Weak convergence on Douglas-Rachford method

B. F. Svaiter* †

Abstract

We prove that the sequences generate by the Douglas-Rachford method converge weakly to a solution of the inclusion problem

2000 Mathematics Subject Classification: 47H05, 49J52, 47N10.

Key words: Douglas-Rachford, weak convergence. Douglas-Rachford method, weak convergence.

From now on H is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and associated norm $\| \cdot \|$. A *point-to-set operator* $T : H \rightrightarrows H$ is a relation $T \subset H \times H$ and for $x \in H$,

$$T(x) = \{v \in H \mid (x, v) \in T\}.$$

An operator $T : H \rightrightarrows H$ is *monotone* if

$$\langle x_1 - x_2, v_1 - v_2 \rangle \geq 0, \forall (x_1, v_1), (x_2, v_2) \in T$$

and it is *maximal monotone* if it is monotone and maximal in the family of monotone operators with respect to the partial order of inclusion.

Let A, B be maximal monotone operators in H. Consider the problem of finding $x \in H$ such that

$$0 \in A(x) + B(x). \tag{1}$$

Douglas-Rachford method (with sumable residual error) works as follows. Let $\{\alpha_k\}$, $\{\beta_k\}$ be sequences of positive error tolerance such that

$$\sum_{k=1}^{\infty} \alpha_k < \infty, \quad \sum_{k=1}^{\infty} \beta_k < \infty.$$

*IMPA, Estrada Dona Castorina 110, 22460-320 Rio de Janeiro, Brazil (`benar@impa.br`) tel: 55 (21) 25295112, fax: 55 (21)25124115.

†Partially supported by CNPq grants 300755/2005-8, 475647/2006-8 and by PRONEX-Optimization

Take $\lambda > 0$, $(x_0, b_0) \in B$ and for $k = 1, 2, \ldots$

(a) Find (y_k, a_k) such that

$$(y_k, a_k) \in A, \quad \|y_k + \lambda a_k - (x_{k-1} - \lambda b_{k-1})\| \leq \alpha_k \tag{2}$$

(b) Find (x_k, b_k) such that

$$(x_k, b_k) \in B, \quad \|x_k + \lambda b_k - (y_k + \lambda b_{k-1})\| \leq \beta_k \tag{3}$$

Existence of (y_k, a_k), (x_k, b_k) as above follows from Minty's Theorem [6]. From now on $\{(x_k, b_k)\}_{k=0}^{\infty}$, $\{(y_k, a_k)\}_{k=1}^{\infty}$ are sequences generated by the above algorithm. This algorithm was proposed by Douglas and Rachford [1] for the case where A and B are linear, and was extended to arbitrary maximal monotone operators by Lions and Mercier [5]. These later authors proved that $x_k + \lambda b_k$ converges weakly to a point $\bar{z}$ such that, for some $\bar{x}$, $\bar{b}$,

$$\bar{z} = \bar{x} + \lambda \bar{b}, \quad \bar{b} \in B(\bar{x}), \quad -\bar{b} \in A(\bar{x}).$$

Our aim is to prove the next theorem

Theorem 1. *If A, B are maximal monotone operators and the solution set of*

$$0 \in A(x) + B(x)$$

is non-empty, then the sequences $\{(x_k, b_k)\}$ and $\{(y_k, a_k)\}$ generated by Douglas-Rachford method converges weakly to some $(\bar{x}, \bar{b})$ and $(\bar{x}, -\bar{b})$ respectively, such that

$$\bar{b} \in B(\bar{x}), \qquad -\bar{b} \in A(\bar{x})$$

and hence, $0 \in A(\bar{x}) + B(\bar{x})$.

For each k, there exists unique pairs $(\hat{y}_k, \hat{a}_k)$, $(\hat{x}_k, \hat{b}_k)$ such that

$$(\hat{y}_k, \hat{a}_k) \in A, \quad \hat{y}_k + \lambda \hat{a}_k = x_{k-1} - \lambda b_{k-1} \tag{4}$$

$$(\hat{x}_k, \hat{b}_k) \in B, \quad \hat{x}_k + \lambda \hat{b}_k = \hat{y}_k + \lambda b_{k-1} \tag{5}$$

Existence of $(\hat{y}_k, \hat{a}_k)$, $(\hat{x}_k, \hat{b}_k)$ as above follows from Minty's Theorem [6]. Direct use of (4) and (5) shows that for $k = 1, 2, \ldots$

$$x_{k-1} - \hat{x}_k = \lambda(\hat{a}_k + \hat{b}_k), \quad \lambda(b_{k-1} - \hat{b}_k) = \hat{x}_k - \hat{y}_k. \tag{6}$$

and

$$\hat{x}_k - \hat{y}_k + \lambda(\hat{a}_k + \hat{b}_k) = \lambda(\hat{a}_k + b_{k-1}) = x_{k-1} - \hat{y}_k. \tag{7}$$

Using (2), (4) and the monotonicity of A we conclude that

$$\begin{aligned}\|y_k - \hat{y}_k\|^2 + \lambda^2 \|a_k - \hat{a}_k\| \leq & \|y_k - \hat{y}_k + \lambda(a_k - \hat{a}_k)\|^2 \\ = & \|y_k + \lambda a_k - (x_{k-1} - \lambda b_{k-1})\|^2 \leq \alpha_k^2 \end{aligned} \tag{8}$$

Using (5), (3), the monotonicity of B, triangle inequality and the above inequality we have

$$\begin{aligned}\|x_k - \hat{x}_k\|^2 + \lambda^2 \|b_k - \hat{b}_k\| \leq & \left\| x_k - \hat{x}_k + \lambda(b_k - \hat{b}_k) \right\|^2 \\ \leq & \|x_k + \lambda b_k - (\hat{y}_k + \lambda b_{k-1})\|^2 \\ \leq & (\|x_k + \lambda b_k - (y_k + \lambda b_{k-1})\| + \|\hat{y}_k - y_k\|)^2 \\ \leq & (\alpha_k + \beta_k)^2 \end{aligned} \tag{9}$$

We will use in $H \times H$ the inner product $\langle \cdot, \cdot \rangle_\lambda$ and associated norm $\| \cdot \|_\lambda$,

$$\langle (x, v), (x', v') \rangle_\lambda = \langle x, x' \rangle + \lambda^2 \langle v, v' \rangle, \quad \|p\|_\lambda = \sqrt{\langle p, p \rangle}_\lambda.$$

Note that $H \times H$ endowed with the inner product $\langle \cdot, \cdot \rangle_\lambda$ is a Hilbert space isomorphic to $H \times H$ endowed with the canonical inner product $\langle (x, v), (x', v') \rangle = \langle x, x' \rangle + \langle v, v' \rangle$. Hence, the strong/weak topologies of both spaces are the same. To simplify the exposition, define

$$p_k = (x_k, b_k), \quad k = 0, 1, \dots, \qquad \hat{p}_k = (\hat{x}_k, \hat{b}_k), \quad k = 1, 2, \dots \tag{10}$$

We have just proved in (9) that

$$\|p_k - \hat{p}_k\|_\lambda \leq \alpha_k + \beta_k, \qquad k = 1, 2, \dots \tag{11}$$

The *extended solution set* [2] of problem (1) is

$$S(A, B) = B \cap -A = \{(z, w) \mid (z, w) \in B, (z, -w) \in A\}$$

Lemma 2. *If $p \in S(A, B)$ then,*

$$\|p_{k-1} - p\|_\lambda^2 \geq \|\hat{p}_k - p\|_\lambda^2 + \|x_{k-1} - \hat{y}_k\|^2 = \|\hat{p}_k - p\|_\lambda^2 + \lambda^2 \|\hat{a}_k + b_{k-1}\|^2$$

for $k = 1, 2, \dots$.

Proof. First note that $p = (x, b)$ with $b \in B(x)$ and $-b \in A(x)$. Using the first equality in (6), the monotonicity of B and the monotonicity of A we conclude that

$$
\begin{aligned}
\langle x_{k-1} - \hat{x}_k, \hat{x}_k - x\rangle =& \lambda\langle \hat{a}_k + \hat{b}_k, \hat{x}_k - x\rangle \\
=& \lambda\Big[\langle \hat{a}_k + b, \hat{x}_k - x\rangle + \langle \hat{b}_k - b, \hat{x}_k - x\rangle\Big] \\
\geq& \lambda\langle \hat{a}_k + b, \hat{x}_k - x\rangle \\
=& \lambda\Big[\langle \hat{a}_k + b, \hat{x}_k - \hat{y}_k\rangle + \langle \hat{a}_k + b, \hat{y}_k - x\rangle\Big] \geq \lambda\langle \hat{a}_k + b, \hat{x}_k - \hat{y}_k\rangle
\end{aligned}
$$

Using the above inequality, the second equality in (6) and (10) we have

$$
\begin{aligned}
\langle p_{k-1} - \hat{p}_k, \hat{p}_k - p\rangle_\lambda =& \langle x_{k-1} - \hat{x}_k, \hat{x}_k - x\rangle + \lambda^2\langle b_{k-1} - \hat{b}_k, \hat{b}_k - b\rangle \\
\geq& \lambda\langle \hat{a}_k + b, \hat{x}_k - \hat{y}_k\rangle + \lambda\langle \hat{x}_k - \hat{y}_k, \hat{b}_k - b\rangle \\
=& \lambda\langle \hat{x}_k - \hat{y}_k, \hat{a}_k + \hat{b}_k\rangle
\end{aligned}
$$

Using (10) and (6) we have

$$
\|p_{k-1} - \hat{p}_k\|_\lambda^2 = \lambda^2\|\hat{a}_k + \hat{b}_k\|^2 + \|\hat{x}_k - \hat{y}_k\|^2
$$

Therefore,

$$
\begin{aligned}
\|p_{k-1} - p\|_\lambda^2 =& \|p_{k-1} - \hat{p}_k\|_\lambda^2 + 2\langle p_{k-1} - \hat{p}_k, \hat{p}_k - p\rangle_\lambda + \|\hat{p}_k - p\|_\lambda^2 \\
\geq& \lambda^2\|\hat{a}_k + \hat{b}_k\|^2 + \|\hat{x}_k - \hat{y}_k\|^2 + 2\lambda\langle \hat{x}_k - \hat{y}_k, \hat{a}_k + \hat{b}_k\rangle \\
&+ \|\hat{p}_k - p\|_\lambda^2 \\
=& \|\hat{p}_k - p\|_\lambda^2 + \|\hat{x}_k - \hat{y}_k + \lambda(b_k + a_k)\|^2
\end{aligned}
$$

To end the proof, use the above inequality and (7). □

Corollary 3. *The sequence $\{p_k\}$ is Quasi-Fejer convergent to $S(A, B)$. Therefore it has at most one weak cluster point in this set, and it is bounded if $S(A, B) \neq \emptyset$.*

Proof. Take $p \in S(A, B)$. Using (11) and Lemma 2 we have

$$
\|p_k - p\| \leq \|p_k - \hat{p}_k\| + \|\hat{p}_k - p\| \leq \alpha_k + \beta_k + \|p_{k-1} - p\|
$$

which proves that $\{p_k\}$ is Quasi-Fejer convergent to $S_\lambda(A, B)$. The last part of the corollary follows from this result and also from Opial's Lemma [7]. □

Proof of Theorem 1. We are assuming that problem (1) has a solution. Therefore, $S(A,B) \neq \emptyset$. Using Corollary 3 we conclude that $\{p_k\}$ is bounded. Take

$$p \in S(A,B)$$

and let

$$M = 1 + \sup \|p_k - p\|_\lambda$$

Using Lemma 2 we have, for $k = 1, 2, \dots$

$$\|p_k^* - p\|_\lambda^2 \leq \|p_{k-1} - p\|_\lambda^2 - \|x_{k-1} - y_k^*\|^2.$$

Therefore, using also the concavity of $t \mapsto \sqrt{t}$ we conclude that

$$\|p_k^* - p\|_\lambda \leq \|p_{k-1} - p\|_\lambda - \frac{1}{2M}\|x_{k-1} - \hat{y}_k\|^2$$

Hence, combining the above inequality with (11) and triangle inequality we obtain

$$\begin{aligned} \|p_k - p\|_\lambda \leq & \|p_k - \hat{p}_k\|_\lambda + \|\hat{p}_k - p\|_\lambda \\ \leq & (\alpha_k + \beta_k) + \|p_{k-1} - p\|_\lambda - \frac{1}{2M}\|x_{k-1} - \hat{y}_k\|^2 \end{aligned}$$

Adding the above inequality for $k = 1, 2, \dots, n$ we conclude that

$$\frac{1}{2M}\sum_{k=1}^{n} \|x_{k-1} - \hat{y}_k\|^2 \leq \|p_0 - p\|_\lambda + \sum_{k=1}^{n}(\alpha_k + \beta_k)$$

Therefore

$$\sum_{k=1}^{\infty} \|x_{k-1} - \hat{y}_k\|^2 < \infty$$

and, using also (7), we conclude that

$$\lim_{k\to\infty} x_{k-1} - \hat{y}_k = \lim_{k\to\infty} \hat{a}_k + b_{k-1} = 0$$

Using (8) we have

$$\lim_{k\to\infty} y_k - \hat{y}_k = \lim_{k\to\infty} a_k - \hat{a}_k = 0.$$

Therefore

$$\lim_{k\to\infty} x_{k-1} - y_k = \lim_{k\to\infty} a_k + b_{k-1} = 0 \tag{12}$$

and sequence $\{(y_k, b_k)\}$ is also bounded.

Since the sequence $\{p_k\}$ is bounded and $H \times H$ (endowed with the norm $\|\cdot\|_\lambda$) is reflexive, this sequence has weak cluster points. Let $(\bar{x}, \bar{b})$ be a weak cluster point of the *bounded sequence* $\{p_k = (x_k, b_k)\}$. Using again the fact that $H \times H$ is reflexive, we conclude that there exists a subsequence $\{(x_{k_j}, b_{k_j})\}$ converging weakly to $(\bar{x}, \bar{b})$ and hence

$$x_{k_j} \xrightarrow{w} \bar{x}, \quad b_{k_j} \xrightarrow{w} \bar{b}, \qquad \text{as } j \to \infty,$$

which, together with (12) implies also that

$$y_{k_j - 1} \xrightarrow{w} \bar{x}, \quad a_{k_j - 1} \xrightarrow{w} -\bar{b}, \qquad \text{as } j \to \infty.$$

Using the two above equations, (12) and Lemma 5 (see Appendix A) applied to the subsequences $\{(x_{k_j}, b_{k_j}\}$, $\{(y_{k_j-1}, a_{k_j-1}\}$, we conclude that $(\bar{x}, \bar{b}) \in B$, $(\bar{x}, -\bar{b}) \in A$, that is, $(\bar{x}, \bar{b}) \in S(A, B)$.

We have proved that the sequence $\{p_k\}$ has weak cluster points and that all these weak cluster points are in $S(A, B)$. Using these results and Corollary 3 we conclude that $\{p_k\}$ has only one weak cluster point $(\bar{x}, \bar{b})$, and this (weak cluster) point belongs to $S(A, B)$. As $\{p_k\}$ is bounded and $H \times H$ is reflexive, the sequence $\{p_k\}$ converges weakly to such point $(\bar{x}, \bar{b})$, which is equivalent

$$x_k \xrightarrow{w} \bar{x}, \quad b_k \xrightarrow{w} \bar{b}, \qquad \text{as } k \to \infty,$$

To end the proof, use the above equation and (12) to conclude that

$$y_k \xrightarrow{w} \bar{x}, \quad a_k \xrightarrow{w} -\bar{b}, \qquad \text{as } k \to \infty.$$

□

The convergence analysis presented is are based on the framework and techniques introduced in [2, 3], and becomes more intuitive using this framework and results. To make this note shorter, we do not presented this framework here. For historical reasons, here we used the classical sumable error tolerance.

A An auxiliary result

Let X be a real Banach space with topological dual X^*. For $x \in X$, $x^* \in X^*$ we use the notation $\langle x, x^* \rangle = x^*(x)$. An operator $T : X \rightrightarrows X^*$ is monotone if $\langle x - y, x^* - y^* \rangle \geq 0$ for all $(x, x^*), (y, y^*) \in T$, and is maximal monotone if it is monotone and maximal in the family of monotone operators with respect to the partial order of inclusion.

Lemma 4. *Let X be a real Banach space with topological dual X^*. If $T : X \rightrightarrows X$ is maximal monotone, $\{(x_i, x_i^*)\}_{i\in I}$ is a net in T which converges in the weak×weak∗ topology to $(\bar{x}, \bar{x}^*)$, then*

$$\liminf_{i\to\infty} \langle x_i, x_i^*\rangle \geq \langle \bar{x}, \bar{x}^*\rangle.$$

Moreover, if the above inequality holds as an equality, then $(x, x^*) \in T$.

Proof. Let $\varphi : X \times X^* \to \bar{\mathbb{R}}$,

$$\varphi(x, x^*) = \sup_{(y,y^*)\in T} \langle x, y^*\rangle + \langle y, x^*\rangle - \langle y, y^*\rangle$$

The function φ is Fitzpatrick minimal function [4] of T, it is lower semicontinuous in the weak×weak∗ topology, $\varphi(x, x^*) \geq \langle x, x^*\rangle$ for all (x, x^*) and this inequality holds as an equality if and only if $(x, x^*) \in T$. Therefore $\langle x_i, x_i^*\rangle = \varphi(x_i, x_i^*)$ for all $i \in I$ and

$$\liminf_{i\to\infty} \langle x_i, x_i^*\rangle = \liminf_{i\to\infty} \varphi(x_i, x_i^*) \geq \varphi(\bar{x}, \bar{x}^*) \geq \langle \bar{x}, \bar{x}^*\rangle$$

To end the proof, use the fact that φ is bounded below by the duality product and coincide with the duality product if and only if $(x, x^*) \in T$. $\square$

Lemma 5. *Let X be real Banach space. If $T_1, \dots, T_m : X \rightrightarrows X^*$ are maximal monotone operators and $\{(x_{k,i}, x_{k,i}^*)\}_{i\in I}$ are bounded nets such that $(x_{k,i}, x_{k,i}^*) \in T_k$ for all $k = 1, \dots, m$ $i \in I$, and*

$$\begin{array}{ll} x_{k,i} - x_{j,i} \to 0 \quad j, k = 1, \dots, m & \displaystyle\sum_{k=1}^m x_{k,i}^* \to \bar{x}^* \\ x_{k,i} \xrightarrow{w} \bar{x}, & x_{k,i}^* \xrightarrow{w*} \bar{x}_k^* \qquad k = 1, \dots, m \end{array}$$

as $i \to \infty$, then $(\bar{x}, \bar{x}_k^*) \in T_k$ for $k = 1, \dots, m$.

Proof. In view of the above assumptions,

$$\sum_{k=1}^m \bar{x}_k^* = \bar{x}^*$$

Define

$$\alpha_{k,i} = \langle x_{k,i}, x_{k,i}^*\rangle - \langle \bar{x}, \bar{x}_k^*\rangle, \quad k = 1, \dots, m \; i \in I.$$

Direct algebraic manipulations yield

$$\begin{aligned}\sum_{k=1}^{m}\alpha_{k,i} &= \left(\sum_{k=0}^{m}\langle x_{k,i}, x^*_{k,i}\rangle\right) - \langle \bar{x}, \bar{x}^*\rangle\\ &= \left(\sum_{k=0}^{m}\langle x_{k,i}-x_{1,i}, x^*_{k,i}\rangle\right) + \left\langle x_{1,i}, \left(\sum_{k=0}^{m} x^*_{k,i}\right) - \bar{x}^*\right\rangle + \langle x_{1,i}-\bar{x}, \bar{x}^*\rangle,\end{aligned}$$

which ready implies, in view of the assumptions of the lemma, that

$$\lim_{i\to\infty}\sum_{k=1}^{m}\alpha_{k,i} = 0$$

Using the first part of Lemma 4 we have

$$\limsup_{i\to\infty}\alpha_{k,i} \geq 0, \qquad k = 1,\dots,m.$$

Combining the two above equations we conclude that $\lim_{i\to\infty}\alpha_{k,i} = 0$ for $k = 1,\dots,m$, that is

$$\lim_{i\to\infty}\langle x_{k,i}, x^*_{k,i}\rangle = \langle \bar{x}, \bar{x}^*_k\rangle$$

To end the proof, use the above equation and the second part of Lemma 4.
□

References

[1] Jim Douglas, Jr. and H. H. Rachford, Jr. On the numerical solution of heat conduction problems in two and three space variables. *Trans. Amer. Math. Soc.*, 82:421–439, 1956.

[2] Jonathan Eckstein and B. F. Svaiter. A family of projective splitting methods for the sum of two maximal monotone operators. *Math. Program.*, 111(1-2, Ser. B):173–199, 2008.

[3] Jonathan Eckstein and B. F. Svaiter. General projective splitting methods for sums of maximal monotone operators. *SIAM J. Control Optim.*, 48(2):787–811, 2009.

[4] Simon Fitzpatrick. Representing monotone operators by convex functions. In *Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988)*, volume 20 of *Proc. Centre Math. Anal. Austral. Nat. Univ.*, pages 59–65. Austral. Nat. Univ., Canberra, 1988.

[5] P.-L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.*, 16(6):964–979, 1979.

[6] George J. Minty. Monotone (nonlinear) operators in Hilbert space. *Duke Math. J.*, 29:341–346, 1962.

[7] Zdzisław Opial. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Amer. Math. Soc.*, 73:591–597, 1967.