

# Monotone Operator Splitting Methods in Signal and Image Recovery

**P.L. Combettes<sup>1</sup>, J.-C. Pesquet<sup>2</sup>, and N. Pustelnik<sup>3</sup>**

<sup>2</sup> Univ. Pierre et Marie Curie, Paris 6 – LJLL– CNRS UMR 7598

<sup>2</sup> Univ. Paris-Est – LIGM – CNRS UMR 8049

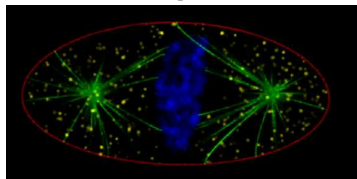
<sup>3</sup> ENS Lyon – Laboratoire de Physique – CNRS UMR 5672

**IEEE ICASSP Tutorial**

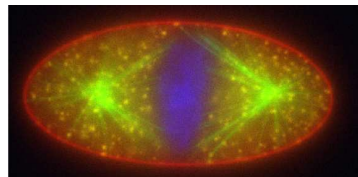
May 2014

# Monotone operators and inverse problems

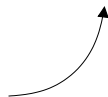
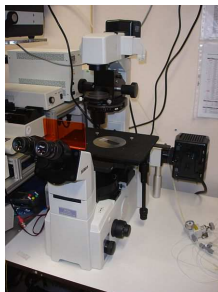
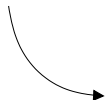
[Microscopy, ISBI Challenge 2013, F. Soulez]



Original image

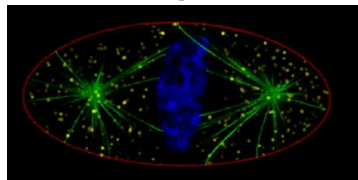


Degraded image



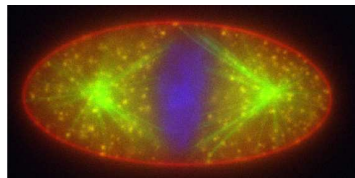
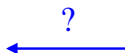
# Monotone operators and inverse problems

[Microscopy, ISBI Challenge 2013, F. Soulez]



Original image

$$\bar{x} \in \mathbb{R}^N$$



Degraded image

$$z = \mathcal{P}_\alpha(H\bar{x}) \in \mathbb{R}^M$$

- ▶  $H \in \mathbb{R}^{M \times N}$ : matrix associated with the degradation operator.
- ▶  $\mathcal{P}_\alpha: \mathbb{R}^M \rightarrow \mathbb{R}^M$ : noise degradation with parameter  $\alpha$  (e.g. Poisson noise).

→ Find a good estimate of  $\bar{x}$  from the observations  $z$ , using some a priori knowledge on  $H$  and on the noise statistics.

## Monotone operators and inverse problems

### Inverse problem:

Find an estimate  $\hat{x}$  close to  $\bar{x}$  from the observations  $z = \mathcal{P}_\alpha(H\bar{x})$ .

- ▶ Inverse filtering (if  $M = N$  and  $H$  is invertible)

$$\begin{aligned}\hat{x} &= H^{-1}z \\ &= H^{-1}(H\bar{x} + b) \quad \leftarrow \text{if } b \in \mathbb{R}^M \text{ is an additive noise} \\ &= \bar{x} + H^{-1}b\end{aligned}$$

→ Closed form expression, but amplification of the noise if  $H$  is ill-conditioned (*ill-posed problem*).

## Monotone operators and inverse problems

### Inverse problem:

Find an estimate  $\hat{x}$  close to  $\bar{x}$  from the observations  $z = \mathcal{P}_\alpha(H\bar{x})$ .

- ▶ Inverse filtering (if  $M \geq N$  and the rank of  $H$  is  $N$ )

$$\begin{aligned}\hat{x} &= (H^\top H)^{-1} H^\top z \\ &= (H^\top H)^{-1} H^\top (H\bar{x} + b) \quad \leftarrow \text{if } b \in \mathbb{R}^M \text{ is an additive noise} \\ &= \bar{x} + (H^\top H)^{-1} H^\top b\end{aligned}$$

→ Closed form expression, but amplification of the noise if  $H$  is ill-conditioned (*ill-posed problem*).

## Monotone operators and inverse problems

### Inverse problem:

Find an estimate  $\hat{x}$  close to  $\bar{x}$  from the observations  $z = \mathcal{P}_\alpha(H\bar{x})$ .

#### ► Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \quad \underbrace{f_1(x)}_{\substack{\text{Data fidelity term} \\ \text{e.g. } \|Hx - z\|_2^2}} + \underbrace{f_2(x)}_{\substack{\text{Regularization term} \\ \text{e.g. } \lambda \|x\|_p^p \text{ with } \begin{cases} p \geq 1 \\ \lambda \in ]0, +\infty[ \end{cases}}}$$

→ Often no closed form expression (e.g. if  $p \neq 2$  and  $H \neq \text{Id}$ )  
 or solution expensive to compute (e.g. if  $p = 2$ ,  $H \neq \text{Id}$  and  $N \gg 1$ )

→ Iterative strategy.

## Monotone operators and inverse problems

### Inverse problem:

Find an estimate  $\hat{x}$  close to  $\bar{x}$  from the observations  $z = \mathcal{P}_\alpha(H\bar{x})$ .

- ▶ Variational approach (more general context)

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \sum_{i=1}^m f_i(x)$$

where  $f_i$  may denote a data fidelity term / a (hybrid) regularization term / constraint.

→ Iterative strategy.

## Monotone operators and inverse problems

Iterative strategy = Optimization algorithm:

Construct a sequence  $(x_n)_{n \in \mathbb{N}}$  that converges to  $\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \sum_{i=1}^m f_i(x)$ .

- ▶ Sequence such that  $(\forall n \in \mathbb{N}) x_{n+1} = Tx_n$  where  $T$  denotes an operator from  $\mathbb{R}^N$  to  $\mathbb{R}^N$ .

→ How can we build  $T$  from the functionals  $(f_i)_{1 \leq i \leq m}$  involved in the minimization problem ?

→ Which properties are required by  $T$  in order to ensure the convergence of  $(x_n)_{n \in \mathbb{N}}$  to  $\hat{x}$  ?



## Naive answer

### Fixed point theorem (E. Picard, 1856-1941)

If

- ▶  $\hat{x}$  is a fixed point of  $T$ , i.e.  $\hat{x} = T\hat{x}$
- ▶  $T$  is a strict contraction, i.e. there exists  $\rho \in [0, 1[$  such that

$$(\forall (x, x') \in \mathbb{R}^N \times \mathbb{R}^N) \quad \|Tx - Tx'\| \leq \rho \|x - x'\|$$

then  $(x_n)_{n \in \mathbb{N}}$  converges to  $\hat{x}$ .



Proof: For all  $n \in \mathbb{N}$ ,

$$\begin{aligned} \|x_{n+1} - \hat{x}\| &= \|Tx_n - T\hat{x}\| \\ &\leq \rho \|x_n - \hat{x}\|. \end{aligned}$$

Consequently,  $\|x_n - \hat{x}\| \leq \rho^n \|x_0 - \hat{x}\|$ . Hence, we have proved that  $(x_n)_{n \in \mathbb{N}}$  converges linearly to  $\hat{x}$ .

## Why do we need to go further ?

### Limitations:

- ▶ It is difficult (even sometimes impossible) to build a *strictly* contractive operator  $T$ .
- ▶ One may prefer **iterations** built as  $(\forall n \in \mathbb{N}) x_{n+1} = T_n x_n$  where  $T_n$  denotes an operator from  $\mathbb{R}^N$  to  $\mathbb{R}^N$ .
- ▶ It is often intricate to build  $T_n$ , while it may be easier to write  $T_n$  as a **composition of simpler operators** (**splitting techniques**).
- ▶  $T_n$  can be multivalued, i.e.  $(\forall n \in \mathbb{N}) x_{n+1} \in T_n x_n$ .



## Tutorial philosophy

- ▶ Illustrate the performance of the algorithms on **inverse problems examples** (without giving too many details).
- ▶ Do not explore **all** the applications of monotone operators.

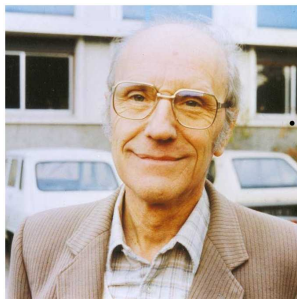
## Tutorial philosophy

- ▶ Illustrate the performance of the algorithms on **inverse problems examples** (without giving too many details).
- ▶ Do not explore **all** the applications of monotone operators. (Denoising, restoration, reconstruction, machine learning, ressource allocation, networking, communications,...)

## Tutorial philosophy

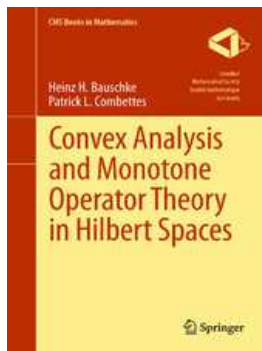
- ▶ Illustrate the performance of the algorithms on **inverse problems examples** (without giving too many details).
- ▶ Do not explore **all** the applications of monotone operators. (Denoising, restoration, reconstruction, machine learning, resource allocation, networking, communications,...)
- ▶ Focus on the convergence of **the iterates**  $(x_n)_{n \in \mathbb{N}}$  rather than on the convergence of criterion  $(\sum_{i=1}^m f_i(x_n))_{n \in \mathbb{N}}$ .

## A pioneer



Jean-Jacques Moreau  
(1923–2014)

# Reference book



– H.H. Bauschke and P.L. Combettes –



# Outline

1. Background on **monotone** and **maximally monotone** operators  
→ *Inversion, subdifferential, conjugate of a convex function*
2. **Nonexpansive** operators  
→ *Taxinomy, resolvent, and proximity operator*
3. Search for a **zero of a maximally monotone** operator  
→ *Fixed points, Fejér monotonicity, Douglas-Rachford, Forward-Backward*
4. **Duality**  
→ *Main theorems, ADMM, primal-dual methods*

# Part 1: Background

## 1. Monotone operators

- ▶ Definition
- ▶ Properties
- ▶ Basic operations
- ▶ Inversion
- ▶ Maximality
- ▶ Usefulness for convex optimization (subdifferential)

## 2. Maximally monotone operators

- ▶ Properties
- ▶ Basic operations
- ▶ Inversion
- ▶ Usefulness of inversion for convex optimization

## Hilbert spaces

A (real) **Hilbert space**  $\mathcal{H}$  is a complete real vector space endowed with an inner product  $\langle \cdot | \cdot \rangle$  whose associated norm is

$$(\forall x \in \mathcal{H}) \quad \|x\| = \sqrt{\langle x | x \rangle}.$$

- ▶ Particular case:  $\mathcal{H} = \mathbb{R}^N$  (Euclidean space with dimension  $N$ ).

## Hilbert spaces

A (real) **Hilbert space**  $\mathcal{H}$  is a complete real vector space endowed with an inner product  $\langle \cdot | \cdot \rangle$  whose associated norm is

$$(\forall x \in \mathcal{H}) \quad \|x\| = \sqrt{\langle x | x \rangle}.$$

- ▶ Particular case:  $\mathcal{H} = \mathbb{R}^N$  (Euclidean space with dimension  $N$ ).

$2^{\mathcal{H}}$  is the power set of  $\mathcal{H}$ , i.e. the family of all subsets of  $\mathcal{H}$ .

## Hilbert spaces

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

A linear operator  $L: \mathcal{H} \rightarrow \mathcal{G}$  is **bounded** (or continuous) if

$$\|L\| = \sup_{\|x\|_{\mathcal{H}} \leq 1} \|Lx\|_{\mathcal{G}} < +\infty$$

## Hilbert spaces

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

A linear operator  $L: \mathcal{H} \rightarrow \mathcal{G}$  is **bounded** (or continuous) if

$$\|L\| = \sup_{\|x\| \leq 1} \|Lx\| < +\infty$$

## Hilbert spaces

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

A linear operator  $L: \mathcal{H} \rightarrow \mathcal{G}$  is **bounded** (or continuous) if

$$\|L\| = \sup_{\|x\| \leq 1} \|Lx\| < +\infty$$

- ▶ In finite dimension, every linear operator is bounded.

## Hilbert spaces

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

A linear operator  $L: \mathcal{H} \rightarrow \mathcal{G}$  is **bounded** (or continuous) if

$$\|L\| = \sup_{\|x\| \leq 1} \|Lx\| < +\infty$$

- ▶ In finite dimension, every linear operator is bounded.

**$\mathcal{B}(\mathcal{H}, \mathcal{G})$** : Banach space of bounded linear operators from  $\mathcal{H}$  to  $\mathcal{G}$ .



## Hilbert spaces

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ . Its **adjoint  $L^*$**  is the operator in  $\mathcal{B}(\mathcal{G}, \mathcal{H})$  defined as

$$(\forall (x, y) \in \mathcal{H} \times \mathcal{G}) \quad \langle y \mid Lx \rangle_{\mathcal{G}} = \langle L^*y \mid x \rangle_{\mathcal{H}}.$$

## Hilbert spaces

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ . Its **adjoint  $L^*$**  is the operator in  $\mathcal{B}(\mathcal{G}, \mathcal{H})$  defined as

$$(\forall (x, y) \in \mathcal{H} \times \mathcal{G}) \quad \langle y \mid Lx \rangle = \langle L^*y \mid x \rangle.$$

## Hilbert spaces

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ . Its **adjoint  $L^*$**  is the operator in  $\mathcal{B}(\mathcal{G}, \mathcal{H})$  defined as

$$(\forall (x, y) \in \mathcal{H} \times \mathcal{G}) \quad \langle Lx \mid y \rangle = \langle x \mid L^*y \rangle.$$

# Hilbert spaces

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ . Its **adjoint  $L^*$**  is the operator in  $\mathcal{B}(\mathcal{G}, \mathcal{H})$  defined as

$$(\forall (x, y) \in \mathcal{H} \times \mathcal{G}) \quad \langle Lx \mid y \rangle = \langle x \mid L^*y \rangle.$$

Example:

If  $L: \mathcal{H} \rightarrow \mathcal{H}^n: y \mapsto (y, \dots, y)$

then  $L^*: \mathcal{H}^n \rightarrow \mathcal{H}: x = (x_1, \dots, x_n) \mapsto \sum_{i=1}^n x_i$

Proof:

$$\langle Ly \mid x \rangle = \langle (y, \dots, y) \mid (x_1, \dots, x_n) \rangle = \sum_{i=1}^n \langle y \mid x_i \rangle = \left\langle y \mid \sum_{i=1}^n x_i \right\rangle$$

## Hilbert spaces

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ . Its **adjoint  $L^*$**  is the operator in  $\mathcal{B}(\mathcal{G}, \mathcal{H})$  defined as

$$(\forall (x, y) \in \mathcal{H} \times \mathcal{G}) \quad \langle Lx \mid y \rangle = \langle x \mid L^*y \rangle.$$

- ▶ We have  $\|L^*\| = \|L\|$ .
- ▶ If  $L$  is bijective (i.e. an **isomorphism**) then  $L^{-1} \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  and  $(L^{-1})^* = (L^*)^{-1}$ .
- ▶ If  $\mathcal{H} = \mathbb{R}^N$  and  $\mathcal{G} = \mathbb{R}^M$  then  $L^* = L^\top$ .

# Hilbert spaces

Let  $\mathcal{H}$  be a Hilbert space and  $L \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ .

- ▶  $L$  is **self-adjoint** if  $L^* = L$ .
- ▶  $L$  is **positive** if  $(\forall x \in \mathcal{H}) \langle x | Lx \rangle \geq 0$ .
- ▶  $L$  is **strictly positive** if  $L$  is positive and if  $(\forall x \in \mathcal{H}) \langle x | Lx \rangle = 0 \Leftrightarrow x = 0$ .

## Mappings versus multivalued operators

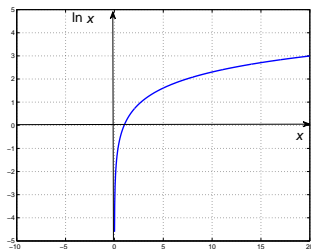
Let  $\mathcal{H}$  be a real Hilbert space.

$A$  is an  $\mathcal{H}$ -valued **mapping** defined on  $D \subset \mathcal{H}$  if

$$\begin{aligned} A: D &\rightarrow \mathcal{H} \\ x &\mapsto A(x) \end{aligned}$$

► Example:

$$\begin{aligned} A: ]0, +\infty[ &\rightarrow \mathbb{R} \\ x &\mapsto \ln x \end{aligned}$$



## Mappings versus multivalued operators

Let  $\mathcal{H}$  be a real Hilbert space.

$A$  is a (multivalued) operator if

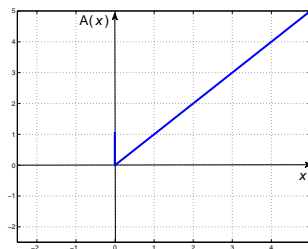
$$A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \mapsto \{A_i(x) \mid i \in I_x \subset \mathbb{R}\}$$

► Example:

$$A: \mathbb{R} \rightarrow 2^{\mathbb{R}}$$

$$x \mapsto \begin{cases} \{x\} & \text{if } x > 0 \\ [0, 1] & \text{if } x = 0 \\ \emptyset & \text{if } x < 0 \end{cases}$$





# Graph

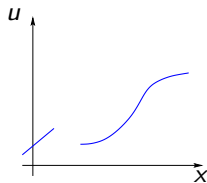
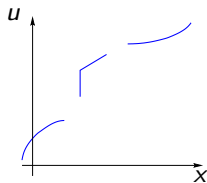
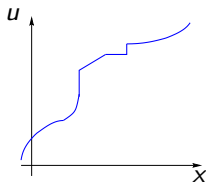
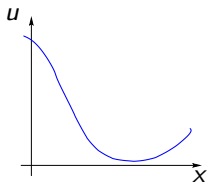
Let  $\mathcal{H}$  be a real Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

The **graph** of  $A$  is

$$\text{gra}A = \{(x, u) \in \mathcal{H}^2 \mid u \in Ax\}.$$

► Graph examples:



# Graph

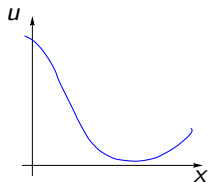
Let  $\mathcal{H}$  be a real Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

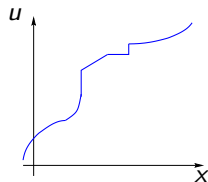
The **graph** of  $A$  is

$$\text{gra}A = \{(x, u) \in \mathcal{H}^2 \mid u \in Ax\}.$$

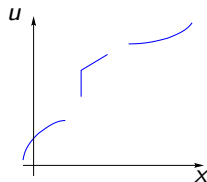
► Graph examples:



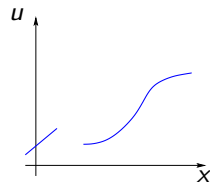
Single-valued



Multivalued



Multivalued



Single-valued

## Monotone operator: definition

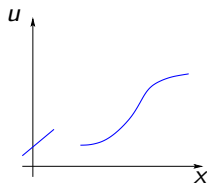
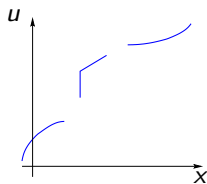
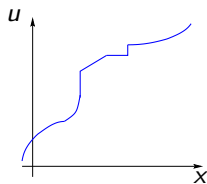
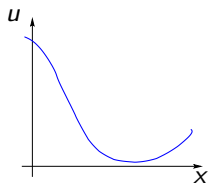
Let  $\mathcal{H}$  be a real Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

$A$  is **monotone** if

$$(\forall (x_1, u_1) \in \text{gra}A) (\forall (x_2, u_2) \in \text{gra}A) \quad \langle u_1 - u_2 \mid x_1 - x_2 \rangle \geq 0.$$

### ► Monotone operators ?



## Monotone operator: definition

Let  $\mathcal{H}$  be a real Hilbert space.

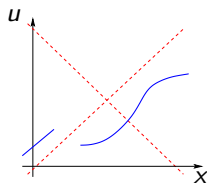
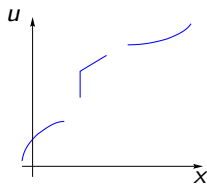
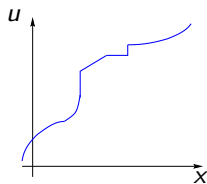
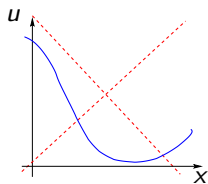
Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

$A$  is **monotone** if

$$(\forall (x_1, u_1) \in \text{gra}A) (\forall (x_2, u_2) \in \text{gra}A)$$

$$\langle u_1 - u_2 \mid x_1 - x_2 \rangle \geq 0.$$

### ► Monotone operators ?



## Monotone operator: example

Let  $\mathcal{H}$  be a real Hilbert space.

Let  $A \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ .

- ▶  $A$  is monotone  $\Leftrightarrow A$  is positive
- ▶  $A$  monotone  $\Leftrightarrow A + A^*$  monotone  $\Leftrightarrow A^*$  monotone.

## Monotone operator: example

Let  $\mathcal{H}$  be a real Hilbert space.

Let  $A \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ .

- ▶  $A$  is monotone  $\Leftrightarrow A$  is positive
- ▶  $A$  monotone  $\Leftrightarrow A + A^*$  monotone  $\Leftrightarrow A^*$  monotone.

Proof:

$$\begin{aligned}
 A \text{ monotone} &\Leftrightarrow (\forall (x_1, x_2) \in \mathcal{H}^2) \quad \langle x_1 - x_2 \mid Ax_1 - Ax_2 \rangle \geq 0 \\
 &\Leftrightarrow (\forall x \in \mathcal{H}) \quad 2 \langle x \mid Ax \rangle \geq 0 \\
 &\Leftrightarrow (\forall x \in \mathcal{H}) \quad \langle x \mid Ax \rangle + \langle A^*x \mid x \rangle \geq 0 \\
 &\Leftrightarrow (\forall x \in \mathcal{H}) \quad \langle x \mid (A + A^*)x \rangle \geq 0 \\
 &\Leftrightarrow A + A^* \text{ monotone}
 \end{aligned}$$

## Monotone operator: example

Let  $\mathcal{H}$  be a real Hilbert space.

Let  $A \in \mathcal{B}(\mathcal{H}, \mathcal{H})$ .

- ▶  $A$  is monotone  $\Leftrightarrow A$  is positive
- ▶  $A$  monotone  $\Leftrightarrow A + A^*$  monotone  $\Leftrightarrow A^*$  monotone.

- ▶ For  $A \in \mathcal{B}(\mathcal{H}, \mathcal{H})$  to be monotone,  $A$  is not required to be self-adjoint.  
Example :  $A \in \mathcal{B}(\mathcal{H}, \mathcal{H})$  skewed (i.e.  $A^* = -A$ ) is monotone.

# Domain

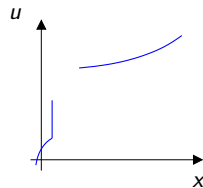
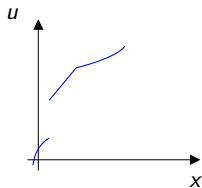
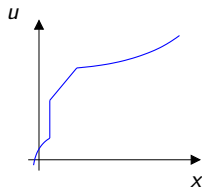
Let  $\mathcal{H}$  be a real Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

The **domain** of  $A$  is

$$\text{dom } A = \{x \in \mathcal{H} \mid Ax \neq \emptyset\}.$$

► Which domain ?





# Domain

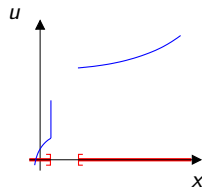
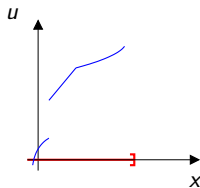
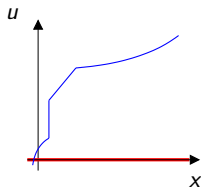
Let  $\mathcal{H}$  be a real Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

The **domain** of  $A$  is

$$\text{dom } A = \{x \in \mathcal{H} \mid Ax \neq \emptyset\}.$$

► Which domain ?



# Domain

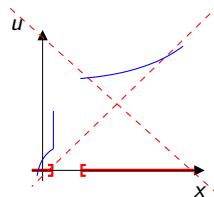
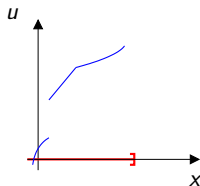
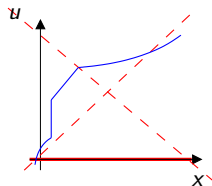
Let  $\mathcal{H}$  be a real Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

The **domain** of  $A$  is

$$\text{dom } A = \{x \in \mathcal{H} \mid Ax \neq \emptyset\}.$$

- ▶ Let  $C \subset \mathcal{H}$ . If  $\text{dom } A = C$  and for every  $x \in C$ ,  $Ax$  is a singleton, we view  $A$  as **a mapping from  $C$  to  $\mathcal{H}$** .



## Monotone operator: properties

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and  $B: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be two monotone operators.

The following operators are monotone:

- ▶  $x \mapsto y + \gamma \rho A(\rho x + z) = \{y + \gamma \rho u \mid u \in A(\rho x + z)\}$   
where  $(y, z) \in \mathcal{H}^2$ ,  $\gamma \in [0, +\infty[$  and  $\rho \in \mathbb{R}$ .
- ▶  $A \times B : \mathcal{H} \times \mathcal{G} \rightarrow 2^{\mathcal{H} \times \mathcal{G}}$   
 $(x, y) \mapsto Ax \times Ay = \{(u, v) \mid u \in Ax, v \in Bx\}$ .
- ▶  $A + B : x \mapsto \{u + v \mid u \in Ax, v \in Bx\}$  if  $\mathcal{G} = \mathcal{H}$ .
- ▶  $L^*BL : x \mapsto \{L^*v \mid v \in B(Lx)\}$  if  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

## Monotone operator: inversion

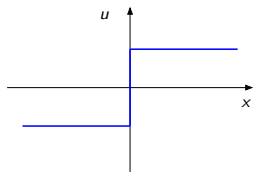
Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

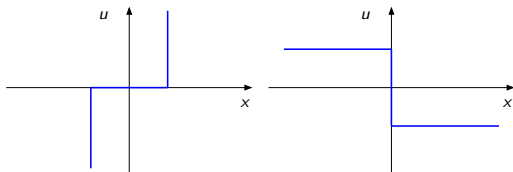
$A^{-1}$  is the operator from  $\mathcal{H}$  to  $2^{\mathcal{H}}$  the graph of which is

$$\text{gra}(A^{-1}) = \{(u, x) \mid (x, u) \in \text{gra}A\}.$$

Graph of  $A$



Graph of  $A^{-1}$  ?



## Monotone operator: inversion

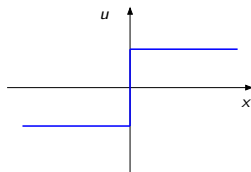
Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

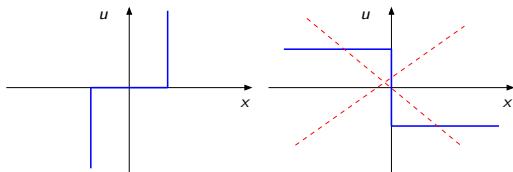
$A^{-1}$  is the operator from  $\mathcal{H}$  to  $2^{\mathcal{H}}$  the graph of which is

$$\text{gra}(A^{-1}) = \{(u, x) \mid (x, u) \in \text{gra}A\}.$$

Graph of  $A$



Graph of  $A^{-1}$  ?



## Monotone operator: inversion

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

$A^{-1}$  is the operator from  $\mathcal{H}$  to  $2^{\mathcal{H}}$  the graph of which is

$$\text{gra}(A^{-1}) = \{(u, x) \mid (x, u) \in \text{gra}A\}.$$

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a monotone operator.

$A^{-1}$  is monotone .

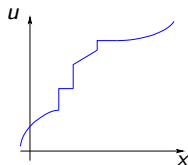
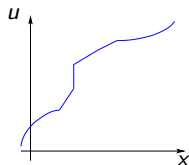
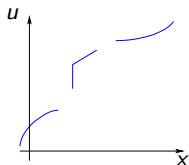
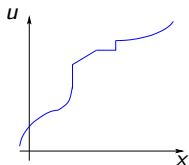
## Maximally monotone operator: definition

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

$A$  is **maximally monotone** if  $A$  is monotone and if there exists no monotone operator  $B : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  (different from  $A$ ) such that  $\text{gra}B$  properly contains  $\text{gra}A$ .

Maximally monotone operator ?



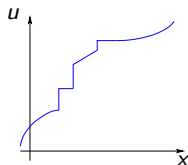
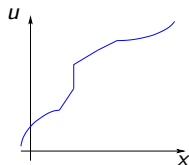
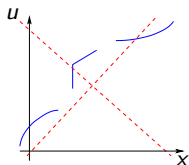
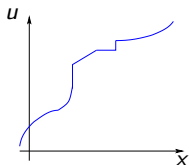
## Maximally monotone operator: definition

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

$A$  is **maximally monotone** if  $A$  is monotone and if there exists no monotone operator  $B : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  (different from  $A$ ) such that  $\text{gra}B$  properly contains  $\text{gra}A$ .

Maximally monotone operator ?





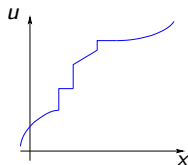
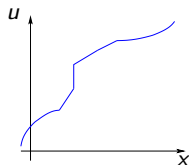
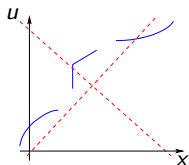
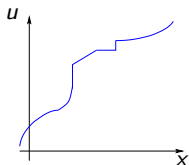
## Maximally monotone operator: definition

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

$A$  is **maximally monotone** if  $A$  is monotone and if there exists no monotone operator  $B : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  (different from  $A$ ) such that  $\text{gra}B$  properly contains  $\text{gra}A$ .

Maximally monotone operator ?



## Maximally monotone operator: second definition

Let  $\mathcal{H}$  be a Hilbert space.

$A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  is maximally monotone if one of the following equivalent conditions is satisfied:

- (i)  $A$  is monotone and there exists no monotone operator  $B : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  such that  $\text{gra}B$  properly contains  $\text{gra}A$ .
- (ii) For every  $(x_1, u_1) \in \mathcal{H}^2$ ,

$$(x_1, u_1) \in \text{gra}A \Leftrightarrow (\forall (x_2, u_2) \in \text{gra}A) \quad \langle x_1 - x_2 \mid u_1 - u_2 \rangle \geq 0.$$

## Continuous functions

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow \mathcal{H}$  be monotone and continuous. Then  $A$  is maximally monotone.

Example :

If  $L \in \mathcal{B}(\mathcal{H}, \mathcal{H})$  is positive, then  $L$  is maximally monotone.

Maximally  
monotone  
operator

Maximally  
monotone  
operator



Maximally  
monotone  
operator



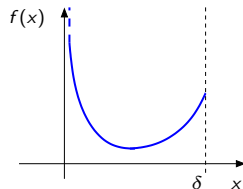
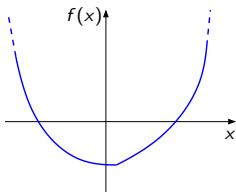
Usefulness in  
convex  
optimization

## Convex analysis: definitions

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  where  $\mathcal{H}$  is a Hilbert space.

- ▶ The **domain** of  $f$  is  $\text{dom } f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$ .
- ▶ The function  $f$  is **proper** if  $\text{dom } f \neq \emptyset$ .

### Domains of the functions ?

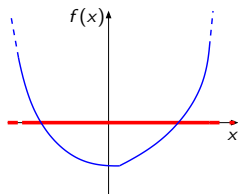


## Convex analysis: definitions

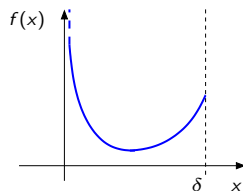
Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  where  $\mathcal{H}$  is a Hilbert space.

- ▶ The **domain** of  $f$  is  $\text{dom } f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$ .
- ▶ The function  $f$  is **proper** if  $\text{dom } f \neq \emptyset$ .

### Domains of the functions ?



$\text{dom } f = \mathbb{R}$   
(proper)



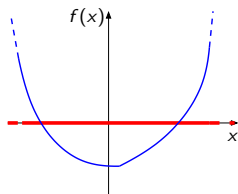


## Convex analysis: definitions

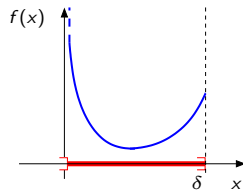
Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  where  $\mathcal{H}$  is a Hilbert space.

- ▶ The **domain** of  $f$  is  $\text{dom } f = \{x \in \mathcal{H} \mid f(x) < +\infty\}$ .
- ▶ The function  $f$  is **proper** if  $\text{dom } f \neq \emptyset$ .

### Domains of the functions ?



$\text{dom } f = \mathbb{R}$   
(proper)



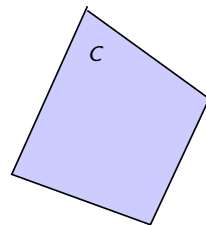
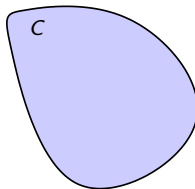
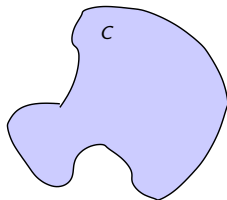
$\text{dom } f = ]0, \delta]$   
(proper)

## Convex analysis: definitions

$C \subset \mathcal{H}$  is a **convex set** if

$$(\forall (x, y) \in C^2)(\forall \alpha \in ]0, 1[) \quad \alpha x + (1 - \alpha)y \in C$$

Convex sets ?

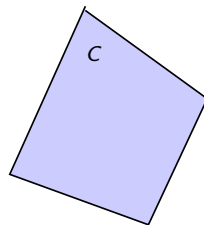
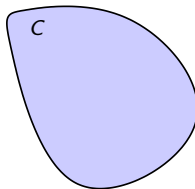
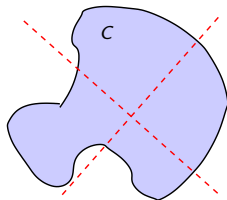


## Convex analysis: definitions

$C \subset \mathcal{H}$  is a **convex set** if

$$(\forall (x, y) \in C^2)(\forall \alpha \in ]0, 1[) \quad \alpha x + (1 - \alpha)y \in C$$

Convex sets ?



## Convex analysis: definitions

$f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  is a **convex function** if

$$(\forall (x, y) \in \mathcal{H}^2)(\forall \alpha \in ]0, 1[)$$

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

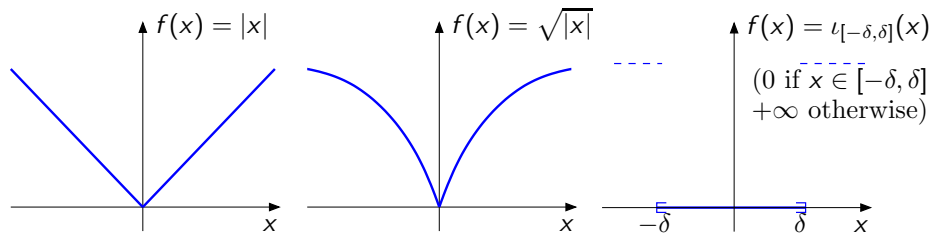
## Convex analysis: definitions

$f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  is a **convex function** if

$$(\forall (x, y) \in \mathcal{H}^2)(\forall \alpha \in ]0, 1[)$$

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

### Convex functions ?



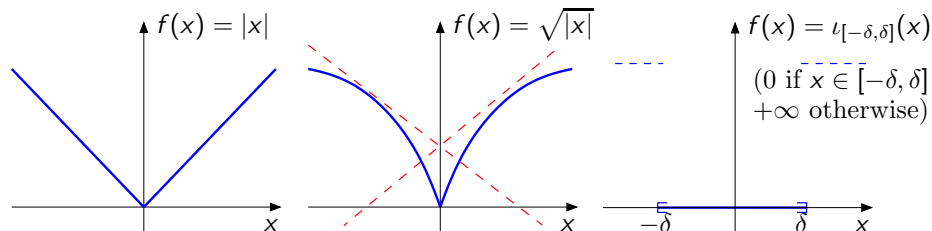
# Convex analysis: definitions

$f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  is a **convex function** if

$$(\forall (x, y) \in \mathcal{H}^2)(\forall \alpha \in ]0, 1[)$$

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

## Convex functions ?



## Convex analysis: definitions

$f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  is convex  $\Leftrightarrow$  the **epigraph** of  $f$ , i.e.

$$\text{epi } f = \{(x, \zeta) \in \text{dom } f \times \mathbb{R} \mid f(x) \leq \zeta\}$$

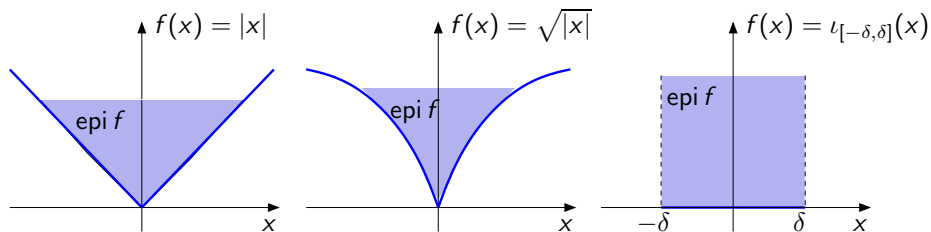
is convex.

## Convex analysis: definitions

$f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  is convex  $\Leftrightarrow$  the **epigraph** of  $f$ , i.e.

$$\text{epi } f = \{(x, \zeta) \in \text{dom } f \times \mathbb{R} \mid f(x) \leq \zeta\}$$

is convex.



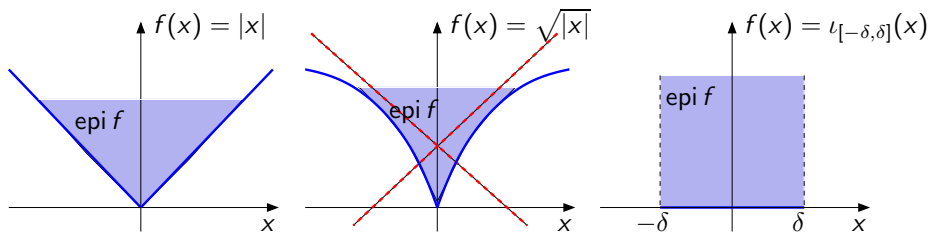


## Convex analysis: definitions

$f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  is convex  $\Leftrightarrow$  the **epigraph** of  $f$ , i.e.

$$\text{epi } f = \{(x, \zeta) \in \text{dom } f \times \mathbb{R} \mid f(x) \leq \zeta\}$$

is convex.

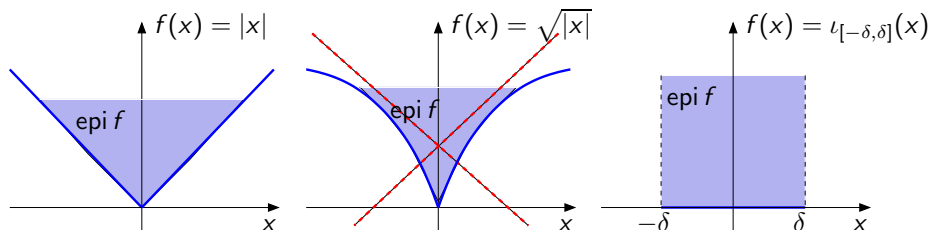


## Convex analysis: definitions

$f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  is convex  $\Leftrightarrow$  the **epigraph** of  $f$ , i.e.

$$\text{epi } f = \{(x, \zeta) \in \text{dom } f \times \mathbb{R} \mid f(x) \leq \zeta\}$$

is convex.



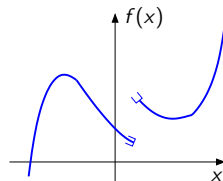
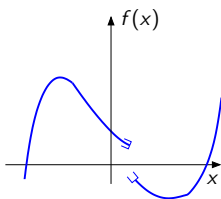
►  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  is concave if  $-f$  is convex.

## Convex analysis: definitions

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

$f$  is a **lower semi-continuous** (l.s.c.) function on  $\mathcal{H}$  if  $\text{epi } f$  is closed

► l.s.c. functions ?

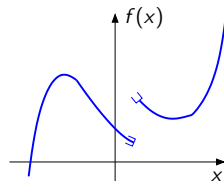
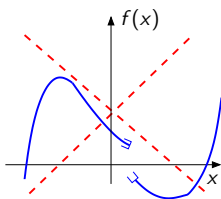


## Convex analysis: definitions

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

$f$  is a **lower semi-continuous** (l.s.c.) function on  $\mathcal{H}$  if  $\text{epi } f$  is closed

► l.s.c. functions ?

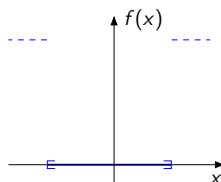
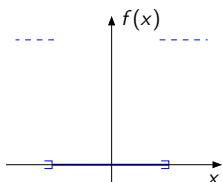


## Convex analysis: definitions

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

$f$  is a **lower semi-continuous** (l.s.c.) function on  $\mathcal{H}$  if  $\text{epi } f$  is closed

► l.s.c. functions ?

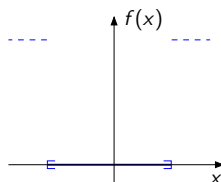
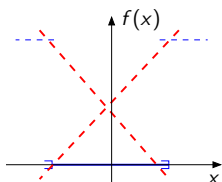


## Convex analysis: definitions

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

$f$  is a **lower semi-continuous** (l.s.c.) function on  $\mathcal{H}$  if  $\text{epi } f$  is closed

► l.s.c. functions ?

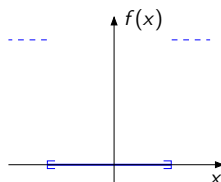
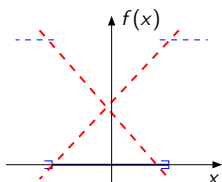


## Convex analysis: definitions

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

$f$  is a **lower semi-continuous** (l.s.c.) function on  $\mathcal{H}$  if  $\text{epi } f$  is closed

► l.s.c. functions ?



## Convex analysis: definitions/properties

- ▶  $\Gamma_0(\mathcal{H})$ : class of convex, l.s.c., and proper functions from  $\mathcal{H}$  to  $] -\infty, +\infty]$ .
- ▶ Every continuous function on  $\mathcal{H}$  is l.s.c.
- ▶ Every finite sum of l.s.c. (convex) functions is l.s.c. (convex).
- ▶ Let  $(f_i)_{i \in I}$  be a family of l.s.c. (convex) functions.  $\sup_{i \in I} f_i$  is l.s.c. (convex).



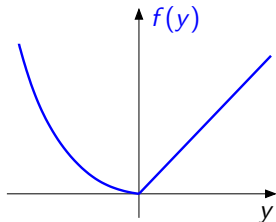
## Subdifferential of a convex function: definition

The (Moreau) subdifferential of  $f$ , denoted by  $\partial f$ ,

## Subdifferential of a convex function: definition

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) **subdifferential of  $f$** , denoted by  $\partial f$ ,



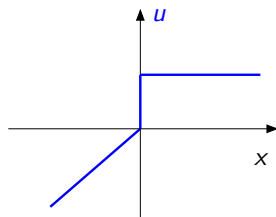
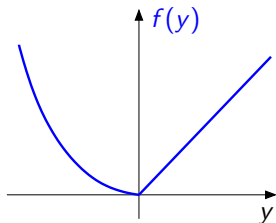
## Subdifferential of a convex function: definition

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) **subdifferential of  $f$** , denoted by  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x | u \rangle + f(x) \leq f(y)\}$$



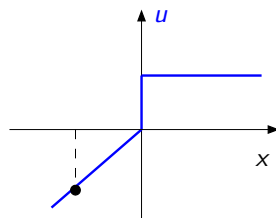
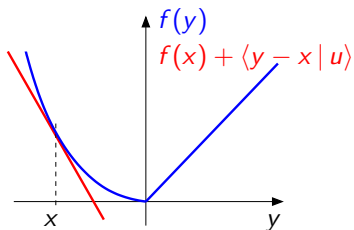
## Subdifferential of a convex function: definition

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) **subdifferential of  $f$** , denoted by  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x \mid u \rangle + f(x) \leq f(y)\}$$



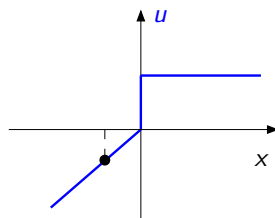
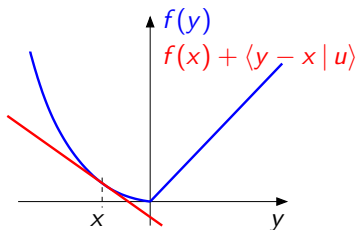
## Subdifferential of a convex function: definition

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) **subdifferential of  $f$** , denoted by  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x \mid u \rangle + f(x) \leq f(y)\}$$



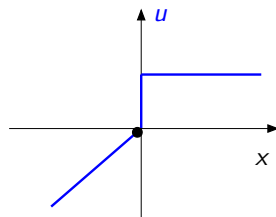
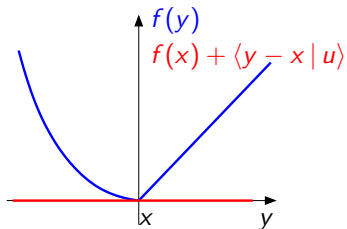
## Subdifferential of a convex function: definition

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) **subdifferential of  $f$** , denoted by  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x \mid u \rangle + f(x) \leq f(y)\}$$



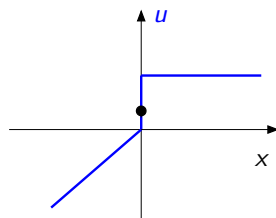
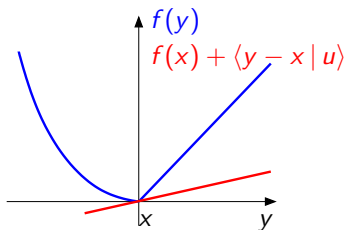
## Subdifferential of a convex function: definition

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) **subdifferential of  $f$** , denoted by  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x \mid u \rangle + f(x) \leq f(y)\}$$



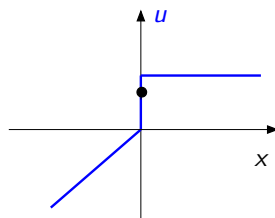
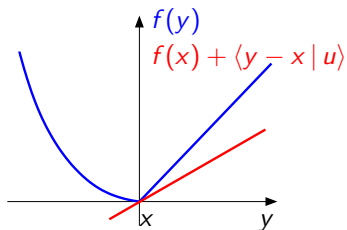
## Subdifferential of a convex function: definition

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) **subdifferential of  $f$** , denoted by  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x \mid u \rangle + f(x) \leq f(y)\}$$





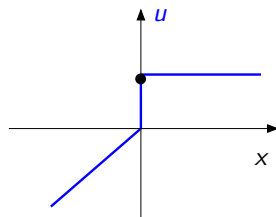
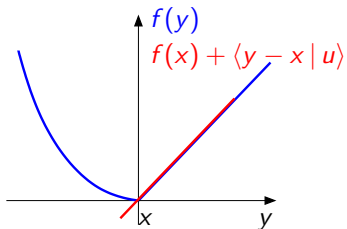
## Subdifferential of a convex function: definition

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) **subdifferential of  $f$** , denoted by  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x \mid u \rangle + f(x) \leq f(y)\}$$



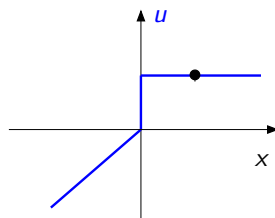
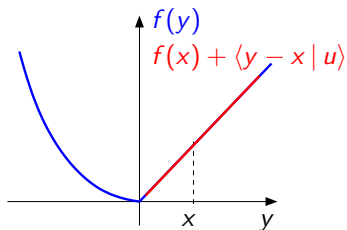
## Subdifferential of a convex function: definition

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) **subdifferential of  $f$** , denoted by  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x \mid u \rangle + f(x) \leq f(y)\}$$



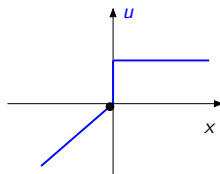
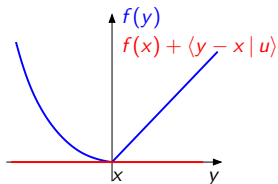
## Subdifferential of a convex function: properties

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) subdifferential of  $f$ , denoted by  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x \mid u \rangle + f(x) \leq f(y)\}$$



► Fermat rule:

$$\begin{aligned} 0 \in \partial f(x) &\Leftrightarrow (\forall y \in \mathcal{H}) \langle y - x \mid 0 \rangle + f(x) \leq f(y) \\ &\Leftrightarrow x \in \text{Argmin} f \end{aligned}$$

## Subdifferential of a convex function: properties

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) subdifferential of  $f$ , denoted  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x, u \rangle + f(x) \leq f(y)\}$$

- ▶  $u \in \partial f(x)$  is a **subgradient** of  $f$  at  $x$ .

## Subdifferential of a convex function: properties

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) subdifferential of  $f$ , denoted  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x, u \rangle + f(x) \leq f(y)\}$$

- ▶  $u \in \partial f(x)$  is a **subgradient** of  $f$  at  $x$ .
- ▶ If  $x \notin \text{dom } f$ , then  $\partial f(x) = \emptyset$ .

## Subdifferential of a convex function: properties

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) subdifferential of  $f$ , denoted  $\partial f$ , is such that

$$\begin{aligned} \partial f : \mathcal{H} &\rightarrow 2^{\mathcal{H}} \\ x &\rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x, u \rangle + f(x) \leq f(y)\} \end{aligned}$$

- ▶  $u \in \partial f(x)$  is a **subgradient** of  $f$  at  $x$ .
- ▶ If  $x \notin \text{dom } f$ , then  $\partial f(x) = \emptyset$ .
- ▶ For every  $x \in \text{dom } f$ ,  $\partial f(x)$  is a closed and convex set.

## Subdifferential of a convex function: properties

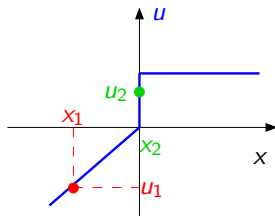
Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) subdifferential of  $f$ , denoted  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x \mid u \rangle + f(x) \leq f(y)\}$$

- ▶  $\partial f$  is a monotone operator :



## Subdifferential of a convex function: properties

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

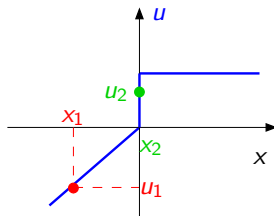
The (Moreau) subdifferential of  $f$ , denoted  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x \mid u \rangle + f(x) \leq f(y)\}$$

- ▶  $\partial f$  is a monotone operator :

Let  $u_1 \in \partial f(x_1)$  and  $u_2 \in \partial f(x_2)$ .





## Subdifferential of a convex function: properties

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) subdifferential of  $f$ , denoted  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x \mid u \rangle + f(x) \leq f(y)\}$$

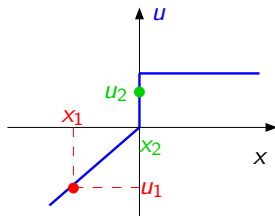
- ▶  $\partial f$  is a monotone operator :

Let  $u_1 \in \partial f(x_1)$  and  $u_2 \in \partial f(x_2)$ .

By using the subdifferential definition:

$$\langle x_2 - x_1 \mid u_1 \rangle + f(x_1) \leq f(x_2)$$

$$\langle x_1 - x_2 \mid u_2 \rangle + f(x_2) \leq f(x_1)$$



## Subdifferential of a convex function: properties

Let  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

The (Moreau) subdifferential of  $f$ , denoted  $\partial f$ , is such that

$$\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}}$$

$$x \rightarrow \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \langle y - x \mid u \rangle + f(x) \leq f(y)\}$$

- ▶  $\partial f$  is a monotone operator :

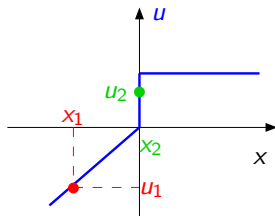
Let  $u_1 \in \partial f(x_1)$  and  $u_2 \in \partial f(x_2)$ .

By using the subdifferential definition:

$$\langle x_2 - x_1 \mid u_1 \rangle + f(x_1) \leq f(x_2)$$

$$\langle x_1 - x_2 \mid u_2 \rangle + f(x_2) \leq f(x_1)$$

and thus  $\langle x_1 - x_2 \mid u_1 - u_2 \rangle \geq 0$ .



## Subdifferential of a convex function: properties

- ▶ The subdifferential of a convex and proper function is:
  - ▶ Monotone
  - ▶ If  $f$  is Gâteaux differentiable at  $x$ , then  $\partial f(x) = \{\nabla f(x)\}$

## Subdifferential of a convex function: properties

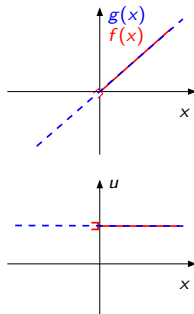
- ▶ The subdifferential of a convex and proper function is:
  - ▶ Monotone
  - ▶ If  $f$  is Gâteaux differentiable at  $x$ , then  $\partial f(x) = \{\nabla f(x)\}$
  - ▶ Non necessarily maximally monotone

Counterexample: For every  $x \in \mathcal{H}$ ,

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ +\infty & \text{otherwise} \end{cases}, \quad g(x) = x$$

$$\Rightarrow \partial f(x) = \begin{cases} \{1\} & \text{if } x > 0 \\ \emptyset & \text{otherwise} \end{cases}, \quad \partial g(x) = \{1\}.$$

Consequently,  $\text{gra} \partial f = ]0, +\infty[ \times \{1\} \subset \mathbb{R} \times \{1\}$   
 $\subset \text{gra} \partial g$



## Subdifferential of a convex function: properties

- ▶ The subdifferential of a convex, proper and l.s.c. function is
  - ▶ Maximally monotone

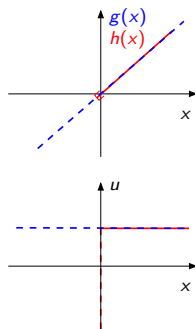
## Subdifferential of a convex function: properties

- ▶ The subdifferential of a convex, proper and l.s.c. function is
  - ▶ Maximally monotone

Example: For every  $x \in \mathcal{H}$ ,

$$h(x) = \begin{cases} x & \text{if } x \geq 0 \\ +\infty & \text{otherwise} \end{cases}, \quad g(x) = x$$

$$\Rightarrow \partial h(x) = \begin{cases} \{1\} & \text{if } x > 0 \\ ]-\infty, 1] & \text{if } x = 0 \\ \emptyset & \text{otherwise} \end{cases}, \quad \partial g(x) = \{1\}.$$



Consequently,  $\text{grad} \partial h \not\subset \text{grad} \partial g$ .

## Subdifferential of a convex function: properties

- ▶ The subdifferential of a convex, proper and l.s.c. function is
  - ▶ Maximally monotone
  - ▶ If  $\mathcal{H} = \mathbb{R}$ , equivalence between both properties.

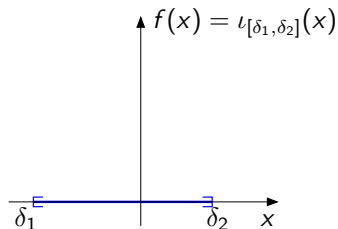
## Subdifferential of a convex function: example

Soit  $C \subset \mathcal{H}$ .

The indicator function of  $C$  is

$$(\forall x \in \mathcal{H}) \quad \iota_C(x) = \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{otherwise.} \end{cases}$$

Example :  $C = [\delta_1, \delta_2]$

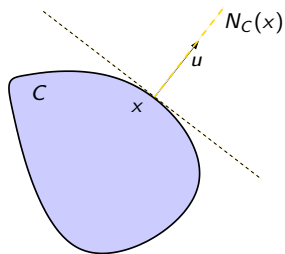
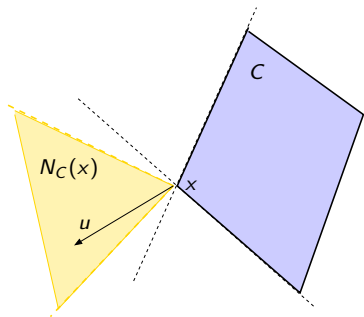




## Subdifferential of a convex function: example

For every  $x \in \mathcal{H}$ ,  $\partial \iota_C(x)$  is the **normal cone** to  $C$  at  $x$  defined by

$$N_C(x) = \begin{cases} \{u \in \mathcal{H} \mid (\forall y \in C) \langle u \mid y - x \rangle \leq 0\} & \text{if } x \in C \\ \emptyset & \text{otherwise.} \end{cases}$$

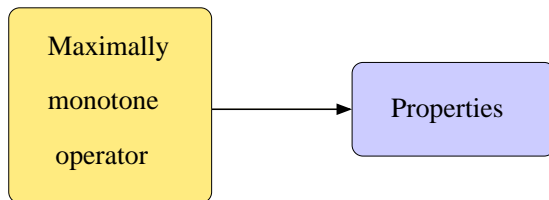


## Subdifferential of a convex function: example

For every  $x \in \mathcal{H}$ ,  $\partial \iota_C(x)$  is the **normal cone** to  $C$  at  $x$  defined by

$$N_C(x) = \begin{cases} \{u \in \mathcal{H} \mid (\forall y \in C) \langle u \mid y - x \rangle \leq 0\} & \text{if } x \in C \\ \emptyset & \text{otherwise.} \end{cases}$$

- ▶ If  $x \in \text{int } C$  then  $N_C(x) = \{0\}$ .
- ▶ If  $C$  is a vector space then, for every  $x \in C$ ,  $N_C(x) = C^\perp$ .



## Maximally monotone operator: properties

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator.

For every  $x \in \mathcal{H}$ ,  $Ax$  is a closed convex.

Proof:

$$Ax = \bigcap_{(x', u') \in \text{gra}A} \{u \in \mathcal{H} \mid \langle x - x' \mid u - u' \rangle \geq 0\}.$$

Consequently,  $Ax$  is an intersection of closed convex sets.

## Maximally monotone operator: properties

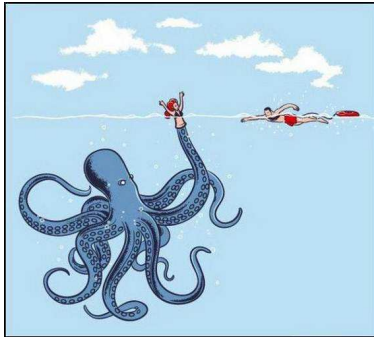
Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and  $B: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be two maximally monotone operators.

The following operators are maximally monotone:

- ▶  $y + \gamma\rho A(\rho \cdot + z)$  where  $(y, z) \in \mathcal{H}^2$ ,  $\gamma \in [0, +\infty[$  and  $\rho \in \mathbb{R}$ ;
- ▶  $A \times B$ ,
- ▶  $A^{-1}$ .

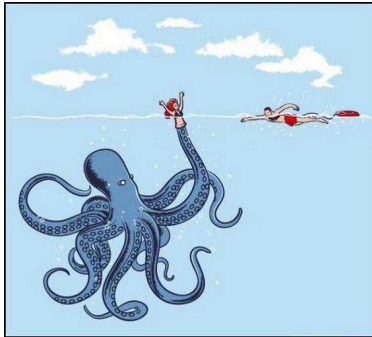
Inverse of a maximally  
monotone operator



Inverse of a maximally  
monotone operator



Usefulness ?



## Conjugate: definition

Let  $\mathcal{H}$  be a Hilbert space and  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

The **conjugate** of  $f$  is the function  $f^*: \mathcal{H} \rightarrow [-\infty, +\infty]$  such that

$$(\forall u \in \mathcal{H}) \quad f^*(u) = \sup_{x \in \mathcal{H}} (\langle x | u \rangle - f(x)).$$

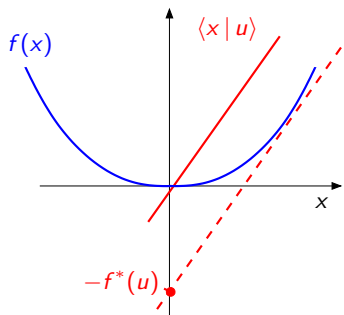


## Conjugate: definition

Let  $\mathcal{H}$  be a Hilbert space and  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

The **conjugate** of  $f$  is the function  $f^*: \mathcal{H} \rightarrow [-\infty, +\infty]$  such that

$$(\forall u \in \mathcal{H}) \quad f^*(u) = \sup_{x \in \mathcal{H}} (\langle x | u \rangle - f(x)).$$

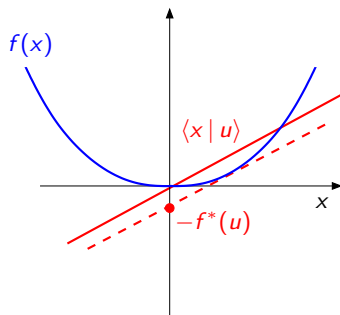


## Conjugate: definition

Let  $\mathcal{H}$  be a Hilbert space and  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

The **conjugate** of  $f$  is the function  $f^*: \mathcal{H} \rightarrow [-\infty, +\infty]$  such that

$$(\forall u \in \mathcal{H}) \quad f^*(u) = \sup_{x \in \mathcal{H}} (\langle x | u \rangle - f(x)).$$

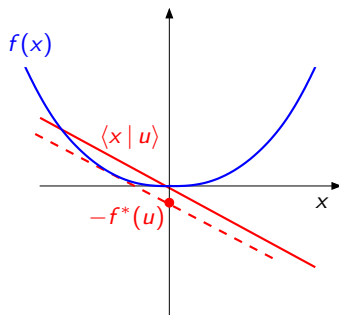


## Conjugate: definition

Let  $\mathcal{H}$  be a Hilbert space and  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

The **conjugate** of  $f$  is the function  $f^*: \mathcal{H} \rightarrow [-\infty, +\infty]$  such that

$$(\forall u \in \mathcal{H}) \quad f^*(u) = \sup_{x \in \mathcal{H}} (\langle x | u \rangle - f(x)).$$



## Conjugate: definition

Let  $\mathcal{H}$  be a Hilbert space and  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

The **conjugate** of  $f$  is the function  $f^*: \mathcal{H} \rightarrow [-\infty, +\infty]$  such that

$$(\forall u \in \mathcal{H}) \quad f^*(u) = \sup_{x \in \mathcal{H}} (\langle x | u \rangle - f(x)).$$

Example :

►  $(\forall x \in \mathbb{R}^N) f(x) = \frac{1}{q} \|x\|_q^q$  with  $q \in ]1, +\infty[$

$\Rightarrow (\forall u \in \mathbb{R}^N) f^*(u) = \frac{1}{q^*} \|u\|_{q^*}^{q^*}$  with  $\frac{1}{q} + \frac{1}{q^*} = 1$

## Conjugate: definition

Let  $\mathcal{H}$  be a Hilbert space and  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

The **conjugate** of  $f$  is the function  $f^*: \mathcal{H} \rightarrow [-\infty, +\infty]$  such that

$$(\forall u \in \mathcal{H}) \quad f^*(u) = \sup_{x \in \mathcal{H}} (\langle x | u \rangle - f(x)).$$

- ▶  $f^*$  is l.s.c. and convex.

## Conjugate: definition

Let  $\mathcal{H}$  be a Hilbert space and  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

The **conjugate** of  $f$  is the function  $f^*: \mathcal{H} \rightarrow [-\infty, +\infty]$  such that

$$(\forall u \in \mathcal{H}) \quad f^*(u) = \sup_{x \in \mathcal{H}} (\langle x | u \rangle - f(x)).$$

### Moreau-Fenchel theorem

Let  $\mathcal{H}$  be a Hilbert space and  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

$$f \text{ is l.s.c. and convex} \Leftrightarrow f^{**} = f.$$

- Consequence: If  $f \in \Gamma_0(\mathcal{H})$  then  $f^* \in \Gamma_0(\mathcal{H})$ .

## Conjugate: definition

Let  $\mathcal{H}$  be a Hilbert space and  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

The **conjugate** of  $f$  is the function  $f^*: \mathcal{H} \rightarrow [-\infty, +\infty]$  such that

$$(\forall u \in \mathcal{H}) \quad f^*(u) = \sup_{x \in \mathcal{H}} (\langle x | u \rangle - f(x)).$$

Let  $f \in \Gamma_0(\mathcal{H})$ .

$$(\partial f)^{-1} = \partial f^*$$

# Conjugate versus Fourier transform

Property	conjugate		Fourier transform	
	$h(x)$	$h^*(u)$	$h(x)$	$\hat{h}(\nu)$
invariant function	$\frac{1}{2} \ x\ ^2$	$\frac{1}{2} \ u\ ^2$	$\exp(-\pi \ x\ ^2)$	$\exp(-\pi \ \nu\ ^2)$
translation $c \in \mathcal{H}$	$f(x - c)$	$f^*(u) + \langle u   c \rangle$	$f(x - c)$	$\exp(-j2\pi \langle \nu   c \rangle) \hat{f}(\nu)$
dual translation $c \in \mathcal{H}$	$f(x) + \langle x   c \rangle$	$f^*(u - c)$	$\exp(j2\pi \langle x   c \rangle) f(x - c)$	$\hat{f}(\nu - c)$
scalar multiplication $\alpha \in ]0, +\infty[$	$\alpha f(x)$	$\alpha f^*\left(\frac{u}{\alpha}\right)$	$\alpha f(x)$	$\alpha \hat{f}(\nu)$
scaling $\alpha \in \mathbb{R}^*$	$f\left(\frac{x}{\alpha}\right)$	$f^*(\alpha u)$	$f\left(\frac{x}{\alpha}\right)$	$ \alpha  \hat{f}(\alpha \nu)$
isomorphism $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$	$f(Lx)$	$f^*((L^{-1})^\top u)$	$f(Lx)$	$\frac{1}{ \det(L) } \hat{f}((L^{-1})^\top \nu)$
reflection	$f(-x)$	$f^*(-u)$	$f(-x)$	$\hat{f}(-\nu)$
separability	$\sum_{n=1}^N \varphi_n(x^{(n)})$ $x = (x^{(n)})_{1 < n < N}$	$\sum_{n=1}^N \varphi_n^*(u^{(n)})$ $u = (u^{(n)})_{1 < n < N}$	$\prod_{n=1}^N \varphi_n(x^{(n)})$ $x = (x^{(n)})_{1 < n < N}$	$\prod_{n=1}^N \hat{\varphi}_n(\nu^{(n)})$ $\nu = (\nu^{(n)})_{1 < n < N}$
isotropy	$\psi(\ x\ )$	$\psi^*(\ u\ )$	$\psi(\ x\ )$	$\psi(\ \nu\ )$
identity element of convolution	$\iota_{\{0\}}(x)$	0	$\delta(x)$	1
identity element of addition/product	0	$\iota_{\{0\}}(u)$	1	$\delta(\nu)$

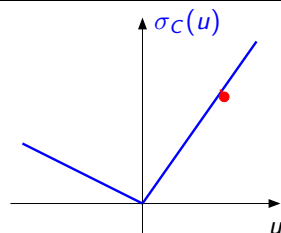
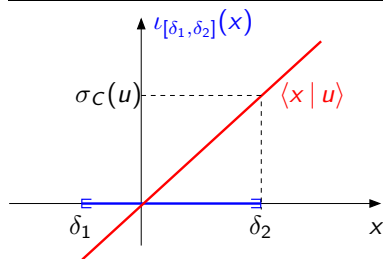


## Conjugate: example

Let  $\mathcal{H}$  be a Hilbert space and  $C \subset \mathcal{H}$ .

$\sigma_C$  is the **support function** of  $C$  if

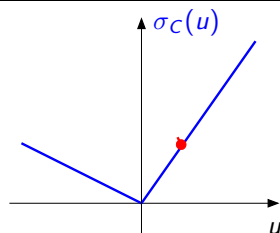
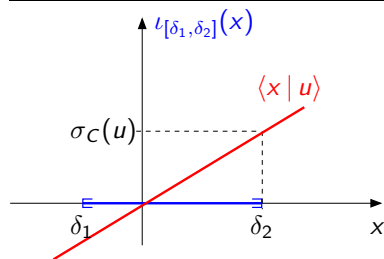
$$\begin{aligned}
 (\forall u \in \mathcal{H}) \quad \sigma_C(u) &= \sup_{x \in C} \langle x | u \rangle \\
 &= \iota_C^*(u).
 \end{aligned}$$



## Conjugate: example

Let  $\mathcal{H}$  be a Hilbert space and  $C \subset \mathcal{H}$ .  
 $\sigma_C$  is the **support function** of  $C$  if

$$\begin{aligned}
 (\forall u \in \mathcal{H}) \quad \sigma_C(u) &= \sup_{x \in C} \langle x | u \rangle \\
 &= \iota_C^*(u).
 \end{aligned}$$

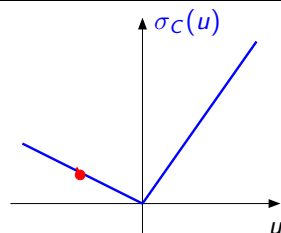
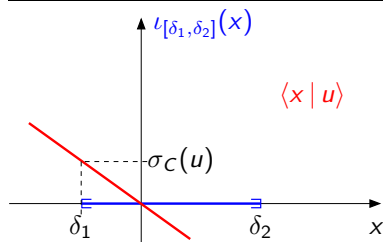


## Conjugate: example

Let  $\mathcal{H}$  be a Hilbert space and  $C \subset \mathcal{H}$ .

$\sigma_C$  is the **support function** of  $C$  if

$$\begin{aligned}
 (\forall u \in \mathcal{H}) \quad \sigma_C(u) &= \sup_{x \in C} \langle x | u \rangle \\
 &= \iota_C^*(u).
 \end{aligned}$$



## Conjugate: example

Let  $\mathcal{H}$  be a Hilbert space.

$f: \mathcal{H} \rightarrow ]-\infty, +\infty]$  is **positively homogeneous** if

$$(\forall x \in \mathcal{H})(\forall \alpha \in ]0, +\infty[) \quad f(\alpha x) = \alpha f(x).$$

$f$  is positively homogeneous and belongs to  $\Gamma_0(\mathcal{H})$



$f = \sigma_C$  where  $C$  is a nonempty closed convex subset of  $\mathcal{H}$ .

## Conjugate: example

Let  $\mathcal{H}$  be a Hilbert space.

$f: \mathcal{H} \rightarrow ]-\infty, +\infty]$  is **positively homogeneous** if

$$(\forall x \in \mathcal{H})(\forall \alpha \in ]0, +\infty[) \quad f(\alpha x) = \alpha f(x).$$

$f$  is positively homogeneous and belongs to  $\Gamma_0(\mathcal{H})$



$f = \sigma_C$  where  $C$  is a nonempty closed convex subset of  $\mathcal{H}$ .

► Example 1: Let  $f: \mathbb{R} \rightarrow ]-\infty, +\infty] : x \mapsto \begin{cases} \delta_1 x & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ \delta_2 x & \text{if } x > 0 \end{cases}$

with  $-\infty \leq \delta_1 < \delta_2 \leq +\infty$ . Then,  $f = \sigma_C$  where  $C$  is the closed real interval such that  $\inf C = \delta_1$  and  $\sup C = \delta_2$ .

## Conjugate: example

Let  $\mathcal{H}$  be a Hilbert space.

$f: \mathcal{H} \rightarrow ]-\infty, +\infty]$  is **positively homogeneous** if

$$(\forall x \in \mathcal{H})(\forall \alpha \in ]0, +\infty[) \quad f(\alpha x) = \alpha f(x).$$

$f$  is positively homogeneous and belongs to  $\Gamma_0(\mathcal{H})$



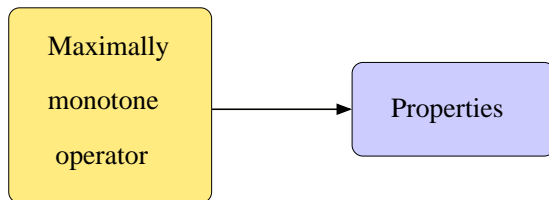
$f = \sigma_C$  where  $C$  is a nonempty closed convex subset of  $\mathcal{H}$ .

- ▶ Example 2: Let  $f$  be a  $\ell^q$  norm of  $\mathbb{R}^N$  with  $q \in [1, +\infty]$ .

We have  $f = \sigma_C$  where

$$C = \{y \in \mathbb{R}^N \mid \|y\|_{q^*} \leq 1\} \quad \text{with } \frac{1}{q} + \frac{1}{q^*} = 1.$$

Particular case :  $\ell^1$  norm of  $\mathbb{R}^N \Rightarrow C = [-1, 1]^N$ .



## Maximally monotone operator: sum

Let  $A$  and  $B$  be two maximally monotone operators.  
 $A + B$  is monotone but may not be maximally monotone.



## Maximally monotone operator: sum

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A$  and  $B$  be two maximally monotone operators from  $\mathcal{H}$  to  $2^{\mathcal{H}}$  such that one of the following assumptions is satisfied:

- ▶  $\text{dom } B = \mathcal{H}$
- ▶  $\text{dom } A \cap \text{int}(\text{dom } B) \neq \emptyset$
- ▶  $0 \in \text{int}(\text{dom } A - \text{dom } B)$

then  $A + B$  is maximally monotone.

## Maximally monotone operator: sum

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A$  and  $B$  be two maximally monotone operators from  $\mathcal{H}$  to  $2^{\mathcal{H}}$  such that one of the following assumptions is satisfied:

- ▶  $\text{dom } B = \mathcal{H}$
- ▶  $\text{dom } A \cap \text{int}(\text{dom } B) \neq \emptyset$
- ▶  $0 \in \text{int}(\text{dom } A - \text{dom } B)$

then  $A + B$  is maximally monotone.

Consequence: Let  $\alpha \in [0, +\infty[$ . If  $A$  is maximally monotone, then  $A + \alpha \text{Id}$  is maximally monotone.

## Maximally monotone operator: linear transform

Let  $\mathcal{H}$  and  $\mathcal{G}$  two Hilbert spaces.

Let  $B: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be a maximally monotone operator and  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that one of the following assumptions is satisfied:

- ▶  $L$  surjective
- ▶  $0 \in \text{int}(\text{dom } B - \text{ran } L)$

then  $L^*BL$  is maximally monotone.

## Maximally monotone operator: linear transform

Let  $\mathcal{H}$  and  $\mathcal{G}$  two Hilbert spaces.

Let  $B: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be a maximally monotone operator and  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that one of the following assumptions is satisfied:

- ▶  $L$  surjective
- ▶  $0 \in \text{int}(\text{dom } B - \text{ran } L)$

then  $L^*BL$  is maximally monotone.

Consequence: Let  $\mu \in ]0, +\infty[$ .

If  $A$  is maximally monotone and  $LL^* = \mu\text{Id}$ , then  $L^*AL$  is maximally monotone.

Proof:  $LL^* = \mu\text{Id} \Rightarrow \text{ran } L = \mathcal{H}$ .

## Part 2: Nonexpansive operators

1. Background on nonexpansive operators
  - ▶ Definition
  - ▶ Properties
  - ▶ Examples
  - ▶ Resolvent
2. Proximal operator
  - ▶ Definition
  - ▶ Properties
  - ▶ Examples

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$ .

$A$  is **nonexpansive** if  $(\forall(x, y) \in C^2) \quad \|Ax - Ay\| \leq \|x - y\|$ .

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$  and  $\nu \in ]0, +\infty[$

$\nu^{-1}A$  is nonexpansive if  $(\forall (x, y) \in C^2) \quad \|Ax - Ay\| \leq \nu \|x - y\|$ .

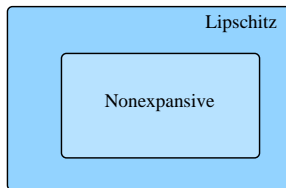
## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$  and  $\nu \in ]0, +\infty[$

$\nu^{-1}A$  is nonexpansive if  $(\forall (x, y) \in C^2) \quad \|Ax - Ay\| \leq \nu \|x - y\|$ .

$\nu^{-1}A$  is nonexpansive  $\Leftrightarrow A$  is  $\nu$ -Lipschitzian .





## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$

$A$  is **firmly nonexpansive** if

$$(\forall (x, u) \in \text{gra}A)(\forall (y, v) \in \text{gra}A) \quad \|u - v\|^2 \leq \langle u - v \mid x - y \rangle .$$

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$ .

$A$  is **firmly nonexpansive** if

$$(\forall x \in C)(\forall y \in C) \quad \|Ax - Ay\|^2 \leq \langle Ax - Ay \mid x - y \rangle .$$

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$ .

$A$  is **firmly nonexpansive** if

$$(\forall (x, y) \in C^2) \quad \|Ax - Ay\|^2 + \|(\text{Id} - A)x - (\text{Id} - A)y\|^2 \leq \|x - y\|^2.$$

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$ .

$A$  is **firmly nonexpansive** if

$$(\forall (x, y) \in C^2) \quad \|Ax - Ay\|^2 + \|(\text{Id} - A)x - (\text{Id} - A)y\|^2 \leq \|x - y\|^2.$$

- ▶  $A$  is firmly nonexpansive  $\Leftrightarrow$   $\text{Id} - A$  is firmly nonexpansive.
- ▶  $A$  is firmly nonexpansive  $\Leftrightarrow$   $2A - \text{Id}$  is nonexpansive.

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$ .

$A$  is **firmly nonexpansive** if

$$(\forall (x, y) \in C^2) \quad \|Ax - Ay\|^2 + \|(\text{Id} - A)x - (\text{Id} - A)y\|^2 \leq \|x - y\|^2.$$

- ▶  $A$  is firmly nonexpansive  $\Leftrightarrow$   $\text{Id} - A$  is firmly nonexpansive.
- ▶  $A$  is firmly nonexpansive  $\Leftrightarrow$   $\underbrace{2A - \text{Id}}$  is nonexpansive.  
**Reflection of A**

## Nonexpansive operator: definition

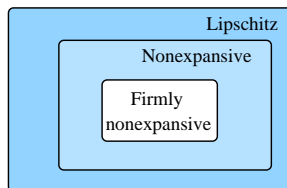
Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$ .

$A$  is **firmly nonexpansive** if

$$(\forall (x, y) \in C^2) \quad \|Ax - Ay\|^2 + \|(\text{Id} - A)x - (\text{Id} - A)y\|^2 \leq \|x - y\|^2.$$

$A$  is firmly nonexpansive  $\Rightarrow A$  is nonexpansive.



## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$  and  $\beta \in ]0, +\infty[$ .

$A$  is  $\beta$ -cocoercive if  $\beta A$  is firmly nonexpansive, i.e.,

$$(\forall x \in C)(\forall y \in C) \quad \beta \|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay \rangle .$$

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$  and  $\beta \in ]0, +\infty[$ .

$A$  is  $\beta$ -cocoercive if  $\beta A$  is firmly nonexpansive, i.e.,

$$(\forall x \in C)(\forall y \in C) \quad \beta \|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay \rangle .$$

- ▶ Let  $\mathcal{H}$  and  $\mathcal{G}$  be two real Hilbert spaces,  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  nonzero, and  $A: \mathcal{H} \rightarrow \mathcal{H}$ .  $A$  is  $\beta$ -cocoercive  $\Rightarrow L^*AL$  is  $\|L\|^{-2}\beta$ -cocoercive.



## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$  and  $\beta \in ]0, +\infty[$ .

$A$  is  $\beta$ -cocoercive if  $\beta A$  is firmly nonexpansive, i.e.,

$$(\forall x \in C)(\forall y \in C) \quad \beta \|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay \rangle .$$

- ▶ Let  $\mathcal{H}$  and  $\mathcal{G}$  be two real Hilbert spaces,  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  nonzero, and  $A: \mathcal{H} \rightarrow \mathcal{H}$ .  $A$  is  $\beta$ -cocoercive  $\Rightarrow L^*AL$  is  $\|L\|^{-2}\beta$ -cocoercive.
- ▶  $A$  is  $\beta$ -cocoercive  $\Rightarrow A$  is  $\beta^{-1}$ -Lipschitzian.

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$  and  $\beta \in ]0, +\infty[$ .

$A$  is  $\beta$ -cocoercive if  $\beta A$  is firmly nonexpansive, i.e.,

$$(\forall x \in C)(\forall y \in C) \quad \beta \|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay \rangle .$$

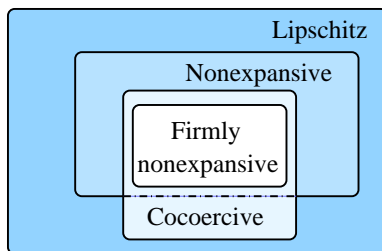
- ▶ Let  $\mathcal{H}$  and  $\mathcal{G}$  be two real Hilbert spaces,  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  nonzero, and  $A: \mathcal{H} \rightarrow \mathcal{H}$ .  $A$  is  $\beta$ -cocoercive  $\Rightarrow L^*AL$  is  $\|L\|^{-2}\beta$ -cocoercive.
- ▶  $A$  is  $\beta$ -cocoercive  $\Rightarrow A$  is  $\beta^{-1}$ -Lipschitzian.
- ▶  $A: \mathcal{H} \rightarrow \mathcal{H}$  is  $\beta$ -cocoercive  $\Rightarrow A$  is maximally monotone.

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .  
Let  $A: C \rightarrow \mathcal{H}$  and  $\beta \in ]0, +\infty[$ .

$A$  is  $\beta$ -cocoercive if  $\beta A$  is firmly nonexpansive, i.e.,

$$(\forall x \in C)(\forall y \in C) \quad \beta \|Ax - Ay\|^2 \leq \langle x - y \mid Ax - Ay \rangle .$$



## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A : C \rightarrow \mathcal{H}$  and let  $\alpha \in ]0, 1[$ .

$A$  is  $\alpha$ -averaged if there exists a nonexpansive operator  $R : C \rightarrow \mathcal{H}$  such that

$$A = (1 - \alpha)\text{Id} + \alpha R .$$

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A : C \rightarrow \mathcal{H}$  and let  $\alpha \in ]0, 1[$ .

$A$  is  $\alpha$ -averaged if

$$(\forall (x, y) \in C^2) \quad \|Ax - Ay\|^2 + \frac{1 - \alpha}{\alpha} \|(\text{Id} - A)x - (\text{Id} - A)y\|^2 \leq \|x - y\|^2.$$

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .  
Let  $A : C \rightarrow \mathcal{H}$  and let  $\alpha \in ]0, 1[$ .

$A$  is  $\alpha$ -averaged if

$$(\forall (x, y) \in C^2) \quad \|Ax - Ay\|^2 + \frac{1 - \alpha}{\alpha} \|(\text{Id} - A)x - (\text{Id} - A)y\|^2 \leq \|x - y\|^2.$$

- ▶  $A$  is  $\alpha$ -averaged  $\Rightarrow A$  is nonexpansive.
- ▶  $A$  is  $\frac{1}{2}$ -averaged  $\Leftrightarrow A$  is firmly nonexpansive.
- ▶  $A$  is  $\alpha$ -averaged  $\Rightarrow A$  is  $\alpha'$ -averaged for every  $\alpha' \in [\alpha, 1[$ .
- ▶ Let  $\lambda \in ]0, 1/\alpha[$ .  $A$  is  $\alpha$ -averaged  $\Rightarrow (1 - \lambda)\text{Id} + \lambda A$  is  $\lambda\alpha$ -averaged.

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .  
Let  $A : C \rightarrow \mathcal{H}$  and let  $\alpha \in ]0, 1[$ .

$A$  is  $\alpha$ -averaged if

$$(\forall (x, y) \in C^2) \quad \|Ax - Ay\|^2 + \frac{1 - \alpha}{\alpha} \|(\text{Id} - A)x - (\text{Id} - A)y\|^2 \leq \|x - y\|^2.$$

- ▶ Let  $(\omega_i)_{1 \leq i \leq n} \in ]0, 1]^n$  be such that  $\sum_{i=1}^n \omega_i = 1$  and let  $(\alpha_i)_{1 \leq i \leq n} \in ]0, 1[^n$ . If, for every  $i \in \{1, \dots, n\}$ ,  $A_i : C \rightarrow \mathcal{H}$  is  $\alpha_i$ -averaged, then  $\sum_{i=1}^n \omega_i A_i$  is  $\alpha$ -averaged with  $\alpha = \max_{1 \leq i \leq n} \alpha_i$ .
- ▶ Let  $(\alpha_i)_{1 \leq i \leq n} \in ]0, 1[^n$ . If, for every  $i \in \{1, \dots, n\}$ ,  $A_i : C \rightarrow C$  is  $\alpha_i$ -averaged, then  $A_1 \cdots A_n$  is  $\alpha$ -averaged with

$$\alpha = \frac{n}{n - 1 + \frac{1}{\max_{1 \leq i \leq n} \alpha_i}}.$$

## Nonexpansive operator: definition

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .  
Let  $A : C \rightarrow \mathcal{H}$  and let  $\alpha \in ]0, 1[$ .

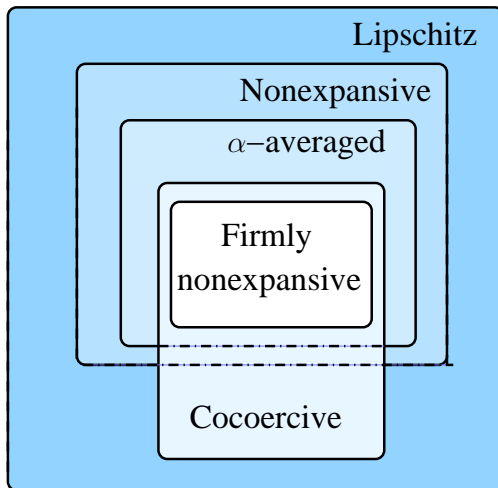
$A$  is  $\alpha$ -averaged if

$$(\forall (x, y) \in C^2) \quad \|Ax - Ay\|^2 + \frac{1 - \alpha}{\alpha} \|(\text{Id} - A)x - (\text{Id} - A)y\|^2 \leq \|x - y\|^2.$$

$A : \mathcal{H} \rightarrow \mathcal{H}$  is  $\alpha$ -averaged with  $\alpha \in ]0, 1/2]$   $\Rightarrow A$  is maximally monotone.

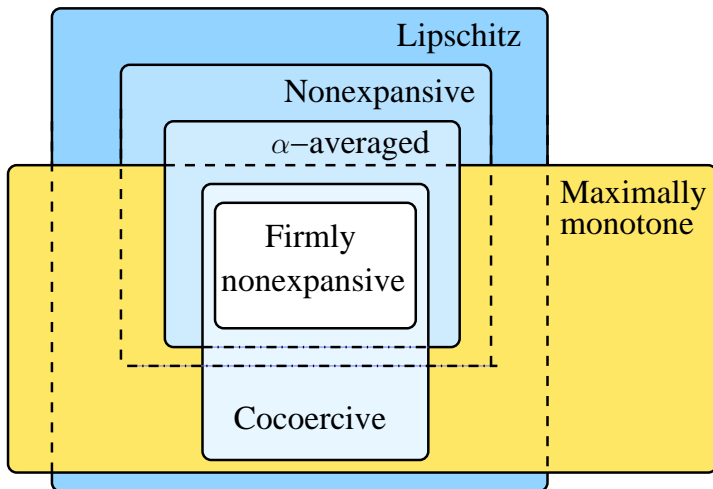


## Nonexpansive operator: recap



## Nonexpansive operator: recap

(if the domain  $C$  is equal to  $\mathcal{H}$ )



## Nonexpansive operator: properties

Let  $\mathcal{H}$  be a real Hilbert space and let  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $A: C \rightarrow \mathcal{H}$ .

Let  $\beta \in ]0, +\infty[$  and  $\gamma \in ]0, 2\beta[$ .

If  $A$  is  $\beta$ -cocoercive, then  $\text{Id} - \gamma A$  is  $\gamma/(2\beta)$ -averaged.

Proof :

$A$   $\beta$ -cocoercive  $\Leftrightarrow \beta A$  firmly nonexpansive.

There exists a nonexpansive operator  $R: C \rightarrow \mathcal{H}$  such that

$$\beta A = (\text{Id} + R)/2.$$

Thus

$$\text{Id} - \gamma A = \left(1 - \frac{\gamma}{2\beta}\right)\text{Id} + \frac{\gamma}{2\beta}(-R).$$

$(-R)$  being nonexpansive,  $\text{Id} - \gamma A$  is  $\gamma/(2\beta)$ -averaged.

## Nonexpansive operators



Nonexpansive operators



What is their use ?



## Nonexpansive operator: example

### Descent lemma

Let  $\mathcal{H}$  be a real Hilbert space,  $f: \mathcal{H} \rightarrow \mathbb{R}$  and  $\nu \in ]0, +\infty[$ .

If  $f$  is differentiable and its gradient is  $\nu$ -Lipschitzian, then

$$(\forall (x, y) \in \mathcal{H}^2) \quad f(y) \leq f(x) + \langle y - x \mid \nabla f(x) \rangle + \frac{\nu}{2} \|y - x\|^2.$$

### Baillon-Haddad theorem

Let  $\mathcal{H}$  be a real Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $\nu \in ]0, +\infty[$ .

If  $f$  is differentiable, then  $\nabla f$   $\nu$ -Lipschitzian  $\Leftrightarrow \nabla f$   $\nu^{-1}$ -cocoercive.

## Nonexpansive operator: example

### Descent lemma

Let  $\mathcal{H}$  be a real Hilbert space,  $f: \mathcal{H} \rightarrow \mathbb{R}$  and  $\nu \in ]0, +\infty[$ .

If  $f$  is differentiable and its gradient is  $\nu$ -Lipschitzian, then

$$(\forall (x, y) \in \mathcal{H}^2) \quad f(y) \leq f(x) + \langle y - x \mid \nabla f(x) \rangle + \frac{\nu}{2} \|y - x\|^2.$$

### Baillon-Haddad theorem

Let  $\mathcal{H}$  be a real Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $\nu \in ]0, +\infty[$ .

If  $f$  is differentiable, then  $\nabla f$   $\nu$ -Lipschitzian  $\Leftrightarrow \nabla f$   $\nu^{-1}$ -cocoercive.

Let  $\mathcal{H}$  be a Hilbert space,  $f \in \Gamma_0(\mathcal{H})$ ,  $\nu \in ]0, +\infty[$  and  $\gamma \in ]0, 2\nu^{-1}[$ .  
 $f$  differentiable and  $\nabla f$   $\nu$ -Lipschitzian  $\Rightarrow \text{Id} - \gamma \nabla f$  is  $\gamma\nu/2$ -averaged.

## Nonexpansive operator: example

### Descent lemma

Let  $\mathcal{H}$  be a real Hilbert space,  $f: \mathcal{H} \rightarrow \mathbb{R}$  and  $\nu \in ]0, +\infty[$ .

If  $f$  is differentiable and its gradient is  $\nu$ -Lipschitzian, then

$$(\forall (x, y) \in \mathcal{H}^2) \quad f(y) \leq f(x) + \langle y - x \mid \nabla f(x) \rangle + \frac{\nu}{2} \|y - x\|^2.$$

### Baillon-Haddad theorem

Let  $\mathcal{H}$  be a real Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $\nu \in ]0, +\infty[$ .

If  $f$  is differentiable, then  $\nabla f$   $\nu$ -Lipschitzian  $\Leftrightarrow \nabla f$   $\nu^{-1}$ -cocoercive.

Let  $\mathcal{H}$  be a Hilbert space,  $f \in \Gamma_0(\mathcal{H})$ ,  $\nu \in ]0, +\infty[$  and  $\gamma \in ]0, 2\nu^{-1}[$ .

$f$  differentiable and  $\nabla f$   $\nu$ -Lipschitzian  $\Rightarrow$   $\underbrace{\text{Id} - \gamma \nabla f}_{\text{gradient descent operator}}$  is  $\gamma\nu/2$ -

averaged.



## Nonexpansive operator: example

### Descent lemma

Let  $\mathcal{H}$  be a real Hilbert space,  $f: \mathcal{H} \rightarrow \mathbb{R}$  and  $\nu \in ]0, +\infty[$ .

If  $f$  is differentiable and its gradient is  $\nu$ -Lipschitzian, then

$$(\forall (x, y) \in \mathcal{H}^2) \quad f(y) \leq f(x) + \langle y - x \mid \nabla f(x) \rangle + \frac{\nu}{2} \|y - x\|^2.$$

### Baillon-Haddad theorem

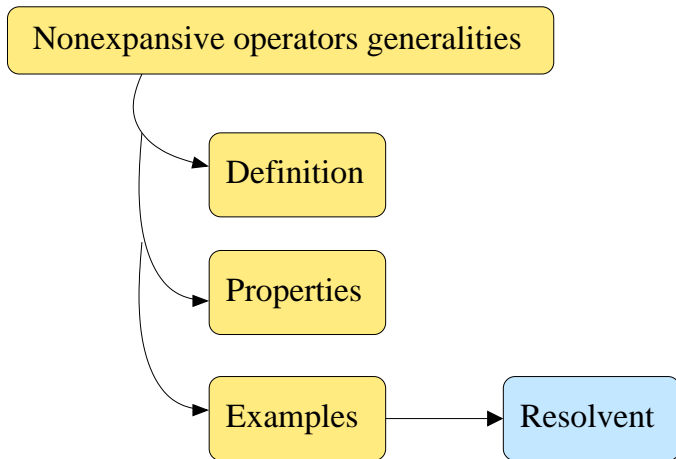
Let  $\mathcal{H}$  be a real Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $\nu \in ]0, +\infty[$ .

If  $f$  is differentiable, then  $\nabla f$   $\nu$ -Lipschitzian  $\Leftrightarrow \nabla f$   $\nu^{-1}$ -cocoercive.

Let  $\mathcal{H}$  be a Hilbert space,  $f \in \Gamma_0(\mathcal{H})$ , and  $\nu \in ]0, +\infty[$

$f$  differentiable and  $\nabla f$   $\nu$ -Lipschitzian  $\Leftrightarrow f^*$  is  $\nu^{-1}$ -strongly convex.

Remark :  $f^*$  is  $\nu^{-1}$ -strongly convex if  $f^* - \nu^{-1} \|\cdot\|^2/2$  is convex.



## Resolvent: definition

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

The **resolvent** of  $A$  is

$$J_A = (\text{Id} + A)^{-1}.$$

## Resolvent: definition

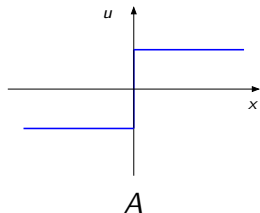
Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

The **resolvent** of  $A$  is

$$J_A = (\text{Id} + A)^{-1}.$$

► Example :



$A + \text{Id} ?$

$J_A ?$

## Resolvent: definition

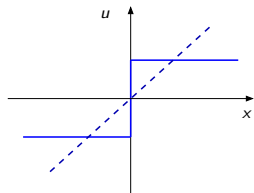
Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

The **resolvent** of  $A$  is

$$J_A = (\text{Id} + A)^{-1}.$$

► Example :



$A$  and  $\text{Id}$

$A + \text{Id} ?$

$J_A ?$

## Resolvent: definition

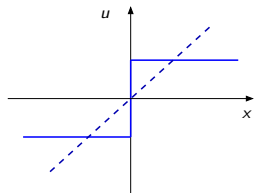
Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

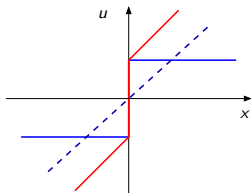
The **resolvent** of  $A$  is

$$J_A = (\text{Id} + A)^{-1}.$$

► Example :



$A$  and  $\text{Id}$



$A + \text{Id}$

$J_A ?$

## Resolvent: definition

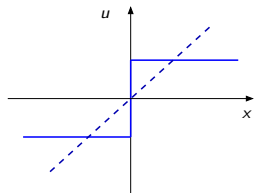
Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

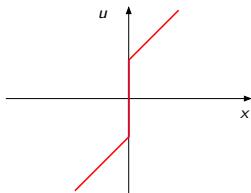
The **resolvent** of  $A$  is

$$J_A = (\text{Id} + A)^{-1}.$$

► Example :



$A$  and  $\text{Id}$



$A + \text{Id}$

$J_A ?$

## Resolvent: definition

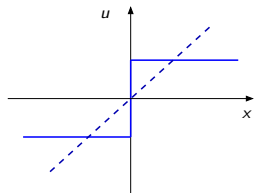
Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

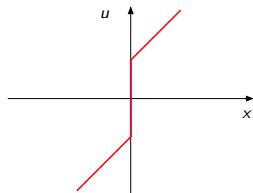
The **resolvent** of  $A$  is

$$J_A = (\text{Id} + A)^{-1}.$$

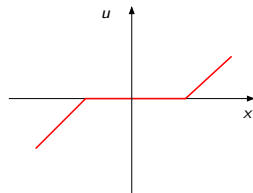
► Example :



$A$  and Id



$A + \text{Id}$



$J_A$



## Resolvent: definition

The **range of an operator**  $B: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  is

$$\text{ran } B = \{u \in \mathcal{H} \mid \exists x \in \mathcal{H}, u \in Bx\}.$$

### Minty theorem

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a monotone operator.

$$\text{ran}(\text{Id} + A) = \mathcal{H} \quad \Leftrightarrow \quad A \text{ is maximally monotone.}$$

## Resolvent: properties

Let  $\mathcal{H}$  be a real Hilbert space. Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .  
A is monotone  $\Leftrightarrow J_A$  is firmly nonexpansive.

Remark :  $J_A : \text{ran}(\text{Id} + A) \rightarrow \mathcal{H}$ .

## Resolvent: properties

Let  $\mathcal{H}$  be a real Hilbert space. Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .  
A is monotone  $\Leftrightarrow J_A$  is firmly nonexpansive.

Let  $\mathcal{H}$  be a real Hilbert space. Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .  
A is maximally monotone  $\Leftrightarrow J_A : \mathcal{H} \rightarrow \mathcal{H}$  is firmly nonexpansive.

Proof: A monotone  $\Leftrightarrow J_A : \text{ran}(\text{Id} + A) \rightarrow \mathcal{H}$  firmly nonexpansive  
+ Minty theorem.

## Resolvent: properties

Let  $\mathcal{H}$  be a real Hilbert space. Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .  
A is monotone  $\Leftrightarrow J_A$  is firmly nonexpansive.

Let  $\mathcal{H}$  be a real Hilbert space. Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .  
A is maximally monotone  $\Leftrightarrow J_A : \mathcal{H} \rightarrow \mathcal{H}$  is firmly nonexpansive.

Let  $\mathcal{H}$  be a Hilbert space. Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  maximally monotone and  $\gamma \in ]0, +\infty[$ . For every  $x \in \mathcal{H}$ , there exists a unique  $p \in \mathcal{H}$  such that  $x - p \in \gamma A p$  and thus  $p = J_{\gamma A} x$ .

Proof:  $x \in (\text{Id} + \gamma A)(p) \Leftrightarrow p \in (\text{Id} + \gamma A)^{-1} x \Leftrightarrow p = J_{\gamma A} x$

## Resolvent: properties

Let  $\mathcal{H}$  be a real Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone and  $\gamma \in ]0, +\infty[$ .

- ▶  $J_{\gamma A}$  and  $\text{Id} - J_{\gamma A}$  are firmly nonexpansive.
- ▶ The **reflected resolvent**  $R_{\gamma A} = 2J_{\gamma A} - \text{Id}$  is nonexpansive.
- ▶  $\gamma A$  is  $\gamma$ -cocoercive.

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and  $\gamma \in ]0, +\infty[$ .

The **Yosida approximation** of  $A$  of index  $\gamma$  is

$$\gamma A = \frac{1}{\gamma}(\text{Id} - J_{\gamma A}).$$

## Resolvent: properties

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator.

- ▶ Let  $z \in \mathcal{H}$  and  $B = A(\cdot - z)$ . Then  $J_B = z + J_A(\cdot - z)$ .
- ▶ Let  $z \in \mathcal{H}$  and  $B = z + A$ . Then  $J_B = J_A(\cdot - z)$ .
- ▶ Let  $\alpha \in [0, +\infty[$  and  $B = A + \alpha \text{Id}$ . Then  $J_B = J_{\frac{A}{1+\alpha}} \left( \frac{\cdot}{1+\alpha} \right)$

## Resolvent: properties

For every  $i \in \{1, \dots, n\}$ , let  $\mathcal{H}_i$  be a Hilbert space and  $A_i: \mathcal{H}_i \rightarrow 2^{\mathcal{H}_i}$  be a maximally monotone operator.

$$J_{A_1 \times \dots \times A_n} = J_{A_1} \times \dots \times J_{A_n}: \mathcal{H}_1 \times \dots \times \mathcal{H}_n \rightarrow \mathcal{H}_1 \times \dots \times \mathcal{H}_n$$
$$(x_1, \dots, x_n) \mapsto (J_{A_1} x_1, \dots, J_{A_n} x_n).$$

## Resolvent: properties

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator and  $\gamma \in ]0, +\infty[$ .

$$J_{\gamma A^{-1}} = \text{Id} - \gamma J_{\gamma^{-1}A}(\gamma^{-1}\cdot)$$



## Resolvent: properties

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator and  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  such that  $LL^* = \mu \text{Id}$  where  $\mu \in ]0, +\infty[$ . Then

$$J_{L^*AL} = \text{Id} - L^* \circ {}^{\mu}A \circ L .$$

## Resolvent: properties

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator and  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  such that  $LL^* = \mu \text{Id}$  where  $\mu \in ]0, +\infty[$ . Then

$$J_{L^*AL} = \text{Id} - L^* \circ \mu A \circ L.$$

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator and  $L \in \mathcal{B}(\mathcal{H}, \mathcal{H})$  be a unitary operator. Then

$$J_{L^*AL} = L^* J_A L.$$

## Resolvent: properties

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator and  $B = \rho A(\rho \cdot)$  where  $\rho \in \mathbb{R}^*$ . Then

$$J_B = \rho^{-1} J_{\rho^2 A}(\rho \cdot).$$

Preuve: Set  $L = \rho \text{Id}$  and apply formula

$$J_{L^* A L} = \text{Id} - L^* \circ \mu A \circ L.$$

## Resolvent: properties

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator and  $B = \rho A(\rho \cdot)$  where  $\rho \in \mathbb{R}^*$ . Then

$$J_B = \rho^{-1} J_{\rho^2 A}(\rho \cdot).$$

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator and  $B = -A(-\cdot)$ . Then

$$J_B = -J_A(-\cdot).$$

Resolvent



Wowai?



Resolvent



Wowai?!



Proximity  
operator

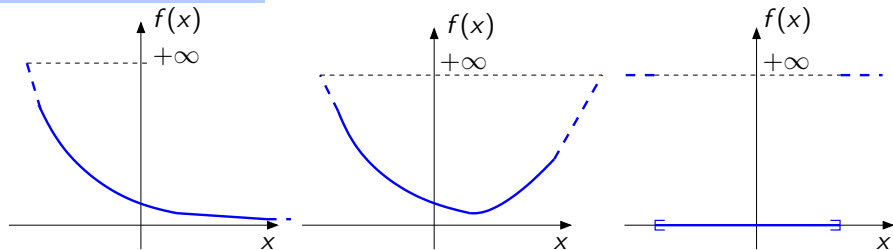
## Convex analysis

Let  $\mathcal{H}$  be a Hilbert space. Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .  
 $f$  is **coercive** if  $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$ .

# Convex analysis

Let  $\mathcal{H}$  be a Hilbert space. Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .  
 $f$  is **coercive** if  $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$ .

Coercive functions ?

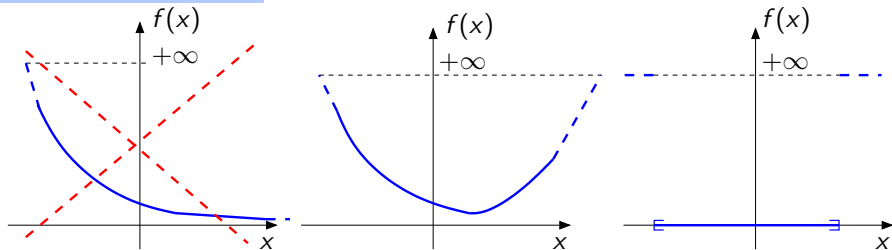




# Convex analysis

Let  $\mathcal{H}$  be a Hilbert space. Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .  
 $f$  is **coercive** if  $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$ .

Coercive functions ?



## Convex analysis

Let  $\mathcal{H}$  be a Hilbert space. Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

$f$  is **strictly convex** if

$$(\forall x \in \text{dom } f)(\forall y \in \text{dom } f)(\forall \alpha \in ]0, 1[)$$

$$x \neq y \quad \Rightarrow \quad f(\alpha x + (1 - \alpha)y) < \alpha f(x) + (1 - \alpha)f(y).$$

# Convex analysis

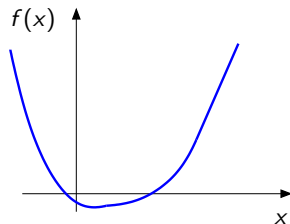
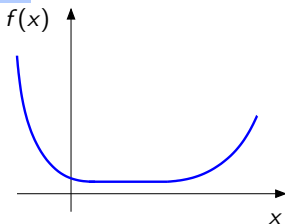
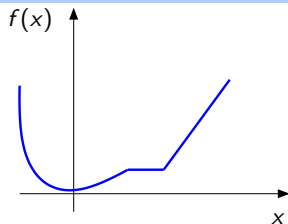
Let  $\mathcal{H}$  be a Hilbert space. Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

$f$  is **strictly convex** if

$$(\forall x \in \text{dom } f)(\forall y \in \text{dom } f)(\forall \alpha \in ]0, 1[)$$

$$x \neq y \Rightarrow f(\alpha x + (1 - \alpha)y) < \alpha f(x) + (1 - \alpha)f(y).$$

Strictly convex functions ?



# Convex analysis

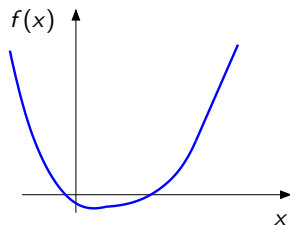
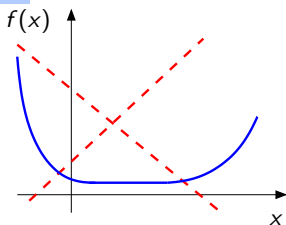
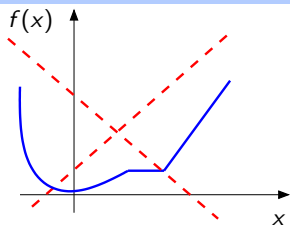
Let  $\mathcal{H}$  be a Hilbert space. Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

$f$  is **strictly convex** if

$$(\forall x \in \text{dom } f)(\forall y \in \text{dom } f)(\forall \alpha \in ]0, 1[)$$

$$x \neq y \Rightarrow f(\alpha x + (1 - \alpha)y) < \alpha f(x) + (1 - \alpha)f(y).$$

Strictly convex functions ?



## Convex analysis

Let  $\mathcal{H}$  be a Hilbert space and  $C$  be a closed convex of  $\mathcal{H}$ .

Let  $f \in \Gamma_0(\mathcal{H})$  such that  $\text{dom } f \cap C \neq \emptyset$ .

If  $f$  is coercive or  $C$  is bounded, then there exists  $p \in C$  such that

$$f(p) = \inf_{x \in C} f(x).$$

Moreover, if  $f$  is strictly convex, this minimizer  $p$  is unique.

## Proximity operator: definition

Let  $\mathcal{H}$  be a Hilbert space. Let  $f \in \Gamma_0(\mathcal{H})$  and  $\gamma \in ]0, +\infty[$ .  
For every  $x \in \mathcal{H}$ , there exists a unique  $p \in \mathcal{H}$  such that

$$f(p) + \frac{1}{2\gamma} \|p - x\|^2 = \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma} \|y - x\|^2.$$

## Proximity operator: definition

Let  $\mathcal{H}$  be a Hilbert space. Let  $f \in \Gamma_0(\mathcal{H})$  and  $\gamma \in ]0, +\infty[$ .  
For every  $x \in \mathcal{H}$ , there exists a unique  $p \in \mathcal{H}$  such that

$$f(p) + \frac{1}{2\gamma} \|p - x\|^2 = \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma} \|y - x\|^2.$$

Let  $\mathcal{H}$  be a Hilbert space. Let  $f \in \Gamma_0(\mathcal{H})$ .

- ▶ The **Moreau envelope** of  $f$  of parameter  $\gamma \in ]0, +\infty[$  is

$$\gamma f: \mathcal{H} \rightarrow \mathbb{R}: x \mapsto \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma} \|y - x\|^2.$$

- ▶ The **proximity operator** of  $f$  is

$$\text{prox}_f: \mathcal{H} \rightarrow \mathcal{H}: x \mapsto \underset{y \in \mathcal{H}}{\text{argmin}} f(y) + \frac{1}{2} \|y - x\|^2.$$

## Proximity operator: definition

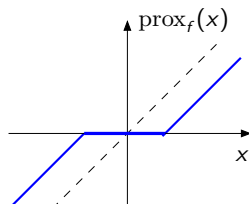
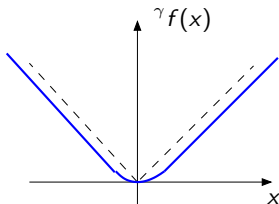
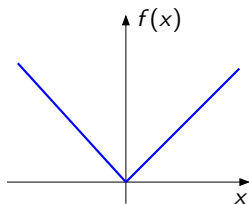
Let  $\mathcal{H}$  be a Hilbert space. Let  $f \in \Gamma_0(\mathcal{H})$ .

- ▶ The **Moreau envelope** of  $f$  of parameter  $\gamma \in ]0, +\infty[$  is

$$\gamma f: \mathcal{H} \rightarrow \mathbb{R}: x \mapsto \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma} \|y - x\|^2.$$

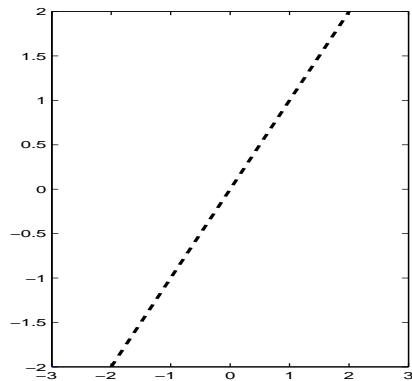
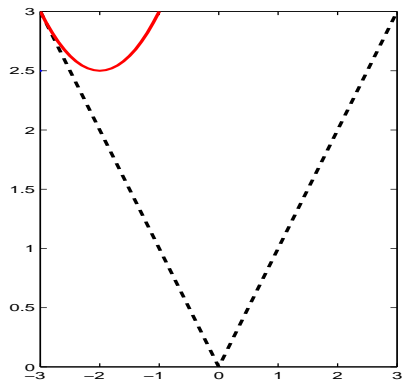
- ▶ The **proximity operator** of  $f$  is

$$\text{prox}_f: \mathcal{H} \rightarrow \mathcal{H}: x \mapsto \underset{y \in \mathcal{H}}{\text{argmin}} f(y) + \frac{1}{2} \|y - x\|^2.$$

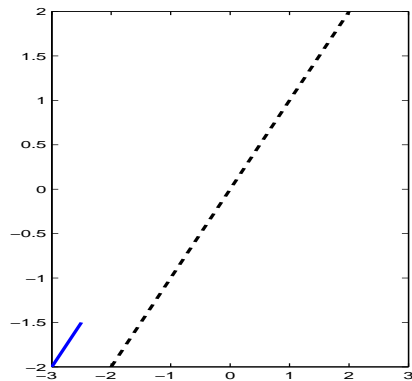
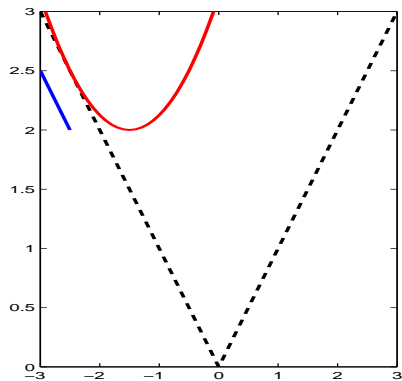




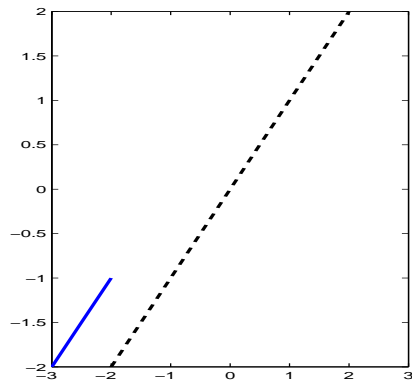
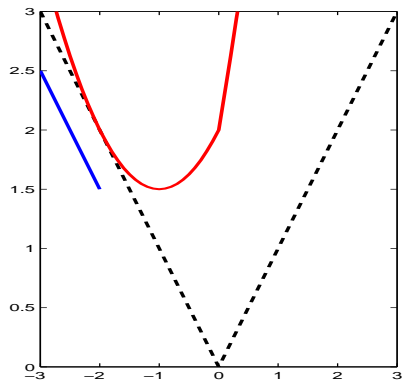
# Proximity operator: definition



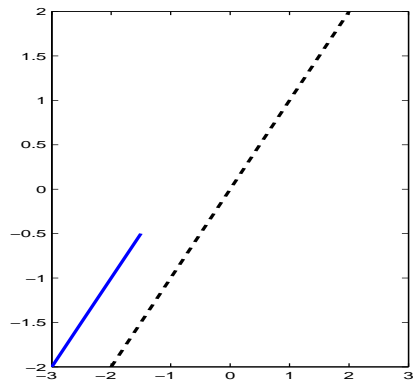
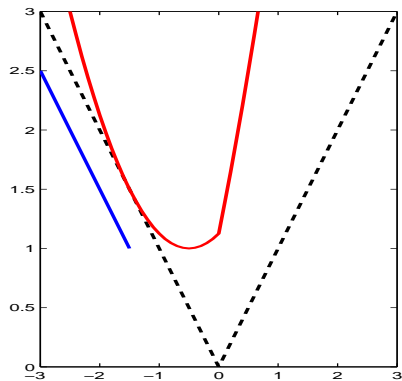
# Proximity operator: definition



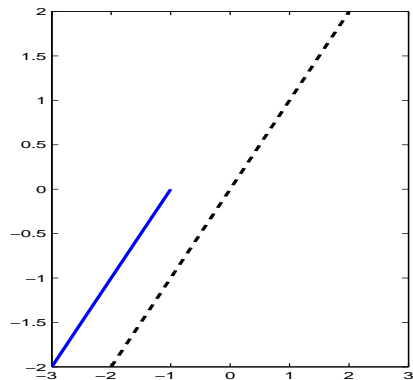
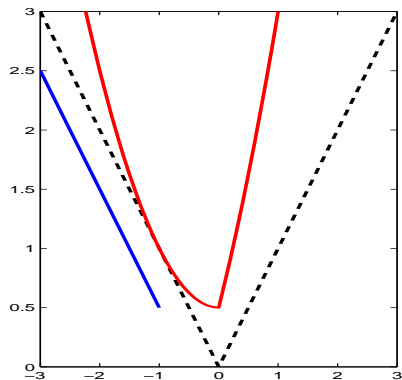
# Proximity operator: definition



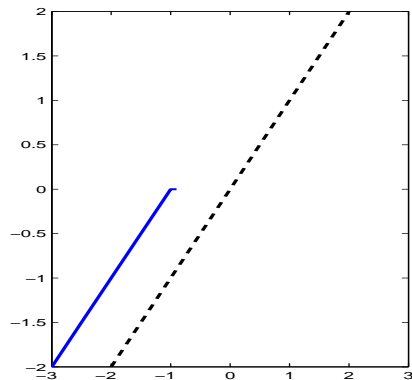
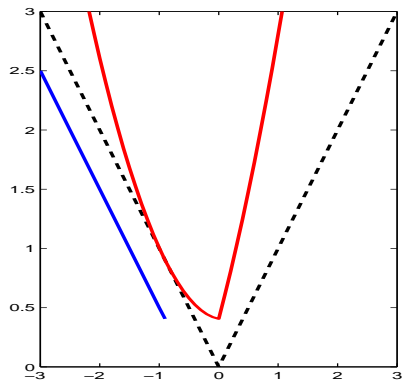
# Proximity operator: definition



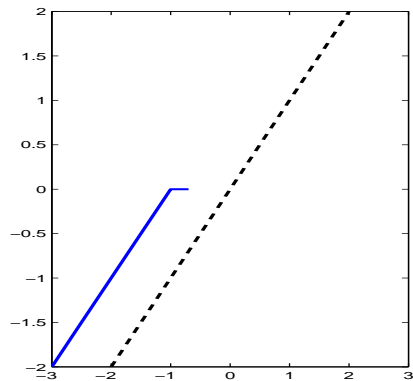
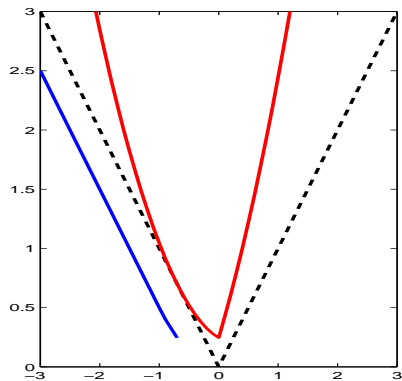
# Proximity operator: definition



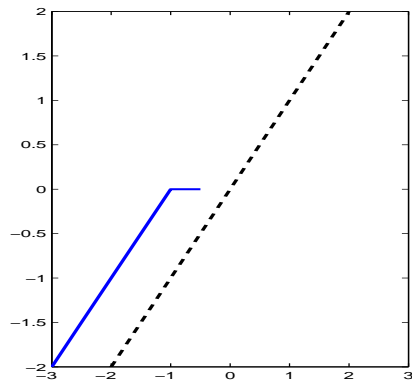
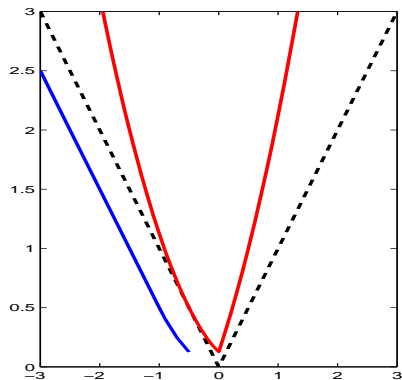
# Proximity operator: definition



# Proximity operator: definition

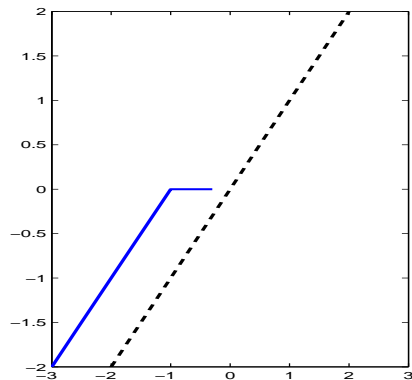
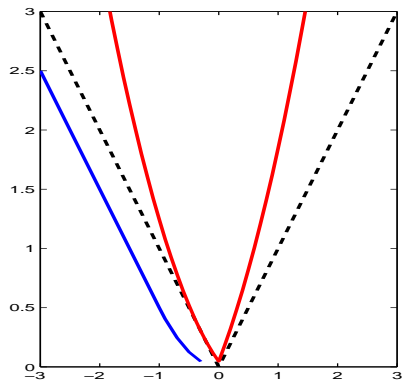


# Proximity operator: definition

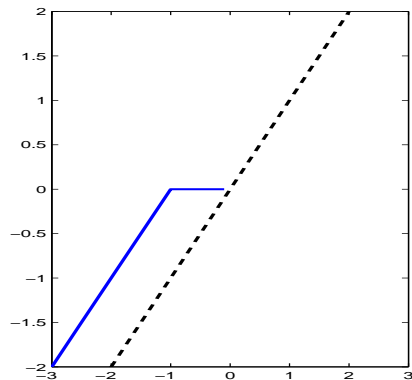
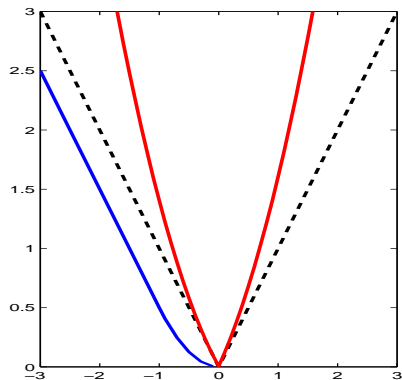




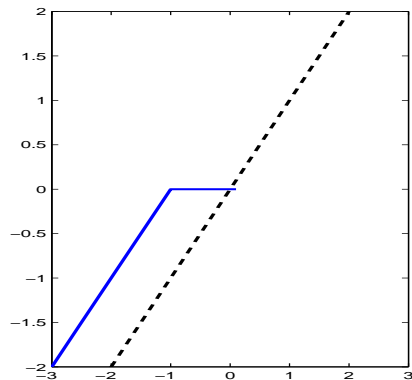
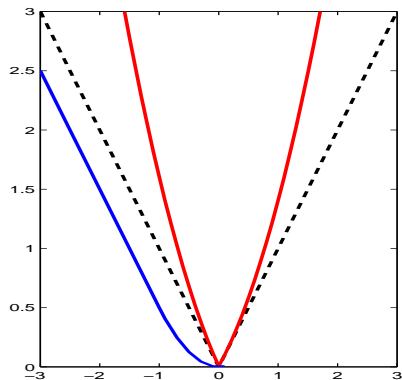
# Proximity operator: definition



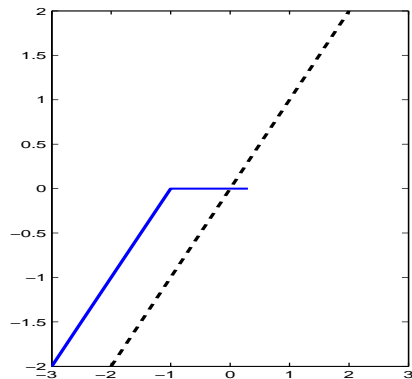
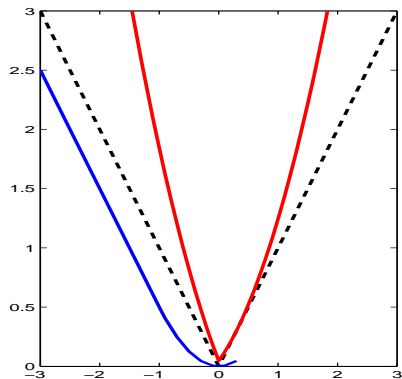
# Proximity operator: definition



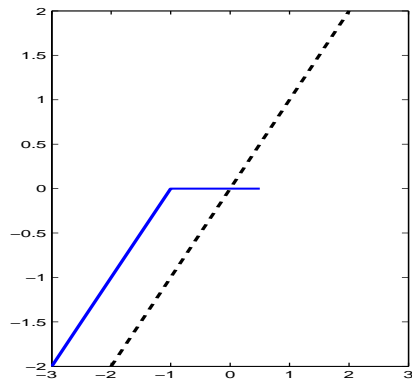
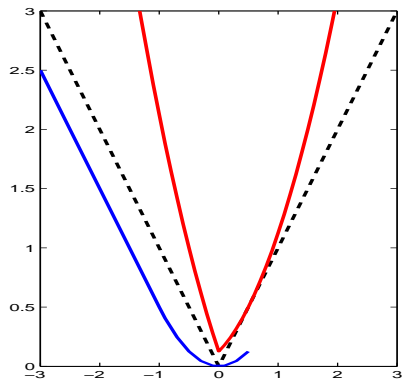
# Proximity operator: definition



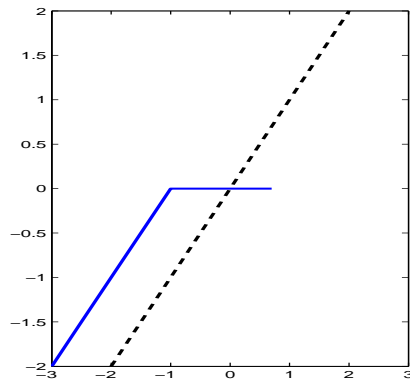
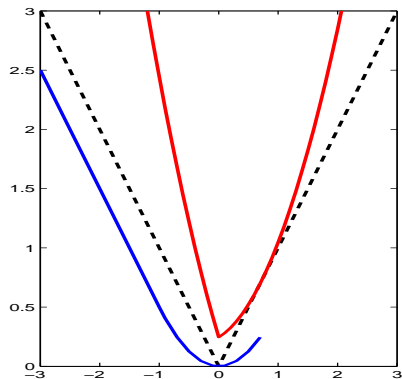
# Proximity operator: definition



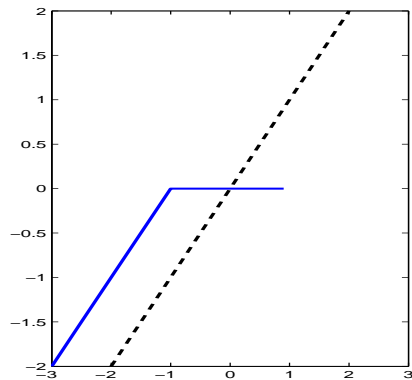
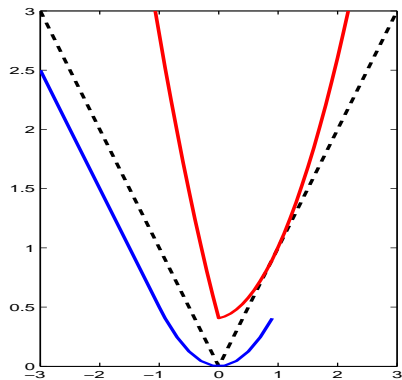
# Proximity operator: definition



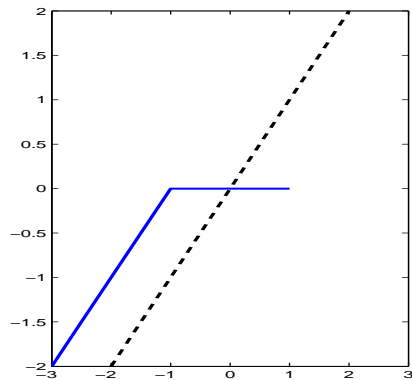
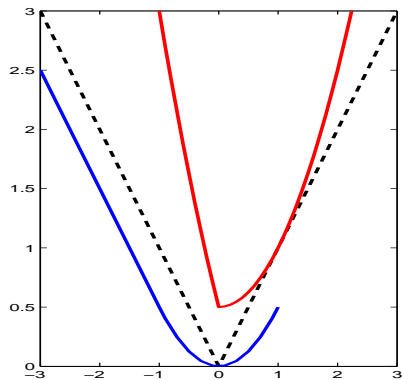
# Proximity operator: definition



# Proximity operator: definition

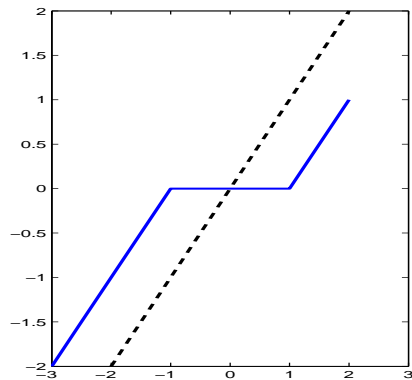
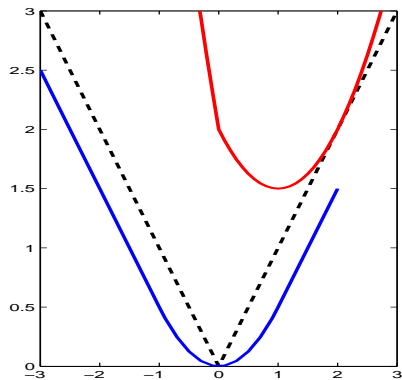


# Proximity operator: definition

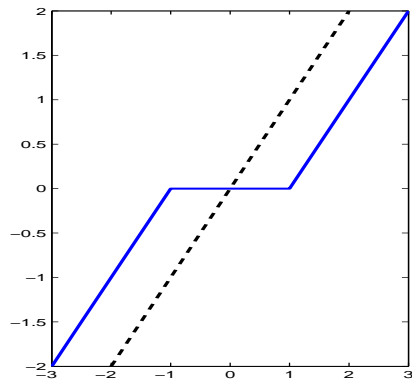
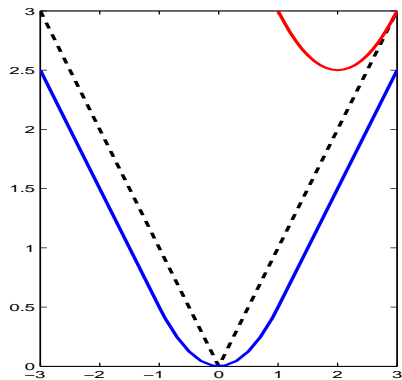




# Proximity operator: definition



# Proximity operator: definition



## Proximity operator: definition

Let  $\mathcal{H}$  be a Hilbert space. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{H})$ .  
If  $\text{dom } f \cap \text{int}(\text{dom } g) \neq \emptyset$  then  $\partial(f + g) = \partial f + \partial g$ .

Let  $\mathcal{H}$  be a Hilbert space and  $f \in \Gamma_0(\mathcal{H})$ .

$$\text{prox}_f = J_{\partial f} .$$

## Proximity operator: definition

Let  $\mathcal{H}$  be a Hilbert space. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{H})$ .  
 If  $\text{dom } f \cap \text{int}(\text{dom } g) \neq \emptyset$  then  $\partial(f + g) = \partial f + \partial g$ .

Let  $\mathcal{H}$  be a Hilbert space and  $f \in \Gamma_0(\mathcal{H})$ .

$$\text{prox}_f = J_{\partial f}.$$

Proof: By using Fermat's rule, for every  $x \in \mathcal{H}$ ,

$$\begin{aligned} p = \arg \min f + (2\gamma)^{-1} \|\cdot - x\|^2 &\Leftrightarrow 0 \in \partial\left(f + \frac{1}{2} \|\cdot - x\|^2\right)(p) \\ &\Leftrightarrow 0 \in \partial f(p) + p - x \\ &\Leftrightarrow x \in (\text{Id} + \partial f)(p) \\ &\Leftrightarrow p = (\text{Id} + \partial f)^{-1}(x). \end{aligned}$$

## Proximity operator: definition

Let  $\mathcal{H}$  be a Hilbert space. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{H})$ .  
If  $\text{dom } f \cap \text{int}(\text{dom } g) \neq \emptyset$  then  $\partial(f + g) = \partial f + \partial g$ .

Let  $\mathcal{H}$  be a Hilbert space and  $f \in \Gamma_0(\mathcal{H})$ .

$$\text{prox}_f = J_{\partial f}.$$

Remark: As  $\text{dom}(\text{prox}_f) = \mathcal{H}$ , this provides a proof that  $\partial f$  is maximally monotone !

## Proximity operator: properties

Let  $\mathcal{H}$  be a Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $(x, p) \in \mathcal{H}^2$ .

$$p = \text{prox}_{\gamma f} x \quad \Leftrightarrow \quad (\forall y \in \mathcal{H}) \quad \langle y - p \mid x - p \rangle + f(p) \leq f(y).$$

## Proximity operator: properties

Let  $\mathcal{H}$  be a Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $(x, p) \in \mathcal{H}^2$ .

$$p = \text{prox}_{\gamma f} x \quad \Leftrightarrow \quad (\forall y \in \mathcal{H}) \quad \langle y - p \mid x - p \rangle + f(p) \leq f(y).$$

Let  $\mathcal{H}$  be a Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $\gamma \in ]0, +\infty[$ .

$\gamma f$  is differentiable and  $\nabla \gamma f$  is  $\gamma^{-1}$ -Lipschitzian

$$(\forall x \in \mathcal{H}) \quad \nabla \underbrace{\gamma f}_{\text{Moreau envelope}} = \gamma^{-1}(\text{Id} - \text{prox}_{\gamma f}) = \underbrace{\gamma \partial f}_{\text{Yosida approximation}}.$$

## Proximity operator: properties

Let  $\mathcal{H}$  be a Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $(x, p) \in \mathcal{H}^2$ .

$$p = \text{prox}_{\gamma f} x \quad \Leftrightarrow \quad (\forall y \in \mathcal{H}) \quad \langle y - p \mid x - p \rangle + f(p) \leq f(y).$$

Let  $\mathcal{H}$  be a Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $\gamma \in ]0, +\infty[$ .

$\gamma f$  is differentiable and  $\nabla \gamma f$  is  $\gamma^{-1}$ -Lipschitzian

$$(\forall x \in \mathcal{H}) \quad \nabla \underbrace{\gamma f}_{\text{Moreau envelope}} = \gamma^{-1}(\text{Id} - \text{prox}_{\gamma f}) = \underbrace{\gamma \partial f}_{\text{Yosida approximation}}.$$

Proof: Previous property + ... calculations.



## Proximity operator: properties

Let  $\mathcal{H}$  be a Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $(x, p) \in \mathcal{H}^2$ .

$$p = \text{prox}_{\gamma f} x \quad \Leftrightarrow \quad (\forall y \in \mathcal{H}) \quad \langle y - p \mid x - p \rangle + f(p) \leq f(y).$$

Let  $\mathcal{H}$  be a Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $\gamma \in ]0, +\infty[$ .

$\gamma f$  is differentiable and  $\nabla \gamma f$  is  $\gamma^{-1}$ -Lipschitzian

$$(\forall x \in \mathcal{H}) \quad \underbrace{\nabla \gamma f}_{\text{Moreau envelope}} = \gamma^{-1}(\text{Id} - \text{prox}_{\gamma f}) = \underbrace{\gamma \partial f}_{\text{Yosida approximation}}.$$

Interpretation:  $\gamma f$  is a smooth approximation of  $f$ .

## Proximity operator: properties

Let  $\mathcal{H}$  be a Hilbert space,  $x \in \mathcal{H}$  and  $f \in \Gamma_0(\mathcal{H})$ .

Properties	$g(x)$	$\text{prox}_{g,x}$
Translation	$f(x - z), z \in \mathcal{H}$	$z + \text{prox}_f(x - z)$
Quadratic perturbation	$f(x) + \alpha \ x\ ^2 / 2 + \langle z   x \rangle + \gamma$ $z \in \mathcal{H}, \alpha > 0, \gamma \in \mathbb{R}$	$\text{prox}_{\frac{f}{\alpha+1}}\left(\frac{x-z}{\alpha+1}\right)$
Scale change	$f(\rho x), \rho \in \mathbb{R}^*$	$\frac{1}{\rho} \text{prox}_{\rho^2 f}(\rho x)$
Reflection	$f(-x)$	$-\text{prox}_f(-x)$
Moreau envelope	$\gamma f(x) = \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma} \ x - y\ ^2$ $\gamma > 0$	$\frac{1}{1+\gamma} \left( \gamma x + \text{prox}_{(1+\gamma)f}(x) \right)$

## Proximity operator: properties

For every  $i \in \{1, \dots, n\}$ , let  $\mathcal{H}_i$  be a Hilbert space and  $f_i \in \Gamma_0(\mathcal{H}_i)$ .

For all  $(x_1, \dots, x_n) \in \mathcal{H}_1 \times \dots \times \mathcal{H}_n$ ,

if

$$f(x_1, \dots, x_n) = \sum_{i=1}^n f_i(x_i).$$

then

$$\text{prox}_f(x_1, \dots, x_n) = (\text{prox}_{f_i}(x_i))_{1 \leq i \leq n}.$$

## Proximity operator: properties

Let  $\mathcal{H}$  be a separable Hilbert space.

Let  $(b_i)_{i \in I}$  be an orthonormal basis of  $\mathcal{H}$ .

For every  $i \in I$ , let  $\varphi_i \in \Gamma_0(\mathbb{R})$  such that  $\varphi_i \geq 0$ . For every  $x \in \mathcal{H}$ , if

$$f(x) = \sum_{i \in I} \varphi_i(\langle x | b_i \rangle)$$

then

$$\text{prox}_f(x) = \sum_{i \in I} \text{prox}_{\varphi_i}(\langle x | b_i \rangle) b_i.$$

Remark: The assumption  $(\forall i \in I) \varphi_i \geq 0$  can be relaxed if  $\mathcal{H}$  is finite dimensional.

## Proximity operator: properties

Let  $\mathcal{H}$  be a separable Hilbert space.

Let  $(b_i)_{i \in I}$  be an orthonormal basis of  $\mathcal{H}$ .

For every  $i \in I$ , let  $\varphi_i \in \Gamma_0(\mathbb{R})$  such that  $\varphi_i \geq 0$ . For every  $x \in \mathcal{H}$ , if

$$f(x) = \sum_{i \in I} \varphi_i(\langle x | b_i \rangle)$$

then

$$\text{prox}_f(x) = \sum_{i \in I} \text{prox}_{\varphi_i}(\langle x | b_i \rangle) b_i.$$

Example:  $\mathcal{H} = \mathbb{R}^N$ ,  $(b_i)_{1 \leq i \leq N}$  canonical basis of  $\mathbb{R}^N$ ,  $f = \lambda \|\cdot\|_1$  with  $\lambda \in [0, +\infty[$ .

$$(\forall x = (x^{(i)})_{1 \leq i \leq N}) \in \mathbb{R}^N \quad \text{prox}_{\lambda \|\cdot\|_1}(x) = (\text{prox}_{\lambda|\cdot|}(x^{(i)}))_{1 \leq i \leq N}$$

## Proximity operator: properties

### Moreau decomposition formula

Let  $\mathcal{H}$  be a Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $\gamma \in ]0, +\infty[$ .

$$(\forall x \in \mathcal{H}) \quad \text{prox}_{\gamma f^*} x = x - \gamma \text{prox}_{\gamma^{-1} f}(\gamma^{-1} x).$$

## Proximity operator: properties

### Moreau decomposition formula

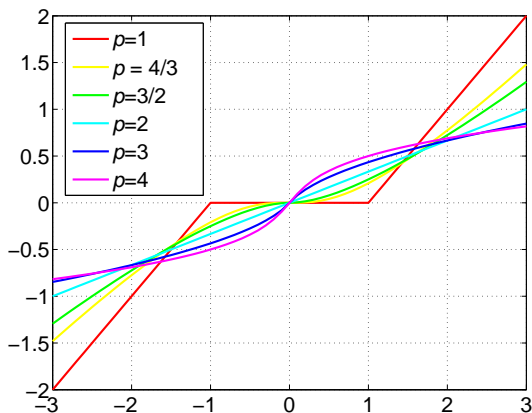
Let  $\mathcal{H}$  be a Hilbert space,  $f \in \Gamma_0(\mathcal{H})$  and  $\gamma \in ]0, +\infty[$ .

$$(\forall x \in \mathcal{H}) \quad \text{prox}_{\gamma f^*} x = x - \gamma \text{prox}_{\gamma^{-1} f}(\gamma^{-1} x).$$

Example: If  $\mathcal{H} = \mathbb{R}^N$ ,  $f = \frac{1}{q} \|\cdot\|_q^q$  with  $q \in ]1, +\infty[$ , then  $f^* = \frac{1}{q^*} \|\cdot\|_{q^*}^{q^*}$  with  $1/q + 1/q^* = 1$ , and

$$(\forall x \in \mathbb{R}^N) \quad \text{prox}_{\frac{\gamma}{q^*} \|\cdot\|_{q^*}^{q^*}} x = x - \gamma \text{prox}_{\frac{1}{\gamma q} \|\cdot\|_q^q}(\gamma^{-1} x).$$

# Proximity operator: properties





## Proximity operator: properties

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  such that  $\text{ran } L = \mathcal{H}$ . Then

$$\partial(f \circ L) = L^* \partial f L.$$

## Proximity operator: properties

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  such that  $\text{ran } L = \mathcal{H}$ . Then

$$\partial(f \circ L) = L^* \partial f L.$$

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  such that  $LL^* = \mu \text{Id}$  where  $\mu \in ]0, +\infty[$ . Then

$$\text{prox}_{f \circ L} = \text{Id} - \mu^{-1} L^* \circ (\text{Id} - \text{prox}_{\mu f}) \circ L.$$

## Proximity operator: properties

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  such that  $\text{ran } L = \mathcal{H}$ . Then

$$\partial(f \circ L) = L^* \partial f L.$$

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  such that  $LL^* = \mu \text{Id}$  where  $\mu \in ]0, +\infty[$ . Then

$$\text{prox}_{f \circ L} = \text{Id} - \mu^{-1} L^* \circ (\text{Id} - \text{prox}_{\mu f}) \circ L.$$

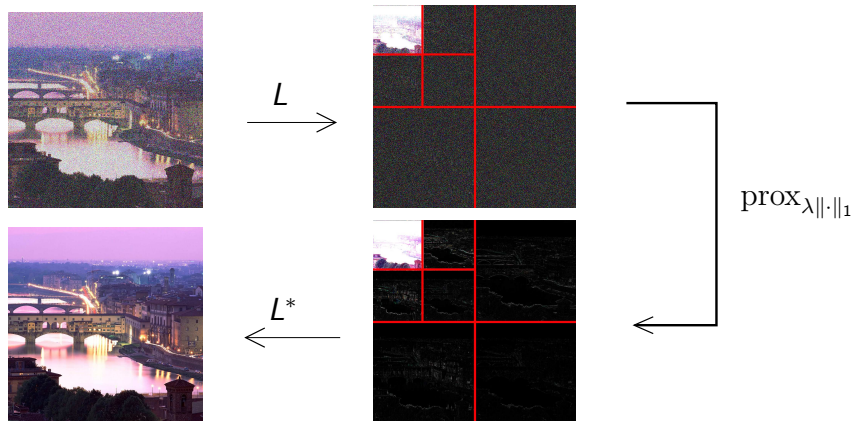
Remark :

Useful property for data fidelity terms involving a neg-log-likelihood  $f$  and a synthesis tight frame operator  $L$ .

## Proximity operator: properties

Particular case :  $L \in \mathcal{B}(\mathcal{H}, \mathcal{H})$  unitary,  $\text{prox}_{f \circ L} = L^* \text{prox}_f L$ .

- ▶ Illustration: denoising using an  $\ell_1$  penalty on the coefficients resulting from an orthogonal wavelet transform  $L$ .

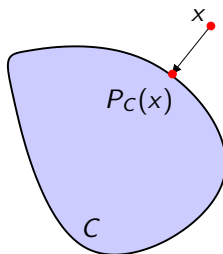


## Proximity operator: examples

### Projection :

Let  $\mathcal{H}$  be a Hilbert space. Let  $C$  be a nonempty closed convex subset of  $\mathcal{H}$ .

$$(\forall x \in \mathcal{H}) \quad \text{prox}_{\iota_C}(x) = \operatorname{argmin}_{y \in C} \frac{1}{2} \|y - x\|^2 = P_C(x).$$



## Proximity operator: examples

### Projection :

Let  $\mathcal{H}$  be a Hilbert space. Let  $C$  be a nonempty closed convex subset of  $\mathcal{H}$ .

$$(\forall x \in \mathcal{H}) \quad \text{prox}_{\iota_C}(x) = \underset{y \in C}{\text{argmin}} \frac{1}{2} \|y - x\|^2 = P_C(x).$$

### Remark :

- ▶  $p = P_C(x) \Leftrightarrow x - p \in \partial \iota_C(p) = N_C(p)$   
 $\Leftrightarrow (\forall y \in C) \langle y - p \mid x - p \rangle \leq 0$ .

Particular case: if  $C$  is a vector space:  $p = P_C(x) \Leftrightarrow x - p \in C^\perp$ .

- ▶  $\gamma_{\iota_C} = (2\gamma)^{-1} d_C^2$  where  $d_C$  distance to the convex set  $C$  is defined by  $d_C: x \mapsto \inf_{y \in C} \|y - x\| = \|x - P_C x\|$ . We have then  $\nabla d_C^2 = \nabla(\frac{1}{2} \iota_C) = 2(\text{Id} - P_C)$ .

## Proximity operator: examples

Quadratic function :

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ ,  $\gamma \in ]0, +\infty[$  and  $z \in \mathcal{G}$ .

$$f = \gamma \|L \cdot - z\|^2 / 2 \quad \Rightarrow \quad \text{prox}_f = (\text{Id} + \gamma L^* L)^{-1}(\cdot + \gamma L^* z).$$

## Proximity operator: examples

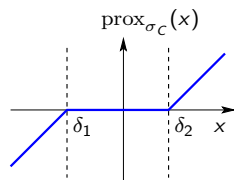
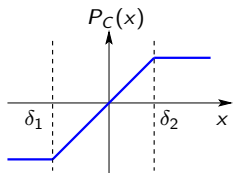
**Support function** :

Let  $\mathcal{H}$  be a Hilbert space and  $C \subset \mathcal{H}$  be nonempty closed convex.

$$(\forall x \in \mathcal{H}) \quad \text{prox}_{\sigma_C} = \text{Id} - P_C.$$

Soft-thresholding :  $\mathcal{H} = \mathbb{R}$ ,  $\delta_1 = \inf C$  and  $\delta_2 = \sup C$ . For every  $x \in \mathbb{R}$ ,

$$\sigma_C(x) = \begin{cases} \delta_1 x & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ \delta_2 x & \text{if } x > 0 \end{cases} \Rightarrow \text{prox}_{\sigma_C}(x) = \text{soft}_C(x) = \begin{cases} x - \delta_1 & \text{if } x < \delta_1 \\ 0 & \text{if } x \in C \\ x - \delta_2 & \text{if } x > \delta_2. \end{cases}$$





## Part 3: Search for a zero

1. Zeros of a (maximally) monotone operator
2. Fixed points
3. Convergence
  - ▶ Definition
  - ▶ Fejér monotonicity
  - ▶ Demiclosedness principle
4. Algorithms
  - ▶ Krasnosel'skii Mann
  - ▶ Douglas-Rachford
  - ▶ PPXA
  - ▶ *Forward-Backward*
  - ▶ *Forward-Backward-Forward*

## Monotone operator: zeros

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a monotone operator.

The set of zeros of  $A$ , denoted  $\text{zer } A$ , is

$$\text{zer } A = \{x \in \mathcal{H} \mid 0 \in Ax\}.$$

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator.

$\text{zer } A$  is closed and convex.

## Monotone operator: zeros

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a monotone operator.

The **set of zeros** of  $A$ , denoted  $\text{zer } A$ , is

$$\text{zer } A = \{x \in \mathcal{H} \mid 0 \in Ax\}.$$

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator.

If one of the following assumptions is satisfied

- ▶  $A$  is surjective
- ▶  $\text{dom } A$  is bounded,

then  $\text{zer } A \neq \emptyset$ .

## Monotone operator: zeros

Let  $\mathcal{H}$  be a Hilbert space.

$A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  is **strictly monotone** if

$$(\forall (x_1, u_1) \in \text{gra}A) (\forall (x_2, u_2) \in \text{gra}A) \quad x_1 \neq x_2 \Rightarrow \langle u_1 - u_2 \mid x_1 - x_2 \rangle > 0.$$

## Monotone operator: zeros

Let  $\mathcal{H}$  be a Hilbert space.

$A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  is strictly monotone if

$$(\forall (x_1, u_1) \in \text{gra}A) (\forall (x_2, u_2) \in \text{gra}A) \quad x_1 \neq x_2 \Rightarrow \langle u_1 - u_2 \mid x_1 - x_2 \rangle > 0.$$

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a strictly monotone operator.

$\text{zer} A$  is at most a singleton.

Proof:

If  $(x_1, x_2) \in (\text{zer} A)^2$  and  $x_1 \neq x_2$  then  $0 = \langle x_1 - x_2 \mid 0 - 0 \rangle \leq 0!$

## Monotone operator: zeros

Let  $\mathcal{H}$  be a Hilbert space and  $\beta \in ]0, +\infty[$ .

$A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  is  $\beta$ -strongly monotone if

$$(\forall (x_1, u_1) \in \text{gra}A) (\forall (x_2, u_2) \in \text{gra}A) \quad \langle u_1 - u_2 \mid x_1 - x_2 \rangle \geq \beta \|x_1 - x_2\|^2.$$

## Monotone operator: zeros

Let  $\mathcal{H}$  be a Hilbert space and  $\beta \in ]0, +\infty[$ .

$A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  is  $\beta$ -strongly monotone if

$$(\forall (x_1, u_1) \in \text{gra}A) (\forall (x_2, u_2) \in \text{gra}A) \quad \langle u_1 - u_2 \mid x_1 - x_2 \rangle \geq \beta \|x_1 - x_2\|^2.$$

Let  $\mathcal{H}$  be a Hilbert space and  $\beta \in ]0, +\infty[$ .

$A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  is  $\beta$ -strongly monotone  $\Leftrightarrow A^{-1}: \mathcal{H} \rightarrow \mathcal{H}$  is  $\beta$ -cocoercive.

Let  $\mathcal{H}$  be a Hilbert space and  $\beta \in ]0, +\infty[$ .

$A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  is  $\beta$ -strongly monotone  $\Rightarrow A$  is strictly monotone.

## Monotone operator: zeros

Let  $\mathcal{H}$  be a Hilbert space and  $\beta \in ]0, +\infty[$ .

$A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  is  $\beta$ -strongly monotone if

$$(\forall (x_1, u_1) \in \text{gra}A) (\forall (x_2, u_2) \in \text{gra}A) \quad \langle u_1 - u_2 \mid x_1 - x_2 \rangle \geq \beta \|x_1 - x_2\|^2.$$

Let  $\mathcal{H}$  be a Hilbert space and  $\beta \in ]0, +\infty[$ .

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator.

$A$   $\beta$ -strongly monotone  $\Rightarrow$   $\text{zer } A$  is a singleton.



## Monotone operator: zeros and minimizer

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

$f$  is strictly convex  $\Rightarrow \partial f$  is strictly monotone.

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

$f$  is strictly convex  $\Rightarrow f$  has at most one minimizer.

## Monotone operator: zeros and minimizer

Let  $\mathcal{H}$  be a Hilbert space and  $\beta \in ]0, +\infty[$ .

Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$  be a proper function.

$f$  is  $\beta$ -strongly convex  $\Rightarrow \partial f$  is  $\beta$ -strongly monotone.

Let  $\mathcal{H}$  be a Hilbert space and  $\beta \in ]0, +\infty[$ .

Let  $f \in \Gamma_0(\mathcal{H})$ .

$f$  is  $\beta$ -strongly convex  $\Rightarrow f$  has a unique minimizer.

## Fixed point algorithms: zeros and fixed points

Let  $\mathcal{H}$  be a Hilbert space. Let  $B : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

The set of fixed points of  $B$  is denoted by

$$\text{Fix}B = \{x \in \mathcal{H} \mid x \in Bx\}$$

## Fixed point algorithms: zeros and fixed points

Let  $\mathcal{H}$  be a Hilbert space. Let  $B : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

The set of fixed points of  $B$  is denoted by

$$\text{Fix}B = \{x \in \mathcal{H} \mid x \in Bx\}$$

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a monotone operator and let  $\gamma \in ]0, +\infty[$ .

Then

$$\text{Fix}J_{\gamma A} = \text{zer} A$$

## Fixed point algorithms: zeros and fixed points

Let  $\mathcal{H}$  be a Hilbert space. Let  $B : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

The set of fixed points of  $B$  is denoted by

$$\text{Fix}B = \{x \in \mathcal{H} \mid x \in Bx\}$$

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a monotone operator and let  $\gamma \in ]0, +\infty[$ .

Then

$$\text{Fix}J_{\gamma}A = \text{zer} A$$

$$\begin{aligned}
 \text{Proof : } (\forall x \in \mathcal{H}) \quad 0 \in Ax &\Leftrightarrow \dots \\
 &\Leftrightarrow \dots \\
 &\Leftrightarrow \dots \\
 &\Leftrightarrow x = J_{\gamma}Ax
 \end{aligned}$$

## Fixed point algorithms: zeros and fixed points

Let  $\mathcal{H}$  be a Hilbert space. Let  $B : \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

The set of fixed points of  $B$  is denoted by

$$\text{Fix}B = \{x \in \mathcal{H} \mid x \in Bx\}$$

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a monotone operator and let  $\gamma \in ]0, +\infty[$ .

Then

$$\text{Fix}J_{\gamma A} = \text{zer} A$$

$$\begin{aligned}
 \text{Proof : } (\forall x \in \mathcal{H}) \quad 0 \in Ax &\Leftrightarrow 0 \in \gamma Ax \\
 &\Leftrightarrow x \in (\text{Id} + \gamma A)x \\
 &\Leftrightarrow x \in (\text{Id} + \gamma A)^{-1}x \\
 &\Leftrightarrow x = J_{\gamma A}x
 \end{aligned}$$

## Fixed point algorithm: convergence

Let  $\mathcal{H}$  be a Hilbert space.

Let  $(x_n)_{n \in \mathbb{N}}$  be a sequence in  $\mathcal{H}$  and  $\hat{x} \in \mathcal{H}$ .

- ▶  $(x_n)_{n \in \mathbb{N}}$  converges strongly to  $\hat{x}$  if

$$\lim_{n \rightarrow +\infty} \|x_n - \hat{x}\| = 0.$$

It is denoted by  $x_n \rightarrow \hat{x}$ .

- ▶  $(x_n)_{n \in \mathbb{N}}$  converges weakly to  $\hat{x}$  if

$$(\forall y \in \mathcal{H}) \quad \lim_{n \rightarrow +\infty} \langle y | x_n - \hat{x} \rangle = 0.$$

It is denoted by  $x_n \rightharpoonup \hat{x}$ .

Remark: In a finite dimensional Hilbert space, strong and weak convergences are equivalent.

## Fixed point algorithm: convergence

Let  $\mathcal{H}$  be a Hilbert space.

Let  $(x_n)_{n \in \mathbb{N}}$  be a sequence of  $\mathcal{H}$ .

$(x_n)_{n \in \mathbb{N}}$  converges weakly if and only if

- ▶  $(x_n)_{n \in \mathbb{N}}$  is bounded  
and
  - ▶  $(x_n)_{n \in \mathbb{N}}$  possesses at most one sequential cluster point in the weak topology.
- 
- ▶  $\hat{x}$  is a sequential cluster point of  $(x_n)_{n \in \mathbb{N}}$  in the weak topology if there exists a sub-sequence  $(x_{n_k})_{k \in \mathbb{N}}$  of  $(x_n)_{n \in \mathbb{N}}$  that converges weakly to  $\hat{x}$ .



## Fixed point algorithm: convergence

Let  $\mathcal{H}$  be a Hilbert space.

Let  $(x_n)_{n \in \mathbb{N}}$  be a sequence of  $\mathcal{H}$ .

$(x_n)_{n \in \mathbb{N}}$  converges weakly if and only if

▶  $(x_n)_{n \in \mathbb{N}}$  is bounded

and

▶  $(x_n)_{n \in \mathbb{N}}$  possesses at most one sequential cluster point in the weak topology.

Illustration:

$x_0$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$\dots$
1	-1	1	-1	1	-1	$\dots$

→  $(x_n)_{n \in \mathbb{N}}$  is bounded but it has 2 sequential cluster points:  $-1$  and  $1$ .

→  $(x_n)_{n \in \mathbb{N}}$  does not converge.

## Fixed point algorithm: Fejér-monotone sequence

Let  $\mathcal{H}$  be a Hilbert space and  $D$  be a nonempty subset of  $\mathcal{H}$ .

Let  $(x_n)_{n \in \mathbb{N}}$  be a sequence in  $\mathcal{H}$ .

$(x_n)_{n \in \mathbb{N}}$  is **Fejér-monotone** with respect to  $D$  if

$$(\forall x \in D)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - x\| \leq \|x_n - x\|.$$

Let  $\mathcal{H}$  be a Hilbert space and  $D$  be a nonempty subset of  $\mathcal{H}$ .

Let  $(x_n)_{n \in \mathbb{N}}$  be Fejér-monotone with respect to  $D$  then

- ▶  $(x_n)_{n \in \mathbb{N}}$  is bounded .
- ▶ for every  $x \in D$ ,  $(\|x_n - x\|)_{n \in \mathbb{N}}$  converges.

## Fixed point algorithm: Fejér-monotone sequence

### Fejér-monotone convergence

Let  $\mathcal{H}$  be a Hilbert space and let  $D$  be a nonempty subset of  $\mathcal{H}$ .

Let  $(x_n)_{n \in \mathbb{N}}$  be a sequence in  $\mathcal{H}$ .

$(x_n)_{n \in \mathbb{N}}$  converges weakly to a point in  $D$  if

- ▶  $(x_n)_{n \in \mathbb{N}}$  is Fejér-monotone with respect to  $D$   
and
- ▶ every sequential cluster point of  $(x_n)_{n \in \mathbb{N}}$  in the weak topology lies in  $D$ .

## Fixed point algorithm: Fejér-monotone sequence

### Demiclosedness principle

Let  $\mathcal{H}$  be a Hilbert space and  $C$  be a nonempty closed convex subset of  $\mathcal{H}$ .  
Let  $T: C \rightarrow \mathcal{H}$  be a nonexpansive operator.

If  $(x_n)_{n \in \mathbb{N}}$  is a sequence in  $C$  that converges weakly to  $\hat{x}$  and if  $\|Tx_n - x_n\| \rightarrow 0$  then

$$\hat{x} \in \text{Fix } T.$$

## Fixed point algorithm: Fejér-monotone sequences

Let  $\mathcal{H}$  be a Hilbert space and  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $T: C \rightarrow C$  be a nonexpansive operator

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = Tx_n.$$

## Fixed point algorithm: Fejér-monotone sequences

Let  $\mathcal{H}$  be a Hilbert space and  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $T: C \rightarrow C$  be a nonexpansive operator

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = Tx_n.$$

If  $x_n - Tx_n \rightarrow 0$ ,

## Fixed point algorithm: Fejér-monotone sequences

Let  $\mathcal{H}$  be a Hilbert space and  $C$  be a nonempty subset of  $\mathcal{H}$ .

Let  $T: C \rightarrow C$  be a **nonexpansive operator** such that  $\text{Fix } T \neq \emptyset$ .

Let  $x_0 \in C$ ,

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = Tx_n.$$

If  **$x_n - Tx_n \rightarrow 0$** , then  $(x_n)_{n \in \mathbb{N}}$  converges weakly to a point in  $\text{Fix } T$ .

Proof :  $(x_n)_{n \in \mathbb{N}}$  Fejér-monotone with respect to  $\text{Fix } T$  +  
demiclosedness principle.

## Fixed point algorithm: proximal point algorithm

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J_A x_n,$$



## Fixed point algorithm: proximal point algorithm

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator

Let  $\gamma \in ]0, +\infty[$

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J_{\gamma A} x_n,$$

## Fixed point algorithm: proximal point algorithm

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator such that  $\text{zer } A \neq \emptyset$ .

Let  $\gamma \in ]0, +\infty[$  et  $x_0 \in \mathcal{H}$ .

If

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J_{\gamma A} x_n,$$

then  $(x_n)_{n \in \mathbb{N}}$  converges weakly to a point in  $\text{zer } A$ .

Proof:  $J_{\gamma A} : \mathcal{H} \rightarrow \mathcal{H}$  is nonexpansive and  $x_n - J_{\gamma A} x_n \rightarrow 0$ .

## Fixed point algorithm: proximal point algorithm

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A : \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator such that  $\text{zer } A \neq \emptyset$ .

Let  $(\gamma_n)_{n \in \mathbb{N}}$  a sequence in  $]0, +\infty[$  such that  $\sum_{n=0}^{+\infty} \gamma_n^2 = +\infty$  and  $x_0 \in \mathcal{H}$ .

If

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = J_{\gamma_n A} x_n,$$

then  $(x_n)_{n \in \mathbb{N}}$  converges weakly to a point in  $\text{zer } A$ .

Proof: ... more challenging.

## Optimization method: proximal point algorithm

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f \in \Gamma_0(\mathcal{H})$  such that  $\text{Argmin} f \neq \emptyset$ .

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \text{prox}_{\gamma_n f} x_n$$

## Optimization method: proximal point algorithm

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f \in \Gamma_0(\mathcal{H})$  such that  $\text{Argmin} f \neq \emptyset$ .

Let  $(\gamma_n)_{n \in \mathbb{N}}$  be a sequence in  $]0, +\infty[$  such that  $\sum_{n=0}^{+\infty} \gamma_n^2 = +\infty$  and  $x_0 \in \mathcal{H}$ .

If

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \text{prox}_{\gamma_n f} x_n$$

then  $(x_n)_{n \in \mathbb{N}}$  converges weakly to a (global) minimizer of  $f$ .

## Fixed point algorithm: Fejér-monotone sequence

### Krasnosel'skii-Mann algorithm

Let  $\mathcal{H}$  be a Hilbert space and  $C$  be a nonempty closed convex subset of  $\mathcal{H}$ .

Let  $T: C \rightarrow C$  be a nonexpansive operator

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n).$$

## Fixed point algorithm: Fejér-monotone sequence

### Krasnosel'skii-Mann algorithm

Let  $\mathcal{H}$  be a Hilbert space and  $C$  be a nonempty closed convex subset of  $\mathcal{H}$ .

Let  $T: C \rightarrow C$  be a nonexpansive operator

Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $[0, 1]$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(1 - \lambda_n) = +\infty$ .

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n).$$

## Fixed point algorithm: Fejér-monotone sequence

### Krasnosel'skii-Mann algorithm

Let  $\mathcal{H}$  be a Hilbert space and  $C$  be a nonempty closed convex subset of  $\mathcal{H}$ .  
Let  $T: C \rightarrow C$  be a nonexpansive operator such that  $\text{Fix} T \neq \emptyset$ .

Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $[0, 1]$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(1 - \lambda_n) = +\infty$ .

Let  $x_0 \in C$  and

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n).$$

Then,  $(x_n)_{n \in \mathbb{N}}$  converges weakly to a point in  $\text{Fix} T$

### Proof:

- ▶  $(x_n)_{n \in \mathbb{N}}$  is Fejér-monotone with respect to  $\text{Fix} T$ .
- ▶  $(Tx_n - x_n)_{n \in \mathbb{N}}$  converges strongly to 0.



## Fixed point algorithm: Douglas-Rachford

Let  $\mathcal{H}$  be a finite dimensional Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and  $B: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be two maximally monotone operators.

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

## Fixed point algorithm: Douglas-Rachford

Let  $\mathcal{H}$  be a finite dimensional Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and  $B: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be two maximally monotone operators.

Let  $\gamma \in ]0, +\infty[$  and let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $[0, 2]$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$ .

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

## Fixed point algorithm: Douglas-Rachford

Let  $\mathcal{H}$  be a finite dimensional Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and  $B: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be two maximally monotone operators.

Let  $\gamma \in ]0, +\infty[$  and let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $[0, 2]$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$ .

We assume that  $\text{zer}(A + B) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

The following properties are satisfied:

- ▶  $x_n \rightarrow \hat{x}$
- ▶  $(y_n)_{n \in \mathbb{N}}$  and  $(z_n)_{n \in \mathbb{N}}$  converge to  $J_{\gamma B} \hat{x} \in \text{zer}(A + B)$ .

## Fixed point algorithm: Douglas-Rachford

Proof: We set  $T = R_{\gamma A}R_{\gamma B}$ .

1.  $T$  is nonexpansive.
2.  $\text{zer}(A + B) = J_{\gamma B}(\text{Fix } T)$ .
3.  $\emptyset \neq \text{zer}(A + B) = J_{\gamma B}(\text{Fix } T) \Rightarrow \text{Fix } T \neq \emptyset$ .
4. We have

$$\begin{aligned}(\forall n \in \mathbb{N}) \quad x_{n+1} &= x_n + \lambda_n (J_{\gamma A}(2J_{\gamma B}x_n - x_n) - J_{\gamma B}x_n) \\ &= x_n + \frac{\lambda_n}{2} (Tx_n - x_n).\end{aligned}$$

$\Rightarrow$  Krasnosel'skii-Mann algorithm with relaxation steps  $(\lambda_n/2)_{n \in \mathbb{N}}$ .

5. Invoke continuity of  $J_{\gamma B}$ .

## Fixed point algorithm: Douglas-Rachford

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and  $B: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be two maximally monotone operators.

Let  $\gamma \in ]0, +\infty[$  and let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $[0, 2]$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$ .

We assume that  $\text{zer}(A + B) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

The following properties are satisfied:

- ▶  $x_n \rightharpoonup \hat{x}$
- ▶  $(y_n)_{n \in \mathbb{N}}$  and  $(z_n)_{n \in \mathbb{N}}$  converge weakly to  $J_{\gamma B} \hat{x} \in \text{zer}(A + B)$ .

## Optimization algorithm: Douglas-Rachford

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{H})$ .

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \text{prox}_{\gamma g} x_n \\ z_n = \text{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

## Optimization algorithm: Douglas-Rachford

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{H})$ .

Let  $\gamma \in ]0, +\infty[$  and let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $[0, 2]$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$ .

We assume that  $\text{zer}(\partial f + \partial g) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \text{prox}_{\gamma g} x_n \\ z_n = \text{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

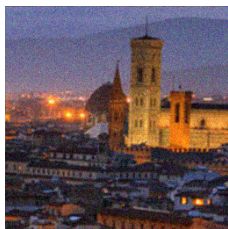
The following properties are satisfied:

- ▶  $x_n \rightarrow \hat{x}$
- ▶  $z_n - y_n \rightarrow 0$ ,  $y_n \rightarrow \hat{y}$ ,  $z_n \rightarrow \hat{y}$  where  $\hat{y} = \text{prox}_{\gamma g} \hat{x} \in \text{Argmin}(f + g)$ .

# Optimization algorithm: Douglas-Rachford

Image restoration :  $z = A\bar{x} + n$

- ▶  $\bar{x} \in \mathbb{R}^N$  : original image (**unknown**),
- ▶  $A \in \mathbb{R}^{N \times N}$  : blur operator,
- ▶  $n \in \mathbb{R}^N$  : additive noise (white zero-mean Gaussian),
- ▶  $z \in \mathbb{R}^N$  : observation = blur + noise



Degraded image  $z$

$\Rightarrow$  Find an image  $\hat{x}$  close to  $\bar{x}$   
using  $z$



## Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|Wx\|_1 \quad \text{with} \quad \eta \in ]0, +\infty[ \quad \text{and} \quad W \in \mathbb{R}^{N \times N}$$

► Douglas-Rachford algorithm with  $g = \|A \cdot - z\|_2^2$  and  $f = \eta \|W \cdot \|_1$

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \operatorname{prox}_{\gamma g} x_n \\ z_n = \operatorname{prox}_{\gamma f} (2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n (z_n - y_n). \end{cases}$$

## Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|Wx\|_1 \quad \text{with } \eta \in ]0, +\infty[ \quad \text{and } W \in \mathbb{R}^{N \times N}$$

► Douglas-Rachford algorithm with  $g = \|A \cdot - z\|_2^2$  and  $f = \eta \|W \cdot \|_1$

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \operatorname{prox}_{\gamma g} x_n & \rightarrow \text{Closed form} \\ z_n = \operatorname{prox}_{\gamma f} (2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n (z_n - y_n). \end{cases}$$

## Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|Wx\|_1 \quad \text{with} \quad \eta \in ]0, +\infty[ \quad \text{and} \quad W \in \mathbb{R}^{N \times N}$$

► Douglas-Rachford algorithm with  $g = \|A \cdot - z\|_2^2$  and  $f = \eta \|W \cdot \|_1$

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \operatorname{prox}_{\gamma g} x_n \\ z_n = \operatorname{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases} \quad \rightarrow \quad \text{Closed form}$$

$$\operatorname{prox}_{\gamma \eta \|\cdot\|_1} = \operatorname{soft}_{[-\gamma \eta, \gamma \eta]}$$

## Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|Wx\|_1 \quad \text{with} \quad \eta \in ]0, +\infty[ \quad \text{and} \quad W \in \mathbb{R}^{N \times N}$$

► Douglas-Rachford algorithm with  $g = \|A \cdot - z\|_2^2$  and  $f = \eta \|W \cdot \|_1$

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \operatorname{prox}_{\gamma g} x_n \\ z_n = \operatorname{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases} \rightarrow \begin{array}{l} \text{Closed form if } W^{-1} = W^* \\ \text{(e.g. wavelet transform)} \end{array}$$

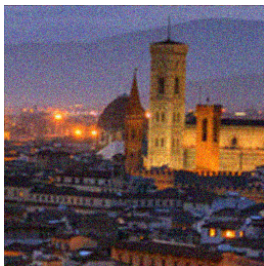
$$\operatorname{prox}_{f \circ W} = W^* \operatorname{prox}_f(W \cdot)$$

# Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \|Wx\|_1 \quad \text{with} \quad \eta \in ]0, +\infty[ \quad \text{and} \quad W \in \mathbb{R}^{N \times N}$$

## ► Douglas-Rachford algorithm



Degraded image  $z$



Restored image  $\hat{x}$  [DR - DWT]

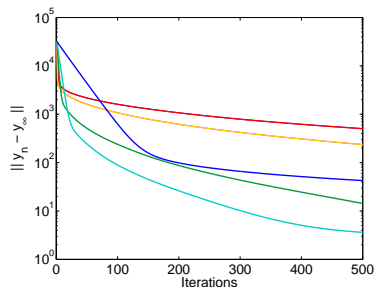
# Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\operatorname{Argmin}} \|Ax - z\|_2^2 + \eta \|Wx\|_1 \quad \text{with} \quad \eta \in ]0, +\infty[ \quad \text{and} \quad W \in \mathbb{R}^{N \times N}$$

## ► Douglas-Rachford algorithm

$$\begin{cases} y_n = \operatorname{prox}_{\gamma g} x_n \\ z_n = \operatorname{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$



$$\gamma = \{50, 10^2, 5 \cdot 10^2, 10^3, 5 \cdot 10^3\}$$

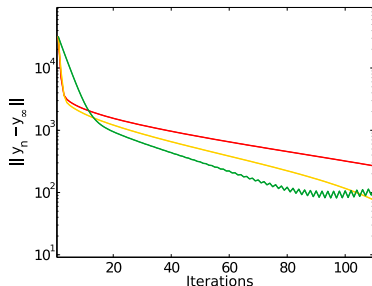
# Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \|Wx\|_1 \quad \text{with} \quad \eta \in ]0, +\infty[ \quad \text{and} \quad W \in \mathbb{R}^{N \times N}$$

► Douglas-Rachford algorithm

$$\begin{cases} y_n = \text{prox}_{\gamma g} x_n \\ z_n = \text{prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$



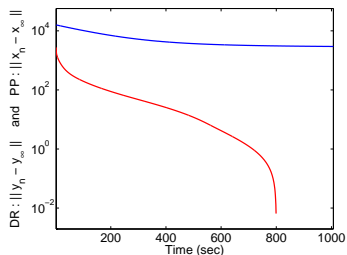
$$\lambda_n \equiv \{1, 1.5, 2.1\}$$

# Optimization algorithm: Douglas-Rachford

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \|Wx\|_1 \quad \text{with} \quad \eta \in ]0, +\infty[ \quad \text{and} \quad W \in \mathbb{R}^{N \times N}$$

## ► Douglas-Rachford algorithm



→ DR (red)

→ Proximal point algo. (blue)



## Optimization algorithm: Douglas-Rachford

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $g \in \Gamma_0(\mathcal{H})$  and  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  such that  $\text{ran } L$  is closed and  $L^*L$  is an isomorphism .

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \text{prox}_{\gamma g} x_n \\ c_n = (L^*L)^{-1}L^*y_n \\ x_{n+1} = x_n + \lambda_n(L(2c_n - v_n) - y_n) \\ v_{n+1} = v_n + \lambda_n(c_n - v_n). \end{cases}$$

## Optimization algorithm: Douglas-Rachford

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $g \in \Gamma_0(\mathcal{H})$  and  $L \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  such that  $\text{ran } L$  is closed and  $L^*L$  is an isomorphism.

Let  $\gamma \in ]0, +\infty[$  and let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $[0, 2]$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(2 - \lambda_n) = +\infty$ .

We assume that  $\text{zer}(L^* \circ \partial g \circ L) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$ ,  $v_0 = (L^*L)^{-1}L^*x_0$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \text{prox}_{\gamma g} x_n \\ c_n = (L^*L)^{-1}L^*y_n \\ x_{n+1} = x_n + \lambda_n(L(2c_n - v_n) - y_n) \\ v_{n+1} = v_n + \lambda_n(c_n - v_n). \end{cases}$$

We have then:

$v_n \rightarrow \hat{v}$  where  $\hat{v} \in \text{Argmin}(g \circ L)$ .

## Optimization algorithm: Douglas-Rachford

Sketch of proof:

$$\underset{v \in \mathcal{G}}{\text{minimize}} \ g(Lv) \Leftrightarrow \underset{x \in \mathcal{H}}{\text{minimize}} \ \iota_E(x) + g(x)$$

where  $E = \text{ran } L$ .

We apply Douglas-Rachford algorithm with  $f = \iota_E$ .

## Optimization algorithm: Douglas-Rachford

Particular case of Douglas-Rachford algorithm: **PPXA+**

$\mathcal{H} = \mathcal{H}_1 \times \cdots \times \mathcal{H}_m$  where  $\mathcal{H}_1, \dots, \mathcal{H}_m$  Hilbert spaces

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{n,i} = \text{prox}_{\gamma g_i} x_{n,i}, & i \in \{1, \dots, m\} \\ c_n = (\sum_{i=1}^m L_i^* L_i)^{-1} \sum_{i=1}^m L_i^* y_{n,i} \\ x_{n+1,i} = x_{n,i} + \lambda_n (L_i (2c_n - v_n) - y_{n,i}), & i \in \{1, \dots, m\} \\ v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{cases}$$

## Optimization algorithm: Douglas-Rachford

Particular case of Douglas-Rachford algorithm: **PPXA+**

$\mathcal{H} = \mathcal{H}_1 \times \cdots \times \mathcal{H}_m$  where  $\mathcal{H}_1, \dots, \mathcal{H}_m$  Hilbert spaces

( $\forall x = (x_1, \dots, x_m) \in \mathcal{H}$ )  $g(x) = \sum_{i=1}^m g_i(x_i)$

where ( $\forall i \in \{1, \dots, m\}$ )  $g_i \in \Gamma_0(\mathcal{H}_i)$

$L: v \mapsto (L_1 v, \dots, L_m v)$  where ( $\forall i \in \{1, \dots, m\}$ )  $L_i \in \mathcal{B}(\mathcal{G}, \mathcal{H}_i)$ .

Let  $(x_{0,i})_{1 \leq i \leq m} \in \mathcal{H}$ ,  $v_0 = (\sum_{i=1}^m L_i^* L_i)^{-1} \sum_{i=1}^m L_i^* x_{0,i}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{n,i} = \text{prox}_{\gamma g_i} x_{n,i}, & i \in \{1, \dots, m\} \\ c_n = (\sum_{i=1}^m L_i^* L_i)^{-1} \sum_{i=1}^m L_i^* y_{n,i} \\ x_{n+1,i} = x_{n,i} + \lambda_n (L_i (2c_n - v_n) - y_{n,i}), & i \in \{1, \dots, m\} \\ v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{cases}$$

We have then  $v_n \rightarrow \hat{v} \in \text{Argmin} \sum_{i=1}^m g_i \circ L_i$ .

## Optimization algorithm: Douglas-Rachford

Particular case of Douglas-Rachford algorithm: **PPXA**

$\mathcal{H} = \mathcal{H}_1 \times \dots \times \mathcal{H}_m$  where  $\mathcal{H}_1, \dots, \mathcal{H}_m$  Hilbert spaces

( $\forall x = (x_1, \dots, x_m) \in \mathcal{H}$ )  $g(x) = \sum_{i=1}^m g_i(x_i)$

where ( $\forall i \in \{1, \dots, m\}$ )  $g_i \in \Gamma_0(\mathcal{H}_i)$

$L: v \mapsto (L_1 v, \dots, L_m v)$  where ( $\forall i \in \{1, \dots, m\}$ )  $L_i \in \mathcal{B}(\mathcal{G}, \mathcal{H}_i)$ .

If  $\mathcal{H}_1 = \dots = \mathcal{H}_m$  and  $L_1 = \dots = L_m = \text{Id} \Rightarrow$  Consensus trick

Let  $(x_{0,i})_{1 \leq i \leq m} \in \mathcal{H}$ ,  $v_0 = \frac{1}{m} \sum_{i=1}^m x_{0,i}$  et

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{n,i} = \text{prox}_{\gamma g_i} x_{n,i}, & i \in \{1, \dots, m\} \\ c_n = \frac{1}{m} \sum_{i=1}^m y_{n,i} \\ x_{n+1,i} = x_{n,i} + \lambda_n (2c_n - v_n - y_{n,i}), & i \in \{1, \dots, m\} \\ v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{cases}$$

We have then  $v_n \rightarrow \hat{v} \in \text{Argmin} \sum_{i=1}^m g_i$ .

# Optimization algorithm: PPXA+

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \| [H^* \ V^*]^* x \|_{2,1} + \iota_C(x) \text{ with } \begin{cases} \eta \in ]0, +\infty[ \\ H, V \in \mathbb{R}^{N \times N} \\ C = [0, 255]^N \end{cases}$$

- ▶ **PPXA+** with  $g_1 = \|A \cdot -z\|_2^2$  and  $L_1 = \text{Id}$   
 $g_2 = \eta \| \cdot \|_{1,2}$  and  $L_2 = [H^* \ V^*]^*$   
 $g_3 = \iota_C$  and  $L_3 = \text{Id}$

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{n,i} = \text{prox}_{\gamma g_i} x_{n,i}, i \in \{1, 2, 3\} \\ c_n = (\sum_{i=1}^3 L_i^* L_i)^{-1} \sum_{i=1}^3 L_i^* y_{n,i} \\ x_{n+1,i} = x_{n,i} + \lambda_n (L_i (2c_n - v_n) - y_{n,i}), i \in \{1, 2, 3\} \\ v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{cases}$$

# Optimization algorithm: PPXA+

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \| [H^* \ V^*]^* x \|_{2,1} + \iota_C(x) \quad \text{with} \quad \begin{cases} \eta \in ]0, +\infty[ \\ H, V \in \mathbb{R}^{N \times N} \\ C = [0, 255]^N \end{cases}$$

- ▶ **PPXA+** with  $g_1 = \|A \cdot -z\|_2^2$  and  $L_1 = \text{Id}$   
 $g_2 = \eta \| \cdot \|_{1,2}$  and  $L_2 = [H^* \ V^*]^*$   
 $g_3 = \iota_C$  and  $L_3 = \text{Id}$

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{n,i} = \text{prox}_{\gamma g_i} x_{n,i}, i \in \{1, 2, 3\} & \rightarrow \text{Closed form} \\ c_n = (\sum_{i=1}^3 L_i^* L_i)^{-1} \sum_{i=1}^3 L_i^* y_{n,i} & \rightarrow \text{Closed form} \\ x_{n+1,i} = x_{n,i} + \lambda_n (L_i (2c_n - v_n) - y_{n,i}), i \in \{1, 2, 3\} \\ v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{cases}$$

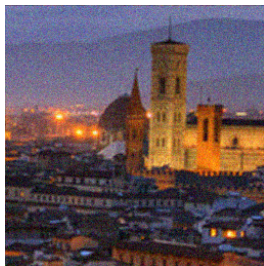


# Optimization algorithm: PPXA+

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \| [H^* \ V^*]^* x \|_{2,1} + \iota_C(x) \quad \text{with} \quad \begin{cases} \eta \in ]0, +\infty[ \\ H, V \in \mathbb{R}^{N \times N} \\ C = [0, 255]^N \end{cases}$$

## ► PPXA+



Degraded image  $z$



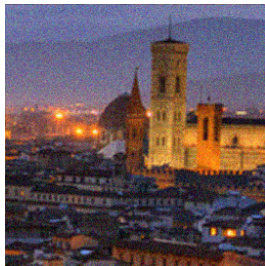
Restored image  $\hat{x}$  [PPXA - TV]

# Optimization algorithm: PPXA+

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \| [H^* \ V^*]^* x \|_{2,1} + \iota_C(x) \text{ with } \begin{cases} \eta \in ]0, +\infty[ \\ H, V \in \mathbb{R}^{N \times N} \\ C = [0, 255]^N \end{cases}$$

## ► PPXA+



Degraded image  $z$



Restored image  $\hat{x}$  [DR – DWT]

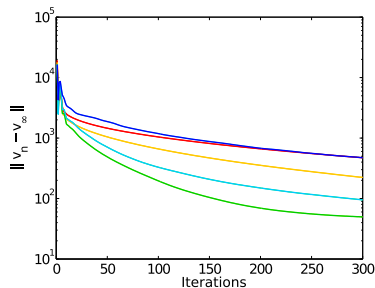
# Optimization algorithm: PPXA+

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \| [H^* \ V^*]^* x \|_{2,1} + \iota_C(x) \text{ with } \begin{cases} \eta \in ]0, +\infty[ \\ H, V \in \mathbb{R}^{N \times N} \\ C = [0, 255]^N \end{cases}$$

► PPXA+

$$\begin{cases} y_{n,i} = \text{prox}_{\gamma g_i} x_{n,i}, \\ c_n = (\sum_{i=1}^3 L_i^* L_i)^{-1} \sum_{i=1}^2 L_i^* y_{n,i} \\ x_{n+1,i} = x_{n,i} + \lambda_n (L_i(2c_n - v_n) - y_{n,i}), \\ v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{cases}$$



$$\gamma = \{5.10^2, 10^3, 5.10^3, 10^4, 5.10^4\}$$

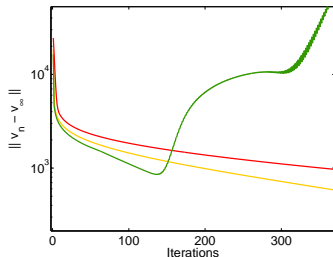
# Optimization algorithm: PPXA+

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \| [H^* \ V^* ]^* x \|_{2,1} + \iota_C(x) \text{ with } \begin{cases} \eta \in ]0, +\infty[ \\ H, V \in \mathbb{R}^{N \times N} \\ C = [0, 255]^N \end{cases}$$

► PPXA+

$$\begin{cases} y_{n,i} = \text{prox}_{\gamma g_i} x_{n,i}, \\ c_n = (\sum_{i=1}^3 L_i^* L_i)^{-1} \sum_{i=1}^3 L_i^* y_{n,i} \\ x_{n+1,i} = x_{n,i} + \lambda_n (L_i (2c_n - v_n) - y_{n,i}), \\ v_{n+1} = v_n + \lambda_n (c_n - v_n). \end{cases}$$



$$\lambda_n \equiv \{1, 1.8, 2.1\}$$

## Fixed point algorithm: $\alpha$ -averaged operator

Let  $\mathcal{H}$  be a Hilbert space.

Let  $T: \mathcal{H} \rightarrow \mathcal{H}$  be a  $\alpha$ -averaged operator with  $\alpha \in ]0, 1[$

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n (Tx_n - x_n).$$

## Fixed point algorithm: $\alpha$ -averaged operator

Let  $\mathcal{H}$  be a Hilbert space.

Let  $T: \mathcal{H} \rightarrow \mathcal{H}$  be a  $\alpha$ -averaged operator with  $\alpha \in ]0, 1[$

Let  $(\lambda_n)_{n \in \mathbb{N}}$  a sequence in  $[0, 1/\alpha]$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(1 - \alpha\lambda_n) = +\infty$ .

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n).$$

## Fixed point algorithm: $\alpha$ -averaged operator

Let  $\mathcal{H}$  be a Hilbert space.

Let  $T: \mathcal{H} \rightarrow \mathcal{H}$  be a  $\alpha$ -averaged operator with  $\alpha \in ]0, 1[$  such that  $\text{Fix} T \neq \emptyset$ .

Let  $(\lambda_n)_{n \in \mathbb{N}}$  a sequence in  $[0, 1/\alpha]$  such that  $\sum_{n \in \mathbb{N}} \lambda_n (1 - \alpha \lambda_n) = +\infty$ .

Let  $x_0 \in \mathcal{H}$  and

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n (Tx_n - x_n).$$

Then,  $(x_n)_{n \in \mathbb{N}}$  converges weakly to a point in  $\text{Fix} T$ .

Proof : Use properties of Krasnosel'skii-Mann algorithm.

## Fixed point algorithm: Forward-Backward

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator .

Let  $B: \mathcal{H} \rightarrow \mathcal{H}$  be a  $\beta$ -cocoercive operator with  $\beta \in ]0, +\infty[$  .

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = x_n - \gamma Bx_n \\ x_{n+1} = x_n + \lambda_n (J_{\gamma A} y_n - x_n). \end{cases}$$



## Fixed point algorithm: Forward-Backward

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator .

Let  $B: \mathcal{H} \rightarrow \mathcal{H}$  be a  $\beta$ -cocoercive operator with  $\beta \in ]0, +\infty[$  .

Let  $\gamma \in ]0, 2\beta[$  and  $\delta = \min\{1, \beta/\gamma\} + 1/2$ .

Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $[0, \delta[$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$ .

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = x_n - \gamma Bx_n \\ x_{n+1} = x_n + \lambda_n(J_{\gamma A}y_n - x_n). \end{cases}$$

## Fixed point algorithm: Forward-Backward

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator .

Let  $B: \mathcal{H} \rightarrow \mathcal{H}$  be a  $\beta$ -cocoercive operator with  $\beta \in ]0, +\infty[$  .

Let  $\gamma \in ]0, 2\beta[$  and  $\delta = \min\{1, \beta/\gamma\} + 1/2$ .

Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $[0, \delta[$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$ .

We assume that  $\text{zer}(A + B) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$  and

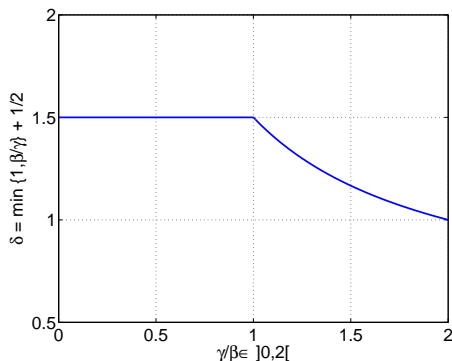
$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = x_n - \gamma Bx_n \\ x_{n+1} = x_n + \lambda_n(J_{\gamma A}y_n - x_n). \end{cases}$$

$(x_n)_{n \in \mathbb{N}}$  converges weakly to a point in  $\text{zer}(A + B)$ .

## Fixed point algorithm: Forward-Backward

Let  $\gamma \in ]0, 2\beta[$  and  $\delta = \min\{1, \beta/\gamma\} + 1/2$ .

Let  $(\lambda_n)_{n \in \mathbb{N}}$  a sequence in  $[0, \delta[$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$ .



## Fixed point algorithm: Forward-Backward

Proof: Let  $T = J_{\gamma A}(\text{Id} - \gamma B)$ .

1.  $\text{Fix } T = \text{zer}(A + B) \neq \emptyset$ :

$$(\forall x \in \mathcal{H}) \quad x \in \text{Fix } T \Leftrightarrow x - \gamma Bx \in (\text{Id} + \gamma A)x \Leftrightarrow 0 \in Ax + Bx.$$

2. The iterations can be written as

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = x_n + \lambda_n(Tx_n - x_n)$$

3.  $T$  is  $\alpha$ -averaged :  $B$   $\beta$ -cocoercive and  $\gamma \in ]0, 2\beta[ \Rightarrow \text{Id} - \gamma B$  is  $\gamma/(2\beta)$ -averaged and  $J_{\gamma A}$  is  $1/2$ -averaged.

Then,  $T$  is  $\alpha$ -averaged with

$$\alpha = \frac{2}{1 + \frac{1}{\max\{\frac{1}{2}, \frac{\gamma}{2\beta}\}}} \Leftrightarrow \alpha^{-1} = \delta.$$

## Fixed point algorithm: Forward-Backward

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator .

Let  $B: \mathcal{H} \rightarrow \mathcal{H}$  be  $\beta$ -cocoercive where  $\beta \in ]0, +\infty[$  .

Let  $(\gamma_n)_{n \in \mathbb{N}}$  be a sequence in  $[\underline{\gamma}, \bar{\gamma}]$  where  $0 < \underline{\gamma} \leq \bar{\gamma} < 2\beta$ .

Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $[\underline{\lambda}, 1]$  where  $\underline{\lambda} \in ]0, 1]$ .

We assume that  $\text{zer}(A + B) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = x_n - \gamma_n Bx_n \\ x_{n+1} = x_n + \lambda_n (J_{\gamma_n A} y_n - x_n). \end{cases}$$

$(x_n)_{n \in \mathbb{N}}$  converges weakly to a point in  $\text{zer}(A + B)$ .

Remark: When  $B = 0$  and  $\lambda_n \equiv 1$ , the algorithm reduces to the proximal point algorithm.

## Optimization algorithm: Forward-Backward

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{H})$  such that  $\nabla g$  is  $\frac{1}{\beta}$ -Lipschitz where  $\beta \in ]0, +\infty[$ .

Let  $(\gamma_n)_{n \in \mathbb{N}}$  in  $[\underline{\gamma}, \bar{\gamma}]$  where  $0 < \underline{\gamma} \leq \bar{\gamma} < 2\beta$ .

Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $[\underline{\lambda}, 1]$  where  $\underline{\lambda} \in ]0, 1]$ .

We assume that  $\text{Argmin}(f + g) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n (\text{prox}_{\gamma_n f} y_n - x_n). \end{cases}$$

$(x_n)_{n \in \mathbb{N}}$  converges weakly to a minimizer of  $f + g$ .

## Optimization algorithm: Forward-Backward

### MM (Majoration-Minimization) interpretation

For all  $n \in \mathbb{N}$ , let  $p_n = \text{prox}_{\gamma_n f}(x_n - \gamma_n \nabla g(x_n))$ . We have:

$$f(x_{n+1}) = f((1 - \lambda_n)x_n + \lambda_n p_n) \leq (1 - \lambda_n)f(x_n) + \lambda_n f(p_n).$$

We use the descent lemma,

$$g(x_{n+1}) \leq g(x_n) + \langle \nabla g(x_n) \mid x_{n+1} - x_n \rangle + \frac{1}{2\beta} \|x_{n+1} - x_n\|^2,$$

the proximity operator definition

$$x_n - \gamma_n \nabla g(x_n) - p_n \in \partial f(p_n)$$

and, the subdifferential definition

$$\gamma_n f(p_n) + \langle x_n - \gamma_n \nabla g(x_n) - p_n \mid x_n - p_n \rangle \leq \gamma_n f(x_n)$$

$$\Leftrightarrow \lambda_n f(p_n) + \langle \nabla g(x_n) \mid x_{n+1} - x_n \rangle + \gamma_n^{-1} \lambda_n^{-1} \|x_{n+1} - x_n\|^2 \leq \lambda_n f(x_n)$$

since  $x_{n+1} - x_n = \lambda_n(p_n - x_n)$ .

## Optimization algorithm: Forward-Backward

### MM (Majoration-Minimization) interpretation

$$\Rightarrow f(x_{n+1}) + g(x_{n+1}) + \left( \frac{1}{\gamma_n \lambda_n} - \frac{1}{2\beta} \right) \|x_{n+1} - x_n\|^2 \leq f(x_n) + g(x_n).$$

Thus, if  $\gamma_n^{-1} \lambda_n^{-1} - \frac{1}{2} \beta^{-1} \geq 0 \Rightarrow \gamma_n \lambda_n \leq 2\beta$ ,  $(f(x_n) + g(x_n))_{n \in \mathbb{N}}$  is a decreasing sequence.

Because  $(\forall n \in \mathbb{N}) p_n = \text{prox}_{\gamma_n f} y_n \in \text{dom } f$ , if  $x_0 \in \text{dom } f$ , then  $(\forall n \in \mathbb{N}) x_{n+1} = (1 - \lambda_n)x_n + \lambda_n p_n \in \text{dom } f$ .

Thus,  $(f(x_n) + g(x_n))_{n \in \mathbb{N}}$  is a real decreasing sequence.



## Optimization algorithm: Forward-Backward

Beck-Teboule proximal gradient algorithm (Nesterov acceleration)

Let  $x_0 = z_0 \in \mathcal{H}$ ,  $t_0 = 1$ , and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = z_n - \beta \nabla g(z_n) \\ x_{n+1} = \text{prox}_{\beta f} y_n \\ t_{n+1} = \frac{1 + \sqrt{4t_n^2 + 1}}{2} \\ \lambda_n = 1 + \frac{t_n - 1}{t_{n+1}} \\ z_{n+1} = z_n + \lambda_n (x_{n+1} - x_n). \end{cases}$$

Convergence of  $(f(x_n))_{n \in \mathbb{N}}$  at optimal  $O(1/n^2)$  rate, but convergence of  $(x_n)_{n \in \mathbb{N}}$  not secured theoretically.

## Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\hat{u} \in \underset{u \in \mathbb{R}^N}{\text{Argmin}} \|AF^*u - z\|_2^2 + \eta \|u\|_1 \text{ with } \eta \in ]0, +\infty[ \text{ and } F \in \mathbb{R}^{N \times M}$$

► FB with  $g = \|AF^* \cdot - z\|_2^2$ ,  $f = \eta \| \cdot \|_1$  and  $\hat{x} = F^* \hat{u}$ .

$$\begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n (\text{prox}_{\gamma_n f} y_n - x_n) \end{cases}$$

## Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\hat{u} \in \underset{u \in \mathbb{R}^N}{\text{Argmin}} \|AF^*u - z\|_2^2 + \eta \|u\|_1 \text{ with } \eta \in ]0, +\infty[ \text{ and } F \in \mathbb{R}^{N \times M}$$

► FB with  $g = \|AF^* \cdot - z\|_2^2$ ,  $f = \eta \|\cdot\|_1$  and  $\hat{x} = F^*\hat{u}$ .

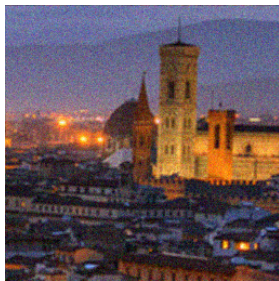
$$\begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) & \rightarrow \text{Closed form: } 2FA^*(AF^* \cdot - z) \\ x_{n+1} = x_n + \lambda_n (\text{prox}_{\gamma_n f} y_n - x_n) & \rightarrow \text{Closed form} \end{cases}$$

# Optimization algorithm: Forward-Backward

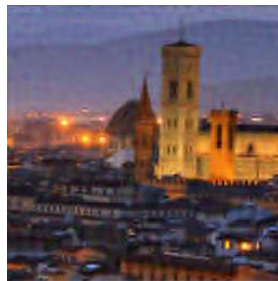
Image restoration : Variational approach

$$\hat{u} \in \underset{u \in \mathbb{R}^N}{\text{Argmin}} \|AF^*u - z\|_2^2 + \eta \|u\|_1 \text{ with } \eta \in ]0, +\infty[ \text{ and } F \in \mathbb{R}^{N \times M}$$

► FB



Degraded image  $z$



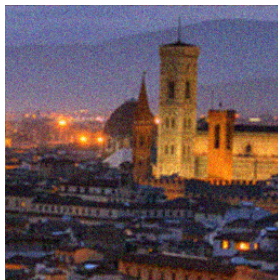
Restored image [FB - DTT]

# Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\hat{u} \in \underset{u \in \mathbb{R}^N}{\text{Argmin}} \|AF^*u - z\|_2^2 + \eta \|u\|_1 \text{ with } \eta \in ]0, +\infty[ \text{ and } F \in \mathbb{R}^{N \times M}$$

► FB



Degraded image  $z$



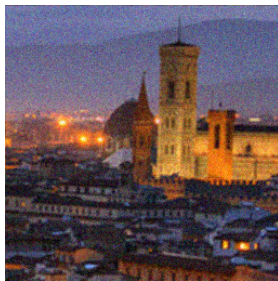
Restored image [PPXA - TV]

# Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\hat{u} \in \underset{u \in \mathbb{R}^N}{\text{Argmin}} \|AF^*u - z\|_2^2 + \eta \|u\|_1 \text{ with } \eta \in ]0, +\infty[ \text{ and } F \in \mathbb{R}^{N \times M}$$

► FB



Degraded image  $z$



Restored image [DR - DWT]

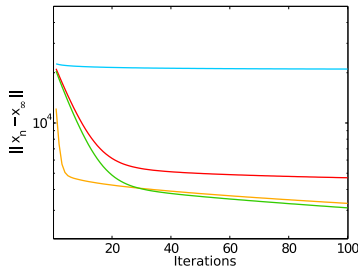
# Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\hat{u} \in \underset{u \in \mathbb{R}^N}{\text{Argmin}} \|AF^*u - z\|_2^2 + \eta \|u\|_1 \text{ with } \eta \in ]0, +\infty[ \text{ and } F \in \mathbb{R}^{N \times M}$$

► FB

$$\begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n (\text{prox}_{\gamma_n f} y_n - x_n) \end{cases}$$



$$2\|AF^*\|^2 \gamma_n \equiv \{0.1, 1.5, 1.9, 2\}$$

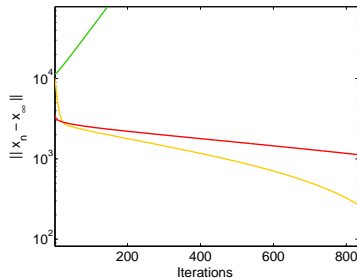
# Optimization algorithm: Forward-Backward

Image restoration : Variational approach

$$\hat{u} \in \underset{u \in \mathbb{R}^N}{\text{Argmin}} \|AF^*u - z\|_2^2 + \eta \|u\|_1 \text{ with } \eta \in ]0, +\infty[ \text{ and } F \in \mathbb{R}^{N \times M}$$

► FB

$$\begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n (\text{prox}_{\gamma_n f} y_n - x_n) \end{cases}$$



$$\lambda_n \equiv \{0.5, 1, 1.1\}$$



## Optimization algorithm: projected gradient

Let  $\mathcal{H}$  be a Hilbert space.

Let  $C$  be a nonempty closed convex subset of  $\mathcal{H}$ .

Let  $g \in \Gamma_0(\mathcal{H})$  such that  $\nabla g$  is  $\frac{1}{\beta}$ -Lipschitz where  $\beta \in ]0, +\infty[$ .

Let  $(\gamma_n)_{n \in \mathbb{N}}$  a sequence in  $[\underline{\gamma}, \bar{\gamma}]$  where  $0 < \underline{\gamma} \leq \bar{\gamma} < 2\beta$ .

Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $[\underline{\lambda}, 1]$  where  $\underline{\lambda} \in ]0, 1]$ .

We assume that  $\text{Argmin}_{x \in C} g(x) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n (P_C y_n - x_n). \end{cases}$$

$(x_n)_{n \in \mathbb{N}}$  converges weakly to a minimizer of  $g$  on  $C$ .

## Optimization algorithm: projected gradient

Image restoration : Variational approach – regularized

$$\hat{u} \in \underset{u \in \mathbb{R}^N}{\text{Argmin}} \|AF^*u - z\|_2^2 + \eta \|u\|_1 \quad \text{with} \quad \begin{cases} \eta \in ]0, +\infty[ \\ F \in \mathbb{R}^{N \times M} \end{cases}$$

Image restoration : Variational approach – constrained

$$\hat{u} \in \underset{u \in \mathbb{R}^N, \|u\|_1 \leq \epsilon}{\text{Argmin}} \|AF^*u - z\|_2^2 \quad \text{with} \quad \begin{cases} \epsilon \in ]0, +\infty[ \\ F \in \mathbb{R}^{N \times M} \end{cases}$$

## Fixed point algorithm: *Forward-Backward-Forward*

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  be a maximally monotone operator .

Let  $B: \mathcal{H} \rightarrow \mathcal{H}$  be a monotone operator,  $\frac{1}{\beta}$ -Lipschitz where  $\beta \in ]0, +\infty[$  .

Let  $\varepsilon \in ]0, \beta/(1 + \beta)[$  and  $(\gamma_n)_{n \in \mathbb{N}}$  in  $[\varepsilon, (1 - \varepsilon)\beta]$ .

We assume that  $\text{zer}(A + B) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = x_n - \gamma_n Bx_n \\ p_n = J_{\gamma_n A} y_n \\ q_n = p_n - \gamma_n Bp_n \\ x_{n+1} = x_n - y_n + q_n. \end{cases}$$

$(x_n)_{n \in \mathbb{N}}$  and  $(p_n)_{n \in \mathbb{N}}$  converge weakly to  $\hat{x} \in \text{zer}(A + B)$ .

## Fixed point algorithm: *Forward-Backward-Forward*

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f \in \Gamma_0(\mathcal{H})$ .

Let  $g \in \Gamma_0(\mathcal{H})$  such that  $\nabla g$  is  $\frac{1}{\beta}$ -Lipschitz where  $\beta \in ]0, +\infty[$ .

Let  $\varepsilon \in ]0, \beta/(1 + \beta)[$  and  $(\gamma_n)_{n \in \mathbb{N}}$  in  $[\varepsilon, (1 - \varepsilon)\beta]$ .

We assume that  $\text{Argmin}(f + g) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = x_n - \gamma_n \nabla g(x_n) \\ p_n = \text{prox}_{\gamma_n f} y_n \\ q_n = p_n - \gamma_n \nabla g(p_n) \\ x_{n+1} = x_n - y_n + q_n. \end{cases}$$

$(x_n)_{n \in \mathbb{N}}$  and  $(p_n)_{n \in \mathbb{N}}$  converge weakly to  $\hat{x} \in \text{Argmin}(f + g)$ .

## Part 4: Duality

1. General duality concepts
  - ▶ Primal and dual problems
  - ▶ Duality theorems
  - ▶ Inf-convolution
2. Augmented Lagrangian algorithms
  - ▶ ADMM
  - ▶ SDMM
3. Primal-dual algorithms
  - ▶ FB-based PD algorithm
  - ▶ M+LFBF algorithm

# Duality

Let  $\mathcal{H}$  be a Hilbert space.

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and  $B: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ .

The **parallel sum** of  $A$  and  $B$  is

$$A \square B = (A^{-1} + B^{-1})^{-1}.$$

- ▶ If  $A$  and  $B$  are monotone, then  $A \square B$  is monotone.
- ▶  $A \square N_{\{0\}} = A$  where  $N_{\{0\}} = \partial \iota_{\{0\}}$  and  $\text{gra} N_{\{0\}} = \{0\} \times \mathcal{H}$ .
- ▶  $A \square B = B \square A$ .

# Duality

## Primal problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ ,  $C: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ ,  $B: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  and  $D: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be monotone operators. Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ . Find  $\hat{x} \in \mathcal{H}$  such that

$$0 \in A\hat{x} + C\hat{x} + L^*(B \square D)L\hat{x}.$$

## Dual problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ ,  $C: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ ,  $B: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  and  $D: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be monotone operators. Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ . Find  $\hat{v} \in \mathcal{G}$  such that

$$0 \in -L(A^{-1} \square C^{-1})(-L^*\hat{v}) + B^{-1}\hat{v} + D^{-1}\hat{v}.$$

## Duality theorem

Let  $\hat{x} \in \mathcal{H}$  and  $\hat{v} \in \mathcal{G}$ .  $(\hat{x}, \hat{v})$  is a **Kuhn-Tucker point** if

$$\begin{cases} -L^*\hat{v} \in (A + C)\hat{x} \\ L\hat{x} \in (B^{-1} + D^{-1})\hat{v}. \end{cases}$$

If  $\hat{x} \in \mathcal{H}$  is a solution to the primal problem, then there exists a solution  $\hat{v}$  to the dual problem such that  $(\hat{x}, \hat{v})$  is a Kuhn-Tucker point.



## Duality theorem

Let  $\hat{x} \in \mathcal{H}$  and  $\hat{v} \in \mathcal{G}$ .  $(\hat{x}, \hat{v})$  is a **Kuhn-Tucker point** if

$$\begin{cases} -L^*\hat{v} \in (A + C)\hat{x} \\ L\hat{x} \in (B^{-1} + D^{-1})\hat{v}. \end{cases}$$

If  $\hat{v} \in \mathcal{G}$  is a solution to the dual problem, then there exists a solution  $\hat{x}$  to the primal problem such that  $(\hat{x}, \hat{v})$  is a Kuhn-Tucker point.

## Duality theorem

Let  $\hat{x} \in \mathcal{H}$  and  $\hat{v} \in \mathcal{G}$ .  $(\hat{x}, \hat{v})$  is a **Kuhn-Tucker point** if

$$\begin{cases} -L^*\hat{v} \in (A + C)\hat{x} \\ L\hat{x} \in (B^{-1} + D^{-1})\hat{v}. \end{cases}$$

If  $(\hat{x}, \hat{v})$  is a Kuhn-Tucker point, then  $\hat{x}$  is a solution to the primal problem and  $\hat{v}$  is a solution to the dual problem.

## Inf-convolution

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$  and  $g: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

The **inf-convolution** of  $f$  and  $g$  is

$$f \square g: \mathcal{H} \rightarrow [-\infty, +\infty]: x \mapsto \inf_{y \in \mathcal{H}} f(y) + g(x - y)$$

## Inf-convolution

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$  and  $g: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

The **inf-convolution** of  $f$  and  $g$  is

$$f \square g: \mathcal{H} \rightarrow [-\infty, +\infty]: x \mapsto \inf_{\substack{(u,v) \in \mathcal{H}^2 \\ u+v=x}} f(u) + g(v)$$

## Inf-convolution

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$  and  $g: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

The **inf-convolution** of  $f$  and  $g$  is

$$f \square g: \mathcal{H} \rightarrow [-\infty, +\infty]: x \mapsto \inf_{\substack{(u,v) \in \mathcal{H}^2 \\ u+v=x}} f(u) + g(v)$$

- ▶  $f \square \iota_{\{0\}} = f$
- ▶  $f \square g = g \square f$
- ▶  $\text{dom}(f \square g) = \text{dom } f + \text{dom } g$
- ▶  $\gamma f = f \square \frac{1}{2\gamma} \|\cdot\|^2, \gamma \in ]0, +\infty[.$

## Inf-convolution

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$  and  $g: \mathcal{H} \rightarrow ]-\infty, +\infty]$ .

The **inf-convolution** of  $f$  and  $g$  is

$$f \square g: \mathcal{H} \rightarrow [-\infty, +\infty]: x \mapsto \inf_{y \in \mathcal{H}} f(y) + g(x - y)$$

If  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$  and  $g: \mathcal{H} \rightarrow ]-\infty, +\infty]$  are convex, then  $f \square g$  is convex.

# Inf-convolution

	conjugate		Fourier transform ( $\mathcal{H}$ finite dimensional)	
Property	$h(x)$	$h^*(u)$	$h(x)$	$\widehat{h}(\nu)$
inf-convolution /convolution	$(f \square g)(x)$ $= \inf_{y \in \mathcal{H}} f(y) + g(x - y)$	$f^*(u) + g^*(u)$	$(f \star g)(x)$ $= \int_{\mathcal{H}} f(y)g(x - y)dy$	$\widehat{f}(\nu)\widehat{g}(\nu)$
sum/product	$f(x) + g(x)$ $f \in \Gamma_0(\mathcal{H})$ $g \in \Gamma_0(\mathcal{H})$ $\text{dom } f \cap \text{dom } g \neq \emptyset$	$(f^* \square g^*)(u)$	$f(x)g(x)$	$(\widehat{f} \star \widehat{g})(\nu)$

## Inf-convolution

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{H})$ .

If  $\text{dom } f^* \cap \text{int}(\text{dom } g^*) \neq \emptyset$ , then

$$\underbrace{\partial(f \square g)}_{\text{inf-convolution}} = \underbrace{\partial f \square \partial g}_{\text{parallel sum}}.$$



## Inf-convolution

Let  $\mathcal{H}$  be a Hilbert space.

Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{H})$ .

If  $\text{dom } f^* \cap \text{int}(\text{dom } g^*) \neq \emptyset$ , then

$$\underbrace{\partial(f \square g)}_{\text{inf-convolution}} = \underbrace{\partial f \square \partial g}_{\text{parallel sum}}$$

Proof:

$$\begin{aligned} \partial(f \square g) &= \partial(f^* + g^*)^* \\ &= (\partial(f^* + g^*))^{-1} \\ &= (\partial f^* + \partial g^*)^{-1} \\ &= ((\partial f)^{-1} + (\partial g)^{-1})^{-1} \\ &= \partial f \square \partial g. \end{aligned}$$

# Fenchel-Rockafellar duality

## Primal problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ ,  $g: \mathcal{G} \rightarrow ]-\infty, +\infty]$ . Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

We want to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + g(Lx).$$

## Dual problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $f: \mathcal{H} \rightarrow ]-\infty, +\infty]$ ,  $g: \mathcal{G} \rightarrow ]-\infty, +\infty]$ . Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

We want to

$$\underset{v \in \mathcal{G}}{\text{minimize}} \quad f^*(-L^*v) + g^*(v).$$

## Fenchel-Rockafellar duality

### Weak duality

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces.

Let  $f$  be a proper function from  $\mathcal{H}$  to  $]-\infty, +\infty]$ ,  $g$  be a proper function from  $\mathcal{G}$  to  $]-\infty, +\infty]$  and  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ . Let

$$\mu = \inf_{x \in \mathcal{H}} f(x) + g(Lx) \quad \text{and} \quad \mu^* = \inf_{v \in \mathcal{G}} f^*(-L^*v) + g^*(v).$$

We have  $\mu \geq -\mu^*$ . If  $\mu \in \mathbb{R}$ ,  $\mu + \mu^*$  is called the **duality gap**.

## Fenchel-Rockafellar duality

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces,  $f \in \Gamma_0(\mathcal{H})$ ,  $g \in \Gamma_0(\mathcal{G})$  and  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .  
If  $\text{int}(\text{dom } g) \cap L(\text{dom } f) \neq \emptyset$  and  $\text{dom } g \cap \text{int}(L(\text{dom } f)) \neq \emptyset$ , then

$$\partial f + L^* \partial g L = \partial(f + g \circ L)$$

## Fenchel-Rockafellar duality

### Duality theorem (1)

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces,  $f \in \Gamma_0(\mathcal{H})$ ,  $g \in \Gamma_0(\mathcal{G})$  and  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

$$\text{zer}(\partial f + L\partial gL^*) \neq \emptyset \quad \Leftrightarrow \quad \text{zer}((-L)\partial f^*(-L^*) + \partial g^*) \neq \emptyset$$

# Fenchel-Rockafellar duality

## Duality theorem (2)

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces,  $f \in \Gamma_0(\mathcal{H})$ ,  $g \in \Gamma_0(\mathcal{G})$  and  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

- ▶ If there exists  $\hat{x} \in \mathcal{H}$  such that  $0 \in \partial f(\hat{x}) + L^* \partial g(L\hat{x})$  then  $\hat{x}$  is a solution to the primal problem. Moreover,  $-L^* \hat{v} \in \partial f(\hat{x})$  and  $L\hat{x} \in \partial g^*(\hat{v})$  where  $\hat{v}$  is a solution to the dual problem.
- ▶ If there exists  $(\hat{x}, \hat{v}) \in \mathcal{H} \times \mathcal{G}$  such that  $-L^* \hat{v} \in \partial f(\hat{x})$  and  $L\hat{x} \in \partial g^*(\hat{v})$  then  $\hat{x}$  (resp.  $\hat{v}$ ) is a solution to the primal (resp. dual) problem.

Particular case: If  $f = \varphi + \frac{1}{2} \|\cdot - z\|^2$  where  $\varphi \in \Gamma_0(\mathcal{H})$  and  $z \in \mathcal{H}$ , then

$$-L^* \hat{v} \in \partial f(\hat{x}) \Leftrightarrow -L^* \hat{v} \in \partial \varphi(\hat{x}) + \hat{x} - z.$$

We have then  $\hat{x} = \text{prox}_{\varphi}(-L^* \hat{v} + z)$ .

## Fenchel-Rockafellar duality

### Duality theorem (3): strong duality

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces,  $f \in \Gamma_0(\mathcal{H})$ ,  $g \in \Gamma_0(\mathcal{G})$  and  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .  
If  $\text{int}(\text{dom } g) \cap L(\text{dom } f) \neq \emptyset$  or  $\text{dom } g \cap \text{int}(L(\text{dom } f)) \neq \emptyset$ , then

$$\mu = \inf_{x \in \mathcal{H}} f(x) + g(Lx) = - \min_{v \in \mathcal{G}} f^*(-L^*v) + g^*(v) = -\mu^*.$$

# Fenchel-Rockafellar duality

## Extension of the duality theorem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert space,  $f \in \Gamma_0(\mathcal{H})$ ,  $h \in \Gamma_0(\mathcal{H})$ ,  $g \in \Gamma_0(\mathcal{G})$ ,  $\ell \in \Gamma_0(\mathcal{G})$  such that  $\text{dom } h = \mathcal{H}$  and  $\text{dom } \ell^* = \mathcal{G}$ ,  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

If there exists  $\hat{x} \in \mathcal{H}$  such that  $0 \in \partial f(\hat{x}) + L^*(\partial g \square \partial \ell)(L\hat{x}) + \partial h(\hat{x})$  then

- ▶  $\hat{x}$  is a solution to the primal problem

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + (g \square \ell)(Lx) + h(x)$$

- ▶  $-L^*\hat{v} \in \partial f(\hat{x}) + \partial h(\hat{x})$  and  $L\hat{x} \in \partial g^*(\hat{v}) + \partial \ell^*(\hat{v})$  where  $\hat{v}$  is a solution to the dual problem

$$\underset{v \in \mathcal{G}}{\text{minimize}} \quad (f^* \square h^*)(-L^*v) + g^*(v) + \ell^*(v).$$



# Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

≡ Douglas-Rachford for the dual problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ .  
Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that  $L^*L$  is an isomorphism and let  $\gamma \in ]0, +\infty[$ .  
The Douglas-Rachford iterations to minimize  $f^* \circ (-L^*) + g^*$  are

$$(\forall n \in \mathbb{N}) \quad \begin{cases} v_n = \text{prox}_{\gamma g^*} u_n \\ w_n = \text{prox}_{\gamma f^* \circ (-L^*)} (2v_n - u_n) \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

# Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

≡ Douglas-Rachford for the dual problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ .  
Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that  $L^*L$  is an isomorphism and let  $\gamma \in ]0, +\infty[$ .  
The Douglas-Rachford iterations to minimize  $f^* \circ (-L^*) + g^*$  are

$$(\forall n \in \mathbb{N}) \quad \begin{cases} v_n = u_n - \gamma \operatorname{prox}_{\gamma^{-1}g}(\gamma^{-1}u_n) \\ 2v_n - u_n - w_n \in \gamma \partial(f^* \circ (-L^*)) w_n \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

# Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

≡ Douglas-Rachford for the dual problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ .  
Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that  $L^*L$  is an isomorphism and let  $\gamma \in ]0, +\infty[$ .  
The Douglas-Rachford iterations to minimize  $f^* \circ (-L^*) + g^*$  are

$$(\forall n \in \mathbb{N}) \quad \begin{cases} v_n = u_n - \gamma \operatorname{prox}_{\gamma^{-1}g}(\gamma^{-1}u_n) \\ 2v_n - u_n - w_n \in -\gamma L \circ \partial f^* \circ (-L^*)w_n \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

# Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

≡ Douglas-Rachford for the dual problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ .  
 Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that  $L^*L$  is an isomorphism and let  $\gamma \in ]0, +\infty[$ .  
 The Douglas-Rachford iterations to minimize  $f^* \circ (-L^*) + g^*$  are

$$(\forall n \in \mathbb{N}) \quad \begin{cases} v_n = u_n - \gamma \text{prox}_{\gamma^{-1}g}(\gamma^{-1}u_n) \\ 2v_n - u_n - w_n \in -\gamma L \circ \underbrace{\partial f^* \circ (-L^*)}_{x_n \in} w_n \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

# Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

≡ Douglas-Rachford for the dual problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ . Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that  $L^*L$  is an isomorphism and let  $\gamma \in ]0, +\infty[$ . The Douglas-Rachford iterations to minimize  $f^* \circ (-L^*) + g^*$  are

$$(\forall n \in \mathbb{N}) \quad \begin{cases} v_n = u_n - \gamma \operatorname{prox}_{\gamma^{-1}g}(\gamma^{-1}u_n) \\ 2v_n - u_n - w_n = -\gamma Lx_n \\ x_n \in \partial f^*(-L^*w_n) \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

# Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

≡ Douglas-Rachford for the dual problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ . Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that  $L^*L$  is an isomorphism and let  $\gamma \in ]0, +\infty[$ . The Douglas-Rachford iterations to minimize  $f^* \circ (-L^*) + g^*$  are

$$(\forall n \in \mathbb{N}) \quad \begin{cases} v_n = u_n - \gamma \operatorname{prox}_{\gamma^{-1}g}(\gamma^{-1}u_n) \\ 2v_n - u_n - w_n = -\gamma Lx_n \\ -L^*w_n \in \partial f(x_n) \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

# Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

≡ Douglas-Rachford for the dual problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ .  
Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that  $L^*L$  is an isomorphism and let  $\gamma \in ]0, +\infty[$ .  
The Douglas-Rachford iterations to minimize  $f^* \circ (-L^*) + g^*$  are

$$(\forall n \in \mathbb{N}) \quad \begin{cases} v_n = u_n - \gamma \text{prox}_{\gamma^{-1}g}(\gamma^{-1}u_n) \\ L^*(u_n - 2v_n - \gamma Lx_n) \in \partial f(x_n) \\ u_{n+1} = u_n + w_n - v_n \end{cases}$$

using  $y_n = \gamma^{-1}(u_n - v_n)$  and  $z_n = \gamma^{-1}v_n$

# Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

≡ Douglas-Rachford for the dual problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ . Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that  $L^*L$  is an isomorphism and let  $\gamma \in ]0, +\infty[$ . The Douglas-Rachford iterations to minimize  $f^* \circ (-L^*) + g^*$  are

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \text{prox}_{\gamma^{-1}g}(\gamma^{-1}u_n) \\ L^*(y_n - z_n - Lx_n) \in \frac{1}{\gamma}\partial f(x_n) \\ u_{n+1} = \gamma z_n + \gamma Lx_n \end{cases}$$



## Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

≡ Douglas-Rachford for the dual problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ .  
Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that  $L^*L$  is an isomorphism and let  $\gamma \in ]0, +\infty[$ .  
The Douglas-Rachford iterations to minimize  $f^* \circ (-L^*) + g^*$  are

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \text{prox}_{\gamma^{-1}g}(\gamma^{-1}u_n) \\ x_n = \underset{x \in \mathcal{H}}{\text{argmin}} \frac{1}{2} \|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ u_{n+1} = \gamma z_n + \gamma Lx_n \end{cases}$$

# Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

≡ Douglas-Rachford for the dual problem

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ .  
Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that  $L^*L$  is an isomorphism and let  $\gamma \in ]0, +\infty[$ .  
The Douglas-Rachford iterations to minimize  $f^* \circ (-L^*) + g^*$  are

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_n = \operatorname{argmin}_{x \in \mathcal{H}} \frac{1}{2} \|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ s_n = Lx_n \\ y_{n+1} = \operatorname{prox}_{\frac{g}{\gamma}}(z_n + s_n) \\ z_{n+1} = z_n + s_n - y_{n+1}. \end{cases}$$

## Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

⇒ Lagrangian interpretation

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + g(Lx) \quad \Leftrightarrow \quad \underset{\substack{x \in \mathcal{H}, y \in \mathcal{G} \\ Lx=y}}{\text{minimize}} \quad f(x) + g(y)$$

Lagrange function:

$$(\forall (x, y, v) \in \mathcal{H} \times \mathcal{G}^2) \quad \mathcal{L}(x, y, z) = f(x) + g(y) + \langle v \mid Lx - y \rangle$$

where  $v \in \mathcal{G}$  denotes the Lagrange multiplier.

## Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

⇒ Lagrangian interpretation

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + g(Lx) \quad \Leftrightarrow \quad \underset{\substack{x \in \mathcal{H}, y \in \mathcal{G} \\ Lx=y}}{\text{minimize}} \quad f(x) + g(y)$$

Augmented Lagrange function: let  $\gamma \in ]0, +\infty[$ , we define

$$(\forall (x, y, z) \in \mathcal{H} \times \mathcal{G}^2) \quad \tilde{\mathcal{L}}(x, y, z) = f(x) + g(y) + \gamma \langle z \mid Lx - y \rangle + \frac{\gamma}{2} \|Lx - y\|^2$$

$$\Rightarrow \begin{cases} \partial_x \tilde{\mathcal{L}}(x, y, z) = \partial f(x) + \gamma L^* z + \gamma L^*(Lx - y) \\ \partial_y \tilde{\mathcal{L}}(x, y, z) = \partial g(y) - \gamma z + \gamma(y - Lx) \\ \partial_z \tilde{\mathcal{L}}(x, y, z) = Lx - y. \end{cases}$$

# Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

⇒ Lagrangian interpretation

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + g(Lx) \quad \Leftrightarrow \quad \underset{\substack{x \in \mathcal{H}, y \in \mathcal{G} \\ Lx=y}}{\text{minimize}} \quad f(x) + g(y)$$

Augmented Lagrange function:

Thus  $(0, 0, 0) \in \partial_x \mathcal{L}(\hat{x}, \hat{y}, \hat{z}) \times \partial_y \mathcal{L}(\hat{x}, \hat{y}, \hat{z}) \times \partial_z \mathcal{L}(\hat{x}, \hat{y}, \hat{z})$

$$\Leftrightarrow \begin{cases} -\gamma L^* \hat{z} \in \partial f(\hat{x}) \\ \gamma \hat{z} \in \partial g(L\hat{x}) \Leftrightarrow L\hat{x} \in \partial g^*(\gamma \hat{z}) \\ \hat{y} = L\hat{x}. \end{cases}$$

## Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

⇒ Lagrangian interpretation

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + g(Lx) \quad \Leftrightarrow \quad \underset{\substack{x \in \mathcal{H}, y \in \mathcal{G} \\ Lx=y}}{\text{minimize}} \quad f(x) + g(y)$$

Augmented Lagrange function: To sum up, if

$$(\forall (x, y, z) \in \mathcal{H} \times \mathcal{G}^2) \quad \tilde{\mathcal{L}}(x, y, z) = f(x) + g(y) + \gamma \langle z \mid Lx - y \rangle + \frac{\gamma}{2} \|Lx - y\|^2$$

then  $(\hat{x}, \hat{y}, \hat{z})$  critical point of  $\tilde{\mathcal{L}} \Rightarrow \hat{y} = L\hat{x}$ ,  $\hat{x}$  solution to the primal problem and  $\gamma\hat{z}$  solution to the dual problem.

Moreover,  $(\hat{x}, \hat{y}, \hat{z})$  is a critical point of  $\tilde{\mathcal{L}} \Leftrightarrow (\hat{x}, \hat{y}, \hat{z})$  saddle point of  $\tilde{\mathcal{L}}$ , i.e.  $\tilde{\mathcal{L}}(\hat{x}, \hat{y}, z) \leq \tilde{\mathcal{L}}(\hat{x}, \hat{y}, \hat{z}) \leq \tilde{\mathcal{L}}(x, \hat{y}, \hat{z})$ .

# Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

⇒ Lagrangian interpretation

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + g(Lx) \quad \Leftrightarrow \quad \underset{\substack{x \in \mathcal{H}, y \in \mathcal{G} \\ Lx=y}}{\text{minimize}} \quad f(x) + g(y)$$

Augmented Lagrange function: Algorithm to find  $(\hat{x}, \hat{y}, \hat{z})$ :

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\text{argmin}} \quad \tilde{\mathcal{L}}(x, y_n, z_n) \\ y_{n+1} = \underset{y \in \mathcal{G}}{\text{argmin}} \quad \tilde{\mathcal{L}}(x_n, y, z_n) \\ z_{n+1} \text{ such that } \tilde{\mathcal{L}}(x_n, y_{n+1}, z_{n+1}) \geq \tilde{\mathcal{L}}(x_n, y_{n+1}, z_n). \end{cases}$$

# Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

⇒ Lagrangian interpretation

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + g(Lx) \quad \Leftrightarrow \quad \underset{\substack{x \in \mathcal{H}, y \in \mathcal{G} \\ Lx=y}}{\text{minimize}} \quad f(x) + g(y)$$

Augmented Lagrange function: Algorithm to find  $(\hat{x}, \hat{y}, \hat{z})$ :

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\operatorname{argmin}} \quad f(x) + \gamma \langle z_n | Lx - y_n \rangle + \frac{\gamma}{2} \|Lx - y_n\|^2 \\ y_{n+1} = \underset{y \in \mathcal{G}}{\operatorname{argmin}} \quad g(y) + \gamma \langle z_n | Lx_n - y \rangle + \frac{\gamma}{2} \|Lx_n - y\|^2 \\ z_{n+1} = z_n + \frac{1}{\gamma} \nabla_z \tilde{\mathcal{L}}(x_n, y_{n+1}, z_n). \end{cases}$$



## Augmented Lagrangian method

ADMM algorithm (*Alternating-direction method of multipliers*)

⇒ Lagrangian interpretation

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + g(Lx) \quad \Leftrightarrow \quad \underset{\substack{x \in \mathcal{H}, y \in \mathcal{G} \\ Lx=y}}{\text{minimize}} \quad f(x) + g(y)$$

Augmented Lagrange function: Algorithm to find  $(\hat{x}, \hat{y}, \hat{z})$ :

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\text{argmin}} \quad \frac{1}{2} \|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ y_{n+1} = \underset{\gamma}{\text{prox}}_{\mathcal{G}} (z_n + Lx_n) \\ z_{n+1} = z_n + Lx_n - y_{n+1}. \end{cases}$$

## Augmented Lagrange method

ADMM algorithm (*Alternating-direction method of multipliers*)

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  et  $g \in \Gamma_0(\mathcal{G})$ .

Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that  $L^*L$  is an isomorphism and let  $\gamma \in ]0, +\infty[$ .

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_n = \operatorname{argmin}_{x \in \mathcal{H}} \frac{1}{2} \|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ s_n = Lx_n \\ y_{n+1} = \operatorname{prox}_{\frac{g}{\gamma}}(z_n + s_n) \\ z_{n+1} = z_n + s_n - y_{n+1}. \end{cases}$$

## Augmented Lagrange method

### ADMM algorithm (*Alternating-direction method of multipliers*)

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $f \in \Gamma_0(\mathcal{H})$  et  $g \in \Gamma_0(\mathcal{G})$ .

Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$  such that  $L^*L$  is an isomorphism and let  $\gamma \in ]0, +\infty[$ .

We assume that  $\text{int}(\text{dom } g) \cap L(\text{dom } f) \neq \emptyset$  or  $\text{dom } g \cap \text{int}(L(\text{dom } f)) \neq \emptyset$  and that  $\text{Argmin}(f + g \circ L) \neq \emptyset$ . Let

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_n = \underset{x \in \mathcal{H}}{\text{argmin}} \frac{1}{2} \|Lx - y_n + z_n\|^2 + \frac{1}{\gamma} f(x) \\ s_n = Lx_n \\ y_{n+1} = \text{prox}_{\frac{g}{\gamma}}(z_n + s_n) \\ z_{n+1} = z_n + s_n - y_{n+1}. \end{cases}$$

We have:

- ▶  $x_n \rightarrow \hat{x}$  where  $\hat{x} \in \text{Argmin}(f + g \circ L)$
- ▶  $\gamma z_n \rightarrow \hat{v}$  where  $\hat{v} \in \text{Argmin}(f^* \circ (-L^*) + g^*)$ .

# Augmented Lagrangian method

## SDMM algorithm (*Simultaneous-direction method of multipliers*)

Let  $\mathcal{H}$  and  $\mathcal{G}_1, \dots, \mathcal{G}_m$  be Hilbert spaces.

Let  $(\forall i \in \{1, \dots, m\})$   $g_i \in \Gamma_0(\mathcal{G}_i)$  and  $L_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$ . Let  $\gamma \in ]0, +\infty[$ .

We assume that  $\sum_{i=1}^m L_i^* L_i$  is an isomorphism

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_n = \left( \sum_{i=1}^m L_i^* L_i \right)^{-1} \sum_{i=1}^m L_i^* (y_{n,i} - z_{n,i}) \\ s_{n,i} = L_i x_n, \quad i \in \{1, \dots, m\} \\ y_{n+1,i} = \text{prox}_{\frac{g_i}{\gamma}}(z_{n,i} + s_{n,i}), \quad i \in \{1, \dots, m\} \\ z_{n+1,i} = z_{n,i} + s_{n,i} - y_{n+1,i}, \quad i \in \{1, \dots, m\}. \end{cases}$$

## Augmented Lagrangian method

### SDMM algorithm (*Simultaneous-direction method of multipliers*)

Let  $\mathcal{H}$  and  $\mathcal{G}_1, \dots, \mathcal{G}_m$  be Hilbert spaces.

Let  $(\forall i \in \{1, \dots, m\})$   $g_i \in \Gamma_0(\mathcal{G}_i)$  and  $L_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$ . Let  $\gamma \in ]0, +\infty[$ .

We assume that  $\sum_{i=1}^m L_i^* L_i$  is an isomorphism and there exists  $\tilde{x} \in \mathcal{H}$  such that  $(\forall i \in \{1, \dots, m\})$   $L_i \tilde{x} \in \text{int}(\text{dom } g_i)$ . Let

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_n = \left( \sum_{i=1}^m L_i^* L_i \right)^{-1} \sum_{i=1}^m L_i^* (y_{n,i} - z_{n,i}) \\ s_{n,i} = L_i x_n, \quad i \in \{1, \dots, m\} \\ y_{n+1,i} = \text{prox}_{\frac{g_i}{\gamma}}(z_{n,i} + s_{n,i}), \quad i \in \{1, \dots, m\} \\ z_{n+1,i} = z_{n,i} + s_{n,i} - y_{n+1,i}, \quad i \in \{1, \dots, m\}. \end{cases}$$

We have:

- ▶  $x_n \rightarrow \hat{x}$  where  $\hat{x} \in \text{Argmin}(\sum_{i=1}^m g_i \circ L_i)$
- ▶  $\gamma z_n = \gamma (z_{n,i})_{1 \leq i \leq m} \rightarrow \hat{v}$  where  $\hat{v} = (\hat{v}_i)_{1 \leq i \leq m} \in \underset{\substack{v=(v_i)_{1 \leq i \leq m} \\ \sum_{i=1}^m L_i^* v_i = 0}}{\text{Argmin}} (\sum_{i=1}^m g_i^*(v_i))$ .

# Primal-dual method

## FB-based PD algorithm (*Condat-Vũ-...*)

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and  $B: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be two maximally monotone operators.

Let  $C: \mathcal{H} \rightarrow \mathcal{H}$  be a  $\mu$ -cocoercive operator with  $\mu \in ]0, +\infty[$ .

Let  $D: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be a  $\nu$ -strongly monotone operator with  $\nu \in ]0, +\infty[$ .

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_n = J_{\tau A}(x_n - \tau(Cx_n + L^*v_n)) \\ q_n = J_{\sigma B^{-1}}(v_n + \sigma(L(2p_n - x_n) - D^{-1}v_n)) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n((p_n, q_n) - (x_n, v_n)). \end{cases}$$

## Primal-dual method

### FB-based PD algorithm (*Condat-Vũ-...*)

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and  $B: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be two maximally monotone operators.

Let  $C: \mathcal{H} \rightarrow \mathcal{H}$  be a  $\mu$ -cocoercive operator with  $\mu \in ]0, +\infty[$ .

Let  $D: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be a  $\nu$ -strongly monotone operator with  $\nu \in ]0, +\infty[$ .

Let  $\beta = \min\{\mu, \nu\}$ ,  $\tau \in ]0, +\infty[$ ,  $\sigma \in ]0, +\infty[$ ,  $\rho = \min\{\tau^{-1}, \sigma^{-1}\}(1 - \sqrt{\tau\sigma}\|L\|)$  and  $\delta = \min\{1, \rho\beta\} + 1/2$ . We assume that  $2\rho\beta > 1$ .

Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $]0, \delta[$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$ .

We assume that  $\text{zer}(A + C + L^*(B \square D)L) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$ ,  $v_0 \in \mathcal{G}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_n = J_{\tau A}(x_n - \tau(Cx_n + L^*v_n)) \\ q_n = J_{\sigma B^{-1}}(v_n + \sigma(L(2p_n - x_n) - D^{-1}v_n)) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n((p_n, q_n) - (x_n, v_n)). \end{cases}$$

We have:  $x_n \rightarrow \hat{x} \in \text{zer}(A + C + L^*(B \square D)L)$

and  $v_n \rightarrow \hat{v} \in \text{zer}((-L)(A^{-1} \square C^{-1})(-L^*) + B^{-1} + D^{-1})$ .

## Primal-dual method

Proof:

$(\widehat{x}, \widehat{v})$  Kuhn-Tucker point iff  $(\widehat{x}, \widehat{v}) \in \text{zer}(P + Q)$  where

- ▶  $P$  maximally monotone such that  $P = M + S$  with

$$M: (x, v) \mapsto (Ax, B^{-1}v)$$

$$S: (x, v) \mapsto \begin{bmatrix} 0 & L^* \\ -L & 0 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix}$$

- ▶  $Q: (x, v) \mapsto (Cx, D^{-1}v)$   $\beta$ -cocoercive with  $\beta = \min\{\mu, \nu\}$

$\Rightarrow$  Forward-Backward algorithm.



# Primal-dual optimization algorithm

## FB-based PD algorithm

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ .

Let  $h \in \Gamma_0(\mathcal{H})$  having a  $1/\mu$ -Lipschitzian gradient,  $\ell \in \Gamma_0(\mathcal{G})$   $\nu$ -strongly convex

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_n = \text{prox}_{\tau f}(x_n - \tau(\nabla h(x_n) + L^* v_n)) \\ q_n = \text{prox}_{\sigma g^*}(v_n + \sigma(L(2p_n - x_n) - \nabla \ell^*(v_n))) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n((p_n, q_n) - (x_n, v_n)). \end{cases}$$

# Primal-dual optimization algorithm

## FB-based PD algorithm

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ .

Let  $h \in \Gamma_0(\mathcal{H})$  having a  $1/\mu$ -Lipschitzian gradient,  $\ell \in \Gamma_0(\mathcal{G})$   $\nu$ -strongly convex and  $\beta = \min\{\mu, \nu\}$ .

Let  $\tau \in ]0, +\infty[$ ,  $\sigma \in ]0, +\infty[$ ,  $\rho = \min\{\tau^{-1}, \sigma^{-1}\}(1 - \sqrt{\tau\sigma}\|L\|)$  et  $\delta = \min\{1, \rho\beta\} + 1/2$ . We assume that  $2\rho\beta > 1$ .

Let  $(\lambda_n)_{n \in \mathbb{N}}$  a sequence in  $]0, \delta[$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$ .

We assume that  $\text{zer}(\partial f + \nabla h + L^*(\partial g \square \partial \ell)L) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$ ,  $v_0 \in \mathcal{G}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_n = \text{prox}_{\tau f}(x_n - \tau(\nabla h(x_n) + L^*v_n)) \\ q_n = \text{prox}_{\sigma g^*}(v_n + \sigma(L(2p_n - x_n) - \nabla \ell^*(v_n))) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n((p_n, q_n) - (x_n, v_n)). \end{cases}$$

We have:

- ▶  $x_n \rightarrow \hat{x} \in \text{Argmin}(f + h + (g \square \ell) \circ L)$
- ▶  $v_n \rightarrow \hat{v} \in \text{Argmin}((f^* \square h^*) \circ (-L^*) + g^* + \ell^*)$

# Primal-dual optimization algorithm

## FB-based PD algorithm

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_n = \text{prox}_{\tau f}(x_n - \tau(\nabla h(x_n) + L^* v_n)) \\ q_n = \text{prox}_{\sigma g^*}(v_n + \sigma(L(2p_n - x_n) - \nabla \ell^*(v_n))) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n((p_n, q_n) - (x_n, v_n)). \end{cases}$$

► Remark:

- \* No operator inversion.
- \* Allow the use of proximable or/and differentiable functions.

# Primal-dual optimization algorithm

FB-based PD algorithm  $\Rightarrow$  CP algorithm

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_n = \text{prox}_{\tau f}(x_n - \tau(\nabla h(x_n) + L^* v_n)) \\ q_n = \text{prox}_{\sigma g^*}(v_n + \sigma(L(2p_n - x_n) - \nabla \ell^*(v_n))) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n((p_n, q_n) - (x_n, v_n)). \end{cases}$$

► Remark:

\* When  $h = 0$ ,  $\ell = \iota_{\{0\}}$ ,  $\lambda_n \equiv 1$  and  $\sigma\tau\|L\|^2 < 1$ , this yields the Chambolle-Pock algorithm.

# Primal-dual optimization algorithm

FB-based PD algorithm  $\Rightarrow$  CP algorithm

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_{n+1} = \text{prox}_{\tau f}(x_n - \tau L^* v_n) \\ y_n = 2x_{n+1} - x_n \\ v_{n+1} = \text{prox}_{\sigma g^*}(v_n + \sigma L y_n). \end{cases}$$

► Remark:

\* When  $h = 0$ ,  $\ell = \iota_{\{0\}}$ ,  $\lambda_n \equiv 1$  and  $\sigma\tau\|L\|^2 < 1$ , this yields the Chambolle-Pock algorithm.

# Primal-dual optimization algorithm

FB-based PD algorithm  $\Rightarrow$  FB algorithm

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_n = \text{prox}_{\tau f}(x_n - \tau(\nabla h(x_n) + L^* v_n)) \\ q_n = \text{prox}_{\sigma g^*}(v_n + \sigma(L(2p_n - x_n) - \nabla \ell^*(v_n))) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n((p_n, q_n) - (x_n, v_n)). \end{cases}$$

► Remark:

\* When  $h = 0$ ,  $\ell = \iota_{\{0\}}$ ,  $\lambda_n \equiv 1$  and  $\sigma\tau\|L\|^2 < 1$ , this yields the Chambolle-Pock algorithm.

\* When  $g = 0$ ,  $\ell = \iota_{\{0\}}$  and  $L = 0$ , this yields the forward-backward algorithm.

# Primal-dual optimization algorithm

FB-based PD algorithm  $\Rightarrow$  FB algorithm

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_n = \text{prox}_{\tau f}(x_n - \tau \nabla h(x_n)) \\ x_{n+1} = x_n + \lambda_n (p_n - x_n). \end{cases}$$

► Remark:

\* When  $h = 0$ ,  $\ell = \iota_{\{0\}}$ ,  $\lambda_n \equiv 1$  and  $\sigma\tau\|L\|^2 < 1$ , this yields the Chambolle-Pock algorithm.

\* When  $g = 0$ ,  $\ell = \iota_{\{0\}}$  and  $L = 0$ , this yields the forward-backward algorithm.

# Primal-dual optimization algorithm

FB-based PD algorithm  $\Rightarrow$  DR algorithm

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_n = \text{prox}_{\tau f}(x_n - \tau(\nabla h(x_n) + L^* v_n)) \\ q_n = \text{prox}_{\sigma g^*}(v_n + \sigma(L(2p_n - x_n) - \nabla \ell^*(v_n))) \\ (x_{n+1}, v_{n+1}) = (x_n, v_n) + \lambda_n((p_n, q_n) - (x_n, v_n)). \end{cases}$$

► Remark:

\* When  $h = 0$ ,  $\ell = \iota_{\{0\}}$ ,  $\lambda_n \equiv 1$  and  $\sigma\tau\|L\|^2 < 1$ , this yields the Chambolle-Pock algorithm.

\* When  $g = 0$ ,  $\ell = \iota_{\{0\}}$  and  $L = 0$ , this yields the forward-backward algorithm.

\* In the limit case when  $h = 0$ ,  $\ell = \iota_{\{0\}}$ ,  $\lambda_n \equiv 1$ ,  $L = \text{Id}$  and  $\sigma = 1/\tau$ , this yields the Douglas-Rachford algorithm.



# Primal-dual optimization algorithm

FB-based PD algorithm  $\Rightarrow$  DR algorithm

$$(\forall n \in \mathbb{N}) \quad \begin{cases} x_{n+1} = \text{prox}_{\tau f}(x_n - \tau v_n) \\ s_n = \text{prox}_{\tau g}(2x_{n+1} - (x_n - \tau v_n)) \\ x_{n+1} - \tau v_{n+1} = (x_n - \tau v_n) + s_n - x_{n+1} \end{cases}$$

► Remark:

\* When  $h = 0$ ,  $\ell = \iota_{\{0\}}$ ,  $\lambda_n \equiv 1$  and  $\sigma\tau\|L\|^2 < 1$ , this yields the Chambolle-Pock algorithm.

\* When  $g = 0$ ,  $\ell = \iota_{\{0\}}$  and  $L = 0$ , this yields the forward-backward algorithm.

\* In the limit case when  $h = 0$ ,  $\ell = \iota_{\{0\}}$ ,  $\lambda_n \equiv 1$ ,  $L = \text{Id}$  and  $\sigma = 1/\tau$ , this yields the Douglas-Rachford algorithm.

## FB-based PD algorithm

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \| [W_1^* H_1^* \dots W_J^* H_J^*]^* x \|_{2,1} + \iota_C(x)$$

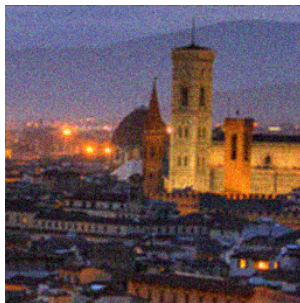
$$\text{with } \begin{cases} \eta > 0 \\ H_1, \dots, H_J \in \mathbb{R}^{N \times N} \text{ (e.g. operators associated with filters)} \\ W_1, \dots, W_J \in \mathbb{R}^{N \times N} \text{ (e.g. weights)} \\ C = [0, 255]^N \end{cases}$$

- ▶ PPXA+ :  $(2\text{Id} + W_1^* H_1^* H_1 W_1 + \dots + W_J^* H_J^* H_J W_J)^{-1}$  complicated
- ▶ FB-based PD algorithm : implementation without inversion

## FB-based PD algorithm

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \| [W_1^* H_1^* \dots W_J^* H_J^*]^* x \|_{2,1} + \iota_C(x)$$



Degraded image  $z$

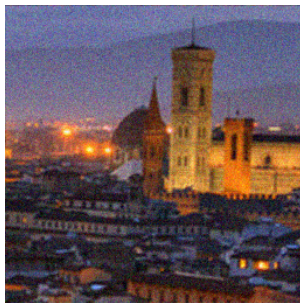


Restored image  
[PD - NLTV]

## FB-based PD algorithm

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \| [W_1^* H_1^* \dots W_J^* H_J^*]^* x \|_{2,1} + \iota_C(x)$$



Degraded image  $z$

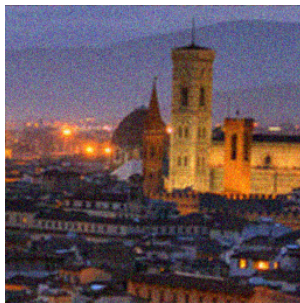


Restored image  
[FB - DTT]

## FB-based PD algorithm

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \| [W_1^* H_1^* \dots W_J^* H_J^*]^* x \|_{2,1} + \iota_C(x)$$



Degraded image  $z$



Restored image  
[PPXA - TV]

## FB-based PD algorithm

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \| [W_1^* H_1^* \dots W_J^* H_J^*]^* x \|_{2,1} + \iota_C(x)$$



Degraded image  $z$



Restored image  
[DR - DWT]

# Primal-dual method

## Parallel FB-based PD algorithm

Let  $\mathcal{H}$  and  $\mathcal{G}_1, \dots, \mathcal{G}_m$  be Hilbert spaces. For every  $i \in \{1, \dots, m\}$ , let  $L_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$ .

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and, for every  $i \in \{1, \dots, m\}$ ,  $B_i: \mathcal{G}_i \rightarrow 2^{\mathcal{G}_i}$  maximally monotone operators.

Let  $C: \mathcal{H} \rightarrow \mathcal{H}$  be a  $\mu$ -cocoercive operator where  $\mu \in ]0, +\infty[$ ,

$(\forall i \in \{1, \dots, m\}) D_i: \mathcal{G}_i \rightarrow 2^{\mathcal{G}_i}$  be  $\nu_i$ -strongly monotone where  $\nu_i \in ]0, +\infty[$ .

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_n = J_{\tau A}(x_n - \tau(Cx_n + \sum_{i=1}^m L_i^* v_{n,i})) \\ q_{n,i} = J_{\sigma B_i^{-1}}(v_{n,i} + \sigma(L_i(2p_n - x_n) - D_i^{-1}v_{n,i})), \quad i \in \{1, \dots, m\} \\ x_{n+1} = x_n + \lambda_n(p_n - x_n) \\ v_{n+1,i} = v_{n,i} + \lambda_n(q_{n,i} - v_{n,i}), \quad i \in \{1, \dots, m\}. \end{cases}$$

# Primal-dual method

## Parallel FB-based PD algorithm

Let  $\mathcal{H}$  and  $\mathcal{G}_1, \dots, \mathcal{G}_m$  be Hilbert spaces. For every  $i \in \{1, \dots, m\}$ , let  $L_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$ .

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and, for every  $i \in \{1, \dots, m\}$ ,  $B_i: \mathcal{G}_i \rightarrow 2^{\mathcal{G}_i}$  maximally monotone operators.

Let  $C: \mathcal{H} \rightarrow \mathcal{H}$  be a  $\mu$ -cocoercive operator where  $\mu \in ]0, +\infty[$ ,

( $\forall i \in \{1, \dots, m\}$ )  $D_i: \mathcal{G}_i \rightarrow 2^{\mathcal{G}_i}$  be  $\nu_i$ -strongly monotone where  $\nu_i \in ]0, +\infty[$ .

Let  $\beta = \min\{\mu, \nu_1, \dots, \nu_m\}$ ,  $\tau > 0$ ,  $\sigma > 0$ ,  $\rho = \min\{\tau^{-1}, \sigma^{-1}\}(1 - \sqrt{\tau\sigma \sum_{i=1}^m \|L_i\|^2})$

and  $\delta = \min\{1, \rho\beta\} + 1/2$ . We assume that  $2\rho\beta > 1$ .

Let  $(\lambda_n)_{n \in \mathbb{N}}$  be a sequence in  $]0, \delta[$  such that  $\sum_{n \in \mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$ .

We assume that  $\text{zer}(A + C + \sum_{i=1}^m L_i^*(B_i \square D_i)L_i) \neq \emptyset$ .

Let  $x_0 \in \mathcal{H}$ ,  $(v_{0,1}, \dots, v_{0,m}) \in \mathcal{G}_1 \times \dots \times \mathcal{G}_m$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} p_n = J_{\tau A}(x_n - \tau(Cx_n + \sum_{i=1}^m L_i^* v_{n,i})) \\ q_{n,i} = J_{\sigma B_i^{-1}}(v_{n,i} + \sigma(L_i(2p_n - x_n) - D_i^{-1}v_{n,i})), \quad i \in \{1, \dots, m\} \\ x_{n+1} = x_n + \lambda_n(p_n - x_n) \\ v_{n+1,i} = v_{n,i} + \lambda_n(q_{n,i} - v_{n,i}), \quad i \in \{1, \dots, m\}. \end{cases}$$

We have:  $x_n \rightarrow \hat{x} \in \text{zer}(A + C + \sum_{i=1}^m L_i^*(B_i \square D_i)L_i)$

and  $v_{n,i} \rightarrow \hat{v}_i \in \text{zer}((-L_i)(A^{-1} \square C^{-1})(-L_i^*) + B_i^{-1} + D_i^{-1}), \quad i \in \{1, \dots, m\}$ .



# Primal-dual method

M+LFBF algorithm (Monotone+Lipschitz *Forward Backward Forward*)

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and  $B: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be two maximally monotone operators.

Let  $C: \mathcal{H} \rightarrow \mathcal{H}$  be a monotone operator  $\mu^{-1}$ -Lipschitzian with  $\mu \in ]0, +\infty[$ ,

$D: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be  $\nu$ -strongly monotone where  $\nu \in ]0, +\infty[$  and  $\beta^{-1} = \max\{\mu^{-1}, \nu^{-1}\} + \|L\|$ .

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{1,n} = x_n - \gamma_n(Cx_n + L^*v_n) \\ y_{2,n} = v_n + \gamma_n(Lx_n - D^{-1}v_n) \\ p_{1,n} = J_{\gamma_n A} y_{1,n}, p_{2,n} = J_{\gamma_n B^{-1}} y_{2,n} \\ q_{1,n} = p_{1,n} - \gamma_n(Cp_{1,n} + L^*p_{2,n}) \\ q_{2,n} = p_{2,n} + \gamma_n(Lp_{1,n} - D^{-1}p_{2,n}) \\ (x_{n+1}, v_{n+1}) = (x_n - y_{1,n} + q_{1,n}, v_n - y_{2,n} + q_{2,n}). \end{cases}$$

# Primal-dual method

## M+LFBF algorithm (Monotone+Lipschitz *Forward Backward Forward*)

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

Let  $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$  and  $B: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be two maximally monotone operators.

Let  $C: \mathcal{H} \rightarrow \mathcal{H}$  be a monotone operator  $\mu^{-1}$ -Lipschitzian with  $\mu \in ]0, +\infty[$ ,

$D: \mathcal{G} \rightarrow 2^{\mathcal{G}}$  be  $\nu$ -strongly monotone where  $\nu \in ]0, +\infty[$  and  $\beta^{-1} = \max\{\mu^{-1}, \nu^{-1}\} + \|L\|$ .

Let  $\varepsilon \in ]0, \beta/(1 + \beta)[$  and  $(\gamma_n)_{n \in \mathbb{N}}$  a sequence in  $[\varepsilon, (1 - \varepsilon)\beta]$ .

We assume  $\text{zer}(A + C + L^*(B \square D)L) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$ ,  $v_0 \in \mathcal{G}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{1,n} = x_n - \gamma_n(Cx_n + L^*v_n) \\ y_{2,n} = v_n + \gamma_n(Lx_n - D^{-1}v_n) \\ p_{1,n} = J_{\gamma_n A} y_{1,n}, p_{2,n} = J_{\gamma_n B^{-1}} y_{2,n} \\ q_{1,n} = p_{1,n} - \gamma_n(Cp_{1,n} + L^*p_{2,n}) \\ q_{2,n} = p_{2,n} + \gamma_n(Lp_{1,n} - D^{-1}p_{2,n}) \\ (x_{n+1}, v_{n+1}) = (x_n - y_{1,n} + q_{1,n}, v_n - y_{2,n} + q_{2,n}). \end{cases}$$

We have:  $x_n \rightarrow \hat{x}$  and  $p_{1,n} \rightarrow \hat{x}$  where  $\hat{x} \in \text{zer}(A + C + L^*(B \square D)L)$

and  $v_n \rightarrow \hat{v}$  and  $p_{2,n} \rightarrow \hat{v}$  where  $\hat{v} \in \text{zer}((-L)(A^{-1} \square C^{-1})(-L^*) + B^{-1} + D^{-1})$ .

# Primal-dual optimization method

M+LFBF algorithm (Monotone+Lipschitz *Forward Backward Forward*)

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ .

Let  $h \in \Gamma_0(\mathcal{H})$  having a  $\mu^{-1}$ -Lipschitz gradient with  $\mu \in ]0, +\infty[$ ,

$\ell \in \Gamma_0(\mathcal{G})$   $\nu$ -strongly convex

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{1,n} = x_n - \gamma_n(\nabla h(x_n) + L^* v_n) \\ y_{2,n} = v_n + \gamma_n(Lx_n - \nabla \ell^*(v_n)) \\ p_{1,n} = \text{prox}_{\gamma_n f} y_{1,n}; p_{2,n} = \text{prox}_{\gamma_n g^*} y_{2,n} \\ q_{1,n} = p_{1,n} - \gamma_n(\nabla h(p_{1,n}) + L^* p_{2,n}) \\ q_{2,n} = p_{2,n} + \gamma_n(Lp_{1,n} - \nabla \ell^*(p_{2,n})) \\ (x_{n+1}, v_{n+1}) = (x_n - y_{1,n} + q_{1,n}, v_n - y_{2,n} + q_{2,n}). \end{cases}$$

# Primal-dual optimization method

## M+LFBF algorithm (Monotone+Lipschitz Forward Backward Forward)

Let  $\mathcal{H}$  and  $\mathcal{G}$  be two Hilbert spaces. Let  $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ .

Let  $f \in \Gamma_0(\mathcal{H})$  and  $g \in \Gamma_0(\mathcal{G})$ .

Let  $h \in \Gamma_0(\mathcal{H})$  having a  $\mu^{-1}$ -Lipschitz gradient with  $\mu \in ]0, +\infty[$ ,

$\ell \in \Gamma_0(\mathcal{G})$   $\nu$ -strongly convex where  $\nu \in ]0, +\infty[$  and  $\beta^{-1} = \max\{\mu^{-1}, \nu^{-1}\} + \|L\|$ .

Let  $\varepsilon \in ]0, \beta/(1 + \beta)[$  and  $(\gamma_n)_{n \in \mathbb{N}}$  a sequence in  $[\varepsilon, (1 - \varepsilon)\beta]$ .

We assume that  $\text{Argmin}(f + g + (h \square \ell) \circ L) \neq \emptyset$ . Let  $x_0 \in \mathcal{H}$ ,  $v_0 \in \mathcal{G}$  and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_{1,n} = x_n - \gamma_n(\nabla h(x_n) + L^* v_n) \\ y_{2,n} = v_n + \gamma_n(Lx_n - \nabla \ell^*(v_n)) \\ p_{1,n} = \text{prox}_{\gamma_n f} y_{1,n}; p_{2,n} = \text{prox}_{\gamma_n g^*} y_{2,n} \\ q_{1,n} = p_{1,n} - \gamma_n(\nabla h(p_{1,n}) + L^* p_{2,n}) \\ q_{2,n} = p_{2,n} + \gamma_n(Lp_{1,n} - \nabla \ell^*(p_{2,n})) \\ (x_{n+1}, v_{n+1}) = (x_n - y_{1,n} + q_{1,n}, v_n - y_{2,n} + q_{2,n}). \end{cases}$$

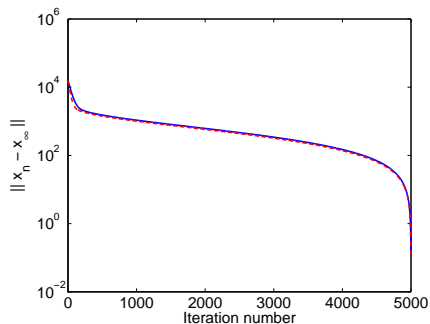
We have:  $x_n \rightarrow \hat{x}$  and  $p_{1,n} \rightarrow \hat{x}$  where  $\hat{x} \in \text{Argmin}(f + h + (g \square h) \circ L)$

and  $v_n \rightarrow \hat{v}$  et  $p_{2,n} \rightarrow \hat{v}$  where  $\hat{v} \in \text{Argmin}((f^* \square h^*) \circ (-L^*) + g^* + \ell^*)$ .

# Primal-dual optimization method

Image restoration : Variational approach

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \|Ax - z\|_2^2 + \eta \| [W_1^* H_1^* \dots W_J^* H_J^*]^* x \|_{2,1} + \iota_C(x)$$



→ FB-based PD (red)

→ M+LFBF (blue)

## Conclusions

- ▶ Flexible framework unifying several problems:
  - ▶ regularized approaches

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + \sum_{i=1}^m g_i(x)$$

where  $f: \mathcal{H} \rightarrow \mathbb{R}$ ,  $(\forall i \in \{1, \dots, m\}) g_i: \mathcal{H} \rightarrow \mathbb{R}$  are convex.

- ▶ feasibility approaches

$$\text{Find } x \in \mathcal{H} \text{ such that } x \in \bigcap_{i=1}^m C_i$$

where  $(\forall i \in \{1, \dots, m\}) C_i$  is a nonempty closed convex subset of  $\mathcal{H}$ .

## Conclusions

- ▶ Flexible framework unifying several problems:

- ▶ sparse problems

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(x) + \sigma_C(x)$$

where  $f \in \Gamma_0(\mathcal{H})$  and  $C$  is a nonempty closed convex subset of  $\mathcal{H}$ .

- ▶ constrained problems

$$\underset{x \in C}{\text{minimize}} \quad g(x)$$

where  $g \in \Gamma_0(\mathcal{H})$ .

## Conclusions

- ▶ Results in infinite dimension (continuous problems).
- ▶ Splitting is the key.  
Parallel methods adapted to multi-core architectures.  
Can be extended to distributed/stochastic methods.
- ▶ Robustness to computational errors.

$$\text{prox}_f x_n \rightarrow \text{prox}_f x_n + e_n$$

where  $(e_n)_{n \in \mathbb{N}}$  is a sequence in  $\mathcal{H}$  such that  $\sum_{n \in \mathbb{N}} \|e_n\| < +\infty$ .

- ▶ Applications to other fields: game theory, PDEs, ...  
Find maximally monotone operators in your favorite area !
- ▶ Extension to the nonconvex case.



## A few references

P.L. Combettes and J.-C. Pesquet, “Proximal splitting methods in signal processing”, in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Springer-Verlag, pp. 185–212, 2010.

F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, “Optimization with sparsity-inducing penalties”, *Foundations and Trends in Machine Learning*, vol. 4, no. 1, pp. 1–106, 2012.

N. Parikh and S. Boyd, “Proximal algorithms”, *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 123–231, 2013.