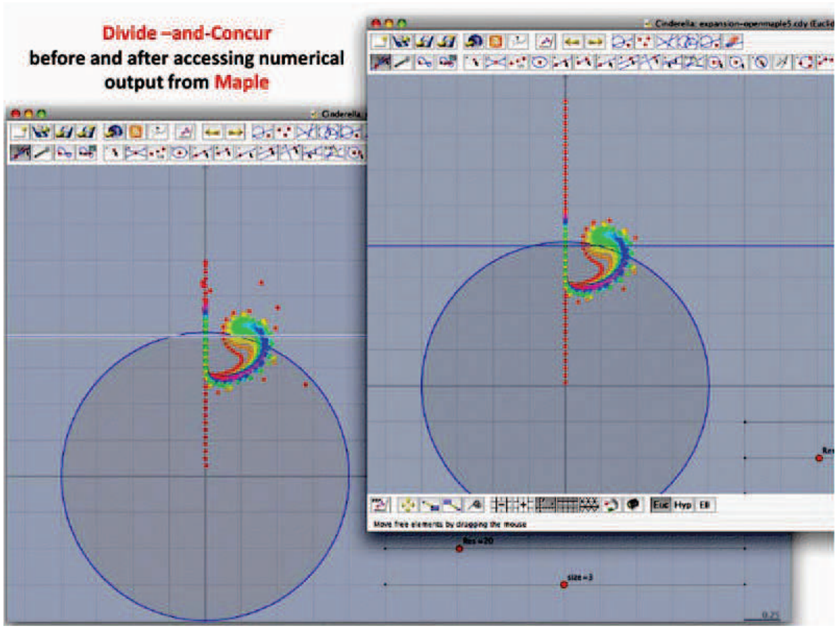

Fig. 6.8 Multiple iterations in Cinderella.

Conjecture 6.1. In the simple example of a sphere and a line with two intersection points, the basins of attraction are the two open half-spaces forming the complement of the singular manifold.

Remark 6.5 (The case of a sphere and a proper affine subset of X). If we replace the line L by a proper affine subset, say $A := \{\lambda_1 a_1 + \lambda_2 a_2 + \cdots \lambda_K a_K + \alpha b : \lambda_1, \cdots, \lambda_K \in \mathbf{R}\}$, where $0 \le \alpha$, $1 < K < N$, and $a_1, a_2, \cdots, a_K, b$ are mutually orthogonal norm one elements, then when $\alpha < 1$ the feasible points are no longer isolated, so Theorem 6.1 no longer applies, indeed local convergence in the sense described above is impossible. Nonetheless, *all our results appropriately viewed continue to hold* and we shall sketch the argument. Details will be given elsewhere.

Indeed, if for any non-feasible point $q \neq 0$ we let $Q := A_0^{\perp} + \mathbf{R}q$, where $A_0^{\perp}$ is the orthogonal complement of the subspace $A_0 := A - \alpha b$, then we see that for any initial point $x_0 \in Q$ the sequence of iterates, $x_n = T_{S,A}^n(x_0)$ remains confined to the subspace Q. So, if the Douglas-Rachford scheme converges it will converge to a point in $S \cap A \cap Q$. Further, the fixed points of $T_{S,A}|_Q$ consists of two isolated points comprising $S \cap A \cap Q$; namely, $p = (kq(1), kq(2), \cdots, kq(K), \alpha, 0, \cdots, 0)$, where,

$$k := \pm\sqrt{\frac{1-\alpha^2}{q(1)^2 + \cdots + q(K)^2}}.$$

And so we have 'local convergence' in the following sense. For either feasible point $p \in S \cap A \cap Q$ there is a neighbourhood, N_p of p in the subspace Q such that starting from any point x_0 in N_p the iterates converge to p.

Additionally, we may derive similar conclusions to those obtained above in the cases when $\alpha = 0$, $\alpha = 1$ and $\alpha > 1$. Further, in this case the singular manifold is the subspace $A_0^{\perp}$. $\diamond$

In conclusion, our analysis sheds some significant light on the behaviour of non-convex Douglas-Rachford schemes but much remains to be studied.

Example 6.2 (Other regions). For example, we observe that neither convexity nor so much symmetry is essential to the behaviour exhibited in Theorem 6.1. Figure 6.9 shows the situation for a line and a nonconvex p-sphere, where $S(p) := \{(x,y) \colon |x|^p + |y|^p = 1\}$, in the plane. The details of such analysis remain to be performed. $\diamond$

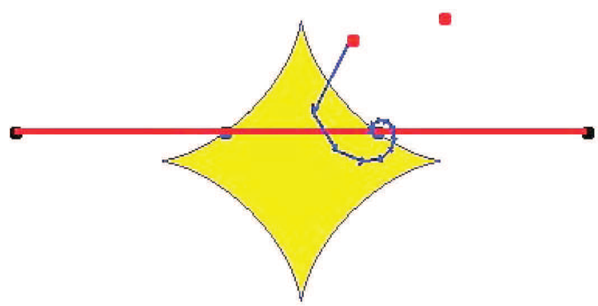

Fig. 6.9 Spiralling with the $1/2$-sphere.

Acknowledgements This research was supported by the Australian Research Council. We also express our thanks to Chris Maitland, Matt Skerritt and Ulli Kortenkamp for helping us exploit the full resources of *Cinderella*.

References

1. Andres, J., Fürst, T., Pastor, K.: Full analogy of Sharkovskii's theorem for lower semicontinuous maps. J. Math. Anal. Appl. **340**, 1132–1144, (2008)
2. Andres, J., Pastor, K., Šnyrychová: A multivalued version of Sharkovskii's theorem holds with at most two exceptions. J. Fixed Point Theory Appl. **2**, 153–170 (2007)
3. Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. SIAM Review **38**, 367–426 (1996)

4. Bauschke, H.H., Combettes, P.L., Luke, D.R..: Finding best approximation pairs relative to two closed convex sets in Hilbert spaces. J. Approx. Theory **127**, 178–192 (2004)
5. Bauschke, H.H., Combettes, P.L., Luke, D.R.: Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization. J. Opt. Soc. Amer. A **19**, 1334–1345 (2002)
6. Bauschke, H.H., Combettes, P.L., Luke, D.R.: A strongly convergent reflection method for finding the projection onto the intersection of two closed convex sets in a Hilbert space. J. Approx. Theory **141**, 63–69 (2006)
7. Douglas, J., Rachford, H.H.: On the numerical solution of heat conduction problems in two or three space variables. Trans. Am. Math. Soc. **82**, 421–439 (1956)
8. Elser, V., Rankenburg, I., Thibault, P.: Searching with iterated maps. Proceedings of the National Academy of Sciences **104**, 418–423 (2007)
9. Gravel, S., Elser, V.: Divide and concur: A general approach constraint satisfaction. Phys. Rev. E **78** 036706, pp. 5 (2008), `http://link.aps.org/doi/10.1103/PhysRevE.78.036706`
10. Pierra, G.: Eclatement de contraintes en parallèle pour la miniminisation d'une forme quadratique. Lecture Notes in Computer Science, Springer Verlag, **41** 200–218 (1976)
11. Lakshmikantham, V., Trigiante, D.: Theory of Difference Equations - Numerical Methods and Applications. Marcel Dekker (2002)
12. Lions, P.-L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**, 964–979 (1979)