الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي و البحث العلمي

People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Researche

Ahmed Zabana University of Relizane
Faculty of Science and Technology
Department of Mathematics

To obtain a Master's degree

THESES

Differential Geometry

Titled

Difference of Two Maximal Monotone Operators

Présenté par :

Miss: BESSEDIK Khadidja

Before the jury members:

- **President: Mr Bahri Noureddine** Associate Professor (A), University of Chlef
- **Supervisor: Mr Beddani Abdallah** Associate Professor (A), University of Relizane
- **Examiner: Mrs Kadri Halima** Assistant Professor (B), University of Relizane

Academic Year: 2024/2025

Thanks

I would first like to thank Almighty and Merciful Allah who granted us the health, courage, strength, and patience to accomplish this modest task.

I would like to express my gratitude and sincere thanks to my supervisor, Mr. Abdallah Beddani, for agreeing to supervise this thesis with great patience, diligence, and competence. I sincerely thank him for his advice, guidance, and assistance.

I would like to thank the members of the jury who agreed to evaluate my thesis.

I would like to thank all the professors in the mathematics department at Ahmed Zabana-Relizane University, not to mention my colleagues and friends, as well as all those who participated, directly or indirectly, in the development of this thesis.

I cannot close my thanks without turning to those who are dearest to me, my family, who have played a vital and ongoing role in my success.

THANKS

Dedications

Invocations to ALLAH, for all power.

To my dearest mother,

To my dearest father,

To my brother and sisters,

To my entire family,

To my best friends

And to all those who are dear to me

I dedicate this work.

Khadidja

ABSTRACT

The primary objective of this study is to investigate the properties of the zero of the difference of two maximal monotone operators, with a focus on the existence and uniqueness of solutions.

We explore the mathematical structure of the zero set of such differences and analyze sufficient conditions that guarantee the existence of a unique solution to the associated inclusion problems. Additionally, we examine the zero of the sum of two maximal monotone operators and study the fundamental properties of maximal monotone operators. Furthermore, we analyze the relationship between maximal monotonicity properties and functional analysis techniques used to study this class of operators.

We also discuss practical applications of these results in optimization problems, variational inequalities, and iterative algorithms, contributing to a deeper theoretical and practical understanding of this mathematical framework.

CONTENTS

0.1 General Introduction 8

1 Set-Valued Maps **9**
1.1 First Definitions 9
1.2 Examples 10
1.3 Some Algebraic Operations 13
1.3.1 Sum of Set-Valued Maps 13
1.3.2 Composition of Set-Valued Maps 13
1.3.3 Scalar Multiplication 14
1.3.4 Inverse of a Set-Valued Map 14

2 Maximal Monotone Operators in Hilbert Spaces **15**
2.1 Monotone Operators 15
2.1.1 Strictly Monotone Operator 16
2.1.2 Strong Monotonicity 16
2.2 Concepts of Convex Analysis 17
2.2.1 Convex Sets 17
2.2.2 Convex Cones 18
2.2.3 Normal Cones 19
2.3 Convex, Lower Semi-Continuous, Absolutely Continuous Functions 19
2.3.1 Convex function : 19
2.3.2 Lower Semi-Continuous : 21
2.3.3 Absolutely Continuous Functions: 23
2.4 Proximal and Subdifferential of a Convex Function 24
2.4.1 Subdifferential of a Convex Function: 24
2.4.2 Proximal Operator of a Convex Function 25
2.5 Contraction Operators on Hilbert Spaces 27
2.6 Maximal Monotone Operators 27
2.7 Resolvent of a Maximal Monotone Operator 29
2.7.1 Nonexpansiveness of the Resolvent: 30
2.8 Yosida Approximation of a Maximal Monotone Operator 31

3 Results and Applications **34**
3.1 Some Results on the Difference Between Two Maximal Monotone Operators . . 34
3.1.1 The difference between two maximal monotone operators : 34

3.2 Applications in Optimization Problems 43
3.2.1 DC (Difference of Convex) Programming: 43
3.2.2 Variational Inequalities & Equilibrium Problems: 44
3.2.3 Saddle-Point & Minimax Problems 44
3.3 Conclusion 46

LIST OF FIGURES

1.1 Set-Valued Map Example 4 12
1.2 Set-Valued Map example 5 13

2.1 Example of a convex set (left) and a non-convex set (right) 17
2.2 (a) Representation of a convex set as the convex hull of a set of points. (b) Representation of a convex set as the intersection of a (possibly infinite) number of halfspaces. 18

0.1 General Introduction

The study of set-valued maps and maximal monotone operators plays a crucial role in various fields of mathematical analysis, particularly in optimization theory, convex ,analysis and functional analysis.

This thesis explores these fundamental concepts by providing definitions, examples, and essential algebraic operations. The first chapter introduces set-valued maps, presenting basic definitions, illustrative exam- ples, and fundamental algebraic operations. The second chapter is dedicated to the theory of maximal monotone operators in Hilbert spaces.

It begins with an introduction to monotone operators, followed by an exploration of key concepts in convex analysis, including convex sets, convex cones, and normal cones. Special attention is paid to convex, lower semicontinuous, and absolutely continuous functions, along with their associated operators, such as the proximal operator and subdifferential operator. Additionally, this chapter discusses contraction operators, resolvents, and Yosida approximations of maximal monotone operators, which are fundamental in the study of nonlinear functional analysis. The third chapter presents important results regarding the difference between two maxi- mal monotone operators and highlights applications in optimization problems.

These results provide valuable insights into the interaction between monotonicity and convexity in various mathematical and applied contexts. The objective of this thesis is to provide a rigorous yet accessible exposition of these topics, serving as a useful reference for researchers and students in mathematical analysis and related fields.

CHAPTER 1

SET-VALUED MAPS

Set-valued maps, or multifunctions, have properties that generalize those of single-valued functions. These properties are crucial for analyzing their behavior and applications in optimization, control theory, economics, and other fields. Below, I will outline the key properties of set-valued maps, categorized into topological, algebraic, and analytical properties.

1.1 First Definitions

1.1.1 Definition :

Given two sets X and Y. A set-valued map[8] F from X to Y is a rule that assigns to each $x \in X$ a subset $F(x) \subseteq Y$. Define as:

$$F : X \rightrightarrows Y \quad \text{or} \quad F : X \to 2^Y$$

where 2^Y is the power set of Y, i.e., the collection of all subsets of Y.
Other names: *point-to-set mapping*, *set-valued operator*, *multi-valued function*, *correspondence*. For simplicity, write $Fx = F(x)$.

1.1.2 Continuity

X and Y two sets and $S : X \to 2^Y$ a set-valued map.

1. S is said to be upper semicontinuous (u.s.c.) at $x \in X$ if for every open set $V \subseteq Y$ and $F(x) \subseteq V$, there exists neighborhood $U(x)$ such that:

$$S(U(x)) := \bigcup_{u \in U(x)} S(u) \subseteq V.$$

   S is u.s.c. if it is u.s.c. at every $x \in X$.

2. S is said lower semicontinuous (l.s.c.) at $x \in X$ if for every $y \in S(x)$ and for every neighborhood $V(y)$ of y, there exists is neighborhood $U(x)$ such that:

$$\forall u \in U(x), \quad S(u) \cap V(y) \neq \emptyset.$$

   S is l.s.c. if it is l.s.c. at all $x \in X$.

3. S is said to be continuous whenever S is both upper semicontinuous and lower semicontinuous.

1.1.2 Key concepts :

1. **Graph:** The graph of set-valued mapping $F : X \to 2^Y$, the graph of F is given by:

$$\text{Graph}(F) = \{(x,y) \mid x \in X, y \in F(x)\}.$$

The graph of F is represented by the set Graph(F).

2. **Domain:** The domain of set-valued mapping F is the set of all $x \in X$ for which $F(x)$ is nonempty::

$$\text{Dom}(F) = \{x \in X \mid F(x) \neq \emptyset\}.$$

3. **Image:** The image of F is the union of all $F(x)$ for $x \in X$:

$$\text{Im}(F) = F(X) = \bigcup_{x \in X} F(x).$$

4. **Range :**

$$\text{R}(F) := \{F(x) \mid x \in \text{Dom}(F)\}$$

5. **Inverse and Core:** Let $X = \text{Dom}(F)$. For every set-valued map $F : X \to 2^Y$, there are two possible definitions of the inverse image. For a subset $M \subset Y$, we define:

$$F^{-1}(M) = \{x \in X \mid F(x) \cap M \neq \emptyset\},$$
$$F^{+1}(M) = \{x \in X \mid F(x) \subseteq M\}.$$

The subset $F^{-1}(M)$ is written as the inverse of M under F, and $F^{+1}(M)$ the center of$M$ under F.

1.2 Examples

Here are some examples of set-valued maps that exhibit complex behaviors:

Example 1. *Simple Set-Valued Map*
Let $X = \{1,2,3\}$ *and* $Y = \{a,b,c\}$. *A set-valued map* $F : X \to 2^Y$ *can be defined as:*

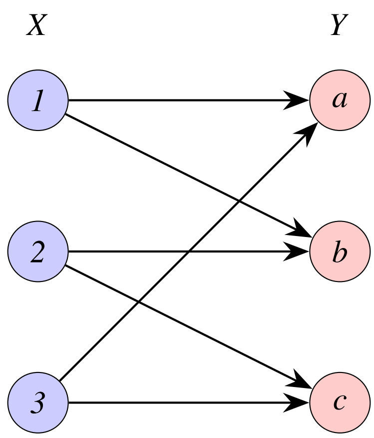
$$F(1) = \{a,b\}, \quad F(2) = \{b,c\}, \quad F(3) = \{a,c\}.$$

The graph of F is:

$$\text{Graph}(F) = \{(1,a),(1,b),(2,b),(2,c),(3,a),(3,c)\}.$$

The inverse mapping $F^{-1} : Y \to 2^X$ *is:*

$$F^{-1}(a) = \{1,3\}, \quad F^{-1}(b) = \{1,2\}, \quad F^{-1}(c) = \{2,3\}.$$

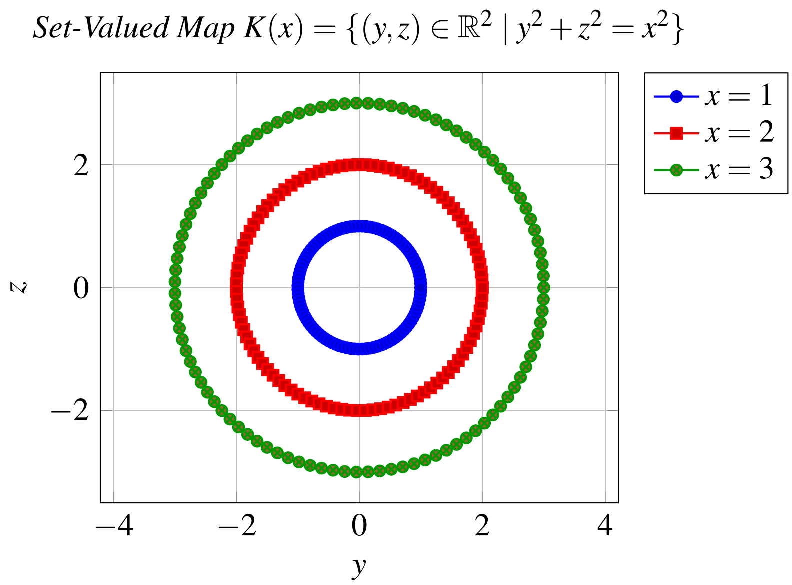


Example 2. *Set-Valued Map whose Range is Not Convex Let the set-valued map $K : \mathbb{R} \to 2^{\mathbb{R}^2}$ be given by :*

$$K(x) = \{(y,z) \in \mathbb{R}^2 \mid y^2 + z^2 = x^2\}.$$

The set $K(x)$ for any x is the circle of radius $|x|$ about the origin.

- *$K(x)$ is not convex for any x other than 0.*
- *For $x = 0$, the set $K(0)$ is collapsed to a point at the origin $(0,0)$.*
- *It is a geometric counterexample where set-valued mappings are curves rather than convex sets.*

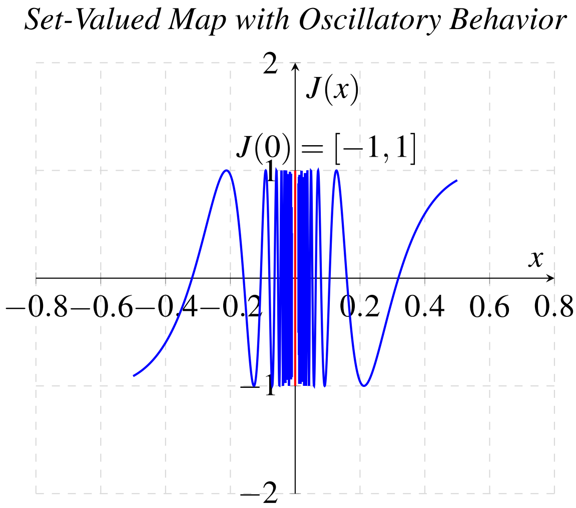


Example 3. *A Set-Valued Map with an Oscillatory Behavior Define the set-valued function $J : \mathbb{R} \to 2^{\mathbb{R}}$ as:*

$$J(x) = \begin{cases} \{\sin(\frac{1}{x})\}, & x \neq 0 \\ [-1,1], & x = 0 \end{cases}$$

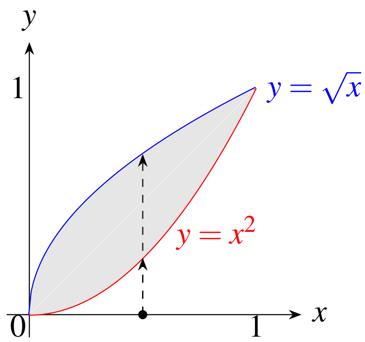
- *As $x \to 0$, the function $J(x)$ oscillates rapidly between -1 and 1.*
- *At $x = 0$, the function fills the entire interval $[-1,1]$.*

- *This is an example of an* ***upper semi-continuous*** *set-valued function.*



Example 4. *The set-valued map* $F : [0,1] \rightrightarrows [0,1]$ *is defined as follows:*

$$F(x) = \{y \in [0,1] \mid x^2 \leq y \leq \sqrt{x}\}$$



Set-Valued Map $F : [0,1] \Rightarrow [0,1]$ defined by $F(x) = \{y \mid x^2 \leq y \leq \sqrt{x}\}$

Figure 1.1: Set-Valued Map Example 4

Another example is $G : [0,1] \rightrightarrows [0,1]$*defined as follows*

$$G(x) = \begin{cases} \left[\frac{1}{4}, \frac{3}{4}\right], & \text{if } x \in \left[0, \frac{1}{4}\right) \\ [0,1], & \text{if } x \in \left[\frac{1}{4}, \frac{3}{4}\right] \\ \left[\frac{1}{4}, \frac{3}{4}\right], & \text{if } x \in \left(\frac{3}{4}, 1\right] \end{cases}$$

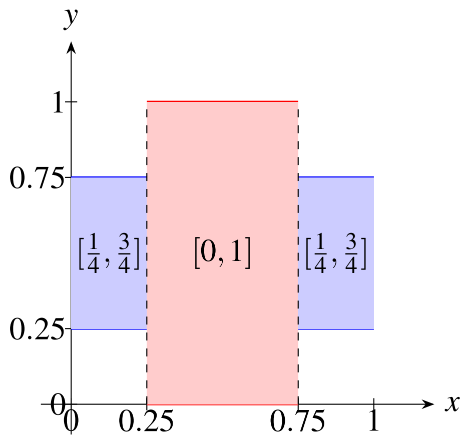


Figure 1.2: Set-Valued Map example 5

1.3 Some Algebraic Operations

1.3.1 Sum of Set-Valued Maps

Definition

Let $F : X \to 2^Y$ and $G : X \to 2^Y$ be two set-valued maps[1], and let X and Y be spaces. The sum of F and G is a new set-valued map $(F+G) : X \rightrightarrows Y$ given by:

$$(F+G)(x) = F(x) + G(x) = \{y_1 + y_2 \mid y_1 \in F(x), y_2 \in G(x)\}.$$

Here, $F(x) + G(x)$meanes Minkowski sum of the sets $F(x)$ and $G(x)$.

Example 5. *Let $F : \mathbb{R} \to 2^{\mathbb{R}}$ and $G : \mathbb{R} \to 2^{\mathbb{R}}$ set-valued mappings are given as:*

$$F(x) = [0,x], \quad G(x) = [-x,x].$$

The sum $F+G$ is given by:

$$(F+G)(x) = [0,x] + [-x,x] = [-x,2x].$$

1.3.2 Composition of Set-Valued Maps

Definition

Let $F : X \to 2^Y$ and $G : Y \to 2^Z$ be two set-valued functions . The composition $(G \circ F) : X \to 2^Z$ is given by:

$$(G \circ F)(x) = \bigcup_{y \in F(x)} G(y).$$

That is, for every $x \in X$, the composition $G \circ F$ maps x to the union of all sets $G(y)$ where y belongs to $F(x)$.

Example 6. *Let $F : \mathbb{R} \to 2^{\mathbb{R}}$ and $G : \mathbb{R} \to 2^{\mathbb{R}}$ be set-valued maps defined as:*

$$F(x) = [0,x], \quad G(y) = [-y,y].$$

The composition $G \circ F$ is given by:

$$(G \circ F)(x) = \bigcup_{y \in [0,x]} [-y,y].$$

Since the union of all intervals $[-y,y]$ for $y \in [0,x]$ results in $[-x,x]$, we conclude:

$$(G \circ S)(x) = [-x,x].$$

1.3.3 Scalar Multiplication

Definition

Let $F : X \to 2^Y$ be a set-valued map[1], where X and Y are spaces (e.g., vector spaces), and let $\lambda \in \mathbb{R}$ The scalar multiplication λF is defined by:

$$(\lambda F)(x) = \lambda F(x) = \{\lambda y \mid y \in F(x)\}.$$

Here, $\lambda F(x)$ is the set obtained by multiplying each element of $F(x)$ by λ.

Example 7. *Let $F : \mathbb{R} \to 2^{\mathbb{R}}$ be a set-valued map defined as:*

$$F(x) = [0,x].$$

For $\lambda = 2$, the scalar multiplication λF is given by:

$$(2F)(x) = 2 \cdot [0,x] = [0,2x].$$

1.3.4 Inverse of a Set-Valued Map

Definition

Let $F : X \to 2^Y$ be a set-valued map[1].The inverse of F, denoted
$F^{-1} : Y \to 2^X$, is defined as:

$$F^{-1}(y) = \{x \in X \mid y \in S(x)\}.$$

Example 8. *Consider the set-valued map:*

$$F(x) = \{|x|\}, \quad x \in \mathbb{R}.$$

The inverse set-valued map is:

$$F^{-1}(y) = \{-y,y\}, \quad y \geq 0.$$

For example:

$$F^{-1}(3) = \{-3,3\}.$$

$$F^{-1}(0) = \{0\}.$$

CHAPTER 2 MAXIMAL MONOTONE OPERATORS IN HILBERT SPACES

Maximal monotone operators are central to nonlinear functional analysis, optimization, and differential inclusions. They generalize the notion of monotonicity to set-valued mappings while preserving key properties crucial for solving variational problems.

2.1 Monotone Operators

Definition :

Let $\mathcal{H}$ be a Hilbert space[3]. The inner product in $\mathcal{H} \times \mathcal{H}$ will be denoted by $\langle \cdot, \cdot \rangle$, and the corresponding norm by $\| \cdot \|$.
A set-valued map $T : \mathcal{H} \to 2^{\mathcal{H}}$ is *monotone* if

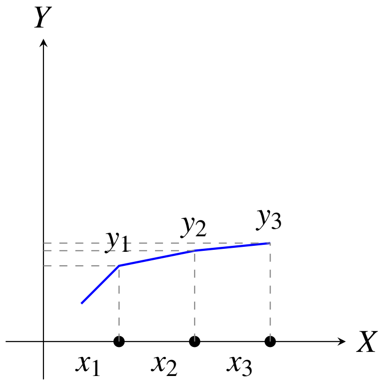
$$\langle x - y, u - v \rangle \geq 0, \quad \forall (x,u),(y,v) \in Graph(T),$$

where the graph of T is defined as

$$G(T) := \{(x,y) \in \mathcal{H} \times \mathcal{H} \; ; \; y \in T(x)\}.$$

Let $D(T)$ be the domain of T, i.e.,

$$D(T) := \{x \in \mathcal{H} \; ; \; T(x) \neq \emptyset\}.$$



Monotone operator (not maximal)

Example 9. 1. *If $\mathcal{H}$ is a Hilbert space and $T : \mathcal{H} \mapsto \mathcal{H}^* \equiv \mathcal{H}$ is a linear map, then T is monotone if and only if T is a positive operator.*
That is $\langle x, Tx \rangle \geq 0$, for all $x \in \mathcal{H}$.

2. *The mapping $A : \mathbb{R} \mapsto 2^{\mathbb{R}}$ defined by*

$$A(x) = \begin{cases} x - a & \text{if } x < 0 \\ \text{any subset of } [-a, a] & \text{if } x = 0 \\ x + a & \text{if } x > 0 \end{cases}$$

is a monotone mapping.

2.1.1 Strictly Monotone Operator

Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be an operator in a Hilbert space. We define A to be *strictly monotone* if:

$$\langle x - y, u - v \rangle > 0 \quad \text{for all } x \neq y,\ u \in A(x),\ v \in A(y).$$

This is a stronger condition than the usual (non-strict) monotonicity, which only requires the inner product to be nonnegative.

Example 10. *Let the function $f(x) = \frac{1}{2}x^2$. Then its gradient is $\nabla f(x) = x$. Take the operator $A(x) = x$. To prove strict monotonicity, consider:*

$$\langle A(x) - A(y), x - y \rangle = \langle x - y, x - y \rangle = \|x - y\|^2 > 0$$

for all $x \neq y$.

Therefore, we conclude that the operator $A(x) = x$ is strictly monotone.

2.1.2 Strong Monotonicity

Let A be an operator on a Hilbert space $\mathcal{H}$. We say that A is strongly monotone if there exists a constant $\alpha > 0$ such that:

For all $x, y \in H$, and all $u \in A(x)$, $v \in A(y)$, the following inequality holds:

$$\langle x - y, u - v \rangle \geq \alpha \|x - y\|^2$$

The constant μ is called the strong monotonicity parameter.

Example 11. *Let $A(x) = \alpha x$, where $\alpha > 0$. Then:*

$$\begin{aligned}\langle A(x) - A(y), x - y\rangle &= \langle \alpha x - \alpha y, x - y\rangle \\ &= \alpha\langle x - y, x - y\rangle \\ &= \alpha\|x - y\|^2\end{aligned}$$

So $A(x)$ is strongly monotone with strong monotonicity parameter $\mu = \alpha$.

2.2 Concepts of Convex Analysis

The concept of convexity plays a deep role in variational analysis. In maximization and minimization analyses, the division among problems of convex or nonconvex type is as significant as the division in other areas of mathematics among problems of linear or nonlinear nature. Additionally, convexity can often be defined or utilized in a local manner and in this way serves many theoretical purposes.

2.2.1 Convex Sets

Definition :

A set S is *convex*[9] if the line segment between any two points in S lies in S, i.e., if for all $x_1, x_2 \in S$ and for all λ with $0 \leq \lambda \leq 1$, we have

$$\lambda x_1 + (1 - \lambda)x_2 \in S.$$

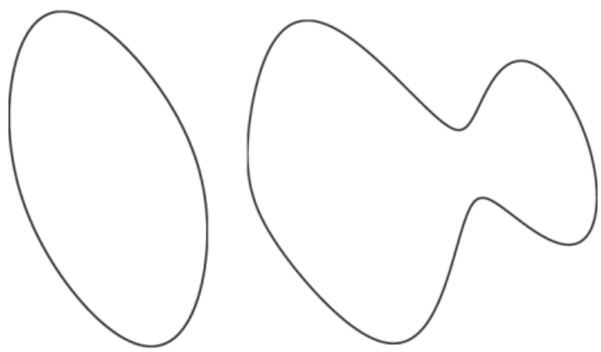


Figure 2.1: Example of a convex set (left) and a non-convex set (right)

Examples:

1. *Euclidean Space* $\mathbb{R}^n$: The entire space is trivially convex.
2. *Line and Plane*: Any line or plane in $\mathbb{R}^n$ is convex.

3. *Half-Spaces*: A half-space defined by

$$H = \{x \in \mathbb{R}^n \mid a^T x \leq b\}$$

is convex.

4. *Norm Balls*: The set

$$B = \{x \in \mathbb{R}^n \mid \|x - c\| \leq r\}$$

is convex because it contains all line segments between points within it.

5. *Polyhedra*: Any intersection of finitely many half-spaces, such as polytopes, is convex.

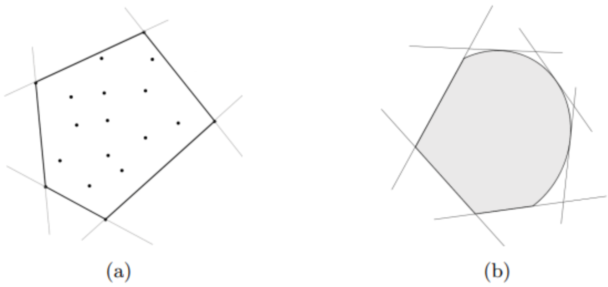


Figure 2.2: (a) Representation of a convex set as the convex hull of a set of points. (b) Representation of a convex set as the intersection of a (possibly infinite) number of halfspaces.

2.2.2 Convex Cones

Convex cones are a special class of convex sets that play a central role in convex analysis

Definition (Convex cones):

A set $C \subseteq \mathbb{R}^n$ is called a **convex cone** [2]if for all vectors $x, y \in C$ and all non-negative scalars $\lambda, \mu \geq 0$, the following holds:

$$\lambda x + \mu y \in C.$$

Example 12. *Nonnegative Orthant*
$\mathbb{R}^n_+$ *The set of all vectors with nonnegative components in* $\mathbb{R}^n$*:*

$$\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0, \forall i\}$$

is a convex cone since nonnegative combinations of nonnegative vectors remain nonnegative.

Solution: Given two vectors $x, y \in \mathbb{R}^n_+$, if

$$x = (x_1, x_2, \ldots, x_n) \quad \text{and} \quad y = (y_1, y_2, \ldots, y_n),$$

then for any $\lambda, \mu \geq 0$:

$$\lambda x + \mu y = (\lambda x_1 + \mu y_1, \ldots, \lambda x_n + \mu y_n).$$

Thus, $\mathbb{R}^n_+$is a convex cone.

2.2.3 Normal Cones

Definition (normal cones):

Let $C \subset \mathbb{R}^n$ be a convex set and $x \in C$[2]. The normal cone to C at $\bar{x}$, denoted $N_C(\bar{x})$, is defined as:

$$N_C(\bar{x}) = \{v \in \mathbb{R}^n \mid \langle v, x - \bar{x} \rangle \leq 0 \text{ for all } x \in C\}$$

Proposition 13. *For any convex set $C \subseteq \mathbb{R}^n$ and point $\bar{x} \in C$, the normal cone $N_C(\bar{x})$ is:*

1. *A closed convex cone*
2. *Non-empty (contains at least 0)*

Proof.

1. *convex cone :* For any $v_1, v_2 \in N_C(\bar{x})$ and $\lambda, \mu \geq 0$:

$$\begin{aligned} \langle \lambda v + \mu v_2, x - \bar{x} \rangle &= \lambda \langle v_1, x - \bar{x} \rangle + \mu \langle v_2, x - \bar{x} \rangle \\ &\leq \lambda \cdot 0 + \mu \cdot 0 = 0 \quad \forall x \in C \end{aligned}$$

   Thus $\lambda v_1 + \mu v_2 \in N_C(\bar{x})$, proving it's a convex cone. Closedness follows from continuity of the inner product.

2. *Non-empty:* The zero vector 0 always satisfies $\langle 0, x - \bar{x} \rangle = 0 \leq 0$ for all $x \in C$.

□

2.3 Convex, Lower Semi-Continuous, Absolutely Continuous Functions

2.3.1 Convex function :

Definition (Convex function):

Let $\mathcal{H}$ be a real Hilbert space[2]. A function $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is called to be **convex** if its domain $\text{dom}(f) \subseteq \mathcal{H}$ is convex, and for all $x, y \in \text{dom}(f)$ and all $\lambda \in [0,1]$, the following inequality holds:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

That is, the function value at any point of the line segment between x and y is on or below the chord joining$f(x)$ and $f(y)$.

Examples of convex functions

- **Univariate functions:**
  - *Exponential function:* The exponential function e^{ax} is convex for any $a \in \mathbb{R}$.
  - *Power function:* The power function x^a is convex for $a \geq 1$ or $a \leq 0$ and is concave for $0 \leq a \leq 1$.

  - *Logarithmic function:* The logarithmic function $\log(x)$ is always concave (and thus $-\log(x)$ is convex).

- **Affine function:** The affine function $a^T x + b$ is both convex and concave.
- **Quadratic function:** The quadratic function $\frac{1}{2}x^T Qx + b^T x + c$ is convex provided that $Q \succeq 0$ (i.e., Q is positive semidefinite).

Key properties of convex functions [2]:

- **Restriction to lines:** A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if its restriction to any line is convex. For example, let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and $x_0, a \in \mathbb{R}^n$ be fixed. Define $g(t) = f(x_0 + ta)$. Then f is convex if and only if g is convex for every choice of x_0 and a. This property is useful for proving the convexity of certain functions.
- **Epigraph characterization:** A function f is convex if and only if its epigraph is a convex set, where the epigraph is defined as:

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

  Intuitively, the epigraph is the set of points that lie above the graph of the function.

- **Convex sublevel sets:** If f is convex, then every sublevel set of f is convex, where a sublevel set is defined as

$$\{x \in \text{dom}(f) : f(x) \leq t\}$$

  for some parameter $t \in \mathbb{R}$. However, the converse is not true. For example, $f(x) = \sqrt{|x|}$ is not convex but all its sublevel sets are convex.

- **First-order characterization:** If f is differentiable, then f is convex if and only if $\text{dom}(f)$ is convex, and

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

  for all $x, y \in \text{dom}(f)$. Geometrically, this means the graph of f lies above all its tangent hyperplanes. This also implies that for differentiable f, a point x minimizes f if and only if $\nabla f(x) = 0$.

Example 14. ***Quadratic Function***

A function of the form:

$$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle$$

where $A : \mathcal{H} \to 2^{\mathcal{H}}$ is a positive semi-definite operator and $b \in \mathcal{H}$, is convex.

Solution : For convexity, we check if:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y), \quad \forall x, y \in \mathcal{H}, \lambda \in [0,1].$$

Expanding the function:

$$f(\lambda x + (1-\lambda)y) = \frac{1}{2}\langle A(\lambda x + (1-\lambda)y), \lambda x + (1-\lambda)y\rangle + \langle b, \lambda x + (1-\lambda)y\rangle.$$

Using bilinearity:

$$\langle A(\lambda x + (1-\lambda)y), \lambda x + (1-\lambda)y\rangle = \lambda^2\langle Ax, x\rangle + (1-\lambda)^2\langle Ay, y\rangle + 2\lambda(1-\lambda)\langle Ax, y\rangle.$$

Since A is positive semi-definite, we obtain:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

Thus, $f(x)$ is convex.

2.3.2 Lower Semi-Continuous :

Definition (L.S.C):

Let $\mathcal{H}$ be a Hilbert space, A function $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is called lower semicontinuous at a point $x \in \mathcal{H}$ if:

$$\liminf_{y \to x} f(y) \geq f(x).$$

Equivalently, for every sequence $x_k \to x$ in $\mathcal{H}$:

$$\liminf_{k\to\infty} f(x_k) \geq f(x).$$

Proposition 15 (Epigraph Characterization)**.** *A function is lsc if and only if its epigraph*

$$\text{epi}(f) = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq \alpha\}$$

is a closed set.

Proof of Epigraph Characterization. ($\Rightarrow$) Suppose f is lsc and let $(x_k, \alpha_k) \in \text{epi}(f)$ converge to $(\bar{x}, \bar{\alpha})$. Then:

$$f(\bar{x}) \leq \liminf_{k\to\infty} f(x_k) \leq \liminf_{k\to\infty} \alpha_k = \bar{\alpha}$$

Thus $(\bar{x}, \bar{\alpha}) \in \text{epi}(f)$.

($\Leftarrow$) For any $x_k \to \bar{x}$, take $\alpha_k = f(x_k) + \varepsilon$. Since $(x_k, \alpha_k) \in \text{epi}(f)$ converges to $(\bar{x}, \liminf f(x_k) + \varepsilon)$, closedness implies:

$$f(\bar{x}) \leq \liminf f(x_k) + \varepsilon \quad \forall \varepsilon > 0$$

□

Theorem 16 (Properties of lsc Functions)**.**

1. *If $\{f_i\}_{i\in I}$ is a family of lsc functions, then $\sup_{i\in I} f_i$ is lsc*
2. *The sum $f_1 + f_2$ of two lsc functions is lsc*
3. *If f is lsc and g is continuous, then $f \circ g$ is lsc*
4. *(Weierstrass Theorem) An lsc function on a compact set attains its infimum*

Properties

1. **Open Set Characterization:** A function f is lower semi-continuous if and only if the set:

$$\{x \mid f(x) > c\}$$

is open for all $c \in \mathbb{R}$.

2. **Epigraph Characterization:** A function f is lower semi-continuous if and only if its epigraph,

$$\text{epi}(f) = \{(x,y) \mid y \geq f(x)\}$$

is a closed set in $\mathbb{R}^{n+1}$.

Proof. 1. **(⇒) Assume f is lower semi-continuous, demonstrate $\{x \mid f(x) > c\}$ is open.** Choose any x_0 for which $f(x_0) > c$. By the definition of lower semi-continuity, we have:

$$\liminf_{x_k \to x_0} f(x_k) \geq f(x_0).$$

This means that there is a small neighborhood around x_0 where $f(x) > c$, and thus x_0 is an interior point of the set $\{x \mid f(x) > c\}$. Since this holds for any x_0, the set is open. **(⇐) Assume $\{x \mid f(x) > c\}$ is open for all c, prove f is lower semi-continuous.** Consider any sequence $x_k \to x_0$. We have to prove:

$$\liminf_{x_k \to x_0} f(x_k) \geq f(x_0).$$

Suppose for contradiction otherwise. Then there is a sequence $x_k \to x_0$ such that:

$$\liminf f(x_k) < f(x_0).$$

Thus there is some $c < f(x_0)$ such that infinitely many x_k have $f(x_k) \leq c$, which contradicting the openness of $\{x \mid f(x) > c\}$. thus, f is lower semi-continuous.

2. **(⇒) Assume f is lower semi-continuous, prove** epi(f) **is closed.** Let $(x_k, y_k) \to (x, y)$ be a convergent sequence in $\mathbb{R}^{n+1}$, with each $(x_k, y_k) \in \text{epi}(f)$, i.e.:

$$y_k \geq f(x_k).$$

Taking limits ,

$$y = \lim y_k, \quad x = \lim x_k.$$

Since f is lower semi-continuous, we know:

$$\liminf_{x_k \to x} f(x_k) \geq f(x).$$

Since $y_k \geq f(x_k)$, taking limits,

$$y \geq \liminf f(x_k) \geq f(x).$$

This shows $(x,y) \in \text{epi}(f)$, proving that the epigraph is closed.
($\Leftarrow$) Assume epi(f) is closed, show f is lower semi-continuous. pick any sequence $x_k \to x$. Consider the points $(x_k, f(x_k)) \in \text{epi}(f)$. If the epigraph is closed, then the limit point $(x, \liminf f(x_k))$ must also be in the epigraph:

$$\liminf f(x_k) \geq f(x).$$

This proves f to be lower semi-continuous.

□

2.3.3 Absolutely Continuous Functions:

Definition(A.C.F):

Let $\mathcal{H}$ be a Hilbert space,A function $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$is said to be absolutely continuous on $[a,b]$ if for all $\varepsilon > 0$, there exists $\delta > 0$ such that for any finite set of disjoint intervals (x_k, y_k) in $[a,b]$,

$$\sum_k |y_k - x_k| < \delta \quad \Rightarrow \quad \sum_k \|f(y_k) - f(x_k)\| < \varepsilon.$$

This definition generalizes absolute continuity from real functions to vector-valued functions in Hilbert spaces.

Properties

Differentiability Almost Everywhere: If f is absolutely continuous, then it is differentiable almost everywhere on $[a,b]$.
Integration Formula: If f is absolutely continuous, then for all $x, y \in [a,b]$,

$$f(y) = f(x) + \int_x^y f'(t)\,dt.$$

Lipschitz Continuity Implies Absolute Continuity: If f is Lipschitz continuous, i.e.,

$$\|f(y) - f(x)\| \leq L|y - x|, \quad \forall x, y \in [a,b],$$

then f is absolutely continuous.

2.4 Proximal and Subdifferential of a Convex Function

2.4.1 Subdifferential of a Convex Function:

Definition:

A function$f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex, and lower semicontinuous.
The **subdifferential** of f at $x \in \mathrm{Dom}(f)$is defined as:

$$\partial f(x) = \{x^* \in \mathcal{H} \mid f(y) \geq f(x) + \langle x^*, y - x \rangle, \quad \forall y \in \mathcal{H}\}.$$

Any element $x^* \in \partial f(x)$ is known as *subgradient* of f at x.

1. Whenever f is differentiable at x, then the subdifferentialis simply to the standard gradient:
$$\partial f(x) = \{\nabla f(x)\}.$$
2. When f is not differentiable, $\partial f(x)$ will be a set, maybe containing more than one subgradient.

Monotonicity of subdifferentials

The operator ∂f is monotone in the sense that for every $x, y \in X$ and for $x^* \in \partial f(x)$, $y^* \in \partial f(y)$, it holds that:

$$\langle x^* - y^*, x - y \rangle \geq 0.$$

This at once follows from the definition of the subdifferential and the convexity of f. It captures the fact that the graph of ∂f has points that are not inconsistent with monotonicity,, therefore ∂f is a monotone operator.

Proof. The operator ∂f is monotone in the sense that for any $x, y \in X$ and for any $x^* \in \partial f(x)$, $y^* \in \partial f(y)$, we have :

$$\langle x^* - y^*, x - y \rangle \geq 0.$$

From the definition , we have:

$$f(y) \geq f(x) + \langle \partial f(x), y - x \rangle,$$
$$f(x) \geq f(y) + \langle \partial f(y), x - y \rangle.$$

It results from combining both inequalities gives as:

$$\langle \partial f(x) - \partial f(y), x - y \rangle \geq 0.$$

As a direct result of the convexity of f and the implication that the graph of ∂f is composed of points not in a way that does not satisfy monotonicity so that ∂f is a monotone operator.

□

Proposition 17. *Let $\mathcal{H}$ be a Hilbert space and let $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a proper, convex, and lower semicontinuous function.*

If $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ *satisfies*

$$\langle \partial f(x), x \rangle \geq 0, \quad \forall x \in \operatorname{Dom} \partial f, \tag{2}$$

then $0 \in \partial f(0)$.

Theorem 18. *Let X be a set [9] and let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper, convex, and lower semicontinuous function. Then, the subdifferential ∂f is a maximal monotone mapping from X to X^*.*

Proof. Let $(u, v) \in X \times X^*$ be in monotone relation to the graph of ∂f, that is,

$$\langle \partial f(x) - v, x - u \rangle \geq 0, \quad \forall x \in \operatorname{dom} \partial f. \tag{1}$$

Consider the function

$$\bar{f}(x) := f(x+u) - v(x). \tag{2}$$

It is immediate to check that (1) implies (2) for $\bar{f}$. By Proposition 1, we get

$$0 \in \partial \bar{f}(0),$$

which easily translates to $v \in \partial f(u)$. Therefore, ∂f cannot be properly extended in a monotone way. □

2.4.2 Proximal Operator of a Convex Function

Definition:

Let $f : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ be a proper, convex, and lower semicontinuous function.
The **proximal mapping** (or **prox operator**) of f with parameter $\lambda > 0$ is defined as:

$$\operatorname{prox}_{\lambda f}(x) = \arg\min_{y \in H} \left\{ f(y) + \frac{1}{2\lambda} \|y - x\|^2 \right\}.$$

In words, the proximal operator returns the point y that minimizes the sum of:

- The function $f(y)$ (the original function we wish to regularize).
- A quadratic term $\frac{1}{2\lambda}\|y - x\|^2$ that keeps y close to x.

Properties

1. *Non-expansiveness* The proximal operator is 1-Lipschitz, meaning:

$$\|\operatorname{prox}_{\lambda f}(x) - \operatorname{prox}_{\lambda f}(y)\| \leq \|x - y\|.$$

2. *Moreau Decomposition* The proximal operator is related to the convex conjugate f^* via:

$$x = \operatorname{prox}_{\lambda f}(x) + \lambda \operatorname{prox}_{f^*/\lambda}(x/\lambda).$$

3. *Connection to Subgradients*

$$x - \operatorname{prox}_{\lambda f}(x) \in \lambda \partial f(\operatorname{prox}_{\lambda f}(x)).$$

This means that the proximal operator finds a point where the scaled subgradient of f balances the distance to x.

Exemple: Let

$$f(x) = \frac{1}{2}\alpha\|x\|^2, \quad \alpha > 0.$$

The proximal mapping is defined as:

$$\operatorname{prox}_{\lambda f}(x) = \arg\min_{y \in H} \left\{ f(y) + \frac{1}{2\lambda}\|y - x\|^2 \right\}.$$

Solution: We need to solve:

$$\min_y \left\{ \frac{1}{2}\alpha\|y\|^2 + \frac{1}{2\lambda}\|y - x\|^2 \right\}.$$

Taking the derivative with respect to y and setting it to zero:

$$\begin{aligned} \alpha y + \frac{1}{\lambda}(y - x) = 0, \\ y\left(\alpha + \frac{1}{\lambda}\right) = \frac{x}{\lambda}, \\ y = \frac{x}{1 + \lambda\alpha}, \\ \operatorname{prox}_{\lambda f}(x) = \frac{x}{1 + \lambda\alpha}. \end{aligned}$$

This mapping shrinks x by a factor of $\frac{1}{1+\lambda\alpha}$, which acts as a scaled identity operator.

2.5 Contraction Operators on Hilbert Spaces

Definition:

Let $\mathcal{H}$ be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. An operator $T : \mathcal{H} \to \mathcal{H}$ is called a contraction if there exists a constant $\alpha \in [0,1)$ such that for all $x, y \in \mathcal{H}$,

$$\|T(x) - T(y)\| \leq \alpha \|x - y\|.$$

If $\alpha = 1$, T is called nonexpansive.
If $\alpha \in [0,1)$, T is called a strict contraction
This means that T does not increase the norm of any vector in $\mathcal{H}$.

2.6 Maximal Monotone Operators

Definition:

The operator $A : \mathcal{H} \to 2^{\mathcal{H}}$ is *maximally monotone* if it is monotone and its graph is a maximal element with respect to the inclusion order in $\mathcal{H} \times \mathcal{H}$.

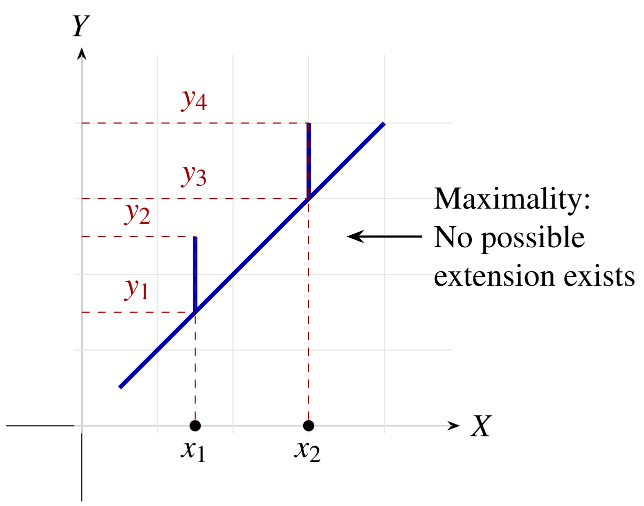
Equivalently,
A is maximally monotone if it is monotone and, in addition,

$$\langle x - y, x^* - y^* \rangle \geq 0 \quad \text{for all } (y, y^*) \in \text{Graph}(A)$$

implies $(x, x^*) \in \text{Graph}(A)$.

The graph of a maximally monotone operator $T : \mathcal{H} \to 2^{\mathcal{H}}$ is given by:

$$\text{Graph}(A) = \{(x, x^*) \in \mathcal{H} \times \mathcal{H} \mid x^* \in T(x)\}.$$



Maximal Monotone Operator (Multivalued Case)

Definition:(Zero set of a maximal monotone operator)

Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximal monotone operator on a Banach space $\mathcal{H}$. The zero set of A is defined as:

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

This means that for each $x \in \text{zer}(A)$, there exists an element 0 in the set $A(x)$, i.e.,$A(x) \ni 0$.

Lemma 19. *Let T be a maximal monotone operator[7]. The point $[x,x^*] \in \mathcal{H} \times \mathcal{H}$ belongs to the graph of T if, and only if,*

$$\langle x^* - u^*, x - u\rangle \geq 0 \quad \textit{for all } [u,u^*] \in \text{graph}(T).$$

Proof. If $[x,x^*] \in T$, the inequality $\langle x^* - u^*, x - u\rangle \geq 0$ follows immediately by the monotonicity of T.

If $[x,x^*] \notin A$, then the set $A \cup \{[x,x^*]\}$ would be the graph of a monotone operator strictly extending A, which contradicts the maximality of A. □

An operator $A\colon \mathcal{H} \rightrightarrows \mathcal{H}$ is called **nonexpansive** if for all $x^*, y^* \in \mathcal{H}$,

$$\|x^* - y^*\| \leq \|x - y\|$$

where $x^* \in A(x)$ and $y^* \in A(y)$.

Example 20. *Let $\Gamma_0(\mathcal{H})$ denote the set of all proper, lower semicontinuous convex functions $f\colon \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$. For $f \in \Gamma_0(\mathcal{H})$, the subdifferential of f is the operator $\partial f\colon \mathcal{H} \rightrightarrows \mathcal{H}$ defined by*

$$\partial f(x) = \{x^* \in H : f(z) \geq f(x) + \langle x^*, z - x\rangle \textit{ for all } z \in H\}.$$

To see that it is monotone, take $x^* \in \partial f(x)$ and $y^* \in \partial f(y)$. Thus

$$f(y) \geq f(x) + \langle x^*, y - x\rangle,$$
$$f(x) \geq f(y) + \langle y^*, x - y\rangle.$$

Adding these two inequalities we obtain $\langle x^* - y^*, x - y\rangle \geq 0$.

For maximality, according to Theorem 2 it suffices to prove that for each $y \in H$ and each $\lambda > 0$ there exists $x_\lambda \in D(\partial f)$ such that $y \in x_\lambda + \lambda \partial f(x_\lambda)$. Indeed, consider the Moreau-Yosida approximation of f at y, which is the function f_λ defined by

$$f_\lambda(x) = f(x) + \frac{1}{2\lambda}\|x - y\|^2. \tag{4}$$

It is proper, lower semicontinuous, strongly convex and coercive (due to the quadratic term and the fact that f has an affine minorant). Its unique minimizer x_λ satisfies

$$0 \in \partial f_\lambda(x_\lambda) = \partial f(x_\lambda) + \frac{1}{\lambda}(x_\lambda - y).$$

That is, $y \in x_\lambda + \lambda \partial f(x_\lambda)$.

2.7 Resolvent of a Maximal Monotone Operator

Definition :

Let $T : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximal monotone operator . For all $\lambda > 0$, the resolvent of T with index λ is defined by:

$$J_\lambda^T = (I + \lambda T)^{-1}. \tag{2.1}$$

Therefore, for each $x \in \mathcal{H}$, the resolvent $J_\lambda^T(x)$ is given by the unique solution $y \in \mathcal{H}$ of the inclusion: $x \in y + \lambda T(y)$.

Definition :(Reflected resolvent)

The reflected resolvent of a maximal monotone operator T is given by:
$R_\lambda^T = 2J_\lambda^T - I$.
It maps any point $x$to a "reflected" point with respect to the resolvent $J_\lambda^T(x)$:

$$R_\lambda^T(x) = 2J_\lambda^T(x) - x.$$

Proposition 21. *Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximal monotone operator on a real Hilbert space $\mathcal{H}$, and let $\lambda > 0$. Then:*

1. *The resolvent operator $J_\lambda^A = (I + \lambda A)^{-1}$ is single-valued.*
2. *The domain of J_λ^A is the entire space X, i.e., $\mathrm{dom}(J_\lambda^A) = \mathcal{H}$.*

Proof. 1. By definition, 2.1 which means $J_\lambda^A(x)$ is the set of solutions $y \in X$ to the inclusion:

$$x \in y + \lambda A(y).$$
$$x - y \in \lambda A(y).$$

Since A is maximal monotone, the operator $I + \lambda A$ is strongly monotone. Specifically, for any $y_1, y_2 \in \mathcal{H}$, if

$$x \in y_1 + \lambda A(y_1) \quad \text{and} \quad x \in y_2 + \lambda A(y_2),$$
$$\Longrightarrow x - y_1 \in \lambda A(y_1) \quad \text{and} \quad x - y_2 \in \lambda A(y_2).$$

Subtracting these two inclusions and using the monotonicity of A, we get:

$$\langle (x - y_1) - (x - y_2), y_1 - y_2 \rangle \geq 0.$$

This becomes:

$$\langle y_2 - y_1, y_1 - y_2 \rangle \geq 0 \Rightarrow -\|y_1 - y_2\|^2 \geq 0.$$

This implies $y_1 = y_2$, so the solution y is unique. Thus, J_λ^A is single-valued.

2. To show that $\mathrm{Dom}(J_\lambda^A) = \mathcal{H}$, we need to prove that for every $x \in \mathcal{H}$, there exists a $y \in \mathcal{H}$ such that:

$$x \in y + \lambda A(y).$$

Equivalently, we need to show that the equation:

$$x = y + \lambda v, \quad \text{where } v \in A(y),$$

has a solution $y \in \mathcal{H}$ for every $x \in \mathcal{H}$.

This is equivalent to showing that the operator $I + \lambda A$ is surjective, i.e.,

$$\mathrm{R}(I + \lambda A) = \mathcal{H}.$$

The surjectivity of $I + \lambda A$ is a direct consequence of Minty's theorem (or Rockafellar's theorem), which states that for a maximal monotone operator A, the operator $I + \lambda A$ is surjective for all $\lambda > 0$.

Therefore, for every $x \in \mathcal{H}$, there exists a unique $y \in \mathcal{H}$ such that:

$$x = y + \lambda v, \quad \text{where } v \in A(y).$$

This means $y = J_\lambda^A(x)$, and J_λ^A is defined everywhere on $\mathcal{H}$. Thus,

$$\mathrm{Dom}(J_\lambda^A) = \mathcal{H}.$$

□

2.7.1 Nonexpansiveness of the Resolvent:

The resolvent J_λ^A of a maximal monotone operator A is nonexpansive. This means:

$$\|J_\lambda^A(x) - J_\lambda^A(y)\| \leq \|x - y\|, \quad \forall x, y \in \mathcal{H}. \tag{2.2}$$

Proof. Let $x, y \in \mathcal{H}$, and define $u = J_\lambda^A(x)$ and $v = J_\lambda^A(y)$. By equation 2.1:

$$\begin{aligned} u = (I + \lambda A)^{-1}(x) \quad &\text{and} \quad v = (I + \lambda A)^{-1}(y). \\ \Longrightarrow x \in u + \lambda A(u) \quad &\text{and} \quad y \in v + \lambda A(v). \\ \Longrightarrow x - u \in \lambda A(u) \quad &\text{and} \quad y - v \in \lambda A(v). \end{aligned}$$

Since A is monotone, we have:

$$\langle (x - u) - (y - v), u - v \rangle \geq 0.$$

Substituting $x - u = \lambda a$ and $y - v = \lambda b$ for some $a \in A(u)$ and $b \in A(v)$, the monotonicity of A ensures:

$$\langle a - b, u - v \rangle \geq 0.$$

From the above, we derive:

$$\begin{aligned} &\langle (x - u) - (y - v), u - v \rangle \geq 0. \\ &\langle x - y, u - v \rangle - \|u - v\|^2 \geq 0. \\ &\langle x - y, u - v \rangle \geq \|u - v\|^2. \end{aligned}$$

By the Cauchy-Schwarz inequality:

$$\|x-y\|\cdot\|u-v\| \geq \langle x-y, u-v\rangle \geq \|u-v\|^2.$$

Dividing both sides by $\|u-v\|$ (assuming $u \neq v$):

$$\|x-y\| \geq \|u-v\|.$$

This shows:

$$\|J_\lambda^A(x) - J_\lambda^A(y)\| \leq \|x-y\|.$$

□

Theorem 22. *[7]Let $T : \mathcal{H} \to 2^{\mathcal{H}}$.Then:*

1. *T is monotone if and only if J_λ^T is nonexpansive for each $\lambda > 0$.*
2. *A monotone operator T is maximal if and only if $I + \lambda A$ is surjective for each $\lambda > 0$.*

2.8 Yosida Approximation of a Maximal Monotone Operator

Definition(Yosida Approximation):

An operator's Yosida approximation, corresponding to $\lambda > 0$, is defined as:

$$T_\lambda = \frac{1}{\lambda}\left(I - (I+\lambda A)^{-1}\right).$$

where I denotes the identity operator on H. It is beneficial to define the resolvent of A:

$$J_\lambda^A = (I+\lambda T)^{-1}.$$

This is known as the resolvent of T for $\lambda > 0$. To avoid ambiguity, we sometimes write it as $J_{T,\lambda}$, but we can simply write:$T_\lambda = \frac{1}{\lambda}(I - J_\lambda^T)$.
The domains satisfy: $Dom(J_\lambda^T) = Dom(T_\lambda) = R(I+\lambda T)$.

Example 23. *Consider the subdifferential operator of the function $f(x) = |x|$, given by:*

$$\partial f(x) = \begin{cases} \{-1\}, & x < 0, \\ [-1,1], & x = 0, \\ \{1\}, & x > 0. \end{cases}$$

Solution : By definition, the Yosida approximation of $\partial f(x)$ is given by:

$$\partial f_\lambda(x) = \frac{1}{\lambda}(x - J_\lambda^{\partial f}(x)).$$

Using the values of $J_\lambda^{\partial f(x)}$:
- For $x > \lambda$:

$$\partial f_\lambda(x) = \frac{1}{\lambda}(x - (x-\lambda)) = \frac{1}{\lambda}\lambda = 1.$$

- For $x < -\lambda$:

$$\partial f_\lambda(x) = \frac{1}{\lambda}(x - (x+\lambda)) = \frac{1}{\lambda}(-\lambda) = -1.$$

- For $|x| \leq \lambda$:

$$\partial f_\lambda(x) = \frac{1}{\lambda}(x - 0) = \frac{x}{\lambda}.$$

So:

$$\partial f_\lambda(x) = \begin{cases} 1, & x > \lambda, \\ \frac{x}{\lambda}, & |x| \leq \lambda, \\ -1, & x < -\lambda. \end{cases}$$

Theorem 24. *[7] If $T : \mathcal{H} \to 2^{\mathcal{H}}$ is maximal monotone, then for any $\lambda > 0$:*

(i) The ***resolvent*** *$J_\lambda^T := (I + \lambda T)^{-1}$ is single-valued*

(ii) J_λ^T is ***firmly nonexpansive****:*

$$\|J_\lambda^T x - J_\lambda^T y\|^2 \leq \langle x - y, J_\lambda^T x - J_\lambda^T y \rangle \quad \forall x, y \in \mathcal{H}$$

(iii) The ***Yosida approximation*** *$T_\lambda := \frac{I - J_\lambda^T}{\lambda}$ is $\frac{1}{\lambda}$-Lipschitz continuous*

Properties

1. *Single-Valued and Lipschitz Continuous*
   Unlike the original operator T, which may be multi-valued, the Yosida approximation T_λ is always single-valued and Lipschitz continuous:

$$\|T_\lambda(x) - T_\lambda(y)\| \leq \frac{1}{\lambda}\|x - y\|.$$

   This makes it useful for regularization and numerical computations.

2. *Monotonicity*
   The Yosida approximation preserves the monotonicity of T:

$$\langle T_\lambda(x) - T_\lambda(y), x - y \rangle \geq 0, \quad \forall x, y \in H.$$

   This ensures that it behaves like T in terms of order structure.

3. *Relation to the Resolvent*
   By definition:

$$x - \lambda T_\lambda(x) = J_\lambda^T(x).$$

   This means $T_\lambda(x)$ measures the difference between x and its resolvent mapping.

Proof. 1. The Yosida approximation is defined as:

$$T_\lambda(x) = \frac{1}{\lambda}\left(x - J_\lambda^T(x)\right),$$

where $J_T^\lambda = (I + \lambda T)^{-1}$ is the resolvent of T.

From 2.2, J_T^λ is firmly nonexpansive, meaning:

$$\|J_\lambda^T(x) - J_\lambda^T(y)\| \leq \|x - y\|.$$

Using the definition of T_λ,

$$\begin{aligned}
\|T_\lambda(x) - T_\lambda(y)\| &= \|\frac{1}{\lambda}((x - J_\lambda^T(x)) - (y - J_\lambda^T(y))\|. \\
&= \frac{1}{\lambda}\|(x-y) - (J_\lambda^T(x) - J_\lambda^T(y))\| \\
&\leq \frac{1}{\lambda}\|x - y\|.
\end{aligned}$$

Thus, T_λ is $\frac{1}{\lambda}$-Lipschitz continuous and single-valued.

2. From the resolvent identity:

$$T_\lambda(x) = \frac{1}{\lambda}\left(x - J_\lambda^T(x)\right).$$

Using the firmly nonexpansive property of the resolvent:

$$\begin{aligned}
\langle T_\lambda(x) - T_\lambda(y), x - y\rangle &= \left\langle \frac{1}{\lambda}\left(x - J_\lambda^T(x) - y + J_\lambda^T(y)\right), x - y \right\rangle \\
&= \frac{1}{\lambda}\left\langle (x-y) - (J_\lambda^T(x) - J_\lambda^T(y)), x - y\right\rangle \\
&= \frac{1}{\lambda}\left(\|x-y\|^2 - \langle J_\lambda^T(x) - J_\lambda^T(y), x - y\rangle\right) \\
&= \frac{1}{\lambda}(\|x-y\|^2 - \|J_\lambda^T(x) - J_\lambda^T(y)\|^2) \\
&\geq 0
\end{aligned}$$

Thus, T_λ is monotone.

3. by definition:

$$T_\lambda(x) = \frac{1}{\lambda}\left(x - J_\lambda^T(x)\right) \implies x - \lambda T_\lambda(x) = J_\lambda^T(x).$$

This equation shows that the Yosida approximation measures the deviation from the resolvent.

□

CHAPTER 3

RESULTS AND APPLICATIONS

3.1 Some Results on the Difference Between Two Maximal Monotone Operators

Take$A, B : X \to 2^{X^*}$ The two together create maximal monotone operators on a Banach space X. Gap $A - B$ is present in variational analysis [5], optimization, and PDEs. For the most part disjunctive maximal monotone operators may not be necessarily monotone or maximal monotone. The main results and prerequisites are categorized below.

Definition(Monotone positive linear relation):

T is said to be **monotone positive linear relation** if:

1. **Linear:**
   For any $(x, x^*), (y, y^*) \in T$ and scalars $\alpha, \beta \in \mathbb{R}$, the linear combination

   $$(\alpha x + \beta y, \alpha x^* + \beta y^*) \in T.$$

   T is a linear subspace of $\mathcal{H} \times \mathcal{H}$.
2. **Monotone:** $\langle x - y, x^* - y^* \rangle \geq 0 \quad$ for all $(x, x^*), (y, y^*) \in T$.
3. **Positive:** $\langle x, x^* \rangle \geq 0 \quad$ for all $(x, x^*) \in T$. This requires that each element of the relation be positively paired under the duality pairing.

Definition(Skew Linear Relation):

A linear relation $T \subseteq \mathcal{H} \times \mathcal{H}$ is called **skew** if $\langle x, x^* \rangle = 0 \quad$ for all $(x, x^*) \in T$.

3.1.1 The difference between two maximal monotone operators :

Let us take two maximal monotone operators, A and B, and let us define their difference as follows: $K = A - B$ in general, K is not maximally monotone or monotone. But it is possible to find examples where some features are guaranteed. A maximal monotone operator and $\lambda > 0$. we have the resolvent of A By definition2.1

Theorem 25. *For all $\lambda > 0$, $0 \in A(x)$ if and only if $J_\lambda^A(x) = x$.*

Proof.

$$\begin{aligned} 0 \in \lambda A(x) &\iff x \in x + \lambda A(x) \\ &\iff x \in (I + \lambda A)(x) \\ &\iff x = (I + \lambda A)^{-1}(x) \\ &\iff x = J_\lambda^A(x) \end{aligned}$$

□

We have $J_\lambda^A(x) = (I + \lambda A)^{-1}(x)$ and the Yosida approximation is defined as

$$A_\lambda(x) = \frac{1}{\lambda}(x - (J_\lambda^A(x))$$

Proposition 26. *For all $\lambda > 0$,*

$$A_\lambda(x) \in A(J_\lambda^A(x))$$

Proof. We have $J_\lambda^A(x) = (I + \lambda A)^{-1}(x)$ and the Yosida approximation is defined as $A_\lambda(x) = \frac{1}{\lambda}(x - (J_\lambda^A(x))$

$$\begin{aligned} &\iff x \in (I + \lambda A)((J_\lambda^A(x)) \\ &\iff x \in J_\lambda^A(x) + \lambda A(J_\lambda^A(x)) \\ &\iff x - J_\lambda^A(x) \in \lambda A(J_\lambda^A(x)) \\ &\iff \frac{1}{\lambda}(x - (J_\lambda^A(x)) \in A(J_\lambda^A(x)) \\ &\iff A_\lambda(x) \in A(J_\lambda^A(x)) \end{aligned}$$

□

Theorem 27. *For all $\lambda > 0$, if $0 \in A(J_\lambda^A(x)) - B(J_\lambda^A(x))$ then $J_\lambda^A(x) = J_\lambda^B(x)$.*

Proof.

$$\begin{aligned} 0 \in A(J_\lambda^A(x)) - B(J_\lambda^A(x)) &\implies 0 \in A_\lambda(x) - B(J_\lambda^A(x)) \\ &\implies A_\lambda(x) \in B(J_\lambda^A(x)) \\ &\implies \lambda A_\lambda(x) \in \lambda B(J_\lambda^A(x)) \\ &\implies J_\lambda^A(x) + \lambda A_\lambda(x) \in (I + \lambda B)(J_\lambda^A(x)) \\ &\implies x \in (I + \lambda B)(J_\lambda^A(x)) \\ &\implies J_\lambda^A(x) = (I + \lambda B)^{-1}(x) \\ &\implies J_\lambda^A(x) = J_\lambda^B(x). \end{aligned}$$

□

Theorem 28. *[4] Suppose $A : \mathcal{H} \to 2^{\mathcal{H}}$ is maximal monotone and $B : \mathcal{H} \to 2^{\mathcal{H}}$ is monotone with $Dom(B) = \mathcal{H}$. Then exactly one of the following statements holds:*

(a) $A - B$ is not monotone.

(b) $A - B$ is maximal monotone.

Proof. Suppose (a) does not hold, then $A - B$ is monotone. We aim to prove that $A - B$ is maximal monotone. To this end, assume that $(y, y^*) \in \mathcal{H} \times \mathcal{H}$ is monotonically related to $\text{Graph}(A - B)$. For each $(x, x^*) \in \text{Graph}(A)$ and each $(x, z^*) \in \text{Gph}(B)$, we get $(x, x^* - z^*) \in \text{Graph}(A - B)$, and hence

$$\langle x^* - z^* - y^*, x - y \rangle \geq 0.$$

Since B is monotone and $B(y) = \emptyset$ by assumption, choosing $b^* \in B(y)$, we get

$$\langle x^* - b^* - y^*, x - y \rangle = \langle x^* - z^* - y^*, x - y \rangle + \langle z^* - b^*, x - y \rangle \geq 0;$$

i.e., $(y, y^* + b^*)$ is monotonically related to $\text{Graph}(A)$. The maximality of A guarantees that $(y, y^* + b^*)$ belongs to $\text{Graph}(A)$, which implies that $(y, y^*) \in \text{Graph}(A - B)$. Consequently , $A - B$ is a maximal monotone operator. □

Example 29.

$$A(x) := \begin{cases} \{0\}, & x = 0, \\ \emptyset, & x \neq 0. \end{cases}$$

$$B(x) := \begin{cases} \{0\}, & x < 0, \\ [0, +\infty), & x = 0, \\ \emptyset, & x > 0. \end{cases}$$

Then, A is maximal monotone, B is monotone, and $A - B$ is monotone but not maximal monotone, since

$$\text{Graph}(A - B) = \{0\} \times [0, \infty).$$

Therefore, in theorem 28, the condition $\text{Dom}(B) = \mathcal{H}$ *is necessary condition..*

Corollary 30. *[4] Let $\mathcal{H}$ be a real Hilbert space, $A : \mathcal{H} \to 2^{\mathcal{H}}$ be maximal monotone, and $B : \mathcal{H} \to 2^{\mathcal{H}}$ be a monotone positive linear relation and* $\text{dom}(B) = \mathcal{H}$. *Then exactly one of the following statements holds true:*

(a) $A - B$ is not monotone.

(b) $A - B$ is maximal monotone.

Proof. Assume that $x^* \in B(x)$ and $y^* \in B(y)$. By , we have

$$B(x) - B(y) = B(x - y).$$

So there exists $z^* \in B(z)$ such that $x^* - y^* = z^*$, where $z = x - y$. Then positivity of B implies that

$$\langle x^* - y^*, x - y \rangle = \langle z^*, z \rangle \geq 0,$$

i.e., B is monotone. Now theorem 28 completes the proof. □

Example 31. *Let* $A : \mathbb{R} \to 2^{\mathbb{R}}$ *be defined by* $A(x) = 2x$ *for each* $x \in \mathbb{R}$*. Consider the mapping* $B : \mathbb{R} \to 2^{\mathbb{R}}$ *defined by*

$$B(x) = \begin{cases} x+1, & x \leq 0, \\ x^2+1, & x > 0. \end{cases}$$

for each $x \in \mathbb{R}$*. Clearly,* A *is maximal monotone,* B *is monotone but it is not positive and linear, whereas* $A - B$ *is maximal monotone.*

The linear relation $B : \mathcal{H} \to 2^{\mathcal{H}}$ *is called a skew linear relation if* $\langle x, x^* \rangle = 0, \quad$ *for each* $(x, x^*) \in \operatorname{Graph}(B)$.

Corollary 32. *Let* $\mathcal{H}$ *be a real Hilbert space,* $A : \mathcal{H} \to 2^{\mathcal{H}}$ *be maximal monotone, and* $B : \mathcal{H} \to 2^{\mathcal{H}}$ *is skew and linear ,* $\operatorname{dom}(B) = \mathcal{H}$.
Then: $A - B$ *is maximal monotone*

Proof. The mapping B is skew and linear, so $-B$ is skew and linear too. Then $\pm B$ and so $A - (\pm B)$ is monotone. That $A \pm B$ is maximal monotone follows from Theorem 4. □

When B is single-valued, we have the following special case:

Corollary 33. *[4] Suppose* $A : X \to 2^{X^*}$ *is maximal monotone and* $B : X \to 2^{X^*}$ *is skew and linear. Then* $A \pm B$ *is maximal monotone.*

We consider the following problem

$$\text{find } x \in H \text{ such that } 0 \in A(x) - B(x). \tag{$\mathcal{K}$}$$

$$\text{find } y \in H \text{ such that } 0 = A_\lambda(x) - B_\lambda(x). \tag{$\mathcal{L}$}$$

$$\text{find } x \in H \text{ such that } 0 \in A^{-1}(y) - B^{-1}(y) \tag{$\mathcal{M}$}$$

Let us first compare the zeroes of $\mathcal{K}$ and $\mathcal{L}$. It is easy to verify that

Proposition 34. *If x is a solution to$\mathcal{L}$, then $y = J_\lambda^A(x)$ is a solution to $\mathcal{K}$.*

Proof. Since x is a solution of the regularized problem $\mathcal{L}$, we have :$A_\lambda(x) = B_\lambda(x)$. Using the definition of the Yosida approximation, this is equivalent to:$J_\lambda^A(x) = J_\lambda^B(x)$. Let:$y = J_\lambda^A(x)$ so: $y = J_\lambda^B(x)$ as well.

From the properties 26, we know: $A_\lambda(x) \in A(J_\lambda^A(x))$, and $\quad B_\lambda(x) \in B(J_\lambda^B(x))$. Since

$$J_\lambda^A(x) = J_\lambda^B(x) = y \implies A_\lambda(x) \in A(y), \quad B_\lambda(x) \in B(y).$$

$$A_\lambda(x) = B_\lambda(x),$$
$$0 \in A(y) - B(y)$$

This means that y satisfies the original problem $A_\lambda(x) - B_\lambda(x)$ ...($\mathcal{K}$):

$$A_\lambda(x) \in A(J_\lambda^B(x)) \quad \text{and} \quad A_\lambda(x) = B_\lambda(x) \in B(J_\lambda^B(x)) = B(J_\lambda^A(x)). \tag{3.1}$$

□

It should be mentioned that in the particular case, where $\lambda = 1$, F is a single-valued continuous mapping, $A = N_C$ (the normal cone to a closed convex set C) and $P_C := (I + N_C)^{-1}$ is the metric projector on C, relation 3.1 is nothing but

$$C(P_C(x)) + (x - P_C(x)) = 0, \tag{3.2}$$

a problem which arises frequently in optimization and equilibrium analysis and is related to the so-called WienerHopf equations (normal mapping).

Now, if F is a set-valued map, $\mathcal{K}$ can be viewed as a modification of the problem of finding the zeroes of F in such a way that the operator $N_C - F$ satisfies Aubins tangential condition, namely

$$\forall x \in C, \quad (F(x) - N_C(x)) \cap T_C(x) \neq \emptyset, \tag{3.3}$$

where $T_C(x)$ is the negative polar cone of $N_C(x)$.

Moreover, by the definition of the normal cone to C at x, the inclusion $\mathcal{K}$ in this case is equivalent to the variational inequality

$$\text{find } x \in C, \text{ such that } \exists v \in B(x) \text{ with } \langle v, z - x \rangle \leq 0, \quad \forall z \in C. \tag{3.4}$$

For the finite-dimensional situation, B is an upper semicontinuous set-valued map on int$C$ with convex compact values. Thus, if B is compact, by Aubin's theorem ([1], theorem 9.9), the variational inequality possesses a solution on int$C$ which inour situation is an operator zero B. The following result is immediate in proof.

Proposition 35. *[6] If x is a solution to $\mathcal{K}$, then $x+\lambda z$ is a solution to $\mathcal{L}$, where z is a solution to $\mathcal{M}$. Thus, the solutions $x_{\mathcal{K}}, x_{\mathcal{L}}$, and $x_{\mathcal{M}}$ of problems $(\mathcal{K}), (\mathcal{L})$, and $(\mathcal{M})$ are related by*

$$x_{\mathcal{L}} = \lambda x_{\mathcal{M}} + x_{\mathcal{K}}. \tag{3.5}$$

This clearly shows that $\lim_{\lambda\to 0} x_{\mathcal{L}} = x_{\mathcal{K}}$.

Now, let us say something about the uniqueness of solutions.

Proposition 36. *[6] If $(\mathcal{M})$ admits a unique solution and, in addition, A is strongly monotone with constant α, then $(\mathcal{K})$ has a unique solution, and so does $(\mathcal{L})$.*

Proof. Let x be the solution of $(\mathcal{M})$, namely

$$0 \in B^{-1}(x) - A^{-1}(x).$$

Then there exists $z \in \mathcal{H}$, such that $z \in B^{-1}(x) \cap A^{-1}(x)$, or equivalently,

$$x \in B(z) \cap A(z),$$

such that z solves $(\mathcal{K})$. Consequently, $\lambda x + z$ solves $(\mathcal{L})$.

Assume that $(\mathcal{L})$ has two solutions

$$y_1 = \lambda x + z_1 \quad \text{and} \quad y_2 = \lambda x + z_2.$$

Then $J^B_\lambda y_1$ and $J^B_\lambda y_2$ are the solutions of $(\mathcal{K})$, and $B_\lambda(y_1)$ and $B_\lambda(y_2)$ solve $(\mathcal{M})$.

The uniqueness of the solution of $(\mathcal{M})$ ensures that

$$B_\lambda(y_1) = B_\lambda(y_2),$$

or equivalently,

$$y_1 - y_2 = J^B_\lambda y_1 - J^B_\lambda y_2.$$

Using the α-strong monotonicity of B and taking into account the fact that its resolvent is $\frac{1}{1+\lambda\alpha}$-Lipschitz continuous, we infer that

$$\|y_1 - y_2\| \leq \frac{1}{1+\lambda\alpha}\|y_1 - y_2\|.$$

This implies that

$$y_1 = y_2 \quad \text{and} \quad z_1 = z_2.$$

This completes the proof. □

Example 37. *Let $\mathcal{H} = \mathbb{R}$, $A = I$, and $B = \partial|x|$, namely:*

$$B(x) = \operatorname{sgn}(x)$$

in particular,

$$B(0) = \{-1, +1\}.$$

The critical points of

$$C := I - \partial|x|$$

are $-1, 0$, *and* 1.
The resolvent operators of A and B are given, respectively, by:

$$\forall x \in \mathbb{R}, \quad J_\lambda^A(x) = \frac{x}{1+\lambda}$$

and

$$J_\lambda^B(x) = \begin{cases} 0, & |x| \leq \lambda, \\ x-\lambda, & x \geq \lambda, \\ x+\lambda, & x \leq -\lambda. \end{cases}$$

Thus, the critical points of $A_\lambda B_\lambda$ *are* $-1-\lambda, 0$, *and* $1+\lambda$, *and those of the dual problem are* $-1, 0, 1$.

Example 38. *Let C be a nonempty closed convex subset of* $\mathcal{H}$, *let A be a maximal monotone operator, and consider the problem:*

$$0 \in N_C(x) - A(x).$$

According to our methodology, a way of approximating these critical points is to consider the critical points of the regularized version of $N_C(x) - A(x)$. *That is,*

$$\frac{1}{2\lambda}\partial d_C^2(x) + A_\lambda(x),$$

where d_λ *denotes the distance function to the set C. The point x is a critical point of the initial operator if and only if:*

$$J_A^\lambda(x) = P_C(x),$$

where $P_C(x)$ *is the projection of x onto C.*

Theorem 39 (Non-Monotonicity of the Difference)**.** *Let A and B be two maximal monotone operators on a Hilbert space* $\mathcal{H}$. *Then, the difference* $A-B$ *is not necessarily monotone, even if* $A = \partial f$ *and* $B = \partial g$ *for convex, lower semicontinuous (l.s.c.) functions f and g.*

Proof. The standard counter example takes $\mathcal{H} = \mathbb{R}$, $A = \partial(\frac{1}{2}x^2)$ and $B = \partial|x|$. Then $(A-B)(x) = x - \mathrm{Sgn}(x)$ fails to be monotone at $x = 0$. □

Example 40. *Define* $f, g : \mathbb{R}^2 \to \mathbb{R}$ *as:*

$$f(x,y) = \frac{x^2}{2} + \frac{y^2}{2}, \quad g(x,y) = |x| + |y|.$$

Their subdifferentials are:

$$\partial f(x,y) = \{(x,y)\}, \quad \partial g(x,y) = \partial|x| \times \partial|y|.$$

the difference $\partial f - \partial g$ *is:*

$$(\partial f - \partial g)(x,y) = \begin{cases} (x - sgn(x),\, y - sgn(y)) & \text{if } (x,y) \neq (0,0), \\ (x - [-1,1],\, y - [-1,1]) & \text{if } (x,y) = (0,0). \end{cases}$$

for Monotonicity
Let $a = (1,0)$ and $b = (0,1)$:

$$\begin{aligned}(\partial f - \partial g)(1,0) &= (1-1, 0-0)\\ &= (0,0),\\ (\partial f - \partial g)(0,1) &= (0-0, 1-1)\\ &= (0,0).\end{aligned}$$

This clearly satisfies monotonicity.
Now let $a = (1,1)$ and $b = (-1,-1)$:

$$\begin{aligned}(\partial f - \partial g)(1,1) &= (1-1, 1-1)\\ &= (0,0),\\ (\partial f - \partial g)(-1,-1) &= (-1+1, -1+1)\\ &= (0,0).\end{aligned}$$

Monotonicity holds once more.

Theorem 41. *Let $\mathcal{H}$ be a Hilbert space, and let $A : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximal monotone operator. Let $B : \mathcal{H} \to \mathcal{H}$ satisfy the following conditions:*

- *B is L-Lipschitz continuous: $\|Bx - By\| \leq L\|x-y\| \quad$ for all $x, y \in \mathcal{H}$,*
- *B is monotone: $\langle Bx - By, x-y\rangle \geq 0 \quad$ for all $x, y \in \mathcal{H}$.*

Then, the operator $A - B$ is maximal monotone.

Proof. We prove maximal monotonicity by verifying Minty's surjectivity condition.

Lemma 42 (Minty's Theorem)**.** *An operator T is maximal monotone if and only if $R(I + \lambda T) = \mathcal{H}$ for all $\lambda > 0$.*

Let $\lambda > 0$ and $y \in \mathcal{H}$. We need to show there exists $x \in \mathcal{H}$ such that

$$y \in (I + \lambda(A - B))x \tag{3.6}$$

which is equivalent to

$$y + \lambda Bx \in (I + \lambda A)x \tag{3.7}$$

Define the operator $T_\lambda(x) := (I + \lambda A)^{-1}(y + \lambda Bx)$.

1. **Contraction Property**: For any $x_1, x_2 \in \mathcal{H}$,

$$\begin{aligned}\|T_\lambda(x_1) - T_\lambda(x_2)\| &= \|(I+\lambda A)^{-1}(y + \lambda Bx_1) - (I+\lambda A)^{-1}(y+\lambda Bx_2)\|\\ &\leq \|\lambda Bx_1 - \lambda Bx_2\| \quad \text{(since } (I+\lambda A)^{-1} \text{ is nonexpansive)}\\ &\leq \lambda L\|x_1 - x_2\|\end{aligned}$$

Thus T_λ is a contraction when $\lambda L < 1$.

2. **Fixed Point Existence**: For $\lambda < 1/L$, by the Banach fixed-point theorem, there exists a unique $x_\lambda \in \mathcal{H}$ such that
$$x_\lambda = (I + \lambda A)^{-1}(y + \lambda B x_\lambda)$$
This gives $y \in (I + \lambda(A - B))x_\lambda$.

3. **Extension to All $\lambda > 0$**: For $\lambda \geq 1/L$, take $\mu < 1/L$ and iterate:
$$(A - B) = (A - B_\mu) - (B - B_\mu)$$
where B_μ is a μ-averaged version of B. The result follows by perturbation arguments.

Therefore, $R(I + \lambda(A - B)) = \mathcal{H}$ for all $\lambda > 0$, proving $A - B$ is maximal monotone. □

Theorem 43 (Strong Monotonicity Preservation). *Let A and B be maximal monotone operators on a Hilbert space $\mathcal{H}$. If A is μ-strongly monotone for some $\mu > 0$, then:*

1. *$A - B$ is monotone*
2. *$A - B$ is coercive when B is L-Lipschitz continuous with $L < \mu$*

Proof. **Part 1: Monotonicity of $A - B$**
Take arbitrary $(x, a^* - b^*), (x', a'^* - b'^*) \in \text{Graph}(A - B)$ where:

- $a^* \in A(x)$, $a'^* \in A(x')$
- $b^* \in B(x)$, $b'^* \in B(x')$

The monotonicity condition requires:

$$\langle (a^* - b^*) - (a'^* - b'^*), x - x' \rangle = \underbrace{\langle a^* - a'^*, x - x' \rangle}_{\text{(strong monotonicity)}} - \underbrace{\langle b^* - b'^*, x - x' \rangle}_{\text{(monotonicity of } B\text{)}}$$
$$\geq \mu \|x - x'\|^2 \geq 0$$

Part 2: Coercivity when B is Lipschitz
Assume B is L-Lipschitz ($L < \mu$). For any $x \in \text{dom} A$:

$$\begin{aligned}
\frac{\langle (A - B)x, x \rangle}{\|x\|} &= \frac{\langle a^* - b^*, x \rangle}{\|x\|} \quad \text{for some } a^* \in A(x), b^* \in B(x) \\
&\geq \frac{\langle a^*, x \rangle - \|b^*\| \|x\|}{\|x\|} \\
&\geq \frac{\mu \|x\|^2 - L\|x\|^2}{\|x\|} \quad \text{(using strong monotonicity and Lipschitzness)} \\
&= (\mu - L)\|x\| \to \infty \text{ as } \|x\| \to \infty
\end{aligned}$$

Lemma 44 (Technical Justification). *The inequality $\langle a^*, x \rangle \geq \mu \|x\|^2$ follows from strong monotonicity by taking $x' = 0$ and $a'^* \in A(0)$. The Lipschitz condition gives $\|b^* - b'^*\| \leq L\|x - x'\|$.*

□

3.2 Applications in Optimization Problems

Here, you connect the theory to real-world and algorithmic applications, especially in convex optimization, variational inequalities, and evolution equations.

3.2.1 DC (Difference of Convex) Programming:

A Difference-of-Convex (DC) program is an optimization problem of the following form:

$$\min_{x\in\mathbb{R}^n} f(x) = g(x) - h(x) \quad \text{subject to } x \in C,$$

where:

- $g(x)$ and $h(x)$ are convex functions,
- $C \subseteq \mathbb{R}^n$ is a convex set.

Key Properties

- DC problems are generally non-convex but have a special structure that can be exploited by specialized algorithms, such as the DC Algorithm (DCA).
- Any continuous function with a convex domain can be approximated by a DC function through DC decomposition.

Proximal DC Algorithm (DCA)

Let $\gamma > 0$ be a step size (or implicitly taken up by prox). Beginning at. Starting from an initial point x_0, we iterate using:

$$x_{k+1} = \text{prox}_{\gamma f}\left(x_k - \gamma\nabla g(x_k)\right)$$

This update step solves the following optimization problem:

$$x_{k+1} = \arg\min_x \left\{ f(x) + \frac{1}{2\gamma}\|x - (x_k - \gamma\nabla g(x_k))\|^2 \right\}$$

This is a proximal step on$f$, with a forward (gradient) step on g.

Example 45. *Let:*

$$f(x) = \|x\|_1, \quad g(x) = \|x\|_2$$

and

$$\min_x \|x\|_1 - \|x\|_2$$

Then the gradient of g is:

$$\nabla g(x) = \frac{x}{\|x\|_2} \quad \textit{for } x \neq 0$$

Using the proximal gradient update, the iteration becomes:

$$x_{k+1} = \text{prox}_{\gamma\|\cdot\|_1}\left(x_k + \gamma\frac{x_k}{\|x_k\|_2}\right)$$

3.2.2 Variational Inequalities & Equilibrium Problems:

The difference of two maximal monotone operators is commonly used to express variational inequalities (VIs) and equilibrium problems. These problems often involve finding a point where two competing forces (represented by the two operators) balance out

(a) **Variational Inequalities**

Let $C \subseteq \mathbb{R}^n$ be a nonempty, closed, convex set and let $F : C \to \mathbb{R}^n$ be a given operator (often a vector field or gradient).
Find $x^* \in C$ such that:

$$\langle F(x^*), y - x^* \rangle \geq 0 \quad \forall y \in C$$

In simple terms, a variational inequality looks for a point $x \in C$ such that no point y in C can make the quantity $\langle F(x), y - x \rangle$ negative.

(b) **Equilibrium Problems**

Equilibrium problems are a class of mathematical problems that focus on finding a balance or equilibrium between various interacting components or agents in a system. An equilibrium point is one where no component or agent has an incentive to change their strategy or position. An equilibrium problem can be formulated as finding a point x such that:

$$\langle F(x), y - x \rangle \geq 0 \quad \forall y \in C,$$

where $F(x)$ represents the forces or responses in the system C is the set of feasible solutions or constraints

(c) **Relationship Between Variational Inequalities and Equilibrium Problems:**

Equilibrium problems are a special class of variational inequalities. Assuming the system in equilibrium, let the function $F(x)$ represent the balancing response or forces of the system. Then the condition of equilibrium as a variational inequality can be stated.

namely, each equilibrium problem has the form as a variational inequality, however, a variational inequality in general is not an equilibrium problem.

3.2.3 Saddle-Point & Minimax Problems

(a) Saddle-Point problems

Let $L(x,y)$ be a function defined in a product space $X \times Y$. A saddle point (x^*, y^*) satisfies the inequality:

$$L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*) \quad \forall x \in X, y \in Y$$

That is :

- For fixed x^*, $L(x^*, y)$ has its maximum value when $y = y^*$
- For fixed y^*, $L(x, y^*)$ has its minimum value when $x = x^*$

Therefore, (x^*, y^*) is a saddle point of the function L.

(b) Minimax Problems

Let $f(x,y)$ be a real-valued function over sets $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$. The minimax problem seeks to solve:

$$\min_{x\in X}\max_{y\in Y} f(x,y)$$

This is often encountered in game theory, where x and y are strategies for two players. The maximin problem is defined as:

$$\max_{y\in Y}\min_{x\in X} f(x,y)$$

3.3 Conclusion

It has been shown that the difference of two maximal monotone operators is, in general, neither monotone nor maximal. This highlights the subtle structure of monotonicity under algebraic operations. Nevertheless, certain special cases preserve this property, particularly when one of the operators is linear or when their domains are appropriately aligned. The behavior of the graphs of these operators under subtraction reveals intricate interactions between their domains and ranges. This study naturally finds applications in the analysis of differential inclusions and evolution equations involving differences of maximal monotone operators. Moreover, regularity conditions such as Lipschitz continuity, boundedness, or coercivity may ensure that the difference retains behavior close to that of a maximal monotone operator. Finally, tools such as Fitzpatrick functions, as well as considerations of convexity and closure of operator graphs, offer promising avenues for further analysis. These results pave the way for future research, particularly in extending the framework to non-reflexive Banach spaces or in studying the behavior of associated resolvent approximations.

BIBLIOGRAPHY

[1] J. P. Aubin and H. Frankowska. *Set-Valued Analysis*. Birkhäuser, 1990.

[2] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[3] H. Brezis. *Opérateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert*, volume 5 of *Math. Studies*. North-Holland Ameri. Elsevier, 1973.

[4] A. Mohammady, M. Ramazannejad, Z. Bagheri, and R.J. Shahkoohi. Common zero points of two finite families of maximal monotone operators via iteration methods. *Iran. J. Math. Sci. Inform.*, 12(2):73–99, 2017.

[5] A. Moudafi. On the difference of two maximal monotone operators: Regularization and algorithmic approaches. *Applied Mathematics and Computation*, 222(2):446–456, 2008.

[6] A. Moudafi and M. Thera. Finding a zero of the sum of two maximal monotone operators. *J Optim Theory Appl.*, 94(2):425–448, 1997.

[7] Juan Peypouquet and Sylvain Sorin. Evolution equations for maximal monotone operators: asymptotic analysis in continuous and discrete time. *Journal of Convex Analysis*, 7(3):1113–1163, 2010.

[8] R. T. Rockafellar and R. J-B. Wets. *Variational Analysis*. Grundlehren der mathematischen Wissenschaften. Springer, 1998.

[9] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.