Cyclic Douglas–Rachford Iterations

Matthew K. Tam
Joint work with Laureate Prof Jon Borwein

School of Mathematical and Physical Sciences
University of Newcastle, Australia

57th AustMS Meeting, 30th September to 3rd October 2013

With support from the AustMS Student Support Scheme

A Common Problem

The feasibility problem asks

$$\text{Find } x \in C = \bigcap_{i=1}^{N} C_i,$$

where C and the C_i's are subsets of a Hilbert space, $\mathcal{H}$. Examples are:

- Linear systems of equations; i.e. affine C_i's.
- Matrix completion problems;[1] e.g. PSD matrices, protein structure.
- 3-SAT, TetraVex, Sudoku, nonograms;[2] (NP-complete, combin.)
- Various inverse problems; e.g. phase retrieval.

Projection algorithms are frequently used to solve such problems. At each step, these methods utilise the nearest point projections onto the C_i's (rather than directly onto C).

[1] **Douglas–Rachford feasibility methods for matrix completion problems** with F.J. Aragón Artacho and J.M. Borwein. Submitted Aug. 2013. arXiv:1308.4243

[2] **Recent results on Douglas–Rachford methods for combinatorial optimization** with F.J. Aragón Artacho and J.M. Borwein. Submitted 2013 arXiv:1305.2657

A Variational Toolkit

Let $S \subseteq \mathcal{H}$. The (nearest point) projection onto S is the (set-valued) mapping,

$$P_S x := \operatorname*{argmin}_{s \in S} \|s - x\|.$$

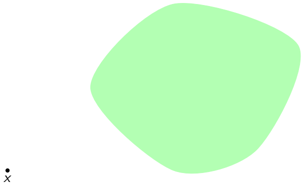
Variational characterisation of convex projections

Let $C \subseteq \mathcal{H}$ be closed and convex. Then $P_C x$ exist uniquely, $\forall x \in \mathcal{H}$, and

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

The reflection w.r.t. S is the (set-valued) mapping,

$$R_S := 2P_S - I.$$



A Variational Toolkit

Let $S \subseteq \mathcal{H}$. The (nearest point) projection onto S is the (set-valued) mapping,

$$P_S x := \operatorname*{argmin}_{s \in S} \|s - x\|.$$

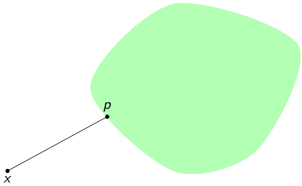
Variational characterisation of convex projections

Let $C \subseteq \mathcal{H}$ be closed and convex. Then $P_C x$ exist uniquely, $\forall x \in \mathcal{H}$, and

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

The reflection w.r.t. S is the (set-valued) mapping,

$$R_S := 2P_S - I.$$



A Variational Toolkit

Let $S \subseteq \mathcal{H}$. The (nearest point) projection onto S is the (set-valued) mapping,

$$P_S x := \operatorname*{argmin}_{s \in S} \|s - x\|.$$

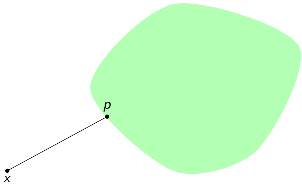
Variational characterisation of convex projections

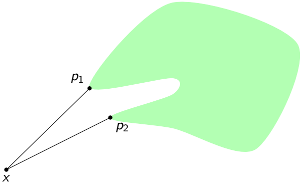
Let $C \subseteq \mathcal{H}$ be closed and convex. Then $P_C x$ exist uniquely, $\forall x \in \mathcal{H}$, and

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

The reflection w.r.t. S is the (set-valued) mapping,

$$R_S := 2P_S - I.$$





A Variational Toolkit

Let $S \subseteq \mathcal{H}$. The (nearest point) projection onto S is the (set-valued) mapping,

$$P_S x := \operatorname*{argmin}_{s \in S} \|s - x\|.$$

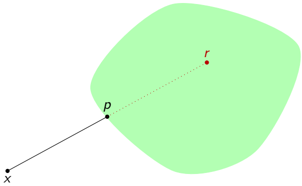
Variational characterisation of convex projections

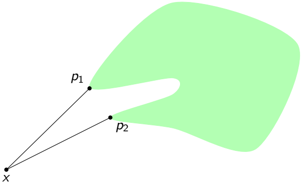
Let $C \subseteq \mathcal{H}$ be closed and convex. Then $P_C x$ exist uniquely, $\forall x \in \mathcal{H}$, and

$$p = P_C x \iff p \in C \text{ and } \langle x - p, c - p \rangle \le 0, \; \forall c \in C.$$

The reflection w.r.t. S is the (set-valued) mapping,

$$R_S := 2P_S - I.$$





A Variational Toolkit

Let $S \subseteq \mathcal{H}$. The (nearest point) projection onto S is the (set-valued) mapping,

$$P_S x := \operatorname*{argmin}_{s \in S} \|s - x\|.$$

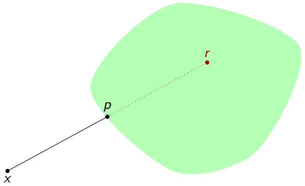
Variational characterisation of convex projections

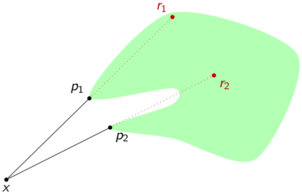
Let $C \subseteq \mathcal{H}$ be closed and convex. Then $P_C x$ exist uniquely, $\forall x \in \mathcal{H}$, and

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

The reflection w.r.t. S is the (set-valued) mapping,

$$R_S := 2P_S - I.$$





Some Common Projection Methods

A significant portion of the literature focuses on results like the following:

Theorems

Let $x_0 \in \mathcal{H}$. If $C_1, \ldots, C_N$ have certain properties then (x_n) converges in some sense to a point x having some properties.

Scheme	Iteration
Cyclic Projections	$x_{n+1} := \prod_{i=1}^{N} P_{C_i} x_n$
Averaged Projections	$x_{n+1} := \frac{1}{N} \sum_{i=1}^{N} P_{C_i} x_n$
Relaxed projections	$x_{n+1} := \prod_{i=1}^{N} (\lambda I + (1-\lambda) P_{C_i}) x_n$
Project-project-average	$x_{n+1} := \frac{1}{2}(I + \prod_{i=1}^{N} P_{C_i}) x_n$
Douglas–Rachford	$x_{n+1} := \frac{1}{2}(I + R_{C_2} R_{C_1}) x_n$
Dykstra's method	$x_n^i := P_{C_i}(x_n^{i-1} - I_{n-1}^i)$, $I_n^i := x_n^i - (x_n^{i-1} - I_{n-1}^i)$

- In the convex setting, projection methods are *fairly* well understood.
- In the non-convex setting, there are some useful beginnings.
- There also exists a large literature addressing convergence rates, etc.

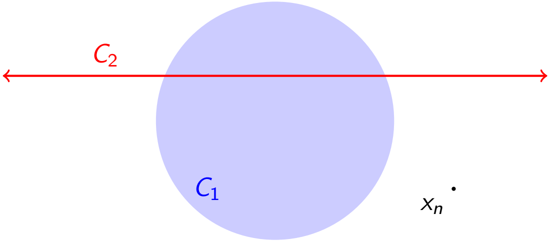
The Classical Douglas–Rachford Scheme

Theorem (Douglas–Rachford, Lions–Mercier)

Suppose $C_1, C_2 \subseteq \mathcal{H}$ are closed and convex with nonempty intersection. For any $x_0 \in \mathcal{H}$ define

$$x_{n+1} := T_{C_1,C_2} x_n \text{ where } T_{C_1,C_2} := \frac{I + R_{C_2} R_{C_1}}{2}.$$

Then $x_n \overset{w}{\rightharpoonup} x$ such that $P_{C_1} x \in C_1 \cap C_2$.



$$C_1 = \{x \in \mathcal{H} : \|x\| \leq 1\}, \quad C_2 = \{x \in \mathcal{H} : \langle a, x \rangle = b\}.$$

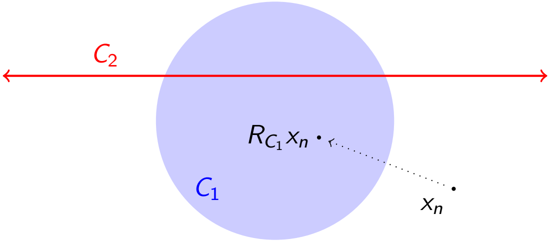
The Classical Douglas–Rachford Scheme

Theorem (Douglas–Rachford, Lions–Mercier)

Suppose $C_1, C_2 \subseteq \mathcal{H}$ are closed and convex with nonempty intersection. For any $x_0 \in \mathcal{H}$ define

$$x_{n+1} := T_{C_1,C_2} x_n \text{ where } T_{C_1,C_2} := \frac{I + R_{C_2} R_{C_1}}{2}.$$

Then $x_n \overset{w}{\rightharpoonup} x$ such that $P_{C_1} x \in C_1 \cap C_2$.



$$C_1 = \{x \in \mathcal{H} : \|x\| \leq 1\}, \quad C_2 = \{x \in \mathcal{H} : \langle a, x \rangle = b\}.$$

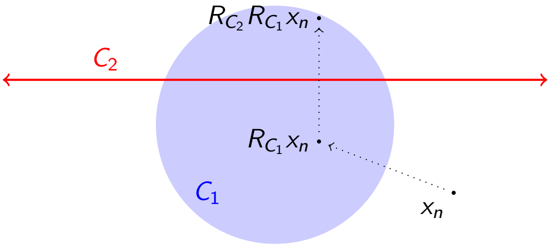
The Classical Douglas–Rachford Scheme

Theorem (Douglas–Rachford, Lions–Mercier)

Suppose $C_1, C_2 \subseteq \mathcal{H}$ are closed and convex with nonempty intersection. For any $x_0 \in \mathcal{H}$ define

$$x_{n+1} := T_{C_1,C_2} x_n \text{ where } T_{C_1,C_2} := \frac{I + R_{C_2} R_{C_1}}{2}.$$

Then $x_n \overset{w}{\rightharpoonup} x$ such that $P_{C_1} x \in C_1 \cap C_2$.



$$C_1 = \{x \in \mathcal{H} : \|x\| \leq 1\}, \quad C_2 = \{x \in \mathcal{H} : \langle a, x \rangle = b\}.$$

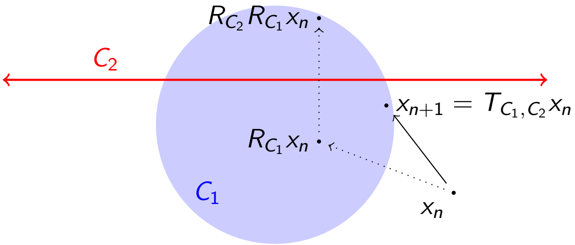
The Classical Douglas–Rachford Scheme

Theorem (Douglas–Rachford, Lions–Mercier)

Suppose $C_1, C_2 \subseteq \mathcal{H}$ are closed and convex with nonempty intersection. For any $x_0 \in \mathcal{H}$ define

$$x_{n+1} := T_{C_1,C_2} x_n \text{ where } T_{C_1,C_2} := \frac{I + R_{C_2} R_{C_1}}{2}.$$

Then $x_n \overset{w}{\rightharpoonup} x$ such that $P_{C_1} x \in C_1 \cap C_2$.



$$C_1 = \{x \in \mathcal{H} : \|x\| \leq 1\}, \quad C_2 = \{x \in \mathcal{H} : \langle a, x \rangle = b\}.$$

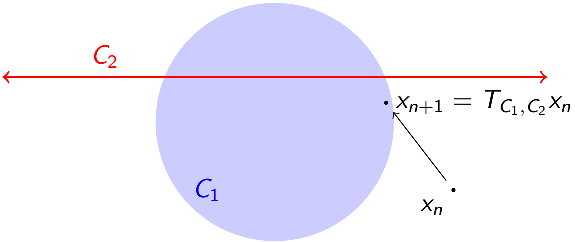
The Classical Douglas–Rachford Scheme

Theorem (Douglas–Rachford, Lions–Mercier)

Suppose $C_1, C_2 \subseteq \mathcal{H}$ are closed and convex with nonempty intersection. For any $x_0 \in \mathcal{H}$ define

$$x_{n+1} := T_{C_1,C_2} x_n \text{ where } T_{C_1,C_2} := \frac{I + R_{C_2} R_{C_1}}{2}.$$

Then $x_n \xrightarrow{w} x$ such that $P_{C_1} x \in C_1 \cap C_2$.



$$C_1 = \{x \in \mathcal{H} : \|x\| \leq 1\}, \quad C_2 = \{x \in \mathcal{H} : \langle a, x \rangle = b\}.$$

The Douglas–Rachford Scheme (cont.)

A significant portion of the literature focuses on results like the following:

Theorems

Let $x_0 \in \mathcal{H}$. If $C_1, \dots, C_N$ have certain properties then (x_n) converges in some sense to a point x having some properties.

Scheme	Iteration
Cyclic Projections	$x_{n+1} := \prod_{i=1}^{N} P_{C_i} x_n$
Averaged Projections	$x_{n+1} := \frac{1}{N} \sum_{i=1}^{N} P_{C_i} x_n$
Relaxed projections	$x_{n+1} := \prod_{i=1}^{N} (\lambda I + (1-\lambda) P_{C_i}) x_n$
Project-project-average	$x_{n+1} := \frac{1}{2}(I + \prod_{i=1}^{N} P_{C_i}) x_n$
Douglas–Rachford	$x_{n+1} := \frac{1}{2}(I + R_{C_2} R_{C_1}) x_n$
Dykstra's method	$x_n^i := P_{C_i}(x_n^{i-1} - I_{n-1}^i)$, $I_n^i := x_n^i - (x_n^{i-1} - I_{n-1}^i)$

- In the convex setting, projection methods are *fairly* well understood.
- In the non-convex setting, there are some useful beginnings.
- There also exists a large literature addressing convergence rates, etc.

The Douglas–Rachford Scheme (cont.)

1. Despite insufficient theoretical justification, the Douglas–Rachford scheme has been successfully applied to various hard non-convex feasibility problems.

The Douglas–Rachford Scheme (cont.)

1. Despite insufficient theoretical justification, the Douglas–Rachford scheme has been successfully applied to various hard non-convex feasibility problems. Other projection methods often get “stuck”.

The Douglas–Rachford Scheme (cont.)

1. Despite insufficient theoretical justification, the Douglas–Rachford scheme has been successfully applied to various hard non-convex feasibility problems. Other projection methods often get "stuck". An illustrative example: protein reconstruction (from Ref. 1).

The Douglas–Rachford Scheme (cont.)

1. Despite insufficient theoretical justification, the Douglas–Rachford scheme has been successfully applied to various hard non-convex feasibility problems. Other projection methods often get “stuck”. An illustrative example: protein reconstruction (from Ref. 1).

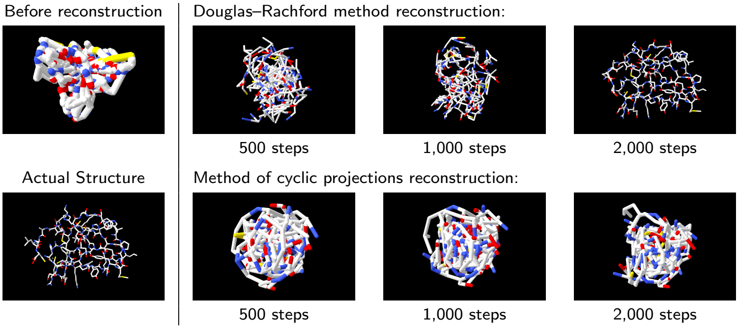


Table: Reconstructions of the protein 1PTQ (404 atoms) from “NMR” data.

- The method of cyclic projections works well in optical aberration correction (Hubble) (a non-convex feasibility problem) why not here?

The Douglas–Rachford Scheme (cont.)

2. Most projection algorithms can be naturally extended to handle feasibility problems involving a large (finite) number of sets.

The Douglas–Rachford Scheme (cont.)

2. Most projection algorithms can be naturally extended to handle feasibility problems involving a large (finite) number of sets. The Douglas–Rachford method can only directly handle two set problems. For more sets, one must use an equivalent problem posed in $\mathcal{H}^N$.

The Douglas–Rachford Scheme (cont.)

2. Most projection algorithms can be naturally extended to handle feasibility problems involving a large (finite) number of sets. The Douglas–Rachford method can only directly handle two set problems. For more sets, one must use an equivalent problem posed in $\mathcal{H}^N$. The "obvious" extensions fail.

Failure of three set Douglas–Rachford (Example due to Brailey Sims)

$$T_{A,B,C} := \frac{1}{2}(I + R_C R_B R_A).$$

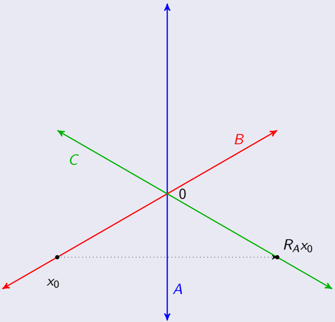
Let $x_0 = (-\sqrt{3}, -1), \alpha \in [2, \infty]$.

$$A := \{\lambda(0,1) : |\lambda| \leq \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \leq \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \leq \alpha\}.$$

Then $A \cap B \cap C = \{0\}$.

We have $x_0 \in \text{Fix}\, T_{A,B,C}$. However,

$$P_A x_0, P_B x_0, P_C x_0 \neq 0.$$



The Douglas–Rachford Scheme (cont.)

2. Most projection algorithms can be naturally extended to handle feasibility problems involving a large (finite) number of sets. The Douglas–Rachford method can only directly handle two set problems. For more sets, one must use an equivalent problem posed in $\mathcal{H}^N$. The "obvious" extensions fail.

Failure of three set Douglas–Rachford (Example due to Brailey Sims)

$$T_{A,B,C} := \frac{1}{2}(I + R_C R_B R_A).$$

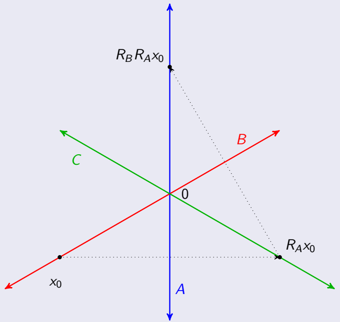
Let $x_0 = (-\sqrt{3}, -1), \alpha \in [2, \infty]$.

$$A := \{\lambda(0,1) : |\lambda| \leq \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \leq \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \leq \alpha\}.$$

Then $A \cap B \cap C = \{0\}$.

We have $x_0 \in \operatorname{Fix} T_{A,B,C}$. However,

$$P_A x_0, P_B x_0, P_C x_0 \neq 0.$$



2. Most projection algorithms can be naturally extended to handle feasibility problems involving a large (finite) number of sets. The Douglas–Rachford method can only directly handle two set problems. For more sets, one must use an equivalent problem posed in $\mathcal{H}^N$. The "obvious" extensions fail.

Failure of three set Douglas–Rachford (Example due to Brailey Sims)

$$T_{A,B,C} := \frac{1}{2}(I + R_C R_B R_A).$$

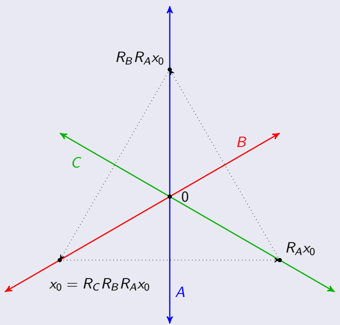
Let $x_0 = (-\sqrt{3}, -1), \alpha \in [2, \infty]$.

$$A := \{\lambda(0,1) : |\lambda| \leq \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \leq \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \leq \alpha\}.$$

Then $A \cap B \cap C = \{0\}$.

We have $x_0 \in \operatorname{Fix} T_{A,B,C}$. However,

$$P_A x_0, P_B x_0, P_C x_0 \neq 0.$$



The Douglas–Rachford Scheme (cont.)

2. Most projection algorithms can be naturally extended to handle feasibility problems involving a large (finite) number of sets. The Douglas–Rachford method can only directly handle two set problems. For more sets, one must use an equivalent problem posed in $\mathcal{H}^N$. The "obvious" extensions fail.

Failure of three set Douglas–Rachford (Example due to Brailey Sims)

$$\boxed{T_{A,B,C} := \tfrac{1}{2}(I + R_C R_B R_A).}$$

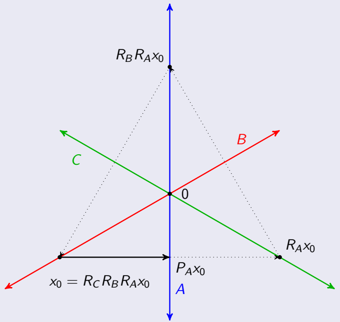
Let $x_0 = (-\sqrt{3}, -1), \alpha \in [2, \infty]$.

$$A := \{\lambda(0,1) : |\lambda| \le \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \le \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \le \alpha\}.$$

Then $A \cap B \cap C = \{0\}$.

We have $x_0 \in \operatorname{Fix} T_{A,B,C}$. However,

$$P_A x_0, P_B x_0, P_C x_0 \neq 0.$$



The Douglas–Rachford Scheme (cont.)

2. Most projection algorithms can be naturally extended to handle feasibility problems involving a large (finite) number of sets. The Douglas–Rachford method can only directly handle two set problems. For more sets, one must use an equivalent problem posed in $\mathcal{H}^N$. The "obvious" extensions fail.

Failure of three set Douglas–Rachford (Example due to Brailey Sims)

$$T_{A,B,C} := \frac{1}{2}(I + R_C R_B R_A).$$

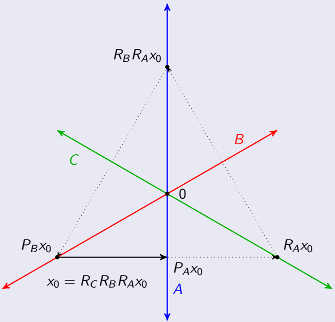
Let $x_0 = (-\sqrt{3}, -1), \alpha \in [2, \infty]$.

$$A := \{\lambda(0,1) : |\lambda| \leq \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \leq \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \leq \alpha\}.$$

Then $A \cap B \cap C = \{0\}$.

We have $x_0 \in \operatorname{Fix} T_{A,B,C}$. However,

$$P_A x_0, P_B x_0, P_C x_0 \neq 0.$$



The Douglas–Rachford Scheme (cont.)

2. Most projection algorithms can be naturally extended to handle feasibility problems involving a large (finite) number of sets. The Douglas–Rachford method can only directly handle two set problems. For more sets, one must use an equivalent problem posed in $\mathcal{H}^N$. The "obvious" extensions fail.

Failure of three set Douglas–Rachford (Example due to Brailey Sims)

$$T_{A,B,C} := \frac{1}{2}(I + R_C R_B R_A).$$

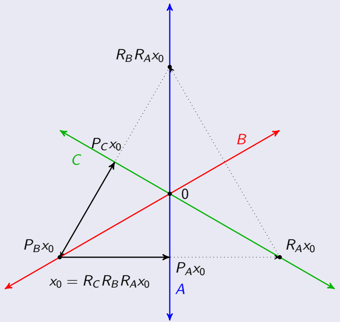
Let $x_0 = (-\sqrt{3}, -1), \alpha \in [2, \infty]$.

$$A := \{\lambda(0,1) : |\lambda| \leq \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \leq \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \leq \alpha\}.$$

Then $A \cap B \cap C = \{0\}$.

We have $x_0 \in \text{Fix}\, T_{A,B,C}$. However,

$$P_A x_0, P_B x_0, P_C x_0 \neq 0.$$



The Douglas–Rachford Scheme (cont.)

Our investigation was motivated by the classical Douglas–Rachford scheme's good behaviour on various non-convex problems, and the absence of an obvious extension to feasibility problems with more than two sets.

The Douglas–Rachford Scheme (cont.)

Our investigation was motivated by the classical Douglas–Rachford scheme's good behaviour on various non-convex problems, and the absence of an obvious extension to feasibility problems with more than two sets. In the remainder of this talk I will discuss our findings. In particular, I will discuss the content of our recent paper:

A Cyclic Douglas–Rachford Iteration Scheme with J.M. Borwein.
Published online in *J. Optim. Theory. Appl.*, August 2013.
DOI: 10.1007/s10957-013-0381-x

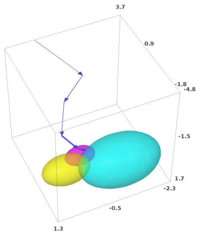


Fig. A cyclic Douglas–Rachford iteration for three balls constraints drawn in *Sage*.

Tools from Nonexpansive Mapping Theory

Let $T : \mathcal{H} \to \mathcal{H}$. Then T is:

- nonexpansive if

$$\|Tx - Ty\| \leq \|x - y\|, \quad \forall x, y \in \mathcal{H}.$$

- firmly nonexpansive if

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

Tools from Nonexpansive Mapping Theory

Let $T : \mathcal{H} \to \mathcal{H}$. Then T is:

- nonexpansive if

$$\|Tx - Ty\| \leq \|x - y\|, \quad \forall x, y \in \mathcal{H}.$$

- firmly nonexpansive if

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

Proposition (Nonexpansive properties)

The following are equivalent.

- T is firmly nonexpansive.
- $I - T$ is firmly nonexpansive.
- $2T - I$ is nonexpansive.
- $T = \alpha I + (1 - \alpha)R$, for $\alpha \in (0, 1/2]$ and some nonexpansive R.
- Many other characterisations.

Tools from Nonexpansive Mapping Theory

Let $T : \mathcal{H} \to \mathcal{H}$. Then T is:

- nonexpansive if

$$\|Tx - Ty\| \leq \|x - y\|, \quad \forall x, y \in \mathcal{H}.$$

- firmly nonexpansive if

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

Nonexpansive properties of projections

Let $A, B \subseteq \mathcal{H}$ be closed and convex. Then

- $P_A := \mathrm{argmin}_{s \in S} \|\cdot - s\|$ is firmly nonexpansive.
- $R_A := 2P_A - I$ is nonexpansive.
- $T_{A,B} := \frac{1}{2}(I + R_B R_A)$ is firmly nonexpansive.

Nonexpansive maps are closed under composition, convex combinations, etc. Firmly nonexpansive maps need not be. E.g., Composition of two projections onto subspace in $\mathbb{R}^2$ (Bauschke–Borwein–Lewis, 1997).

Tools from Nonexpansive Mapping Theory (cont.)

- asymptotically regular if, for all $x \in \mathcal{H}$,

$$\|T^{n+1}x - T^n x\| \to 0.$$

Any firmly nonexpansive mapping with at least one fixed point is asymptotically regular.

Tools from Nonexpansive Mapping Theory (cont.)

- asymptotically regular if, for all $x \in \mathcal{H}$,

$$\|T^{n+1}x - T^n x\| \to 0.$$

Any firmly nonexpansive mapping with at least one fixed point is asymptotically regular.

A useful Theorem for building iterative schemes:

Theorem (Opial, 1967)

Let $T : \mathcal{H} \to \mathcal{H}$ be nonexpansive and asymptotically regular. Set $x_{n+1} = Tx_n$. Then $x_n \overset{w}{\rightharpoonup} x$ such that $x \in \operatorname{Fix} T$.

Cyclic Douglas–Rachford Scheme

- In some sense, the classical Douglas–Rachford scheme is "unfair".
  - Reflection is always performed first with respect to the same set.
- A "fair" scheme might change the reflection order at each step.
- For two sets,

$$x_{n+1} := T_{C_2,C_1} T_{C_1,C_2} x_n = \left(\frac{I + R_{C_1} R_{C_2}}{2}\right) \left(\frac{I + R_{C_2} R_{C_1}}{2}\right) x_n.$$

Cyclic Douglas–Rachford Scheme

- In some sense, the classical Douglas–Rachford scheme is "unfair".
  - Reflection is always performed first with respect to the same set.
- A "fair" scheme might change the reflection order at each step.
- For two sets,

$$x_{n+1} := T_{C_2,C_1} T_{C_1,C_2} x_n = \left(\frac{I + R_{C_1} R_{C_2}}{2}\right)\left(\frac{I + R_{C_2} R_{C_1}}{2}\right) x_n.$$

- For three sets,

$$\begin{aligned} x_{n+1} &:= T_{C_3,C_1} T_{C_2,C_3} T_{C_1,C_2} x_n \\ &= \left(\frac{I + R_{C_1} R_{C_3}}{2}\right)\left(\frac{I + R_{C_3} R_{C_2}}{2}\right)\left(\frac{I + R_{C_2} R_{C_1}}{2}\right) x_n. \end{aligned}$$

- And so on . . .

Cyclic Douglas–Rachford Scheme (cont.)

Theorem (Borwein–T 2013)

Let $C_1, \dots, C_N \subseteq \mathcal{H}$ be closed and convex with nonempty intersection. For any $x_0 \in \mathcal{H}$, define[3]

$$x_{n+1} = T_{[C_1\ C_2\ \dots\ C_N]} x_n \text{ where } T_{[C_1\ C_2\ \dots\ C_N]} := \prod_{i=1}^{N} T_{C_i, C_{i+1}}.$$

Then $x_n \xrightarrow{w} x$ such that $P_{C_i} x = P_{C_j} x$, for all indices i, j. In particular,

$$P_{C_j} x \in \bigcap_{i=1}^{N} C_i, \text{ for each index } j.$$

[3] Here and elsewhere, indices are understood modulo N.

Cyclic Douglas–Rachford Scheme (cont.)

Theorem (Borwein–T 2013)

Let $C_1, \ldots, C_N \subseteq \mathcal{H}$ be closed and convex with nonempty intersection. For any $x_0 \in \mathcal{H}$, define[3]

$$x_{n+1} = T_{[C_1\ C_2\ \ldots\ C_N]} x_n \text{ where } T_{[C_1\ C_2\ \ldots\ C_N]} := \prod_{i=1}^{N} T_{C_i, C_{i+1}}.$$

Then $x_n \stackrel{w}{\rightharpoonup} x$ such that $P_{C_i} x = P_{C_j} x$, for all indices i, j. In particular,

$$P_{C_j} x \in \bigcap_{i=1}^{N} C_i, \text{ for each index } j.$$

Proof.

First show $\mathrm{Fix}\, T_{[C_1 \ldots, C_N]} = \cap_{i=1}^{N} \mathrm{Fix}\, T_{C_i, C_{i+1}} \neq \emptyset$. Establish weak convergence to a fixed point, and use the variational characterisation of convex projections. □

[3] Here and elsewhere, indices are understood modulo N.

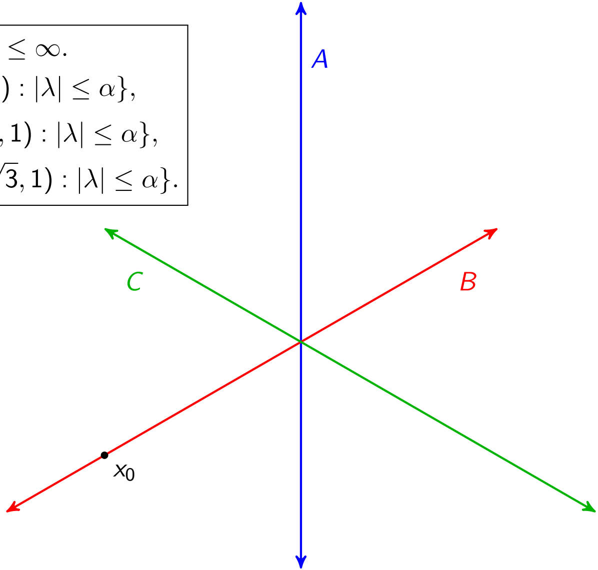
Example Revisited

$$2 \le \alpha \le \infty.$$

$$A := \{\lambda(0,1) : |\lambda| \le \alpha\},$$

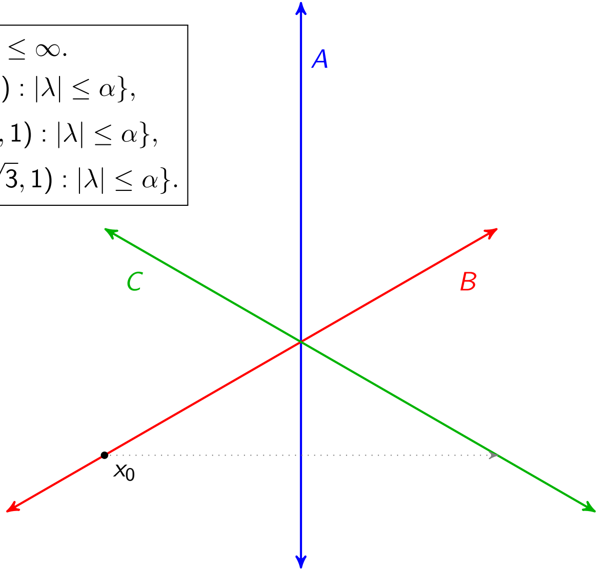
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \le \alpha\},$$

$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \le \alpha\}.$$



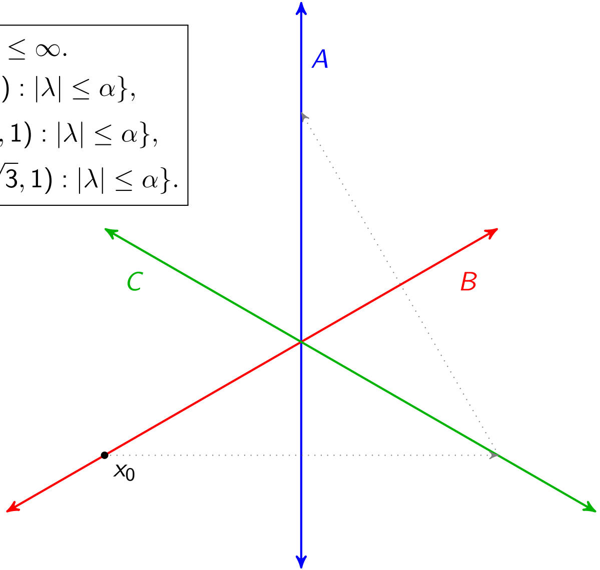
Example Revisited

$$2 \le \alpha \le \infty.$$
$$A := \{\lambda(0,1) : |\lambda| \le \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \le \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \le \alpha\}.$$



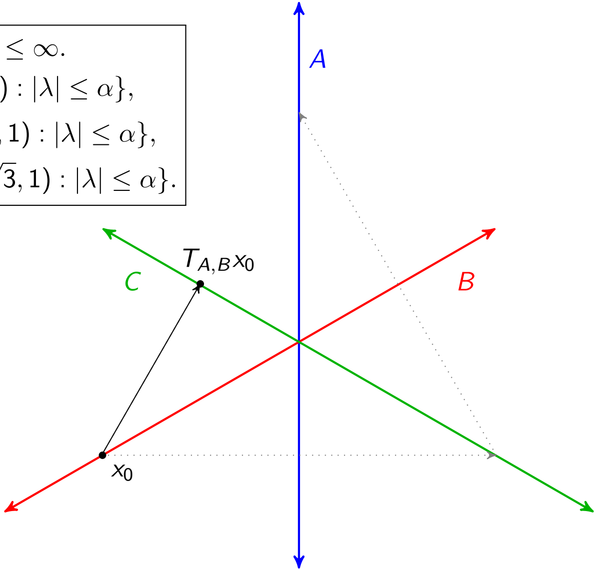
Example Revisited

$$2 \leq \alpha \leq \infty.$$
$$A := \{\lambda(0,1) : |\lambda| \leq \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \leq \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \leq \alpha\}.$$



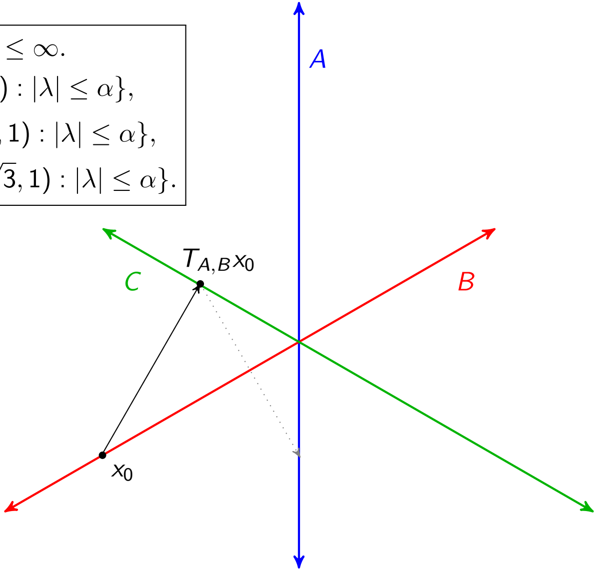
Example Revisited

$$2 \leq \alpha \leq \infty.$$
$$A := \{\lambda(0,1) : |\lambda| \leq \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \leq \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \leq \alpha\}.$$



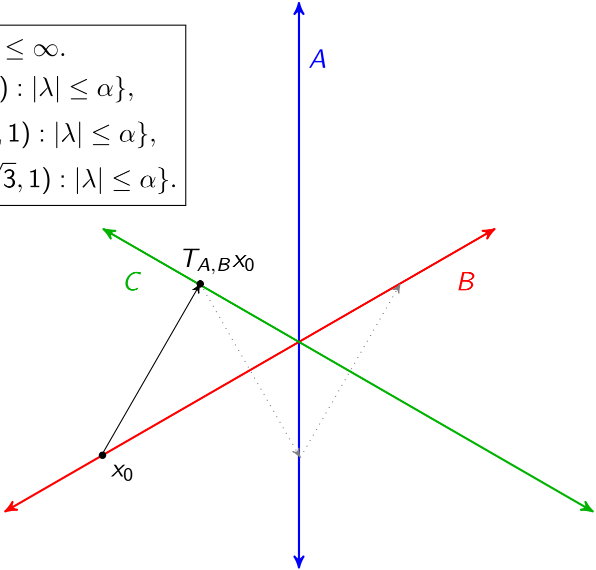
Example Revisited

$$2 \le \alpha \le \infty.$$
$$A := \{\lambda(0,1) : |\lambda| \le \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \le \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \le \alpha\}.$$



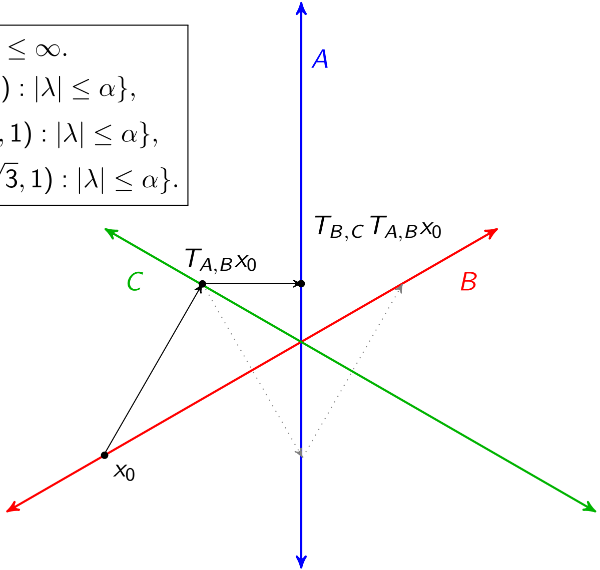
Example Revisited

$$2 \leq \alpha \leq \infty.$$
$$A := \{\lambda(0,1) : |\lambda| \leq \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \leq \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \leq \alpha\}.$$



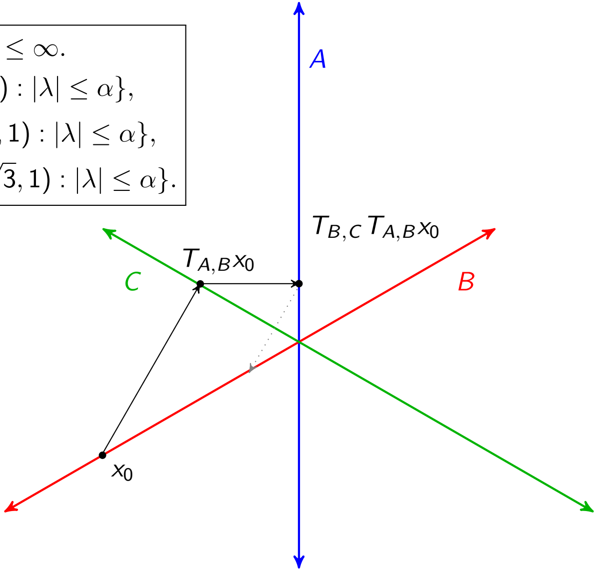
Example Revisited

$$2 \le \alpha \le \infty.$$
$$A := \{\lambda(0,1) : |\lambda| \le \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \le \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \le \alpha\}.$$



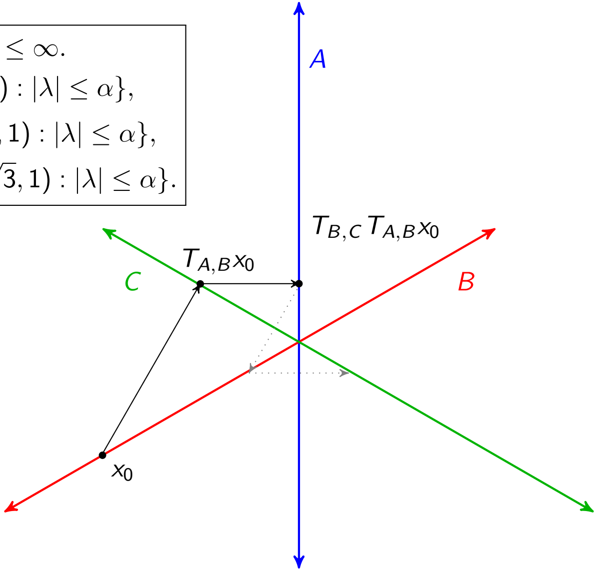
Example Revisited

$$2 \le \alpha \le \infty.$$
$$A := \{\lambda(0,1) : |\lambda| \le \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \le \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \le \alpha\}.$$



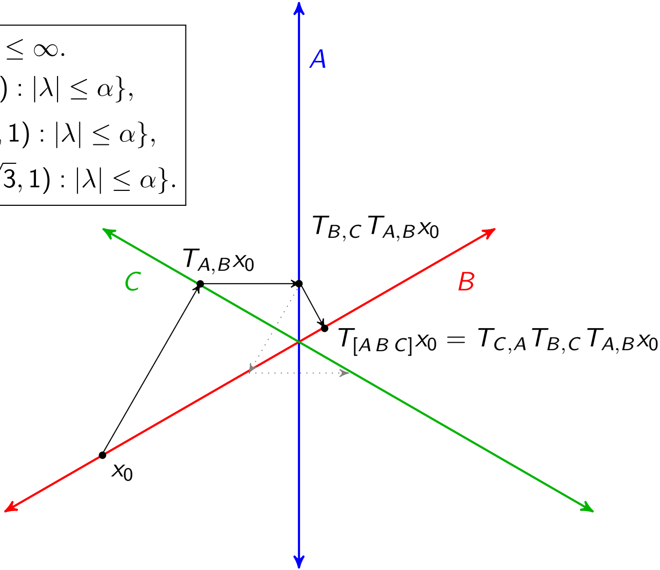
Example Revisited

$$2 \le \alpha \le \infty.$$
$$A := \{\lambda(0,1) : |\lambda| \le \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \le \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \le \alpha\}.$$



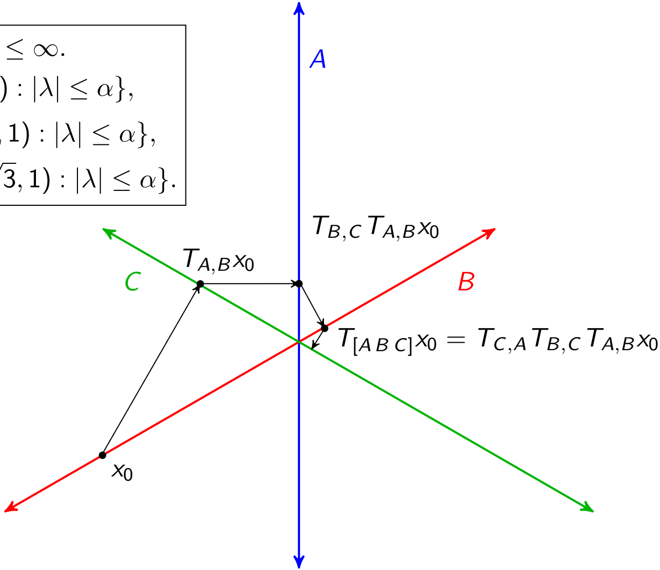
Example Revisited

$$2 \le \alpha \le \infty.$$
$$A := \{\lambda(0,1) : |\lambda| \le \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \le \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \le \alpha\}.$$



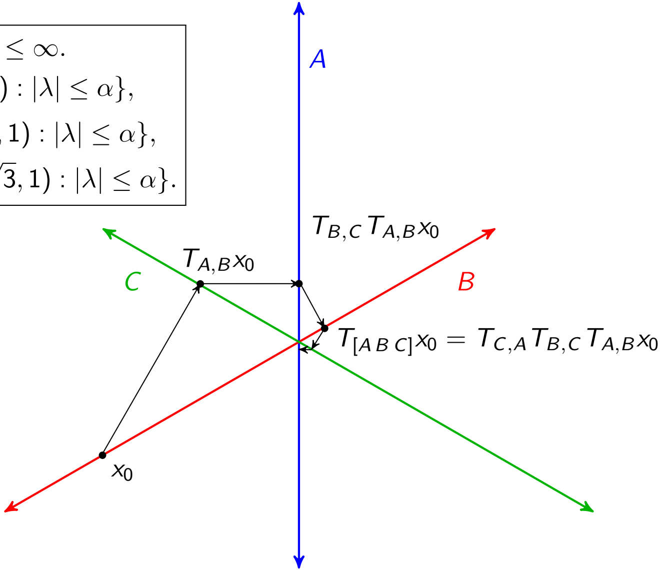
Example Revisited

$$2 \leq \alpha \leq \infty.$$
$$A := \{\lambda(0,1) : |\lambda| \leq \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \leq \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \leq \alpha\}.$$



Example Revisited

$$2 \le \alpha \le \infty.$$
$$A := \{\lambda(0,1) : |\lambda| \le \alpha\},$$
$$B := \{\lambda(\sqrt{3},1) : |\lambda| \le \alpha\},$$
$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \le \alpha\}.$$



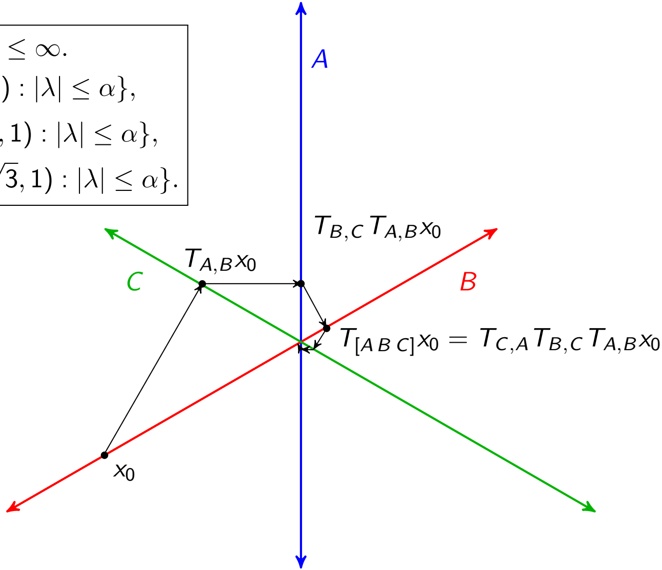
Example Revisited

$$2 \leq \alpha \leq \infty.$$

$$A := \{\lambda(0,1) : |\lambda| \leq \alpha\},$$

$$B := \{\lambda(\sqrt{3},1) : |\lambda| \leq \alpha\},$$

$$C := \{\lambda(-\sqrt{3},1) : |\lambda| \leq \alpha\}.$$



Cyclic Douglas–Rachford (cont.)

Can the iteration fail to converge in norm?

Cyclic Douglas–Rachford (cont.)

Can the iteration fail to converge in norm?

- Yes – we modify an example originally due to Hundal (2004).

Cyclic Douglas–Rachford (cont.)

Can the iteration fail to converge in norm?

- Yes – we modify an example originally due to Hundal (2004).

Failure of Norm Convergence (Hundal, Matoušková–Reich, Kopecká)

Let $\mathcal{H} = \ell_2$ and $\{e_i\}$ denote the standard basis. Define

$$C_1 = \{x \in \mathcal{H} : \langle e_1, x \rangle = x_1 = 0\}, \quad C_2 = \text{an "unnatural" cone}.$$

Then $C_1 \cap C_2 = \{0\}$. There exists $x_0 \in \mathcal{H}$ such that $T^n_{[C_1\ C_2]} x_0$ does not converge in norm

- C_1 is a closed subspace, C_2 a closed convex cone.
- For appropriate initial points, the cyclic Douglas–Rachford iterations and the alternating projections iterations coincide.
- Both converge weakly to 0, the unique point in the intersection.
- (Bauschke–Borwein 1993) Conjecture norm convergence if C_1 is affine, finite codimension, and $C_2 = L_2^+(\Omega, \mu)$. True for codim. 1.

General Framework

The cyclic Douglas–Rachford method framework applies more generally.

Theorem (Borwein–T 2013)

Let $C_1, \dots, C_N \subseteq \mathcal{H}$ be closed and convex with nonempty intersection. For any $x_0 \in \mathcal{H}$, define

$$x_{n+1} := Tx_n \text{ where } T := \prod_{i=1}^{M} T_i.$$

Further, suppose

1. T is nonexpansive and asymptotically regular,
2. $\operatorname{Fix} T = \bigcap_{i=1}^{M} \operatorname{Fix} T_i \neq \emptyset$,
3. $P_{C_i} \operatorname{Fix} T_i \subseteq C_{i+1}$, for each index i.

Then $x_n \overset{w}{\rightharpoonup} x$ such that $P_{C_i}x = P_{C_j}x$, for all indices i, j. In particular,

$$P_{C_j}x \in \bigcap_{i=1}^{N} C_i, \text{ for each index } j.$$

Averaged Douglas–Rachford Scheme

Theorem (Borwein–T 2013)

Let $C_1, C_2, \ldots, C_N \subseteq \mathcal{H}$ be closed and convex with nonempty intersection. For any $x_0 \in \mathcal{H}$, define

$$x_{n+1} := \frac{1}{N}\left(\sum_{i=1}^{N} T_{C_i,C_{i+1}}\right) x_n.$$

Then $x_n \overset{w}{\rightharpoonup} x$ such that $P_{C_i}x = P_{C_j}x$, for all indexes i,j. In particular,

$$P_{C_j}x \in \bigcap_{i=1}^{N} C_i, \text{ for each index } j.$$

Averaged Douglas–Rachford Scheme

Theorem (Borwein–T 2013)

Let $C_1, C_2, \ldots, C_N \subseteq \mathcal{H}$ be closed and convex with nonempty intersection. For any $x_0 \in \mathcal{H}$, define

$$x_{n+1} := \frac{1}{N}\left(\sum_{i=1}^{N} T_{C_i,C_{i+1}}\right) x_n.$$

Then $x_n \xrightarrow{w} x$ such that $P_{C_i}x = P_{C_j}x$, for all indexes i, j. In particular,

$$P_{C_j}x \in \bigcap_{i=1}^{N} C_i, \text{ for each index } j.$$

Proof. (Performed in $\mathcal{H}^N$).

Apply the previous Theorem to the sequence defined by

$$\mathbf{x}_{n+1} := P_D(T_1, T_2, \ldots, T_N)\mathbf{x}_n,$$

where $D = \{(x, x, \ldots, x) \in \mathcal{H}^N : x \in \mathcal{H}\}$. □

- Other applicable variants! e.g. cyclic project-project-average.

Infeasible Iterations (Alternating Projections)

Consider $A, B \subseteq \mathcal{H}$ with possibly empty intersection. For convenience, we introduce the sequences (a_n) and (b_n) where

$$x_0 \overset{P_A}{\mapsto} a_1 \overset{P_B}{\mapsto} b_1 \overset{P_A}{\mapsto} a_2 \overset{P_B}{\mapsto} b_2 \overset{P_A}{\mapsto} a_3 \overset{P_B}{\mapsto} \dots$$

Further define

$$E := \{x \in A : d(x, B) = d(A, B)\}, \; F := \{x \in B : d(x, A) = d(A, B)\}.$$

Infeasible Iterations (Alternating Projections)

Consider $A, B \subseteq \mathcal{H}$ with possibly empty intersection. For convenience, we introduce the sequences (a_n) and (b_n) where

$$x_0 \overset{P_A}{\mapsto} a_1 \overset{P_B}{\mapsto} b_1 \overset{P_A}{\mapsto} a_2 \overset{P_B}{\mapsto} b_2 \overset{P_A}{\mapsto} a_3 \overset{P_B}{\mapsto} \dots$$

Further define

$$E := \{x \in A : d(x, B) = d(A, B)\}, \ F := \{x \in B : d(x, A) = d(A, B)\}.$$

Theorem (Bauschke–Borwein 1994)

Let $A, B \subseteq \mathcal{H}$ be closed and convex. Exactly one of the following alternatives hold.

(a) $E, F = \emptyset$, $\|a_n\|, \|b_n\| \to \infty$.

(b) $E, F \neq \emptyset$, $a_n \overset{w}{\rightharpoonup} a \in E$, $b_n \overset{w}{\rightharpoonup} b \in F$ where $b = P_B a$ and $a = P_A b$. Furthermore, $\|a - b\| = d(A, B)$ and $b_n - a_n, \ b_n - a_{n+1} \to b - a$.

Infeasible Iterations (Alternating Projections)

Consider $A, B \subseteq \mathcal{H}$ with possibly empty intersection. For convenience, we introduce the sequences (a_n) and (b_n) where

$$x_0 \overset{P_A}{\mapsto} a_1 \overset{P_B}{\mapsto} b_1 \overset{P_A}{\mapsto} a_2 \overset{P_B}{\mapsto} b_2 \overset{P_A}{\mapsto} a_3 \overset{P_B}{\mapsto} \dots$$

Further define

$$E := \{x \in A : d(x, B) = d(A, B)\}, \; F := \{x \in B : d(x, A) = d(A, B)\}.$$

Theorem (Bauschke–Borwein 1994)

Let $A, B \subseteq \mathcal{H}$ be closed and convex. Exactly one of the following alternatives hold.

(a) $E, F = \emptyset$, $\|a_n\|, \|b_n\| \to \infty$.

(b) $E, F \neq \emptyset$, $a_n \overset{w}{\rightharpoonup} a \in E$, $b_n \overset{w}{\rightharpoonup} b \in F$ where $b = P_B a$ and $a = P_A b$. Furthermore, $\|a - b\| = d(A, B)$ and $b_n - a_n, \; b_n - a_{n+1} \to b - a$.

- Does not generalise to more than two sets: "*There is no variational characterization of the cycles in the method of periodic projections*", Baillion–Combettes–Cominetti (2012).

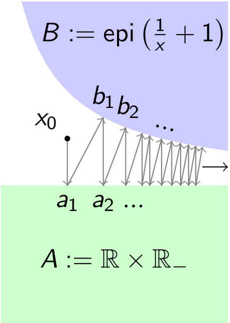
Infeasible Iterations (cont.)

Theorem (Bauschke–Borwein 1994)

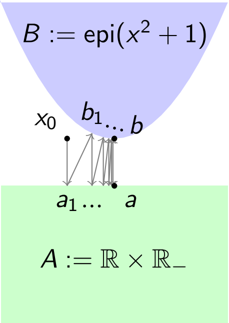
Let $A, B \subseteq \mathcal{H}$ be closed and convex. Exactly one of the following alternatives hold.

(a) $E, F = \emptyset$, $\|a_n\|, \|b_n\| \to \infty$.

(b) $E, F \neq \emptyset$, $a_n \overset{w}{\rightharpoonup} a \in E$, $b_n \overset{w}{\rightharpoonup} b \in F$ where $b = P_B a$ and $a = P_A b$. Furthermore, $\|a - b\| = d(A, B)$ and $b_n - a_n$, $b_n - a_{n+1} \to b - a$.



Example of (a).



Example of (b).

Infeasible Iterations (Alternating Douglas-Rachford)

Similarly we introduce the sequences (α_n) and (β_n) where

$$x_0 \overset{T_{A,B}}{\mapsto} \beta_1 \overset{T_{B,A}}{\mapsto} \alpha_1 \overset{T_{A,B}}{\mapsto} \beta_2 \overset{T_{B,A}}{\mapsto} \alpha_2 \overset{T_{A,B}}{\mapsto} \beta_3 \overset{T_{B,A}}{\mapsto} \dots$$

The difficulty is $\text{Fix } T_{A,B} \neq \emptyset \iff A \cap B \neq \emptyset$. In the empty case,

$$\text{Fix } T_{[C_1 \dots, C_N]} \supseteq \bigcap_{i=1}^{N} \text{Fix } T_{C_i, C_{i+1}} = \emptyset.$$

Theorem (Borwein–T 201?)

Let $A, B \subseteq \mathcal{H}$ be closed and convex. Exactly one of the following alternatives hold.

(a) $E, F, \text{Fix } T_{[A\,B]}, \text{Fix } T_{[B\,A]} = \emptyset$, and $\|\alpha_n\|, \|\beta_n\| \to \infty$.

(b) $E, F, \text{Fix } T_{[A\,B]}, \text{Fix } T_{[B\,A]} \neq \emptyset$, and

$$\alpha_n \overset{w.}{\rightharpoonup} \alpha \in \text{Fix } T_{[A\,B]}, \quad \beta_n \overset{w.}{\rightharpoonup} \beta \in \text{Fix } T_{[B\,A]},$$

where $\beta = T_{A,B}\alpha$ and $\alpha = T_{B,A}\beta$. Furthermore,

$$\beta - \alpha = P_B\beta - P_A\alpha, \quad \|P_B\beta - P_A\alpha\| = d(A, B),$$

and $\beta_n - \alpha_n, \beta_{n+1} - \alpha \to \beta - \alpha$.

Infeasible Iterations (Alternating Douglas-Rachford)

Similarly we introduce the sequences (α_n) and (β_n) where

$$x_0 \overset{T_{A,B}}{\mapsto} \beta_1 \overset{T_{B,A}}{\mapsto} \alpha_1 \overset{T_{A,B}}{\mapsto} \beta_2 \overset{T_{B,A}}{\mapsto} \alpha_2 \overset{T_{A,B}}{\mapsto} \beta_3 \overset{T_{B,A}}{\mapsto} \dots$$

The difficulty is $\text{Fix } T_{A,B} \neq \emptyset \iff A \cap B \neq \emptyset$. In the empty case,

$$\text{Fix } T_{[C_1 \dots, C_N]} \supseteq \bigcap_{i=1}^{N} \text{Fix } T_{C_i, C_{i+1}} = \emptyset.$$

Theorem (Borwein–T 201?)

Let $A, B \subseteq \mathcal{H}$ be closed and convex. Exactly one of the following alternatives hold.

(a) $E, F, \text{Fix } T_{[A\,B]}, \text{Fix } T_{[B\,A]} = \emptyset$, and $\|\alpha_n\|, \|\beta_n\| \to \infty$.

(b) $E, F, \text{Fix } T_{[A\,B]}, \text{Fix } T_{[B\,A]} \neq \emptyset$, and

$$\alpha_n \overset{w.}{\rightharpoonup} \alpha \in \text{Fix } T_{[A\,B]}, \quad \beta_n \overset{w.}{\rightharpoonup} \beta \in \text{Fix } T_{[B\,A]},$$

where $\beta = T_{A,B}\alpha$ and $\alpha = T_{B,A}\beta$. Furthermore,

$$\beta - \alpha = P_B\beta - P_A\alpha, \quad \|P_B\beta - P_A\alpha\| = d(A, B),$$

and $\beta_n - \alpha_n, \beta_{n+1} - \alpha \to \beta - \alpha$.

- cf. Classical Douglas–Rachford: If $A \cap B = \emptyset$ then $\|x_n\| \to \infty$.

Closing Remarks and Future Work

Some avenues for future investigation include:

1. Better understand asymptotics in the (two set) infeasible case.
   - Is there a variational characterisation for more than two sets?
2. Norm convergence assuming regularity *a lá* Bauschke–Borwein.
3. Non-convex settings:
   - Euclidean sphere and affine subspace: Aragón–Borwein–Sims.
   - Local relaxations of firm nonexpansivity: Hesse–Luke.
4. Applications & computational studies: Initial results are promising!
   - 200 ball constraints in $\mathbb{R}^{2000}$, implemented in *Python*:
     - Classical Douglas–Rachford: $\sim$ 30s for a solution with error $\sim 10^{-4}$.
     - Cyclic Douglas–Rachford: $\sim$ 0.5s for a solution with error $\sim 10^{-25}$.

A Cyclic Douglas–Rachford Iteration Scheme with J.M. Borwein. Published online in *J. Optim. Theory. Appl.*, August 2013. DOI: 10.1007/s10957-013-0381-x

Many resources can be found at:

http://carma.newcastle.edu.au/DRmethods