A parameterized Douglas-Rachford algorithm: theory and applications

by

Dongying Wang

B.Sc., Xi'an Jiaotong-Liverpool University, China, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE COLLEGE OF GRADUATE STUDIES

(Mathematics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

August 2017

The undersigned certify that they have read, and recommend to the College of Graduate Studies for acceptance, a thesis entitled: A PARAMETERIZED DOUGLAS-RACHFORD ALGORITHM: THEORY AND APPLICATIONS submitted by DONGYING WANG in partial fulfilment of the requirements of the degree of Master of Science

Dr. Shawn Wang, Irving K. Barber School of Arts and Sciences
Supervisor, Professor

Dr. Heinz Bauschke, Irving K. Barber School of Arts and Sciences
Supervisory Committee Member, Professor

Dr. Julian Cheng, School of Engineering
Supervisory Committee Member, Professor

Dr. Philip D. Loewen, University of British Columbia (Vancouver)
University Examiner, Professor

August 17, 2017
(Date Submitted to Grad Studies)

Abstract

Douglas-Rachford algorithm is important due to its applications on the Heron problem and on the image denoising. Mathematically, it can be considered as finding a point such that the point belongs to a zero set of the sum of two maximally monotone operators.
In this thesis, previous work on Douglas-Rachford algorithm is presented and the Douglas-Rachford algorithm with a changed parameter is considered. I give it the name "α-Douglas-Rachford algorithm". The new algorithm which has the changed parameter is shown to have a convergent result and other conclusions similar to those of the classic Douglas-Rachford algorithm. At the same time, it has been shown that the application of the α-Douglas-Rachford algorithm is wider than the application of the classic one.
Later on, the α-Douglas-Rachford algorithm is proved to converge to the solution of the composited monotone inclusion problems, and in a special-limit case, it has some other properties. The numerical experiments confirm that the α-Douglas-Rachford algorithm does have the properties that I proved theoretically.

Table of Contents

List of Tables

List of Figures

Acknowledgements

First and foremost, I thank my supervisor Dr. Shawn Wang for his patient instruction in mathematics and his support for my study in UBC (Okanagan campus). His support and patience go beyond the work in this thesis. I would also like to thank Dr. Heinz Bauschke and Julian Cheng for agreeing to serve on my thesis committee and for their constructive comments.
I thank Dr. Andrew Jirasek, Dr. John Braun, Dr. Qiduan Yang, and Dr. Rebecca Tyson, for sharing their knowledge and experience in statistics, algebra, and mathematical biology.
I recognize the support of Unit 5 members and my colleagues in the COCANA lab at UBC (Okanagan campus), who had helped me during my study. In particular, I thank Chayne Planiden for his comments on my English.
I thank UBC (Okanagan campus) for providing me with financial support over the years of my master's studies.
Finally, I thank my parents and my grandparents, whose love and belief in me are the ultimate encouragement I have.

Dedication

To my parents and my grandparents.

Chapter 1

Introduction

In this chapter, we will introduce some background materials on inner product spaces and some necessary convex analysis used later in the thesis.

1.1 Inner product space

Definition 1.1.1. A vector space consists of a set V with elements called vectors, along with two operations such that the following properties hold:

1 Vector addition: Let $u, v \in V$, then there is a vector $u + v \in V$ and the following are satisfied.

  i Commutativity: $u + v = v + u, \quad \forall u, v, w \in V$.

  ii Associativity: $u + (v + w) = (u + v) + w, \forall u, v, w \in V$.

  iii Zero: there is a vector $\mathbf{0} \in V$ such that $\mathbf{0} + u = u = u + \mathbf{0}, \forall u \in V$.

  iv Inverses, for each $u \in V$, there is a vector $-u$ such that $u + (-u) = \mathbf{0}$.

2 Scalar multiplication: Let $u, v \in V$ and $r, s \in \mathbb{R}$, then the following are satisfied.

  i Left distributivity: $(r + s)v = rv + sv$.

  ii Associativity: $r(sv) = (rs)v$.

  iii Right distributivity: $r(u + v) = ru + rv$.

  iv Neutral element: $1v = v$.

  v Absorbing element: $0v = \mathbf{0}$.

  vi Inverse neutral element:$(-1)v = -v$.

Example 1.1.2. The space $\mathbb{R}^n$ consists of vectors $v = (v_1, \dots, v_n)$ with $v_i \in \mathbb{R}$ for $1 \leq i \leq n$ and operations defined by

$$(u_1, \dots, u_n) + (v_1, \dots, v_n) := (u_1 + v_1, \dots, u_n + v_n);$$

$$r(v_1, \dots, v_n) := (rv_1, \dots, rv_n).$$

where $r \in \mathbb{R}$.

Definition 1.1.3. Let V be a vector space over $\mathbb{R}$, and let W be a subset of V. Then W is a subspace if:

(1) The zero vector, $\mathbf{0}$, is in W.

(2) If $\boldsymbol{u}$ and $\boldsymbol{v}$ are elements of W, then $\boldsymbol{u} + \boldsymbol{v}$ is an element of W.

(3) If $\boldsymbol{u}$ is an element of W and c is a scalar from $\mathbb{R}$, then the scalar multiple $c\boldsymbol{u}$ is an element of W.

Definition 1.1.4. A norm $\|\cdot\|$ on a vector space V over the field $\mathbb{R}$ is a function $V \to \mathbb{R}$ with the following properties:

(1) Positive definite: $\|x\| \geq 0$ for all $x \in V$ and $\|x\| = 0$ if and only if $x = 0$.

(2) Homogeneous: $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{R}$ and $x \in V$.

(3) Triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$.

Here, $(V, \|\cdot\|)$ is called a normed space.

Definition 1.1.5. An inner product on a vector space V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ with the following properties.

(1) Positive definite: $\langle x, x \rangle \geq 0$, and $\langle x, x \rangle = 0$ if and only if $x = 0$.

(2) Symmetry: $\langle x, y \rangle = \langle y, x \rangle$.

(3) Bilinearity: $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$ for $\alpha, \beta \in \mathbb{R}, x, y, z \in V$.

We call a vector space paired with an inner product and norm induced by $\|x\| := \sqrt{\langle x, x \rangle}$ an inner product space.

Example 1.1.6. In $\mathbb{R}^n$, define

$$(\forall x \in \mathbb{R}^n)(\forall y \in \mathbb{R}^n), \langle x, y \rangle := \sum_{i=1}^{n} x_i y_i.$$

Then $\mathbb{R}^n$ is an inner product space.

Fact 1.1.7. Let $x, y \in \mathbb{R}^n$, and $\langle x, z \rangle = \langle y, z \rangle$ for all $z \in \mathbb{R}^n$. Then $x = y$.

Proof. For all $z \in \mathbb{R}^n$, $\langle x, z \rangle = \langle y, z \rangle$ implies that $\langle x, x - y \rangle = \langle y, x - y \rangle$ (by setting $z = x - y$). Moreover,

$$\begin{aligned} & 0 = \langle x, x - y \rangle - \langle y, x - y \rangle \\ \Rightarrow\ & 0 = \langle x - y, x - y \rangle \\ \Rightarrow\ & 0 = \|x - y\|^2. \end{aligned}$$

Thus, $x = y$. $\square$

In this thesis, we use the Euclidean norm, which is given by

$$\|x\| = \sqrt{x^\intercal x} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

In the following thesis, $\|x\|$ refers to the Euclidean norm of x.

Fact 1.1.8. (Cauchy Schwarz Inequality) Let x and y be in $\mathbb{R}^n$. Then

$$|\langle x, y\rangle| \le \|x\| \cdot \|y\|, \quad i.e. \quad \left|\sum_{i=1}^{n} x_i y_i\right| \le \sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}.$$

Moreover, $\langle x, y\rangle = \|x\| \cdot \|y\| \Leftrightarrow \exists \alpha \in [0, +\infty)$ such that $x = \alpha y$ or $y = \alpha x$.

Definition 1.1.9. In a normed vector space $(V, \|\cdot\|)$, a sequence $(v_n)_{n=1}^{+\infty}$ is said to converge (or strongly converge) to a point $v \in V$ if $\forall \epsilon > 0, \exists N > 0$ such that $\|v_n - v\| < \epsilon$ for all $n \ge N$.

In the following thesis, $(x_n)_{n=1}^{+\infty} \to x$ denotes the sequence $(x_n)_{n=1}^{+\infty}$ converges (or strongly converges) to x. We also write it as $x_n \to x$.

Definition 1.1.10. Let $(x_n)_{n=1}^{+\infty}$ be a sequence in a vector space $(V, \|\cdot\|)$, let $(n_k)_{k=1}^{+\infty}$ be a strictly increasing sequence in $\mathbb{N}$. Then the sequence $(x_{n_k})_{k=1}^{+\infty}$ is called a subsequence of $(x_n)_{n=1}^{+\infty}$.

Definition 1.1.11. A sequence $(x_n)_{n=1}^{+\infty}$ is bounded if $\exists M > 0$ such that $\|x_n\| \le M, \forall n \ge 1$.

Fact 1.1.12. (Bolzano-Weierstrass Theorem) Every bounded sequence $(x_n)_{n=1}^{+\infty}$ in $\mathbb{R}^m$ has a convergent subsequence, i.e., there exists a subsequence $(x_{n_k})_{k=1}^{+\infty}$ of $(x_n)_{n=1}^{+\infty}$ such that $x_{n_k} \to x$, for some $x \in \mathbb{R}^m$.

Definition 1.1.13. A sequence $(x_n)_{n=1}^{+\infty}$ is called a Cauchy sequence if for every $\epsilon > 0$, there exists an integer $N > 0$ such that $\|x_n - x_k\| < \epsilon$ for all $n, k > N$.

Fact 1.1.14. Let $(x_n)_{n=1}^{+\infty}$ be a Cauchy sequence in a normed vector space $(V, \|\cdot\|)$. Let $x \in V$. Then $(x_n)_{n=1}^{+\infty}$ converges to x if and only if it has a subsequence that converges to x.

Proof. We separate this proof into two parts:

(1) " $\Rightarrow$ " If the Cauchy sequence $(x_n)_{n=1}^{+\infty} \to x$, it follows that $(x_n)_{n=1}^{+\infty}$ is a subsequence of itself which converges to x.

(2) " $\Leftarrow$ " Suppose $(x_{n_k})_{k=1}^{+\infty}$ is a subsequence of $(x_n)_{n=1}^{+\infty}$ and converges to x. Then for all $\epsilon > 0, \exists N_1 > 0 (N_1 \in \mathbb{N})$ such that $\|x_{n_k} - x\| < \frac{\epsilon}{2}, \forall k > N_1$.

Since $(x_n)_{n=1}^{+\infty}$ is a Cauchy sequence, $\forall \epsilon > 0, \exists N_2 > 0 (N_2 \in \mathbb{N})$ such that $\|x_m - x_{n_k}\| < \frac{\epsilon}{2}, \forall m, k > N_2$.

Let $M = \max(N_1, N_2)$, we have $n_M > M$. Then $\forall \epsilon > 0, m > M$:

$$\begin{aligned}\|x_m - x\| &\leq \|x_m - x_{n_M}\| + \|x_{n_M} - x\| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon.\end{aligned}$$

That is, $(x_n)_{n=1}^{+\infty}$ converges to x.

□

Definition 1.1.15. An inner product space $\mathcal{H}$ is called complete, or a Hilbert Space, if each Cauchy sequence in $\mathcal{H}$ converges to a point in $\mathcal{H}$.

In the following thesis, $\mathcal{H}$ denotes a Hilbert space.

Example 1.1.16. $\mathbb{R}^m$ is complete. Thus, $\mathbb{R}^m$ is a Hilbert space.

Proof. Suppose $(x_n)_{n=1}^{+\infty}$ is a Cauchy sequence in $\mathbb{R}^m$, we want to prove its convergence. According to Fact 1.1.14, we only need to prove it has a subsequence which is convergent.
Since $(x_n)_{n=1}^{+\infty}$ is a Cauchy sequence, let $\epsilon = 1$. Then, there exists $N \in \mathbb{N}$ such that for all $m, k > N$, $\|x_m - x_k\| < 1$. Thus, for all $k > N$,

$$\begin{aligned}\|x_k\| &= \|x_k - x_{N+1} + x_{N+1}\| \\ &\leq \|x_k - x_{N+1}\| + \|x_{N+1}\| \\ &< 1 + \|x_{N+1}\|.\end{aligned}$$

Let $M = \max(\|x_1\|, \|x_2\|, \ldots, \|x_N\|, \|x_{N+1}\|, 1 + \|x_{N+1}\|)$, then for all $k \geq 1$, $\|x_k\| \leq M$. According to the definition of the bounded sequence, we find $(x_n)_{n=1}^{+\infty}$ is bounded. By using Bolzano-Weierstrass theorem, $(x_n)_{n=1}^{+\infty}$ has a convergent subsequence, which completes the proof. □

Definition 1.1.17. A sequence $(x_n)_{n=1}^{+\infty}$ in a Hilbert space $\mathcal{H}$ is said to converge weakly to a point $v \in \mathcal{H}$ if for all $y \in \mathcal{H}$, $\langle v_n, y \rangle \to \langle v, y \rangle$.

Fact 1.1.18. [4, Lemma 2.51] Let $(x_n)_{n=1}^{+\infty}$ and $(u_n)_{n=1}^{+\infty}$ be sequences in $\mathcal{H}$, and let x and u be points in $\mathcal{H}$. Then the following hold:

(1) Suppose that $\mathcal{H}$ is finite-dimensional. Then $x_n \rightharpoonup x \Leftrightarrow x_n \to x$.

(2) Suppose that $x_n \rightharpoonup x$ and $u_n \to u$. Then $\langle x_n, u_n\rangle \to \langle x, u\rangle$.

Fact 1.1.19. [4, Lemma 2.46] Let $(x_n)_{n=1}^{+\infty}$ be a sequence in $\mathbb{R}^m$. Then $(x_n)_{n=1}^{+\infty}$ converges if and only if it is bounded and possesses at most one sequential cluster point.

Definition 1.1.20. Let l^2 be a space such that each element in it is a sequence $x = (\zeta_j)_{j=1}^{+\infty} = (\zeta_1, \zeta_2, \ldots)$ of numbers such that

$$\sum_{j=1}^{+\infty} |\zeta_j|^2 < +\infty,$$

and its distance function is defined by

$$d(x, y) = \sqrt{\sum_{j=1}^{+\infty} |\zeta_j - \mu_j|^2},$$

where $y = (\mu_j)_{j=1}^{+\infty}$ and

$$\sum_{j=1}^{+\infty} |\mu_j|^2 < +\infty.$$

Remark 1.1. Let $(x_n)_{n=1}^{+\infty}$ and $(u_n)_{n=1}^{+\infty}$ be weakly convergent sequences in l^2, $x_n \rightharpoonup x$ and $u_n \rightharpoonup u$ do not imply $\langle x_n, u_n\rangle \to \langle x, u\rangle$. A counter example is: let $(x_n)_{n=1}^{+\infty} = (e_n)_{n=1}^{+\infty}$ and $(u_n)_{n=1}^{+\infty} = (e_n)_{n=1}^{+\infty}$, where $(e_n)_{n=1}^{+\infty}$ is an orthonormal sequence in l^2. We have

$$\langle x_n, u_n\rangle = \langle e_n, e_n\rangle = 1;$$

while

$$x_n \rightharpoonup 0, u_n \rightharpoonup 0 \text{ and } \langle 0, 0\rangle = 0.$$

Remark 1.2. We show $e_n \rightharpoonup 0$. We need $\langle e_n, x\rangle \to \langle 0, x\rangle\ \forall x \in l^2$ as $n \to +\infty$. Because $x \in l^2$, let $x = (\zeta_n)_{n=1}^{+\infty}$, we have

$$\sum_{n=1}^{+\infty} |\zeta_n|^2 < +\infty \Rightarrow \zeta_n^2 \to 0 \Rightarrow \zeta_n \to 0 \text{ as } n \to +\infty.$$

Therefore, $\lim\limits_{n\to+\infty} \zeta_n = 0$. So

$$\lim_{n\to+\infty} \langle e_n, x\rangle = \zeta_n = 0 = \langle 0, x\rangle$$

for all $x \in l^2$. Hence $e_n \rightharpoonup 0$.

1.1.1 Sets in vector spaces

Definition 1.1.21. A set $C \subseteq \mathbb{R}^m$ is ***closed*** if it contains all limit points, i.e., whenever there exists a sequence $(x_k)_{k=1}^{\infty}$ in C and $x_k \to x$, then $x \in C$.

Fact 1.1.22. For a Hilbert space $\mathcal{H}$, let S be a subspace of $\mathcal{H}$. Then S is closed if and only if S is complete.

Proof. (1) $\Rightarrow$ Let $(x_n)_{n=1}^{+\infty}$ be a Cauchy sequence in S. Because the Hilbert space $\mathcal{H}$ is complete, $(x_n)_{n=1}^{+\infty}$ must converge to some $x \in \mathcal{H}$. However, as S is closed, $x \in S$. Thus, S is complete.

(2) $\Leftarrow$ Let $x \in \bar{S}$. Then there exists a sequence $(x_n)_{n=1}^{+\infty} \in S$ that converges to x. Since a convergent sequence must be a Cauchy sequence, moreover, since S is complete, $(x_n)_{n=1}^{+\infty} \in S$ must converge to a point included in S. Because a convergent sequence cannot converge to more than one point, we have $x \in S$. Thus, S is closed.

□

Definition 1.1.23. A set $O \subseteq \mathbb{R}^m$ is ***open*** if $\forall x \in O, \exists r > 0$ such that the open ball $\mathbb{B}(x; r) \subseteq O$, where

$$\mathbb{B}(x; r) = \{y \in \mathbb{R}^m : \|y - x\| < r\}.$$

Definition 1.1.24. The interior of a subset C of $\mathcal{H}$ can be expressed as

$$\operatorname{int} C := \{x \in C | (\exists r \in \mathbb{R}_{++}) \quad \mathbb{B}(0; r) \subset C - x\}.$$

Definition 1.1.25. The orthogonal complement of a subset C of $\mathcal{H}$ is denoted by $C^{\perp}$, i.e.,

$$C^{\perp} := \{u \in \mathcal{H} | (\forall x \in C) \quad \langle x, u \rangle = 0\}.$$

Example 1.1.26. For any Hilbert space $\mathcal{H}$, $\mathcal{H}^{\perp} = \{0\}$.

Proof. (1) $\{0\} \subseteq \mathcal{H}^{\perp}$ is clear.

(2) Suppose there exists a $u \neq 0$ such that $u \in \mathcal{H}^{\perp}$. According to the definition of $\mathcal{H}^{\perp}$, for any $x \in \mathcal{H}$, $\langle x, u \rangle = 0$. However, since $u \in \mathcal{H}$ and $u \neq 0$, we have $\langle u, u \rangle \neq 0$, which is a contradiction. Therefore, u must equal 0.

Altogether, we have $\mathcal{H}^{\perp} = \{0\}$. □

Fact 1.1.27. Let C and D be two subsets of $\mathcal{H}$. Then $D^{\perp} \subseteq C^{\perp}$ if $C \subseteq D$.

Proof. Let $u \in D^{\perp}$. Then u is a vector in $\mathcal{H}$ such that for all $x \in D$, $\langle x, u \rangle = 0$. Since $C \subseteq D$, all $y \in C$ also contained in D. Thus we have $\forall y \in C$, $\langle y, u \rangle = 0$. Therefore, $u \in C^{\perp}$ and so $D^{\perp} \subseteq C^{\perp}$. □

Fact 1.1.28. [12, Lemma 3.3-6] If S is a closed subspace of a Hilbert space $\mathcal{H}$, then

$$S = S^{\perp\perp}.$$

Definition 1.1.29. A set $C \subseteq \mathbb{R}^m$ is convex if for any $x, y \in C$ and $\alpha \in (0,1)$ we have

$$\alpha x + (1-\alpha)y \in C.$$

Graphically, a set C is convex if the line segment between any two points in C is also contained in C, see Figure 1.1.

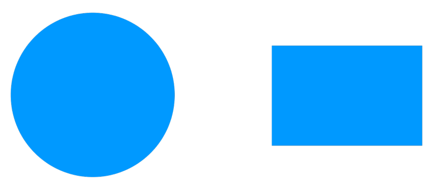

Figure 1.1: Examples of convex sets

Figure 1.2 shows two nonconvex sets.

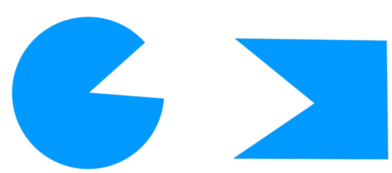

Figure 1.2: Examples of nonconvex sets

Example 1.1.30. Let $r \in \mathbb{R}_{++}$. Then the closed ball

$$B(c;r) = \{x \in \mathbb{R}^m : \|x - c\| \leq r\}$$

is convex.

Proof. Let $x, y \in B(c; r)$ and $a \in (0, 1)$. Then

$$\begin{aligned}\|ax + (1-a)y - c\| =& \|a(x-c) + (1-a)(y-c)\| \\ \leq& a\|x-c\| + (1-a)\|y-c\| \\ \leq& ar + (1-a)r = r.\end{aligned}$$

Therefore, $ax + (1-a)y \in B(c; r)$, which implies $B(c; r)$ is convex. □

Example 1.1.31. Let $a_i \leq b_i$ for all $i \in \{1, 2, \dots, m\}$. Then the box

$$C = \{x \in \mathbb{R}^m : a_i \leq x_i \leq b_i\}$$

is convex.

Proof. Let $x, y \in C$ and $a \in (0, 1)$. Then for all $i \in \{1, 2, \dots, m\}$,

$$\begin{aligned}& aa_i + (1-a)a_i \leq ax_i + (1-a)y_i \leq ab_i + (1-a)b_i \\ \Leftrightarrow & a_i \leq ax_i + (1-a)y_i \leq b_i.\end{aligned}$$

Therefore, $ax + (1-a)y \in C$, which implies C is convex. □

Definition 1.1.32. Let $A \subseteq \mathbb{R}$. The infimum of A is the largest lower bound and denoted by $\inf A$; the supremum of A is the smallest upper bound and denoted by $\sup A$.

Remark 1.3. When $A = \emptyset$, $\inf A = +\infty$ and $\sup A = -\infty$.

Definition 1.1.33. Let C be a nonempty convex subset of $\mathcal{H}$ and let $x \in \mathcal{H}$. The normal cone to C at x is

$$N_C x = \begin{cases} \{u \in \mathcal{H} | \sup\langle C - x, u\rangle \leq 0\}, & \text{if } x \in C; \\ \emptyset & \text{otherwise.} \end{cases}$$

Example 1.1.34. [4, Example 6.39] Let $C = B(0; 1)$ and let $x \in C$. Then

$$N_C x = \begin{cases} \mathbb{R}_+ x, & \text{if } \|x\| = 1; \\ \{0\} & \text{if } \|x\| < 1; \\ \emptyset & \text{if } \|x\| > 1. \end{cases}$$

Lemma 1.1.35. In $\mathbb{R}^m$, we have

$$N_{\{0\}} x = \begin{cases} \mathbb{R}^m, & \text{if } x = 0; \\ \emptyset & \text{if } x \neq 0. \end{cases}$$

Then, $\operatorname{dom} N_{\{0\}} = \{0\}$ and $\operatorname{ran} N_{\{0\}} = \mathbb{R}^m$.

Proof. According to the definition of the normal cone, $y \in N_{\{0\}}x$ if and only if $\sup\langle 0 - x, y\rangle \leq 0$. When $x = 0$, the inequality

$$\sup\langle 0 - 0, y\rangle \leq 0$$

is satisfied by all $y \in \mathbb{R}^m$. Therefore,

$$N_{\{0\}}x = \mathbb{R}^m \text{ if } x = 0.$$

That is, $\operatorname{dom} N_{\{0\}} = \{0\}$ and $\operatorname{ran} N_{\{0\}} = \mathbb{R}^m$. □

1.2 Operators

Definition 1.2.1. Let $M : \mathcal{H} \to 2^{\mathcal{H}}$ be a set-valued operator, the domain of M is

$$\operatorname{dom} M := \{x \in \mathcal{H} : Mx \neq \emptyset\};$$

the range of M is

$$\operatorname{ran} M := \{u \in \mathcal{H} : \exists x \in \mathcal{H}, u \in Mx\};$$

the graph of M is

$$\operatorname{gra} M := \{(x, u) \in \mathcal{H} \times \mathcal{H} : u \in Mx\};$$

the set of zeros of M is:

$$\operatorname{zer} M := \{x \in \mathcal{H} : 0 \in Mx\};$$

the set of fixed points of M is

$$\operatorname{Fix} M := \{x \in \mathcal{H} : x \in Mx\};$$

the inverse of M is

$$M^{-1} : \mathcal{H} \to 2^{\mathcal{H}} : u \mapsto \{x \in \mathcal{H} : u \in Mx\}.$$

Definition 1.2.2. Let $M_1, M_2 : \mathcal{H} \to 2^{\mathcal{H}}$,

(1) The sum of M_1, M_2 is defined as $(M_1 + M_2)(x) := M_1(x) + M_2(x)$ for all $x \in \mathcal{H}$;

(2) The parallel sum of M_1, M_2 is $M_1 \Box M_2 : \mathcal{H} \to 2^{\mathcal{H}}$, defined by

$$M_1 \Box M_2 := (M_1^{-1} + M_2^{-1})^{-1}.$$

Definition 1.2.3. The identity operator is denoted by $\operatorname{Id} : \mathbb{R}^m \to \mathbb{R}^m$, for which we have

$$(\forall x \in \mathbb{R}^m) \operatorname{Id} x = x.$$

Definition 1.2.4. Let $A : \mathbb{R}^m \to \mathbb{R}^n$. The induced norm or operator norm on $\mathbb{R}^{n \times m}$ is given by :

$$\|A\| := \sup \left\{ \frac{\|Ax\|}{\|x\|} : x \in \mathbb{R}^m \text{ with } x \neq 0 \right\}.$$

Example 1.2.5. On space $\mathbb{R}^m$,

$$\| \operatorname{Id} \| = \sup \left\{ \frac{\|x\|}{\|x\|} : x \in \mathbb{R}^m \text{ with } x \neq 0 \right\} = 1.$$

Definition 1.2.6. The distance to a set $C \subset \mathcal{H}$ is the function

$$d_C : \mathcal{H} \to [0, +\infty] : x \mapsto \inf_{y \in C} \|x - y\|.$$

Note that if $C = \emptyset$ then $d_C \equiv +\infty$.

Definition 1.2.7. Let C be a subset of $\mathcal{H}$, let $x \in \mathcal{H}$, and let $p \in C$. Then p is a projection of x onto C if $\|x - p\|$ equals to the distance between x and C, which denoted by d_C. If every point in $\mathcal{H}$ has at least one projection onto C, then C is proximinal. If every point in $\mathcal{H}$ has exactly one projection onto C, then C is a Chebyshev set. In this case, the projector onto C is the operator, denoted by P_C, that maps every point in $\mathcal{H}$ to its unique projection onto C.

Definition 1.2.8. An operator $M : \mathbb{R}^m \to \mathbb{R}^m$ is called ρ-strongly positive ($\rho \in \mathbb{R}_+$) if $\langle Mx, x \rangle \geq \rho \|x\|^2$.

Example 1.2.9. For all $\alpha \in \mathbb{R}_{++}$, the following two operators are ρ-strongly positive:

(1) $\alpha \operatorname{Id}$ is ρ-strongly positive for any $0 < \rho \leq \alpha$.

(2) Let $V : \mathbb{R}^2 \to \mathbb{R}^2 : (x_1, x_2) \to (\frac{x_1 + x_2}{\alpha}, \frac{x_2 - x_1}{\alpha})$. V is ρ-strongly positive for any $0 < \rho \leq \frac{1}{\alpha}$.

Proof. (1) For any $x \in \mathbb{R}^m$,

$$\langle \alpha \operatorname{Id} x, x \rangle = \alpha \|x\|^2.$$

Therefore, $\alpha \operatorname{Id}$ is ρ-strongly positive for any $0 < \rho \leq \alpha$.

(2) For any $x_1 \in \mathbb{R}, x_2 \in \mathbb{R}$,

$$\begin{aligned}\langle V(x_1,x_2),(x_1,x_2)\rangle &= \langle (\frac{x_1+x_2}{\alpha}, \frac{x_2-x_1}{\alpha}),(x_1,x_2)\rangle \\ &= \frac{x_1^2+x_2^2}{\alpha} \\ &= \frac{1}{\alpha}\|x\|^2.\end{aligned}$$

Therefore, V is ρ-strongly positive for any $0 < \rho \leq \frac{1}{\alpha}$.

□

1.2.1 Linear operators

Definition 1.2.10. An operator $L : \mathbb{R}^n \to \mathbb{R}^m$ is said to be linear if and only if

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$$

for all $x, y \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$.

Definition 1.2.11. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear operator. The adjoint of L is the unique linear operator $L^\star : \mathbb{R}^m \to \mathbb{R}^n$ that satisfies

$$(\forall x \in \mathbb{R}^n)(\forall y \in \mathbb{R}^m) \quad \langle Lx, y\rangle = \langle x, L^\star y\rangle.$$

Fact 1.2.12. [12, Theorem 3.9-2] The Hilbert-adjoint operator $L^\star$ of L in Definition 1.2.11 exists, is unique and is a bounded linear operator with norm

$$\|L^\star\| = \|L\|.$$

Fact 1.2.13. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear operator, let $\lambda \in \mathbb{R}$. Then

(1) λL is a linear operator.

(2) $(\lambda L)^\star = \lambda L^\star$.

Proof. (1) Clear.

(2)

$$\begin{aligned}\langle x, (\lambda L)^\star y\rangle &= \langle \lambda L x, y\rangle \\ &= \lambda \langle x, L^\star y\rangle \\ &= \langle x, \lambda L^\star y\rangle.\end{aligned}$$

Thus, $(\lambda L)^\star = \lambda L^\star$.

□

Fact 1.2.14. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear operator. Then $L^\star = L^\intercal$. In this thesis, we use $A^\intercal$ to denote the transpose of a matrix A.

Proof. Since $L : \mathbb{R}^n \to \mathbb{R}^m$, we set

$$L = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Then, let $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$; let $y = (y_1, y_2, \dots, y_m) \in \mathbb{R}^m$. We have:

$$\begin{aligned}
\langle Lx, y \rangle &= \left\langle \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \right\rangle \\
&= \left\langle \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \right\rangle \\
&= \sum_{i=1}^{m} (a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n) y_i \\
&= \sum_{j=1}^{n} (a_{1j}y_1 + a_{2j}y_2 + \dots + a_{mj}y_m) x_j \\
&= \left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} a_{11}y_1 + a_{21}2_2 + \dots + a_{m1}y_m \\ a_{12}y_1 + a_{22}y_2 + \dots + a_{m2}y_m \\ \vdots \\ a_{1n}y_1 + a_{2n}y_2 + \dots + a_{mn}y_m \end{pmatrix} \right\rangle \\
&= \left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \right\rangle.
\end{aligned}$$

Thus,

$$L^\star = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix} = L^\intercal.$$

□

Fact 1.2.15. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear operator. If L^{-1} exists, then L^{-1} is also a linear operator.

Proof. Let $x, y \in \mathbb{R}^n$, and let $\alpha, \beta \in \mathbb{R}$. Since L is a linear operator and $L^{-1}L = \mathrm{Id}$, we have

$$\begin{aligned} L^{-1}[\alpha L(x) + \beta L(y)] &= L^{-1}[L(\alpha x + \beta y)] \\ &= \alpha x + \beta y. \end{aligned} \tag{1.1}$$

Let $x' = L(x), y' = L(y)$, equation (1.1) becomes

$$L^{-1}(\alpha x' + \beta y') = \alpha L^{-1}(x') + \beta L^{-1}(y'),$$

that is, L^{-1} is a linear operator. □

Fact 1.2.16. [12, Theorem 2.4-2] Every finite dimensional subspace of a normed space is complete.

Fact 1.2.17. Let $T : \mathbb{R}^m \to \mathbb{R}^m$ be a linear operator. Then $\operatorname{ran} T$ is a subspace, so is closed. In other words, $\overline{\operatorname{ran}} T = \operatorname{ran} T$. Here we use $\overline{\operatorname{ran}} T$ to denote the closure of $\operatorname{ran} T$.

Proof. We set this proof into two parts.

(1) Prove $\operatorname{ran} T$ is a subspace of $\mathbb{R}^m$.

   (a) Since T is a linear operator, $T(0) = 0$. Thus, $0 \in \operatorname{ran} T$.

   (b) For any $u, v \in \operatorname{ran} T$, there must exists $x, y \in \mathbb{R}^m$ such that $Tx = u, Ty = v$. Meanwhile, $T(x + y) = Tx + Ty = u + v$. Thus, $u + v \in \operatorname{ran} T$.

   (c) For any $u \in \operatorname{ran} T$, there must exists $x \in \mathbb{R}^m$ such that $Tx = u$. Meanwhile, for any scalar c, $T(cx) = cT(x) = cu$. Thus, $cu \in \operatorname{ran} T$.

   Since $\operatorname{ran} T$ satisfies all the conditions of being a subspace, $\operatorname{ran} T$ is a subspace of $\mathbb{R}^m$.

(2) Since $\mathbb{R}^m$ is a finite dimensional complete space and $\operatorname{ran} T$ is a subspace of it, by using Fact 1.2.16 and Fact 1.1.22, we get $\operatorname{ran} T$ is closed.

□

Fact 1.2.18. [4, Fact 2.25] Let $T : \mathbb{R}^m \to \mathbb{R}^m$ be a linear operator. Then $(\operatorname{ran} T)^\perp = \operatorname{zer} T^\star$.

1.2.2 Nonexpansive operators

Definition 1.2.19. Let D be a nonempty subset of $\mathbb{R}^m$. Let $T : D \to \mathbb{R}^m$. Then T is

(1) nonexpansive if

$$\forall x \in D, \forall y \in D, \quad \|Tx - Ty\| \leq \|x - y\|;$$

(2) firmly nonexpansive if

$$\forall x \in D, \forall y \in D, \quad \|Tx - Ty\|^2 \leq \|x - y\|^2 - \|(\operatorname{Id} - T)x - (Id - T)y\|^2.$$

Example 1.2.20. For any $0 \leq \alpha \leq 1, \alpha \operatorname{Id}$ is firmly nonexpansive.

Proof. Since $0 \leq \alpha \leq 1$, we have $\alpha^2 - \alpha \leq 0$, which also implies

$$\alpha^2 + 1 + \alpha^2 - 2\alpha \leq 1.$$

That is, for any $x, y \in \mathbb{R}^m$, and $x \neq y$,

$$\begin{aligned} \alpha^2 + (1-\alpha)^2 \leq 1 &\Leftrightarrow \alpha^2\|x - y\|^2 \leq \|x - y\|^2 - (1-\alpha)^2\|x - y\|^2 \\ &\Leftrightarrow \|\alpha x - \alpha y\|^2 \leq \|x - y\|^2 - \|(1-\alpha)x - (1-\alpha)y\|^2. \end{aligned}$$

Thus, for any $0 < \alpha \leq 1, \alpha \operatorname{Id}$ is firmly nonexpansive. □

Fact 1.2.21. [4, Proposition 4.2] Let D be a nonempty subset of $\mathbb{R}^m$, let $T : D \to \mathcal{H}$.Then the following are equivalent:

(1) T is firmly nonexpansive.

(2) $\operatorname{Id} - T$ is firmly nonexpansive.

(3) $2T - \operatorname{Id}$ is firmly nonexpansive.

(4) $(\forall x \in D) \quad (\forall y \in D) \quad \|Tx - Ty\|^2 \leq \langle x - y, Tx - Ty\rangle$.

(5) $(\forall x \in D) \quad (\forall y \in D) \quad 0 \leq \langle Tx - Ty, (\operatorname{Id} - T)x - (\operatorname{Id} - T)y \rangle$.

(6) $(\forall x \in D) \quad (\forall y \in D) \quad (\forall \alpha \in [0,1]) \quad \|Tx - Ty\| \leq \|\alpha(x-y) + (1-\alpha)(Tx - Ty)\|$.

Fact 1.2.22. [4, Theorem 5.15] Let D be a nonempty closed convex subset of $\mathbb{R}^m$, let $T : D \to D$ be a nonexpansive operator such that $\operatorname{Fix} T \neq \emptyset$, where the fixed points set

$$\operatorname{Fix} T = \{x \in \mathbb{R}^m : Tx = x\}.$$

Let $(\lambda_n)_{n=1}^{+\infty}$ be a sequence in $[0,1]$ such that $\sum_{n=1}^{+\infty} \lambda_n(1-\lambda_n) = +\infty$, and let $x_0 \in D$. Set

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

Then the following hold:

(1) $(Tx_n - x_n)_{n=1}^{+\infty}$ converges to 0.

(2) $(x_n)_{n=1}^{+\infty}$ converges to a point in $\operatorname{Fix} T$.

Definition 1.2.23. Let D be a nonempty subset of $\mathcal{H}$, let $T : D \to \mathcal{H}$ be nonexpansive, and let $\gamma \in (0,1)$. Then T is averaged with constant γ, or $\gamma - averaged$, if there exists a nonexpansive operator $R : D \to \mathcal{H}$ such that

$$T = (1-\gamma) \operatorname{Id} + \gamma R.$$

1.2.3 Monotone operators

Definition 1.2.24. [4] An operator $M : \mathcal{H} \to 2^{\mathcal{H}}$ is *monotone* if $\langle x-y, u-v \rangle \geq 0$ for all $(x,u), (y,v) \in \operatorname{gra} M$.
M is a *maximally monotone* operator if there is no monotone operator whose graph properly contains $\operatorname{gra} M$.
M is a *strictly monotone* operator if

$$(\forall (x,u),(y,v) \in \operatorname{gra} M) \quad x \neq y \Rightarrow \langle x-y, u-v \rangle > 0 \tag{1.2}$$

M is a *uniformly monotone* operator if there exists an increasing function $\phi_M : \mathbb{R}_+ \to [0,+\infty]$ with $\phi_M(0) = 0$, and $\langle x-y, u-v \rangle \geq \phi_M(\|x-y\|)$ for all $(x,u),(y,v) \in \operatorname{gra} M$. When $\phi_M(\|x-y\|) = \|x-y\|^2$, M is called strongly monotone.

Remark 1.2.25. [4, Remark 22.3] The notions of strict, uniform, and strong monotonicity of $A : \mathcal{H} \to 2^{\mathcal{H}}$ can naturally be localized to a subset C of $\operatorname{dom} A$.

Proposition 1.2.26. If M is uniformly monotone on $\mathcal{H}$, then A is strictly monotone.

Proof. M is uniformly monotone on $\mathcal{H}$ implies there exists a continuous increasing function $\phi_M : \mathbb{R}_+ \to [0, +\infty]$ with $\phi_M(0) = 0$, and $\langle x-y, u-v\rangle \geq \phi_M(\|x-y\|)$ for all $(x, u), (y, v) \in \operatorname{gra} M$. Thus, in the case $x \neq y$, for all $(x, u), (y, v) \in \operatorname{gra} M$, we have

$$\langle x - y, u - v\rangle \geq \phi_M(\|x - y\|) \tag{1.3}$$

According to the definition of ϕ_M, (1.3) means $\phi_M(\|x - y\|) > 0$, so M is a strictly monotone operator. $\square$

Fact 1.2.27. [4, Proposition 23.35] Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be strictly monotone. Then $\operatorname{zer} A$ is at most a singleton.

Proof. Suppose $x \in \operatorname{zer} A, y \in \operatorname{zer} A, x \neq y$, i.e., $0 \in Ax, 0 \in Ay$. Since A is strictly monotone, we have

$$\begin{aligned} \langle x - y, 0 - 0\rangle &> 0 \\ \Rightarrow 0 &> 0 \end{aligned}$$

which is a contradiction. Thus, $\operatorname{zer} A$ is at most a singleton. $\square$

Fact 1.2.28. [4, Corollary 25.5] Let A and B be maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$ such that one of the following holds:

(1) $\operatorname{dom} A \cap \operatorname{intdom} B \neq \emptyset$.

(2) $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$.

Then $A+B$ is maximally monotone. In particular, (1) and (2) hold when $\operatorname{dom} B = \mathcal{H}$.

Fact 1.2.29. [4, Propositions 20.22] Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone, let $z \in \mathcal{H}, u \in \mathcal{H}$, and $\gamma \in \mathbb{R}_{++}$. Then A^{-1}, and $C : x \mapsto u + \gamma A(x + z)$ are maximally monotone.

Fact 1.2.30. [4, Propositions 20.22, 20.23] Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ and $B : \mathcal{G} \to 2^{\mathcal{G}}$ be maximally monotone. Then $A \times B : \mathcal{H} \times \mathcal{G} \to 2^{\mathcal{H}\times\mathcal{G}}$:$(x, y) \mapsto Ax \times By$ is maximally monotone.

Fact 1.2.31. [4, Example 20.35] Let $A : \mathcal{H} \to \mathcal{H}$ be a bounded linear operator such that $A^\star = -A$. Then A is maximally monotone.

Fact 1.2.32. [4, Minty's Theorem] Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be monotone. Then A is maximally monotone if and only if $\operatorname{ran}(\operatorname{Id} + A) = \mathcal{H}$.

Example 1.2.33. The following are maximally monotone operators:

(1) Let C be a nonempty convex set in $\mathbb{R}^m$. Then N_C, N_C^{-1}, $N_C^{-1} + \mathrm{Id}$, and $(N_C^{-1} + \mathrm{Id})^{-1}$ are all maximally monotone.

(2) For any $\gamma \in \mathbb{R}_{++}$, $\gamma \,\mathrm{Id}$ is maximally monotone.

Proof. (1) C is closed and convex, so due to Example 1.3.13, $\iota_C \in \Gamma_0(\mathbb{R}^m)$. As $\partial\iota_C$ is maximally monotone by Fact 1.3.27 and $\partial\iota_C = N_C$ by Example 1.3.25, we get N_C is maximally monotone. Because $\operatorname{dom}\mathrm{Id} = \mathbb{R}^m$ and N_C is maximally monotone, by Fact 1.2.28 and Fact 1.2.29, N_C^{-1}, $N_C^{-1} + \mathrm{Id}$, and $(N_C^{-1} + \mathrm{Id})^{-1}$ are all maximally monotone.

(2) For any $x \in \mathbb{R}^m$, we have $\gamma \,\mathrm{Id}\, x = \gamma x$. Therefore, for any $x, y \in \mathbb{R}^m$,

$$\begin{aligned}\langle y - x, \gamma \,\mathrm{Id}\, y - \gamma \,\mathrm{Id}\, x\rangle =& \langle y - x, \gamma y - \gamma x\rangle \\ =& \gamma\|y - x\|^2 \\ \geq& 0.\end{aligned}$$

Moreover, we have

$$\operatorname{ran}(\mathrm{Id} + \gamma \,\mathrm{Id}) = \operatorname{ran}(1 + \gamma)\,\mathrm{Id} = \mathbb{R}^m.$$

Therefore, by Fact 1.2.32, $\gamma \,\mathrm{Id}$ is maximally monotone.

□

Lemma 1.2.34. Let $A : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ be maximally monotone. Then $-A(-\cdot)$ is also maximally monotone.

Proof. For any $y_1 \in -A(-x_1)$ and $y_2 \in -A(-x_2)$, we have

$$\begin{cases} -y_1 \in A(-x_1) \\ -y_2 \in A(-x_2). \end{cases}$$

Since A is maximally monotone,

$$\langle -y_1 - (-y_2), -x_1 - (-x_2)\rangle \geq 0,$$

which is equivalent to

$$\langle y_2 - y_1, x_2 - x_1\rangle \geq 0.$$

Therefore, $-A(-\cdot)$ is monotone. Since A is maximally monotone, we have $-A(-\cdot)$ is maximally monotone. □

1.2.4 Resolvent

Definition 1.2.35. [4] Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone. The resolvent of A is defined as

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

Fact 1.2.36. Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be monotone and let $\gamma \in \mathbb{R}_{++}$. Then $J_{\gamma A}$ is single-valued.

Proof. Suppose $J_{\gamma A}$ is not single-valued. Then there exists $x, y_1, y_2 \in \mathcal{H}$ and $y_1 \neq y_2$ such that

$$y_1 \in J_{\gamma A}(x) \text{ and } y_2 \in J_{\gamma A}(x).$$

That is,

$$\begin{cases} x \in (\operatorname{Id} + \gamma A) y_1 \\ x \in (\operatorname{Id} + \gamma A) y_2, \end{cases}$$

which implies

$$y_2 - y_1 \in \gamma(A y_1 - A y_2).$$

Therefore, there exists $(y_1, u) \in \operatorname{gra} A$ and $(y_2, v) \in \operatorname{gra} A$ such that $u - v = \frac{1}{\gamma}(y_2 - y_1)$. That implies

$$\langle y_1 - y_2, u - v \rangle = -\frac{1}{\gamma}\|y_1 - y_2\|^2,$$

which is less than 0 as $y_1 \neq y_2$. This contradicts the assumption that A is monotone. Therefore, $J_{\gamma A}$ must be single-valued. $\square$

Fact 1.2.37. [4, Proposition 23.20] Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone and let $\gamma \in \mathcal{R}_{++}$. Then

$$\operatorname{Id} = J_{\gamma A} + \gamma J_{\gamma^{-1} A^{-1}} \circ \gamma^{-1} \operatorname{Id}.$$

In particular,

$$J_{A^{-1}} = \operatorname{Id} - J_A.$$

Fact 1.2.38. [4, Proposition 23.10] Let D be a nonempty subset of $\mathcal{H}$, let $T : D \to \mathcal{H}$, and set $A = T^{-1} - \operatorname{Id}$. Then the following hold:

(1) $T = J_A$.

(2) T is firmly nonexpansive if and only if A is monotone.

(3) T is firmly nonexpansive and $D = \mathcal{H}$ if and only if A is maximally monotone.

Definition 1.2.39. An operator $T : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz continous with constant $\beta \in [0, \infty)$ if

$$(\forall x \in \mathbb{R}^n)(\forall y \in \mathbb{R}^n) \quad \|Tx - Ty\| \leq \beta\|x - y\|.$$

The operator T is locally Lipschitz continuous near a point $x_0 \in \mathbb{R}^n$ if there exists $r \in \mathbb{R}_{++}$ such that $T|_{B(x_0;r)}$, which means the restriction of T to $B(x_0; r)$, is Lipschitz continuous.

Fact 1.2.40. [4, Corollary 23.11] Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be maximally monotone and let $\gamma \in \mathbb{R}_{++}$. Then the following hold:

(1) $J_{\gamma A} : \mathcal{H} \to \mathcal{H}$ and $\mathrm{Id} - J_{\gamma A} : \mathcal{H} \to \mathcal{H}$ are firmly nonexpansive and maximally monotone.

(2) The reflected resolvent $R_{\gamma A} : \mathcal{H} \to \mathcal{H} : x \mapsto 2J_{\gamma A}x - x$ is nonexpansive.

1.3 Functions

Definition 1.3.1. Let $f : \mathbb{R}^m \to [-\infty, +\infty]$. The domain of f is

$$\operatorname{dom} f := \{x \in \mathbb{R}^m : f(x) < +\infty\},$$

the epigraph of f is

$$\operatorname{epi} f := \{(x, \xi) \in \mathbb{R}^m \times \mathbb{R} : f(x) \leq \xi\},$$

and the reversal of f is

$$f^{\vee} := \{x \in \mathbb{R}^m : f^{\vee}(x) := f(-x)\}.$$

Definition 1.3.2. A function $f : \mathbb{R}^m \to [-\infty, +\infty]$ is proper if its domain is nonempty and $-\infty \notin f(\mathbb{R}^m)$.

1.3.1 Convex functions

Definition 1.3.3. Let $f : \mathcal{H} \to [-\infty, +\infty]$. Then $f(x)$ is ***convex*** if its epigraph $\{(x, r) : f(x) \leq r\}$ is a convex subset of $\mathcal{H} \times \mathbb{R}$. Moreover, f is concave if $-f$ is convex.

Fact 1.3.4. [4, Proposition 8.4] Let $f : \mathcal{H} \to [-\infty, +\infty]$. Then $f(x)$ is ***convex*** if and only if for all $x \in \operatorname{dom} f$, for all $y \in \operatorname{dom} f$, for all $\alpha \in (0, 1)$

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

Fact 1.3.5. Let $f : \mathbb{R}^m \to [-\infty, +\infty]$ be convex. Then its domain $\operatorname{dom} f = \{x \in \mathbb{R}^m : f(x) < +\infty\}$ is convex.

Proof. For any $x, y \in \operatorname{dom} f$, we have $f(x) < +\infty, f(y) < +\infty$. Thus, for all $\alpha \in (0,1), \alpha f(x) + (1-\alpha)f(y) < +\infty$. By Fact 1.3.4, for all $x \in \operatorname{dom} f$, for all $y \in \operatorname{dom} f$, for all $\alpha \in (0,1)$

$$\begin{aligned} f(\alpha x + (1-\alpha)y) &\leq \alpha f(x) + (1-\alpha)f(y) \\ &< +\infty. \end{aligned}$$

That is, $\alpha x + (1-\alpha)y \in \operatorname{dom} f$. According to the definition of the convex set, we find $\operatorname{dom} f$ is a convex set. $\square$

Definition 1.3.6. Let $f : \mathbb{R}^m \to [-\infty, +\infty]$ be a proper function. Then $f(x)$ is ***strictly convex*** if $\forall x \in \operatorname{dom} f, \forall y \in \operatorname{dom} f, \forall \alpha \in (0,1)$, and for $x \neq y$, we have

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

Now let C be a nonempty subset of $\operatorname{dom} f$. Then f is convex on C if $\forall x \in C, \forall y \in C, \forall \alpha \in (0,1)$, and for $x \neq y$, we have

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

and f is strictly convex on C if $\forall x \in C, \forall y \in C, \forall \alpha \in (0,1)$, and for $x \neq y$, we have

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

Example 1.3.7. The function $\|\cdot\|^2$ is strictly convex.

Proof. Let $x, y \in \mathbb{R}^m$, and $x \neq y$. Let $0 < a < 1$. Then

$$\begin{aligned} &\|ax + (1-a)y\|^2 - a\|x\|^2 - (1-a)\|y\|^2 \\ =&a^2x^2 + (1-a)^2y^2 + 2a(1-a)\langle x, y\rangle - ax^2 - (1-a)y^2 \\ =&-a(1-a)(x^2 + y^2 - 2\langle x, y\rangle) \\ =&-a(1-a)(x-y)^2. \end{aligned}$$

Because $x \neq y$ and $0 < a < 1$, we have $-a(1-a)(x-y)^2 < 0$. That is,

$$\|ax + (1-a)y\|^2 < a\|x\|^2 + (1-a)\|y\|^2.$$

Therefore, the function $\|\cdot\|^2$ is strictly convex. $\square$

Fact 1.3.8. Let $f : \mathbb{R}^m \to \mathbb{R}$ be a convex function, $A : \mathbb{R}^m \to \mathbb{R}^m$ be a linear operator. Then $g = f \circ A$ is convex.

Proof. For all $x \in \operatorname{dom} g$, for all $y \in \operatorname{dom} g$, for all $\alpha \in (0,1)$. Since $g = f \circ A$, we have

$$g(\alpha x + (1-\alpha)y) = f(A[\alpha x + (1-\alpha)y]).$$

Because A is a linear operator, we have

$$f(A[\alpha x + (1-\alpha)y]) = f(\alpha Ax + (1-\alpha)Ay). \tag{1.4}$$

As f is convex, equation (1.4) implies

$$\begin{aligned} g(\alpha x + (1-\alpha y)) &\leq \alpha f(Ax) + (1-\alpha)f(Ay) \\ &= \alpha g(x) + (1-\alpha)g(y). \end{aligned}$$

Therefore, g is a convex function. $\square$

Fact 1.3.9. Let $f: \mathbb{R}_+ \to \mathbb{R}$ be convex and increasing, $g: \mathbb{R}^m \to \mathbb{R}_+$ be convex. Then $h = f \circ g$ is convex.

Proof. For all $x \in \operatorname{dom} h$, for all $y \in \operatorname{dom} h$, for all $\alpha \in (0,1)$. Since $h = f \circ g$, we have

$$h(\alpha x + (1-\alpha)y) = f(g(\alpha x + (1-\alpha)y)). \tag{1.5}$$

As f is convex and increasing, g is convex, equation (1.5) implies that

$$\begin{aligned} h(\alpha x + (1-\alpha)y) &\leq f(\alpha g(x) + (1-\alpha)y) \\ &\leq \alpha f(g(x)) + (1-\alpha)f(g(y)) \\ &= \alpha h(x) + (1-\alpha)h(y). \end{aligned}$$

Therefore, h is a convex function. $\square$

1.3.2 Lower semicontinuous functions

In the following thesis, I shall use $B(x;r)$ to denote the closed ball with center at x and radius $r \in \mathbb{R}_{++}$.

Definition 1.3.10. The lower limit of a function $f: \mathbb{R}^m \to \mathbb{R}$ at $\bar{x}$ is the value in $\mathbb{R}^m$ defined by

$$\begin{aligned} \liminf_{x \to \bar{x}} f(x) &:= \lim_{\delta \searrow 0} \inf_{x \in B(\bar{x};\delta)} f(x) \\ &= \sup_{\delta > 0} \inf_{x \in B(\bar{x};\delta)} f(x). \end{aligned}$$

Definition 1.3.11. A function f is lower semicontinous at a point x_0 if

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

The function is said to be lower semicontinuous on $\mathcal{H}$ if it is lower semicontinous at every point $x_0 \in \mathcal{H}$.

Example 1.3.12. The following functions are lower semicontinous:

(1) All continuous functions are lower semicontinous.

(2) The piecewise function

$$f(x) = \begin{cases} \sin(x) & \text{if } x \leq \frac{\pi}{2}, \\ \sin(x) + 1 & \text{if } x > \frac{\pi}{2} \end{cases}$$

is lower semicontinuous.

Example 1.3.13. [4, Example 1.25] The indicator function of a set $C \in \mathcal{H}$, i.e., the function

$$\iota_C : \mathcal{H} \to [-\infty, +\infty] : x \mapsto \begin{cases} 0 & \text{if } x \in C, \\ +\infty & \text{otherwise} \end{cases}$$

is lower semicontinuous if and only if C is closed. Moreover, if C is closed and convex, then ι_C is a proper, lower semicontinuous, and convex function.

In the following thesis, I use $\Gamma(\mathcal{H})$ to denote the set of lower semicontinuous convex functions from $\mathcal{H}$ to $[-\infty, +\infty]$, and use $\Gamma_0(\mathcal{H})$ to denote the set of proper lower semicontinuous convex functions from $\mathcal{H}$ to $(-\infty, +\infty]$.

Fact 1.3.14. [4, Corollary 9.4] Let $(f_i)_{i \in I}$ be a family in $\Gamma(\mathcal{H})$. If I is finite and $-\infty \notin \bigcup_{i \in I} f_i(\mathcal{H})$. Then $\sum_{i \in I} f_i \in \Gamma(\mathcal{H})$.

Lemma 1.4. *Let* $f, g \in \Gamma_0(\mathcal{H})$, *and* $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$. *Then* $f + g \in \Gamma_0(\mathcal{H})$.

Proof. Since $f, g \in \Gamma_0(\mathcal{H})$, that is, $-\infty \notin f(\mathcal{H}) \cup g(\mathcal{H})$, we have $f + g > -\infty$. As $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$, there must exists at least an x such that $f(x) + g(x) < +\infty$. Therefore, combining with Fact 1.3.14, $f + g \in \Gamma_0(\mathcal{H})$. $\square$

Fact 1.3.15. [4, Theorem 9.20] Let $f \in \Gamma_0(\mathcal{H})$. Then for any $x \in \mathcal{H}$, there exists a $u \in \mathcal{H}$ and an $\eta \in \mathbb{R}$ such that $f(x) \geq \langle x, u \rangle + \eta$.

1.3.3 Coercive and supercoercive functions

Definition 1.3.16. Let $f : \mathcal{H} \to [-\infty, +\infty]$. Then f is coercive if

$$\lim_{\|x\|\to+\infty} f(x) = +\infty,$$

and supercoercive if

$$\lim_{\|x\|\to+\infty} \frac{f(x)}{\|x\|} = +\infty.$$

By convention, we say f is coercive and supercoercive if $\mathcal{H} = \{0\}$.

Example 1.3.17. The function $\|\cdot\|^2$ is supercoercive.

Fact 1.3.18. Let f be in $\Gamma_0(\mathcal{H})$, and let $g : \mathcal{H} \to (-\infty, +\infty]$ be supercoercive. Then $f + g$ is supercoercive.

Proof. According to Fact 1.3.15, there exists a $u \in \mathcal{H}$ and an $\eta \in \mathbb{R}$ such that for all $x \in \mathcal{H}$,

$$f(x) \geq \langle x, u\rangle + \eta.$$

Then we have

$$\begin{aligned}
\lim_{\|x\|\to+\infty} \frac{f(x)+g(x)}{\|x\|} &\geq \lim_{\|x\|\to+\infty} \frac{\langle x, u\rangle + \eta + g(x)}{\|x\|} \\
&\geq \lim_{\|x\|\to+\infty} \frac{-\|u\|\|x\| + \eta + g(x)}{\|x\|} \\
&= \lim_{\|x\|\to+\infty} (-\|u\| + \frac{\eta + g(x)}{\|x\|}) \\
&\geq -\|u\| - |\eta| + \lim_{\|x\|\to+\infty} \frac{g(x)}{\|x\|} \\
&\to +\infty.
\end{aligned}$$

Thus, $f + g$ is supercoercive. □

Fact 1.3.19. [4, Corollary 11.16] Let f and g be in $\Gamma_0(\mathcal{H})$. Suppose that $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$ and f is supercoercive. Then $f + g$ is coercive and it has a minimizer over $\mathcal{H}$. If f or g is strictly convex, then $f + g$ has exactly one minimizer over $\mathcal{H}$.

1.3.4 Subgradient and subdifferential

Definition 1.3.20. Let $f : \mathbb{R}^m \to [-\infty, +\infty]$, and let x be a point such that $|f(x)| < +\infty$. We say that f is differentiable (or Fréchet differentiable) at x if and only if there exists a vector x^* with the property

$$\lim_{y \to x} \frac{f(y) - f(x) - \langle x^*, y - x\rangle}{\|y - x\|} = 0.$$

If such x^* exists, it is called the gradient of f at x and is denoted by $\nabla f(x)$.

Definition 1.3.21. A vector $u \in \mathbb{R}^m$ is said to be a subgradient of a convex function $f : \mathbb{R}^m \to \mathbb{R}$ at the point x if we have

$$\forall y \in \mathbb{R}^m \quad \langle y - x, u\rangle + f(x) \leq f(y).$$

The set of all subgradients of f at x is called the subdifferential of f at x and is denoted by $\partial f(x)$.

Definition 1.3.22. (Fenchel Subdifferential) For a (not necessarily convex) $f : \mathbb{R}^m \to \mathbb{R}$, define its Fenchel subdifferential at x

$$\partial f(x) := \{v \in \mathbb{R}^m : f(y) \geq f(x) + \langle v, y - x\rangle \text{ for all } y \in \mathbb{R}^m\}.$$

When f is convex, $\partial f(x)$ is the usual subdifferential.

Fact 1.3.23. [15, Proposition 2.36] Let $f : \mathbb{R}^m \to (-\infty, +\infty]$ be proper and convex, and let $x \in \operatorname{dom} f$. Suppose that f is differentiable at x. Then

$$\partial f(x) = \{\nabla f(x)\}.$$

Example 1.3.24. Let $f(x) = \frac{1}{2}\|x\|^2$. Then $\partial f(x) = \{\nabla f(x)\} = \{x\}$.

Proof. We already proved $\|\cdot\|^2$ is strictly convex in Example 1.3.7, therefore in the case $x, y \in \mathbb{R}^m$, $x \neq y$ and $0 < a < 1$,

$$\begin{aligned}\frac{1}{2}\|ax + (1-a)y\|^2 <& \frac{1}{2}(a\|x\|^2 + (1-a)\|y\|^2)\\ =& \frac{a}{2}\|x\|^2 + \frac{1-a}{2}\|y\|^2.\end{aligned}$$

Therefore, $f(x) = \frac{1}{2}\|x\|^2$ is strictly convex. Since $f(x)$ is proper, convex, and differentiable on $\mathbb{R}^m$, applying Fact 1.3.23 here, we have

$$\partial f(x) = \{\nabla f(x)\} = \{x\}.$$

$\square$

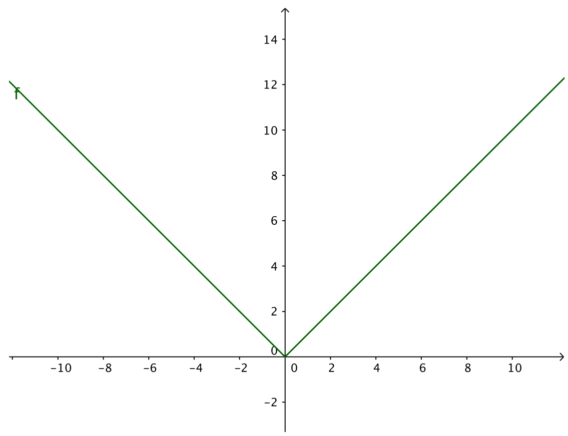

Figure 1.3: Figure of $f(x) = |x|$

Example 1.3.25. [4, Example 16.13] Let C be a nonempty convex subset of $\mathcal{H}$. Then $\partial\iota_C = N_C$.

Example 1.3.26. Let $f : \mathbb{R} \to \mathbb{R}$ defined as $f(x) = |x|$.

Through the graph of this fuction, we can easily see its global minimizer is at $\bar{x} = 0$. However, this function is not differentiable at $x = 0$. At points other than 0, f is differentiable. According to the definition of the subgradient of a convex function, we can get the the subdifferential of $f(x) = |x|$ is

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

Proof. Because f is differentiable when $x < 0$ (or $x > 0$). According to Fact 1.3.23,

$$\begin{aligned} \partial f(x) = \{\nabla f(x)\} = \{1\} \quad & \text{for } x > 0; \\ \partial f(x) = \{\nabla f(x)\} = \{-1\} \quad & \text{for } x < 0. \end{aligned}$$

For $\bar{x} = 0$, let $v \in \partial f(\bar{x})$. Then we have

$$\begin{aligned} & \langle v, x - \bar{x}\rangle \leq f(x) - f(\bar{x}) \; \forall x \in \mathbb{R} \\ \Leftrightarrow & \langle v, x\rangle \leq f(x) \text{ for } \; \bar{x} = 0, \forall x \in \mathbb{R} \\ \Leftrightarrow & \begin{cases} v \geq \frac{f(x)}{x} = \frac{-x}{x} = -1 & \text{if } \; x < 0 \\ v \leq \frac{f(x)}{x} = \frac{x}{x} = 1 & \text{if } \; x > 0 \end{cases} \\ \Leftrightarrow & \, v \in [-1, 1]. \end{aligned}$$

Thus, we have $\partial f(0) = [-1, 1]$. □

Fact 1.3.27. [4, Theorem 20.25] Let $f \in \Gamma_0(\mathcal{H})$. Then ∂f is maximally monotone.

Fact 1.3.28. Let $f \in \Gamma_0(\mathbb{R}^m)$ Then $\operatorname{ran}(\operatorname{Id} + \partial f) = \mathbb{R}^m$.

Proof. Since $f \in \Gamma_0(\mathbb{R}^m)$, by Fact 1.3.27, ∂f is maximally monotone. According to Fact 1.2.32, we have $\operatorname{ran}(\operatorname{Id} + \partial f) = \mathbb{R}^m$. □

Fact 1.3.29. [4, Theorem 16.47] Let $\mathcal{K}$ be a real Hilbert space, let $f \in \Gamma_0(\mathcal{H})$, let $g \in \Gamma_0(\mathcal{K})$, and let $L : \mathcal{H} \to \mathcal{K}$ be a nonzero bounded linear operator. Suppose $L(\operatorname{dom} f) \cap \operatorname{intdom} g \neq \emptyset$. Then $\partial(f + g \circ L) = \partial f + L^\star \circ (\partial g) \circ L$.

Fact 1.3.30. [10, Page 20] Let $f : \mathbb{R}^m \to \mathbb{R}$ be convex, and let $L : \mathcal{H} \to \mathcal{K}$ be a nonzero bounded linear operator. If $h(x) = f(Lx + b)$, where $b \in \mathbb{R}$, then $\partial h(x) = L^\star \circ \partial f \circ (Lx + b)$.

Fact 1.3.31. [4, Corollary 16.50] Let m be an integer such that $m \geq 2$, set $I = \{1, \ldots, m\}$, and let $(f_i)_{i \in I}$ be functions in $\Gamma_0(\mathcal{H})$ such that

$$\operatorname{dom} f_m \cap_{i=1}^{m-1} \operatorname{intdom} f_i \neq \emptyset.$$

Then $\partial(\sum_{i=1}^m f_i) = \sum_{i=1}^m \partial f_i$.

Fact 1.3.32. Let $f \in \Gamma_0(\mathcal{H})$ and let $\gamma \in \mathbb{R}_{++}$. Then $\partial(f + (\gamma/2)\|\cdot\|^2) = \partial f + \gamma \operatorname{Id}$.

Proof. Since $\operatorname{dom} \|\cdot\|^2 = \mathcal{H}$, and $f \in \Gamma_0(\mathcal{H})$, we have

$$\operatorname{dom} f \cap \operatorname{intdom}[(\gamma/2)\|\cdot\|^2] \neq \emptyset.$$

Therefore, by using Fact 1.3.31,

$$\partial(f + (\gamma/2)\|\cdot\|^2) = \partial f + \partial[(\gamma/2)\|\cdot\|^2] = \partial f + \gamma \operatorname{Id}.$$

□

Definition 1.3.33. The set of global minimizers of a function f is denoted as $\operatorname{Argmin} f$.

Fact 1.3.34. Let $f : \mathbb{R}^m \to (-\infty, +\infty]$ be proper. Then

$$\operatorname{Argmin} f = \operatorname{zer} \partial f = \{x \in \mathbb{R}^m : 0 \in \partial f(x)\}.$$

Proof. Let $x \in \operatorname{Argmin} f$. Then for all $y \in \mathbb{R}^m$, $f(x) \leq f(y)$. That implies $f(x) + \langle y - x, 0 \rangle \leq f(y)$, which is equivalent to $0 \in \partial f(x)$. Thus,

$$\operatorname{Argmin} f = \operatorname{zer} \partial f.$$

□

1.3.5 Conjugation

Definition 1.3.35. Let $f : \mathcal{H} \to [-\infty, +\infty]$. The ***conjugate*** of f is

$$f^* : \mathcal{H} \to [-\infty, +\infty] : u \mapsto \sup_{x\in\mathcal{H}}(\langle x, u\rangle - f(x)),$$

and the ***biconjugate*** of f is $f^{**} = (f^*)^*$.

Fact 1.3.36. [4, Proposition 13.19] Let $f : \mathcal{H} \to [-\infty, +\infty]$. Then

$$f = \frac{1}{2}\|\cdot\|^2 \Leftrightarrow f^* = f.$$

Fact 1.3.37. [4, Proposition 13.23] Let $f : \mathcal{H} \to (-\infty, +\infty]$. Then for any $\alpha \in \mathbb{R}_{++}$, $(\alpha f)^* = \alpha f^*(\cdot/\alpha)$.

Example 1.3.38. Let $x \in \mathbb{R}^m$, let $\lambda \in \mathbb{R}_{++}$. If $f(x) = \lambda\|x\|^2$, then $f^*(u) = \frac{\|u\|^2}{4\lambda}$.

Proof. By applying Fact 1.3.37 together with Fact 1.3.36, we have

$$\begin{aligned}
(\lambda\|\cdot\|^2)^* &= \left(2\lambda\frac{1}{2}\|\cdot\|^2\right)^* \\
&= 2\lambda(\frac{1}{2}\|\cdot\|^2)^*\left(\frac{\cdot}{2\lambda}\right) \\
&= \lambda\left\|\frac{\cdot}{2\lambda}\right\|^2 \\
&= \frac{\|\cdot\|^2}{4\lambda}.
\end{aligned}$$

□

Fact 1.3.39. [4, Proposition 13.13] Let $f : \mathcal{H} \to [-\infty, +\infty]$. Then $f^* \in \Gamma(\mathcal{H})$.

Fact 1.3.40. [4, Theorem 13.37] Let $f : \mathcal{H} \to (-\infty, +\infty]$ be proper. Then $f \in \Gamma_0(\mathcal{H})$ if and only if $f = f^{**}$. In this case, f^* is proper as well.

Theorem 1.3.41. Let $f, g \in \Gamma_0(\mathcal{H})$ and $\operatorname{dom} f^* \cap \operatorname{dom} g^* \neq \emptyset$. Then $f^* + g^* \in \Gamma_0(\mathcal{H})$.

Proof. Combine Fact 1.3.39 and Fact 1.3.40, $f, g \in \Gamma_0(\mathcal{H})$ implies $f^*, g^* \in \Gamma_0(\mathcal{H})$. Moreover, as $\operatorname{dom} f^* \cap \operatorname{dom} g^* \neq \emptyset$, according to Lemma 1.4, $f^* + g^* \in \Gamma_0(\mathcal{H})$. □

Fact 1.3.42. [4, Proposition 16.10] Let f be a proper function on $\mathcal{H}$, let $x \in \mathcal{H}$, and $u \in \mathcal{H}$. Then $u \in \partial f(x)$ if and only if $f(x) + f^*(u) = \langle x, u\rangle \Rightarrow x \in \partial f^*(u)$.

Fact 1.3.43. [4, Proposition 16.49] Let $f \in \Gamma_0(\mathcal{H})$. Then $\operatorname{intdom} f \subset \operatorname{dom} \partial f \subset \operatorname{dom} f$.

Example 1.3.44. Let $f : \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = \begin{cases} x \ln x & \text{if } x \geq 0, \\ +\infty & \text{if } x < 0. \end{cases}$$

We have $\operatorname{dom} \partial f = (0, +\infty)$ because $\partial f(0) = \emptyset$. As $\operatorname{dom} f = [0, +\infty)$, in this example $\operatorname{dom} \partial f \subsetneq \operatorname{dom} f$.

Example 1.3.45. Let $f(x) = \iota_{[2,3]}(x)$. Then $2 \in \operatorname{dom} \partial f$ but $2 \notin \operatorname{intdom} f$. Therefore, $\operatorname{intdom} f \subsetneq \operatorname{dom} \partial f$.

Fact 1.3.46. [4, Corollary 16.30] Let $f \in \Gamma_0(\mathcal{H})$. Then $(\partial f)^{-1} = \partial f^*$.

1.3.6 Infimal convolution

Definition 1.3.47. Let f and g be functions from $\mathbb{R}^m$ to $[-\infty, +\infty]$. The infimal convolution of f and g is

$$f \Box g : \mathbb{R}^m \to [-\infty, +\infty] : x \mapsto \inf_{y \in \mathbb{R}^m} (f(y) + g(x - y)),$$

and it is exact at a point $x \in \mathbb{R}^m$ if $(f \Box g)(x) = \min_{y \in \mathbb{R}^m} \{f(y) + g(x - y)\}$, i.e.,

$$\exists y \in \mathcal{H} : (f \Box g)(x) = f(y) + g(x - y) \in (-\infty, +\infty];$$

$f \Box g$ is exact if it is exact at every point of its domain, in which case it is denoted by $f \boxdot g$.

Fact 1.3.48. [4, Proposition 12.6] Let f and g be functions from $\mathcal{H} \to (-\infty, +\infty]$. Then the following hold:

(1) $\operatorname{dom}(f \Box g) = \operatorname{dom} f + \operatorname{dom} g$.

(2) $f \Box g = g \Box f$.

Fact 1.3.49. [4, Proposition 13.24] Let f and g be functions in from $\mathcal{H}$ to $(-\infty, +\infty]$. Then $(f \Box g)^* = f^* + g^*$.

Fact 1.3.50. [4, Proposition 15.2] Let f and g be functions in $\Gamma_0(\mathcal{H})$ such that $0 \in \operatorname{int}(\operatorname{dom} f - \operatorname{dom} g)$. Then $(f + g)^* = f^* \boxdot g^*$.

Fact 1.3.51. [4, Proposition 15.7] Let f and g be in $\Gamma_0(\mathcal{H})$. Suppose

$$0 \in \operatorname{int}(\operatorname{dom} f^* - \operatorname{dom} g^*).$$

Then $f \Box g = f \boxdot g \in \Gamma_0(\mathcal{H})$.

Fact 1.3.52. [4, Proposition 16.61] Let f and g be in $\Gamma_0(\mathcal{H})$, let $x \in \operatorname{dom}(f\Box g)$, and let $y \in \mathcal{H}$. Then the following hold:

(1) Suppose that $(f\Box g)(x) = f(y) + g(x-y)$. Then

$$\partial(f\Box g)(x) = \partial f(y) \cap \partial g(x-y).$$

(2) Suppose that $\partial f(y) \cap \partial g(x-y) \neq \emptyset$. Then $(f\Box g)(x) = f(y) + g(x-y)$.

Definition 1.3.53. Let $f : \mathcal{H} \to (-\infty, +\infty]$ be proper and let $\lambda \in \mathbb{R}_{++}$. The Moreau envelope of f with parameter λ is

$$e_\lambda f := f\Box(\frac{1}{2\lambda}\|\cdot\|^2).$$

Example 1.3.54. [4, Example 12.21] Let $C \subset \mathcal{H}$ and let $\lambda \in \mathbb{R}_{++}$. Then $e_\lambda \iota_C = (2\lambda)^{-1} d_C^2$.

Proof. We have

$$\begin{aligned} e_\lambda \iota_C(x) =& [\iota_C \Box(\frac{1}{2\lambda}\|\cdot\|^2)](x) \\ =& \inf_{y\in\mathbb{R}^m}(\iota_C(y) + \frac{1}{2\lambda}\|x-y\|^2) \\ =& \inf_{y\in C}(\frac{1}{2\lambda}\|x-y\|^2) \\ =& \frac{1}{2\lambda} d_C^2(x). \end{aligned}$$

□

Fact 1.3.55. Let $f : \mathbb{R}^m \to (-\infty, +\infty]$ be proper and let $\lambda \in \mathbb{R}_{++}$. Then $e_\lambda f$ is full domain, i.e., $\operatorname{dom} e_\lambda f = \mathbb{R}^m$.

Proof. Combining the definition of $e_\lambda f$ with the Fact 1.3.48, we have

$$\begin{aligned} \operatorname{dom} e_\lambda f &= \operatorname{dom} f\Box(\frac{1}{2\lambda}\|\cdot\|^2) \\ &= \operatorname{dom} f + \mathbb{R}^m \\ &= \mathbb{R}^m. \end{aligned}$$

Thus, $e_\lambda f$ is full domain. □

Definition 1.3.56. Let $f \in \Gamma_0(\mathcal{H})$ and let $x \in \mathcal{H}$. Then $\operatorname{Prox}_f x$ is the unique point in $\mathcal{H}$ that satisfies

$$e_1 f(x) = \min_{y \in \mathcal{H}} \left(f(y) + \frac{1}{2}\|x - y\|^2 \right) = f(\operatorname{Prox}_f x) + \frac{1}{2}\|x - \operatorname{Prox}_f x\|^2.$$

Fact 1.3.57. [4, Proposition 16.44] Let $f \in \Gamma_0(\mathcal{H})$, and let x and p be in $\mathcal{H}$. Then

$$p = \operatorname{Prox}_f x \Leftrightarrow x - p \in \partial f(p).$$

In other words,

$$\operatorname{Prox}_f = (\operatorname{Id} + \partial f)^{-1} = J_{\partial f}.$$

1.3.7 Fenchel-Rockafellar duality

Definition 1.3.58. [4, Definition 15.19] Let $f : \mathcal{H} \to (-\infty, +\infty]$, let $g : \mathcal{K} \to (-\infty, +\infty]$, and let $L : \mathcal{H} \to \mathcal{K}$ be a nonzero bounded linear operator. The primal problem associated with the composition function $f + g \circ L$ is

$$\min_{x \in \mathcal{H}} \{f(x) + g(Lx)\}, \tag{1.6}$$

its dual problem is

$$\min_{v \in \mathcal{K}} \{f^*(L^\star v) + g^*(-v)\}, \tag{1.7}$$

the primal optimal value is $\mu = \inf(f + g \circ L)(\mathcal{H})$, the dual optimal value is $\mu^* = \inf(f^* \circ L^\star + g^{*\vee})(\mathcal{K})$, and the duality gap is

$$\Delta(f, g, L) = \begin{cases} 0, \text{ if } \mu = -\mu^* \in \{-\infty, +\infty\} \\ \mu + \mu^* \text{otherwise.} \end{cases}$$

Fact 1.3.59. [4, Proposition 15.21] Let $f : \mathcal{H} \to (-\infty, +\infty]$ and $g : \mathcal{K} \to (-\infty, +\infty]$ be proper, and let $L : \mathcal{H} \to \mathcal{K}$ be a nonzero bounded linear operator. Set $\mu = \inf(f + g \circ L)(\mathcal{H})$ and $\mu^* = \inf(f^{*\vee} \circ L^\star + g^*)(\mathcal{K})$. Then

$$\mu = -\mu^* \Leftrightarrow \Delta(f, g, L) = 0.$$

Remark 1.5. There exists a solution to problem (1.6) implies there must exists a solution to problem (1.7), and vice versa. Therefore, solving problem (1.6) is equivalent to solving problem (1.7).

Chapter 2

Classic Douglas-Rachford algorithm

2.1 Overview

In this chapter, the history of the Douglas-Rachford algorithm is reviewed. There is a relation between the composited monotone inclusion problem and the Douglas-Rachford algorithm, and there is also a relation between the composited monotone inclusion problem and the optimization problems. Those two relations are roughly given by Bot and Hendrich [6] in 2013. Here, I will show those relations in details.

2.2 Douglas-Rachford splitting problem and the brief history of Douglas-Rachford algorithm

The Douglas-Rachford splitting problem is the problem of finding a point $x \in \mathcal{H}$ such that

$$0 \in Ax + Bx,$$

where A and B are maximally monotone operators. Naturally, this approach is numerically viable only in those cases in which it is easy to compute $J_{\gamma(A+B)}$, where $\gamma \in \mathbb{R}_{++}$. However, the Douglas-Rachford algorithm, in which the operators A and B are employed in separate steps, can be seen as a widely applicable alternative.

The Douglas-Rachford algorithm was first be proposed by J. Douglas and H. H. Rachford [9] in 1956 as a method for solving certain matrix equations. In 1969 Lieutaud (see [13]) extended their method to deal with (possibly nonlinear) maximally monotone operators that are defined everywhere. Lions and Mercier, in their paper [14] from 1979, presented a broad and powerful generalization to its current form, i.e., to handle the sum of any two maximally monotone operators that are possibly nonlinear, possibly set-valued and not necessarily defined everywhere. With the joint work of Eckstein and Bertsekas (see [11, Theorem 5]) from

1992, the inclusion $J_B(\operatorname{Fix} T) \subseteq \operatorname{zer}(A+B)$ has been proved. Later on, in 2004, Combettes (see [7, Lemma 2.6(iii)]) refined the results by Eckstein and Bertsekas. Together with the earlier results by Lions and Mercier [14], the work by Eckstein and Bertsekas and later by Combettes complete the following Douglas-Rachford algorithm in the finite dimensional setting.

Lemma 2.2.1. [4, Douglas-Rachford algorithm] Let A and B be maximally monotone operators from $\mathbb{R}^m$ to $2^{\mathbb{R}^m}$ such that $\operatorname{zer}(A+B) \neq \emptyset$. Let $(\lambda_n)_{n=1}^{+\infty}$ be a sequence in $[0,2]$ such that $\sum_{n=1}^{+\infty} \lambda_n(2-\lambda_n) = +\infty$, let $\gamma \in \mathbb{R}_{++}$, and let $x_0 \in \mathcal{H}$. Set

$$\text{(DR)} \qquad (\forall n \in \mathbb{N}) \quad \begin{cases} y_n = J_{\gamma B} x_n \\ z_n = J_{\gamma A}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

Then there exists $x \in \operatorname{Fix} R_{\gamma A} \circ R_{\gamma B}$ such that the following hold:

(1) $J_{\gamma B} x \in \operatorname{zer}(A+B)$.

(2) $(y_n - z_n)_{n=1}^{+\infty}$ converges to 0.

(3) $(x_n)_{n=1}^{+\infty}$ converges to x.

(4) $(y_n)_{n=1}^{+\infty}$ converges to $J_{\gamma B} x$.

(5) $(z_n)_{n=1}^{+\infty}$ converges to $J_{\gamma B} x$.

(6) Suppose that one of the following holds:

   (a) A is uniformly monotone on every nonempty bounded subset of $\operatorname{dom} A$.

   (b) B is uniformly monotone on every nonempty bounded subset of $\operatorname{dom} B$.

   Then $(y_n)_{n=1}^{+\infty}$ and $(z_n)_{n=1}^{+\infty}$ converge to the unique point in $\operatorname{zer}(A+B)$.

Later on, in 2009, Combettes [8] proved that Douglas-Rachford algorithm is error-tolerant. Relying on the work of Combettes, in 2013, the joint work of Bot and Hendrich [6] showed that there are two different primal-dual iterative error-tolerant methods for solving inclusions with mixtures of composite and parallel-sum type monotone operators.

Remark 2.1. In 2011, Svaiter (see [16]) demonstrated that $A+B$ does not have to be maximally monotone and $(J_B(T^n x))_{n=1}^{+\infty}$ converges weakly to a point in $\operatorname{zer}(A+B)$ in the general Hilbert space. In 2017, Bauschke and Moursi [5] gave a simpler proof of the weakly convergence of the sequence $(J_B(T^n x))_{n=1}^{+\infty}$. But this is beyond the scope of this thesis.

2.3 The composited monotone inclusion problem

The composited monotone inclusion problems in this thesis are all considered in the $\mathbb{R}^m$ space. All of them contain two parts: the primal inclusion problem; and the dual inclusion problem. Here is the general case: Let $A : \mathbb{R}^m \to 2^{\mathbb{R}^m}$, $B : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ and $D : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ be maximally monotone operators. Let $r \in \mathbb{R}^m$, and let $L : \mathbb{R}^m \to \mathbb{R}^m$ be a nonzero linear invertible operator.

Let $z \in \mathbb{R}^m$, the primal inclusion problem is to find a point $\bar{x} \in \mathbb{R}^m$ such that

$$\text{(P)} \qquad z \in A\bar{x} + L^\star(B\Box D)(L\bar{x} - r).$$

The dual inclusion problem is to find a point $\bar{v} \in \mathbb{R}^m$ such that

$$\text{(D)} \qquad (\exists x \in \mathbb{R}^m) \quad \begin{cases} z - L^\star\bar{v} \in Ax \\ \bar{v} \in (B\Box D)(Lx - r) \end{cases}$$

Lemma 2.3.1. The primal inclusion problem (P) is equivalent to the dual inclusion problem (D).

Proof. Suppose $\bar{x} \in \mathbb{R}^m$ is the solution of the primal inclusion problem. That is,

$$z \in A\bar{x} + L^\star(B\Box D)(L\bar{x} - r),$$

which is equivalent to

$$0 \in -z + A\bar{x} + L^\star(B\Box D)(L\bar{x} - r). \tag{2.1}$$

Equation (2.1) implies that some $\bar{v} \in (B\Box D)(L\bar{x} - r)$ obeys $0 \in -z + A\bar{x} + L^\star\bar{v}$. In other words, that means $\bar{v} \in \mathbb{R}^m$ obeys

$$\begin{cases} 0 \in -z + A\bar{x} + L^\star\bar{v} \\ \bar{v} \in (B\Box D)(L\bar{x} - r). \end{cases} \tag{2.2}$$

This $\bar{v}$ solves (D), because $x = \bar{x}$ satisfies the required conditions. Since L is a nonzero linear operator, $0 \in -z + Ax + L^\star\bar{v}$ means $z - L^\star\bar{v} \in Ax$. Therefore, problem (2.2) is the dual problem. Thus, finding the solution of the primal inclusion problem is equivalent to finding the solution of the dual inclusion problem. In another words, if we can find a solution to the primal inclusion problem, there must exists a solution to the dual inclusion problem. Conversely, if we can find a solution to the dual inclusion problem, there must exists a solution to the primal inclusion problem. $\square$

Remark 2.2. We say $(\bar{x}, \bar{v})$ is a primal-dual solution to problem (P) together with (D), if

$$z - L^\star \bar{v} \in A\bar{x} \text{ and } \bar{v} \in (B \Box D)(L\bar{x} - r).$$

Here, $\bar{x}$ is a solution to (P) and $\bar{v}$ is a solution to (D), see Bot and Hendrich [6].

Before we get a corollary of Lemma 2.3.1, we need the following lemma.

Lemma 2.3.2. Let $B : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ and $D = N_{\{0\}}$. Then

$$B \Box D = B.$$

Proof. Since $D = N_{\{0\}}$, according to Lemma 1.1.35, ${N_{\{0\}}}^{-1} y = 0$, for any $y \in \mathbb{R}^m$. Suppose $B \Box D \neq \emptyset$, then there exists a pair of $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^m$ such that

$$y \in (B \Box D)(x).$$

That is,

$$\begin{aligned} y \in (B^{-1} + {N_{\{0\}}}^{-1})^{-1}(x) &\Leftrightarrow x \in (B^{-1} + {N_{\{0\}}}^{-1})(y) \\ &\Leftrightarrow x \in B^{-1} y \\ &\Leftrightarrow y \in Bx. \end{aligned}$$

Therefore,

$$B \Box N_{\{0\}} = B.$$

$\square$

Corollary 2.3.3. Let $A : \mathbb{R}^m \to 2^{\mathbb{R}^m}$, $B : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ be maximally monotone operators, let $D = N_{\{0\}}$. Let $r = 0, z = 0$, and let $L = \mathrm{Id}$. Then the following problems are equivalent:

(1) the primal inclusion problem:

$$\text{find a point } \bar{x} \in \mathbb{R}^m \text{ such that } 0 \in A\bar{x} + B\bar{x}, \tag{2.3}$$

(2) the dual inclusion problem:

$$\text{find a point } \bar{v} \in \mathbb{R}^m \text{ such that } (\exists x \in \mathbb{R}^m) \quad \begin{cases} -\bar{v} \in Ax \\ \bar{v} \in Bx. \end{cases} \tag{2.4}$$

Therefore, in this case, the dual inclusion problem (D) becomes: find v' such that

$$0 \in A^{-1}(v') - B^{-1}(-v'). \tag{2.5}$$

Proof. (1) Plugging $D = N_{\{0\}}$, $r = 0, z = 0$, and $L = \mathrm{Id}$ into the primal inclusion problem (P), we get

$$\text{find a point } \bar{x} \in \mathbb{R}^m \text{ such that } 0 \in A\bar{x} + (B\Box N_{\{0\}})\bar{x}.$$

By Lemma 2.3.2, we get $0 \in A\bar{x} + (B\Box N_{\{0\}})\bar{x}$ is equivalent to

$$0 \in A\bar{x} + B\bar{x}.$$

(2) Again, we plug $D = N_{\{0\}}$, $r = 0, z = 0$, and $L = \mathrm{Id}$ into the dual inclusion problem (D), we get (2.4). Since there exists $x \in \mathbb{R}^m$ such that $-\bar{v} \in Ax, \bar{v} \in Bx$, x should be a solution of (2.3). Therefore, (2.4) is equivalent to (2.3). Since $-\bar{v} \in Ax \Rightarrow x \in A^{-1}(-\bar{v})$ and $\bar{v} \in Bx \Rightarrow x \in B^{-1}(\bar{v})$, the inclusion problem (2.4) is equivalent to find $\bar{v}$ such that

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

Now we let $v' = -\bar{v}$, then the inclusion problem (2.4) becomes: find v' such that

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

$\square$

Remark 2.3. (2.5) is called the Attouch-Théra duality [1] of (2.3).

Lemma 2.3.4. [6, Theorem 2.1] Let $A : \mathbb{R}^m \to 2^{\mathbb{R}^m}$, $B : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ and $D : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ be maximally monotone operators. Let z and $r \in \mathbb{R}^m$, let $L : \mathbb{R}^m \to \mathbb{R}^m$ be a nonzero linear operator. Let $\mathcal{K} = \mathbb{R}^m \times \mathbb{R}^m$. If we define three set-valued operators M, Q and S as follows:

$$M : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (-z + Ax, r + B^{-1}v); \tag{M}$$

$$Q : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (0, D^{-1}v); \tag{Q}$$

$$S : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (L^\star v, -Lx). \tag{S}$$

Moreover, define an bounded linear operator

$$V : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (\frac{x}{\tau} - \frac{1}{2}L^\star v, \frac{v}{\sigma} - \frac{1}{2}Lx), \tag{V}$$

where $\tau, \sigma \in \mathbb{R}_{++}$, and $\tau\sigma\|L\|^2 < 4$.
Finally, define two operators on $\mathcal{K}V$:

$$\boldsymbol{A} := V^{-1}(\frac{1}{2}S + Q). \tag{A}$$

$$\boldsymbol{B} := V^{-1}(\frac{1}{2}S + M). \tag{B}$$

Here, the space $\mathcal{K}V$ is a vector space with inner product $\langle x, y\rangle_{\mathcal{K}V} = \langle x, Vy\rangle_{\mathcal{K}}$ and norm $\|x\|_{\mathcal{K}V} = \sqrt{\langle x, Vx\rangle_{\mathcal{K}}}$. Then any

$$(\bar{x}, \bar{v}) \in \mathrm{zer}(\boldsymbol{A} + \boldsymbol{B}).$$

is a pair of primal-dual solution to problem (P) and (D) and vice versa, while $\bar{x}$ is the solution of the primal inclusion problem (P) and $\bar{v}$ is the solution of the dual inclusion problem (D)

For the completeness of the thesis, we show the proof of this lemma here.

Proof. We split this proof into three steps.

Step 1: Prove the set-valued operators M, Q, and S are maximally monotone.

Since L is a nonzero linear operator, $A, B,$ and D are maximally monotone operators, and operators M and Q are maximally monotone on $\mathcal{K}$ by Fact 1.2.29 and Fact 1.2.30.
For operator S, let $a = (x, v), b = (y, u), a \in \mathcal{K}, b \in \mathcal{K}$. Then

$$\begin{aligned}
\langle Sa, b\rangle &= \langle (L^\star v, -Lx), (y, u)\rangle \\
&= \langle L^\star v, y\rangle + \langle -Lx, u\rangle \\
&= \langle v, Ly\rangle + \langle x, -L^\star u\rangle \\
&= \langle (x, v), (-L^\star u, Ly)\rangle \\
&= \langle a, -Sb\rangle
\end{aligned}$$

i.e., $S^\star = -S$. From Fact 1.2.31, it follows that S is maximally monotone.

Step 2: Show V is maximally monotone, and prove V^{-1} exists.

Let $a = (x, v), b = (y, u), a \in \mathcal{K}, b \in \mathcal{K}$. Then

$$\begin{aligned}
\langle Va, b\rangle &= \left\langle \left(\frac{x}{\tau} - \frac{1}{2}L^\star v, \frac{v}{\sigma} - \frac{1}{2}Lx\right), (y, u)\right\rangle \\
&= \langle \frac{x}{\tau}, y\rangle + \langle -\frac{1}{2}L^\star v, y\rangle + \langle \frac{v}{\sigma} - \frac{1}{2}Lx, u\rangle \\
&= \langle x, \frac{y}{\tau}\rangle + \frac{1}{2}\langle v, -Ly\rangle + \langle v, \frac{u}{\sigma}\rangle - \frac{1}{2}\langle x, L^\star u\rangle \\
&= \langle x, \frac{y}{\tau} - \frac{1}{2}L^\star u\rangle + \langle v, -\frac{1}{2}Ly + \frac{u}{\sigma}\rangle
\end{aligned}$$

$$= \left\langle (x, v), (\frac{y}{\tau} - \frac{1}{2}L^\star u, -\frac{1}{2}Ly + \frac{u}{\sigma}) \right\rangle$$
$$= \langle a, Vb \rangle.$$

That means V is self-adjoint, i.e., $V^\star = V$.
Let $a = (x, v), a \in \mathcal{K}$,

$$\begin{aligned}\langle Va, a \rangle &= \langle (\frac{x}{\tau} - \frac{1}{2}L^\star v, \frac{v}{\sigma} - \frac{1}{2}Lx), (x, v) \rangle \\ &= \frac{\|x\|^2}{\tau} - \langle v, Lx \rangle + \frac{\|v\|^2}{\sigma} \\ &\geq \frac{\|x\|^2}{\tau} - \|v\|\|L\|\|x\| + \frac{\|v\|^2}{\sigma}.\end{aligned}$$

For any $\lambda \in \mathbb{R}_{++}$,

$$\frac{\|x\|^2}{\lambda} + \lambda\|v\|^2 - 2\|x\|\|v\| \geq 0.$$

Then we get

$$\begin{aligned}&\frac{\|x\|^2}{\tau} - \|v\|\|L\|\|x\| + \frac{\|v\|^2}{\sigma} \\ \geq &\frac{\|x\|^2}{\tau} + \frac{\|v\|^2}{\sigma} - \frac{1}{2}(\frac{\sigma\|L\|^2}{\sqrt{\tau\sigma\|L\|^2}}\|x\|^2 + \frac{\sqrt{\tau\sigma\|L\|^2}}{\sigma}\|v\|^2) \\ \geq &\frac{\|x\|^2}{\tau} + \frac{\|v\|^2}{\sigma} - \frac{1}{2}(\frac{\sqrt{\tau\sigma\|L\|^2}}{\tau}\|x\|^2 + \frac{\sqrt{\tau\sigma\|L\|^2}}{\sigma}\|v\|^2) \\ \geq &(1 - \frac{1}{2}\sqrt{\tau\sigma\|L\|^2})(\frac{\|x\|^2}{\tau} + \frac{\|v\|^2}{\sigma}) \\ \geq &(1 - \frac{1}{2}\sqrt{\tau\sigma\|L\|^2}) \min\{\frac{1}{\tau}, \frac{1}{\sigma}\}\|a\|^2.\end{aligned}$$

Let $\rho = (1 - \frac{1}{2}\sqrt{\tau\sigma\|L\|^2}) \min\{\frac{1}{\tau}, \frac{1}{\sigma}\}$. Then $\langle Va, a \rangle \geq \rho\|a\|^2$. Since τ and σ satisfy the condition $\tau\sigma\|L\|^2 < 4$, we have $\rho > 0$. That means, V is $\rho-$strongly positive.
Let $a = (x, v), b = (y, u), a \in \mathcal{K}, b \in \mathcal{K}$. Since V is a bounded linear operator, we have :

$$\begin{aligned}\langle Va - Vb, a - b \rangle &= \langle V(a - b), a - b \rangle \\ &\geq \rho\|a - b\|^2 \\ &\geq 0.\end{aligned}$$

Thus, V is maximally monotone.
To prove the existence of V^{-1}, we only need to prove V is one-to-one (in other

words, $\operatorname{zer} V = \{0\}$), and is onto (in other words, $\operatorname{ran} V = \mathcal{K}$).
Because V is ρ-strongly positive, $\operatorname{zer} V = \{0\}$. Otherwise, suppose we have $x \in \operatorname{zer} V$ with $x \neq 0$. Then $\langle x, 0\rangle = \langle x, Vx\rangle \geq \rho\|x\|^2 > 0$, which is impossible.
According to Fact 1.2.17, since V is a bounded linear operator, $\operatorname{ran} V$ is a closed subspace, that is, $\operatorname{ran} V = \overline{\operatorname{ran}} V$. By Fact 1.2.18,

$$(\operatorname{ran} V)^{\perp} = \operatorname{zer} V^{\star}. \tag{2.6}$$

As V is self-adjoint, $V^{\star} = V$. Thus,

$$\operatorname{zer} V^{\star} = \operatorname{zer} V = \{0\}. \tag{2.7}$$

Combining the result of (2.6) and (2.7), we get

$$(\operatorname{ran} V)^{\perp} = \{0\}.$$

Because $(\operatorname{ran} V)^{\perp} = \{0\} \subseteq \mathcal{K}^{\perp}$, by Fact 1.1.27,

$$\mathcal{K}^{\perp\perp} \subseteq (\operatorname{ran} V)^{\perp\perp}. \tag{2.8}$$

Since $\operatorname{ran} V$ and $\mathcal{K}$ are closed subspaces, by Fact 1.1.28, $\mathcal{K}^{\perp\perp} = \mathcal{K}$, $(\operatorname{ran} V)^{\perp\perp} = \operatorname{ran} V$. Thus, (2.8) implies that $\mathcal{K} \subseteq \operatorname{ran} V$. Therefore,

$$\operatorname{ran} V = \mathcal{K}.$$

That means, V^{-1} exists.

Step 3: Show that $(\bar{x}, \bar{v}) \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B})$ if and only if $\bar{x}$ is a primal solution of (P) and $\bar{v}$ is a dual solution of (D)

Since S, M and Q are maximally monotone and $\operatorname{dom} S = \mathcal{K}$, according to Fact 1.2.28, $\frac{1}{2}S + Q$ and $\frac{1}{2}S + M$ are maximally monotone on $\mathcal{K}$.
Take $\frac{1}{2}S + Q$ as an example. Let

$$(x, u) \in \operatorname{gra}(\frac{1}{2}S + Q); \quad (y, v) \in \operatorname{gra}(\frac{1}{2}S + Q).$$

As $\frac{1}{2}S + Q$ is maximally monotone on space $\mathcal{K}$, we have

$$\langle x - y, u - v\rangle_{\mathcal{K}} \geq 0. \tag{2.9}$$

Because V^{-1} exists, (2.9) can be written as $\langle x - y, VV^{-1}(u - v)\rangle_{\mathcal{K}} \geq 0$, which equals

$$\langle x - y, V^{-1}(u - v)\rangle_{\mathcal{K}V} \geq 0. \tag{2.10}$$

By Fact 1.2.15,

$$(2.10) \Leftrightarrow \langle x-y, V^{-1}u - V^{-1}v\rangle_{\mathcal{K}V} \geq 0.$$

That means $\boldsymbol{A} := V^{-1}(\frac{1}{2}S+Q)$ is a maximally monotone operator on space $\mathcal{K}V$. By using the same method, we can prove $\boldsymbol{B} := V^{-1}(\frac{1}{2}S+M)$ is a maximally monotone operator on space $\mathcal{K}V$ too.

Again, since V^{-1} is linear,

$$\begin{aligned} \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) &= \operatorname{zer}(V^{-1}(\frac{1}{2}S+M) + V^{-1}(\frac{1}{2}S+Q)) \\ &= \operatorname{zer}(V^{-1}(S+M+Q)). \end{aligned}$$

(1) On the one hand, let $x \in \operatorname{zer}(V^{-1}(S+M+Q))$. Then

$$\begin{aligned} (V^{-1}(S+M+Q))(x) &= 0 \\ \text{i.e.,} \quad (S+M+Q)(x) &= V(0) \\ &= 0. \end{aligned}$$

Thus, $x \in \operatorname{zer}(S+M+Q)$.

(2) On the other hand, let $x \in \operatorname{zer}(S+M+Q)$. Then we have

$$\begin{aligned} (V^{-1}(S+M+Q))(x) &= V^{-1}((S+M+Q)(x)) \\ &= V^{-1}(0) \\ &= 0. \end{aligned}$$

Thus, $x \in \operatorname{zer}(V^{-1}(S+M+Q))$.

Altogether, we have

$$\operatorname{zer}(V^{-1}(S+M+Q)) = \operatorname{zer}(S+M+Q).$$

Consequently, one has

$$\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B}) = \operatorname{zer}(S+M+Q).$$

If $\operatorname{zer}(M+S+Q) \neq \emptyset$, then according to the definition of M, S, Q, there exists $(\bar{x}, \bar{v}) \in \mathcal{K}$ such that

$$\begin{cases} 0 \in -z + A\bar{x} + L^{\star}\bar{v}, \\ 0 \in r + B^{-1}\bar{v} + D^{-1}\bar{v} - L\bar{x}, \end{cases}$$

which is equivalent to

$$\begin{cases} z - L^\star \bar{v} \in A\bar{x}, \\ \bar{v} \in (B\Box D)(L\bar{x} - r). \end{cases} \tag{2.11}$$

That is, $\bar{v}$ is the solution of the dual inclusion problem. Moreover, since L is a nonzero linear operator, $z - L^\star \bar{v} \in A\bar{x}$ means $0 \in -z + A\bar{x} + L^\star \bar{v}$. Then, those two inclusions in (2.11) implies

$$z \in A\bar{x} + L^\star (B\Box D)(L\bar{x} - r),$$

i.e., $\bar{x}$ is the solution of the primal inclusion problem. Altogether, we call $(\bar{x}, \bar{v})$ the primal-dual solution.
Therefore, we deduce that $(\bar{x}, \bar{v}) \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B})$ if and only if $\bar{x}$ is a solution of the primal inclusion problem (P) and $\bar{v}$ is a solution of the dual inclusion problem (D). $\square$

Remark 2.4. For the operator $V : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (\frac{x}{\tau} - \frac{1}{2}L^\star v, \frac{v}{\sigma} - \frac{1}{2}Lx)$, it is conjectured that we can consider it as a matrix and get $\det V = \frac{1}{\tau\sigma} - \frac{\|L\|^2}{4}$. Therefore, if we have $\tau\sigma\|L\|^2 < 4$, V^{-1} exists.

Remark 2.5. According to the definition of the Douglas-Rachford Splitting Problem, Lemma 2.3.4 implies that the primal inclusion problem (P) and the dual inclusion problem (D) can be solved by using the Douglas-Rachford algorithm.

2.4 Application to proper, lower-semicontinuous convex functions

Before we show the relationship between the primal dual inclusion problems and the optimization problems, we must get the following results, which can be found in [4].

Theorem 2.4.1. Let C be a convex subset of $\mathcal{H}$, let $\mathcal{K}$ be a real Hilbert space, let $L : \mathcal{H} \to \mathcal{K}$ be linear and continuous, and let D be a convex subset of $\mathcal{K}$. If $D \cap \operatorname{int} L(C) \neq \emptyset$ or $\operatorname{int} D \cap L(C) \neq \emptyset$, then $0 \in \operatorname{int}(D - L(C))$.

Proof. Suppose that $y \in D \cap \operatorname{int} L(C)$. Then there exists an open ball $\mathbb{B}(y; r) \subseteq L(C)$ for some $r \in \mathbb{R}_{++}$. Clearly

$$\mathbb{B}(0; r) = y - \mathbb{B}(y; r). \tag{2.12}$$

Since $y \in D$, $\mathbb{B}(y; r) \subseteq L(C)$, equation (2.12) implies $\mathbb{B}(0; r) \subseteq D - \mathbb{B}(y; r)$. Therefore,

$$0 \in \operatorname{int}(D - L(C)).$$

In the case of $\operatorname{int} D \cap L(C) \neq \emptyset$, we let $x \in \operatorname{int} D \cap L(C)$. Then there exists an open ball $\mathbb{B}(x;r) \subseteq D$ for some $r \in \mathbb{R}_{++}$. Clearly

$$\mathbb{B}(0;r) = x - \mathbb{B}(x;r). \tag{2.13}$$

Since $x \in L(C)$, $\mathbb{B}(x;r) \subseteq D$, equation (2.13) implies $\mathbb{B}(0;r) \subseteq L(C) - D$. Therefore,

$$0 \in \operatorname{int}(D - L(C)).$$

$\square$

Theorem 2.4.2. Let $f \in \Gamma_0(\mathcal{H}), g \in \Gamma_0(\mathcal{H})$. Then

$$\partial f \Box \partial g \subseteq \partial(f \Box g).$$

Proof. If $(\partial f \Box \partial g)(z) = \emptyset$, clearly the inclusion holds. Assume $(\partial f \Box \partial g)(z) \neq \emptyset$. Let $v \in (\partial f \Box \partial g)(z)$. Since $f \in \Gamma_0(\mathcal{H}), g \in \Gamma_0(\mathcal{H})$, by Fact 1.3.27, ∂f and ∂g are maximally monotone. According to the definition of the parallel sum between operators (Definition 1.2.2), $v \in (\partial f \Box \partial g)(z)$ implies $v \in ((\partial f)^{-1} + (\partial g)^{-1})^{-1}(z)$, i.e., $z \in ((\partial f)^{-1} + (\partial g)^{-1})(v)$. In other words, $z \in (\partial f)^{-1}(v) + (\partial g)^{-1}(v)$.
Let $a_1 \in (\partial f)^{-1}(v)$ and $a_2 \in (\partial g)^{-1}(v)$ such that $z = a_1 + a_2$. Then

$$\begin{cases} v \in \partial f(a_1), \\ v \in \partial g(a_2), \end{cases}$$

so

$$v \in \partial f(a_1) \cap \partial g(a_2) \quad (a_1 + a_2 = z).$$

According to the Fact 1.3.52,

$$v \in \partial f(a_1) \cap \partial g(a_2) \Leftrightarrow v \in \partial(f \Box g)(z).$$

Thus,

$$\partial f \Box \partial g \subseteq \partial(f \Box g).$$

$\square$

Theorem 2.4.3. Let $f \in \Gamma_0(\mathbb{R}^m), g \in \Gamma_0(\mathbb{R}^m)$. If $\operatorname{dom} f^* \cap \operatorname{int} \operatorname{dom} g^* \neq \emptyset$, then

(1) $0 \in \operatorname{int}(\operatorname{dom} f^* - \operatorname{dom} g^*)$.

(2) $f \Box g = f \boxdot g \in \Gamma_0(\mathbb{R}^m)$.

(3) $\partial(f \Box g) = \partial f \Box \partial g$.

Proof. (1) Since $f \in \Gamma_0(\mathbb{R}^m), g \in \Gamma_0(\mathbb{R}^m)$, by Fact 1.3.39 and Fact 1.3.40, f^* and g^* are in $\Gamma_0(\mathbb{R}^m)$. Thus, $\operatorname{dom} f^*$ and $\operatorname{dom} g^*$ are convex.

Because $\operatorname{dom} f^* \cap \operatorname{int} \operatorname{dom} g^* \neq \emptyset$, due to Theorem 2.4.1,

$$0 \in \operatorname{int}(\operatorname{dom} f^* - \operatorname{dom} g^*).$$

(2) Using the result of (1) with the Fact 1.3.51 to complete the proof that $f \Box g = f \boxdot g \in \Gamma_0(\mathbb{R}^m)$.

(3) Since we proved $f \Box g = f \boxdot g \in \Gamma_0(\mathbb{R}^m)$ above, according to Fact 1.3.40, $f \Box g = (f \Box g)^{**}$. Thus,

$$\begin{aligned} \partial(f \Box g) &= \partial(f \Box g)^{**} \\ &= \partial[(f \Box g)^*]^*. \end{aligned} \tag{2.14}$$

Since $f \Box g \in \Gamma_0(\mathbb{R}^m)$, combine Fact 1.3.39 and Fact 1.3.40, we have

$$(f \Box g)^* \in \Gamma_0(\mathbb{R}^m).$$

Therefore, by Fact 1.3.46, $\partial[(f \Box g)^*]^* = [\partial(f \Box g)^*]^{-1}$.

By Fact 1.3.49, $[\partial(f \Box g)^*]^{-1} = [\partial(f^* + g^*)]^{-1}$. As $\operatorname{dom} f^* \cap \operatorname{int} \operatorname{dom} g^* \neq \varnothing$, the sum rule for subdifferentials (Fact 1.3.31) gives

$$[\partial(f^* + g^*)]^{-1} = [\partial f^* + \partial g^*]^{-1}. \tag{2.15}$$

Again, because $f \in \Gamma_0(\mathbb{R}^m), g \in \Gamma_0(\mathbb{R}^m)$, by Fact 1.3.46,

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

Thogether with (2.15) yeilds

$$\begin{aligned} [\partial f^* + \partial g^*]^{-1} &= [(\partial f)^{-1} + (\partial g)^{-1}]^{-1} \\ &= \partial f \Box \partial g. \end{aligned}$$

□

Lemma 2.6. *Let $\mathcal{K}$ be a real Hilbert space, let $f \in \Gamma_0(\mathcal{H})$, let $g \in \Gamma_0(\mathcal{K})$, and let $L : \mathcal{H} \to \mathcal{K}$ be a nonzero bounded linear invertible operator. Suppose $[L(\operatorname{dom} f) + b] \cap \operatorname{intdom} g \neq \emptyset$. Then $\partial[f(x) + g(Lx + b)] = \partial f(x) + L^\star \circ \partial g \circ (Lx + b)$.*

Proof. Let $h(x) = g(Lx + b)$, which can implies

$$\operatorname{dom} h = L^{-1}(\operatorname{dom} g - b).$$

Since $[L(\operatorname{dom} f) + b] \cap \operatorname{intdom} g \neq \emptyset$, there exists an $x_0 \in \operatorname{dom} f$ such that

$$Lx_0 + b \in \operatorname{intdom} g,$$

i.e.,

$$x_0 \in L^{-1}(\operatorname{intdom} g - b) \subset \operatorname{dom} h.$$

Because $L^{-1}(\operatorname{intdom} g - b)$ is open, $x_0 \in \operatorname{intdom} h$. Therefore,

$$x_0 \in \operatorname{dom} f \cap \operatorname{intdom} h.$$

By Fact 1.3.31,

$$\partial(f + h) = \partial f + \partial h.$$

Because $h(x) = g(Lx + b)$, we have

$$\partial h(x) = L^\star \circ \partial g \circ (Lx + b)$$

by Fact 1.3.30. Therefore,

$$\partial[f(x) + g(Lx + b)] = \partial f(x) + L^\star \circ \partial g \circ (Lx + b).$$

□

Theorem 2.4.4. By the definition of the primal inclusion problem (P), with $A = \partial f, B = \partial g, D = \partial l$, where $f, g, l \in \Gamma_0(\mathbb{R}^m)$. Then

(1) We obtain the primal inclusion problem

$$\text{find } \bar{x} \in \mathbb{R}^m \text{ such that } z \in \partial f(\bar{x}) + L^\star \circ (\partial g \Box \partial l) \circ (L\bar{x} - r) \tag{2.16}$$

(2) Every solution of (2.16) is also a solution of the optimization problem

$$\operatorname{Argmin}_{x \in \mathbb{R}^m}\{f(x) + ((g \Box l)(Lx - r)) - \langle z, x\rangle\}. \tag{2.17}$$

(3) If in addition, $[L(\operatorname{dom} f) - r] \cap \operatorname{intdom}(g \Box l) \neq \emptyset$, and $\operatorname{dom} g^* \cap \operatorname{int} \operatorname{dom} l^* \neq \emptyset$, then (2.16) and (2.17) are equivalent.

Proof. (1) By Fact 1.3.27, $\partial f, \partial g$ and ∂l are maximally monotone. Thus, once we plug them into the dual problem (P) by letting $A = \partial f, B = \partial g, D = \partial l$, we obtain equation (2.16).

(2) Now we show every solution $\bar{x}$ of (2.16) is also a solution of (2.17).
First, let's move z from the left side of equation (2.16) to its right side. We obtain
$$0 \in \partial f(\bar{x}) + L^{\star} \circ (\partial g \Box \partial l) \circ (L\bar{x} - r) - \partial\langle z, \bar{x}\rangle. \tag{2.18}$$
Due to Theorem 2.4.2, since $g, l \in \Gamma_0(\mathbb{R}^m)$, equation (2.18) implies
$$0 \in \partial f(\bar{x}) + L^{\star} \circ \partial(g \Box l) \circ (L\bar{x} - r) - \partial\langle z, \bar{x}\rangle.$$
Take
$$\begin{cases} v_1 \in \partial f(\bar{x}), \\ v_2 \in \partial(g \Box l) \circ (L\bar{x} - r), \\ v_3 \in \partial\langle -z, \bar{x}\rangle, \end{cases}$$
such that $v_1 + L^{\star}v_2 + v_3$ is a generic point in $\partial f(\bar{x}) + L^{\star} \circ \partial(g \Box l) \circ (L\bar{x} - r) + \partial\langle -z, \bar{x}\rangle$. By the definition of subdifferential, for all $y \in \mathbb{R}^m$
$$\begin{cases} \langle v_1, y - \bar{x}\rangle \leq f(y) - f(\bar{x}), \\ \langle v_2, Ly - r - (L\bar{x} - r)\rangle \leq (g \Box l)(Ly - r) - (g \Box l)(L\bar{x} - r), \\ \langle v_3, y - \bar{x}\rangle \leq \langle z, \bar{x}\rangle - \langle z, y\rangle. \end{cases}$$
Due to $\langle v_2, Ly - r - (L\bar{x} - r)\rangle = \langle L^{\star}v_2, y - \bar{x}\rangle$, we have
$$\begin{aligned} \langle v_1 + L^{\star}v_2 + v_3, y - \bar{x}\rangle &\leq f(y) - f(\bar{x}) + [(g \Box l)(Ly - r) \\ &\quad - (g \Box l)(L\bar{x} - r)] - \langle z, y\rangle + \langle z, \bar{x}\rangle \\ &= f(y) + (g \Box l)(Ly - r) - \langle z, y\rangle \\ &\quad - f(\bar{x}) - (g \Box l)(L\bar{x} - r) + \langle z, \bar{x}\rangle. \end{aligned}$$
In turn,
$$v_1 + L^{\star}v_2 + v_3 \in \partial(f(\bar{x}) + (g \Box l)(L\bar{x} - r) - \langle z, \bar{x}\rangle).$$
Therefore,
$$\partial f(\bar{x}) + L^{\star} \circ \partial(g \Box l) \circ (L\bar{x} - r) + \partial\langle -z, \bar{x}\rangle \subseteq \partial(f(\bar{x}) + (g \Box l)(L\bar{x} - r) - \langle z, \bar{x}\rangle). \tag{2.19}$$
Since $0 \in \partial f(\bar{x}) + L^{\star} \circ \partial(g \Box l) \circ (L\bar{x} - r) - \partial\langle z, \bar{x}\rangle$, (2.19) implies that
$$0 \in \partial(f(\bar{x}) + (g \Box l)(L\bar{x} - r) - \langle z, \bar{x}\rangle). \tag{2.20}$$
By Fact 1.3.34, the $\bar{x}$ which satisfies the inclusion (2.20) is also an element of the set
$$\operatorname{Argmin}_{x \in \mathbb{R}^m}\{f(x) + ((g \Box l)(Lx - r)) - \langle z, x\rangle\},$$
and vice versa.

(3) It suffices prove

$$\partial f(\bar{x})+L^{\star}\circ\partial(g\Box l)\circ(L\bar{x}-r)+\partial\langle -z,\bar{x}\rangle = \partial(f(\bar{x})+\partial(g\Box l)(L\bar{x}-r)-\langle z,\bar{x}\rangle). \tag{2.21}$$

Because $g, l \in \Gamma_0(\mathbb{R}^m)$, and $\operatorname{dom} g^* \cap \operatorname{int}\operatorname{dom} l^* \neq \emptyset$, using the Theorem 2.4.3 we have

$$\partial f(\bar{x})+L^{\star}(\partial g\Box\partial l)(L\bar{x}-r)-\partial\langle z,\bar{x}\rangle = \partial f(\bar{x})+L^{\star}\partial(g\Box l)(L\bar{x}-r)-\partial\langle z,\bar{x}\rangle.$$

Again, by using the same theorem, we have $g\Box l \in \Gamma_0(\mathbb{R}^m)$. Then we can apply the Lemma 2.6 to get the conclusion that

$$\partial f(\bar{x}) + L^{\star}\circ\partial(g\Box l)\circ(L\bar{x}-r) = \partial[f(\bar{x}) + (g\Box l)(L\bar{x}-r)]$$

since we have the condition $[L(\operatorname{dom} f) - r] \cap \operatorname{intdom}(g\Box l) \neq \emptyset$. As

$$\operatorname{dom}\langle z, x\rangle = \mathbb{R}^m,$$

we have

$$\begin{aligned}
&[L(\operatorname{dom} f) - r] \cap \operatorname{intdom}(g\Box l) \cap \operatorname{intdom}\langle z,\bar{x}\rangle\\
=&[L(\operatorname{dom} f) - r] \cap \operatorname{intdom}(g\Box l) \cap \mathbb{R}^m\\
=&[L(\operatorname{dom} f) - r] \cap \operatorname{intdom}(g\Box l) \neq \emptyset.
\end{aligned}$$

Thus, by Fact 1.3.31, we get equation (2.21).

□

Theorem 2.4.5. By the definition of the dual problem (D), with $A = \partial f, B = \partial g, D = \partial l$, where $f, g, l \in \Gamma_0(\mathbb{R}^m)$. Then

(1) We obtain the dual inclusion problem

$$\text{find } \bar{v} \in \mathbb{R}^m \text{ such that } (\exists \bar{x} \in \mathbb{R}^m) \begin{cases} z - L^{\star}\bar{v} \in \partial f(\bar{x}) \\ \bar{v} \in (\partial g\Box\partial l)(L\bar{x}-r). \end{cases} \tag{2.22}$$

(2) Every solution $\bar{v}$ of (2.22) is also a solution of the the optimization problem

$$\operatorname{Argmin}_{v\in\mathbb{R}^m}\{(g^*+l^*)(v) + f^*(z - L^{\star}v) + \langle r, v\rangle\}. \tag{2.23}$$

(3) If in addition, $[-L^{\star}\operatorname{dom}(g^*+l^*) + z] \cap \operatorname{intdom}(f^*) \neq \emptyset$, and $\operatorname{dom} g^* \cap \operatorname{int}\operatorname{dom} l^* \neq \emptyset$, then (2.22) and (2.23) are equivalent.

Proof. (1) By Fact 1.3.27, $\partial f, \partial g$ and ∂l are maximally monotone. Thus, once we plug them into the dual problem (D) by letting $A = \partial f, B = \partial g, D = \partial l$, we obtain equation (2.22).

(2) Now we show every solution of (2.22) is also a solution of (2.23). Due to Theorem 2.4.2, since $g, l \in \Gamma_0(\mathbb{R}^m)$,

$$\bar{v} \in (\partial g \Box \partial l)(L\bar{x} - r) \text{ implies } \bar{v} \in [\partial(g \Box l)](L\bar{x} - r). \tag{2.24}$$

Moreover, by Fact 1.3.42, (2.24) implies $L\bar{x} - r \in \partial((g \Box l)^*)(\bar{v})$. In general,

$$(2.22) \Leftrightarrow \begin{cases} \bar{x} \in \partial f^* \circ (z - L^\star \bar{v}) & (2.25\text{a}) \\ L\bar{x} - r \in \partial((g \Box l)^*)(\bar{v}) & (2.25\text{b}) \end{cases}$$

Multiplying (2.25a) by L, we obtain

$$L\bar{x} \in L \circ \partial f^* \circ (z - L^\star \bar{v}). \tag{2.26}$$

$$(2.25\text{b}) - (2.26) \Rightarrow 0 \in \partial[(g \Box l)^*](\bar{v}) + \partial\langle r, \bar{v}\rangle - L \circ \partial f^* \circ (z - L^\star \bar{v}).$$

Since $g \in \Gamma_0(\mathbb{R}^m)$ and $l \in \Gamma_0(\mathbb{R}^m)$, by Fact 1.3.49, $(g \Box l)^* = g^* + l^*$. That is,

$$0 \in \partial(g^* + l^*)(\bar{v}) - L \circ \partial f^* \circ (z - L^\star \bar{v}) + \partial\langle r, \bar{v}\rangle.$$

Take

$$\begin{cases} v_1 \in \partial(g^* + l^*)(\bar{v}), \\ v_2 \in \partial f^* \circ (z - L^\star \bar{v}), \\ v_3 \in \partial\langle r, \bar{v}\rangle, \end{cases}$$

such that $v_1 - Lv_2 + v_3$ is a generic point in $\partial(g^* + l^*)(\bar{v}) - L \circ \partial f^* \circ (z - L^\star \bar{v}) + \partial\langle r, \bar{v}\rangle$. By the definition of subdifferential, for all $y \in \mathbb{R}^m$

$$\begin{cases} \langle v_1, y - \bar{v}\rangle \leq (g^* + l^*)(y) - (g^* + l^*)(\bar{v}), \\ \langle v_2, z - L^\star y - (z - L^\star \bar{v})\rangle \leq f^*(z - L^\star y) - f^*(z - L^\star \bar{v}), \\ \langle v_3, y - \bar{v}\rangle \leq \langle r, y\rangle - \langle r, \bar{v}\rangle. \end{cases}$$

Due to $\langle v_2, z - L^\star y - (z - L^\star \bar{v})\rangle = \langle -Lv_2, y - \bar{v}\rangle$, for all $y \in \mathbb{R}^m$:

$$\begin{aligned} \langle v_1 - Lv_2 + v_3, y - \bar{v}\rangle &\leq (g^* + l^*)(y) - (g^* + l^*)(\bar{v}) + f^*(z - L^\star y) \\ &\quad - f^*(z - L^\star \bar{v}) + \langle r, y\rangle - \langle r, \bar{v}\rangle \\ &= (g^* + l^*)(y) + f^*(z - L^\star y) + \langle r, y\rangle - \\ &\quad ((g^* + l^*)(\bar{v}) + f^*(z - L^\star \bar{v}) + \langle r, \bar{v}\rangle) \end{aligned}$$

in turn,

$$v_1 - Lv_2 + v_3 \in \partial((g^* + l^*)(\bar{v}) + f^*(z - L^\star \bar{v}) + \langle r, \bar{v}\rangle).$$

Therefore,

$$\begin{aligned}&\partial(g^* + l^*)(\bar{v}) - L \circ \partial f^* \circ (z - L^\star \bar{v}) + \partial\langle r, \bar{v}\rangle \\ &\quad \subseteq \partial((g^* + l^*)(\bar{v}) + f^*(z - L^\star \bar{v}) + \langle r, \bar{v}\rangle).\end{aligned} \tag{2.27}$$

Since $0 \in \partial(g^* + l^*)(\bar{v}) - L \circ \partial f^* \circ (z - L^\star \bar{v}) + \partial\langle r, \bar{v}\rangle$, (2.27) implies that

$$0 \in \partial((g^* + l^*)(\bar{v}) + f^*(z - L^\star \bar{v}) + \langle r, \bar{v}\rangle). \tag{2.28}$$

By Fact 1.3.34, the $\bar{v}$ which satisfies the inclusion (2.28) is also an element of the set

$$\operatorname{Argmin}_{v \in \mathbb{R}^m}\{(g^* + l^*)(v) + f^*(z - L^\star v) + \langle r, v\rangle\},$$

and vice versa.

(3) It suffices to prove

$$\begin{aligned}&\partial(g^* + l^*)(\bar{v}) - L \circ \partial f^* \circ (z - L^\star \bar{v}) + \partial\langle r, \bar{v}\rangle \\ &\quad = \partial((g^* + l^*)(\bar{v}) + f^*(z - L^\star \bar{v}) + \langle r, \bar{v}\rangle).\end{aligned} \tag{2.29}$$

Because $g, l \in \Gamma_0(\mathbb{R}^m)$, and $\operatorname{dom} g^* \cap \operatorname{int}\operatorname{dom} l^* \neq \emptyset$, using the Theorem 1.3.41 we have $(g^* + l^*) \in \Gamma_0(\mathbb{R}^m)$. Then we can apply the Lemma 2.6 to get the conclusion that

$$\partial(g^* + l^*)(\bar{v}) - L \circ \partial f^* \circ (z - L^\star \bar{v}) = \partial((g^* + l^*)(\bar{v}) + f^*(z - L^\star \bar{v}))$$

since we have the condition $[-L^\star \operatorname{dom}(g^* + l^*) + z] \cap \operatorname{intdom}(f^*) \neq \emptyset$. As $\operatorname{dom}\langle r, v\rangle = \mathbb{R}^m$, we have

$$\begin{aligned}&[-L^\star \operatorname{dom}(g^* + l^*) + z] \cap \operatorname{intdom}(f^*) \cap \operatorname{intdom}\langle r, \bar{v}\rangle \\ =&[-L^\star \operatorname{dom}(g^* + l^*) + z] \cap \operatorname{intdom}(f^*) \cap \mathbb{R}^m \\ =&[-L^\star \operatorname{dom}(g^* + l^*) + z] \cap \operatorname{intdom}(f^*) \neq \emptyset.\end{aligned}$$

Thus, by Fact 1.3.31, we get equation (2.29).

□

Chapter 3

Douglas-Rachford algorithm with a changed parameter

3.1 Overview

As we saw in the previous chapter, the classic Douglas-Rachford algorithm has a parameter 2 in its iteration. That parameter gives me an inspiration: what if we changed the value of that parameter? In this chapter, a new algorithm which based on the classic Douglas-Rachford algorithm is constructed. I will prove the properties of this algorithm and then show that it can be applied on the composited monotone inclusion problems and on the optimization problems.

3.2 α-Douglas-Rachford algorithm, with parameter $\alpha \in [1, 2)$

First, let's see some theorems and facts.

Theorem 3.2.1. If A and B are maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$, and $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$, then $\operatorname{zer}(A + B + \gamma \operatorname{Id}) \neq \emptyset$ when $\gamma \in \mathbb{R}_{++}$.

Proof. As A and B are maximally monotone operators and $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$, according to Fact 1.2.28, $A + B$ is a maximally monotone operator. By Fact 1.2.29, $\frac{1}{\gamma}(A+B)$ is also maximally monotone. Let $\bar{A} = \frac{1}{\gamma}(A+B)$. According to Fact 1.2.32,

$$\operatorname{ran}(\bar{A} + \operatorname{Id}) = \mathcal{H} \Rightarrow 0 \in \operatorname{ran}(\bar{A} + \operatorname{Id}),$$

Then, $\operatorname{zer}(\bar{A} + \operatorname{Id}) \neq \emptyset$. Because

$$\begin{aligned} \operatorname{zer}(\bar{A} + \operatorname{Id}) &= \operatorname{zer}[\gamma(\bar{A} + \operatorname{Id})], \\ &= \operatorname{zer}(A + B + \gamma \operatorname{Id}), \end{aligned}$$

we have $\operatorname{zer}(A + B + \gamma \operatorname{Id}) \neq \emptyset$. ☐

Fact 3.2.2. [4, Corollary 26.8](see also [2, Corollary 3]) Let m be an integer such that $m \geq 2$, set $I = \{1, \ldots, m\}$, and let

$$(A_i)_{i\in I} : \mathcal{H} \to 2^{\mathcal{H}}$$

be maximally monotone operators. For every $i \in I$, let $(x_{i,n}, u_{i,n})_{n=1}^{+\infty}$ be a sequence in $\operatorname{gra} A_i$ and let $(x_i, u_i) \in \mathcal{H} \times \mathcal{H}$. Suppose that

$$\sum_{i\in I} u_{i,n} \to 0 \quad \text{and } (\forall i \in I) \begin{cases} x_{i,n} \rightharpoonup x_i \\ u_{i,n} \rightharpoonup u_i \\ m x_{i,n} - \sum_{j\in I} x_{j,n} \to 0. \end{cases}$$

Then there exists $x \in \operatorname{zer} \sum_{i\in I} A_i$ such that the following hold:

(i) $x = x_1 = \cdots = x_m$.

(ii) $\sum_{i\in I} u_i = 0$.

(iii) $(\forall i \in I) \quad (x, u_i) \in \operatorname{gra} A_i$.

(iv) $\sum_{i\in I}\langle x_{i,n}, u_{i,n}\rangle \to \langle x, \sum_{i\in I} u_i\rangle = 0$.

Theorem 3.2.3. Let A be a maximally monotone operator from $\mathcal{K}$ to $2^{\mathcal{K}}$, let $\alpha \in [1, 2)$. Define $R_A^\alpha = \alpha J_A - \operatorname{Id}$. Then R_A^α is nonexpansive.

Proof. Let $x, y \in \mathcal{K}$. We derive an equivalent criterion for nonexpansivity:

$$\begin{aligned}
&R_A^\alpha \text{ is nonexpansive}\\
\Leftrightarrow&\|R_A^\alpha x - R_A^\alpha y\| \leq \|x - y\|\\
\Leftrightarrow&\|(\alpha J_A - \operatorname{Id})x - (\alpha J_A - \operatorname{Id})y\| \leq \|x - y\|\\
\Leftrightarrow&\|\alpha(J_A x - J_A y) - (x - y)\| \leq \|x - y\|\\
\Leftrightarrow&\|\alpha(J_A x - J_A y) - (x - y)\|^2 \leq \|x - y\|^2\\
\Leftrightarrow&\langle\alpha(J_A x - J_A y) - (x - y), \alpha(J_A x - J_A y) - (x - y)\rangle \leq \|x - y\|^2\\
\Leftrightarrow&\alpha^2\|(J_A x - J_A y)\|^2 + \|x - y\|^2 - 2\alpha\langle J_A x - J_A y, x - y\rangle \leq \|x - y\|^2\\
\Leftrightarrow&\alpha^2\|(J_A x - J_A y)\|^2 \leq 2\alpha\langle J_A x - J_A y, x - y\rangle\\
\Leftrightarrow&\frac{\alpha}{2}\|(J_A x - J_A y)\|^2 \leq \langle J_A x - J_A y, x - y\rangle
\end{aligned}$$

Since A is maximally monotone, it follows from Fact 1.2.38 that J_A is firmly nonexpansive. According to Fact 1.2.21, we have

$$(\forall x \in \mathcal{H}) \quad (\forall y \in \mathcal{H}) \quad \|J_A x - J_A y\|^2 \leq \langle x - y, J_A x - J_A y\rangle.$$

As $\alpha \in [1,2]$, $\frac{\alpha}{2} \in [\frac{1}{2}, 1]$,

$$\frac{\alpha}{2}\|(J_A x - J_A y)\|^2 \leq \|J_A x - J_A y\|^2 \leq \langle J_A x - J_A y, x - y\rangle.$$

Thus, R_A^α is nonexpansive. □

Remark 3.1. Theorem 3.2.3 holds whenever $0 \leq \alpha \leq 2$.

Similarly, if we let B be a maximally monotone operator from $\mathcal{K}$ to $2^{\mathcal{K}}$, and define $R_B^\alpha = \alpha J_B - \mathrm{Id}$, R_B^α is nonexpansive.

Theorem 3.2.4. Let $\alpha \in [1,2)$, let A, B be maximally monotone operators from $\mathcal{K}$ to $2^{\mathcal{K}}$, and $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$. Let $T = R_A^\alpha \circ R_B^\alpha$. Then

(1) T is nonexpansive;

(2) $J_B(\operatorname{Fix} T) = \operatorname{zer}(A + B + (2-\alpha)\operatorname{Id})$;

(3) $\operatorname{Fix} T \neq \emptyset$.

Proof. (1) According to Theorem 3.2.3, both R_A^α and R_B^α are nonexpansive. Thus, for any $x, y \in \mathcal{K}$,

$$\begin{aligned}\|Tx - Ty\| &= \|R_A^\alpha \circ R_B^\alpha x - R_A^\alpha \circ R_B^\alpha y\|\\ &\leq \|R_B^\alpha x - R_B^\alpha y\|\\ &\leq \|x - y\|,\end{aligned}$$

that is, T is nonexpansive.

(2) Since $A + B$ is maximally monotone, according to Theorem 3.2.1,

$$\operatorname{zer}(A + B + (2-\alpha)\operatorname{Id}) \neq \emptyset.$$

Consider an arbitrary $x \in \mathcal{K}$ from this set, i.e.,

$$0 \in Ax + Bx + (2-\alpha)x.$$

Therefore, there exists $y \in \mathcal{K}$ such that

$$x - y \in Ax + (2-\alpha)x \text{ and } y - x \in Bx,$$

which is equivalent to

$$(\alpha - 1)x - y \in Ax \text{ and } x = J_B y.$$

Thus,

$$\alpha J_B y - y \in A \circ J_B y + J_B y.$$

which implies

$$J_B y = J_A(\alpha J_B y - y).$$

Therefore,

$$\begin{aligned}
&0 = \alpha J_A(\alpha J_B y - y) - \alpha J_B y \\
\Leftrightarrow &y = \alpha J_A(\alpha J_B y - y) - (\alpha J_B y - y) \\
\Leftrightarrow &y = (\alpha J_A - \mathrm{Id}) \circ (\alpha J_B - \mathrm{Id}) y \\
\Leftrightarrow &y = R_A^\alpha(R_B^\alpha y).
\end{aligned}$$

Note that $x = J_B y$ and $x \in \operatorname{zer}(A + B + (2-\alpha)\,\mathrm{Id})$. Consequently, we have

$$J_B(\operatorname{Fix} T) \in \operatorname{zer}(A + B + (2-\alpha)\,\mathrm{Id}).$$

Since $2 - \alpha > 0$, $A + B + (2-\alpha)\,\mathrm{Id}$ is strictly monotone, we obtain that

$$\operatorname{zer}(A + B + (2-\alpha)\,\mathrm{Id})$$

is a singleton by using Fact 1.2.27. Hence,

$$J_B(\operatorname{Fix} T) = \operatorname{zer}(A + B + (2-\alpha)\,\mathrm{Id}).$$

(3) In the proof of (2), we have $y \in \operatorname{Fix} T$, so $\operatorname{Fix} T \neq \emptyset$.

□

The following result is well-known.

Lemma 3.2. *Consider the Douglas-Rachford algorithm with $\lambda_n = 1$ for all n and $\gamma = 1$. Then* (DR) *becomes*

$$\begin{cases} y_n = J_B x_n, \\ z_n = J_A(2y_n - x_n), \\ x_{n+1} = x_n + (z_n - y_n). \end{cases}$$

This algorithm can also be written as

$$x_{n+1} = x_n + \frac{1}{2}(R_A \circ R_B x_n - x_n),$$

in terms of

$$x_{n+1} = D_{A,B}(x_n), \tag{3.1}$$

where

$$D_{A,B} = \frac{\mathrm{Id} + R_A \circ R_B}{2} = \frac{1}{2}\,\mathrm{Id} + \frac{1}{2} R_A \circ R_B.$$

Proof.

$$\begin{aligned} x_{n+1} &= x_n + (z_n - y_n) \\ &= x_n + J_A(2J_Bx_n - x_n) - J_Bx_n \\ &= \frac{x_n + (x_n - 2J_Bx_n) + 2J_A(2J_Bx_n - x_n)}{2} \\ &= \frac{x_n - R_Bx_n + 2J_A(R_Bx_n)}{2} \\ &= \frac{x_n + R_A \circ R_Bx_n}{2} \\ &= \frac{\operatorname{Id} + R_A \circ R_B}{2} x_n \end{aligned}$$

That is, $x_{n+1} = x_n + \frac{1}{2}(R_A \circ R_Bx_n - x_n)$. Hence, (3.1) holds. $\square$

We now introduce α-Douglas-Rachford algorithm.

Lemma 3.3. *Changing the parameter 2 of the algorithm* (DR) *into α, where $\alpha \in [1,2)$, we propose the α-DR algorithm*

$$(\alpha\text{-DR}) \qquad \begin{cases} y_n = J_{\gamma B}x_n \\ z_n = J_{\gamma A}(\alpha y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n). \end{cases}$$

If we keep the assumption that $\lambda_n = 1$ for all n and $\gamma = 1$, then following holds:

(1) (α-DR) *can also be written as*

$$x_{n+1} = x_n + \frac{1}{\alpha}(R_A^\alpha \circ R_B^\alpha x_n - x_n),$$

in terms of

$$x_{n+1} = D_{A,B}^\alpha(x_n), \tag{3.2}$$

where

$$D_{A,B}^\alpha = (1 - \frac{1}{\alpha})\operatorname{Id} + \frac{1}{\alpha}R_A^\alpha \circ R_B^\alpha.$$

(2) $D_{A,B}^\alpha$ *is an averaged mapping.*

Proof. (1) According to the definition of the α-DR algorithm,

$$\begin{aligned} x_{n+1} &= x_n + (J_A(\alpha y_n - x_n) - y_n) \\ &= x_n + (J_A(\alpha J_B x_n - x_n) - J_B x_n) \\ &= \frac{\alpha x_n - \alpha J_B x_n + \alpha J_A(\alpha J_B x_n - x_n)}{\alpha} \\ &= \frac{(\alpha - 1)x_n + x_n - \alpha J_B x_n + \alpha J_A(\alpha J_B x_n - x_n)}{\alpha} \\ &= \frac{(\alpha - 1)x_n + (\alpha J_A - \mathrm{Id}) \circ (\alpha J_B - \mathrm{Id})x_n}{\alpha} \\ &= \frac{(\alpha - 1)x_n + R_A^\alpha \circ R_B^\alpha x_n}{\alpha} \\ &= (1 - \frac{1}{\alpha})x_n + \frac{1}{\alpha} R_A^\alpha R_B^\alpha x_n. \end{aligned}$$

It follows that

$$x_{n+1} = \left[(1 - \frac{1}{\alpha})\,\mathrm{Id} + \frac{1}{\alpha} R_A^\alpha \circ R_B^\alpha\right] x_n.$$

So (3.2) holds.

(2) Because $R_A^\alpha \circ R_B^\alpha$ is nonexpansive, as $1 \leq \alpha < 2$, $D_{A,B}^\alpha$ is an averaged operator.

□

Remark 3.4. [4, Remark 4.34] Let D be a nonempty subset of $\mathcal{H}$, let $T : D \to \mathcal{H}$.

(1) If T is averaged, then it is nonexpansive.

(2) If T is nonexpansive, it is not necessarily averaged: consider $T = -\mathrm{Id} : \mathcal{H} \to \mathcal{H}$ when $\mathcal{H} \neq \{0\}$.

(3) T is firmly nonexpansive if and only if it is 1/2-averaged.

Theorem 3.2.5. Let $\alpha \in [1, 2)$, let A and B be maximally monotone operators from $\mathcal{K}$ to $2^{\mathcal{K}}$ with $0 \in \mathrm{int}(\mathrm{dom}\, A - \mathrm{dom}\, B)$. Let $\lambda_n = 1$ for all n, let $\gamma = 1$, and let $x_0 \in \mathbb{R}^m$. Set

$$\begin{cases} y_n = J_B x_n \\ z_n = J_A(\alpha y_n - x_n) \\ x_{n+1} = x_n + (z_n - y_n). \end{cases} \tag{3.3}$$

Then there exists $x \in \mathrm{Fix}\, R_A^\alpha \circ R_B^\alpha$ such that the following hold:

(1) $J_B x = \mathrm{zer}(A + B + (2 - \alpha)\,\mathrm{Id})$. Moreover, the answer is unique.

(2) $(y_n - z_n)_{n=1}^{+\infty}$ converges to 0.

(3) $(x_n)_{n=1}^{+\infty}$ converges to x.

(4) $(y_n)_{n=1}^{+\infty}$ converges to $J_B x$.

(5) $(z_n)_{n=1}^{+\infty}$ converges to $J_B x$.

Proof. (1) Let $T = R_A^\alpha \circ R_B^\alpha$. According to Theorem 3.2.4, $\operatorname{Fix} T \neq \emptyset$. Then for any $x \in \operatorname{Fix} T$, we have $x = R_A^\alpha(R_B^\alpha x)$, and this together with $R_A^\alpha = \alpha J_A - \mathrm{Id}$ and $R_B^\alpha = \alpha J_B - \mathrm{Id}$, yields that

$$(\alpha - 1)J_B x - x \in AJ_B x,$$

that is,

$$J_B x - x \in AJ_B x + (2-\alpha)J_B x.$$

Thus,

$$0 \in AJ_B x + (2-\alpha)J_B x + (x - J_B x). \tag{3.4}$$

By the definition of the resolvent, we have

$$x \in (B + \mathrm{Id})J_B x,$$

and so,

$$x - J_B x \in B \circ J_B x.$$

Combining with (3.4), one has

$$0 \in AJ_B x + (2-\alpha)J_B x + B \circ J_B x,$$

that is

$$0 \in [A + B + (2-\alpha)\,\mathrm{Id}] \circ J_B x.$$

Consequently, we have

$$J_B x \in \operatorname{zer}(A + B + (2-\alpha)\,\mathrm{Id}).$$

Since $2 - \alpha > 0$, $A + B + (2-\alpha)\,\mathrm{Id}$ is strictly monotone, we obtain that

$$\operatorname{zer}(A + B + (2-\alpha)\,\mathrm{Id})$$

is a singleton by using Fact 1.2.27. Therefore

$$J_B x = \operatorname{zer}(A + B + (2-\alpha)\,\mathrm{Id}).$$

(2) From (3.3), it follows that

$$\begin{aligned}
z_n - y_n &= J_A(\alpha y_n - x_n) - J_B x_n \\
&= J_A(\alpha J_B x_n - x_n) - J_B x_n \\
&= \frac{\alpha J_A(\alpha J_B x_n - x_n) - \alpha J_B x_n}{\alpha} \\
&= \frac{\alpha J_A(\alpha J_B x_n - x_n) - (\alpha J_B x_n - x_n) - x_n}{\alpha} \\
&= \frac{\alpha J_A(R_B^\alpha x_n) - R_B^\alpha x_n - x_n}{\alpha} \\
&= \frac{R_A^\alpha \circ R_B^\alpha x_n - x_n}{\alpha} \\
&= \frac{1}{\alpha}(T x_n - x_n).
\end{aligned}$$

Thus, $x_{n+1} = x_n + \frac{1}{\alpha}(Tx_n - x_n)$. By Fact 1.2.22(1), we have $Tx_n - x_n \to 0$. Therefore, $z_n - y_n \to 0$.

(3) Since $1 \leq \alpha < 2$, we can apply Fact 1.2.22(2) to complete this proof.

(4) It follows from (3.3) and $J_B = (\mathrm{Id} + B)^{-1}$ that $x_n \in (B + \mathrm{Id})y_n$, i.e., $x_n - y_n \in By_n$. Hence, $(y_n, x_n - y_n) \in \operatorname{gra} B$.

Similarly, $(z_n, \alpha y_n - x_n - z_n) \in \operatorname{gra} A$. Set

$$\forall n \in \mathbb{N}, \quad v_n := x_n - y_n, \quad w_n := \alpha y_n - x_n - z_n. \tag{3.5}$$

Then we have

$$\begin{cases} (z_n, w_n) \in \operatorname{gra} A \\ (y_n, v_n) \in \operatorname{gra} B \\ w_n + v_n = (\alpha - 1)y_n - z_n \end{cases} \tag{3.6}$$

Since result (3) tells us that $(x_n)_{n=1}^{+\infty}$ converges to x, $(x_n)_{n=1}^{+\infty}$ is bounded. According to the algorithm (3.3), $y_n = J_B x_n$. By Fact 1.2.40 (1), J_A and J_B are firmly nonexpansive. Thus,

$$\forall n \in \mathbb{N} \quad \|y_n - y_0\| = \|J_B x_n - J_B x_0\| \leq \|x_n - x_0\|,$$

which implies $(y_n)_{n=1}^{+\infty}$ is bounded.
Then there exists a subsequence $(y_{n_k})_{k=1}^{+\infty}$ of $(y_n)_{n=1}^{+\infty}$ such that $(y_{n_k})_{k=1}^{+\infty}$ is a convergent sequence. Suppose $y_{n_k} \to \bar{y}$, as $z_n - y_n \to 0$ has been proved, there exists a corresponding subsequence $(z_{n_k})_{k=1}^{+\infty}$ of $(z_n)_{n=1}^{+\infty}$ such that $z_{k_n} \to \bar{y}$. By (3.5), we have $v_{n_k} \to x - \bar{y}$ and $w_{n_k} \to (\alpha - 1)\bar{y} - x$.

Let's set $A_1 := A + (2 - \alpha)\,\mathrm{Id}, A_2 := B$. It follows from Fact 1.2.28 that A_1 and A_2 are maximally monotone. Since

$$(z_n, w_n) \in \operatorname{gra} A \quad \text{and} \quad (y_n, v_n) \in \operatorname{gra} B$$

as showed in (3.6), we have

$$\begin{cases} (z_{n_k}, w_{n_k} + (2-\alpha)z_{n_k}) \in \operatorname{gra} A_1 \\ (y_{n_k}, v_{n_k}) \in \operatorname{gra} A_2, \end{cases}$$

with $z_{n_k} \to \bar{y}, y_{n_k} \to \bar{y}, w_{n_k} + (2-\alpha)z_{n_k} \to \bar{y} - x, v_{n_k} \to x - \bar{y}, z_{n_k} - y_{n_k} \to 0, y_{n_k} - z_{n_k} \to 0$. Again, from (3.5), we have

$$\begin{aligned} w_{n_k} + (2-\alpha)z_{n_k} + v_{n_k} &= \alpha y_{n_k} - x_{n_k} - z_{n_k} + (2-\alpha)z_{n_k} + x_{n_k} - y_{n_k} \\ &= (\alpha - 1)(y_{n_k} - z_{n_k}) \to 0 \end{aligned}$$

This combining with Fact 3.2.2, yields that there exists $a \in \operatorname{zer}(A_1 + A_2)$ such that

$$a = \bar{y} \text{ and } \begin{cases} (a, \bar{y} - x) \in \operatorname{gra} A_1, \\ (a, x - \bar{y}) \in \operatorname{gra} A_2. \end{cases}$$

That is,

$$\begin{cases} (\bar{y}, \bar{y} - x) \in \operatorname{gra}(A + (2-\alpha)\,\mathrm{Id}), \\ (\bar{y}, x - \bar{y}) \in \operatorname{gra} B. \end{cases}$$

In view of $(\bar{y}, \bar{y} - x) \in \operatorname{gra}(A + (2-\alpha)\,\mathrm{Id})$, one has

$$\bar{y} - x \in [A + (2-\alpha)]\bar{y}.$$

which is equivalent to

$$(\alpha - 1)\bar{y} - x \in A\bar{y}.$$

In view of $(\bar{y}, x - \bar{y}) \in \operatorname{gra} B$, one has

$$\bar{y} = J_B x.$$

Together, we obtain

$$\bar{y} = J_B x \quad \text{and} \quad \bar{y} \in \operatorname{dom} A.$$

Thus, $J_B x$ is a sequential cluster point of $(y_n)_{n=1}^{+\infty}$. Because $J_B x$ emerges as the limit of every convergent subsequence for $(y_n)_{n=1}^{+\infty}$, we conclude that $J_B x$ is the unique sequential cluster point. Since y_n is bounded and has only one sequential cluster point, by Fact 1.1.19, $y_n \to J_B x$.

(5) Combining result (2) and result (4), we have $z_n = (z_n - y_n) + y_n \to 0 + J_B x$, i.e., $z_n \to J_B x$.

□

Remark 3.5. The proof of (4) works in a Hilbert space by replacing strong convergence with weak convergence in appropriate places. In particular, we have $y_n \rightharpoonup J_B x$.

Remark 3.6. In $\mathbb{R}^m$, the proof of (4) is simple by using that J_B is Lipschitz continuous:

$$\lim_{n \to +\infty} y_n = \lim_{n \to +\infty} J_B(x_n) = J_B x.$$

Remark 3.7. The application of the α-Douglas-Rachford algorithm is wider than the classic Douglas-Rachford algorithm since there is a requirement on the classic one that $\operatorname{zer}(A + B) \neq \emptyset$, while the α-Douglas-Rachford algorithm does not need a strict condition like that. Because according to Theorem 3.2.1, once $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$, $\operatorname{zer}[A + B + (2 - \alpha) \operatorname{Id}] \neq \emptyset$ for sure.

3.2.1 Application to composited monotone inclusion problems

Definition 3.2.6. Let $B : \mathcal{H} \to 2^{\mathcal{H}}$, $D : \mathcal{H} \to 2^{\mathcal{H}}$. Define $B \overset{\beta}{\square} D = (B^{-1} + D^{-1} + \beta \operatorname{Id})^{-1}$, where $\beta \in \mathbb{R}$.

Lemma 3.2.7. Let $B : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ and $D = \beta_2 \operatorname{Id}$ with $\beta_1, \beta_2 \in \mathbb{R}_{++}$. Then

$$(B \overset{\beta_1}{\square} D) = B \square \left(\frac{\beta_2}{1 + \beta_1 \beta_2} \operatorname{Id} \right).$$

Proof. Since $D = \beta_2 \operatorname{Id}$,

$$\begin{aligned} (B \overset{\beta_1}{\square} D) =& [B^{-1} + (\beta_2 \operatorname{Id})^{-1} + \beta_1 \operatorname{Id}]^{-1} \\ =& [B^{-1} + \frac{1 + \beta_1 \beta_2}{\beta_2} \operatorname{Id}]^{-1}. \end{aligned}$$

Therefore,

$$(B \overset{\beta_1}{\square} D) = B \square \left(\frac{\beta_2}{1 + \beta_1 \beta_2} \operatorname{Id} \right).$$

□

Lemma 3.2.8. Let $B : \mathbb{R}^m \to 2^{\mathbb{R}^m}$, $D = N_{\{0\}}$, and $\beta \in \mathbb{R}_{++}$. Then

$$(B \overset{\beta}{\square} D) = B \square \left(\frac{1}{\beta} \operatorname{Id} \right).$$

Proof. Since $D = N_{\{0\}}$,

$$(B \overset{\beta}{\square} D) = (B^{-1} + {N_{\{0\}}}^{-1} + \beta \operatorname{Id})^{-1}. \tag{3.7}$$

According to Lemma 1.1.35, ${N_{\{0\}}}^{-1} y = 0$, for any $y \in \mathbb{R}^m$. Thus,

$$(B^{-1} + {N_{\{0\}}}^{-1} + \beta \operatorname{Id})^{-1} = (B^{-1} + \beta \operatorname{Id})^{-1}$$

Therefore, (3.7) is equivalent to

$$(B \overset{\beta}{\square} N_{\{0\}}) = B \square \left(\frac{1}{\beta} \operatorname{Id} \right).$$

$\square$

Now suppose $A : \mathbb{R}^m \to 2^{\mathbb{R}^m}$, $B : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ and $D : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ are maximally monotone operators, and $L : \mathbb{R}^m \to \mathbb{R}^m$ is a nonzero linear invertible operator.

Also recall that $M, Q, S, V, \boldsymbol{A}$, and $\boldsymbol{B}$: Let $\mathcal{K} = \mathbb{R}^m \times \mathbb{R}^m$, $\tau, \sigma \in \mathbb{R}_{++}$, and $\tau\sigma \|L\|^2 < 4$.

(M) $\quad M : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (-z + Ax, r + B^{-1}v);$

(Q) $\quad Q : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (0, D^{-1}v);$

(S) $\quad S : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (L^\star v, -Lx);$

(V) $\quad V : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (\frac{x}{\tau} - \frac{1}{2} L^\star v, \frac{v}{\sigma} - \frac{1}{2} Lx);$

(A) $\quad \boldsymbol{A} := V^{-1}(\frac{1}{2} S + Q);$

(B) $\quad \boldsymbol{B} := V^{-1}(\frac{1}{2} S + M).$

Theorem 3.2.9. Suppose $M, Q, S, V, \boldsymbol{A}$, and $\boldsymbol{B}$ are constructed by (M), (Q), (S), (V), (A) and (B) respectively. Let $\alpha \in [1, 2)$. Then the following two inclusion problems are equivalent:

(1) Find $(x, v) \in \mathbb{R}^m \times \mathbb{R}^m$ such that $(x, v) \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B} + (2 - \alpha) \operatorname{Id})$.

(2) Solve the problem with primal inclusion: find $x \in \mathbb{R}^m$ such that

$$z \in Ax + \frac{2 - \alpha}{\tau} x + \frac{\alpha}{4 - \alpha} \boldsymbol{L}^\star \circ (B \overset{\frac{2-\alpha}{\sigma}}{\square} D) \circ (\boldsymbol{L}x - r) \tag{3.8}$$

where $\boldsymbol{L} = \frac{4-\alpha}{2}L$, $\tau \in \mathbb{R}_{++}$ and $\sigma \in \mathbb{R}_{++}$, together with the dual inclusion: find v such that there exists an $x \in \mathbb{R}^m$ that

$$\begin{cases} z - \frac{\alpha}{4-\alpha}\boldsymbol{L}^\star v \in Ax + \frac{(2-\alpha)}{\tau}x \\ v \in (B \overset{\frac{2-\alpha}{\sigma}}{\square} D) \circ (\boldsymbol{L}x - r). \end{cases} \tag{3.9}$$

Proof. By the definitions of $M, Q, S, V, \boldsymbol{A}$, and $\boldsymbol{B}$, and step 3 of the proof of Lemma 2.3.4, we have $\boldsymbol{B} := V^{-1}(\frac{1}{2}S + M)$ and $\boldsymbol{A} := V^{-1}(\frac{1}{2}S + Q)$. Note that

$$\begin{aligned} \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B} + (2-\alpha)\operatorname{Id}) &= \operatorname{zer}(V^{-1}(M + S + Q) + (2-\alpha)\operatorname{Id}) \\ &= \operatorname{zer}(V^{-1}(M + S + Q + (2-\alpha)V)) \\ &= \operatorname{zer}(M + S + Q + (2-\alpha)V)) \end{aligned}$$

For all $(x, v) \in \operatorname{zer}(M + S + Q + (2-\alpha)V))$, we have

$$\begin{aligned} (0,0) \in &(M + S + Q + (2-\alpha)V)(x, v) \\ =&(-z + Ax + L^\star v + (2-\alpha)\frac{x}{\tau} - \frac{2-\alpha}{2}L^\star v, \\ &r + B^{-1}v + D^{-1}v - Lx + (2-\alpha)\frac{v}{\sigma} - \frac{2-\alpha}{2}Lx). \end{aligned}$$

That is

$$\begin{aligned} &\begin{cases} 0 \in -z + Ax + L^\star v + (2-\alpha)\frac{x}{\tau} - \frac{2-\alpha}{2}L^\star v \\ 0 \in r + B^{-1}v + D^{-1}v - Lx + (2-\alpha)\frac{v}{\sigma} - \frac{2-\alpha}{2}Lx. \end{cases} \\ \Leftrightarrow &\begin{cases} 0 \in -z + Ax + (2-\alpha)\frac{x}{\tau} + \frac{\alpha}{2}L^\star v \\ \frac{4-\alpha}{2}Lx - r \in B^{-1}v + D^{-1}v + \frac{(2-\alpha)}{\sigma}v. \end{cases} \end{aligned}$$

According to Definition 3.2.6, $\frac{4-\alpha}{2}Lx - r \in B^{-1}v + D^{-1}v + \frac{(2-\alpha)}{\sigma}v$ can be written as

$$v \in (B \overset{\frac{2-\alpha}{\sigma}}{\square} D) \circ (\frac{4-\alpha}{2}Lx - r).$$

Thus, we have

$$\begin{cases} 0 \in -z + Ax + \frac{(2-\alpha)}{\tau}x + \frac{\alpha}{2}L^\star v \\ v \in (B \overset{\frac{2-\alpha}{\sigma}}{\square} D) \circ (\frac{4-\alpha}{2}Lx - r). \end{cases}$$

Since $\boldsymbol{L} = \frac{4-\alpha}{2}L$, $L^\star = \frac{2}{4-\alpha}\boldsymbol{L}^\star$. Then we have

$$z \in Ax + \frac{2-\alpha}{\tau}x + \frac{\alpha}{4-\alpha}\boldsymbol{L}^\star \circ (B \overset{\frac{2-\alpha}{\sigma}}{\square} D) \circ (\boldsymbol{L}x - r).$$

Therefore, for any $(x, v) \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B} + (2-\alpha)\operatorname{Id})$, x is also the solution of the primal inclusion problem (3.8), and v is the solution of the dual inclusion problem (3.9). $\square$

Theorem 3.2.10. Suppose $M, Q, S, V, \boldsymbol{A}$, and $\boldsymbol{B}$ are constructed by (M), (Q), (S), (V), (A) and (B) respectively. Let $\alpha \in [1, 2)$. The inclusion problem (3.8) together with inclusion problem (3.9) can be solved by using α-Douglas-Rachford algorithm if $\operatorname{dom} D^{-1} = \mathbb{R}^m$. In particular, $\operatorname{dom} D^{-1} = \mathbb{R}^m$ if one of the following holds:

(1) $D = N_{\{0\}}$.

(2) $D = \operatorname{Id}$.

Proof. Because $\operatorname{dom} D^{-1} = \mathbb{R}^m$, we get $\operatorname{dom} Q = \mathcal{K}$. Since $\operatorname{dom} S = \mathcal{K}$, we conclude that

$$\operatorname{dom}(\frac{1}{2}S + Q) = \mathcal{K}.$$

As Lemma 2.3.4 shows V is invertible (one-to-one, onto), one has

$$\begin{aligned}\operatorname{dom} \boldsymbol{A} &= \operatorname{dom} V^{-1}(\frac{1}{2}S + Q)\\ &= \operatorname{dom}(\frac{1}{2}S + Q)\\ &= \mathcal{K}.\end{aligned}$$

Then

$$\operatorname{dom} A - \operatorname{dom} B = \mathcal{K}.$$

Therefore, we get

$$0 \in \operatorname{int}(\operatorname{dom} \boldsymbol{A} - \operatorname{dom} \boldsymbol{B}).$$

As Lemma 2.3.4 also shows $\boldsymbol{A}$ and $\boldsymbol{B}$ are maximally monotone, by Theorem 3.2.9 and Theorem 3.2.5, the composited monotone inclusion problem (3.8) together with inclusion problem (3.9) can be solved by using α-Douglas-Rachford algorithm. $\square$

3.2.2 The application to proper, lower-semicontinuous convex functions

As $f, g, l \in \Gamma_0(\mathbb{R}^m)$, we consider the maximal monotone operators: $A = \partial f, B = \partial g, D = \partial l$. Thus the primal inclusion problem:

$$\text{find } x \in \mathbb{R} \text{ such that } z \in Ax + \frac{2-\alpha}{\tau}x + \frac{\alpha}{4-\alpha}\boldsymbol{L}^\star \circ (B \overset{\frac{2-\alpha}{\sigma}}{\square} D) \circ (\boldsymbol{L}x - r)$$

is equivalent to the following inclusion problem:

$$\text{find } x \in \mathbb{R} \text{ such that } z \in \partial f(x) + \frac{2-\alpha}{\tau}x + \frac{\alpha}{4-\alpha}\boldsymbol{L}^\star \circ (\partial g \overset{\frac{2-\alpha}{\sigma}}{\square} \partial l) \circ (\boldsymbol{L}x - r).$$

Theorem 3.2.11. Let $f, g, l \in \Gamma_0(\mathbb{R}^m)$, let $\boldsymbol{L}$ be a nonzero linear invertible operator, let $z, r \in \mathbb{R}^m$, let $\tau \in \mathbb{R}_{++}$ and $\sigma \in \mathbb{R}_{++}$, and $\alpha \in [1, 2)$. If $\operatorname{dom} g^* \cap \operatorname{intdom} l^* \neq \emptyset$, then the following primal inclusion problem

$$\text{Find } x \in \mathbb{R}^m \text{ such that } z \in \partial f(x) + \frac{2-\alpha}{\tau}x + \frac{\alpha}{4-\alpha}\boldsymbol{L}^\star \circ (\partial g \overset{\frac{2-\alpha}{\sigma}}{\square} \partial l) \circ (\boldsymbol{L}x - r) \tag{3.10}$$

can be characterized by the following optimization problem

$$\operatorname{Argmin}\{f(\cdot) + \frac{2-\alpha}{2\tau}\|\cdot\|^2 + \frac{\alpha}{4-\alpha}[(g \,\square\, l) \,\square\, (\frac{\sigma}{2(2-\alpha)}\|\cdot\|^2)](\boldsymbol{L}\cdot - r) - \langle z, \cdot\rangle\}. \tag{3.11}$$

Proof. According to Definition 3.2.6,

$$\partial g \overset{\frac{2-\alpha}{\sigma}}{\square} \partial l = [(\partial g)^{-1} + (\partial l)^{-1} + \frac{2-\alpha}{\sigma}\operatorname{Id}]^{-1}.$$

Since $g, l \in \Gamma_0(\mathbb{R}^m)$, by Fact 1.3.46,

$$[(\partial g)^{-1} + (\partial l)^{-1} + \frac{2-\alpha}{\sigma}\operatorname{Id}]^{-1} = (\partial g^* + \partial l^* + \frac{2-\alpha}{\sigma}\operatorname{Id})^{-1}.$$

Because $\operatorname{dom} g^* \cap \operatorname{intdom} l^* \neq \emptyset$, we can use the sum rule for subdifferentials (Fact 1.3.31) to get

$$(\partial g^* + \partial l^* + \frac{2-\alpha}{\sigma}\operatorname{Id})^{-1} = (\partial(g^* + l^*) + \frac{2-\alpha}{\sigma}\operatorname{Id})^{-1}.$$

Again, since $g, l \in \Gamma_0(\mathbb{R}^m)$ and $\operatorname{dom} g^* \cap \operatorname{intdom} l^* \neq \emptyset$, Theorem 1.3.41 implies that $g^* + l^* \in \Gamma_0(\mathbb{R}^m)$. Therefore, we can use Fact 1.3.32 to get

$$(\partial(g^* + l^*) + \frac{2-\alpha}{\sigma}\operatorname{Id})^{-1} = [\partial(g^* + l^* + \frac{2-\alpha}{2\sigma}\|\cdot\|^2)]^{-1},$$

which is equivalent to

$$\partial(g^* + l^* + \frac{2-\alpha}{2\sigma}\|\cdot\|^2)^* \tag{3.12}$$

as $g^* + l^* + \frac{2-\alpha}{2\sigma}\|\cdot\|^2 \in \Gamma_0(\mathbb{R}^m)$. By using the same method, as $f \in \Gamma_0(\mathbb{R}^m)$, we have

$$\partial f + \frac{2-\alpha}{\tau}\operatorname{Id} = \partial(f + \frac{2-\alpha}{2\tau}\|\cdot\|^2). \tag{3.13}$$

Obviously, $\operatorname{dom} \|\cdot\|^2 = \mathbb{R}^m$ and $\|\cdot\|^2 \in \Gamma_0(\mathbb{R}^m)$. Thus, we have $0 \in \operatorname{int}[\operatorname{dom}(g^* + l^*) - \operatorname{dom} \frac{2-\alpha}{2\sigma}\|\cdot\|^2]$.
By Fact 1.3.50, equation (3.12) can be rewritten as $\partial[(g^* + l^*)^* \boxdot (\frac{2-\alpha}{2\sigma}\|\cdot\|^2)^*]$. Again, by using the Theorem 2.4.1 and Fact 1.3.50 with the reason that $\operatorname{dom} g^* \cap \operatorname{intdom} l^* \neq \emptyset$, the set in line (3.12) can be written as

$$\partial[(g^{**} \boxdot l^{**}) \boxdot (\frac{2-\alpha}{2\sigma}\|\cdot\|^2)^*].$$

Because $g, l \in \Gamma_0(\mathbb{R}^m)$, by Fact 1.3.40, $g^{**} = g$ and $l^{**} = l$. Moreover, as Example 1.3.38 says $(\frac{2-\alpha}{2\sigma}\|\cdot\|^2)^* = (\frac{\sigma}{2(2-\alpha)}\|\cdot\|^2)$, expression (3.12) finally equals

$$\partial[(g \boxdot l) \boxdot (\frac{\sigma}{2(2-\alpha)}\|\cdot\|^2)]. \tag{3.14}$$

Combining the result of (3.13) and (3.14), and moving z to the end of the right hand side, the inclusion problem showed by (3.10) becomes

$$0 \in \partial(f + \frac{2-\alpha}{2\sigma}\|\cdot\|^2) + \frac{\alpha}{4-\alpha}\boldsymbol{L}^\star \circ \partial[(g \boxdot l) \boxdot (\frac{\sigma}{2(2-\alpha)}\|\cdot\|^2)] \circ (\boldsymbol{L}x - r) + \partial\langle -z, x\rangle \tag{3.15}$$

Since $(g \boxdot l) \boxdot \frac{\sigma}{2(2-\alpha)}\|\cdot\|^2$ can be considered as a Moreau envelope with the function $g \boxdot l$ and $\lambda = \frac{2-\alpha}{\sigma}$. Thus, by using Fact 1.3.55, we have

$$\operatorname{dom}[(g \boxdot l) \boxdot \frac{\sigma}{2(2-\alpha)}\|\cdot\|^2] = \mathbb{R}^m.$$

For $f + \frac{2-\alpha}{2\tau}\|\cdot\|^2$, we have

$$\operatorname{dom}(f + \frac{2-\alpha}{2\tau}\|\cdot\|^2) = \operatorname{dom} f \neq \emptyset.$$

Since $\boldsymbol{L}$ is a nonempty linear operator and $r \in \mathbb{R}^m$, once $\operatorname{dom} f \neq \emptyset$,

$$[\boldsymbol{L}\operatorname{dom}(f + \frac{2-\alpha}{2\tau}\|\cdot\|^2)] - r \neq \emptyset. \tag{3.16}$$

Combining the result that $\operatorname{dom}[(g \boxdot l) \boxdot \frac{\sigma}{2(2-\alpha)}\|\cdot\|^2] = \mathbb{R}^m$ with (3.16), we have

$$[\boldsymbol{L}\operatorname{dom}(f + \frac{2-\alpha}{2\tau}\|\cdot\|^2) - r] \cap \operatorname{intdom}[(g \boxdot l) \boxdot \frac{\sigma}{2(2-\alpha)}\|\cdot\|^2] \neq \emptyset. \tag{3.17}$$

Due to Lemma 2.6

$$\begin{aligned}
&\partial(f + \frac{2-\alpha}{2\tau}\|\cdot\|^2) + \frac{\alpha}{4-\alpha}\boldsymbol{L}^\star \circ \partial[(g \boxdot l) \boxdot (\frac{\sigma}{2(2-\alpha)}\|\cdot\|^2)] \circ (\boldsymbol{L}\cdot - r) \\
=&\partial\{f + \frac{2-\alpha}{2\tau}\|\cdot\|^2 + \frac{\alpha}{4-\alpha}[(g \boxdot l) \boxdot (\frac{\sigma}{2(2-\alpha)}\|\cdot\|^2)] \circ (\boldsymbol{L}\cdot - r)\}
\end{aligned} \tag{3.18}$$

Combining with the fact that $\operatorname{intdom}\langle -z, x\rangle = \mathbb{R}^m$, we get

$$\begin{aligned}&\partial\{f + \frac{2-\alpha}{2\tau}\|\cdot\|^2 + \frac{\alpha}{4-\alpha}[(g \boxdot l) \boxdot (\frac{\sigma}{2(2-\alpha)}\|\cdot\|^2)] \circ (\boldsymbol{L}\cdot - r)\} + \partial\langle -z, x\rangle \\ =&\partial\{f + \frac{2-\alpha}{2\tau}\|\cdot\|^2 + \frac{\alpha}{4-\alpha}[(g \boxdot l) \boxdot (\frac{\sigma}{2(2-\alpha)}\|\cdot\|^2)] \circ (\boldsymbol{L}\cdot - r) + \langle -z, \cdot\rangle\}\end{aligned} \tag{3.19}$$

by Fact 1.3.31. Combining the result of (3.15), (3.18) and (3.19), we get

$$0 \in \partial\{f + \frac{2-\alpha}{2\tau}\|\cdot\|^2 + \frac{\alpha}{4-\alpha}[(g \boxdot l) \boxdot (\frac{\sigma}{2(2-\alpha)}\|\cdot\|^2)] \circ (\boldsymbol{L}\cdot - r) - \langle z, \cdot\rangle\}(x). \tag{3.20}$$

By Fact 1.3.34, the inclusion (3.20) is equivalent to the statement that x belongs to

$$\operatorname{Argmin}\{f(\cdot) + \frac{2-\alpha}{2\tau}\|\cdot\|^2 + \frac{\alpha}{4-\alpha}[(g \boxdot l) \boxdot (\frac{\sigma}{2(2-\alpha)}\|\cdot\|^2)] \circ (\boldsymbol{L}\cdot - r) - \langle z, \cdot\rangle\},$$

which is exactly the optimization problem (3.11). Thus, the primal inclusion problem (3.10) is equivalent to the optimization problem (3.11). □

Remark 3.8. Combining Theorem 3.2.11 and Theorem 3.2.10, the optimization problem (3.11) can be solved by using α-Douglas-Rachford algorithm if we have $\operatorname{dom}(\partial l)^{-1} = \mathbb{R}^m$. This holds when $l = \iota_{\{0\}}$ or $l = \frac{1}{2}\|\cdot\|^2$.

Chapter 4

Special cases of the composited monotone inclusion problems and the double regularization

4.1 Overview

In this chapter, we still let $A : \mathbb{R}^m \to 2^{\mathbb{R}^m}$, $B : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ and $D : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ be maximally monotone operators. Let z, $r \in \mathbb{R}^m$ and $\alpha \in [1, 2)$, let $L : \mathbb{R}^m \to \mathbb{R}^m$ be a nonzero linear invertible operator, let $\tau, \sigma \in \mathbb{R}_{++}$ and $\tau\sigma\|L\|^2 < 4$. Setting the α inclusion problem in the general case: find $x \in \mathbb{R}^m$ such that

$$z \in Ax + \frac{2-\alpha}{\tau}x + \frac{\alpha}{4-\alpha}\boldsymbol{L}^\star(B \overset{\frac{2-\alpha}{\sigma}}{\square} D)(\boldsymbol{L}x - r), \tag{4.1}$$

where $\boldsymbol{L} = \frac{4-\alpha}{2}L$. The general case of the α inclusion problem (4.1) is too complex to get some direct result. In this chapter, we will consider the α inclusion problem (4.1) in some special cases, and then get some particular results. In those special cases, operator D satisfies $\operatorname{dom} D^{-1} = \mathbb{R}^m$, that means, all of the special cases that will be showed below can be solved by using α-Douglas-Rachford algorithm.

4.2 The special cases

Theorem 4.2.1. Let $a \in \mathbb{R}_{++}, \boldsymbol{L} = a\,\mathrm{Id}, D = \mathrm{Id}, z = 0, r = 0$. Then

(1) The α inclusion problem (4.1) becomes: find $x \in \mathbb{R}^m$ such that

$$0 \in Ax + \frac{2-\alpha}{\tau}x + \frac{\alpha a}{4-\alpha}[B\square(\frac{\sigma}{2-\alpha+\sigma}\,\mathrm{Id})](ax). \tag{4.2}$$

(2) If in addition, $A = \partial f, B = \partial g$, where $f, g \in \Gamma_0(\mathbb{R}^m)$, then the primal inclusion problem (4.2) can be characterized by the following optimization

problem

$$\text{Argmin}_{x\in\mathbb{R}^m}\{f(x)+\frac{2-\alpha}{2\tau}\|x\|^2+\frac{\alpha}{4-\alpha}(g\Box\frac{\gamma}{2}\|\cdot\|^2)(ax)\} \tag{4.3}$$

where $\gamma = \frac{\sigma}{2-\alpha+\sigma}$.

Proof. (1) Since $\boldsymbol{L} = a\,\text{Id}$, $z = 0, r = 0$, the α inclusion problem (4.1) becomes: find $x \in \mathbb{R}^m$ such that

$$0 \in Ax + \frac{2-\alpha}{\tau}x + \frac{\alpha a}{4-\alpha}(B \overset{\frac{2-\alpha}{\sigma}}{\Box} D)(ax), \tag{4.4}$$

Since $D = \text{Id}$, according to Lemma 3.2.7,

$$(B \overset{\frac{2-\alpha}{\sigma}}{\Box} D) = B\Box(\frac{\sigma}{2-\alpha+\sigma}\,\text{Id}).$$

Thus, (4.1) is equivalent to

$$0 \in Ax + \frac{2-\alpha}{\tau}x + \frac{\alpha a}{4-\alpha}[B\Box(\frac{\sigma}{2-\alpha+\sigma}\,\text{Id})](ax)$$

(2) Let $\gamma = \frac{\sigma}{2-\alpha+\sigma}$, since $\partial f = A, \partial g = B$, (4.2) is equivalent to

$$0 \in \partial f(x) + \frac{2-\alpha}{\tau}x + \frac{\alpha a}{4-\alpha}(\partial g\Box\gamma\,\text{Id})(ax). \tag{4.5}$$

Because $f \in \Gamma_0(\mathbb{R}^m)$, we have

$$\partial f(x) + \frac{2-\alpha}{\tau}x = \partial[f(x) + \frac{2-\alpha}{2\tau}\|x\|^2]$$

by Example 1.3.32. Moreover, since $\partial(\gamma f) = \gamma\partial f$ for any $\gamma \in \mathbb{R}_{++}$, combining this with the Example 1.3.24, we have

$$\gamma\,\text{Id} = \gamma\partial(\frac{1}{2}\|x\|^2) = \partial(\frac{\gamma}{2}\|x\|^2).$$

Therefore, (4.5) becomes

$$0 \in \partial[f(x) + \frac{2-\alpha}{2\tau}\|x\|^2] + \frac{\alpha a}{4-\alpha}[\partial g\Box\partial(\frac{\gamma}{2}\|\cdot\|^2)](ax). \tag{4.6}$$

As shown in Example 1.3.38, $(\frac{\gamma}{2}\|x\|^2)^* = \frac{\|x\|^2}{4\gamma/2} = \frac{\|x\|^2}{2\gamma}$. Because $g \in \Gamma_0(\mathbb{R}^m)$, by Fact 1.3.40, $\operatorname{dom} g^* \neq \emptyset$. Therefore,

$$\begin{aligned}\operatorname{dom} g^* \cap \operatorname{intdom}(\frac{\gamma}{2}\|\cdot\|^2)^* &= \operatorname{dom} g^* \cap \operatorname{intdom}(\frac{\|\cdot\|^2}{2\gamma})\\ &= \operatorname{dom} g^* \cap \mathbb{R}^m\\ &= \operatorname{dom} g^* \neq \emptyset.\end{aligned}$$

Combing the above result with the fact that $g \in \Gamma_0(\mathbb{R}^m)$, we get

$$\partial g \Box \partial(\frac{\gamma}{2}\|\cdot\|^2) = \partial(g \Box \frac{\gamma}{2}\|\cdot\|^2)$$

by Fact 2.4.3.
Again, we use the fact $\partial(\gamma f) = \gamma \partial f$, (4.6) is equivalent to

$$0 \in \partial[f(x) + \frac{2-\alpha}{2\tau}\|x\|^2] + a\partial[\frac{\alpha}{4-\alpha}(g \Box \frac{\gamma}{2}\|\cdot\|^2)](ax). \tag{4.7}$$

Because $(g \Box \frac{\gamma}{2}\|\cdot\|^2)$ can be considered as a Moreau envelope with the function g and $\lambda = \frac{1}{\gamma}$. Thus, by using Fact 1.3.55, we have

$$\operatorname{dom}(g \Box \frac{\gamma}{2}\|\cdot\|^2) = \mathbb{R}^m.$$

Thus,

$$\begin{aligned}&a \operatorname{dom}[f(\cdot) + \frac{2-\alpha}{2\tau}\|\cdot\|^2] \cap \operatorname{intdom}(g \Box \frac{\gamma}{2}\|\cdot\|^2)\\ =&a \operatorname{dom} f \cap \mathbb{R}^m\\ =&a \operatorname{dom} f \neq \emptyset.\end{aligned}$$

Thus, by Fact 1.3.29, (4.7) is equivalent to

$$0 \in \partial[f(x) + \frac{2-\alpha}{2\tau}\|x\|^2 + \frac{\alpha}{4-\alpha}(g \Box \frac{\gamma}{2}\|\cdot\|^2)(ax)]. \tag{4.8}$$

Due to Fact 1.3.34, finding x such that

$$0 \in \partial[f(x) + \frac{2-\alpha}{2\tau}\|x\|^2 + \frac{\alpha}{4-\alpha}(g \Box \frac{\gamma}{2}\|\cdot\|^2)(ax)]$$

is equivalent to x being a solution of

$$\operatorname{Argmin}_{x\in\mathbb{R}^m}\{f(x) + \frac{2-\alpha}{2\tau}\|x\|^2 + \frac{\alpha}{4-\alpha}(g \Box \frac{\gamma}{2}\|\cdot\|^2)(ax)\}$$

Therefore, we completed the proof.

□

Theorem 4.2.2. Let $a \in \mathbb{R}_{++}, \boldsymbol{L} = a\,\mathrm{Id}, D = N_{\{0\}}, z = 0, r = 0$. Then

(1) the α inclusion problem (4.1) becomes: find $x \in \mathbb{R}^m$ such that

$$0 \in Ax + \frac{2-\alpha}{\tau}x + \frac{\alpha a}{4-\alpha}(B\Box\frac{\sigma}{2-\alpha}\,\mathrm{Id})(ax). \tag{4.9}$$

(2) If in addition, $A = \partial f, B = \partial g$, where $f, g \in \Gamma_0(\mathbb{R}^m)$, then the primal inclusion problem (4.9) can be characterized by the following optimization problem

$$\mathrm{Argmin}_{x\in\mathbb{R}^m}\{f(x) + \frac{2-\alpha}{2\tau}\|x\|^2 + \frac{\alpha}{4-\alpha}(g\Box\frac{\gamma}{2}\|\cdot\|^2)(ax)\}, \tag{4.10}$$

where $\gamma = \frac{\sigma}{2-\alpha}$.

Proof. (1) Since $\boldsymbol{L} = a\,\mathrm{Id}$, $z = 0, r = 0$, the α inclusion problem (4.1) becomes: find $x \in \mathbb{R}^m$ such that

$$0 \in Ax + \frac{2-\alpha}{\tau}x + \frac{\alpha a}{4-\alpha}(B \overset{\frac{2-\alpha}{\sigma}}{\Box} D)(ax), \tag{4.11}$$

Since $D = N_{\{0\}}$, by Lemma 3.2.8

$$(B \overset{\frac{2-\alpha}{\sigma}}{\Box} N_{\{0\}}) = B\Box\frac{\sigma}{2-\alpha}\,\mathrm{Id}\,.$$

Thus, (4.11) is equivalent to

$$0 \in Ax + \frac{2-\alpha}{\tau}x + \frac{\alpha a}{4-\alpha}(B\Box\frac{\sigma}{2-\alpha}\,\mathrm{Id})(ax).$$

(2) Since the only difference between (4.2) and (4.9) is the parameter of the identity function. Thus, as we proved in 4.2.1(2), if we let $\gamma = \frac{\sigma}{2-\alpha}$ here, the primal inclusion problem (4.9) can be characterized by the following optimization problem

$$\mathrm{Argmin}_{x\in\mathbb{R}^m}\{f(x) + \frac{2-\alpha}{2\tau}\|x\|^2 + \frac{\alpha}{4-\alpha}(g\Box\frac{\gamma}{2}\|\cdot\|^2)(ax)\}.$$

$\square$

4.3 Double regularization: Moreau regularization and Tychonov regularization combined

Let $f, g : \mathbb{R}^m \to \mathbb{R}$ be proper, lower-semicontinuous and convex. Often we must find the minimizer of $f + g$. The functions f, g might not be differentiable. We can first make one of the function differentiable by using the Moreau envelope. In order to obtain the least norm solution, we can do Tychonov regularization. Therefore, we can consider minimization of the following function:

$$f + \beta_1 \,\mathrm{q} + \beta_2[g\Box(\beta_3 \,\mathrm{q})] \tag{4.12}$$

where $\beta_i > 0$ $(i \in \{1, 2, 3\})$, and $\mathrm{q} = \frac{1}{2}\|\cdot\|^2$. Put $h = f + \beta_2[g\Box(\beta_3 \,\mathrm{q})]$. While $g\Box(\beta_3 \,\mathrm{q})$ is the Moreau regularization with the function being g and $\lambda = \frac{1}{\beta_3}$, and $h + \beta_1 \,\mathrm{q}$ is the Tychonov regularization. Therefore, Problem (4.12) is a double regularization.

Remark 4.1. In (4.12), we usually require $\beta_1 \downarrow 0, \beta_2 \to 1$, and $\beta_3 \downarrow 0$.

4.3.1 Subdifferential of infimal convolutions

Let $\mathrm{q}(x) = \frac{1}{2}\|x\|^2$.

Corollary 4.3.1. Let $f : \mathbb{R}^m \to (-\infty, +\infty]$ be proper, lower semicontinuous and convex. Then for any $\beta_3 \in \mathbb{R}_{++}$,

$$\partial[f\Box(\beta_3 \,\mathrm{q})] = \partial f \Box \beta_3 \,\mathrm{Id} = [\frac{1}{\beta_3} \,\mathrm{Id} + (\partial f)^{-1}]^{-1}.$$

Proof. By Fact 1.3.39 and Fact 1.3.40, $f \in \Gamma_0(\mathbb{R}^m)$ implies $f^* \in \Gamma_0(\mathbb{R}^m)$; in particular, $\operatorname{dom} f^* \neq \emptyset$. According to Example 1.3.38, we have

$$\begin{aligned}
[\beta_3 \,\mathrm{q}(x)]^* &= \left(\frac{\beta_3}{2}\|x\|^2\right)^* \\
&= \frac{\|x\|^2}{2\beta_3}.
\end{aligned}$$

Therefore, $\operatorname{dom}(\beta_3 \,\mathrm{q})^* = \mathbb{R}^m = \operatorname{int} \operatorname{dom}(\beta_3 \,\mathrm{q})^*$. Thus, $\operatorname{dom} f^* \cap \operatorname{int} \operatorname{dom}(\beta_3 \,\mathrm{q})^* = \operatorname{dom} f^* \neq \emptyset$. Applying Fact 2.4.3, we get

$$\begin{aligned}
\partial(f\Box\beta_3 \,\mathrm{q}) &= \partial f \Box \partial(\beta_3 \,\mathrm{q}) \\
&= \partial f \Box \beta_3 \,\mathrm{Id} \\
&= [(\partial f)^{-1} + (\beta_3 \,\mathrm{Id})^{-1}]^{-1} \\
&= [(\partial f)^{-1} + \frac{1}{\beta_3} \,\mathrm{Id}]^{-1}.
\end{aligned}$$

□

4.3.2 Main results

Theorem 4.2. *Let $f, g : \mathbb{R}^m \to (-\infty, +\infty]$ be proper, lower semicontinuous and convex, let $\mathrm{q}(x) = \frac{1}{2}\|x\|^2$. For every $\beta_1 > 0, \beta_2 > 0, \beta_3 > 0$, consider*

$$\min_{x\in\mathbb{R}^m} \{f + \beta_1 \mathrm{q} + \beta_2[g\Box(\beta_3 \mathrm{q})]\}. \tag{p}$$

Then the following hold:

(1)
$$\partial\{f + \beta_1 \mathrm{q} + \beta_2[g\Box(\beta_3 \mathrm{q})]\} = \partial f + \beta_1 \mathrm{Id} + \beta_2[\partial g\Box(\beta_3 \mathrm{Id})].$$

(2) (p) *always has a unique solution.*

(3) The Fenchel Dual of (p) *is*

$$\min_{v\in\mathbb{R}^m} \left\{\left(f^* \boxdot \frac{\mathrm{q}}{\beta_1}\right)(v) + \frac{1}{\beta_2\beta_3}\mathrm{q}(v) + \beta_2 g^*\left(\frac{-v}{\beta_2}\right)\right\}, \tag{d}$$

and it also has a unique solution.

Proof. (1) Because f and g are proper, $\operatorname{dom} f \neq \emptyset$ and $\operatorname{dom} g \neq \emptyset$. Since $g\Box(\beta_3 \mathrm{q})$ is actually a Moreau envelope with the function g and $\lambda = \frac{1}{\beta_3}$, according to Fact 1.3.55,

$$\begin{aligned}\operatorname{intdom}[g\Box(\beta_3 \mathrm{q})] &= \operatorname{int}\mathbb{R}^m \\ &= \mathbb{R}^m.\end{aligned} \tag{4.13}$$

For $\beta_1 \mathrm{q}$, we have

$$\operatorname{intdom}(\beta_1 \mathrm{q}) = \operatorname{int}\mathbb{R}^m = \mathbb{R}^m.$$

Thus,

$$\begin{aligned}&\operatorname{dom} f \cap \operatorname{intdom} \beta_2[g\Box(\beta_3 \mathrm{q})] \cap \operatorname{intdom}(\beta_1 \mathrm{q}) \\ =& \operatorname{dom} f \cap \mathbb{R}^m \cap \mathbb{R}^m \\ =& \operatorname{dom} f \\ \neq& \emptyset.\end{aligned}$$

Additionally, $f \in \Gamma_0(\mathbb{R}^m), g \in \Gamma_0(\mathbb{R}^m)$. Thus, by Fact 1.3.31,

$$\partial\{f + \beta_1 \mathrm{q} + \beta_2[g\Box(\beta_3 \mathrm{q})]\} = \partial f + \partial(\beta_1 \mathrm{q}) + \partial[\beta_2(g\Box(\beta_3 \mathrm{q}))]. \tag{4.14}$$

Again, since $g \in \Gamma_0(\mathbb{R}^m)$, by using Corollary 4.3.1, we have $\partial[g\Box(\beta_3 \mathrm{q})] = \partial g\Box(\beta_3 \mathrm{Id})$. Therefore, (4.14) yields

$$\partial\{f + \beta_2[g\Box(\beta_3 \mathrm{q})] + \beta_1 \mathrm{q}\} = \partial f + \beta_1 \mathrm{Id} + \beta_2[\partial g\Box(\beta_3 \mathrm{Id})].$$

(2) Because $g \in \Gamma_0(\mathbb{R}^m)$, by Fact 1.3.40, $\operatorname{dom} g^* \neq \emptyset$. As noted in the proof of Corollary 4.3.1, $\operatorname{dom}(\beta_3\,\mathrm{q})^* = \mathbb{R}^m$. Consequently,

$$\operatorname{dom} g^* \cap \operatorname{int}\operatorname{dom}(\beta_3\,\mathrm{q})^* = \operatorname{dom} g^* \neq \emptyset.$$

Thus, by Fact 2.4.3(2),

$$[g\Box(\beta_3\,\mathrm{q})] = [g \boxdot (\beta_3\,\mathrm{q})] \in \Gamma_0(\mathbb{R}^m).$$

Because $f \in \Gamma_0(\mathbb{R}^m)$, $\operatorname{dom} f \neq \emptyset$. According to (4.13),

$$\operatorname{dom} f \cap \operatorname{dom}[\beta_2 g\Box(\beta_3\,\mathrm{q})] = \operatorname{dom} f \neq \emptyset.$$

Thus, by using Lemma 1.4, we have

$$f + \beta_2[g\Box(\beta_3\,\mathrm{q})] \in \Gamma_0(\mathbb{R}^m).$$

Since $\operatorname{dom} f \cap \operatorname{dom}\{\beta_2[g\Box(\beta_3\,\mathrm{q})]\} = \operatorname{dom} f$,

$$\begin{aligned} &\operatorname{dom}(\beta_1\,\mathrm{q}) \cap \{\operatorname{dom} f \cap \operatorname{dom}[\beta_2(g\Box(\beta_3\,\mathrm{q}))]\} \\ =&\mathbb{R}^m \cap \operatorname{dom} f \\ =&\operatorname{dom} f \\ \neq&\emptyset. \end{aligned}$$

In addition, according to Example 1.3.17 and Example 1.3.7, q is supercoercive and strictly convex. By Fact 1.3.19, $\beta_1\,\mathrm{q} + f + \beta_2[g\Box(\beta_3\,\mathrm{q})]$ has exactly one minimizer over $\mathbb{R}^m$. In another words, (p) always has a unique solution.

(3) According to Fenchel duality (see Definition 1.3.58), we have

$$\begin{aligned} &\min_{x\in\mathbb{R}^m}\{f(x) + \beta_1\,\mathrm{q}(x) + \beta_2[g\Box(\beta_3\,\mathrm{q})](x)\} \\ =&\min_{v\in\mathbb{R}^m}\{(f+\beta_1\,\mathrm{q})^*(v) + [\beta_2(g\Box(\beta_3\,\mathrm{q}))]^*(-v)\}. \end{aligned} \tag{4.15}$$

Since $\operatorname{dom}\mathrm{q} = \mathbb{R}^m$ and $f \in \Gamma_0(\mathbb{R}^m)$, we have

$$\operatorname{dom} f \cap \operatorname{intdom}\mathrm{q} = \operatorname{dom} f \neq \emptyset.$$

Thus, by Theorem 2.4.1 and Fact 1.3.50,

$$(f + \beta_1\,\mathrm{q})^* = f^* \boxdot (\beta_1\,\mathrm{q})^*. \tag{4.16}$$

As we showed in Example 1.3.38, for $\lambda > 0$, $(\lambda\|x\|^2)^* = \frac{\|x\|^2}{4\lambda}$. Thus,

$$(\beta_1\,\mathrm{q})^* = (\frac{\beta_1}{2}\|x\|^2)^* = \frac{\|x\|^2}{2\beta_1} = \frac{\mathrm{q}}{\beta_1}. \tag{4.17}$$

That is,

$$f^* \boxdot (\beta_1\,\mathrm{q})^* = f^* \boxdot \frac{\mathrm{q}}{\beta_1}. \tag{4.18}$$

For $[\beta_2(g\Box(\beta_3\,\mathrm{q}))]^*(-v)$, according to Fact 1.3.37, we have

$$[\beta_2(g\Box(\beta_3\,\mathrm{q}))]^*(-v) = \beta_2[g\Box(\beta_3\,\mathrm{q})]^*\left(\frac{-v}{\beta_2}\right).$$

Due to Fact 1.3.49 and equation (4.17)

$$\begin{aligned}[g\Box(\beta_3\,\mathrm{q})]^*\left(\frac{-v}{\beta_2}\right) &= [g^* + (\beta_3\,\mathrm{q})^*]\left(\frac{-v}{\beta_2}\right)\\ &= g^*\left(\frac{-v}{\beta_2}\right) + \frac{1}{\beta_3}\,\mathrm{q}\left(\frac{-v}{\beta_2}\right)\\ &= g^*\left(\frac{-v}{\beta_2}\right) + \frac{1}{\beta_2^2\beta_3}\,\mathrm{q}(v). \end{aligned} \tag{4.19}$$

Combining the result of (4.18) and (4.19), we have

$$\begin{aligned}&\min_{v\in\mathbb{R}^m}\{(f+\beta_1\,\mathrm{q})^*(v) + [\beta_2(g\Box(\beta_3\,\mathrm{q}))]^*(-v)\}\\ =&\min_{v\in\mathbb{R}^m}\left\{\left(f^*\boxdot\frac{\mathrm{q}}{\beta_1}\right)(v) + \frac{1}{\beta_2\beta_3}\,\mathrm{q}(v) + \beta_2 g^*\left(\frac{-v}{\beta_2}\right)\right\}.\end{aligned}$$

Therefore, we complete the proof that the Fenchel Dual of (p) is (d). For the similar reason as (2), (d) has a unique solution.

□

Remark 4.3. We are interested in considering: $\beta_1 = \frac{2-\alpha}{\tau}$ and $\beta_2 = \frac{\alpha}{4-\alpha}$, where $\tau \in \mathbb{R}_{++}$ and $\alpha \in [1,2)$.

If we apply Theorem 4.2 to the optimization problem (4.3), we can get the following corollary.

Corollary 4.3.2. Let $f, g \in \Gamma_0(\mathbb{R}^m), let\, \mathrm{q} = \frac{1}{2}\|\cdot\|^2$. Consider the optimization problem (4.3) where $a = 1, \gamma = \frac{\sigma}{2-\alpha+\sigma}$. Then the following hold:

(1) For any $x \in \mathbb{R}^m$,

$$\begin{aligned}&\partial\{f(x) + \frac{2-\alpha}{2\tau}\|x\|^2 + \frac{\alpha}{4-\alpha}(g\Box\frac{\gamma}{2}\|\cdot\|^2)(x)\}\\ =&\partial f(x) + \frac{2-\alpha}{\tau}x + \frac{\alpha}{4-\alpha}[\partial g(x)\Box(\gamma x)].\end{aligned}$$

(2) The following problem, where $\gamma = \frac{\sigma}{2-\alpha+\sigma}$, always has a unique solution:

$$\min_{x\in\mathbb{R}^m}\{f(x)+\frac{2-\alpha}{2\tau}\|x\|^2+\frac{\alpha}{4-\alpha}(g\Box\frac{\gamma}{2}\|\cdot\|^2)(x)\}. \tag{4.20}$$

(3) The Fenchel Dual of (4.20) is

$$\min_{v\in\mathbb{R}^m}\left\{\left[f^*\boxdot\left(\frac{\tau}{2-\alpha}\,\mathrm{q}\right)\right](v)+\frac{4-\alpha}{\alpha\gamma}\,\mathrm{q}(v)+\frac{\alpha}{4-\alpha}g^*\left(\frac{\alpha-4}{\alpha}v\right)\right\}.$$

Moreover, this Fenchel Dual has a unique solution.

If we apply Theorem 4.2 to the optimization problem (4.10), we can get the following corollary.

Corollary 4.3.3. Let $f, g \in \Gamma_0(\mathbb{R}^m), let\,\mathrm{q} = \frac{1}{2}\|\cdot\|^2$. Consider the optimization problem (4.10) where $a = 1, \gamma = \frac{\sigma}{2-\alpha}$. Then the following hold:

(1) For any $x \in \mathbb{R}^m$,

$$\begin{aligned}&\partial\{f(x)+\frac{2-\alpha}{2\tau}\|x\|^2+\frac{\alpha}{4-\alpha}(g\Box\frac{\gamma}{2}\|\cdot\|^2)(x)\}\\=&\partial f(x)+\frac{2-\alpha}{\tau}x+\frac{\alpha}{4-\alpha}[\partial g(x)\Box(\gamma x)].\end{aligned}$$

(2) The following problem, where $\gamma = \frac{\sigma}{2-\alpha}$, always has a unique solution:

$$\min_{x\in\mathbb{R}^m}\{f(x)+\frac{2-\alpha}{2\tau}\|x\|^2+\frac{\alpha}{4-\alpha}(g\Box\frac{\gamma}{2}\|\cdot\|^2)(x)\}. \tag{4.21}$$

(3) The Fenchel Dual of (4.21) is

$$\min_{v\in\mathbb{R}^m}\left\{\left[f^*\boxdot\left(\frac{\tau}{2-\alpha}\,\mathrm{q}\right)\right](v)+\frac{4-\alpha}{\alpha\gamma}\,\mathrm{q}(v)+\frac{\alpha}{4-\alpha}g^*\left(\frac{\alpha-4}{\alpha}v\right)\right\}.$$

Moreover, this Fenchel Dual has a unique solution.

Chapter 5

The α-Douglas-Rachford algorithm with $\alpha \to 2$

5.1 Overview

As we set $\alpha \in [1, 2)$ in the α-Douglas-Rachford algorithm, we want to consider the properties of that algorithm when α is in its limit case. Therefore, the α-Douglas-Rachford algorithm is considered in a special-limit case in this chapter.

5.2 Parameter $\alpha \to 2$

Fact 5.2.1. [4, Theorem 23.44] Let $A : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximally monotone operator and let $x \in \mathcal{H}$. Then the inclusions

$$(\forall \gamma \in (0, 1)) \quad 0 \in Ax_\gamma + \gamma(x_\gamma - x)$$

define a unique curve $(x_\gamma)_{\gamma \in (0,1)}$. Moreover, exactly one of the following holds:

(1) $\operatorname{zer} A \neq \emptyset$ and $x_\gamma \to P_{\operatorname{zer} A} x$ as $\gamma \downarrow 0$.

(2) $\operatorname{zer} A = \emptyset$ and $\|x_\gamma\| \to +\infty$ as $\gamma \downarrow 0$.

Fact 5.2.1 and Theorem 3.2.5 yield the following characterizations of the α-Douglas-Rachford algorithm. However, before we go to the theorem, we need to get the following lemma.

Lemma 5.2.2. Let A be a maximally monotone operator from $\mathcal{K}$ to $2^{\mathcal{K}}$, $(\alpha_k)_{k=1}^{+\infty}$ be an increasing sequence in $[1, 2)$ such that $\lim\limits_{k \to +\infty} \alpha_k = 2$. Then,

$$\lim_{k \to +\infty} R_A^{\alpha_k} = R_A^2.$$

Proof. As $R_A^{\alpha_k} = \alpha_k J_A - \operatorname{Id}$ and $R_A^2 = 2J_A - \operatorname{Id}$, we have for fixed x,

$$\|R_A^{\alpha_k} x - R_A x\| = \|\alpha_k J_A x - 2J_A x\| \leq |\alpha_k - 2| \|J_A x\|.$$

Clearly the right side tends to 0 as $k \to +\infty$, so $R_A^{\alpha_k}x \to R_A x$. This holds for every x.

□

Theorem 5.2.3. Let A and B be maximally monotone operators from $\mathcal{K}$ to $2^{\mathcal{K}}$, $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$ and $\operatorname{zer}(A+B) \neq \emptyset$. Let $(\alpha_k)_{k=1}^{+\infty}$ be an increasing sequence in $[1,2)$ such that $\lim\limits_{k\to+\infty} \alpha_k = 2$. for each k, consider the sequences

$$\begin{cases} y_n = J_B x_n \\ z_n = J_A(\alpha_k y_n - x_n) \\ x_{n+1} = x_n + (z_n - y_n). \end{cases} \tag{5.1}$$

Then the sequence x_n converges to some $x_k^* \in \operatorname{Fix} R_A^{\alpha_k} \circ R_B^{\alpha_k}$ such that

$$J_B x_k^* = \operatorname{zer}(A + B + (2-\alpha_k)\operatorname{Id}).$$

For any resulting sequence $(x_k^*)_{k=1}^{+\infty}$,

(1) $\lim\limits_{k\to+\infty} J_B x_k^* = \mathrm{P}_{\operatorname{zer}(A+B)}(0)$.

(2) Suppose $(x_k^*)_{k=1}^{+\infty}$ is a convergent sequence. Let $\lim\limits_{k\to+\infty} x_k^* = x^*$. Then $J_B x^*$ is a solution to $0 \in Ax + Bx$, and $\|J_B x^*\| \leq \|y\|$ for any $y \in \operatorname{zer}(A+B)$.

Proof. The existence of x_k^* follows from Theorem 3.2.5.

(1) Because A, B are maximally monotone and $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$, Fact 1.2.28 implies that $A+B$ is also maximally monotone. According to Theorem 3.2.5 (1), we have $J_B x_k^* \in \operatorname{zer}(A + B + (2-\alpha_k)\operatorname{Id})$. That means:

$$0 \in (A+B)J_B x_k^* + (2-\alpha_k)J_B x_k^*,$$

which can also be written as

$$0 \in (A+B)J_B x_k^* + (2-\alpha_k)(J_B x_k^* - 0).$$

As $\operatorname{zer}(A+B) \neq \emptyset$, according to Fact 5.2.1,

$$J_B x_k^* \to \mathrm{P}_{\operatorname{zer}(A+B)}(0) \text{ as } (2-\alpha_k) \downarrow 0.$$

Since $\lim\limits_{k\to+\infty} \alpha_k = 2$, we can also write this as

$$\lim_{k\to+\infty} J_B x_k^* = \mathrm{P}_{\operatorname{zer}(A+B)}(0).$$

(2) Here we want to prove $\lim\limits_{k\to+\infty} J_B(x_k^*) = J_B(x^*)$, where $\lim\limits_{k\to+\infty} x_k^* = x^*$. For any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for any $l \geq N$,

$$\|x_l^* - x^*\| < \epsilon.$$

Because B is maximally monotone, by Fact 1.2.38, J_B is firmly nonexpansive on $\mathbb{R}^m$. Thus,

$$\|J_B(x_l^*) - J_B(x^*)\| \leq \|x_l^* - x^*\| < \epsilon,$$

that is,

$$\lim_{k\to+\infty} J_B(x_k^*) = J_B(x^*).$$

As we already proved $\lim\limits_{k\to+\infty} J_B(x_k^*) = \mathrm{P}_{\mathrm{zer}(A+B)}(0)$, we have

$$J_B(x^*) = \mathrm{P}_{\mathrm{zer}(A+B)}(0).$$

Therefore, J_Bx^* is a solution to $0 \in Ax + Bx$, and $\|J_Bx^*\| \leq \|y\|$ for any $y \in \mathrm{zer}(A+B)$.

$\square$

This theorem shows that when $\mathrm{zer}(A+B) \neq \emptyset$, we can either use the Douglas-Rachford algorithm or the α-Douglas-Rachford algorithm to get a solution of it. Moreover, when $\mathrm{zer}(A+B)$ has multiple solutions, we can use α-Douglas-Rachford algorithm to get the one which has the shortest norm.

Theorem 5.2.4. Let $f, g, l \in \Gamma_0(\mathbb{R}^m)$, let $L : \mathbb{R}^m \to \mathbb{R}^m$ be a nonzero invertible linear operator where $[L(\mathrm{dom}\, f) - r] \cap \mathrm{intdom}(g \square l) \neq \emptyset$, and $\mathrm{dom}\, g^* \cap \mathrm{int\,dom}\, l^* \neq \emptyset$. Let z and $r \in \mathbb{R}^m$, let $\tau \in \mathbb{R}_{++}$ and $\sigma \in \mathbb{R}_{++}$. We set $A = \partial f, B = \partial g, D = \partial l$ which $\mathrm{dom}\, D^{-1} = \mathbb{R}^m$, and let α_k be a increasing convergent sequence in $[1, 2)$ such that $\lim\limits_{k\to+\infty} \alpha_k = 2$, the following holds:

(1) The sequence of problems: find $x \in \mathbb{R}^m$ such that

$$z \in Ax + \frac{2-\alpha_k}{\tau}x + \frac{\alpha_k}{4-\alpha_k}\boldsymbol{L}^\star \circ (B \overset{\frac{2-\alpha_k}{\sigma}}{\square} D) \circ (\boldsymbol{L}x - r) \tag{5.2}$$

where $\boldsymbol{L} = \frac{4-\alpha_k}{2}L$, together with the sequence of duals: find v such that there exists an $x \in \mathbb{R}^m$ that

$$\begin{cases} z - \frac{\alpha_k}{4-\alpha_k}\boldsymbol{L}^\star v \in Ax + \frac{(2-\alpha_k)}{\tau}x \\ v \in (B \overset{\frac{2-\alpha_k}{\sigma}}{\square} D) \circ (\boldsymbol{L}x - r). \end{cases} \tag{5.3}$$

have solution pairs (x_k, v_k) that converge to (x, v) satisfying the primal inclusion:

$$z \in Ax + L^\star(B\Box D)(Lx - r) \tag{5.4}$$

together with the dual inclusion:

$$\begin{cases} z - L^\star v \in Ax \\ v \in (B\Box D)(Lx - r). \end{cases} \tag{5.5}$$

(2) The sequence of optimization problems

$$\begin{aligned} &\mathrm{Argmin}_{x\in\mathbb{R}^m}\{f(x) + \frac{2-\alpha_k}{2\tau}\|x\|^2 + \\ &\qquad \frac{\alpha_k}{4-\alpha_k}[(g \boxdot l) \boxdot (\frac{\sigma}{2(2-\alpha_k)}\|x\|^2)] \circ (\boldsymbol{L}x - r) - \langle z, x\rangle\}, \end{aligned} \tag{5.6}$$

where $\boldsymbol{L} = \frac{4-\alpha_k}{2}L$, has a sequence of solutions x_k that converges to an element x of

$$\mathrm{Argmin}_{x\in\mathbb{R}^m}\{f(x) + (g \boxdot l) \circ (Lx - r) - \langle z, x\rangle\}, \tag{5.7}$$

when $\alpha_k \to 2$.

Proof. (1) Because $f, g, l \in \Gamma_0(\mathbb{R}^m)$, by Fact 1.3.27, A, B, D are maximally monotone. Suppose $\bar{x}_k$ is the solution of (5.2), and $\bar{v}_k$ is the solution of (5.3). Then by Theorem 3.2.9, $(\bar{x}_k, \bar{v}_k)$ is the solution of the inclusion problem:

$$\text{find } (x, v) \in \mathbb{R}^m \times \mathbb{R}^m \text{ such that } (x, v) \in \mathrm{zer}(\boldsymbol{A} + \boldsymbol{B} + (2 - \alpha_k)\,\mathrm{Id}), \tag{5.8}$$

and vice versa. Here $\boldsymbol{A} := V^{-1}(\frac{1}{2}S + Q)$, and $\boldsymbol{B} := V^{-1}(\frac{1}{2}S + M)$, where

$$\begin{aligned} &M : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (-z + Ax, r + B^{-1}v); \\ &Q : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (0, D^{-1}v); \\ &S : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (L^\star v, -Lx); \\ &V : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (\frac{x}{\tau} - \frac{1}{2}L^\star v, \frac{v}{\sigma} - \frac{1}{2}Lx). \end{aligned}$$

We proved that $\boldsymbol{A}$ and $\boldsymbol{B}$ are maximally monotone in Lemma 2.3.4. Since $\mathrm{dom}\, D^{-1} = \mathbb{R}^m$, by Theorem 3.2.10 and Theorem 3.2.5 (1), there exists an $x_k^* \in \mathrm{Fix}\, R_{\boldsymbol{A}}^{\alpha_k} \circ R_{\boldsymbol{B}}^{\alpha_k}$ such that $J_{\boldsymbol{B}} x_k^* \in \mathrm{zer}(\boldsymbol{A} + \boldsymbol{B} + (2 - \alpha_k)\,\mathrm{Id})$. Suppose $\mathrm{zer}(\boldsymbol{A} + \boldsymbol{B}) \neq \emptyset$. By Theorem 5.2.3, one has

$$\lim_{k\to+\infty} J_{\boldsymbol{B}} x_k^* = \mathrm{P}_{\mathrm{zer}(\boldsymbol{A}+\boldsymbol{B})}(0).$$

In other words, if we set $(x^*, v^*) = \lim_{k \to +\infty} J_{\boldsymbol{B}} x_k^*$, which $x^*, v^* \in \mathbb{R}^m$, from Lemma 2.3.4, it follows that x^* is the solution of the primal inclusion problem: find $x \in \mathbb{R}^m$ such that

$$z \in Ax + L^\star (B \Box D)(Lx - r),$$

and v^* is the solution of the dual inclusion problem: find v such that there exists an $x \in \mathbb{R}^m$ that

$$\begin{cases} z - L^\star v \in Ax \\ v \in (B \Box D)(Lx - r). \end{cases}$$

Therefore, we complete the proof.

(2) According to Theorem 3.2.11, we get that the optimization problem (5.6) is equivalent to the primal inclusion problem (5.2). Moreover, by using the Theorem 2.4.4, the inclusion problem (5.7) is equivalent to the optimization problem (5.4). Therefore, by using the result of (1), we complete the proof. □

Remark 5.1. If we let $\boldsymbol{A} := V^{-1}(\frac{1}{2}S + Q)$, and $\boldsymbol{B} := V^{-1}(\frac{1}{2}S + M)$, where

$$\begin{aligned} &M : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (-z + Ax, r + B^{-1}v); \\ &Q : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (0, D^{-1}v); \\ &S : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (L^\star v, -Lx); \\ &V : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (\frac{x}{\tau} - \frac{1}{2}L^\star v, \frac{v}{\sigma} - \frac{1}{2}Lx). \end{aligned}$$

According to Lemma 2.3.4, $(x, v) \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B})$ if and only if $(x, v) \in \operatorname{zer}(M + Q + S)$, i.e.,

$$(0, 0) \in (-z + Ax + L^\star v, r + B^{-1}v + D^{-1}v - Lx). \tag{5.9}$$

(5.9) is the exactly: find (x, v) such that

$$\begin{cases} z \in Ax + L^\star v \\ v \in (B \Box D)(Lx - r). \end{cases}$$

From Theorem 5.2.3, we know if we let $(x^*, v^*) = \lim_{\alpha_k \to 2} J_{\boldsymbol{B}} x_k^*$, $\sqrt{\|x^*\|^2 + \|v^*\|^2}$ is the shortest norm for all $(x, v) \in \operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$, since $\lim_{\alpha_k \to 2} J_{\boldsymbol{B}} x_k^* = \mathrm{P}_{\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})}(0)$.

Remark 5.2. In 2011, Wang [17] gave two self-dual regularizations of maximal monotone operators on $\mathcal{H}$, which can be effectively used to find the least norm solution to maximally monotone operators.

Remark 5.3. Dykstra method can also be used to find the least norm solution. See Bauschke and Borwein's paper [3].

Chapter 6

The application of the α-Douglas-Rachford algorithm

6.1 Overview

In this chapter, we are going to use the α-Douglas-Rachford algorithm to solve the inclusion problem

$$0 \in Ax + Bx,$$

where $A : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ and $B : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ are maximally monotone operators, in two different ways.

6.2 Least norm solution of the primal problem

Theorem 6.2.1. Let $A : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ and $B : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ be maximally monotone operators, $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$ and $\operatorname{zer}(A+B) \neq \emptyset$. In order to solve

$$0 \in Ax + Bx, \tag{6.1}$$

we can let α_k be an increasing convergent sequence in $[1, 2)$ such that $\lim\limits_{k\to+\infty} \alpha_k = 2$. Then we use the α-Douglas-Rachford algorithm to solve the problem

$$0 \in Ax + Bx + (2 - \alpha_k) \operatorname{Id}. \tag{6.2}$$

When $\alpha_k \to 2$, the answers of problem (6.2) converge to the shortest norm solution of problem (6.1).

Proof. Since A, B are maximally monotone operators, $0 \in \operatorname{int}(\operatorname{dom} A - \operatorname{dom} B)$ and $\operatorname{zer}(A+B) \neq \emptyset$, by Theorem 5.2.3, if we use the α-Douglas-Rachford algorithm to solve problem (6.2), for any fixed α_k, there exists a corresponding $x_k^* \in \operatorname{Fix} R_A^{\alpha_k} \circ R_B^{\alpha_k}$ such that $J_B x_k^* \in \operatorname{zer}(A + B + (2-\alpha_k)\operatorname{Id})$. Moreover, we have $\lim\limits_{k\to+\infty} J_B x_k^* = \mathrm{P}_{\operatorname{zer}(A+B)}(0)$. That is,

$$\|\lim_{k\to+\infty} J_B x_k^*\| \leq \|y\|$$

for any $y \in \operatorname{zer}(A+B)$. $\square$

6.3 Least norm solution of the primal-dual problem

Theorem 6.3.1. Let $A : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ and $B : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ be maximally monotone operators, and $\operatorname{zer}(A+B) \neq \emptyset$, let $L = \operatorname{Id}$. In order to solve the problem with primal inclusion: find $x \in \mathbb{R}^m$ such that

$$0 \in Ax + Bx, \tag{6.3}$$

together with the dual inclusion: find $v \in \mathbb{R}^m$ such that for some x,

$$\begin{cases} -v \in Ax \\ v \in Bx, \end{cases} \tag{6.4}$$

we can let α_k be an increasing convergent sequence in $[1, 2)$ such that

$$\lim_{k \to +\infty} \alpha_k = 2.$$

Let $\boldsymbol{L} = \frac{4-\alpha_k}{2} \operatorname{Id}$. Then we use the α-Douglas-Rachford algorithm to solve the problem with primal inclusion: find $x \in \mathbb{R}^m$ such that

$$0 \in Ax + \frac{2-\alpha_k}{\tau} x + \frac{\alpha_k}{4-\alpha_k} \boldsymbol{L}^\star (B \Box \frac{\sigma}{2-\alpha_k} \operatorname{Id})(\boldsymbol{L}x), \tag{6.5}$$

where $\tau \in \mathbb{R}_{++}$, $\sigma \in \mathbb{R}_{++}$, and $\tau\sigma < 4$, together with the dual inclusion: find v such that there exists an $x \in \mathbb{R}^m$ that

$$\begin{cases} -\frac{\alpha_k}{4-\alpha_k} \boldsymbol{L}^\star v \in Ax + \frac{2-\alpha_k}{\tau} x \\ v \in (B \Box \frac{\sigma}{2-\alpha_k} \operatorname{Id})(\boldsymbol{L}x). \end{cases} \tag{6.6}$$

When $\alpha_k \to 2$, the sequence of solutions of the primal-dual problem (6.5) together with (6.6) converge to the primal-dual shortest norm solution of problem (6.3) together with (6.4).

Proof. Let

$$\begin{aligned}
&M : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (Ax, B^{-1}v); \\
&Q : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (0, D^{-1}v) \text{ where } D = N_{\{0\}}; \\
&S : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (v, -x); \\
&V : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (\frac{x}{\tau} - \frac{1}{2}v, \frac{v}{\sigma} - \frac{1}{2}x).
\end{aligned}$$

$\boldsymbol{A} := V^{-1}(\frac{1}{2}S + Q)$, and $\boldsymbol{B} := V^{-1}(\frac{1}{2}S + M)$. According to Lemma 2.3.4, we get $\boldsymbol{A}$ and $\boldsymbol{B}$ are maximally monotone and

$$\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B}) = \operatorname{zer}(M + S + Q) = \operatorname{zer}(M + S),$$

as for any $v \in \mathbb{R}^m, D^{-1}v = 0$ by the definition of $N_{\{0\}}$. That means for all $(x, v) \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B})$, we have

$$\begin{aligned}(0,0) \in &(M+S)(x,v)\\ =&(Ax+v, B^{-1}v - x).\end{aligned}$$

That is

$$\begin{cases} 0 \in Ax + v \\ 0 \in B^{-1}v - x, \end{cases}$$

which is equivalent to

$$0 \in Ax + Bx$$

together with

$$\begin{cases} -v \in Ax \\ v \in Bx. \end{cases}$$

Therefore, solving $(x, v) \in \operatorname{zer}(\boldsymbol{A} + \boldsymbol{B})$ is equivalent to solving the problem with primal inclusion (6.3) together with dual inclusion (6.4). Since we have $\operatorname{zer}(A + B) \neq \emptyset$, according to Lemma 2.3.1, the primal inclusion (6.3) is equivalent to the dual inclusion (6.4). Therefore $\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B}) \neq \emptyset$.
In the proof of Theorem 3.2.10, we showed that once $\operatorname{dom} D^{-1} = \mathbb{R}^m$, we have $0 \in \operatorname{int}(\operatorname{dom} \boldsymbol{A} - \operatorname{dom} \boldsymbol{B})$. Since $D = N_{\{0\}}$, by Lemma 1.1.35, we have

$$\operatorname{dom} D^{-1} = \operatorname{ran} D = \mathbb{R}^m.$$

Therefore, we know $\boldsymbol{A}, \boldsymbol{B}$ are maximally monotone, $0 \in \operatorname{int}(\operatorname{dom} \boldsymbol{A} - \operatorname{dom} \boldsymbol{B})$ and $\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B}) \neq \emptyset$. By Theorem 6.2.1, we can use the α-Douglas-Rachford algorithm to solve each problem

$$\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B} + (2 - \alpha_k) \operatorname{Id}). \tag{6.7}$$

When $\alpha_k \to 2$, the answers from problem (6.7) converge to the shortest norm solution of problem $\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B})$. By Theorem 3.2.9, the solution of problem $\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B} + (2 - \alpha_k) \operatorname{Id})$ is also the primal-dual solution of problem (6.5) together with (6.6). Therefore, when $\alpha_k \to 2$, the sequence of solutions of the primal-dual problem (6.5) together with (6.6) converge to the primal-dual shortest norm solution of problem (6.3) together with (6.4). $\square$

Remark 6.1. The operator

$$M + S : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (Ax + v, B^{-1}v - x)$$

is maximally monotone. Because $M = (A, B^{-1}) : \mathcal{K} \to 2^{\mathcal{K}}$ is maximally monotone according to Fact 1.2.30. And because S is skew and linear, by Fact 1.2.31, S is maximally monotone. Moreover, since $\operatorname{dom} S = \mathcal{K}$, by Fact 1.2.28, we get $M + S$ is maximally monotone.

Remark 6.2. Note that Theorem 6.2.1 gives the primal shortest norm solution, but Theorem 6.3.1 gives the primal-dual shortest norm solution.

Remark 6.3. According to Corollary 2.3.3, the dual inclusion (6.4) is equivalent to the problem: find v' such that

$$0 \in A^{-1}(v') - B^{-1}(-v'), \tag{6.8}$$

which is the Attouch-Théra dual [1] of (6.3).

Chapter 7

Numerical experiments

7.1 Overview

In this chapter, we describe numerical experiments whose results confirms the properties of the α-Douglas-Rachford algorithm derived above.

In the following three numerical experiments, the operator D satisfies

$$\operatorname{dom} D^{-1} = \mathbb{R}^m.$$

Before we go to the numerical examples, the following formulas for proximal points are necessary.

Lemma 7.1.1. Let C be a closed convex set in $\mathbb{R}^m$, let $\tau \in \mathbb{R}_{++}$ and $f = \iota_C$. Then we have:

(1) $\operatorname{Prox}_{\tau f}(x) = \operatorname{P}_C(x)$.

(2) $\operatorname{Prox}_{\tau f^*}(x) = x - \tau \operatorname{P}_C(\frac{x}{\tau})$.

Proof. (1) Since C is a closed convex set, by by Example 1.3.13, $f \in \Gamma_0(\mathbb{R}^m)$. Then, for any $x \in \mathbb{R}^m$, we have

$$\begin{aligned}\operatorname{Prox}_{\tau f}(x) &= \operatorname{Prox}_{\tau \iota_C}(x) \\ &= \operatorname{Argmin}_{u \in \mathbb{R}^m} \{\tau \iota_C(u) + \frac{1}{2}\|u - x\|^2\} \\ &= \operatorname{Argmin}_{u \in C} \frac{1}{2}\|u - x\|^2 \\ &= \operatorname{P}_C(x).\end{aligned}$$

(2) Since $f \in \Gamma_0(\mathbb{R}^m)$, by Fact 1.3.40, $f = f^{**}$. Therefore, we use Fact 1.2.37

to get

$$\begin{aligned}\operatorname{Prox}_{\tau f^*}(x) =& [\operatorname{Id} - \tau \operatorname{Prox}_{\tau^{-1} f} \circ (\tau^{-1} \operatorname{Id})](x)\\ =& x - \tau \operatorname{Prox}_{\tau^{-1}\iota_C}(\frac{x}{\tau})\\ =& x - \tau \operatorname{Argmin}_{u\in C} \frac{1}{2}\|u - \frac{x}{\tau}\|^2\\ =& x - \tau \operatorname{P}_C\left(\frac{x}{\tau}\right).\end{aligned}$$

□

Lemma 7.1.2. Let $g = \|\cdot\|$, let $\tau \in \mathbb{R}_{++}$. Then for any $x \in \mathbb{R}^m$,

$$\operatorname{Prox}_{\tau g^*}(x) = \operatorname{P}_{B(0;1)}(x).$$

Proof. As

$$\begin{aligned}g^*(u) =& \sup_{x\in\mathbb{R}^m} \{\langle u, x\rangle - \|x\|\}\\ =& \iota_{B(0;1)}(u),\end{aligned}$$

by Lemma 7.1.1, we have

$$\operatorname{Prox}_{\tau g^*}(x) = \operatorname{P}_{B(0;1)}(x).$$

□

7.2 A feasibility problem

In this part, we consider solving the inclusion problem

$$z \in Ax + B(x - r), \tag{7.1}$$

where A, B are maximally monotone operators and $z, r \in \mathbb{R}^m$ are given.

Example 7.2.1. Let $f = \iota_{C_1}, g = \iota_{C_2}$, where C_1 is a circle centred at $(5, 0)$ with radius 2, and C_2 is a box centred at $(3, 1.5)$ with radius 1. Let $z = 0, r = 0$. If we let $A = \partial f, B = \partial g$, the problem (7.1) becomes

$$0 \in N_{C_1}(x) + N_{C_2}(x). \tag{7.2}$$

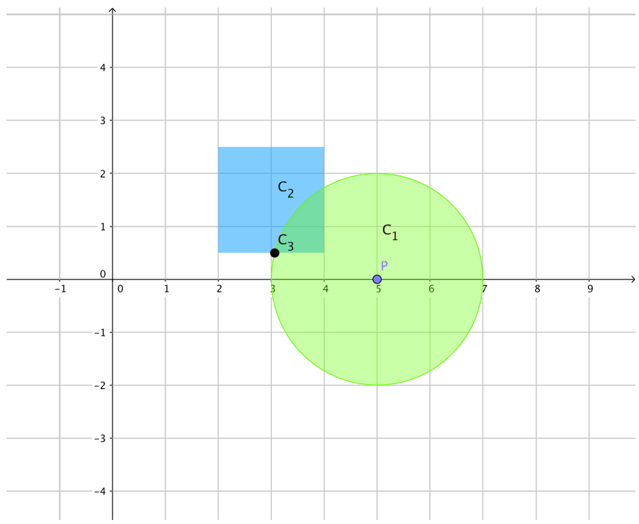

Figure 7.1: The plot of Example 7.2.1

Let α_k be a increasing convergent sequence in $[1, 2)$ such that $\lim_{k \to +\infty} \alpha_k = 2$. Then the following holds:

(1) The inclusion problem: The solution of problem: finding $x \in \mathbb{R}^2$ such that

$$0 \in N_{C_1}(x) + N_{C_2}(x) + (2 - \alpha_k)(x) \tag{7.3}$$

is the shortest norm solution of problem (7.2) when $\alpha_k \to 2$.

(2) The problem (7.3) can be solved by the α-Douglas-Rachford algorithm. Moreover, as $\alpha_k \to 2$, the optimization result which is gotten by the α-Douglas-Rachford algorithm converges to the shortest norm solution of (7.2).

Proof. (1) We apply Theorem 6.2.1 to complete this proof.

Since C_1 and C_2 are closed, bounded and convex sets, according to Example 1.2.33, N_{C_1} and N_{C_2} are maximally monotone operators. Because we also have $\text{int}(C_1 \cap C_2) \neq \emptyset$, according to Theorem 3.2.5, the inclusion problems (7.3) can be solved by the α-Douglas-Rachford algorithm (3.3) by letting $A = N_{C_1}$ and $B = N_{C_2}$. That is:

$$\begin{cases} y_n = J_{N_{C_2}}(x_n) \\ z_n = J_{N_{C_1}}(\alpha_k y_n - x_n) \\ x_{n+1} = x_n + (z_n - y_n). \end{cases} \tag{7.4}$$

According to Example 1.3.25,

$$N_{C_1} = \partial\iota_{C_1}, N_{C_2} = \partial\iota_{C_2}.$$

Therefore, algorithm (7.4) becomes

$$\begin{cases} y_n = J_{\partial\iota_{C_2}}(x_n) \\ z_n = J_{\partial\iota_{C_1}}(\alpha_k y_n - x_n) \\ x_{n+1} = x_n + (z_n - y_n). \end{cases} \tag{7.5}$$

By Fact 1.3.57, we have

$$J_{\partial\iota_{C_2}} = \operatorname{Prox}_{\iota_{C_2}}, \text{ and } J_{\partial\iota_{C_1}} = \operatorname{Prox}_{\iota_{C_1}}.$$

Let's plug $\operatorname{Prox}_{\iota_{C_2}}, \operatorname{Prox}_{\iota_{C_1}}$, into (7.5) instead of $J_{\partial\iota_{C_2}}$, and $J_{\partial\iota_{C_1}}$ respectively, we get

$$\begin{cases} y_n = \operatorname{Prox}_{\iota_{C_2}}(x_n) \\ z_n = \operatorname{Prox}_{\iota_{C_1}}(\alpha_k y_n - x_n) \\ x_{n+1} = x_n + (z_n - y_n). \end{cases} \tag{7.6}$$

According to Lemma 7.1.1(1), we have

$$\operatorname{Prox}_{\iota_{C_1}}(x) = \mathrm{P}_{C_1}(x)$$

and

$$\operatorname{Prox}_{\iota_{C_2}}(x) = \mathrm{P}_{C_2}(x).$$

As C_1 is a circle centred at $(5, 0)$ with radius 2, we can also write

$$\operatorname{Prox}_{\iota_{C_1}}(x) = (5, 0) + \mathrm{P}_{B(0;2)}(x - (5, 0)).$$

Hence, when we choosing $x_0 = (5, 1)$ as starting values, for any fixed k, let $\alpha_k = 2 - \frac{1}{k}$, the iterative scheme algorithm (7.6) becomes:

$$\begin{cases} y_n = \mathrm{P}_{C_2}(x_n) \\ z_n = (5, 0) + \mathrm{P}_{B(0;2)}(\alpha_k y_n - x_n) \\ x_{n+1} = x_n + (z_n - y_n). \end{cases}$$

The stopping criteria of this algorithm is: Let $\epsilon = 10^{-5}$, continue running this iteration until $\|x_{n+1} - x_n\| < \epsilon$.

Table 7.1: Example 1: Six fixed α_k where $\alpha_k = 2 - 1/k$, the corresponding optimization point y^*, and the norm of y^*.

k	y^*	$\|y^*\|$
1	(3.0635,0.5)	3.104
10	(3.0635,0.5)	3.104
50	(3.0635,0.5)	3.104
100	(3.0635,0.5)	3.104
1000	(3.0635,0.5)	3.104
10000	(3.0635,0.5)	3.104

As we can see, the optimization result y^* doesn't change its value when we use different value of k. It is clear that y^* locate at the boundary of C_1 and also locate at the boundary of C_2. With the help of Figure 7.2.1, we get y^* is the smallest norm solution of problem (7.2).

We also tried to run this algorithm from different starting point, and the result shows once we fix the value of k, the result we get does not change if we change its starting point.

However, when we use the classic Douglas-Rachford algorithm to solve (7.2), the answer changes if we choose different starting point. Here is the result:

Table 7.2: Example 1: starting point x_0, the corresponding optimization point y^*, and the norm of y^*.

x_0	y^*	$\|y^*\|$
(5,1)	(4,0.8944)	4.0988
(-3,1)	(3.0785,0.5548)	3.1281
(-4,-6)	(4,0.5)	4.0311
(10,-20)	(4,0.5)	4.0311

That means, in this example, if we directly use the classic Douglas-Rachford algorithm to solve problem (7.2), the answer we get may not have the shortest norm. However, if we use the α-Douglas-Rachford algorithm to solve problem (7.3) and let $\alpha_k \to 2$, the answer we get has the shortest norm.

□

7.3 A Heron problem

In this part, we are going to consider solving a primal-dual inclusion problem by using α-Douglas-Rachford algorithm.

Theorem 7.3.1. Let $A : \mathbb{R}^m \to 2^{\mathbb{R}^m}$, $B : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ and $D : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ be maximally monotone operators and $\operatorname{dom} D^{-1} = \mathbb{R}^m$. Let $z, r \in \mathbb{R}^m$, and let $L : \mathbb{R}^m \to \mathbb{R}^m$ be a nonzero linear invertible operator. Let α_k be an increasing convergent sequence in $[1, 2)$ such that $\lim\limits_{k\to+\infty} \alpha_k = 2$.

(1) The problem with primal inclusion: find $x \in \mathbb{R}^m$ such that

$$z \in Ax + \frac{2-\alpha_k}{\tau}x + \frac{\alpha_k}{4-\alpha_k}\boldsymbol{L}^\star(B \overset{\frac{2-\alpha_k}{\sigma}}{\square} D)(\boldsymbol{L}x - r), \tag{7.7}$$

where $\boldsymbol{L} = \frac{4-\alpha_k}{2}L$, $\tau \in \mathbb{R}_{++}$, $\sigma \in \mathbb{R}_{++}$, and $\tau\sigma\|L\|^2 < 4$, together with the dual inclusion: find v such that there exists an $x \in \mathbb{R}^m$ that

$$\begin{cases} z - \frac{\alpha_k}{4-\alpha_k}\boldsymbol{L}^\star v \in Ax + \frac{(2-\alpha_k)}{\tau}x \\ v \in (B \overset{\frac{2-\alpha_k}{\sigma}}{\square} D) \circ (\boldsymbol{L}x - r) \end{cases} \tag{7.8}$$

can be solved by using the α-Douglas-Rachford algorithm:

$$\begin{cases} y_n = J_{\boldsymbol{B}}x_n \\ z_n = J_{\boldsymbol{A}}(\alpha_k y_n - x_n) \\ x_{n+1} = x_n + (z_n - y_n). \end{cases} \tag{7.9}$$

Here, $\boldsymbol{A} := V^{-1}(\frac{1}{2}S + Q)$, and $\boldsymbol{B} := V^{-1}(\frac{1}{2}S + M)$, where

$$\begin{aligned} &M : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (-z + Ax, r + B^{-1}v); \\ &Q : \mathcal{K} \to 2^{\mathcal{K}} : (x, v) \mapsto (0, D^{-1}v); \\ &S : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (L^\star v, -Lx); \\ &V : \mathcal{K} \to \mathcal{K} : (x, v) \mapsto (\frac{x}{\tau} - \frac{1}{2}L^\star v, \frac{v}{\sigma} - \frac{1}{2}Lx). \end{aligned}$$

(2) The algorithm (7.9) can be rewritten as

$$\begin{cases} y_{1n} = J_{\tau A}(x_{1n} - \frac{\tau}{2}L^\star x_{2n} + \tau z) \\ y_{2n} = J_{\sigma B^{-1}}(x_{2n} - \frac{\sigma}{2}Lx_{1n} + \sigma L y_{1n} - \sigma r) \\ w_{1n} = \alpha_k y_{1n} - x_{1n} \\ w_{2n} = \alpha_k y_{2n} - x_{2n} \\ z_{1n} = w_{1n} - \frac{\tau}{2}L^\star w_{2n} \\ z_{2n} = J_{\sigma D^{-1}}(w_{2n} - \frac{\sigma}{2}Lw_{1n} + \sigma L z_{1n}) \\ x_{1n+1} = x_{1n} + (z_{1n} - y_{1n}) \\ x_{2n+1} = x_{2n} + (z_{2n} - y_{2n}), \end{cases} \tag{7.10}$$

(3) Let $f, g, l \in \Gamma_0(\mathbb{R}^m)$, and $A = \partial f, B = \partial g, D = \partial l$. Then algorithm (7.10) implies the following algorithm

$$\begin{cases} y_{1n} = \mathrm{Prox}_{\tau f}(x_{1n} - \frac{\tau}{2}L^\star x_{2n} + \tau z) \\ y_{2n} = \mathrm{Prox}_{\sigma g^*}(x_{2n} - \frac{\sigma}{2}Lx_{1n} + \sigma L y_{1n} - \sigma r) \\ w_{1n} = \alpha_k y_{1n} - x_{1n} \\ w_{2n} = \alpha_k y_{2n} - x_{2n} \\ z_{1n} = w_{1n} - \frac{\tau}{2}L^\star w_{2n} \\ z_{2n} = \mathrm{Prox}_{\sigma l^*}(w_{2n} - \frac{\sigma}{2}Lw_{1n} + \sigma L z_{1n}) \\ x_{1n+1} = x_{1n} + (z_{1n} - y_{1n}) \\ x_{2n+1} = x_{2n} + (z_{2n} - y_{2n}). \end{cases} \tag{7.11}$$

Proof. (1) For any fixed α_k, we can apply Theorem 3.2.10 to complete this proof.

(2) According to the definition of $\boldsymbol{A}$ and $\boldsymbol{B}$, (7.9) can be rewritten as

$$\begin{cases} x_n = [V^{-1}(\frac{1}{2}S + M) + \mathrm{Id}](y_n), \\ \alpha_k y_n - x_n = [V^{-1}(\frac{1}{2}S + Q) + \mathrm{Id}](z_n), \\ x_{n+1} = x_n + (z_n - y_n). \end{cases}$$

Set $w_n = \alpha_k y_n - x_n$, we have

$$\begin{cases} V(x_n - y_n) = (\frac{1}{2}S + M)(y_n), \\ w_n = \alpha_k y_n - x_n, \\ V(w_n - z_n) = (\frac{1}{2}S + Q)(z_n), \\ x_{n+1} = x_n + (z_n - y_n). \end{cases} \tag{7.12}$$

Here, we let

$$x_n = (x_{1n}, x_{2n}), y_n = (y_{1n}, y_{2n}), w_n = (w_{1n}, w_{2n}), z_n = (x_{1n}, z_{2n}).$$

Since $M, Q, S, V, \boldsymbol{A}$, and $\boldsymbol{B}$ are constructed by (M), (Q), (S), (V), (A) and (B) respectively, (7.12) is equivalent to

$$\begin{cases} \frac{x_{1n} - y_{1n}}{\tau} - \frac{1}{2}L^\star(x_{2n} - y_{2n}) = \frac{1}{2}L^\star y_{2n} - z + Ay_{1n} \\ \frac{x_{2n} - y_{2n}}{\sigma} - \frac{1}{2}L(x_{1n} - y_{1n}) = -\frac{1}{2}Ly_{1n} + r + B^{-1}y_{2n} \\ w_{1n} = \alpha_k y_{1n} - x_{1n} \\ w_{2n} = \alpha_k y_{2n} - x_{2n} \\ \frac{w_{1n} - z_{1n}}{\tau} - \frac{1}{2}L^\star(w_{2n} - z_{2n}) = \frac{1}{2}L^\star z_{2n} + 0 \\ \frac{w_{2n} - z_{2n}}{\sigma} - \frac{1}{2}L(w_{1n} - z_{1n}) = -\frac{1}{2}Lz_{1n} + D^{-1}z_{2n} \\ x_{1n+1} = x_{1n} + (z_{1n} - y_{1n}) \\ x_{2n+1} = x_{2n} + (z_{2n} - y_{2n}). \end{cases}$$

Finally, we simplify each line of the above algorithm and get:

$$\begin{cases} y_{1n} = J_{\tau A}(x_{1n} - \frac{\tau}{2}L^\star x_{2n} + \tau z) \\ y_{2n} = J_{\sigma B^{-1}}(x_{2n} - \frac{\sigma}{2}Lx_{1n} + \sigma L y_{1n} - \sigma r) \\ w_{1n} = \alpha_k y_{1n} - x_{1n} \\ w_{2n} = \alpha_k y_{2n} - x_{2n} \\ z_{1n} = w_{1n} - \frac{\tau}{2}L^\star w_{2n} \\ z_{2n} = J_{\sigma D^{-1}}(w_{2n} - \frac{\sigma}{2}Lw_{1n} + \sigma L z_{1n}) \\ x_{1n+1} = x_{1n} + (z_{1n} - y_{1n}) \\ x_{2n+1} = x_{2n} + (z_{2n} - y_{2n}), \end{cases}$$

which is algorithm (7.10). Therefore, algorithm (7.9) is equivalent to algorithm (7.10)

(3) Since $A = \partial f, B = \partial g, D = \partial l$, we can write (7.10) as

$$\begin{cases} y_{1n} = J_{\tau \partial f}(x_{1n} - \frac{\tau}{2}L^\star x_{2n} + \tau z) \\ y_{2n} = J_{\sigma(\partial g)^{-1}}(x_{2n} - \frac{\sigma}{2}Lx_{1n} + \sigma L y_{1n} - \sigma r) \\ w_{1n} = \alpha_k y_{1n} - x_{1n} \\ w_{2n} = \alpha_k y_{2n} - x_{2n} \\ z_{1n} = w_{1n} - \frac{\tau}{2}L^\star w_{2n} \\ z_{2n} = J_{\sigma(\partial l)^{-1}}(w_{2n} - \frac{\sigma}{2}Lw_{1n} + \sigma L z_{1n}) \\ x_{1n+1} = x_{1n} + (z_{1n} - y_{1n}) \\ x_{2n+1} = x_{2n} + (z_{2n} - y_{2n}). \end{cases} \tag{7.13}$$

Because $g, l \in \Gamma_0(\mathbb{R}^m)$, by Fact 1.3.46,

$$(\partial g)^{-1} = \partial g^* \quad \text{and} \quad (\partial l)^{-1} = \partial l^*.$$

For any $\lambda \in \mathbb{R}_{++}$, we have $\lambda \partial f = \partial(\lambda f)$. By Fact 1.3.57, we have

$$J_{\tau\partial f} = \mathrm{Prox}_{\tau f}; \quad J_{\sigma(\partial g)^{-1}} = \mathrm{Prox}_{\sigma g^*}; \quad \text{and} \quad J_{\sigma(\partial l)^{-1}} = \mathrm{Prox}_{\sigma l^*}.$$

Let's plug $\mathrm{Prox}_{\tau f}, \mathrm{Prox}_{\sigma g^*}$, and $\mathrm{Prox}_{\sigma l^*}$ into (7.13) instead of $J_{\tau\partial f}, J_{\sigma(\partial g)^{-1}}$, and $J_{\sigma(\partial l)^{-1}}$ respectively, we get

$$\begin{cases} y_{1n} = \mathrm{Prox}_{\tau f}(x_{1n} - \frac{\tau}{2}L^\star x_{2n} + \tau z) \\ y_{2n} = \mathrm{Prox}_{\sigma g^*}(x_{2n} - \frac{\sigma}{2}Lx_{1n} + \sigma L y_{1n} - \sigma r) \\ w_{1n} = \alpha_k y_{1n} - x_{1n} \\ w_{2n} = \alpha_k y_{2n} - x_{2n} \\ z_{1n} = w_{1n} - \frac{\tau}{2}L^\star w_{2n} \\ z_{2n} = \mathrm{Prox}_{\sigma l^*}(w_{2n} - \frac{\sigma}{2}Lw_{1n} + \sigma L z_{1n}) \\ x_{1n+1} = x_{1n} + (z_{1n} - y_{1n}) \\ x_{2n+1} = x_{2n} + (z_{2n} - y_{2n}), \end{cases}$$

which is exactly the alorithm (7.11). Therefore, algorithm (7.10) implies (7.11) in this case.

□

Lemma 7.3.2. Let C be a closed convex set in $\mathbb{R}^m$. Then we have

$$d_C = \|\cdot\| \Box \iota_C,$$

and

$$\partial d_C = \partial \|\cdot\| \Box N_C.$$

Proof.

$$\begin{aligned} d_C(x) &= \inf_{y \in C} \|x - y\| \\ &= \inf_{y \in \mathbb{R}^m} \{\|x - y\| + \iota_C(y)\} \\ &= (\iota_C \Box \|\cdot\|)(x). \end{aligned}$$

Since C is a closed bounded convex set, for any $u \in \mathbb{R}^m$,

$$\begin{aligned} \iota_C^*(u) &= \sup_{x \in \mathbb{R}^m} \{\langle x, u\rangle - \iota_C(x)\} \\ &= \sup_{x \in C} \langle x, u\rangle < +\infty. \end{aligned}$$

Therefore, $\operatorname{dom} \iota_C^* = \mathbb{R}^m$. According to Fact 1.3.40, $\operatorname{dom} \|\cdot\|^* \neq \emptyset$, that means $\operatorname{dom} \|\cdot\|^* \cap \operatorname{intdom} \iota_C^* \neq \emptyset$. Moreover, from Example 1.3.25, we have $N_C = \partial \iota_C$. By using Theorem 2.4.3 , we have

$$\begin{aligned} \partial d_C &= \partial(\|\cdot\| \Box \iota_C) \\ &= \partial \|\cdot\| \Box \partial \iota_C \\ &= \partial \|\cdot\| \Box N_C. \end{aligned}$$

□

Example 7.1. Let $f = \iota_{C_1}, l = \iota_{C_2}$, where C_1 is a circle centred at $(5, 0)$ with radius 2, and C_2 is a box centred at $(-2, 4)$ with radius 0.5. A simple Heron problem is to solve

$$\min_{x \in C_1} d_{C_2}(x) = \min_{x \in \mathbb{R}^2} (\iota_{C_1}(x) + d_{C_2}(x)).$$

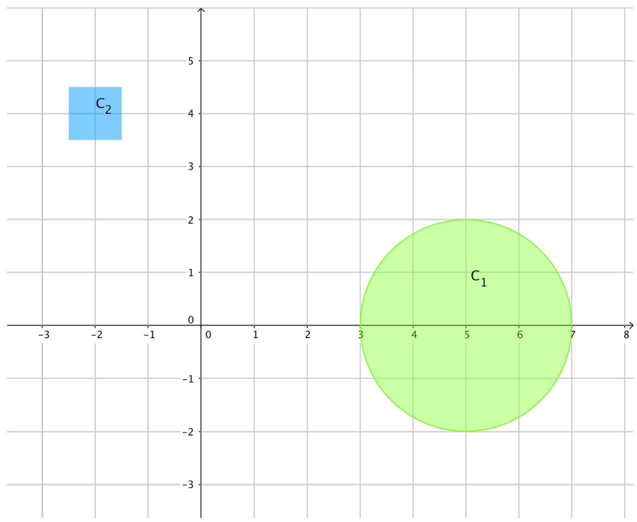

Figure 7.2: The plot of Example 7.1

Let $g = \|\cdot\|$, $z = 0, r = 0$. Then, we aim to solve the problem with the primal inclusion: find $x \in \mathbb{R}^m$ such that

$$0 \in N_{C_1}(x) + (\partial\|\cdot\| \Box N_{C_2})(x) \tag{7.14}$$

together with the dual inclusion: find v such that there exists an $x \in \mathbb{R}^m$ that

$$\begin{cases} -v \in N_{C_1}(x) \\ v \in (\partial\|\cdot\| \Box N_{C_2})(x). \end{cases} \tag{7.15}$$

Here, $L = \mathrm{Id}$. Let α_k be an increasing convergent sequence in $[1, 2)$ such that $\lim\limits_{k\to+\infty} \alpha_k = 2$. For each α_k, let $\boldsymbol{L} = \frac{4-\alpha_k}{2}\mathrm{Id}$. Then the following holds:

(1) When $\alpha_k \to 2$, the sequence of problems: find $x \in \mathbb{R}^m$ such that

$$0 \in N_{C_1}(x) + \frac{2-\alpha_k}{\tau}x + \frac{\alpha_k}{4-\alpha_k}\boldsymbol{L}^\star(\partial\|\cdot\| \overset{\frac{2-\alpha_k}{\sigma}}{\Box} N_{C_2})(\boldsymbol{L}x), \tag{7.16}$$

where $\tau \in \mathbb{R}_{++}$, $\sigma \in \mathbb{R}_{++}$, and $\tau\sigma\|L\|^2 < 4$, together with the sequence of duals: find v such that there exists an $x \in \mathbb{R}^m$ that

$$\begin{cases} -\frac{\alpha_k}{4-\alpha_k}\boldsymbol{L}^\star v \in N_{C_1}(x) + \frac{(2-\alpha_k)}{\tau}x \\ v \in (\partial\|\cdot\| \overset{\frac{2-\alpha_k}{\sigma}}{\Box} N_{C_2})(\boldsymbol{L}x) \end{cases} \tag{7.17}$$

have solution pairs (x_k, v_k) that converge to (x, v) satisfying the primal inclusion (7.14) together with the dual inclusion (7.15).

(2) The problem with primal inclusion (7.16) and dual inclusion (7.17) can be solved by the α-Douglas-Rachford algorithm. Moreover, the optimization result which is obtained by the classic Douglas-Rachford algorithm is the solution of the problem with primal inclusion (7.14) and dual inclusion (7.15).

Proof. (1) Since C_1 and C_2 are closed, bounded and convex sets, according to Example 1.2.33, N_{C_1} and N_{C_2} are maximally monotone operators. Because $\|\cdot\| \in \Gamma_0(\mathbb{R}^2)$, by Fact 1.3.27, $\partial\|\cdot\|$ is maximally monotone. According to Example 1.3.25,

$$N_{C_1} = \partial\iota_{C_1}, N_{C_2} = \partial\iota_{C_2},$$

and according to Example 1.3.13, $\iota_{C_1}, \iota_{C_2} \in \Gamma_0(\mathbb{R}^2)$.

As $g = \|\cdot\|$, and $\operatorname{dom}(g\Box l) = \operatorname{dom} g + \operatorname{dom} l$, we have $L(\operatorname{dom} f) \cap \operatorname{intdom}(g\Box l) \neq \emptyset$. As $l = \iota_{C_2}$ where C_2 is a closed bounded convex set, for any $u \in \mathbb{R}^2$,

$$\begin{aligned} l^*(u) &= \sup_{x\in\mathbb{R}^2} \{\langle x, u\rangle - l(x)\} \\ &= \sup_{x\in C_2} \langle x, u\rangle < +\infty. \end{aligned}$$

Therefore, $\operatorname{dom} l^* = \mathbb{R}^2$. Consequently, $\operatorname{dom} g^* \cap \operatorname{intdom} l^* \neq \emptyset$.

By Fact 1.3.46,

$$D^{-1} = (\partial l)^{-1} = \partial l^*.$$

As $\operatorname{dom} l^* = \mathbb{R}^2$, we can use Fact 1.3.43 to get

$$\operatorname{intdom} l^* = \operatorname{dom} \partial l^* = \operatorname{dom} l^* = \mathbb{R}^2.$$

That is, $\operatorname{dom} D^{-1} = \mathbb{R}^2$.

Now, we can apply Theorem 5.2.4 to complete the proof.

(2) As

$$\operatorname{dom} D^{-1} = \mathbb{R}^2,$$

by Theorem 3.2.10, the problem with primal inclusion: For any fixed α_k, find $x \in \mathbb{R}^2$ such that

$$0 \in N_{C_1}x + \frac{2-\alpha_k}{\tau}x + \frac{\alpha_k}{4-\alpha_k}\boldsymbol{L}^\star(\partial\|\cdot\| \overset{\frac{2-\alpha_k}{\sigma}}{\Box} N_{C_2})(\boldsymbol{L}x). \tag{7.18}$$

together with the dual inclusion: find v such that there exists an $x \in \mathbb{R}^m$ that

$$\begin{cases} -\frac{\alpha_k}{4-\alpha_k}\boldsymbol{L}^\star v \in N_{C_1}x + \frac{(2-\alpha_k)}{\tau}x, \\ v \in (\partial\|\cdot\| \overset{\frac{2-\alpha_k}{\sigma}}{\square} N_{C_2})(\boldsymbol{L}x). \end{cases} \tag{7.19}$$

can be solved by the α-Douglas-Rachford algorithm.

Because $\iota_{C_1}, \iota_{C_2} \in \Gamma_0(\mathbb{R}^2)$, according to Theorem 7.3.1, we can use algorithm (7.11) to solve the problem with primal inclusion (7.18) and dual inclusion (7.19).

When solving that problem with algorithm (7.11), by Lemma 7.1.1(1),

$$\mathrm{Prox}_{\tau f}(x) = \mathrm{P}_{C_1}(x).$$

As C_1 is a circle centred at $(5,0)$ with radius 2, we can also write

$$\mathrm{Prox}_{\tau f}(x) = (5,0) + \mathrm{P}_{B(0;2)}(x-(5,0)).$$

We use Lemma 7.1.1(2) to get

$$\mathrm{Prox}_{\sigma l^*}(x) = x - \sigma\,\mathrm{P}_{C_2}\left(\frac{x}{\sigma}\right).$$

As $g = \|\cdot\|$, by Lemma 7.1.2,

$$\mathrm{Prox}_{\sigma g^*}(x) = \mathrm{P}_{B(0;1)}(x).$$

Because $\tau, \sigma, \|L\|$ must satisfy the relation $\tau\sigma\|L\|^2 < 4$, and $L = \mathrm{Id}$. To be on the safe side, we let $\sigma = 2$, and $\tau = 3/2$. Hence, when we choosing $x_0 = (5,-2), v_0 = (0,0)$ as starting values, for any fixed k, let $\alpha_k = 2 - \frac{1}{k}$, the iterative scheme algorithm (7.11) becomes:

$$\begin{cases} y_{1n} = (5,0) + \mathrm{P}_{B(0;2)}(x_{1n} - \frac{\tau}{2}x_{2n} - (5,0)) \\ y_{2n} = \mathrm{P}_{B(0;1)}(x_{2n} - \frac{\sigma}{2}x_{1n} + \sigma y_{1n}) \\ w_{1n} = \alpha_k y_{1n} - x_{1n} \\ w_{2n} = \alpha_k y_{2n} - x_{2n} \\ z_{1n} = w_{1n} - \frac{\tau}{2}w_{2n} \\ z_{2n} = w_{2n} - \frac{\sigma}{2}w_{1n} + \sigma z_{1n} - \sigma\,\mathrm{P}_{C_2}((w_{2n} - \frac{\sigma}{2}w_{1n} + \sigma z_{1n})/\sigma) \\ x_{1n+1} = x_{1n} + (z_{1n} - y_{1n}) \\ x_{2n+1} = x_{2n} + (z_{2n} - y_{2n}). \end{cases} \tag{7.20}$$

The stopping criteria of this algorithm is: Let $\epsilon = 10^{-5}$, continue running this iteration until $\|x_{1n+1} - x_{1n}\| < \epsilon$, and $\|x_{2n+1} - x_{2n}\| < \epsilon$.

Table 7.3: Example 2: Seven fixed α_k, where $\alpha_k = 2 - 1/k$, the corresponding optimal point y_1^* and y_2^*, together with the case $\alpha = 2$.

α_k	y_1^*	y_2^*
1	(3.0041,0.1277)	(0.8759,-0.4825)
$2 - \frac{1}{10}$	(3.1449,0.7475)	(0.8705,-0.4922)
$2 - \frac{1}{50}$	(3.2156,0.9033)	(0.8781,-0.4786)
$2 - \frac{1}{100}$	(3.2270,0.9254)	(0.8792,-0.4764)
$2 - \frac{1}{1000}$	(3.2378,0.9459)	(0.8803,-0.4743)
$2 - \frac{1}{10000}$	(3.2389,0.9480)	(0.8805,-0.4741)
$2 - \frac{1}{10^6}$	(3.2391,0.9482)	(0.8805,-0.4741)
$\alpha = 2$	(3.2391,0.9482)	(0.8805,-0.4741)

As we can see, the numerical result shows that as k gets larger, or we can say as α_k gets closer to 2, the optimal result y_1^* and y_2^* which are gotten by the α-Douglas-Rachford algorithm gets closer to the optimal result which is gotten by the classic Douglas-Rachford algorithm.

Then we run the algorithm (7.20) again with the same starting point and same stoping criteria, but in this time, we set $\tau = 1, \sigma = 1$. Here is the result:

Table 7.4: Example 2: Seven fixed α_k, where $\alpha_k = 2 - 1/k$, the corresponding optimal point y_1^* and y_2^*, together with the case $\alpha = 2$.

α_k	y_1^*	y_2^*
1	(3.0020,0.0899)	(0.8723,-0.4890)
$2 - \frac{1}{10}$	(3.1196,0.6813)	(0.8639,-0.5037)
$2 - \frac{1}{50}$	(3.2063,0.8847)	(0.8762,-0.4820)
$2 - \frac{1}{100}$	(3.2219,0.9157)	(0.8783,-0.4782)
$2 - \frac{1}{1000}$	(3.2373,0.9449)	(0.8802,-0.4745)
$2 - \frac{1}{10000}$	(3.2389,0.9479)	(0.8804,-0.4741)
$2 - \frac{1}{10^6}$	(3.2391,0.9482)	(0.8805,-0.4741)
$\alpha = 2$	(3.2391,0.9482)	(0.8805,-0.4741)

Again, the numerical result shows that as k gets larger, or we can say as

α_k gets closer to 2, the optimal result y_1^* and y_2^* which are gotten by the α-Douglas-Rachford algorithm gets closer to the optimal result which is gotten by the classic Douglas-Rachford algorithm.

We also tried to run this algorithm with different starting point with fixed τ and σ, and the result shows that once we fix the value of k, the result we get does not change if we change its starting point. This is because when α_k fixed, $0 \in (\boldsymbol{A} + \boldsymbol{B} + (2 - \alpha_k)\,\mathrm{Id})(x)$ has a unique solution. So it does not matter when we change the starting point of the algorithm.

$\square$

7.4 A feasibility problem solved by the primal-dual formulation

Example 7.2. Let $f = \iota_{C_1}, g = \iota_{C_2}$, where C_1 is a circle centred at $(5, 0)$ with radius 2, and C_2 is a box centred at $(3, 1.5)$ with radius 1. Let $z = 0, r = 0$. If we let $A = \partial f, B = \partial g$, then, we aim to solve the problem with the primal inclusion: find $x \in \mathbb{R}^m$ such that

$$0 \in N_{C_1}(x) + N_{C_2}(x). \tag{7.21}$$

together with the dual inclusion: find $v \in \mathbb{R}^m$ such that

$$\begin{cases} -v \in N_{C_1}(x) \\ v \in N_{C_2}(x), \end{cases} \tag{7.22}$$

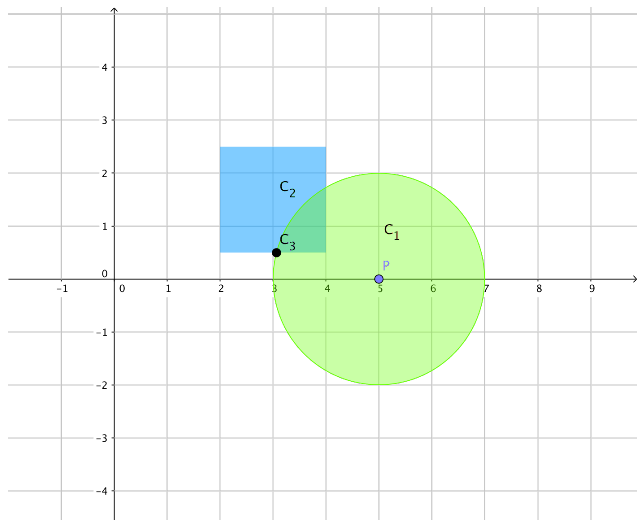

Figure 7.3: The plot of Example 7.2

We can solve (7.22) by our α-Douglas-Rachford method.

Let α_k be a increasing convergent sequence in $[1, 2)$ such that $\lim\limits_{k\to+\infty} \alpha_k = 2$. For each α_k, let $\boldsymbol{L} = \frac{4-\alpha_k}{2} \operatorname{Id}$. Then we use the α-Douglas-Rachford algorithm to solve the problem with primal inclusion: find $x \in \mathbb{R}^m$ such that

$$0 \in N_{C_1}(x) + \frac{2-\alpha_k}{\tau} x + \frac{\alpha_k}{4-\alpha_k} \boldsymbol{L}^\star (N_{C_2} \Box \frac{\sigma}{2-\alpha_k} \operatorname{Id})(\boldsymbol{L}x), \tag{7.23}$$

where $\tau \in \mathbb{R}_{++}$, $\sigma \in \mathbb{R}_{++}$, and $\tau\sigma < 4$, together with the dual inclusion: find v such that there exists an $x \in \mathbb{R}^m$ that

$$\begin{cases} -\frac{\alpha_k}{4-\alpha_k} \boldsymbol{L}^\star v \in N_{C_1}(x) + \frac{2-\alpha_k}{\tau} x \\ v \in (N_{C_2} \Box \frac{\sigma}{2-\alpha_k} \operatorname{Id})(\boldsymbol{L}x). \end{cases} \tag{7.24}$$

When $\alpha_k \to 2$, the sequence of solutions of the primal-dual problem (7.23) together with (7.24) converge to the primal-dual shortest norm solution of problem (7.21) together with (7.22).

7.4.1 The algorithm

Since C_1 and C_2 are closed, bounded and convex sets, according to Example 1.2.33, N_{C_1} and N_{C_2} are maximally monotone operators. Because $\operatorname{int}(C_1 \cap C_2) \neq$

$\emptyset$. we have $\mathrm{zer}(N_{C_1} + N_{C_2}) \neq \emptyset$. Due to Theorem 6.3.1, we can use α-Douglas-Rachford algorithm to solve the problem with primal inclusion (7.23) together with the dual inclusion (7.24).
According to Example 1.3.25,

$$N_{C_1} = \partial \iota_{C_1}, N_{C_2} = \partial \iota_{C_2},$$

and according to Example 1.3.13, $\iota_{C_1}, \iota_{C_2} \in \Gamma_0(\mathbb{R}^2)$.

Because $\iota_{C_1}, \iota_{C_2} \in \Gamma_0(\mathbb{R}^2)$, Theorem 7.3.1 shows that we can use algorithm (7.11) to solve the problem with primal inclusion (7.23) and dual inclusion (7.24).

When solving that problem with algorithm (7.11), according to Lemma 7.1.1 (1),

$$\mathrm{Prox}_{\tau f}(x) = \mathrm{P}_{C_1}(x).$$

As C_1 is a circle centred at $(5, 0)$ with radius 2, we can also write

$$\mathrm{Prox}_{\tau f}(x) = (5, 0) + \mathrm{P}_{B(0;2)}(x - (5, 0)).$$

We use Lemma 7.1.1 (2) to get

$$\mathrm{Prox}_{\sigma g^*}(x) = x - \sigma \, \mathrm{P}_{C_2}\left(\frac{x}{\sigma}\right).$$

Because τ, σ must satisfy the relation $\tau\sigma < 4$, to be on the safe side, we let $\sigma = 2$, and $\tau = 3/2$. Hence, when we choose $x_0 = (5, 1), v_0 = (0, 0)$ as starting values, for any fixed k, let $\alpha_k = 2 - \frac{1}{k}$, the iterative scheme (7.11) becomes:

$$\begin{cases} y_{1n} = (5,0) + \mathrm{P}_{B(0;2)}(x_{1n} - \frac{\tau}{2} x_{2n} - (5,0)) \\ y_{2n} = (x_{2n} - \frac{\sigma}{2} x_{1n} + \sigma y_{1n}) - \sigma \, \mathrm{P}_{C_2}((x_{2n} - \frac{\sigma}{2} x_{1n} + \sigma y_{1n})/\sigma) \\ w_{1n} = \alpha_k y_{1n} - x_{1n} \\ w_{2n} = \alpha_k y_{2n} - x_{2n} \\ z_{1n} = w_{1n} - \frac{\tau}{2} w_{2n} \\ z_{2n} = w_{2n} - \frac{\sigma}{2} w_{1n} + \sigma z_{1n} \\ x_{1n+1} = x_{1n} + (z_{1n} - y_{1n}) \\ x_{2n+1} = x_{2n} + (z_{2n} - y_{2n}). \end{cases} \tag{7.25}$$

The stopping criteria of this algorithm is: Let $\epsilon = 10^{-5}$, continue running this iteration until $\|x_{1n+1} - x_{1n}\| < \epsilon$, and $\|x_{2n+1} - x_{2n}\| < \epsilon$.

7.4.2 Numerical results

We summarize our numerical implementation in three tables.

Table 7.5: Example 3: Six fixed α_k, where $\alpha_k = 2 - 1/k$, the corresponding optimal point y_1^* and y_2^*, and $\sqrt{\|y_1^*\|^2 + \|y_2^*\|^2}$, together with the case $\alpha = 2$.

α_k	y_1^*	y_2^*	$\sqrt{\|y_1^*\|^2 + \|y_2^*\|^2}$
1	(3.0053,0.1460)	(1.0160,-0.5621)	3.2251
$2 - \frac{1}{10}$	(3.0565,0.4721)	(0,-0.0852)	3.0939
$2 - \frac{1}{50}$	(3.0622,0.4949)	(0,-0.0172)	3.1020
$2 - \frac{1}{100}$	(3.0629,0.4975)	(0,-0.0086)	3.1030
$2 - \frac{1}{1000}$	(3.0634,0.4997)	1.0e-03 *(0,-0.8606)	3.1039
$2 - \frac{1}{10000}$	(3.0635,0.5000)	1.0e-04 *(0,-0.8607)	3.1040
$\alpha = 2$	(3.6259,0.6339)	(0,0)	3.6809

Then we run the algorithm (7.25) again with the same starting point and same stoping criteria, but in this time, we set $\tau = 1, \sigma = 1$. Here is the result:

Table 7.6: Example 3: Six fixed α_k, where $\alpha_k = 2 - 1/k$, the corresponding optimal point y_1^* and y_2^*, and $\sqrt{\|y_1^*\|^2 + \|y_2^*\|^2}$, together with the case $\alpha = 2$.

α_k	y_1^*	y_2^*	$\sqrt{\|y_1^*\|^2 + \|y_2^*\|^2}$
1	(3.0014,0.0740)	(0.5021,-0.3890)	3.0687
$2 - \frac{1}{10}$	(3.0546,0.4642)	(0,-0.1256)	3.0922
$2 - \frac{1}{50}$	(3.0621,0.4945)	(0,-0.0258)	3.1019
$2 - \frac{1}{100}$	(3.0628,0.4974)	(0,-0.0129)	3.1030
$2 - \frac{1}{1000}$	(3.0634,0.4997)	(0,-0.0013)	3.1039
$2 - \frac{1}{10000}$	(3.0635,0.5000)	1.0e-03 *(0,-0.1291)	3.1040
$\alpha = 2$	(3.7500,0.7500)	(0,0)	3.8243

We also tried to run this algorithm with the starting point $x_0 = (-4, -6), v_0 = (0, 0)$, and fix $\tau = 1$ and $\sigma = 1$, Here is the result:

Table 7.7: Example 3: Six fixed α_k, where $\alpha_k = 2 - 1/k$, the corresponding optimal point y_1^* and y_2^*, and $\sqrt{\|y_1^*\|^2 + \|y_2^*\|^2}$, together with the case $\alpha = 2$.

α_k	y_1^*	y_2^*	$\sqrt{\|y_1^*\|^2 + \|y_2^*\|^2}$
1	(3.0014,0.0740)	(0.5021,-0.3890)	3.0687
$2 - \frac{1}{10}$	(3.0546,0.4642)	(0,-0.1256)	3.0922
$2 - \frac{1}{50}$	(3.0621,0.4945)	(0,-0.0258)	3.1019
$2 - \frac{1}{100}$	(3.0628,0.4974)	(0,-0.0129)	3.1030
$2 - \frac{1}{1000}$	(3.0634,0.4997)	(0,-0.0013)	3.1039
$2 - \frac{1}{10000}$	(3.0635,0.5000)	1.0e-03 *(0,-0.1291)	3.1040
$\alpha = 2$	(3.3945,0.6448)	(0,0)	3.4552

According to the numerical result above, we can get the following conclusions:

(1) If we let $y^* = (3.0635, 0.5000)$ and $w^* = (0, 0)$,tables 7.5, 7.6, and 7.7 all show that when $\alpha_k \to 2$, we have the smallest norm primal-dual solution (y^*, w^*). y^* solves the primal and w^* solves the dual.

(2) Once we fix the value of k with fixed τ and σ, the result we get by using α-Douglas-Rachford algorithm does not change if we change its starting point.

(3) In three tables 7.5, 7.6, and 7.7, $\alpha = 2$ gives different y_1^* because

$$0 \in N_{C_1}(x) + N_{C_2}(x)$$

has multiple solutions.

Remark 7.3. In Bot and Hendrich's paper (see [6]), they have Douglas-Rachford algorithm applications on denoising problems in image processing. We believe similar work can be done for the α-Douglas-Rachford algorithm.

Chapter 8

Conclusions and future work

In this thesis, the Douglas-Rachford algorithm is studied. This is an algorithm for solving the split problem: find $x \in \mathbb{R}^m$ such that

$$0 \in Ax + Bx,$$

where A and B are maximally monotone operators. The Douglas-Rachford algorithm can be written as

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = J_B x_n \\ z_n = J_A(2y_n - x_n) \\ x_{n+1} = x_n + (z_n - y_n). \end{cases}$$

In this thesis, I built a new algorithm based on Douglas-Rachford algorithm and called it the α-Douglas-Rachford algorithm. This algorithm solves the split problem: find $x \in \mathbb{R}^m$ such that

$$0 \in Ax + Bx + (2 - \alpha)x$$

where $\alpha \in [1, 2)$, A and B are maximally monotone operators. The new algorithm can be written as

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = J_B x_n \\ z_n = J_A(\alpha y_n - x_n) \\ x_{n+1} = x_n + (z_n - y_n). \end{cases}$$

I proved that the α-Douglas-Rachford algorithm has very similar properties to the Douglas-Rachford algorithm, and also showed the connection between those two algorithms when $\alpha \to 2$. One distinctive feature of α-Douglas-Rachford algorithm is that it can be used to find the least norm solution.

Possible future work:

(1) Is the α-Douglas-Rachford algorithm error torlerant?

(2) In the primal-dual problems, what are the optimal parameters τ, σ to implement the α-Douglas-Rachford algorithms?

(3) If we change the space from $\mathbb{R}^m$ to a more general space, like $\mathcal{H}$, a general Hilbert space, does the α-Douglas-Rachford algorithm have the same results and properties?

(4) Comparing with the Douglas-Rachford algorithm, does the α-Douglas-Rachford algorithm converge faster?

(5) More numerical experiments on the α-Douglas-Rachford algorithm for higher dimensions and practical applications are required.

Bibliography

[1] ATTOUCH, H., AND THÉRA, M. A general duality principle for the sum of two operators. *Journal of Convex Analysis* 3, (1996), 1-24. → pages 35, 82

[2] BAUSCHKE, H. H. A note on the paper by Eckstein and Svaiter on "General projective splitting methods for sums of maximal monotone operators" . *SIAM Journal on Control and Optimtimization* 48, 4 (2009), 2513-2515. → pages 49

[3] BAUSCHKE, H. H., AND BORWEIN, J. M. Dykstra's alternating projection algorithm for two sets. *Journal of Approximation Theory* 79, 3 (1994), 418-443. → pages 78

[4] BAUSCHKE, H. H., AND COMBETTES, P. L. *Convex analysis and monotone operator theory in Hilbert spaces*. Springer, 2017. → pages 4, 5, 8, 14, 15, 16, 18, 19, 22, 23, 25, 26, 27, 28, 29, 30, 32, 40, 49, 53, 73

[5] BAUSCHKE, H. H., AND MOURSI, W. M. On the Douglas-Rachford algorithm. *Mathematical Programming*, 164, 1-2 (2017), 263-284. → pages 32

[6] BOT, R. I., AND HENDRICH, C. A Douglas-Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators. *SIAM Journal on Optimization* 23, 4 (2013), 2541-2565. → pages 31, 32, 34, 35, 100

[7] COMBETTES, P. L. Solving monotone inclusions via compositions of nonexpansive averaged operators. *Optimization* 53, 5-6 (2004), 475-504. → pages 32

[8] COMBETTES, P. L. Iterative construction of the resolvent of a sum of maximal monotone operators. *Journal of Convex Analysis* 16, 4 (2009), 727-748. → pages 32

[9] DOUGLAS, J., AND RACHFORD, H. H. On the numerical solution of heat conduction problems in two and three space variables. *Transactions of the American Mathematical Society* 82, 2 (1956), 421-439. → pages 31

[10] DUCHI, J. Introductory lectures on stochastic optimization. *Manuscript, Stanford University*, (2016). → pages 26

[11] ECKSTEIN, J., AND BERTSEKAS, D. P. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming* 55, 1 (1992), 293-318. → pages 31

[12] KREYSZIG, E. *Introductory functional analysis with applications (Vol. 1)*. Wiley New York, 1989. → pages 7, 11, 13

[13] LIEUTAUD, J. Approximations d'opérateurs monotones par des méthodes de splitting. *Doctoral thesis, University of Paris*, 1969. → pages 31

[14] LIONS, P. L., AND MERCIER, B. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis* 16, 6 (1979), 964-979. → pages 31, 32

[15] MORDUKHOVICH, B. S., AND NAM, N. M. *An easy path to convex analysis and applications*. Synthesis Lectures on Mathematics and Statistics, 6:1-218, 2013. → pages 24

[16] SVAITER, B. F. On weak convergence of the Douglas-Rachford method. *SIAM Journal on Control and Optimization* 49, 1 (2011), 280-287. → pages 32

[17] WANG, X. Self-dual regularization of monotone operators via the resolvent average. *SIAM Journal on Optimization* 21, 2 (2011), 438-462. → pages 78