Instituto Nacional de Matemática Pura e Aplicada

Ph.D. Thesis by: Reza Arefidamghani

Circumcentered-Reflection methods for the Convex Feasibility problem and the Common Fixed-Point problem for firmly nonexpansive operators

---

Supervisor:
Alfredo Noel Iusem

Co-Supervisor:
Roger Behling

April 20, 2022

Conselho Nacional de Desenvolvimento Científico e Tecnológico

Dedicated to Mahdi with all my heart.

Acknowledgements

First and foremost I am extremely grateful to my supervisor, Prof. Alfredo Noel Iusem, for your invaluable advice, continuous support, and patience during my Ph.D. study. I am extremely grateful to you for being beside me and pushing me forward through solving problems and writing papers and completing my thesis. Thanks to you that always were available even on weekends and holy days. Quarantine (COVID-19) was tough for everyone, but talking with you (by voice call) about mathematics made it tolerable. I learned a lot from you and I will remember and use them in my academic and personal life.

I would also like to thank Dr. Roger Behling, Dr. Yunier Bello-Cruz, Dr. Luiz-Rafael Santos, for their technical support in my study. This is my honor to work and be co-authored with you, great mathematicians. Special thanks to you dear Rafa, that always being supportive, positive, and available any time that I asked for your help. You taught me how to be an academic teammate.

I would like to thank all the members of IMPA. It is their kind help and support that have made my study and life in Brazil a wonderful time. In special, the Department of Ensino.

Finally, I would like to express my gratitude to my family; My deepest thanks to my parents Karim and Maryam, for their encouragement and love, despite the distance. To my sisters Rezvan and Sedigheh and also tia Tânia, my many thanks for their friendship and constant support. I leave my most sincere gratitude to my fiancee, Oriana, for her infinite patience and for always being there for me. Without your tremendous understanding and encouragement over the past few years, it would be impossible for me to complete my study.

" Last but not least, I wanna thank me
I wanna thank me for, believing in me
I wanna thank me for, doing all this hard work
I wanna thank me for, having no days off
I wanna thank me for, for never quitting
I wanna thank me for, just being me at all times"

Snoop Dogg

Contents

List of Figures

List of Tables

Index of Notations

- CFP : Convex Feasibility Problem;
- FPP : Fixed-Point Problem;
- SPM : Sequential Projection Method;
- PPM : Parallel Projection Method;
- CRM : Circumcentered-Reflection Method;
- MAP : Method of Alternating Projections;
- CARM : Circumcentered Approximate-Reflection Method;
- MAAP : Method of Alternating Approximate Projections;
- $P_M(x) = \operatorname{argmin}_{y \in M} \|x - y\|$: Orthogonal projection onto a closed convex M;
- $\operatorname{dist}(x, M)$: Euclidean distance from a point $x \in \mathbb{R}^n$ to a closed convex $M \subset \mathbb{R}^n$;
- $B(x, \rho)$: Ball centered at x with radius ρ;
- K : Closed convex set in $\mathbb{R}^n$;
- U : Affine manifold in $\mathbb{R}^n$;
- P_U : Orthogonal projection on affine set U;
- P_K : Orthogonal projection on closed convex set K;
- $\widehat{P} := P_{K_m} \circ \cdots \circ P_{K_1}$: Sequential Projection Method operator;
- $\overline{P} := \frac{1}{m} \sum_{i=1}^{m} P_{K_i}$: Parallel Projection Method operator;
- Id : Identity operator in $\mathbb{R}^n$.
- $R_U = 2P_U - \operatorname{Id}$: Reflection operator defined by P_U;
- $R_K = 2P_K - \operatorname{Id}$: Reflection operator defined by P_K;
- $S : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$: The point to set separating operator;
- P^S : Outer-approximate projection.*i.e.*, the orthogonal projection onto superset $S(x) \supset K$;
- $R^S = 2P^S - \operatorname{Id}$: Reflection operator defined by P^S;
- $F(T) = \operatorname{Fix}(T) = \{x \in \mathbb{R}^n : T(x) = x\}$: the fixed-point set of the operator T;
- $\operatorname{Fix}(T_1, \ldots, T_m)$: The common fixed-point set of operators $T_1, \ldots, T_m$;

- $R = 2T - \mathrm{Id}$: Reflection operator defined by T;
- $\mathrm{circ}(x, y, z)$: Circumcenter of three points x, y and z, in $\mathbb{R}^n$.
- $C(\cdot) = \mathrm{circ}\,(\,\cdot\,, R_K(\cdot), R_U(R_K(\cdot)))$: Circumcentered-reflection operator;
- $C^S(\cdot) = \mathrm{circ}\left(\,\cdot\,, R^S(\cdot), R_U(R^S(\cdot))\right)$: Circumcentered approximate-reflection operator;
- $C(\cdot) = \mathrm{circ}\,(\,\cdot\,, R(\cdot), R_U(R(\cdot)))$: Circumcentered-reflection operator induced by a firmly nonexpansive operator;
- $D(\cdot) = P_U(P_K(\cdot))$: Alternating projections operator;
- $D^S(\cdot) = P_U(P^S(\cdot))$: Alternating approximate projections operator;
- $D(\cdot) = P_U(T(\cdot))$: Alternating projections operator induced by a firmly nonexpansive operator;
- $\left\{z^k\right\}_{k\in\mathbb{N}}$: The sequence generated by MAP or MAAP;
- $\bar{z} \in K \cap U$; Limit point of the sequence generated by CRM or CARM;
- $\left\{x^k\right\}_{k\in\mathbb{N}}$:The sequence generated by CRM or CARM;
- $\bar{x} \in K \cap U$: Limit point of the sequence generated by CRM or CARM ;
- EB : Global error bound which ensures linear convergence of MAP and CRM applied to CFP; there exists $\omega > 0$ such that $\mathrm{dist}(x, K) \geq \omega\,\mathrm{dist}(x, K \cap U)$ for all $x \in U$.
- LEB : Local error bound assumption which ensures linear convergence of MAAP and CARM; there exists a set $V \subset \mathbb{R}^n$, and a scalar $\omega > 0$, such that, for all $x \in U \cap V$ $d(x, S(x)) \geq \omega\ d(x, K \cap U)$.
- LEB1 : Local error bound assumption which is a local version of EB: there exist $\rho, \bar{\omega} > 0$ such that $\mathrm{dist}(x, K) \geq \bar{\omega}\,\mathrm{dist}(x, K \cap U)$ for all $x \in U \cap B(x^*, \rho)$.
- EB1; Global error bound assumption which ensures linear convergence of MAP and CRM applied to FPP; there exists $\omega > 0$ such that $\|x - T(x)\| \geq \omega\,\mathrm{dist}(x, \mathrm{Fix}(T, P_U))$ for all $x \in U$.

Abstract

We study the convergence rate of the Circumcentered-Reflection Methods (CRM) for solving the Convex Feasibility Problems (CFP) and compare it with the Method of Alternating Projections (MAP). Under an error bound assumption, we prove that both methods applied to CFP, converge linearly, with asymptotic constants depending on a parameter of the error bound, and that the one derived for CRM is strictly better than the one for MAP. Two rather generic families of examples for which CRM is faster than MAP are presented.

We introduce the circumcentered approximate-reflection method (CARM), which uses outer-approximate projections instead of exact ones. The appeal of CARM is that, in rather general situations, the cost of computing these approximate projections is much lower than the cost of computing exact projections. We prove convergence of CARM and linear convergence under an error bound condition. We also present successful theoretical and numerical comparisons of CARM to the original CRM, to the classical MAP and to a correspondent outer-approximate version of MAP, referred to as MAAP. Along with our results and numerical experiments, we present a couple of illustrative examples.

Last but not least, we apply CRM to solving FPP, consisting of finding a common fixed-point of firmly nonexpansive operators. We prove that CRM is globally convergent to a common fixed-point (supposing that at least one exists). We also establish linear convergence of the sequence generated by CRM applied to FPP, under a not too demanding error bound assumption, and provide an estimate of the asymptotic constant. We provide solid numerical evidence of the superiority of CRM when compared to the classical Parallel Projections Method (PPM) for solving FPP.

Keywords : Convex feasibility problem · fixed-point problem · method of alternating projections · circumcentered-reflection method · convergence rate.

Resumo

Estudamos a taxa de convergência dos Métodos de Reflexão Circuncentrada (CRM) para resolver os Problemas de Viabilidade Convexa (CFP) e comparamos com o Método de Projeções Alternadas (MAP). Sob uma suposição de limite de erro, provamos que ambos os métodos aplicados ao CFP convergem linearmente, com constantes assintóticas dependendo de um parâmetro do limite de erro, e que a constante assintótica de CRM é estritamente melhor que a de MAP. São apresentadas duas famílias bastante genéricas de exemplos para os quais o CRM é mais rápido que o MAP.
Introduzimos o método de reflexão aproximada circuncêntrica (CARM), que usa projeções aproximadas externas em vez de projeções exatas. O apelo do CARM é que, em situações gerais diversas, o custo de calcular essas projeções aproximadas é muito menor do que o custo de calcular projeções exatas. Provamos a convergência de CARM e convergência linear sob uma condição de limite de erro. Apresentamos também comparações teóricas e numéricas bem-sucedidas do CARM com o CRM original, com o MAP clássico e com uma versão aproximada do MAP correspondente, denominada MAAP. Junto com nossos resultados e experimentos numéricos, apresentamos alguns exemplos ilustrativos.
Por último, mas não menos importante, aplicamos o CRM para resolver FPP, consistindo em encontrar um ponto fixo comum de operadores firmemente não expansivos. Provamos que o CRM é globalmente convergente para um ponto fixo (supondo que exista pelo menos um). Também estabelecemos convergência linear da sequência gerada pelo CRM aplicada ao FPP, sob uma suposição de limite de erro não tão exigente e fornecendo uma estimativa da constante assintótica. Fornecemos evidências numéricas sólidas da superioridade do CRM quando comparado ao clássico Método de Projeções Paralelas (PPM) para resolver FPP.

Palavras chave : Problema de viabilidade convexa · problema de ponto fixo · método de projeções alternadas · método de reflexão circuncêntrica · taxa de convergência.

Chapter 1

Introduction

1.1 The Sequential Projection Method and the Parallel Projection Method for the Convex Feasibility Problem

1.1.1 Convex Feasibility Problem

A very common problem in diverse areas of mathematics and physical sciences consists of finding a point in the intersection of convex sets. This problem is referred to as the Convex Feasibility Problem (CFP, from now on); its precise mathematical formulation is as follows. Given closed and convex sets $K_1, \ldots, K_m \subset \mathbb{R}^n$ with $\cap_{i=1}^m K_i \neq \varnothing$, the convex feasibility problem (CFP) consists of:

$$\text{finding } \; \bar{x} \in \bigcap_{i=1}^{m} K_i. \tag{1.1.1}$$

A CFP is said to be *consistent* when it has solution (*i.e.,* when $\cap_{i=1}^m K_i \neq \varnothing$,), otherwise it is said to be *inconsistent.* This seemingly simple problem provides a modeling framework with great flexibility and power. For the purpose of numerical schemes however, devising computationally tractable formulations is often a nontrivial task and some creativity is required. Projection methods are a family of iterative algorithms for solving CFP (1.1.1).
The idea is to involve the projections onto each set K_i (respectively, onto a superset of K_i) to generate a sequence of points that is supposed to converge to a solution of CFP. This is the approach we will investigate.

Definition 1.1.1. Given a nonempty subset $M \subset \mathbb{R}^n$, the orthogonal projection mapping onto M is the possibly set-valued operator, $P_M : \mathbb{R}^n \to M$, defined at each $x \in \mathbb{R}^n$ by

$$P_M(x) = \operatorname*{argmin}_{y \in M} \|x - y\| .$$

It is well known that the orthogonal projection P_M has nonempty value (*i.e.,* $P_M(x) \neq \varnothing$ for all $x \in \mathbb{R}^n$) and it is a singleton when M is closed and convex, respectively. In this case, the orthogonal projection P_M is a single-valued mapping that sends each point $x \in \mathbb{R}^n$ to its unique nearest point in M.

1.1.2 The Sequential Projection Method and the Parallel Projection Method

Two very well-known methods for CFP are the *Sequential Projection Method* (SPM) and the *Parallel Projection Method* (PPM), which can be traced back to [45, 29] respectively, and are defined as follows.
Let $P_{K_i} : \mathbb{R}^n \to K_i$ denote the orthogonal projection onto K_i. Consider the operators $\widehat{P}, \overline{P} : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$\widehat{P} := P_{K_m} \circ \cdots \circ P_{K_1}, \text{ and } \overline{P} := \frac{1}{m}\sum_{i=1}^{m} P_{K_i}. \tag{1.1.2}$$

Starting from an arbitrary $z \in \mathbb{R}^n$, SPM and PPM generate sequences $\left\{\hat{x}^k\right\}_{k\in\mathbb{N}}$ and $\left\{\bar{x}^k\right\}_{k\in\mathbb{N}}$ given by $\hat{x}^{k+1} = \widehat{P}(\hat{x}^k)$, $\bar{x}^{k+1} = \overline{P}(\bar{x}^k)$, respectively, where $\bar{x}^0 = \hat{x}^0 = z$. When $\cap_{i=1}^m K_i \neq \varnothing$, the sequences generated by both methods are known to be globally convergent to points belonging to $\cap_{i=1}^m K_i$, *i.e.*, to solve CFP. Under suitable assumptions, both methods have interesting convergence properties also in the infeasible case, *i.e.*, when $\cap_{i=1}^m K_i = \varnothing$, but we will not deal with this case. See [7] for an in-depth study of these and other projection methods for CFP.
A simple example to visualize how SPM and PPM generate the iterations for solving the CFP (1.1.1), in the case of two convex sets, is illustrated in Figure 1.1.

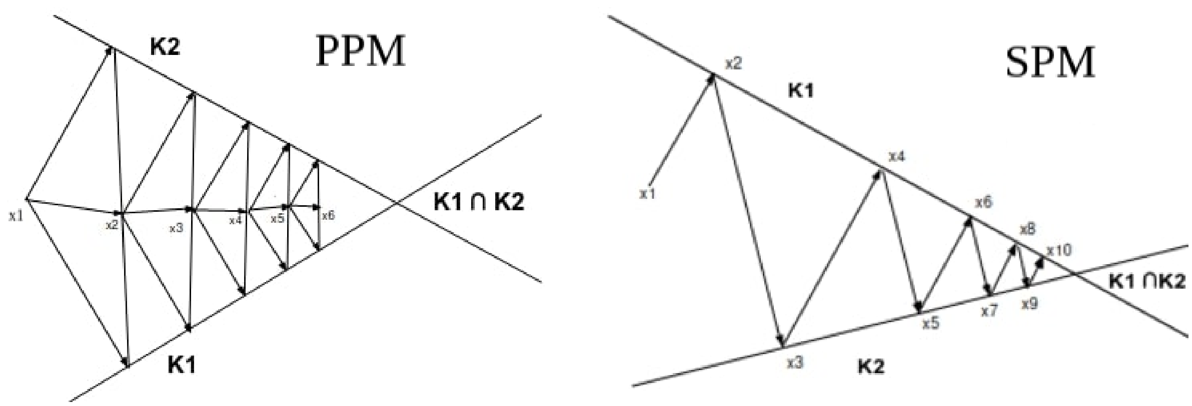

Figure 1.1: SPM-PPM iterations, in the case of two sets

1.1.3 Pierra's product space reformulation

An interesting relation between SPM and PPM was found by Pierra in [57]. Consider the two following closed and convex sets:

$$\mathbf{K} := K_1 \times \cdots \times K_m \subset \mathbb{R}^{nm}, \text{ and } \mathbf{U} := \{(x, \ldots, x) : x \in \mathbb{R}^m\} \subset \mathbb{R}^{nm}. \tag{1.1.3}$$

Note that it is straightforward to prove that $\mathbf{U}$ is a subspace of $\mathbb{R}^{nm}$, called the Diagonal subspace. Moreover, given m arbitrary vectors x^i in R^n , with $i = 1, \ldots, m$, we can construct an arbitrary point in R^{nm} of the form $\mathbf{x} = (x^1, x^2, \ldots, x^m) \in \mathbb{R}^{nm}$ and its projection onto $\mathbf{U}$ is given by

$$P_{\mathbf{U}}(\mathbf{x}) = \frac{1}{m}\left(\sum_{i=1}^{m} x^i, \sum_{i=1}^{m} x^i, \ \ldots \ , \sum_{i=1}^{m} x^i\right). \tag{1.1.4}$$

As for the orthogonal projection of $\mathbf{x} = (x^1, x^2, \ldots, x^m) \in \mathbb{R}^{nm}$ onto $\mathbf{K}$, we have

$$P_{\mathbf{K}}(\mathbf{x}) = \left(P_{K_1}(x^1), P_{K_2}(x^2), \ \ldots \ , P_{K_m}(x^m)\right).$$

Apply SPM to the sets $\mathbf{K}, \mathbf{U}$ in the product space $\mathbb{R}^{nm}$, *i.e.*, take $\mathbf{x}^{k+1} = P_{\mathbf{U}}\left(P_{\mathbf{K}}((\mathbf{x}^k)\right)$ starting from $\mathbf{x}^0 \in \mathbf{U}$. Clearly, $\mathbf{x}^k$ belongs to $\mathbf{U}$ for all $k \in \mathbb{N}$, so that we may write $\mathbf{x}^k = (x^k, \ldots, x^k)$ with $x^k \in \mathbb{R}^n$. It was proved in [57] that $x^{k+1} = \overline{P}(x^k)$, *i.e.*, a step of SPM applied to two convex sets in the product space $\mathbb{R}^{nm}$ is equivalent to a step of PPM in the original space $\mathbb{R}^n$. In fact, $\bar{x} \in \bigcap_{i=1}^{m} K_i$ if and only if $\mathbf{x} = (\bar{x}, \ldots, \bar{x}) \in \mathbf{K} \cap \mathbf{U}$. Thus, if we can solve any intersection problem featuring a closed convex set and an affine subspace, we are able to solve the general CFP. Let us proceed in this direction by considering a closed convex set $K \subset \mathbb{R}^n$ and an affine subspace $U \subset \mathbb{R}^n$ with nonempty intersection. From now on, the CFP we are going to focus on is the one of tracking a point in $K \cap U$. Thus, SPM with just two sets plays a sort of special role and, therefore, carries a name of its own, namely *Method of Alternating Projections* (MAP).

The Method of Alternating Projections is a very simple algorithm for computing a point in the intersection of two convex sets, using a sequence of projections onto the sets. Like a gradient or subgradient method, alternating projections can be slow, but the method can be useful when one has some efficient method, such as an analytical formula, for carrying out the projections.

We consider now two operators A and $B : \mathbb{R}^n \to \mathbb{R}^n$ and define $D = A \circ B$. Under adequate assumptions, the sequence $\{x^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$ defined by

$$x^{k+1} = D(x^k) = A(B(x^k)) \tag{1.1.5}$$

is expected to converge to a common fixed point of A and B.

If the operators A and B are the orthogonal projections onto the convex sets U and K, that is, $A = P_U, B = P_K$, then the sequence generated by (1.1.5) recovers MAP. Moreover, the set of common fixed points of A and B in this case is precisely $K \cap U$, and (1.1.5) becomes

$$x^{k+1} = D(x^k) = P_U(P_K(x^k)) \tag{1.1.6}$$

which converges to a point in $K \cap U$ for any starting point in $\mathbb{R}^n$, provided that $K \cap U \neq \varnothing$ (see, [7, 28]). Moreover, MAP is known to be linearly convergent in several special situations, *e.g.*, when both K and U are affine manifolds (see [48]) or when $K \cap U$ has nonempty interior (see [6]).

1.2 The Circumcentered-Reflection Method for solving CFP

The Circumcentered-Reflection Method (CRM) is the main object of study in this thesis, and we describe it next. We begin the formal definition of the circumcenter of three points in $\mathbb{R}^n$.

Definition 1.2.1. Given $x, y, z \in \mathbb{R}^n$, their circumcenter $\text{circ}(x, y, z) \in \mathbb{R}^n$ is a point satisfying

(i) $\|\text{circ}(x, y, z) - x\| = \|\text{circ}(x, y, z) - y\| = \|\text{circ}(x, y, z) - z\|$ and

(ii) $\text{circ}(x, y, z) \in \text{aff}\{x, y, z\} := \{w \in \mathbb{R}^n \mid w = x + \alpha(y - x) + \beta(z - x),\ \alpha, \beta \in \mathbb{R}\}$.

The point $\text{circ}(x, y, z)$ is well and uniquely defined if the cardinality of the set $\{x, y, z\}$ is one or two. In the case in which the three points are all distinct, $\text{circ}(x, y, z)$ is well and uniquely defined only if x, y and z are not collinear.
Other equivalent definitions for finding $\text{circ}(x, y, z)$ are given in [10] and [17]: the circumcenter of three distinct point (x, y, z) is the intersection point of the perpendicular bisectors of the triangle of vertices (x, y, z) (we recall that a perpendicular bisector is a line that forms a right angle with a segment and cuts the segment in half); the circumcenter can also be considered as the center of the circle in the affine hull of the three points that passes through of all of them (see Figure 1.2, is taken from [18].). For more general notions, definitions and results on circumcenters see [10, 13, 14, 17, 18].

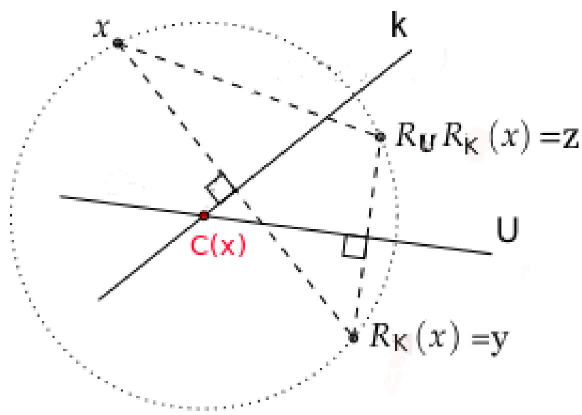

Figure 1.2: Circumcenter on the affine subspace aff{x, y, z}.

We consider now two operators A and $B : \mathbb{R}^n \to \mathbb{R}^n$ and define the reflection operators $A^R, B^R : \mathbb{R}^n \to \mathbb{R}^n$ as $A^R = 2A - \text{Id}$, $B^R = 2B - \text{Id}$, where Id stands for the identity operator in $\mathbb{R}^n$. The CRM operator $C : \mathbb{R}^n \to \mathbb{R}^n$ is defined as

$$C(x) = \text{circ}(x, B^R(x), A^R(B^R(x))), \tag{1.2.1}$$

i.e., the circumcenter of the three points $x, B^R(x), A^R(B^R(x))$. The CRM sequence $\{x^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$, starting at some $x^0 \in \mathbb{R}^n$, is then defined as

$$x^{k+1} = C(x^k) = \text{circ}(x^k, B^R(x^k), A^R(B^R(x^k))). \tag{1.2.2}$$

In particular, if, $A = P_U$ and $B = P_K$, then respectively, $R_U = 2P_U - \text{Id}$, $R_K = 2P_K - \text{Id}$, and the CRM operator $C : \mathbb{R}^n \to \mathbb{R}^n$ is defined as $C(x) = \text{circ}(x, R_K(x), R_U(R_K(x)))$, *i.e.*, the circumcenter of the three points $x, R_K(x), R_U(R_K(x))$. An illustration of CRM iteration, is given in Figure 1.2.

The CRM sequence $\{x^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$, starting at some $x^0 \in U$, is then defined as

$$x^{k+1} = C(x^k) = \operatorname{circ}(x^k, R_K(x^k), R_U(R_K(x^k))). \tag{1.2.3}$$

It converges to a point in $K \cap U$ as long as the initial point lies in U([17]). We anticipate that in our approach, it will be essential to initialize CRM in U. Indeed, it is known that if $x^0 \in U$, then $x^k \in U$ for all k (see [17]).
CRM is in its original form (1.2.3) faces some difficulties when dealing with general convex feasibility problems. Indeed, in the first paper introducing CRM [18], it was pointed out that if both convex sets fail to be affine, then the method could possibly diverge or simply be undefined. There is now an actual example featuring two intersecting balls for which CRM stalls or diverges depending on the initial point [2] (see Figure 1.3, is taken from [2]). These apparent drawbacks are genuinely overcome in [17]. The key is to reformulate the problem in the product space, following [57], as explained in Subsection 1.1.3, in which case one of convex sets in known to be an affine manifold.

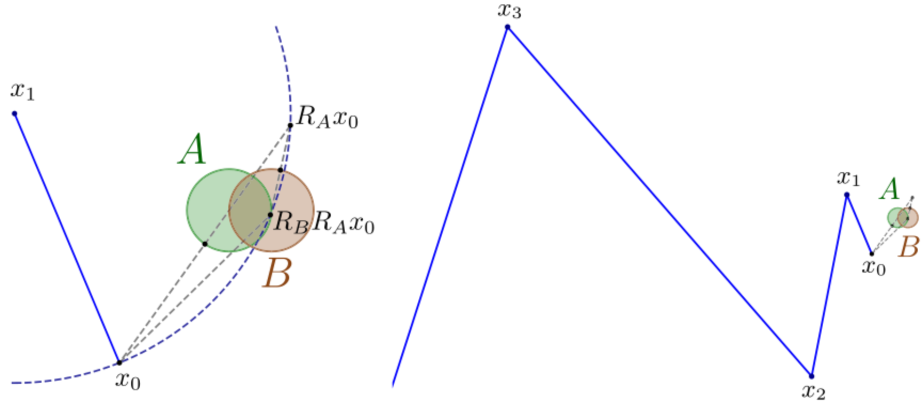

Figure 1.3: Failure of the Circumcentered-Reflection Method when applied to two balls of the same radius $A, B \in \mathbb{R}^2$: The CRM sequence $\{x^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$, diverges. The left figure shows the construction of the first iteration from the right figure in more detail.

The ability of CRM for finding a point in the intersection of a closed convex set K and an affine manifold U is simply impressive in comparison to the classical MAP. Moreover, a CRM iteration is always better in terms of distance to the solution set than MAP iteration (see [17]). Also, it should be noted that the computational effort for calculating a CRM step is essentially the same as the one for computing a MAP step. This is due to the fact that in each iteration, we need to compute the same number of projections for both methods; the additional computations of CRM over MAP reduce to the trivial determination of the reflections and the solution of an elementary system of two linear equations in two real variables.
Now we focus on the alleged acceleration effect of CRM with respect to MAP. There is abundant numerical evidence of this effect (see [14, 17, 18, 32]); in this thesis, we will present some analytical evidence, which strengthens the results from [17].
We will denote the Euclidean distance from a point $x \in \mathbb{R}^n$ to a set $K \subset \mathbb{R}^n$ as $\operatorname{dist}(x, K)$. A first result in the analytical study of the acceleration effect of CRM over MAP was derived

in [17, Theorem 2], where it was proved that $\operatorname{dist}(C(x), K \cap U) \leq \operatorname{dist}(D(x), K \cap U)$ for all $x \in U$, meaning that the point obtained after a CRM step is closer to (or at least no farther from) $K \cap U$ than the one obtained after a MAP step from the same point. This local (or myopic) acceleration does not imply immediately that the CRM sequence converges faster than the MAP one. In order to show global acceleration, we will focus on special situations where the convergence rate of the MAP can be precisely established. One such situation occurs when a certain so-called *error bound* (EB) holds. EB is defined as:

EB) There exists $\bar{\omega} > 0$ such that $\operatorname{dist}(x, K) \geq \bar{\omega} \operatorname{dist}(x, K \cap U)$ for all $x \in U$.

This error bound resembles the regularity conditions presented in [6, 7, 19]. We prove that in this case both the MAP and the CRM sequences converge linearly, with asymptotic constants bounded by $\sqrt{1-\bar{\omega}^2}$ for MAP, and by the strictly better bound $\sqrt{(1-\bar{\omega}^2)/(1+\bar{\omega}^2)}$ for CRM, thus showing that under EB, CRM is in principle faster than MAP. For the case of MAP, linear convergence under the error bound condition with this asymptotic constant is already known (see, for instance, [6]).
Next, we analyze two classes of fairly generic examples. In the first one, the angle between the convex sets approaches zero near the intersection, so that the MAP sequence converges sublinearly, but CRM still enjoys linear convergence. In the second class of examples, the angle between the sets does not vanish and MAP exhibits its standard behavior, *i.e.*, it converges linearly, yet, perhaps surprisingly, CRM attains superlinear convergence.
These results firmly corroborate the already established numerical evidence in [17] of the superiority of CRM over MAP.
We emphasize that in the cases above, MAP exhibits its usual behavior, *i.e.*, linear convergence the examples of the first family were somewhat special because, roughly speaking, the angle between K and U goes to 0 near the intersection. On the other hand, the superlinear convergence of CRM is quite remarkable. As discussed above, the additional computational cost of CRM over MAP is negligeable. Now MAP is a typical first-order method (projections disregard the curvature of the sets), and thus its convergence is generically no better than linear. We show the CRM acceleration, in a rather large class of instances, improves this linear convergence to superlinear.
We conjecture that CRM enjoys superlinear convergence whenever U intersect the interior of K. Our results, firmly support this conjecture.

1.3 Approximate Projections, Approximate MAP and Approximate CRM

CRM iterates by computing a circumcenter upon a composition of reflections with respect to convex sets. Remind that reflections are based on exact projections. Computing exact projections onto general convex sets can be, context depending, too demanding in comparison to solving the given CFP itself. Bearing this in mind, we replace the exact projections onto closed and convex sets by *outer-approximate projections*. These approximate projections still enjoy some of the properties of the exact ones, having the advantage of being potentially more tractable. For instance, they cover the subgradient projections of Fukushima [38]. We introduce approximate versions of MAP and CRM for solving CFP, which we call Method of Approximate Alternating Projections (MAAP, from now on) and Circumcentred Approximate Reflection Method (CARM, from now on). The MAAP and CARM iterations are computed by (1.1.5) and (1.2.2) with A being the exact projection onto U and B an approximate projection onto K. The approximation consists of replacing at each iteration the set K by a larger set separating the current iterate from K. With this purpose, we introduce the separating operator needed for the approximate versions of MAP and CRM, namely MAAP and CARM, in the following way.

Definition 1.3.1. Given a closed and convex set $K \subset \mathbb{R}^n$, a *separating operator* for K is a point-to-set mapping $S : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ satisfying:

A1) $S(x)$ is closed and convex for all $x \in \mathbb{R}^n$.

A2) $K \subset S(x)$ for all $x \in \mathbb{R}^n$.

A3) If a sequence $\{z^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ converges to $z^* \in \mathbb{R}^n$ and $\lim_{k\to\infty} \operatorname{dist}(z^k, S(z^k)) = 0$ then $z^* \in K$.

Several notions of separating operators have been introduced in the literature; see, *e.g.*, [25, Section 2.1.13] and references therein. Our definition is a point-to-set version of the separating operators in [24, Definition 2.1]. It encompasses not only hyperplane-based separators as the ones in the seminal work by Fukushima [38], but also more general situations. We present two particular choices of the separating operator S, which induce easily computable projections. This separating scheme is rather general, and for a large family of convex sets, the separating set is a half-space, or a Cartesian product of half-spaces, in which cases all the involved projections have a very low computational cost. One could fear that this significant reduction in the computational cost per iteration could be nullified by a substantial slowing down of the process as a whole, through a deterioration of the convergence speed. However, we show that this is not necessarily the case. Indeed, under some adequate assumptions, we show that MAAP and CARM enjoy linear convergence rates.
We will prove some properties of approximate methods which follow quite closely the corresponding results for the exact algorithms, the difference consisting in the replacement of the set K by the separating set $S(x)$. However, some care is needed, because K is fixed, while $S(x)$ changes along the algorithm, so that we present the complete analysis for the approximate algorithms MAAP and CARM. The crux of the convergence analysis of CRM, performed in [17], is the remarkable observation that for $x \in U \setminus K$, $C(x)$ is indeed the projection of x onto a half-space, separating x from $K \cap U$. We successfully extend this result to CARM.
We prove the convergence of the MAAP and CARM sequences, using the well known Fejér monotonicity argument.

Definition 1.3.2. A sequence $\left\{x^k\right\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ is Fejér monotone with respect to nonempty closed and convex set $M \subset \mathbb{R}^n$ when

$$\left\|x^{k+1} - y\right\| \leq \left\|x^k - y\right\| \qquad \forall y \in M,\ k \in \mathbb{N}. \tag{1.3.1}$$

We consider a rather standard and not too demanding local error bound, denoted as (LEB) and defined as:

LEB: There exists a set $V \subset \mathbb{R}^n$, and a scalar $\omega > 0$, such that

$$d(x, S(x)) \geq \omega \; d(x, K \cap U), \quad \text{for all } x \in U \cap V,$$

Under LEB, MAAP and CARM enjoy linear convergence rates, with the linear rate of CARM being strictly better than MAAP. We show that, under the LEB condition, both the MAAP and the CARM sequences converge linearly, with asymptotic constants bounded by $\sqrt{1-\omega^2}$ for MAAP, and by the strictly better bound $\sqrt{(1-\omega^2)/(1+\omega^2)}$ for CARM.

We remark that, the set V in the LEB condition, could be expected to be a neighborhood of the limit $\bar{x}$ of the sequence, but in our analysis it will have a slightly more complicated structure: a ball centered at $\bar{x}$ minus a certain "slice ". We will be able to prove that such a set V contains the tail of the sequences generated by CARM and MAAP.

We also analyze two classes of somewhat generic examples for which CARM is faster than MAAP. In the first one, the angle between the convex sets approaches zero near the intersection, so that the MAAP sequence converges sublinearly, but CARM still enjoys linear convergence.

In the second family, K will still be the epigraph of a convex function f, but U will not be a supporting hyperplane of K; rather it will intersect the interior of K. In this case, under not too demanding assumptions on f, the MAAP sequence converges linearly (we will give an explicit bound of its asymptotic constant), while CARM converges superlinearly.

We emphasize once again that, under LEB, we prove that the approximate algorithms MAAP and CARM achieve the same rate of convergence, with exactly the same explicit bound of asymptotic constant, as the exact methods MAP and CRM, but the approximate projections in MAAP and CARM are, generically, much cheaper than the exact projections used in MAP and CRM.

Our numerical experiments confirm these statements, and more than that, they show CARM outperforming MAP, CRM and MAAP in terms of computational time.

1.4 CRM for the problem of finding a common fixed-point

We start by recalling that the Convex Feasibility Problem (CFP), consists of finding a point in the intersection of a finite number of closed convex subsets of $\mathbb{R}^n$. CFP is clearly equivalent to solving a finite system of convex inequalities in $\mathbb{R}^n$, and it can be also rephrased as the problem of finding a common fixed-point of the orthogonal projections onto such subsets. A natural extension of CFP is the problem of finding a common fixed-point of a finite set of operators other than orthogonal projections, but sharing some of their properties. A vast literature on the subject has been developed; we cite just a few references, namely [27, 55, 61, 62]. We consider a particular generalization of orthogonal projections, namely *firmly nonexpansive operators*. The operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be firmly nonexpansive, when

$$\|T(x)-T(y)\|^2 \leq \|x-y\|^2 - \|(T(x)-T(y))-(x-y)\|^2, \quad \text{for all } \; x,y \in \mathbb{R}^n.$$

$F(T) = \{x \in \mathbb{R}^n \mid T(x) = x\}$ denotes the fixed-point set of the operator T.
Let $T_1, \dots, T_m : \mathbb{R}^n \to \mathbb{R}^n$ be firmly nonexpansive operators; The problem of finding a common fixed-point of $T_1, \dots, T_m$, will be denoted by FPP, and defined as:

$$\text{find } \bar{x} \in \bigcap_{i=1}^{m} F(T_i), \tag{1.4.1}$$

i.e., a point $\bar{x} \in \mathbb{R}^n$ such that $T_i(\bar{x}) = \bar{x}$ for all $i \in \{1, \dots, m\}$.
The set of common fixed-points of the T_i's will be denoted as $\text{Fix}(T_1, \dots, T_m)$. Two classical methods for FPP are the *Sequential Projection Method* (SPM) and the *Parallel Projection Method* (PPM), which are defined in Subsection 1.1.2. Starting from an arbitrary $x^0 \in \mathbb{R}^n$, the sequences generated by both methods are known to be globally convergent to points belonging to a point in $\text{Fix}(T_1, \dots, T_m)$, *i.e.*, to solve FPP.
We use Pierra's reduction, explained in Subsection 1.1.3, for FPP. Define the operator $\mathbf{T} : \mathbb{R}^{nm} \to \mathbb{R}^{nm}$ as $\mathbf{T}(x^1, \dots, x^m) = (T_1(x^1), \dots, T_m(x^m))$, with $x^i \in \mathbb{R}^n$ $(1 \leq i \leq m)$. It is rather immediate to check that $\mathbf{T}$ is firmly nonexpansive. Consider the set $\mathbf{U} = \{(x, \dots, x) : x \in \mathbb{R}^n\} \subset \mathbb{R}^{nm}$. Define $\{\mathbf{x}^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^{nm}$ as the sequence resulting from applying SPM to the operators $\mathbf{T}, P_{\mathbf{U}}$, starting from a point $\mathbf{x}^0 = (x^0, \cdots, x^0) \in \mathbf{U}$, *i.e.*, take $\mathbf{x}^{k+1} = P_{\mathbf{U}}(\mathbf{T}(\mathbf{x}^k))$. Again, in view of the formula of $P_{\mathbf{U}}$ given by (1.1.4), if $\mathbf{x}^k = (x^k, \dots x^k)$ is the k-th iterate of this sequence, then $x^k \in \mathbb{R}^n$ is the k-th iterate of the sequence in $\mathbb{R}^n$ defined as $x^{k+1} = \frac{1}{m}\sum_{i=1}^{m} T_i(x^k)$, which is the sequence defined by PPM applied to FPP.
We reckon that the use of the word "projections" in the names of SPM, PPM and MAP applied to FPP is an abuse of notation, since in general there are no projections involved in FPP. Indeed, they correspond to these methods applied to CFP, a particular case of FPP. We keep them because the structure of the methods applied to either CFP and FPP is basically the same.
Now we take two firmly nonexpansive operators $A, B : \mathbb{R}^n \to \mathbb{R}^n$, and recall from (1.1.5) and (1.2.2) that the MAP and CRM iterations are defined as $x^{k+1} = D(x^k) = A(B(x^k))$, and $x^{k+1} = C(x^k) = \text{circ}\left(x^k, B^R(x^k), A^R(B^R(x^k))\right)$, respectively. It is known that, under adequate assumptions, the sequence generated by MAP converges to a common fixed-point of A and B. Note that, if A, B are orthogonal projections onto convex sets K_1, K_2, then MAP turns out to be a special case of this iteration, and $\text{Fix}(A, B) = K_1 \cap K_2$.
CRM can be seen as an acceleration technique for the sequence defined by MAP. CRM was shown in [17] to converge to a solution of CFP. our results prove that, under a not too demanding error bound condition, the sequences generated by MAP and CRM for solving CFP converge linearly, but the asymptotic constant for CRM is better than the one for MAP. This superiority was widely confirmed in the numerical experiences executed in this thesis.
Here we will apply CRM for solving FPP with firmly nonexpansive operators $T_1, \dots T_m : \mathbb{R}^n \to \mathbb{R}^n$ in the following way. We will apply it to two operators in $\mathbb{R}^{nm}$, namely $\mathbf{T}$ and $P_{\mathbf{U}}$ as defined above, starting from a point in $\mathbf{U}$. Note that, since $\mathbf{U}$ is a linear subspace, the operator $P_{\mathbf{U}}$ is affine.
We prove that CRM applied to FPP is globally convergent to a common fixed-point (supposing that at least one exists). We also establish linear convergence of the sequence generated

by CRM applied to FPP, under a not too demanding error bound assumption defined as:

EB1: There exists $\omega > 0$ such that $\|x - T(x)\| \geq \omega \operatorname{dist}(x, \operatorname{Fix}(T, P_U))$ for all $x \in U$.

and provide an estimate of the asymptotic constant, which holds also for MAP. We were not able to prove the superiority of CRM in terms of the asymptotic constant of linear convergence, but our numerical experiments suggest that a theoretical superiority is likely to hold. This issue is left as a subject for future research.

1.5 Historical remarks

We present here some historical bibliography (see [26] for more details). In telegraphic language we recognize von Neumann, Kaczmarz and Cimmino as the forefathers. John von Neumann's 1933 [60], alternating projection method (MAP) is a projection method for finding the projection of a given point onto the intersection of two subspaces in Hilbert space. Stefan Kaczmarz, in a three pages paper published in 1937 [45], (posthumous translation into English in [46]) presented a sequential projection method for solving a (consistent) system of linear equations. Historical information about his life can be found in the papers of Maligranda [53] and Parks [56].

Gianfranco Cimmino proposed in [29], published in 1938, a simultaneous projection method for the same problem in which one projects the current iterate simultaneously on all hyperplanes, representing the linear equations of the system, and then takes a convex combination to form the next iterate. A historical account of Cimmino's work was published by Benzi [21].

In 1954 Agmon [1] and Motzkin and Schoenberg [54] generalized the sequential projection method from hyperplanes to half-spaces, and then Eremin [36] in 1965, Bregman [23] in 1965, and Gubin, Polyak and Raik in 1967 [40] generalized it farther to convex sets. In 1970 Auslender [5], generalized Cimmino's simultaneous projection method to convex sets.

These were the early beginnings that paved the way for the subsequent "explosion" of research in this field that continues to this day and covers many aspects. These include, but are not limited to, developments of new algorithmic structures for projection methods, usage of different types of projections, application of projection methods to new types of feasibility, optimization, or variational inequalities problems, investigations of the above in various spaces, branching into fixed-point theory and other mathematical areas, and using projection methods in significant real world problems with real data sets of humongous dimensions, and more.

Bauschke and Borwein's list of references in [7], with 109 items, is a treasure of knowledge on the field.

We continue with some historical references about CRM. The history of circumcenters dates back to as early as 300 BC, when they were described in Euclid's Elements [37, Book 4, Proposition 5]. More than two thousand years later, in 2018, Behling, Bello Cruz, and Santos [18], discovered that circumcenters are a simple yet effective way of accelerating the prominent Douglas-Rachford method. This work motivated groups of researchers to carefully study the properties of circumcenters from different viewpoints, characterizing them as useful tools for accelerating some classical methods. We are going to mention some of these recent works.

The circumcenter introduced by Behling, Bello Cruz, and Santos [18] led Bauschke, Ouyang, Wang in 2018, to present basic results and properties of the circumcenter of finite sets in Hilbert spaces, [10], and in the same year they studied the circumcenter mappings induced by nonexpansive operators, systematically explorimg the properness of the circumcenter mapping induced by reflections or projections.

Behling, Bello-Cruz and Santos, proved the convergence of the Circumcentered-Reflection Method (CRM) for finding a point in the intersection of a finite number of closed convex sets, see [17]. It accelerates well-known classic projection methods for solving CFP; namely, the Method of Alternating Projections (MAP) [45, 7], the Douglas-Rachford Method (DRM) [34, 52] and the Simultaneous Projection Method [29, 7]. In particular, CRM is an acceleration of the well-known Douglas-Rachford method (DRM) for finding the best approximation onto the intersection of finitely many affine subspaces. In 2019, in a paper celebrating 60 years of DRM, by Lindstrom and Sims [51], CRM was employed for multi-affine set problems, and also mentioned circumcenters as a natural way of dealing with DRM's spiralness.

Inspired by the CRM, Bauschke, Ouyang, Wang [11], introduced the more flexible circumcentered isometry method (CIM). The CIM essentially chooses the closest point to the solution among all of the points in an associated affine hull as its iterate and is a generalization of the CRM. They extended the linear convergence results on CRM in finite-dimensional spaces from reflections to isometries [10, 11].

In 2019, the notion of the best approximation mapping (BAM) with respect to a closed affine subspace in finite-dimensional spaces was introduced by Behling, Bello Cruz and Santos [14]. They show linear convergence of the block-wise circumcentered-reflection method. They introduced the technique of circumcentering in blocks, which, more than just an option over the basic algorithm of circumcenters, turns out to be an elegant manner of generalizing the method of alternating projections. Linear convergence for this novel block-wise circumcenter framework was derived and illustrated numerically. Furthermore, it was proved that the original circumcentered–reflection method essentially finds the best approximation solution in one single step if the given affine subspaces are hyperplanes.

Because the iteration sequence of BAM converges linearly, BAM is interesting in its own right. In 2021, Bauschke, Ouyang, Wang [12] extended the definition of BAM from closed affine subspaces to nonempty closed convex sets and from $\mathbb{R}^n$ to general Hilbert spaces. It was found that the convex set associated with BAM is the set of fixed point set of BAM. Hence, the iteration sequence generated by BAM converges linearly to the nearest fixed point of the BAM.

In 2020, Ouyang studied the finite convergence of locally proper circumcentered methods.

Inspired by some results on circumcentered-reflection method by Behling, Bello-Cruz, and Santos in their recent papers [18, 13, 17], sufficient conditions for one step convergence of circumcentered isometry methods for finding the best approximation point onto the intersection of fixed point sets of related isometries was established.

Very recently, new results on CRM were presented in [15] and [16]. It seems that circumcenters are here to stay.

1.6 Brief description of the contents in this thesis

This thesis is organized as follows.
Chapter 2 collects some standard materials and basic known facts concerning classical projection methods and firmly nonexpansive operators, which are useful in our later proofs. We designate all the known results as propositions with explicit references.

Our main results start in Chapter 3.
In Section 3.1 we study the convergence rate of CRM for solving the CFP and compare it with MAP. Under an error bound assumption, we prove that both methods converge linearly, with asymptotic constants depending on a parameter of the error bound, and that the one derived for CRM is strictly better than the one for MAP.
In Section 3.2 we present two rather generic families of examples for which CRM is faster than MAP. In the first one, MAP converges sublinearly while CRM converges linearly; in the second one, MAP converges linearly and CRM converges superlinearly.
The main results of this chapter are included in:
[4] Arefidamghani, R., Behling, R., Bello-Cruz, Y., Iusem, A. N., and Santos, L.-R.– The circumcentered-reflection method achieves better rates than alternating projections. *Computational Optimization and Applications* 79, 2 (2021), 507–530.

In Chapter 4 we introduce approximate versions of MAP and CRM for solving CFP, which we call MAAP and CARM, respectively. Chapter 4 is organized as follows.
In Section 4.1 we introduce the separating operator needed for MAAP and CARM. We present two particular choices of the separating operator S, which induce easily computable projections. This separating scheme is rather general, and for a large family of convex sets, includes this particular instances, the separating set is a half-space, or a cartesian product of half-spaces, in which cases all the involved projections have a very low computational cost.
In Section 4.2 we prove the convergence of the MAAP and CARM sequences, using the well known Fejér monotonicity argument. Moreover, we successfully extend some useful properties of exact CRM to CARM.
In Section 4.3 we study the linear convergence rate of MAAP and CARM. We prove that under error bound conditions, separating schemes are available so that MAAP and CARM enjoy linear convergence rates, with the linear rate of CARM being strictly better than MAAP.
Section 4.4 is specifically dedicated to examples. We analyze the convergence rate results for CARM and MAAP applied to specific instances of CFP. Two classes of somewhat generic examples are introduced. In the first one, the angle between the convex sets approaches zero near the intersection (U is a supporting hyperplane of K), so that the MAAP sequence converges sublinearly, but CARM still enjoys linear convergence.
In the second family, K is the epigraph of a convex f, but U is not a supporting hyperplane of K; rather it intersects the interior of K. In this case, under not too demanding assumptions on f, the MAAP sequence converges linearly (we will give an explicit bound of its asymptotic constant), while CARM converges superlinearly. We show that the CARM acceleration, in a rather large class of instances, improve the linear convergence to superlinear.
Numerical experiments are implemented in Section 4.5. Performance profiles are used to compare the convergence rates of several algorithms. We applied the Alternating Projection

Method (MAP), the Circumcentred-Reflection Method (CRM), the Approximate Alternating Projection Method (MAAP) and the Approximate Circumcentred-Reflection Method (CARM) for solving the convex feasibility problem and compare the rate of convergences for these methods. Our experimental results are consistent with the theoretical results proved in the previous sections, *i.e.,* our numerical experiments show CARM outperforming MAP, CRM and MAAP in terms of computational time.

The main results of this chapter are included in:
[3] Araujo, G., Arefidamghani, R., Behling, R., Bello-Cruz, Y., Iusem, A., and Santos, L.-R. Circumcentering approximate reflections for solving the convex feasibility problem. *Fixed Point Theory and Algorithms for Sciences and Engineering* 2022, 1 (2022), 1–30

In Chapter 5 we introduce CRM for the problem of finding a common fixed-point of a finite family of firmly nonexpansive operators (FPP).
Chapter 5 is organized as follows.
In Section 5.1 we present certain results, of some interest on its own, on convex combinations of orthogonal projections, which we take as a prototypical family of firmly nonexpansive operators (beyond orthogonal projections themselves).
In Section 5.2 we prove global convergence of CRM applied for solving FPP.
We also establish in Section 5.3 linear convergence of the sequence generated by CRM and MAP applied to FPP, under a not too demanding error bound assumption, and provide an estimate of the asymptotic constant, *i.e.,* we prove that in this case both the MAP and the CRM sequences converge linearly, with the same bound for asymptotic constants.
In Section 5.4 we report numerical comparisons between CRM and PPM for solving FPP with p firmly nonexpansive operators. Performance profiles are used to compare the convergence rates of algorithms; We provide solid numerical evidence of the superiority of CRM when compared to the classical Parallel Projections Method (PPM). Our experimental results are consistent with the theoretical results proved in this chapter.

The main results of this chapter are included in:
Arefidamghani, R., Behling, R., Iusem, A., and Santos, L.-R. A circumcentered-reflection method for finding common fixed points of firmly nonexpansive operators, *Submited in Journal of Applied and Numerical Optimization (JANO), to be published.*

Chapter 2

Background material and preliminaries

In this chapter, we collect some mathematical definitions and results from the literature. For all known results, we explicitly refer to the appropriate references. The Circumcentred-Reflection method, introduced in Section 2.5.3, plays a critical role in this thesis.

2.1 Convergence rates and Fejér monotonicity

We start by recalling the definition of Q-linear and R-linear convergence.

Definition 2.1.1. Consider a sequence $\left\{x^k\right\}_{k\in N} \subset \mathbb{R}^n$ that converges to $\bar{x}$. Assume that $x^k \neq \bar{x}$ for all $k \in \mathbb{N}$. Let $q := \limsup_{k\to\infty} \dfrac{\left\|x^{k+1} - \bar{x}\right\|}{\left\|x^k - \bar{x}\right\|}$. Then the sequence $\left\{x^k\right\}_{k\in N}$ converges

(i) Q-superlinearly, if $q = 0$,

(ii) Q-linearly, if $q \in (0, 1)$,

(iii) Q-sublinearly, if $q \geq 1$.

Let $r := \limsup_{k\to\infty} \left\|x^k - \bar{x}\right\|^{\frac{1}{k}}$. Then the sequence $\left\{x^k\right\}_{k\in N}$ converges

(iv) R-superlinearly, if $r = 0$,

(v) R-linearly, if $r \in (0, 1)$,

(vi) R-sublinearly, if $r \geq 1$.

The values q and r are called asymptotic constant of Q-linear and R-linear convergence respectively.

Proposition 2.1.1. Let $\left\{x^k\right\}_{k\in N}$ be a sequence that converges to $\bar{x}$ Q-linearly. Then it is R-linearly convergent with the same asymptotic constant.

Proof. Elementary. ■

In the following example, we show that the converse of the above statement is not true.

Example 2.1.1. Consider the sequence given by

$$x^k := \begin{cases} \frac{1}{2^k} & \text{k is even ,} \\ \frac{1}{2^{(k+1)}} & \text{k is odd.} \end{cases}$$

Since

$$\lim_{k\to\infty} \sqrt[k]{x^k} = \lim_{k\to\infty} \frac{1}{\sqrt[k]{2^k}} = \frac{1}{2}, \qquad \text{while k is even,}$$

$$\lim_{k\to\infty} \sqrt[k]{x^k} = \lim_{k\to\infty} \frac{1}{\sqrt[k]{2^{k+1}}} = \frac{1}{2}, \qquad \text{while k is odd,}$$

this sequence is R-linear convergent with asymptotic constant $\frac{1}{2}$. Note that $\left\{x^k\right\}_{k\in N}$ is not Q-linear convergent, because when k is odd, we have

$$x^k = \frac{1}{2^{k+1}} \ , \ x^{k+1} = \frac{1}{2^{(k+1)}}.$$

Hence,

$$\limsup_{k\to\infty} \frac{\left\|x^{k+1} - \bar{x}\right\|}{\left\|x^k - \bar{x}\right\|} = \frac{\left|\frac{1}{2^{k+1}}\right|}{\left|\frac{1}{2^{k+1}}\right|} = 1 \tag{2.1.1}$$

along the subsequence of odd iterates. From (2.1.1) ,we conclude that

$$\limsup_{k\to\infty} \frac{\left\|x^{k+1} - \bar{x}\right\|}{\left\|x^k - \bar{x}\right\|} \geq 1$$

along the whole sequence, which means that the sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ is not Q-sublinear convergent.

We remind now the notion of Fejér monotonicity. We will see in the following chapters that Fejér monotonicity plays a crucial role for proving convergence of sequences generated by CRM and MAP for solving either CFP or FPP.

Definition 2.1.2. A sequence $\left\{x^k\right\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ is Fejér monotone with respect to nonempty closed convex set $M \subset \mathbb{R}^n$ when

$$\left\|x^{k+1} - y\right\| \leq \left\|x^k - y\right\| \qquad \forall y \in M, \ k \in \mathbb{N}. \tag{2.1.2}$$

Proposition 2.1.2. Suppose that $\left\{x^k\right\}_{k\in\mathbb{N}}$ is Fejér monotone with respect to closed convex set $M \subset \mathbb{R}^n$. Then

(i) $\left\{x^k\right\}_{k\in\mathbb{N}}$ is bounded and $\operatorname{dist}(x^{k+1}, M) \leq \operatorname{dist}(x^k, M)$.

(ii) if a cluster point $\bar{x}$ of $\left\{x^k\right\}_{k\in\mathbb{N}}$ belongs to M, we have $\lim_{k\to\infty} x^k = \bar{x}$.

(iii) if $\left\{x^k\right\}_{k\in\mathbb{N}}$ converges to $\bar{x} \in M$, we get $\left\|x^k - \bar{x}\right\| \leq 2 \operatorname{dist}(x^k, M)$

Proof. For (i) and ii), see [7, Theorem 2.16]. Regarding the estimate in item (iii), for all $j > k$, we have

$$\begin{aligned} \left\|x^k - x^j\right\| &\leq \left\|x^k - P_M(x^k)\right\| + \left\|x^j - P_M(x^k)\right\| \\ &\leq 2\left\|x^k - P_M(x^k)\right\| = 2 \operatorname{dist}(x^k, M), \end{aligned} \tag{2.1.3}$$

using the definition of Fejér monotonicity with $y = P_M(x^k)$ in the last inequality. Taking limits in (2.1.3) with $j \to \infty$, we get the result. ■

2.2 Orthogonal projections onto convex sets

Next, we address some properties of the orthogonal projection onto a closed and convex set. We recall that given a closed and convex set $M \subset \mathbb{R}^n$, the *orthogonal projection* onto M is the operator $P_M : \mathbb{R}^n \to M$ defined as $P_M(x) = \operatorname{argmin}_{y \in M}\{\|x - y\|\}$. The next proposition establishes that P_M is well defined.

Proposition 2.2.1. Let $M \subseteq \mathbb{R}^n$ be a closed and convex set. Then for any $x \in \mathbb{R}^n$, the projection of x onto M denoted by $P_M(x)$, exists and is unique. Moreover,

$$p = P_M(x) \iff p \in M \text{ and } \langle x - p, y - p \rangle \leq 0, \quad \forall\, y \in M. \tag{2.2.1}$$

Proof. See [2, Proposition 2]. ■

Corollary 2.2.1. Let $U \subseteq \mathbb{R}^n$ be an affine manifold. Then for any point $x \in \mathbb{R}^n$, the projection onto U exist, and it is unique. Moreover, the projection onto an affine manifold satisfies (2.2.1) with equality.

Proof. See [2, Corollary 1]. ■

We continue with some elementary and well known properties of orthogonal projections onto closed and convex sets.

Proposition 2.2.2. Let $M \subset \mathbb{R}^n$ be a nonempty closed and convex set. Then, for all $x, y \in \mathbb{R}^n$ we have

$$\|P_M(x) - P_M(y)\|^2 \leq \|x - y\|^2 - \|(x - P_M(x)) - (y - P_M(y))\|^2, \tag{2.2.2}$$

and moreover $\operatorname{Fix}(P_M) = M$.

Proof. See [8, Proposition 4.8]. ■

Proposition 2.2.2 establishes that orthogonal projections onto closed and convex sets are *firmly nonexpansive*; see Section 2.5.1

Proposition 2.2.3. Let $C \subset \mathbb{R}^n$ be closed and convex. Then,

i) Take $x \in \mathbb{R}^n$ and let $z = P_C(x)$. Then, $P_C(z + \alpha(x - z)) = P_C(x)$ for all $\alpha \geq 0$.

ii) Define $h : \mathbb{R}^n \to \mathbb{R}$ as $h(x) = \|x - P_C(x)\|^2$. Then h is continuously differentiable and $\nabla h(x) = 2\,(x - P_C(x))$.

Proof. Elementary. ■

2.3 Subdifferentials of convex functions

In this section, we recall the definitions of subgradient and subdifferential, and present some of their properties.

Definition 2.3.1. A vector $u \in \mathbb{R}^n$ is called a subgradient of $f : \mathbb{R}^n \to \mathbb{R}$ at x_0 iff:

$$f(x_0) + \langle u, x - x_0 \rangle \leq f(x) \tag{2.3.1}$$

for all $x \in \mathbb{R}^n$.
The set of all subgradients of f at x is called the subdifferential of f at x, denoted as

$$\partial f(x) = \{u : u \text{ is a subgradient of } f \text{ at } x\}.$$

Proposition 2.3.1. Given a convex function $f : \mathbb{R}^n \to \mathbb{R}$, the mapping $\partial f : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ is locally bounded, *i.e.,* the image $\partial f(B)$ of a bounded set $B \subset \mathbb{R}^n$ is bounded in $\mathbb{R}^n$.

Proof. See [41, Proposition 6.2.2]. ■

2.4 Convergence results for MAP and CRM

In this section, we collect some known classical results about Method of Alternating Projection (MAP) and Circumcentered-Reflection method (CRM), in view of their later use. See Chapters 3, 4 and 5, for an in-depth study and farther analyze of these and other projection methods. In order to understand how work, we give a short explanation in the following. Let us start with MAP.
MAP is an algorithm for computing a point in the intersection of two convex sets, using a sequence of projections onto these sets. Suppose that K_1 and K_2 are closed convex sets in $\mathbb{R}^n$, and let P_{K_1} and P_{K_2} denote the projections onto K_1 and K_2, respectively. The algorithm starts with any $z^0 \in \mathbb{R}^n$, and then alternately projects onto K_1 and K_2:

$$z^{k+1} = P_{K_2}(P_{K_1}(z^k)), \quad k = 0, 1, 2, \dots$$

Thus, MAP generates a sequence of points $\{z^k\}_{k\in\mathbb{N}} \subset K_2$.
The next result establishes that the sequence generated by MAP converges to a solution of CFP, starting from any initial point in $\mathbb{R}^n$. The basic result, due to Cheney and Goldstein, is the following.

Proposition 2.4.1. Take closed convex sets $K_1, K_2 \subset \mathbb{R}^n$ such that $K_1 \cap K_2 \neq \varnothing$. Starting from any $z^0 \in \mathbb{R}^n$, the sequence $\{z^k\}_{k\in\mathbb{N}}$ generated by MAP is Fejér monotone with respect to $K_1 \cap K_2$ and converges to a point $\bar{z} \in K_1 \cap K_2$.

Proof. See [28, Theorem 4]. ■

We present now the formal definition of the circumcenter of three points.

Definition 2.4.1. Let $x, y, z \in \mathbb{R}^n$ be given. The circumcenter $\operatorname{circ}(x, y, z) \in \mathbb{R}^n$ is a point satisfying

(i) $\|\mathrm{circ}(x,y,z)-x\| = \|\mathrm{circ}(x,y,z)-y\| = \|\mathrm{circ}(x,y,z)-z\|$ and

(ii) $\mathrm{circ}(x,y,z) \in \mathrm{aff}\{x,y,z\} := \{w \in \mathbb{R}^n \mid w = x + \alpha(y-x) + \beta(z-x),\ \alpha,\beta \in \mathbb{R}\}$.

The point $\mathrm{circ}(x,y,z)$ is well and uniquely defined if the cardinality of the set $\{x,y,z\}$ is one or two. In the case in which the three points are all distinct, $\mathrm{circ}(x,y,z)$ is well and uniquely defined only if x, y and z are not collinear. For more general notions, definitions and results on circumcenters see [18, 17, 10].
Consider now a closed convex set $K \subset \mathbb{R}^n$ and an affine manifold $U \subset \mathbb{R}^n$. Let P_K, P_U be the orthogonal projections onto K, U respectively. Define $R_K, R_U, C : \mathbb{R}^n \to \mathbb{R}^n$ as

$$R_K = 2P_K - \mathrm{Id}, \quad R_U = 2P_U - \mathrm{Id}, \quad C(\cdot) = \mathrm{circ}(\cdot, R_K(\cdot), R_U(R_K(\cdot))). \tag{2.4.1}$$

Starting from $z \in U$, the CRM scheme is defined as:

$$z^{k+1} := C(z^k) = \mathrm{circ}\left(z^k, R_K(z^k), R_U R_K(z^k)\right), \quad z^0 \in U. \tag{2.4.2}$$

It turns out that if $z^0 \in U$, then the whole sequence $\{z^k\}$ generated by (2.4.2) remains in U and converges to a solution of CFP with sets K, U, that is a point in $K \cap U$. Moreover, when the sets in CFP are an affine manifold and a hyperplane, CRM indeed converges in one step (see [17] for more details).
The next results concern the domain of the circumcenter operator for K and U, namely

$$\mathrm{dom}(C) := \{z \in \mathbb{R}^n | C(z) = \mathrm{circ}\left(z, R_K(z), R_U R_K(z)\right) \text{ is well-defined}\}.$$

We will see that $U \subset \mathrm{dom}(C)$ and $C(z) \in U$, whenever $z \in U$.

Proposition 2.4.2. Let $U, H \subset \mathbb{R}^n$ be an affine manifold and a subspace respectively, such that $H \cap U \neq \varnothing$. Denote as $P_H, R_H : \mathbb{R}^n \to \mathbb{R}^n$ the projection and the reflection with respect to H, respectively. Then,

- $P_{H\cap U}(x) = \mathrm{circ}(x, R_U(x), R_H(R_U(x))$ for all $x \in U$,
- $\mathrm{circ}(x, R_U(x), R_H(R_U(x)) \in U$ for all $x \in U$.

Proof. See [17, Lemmas 2 and 3]. ■

This means that when the sets in CFP are an affine manifold and a hyperplane, CRM indeed converges in one step, which is a first indication of its superiority over MAP, which certainly does not enjoy this one-step convergence property, but also points to the main weakness of CRM, namely that for its convergence we may replace H by a general closed convex set, but the other set must be kept as an affine manifold.

Proposition 2.4.3. Take $K, U \subset \mathbb{R}^n$, where K is a closed convex set and U is an affine manifold, and assume that $K \cap U \neq \varnothing$. Then, for all $z \in U$, $C(z) := \mathrm{circ}\left(z, R_K(z), R_U R_K(z)\right)$ is well-defined, and we have $C(z) \in U$. Furthermore $C(z) = P_{H_z \cap U}(z)$, where $H_z := \{x \in \mathbb{R}^n | \langle x - P_K(z), z - P_K(z)\rangle = 0\}$ if $z \notin K$ and $H_z := K$, otherwise.

Proof. See [17, Lemma 3]. ■

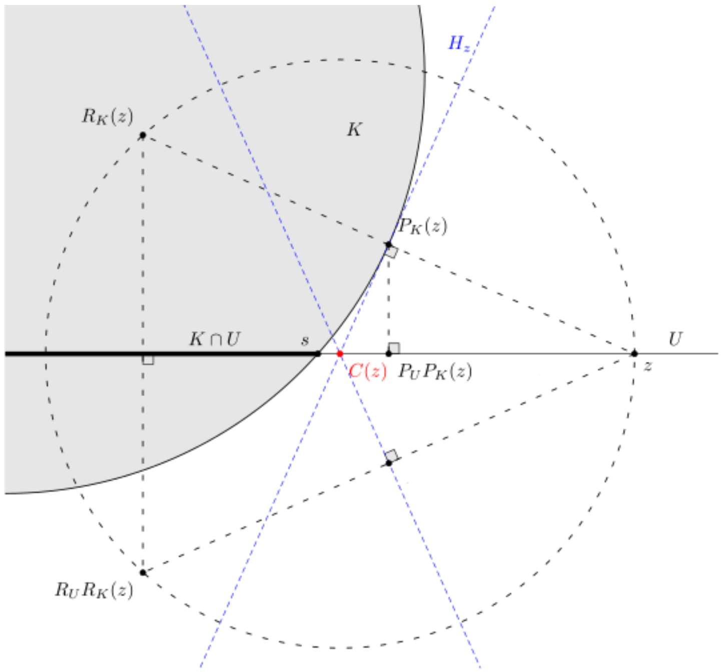

Figure 2.1: Illustration of CRM for the intersection between an affine U and a convex K,

In Figure 2.1(is taken from [17]), we illustrate geometrically what has been established in Proposition 2.4.3. The next result shows that CRM finds a point in $K \cap U$ whenever the initial point is chosen in U.

Proposition 2.4.4. Take $K, U \subset \mathbb{R}^n$, where K is a closed convex set and U is an affine manifold, and assume $K \cap U \neq \varnothing$. Then the sequence $\left\{x^k\right\}_{k\in\mathbb{N}}$ generated by CRM, is Fejér monotone with respect to the convex set $K \cap U$ and converges to a point in $K \cap U$.

Proof. See [17, Theorem 1]. ■

We close this section with a theorem stating that for a given iterate in U, CRM gets us closer to the solution set than MAP.

Proposition 2.4.5. Take $K, U \subset \mathbb{R}^n$, where K is a closed convex set and U is an affine manifold and assume $K \cap U \neq \varnothing$. Recall the operators $D, C : \mathbb{R}^n \to \mathbb{R}^n$, defined as

$$D = P_U \circ P_K \ , \quad C(\cdot) = \operatorname{circ}\left(\cdot, R_K(\cdot), R_U R_K(\cdot)\right). \tag{2.4.3}$$

Then, For $z \in U$, and for any $y \in K \cap U$ we have

(i) $\|C(z) - y\| \leq \|D(z) - y\|$,

(ii) $\operatorname{dist}(C(z), K \cap U) \leq \operatorname{dist}(D(z), K \cap U)$.

(iii) $D(z)$ belongs to the segment between z and $C(z)$.

Proof. See [17, Theorem 2]. ■

2.5 Firmly nonexpansive operators

We present in this section the definition and some elementary properties of firmly nonexpansive operators.

Definition 2.5.1. Let M be a nonempty subset of $\mathbb{R}^n$ and let $T : M \to \mathbb{R}^n$. Then T is
(i) firmly nonexpansive if

$$\|T(x) - T(y)\|^2 \leq \|x - y\|^2 - \|(T(x) - T(y)) - (x - y)\|^2 \qquad \forall\, x, y \in M; \tag{2.5.1}$$

(ii) nonexpansive if it is Lipschitz continuous with constant 1, *i.e.*,

$$\|T(x) - T(y)\| \leq \|x - y\| \qquad \forall\, x, y \in M; \tag{2.5.2}$$

It is clear that firm nonexpansiveness implies nonexpansiveness. For an operator $T : \mathbb{R}^n \to \mathbb{R}^n$, we denote as $F(T)$ the set of its fixed-points, *i.e.*, $F(T) := \{x \in \mathbb{R}^n : T(x) = x\}$.

Example 2.5.1. The operator T_1 and T_2 defined as

$$T_1(x) = \begin{cases} (1 - \frac{1}{\|x\|})x, & \|x\| > 1, \\ 0, & \|x\| \leq 1. \end{cases}$$

$$T_2(x) = \begin{cases} (1 - \frac{2}{\|x\|})x, & \|x\| > 1, \\ -x, & \|x\| \leq 1. \end{cases}$$

T_2 is an example of nonexpansive operator which is not firmly nonexpansive. See [8, Example 4.9] for more details.

Proposition 2.5.1. Let M be a nonempty subset of $\mathbb{R}^n$. Take $T : M \to \mathbb{R}^n$. Let Id denote the identity operator in $\mathbb{R}^N$. The following statements are equivalent.

(i) T is firmly nonexpansive.

(ii) $\mathrm{Id} - T$ is firmly nonexpansive.

(iii) $2T - \mathrm{Id}$ is nonexpansive.

(iv) $\|T(x) - T(y)\|^2 \leq \langle x - y, T(x) - T(y)\rangle \qquad \forall\, x, y \in M$.

(v) $0 \leq \langle T(x) - T(y), (\mathrm{Id} - T)x - (\mathrm{Id} - T)y\rangle \qquad \forall x, y \in D$.

Proof. Simply follows from expanding norms. See [8, Proposition 4.2]. ■

We continue with other elementary properties of firmly nonexpansive operators.

Proposition 2.5.2. Let T_i for $(i = 1, 2, \cdots, n)$ be a finite family of firmly nonexpansive operators, and positive scalars α_i's for $(i = 1, 2, \cdots, n)$ such that $\sum_{i=1}^n \alpha_i = 1$. Then if $\bigcap_{i=1}^m F(T_i) \neq \varnothing$, we have

$$\bigcap_{i=1}^m F(T_i) = F(T_m \circ \cdots \circ T_1) = F\left(\sum_{i=1}^m \alpha_i T_i\right).$$

Proof. See [7, Propositions 2.10(i), 2.12(i)]. ■

Remark 2.5.1. We remark that the nonempty intersection condition in Proposition 2.5.2 is necessary. For example, consider any two parallel lines C, D in $\mathbb{R}^2$, and define $T = \frac{1}{2}(P_D + P_C)$, as the convex combination of P_C and P_D, with coefficients $\frac{1}{2}$. In this case $F(T)$ is the line parallel to C and D, equidistant from both. So $F(T) \neq F(P_C) \cap F(P_D) = C \cap D = \varnothing$.

Proposition 2.5.3. Let M be a nonempty closed convex subset in $\mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^n$ be a firmly nonexpansive operator such that

$$\mathrm{Im}(T) \subseteq F(T) = M,$$

then $T = P_M$.

Proof. Take $x \in \mathbb{R}^n$ and $y \in M$. Then we have $T(x) \in \mathrm{Img}(T) \subseteq M$ and $y = T(y) \in F(T) = M$. Since T is a firmly nonexpansive operator, by Proposition 2.5.1(v), we have

$$0 \leq \langle T(x) - T(y), (x - T(x)) - (y - T(y)) \rangle = \langle T(x) - y, x - T(x) \rangle.$$

Thus we have $\langle x - T(x), y - T(x) \rangle \leq 0$, for any $x \in \mathbb{R}^n$ and $y \in M$. It follows from Proposition 2.2.1 that $T = P_M$. ■

We present next some properties of the set of fixed-point of operators in $\mathbb{R}^n$. They have been proved, *e.g.,* in [8, Propositions 4.13 , 4.14], but we include the proofs for the sake of completeness.

Proposition 2.5.4. Let M be a nonempty closed convex subset of $\mathbb{R}^n$ and let $T : M \to \mathbb{R}^n$ be firmly nonexpansive. Then T is continuous and $F(T)$ is closed convex.

Proof. Let x and y be in $F(T)$, let $\alpha \in (0, 1)$, and set $z = \alpha x + (1 - \alpha)y$. Then $z \in M$ and we have

$$\begin{aligned} \|T(z) - z\|^2 &= \|\alpha(T(z) - x) + (1 - \alpha)(T(z) - y)\|^2 \\ &= \alpha\|T(z) - x\|^2 + (1 - \alpha)\|T(z) - y\|^2 - \alpha(1 - \alpha)\|x - y\|^2 \\ &\leq \alpha\|z - x\|^2 + (1 - \alpha)\|z - y\|^2 - \alpha(1 - \alpha)\|x - y\|^2 \\ &= \|\alpha(z - x) + (1 - \alpha)(z - y)\|^2 = 0. \end{aligned} \tag{2.5.3}$$

Therefore, $z \in F(T)$. Note that the inequality in (2.5.3) holds because T is nonexpansive, and the second and the last equality holds because $\|\alpha x + (1-\alpha)y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1 - \alpha)\|x - y\|^2$.
Since T is firmly nonexpansive, it is Lipschitz with constant 1, and therefore it is continuous. Let x^k be a sequence in $F(T)$ that converges to a point $x \in \mathbb{R}^n$. Then $x \in M$ by closedness of M, while $T(x^k)$ converges to $T(x)$ by continuity of T. On the other hand, since $x^k \subset F(T)$, $T(x^k)$ converges to x. We conclude that $T(x) = x$. ■

Corollary 2.5.1. Let T_i $(i \in \{1, 2, \cdots, n\})$ be a finite family of firmly nonexpansive operators. Then $\cap_{i=1}^m F(T_i)$ is closed and convex.

Proof. By Proposition 2.5.4, the set of fixed-point of each T_i's is closed convex. Hence, the intersection of the fixed-point sets is closed and convex. ■

Proposition 2.5.5. Let M be a nonempty convex subset of $\mathbb{R}^n$ and let $T : M \to \mathbb{R}^n$ be firmly nonexpansive. Then

$$F(T) = \bigcap_{x \in M} \{y \in M : \langle y - T(x), x - T(x) \rangle \leq 0\}. \tag{2.5.4}$$

Proof. Set $C = \bigcap_{x \in M} \{y \in M : \langle y - T(x), x - T(x) \rangle \leq 0\}$. For every $x \in M$ and every $y \in F(T)$, Proposition 2.5.1(v) yields $0 \leq \langle T(x) - y, x - T(x) \rangle$. Hence, $F(T) \subseteq C$.
Conversely, let $x \in C$. Then $x \in \{y \in M | \langle y - T(x), x - T(x) \rangle \leq 0\}$, and therefore

$$\|x - T(x)\|^2 = \langle x - T(x), x - T(x) \rangle \leq 0,$$

i.e., $x = T(x)$. Thus, $C \subseteq F(T)$. ■

To our knowledge, unless otherwise stated, all the results contained in the remainder of this thesis are new.

Chapter 3

The convergence rate of the Circumcentered-Reflection method applied to the convex feasibility problem

We deal in this chapter with the convex feasibility problem (CFP), defined as follows: given closed convex sets $K_1, \dots, K_m \subset \mathbb{R}^n$, find a point in $\bigcap_{i=1}^m K_i$.
We study the convergence rate of the Circumcentered-Reflection Method (CRM) applied for solving CFP, and compare it with the Method of Alternating Projections (MAP). Under an error bound assumption, we prove that both methods converge linearly, with asymptotic constants depending on a parameter of the error bound, and that the one derived for CRM is strictly better than the one for MAP. Next, we analyze two classes of somewhat generic examples. In the first one, the angle between the convex sets approaches zero near the intersection, so that the MAP sequence converges sublinearly, but CRM still enjoys linear convergence. In the second class of examples, the angle between the sets does not vanish and MAP exhibits its standard behavior, *i.e.*, it converges linearly, yet, perhaps surprisingly, CRM attains superlinear convergence.

3.1 Linear convergence of MAP and CRM under an error bound assumption

In view of Pierra's reformulation (Subsection 1.1.3), the general CFP can be seen as a specific convex-affine intersection problem. Hence, from now on, we focus on finding a point common to a given closed convex set $K \subset \mathbb{R}^n$ and an affine manifold $U \subset \mathbb{R}^n$.
Recall that MAP and CRM iterate by means of the operators,

$$D = P_U \circ P_K, \text{ and } C(\cdot) = \operatorname{circ}(\cdot, R_K(\cdot), R_U(R_K(\cdot))) \tag{3.1.1}$$

respectively, where $R_K = 2P_K - \operatorname{Id}$, and $R_U = 2P_U - \operatorname{Id}$, are reflection operators over K and U. The distance of a point $x \in \mathbb{R}^n$ to a set $K \subset \mathbb{R}^n$ will be denoted as $\operatorname{dist}(x, K)$.
A first result in the analytical study of the acceleration effect of CRM over MAP was derived in [17, Theorem 2], where it was proved that $\operatorname{dist}(C(x), K \cap U) \leq \operatorname{dist}(D(x), K \cap U)$ for all

$x \in U$, meaning that the point obtained after a CRM step is closer to (or at least no farther from) $K \cap U$ than the one obtained after a MAP step from the same point. This local (or myopic) acceleration does not imply immediately that the CRM sequence converges faster than the MAP one. In order to show global acceleration, we will focus on special situations where the convergence rate of the MAP can be precisely established.
MAP is known to be linearly convergent in several special situations, *e.g.*, when both K and U are affine manifolds (see [48]) or when $K \cap U$ has nonempty interior (see [6]). In this section, we consider another such case, namely when a certain so-called *error bound* holds. Next, we define an error bound assumption (EB, from now on) on a closed convex set $K \subset \mathbb{R}^n$ and an affine manifold $U \subset \mathbb{R}^n$ which will ensure linear convergence of MAP and CRM.

EB. $K \cap U \neq \varnothing$ and there exists $\omega \in (0,1)$ such that $\operatorname{dist}(x, K) \geq \omega \operatorname{dist}(x, K \cap U)$ for all $x \in U$.

Let us comment on the connection between (EB) and other notions of error bounds which have been introduced in the past, all of them related to regularity assumptions imposed on the solutions of certain problems. If the problem at hand consists of solving $H(x) = 0$ with a smooth $H : \mathbb{R}^n \to \mathbb{R}^m$, a classical regularity condition demands that $m = n$ and the Jacobian matrix of H be nonsingular at a solution x^*, in which case, Newton's method, for instance, is known to enjoy superlinear or quadratic convergence. This condition implies local uniqueness of the solution x^*. For problems with nonisolated solutions, a less demanding assumption is the notion of *calmness* (see [59, Chapter 8, Section F]), which requires that

$$\frac{\|H(x)\|}{\operatorname{dist}(x, S^*)} \geq \omega \tag{3.1.2}$$

for all $x \in \mathbb{R}^n \setminus S^*$ and some $\omega > 0$, where S^* is the solution set, *i.e.*, the set of zeros of H. Calmness, also called upper-Lipschitz continuity (see [58]), is a classical example of error bound, and it holds in many situations (*e.g.*, when H is affine, by virtue of Hoffman's Lemma, [42]). It implies that the solution set is locally a Riemannian manifold (see [20]), and it has been used for establishing superlinear convergence of Levenberg-Marquardt methods in [47].
When dealing with convex feasibility problems, it seems reasonable to replace the numerator of (3.1.2) by the distance from x to some of the convex sets, giving rise to (EB). Similar error bounds for feasibility problems can be found, for instance, in [6, 7, 35, 49].
Assuming that K, U satisfy Assumption EB, we will prove linear convergence of the sequences $\{z^k\}_{k\in\mathbb{N}}$ and $\{x^k\}_{k\in\mathbb{N}}$ generated by MAP and CRM, respectively. We start by proving that, for both methods, both distance sequences $\{\operatorname{dist}(z^k, K \cap U)\}_{k\in\mathbb{N}}$ and $\{\operatorname{dist}(x^k, K \cap U)\}_{k\in\mathbb{N}}$ converge linearly to 0, which will be a corollary of the next proposition.

Proposition 3.1.1. Assume that K, U satisfy EB. Consider $D, C : \mathbb{R}^n \to \mathbb{R}^n$ as (3.1.1). Then, for all $x \in U$,

$$(1 - \omega^2) \|x - P_{K\cap U}(x)\|^2 \geq \|D(x) - P_{K\cap U}(D(x))\|^2 \geq \|C(x) - P_{K\cap U}(C(x))\|^2, \tag{3.1.3}$$

with ω as in Assumption EB.

Proof. It follows easily from Proposition 2.2.2 that

$$\|P_K(x) - y\|^2 \leq \|x - y\|^2 - \|x - P_K(x)\|^2 \tag{3.1.4}$$

for all $x \in \mathbb{R}^n$ and all $y \in K \cap U \subset K$. Invoking again Proposition 2.2.2, we get from (3.1.4)

$$\begin{aligned}\|D(x) - y\|^2 = \|P_U(P_K(x)) - y\|^2 &\leq \|P_K(x) - y\|^2 - \|P_U(P_K(x)) - P_K(x)\|^2 \\ &\leq \|x - y\|^2 - \|x - P_K(x)\|^2 - \|P_U(P_K(x)) - P_K(x)\|^2 \\ &\leq \|x - y\|^2 - \|x - P_K(x)\|^2 = \|x - y\|^2 - \operatorname{dist}^2(x, K) \\ &\leq \|x - y\|^2 - \omega^2 \operatorname{dist}^2(x, K \cap U) \end{aligned} \tag{3.1.5}$$

for all $x \in U, y \in K \cap U$, using Assumption EB in the last inequality. Take now $y = P_{K\cap U}(x)$. Then, in view of (3.1.5),

$$\begin{aligned}\|C(x) - P_{K\cap U}(C(x))\|^2 &\leq \|C(x) - P_{K\cap U}(D(x))\|^2 \\ &\leq \|D(x) - P_{K\cap U}(D(x))\|^2 \leq \|D(x) - P_{K\cap U}(x)\|^2 \\ &\leq \|x - P_{K\cap U}(x)\|^2 - \omega^2 \operatorname{dist}^2(x, K \cap U) \\ &= (1 - \omega^2) \|x - P_{K\cap U}(x)\|^2, \end{aligned} \tag{3.1.6}$$

using the definition of $P_{K\cap U}$ in the first and the third inequality and Proposition 2.4.5(i) in the second inequality. Note that (3.1.3) follows immediately from (3.1.6). ■

Corollary 3.1.1. Let $\{z^k\}_{k\in\mathbb{N}}$ and $\{x^k\}_{k\in\mathbb{N}}$ be the sequences generated by MAP and CRM starting at any $z^0 \in \mathbb{R}^n$ and any $x^0 \in U$, respectively. If K, U satisfy Assumption EB, then the sequences $\{\operatorname{dist}(z^k, K \cap U)\}_{k\in\mathbb{N}}$ and $\{\operatorname{dist}(x^k, K \cap U)\}_{k\in\mathbb{N}}$ converge Q-linearly to 0, and the asymptotic constants are bounded above by $\sqrt{1 - \omega^2}$, with ω as in Assumption EB.

Proof. In view of the definition of $P_{K\cap U}$, (3.1.3) can be rewritten as

$$(1 - \omega^2) \operatorname{dist}^2(x, K \cap U) \geq \operatorname{dist}^2(D(x), K \cap U) \geq \operatorname{dist}^2(C(x), K \cap U), \tag{3.1.7}$$

for all $x \in U$. Since $z^{k+1} = D(z^k)$, we get from the first inequality in (3.1.7),

$$(1 - \omega^2) \operatorname{dist}^2(z^k, K \cap U) \geq \operatorname{dist}^2(z^{k+1}, K \cap U),$$

using the fact that $\{z^k\}_{k\in\mathbb{N}} \subset U$. Hence,

$$\frac{\operatorname{dist}(z^{k+1}, K \cap U)}{\operatorname{dist}(z^k, K \cap U)} \leq \sqrt{1 - \omega^2}. \tag{3.1.8}$$

By the same token, using the second inequality in (3.1.7) and Proposition 2.4.5(ii), we get

$$\frac{\operatorname{dist}(x^{k+1}, K \cap U)}{\operatorname{dist}(x^k, K \cap U)} \leq \sqrt{1 - \omega^2}. \tag{3.1.9}$$

The inequalities in (3.1.8) and (3.1.9) imply the result. ■

We remark that the result for MAP holds when U is any closed convex set, not necessarily an affine manifold. We need U to be an affine manifold in Proposition 2.4.4 (otherwise, $\{x^k\}_{k\in\mathbb{N}}$ may even diverge, see Figure 1.3), but this proposition is used in our proofs only when the CRM sequence is involved.
We will show that, under Assumption EB, CRM achieves a linear rate with an asymptotic constant better than the one given above. To do this a preliminary result, relating $x, C(x)$ and $D(x)$, is needed. We show that $x, C(x)$ and $D(x)$, are collinear. A similar result can be found in [17], where it is proved with a geometrical argument. Here we present an analytical one.

Proposition 3.1.2. Consider the operators $C, D : \mathbb{R}^n \to \mathbb{R}^n$ defined in (3.1.1). Then $D(x)$ belongs to the segment between x and $C(x)$ for all $x \in U$.

Proof. Let E denote the affine manifold spanned by $x, R_K(x)$ and $R_U(R_K(x))$. By definition, the circumcenter of these three points, namely $C(x)$, belongs to E. We claim that $D(x)$ also belongs to E, and next we proceed to prove the claim. Since U is an affine manifold, P_U is an affine operator, so that $P_U(\alpha x + (1-\alpha)x') = \alpha P_U(x) + (1-\alpha)P_U(x')$ for all $\alpha \in \mathbb{R}$ and all $x, x' \in \mathbb{R}^n$. By definition of reflection, $R_U(R_K(x)) = 2P_U(R_K(x)) - R_K(x)$, so that

$$P_U(R_K(x)) = \frac{1}{2}\left(R_U(R_K(x)) + R_K(x)\right). \tag{3.1.10}$$

On the other hand, using the affinity of P_U, the definition of D and the assumption that $x \in U$, we have

$$P_U(R_K(x)) = P_U(2P_K(x) - x) = 2P_U(P_K(x)) - P_U(x) = 2D(x) - x, \tag{3.1.11}$$

so that

$$D(x) = \frac{1}{2}\left(P_U(R_K(x)) + x\right). \tag{3.1.12}$$

Combining (3.1.10) and (3.1.12),

$$D(x) = \frac{1}{2}x + \frac{1}{4}R_U(R_K(x)) + \frac{1}{4}R_K(x),$$

i.e., $D(x)$ is a convex combination of $x, R_U(R_K(x))$ and $R_K(x)$. Since these three points belong to E, the same holds for $D(x)$ and the claim holds. We observe now that $x \in U$ by assumption, $D(x) \in U$ by definition, and $C(x) \in U$ by Proposition 2.4.3. Now we consider three cases: if $\dim(E\cap U) = 0$ then $x, D(x)$ and $C(x)$ coincide and the result holds trivially. If $\dim(E \cap U) = 2$ then $E \subset U$, so that $R_K(x) \in U$ so that $R_U(R_K(x)) = R_K(x)$, in which case $C(x)$ is the midpoint between x and $R_K(x)$, which is precisely $P_K(x)$. Hence, $P_K(x) \in U$, so that $D(x) = P_U(P_K(x)) = P_K(x) = C(x)$, implying that $D(x)$ and $C(x)$ coincide, and the result holds trivially. The interesting case is the remaining one, *i.e.*, $\dim(E \cap U) = 1$. In this case $x, D(x)$ and $C(x)$ lie in a line, so that we can write $C(x) = x + \eta(D(x) - x)$ with $\eta \in \mathbb{R}$, and it suffices to prove that $\eta \geq 1$. By the definition of η,

$$\|C(x) - x\| = |\eta|\ \|D(x) - x\|. \tag{3.1.13}$$

In view of (2.2.2) with $M = U, y = C(x)$ and $x = R_K(x)$,

$$\|C(x) - R_K(x)\| \geq \|C(x) - P_U(R_K(x))\| . \tag{3.1.14}$$

Then

$$\begin{aligned}\|C(x) - x\| &= \|C(x) - R_K(x)\| \geq \|C(x) - P_U(R_K(x))\| \\ &= \|(C(x) - x) - (P_U(R_K(x)) - x)\| \\ &= \|\eta\,(D(x) - x) - 2\,(D(x) - x)\| \\ &= |\eta - 2|\; \|D(x) - x\| ,\end{aligned} \tag{3.1.15}$$

using the definition of the circumcenter in the first equality, (3.1.14) in the inequality, and (3.1.11), as well as the definition of η, in the third equality. Combining (3.1.13) and (3.1.15), we get

$$|\eta|\; \|D(x) - x\| \geq |\eta - 2|\; \|D(x) - x\| ,$$

implying that $|\eta| \geq |2 - \eta|$, which holds only when $\eta \geq 1$, completing the proof. ■

Next, we show that, under Assumption EB, CRM achieves a linear rate with an asymptotic constant better than the one given in Corollary 3.1.1.

Proposition 3.1.3. Let $\{x^k\}_{k\in\mathbb{N}}$ be the sequence generated by CRM starting at any $x^0 \in U$. If K, U satisfy Assumption EB, then the sequence $\{\operatorname{dist}(x^k, K \cap U)\}_{k\in\mathbb{N}}$ converges to 0 with the asymptotic constant bounded above by $\sqrt{(1-\omega^2)/(1+\omega^2)}$, where ω is as in Assumption EB.

Proof. Take $y^* \in K \cap U$ and $x \in U\backslash K$. Note that

$$\begin{aligned}\operatorname{dist}^2(x, K) &= \|x - P_K(x)\|^2 \\ &\leq \|x - y^*\|^2 - \|P_K(x) - y^*\|^2 \\ &= \|x - y^*\|^2 - \|P_K(x) - P_K(y^*)\|^2 \\ &\leq \|x - y^*\|^2 - \|P_U(P_K(x)) - P_U(P_K(y^*))\|^2 - \|P_U(P_K(x)) - P_K(x)\|^2 \\ &= \|x - y^*\|^2 - \|P_U(P_K(x)) - y^*\|^2 - \|P_U(P_K(x)) - P_K(x)\|^2 ,\end{aligned} \tag{3.1.16}$$

using the definition of orthogonal projection onto K and the fact that $y^* \in K$ and Proposition 2.2.2 in the first inequality, and again Proposition 2.2.2 regarding U in the second inequality. Now, we will invoke Proposition 2.4.3 for proving that $C(x)$ is indeed the orthogonal projection of x onto the intersection of U with the half-space $H_x^+ : \{y \in \mathbb{R}^n \mid \langle y - P_K(x), x - P_K(x)\rangle \leq 0\}$ containing K. In view of Proposition 2.4.3, $C(x) = P_{H_x \cap U}(x)$ with $H_x := \{y \in \mathbb{R}^n : \langle y - P_K(x), x - P_K(x)\rangle = 0\} \supset K$.
Using the fact that $P_{H_x\cap U}(x) = P_{H_x^+\cap U}(x)$, we get

$$C(x) = P_{H_x\cap U}(x) = P_{H_x^+\cap U}(x).$$

Hence, the above equality and the fact that $y^* \in K \cap U \subset H_x^+ \cap U$ imply

$$\langle y^* - C(x), x - C(x)\rangle \leq 0,$$

and since x, $P_U(P_K(x))$ and $C(x)$ are collinear (see Proposition 3.1.2), we get

$$\langle y^* - C(x), P_U(P_K(x)) - C(x)\rangle \leq 0.$$

Thus,

$$\|P_U(P_K(x)) - y^*\|^2 \geq \|C(x) - y^*\|^2 + \|C(x) - P_U(P_K(x))\|^2. \tag{3.1.17}$$

Now, (3.1.17) and (3.1.16) imply

$$\begin{aligned}
\operatorname{dist}^2(x, K) &\leq \|x - y^*\|^2 - \|C(x) - y^*\|^2 - \|C(x) - P_U(P_K(x))\|^2 - \|P_U(P_K(x)) - P_K(x)\|^2 \\
&= \|x - y^*\|^2 - \|C(x) - y^*\|^2 - \|C(x) - P_K(x)\|^2 \\
&\leq \|x - y^*\|^2 - \|C(x) - y^*\|^2 - \operatorname{dist}^2(C(x), K) \\
&\leq \|x - y^*\|^2 - \operatorname{dist}^2(C(x), K \cap U) - \operatorname{dist}^2(C(x), K),
\end{aligned} \tag{3.1.18}$$

using the definition of the distance in the last two inequalities. Now, taking $y^* = P_{K\cap U}(x)$ and using the error bound condition for x and $C(x)$, we obtain

$$\begin{aligned}
\omega^2 \operatorname{dist}^2(x, K \cap U) &\leq \operatorname{dist}^2(x, K) \\
&\leq \operatorname{dist}^2(x, K \cap U) - \operatorname{dist}^2(C(x), K \cap U) - \omega^2 \operatorname{dist}^2(C(x), K \cap U) \\
&= \operatorname{dist}^2(x, K \cap U) - (1 + \omega^2) \operatorname{dist}^2(C(x), K \cap U).
\end{aligned} \tag{3.1.19}$$

Rearranging (3.1.19), we get

$$(1 + \omega^2) \operatorname{dist}^2(C(x), K \cap U) \leq (1 - \omega^2) \operatorname{dist}^2(x, K \cap U), \tag{3.1.20}$$

and since $x^{k+1} = C(x^k)$, from (3.1.20) we have

$$\frac{\operatorname{dist}(x^{k+1}, K \cap U)}{\operatorname{dist}(x^k, K \cap U)} \leq \sqrt{\frac{1 - \omega^2}{1 + \omega^2}}, \tag{3.1.21}$$

which implies the result. ■

Propositions 3.1.1 and 3.1.3 do not entail immediately that the sequences $\{z^k\}_{k\in\mathbb{N}}, \{x^k\}_{k\in\mathbb{N}}$ themselves converge linearly; a sequence $\{y^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ may converge to a point $y \in M \subset \mathbb{R}^n$, in such a way that $\{\operatorname{dist}(y^k, M)\}_{k\in\mathbb{N}}$ converges linearly to 0 but $\{y^k\}_{k\in\mathbb{N}}$ itself converges sublinearly. Take for instance $M = \{(s, 0) \in \mathbb{R}^2\}$, $y^k = \left(1/k, 2^{-k}\right)$. This sequence converges to $0 \in M$, $\operatorname{dist}(y^k, M) = 2^{-k}$ converges linearly to 0 with asymptotic constant equal to 1/2, but the first component of y^k converges to 0 sublinearly, and hence the same holds for the sequence $\{y^k\}_{k\in\mathbb{N}}$. The next lemma, possibly of some interest on its own, establishes that this situation cannot occur when $\{y^k\}_{k\in\mathbb{N}}$ is Fejér monotone with respect to M. the result below is similar to [8, Theorem 5.12], but we include its proof for the sake of completeness.

Lemma 3.1.1. Consider a nonempty closed convex set $M \subset \mathbb{R}^n$, and $\{y^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$. Assume that $\{y^k\}_{k\in\mathbb{N}}$ is Fejér monotone with respect to M, and that $\{\operatorname{dist}(y^k, M)\}_{k\in\mathbb{N}}$ converges R-linearly to 0. Then $\{y^k\}_{k\in\mathbb{N}}$ converges R-linearly to some point $y^* \in M$, with asymptotic constant bounded above by the asymptotic constant of $\{\operatorname{dist}(y^k, M)\}_{k\in\mathbb{N}}$.

Proof. Fix $k \in \mathbb{N}$ and note that the Fejér monotonicity hypothesis implies that, for all $j \geq k$,

$$\left\|y^j - P_M(y^k)\right\| \leq \left\|y^k - P_M(y^k)\right\| = \operatorname{dist}(y^k, M). \tag{3.1.22}$$

By Proposition 2.1.2(i), $\{y^k\}_{k\in\mathbb{N}}$ is bounded. Take any cluster point $\bar{y}$ of $\{y^k\}_{k\in\mathbb{N}}$. Taking limits with $j \to \infty$ in (3.1.22) along a subsequence $\{y^{k_j}\}_{j\in\mathbb{N}}$ of $\{y^k\}_{k\in\mathbb{N}}$ converging to $\bar{y}$, we get that $\left\|\bar{y} - P_M(y^k)\right\| \leq \operatorname{dist}(y^k, M)$. Since $\lim_{k\to\infty} \operatorname{dist}(y^k, M) = 0$, we conclude that $\{P_M(y^k)\}_{k\in\mathbb{N}}$ converges to $\bar{y}$, so that there exists a unique cluster point, say y^*. Therefore, $\lim_{k\to\infty} y^k = y^*$, and hence $\left\|y^* - P_M(y^k)\right\| \leq \operatorname{dist}(y^k, M)$. Since $y^* = \lim_{k\to\infty} P_M(y^k)$, we conclude that $y^* \in M$. Observe further that

$$\begin{aligned}\left\|y^k - y^*\right\| &\leq \left\|y^k - P_M(y^k)\right\| + \left\|P_M(y^k) - y^*\right\| \\ &= \operatorname{dist}(y^k, M) + \left\|y^* - P_M(y^k)\right\| \leq 2\operatorname{dist}(y^k, M).\end{aligned} \tag{3.1.23}$$

Taking kth-root and then $\limsup$ with $k \to \infty$ in (3.1.23), and using the R-linearity hypothesis,

$$\limsup_{k\to\infty} \left\|y^k - y^*\right\|^{1/k} \leq \limsup_{k\to\infty} 2^{1/k} \operatorname{dist}(y^k, M)^{1/k} \tag{3.1.24}$$

$$= \limsup_{k\to\infty} \operatorname{dist}(y^k, M)^{1/k} < 1, \tag{3.1.25}$$

establishing both that $\{y^k\}_{k\in\mathbb{N}}$ converges R-linearly to $y^* \in M$ and the statement on the asymptotic constant. ■

With the help of Lemma 3.1.1, we prove next R-linear convergence of the MAP and CRM sequences under Assumption EB, and give bounds for their asymptotic constants.

Theorem 3.1.1. Consider a closed convex set $K \subset \mathbb{R}^n$ and an affine manifold $U \subset \mathbb{R}^n$. Assume that K, U satisfy Assumption EB. Let $\{z^k\}_{k\in\mathbb{N}}$ and $\{x^k\}_{k\in\mathbb{N}}$ be the sequences generated by MAP and CRM, respectively, starting from arbitrary points $z^0 \in \mathbb{R}^n, x^0 \in U$. Then both sequences $\{z^k\}_{k\in\mathbb{N}}$ and $\{x^k\}_{k\in\mathbb{N}}$ converge R-linearly to points in $K \cap U$, and the asymptotic constants are bounded above by $\sqrt{1-\omega^2}$ for MAP, and by $\sqrt{(1-\omega^2)/(1+\omega^2)}$ for CRM, with ω as in Assumption EB.

Proof. In view of Propositions 2.4.1 and 2.4.4, both sequences generated by MAP and CRM, are Fejér monotone with respect to $K \cap U$ and converge to points in $K \cap U$. By Corollary 3.1.1, both sequences $\{\operatorname{dist}(z^k, K \cap U)\}_{k\in\mathbb{N}}$ and $\{\operatorname{dist}(x^k, K \cap U)\}_{k\in\mathbb{N}}$ are Q-linearly convergent to 0, and henceforth R-linearly convergent to 0. Corollary 3.1.1 shows that the asymptotic constant of the sequence $\{\operatorname{dist}(z^k, K \cap U)\}_{k\in\mathbb{N}}$ is bounded above by $\sqrt{1-\omega^2}$, and Proposition 3.1.3 establishes that the asymptotic constant of the sequence $\{\operatorname{dist}(x^k, K \cap U)\}_{k\in\mathbb{N}}$ is bounded above by $\sqrt{(1-\omega^2)/(1+\omega^2)}$. Finally, by Lemma 3.1.1, both $\{z^k\}_{k\in\mathbb{N}}$ and $\{x^k\}_{k\in\mathbb{N}}$ are R-linearly convergent, with the announced bounds for their asymptotic constants. ■

We remark that, in view of Theorem 3.1.1, the upper bound for the asymptotic constant of the CRM sequence is substantially better than the one for the MAP sequence. Note that the CRM bound reduces the MAP one by a factor of $\sqrt{1+\omega^2}$, which increases up to $\sqrt{2}$ when ω approaches 1.

3.2 Two families of examples for which CRM is much faster than MAP

In this section, we will present two rather generic families of examples for which CRM is faster than MAP. In the first one, MAP converges sublinearly while CRM converges linearly; in the second one, MAP converges linearly and CRM converges superlinearly.
In both families, we work in $\mathbb{R}^{n+1}$. K will be the epigraph of a proper convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, and U the hyperplane $\{(x, 0) : x \in \mathbb{R}^n\} \subset \mathbb{R}^{n+1}$. From now on, we consider f to be continuously differentiable in $\operatorname{int}(\operatorname{dom}(f))$, where $\operatorname{dom}(f) := \{x \in \mathbb{R}^n : f(x) < +\infty\}$ and $\operatorname{int}(\operatorname{dom}(f))$ is its topological interior. Next, we make the following assumptions on f:

A1. $0 \in \operatorname{int}(\operatorname{dom}(f))$ is the unique minimizer of f.

A2. $\nabla f(0) = 0$.

A3. $f(0) = 0$.

We will show that under these assumptions, MAP always converges sublinearly, while, adding an additional hypothesis, CRM converges linearly.
Note that under hypotheses A1 to A3, $0 \in \mathbb{R}^n$ is the unique zero of f and hence $K \cap U = \{(0, 0)\} \in \mathbb{R}^{n+1}$. In view of Propositions 2.4.1 and 2.4.4, the sequences generated by MAP and CRM, with arbitrary initial points in $\mathbb{R}^{n+1}$ and U, respectively, both converge to $(0, 0)$, and are Fejér monotone with respect to $\{(0, 0)\}$, so that, in view of A1, for large enough k the iterates of both sequences belong to $\operatorname{int}(\operatorname{dom}(f)) \times \mathbb{R}$. We take now any point $(x, 0) \in U$, with $x \neq 0$ and proceed to compute $P_K(x, 0)$. Since $(x, 0) \notin K$ (because $x \neq 0$ and $K \cap U = \{(0, 0)\}$), $P_K(x, 0)$ must belong to the boundary of K, *i.e.*, it must be of the form $(u, f(u))$, and u is determined by minimizing $\|(x, 0) - (u, f(u))\|^2$, so that $u - x + f(u)\nabla f(u) = 0$, or equivalently

$$x = u + f(u)\nabla f(u). \tag{3.2.1}$$

Note that since $x \neq 0$, $u \neq 0$ by A3. With the notation of Section 3.1 and bearing in mind that D and C are the MAP and CRM operators defined in (3.1.1), it is easy to check that

$$\begin{aligned} P_K(x, 0) &= (u, f(u)), \text{ and} \\ D(x, 0) &= P_U(P_K(x, 0)) = (u, 0), \end{aligned} \tag{3.2.2}$$

with u as in (3.2.1). Moreover,

$$\begin{aligned} R_K(x, 0) &= (2u - x, 2f(u)), \\ P_U\left(R_K(x, 0)\right) &= (2u - x, 0), \text{ and} \\ R_U(R_K(x, 0)) &= (2u - x, -2f(u)). \end{aligned}$$

Next we compute $C(x, 0) = \operatorname{circ}((x, 0), R_K(x, 0), R_U(R_K(x, 0)))$. Suppose that $C(x, 0) = (v, s)$. The conditions

$$\|(v, s) - (x, 0)\| = \|(v, s) - R_K(x, 0)\| = \|(v, s) - R_U(R_K(x, 0))\|$$

give rise to two quadratic equations whose solution is

$$s = 0, \qquad v = u - \left[\frac{f(u)}{\|x-u\|}\right]^2 (x-u) = u - \frac{f(u)}{\|\nabla f(u)\|^2}\nabla f(u), \tag{3.2.3}$$

using (3.2.1) in the last equality.
We proceed to compute the quotients $\|D(x,0)-0\| / \|(x,0)-0\|$, $\|C(x,0)-0\| / \|(x,0)-0\|$. Since both the MAP and the CRM sequences converge to 0, these quotients are needed for determining their convergence rates. In view of (3.2.2) and (3.2.3), these quotients reduce to $\|u\| / \|x\|$, $\|v\| / \|x\|$. We state the result of the computation of these quotients in the next proposition.

Proposition 3.2.1. Take $(x,0) \in U$ with $x \neq 0$. Let $D(x,0) = (u,0)$ and $C(x,0) = (v,0)$. Then,

$$\frac{\|D(x,0)\|}{\|(x,0)\|} = \frac{\|u\|}{\|x\|} = \frac{1}{\left\|\bar{u} + \frac{f(u)}{\|u\|}\nabla f(u)\right\|} \tag{3.2.4}$$

with $\bar{u} = u / \|u\|$,

$$\left[\frac{\|C(x,0)\|}{\|D(x,0)\|}\right]^2 = \left[\frac{\|v\|}{\|u\|}\right]^2 \leq 1 - \left[\frac{f(u)}{\|u\| \, \|\nabla f(u)\|}\right]^2, \tag{3.2.5}$$

and

$$\left[\frac{\|C(x,0)\|}{\|(x,0)\|}\right]^2 \leq \left[1 - \left(\frac{f(u)}{\|u\| \, \|\nabla f(u)\|}\right)^2\right] \left[\frac{\|u\|}{\|x\|}\right]^2. \tag{3.2.6}$$

Proof. In view of (3.2.1),

$$\frac{\|u\|}{\|x\|} = \frac{\|u\|}{\|u + f(u)\nabla f(u)\|}$$

and (3.2.4) follows by dividing the numerator and the denominator by $\|u\|$.
We proceed to establish (3.2.5). In view of (3.2.3), we have

$$\begin{aligned}
\|v\|^2 &= \left\|u - \frac{f(u)}{\|\nabla f(u)\|^2}\nabla f(u)\right\|^2 \\
&= \|u\|^2 + \left[\frac{f(u)}{\|\nabla f(u)\|}\right]^2 - 2\frac{f(u)}{\|\nabla f(u)\|^2}\langle \nabla f(u), u\rangle \\
&\leq \|u\|^2 + \left[\frac{f(u)}{\|\nabla f(u)\|}\right]^2 - 2\left[\frac{f(u)}{\|\nabla f(u)\|}\right]^2 \\
&= \|u\|^2 - \left[\frac{f(u)}{\|\nabla f(u)\|}\right]^2,
\end{aligned} \tag{3.2.7}$$

using the gradient inequality $\langle \nabla f(u), u\rangle \geq f(u)$, which holds because f is convex and $f(0) = 0$. Now, (3.2.5) follows by dividing (3.2.7) by $\|u\|^2$. Finally, (3.2.6) follows by multiplying (3.2.5) by $\frac{\|u\|^2}{\|x\|^2} = \frac{\|D(x,0)\|^2}{\|(x,0)\|^2}$. ■

Next, we compute the limits with $x \to 0$ of the quotients in Proposition 3.2.1.

Proposition 3.2.2. Take $(x, 0) \in U$ with $x \neq 0$. Let $D(x, 0) = (u, 0)$ and $C(x, 0) = (v, 0)$. Then,

$$\limsup_{x\to 0} \frac{\|D(x,0)\|}{\|(x,0)\|} = \lim_{x\to 0} \frac{\|u\|}{\|x\|} = 1 \tag{3.2.8}$$

and

$$\limsup_{x\to 0} \left[\frac{\|C(x,0)\|}{\|(x,0)\|}\right]^2 = \limsup_{x\to 0} \left[\frac{\|v\|}{\|x\|}\right]^2 \leq 1 - \liminf_{x\to 0} \left[\frac{f(x)}{\|x\| \, \|\nabla f(x)\|}\right]^2. \tag{3.2.9}$$

Proof. By convexity of f, using A3, $f(y) \leq \langle \nabla f(y), y\rangle \leq \|\nabla f(y)\| \, \|y\|$ for all $y \in \operatorname{int}(\operatorname{dom}(f))$ sufficiently close to 0. Hence, for all nonzero $y \in \operatorname{int}(\operatorname{dom}(f))$, $0 < f(y)/\|y\| \leq \|\nabla f(y)\|$, using A1 and A3. Since $\lim_{y\to 0} \nabla f(y) = 0$ by A1 and A2 and the convexity of f, it follows that

$$\lim_{y\to 0} f(y)/\|y\| = 0. \tag{3.2.10}$$

Now we take limits with $x \to 0$ in (3.2.4). Since $(u, 0) = P_K((x, 0))$ and using the continuity of projections, $\lim_{x\to 0} u = 0$. Thus,

$$\begin{aligned}\limsup_{x\to 0} \frac{\|D(x,0)\|}{\|(x,0)\|} &= \lim_{x\to 0} \frac{\|D(x,0)\|}{\|(x,0)\|} = \lim_{x\to 0} \frac{1}{\left\|\bar{u} + \dfrac{f(u)}{\|u\|}\nabla f(u)\right\|} \\ &= \lim_{u\to 0} \frac{1}{\left\|\bar{u} + \dfrac{f(u)}{\|u\|}\nabla f(u)\right\|} = \frac{1}{\|\bar{u}\|} = 1,\end{aligned}$$

using (3.2.10) and the fact that $\|\bar{u}\| = \|u/\|u\|\| = 1$. We have proved that (3.2.8) holds. Now we deal with (3.2.9). Taking limits with $x \to 0$ in (3.2.6), we have

$$\limsup_{x\to 0} \left[\frac{\|C(x,0)\|}{\|(x,0)\|}\right]^2 \leq \left[1 - \liminf_{x\to 0} \left(\frac{f(u)}{\|u\| \, \|\nabla f(u)\|}\right)^2\right] \limsup_{x\to 0} \left[\frac{\|u\|}{\|x\|}\right]^2. \tag{3.2.11}$$

The second lim sup on the right-hand side of (3.2.11) is equal to $\lim_{x\to 0} \left[\frac{\|u\|}{\|x\|}\right]^2$ and by (3.2.8) it is equal 1 and so (3.2.9) follows from the already made observation that $\lim_{x\to 0} u = 0$. ■

We proceed to establish the convergence rates of the sequences generated by MAP and CRM for this choice of K and U.

Corollary 3.2.1. Consider $K, U \subset \mathbb{R}^{n+1}$ given by $K = \operatorname{epi}(f)$, with $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ satisfying A1 to A3, and also $U := \{(x, 0) \mid x \in \mathbb{R}^n\} \subset \mathbb{R}^{n+1}$. Let $\{(z^k, 0)\}_{k\in\mathbb{N}}$ and $\{(x^k, 0)\}_{k\in\mathbb{N}}$, be the sequences generated by MAP and CRM, starting from $(z^0, 0) \in \mathbb{R}^{n+1}$ and $(x^0, 0) \in U$, respectively. Then,

$$\limsup_{k\to\infty} \frac{\left\|(z^{k+1}, 0)\right\|}{\|(z^k, 0)\|} = 1$$

and

$$\limsup_{k\to\infty} \frac{\left\|(x^{k+1},0)\right\|}{\|(x^k,0)\|} \leq \sqrt{1-\gamma^2}, \tag{3.2.12}$$

with

$$\gamma := \liminf_{x\to 0} \frac{f(x)}{\|x\| \, \|\nabla f(x)\|}. \tag{3.2.13}$$

Proof. Since $D(z^k,0) = (z^{k+1},0)$ and $C(x^k,0) = (x^{k+1},0)$, and $\lim_{k\to\infty} x^k = \lim_{k\to\infty} z^k = 0$, it suffices to apply Proposition 3.2.2 with $x = z^k$ in (3.2.8) and $x = x^k$ in (3.2.9). ■

We add now an additional hypothesis on f.

A4. f satisfies $\displaystyle\liminf_{x\to\infty} \frac{f(x)}{\|x\| \, \|\nabla f(x)\|} > 0$.

Observe that by using the convexity of f and Cauchy-Schwarz inequality, we have

$$0 \leq f(x) \leq \langle \nabla f(x), x\rangle \leq \|\nabla f(x)\| \, \|x\|.$$

Thus, A1 to A3, imply $f(x)/(\|x\| \, \|\nabla f(x)\|) \in (0,1]$, for all $x \neq 0$, so that A4 just excludes the case in which the lim inf above is equal to 0. Next we rephrase Corollary 3.2.1.

Corollary 3.2.2. Consider $K, U \subset \mathbb{R}^{n+1}$ given by $K = \operatorname{epi}(f)$, with $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ satisfying A1 to A3, and $U := \{(x,0) \mid x \in \mathbb{R}^n\} \subset \mathbb{R}^{n+1}$. Then the sequence generated by MAP from an arbitrary initial point converges sublinearly. If f also satisfies hypothesis A4, then the sequence generated by CRM from an initial point in U converges linearly, and its asymptotic constant is bounded above by $\sqrt{1-\gamma^2} < 1$, with $\gamma > 0$ as in (3.2.13).

Proof. Immediate from Corollary 3.2.1 and hypothesis A4. ■

Next we discuss several situations for which hypothesis A4 holds, showing that it is rather generic. The first case is as follows.

Proposition 3.2.3. Assume that f, besides satisfying A1 to A3, is of class $\mathcal{C}^2$ (around $0 \in \mathbb{R}^n$) and $\nabla^2 f(0)$ is nonsingular. Then, assumption A4 holds, and $\gamma \geq \lambda_{\min}/(2\lambda_{\max}) > 0$ where $\lambda_{\min}, \lambda_{\max}$ are the smallest and largest eigenvalues of $\nabla f^2(0)$, respectively.

Proof. In view of A2, A3 and the hypothesis on $\nabla f^2(0)$, we have

$$f(x) = \frac{1}{2}\left\langle x, \nabla^2 f(0)x\right\rangle + o(\|x\|^2) \geq \frac{\lambda_{\min}}{2}\|x\|^2 + o(\|x\|^2). \tag{3.2.14}$$

Also, using the Taylor expansion of ∇f around $x = 0$, $\nabla f(x) = \nabla^2 f(0)x + o(\|x\|)$, so that

$$\begin{aligned}\|x\| \, \|\nabla f(x)\| &= \|x\| \left\|\nabla^2 f(0)x\right\| + o(\|x\|^2)\\ &\leq \|x\|^2 \left\|\nabla^2 f(0)\right\| + o(\|x\|^2)\\ &\leq \lambda_{\max}\|x\|^2 + o(\|x\|^2).\end{aligned} \tag{3.2.15}$$

By (3.2.14) and (3.2.15),

$$\frac{f(x)}{\|x\|\ \|\nabla f(x)\|} \geq \frac{\lambda_{\min}\|x\|^2 + o(\|x\|^2)}{2\lambda_{\max}\|x\|^2 + o(\|x\|^2)} \tag{3.2.16}$$

the result follows by taking lim inf in (3.2.16), since the right hand converges to $\dfrac{\lambda_{\min}}{2\lambda_{\max}} > 0$ as $x \to 0$. ■

Note that nonsingularity of $\nabla^2 f(0)$ holds when f is of class $\mathcal{C}^2$ and strongly convex.
We consider next other instances for which assumption A4 holds. Now we deal with the case in which $f(x) = \phi(\|x\|)$ with $\phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, satisfying A1 to A3. This case has a one dimensional flavor, and computations are easier. The first point to note is that

$$\liminf_{x\to 0} \frac{f(x)}{\|x\|\ \|\nabla f(x)\|} = \lim_{t\to 0} \frac{\phi(t)}{t\phi'(t)}, \tag{3.2.17}$$

so that assumption A4 becomes:

A4′. ϕ satisfies $\liminf_{t\to 0} \dfrac{\phi(t)}{t\phi'(t)} > 0$.

More importantly, in this case $\nabla f(x)$ and x are collinear, which allows for an improvement in the asymptotic constant: we will have $1 - \gamma$ instead of $\sqrt{1-\gamma^2}$ in (3.2.12), as we show next. We reformulate Propositions 3.2.1 and 3.2.2 for this case.

Proposition 3.2.4. Assume that $f(x) = \phi(\|x\|)$, with $\phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ satisfying A1 to A3 and A4′. Take $(x, 0) \in U$ with $x \neq 0$. Let $C(x, 0) = (v, 0)$. Then,

$$\text{(i)} \qquad \frac{\|C(x,0)\|}{\|D(x,0)\|} = \frac{\|v\|}{\|u\|} = 1 - \frac{\phi(\|u\|)}{\phi'(\|u\|)\,\|u\|}, \tag{3.2.18}$$

$$\text{(ii)} \qquad \limsup_{x\to 0} \frac{\|C(x,0)\|}{\|(x,0)\|} = 1 - \liminf_{x\to 0} \frac{f(x)}{\|x\|\ \|\nabla f(x)\|} = 1 - \liminf_{t\to 0} \frac{\phi(t)}{t\phi'(t)}. \tag{3.2.19}$$

Proof. In this case

$$\nabla f(x) = \frac{\phi'(\|x\|)}{\|x\|} x$$

so that (3.2.1) becomes

$$x = \left(1 + \frac{\phi(u)\phi'(u)}{\|u\|}\right) u,$$

and (3.2.3) can be rewritten as

$$v = \left(1 - \frac{\phi(\|u\|)}{\phi'(\|u\|)\,\|u\|}\right) u.$$

Hence,

$$\frac{\|v\|}{\|u\|} = 1 - \frac{\phi(\|u\|)}{\phi'(\|u\|)\,\|u\|},$$

establishing (3.2.18). Then, (3.2.19) follows from (3.2.18) as in the proofs of Propositions 3.2.1 and 3.2.2, taking into account (3.2.17). ■

Corollary 3.2.3. Let $\{(x^k, 0)\}_{k\in\mathbb{N}}$ be the sequence generated by CRM with $(x^0, 0) \in U$. Assume that $f(x) = \phi(\|x\|)$ with $\phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ satisfying A1 to A3. Then,

$$\limsup_{k\to\infty} \frac{\left\|(x^{k+1}, 0)\right\|}{\|(x^k, 0)\|} = 1 - \hat{\gamma},$$

with

$$\hat{\gamma} := \liminf_{t\to 0} \frac{\phi(t)}{t\phi'(t)}. \tag{3.2.20}$$

If ϕ satisfies hypothesis A4′ then, the CRM sequence is Q-linearly convergent, with asymptotic constant equal to $1 - \hat{\gamma}$.

Proof. It is an immediate consequence of Proposition 3.2.4(ii), in view of the definition of the circumcenter operator C, given in (3.1.1). ■

We verify next that assumption A4′ is rather generic. It holds, *e.g.*, if ϕ is analytic around 0.

Proposition 3.2.5. If ϕ satisfies A1 to A3, and is analytic around 0 then it satisfies A4′, and $\hat{\gamma} = 1 - 1/p$, where $p := \min\{j : \phi^{(j)}(0) \neq 0\}$.

Proof. In this case

$$\phi(t) = (1/p!)\phi^{(p)}(0)t^p + o(t^{p+1})$$

and

$$t\phi'(t) = (1/(p-1)!)\phi^{(p)}(0)t^p + o(t^{p+1}),$$

and the result follows taking limits with $t \to 0$, taking into account (3.2.20). ■

Note that for an analytic ϕ the asymptotic constant is always of the form $1-1/p$ with $p \in \mathbb{N}$. This is not the case in general. Take, *e.g.*, $\phi(t) = |t|^\alpha$ with $\alpha \in \mathbb{R}$, $\alpha > 1$. Then a simple computation shows that $\hat{\gamma} = 1/\alpha$. Note that ϕ is of class $\mathcal{C}^p$, where p is the integer part of α, but not of class $\mathcal{C}^{p+1}$, so that Proposition 3.2.5 does not apply.
Take now

$$f(x) = \begin{cases} 1 - \sqrt{1 - \|x\|^2}, & \text{if } \|x\| \leq 1, \\ +\infty, & \text{otherwise,} \end{cases}$$

i.e., $f(x) = \phi(\|x\|)$ with $\phi(t) = 1 - \sqrt{1 - t^2}$, when $t \in [-1, 1], \phi(t) = +\infty$ otherwise. Note that f satisfies A1 to A3 and its effective domain is the unit ball in $\mathbb{R}^n$. Since ϕ is analytic around 0 and $\phi''(0) \neq 0$, we get from Proposition 3.2.5 that $\hat{\gamma} = 1/2$ and so the asymptotic constant of the CRM sequence is also 1/2. Note that the graph of f is the lower hemisphere of the ball $B \subset \mathbb{R}^{n+1}$ centered at $(0, 1)$ with radius 1. Observe also that the projection onto B of a point of the form $(x, 0) \in \mathbb{R}^{n+1}$ is of the form (u, t) with $t < 1$, so it belongs to $\operatorname{epi}(f)$. Hence, the sequences generated by CRM for the pair K, U with $K = \operatorname{epi}(f)$ and $K = B$ coincide. It follows easily that the sequence generated by CRM for a pair K, U where K is any ball and U is a hyperplane tangent to the ball, converges linearly, with asymptotic constant equal to 1/2. We remark that in all these cases, the sequence generated by MAP converges sublinearly, by virtue of Corollary 3.2.2.

We look now at a case where hypothesis A4′ fails. Define

$$f(x) = \begin{cases} e^{-\|x\|^{-2}}, & \text{if } \|x\| \le \frac{1}{\sqrt{3}}, \\ +\infty, & \text{otherwise.} \end{cases}$$

so that $f(x) = \phi(\|x\|)$ with $\phi(t) = e^{-1/t^2}$, when $t \in (-3^{-1/2}, 3^{-1/2}), \phi(t) = +\infty$ otherwise. Again f satisfies A1 to A3. It is easy to check that $\phi(t)/(t\phi'(t)) = (1/2)t^2$, so that it follows immediately that $\lim_{t\to 0} \phi(t)/(t\phi'(t)) = 0$ and A4′ fails. It is known that this ϕ, which is of class $\mathcal{C}^\infty$ but not analytic, is extremely flat (in fact, $f^{(k)}(0) = 0$ for all k), and not even CRM can overcome so much flatness; in view of Corollary 3.2.3, in this case it converges sublinearly, as MAP does. The examples above are also presented as a study case in [9], illustrating the slow convergence of the proximal point algorithm, Douglas-Rachford algorithm and alternating projections.
Let us abandon such an appalling situation, and move over to other examples where CRM will be able to exhibit again its superiority; next, we deal with our second family of examples. In this case we keep the framework of the first family with just one change, namely in hypothesis A3 on f; now we will request that $f(0) < 0$. With this single trick (and a couple of additional technical assumptions), we will achieve linear convergence of the MAP sequence and superlinear convergence of the CRM one. We will assume also that the effective domain of f is the whole space (differently from the previous section, we don't have now interesting examples with smaller effective domains; also, since now the limit of the sequences can be anywhere, a hypothesis on the effective domain becomes rather cumbersome). We'll also demand that f be of class $\mathcal{C}^2$.
Finally, we will restrict ourselves to the case of $f(x) = \phi(\|x\|)$, with $\phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$. This assumption is not essential, but will considerably simplify our analysis. Thus, we rewrite the assumptions for ϕ, in this new context. We assume that function ϕ is proper, strictly convex and twice continuously differentiable, satisfying

A2′. $\phi'(0) = 0$.

A3′. $\phi(0) < 0$.

In the remainder of this Chapter we will study the behavior of the MAP and CRM sequences for the pair $K, U \subset \mathbb{R}^{n+1}$, where K is the epigraph of $f(x) = \phi(\|x\|)$, with ϕ satisfying hypotheses A2′ and A3′ above, and $U := \{(x, 0) : x \in \mathbb{R}^n\} \subset \mathbb{R}^{n+1}$. As in the previous case, Propositions 2.4.1 and 2.4.4, ensure that both sequences converge to points in $K \cap U$. Since we are dealing with convergence rates, we will exclude the case in which the sequences of interest have finite convergence. We continue with an elementary property of the limit of these sequences.

Proposition 3.2.6. Assume that K, U are as above. Let $(x^*, 0)$ be the limit of either the MAP or the CRM sequences and $t^* := \|x^*\|$. Then, $\phi(t^*) = 0$ and $\phi'(t^*) > 0$.

Proof. Since these sequences stay in U, remain outside K (otherwise convergence would be finite), and converge to points in $K \cap U$, it follows that their limits must belong to $\mathrm{bd}(K) \cap U$, where $\mathrm{bd}(K) := \{(x, f(x) : x \in \mathbb{R}^n\}$ denotes the boundary of K. So, we conclude that $0 = f(x^*) = \phi(t^*)$. Now, since $\phi'(0) = 0$, in view of A2$'$, and ϕ' is strictly increasing, we conclude that $\phi'(t) > 0$ for all $t > 0$. Note that $x^* \neq 0$, because $f(x^*) = 0$ and $f(0) < 0$ by A3$'$. Hence $t^* = \|x^*\| > 0$, so that $\phi'(t^*) > 0$. ■

Now we analyze the behavior of the operators C and D in this case.

Proposition 3.2.7. Assume that $K, U \subset \mathbb{R}^{n+1}$ are defined as

$$U := \{(x, 0) : x \in \mathbb{R}^n\} \subset \mathbb{R}^{n+1},$$

and

$$K = \operatorname{epi}(f),$$

where

$$f(x) = \phi(\|x\|),$$

and ϕ satisfies A2$'$ and A3$'$. Let D and C be the operators associated to MAP and CRM respectively, and $(z^*, 0)$ and $(x^*, 0)$ the limits of the sequences $\{z^k\}_{k\in\mathbb{N}}$ and $\{x^k\}_{k\in\mathbb{N}}$ generated by these methods, starting from some $(z^0, 0) \in \mathbb{R}^{n+1}$, and some $(x^0, 0) \in U$, respectively. Then,

$$\limsup_{x\to z^*} \frac{\|D(x,0) - (z^*,0)\|}{\|(x,0) - (z^*,0)\|} = \frac{1}{1 + \phi'(\|z^*\|)^2} \tag{3.2.21}$$

and

$$\limsup_{x\to z^*} \frac{\|C(x,0) - (x^*,0)\|}{\|(x,0) - (x^*,0)\|} = 0. \tag{3.2.22}$$

Proof. Since, in this case, $\nabla f(x) = \dfrac{\phi'(\|x\|)}{\|x\|} x$ for all $x \neq 0$, we rewrite (3.2.1) and (3.2.3) as

$$x = (1 + \frac{\phi(\|u\|)\phi'(\|u\|)}{\|u\|})u \tag{3.2.23}$$

and

$$v = \left(1 - \frac{\phi(\|u\|)}{\phi'(\|u\|)\,\|u\|}\right) u. \tag{3.2.24}$$

In view of (3.2.23) and (3.2.24), u, v and x are collinear. In terms of the operators C and D, we have that $x, C(x)$ and $D(x)$ are collinear, so the same holds for the whole sequences generated by MAP, CRM and hence also for their limits $(z^*, 0), (x^*, 0)$. This is a consequence of the one-dimensional flavor of this family of examples. So, we define $s := \|z^*\|$, $t := \|x^*\|$, $r := \|u\|$, and therefore we get $u = (r/s)z^* = (r/t)x^*$. We compute next the quotients

$$\frac{\|(D(x),0) - (z^*,0)\|}{\|(x,0) - (z^*,0)\|} = \frac{\|u - z^*\|}{\|x - z^*\|}$$

and

$$\frac{\|(C(x),0) - (x^*,0)\|}{\|(x,0) - (x^*,0)\|} = \frac{\|v - x^*\|}{\|x - x^*\|},$$

needed for determining the convergence rate of the MAP and CRM sequences. We start with the MAP case. $s\left|\frac{r}{s}-1+\frac{\phi(r)\phi'(r)}{s}\right|$

$$\frac{\|D(x,0)-(z^*,0)\|}{\|(x,0)-(z^*,0)\|}=\frac{\|u-z^*\|}{\|x-z^*\|}=\frac{s\left|\frac{r}{s}-1\right|}{s\left|\frac{r}{s}-1+\phi(r)\phi'(r)\right|}$$
$$=\frac{|r-s|}{|r-s+s\phi'(r)\phi(r)|}$$
$$=\frac{1}{\left|1+\phi'(r)\left(\dfrac{\phi(r)-\phi(s)}{r-s}\right)\right|}, \tag{3.2.25}$$

using (3.2.23) in the second equality and the fact that $s=\phi(\|z^*\|)=f(z^*)=0$, established in Proposition 3.2.6, in the fourth one.
Now, we perform a similar computation for the operator C, needed for the CRM sequence.

$$\frac{\|C(x,0)-(x^*,0)\|}{\|(x,0)-(x^*,0)\|}=\frac{\|v-x^*\|}{\|x-x^*\|}=\frac{t\left|\left(1-\dfrac{\phi(r)}{\phi'(r)r}\right)\dfrac{r}{t}-1\right|}{t\left|\dfrac{r}{t}-1+\dfrac{\phi(r)\phi'(r)}{t}\right|}$$
$$=\frac{\left|\left(1-\dfrac{\phi(r)}{r\phi'(r)}\right)r-t\right|}{|r-t+\phi(r)\phi'(r)|}=\frac{\left|r-t-\dfrac{\phi(r)}{\phi'(r)}\right|}{|r-t+\phi(r)\phi'(r)|}$$
$$=\frac{\left|1-\dfrac{1}{\phi'(r)}\left(\dfrac{\phi(r)-\phi(t)}{r-t}\right)\right|}{\left|1+\phi'(r)\left(\dfrac{\phi(r)-\phi(t)}{r-t}\right)\right|}, \tag{3.2.26}$$

using (3.2.24) in the second equality, and Proposition 3.2.6, which implies $\phi(t)=0$, in the fifth one.
Finally, we take limits in (3.2.25) with $x\to z^*$ and in (3.2.26) with $x\to x^*$. Note that, since $u=P_K(x)$, $\lim_{x\to z^*}u=P_K(z^*)=z^*$, because $z^*\in K$. Hence, we take limit with $r\to s$ in the right-hand side of (3.2.25). We also take limits with $x\to x^*$ in (3.2.26). By the same token, taking limit with $r\to t$ in the right-hand side, we get

$$\limsup_{x\to z^*}\frac{\|D(x,0)-(z^*,0)\|}{\|(x,0)-(z^*,0)\|}=\limsup_{r\to s}\frac{1}{\left|1+\phi'(r)\left(\dfrac{\phi(r)-\phi(s)}{r-s}\right)\right|}$$
$$=\frac{1}{\left|1+\lim_{r\to s}\phi'(r)\left(\dfrac{\phi(r)-\phi(s)}{r-s}\right)\right|}$$
$$=\frac{1}{1+\phi'(s)^2} \tag{3.2.27}$$

and

$$\limsup_{x\to x^*}\frac{\|C(x,0)-(x^*,0)\|}{\|(x,0)-(x^*,0)\|}=\limsup_{r\to t}\frac{\left|1-\dfrac{1}{\phi'(r)}\left(\dfrac{\phi(r)-\phi(t)}{r-t}\right)\right|}{\left|1+\phi'(r)\left(\dfrac{\phi(r)-\phi(t)}{r-t}\right)\right|}$$

$$=\frac{\left|1-\lim_{r\to t}\dfrac{1}{\phi'(r)}\left(\dfrac{\phi(r)-\phi(t)}{r-t}\right)\right|}{\left|1+\lim_{r\to t}\phi'(r)\left(\dfrac{\phi(r)-\phi(t)}{r-t}\right)\right|}$$

$$=\frac{\left|1-\dfrac{\phi'(t)}{\phi'(t)}\right|}{|1+\phi'(t)^2|}=0. \qquad (3.2.28)$$

The results follow, in view of the definitions of s and t, from (3.2.27) and (3.2.28), respectively. ■

Note that the denominators in the expressions of (3.2.27) and (3.2.28) are the same; the difference lies in the numerators: in the MAP case it is 1; in the CRM one, the presence of the factor $(\phi(r)-\phi(t))/(r-t)$ makes the numerator go to 0 when r tends to t.

Corollary 3.2.4. Under the assumptions of Proposition 3.2.7 the sequence generated by MAP converges Q-linearly to a point $(z^*,0)\in K\cap U$, with asymptotic constant equal to $1/(1+\phi'(\|z^*\|)^2)$, and the sequence generated by CRM converges superlinearly.

Proof. The result for the MAP sequence follows from (3.2.21) in Proposition 3.2.7, observing that for $x=z^k$, we have $D(x,0)=(z^{k+1},0)$. Note that the asymptotic constant is indeed smaller than 1, because $z^*\neq 0$, and $\phi'(\|z^*\|)\neq 0$ by Proposition 3.2.6. The result for the CRM sequence follows from (3.2.22) in Proposition 3.2.7, observing that for $x=x^k$, we have $C(x,0)=(x^{k+1},0)$. ■

We now present an example that, although very simple, enables one to visualize how fast CRM is in comparison to MAP.

Example 3.2.1. Let $\phi:\mathbb{R}\to\mathbb{R}$, given by $\phi(t)=|t|^\alpha-\beta$, where $\alpha>1$ and $\beta\geq 0$. Consider $K,U\subset\mathbb{R}^2$ such that $K:=\operatorname{epi}(\phi)$ and U is the abscissa axis. Note that, if $\beta=0$, the error bound condition EB between K and U does not hold. For any $\beta>0$, though, it is easily verifiable that EB is valid. Figures 3.1 and 3.2, shows CRM and MAP tracking a point in $K\cap U$ up to a precision $\epsilon>0$, with the same starting point $(1.1,0)\in\mathbb{R}^2$. We fix $\alpha=2$ and take $\beta=0$ in Fig. 3.1, and $\beta=0.06$ in Fig. 3.2. We count and display the iterations of the MAP sequence $\{z^k\}_{k\in\mathbb{N}}$ and the CRM sequence $\{x^k\}_{k\in\mathbb{N}}$ until $\operatorname{dist}(z^k,K\cap U)\leq\epsilon$ and $\operatorname{dist}(x^k,K\cap U)\leq\epsilon$, with $\epsilon=10^{-3}$. The figures below depict the results on MAP and CRM derived in Corollaries 3.2.2 and 3.2.4.

We emphasize that in the cases above, MAP exhibits its usual behavior, *i.e.*, linear convergence the examples of the first family were somewhat special because, roughly speaking, the angle between K and U goes to 0 near the intersection. On the other hand, the superlinear

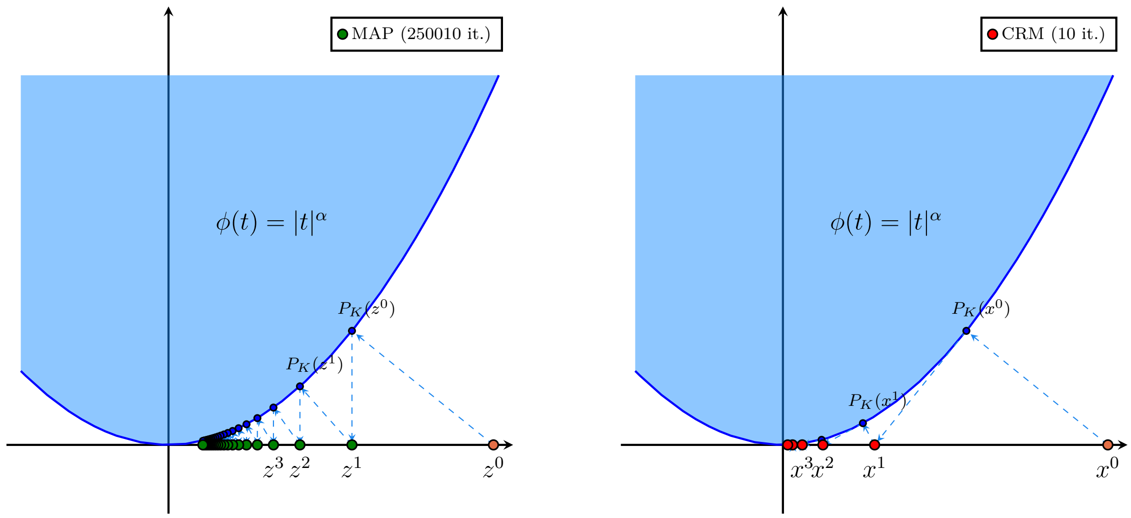

Figure 3.1: Lack of error bound (EB): MAP converges sublinearly and CRM linearly.

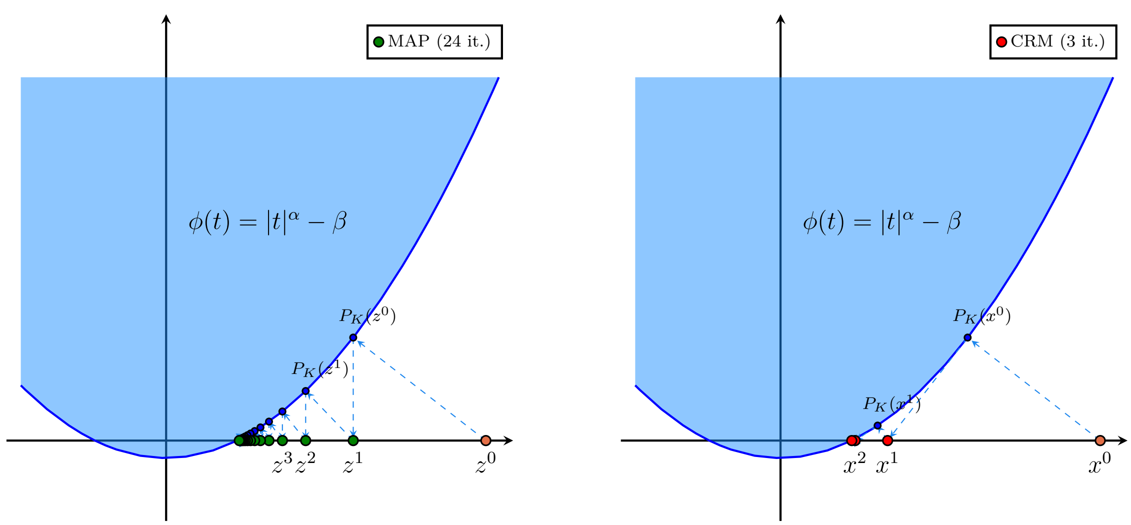

Figure 3.2: Presence of error bound (EB): MAP converges linearly and CRM superlinearly.

convergence of CRM is quite remarkable. The additional computations of CRM over MAP reduce to the trivial determination of the reflections and the solution of an elementary system of two linear equations in two real variables, for finding the circumcenter [10, 18]. Now MAP is a typical first-order method (projections disregard the curvature of the sets), and thus its convergence is generically no better than linear. We have shown that the CRM acceleration, in a rather large class of instances, improves this linear convergence to superlinear.
We conjecture that CRM enjoys superlinear convergence whenever U intersect the interior of K. The results in this section firmly support this conjecture.

Chapter 4

Circumcentering approximate reflections for solving the convex feasibility problem

The circumcentered-reflection method (CRM) has been applied for solving convex feasibility problems. CRM iterates by computing a circumcenter upon a composition of reflections with respect to convex sets. Since reflections are based on exact projections, their computation might be costly. In this regard, we introduce the circumcentered approximate-reflection method (CARM), whose reflections rely on outer-approximate projections. In these approximate methods, the projection onto a closed convex set K is replaced by the projection onto a closed convex set which separates the current iterate x from K. In rather general situations, the separating set can be taken so that the projection onto it is computationally trivial (*e.g.*, a half-space, or a Cartesian product of half-spaces). In Chapter 3, under an error bound assumption on the convex sets, we proved that both the exact CRM and MAP, converge linearly, with asymptotic constants depending on a parameter of the error bound, and that the asymptotic constant for CRM is better than the one for MAP. In this chapter we prove that under an appropriate error bound assumption, involving also the separating set, the same results hold for CARM and to a correspondent outer-approximate version of MAP, referred to as MAAP. In generic situations the separating set can be chosen so that the asymptotic constants for CARM and CRM coincide, and the same happens with the asymptotic constants of MAAP and MAP, so that in these cases the use of computationally inexpensive projections causes no deterioration at all in the convergence rates.
We analyze two families of CFP instances for which the difference between CARM and MAAP is more dramatic: using the prototypical separating operator, in the first one, MAAP converges sublinearly and CARM converges linearly; in the second one, MAAP converges linearly, but CARM converges superlinearly. Similar results on the behavior of MAP and CRM for these two families already established in Section 3.2.
We also present successful numerical comparisons of CARM to MAAP, and also to the original CRM and MAP. The Numerical results show CARM to be much faster than the other methods.

4.1 The separating operator

We start by introducing the separating operator needed for introducing the approximate versions of MAP and CRM, namely MAAP and CARM.

Definition 4.1.1. Given a closed convex set $K \subset \mathbb{R}^n$, a *separating operator* for K is a point-to-set mapping $S : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ satisfying:

A1) $S(x)$ is closed and convex for all $x \in \mathbb{R}^n$.

A2) $K \subset S(x)$ for all $x \in \mathbb{R}^n$.

A3) If a sequence $\{z^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ converges to $z^* \in \mathbb{R}^n$ and $\lim_{k\to\infty} \operatorname{dist}(z^k, S(z^k)) = 0$ then $z^* \in K$.

We have the following immediate result regarding Definition 4.1.1.

Proposition 4.1.1. If S is a separating operator for K then $x \in S(x)$ if and only if $x \in K$.

Proof. The "if" statement follows from A2. For the "only if" statement, take $x \in S(x)$, consider the constant sequence $z^k = x$ for all $k \in \mathbb{N}$, which converges to x, and apply A3. ■

Proposition 4.1.1 implies that if $x \notin K$ then $x \notin S(x)$, which, in view of A2, indicates that the set $S(x)$ separates indeed $x \notin K$ from K. The separating sets $S(x)$ will provide the approximate projections that we are going to employ throughout this chapter.
Several notions of separating operators have been introduced in the literature; see, *e.g.*, [25, Section 2.1.13] and references therein. Our definition is a point-to-set version of the separating operators in [24, Definition 2.1]. It encompasses not only hyperplane-based separators as the ones in the seminal work by Fukushima [38], considered next in Example 4.1.1, but also more general situations. Indeed, in Example 4.1.2, $S(x)$ is the Cartesian product of half-spaces, which is not a half-space.
For the family of convex sets in Examples 4.1.1 and 4.1.2, we get both explicit separating operators complying with Definition 4.1.1 and closed formulas for projections onto them.

Example 4.1.1. Assume that $K = \{x \in \mathbb{R}^n : g(x) \le 0\}$, where $g : \mathbb{R}^n \to \mathbb{R}$ is convex. Define

$$S(x) = \begin{cases} K, & \text{if } x \in K \\ \{z \in \mathbb{R}^n : u^t(z - x) + g(x) \le 0\}, & \text{otherwise,} \end{cases} \tag{4.1.1}$$

where $u \in \partial g(x)$ is an arbitrary subgradient of g at x.

We mention that any closed convex set K can be written as the 0-sublevel set of a convex, and even smooth function g, for instance, $g(x) = \operatorname{dist}(x, K)^2$, but in general this is not advantageous, because for this g it holds that $\nabla g(x) = 2(x - P_K(x))$, so that $P_K(x)$, the exact projection of x onto K, is needed for computing the separating half-space, and nothing has been won. The scheme is interesting when the function g has easily computable gradient or subgradients. For instance, in the quite frequent case in which $K = \{x \in \mathbb{R}^n : g_i(x) \le 0 \ (1 \le i \le \ell)\}$, where the g_i's are convex and smooth, we can take $g(x) = \max_{1\le i\le \ell} g_i(x)$, and the subgradients of g are easily obtained from the gradients of the g_i's.

Example 4.1.2. Assume that $\mathbf{K} = K_1 \times \cdots \times K_m \subset \mathbb{R}^{nm}$, where $K_i \subset \mathbb{R}^n$ is of the form $K_i = \{x \in \mathbb{R}^n : g_i(x) \leq 0\}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$ is convex for $1 \leq i \leq m$. Write $x \in \mathbb{R}^{nm}$ as $x = (x^1, \ldots, x^m)$ with $x^i \in \mathbb{R}^n (1 \leq i \leq m)$. We define the separating operator $\mathbf{S} : \mathbb{R}^{nm} \to \mathcal{P}(\mathbb{R}^{nm})$ as $\mathbf{S}(x) = S_1(x^1) \times \cdots \times S_m(x^m)$, with

$$S_i(x^i) = \begin{cases} K_i, & \text{if } x^i \in K_i, \\ \{z \in \mathbb{R}^n : (u^i)^t(z - x^i) + g_i(x^i) \leq 0\}, & \text{otherwise,} \end{cases} \tag{4.1.2}$$

where $u^i \in \partial g_i(x^i)$ is an arbitrary subgradient of g_i at x^i.

Example 4.1.2 is suited for the reduction of Parallel Projection Method (PPM) for m convex sets in $\mathbb{R}^n$ to MAP regarding two convex sets in $\mathbb{R}^{nm}$. Note that in Example 4.1.1, $S(x)$ is either K or a half-space, and the same holds for the sets $S_i(x^i)$ in Example 4.1.2. We prove next that the separating operators S and $\mathbf{S}$ defined in Examples 4.1.1 and 4.1.2 satisfy assumptions A1–A3.

Proposition 4.1.2. The separating operators S and $\mathbf{S}$ defined in Examples 4.1.1 and 4.1.2 satisfy assumptions A1–A3.

Proof. We start with S as in Example 4.1.1. First we observe that if $x \notin K$ then all subgradient of g at x are nonzero: since K is assumed nonempty, there exists points where g is nonpositive, so that x, which satisfies $g(x) > 0$, cannot be a minimizer of g, and hence $0 \notin \partial g(x)$, *i.e.*, $u \neq 0$ for all $u \in \partial g(x)$. Regarding A1, $S(x)$ is either equal to K or to a half-space, both of which are closed and convex.
Refarding A2, it obviously holds for $x \in K$. If $x \notin K$, we take $z \in K$, and conclude, taking into account the fact that $z \in K$ and the subgradient inequality (2.3.1), that $u^t(z-x)+g(x) \leq g(z) \leq 0$, implying that $z \in S(x)$, in view of (4.1.1).
We deal now with A3. Take a sequence $\{z^k\}_{k\in\mathbb{N}}$ converging to some z^* such that

$$\lim_{k\to\infty} \operatorname{dist}(z^k, S(z^k)) = 0.$$

We must prove that $z^* \in K$. If some subsequence of $\{z^k\}_{k\in\mathbb{N}}$ is contained in K then $z^* \in K$, because K is closed. Otherwise, for large enough k, $S(z^k)$ is a half-space. It is well known, and easy to check, that the projection P_H onto a half-space $H = \{y \in \mathbb{R}^n : a^t y \leq \alpha\} \subset \mathbb{R}^n$, with $a \in \mathbb{R}^n, \alpha \in \mathbb{R}$, is given by

$$P_H(x) = x - \|a\|^{-2} \max\{0, a^t x - \alpha\} a. \tag{4.1.3}$$

Denote by P_{S_k} the projection onto $S(z^k)$. By (4.1.3), $P_{S_k}(z) = z - \left\|u^k\right\|^{-2} \max\{0, g(z)\} u^k$, so that

$$\operatorname{dist}(z^k, S(z^k)) = \left\|z^k - P_{S_k}(z^k)\right\| = \left\|u^k\right\|^{-1} \max\{0, g(z^k)\}.$$

Note that $\{z^k\}_{k\in\mathbb{N}}$ is bounded, because it is convergent. Since the subdifferential operator ∂g is locally bounded in the interior of the domain of g (see Proposition 2.3.1), which here we take as $\mathbb{R}^n$, there exists $\mu > 0$ so that $\left\|u^k\right\| \leq \mu$ for all k and all $u^k \in \partial g(z^k)$. Hence,

$$\operatorname{dist}(z^k, S(z^k)) \geq \mu^{-1} \max\{0, g(z^k)\} \geq 0.$$

Since by assumption $\lim_{k\to\infty} \operatorname{dist}(z^k, S(z^k)) = 0$, and g, being convex, is continuous, we get that

$$0 = \lim_{k\to\infty} \mu^{-1} \max\{0, g(z^k)\} = \mu^{-1} \max\{0, g(z^*)\},$$

implying that $0 = \max\{0, g(z^*)\}$,*i.e.*, $g(z^*) \le 0$, so that $z^* \in K$ and A3 holds.
Now we consider $\mathbf{S}$ as in Example 4.1.2. As before, if $x^i \notin K_i$ then $S_i(x^i)$ is indeed a half-space in $\mathbb{R}^n$. Concerning A1–A3, A1 holds because $\mathbf{S}(x)$ is the Cartesian product of closed convex sets (either K_i or a half-space in $\mathbb{R}^n$). For A2, take $(x^1, \ldots, x^m) \in \mathbf{K}$. If $x^i \in K_i$, then $x^i \in S_i(z^i) = K_i$. Otherwise, we take $z^i \in K_i$, and invoking again the subgradient inequality (2.3.1), we get $(u^i)^t(x^i - z^i) + g(x^i) \le g(z^i) \le 0$ implying that $z^i \in S_i(x^i)$, *i.e.* $K_i \subset S_i(x^i)$ for all i, and the result follows taking into account the definitions of $\mathbf{K}$ and $\mathbf{S}$. For A3, note that $\lim_{k\to\infty} \operatorname{dist}(z^k, \mathbf{S}(z^k)) = 0$ if and only if $\lim_{k\to\infty} \operatorname{dist}(z^{k,i}, S_i(z^{k,i})) = 0$ for $1 \le i \le m$, where $z^k = (z^{k,1}, \ldots, z^{k,m})$ with $z^{k,i} \in \mathbb{R}^n$. Then, the result follows with the same argument as in Example 4.1.1, with $z^{k,i}, S_i, g_i$ substituting for z^k, S, g. ■

4.2 Convergence results for MAAP and CARM

Let us start by recalling the definitions of MAP and CRM. Consider a closed convex set $K \subset \mathbb{R}^n$ and an affine manifold $U \subset \mathbb{R}^n$. We remind that an affine manifold is a set of the form $\{x \in \mathbb{R}^n : Qx = b\}$, for some $Q \in \mathbb{R}^{n\times n}$ and some $b \in \mathbb{R}^n$. Recall that MAP and CRM iterate by means of the operators,

$$D = P_U \circ P_K, \text{ and } C(\cdot) = \operatorname{circ}(\cdot, R_K(\cdot), R_U(R_K(\cdot))) \tag{4.2.1}$$

respectively, where $R_K = 2P_K - \mathrm{Id}$, and $R_U = 2P_U - \mathrm{Id}$, are reflection operators over K and U. Then, starting from any $z^0 \in \mathbb{R}^n$ and $x^0 \in U$, MAP and CRM generate sequences $\{z^k\}_{k\in\mathbb{N}}, \{x^k\}_{k\in\mathbb{N}}$ in $\mathbb{R}^n$, according to

$$z^{k+1} = D(z^k), \text{ and } x^{k+1} = C(x^k). \tag{4.2.2}$$

respectively.
Now, we introduce the formal definitions of the Approximate Circumcentered-Reflection Method (CARM) and the Approximate Method of Alternating Projections (MAAP) applied for solving the Convex Feasibility Problem. For MAAP and CARM, we assume that $S : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^n)$ is a separating operator for K satisfying A1–A3, we take P_U as before, the orthogonal projection onto U, and define P_K as the operator given by $P_K(x) := P_{S(x)}(x)$, where $P_{S(x)}$ is the projection onto $S(x)$. Take R_U as before, and define $R^S, D^S, C^S : \mathbb{R}^n \to \mathbb{R}^n$ as

$$D^S = P_U \circ P^S, \quad R^S = 2P^S - \mathrm{Id}, \quad C^S(x) = \operatorname{circ}\left(\cdot, R^S(\cdot), R_U(R^S(\cdot))\right). \tag{4.2.3}$$

Then, starting from any $z^0 \in \mathbb{R}^n$, MAAP generates a sequence $\{z^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ according to

$$z^{k+1} = D^S(z^k), \tag{4.2.4}$$

and, starting with $x^0 \in U$, CARM generates a sequence $\{x^k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ given by

$$x^{k+1} = C^S(x^k). \tag{4.2.5}$$

We observe now that the "trivial" separating operator $S(x) = K$ for all $x \in \mathbb{R}^n$ satisfies A1-A3, and that in this case we have $D^S = D, C^S = C$, so that MAP, CRM are particular instances of MAAP, CARM respectively. Hence, the convergence analysis of the approximate algorithms encompasses the exact ones. Global convergence of MAP is well known (see, *e.g.,* [28]) and the corresponding result for CRM has been established in Proposition 2.4.4. The following propositions follow quite closely the corresponding results for the exact algorithms (Propositions 2.4.2-2.4.5), the difference consisting in the replacement of the set K by the separating set $S(x)$. However, some care is needed, because K is fixed, while $S(x)$ changes along the algorithm, so that we present the complete analysis for the approximate algorithms MAAP and CARM.

Proposition 4.2.1. For all $z \in K \cap U$ and all $x \in \mathbb{R}^n$, it holds that

$$\left\|D^S(x) - z\right\|^2 \leq \|z - x\|^2 - \left\|D^S(x) - P^S(x)\right\|^2 - \left\|P^S(x) - x\right\|^2 \tag{4.2.6}$$

with D^S as in (4.2.3).

Proof. The projection operator P_M onto any closed convex set M is known to be firmly nonexpansive (Proposition 2.2.2), that is,

$$\|P_M(x) - y\|^2 \leq \|x - y\|^2 - \|P_M(x) - x\|^2 \tag{4.2.7}$$

for all $x \in \mathbb{R}^n$ and all $y \in M$.
Applying consecutively (4.2.7) with $M = U$ and $M = S(x)$, and noting that for $z \in K \cap U$, we get $z \in U$ and also $z \in K \subset S(x)$ (due to Assumption A2), we obtain (4.2.6). ■

A similar result for operator C^S is more delicate due to the presence of the reflections and the circumcenter and requires some intermediate results. We follow closely the analysis for operator C presented in Section 2.4.
The crux of the convergence analysis of CRM, is the remarkable observation that for $x \in U \setminus K$, $C(x)$ is indeed the projection of x onto a half-space $H(x)$ separating x from $K \cap U$(see Proposition 2.4.3). This means that when the sets in CFP are an affine manifold and a hyperplane, CRM indeed converges in one step, which is a first indication of its superiority over MAP, which certainly does not enjoy this one-step convergence property, but also points to the main weakness of CRM, namely that for its convergence we may replace H by a general closed convex set, but the other set must be kept as an affine manifold. Next, we extend this result to C^S.

Lemma 4.2.1. Define $H(x) \subset \mathbb{R}^n$ as,

$$H(x) := \begin{cases} K, & \text{if } x \in K \\ \left\{z \in \mathbb{R}^n : (z - P^S(x))^t(x - P^S(x)) \leq 0\right\}, & \text{otherwise.} \end{cases} \tag{4.2.8}$$

Then, for all $x \in U, C^S(x) = P_{H(x)\cap U}(x)$.

Proof. Take $x \in U$. If $x \in K$, then $x \in S(x)$ by A2, and it follows that $R_U(x) = R^S(x) = x$, so that $C^S(x) = \text{circ}(x, x, x) = x$. Also, $P_{H(x)}(x) = P_K(x) = x$ by (4.2.8), and the result holds. Assume that $x \in U \setminus K$, so that $H(x)$ is the half-space in (4.2.8).

In view of (4.2.8), we get, using (4.1.3) with $a = x - P^S(x), \alpha = (x - P^S(x))^t P^S(x)$, that $P_{H(x)}(x) = P^S(x)$. It follows from the definition of the reflection operator R^S that

$$R^S(x) = R_{H(x)}(x). \tag{4.2.9}$$

Now, by (4.2.3) and (4.2.9),

$$C^S(x) = \operatorname{circ}(x, R^S(x), R_U(R^S(x))) = \operatorname{circ}(x, R_{H(x)}, R_U(R_{H(x)}(x))).$$

Since U is an affine manifold and $H(x)$ is a half-space, we can apply Proposition 2.4.2 and conclude that $C^S(x) = P_{H(x)\cap U}(x)$, proving the last statement of the lemma. ■

This rewriting of the operator C^S as a Projection onto the intersection of a half-space and the affine manifold U. allows us to obtain the result for CARM analogous to Proposition 4.2.1.

Proposition 4.2.2. For all $z \in K \cap U$ and all $x \in U$, it holds that

i) $\left\|C^S(x) - z\right\|^2 \le \|z - x\|^2 - \left\|C^S(x) - x\right\|^2$, with C^S as in (4.2.3).

ii) $C^S(x) \in U$ for all $x \in U$.

Proof. For (i), take $z \in K \cap U$ and $x \in U$. By Lemma 4.2.1, $C^S(x) = P_{H(x)\cap U}(x)$ for all $x \in U$. Since $z \in K \subset H(x)$, we can apply (4.2.7) with $M = H(x)$, obtaining

$$\left\|P_{H(x)\cap U}(x) - z\right\|^2 \le \|x - z\|^2 - \left\|P_{H(x)\cap U}(x) - x\right\|^2,$$

which gives the result, invoking again Lemma 4.2.1. Item (ii) follows from Proposition 2.4.2 and Lemma 4.2.1. ■

Propositions 4.2.1 and 4.2.2 allow us to prove convergence of the MAAP and CARM sequences respectively, using the well known Fejér monotonicity argument.

Theorem 4.2.1. Consider a closed convex set $K \subset \mathbb{R}^n$ and an affine manifold $U \subset \mathbb{R}^n$ such that $K \cap U \neq \varnothing$. Consider also a separating operator S for K satisfying Assumptions A1–A3. Then the sequences generated by either MAAP or CARM, starting from any initial point in the MAAP case, and from a point in U in the CARM case, are well defined, contained in U, Fejér monotone with respect to $K \cap U$, convergent, and their limits belong to $K \cap U$, *i.e.*, they solve CFP.

Proof. Let first $\{z^k\}_{k\in\mathbb{N}}$ be the sequence generated by MAAP, *i.e.* $z^{k+1} = D^S(z^k)$. Take any $z \in K \cap U$. Then, by Proposition 4.2.1,

$$\begin{aligned}\left\|z^{k+1} - z\right\|^2 &\le \left\|z^k - z\right\|^2 - \left\|P_U(P^S(z^k)) - P^S(z^k)\right\|^2 - \left\|P^S(z^k) - z^k\right\|^2 \\ &\le \left\|z^k - z\right\|^2,\end{aligned} \tag{4.2.10}$$

and so $\{z^k\}_{k\in\mathbb{N}}$ is Fejér monotone with respect to $K \cap U$. By the Definition of D^S in (4.2.3), $\{z^k\}_{k\in\mathbb{N}} \subset U$. By Proposition 2.1.2(i), $\{z^k\}_{k\in\mathbb{N}}$ is bounded. Also, $\{\left\|z^k - z\right\|\}_{k\in\mathbb{N}}$

is nonincreasing and nonnegative, therefore convergent, and thus the difference between consecutive iterates converges to 0. Hence, rewriting (4.2.10) as

$$\left\|P_U(P^S(z^k)) - P^S(z^k)\right\|^2 + \left\|P^S(z^k) - z^k\right\|^2 \le \left\|z^k - z\right\|^2 - \left\|z^{k+1} - z\right\|^2,$$

we conclude that

$$\lim_{k\to\infty} \left\|P_U(P^S(z^k)) - P^S(z^k)\right\|^2 = 0, \tag{4.2.11}$$

and

$$\lim_{k\to\infty} \left\|P^S(z^k) - z^k\right\|^2 = 0. \tag{4.2.12}$$

Let $\bar{z}$ be a cluster point of $\{z^k\}_{k\in\mathbb{N}}$ and $\{z^{j_k}\}_{j_k\in\mathbb{N}}$ a subsequence of $\{z^k\}_{k\in\mathbb{N}}$ convergent to $\bar{z}$. By (4.2.12), $\lim_{k\to\infty} \operatorname{dist}(z^{j_k}, S(z^{j_k})) = 0$. By Assumption A3 on the separating operator S, $\bar{z} \in K$. It follows also from (4.2.12) that $\lim_{k\to\infty} P^S(z^{j_k}) = \bar{z}$. By (4.2.11) and continuity of P_U, $P_U(\bar{z}) = \bar{z}$, so that $\bar{z} \in U$ and therefore $\bar{z} \in K \cap U$. By Proposition 2.1.2(ii), $\bar{z} = \lim_{k\to\infty} z^k$, completing the proof for the case of MAAP.
Let now $\{x^k\}_{k\in N}$ be the sequence generated by CARM with $x^0 \in U$. By Lemma 4.2.1, whenever $x^k \in U$, x^{k+1} is the projection onto a closed convex set, namely $H(x^k)$, and hence it is well defined. Since $x^0 \in U$ by assumption, the whole sequence is well defined, and using recursively Proposition 4.2.2(ii), we have that $\{x^k\}_{k\in\mathbb{N}} \subset U$. Now we use Proposition 2.4.2, obtaining, for any $z \in K \cap U$,

$$\left\|x^{k+1} - z\right\|^2 \le \left\|x^k - z\right\|^2 - \left\|C^S(x^k) - x^k\right\|^2 \le \left\|x^k - z\right\|^2,$$

so that again $\{x^k\}_{k\in N}$ is Fejér monotone with respect $K \cap U$, and henceforth bounded (Proposition 2.1.2(i)). Also, with the same argument as before, we get

$$\lim_{k\to\infty} \left\|x^{k+1} - x^k\right\| = \lim_{k\to\infty} \left\|C^S(x^k) - x^k\right\| = 0. \tag{4.2.13}$$

In view of (4.2.13) and the definition of circumcenter, $\left\|x^{k+1} - x^k\right\| = \left\|x^{k+1} - R^S(x^k)\right\|$, so that $\lim_{k\to\infty} \left\|x^{k+1} - R^S(x^k)\right\| = 0$ implying that $\lim_{k\to\infty} \left\|x^{k+1} - P^S(x^k)\right\| = 0$. Thus, since

$$\left\|x^k - P^S(x^k)\right\| \le \left\|x^k - x^{k+1}\right\| + \left\|x^{k+1} - P^S(x^k)\right\|$$

we get that

$$0 = \lim_{k\to\infty} \left\|x^k - P^S(x^k)\right\| = \lim_{k\to\infty} \operatorname{dist}(x^k, S(x^k)). \tag{4.2.14}$$

Let $\bar{x}$ be any cluster point of $\{x^k\}_{k\in\mathbb{N}}$. Looking at (4.2.14) along a subsequence of $\{x^k\}_{k\in\mathbb{N}}$ converging to $\bar{x}$, and invoking Assumption A3 of the separating operator S, we conclude that $\bar{x} \in K$. Since $\{x^k\}_{k\in\mathbb{N}} \subset U$, we get that all cluster points of $\{x^k\}_{k\in\mathbb{N}}$ belong to $K \cap U$, and then, using 2.1.2(ii), we get that $\lim_{k\to\infty} x^k = \bar{x} \in K \cap U$, establishing the convergence result for CARM. ■

4.3 Linear convergence rate of MAAP and CARM under a local error bound assumption

In Section 3.1 the following *global error bound* assumption on the sets K, U, denoted as EB, was considered:

EB) There exists $\bar{\omega} > 0$ such that $\operatorname{dist}(x, K) \geq \bar{\omega} \operatorname{dist}(K \cap U)$ for all $x \in U$.

Under (EB), it was proved in Section 3.1, that MAP converges linearly, with asymptotic constant bounded above by $\sqrt{1 - \bar{\omega}^2}$, and that CRM also converges linearly, with a better upper bound for the asymptotic constant, namely $\sqrt{(1 - \bar{\omega}^2)/(1 + \bar{\omega}^2)}$. In this section, we will prove similar results for MAAP and CARM, assuming a *local error bound* related not just to K, U, but also to the separating operator S. The local error bound, denoted as LEB is defined as:

LEB) There exists a set $V \subset \mathbb{R}^n$ and a scalar $\omega > 0$ such that

$$\operatorname{dist}(x, S(x)) \geq \omega \operatorname{dist}(x, K \cap U) \quad \text{for all} \quad x \in U \cap V.$$

We reckon that (LEB) becomes meaningful, and relevant for establishing convergence rate results, only when the set V contains the tail of the sequence generated by the algorithm; otherwise it might be void (*e.g.,* it holds trivially, with any ω, when $U \cap V = \varnothing$). In order to facilitate the presentation, we opted for introducing additional conditions on V in our convergence results, rather than in the definition of (LEB).
The use of a local error bound instead of a global one makes sense, because the definition of linear convergence rate deals only with iterates x^k of the generated sequence with large enough k, and the convergence of the sequences of interest has already been established in Theorem 4.2.1, so that only points close enough to the limit x^* of the sequence matter. In fact, the convergence rate analysis for MAP and CRM in Section 3.1 holds, without any substantial change, under a local, rather than global error bound.
The set V could be expected to be a neighborhood of the limit x^* of the sequence, but we do not specify it for the time being, because for the prototypical example of separating operator, *i.e.*, the one in Example 4.1.1 (Section 4.1), it will have, as we will show later, a slightly more complicated structure: a ball centered at x^* minus a certain "slice".
We start with the convergence rate analysis for MAAP.

Proposition 4.3.1. Assume that K, U and the separating operator S satisfy (LEB). Consider $D^S : \mathbb{R}^n \to \mathbb{R}^n$ as in (4.2.1). Then, for all $x \in U \cap V$,

$$(1 - \omega^2) \|x - P_{K \cap U}(x)\|^2 \geq \left\|D^S(x) - P_{K \cap U}(D^S(x))\right\|^2, \tag{4.3.1}$$

with ω as in Assumption (LEB).

Proof. By Proposition 4.2.1, for all $z \in K \cap U$ and all $x \in \mathbb{R}^n$,

$$\begin{aligned} \left\|D^S(x) - z\right\|^2 &\leq \|z - x\|^2 - \left\|D^S(x) - P^S(x)\right\|^2 - \left\|P^S(x) - x\right\|^2 \\ &\leq \|x - z\|^2 - \left\|P^S(x) - x\right\|^2. \end{aligned} \tag{4.3.2}$$

Note that $\left\|P^S(x) - x\right\| = \operatorname{dist}(x, S(x))$ and that $\left\|D^S(x) - P_{K \cap U}(D^S(x))\right\| \leq \left\|D^S(x) - z\right\|$ by definition of $P_{K \cap U}$. Take $z = P_{K \cap U}(x)$, and get from (4.3.2)

$$\begin{aligned} \left\|D^S(x) - P_{K \cap U}(D^S(x))\right\|^2 &\leq \left\|D^S(x) - P_{K \cap U}(x)\right\|^2 \\ &\leq \|x - P_{K \cap U}(x)\|^2 - \operatorname{dist}(x, S(x))^2. \end{aligned} \tag{4.3.3}$$

Take now $x \in U \cap V$ and invoke (LEB) to get from (4.3.3)

$$\begin{aligned}\left\|D^S(x) - P_{K\cap U}(D^S(x))\right\|^2 &\le \|x - P_{K\cap U}(x)\|^2 - \omega^2 \operatorname{dist}(x, K\cap U)^2 \\ &= (1-\omega)^2 \|x - P_{K\cap U}(x)\|^2,\end{aligned}$$

which immediately implies the result. ■

Note that Proposition 4.3.1 implies that if $\{x^k\}_{k\in\mathbb{N}}$ is the sequence generated by MAAP and $x^k \in V$ for large enough k, then the sequence $\{\operatorname{dist}(x^k, K\cap U)\}_{k\in\mathbb{N}}$ converges Q-linearly, with asymptotic constant bounded above by $\sqrt{1-\omega^2}$. In order to move from the distance sequence to the sequence $\{x^k\}_{k\in\mathbb{N}}$ itself, we will use Lemma 3.1.1 from Section 3.1.
Next, we establish the linear convergence of MAAP under (LEB).

Theorem 4.3.1. Consider a closed convex set $K \subset \mathbb{R}^n$ and an affine manifold $U \subset \mathbb{R}^n$, such that $K \cap U \neq \varnothing$. Moreover, assume that S is a separating operator for K satisfying Assumptions *A1–A3*. Suppose that K, U and the separating operator S satisfy (LEB). Let $\{z^k\}_{k\in\mathbb{N}}$ be the sequence generated by MAAP from any starting point $z^0 \in \mathbb{R}^n$. If there exists k_0 such that $z^k \in V$ for all $k \ge k_0$, then $\{z^k\}_{k\in\mathbb{N}}$ converges R-linearly to some point $z^* \in K\cap U$, and the asymptotic constant is bounded above by $\sqrt{1-\omega^2}$, with ω and V as in (LEB).

Proof. The fact that $\{z^k\}_{k\in\mathbb{N}}$ converges to some $z^* \in K\cap U$ has been established in Theorem 4.2.1. Take any $k \ge k_0$. By assumption, $z^k \in V$ and by definition of D^S, $z^k \in U$. So, we can take $z = z^k$ in Proposition 4.3.1, in which case $D^S(z) = z^{k+1}$, and rewrite (4.3.1) as,

$$(1-\omega^2)\operatorname{dist}(z^k, K\cap U)^2 \ge \operatorname{dist}(z^{k+1}, K\cap U)^2$$

for $k \ge k_0$, which implies that $\{\operatorname{dist}(z^k, K\cap U)\}_{k\in\mathbb{N}}$ converges Q-linearly, and hence R-linearly, with asymptotic constant bounded by $\sqrt{1-\omega^2}$. The corresponding result for the R-linear convergence of $\{z^k\}_{k\in\mathbb{N}}$ with the same bound for the asymptotic constant follows then from Lemma 3.1.1, since $\{z^k\}_{k\in\mathbb{N}}$ is Fejér monotone with respect to $K\cap U$ by Theorem 4.2.1. ■

Now we analyze the convergence rate of CARM under (LEB), for which a preliminary result, relating $x, C^S(x)$ and $D^S(x)$, is needed. The corresponding result for $x, C(x), D(x)$ can be found in Proposition 3.1.2, where it is proved that $x, C(x), D(x)$ are collinear, and moreover $D(x)$ belongs to the segment between x and $C(x)$ for all $x \in U$. Next, we will extend this result for $x, C^S(x), D^S(x)$ for all $x \in U$.

Proposition 4.3.2. Consider the operators $C^S, D^S : \mathbb{R}^n \to \mathbb{R}^n$ defined in (4.2.3). Then $D^S(x)$ belongs to the segment between x and $C^S(x)$ for all $x \in U$.

Proof. A similar argument as in the proof of the Proposition 3.1.2 establishes the result. ■

We continue with another intermediate result.

Proposition 4.3.3. Assume that (LEB) holds for K, U and S, and take $x \in U$. If $x, C^S(x) \in V$ then

$$(1+\omega^2)\operatorname{dist}(C^S(x), K\cap U)^2 \le (1-\omega^2)\operatorname{dist}(x, K\cap U)^2, \tag{4.3.4}$$

with V, ω as in (LEB).

Proof. Take $z \in K \cap U, x \in V \cap U$. We use Proposition 4.2.1, rewriting (4.2.6) as

$$\left\|x - P^S(x)\right\|^2 \leq \|x - z\|^2 - \left\|P_U(P^S(x)) - z\right\|^2 - \left\|P_U(P^S(x)) - P^S(x)\right\|^2 \tag{4.3.5}$$

for all $x \in \mathbb{R}^n$ and all $z \in K \cap U$. Since $x \in U$, we get from Lemma 4.2.1 that $C^S(x) = P_{H(x)}(x)$. We apply next Proposition 2.2.1 and get

$$\langle x - C^S(x), z - C^S(x)\rangle \leq 0. \tag{4.3.6}$$

In view of Proposition 4.3.2, $P_U(P^S(x))$ is a convex combination of x and $C^S(x)$, meaning that $P_U(P^S(x)) - C^S(x)$ is a nonnegative multiple of $x - C^S(x)$, so that (4.3.6) implies

$$\langle P_U(P^S(x)) - C^S(x), z - C^S(x)\rangle \leq 0. \tag{4.3.7}$$

Expanding the inner product in (4.3.7), we obtain

$$\left\|P_U(P^S(x)) - z\right\|^2 \geq \left\|C^S(x) - z\right\|^2 + \left\|C^S(x) - P_U(P^S(x))\right\|^2. \tag{4.3.8}$$

Combining (4.3.5) and (4.3.8), we have

$$\begin{aligned} \operatorname{dist}(x, S(x))^2 \leq \|x - z\|^2 - \left\|C^S(x) - z\right\|^2 - \left\|C^S(x) - P_U(P^S(x))\right\|^2 \\ - \left\|P_U(P^S(x)) - P^S(x)\right\|^2. \end{aligned} \tag{4.3.9}$$

Now, since U is an affine manifold, $\langle y - P_U(y), w - P_U(y)\rangle = 0$ for all $y \in \mathbb{R}^n$ and all $w \in U$, so that

$$\|w - y\|^2 = \|w - P_U(y)\|^2 + \|P_U(y) - y\|^2. \tag{4.3.10}$$

Since $C^S(x) \in U$ by Lemma 4.2.1, we use (4.3.10) with $y = P^S(x), w = C^S(x)$, getting

$$\left\|C^S(x) - P_U(P^S(x))\right\|^2 + \left\|P_U(P^S(x)) - P^S(x)\right\|^2 = \left\|C^S(x) - P^S(x)\right\|^2. \tag{4.3.11}$$

Replacing (4.3.11) in (4.3.9), we obtain

$$\begin{aligned} \operatorname{dist}(x, S(x))^2 &\leq \|x - z\|^2 - \left\|C^S(x) - z\right\|^2 - \left\|C^S(x) - P^S(x)\right\|^2 \\ &\leq \|x - z\|^2 - \operatorname{dist}(C^S(x), K \cap U)^2 - \operatorname{dist}(C^S(x), S(x))^2, \end{aligned} \tag{4.3.12}$$

using the facts that $P^S(x) \in S(x)$ and $z \in K \cap U$ in the last inequality. Now, we take $z = P_{K\cap U}(x)$, recall that $x, C^S(x) \in V$ by hypothesis, and invoke the (LEB) assumption, together with (4.3.12), in order to get

$$\begin{aligned} \omega^2 \operatorname{dist}(x, K \cap U)^2 &\leq \operatorname{dist}(x, S(x))^2 \\ &\leq \operatorname{dist}(x, K \cap U)^2 - \operatorname{dist}(C^S(x), K \cap U)^2 - \omega^2 \operatorname{dist}(C^S(x), K \cap U)^2 \\ &= \operatorname{dist}(x, K \cap U)^2 - (1 + \omega^2) \operatorname{dist}(C^S(x), K \cap U)^2. \end{aligned} \tag{4.3.13}$$

The result follows rearranging (4.3.13). ■

Next we present our convergence rate result for CARM.

Theorem 4.3.2. Consider a closed convex set $K \subset \mathbb{R}^n$, an affine manifold $U \subset \mathbb{R}^n$, such that $K \cap U \neq \varnothing$, and a separating operator S for K satisfying Assumptions *A1–A3*. Suppose that K, U and the separating operator S satisfy (LEB). Let $\{x^k\}_{k\in\mathbb{N}}$ be the sequence generated by CARM from any starting point $x^0 \in U$. If there exists k_0 such that $x^k \in V$ for all $k \geq k_0$, then $\{x^k\}_{k\in\mathbb{N}}$ converges R-linearly to some point $x^* \in K \cap U$, and the asymptotic constant is bounded above by $\sqrt{(1-\omega^2)/(1+\omega^2)}$, with ω and V as in (LEB).

Proof. The fact that $\{x^k\}_{k\in\mathbb{N}}$ converges to some $x^* \in K \cap U$ has been established in Theorem 4.2.1. Take any $k \geq k_0$. By assumption, $x^k \in V$ and by definition of D^S, $x^k \in U$. By assumption, $x^k \in V$ and and $x^0 \in U$ then from Proposition 4.2.2 x^k and $C^S(x^k)$ belongs to U Also, $k+1 \geq k_0$, so that $C^S(x^k) = x^{k+1} \in V$ So, we can take $x = x^k$ in Proposition 4.3.3, and rewrite (4.3.4) as $(1+\omega^2)\operatorname{dist}(x^{k+1}, K \cap U)^2 \leq (1-\omega^2)\operatorname{dist}(x^k, K\cap U)^2$ or equivalently as

$$\frac{\operatorname{dist}(x^{k+1}, K \cap U)}{\operatorname{dist}(x^k, K \cap U)} \leq \sqrt{\frac{1-\omega^2}{1+\omega^2}},$$

for all $k \geq 0$, which immediately implies that $\{\operatorname{dist}(x^k, K \cap U)\}_{k\in\mathbb{N}}$ converges Q-linearly, and hence R-linearly, with asymptotic constant bounded by $\sqrt{(1-\omega^2)/(1+\omega^2)}$. The corresponding result for the R-linear convergence of $\{x^k\}_{k\in\mathbb{N}}$ with the same bound for the asymptotic constant follows then from Lemma 3.1.1, since $\{x^k\}_{k\in\mathbb{N}}$ is Fejér monotone with respect to $K \cap U$ by Theorem 4.2.1. ■

From now on, given $z \in \mathbb{R}^n, \alpha > 0, B[z, \alpha]$ will denote the closed ball centered at z with radius α.
The results of Theorems 4.3.1 and 4.3.2 become relevant only if we are able to find a separating operator S for K such that (LEB) holds. This is only possible if the "trivial" separating operator satisfies an error bound, *i.e.*, if an error bound holds for the sets K, U themselves. Let $\{x^k\}_{k\in N}$ be a sequence generated by CARM starting at some $x^0 \in U$. By Theorem 4.2.1, $\{x^k\}_{k\in\mathbb{N}}$ converges to some $x^* \in K \cap U$. Without loss of generality, we assume that $x^k \notin K$ for all k, because otherwise the sequence is finite and it makes no sense to deal with convergence rates. In such a case, $x^* \in \partial K$, the boundary of K. We also assume from now on that a local error bound for K, U, say LEB1, holds at some neighborhood of x^*, *i.e.*

LEB1) There exist $\rho, \bar{\omega} > 0$ such that $\operatorname{dist}(x, K) \geq \bar{\omega}\operatorname{dist}(x, K \cap U)$ for all $x \in U \cap B(x^*, \rho)$.

Note that (LEB1) is simply a local version of (EB). Observe further that (LEB1) does not involve the separating operator S, and that it gives a specific form to the set V in (LEB), namely a ball around x^*.
We will analyze the situation for what we call the "prototypical" separating operator, namely the operator S presented in Example 4.1.1, for the case in which K is the 0-sublevel set of a convex function $g : \mathbb{R}^n \to \mathbb{R}$.
We will prove that under some additional mild assumptions on g, for any $\omega < \bar{\omega}$ there exists a set V such that U, K, S satisfy a local error bound assumption, say (LEB), with the pair ω, V.
Indeed, it will not be necessary to assume (LEB) in the convergence rate result; we will prove that under (LEB1), (LEB) will be satisfied for any $\omega < \bar{\omega}$ with an appropriate set V which does contain the tail of the sequence.

Our proof strategy will be as follows: in order to check that the error bounds for K, U and $S(x), U$ are virtually the same for x close to the limit x^* of the CARM sequence, we will prove that the quotient between $\operatorname{dist}(x, S(x))$ and $\operatorname{dist}(x, K)$ approaches 1 when x approaches x^*. Since both distances vanish at $x = x^*$, we will take the quotient of their first order approximations, in a L'Hôspital's rule fashion, and the result will be established, as long as the numerator and denominator of the new quotient are bounded away from 0, because otherwise this quotient remains indeterminate. This "bad" situation occurs when x approaches x^* along a direction almost tangent to $K \cap U$, or equivalently almost normal to $\nabla g(x^*)$. Fortunately, the CARM sequence, being Fejér monotone with respect to $K \cap U$, does not approach x^* in such a tangential way. We will take an adequate value smaller than the angle between $\nabla g(x^*)$ and $x^k - x^*$. Then, we will exclude directions whose angle with $\nabla g(x^*)$ is smaller than such value, and find a ball around x^* such that, given any $\omega < \bar{\omega}$, (LEB) holds with parameter ω in the set V defined as the ball minus the "slice" containing the "bad" directions. Because of the Fejér monotonicity of the CARM sequence, its iterates will remain in V for large enough k, and the results of Theorem 4.3.2 will hold with such ω. We proceed to follow this strategy in detail.
The additional assumptions on g are continuous differentiability and a Slater condition, meaning that there exists $\hat{x} \in \mathbb{R}^n$ such that $g(\hat{x}) < 0$. When g is of class $\mathcal{C}^1$, the separating operator of Example 4.1.1 becomes

$$S(x) = \begin{cases} K, & \text{if } x \in K \\ \{z \in \mathbb{R}^n : \nabla g(x)^t(z - x) + g(x) \leq 0\} & \text{otherwise.} \end{cases} \tag{4.3.14}$$

Proposition 4.3.4. Let $g : \mathbb{R}^n \to \mathbb{R}$ be convex, of class $\mathcal{C}^1$ and such that there exists $\hat{x} \in \mathbb{R}^n$ satisfying $g(\hat{x}) < 0$. Take $K = \{x \in \mathbb{R}^n : g(x) \leq 0\}$. Assume that K, U satisfy (LEB1). Take x^* as in (LEB1), fix $0 < \nu < \|\nabla g(x^*)\|$ (we will establish that $0 \neq \nabla g(x^*)$ in the proof of this proposition), and define $L_\nu := \{z \in \mathbb{R}^n : |\nabla g(x^*)^t(z - x^*)| \leq \nu \|z - x^*\|\}$. Consider the separating operator S defined in (4.3.14). Then, for any $\omega < \bar{\omega}$, with $\bar{\omega}$ as in (LEB1), there exists $\beta > 0$ such that K, U, S satisfy (LEB) with ω and $V := B(x^*, \beta) \setminus L_\nu$.

Proof. The fact that $0 < \nu < \|\nabla g(x^*)\|$ ensures that $V \neq \varnothing$. We will prove that for x close to x^* the quotient $\operatorname{dist}(x, S(x))/\operatorname{dist}(x, K)$ approaches 1, and first we proceed to evaluate $\operatorname{dist}(x, S(x))$. Note that when $x \in K \subset S(x)$, the inequality in LEB1 holds trivially because of A1. Thus, we assume that $x \notin K$, so that $x \notin S(x)$ by Proposition 4.1.1, and hence $g(x) > 0$ and $S(x) = \{z \in \mathbb{R}^n : \nabla g(x)^t(x - z) + g(x) \leq 0\}$, implying, in view of (4.1.3), that

$$\operatorname{dist}(x, S(x)) = \left\|x - P^S(x)\right\| = \frac{g(x)}{\|\nabla g(x)\|}. \tag{4.3.15}$$

Now we look for a more manageable expression for $\operatorname{dist}(x, K) = \|x - P_K(x)\|$. Let $y = P_K(x)$. So, y is the unique solution of the problem $\min \|z - x\|^2$ s.t. $g(z) \leq 0$, whose first order optimality conditions, sufficient by convexity of g, are

$$x - z = \lambda \nabla g(z) \tag{4.3.16}$$

with $\lambda \geq 0$, so that

$$\operatorname{dist}(x, K) = \|x - y\| = \lambda \|\nabla g(y)\|. \tag{4.3.17}$$

Now we observe that the Slater condition implies that the right-hand sides of both (4.3.15) and (4.3.17) are well defined: since $x \notin K$, $g(x) > 0$; since $y = P_K(x) \in \partial K$, $g(y) = 0$. By the Slater condition, $g(x) > g(\hat{x})$ and $g(y) > g(\hat{x})$, so that neither x nor y are minimizers of g, and hence both $\nabla g(y)$ and $\nabla g(x)$ are nonzero. By the same token, $\nabla g(x^*) \neq 0$, because x^* is not a minimizer of g: being the limit of a sequence lying outside K, x^* belongs to the boundary of K, so that $g(x^*) = 0 > g(\hat{x})$.
From (4.3.15), (4.3.17),

$$\frac{\text{dist}(x, S(x))}{\text{dist}(x, K)} = \|\nabla g(y(x))\| \, \|\nabla g(x)\| \left[\frac{\lambda(x)}{g(x)}\right], \tag{4.3.18}$$

where the notation $y(x), \lambda(x)$ emphasizes that both $y = P_K(x)$ and the multiplier λ depend on x.
We look at the right-hand side (4.3.18) for x close to $x^* \in K$, in which case y, by continuity of P_K, approaches $P(x^*) = x^*$, so that $\nabla g(y(x))$ approaches $\nabla g(x^*) \neq 0$, and hence, in view of (4.3.15), $\lambda(x)$ approaches 0. So, the product of the first two factors in the right-hand side of (4.3.18) approaches $\|\nabla g(x^*)\|^2$, but the quotient is indeterminate, because both the numerator and the denominator approach 0, requiring a more precise first order analysis.
Expanding $g(x)$ around x^* and taking into account that $g(x^*) = 0$, we get

$$g(x) = \nabla g(x^*)^t (x - x^*) + o(\|x - x^*\|). \tag{4.3.19}$$

Define $t = \|x - x^*\|, d = t^{-1}(x - x^*)$ so that $\|d\| = 1$, and (4.3.19) becomes

$$g(x) = t \nabla g(x^*)^t d + o(t). \tag{4.3.20}$$

No we look at $\lambda(x)$. Let $\phi(t) = \lambda(x^* + td)$. Note that for $x \in \partial K$ we get $y(x) = x$, so that $0 = \lambda(x) \nabla g(x)$ and hence $\lambda(x) = 0$. Thus, $\phi(0) = 0$ and

$$\lambda(x) = \phi(t) = t\phi'_+(0) + o(t), \tag{4.3.21}$$

where $\phi'_+(0)$ denotes the right derivative of $\phi(t)$ at 0. Since we assume that $x \notin K$, we have $y(x) \in \partial K$ and hence, using (4.3.16),

$$0 = g(y(x)) = g(x - \lambda(x) \nabla g(y(x))) = g(x^* + td - \phi(t) \nabla g(y(x^* + td))) \tag{4.3.22}$$

for all $t > 0$. Let $\sigma(t) = \phi(t) \nabla g(y(x^* + td))$, $\psi(t) = g(x^* + td - \sigma(t))$, so that (4.3.22) becomes $0 = \psi(t) = g(x^* + td - \sigma(t))$ for all $t > 0$ and hence

$$0 = \psi'(t) = \nabla g(y(x^* + td))^t (d - \sigma'(t)) \tag{4.3.23}$$

Taking limits in (4.3.23) with $t \to 0^+$, and noting that $y(x^*) = x^*$ because $x^* \in K$, we get

$$0 = \nabla g(x^*)^t (d - \sigma'_+(0)), \tag{4.3.24}$$

where $\sigma'_+(0)$ denotes the right derivative of $\sigma(t)$ at 0. We compute $\sigma'_+(0)$ directly from the definition, because we assume that g is of class $\mathcal{C}^1$ but perhaps not of class $\mathcal{C}^2$. Recalling that $\phi(0) = 0$, we have that

$$\begin{aligned} \sigma'_+(0) &= \lim_{t \to 0^+} \frac{\phi(t)}{t} \nabla g(y(x^* + td)) \\ &= \lim_{t \to 0^+} \frac{\phi(t)}{t} \lim_{t \to 0^+} \nabla g(y(x^* + td)) = \phi'_+(0) \nabla g(x^*), \end{aligned} \tag{4.3.25}$$

using the facts that g is class $\mathcal{C}^1$ and that $y(x^*) = x^*$. Replacing (4.3.25) in (4.3.24), we get that $0 = \nabla g(x^*)^t(d - \phi'_+(0)\nabla g(x^*))$, and therefore

$$\phi'_+(0) = \frac{\nabla g(x^*)^t d}{\|\nabla g(x^*)\|^2}. \tag{4.3.26}$$

Using (4.3.21) and (4.3.26) we obtain

$$\lambda(x) = \frac{t\nabla g(x^*)^t d}{\|\nabla g(x^*)\|^2} + o(t) = \frac{1}{\|\nabla g(x^*)\|^2}[t\nabla g(x^*)^t d + o(t)]. \tag{4.3.27}$$

Replacing (4.3.27) and (4.3.20) in (4.3.18), we obtain

$$\begin{aligned}\frac{\operatorname{dist}(x, S(x))}{\operatorname{dist}(x, K)} &= \left[\frac{\|\nabla g(y(x))\| \, \|\nabla g(x)\|}{\|\nabla g(x^*)\|^2}\right]\left[\frac{t\nabla g(x^*)^t d + o(t)}{t\nabla g(x^*)^t d + o(t)}\right] \\ &= \left[\frac{\|\nabla g(y(x^* + td))\| \, \|\nabla g(x^* + td)\|}{\|\nabla g(x^*)\|^2}\right]\left[\frac{\nabla g(x^*)^t d + o(t)/t}{\nabla g(x^*)^t d + o(t)/t}\right].\end{aligned} \tag{4.3.28}$$

Now we recall that we must check the inequality of (LEB) only for points in V, and that $V \cap L_\nu = \varnothing$, with $L_\nu = \{z \in \mathbb{R}^n : \nabla g(x^*)^t(z - x^*) \leq \nu \|z - x^*\|\}$. So, for $x \in V$ we have $|\nabla g(x^*)^t(x - x^*)| \geq \nu \|x - x^*\|$, which implies $|\nabla g(x^*)^t d| \geq \nu$, *i.e.*, $\nabla g(x^*)^t d$ is bounded away from 0, independently of the direction d. In this situation, it is clear that the rightmost expression of (4.3.28) tends to 1 when $t \to 0^+$, and so there exists some $\beta > 0$ such that for $t \in (0, \beta)$ such an expression is not smaller than $\omega/\bar{\omega}$, with ω as in (LEB) and $\bar{\omega}$ as in (LEB1). Without loss of generality, we assume that $\beta \leq \rho$, with ρ as in Assumption (LEB1). Since $t = \|x - x^*\|$, we have proved that for $x \in U \cap B(x^*, \beta) \setminus L_\nu = U \cap V$ it holds that

$$\frac{\operatorname{dist}(x, S(x))}{\operatorname{dist}(x, K)} \geq \frac{\omega}{\bar{\omega}}. \tag{4.3.29}$$

It follows from (4.3.29) that

$$\operatorname{dist}(x, S(x)) \geq \operatorname{dist}(x, K)\frac{\omega}{\bar{\omega}} \tag{4.3.30}$$

for all $x \in V \cap U$. Dividing both sides of (4.3.30) by $\operatorname{dist}(x, K \cap U)$, recalling that $\beta \leq \rho$, and invoking Assumption (LEB1), we obtain

$$\frac{\operatorname{dist}(x, S(x))}{\operatorname{dist}(x, K \cap U)} \geq \frac{\operatorname{dist}(x, K)}{\operatorname{dist}(x, K \cap U)}\frac{\omega}{\bar{\omega}} \geq \bar{\omega}\frac{\omega}{\bar{\omega}} = \omega$$

for all $x \in U \cap V$, thus proving that (LEB) holds for any $\omega < \bar{\omega}$, with $V = B(x^*, \beta) \setminus L_\nu$, with $\bar{\omega}$ as (LEB1) for the sets K, U. ■

We have proved that for the prototypical separating operator given by (4.3.14), the result of Proposition 4.3.3 holds. In order to obtain the convergence rate result of Theorem 4.3.2 for this operator, we must prove that in this case the tail of the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by CARM is contained in $V = B(x^*, \beta) \setminus L_\nu$. Note that β depends on ν. Next we will show that if we take ν smaller than a certain constant which depends on x^*, the initial iterate x^0, the Slater point $\hat{x}$ and the parameter $\bar{\omega}$ of (LEB1), then the tail of the sequence $\{x^k\}_{k\in\mathbb{N}}$ will remain outside L_ν. Clearly, this will suffice, because the sequence eventually remains in any ball around its limit, which is x^*, so that its tail will surely be contained in $B(x^*, \beta)$. The fact that $x^k \notin L_\nu$ for large enough k is a consequence of the Fejér monotonicity of the sequence with respect to $K \cap U$, proved in Theorem 4.2.1. In the next proposition, we will prove that indeed $x^k \notin L_\nu$ for large enough k, and so the result of Theorem 4.3.2 holds for this separating operator.

Proposition 4.3.5. Let $g : \mathbb{R}^n \to \mathbb{R}$ be convex, of class $\mathcal{C}^1$ and such that there exists $\hat{x} \in \mathbb{R}^n$ satisfying $g(\hat{x}) < 0$. Take $K = \{x \in \mathbb{R}^n : g(x) \le 0\}$. Assume that K, U satisfy (LEB1). Consider the separating operator S defined in (4.3.14). Let $\{x^k\}_{k\in\mathbb{N}}$ be a sequence generated by (CARM) with starting point $x^0 \in U$ and limit point $x^* \in K \cap U$. Take $\nu > 0$ satisfying

$$\nu < \min\left\{ \frac{\bar{\omega}\,|g(\hat{x})|}{4\left(\|\hat{x} - x^*\| + \|x^* - x^0\|\right)}, \frac{\|\nabla g(x^*)\|}{2} \right\}, \tag{4.3.31}$$

with $\bar{\omega}$ as in (LEB1), and define

$$L_\nu := \{z \in \mathbb{R}^n : \left|\nabla g(x^*)^t (z - x^*)\right| \le \nu \,\|z - x^*\|\}.$$

Then, there exists k_0 such that for all $k \ge k_0$, $x_k \in B(x^*, \beta) \setminus L_\nu$, with β as in Proposition 4.3.4.

Proof. Assume that $x^k \in L_\nu$, *i.e.*,

$$\left|\nabla g(x^*)^t (x^k - x^*)\right| \le \nu \left\|x^k - x^*\right\|. \tag{4.3.32}$$

Using the gradient inequality, the fact that $g(x^*) = 0$ and (4.3.32), we obtain

$$\begin{aligned} g(x^k) &\le g(x^*) - \nabla g(x^k)^t (x^* - x^k) \\ &= [\nabla g(x^*) - \nabla g(x^k) - \nabla g(x^*)]^t (x^* - x^k) \\ &\le \left\|\nabla g(x^*) - \nabla g(x^k)\right\| \left\|x^* - x^k\right\| + \left|\nabla g(x^*)^t (x^k - x^*)\right| \\ &\le \left(\left\|\nabla g(x^*) - \nabla g(x^k)\right\| + \nu\right) \left\|x^k - x^*\right\|. \end{aligned} \tag{4.3.33}$$

By Theorem 4.2.1, $\{x^k\}_{k\in\mathbb{N}}$ is Fejér monotone with respect to $K \cap U$. Thus, we use Proposition 2.1.2(iii) and LEB1, in (4.3.33), obtaining

$$\begin{aligned} g(x^k) &\le 2\left(\left\|\nabla g(x^*) - \nabla g(x^k)\right\| + \nu\right) \operatorname{dist}(x^k, K \cap U) \\ &\le \frac{2\left(\left\|\nabla g(x^*) - \nabla g(x^k)\right\| + \nu\right) \operatorname{dist}(x^k, K)}{\bar{\omega}}. \end{aligned} \tag{4.3.34}$$

Denote $y^k = P_K(x^k)$. Using again the gradient inequality, together with the facts that $g(y^k) = 0$ and that $x^k - y^k$ and $\nabla g(y^k)$ are collinear, which is a consequence of (4.3.16) and the nonnegativity of λ, we get from (4.3.34)

$$\begin{aligned} g(x^k) &\ge g(y^k) + \nabla g(y^k)^t (x^k - y^k) \\ &= \left\|\nabla g(y^k)\right\| \left\|x^k - y^k\right\| = \left\|\nabla g(y^k)\right\| \operatorname{dist}(x^k, K). \end{aligned} \tag{4.3.35}$$

Now we use the Slater assumption on g for finding a lower bound for $\left\|\nabla g(y^k)\right\|$. Take $\hat{x}$ such that $g(\hat{x}) < 0$, and apply once again the gradient inequality.

$$g(\hat{x}) \geq g(y^k) + \nabla g(y^k)^t(\hat{x} - y^k) = \nabla g(y^k)^t(\hat{x} - y^k) \geq - \left\|\nabla g(y^k)\right\| \left\|\hat{x} - y^k\right\|. \tag{4.3.36}$$

Multiplying (4.3.36) by -1, we get

$$\begin{aligned} |g(\hat{x})| &\leq \left\|\nabla g(y^k)\right\| \left\|\hat{x} - y^k\right\| \leq \left\|\nabla g(y^k)\right\| \left(\|\hat{x} - x^*\| + \left\|x^* - y^k\right\|\right) \\ &\leq \left\|\nabla g(y^k)\right\| \left(\|\hat{x} - x^*\| + \left\|x^* - x^k\right\|\right) \\ &\leq \left\|\nabla g(y^k)\right\| \left(\|\hat{x} - x^*\| + \left\|x^* - x^0\right\|\right), \end{aligned} \tag{4.3.37}$$

using the facts that $y^k = P_K(x^k)$ and that $x^* \in K$ in the third inequality and the Féjer monotonicity of $\{x^k\}_{k\in N}$ with respect to $K \cap U$ in the fourth one. Now, since $\lim_{k\to\infty} x^k = x^*$, there exists k_1 such that $\left\|x^k - x^*\right\| \leq \rho$ for $k \geq k_1$, with ρ as in (LEB1). So, in view of (4.3.37), with $k \geq k_1$, $|g(\hat{x})| \leq \left\|\nabla g(y^k)\right\| (\|\hat{x} - x^*\| + \|x^* - x^0\|)$, implying that

$$\left\|\nabla g(y^k)\right\| \geq \frac{|g(\hat{x})|}{\|\hat{x} - x^*\| + \|x^* - x^0\|}. \tag{4.3.38}$$

Combining (4.3.34), (4.3.35), (4.3.38) and (4.3.31), we obtain

$$2\nu < \frac{\bar{\omega}\,|g(\hat{x})|}{2\left(\|\hat{x} - x^*\| + \|x^* - x^0\|\right)} \leq \left\|\nabla g(x^k) - \nabla g(x^*)\right\| + \nu,$$

implying

$$\nu < \left\|\nabla g(x^k) - \nabla g(x^*)\right\|. \tag{4.3.39}$$

The inequality in (4.3.39) has been obtained by assuming that $x^k \in L_\nu$. Now, since $\lim_{k\to\infty} x^k = x^*$ and g is of class $\mathcal{C}^1$, there exists $k_0 \geq k_1$ such that $\left\|\nabla g(x^*) - \nabla g(x^k)\right\| \leq \nu$ for $k \geq k_0$, and hence (4.3.39) implies that for $k \geq k_0$, $x^k \notin L_\nu$. Since $k_0 \geq k_1$, $x^k \in B(x^*, \beta)$ for $k \geq k_0$, meaning that when $k \geq k_0$, $x^k \in B(x^*, \beta) \setminus L_\nu$, establishing the result. ■

Now we conclude the analysis of CARM with the prototypical separating operator, proving that under smoothness of g and a Slater condition, the CARM method achieves linear convergence with precisely the same bound for the asymptotic constant as CRM, thus showing that the approximation of P_K by P^S produces no deterioration in the convergence rate. We emphasize again that for this operator S, P^S has an elementary closed formula, namely the one given by

$$P^S(x) = x - \left(\frac{\max\{0, g(x)\}}{\|\nabla g(x)\|^2}\right) \nabla g(x).$$

Theorem 4.3.3. Let $g : \mathbb{R}^n \to \mathbb{R}$ be convex, of class $\mathcal{C}^1$ and such that there exists $\hat{x} \in \mathbb{R}^n$ satisfying $g(\hat{x}) < 0$. Take $K = \{x \in \mathbb{R}^n : g(x) \leq 0\}$. Assume that K, U satisfy (LEB1). Consider the separating operator S defined in (4.3.14). Let $\{x^k\}_{k\in\mathbb{N}}$ be a sequence generated by CARM with starting point $x^0 \in U$. Then $\{x^k\}_{k\in\mathbb{N}}$ converges to some $x^* \in K \cap U$ with linear convergence rate, and asymptotic constant bounded above by $\sqrt{(1 - \bar{\omega}^2)/(1 + \bar{\omega}^2)}$, with $\bar{\omega}$ as in (LEB1).

Proof. The fact that $\{x^k\}_{k\in\mathbb{N}}$ converges to some $x^* \in K \cap U$ follows from Theorem 4.2.1. Let $\bar{\omega}$ be the parameter in (LEB1). By Proposition 4.3.4, P, K and S satisfy (LEB) with any parameter $\omega \leq \bar{\omega}$ and a suitable V. By Proposition 4.3.5, $x^k \in V$ for large enough k, so that the assumptions of 4.3.2 hold, and hence

$$\limsup_{k\to\infty} \frac{\left\|x^{k+1} - x^*\right\|}{\left\|x^k - x^*\right\|} \leq \sqrt{\frac{1-\omega^2}{1+\omega^2}} \tag{4.3.40}$$

for any $\omega \leq \bar{\omega}$. Taking infimum in the right-hand side of (4.3.40) with $\omega < \bar{\omega}$, we conclude that the inequality holds also for $\bar{\omega}$, *i.e.*

$$\limsup_{k\to\infty} \frac{\left\|x^{k+1} - x^*\right\|}{\left\|x^k - x^*\right\|} \leq \sqrt{\frac{1-\bar{\omega}^2}{1+\bar{\omega}^2}},$$

completing the proof. ■

We mention that the results of Propositions 4.3.4 and 4.3.5 and Theorem 4.3.3 can be extended without any complications to the separating operator **S** in Example 4.1.2, so that they can be applied for accelerating PPM for CFP with m convex sets, presented as 0-sublevel sets of smooth convex functions. We omit the details.

4.4 Convergence rate results for CARM and MAAP applied to specific instances of CFP

In this section we will present two rather generic families of examples for which CARM is faster than MAAP. The results of Section 4.3 indicate that when K, U satisfy an error bound assumption, both CARM and MAAP enjoy linear convergence rates (with a better asymptotic constant for the former). In this section we present two families of CFP instances for which the difference between CARM and MAAP is more dramatic: using the prototypical separating operator, in the first one (for which (LEB) does not hold), MAAP converges sublinearly and CARM converges linearly; in the second one, MAAP converges linearly, as in Section 4.3, but CARM converges superlinearly. Similar results on the behavior of MAP and CRM for these two families already established in section 3.2.

Throughout this section, $K \subset \mathbb{R}^{n+1}$ will be the epigraph of a convex function $f : \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^1$ and U will be the hyperplane $U := \{x \in \mathbb{R}^{n+1} : x_{n+1} = 0\}$. We mention that the specific form of U and the fact that K is an epigraph entail little loss of generality; but the smoothness assumption on f and the fact that U is a hyperplane (*i.e.* an affine manifold of codimension 1), are indeed more restrictive.

First we look at the case when the following assumptions hold:

B1. $f(0) = 0$.

B2. $\nabla f(x) = 0$ if and only if $x = 0$.

Note that under B1–B2, 0 is the unique minimizer of f and that $K \cap U = \{0\}$. It follows from Theorem 4.2.1 that the sequences generated by MAAP and CARM, from any initial iterate in $\mathbb{R}^n$ and U respectively, converge to $x^* = 0$. We prove next that under these assumptions, MAAP converges sublinearly.

Proposition 4.4.1. Assume that $K \subset \mathbb{R}^{n+1}$ is the epigraph of a convex function $f : \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^1$ satisfying B1–B2, and $U := \{x \in \mathbb{R}^{n+1} : x_{n+1} = 0\}$. Consider the separating operator given by (4.3.14) for the function $g : \mathbb{R}^{n+1} \to \mathbb{R}$ defined as $g(x_1, \ldots, x_{n+1}) = f(x_1, \ldots, x_n) - x_{n+1}$. Then the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by MAAP starting at any $x^0 \in \mathbb{R}^{n+1}$ converges sublinearly to $x^* = 0$.

Proof. Convergence of $\{x^k\}_{k\in\mathbb{N}}$ to $x^* = 0$ results from Theorem 4.2.1. We write vectors in $\mathbb{R}^{n+1}$ as (x, s) with $x \in \mathbb{R}^n, s \in \mathbb{R}$. We start by computing the formula for $D^S(x, 0)$. By definition of g, $\nabla g(x, s) = (\nabla f(x), -1)^t$. Let

$$\alpha(x) = \|\nabla f(x)\|^2 + 1. \tag{4.4.1}$$

By (4.1.3),

$$P^S(x, 0) = (x, 0) - \frac{g(x, 0)}{\|\nabla g(x, 0)\|^2} \nabla g(x, 0) = \left(x - \frac{f(x)}{\alpha(x)} \nabla f(x), -\frac{f(x)}{\alpha(x)}\right),$$

which implies, since $P_U(x, s) = (x, 0)$,

$$D^S(x, 0) = P_U(P^S(x)) = \left(x - \frac{f(x)}{\alpha(x)} \nabla f(x), 0\right) \tag{4.4.2}$$

Let $\bar{x} = \|x\|^{-1} x$. From (4.4.2),

$$\left[\frac{\|D^S(x, 0)\|}{\|(x, 0)\|}\right]^2 = 1 - 2\frac{f(x)}{\|x\|}\left(\frac{\nabla f(x)^t \bar{x}}{\alpha(x)}\right) + \left(\frac{f(x)}{\|x\|} \frac{\|\nabla f(x)\|}{\alpha(x)}\right)^2. \tag{4.4.3}$$

Note that $\lim_{x\to 0} \alpha(x) = \alpha(0) = 1$ and that, by B1–B2, $\lim_{x\to 0} \nabla f(x) = \nabla f(0) = 0$, $f(x) = o(\|x\|)$, implying that $\lim_{x\to 0} f(x)/\|x\| = 0$, and conclude from (4.4.3) that

$$\lim_{x\to 0} \frac{\|D^S(x, 0)\|}{\|(x, 0)\|} = 1. \tag{4.4.4}$$

Now, since $x^{k+1} = D^S(x^k)$, $x^k \in U$ for all $k \geq 0$, and $x^* = 0$, we get from (4.4.4)

$$\lim_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = \lim_{x\to 0} \frac{\|D^S(x, 0)\|}{\|(x, 0)\|} = 1,$$

and hence $\{x^k\}_{k\in\mathbb{N}}$ converges sublinearly. ■

Next we study the CARM sequence in the same setting.

Proposition 4.4.2. Assume that $K \subset \mathbb{R}^{n+1}$ is the epigraph of a convex function $f : \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^1$ satisfying *B1–B2*, and $U := \{x \in \mathbb{R}^{n+1} : x_{n+1} = 0\}$. Consider the separating operator given by (4.3.14) for the function $g : \mathbb{R}^{n+1} \to \mathbb{R}$ defined as $g(x_1, \ldots, x_{n+1}) = f(x_1, \ldots, x_n) - x_{n+1}$. For $0 \neq x \in \mathbb{R}^n$, define

$$\theta(x) := \frac{f(x)}{\|x\| \, \|\nabla f(x)\|}. \tag{4.4.5}$$

Then

$$\left[\frac{\left\|C^S(x,0)\right\|}{\|x\|}\right]^2 \leq 1-\theta(x)^2, \tag{4.4.6}$$

with C^S as in (4.2.3).

Proof. Define

$$\beta(x) := \frac{f(x)}{\|\nabla f(x)\|^2}. \tag{4.4.7}$$

By the definition of reflection in (4.2.3),

$$R^S(x,0) = \left(x - 2\frac{f(x)}{\alpha(x)}\nabla f(x), 2\frac{f(x)}{\alpha(x)}\right). \tag{4.4.8}$$

From Proposition 4.3.2,

$$C^S(x,0) = (x,0) + \eta(D^S(x,0)-(x,0)) = \left(x - \eta\frac{f(x)}{\alpha(x)}\nabla f(x), 0\right), \tag{4.4.9}$$

for some $\eta \geq 1$. By the definition of circumcenter, $\left\|C^S(x)-x\right\| = \left\|C^S(x)-R^S(x)\right\|$. Combining this equation with (4.4.8) and (4.4.9), one obtains $\eta = 1 + \|\nabla f(x)\|^{-1}$, which implies, in view of (4.4.7), that

$$\frac{\eta f(x)}{\alpha(x)} = \frac{f(x)}{\|\nabla f(x)\|^2} = \beta(x). \tag{4.4.10}$$

Combining (4.4.9) and (4.4.10),

$$C^S(x,0) = (x - \beta(x)\nabla f(x), 0), \tag{4.4.11}$$

so that

$$\begin{aligned}\left\|C^S(x,0)\right\|^2 &= \|x\|^2 - 2\beta(x)\nabla f(x)^t x + \beta(x)^2\|\nabla f(x)\|^2 \\ &\leq \|x\|^2 - 2\beta(x)f(x) + \beta(x)^2\|\nabla f(x)\|^2,\end{aligned} \tag{4.4.12}$$

using the fact that $f(x) \leq \nabla f(x)^t x$, which follows from that gradient inequality with the points x and 0. It follows from (4.4.12) and the definitions of $\alpha(x)$ and $\beta(x)$, that

$$\begin{aligned}\left[\frac{\left\|C^S(x,0)\right\|}{\|x\|}\right]^2 &\leq 1 - 2\beta(x)\frac{f(x)}{\|x\|^2} + \left(\frac{\beta(x)\|\nabla f(x)\|}{\|x\|}\right)^2 \\ &= 1 - \left(\frac{f(x)}{\|\nabla f(x)\|\,\|x\|}\right)^2 = 1-\theta(x)^2,\end{aligned}$$

using (4.4.5) in the last equality. ■

We prove next the linear convergence of the CARM sequence in this setting under the following additional assumption on f:

B3) $\liminf_{x \to 0} \frac{f(x)}{\|x\| \, \|\nabla f(x)\|} > 0.$

Corollary 4.4.1. Under the assumptions of Proposition 4.4.2, if f satisfies B3 and $\{x^k\}_{k\in\mathbb{N}}$ is the sequence generated by CARM starting at any $x^0 \in U$, then $\lim_{k\to\infty} x^k = x^* = 0$, and

$$\liminf_{k\to\infty} \frac{\left\|x^{k+1} - x^*\right\|}{\left\|x^k - x^*\right\|} \leq \sqrt{1-\delta^2} < 1,$$

with

$$\delta = \liminf_{x\to 0} \frac{f(x)}{\|x\| \, \|\nabla f(x)\|},$$

so that $\{x^k\}_{k\in\mathbb{N}}$ converges linearly, with asymptotic constant bounded by $\sqrt{1-\delta^2}$.

Proof. Convergence of $\{x^k\}_{k\in\mathbb{N}}$ follows from Theorem 4.2.1. Since $x^{k+1} = C^S(x^k)$, we invoke Proposition 4.4.2, observing that $\liminf_{x\to 0} \theta(x) = \delta$, and taking square root and lim sup in (4.4.6):

$$\limsup_{k\to\infty} \frac{\left\|x^{k+1} - x^*\right\|}{\left\|x^k - x^*\right\|} \leq \sqrt{1 - \liminf_{k\to\infty} \theta(x^k)^2} = \sqrt{1-\delta^2} < 1,$$

using (4.4.5) and Assumption B3. ■

In Section 3.2 we showed that Assumption B3 holds in several cases, *e.g.*, when f is of class $\mathcal{C}^2$ and the Hessian $\nabla^2 f(0)$ is positive definite, in which case

$$\delta \geq \frac{1}{2} \frac{\lambda_{\min}}{\lambda_{\max}},$$

where $\lambda_{\max}, \lambda_{\min}$ are the largest and smallest eigenvalues of $\nabla^2 f(0)$, or when $f(x) = \varphi(\|x\|)$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function of class $\mathcal{C}^r$, satisfying $\varphi(0) = \varphi'(0) = 0$, in which case $\delta \geq 1/p$, where $p \leq r$ is defined as $p = \min\{j : \varphi^{(j)} \neq 0\}$.
In all these instances, in view of Proposition 4.4.1 and Corollary 4.4.1, the CARM sequence converges linearly, while the MAAP one converges sublinearly. If we look at the formulae for D^S and C^S, in (4.4.2) and (4.4.11), we note that both operators move from $(x, 0)$ in the direction $(\nabla f(x), 0)$ but with different step-sizes. Looking now at (4.4.3) and (4.4.5), we see that the relevant factors of these step-sizes, for x near 0, are $f(x)/\|x\|$ and $f(x)/(\|x\| \, \|\nabla f(x)\|)$. Since we assume that $\nabla f(0) = 0$, the first one vanishes near 0, inducing the sublinear behavior of MAAP, while the second one, in rather generic situations, will stay away from 0. It is the additional presence of $\|\nabla f(x)\|$ in the denominator of $\theta(x)$ which makes all the difference.
Now we analyze the second family, which is similar to the first one, excepting that condition B1 is replaced by the following one:

B1') $f(0) < 0$.

We also make a further simplifying assumption, which is not essential for the result, but keeps the calculations simpler. We take f of the form $f(x) = \varphi(\|x\|)$ with $\varphi : \mathbb{R} \to \mathbb{R}$. Rewriting B1', B2 in terms of φ, we assume that

(i) $\varphi : \mathbb{R} \to \mathbb{R}$ is strictly convex and of class $\mathcal{C}^1$,

(ii) $\varphi(0) < 0$,

(iii) $\varphi'(0) = 0$.

This form of f gives a one-dimensional flavor to this family. Now, $0 \in \mathbb{R}^{n+1}$ cannot be the limit point of the MAAP or the CARM sequences: 0 is still the unique minimizer of f, but since $f(0) < 0$, $0 \notin \partial K$, while the limit points of the sequences, unless they are finite (in which case convergence rates make no sense), do belong to the boundary of K. Hence, both f and ∇f do not vanish at such limit points, implying that both φ and φ' are nonzero at the norms of the limit points. We have the following result for this family.

Proposition 4.4.3. Assume that $U, K \subset \mathbb{R}^{n+1}$ are defined as $U = \{(x, 0) : x \in \mathbb{R}^n\}$ and $K = \operatorname{epi}(f)$ where $f(x) = \phi(\|x\|)$ and ϕ satisfies (i)–(iii). Let C^S, D^S be as defined in (4.2.3), and $(x^*, 0), (z^*, 0)$ the limits of the sequences $\{x^k\}_{k\in\mathbb{N}}$, $\{z^k\}_{k\in\mathbb{N}}$ generated by CARM and MAAP, starting from some $(x^0, 0) \in U$, and some $(z^0, w) \in \mathbb{R}^{n+1}$, respectively. Then

$$\lim_{x\to z^*} \frac{\left\|D^S(x,0) - (z^*,0)\right\|}{\|(x,0)-(z^*,0)\|} = \frac{1}{1+\phi'(\|z^*\|)^2} \tag{4.4.13}$$

and

$$\lim_{x\to z^*} \frac{\left\|C^S(x,0) - (x^*,0)\right\|}{\|(x,0)-(x^*,0)\|} = 0. \tag{4.4.14}$$

Proof. We start by rewriting the formulae for $C^S(x), D^S(x)$ in terms of φ. We also define $t := \|x\|$. Using (4.4.1), (4.4.2), (4.4.5) and (4.4.6), we obtain

$$D^S(x,0) = \left(\left[1 - \frac{\varphi(\|x\|)\varphi'(\|x\|)}{(\varphi'(\|x\|)^2+1)\|x\|}\right]x, 0\right) = \left(\left[1 - \frac{\varphi(t)\varphi'(t)}{(\varphi'(t)^2+1)t}\right]x, 0\right) \tag{4.4.15}$$

and

$$C^S(x,0) = \left(\left[1 - \frac{\varphi(\|x\|)}{\varphi'(\|x\|)\|x\|}\right]x, 0\right) = \left(\left[1 - \frac{\varphi(t)}{\varphi'(t)t}\right]x, 0\right). \tag{4.4.16}$$

Note that $x, D^S(x), C^S(x)$ are collinear (the one-dimensional flavor!), so that the same happens with x^*, z^*. Let $r := \|x^*\|, s := \|z^*\|$, so that $x^* = (r/t)x, z^* = (s/t)x$. Then, using (4.4.15) and (4.4.16), we get

$$\begin{aligned}\frac{\left\|D^S(x,0) - (z^*,0)\right\|}{\|(x,0)-(z^*,0)\|} &= \frac{t - r - \dfrac{\varphi(t)\varphi'(t)}{\varphi'(t)^2+1}}{t-r} = \left[1 - \frac{\varphi(t)}{t-r}\right]\left[\frac{\varphi'(t)}{\varphi'(t)^2+1}\right] \\ &= \left[1 - \frac{\varphi(t)-\varphi(r)}{t-r}\right]\left[\frac{\varphi'(t)}{\varphi'(t)^2+1}\right],\end{aligned} \tag{4.4.17}$$

and

$$\begin{aligned}\frac{\left\|C^S(x,0) - (x^*,0)\right\|}{\|(x,0)-(x^*,0)\|} &= \frac{t - s - \dfrac{\varphi(t)}{\varphi'(t)}}{t-s} = 1 - \left[\frac{\varphi(t)}{t-s}\right]\frac{1}{\varphi'(t)} \\ &= 1 - \left[\frac{\varphi(t)-\varphi(r)}{t-s}\right]\frac{1}{\varphi'(t)},\end{aligned} \tag{4.4.18}$$

using in the last equalities of (4.4.17) and (4.4.18) the fact that $\varphi(r) = \varphi(s) = 0$, which results from $f(x^*) = f(z^*) = 0$. Now we take limits with $x \to z^*$, $x \to x^*$ in the leftmost expressions of (4.4.17), (4.4.18), which demands limits with $t \to s$, $t \to r$ in the rightmost expressions of them.

$$\lim_{x\to x^*} \frac{\left\|D^S(x,0) - (z^*,0)\right\|}{\|(x,0) - (z^*,0)\|} = \lim_{t\to r}\left[1 - \frac{\varphi(t)-\varphi(r)}{t-r}\right]\left[\frac{\varphi'(t)}{\varphi'(t)^2+1}\right]$$
$$= 1 - \frac{\varphi'(r)^2}{\varphi'(r)^2+1} = \frac{1}{\varphi'(r)^2+1} = \frac{1}{\varphi'(\|z^*\|)^2+1},$$

and

$$\lim_{x\to z^*} \frac{\left\|C^S(x,0) - (x^*,0)\right\|}{\|(x,0) - (x^*,0)\|} = \lim_{t\to s}\left[1 - \frac{\varphi(t)-\varphi(r)}{t-s}\right]\frac{1}{\varphi'(t)} = 1 - \frac{\varphi'(s)}{\varphi'(s)} = 0,$$

completing the proof. ■

Corollary 4.4.2. Under the assumptions of Proposition 4.4.3, the sequence generated by MAAP converges Q-linearly to a point $(x^*,0) \in K \cap U$, with asymptotic constant equal to $1/(1+\varphi'(\|x^*\|)^2)$, and the sequence generated by CARM converges superlinearly.

Proof. Recall that if $\{(x^k,0)\}_{k\in\mathbb{N}}$ is the MAAP sequence, then $(x^{k+1},0) = D^S(x^k,0)$, and if $\{(z^k,0)\}_{k\in\mathbb{N}}$ is the CARM sequence, then $(z^{k+1},0) = C^S(z^k,0)$. Recall also that for both sequences the last components of the iterates vanish because $\{x^k\}_{k\in\mathbb{N}}, \{z^k\}_{k\in\mathbb{N}} \subset U$. Then the result follows immediately from (4.4.13) and (4.4.14) in Proposition 4.4.3. ■

We mention that the results of Corollary 4.4.2 coincide with those obtained in Corollary 3.2.4 in Section 3.2 for the sequences generated by MAP and CRM applied to the same families of instances of CFP, showing that the convergence rate results of the exact methods are preserved without any deterioration also in these cases.

4.5 Numerical comparisons between CARM, MAAP, CRM and MAP

In this section, we perform numerical comparisons between CARM, MAAP, CRM and MAP. These methods are employed for solving the particular CFP of finding a common point in the intersection of finitely many ellipsoids, that is, finding

$$\bar{x} \in \mathcal{E} = \bigcap_{i=1}^{m} \mathcal{E}_i \subset R^n, \tag{4.5.1}$$

with each ellipsoid $\mathcal{E}_i$ being given by

$$\mathcal{E}_i := \{x \in \mathbb{R}^n : g_i(x) \le 0\}, \text{ for } i = 1, \dots, m,$$

where $g_i : \mathbb{R}^n \to \mathbb{R}$ is given by $g_i(x) = x^t A_i x + 2x^t b^i - \alpha_i$, each A_i is a symmetric positive definite matrix, b^i is a n-vector, α_i is a positive scalar.

Problem (4.5.1) has importance of its own (see [44, 50]) and both CRM and MAP are suitable for solving it. Nevertheless, the main motivation for tackling it with approximate projection methods is that the computation of exact projections onto ellipsoids is a formidable burden for any algorithm to bear. Since the gradient of each g_i is easily available, we can consider the separable operators given in Examples 4.1.1 and 4.1.2, and use CARM and MAAP to solve problem (4.5.1) as well. What is more, the experiments illustrate that, in this case, CARM handily outperforms CRM, in terms of CPU time, while still being competitive in terms of iteration count. The exact orthogonal projection onto each ellipsoid is so demanding that even MAAP has a better CPU time result than CRM.

The four methods are employed upon Pierra's product space reformulation (Subsection 1.1.3), that is, we seek a point $x^* \in K \cap D$, where $K := \mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \mathcal{E}_m$ and D is the diagonal space. For each sequence $\{x^k\}_{k\in\mathbb{N}}$ that we generate, we consider the tolerance $\epsilon := 10^{-6}$ and use as stopping criteria the gap distances

$$\|x^k - P_K(x^k)\| < \epsilon, \qquad \text{or} \qquad \|x^k - P_K^S(x^k)\| < \epsilon,$$

where $P_K(x^k)$ is utilized for CRM and MAP, and $P_K^S(x^k)$ is used for CARM and MAAP. We also set the maximum number of iterations as 50 000.

For executing our tests, we randomly generate 250 instances of (4.5.1) in the following way. We range the dimension size n in $\{10, , 30, 50, 100, 200\}$ and for each n we take the number m of underlying sets varying in $\{10, 30, 50, 100, 200\}$. For each of these 25 pairs, (m, n) we build 10 randomly generated instances of (4.5.1). Each matrix A_i is of the form $A_i = \gamma Id + B_i^t B_i$, with $B_i \in \mathbb{R}^{n\times n}$, $\gamma \in \mathbb{R}_{++}$. The matrix B_i is a sparse matrix sampled from the standard normal distribution with sparsity density $p = 2n^{-1}$ and each vector b^i is sampled from the uniform distribution between $[0, 1]$. We then choose each α_i so that $\alpha_i > (b^i)^t A b^i$, which ensures that 0 belongs to every $\mathcal{E}_i$, and thus (4.5.1) is feasible. The initial point x^0 is of the form $(\eta, \eta, \cdots, \eta) \in \mathbb{R}^n$, with η being negative and $|\eta|$ sufficient large, guaranteeing that x^0 is far from all $\mathcal{E}_i$'s.

The computational experiments were performed on an Intel Xeon W-2133 3.60GHz with 32GB of RAM running Ubuntu 20.04 and using `Julia v1.5` programming language [22]. The codes for our experiments are fully available in https://github.com/Mirza-Reza/CFP.

The projections onto ellipsoids are computed using an alternating direction method of multipliers (ADMM), see [30]. We will explain briefly in the following subsection how ADMM works.

The results are summarized in Figure 4.1 using a performance profile [33]. Performance profiles allow one to compare different methods on a problem sets with respect to a performance measure. The vertical axis indicates the percentage of problems solved, while the horizontal axis indicates, in log-scale, the corresponding factor of the performance index used by the best solver. In this case, when looking at CPU time (in seconds), the performance profile shows that CARM always did better than the other three methods. The picture also shows that MAAP took less time than CRM and MAP. We conclude this examination by presenting, in Table 4.1, the following descriptive statistics of the benchmark of CARM, MAAP, CRM and MAP: mean, maximum (max), minimum (min) and standard deviation (std) for iteration count (it) and CPU time in seconds (CPU (s)). In particular, CARM was, in average, almost 300 times faster than CRM.

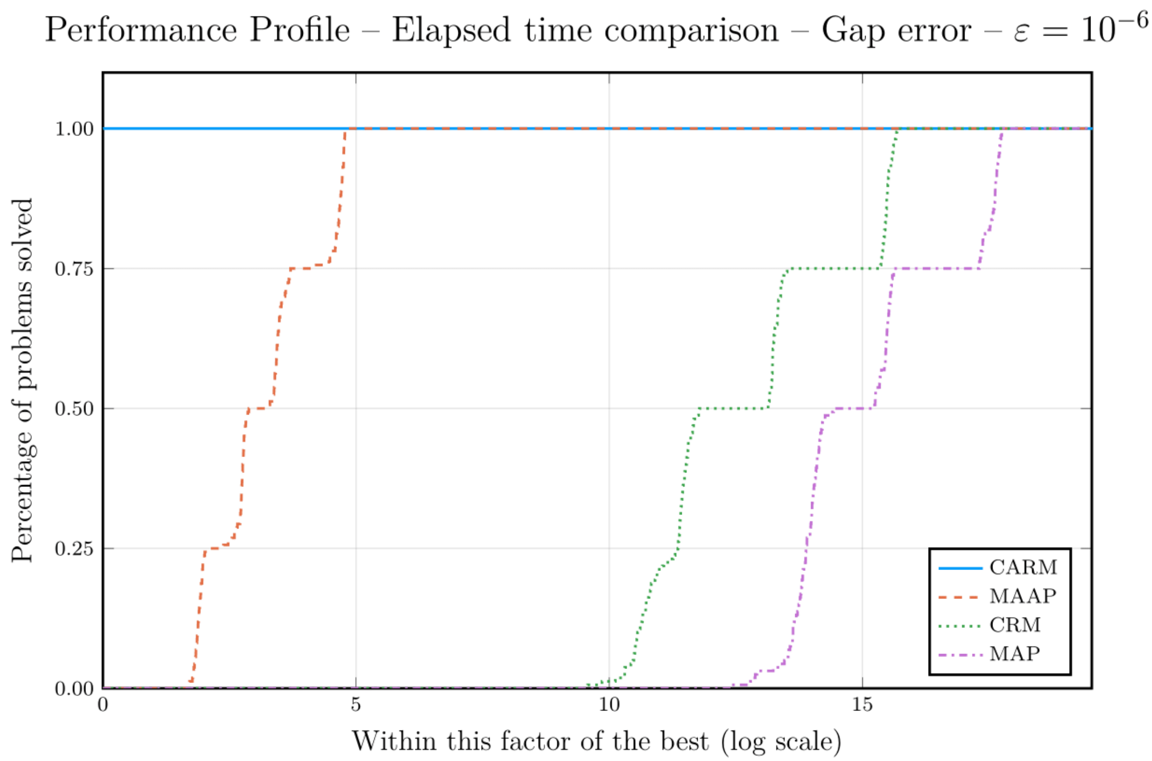

Figure 4.1: Performance profile of experiments with ellipsoidal feasibility – CARM, MAAP, CRM and MAP

Table 4.1: Statistics of the experiments (in number of iterations and CPU time)

Method		**mean**	**max**	**min**	**std**
CARM	`it`	33.4520	346.0000	8.0000	40.9777
	`CPU (s)`	4.6302×10^{-2}	1.7262	6.8472×10^{-4}	0.125 07
MAAP	`it`	926.9000	2631.0000	111.0000	796.7943
	`CPU (s)`	5.4588×10^{-1}	5.4278	3.0145×10^{-3}	1.0006
CRM	`it`	5.4440	8.0000	3.0000	0.8683
	`CPU (s)`	11.8923	84.9785	9.6263×10^{-2}	15.6269
MAP	`it`	922.3360	2578.0000	104.0000	795.5403
	`CPU (s)`	88.4846	706.0387	0.7120	138.4990

4.5.1 Projection onto an ellipsoid

In this section, we describe ADMM, which is the algorithm used for computing the projections onto ellipsoids required in all our numerical examples. Projecting a point onto an ellipsoid is one of the fundamental problems in convex analysis and numerical algorithms. Recently, several fast algorithms were proposed for solving this problem such as Lin-Han algorithm, maximum 2-dimensional inside ball algorithm, sequential 2-dimensional projection algorithm and hybrid projection algorithms of Dai ([30], [50], [44]). The problem of projecting a point onto a general ellipsoid:

$$\begin{aligned} &\min \ \text{dist}(a, x) = ||x - a|| \\ &\text{s.t} \quad x \in \xi := \{x \in \mathbb{R}^n \mid g(x) \le \alpha\}, \end{aligned} \tag{4.5.2}$$

where $a \in \mathbb{R}^n$ is a point to be projected, $g : \mathbb{R}^n \to \mathbb{R}$ is given by $g(x) = x^t Ax + 2x^t b$, where A is a symmetric positive definite matrix, b is a n-vector.

In [44] the problem is considered as a constrained convex optimization problem with a separable objective function, which enables the use of the alternating direction method of multipliers (ADMM). All above mentioned methods converge with a global linear rate. Jia et al. in [44] show that theoretically and numerically ADMM is the most efficient one. We remark that the efficiency of the algorithms was compared not only in terms of the number of iterations, but also in terms of the cost per iteration.

Thus, we chose ADMM for the projections onto ellipsoids in our numerical experiments. In the next subsection, we describe in some detail how actually the alternating direction method works.

ADMM

The alternating direction method of multipliers (ADMM) was proposed and studied in [39] for the following separable convex optimization problem

$$\begin{aligned} &min \quad \theta_1(x) + \theta_2(y) \\ &s.t. \quad A_1 x + A_2 y = c, \\ &\qquad\quad x \in X, y \in Y, \end{aligned} \tag{4.5.3}$$

where $\theta_1 : \mathbb{R}^n \to \mathbb{R}$ and $\theta_2 : \mathbb{R}^m \to \mathbb{R}$ are convex functions, $A_1 \in \mathbb{R}^{l\times n}$ and $A_2 \in \mathbb{R}^{l\times m}$ are two given matrices are $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$ are simple closed convex sets. The iterative scheme of ADMM reads as

$$.\begin{cases} x^{k+1} = \underset{x\in X}{\textbf{argmin}} \left\{ \theta_1(x) - x^t A_1^t \lambda^k + \dfrac{\nu}{2} \left\| A_1 x + A_2 y^k - c \right\|^2 \right\}, & (4.5.4) \\ y^{k+1} = \underset{y\in Y}{\textbf{argmin}} \left\{ \theta_2(x) - y^t A_2^t \lambda^k + \dfrac{\nu}{2} \left\| A_1 x^{k+1} + A_2 y - c \right\|^2 \right\}, & (4.5.5) \\ \lambda^{k+1} = \lambda^k - \nu \left(A_1 x^{k+1} + A_2 y^{k+1} - c \right) & (4.5.6) \end{cases}$$

whereλ^k is the Lagrange multiplier and $\nu > 0$ is the penalty parameter. In our numerical experiments, we took $\nu = 10/||A||$. In order to apply ADMM for the projecting problem, we first need to reformulate (4.5.2) in the format of (4.5.3). Because A is a positive definite symmetric matrix in $R^{n\times n}$, it has a symmetric matrix square root $B = A^{\frac{1}{2}} := \sqrt{A}$, satisfying $B^2 = A$. So we can rewrite (4.5.2) as

$$\begin{aligned} &\min \quad \frac{1}{2} \left\| x - a \right\|^2 \\ &s.t. \quad x \in \xi := \left\{ x \in \mathbb{R}^n : (x^t B)(Bx) + 3 - 2b^t B^{-1}(Bx) \le \alpha \right\}, \\ &\qquad\quad = \left\{ x \in \mathbb{R}^n : \left\| Bx + B^{-t} b \right\|^2 \le \left\| B^{-t} b \right\|^2 \alpha \right\}. \end{aligned}$$

Let $y = Bx + B^{-t}b, \bar{b} = -B^{-t}b$, and $\theta_1(x) := \frac{1}{2}\|x-a\|^2, \theta_2(y) = 0$. We reformulate (4.5.3) as

$$\begin{aligned} & min \quad \theta_1(x) + \theta_2(y) \\ & s.t. \quad Bx - y = \bar{b}, \\ & \qquad x \in X, y \in Y, \end{aligned} \tag{4.5.7}$$

where $X = R^n$ and $Y := \left\{y \in \mathbb{R}^n : \|y\|^2 \leq \alpha + \|b\|^2\right\}$ is a ball in $\mathbb{R}^n$. Then, the iterative scheme (4.5.4) can be specialized to

$$x^{k+1} = \underset{x\in X}{\mathbf{argmin}} \left\{\frac{1}{2}\|x-a\|^2 - x^t B^t \lambda^k + \frac{\nu}{2}\left\|Bx - y^k - \bar{b}\right\|^2\right\}, \tag{4.5.8}$$

amounting to finding the solution of the linear system

$$\left(I + \nu B^t B\right) x^{k+1} = a + B^t\lambda^k + \nu B^t\left(y^k + \bar{b}\right). \tag{4.5.9}$$

The subproblem (4.5.5) reduces to

$$\begin{aligned} y^{k+1} &= \underset{y\in Y}{\mathbf{argmin}} \left\{y^t\lambda^k + \frac{\nu}{2}\left\|Bx^{k+1} - y - \bar{b}\right\|^2\right\}, \\ &= \underset{y\in Y}{\mathbf{argmin}} \left\{\frac{\nu}{2}\left\|y - \left(Bx^{k+1} - \left(\frac{1}{\nu}\right)\lambda^k - \bar{b}\right)\right\|^2\right\}. \end{aligned} \tag{4.5.10}$$

Consequently,

$$y^{k+1} = P_y[Bx^{k+1} - \left(\frac{1}{\nu}\right)\lambda^k - \bar{b},] \tag{4.5.11}$$

where $P_y[\cdot]$ is the Euclidean projection onto Y, taking the form

$$P_Y[y] = \begin{cases} y, & if\ \|y\|^2 \leq \alpha + \left\|\bar{b}\right\|^2 \\ \frac{y}{\|y\|}\sqrt{\alpha + \left\|\bar{b}\right\|^2}, & \text{otherwise.,} \end{cases}$$

Finally, the multiplier update scheme (4.5.6) is

$$\lambda^{k+1} = \lambda^k - \nu\left(Bx^{k+1} - y^{k+1} - \bar{b}\right). \tag{4.5.12}$$

The detail of ADMM for solving (4.5.2) is summarized in Algorithm 6.

Algorithm 6:ADMM for (2.10)

1: Initialize $(y^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^n$ and $\epsilon > 0$, set $k := 0$.
2: Compute $r = \sqrt{\alpha + \left\|\bar{b}\right\|^2}$ and $\bar{A} = \left(I + \nu B^t B\right)^{-1}$.
3: Compute $u^k = \alpha + B^t\lambda^k + \nu B^t\left(y^k + \bar{b}\right)$ and $x^{k+1} = \bar{A}u^k$.
4: Compute $w^k = Bx^{k+1} - \left(\frac{1}{\nu}\right)\lambda^k - \bar{b}$.
If $\left\|w^k\right\| \leq r$, then $y^{k+1} = w^k$; otherwise, $y^{k+1} = \frac{r}{w^k}w^k$. 5: Calculate $\lambda^{k+1} = \lambda^k - \nu\left(Bx^{k+1} - y^{k+1} - \bar{b}\right)$.
6: If $\left\|R\left(x^{k+1}, y^{k+1}, \lambda^{k+1}\right)\right\| \leq \epsilon$ does not hold, set $k := k+1$ and go to step 3.

Here, $R(x, y, \lambda)$ denotes the residual of the optimally condition of (4.5.7) and is defined as

$$R(x, y, z) := \begin{pmatrix} R_x(x, y, \lambda) \\ R_y(x, y, \lambda) \\ R_\lambda(x, y, \lambda) \end{pmatrix} = \begin{pmatrix} x - a - B^t \lambda \\ y - P_Y[y - \lambda] \\ B_x - y - \bar{b} \end{pmatrix} \tag{4.5.13}$$

See [44] for an in-depth study of ADMM and other algorithms to solve (4.5.2).

Chapter 5

A circumcentered-reflection method for finding common fixed points of firmly nonexpansive operators

In this chapter, we apply CRM to the Fixed-Point Problem (denoted as FPP), consisting of finding a common fixed-point of a finite set of firmly nonexpansive operators. We prove that in this setting, CRM is globally convergent to a common fixed-point (supposing that at least one exists). We also establish linear convergence of the sequence generated by CRM applied to FPP, under a not too demanding error bound assumption, and provide an estimate of the asymptotic constant. We provide solid numerical evidence of the superiority of CRM when compared to the classical Parallel Projections Method (PPM).

5.1 Some properties of firmly nonexpansive operators

We start with some elementary properties of nonexpansive operators. By Definition 2.5.1, an operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be firmly nonexpansive when

$$\|T(x) - T(y)\|^2 \leq \|x - y\|^2 - \|(T(x) - T(y)) - (x - y)\|^2 \tag{5.1.1}$$

for all $x, y \in \mathbb{R}^n$. For an operator $T : \mathbb{R}^n \to \mathbb{R}^n$, we denote as $F(T)$ the set of its fixed-points, *i.e.*, $F(T) = \{x \in \mathbb{R}^n : T(x) = x\}$.

Proposition 5.1.1. Convex combinations of firmly nonexpansive operators are firmly nonexpansive.

Proof. Take firmly nonexpansive operators $T_1 \ldots, T_m$ and positive scalars $\alpha_1, \ldots, \alpha_m$ such that $\sum_{i=1}^m \alpha_i = 1$. Let $\overline{T} = \sum_{i=1}^m \alpha_i T_i$. We prove next that $\overline{T}$ is firmly nonexpansive.
Note that by Proposition 2.5.1(iv), (5.1.1) is equivalent to

$$\|T(x) - T(y)\|^2 \leq \langle T(x) - T(y), x - y \rangle. \tag{5.1.2}$$

It suffices to check that $\overline{T}$ satisfies (5.1.2), and we proceed to do so.

$$\left\|\overline{T}(x) - \overline{T}(y)\right\|^2 = \left\|\sum_{i=1}^m \alpha_i (T_i(x) - T_i(y))\right\|^2 \leq \sum_{i=1}^m \alpha_i \|T_i(x) - T_i(y)\|^2$$

$$\leq \sum_{i=1}^{m} \alpha_i \langle T_i(x) - T_i(y), x - y \rangle = \left\langle \sum_{i=1}^{m} \alpha_i \left(T_i(x) - T_i(y)\right), x - y \right\rangle = \langle \overline{T}(x) - \overline{T}(y), x - y \rangle,$$

using the convexity of $\|\cdot\|^2$ in the first inequality and the fact that the T_i's satisfy (5.1.2) in the second one. ■

It is worthwhile to comment at this point that the composition of two firmly nonexpansive operators may fail to be firmly nonexpansive: consider $A = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 0\}$, $B = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = x_1\}$. P_A and P_B are firmly nonexpansive by Proposition 2.2.2, but its composition $P_B \circ P_A$ fails to satisfy (5.1.1) with $x = (0, 0)$ and $y = (2, -1)$.
We present next some properties of the set of fixed-points of combinations of orthogonal projections. They have been proved, e.g. in [31], [43], but we include the proofs for the sake of completeness.

Proposition 5.1.2. Consider closed convex sets $C_1, \ldots, C_m \subset \mathbb{R}^n$ and positive scalars $\alpha_1, \ldots, \alpha_m$ such that $\sum_{i=1}^m \alpha_i = 1$. Denote $P_i = P_{C_i}$ and let $\overline{P} = \sum_{i=1}^m \alpha_i P_i$. Define $g : \mathbb{R}^n \to \mathbb{R}$ as $g(x) = \sum_{i=1}^m \alpha_i \|x - P_i(x)\|^2 = \sum_{i=1}^m \alpha_i \mathrm{dist}(x, C_i)^2$ and let $C = \cap_{i=1}^m C_i$. Then,

i) $F(\overline{P}) = \{x \in \mathbb{R}^n : \nabla g(x) = 0\}$, *i.e.*, since g is convex, the set of fixed-points of $\overline{P}$ (if nonempty) is precisely the set of minimizers of g.

ii) If $C \neq \varnothing$, then $F(\overline{P}) = C$.

Proof.

i) By Proposition 2.2.3(ii),

$$\nabla g(x) = 2 \sum_{i=1}^{m} \alpha_i (x - P_i(x)) = 2 \left(x - \sum_{i=1}^{m} \alpha_i P_i(x) \right) = 2(x - \overline{P}(x)),$$

so that $\nabla g(x) = 0$ iff $x = \overline{P}(x)$ iff $x \in F(\overline{P})$.

ii) Clearly, $C \subset F(\overline{P})$. For the converse inclusion note that when $C \neq \varnothing$, we have $g(x) = 0$ for all $x \in C$, so that the minimum value of g is indeed 0, and the set of minimizers of g coincides with the set of its zeroes, which is C, because $g(x) > 0$ whenever $x \notin C$. The result follows then from item (i).

■

We deal now with the main result of this section, which we describe next. The prototypical examples of firmly nonexpansive operators are the orthogonal projections onto closed convex sets. Proposition 5.1.1 provides a larger class of firmly nonexpansive operators, namely convex combinations of orthogonal projections. It is therefore relevant to check that the second class is indeed larger, *i.e.,* that, generically, convex combinations of orthogonal projections are not orthogonal projections themselves. We will prove that this is indeed the case when the intersection of the convex sets is nonempty. However, when this intersection is empty, a convex combination of orthogonal projections may be itself an orthogonal projection.

Proposition 5.1.3. Consider closed convex sets $C_1, \ldots, C_m \subset \mathbb{R}^n$ and positive scalars $\alpha_1, \ldots, \alpha_m$ such that $\sum_{i=1}^m \alpha_i = 1$. Denote $C = \cap_{i=1}^m C_i$, $P_i = P_{C_i}$ and let $\overline{P} = \sum_{i=1}^m \alpha_i P_i$. Assume that $C \neq \varnothing$. If there exists $E \subset \mathbb{R}^n$ such that $\overline{P} = P_E$ then $E = C_1 = \cdots = C_m$.

Proof. By Propositions 5.1.2(ii) and 2.2.2,

$$C = F(\overline{P}) = F(P_E) = E. \tag{5.1.3}$$

Take $x \in C_i$. Let $\ell = \operatorname{argmax}_{1 \leq j \leq m}\{\|x - P_j(x)\|\}$, $w = \sum_{j=1}^m \alpha_j P_j(x) = \overline{P}(x) = P_E(x)$, so that $w \in \operatorname{Im}(P_E) = E = C$, using (5.1.3), and hence $w \in C_\ell$. It follows that

$$\|x - P_\ell(x)\| \leq \|x - w\| = \left\| \sum_{i=j}^m \alpha_j (x - P_j(x)) \right\| \leq \sum_{j=1}^m \alpha_j \|x - P_j(x)\|$$

$$= \sum_{j=1, j\neq i}^m \alpha_j \|x - P_j(x)\| \leq \sum_{j=1, j\neq i}^m \alpha_j \|x - P_\ell(x)\| =$$

$$\left(\sum_{j=1, j\neq i}^m \alpha_j \right) \|x - P_\ell(x)\| = (1 - \alpha_i) \|x - P_\ell(x)\|, \tag{5.1.4}$$

using the convexity of $\|\cdot\|$ in the first inequality, the fact that $x \in C_i$ in the second equality and the definition of ℓ in the second inequality. It follows from (5.1.4) that $\alpha_i \|x - P_\ell(x)\| \leq 0$, so that $\|x - P_\ell(x)\| = 0$. Since $0 \leq \|x - P_j(x)\| \leq \|x - P_\ell(x)\|$ for all j by definition of ℓ, we conclude that $\|x - P_j(x)\| = 0$ for all j, *i.e.*, $x \in C_j$. Since x is an arbitrary point in C_i, we get that that $C_i \subset C_j$ for all i, j, *i.e.*, $C_1 = \cdots = C_m$, and the result follows immediately from (5.1.3). ■

5.2 Convergence of CRM applied to FPP

In this section we establish convergence of CRM applied to finding a point in $\operatorname{Fix}(T, P_U)$, where $T : \mathbb{R}^n \to \mathbb{R}^n$ is firmly nonexpansive and $P_U : \mathbb{R}^n \to \mathbb{R}^n$ is the orthogonal projection onto an affine manifold $U \subset \mathbb{R}^n$. As explained in Section 1.4, through Pierra's formalism in the product space $\mathbb{R}^{nm}$, this result entails convergence of CRM applied to finding a point in $\operatorname{Fix}(T_1, \ldots, T_m)$, where $T_i : \mathbb{R}^n \to \mathbb{R}^n$ is firmly nonexpansive for $1 \leq i \leq m$.
Our convergence analysis for CRM requires comparing the CRM and the MAP sequences, so that we start by proving convergence of the second one, defined as

$$z^{k+1} = P_U(T(z^k)), \tag{5.2.1}$$

starting at some $z^0 \in \mathbb{R}^n$. This is a classical result, but we include it for the sake of self-containment. We start with the following intermediate result.

Proposition 5.2.1. For all $x \in \mathbb{R}^n$ and all $y \in \operatorname{Fix}(T, P_U)$ it holds that

$$\|P_U(T(x)) - y\|^2 \leq \|x - y\|^2 - \|T(x) - x\|^2 - \|P_U(T(x)) - T(x)\|^2. \tag{5.2.2}$$

Proof. Since P_U is firmly nonexpansive by Proposition 2.2.2, we have

$$\|P_U(x) - y\|^2 \leq \|x - y\|^2 - \|P_U(x) - x\|^2 \tag{5.2.3}$$

for all $x \in \mathbb{R}^n$, using the fact that $y \in U$. Substituting $T(x)$ for x in (5.2.3), we obtain

$$\|P_U(T(x)) - y\|^2 \leq \|T(x) - y\|^2 - \|P_U(T(x)) - T(x)\|^2 . \tag{5.2.4}$$

Since T is firmly nonexpansive,

$$\|T(x) - y\|^2 \leq \|x - y\|^2 - \|T(x) - x\|^2 . \tag{5.2.5}$$

Now combining (5.2.4) with (5.2.5), we get

$$\|P_U(T(x)) - y\|^2 \leq \|x - y\|^2 - \|T(x) - x\|^2 - \|P_U(T(x)) - T(x)\|^2 \tag{5.2.6}$$

which implies the result. ■

Using Proposition 5.2.1 we get convergence of $\{z^k\}_{k\in\mathbb{N}}$ using the classical argument for MAP applied to CFP, as we show next:

Proposition 5.2.2. If $\operatorname{Fix}(T, P_U) \neq \varnothing$, then the sequence $\{z_k\}_{k\in\mathbb{N}}$ defined by (5.2.1) converges to a point $\bar{z} \in \operatorname{Fix}(T, P_U)$.

Proof. Take any $y \in \operatorname{Fix}(T, P_U)$. By (5.2.1), $z^{k+1} = P_U(T(z^k))$. Using (5.2.2), we get

$$\begin{aligned}\left\|z^{k+1} - y\right\|^2 &\leq \left\|z^k - y\right\|^2 - \left\|P_U(T(z^k)) - T(z^k)\right\|^2 - \left\|T(z^k) - z^k\right\|^2 \\ &\leq \left\|z^k - y\right\|^2 .\end{aligned} \tag{5.2.7}$$

It follows from (5.2.7) that $\left\|z^{k+1} - y\right\|^2 \leq \left\|z^k - y\right\|$ for all $k \in \mathbb{N}$, so that $\{z^k\}_{k\in\mathbb{N}}$ is bounded and $\{\left\|z^k - y\right\|\}_{k\in\mathbb{N}}$ is nonincreasing and nonnegative, therefore convergent.
Hence, rewriting (5.2.7) as

$$\left\|P_U(T(z^k)) - T(z^k)\right\|^2 + \left\|T(z^k) - z^k\right\|^2 \leq \left\|z^k - y\right\|^2 - \left\|z^{k+1} - y\right\|^2 ,$$

we conclude that

$$\lim_{k\to\infty} \left\|T(z^k) - z^k\right\| = 0. \tag{5.2.8}$$

Let $\bar{z}$ be a cluster point of the bounded sequence $\{z^k\}_{k\in\mathbb{N}}$. Taking limits in (5.2.8) along a subsequence converging to $\bar{z}$, and using the continuity of T, established in Proposition 2.5.4, we get that $T(\bar{z}) = \bar{z}$. Since $z^k \in U$ for all $k \in \mathbb{N}$ by (5.2.1), we have that $\bar{z} \in U$, so that $\bar{z} \in \operatorname{Fix}(T, P_U)$. Taking now $y = \bar{z}$ in (5.2.7), we conclude that $\{\left\|z^k - \bar{z}\right\|\}_{k\in\mathbb{N}}$ is convergent, and since a subsequence of this sequence converges to 0, the whole sequence $\{\left\|z^k - \bar{z}\right\|\}_{k\in\mathbb{N}}$ converges to 0, *i.e.*, $\lim_{k\to\infty} z^k = \bar{z} \in \operatorname{Fix}(T, P_U)$. ■

Now we proceed to the convergence analysis of CRM applied to FPP. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a firmly nonexpansive operator, $U \subset \mathbb{R}^n$ an affine manifold, and $P_U : \mathbb{R}^n \to \mathbb{R}^n$ the orthogonal projection onto U. We assume that $\operatorname{Fix}(T, P_U) \neq \varnothing$. We denote as R, R_U the reflection operators related to T, P_U respectively, *i.e.*, $R(x) = 2T(x) - x$, $R_U(x) = 2P_U(x) - x$. We recall

that $C : \mathbb{R}^n \to \mathbb{R}^n$ is the CRM operator, *i.e.*, $C(z) = \text{circ}\{z, R(z), R_U(R(z))\}$, where "circ" denotes the circumcenter of three points, as defined in 1.2.1. We also define $D : \mathbb{R}^n \to \mathbb{R}^n$ as $D(x) = P_U(T(x))$, so that D can be seen the MAP operator.
We will prove that, starting from any initial point $x^0 \in U$, the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by CRM, defined as $x^{k+1} = C(x^k)$, converges to a point in $\text{Fix}(T, P_U)$.
Our convergence analysis is close to our results in Chapter 3, for CRM applied to CFP, but with several differences, resulting from the fact that now T is an arbitrary firmly nonexpansive operator, rather than the orthogonal projection onto a convex set. One of the differences is the use of the next property of circumcenters, which will substitute for a specific property of orthogonal projections.

Proposition 5.2.3. For all $x \in \mathbb{R}^n$, $\langle x - T(x), C(x) - T(x)\rangle = 0$.

Proof. By the definition of the reflection, for all $x \in \mathbb{R}^n$,

$$T(x) = \frac{1}{2}(R(x) + x). \tag{5.2.9}$$

By the definition of circumcenter, for all $x \in \mathbb{R}^n$,

$$\|C(x) - x\|^2 = \|C(x) - R(x)\|^2. \tag{5.2.10}$$

Expanding (5.2.10) and rearranging, we get

$$2\langle x - R(x), C(x)\rangle = \|x\|^2 - \|R(x)\|^2. \tag{5.2.11}$$

Substracting $2\langle x - R(x), T(x)\rangle$ from both sides of (5.2.11) and using (5.2.9), we obtain

$$\begin{aligned} 4\langle x - T(x), C(x) - T(x)\rangle &= 2\langle x - R(x), C(x) - T(x)\rangle \\ &= \|x\|^2 - \|R(x)\|^2 - 2\langle x - R(x), T(x)\rangle \\ &= \|x\|^2 - \|R(x)\|^2 - \langle x - R(x), x + R(x)\rangle = 0, \end{aligned}$$

which implies the result. ■

Next we establish a basic property of the circumcenter, which ensures that the CRM sequence, starting at a point in U, remains in U.

Proposition 5.2.4. If $z \in U$ then $C(z) \in U$.

Proof. We consider three cases. If $R(z) \in U$ then $R_U(R(z)) = R(z)$, in which case $z, R(z), R_U(R(z)) \in U$, so that the affine hull of these three points is contained in U. Since by definition $C(z)$ belongs to this affine hull, the result holds. If $z = P_U(R(z))$ then the affine hull of $\{z, R(z), R_U(R(z))\}$ is the line determined by $R(z)$ and $R_U(R(z))$, since $P_U(R(z) = \frac{1}{2}(R(z) + R_U(R(z)))$ and

$$C(z) = \text{circ}\{z, R(z), R_U(R(z))\} = \text{circ}\{R(z), R_U(R(z))\} = P_U(R(z)) = z \in U,$$

so that the result holds.

Assume that $z \neq P_U(R(z))$ and that $R(z) \notin U$. We claim that $C(z)$ belongs to the line passing through z and $P_U(R(z))$. Observe that, since $\|C(z) - R(z)\| = \|C(z) - R_U(R(z))\|$, $C(z)$ belongs to the hyperplane orthogonal to $R(z) - R_U(R(z))$ passing through $\frac{1}{2}(R(z), R_U(R(z)))$ $= P_U(R(z))$, say H. On the other hand, by definition, $C(z)$ belongs to the affine manifold E spanned by $z, R(z), R_U(R(z))$. So, $C(z) \in E \cap U$. Since $R(z) \notin U$, $\dim(E \cap U) < \dim(E) \leq 2$. Note that $P_U(z) = \frac{1}{2}(R(z) + R_U(R(z))) = P_U(R(z))$ belongs to E. Hence the line through $z, P_U(R(z))$, say L, is contained in E, and by a dimensionality argument we conclude that $L = E$. Since $C(z) \in E$, we get that $C(z) \in L$. Since $z, P_U(R(z))$ belong to U, we have that $C(z) \in L \subset U$, completing the proof. ■

We continue with an important intermediate result.

Proposition 5.2.5. Consider the operators $C, D : \mathbb{R}^n \to \mathbb{R}^n$ defined above. Then $D(x)$ belongs to the segment between x and $C(x)$ for all $x \in U$.

Proof. Let E denote the affine manifold spanned by $x, R(x)$ and $R_U(R(x))$. By definition, the circumcenter of these three points, namely $C(x)$, belongs to E. We claim that $D(x)$ also belongs to E. We proceed to prove the claim. Since U is an affine manifold, P_U is an affine operator, so that $P_U(\alpha x + (1-\alpha)x') = \alpha P_U(x) + (1-\alpha)P_U(x')$ for all $\alpha \in \mathbb{R}$ and all $x, x' \in \mathbb{R}^n$. Thus $R_U(R(x)) = 2P_U(R(x)) - R(x)$, so that

$$P_U(R(x)) = \frac{1}{2}\left(R_U(R(x)) + R(x)\right). \tag{5.2.12}$$

On the other hand, using the affinity of P_U, the definition of D and the assumption that $x \in U$, we have

$$P_U(R(x)) = P_U(2T(x) - x) = 2P_U(T(x)) - P_U(x) = 2D(x) - x, \tag{5.2.13}$$

so that

$$D(x) = \frac{1}{2}\left(P_U(R(x)) + x\right). \tag{5.2.14}$$

Combining (5.2.12) and (5.2.14),

$$D(x) = \frac{1}{2}x + \frac{1}{4}R_U(R(x)) + \frac{1}{4}R(x),$$

i.e., $D(x)$ is a convex combination of $x, R_U(R(x))$ and $R(x)$. Since these three points belong to E, the same holds for $D(x)$ and the claim holds.

We observe now that $x \in U$ by assumption, $D(x) \in U$ by definition, and $C(x) \in U$ by Proposition 5.2.4. Now we consider three cases: if $\dim(E \cap U) = 0$ then $x, D(x)$ and $C(x)$ coincide and the result holds trivially. If $\dim(E \cap U) = 2$ then $E \subset U$, so that $R(x) \in U$ and hence $R_U(R(x)) = R(x)$, in which case $C(x)$ is the midpoint between x and $R(x)$, which is precisely $T(x)$. Hence, $T(x) \in U$, so that $D(x) = P_U(T(x)) = T(x) = C(x)$, implying that $D(x)$ and $C(x)$ coincide, and the result holds trivially. The interesting case is the remaining one, *i.e.*, $\dim(E \cap U) = 1$. In this case $x, D(x)$ and $C(x)$ lie in a line, so that we can write $C(x) = x + \eta(D(x) - x)$ with $\eta \in \mathbb{R}$, and it suffices to prove that $\eta \geq 1$.

By the definition of η,

$$\|C(x) - x\| = |\eta| \, \|T(x) - x\|. \tag{5.2.15}$$

Since $C(x) \in U$, nonexpansiveness of P_U implies that

$$\|C(x) - R(x)\| \geq \|C(x) - P_U(R(x))\| . \tag{5.2.16}$$

Then

$$\begin{aligned} \|C(x) - x\| &= \|C(x) - R(x)\| \geq \|C(x) - P_U(R(x))\| \\ &= \|(C(x) - x) - (P_U(R(x)) - x)\| \\ &= \|\eta\,(D(x) - x) - 2\,(D(x) - x)\| \\ &= |\eta - 2|\; \|D(x) - x\| , \end{aligned} \tag{5.2.17}$$

using the definition of the circumcenter in the first equality, (5.2.16) in the inequality, and the definition of η and S in the third equality. Combining (5.2.15) and (5.2.17), we get

$$|\eta|\; \|D(x) - x\| \geq |\eta - 2|\; \|D(x) - x\| ,$$

implying that $|\eta| \geq |2 - \eta|$, which holds only when $\eta \geq 1$, completing the proof. ■

We continue with a key result for the convergence analysis of CRM, comparing the behavior of the CRM and the MAP operators. Again the argument in this proof differs from the case of MAP and CRM applied to CFP, presented in Chapter 3.

Proposition 5.2.6. With the notation of Proposition 5.2.5, for all $y \in \operatorname{Fix}(T, P_U)$ and all $z \in U$, it holds that

i) $\|C(z) - y\| \leq \|D(z) - y\|$,

ii) $\operatorname{dist}(C(z), \operatorname{Fix}(T, P_U)) \leq \operatorname{dist}(D(z), \operatorname{Fix}(T, P_U))$,

Proof.

i) Take $z \in U, y \in \operatorname{Fix}(T, P_U)$. If $z \in F(T)$, then the result follows trivially, because then $P_U(T(z)) = z = C(z)$ and there is nothing to prove. So, assume that $z \in U \setminus F(T)$. We claim that

$$\|P_U(T(z)) - z\| \leq \|T(z) - z\| \leq \|C(z) - z\| . \tag{5.2.18}$$

For proving the first inequality in (5.2.18), we conclude, from the fact that $z \in U$ and nonexpansiveness of orthogonal projections, that

$$\|P_U(T(z)) - z\| \leq \|T(z) - z\| . \tag{5.2.19}$$

Since $R(z) = 2T(z) - z$,we get that

$$\|R(z) - z\| = 2\,\|T(z) - z\| . \tag{5.2.20}$$

Using (5.2.19) and (5.2.20),

$$\begin{aligned}
\|T(z) - z)\| &= \frac{1}{2}\|R(z) - z\| \\
&= \frac{1}{2}\|(R(z) - C(z) + C(z) - z)\| \\
&\leq \frac{1}{2}(\|R(z) - C(z)\| + \|C(z) - z\|) \\
&= \frac{1}{2}(\|z - C(z)\| + \|C(z) - z\|) \qquad (5.2.21)\\
&= \|C(z) - z\|.
\end{aligned}$$

where the third equality holds because $C(z)$ is equidistant from z, $R(z)$, and $R_U(R(z))$. The claim follows then from (5.2.18) and (5.2.21).

By Proposition 5.2.5, $D(z)$ belongs to the segment between z and $C(z)$, *i.e.*, there exists $\alpha \in [0,1]$ such that $D(z) = \alpha C(z) + (1-\alpha)z$ and $\alpha < 1$ because $z \notin F(T)$, so that

$$D(z) - C(z) = \frac{1-\alpha}{\alpha}(z - D(z)). \qquad (5.2.22)$$

Note that

$$\begin{aligned}
\langle z - D(z), C(z) - y\rangle &= \langle z - T(z), C(z) - T(z)\rangle + \langle z - T(z), T(z) - y\rangle \\
&\quad + \langle T(z) - D(z), C(z) - y\rangle. \qquad (5.2.23)
\end{aligned}$$

Now we look at the three terms in the right-hand side of (5.2.23). The first one vanishes as a consequence of Proposition 5.2.3. The third one vanishes because $D(z) = P_U(T(z))$, and U is an affine manifold, so that $T(z)-D(z)$ is orthogonal to any vector in U, as is the case for $C(z)-y$, since $y \in U$ by assumption and $C(z) \in U$ by Proposition 5.2.4. The second term is nonnegative by Proposition 2.5.5. It follows hence from (5.2.23) that

$$\langle z - D(z), C(z) - y\rangle \geq 0. \qquad (5.2.24)$$

Now, (5.2.24) together with (5.2.22) gives us

$$\langle D(z) - C(z), y - C(z)\rangle = \frac{1-\alpha}{\alpha}\langle z - D(z), y - C(z)\rangle \leq 0. \qquad (5.2.25)$$

It follows from (5.2.25) that $\|C(z) - y\| \leq \|D(z) - y\|$ for all $y \in \operatorname{Fix}(T, P_U)$ and all $z \in U$, establishing (i).

ii) Let $\bar{z}, \hat{z} \in \operatorname{Fix}(T, P_U)$ realize the distance from $C(z), D(z)$ to $\operatorname{Fix}(T, P_U)$ respectively. Then, in view of (i),

$$\begin{aligned}
\operatorname{dist}(C(z), \operatorname{Fix}(T, P_U)) &= \|C(z) - \bar{z}\| \leq \|C(z) - \hat{z}\| \\
&\leq \|D(z) - \hat{z}\| = \operatorname{dist}(D(z), \operatorname{Fix}(T, P_U))
\end{aligned}$$

proving (ii).

■

Next, we complete the convergence analysis of CRM applied to FPP. Here again, the proofline differs from the one in Proposition 2.4.3, where a specific property of orthogonal projections was used to characterize $C(z)$ as the projection onto a certain set, which does not work when T is an arbitrary firmly nonexpansive operator.

Theorem 5.2.1. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a firmly nonexpansive operator and $U \subset \mathbb{R}^n$ an affine manifold. Assume that $\operatorname{Fix}(T, P_U) \neq \varnothing$. Let $\{x^k\}_{k\in\mathbb{N}}$ be the sequence generated by CRM for solving $\operatorname{FPP}(T, P_U)$, *i.e.*, $x^{k+1} = C(x^k)$. If $x^0 \in U$, then $\{x^k\}_{k\in\mathbb{N}}$ is contained in U and converges to a point in $\operatorname{Fix}(T, P_U)$.

Proof. The fact that $\{x^k\}_{k\in\mathbb{N}} \subset U$ results from invoking Proposition 5.2.4 in an inductive way, starting with the assumption that $x^0 \in U$.
Take any $y \in \operatorname{Fix}(T, P_U)$ Then,

$$\left\|x^{k+1} - y\right\|^2 = \left\|C(x^k) - y\right\|^2 \leq \left\|D(x^k) - y\right\|^2 \leq \left\|x^k - y\right\|^2 - \left\|D(x^k) - x^k\right\|^2 \tag{5.2.26}$$

where the first inequality follows from Proposition 5.2.6(i), and the second one follows from Proposition 5.2.1, since $P_U(x^k) = x^k$ by Proposition 5.2.4 and $D = P_U \circ T$.
(5.2.26) says that $\{x^k\}_{k\in\mathbb{N}}$ is Fejér monotone with respect to $\operatorname{Fix}(T, P_U)$, and the remainder of the proof is standard. By (5.2.26), $\{x^k\}_{k\in\mathbb{N}}$ is bounded and $\{\left\|x^k - y\right\|\}_{k\in\mathbb{N}}$ is nonincreasing and nonnegative, hence convergent, for all $y \in \operatorname{Fix}(T, P_U)$. It follows also from (5.2.26) that

$$\lim_{k\to\infty} D(x^k) - x^k = 0. \tag{5.2.27}$$

Let $\bar{x}$ be any cluster point of $\{x^k\}_{k\in\mathbb{N}}$. Taking limits in (5.2.27) along a subsequence converging to $\bar{x}$, we conclude that $D(\bar{x}) = \bar{x}$, *i.e.*, $\bar{x} \in F(D) = \operatorname{Fix}(T, P_U)$, so that all cluster points of $\{x^k\}_{k\in\mathbb{N}}$ belong to $\operatorname{Fix}(T, P_U)$. Looking now (5.2.26) with $\bar{x}$ substituting for y, we get that $\{\left\|x^k - \bar{x}\right\|\}_{k\in\mathbb{N}}$ is a nonincreasing sequence with a subsequence converging to 0, so that the whole sequence $\{\left\|x^k - \bar{x}\right\|\}_{k\in\mathbb{N}}$ converges to 0. It follows that $\bar{x}$ is the unique cluster point of $\{x^k\}_{k\in\mathbb{N}}$, so that $\lim_{k\to\infty} x^k = \bar{x} \in \operatorname{Fix}(T, P_U)$. ■

For future reference, we state the Fejér monotonicity of $\{x^k\}_{k\in\mathbb{N}}$ with respect to $\operatorname{Fix}(T, P_U)$ as a corollary.

Corollary 5.2.1. With the notation of Theorem 5.2.1,

$$\left\|x^{k+1} - y\right\|^2 \leq \left\|x^k - y\right\|^2 - \left\|D(x^k) - x^k\right\|^2$$

for all $y \in \operatorname{Fix}(T, P_U)$ and all $k \in \mathbb{N}$.

Proof. The result follows from (5.2.26). ■

5.3 Linear convergence of CRM applied to FPP under an error bound condition

In Chapter 3, when dealing with CFP with two convex sets, namely K, U, the following *global error bound*, which we will call **EB**, was considered:

EB: There exists $\bar{\omega} > 0$ such that $\operatorname{dist}(x, K) \geq \bar{\omega} \operatorname{dist}(K \cap U)$ for all $x \in U$.

In Section 3.1, it was proved that under **EB**, MAP converges linearly, with asymptotic constant bounded above by $\sqrt{1-\bar{\omega}^2}$, and that CRM also converges linearly, with a better upper bound for the asymptotic constant, namely $\sqrt{(1-\bar{\omega}^2)/(1+\bar{\omega}^2)}$. In this section we will prove that in the FPP case both sequences converge linearly, with asymptotic constant bounded by $\sqrt{1-\bar{\omega}^2}$.
In the case of FPP, dealing with a firmly nonexpansive $T : \mathbb{R}^n \to \mathbb{R}^n$, and an affine manifold $U \subset \mathbb{R}^n$, the appropriate error bound turns out to be:

EB1: There exists $\omega > 0$ such that $\|x - T(x)\| \geq \omega \operatorname{dist}(x, \operatorname{Fix}(T, P_U))$ for all $x \in U$.

We mention here that it suffices to consider an error bound less demanding than **EB1**, namely a local one, where the inequality above is requested to hold only for points in $U \cap V$, where V is a given set, e.g., a ball around the limit of the sequence generated by the algorithm, assumed to be convergent. An error bound of this type was used in Chapter 4. We refrain to do so just for the sake of a simpler exposition.

Proposition 5.3.1. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a firmly nonexpansive operator, $U \subset \mathbb{R}^n$ an affine manifold and $C, D : \mathbb{R}^n \to \mathbb{R}^n$ the CRM and the MAP operators respectively. Assume that $\operatorname{Fix}(T, P_U) \neq \varnothing$ and that **EB1** holds. Then

$$\operatorname{dist}(C(x), \operatorname{Fix}(T, P_U))^2 \leq \operatorname{dist}(D(x), \operatorname{Fix}(T, P_U))^2 \leq (1-\omega^2)\operatorname{dist}(x, \operatorname{Fix}(T, P_U))^2, \quad (5.3.1)$$

for all $x \in U$, with ω as in **EB1**.

Proof. First note that if $x \in F(T)$, then (5.3.1) holds trivially, so that we assume from now on that $T(x) \neq x$. Take any $y \in \operatorname{Fix}(T, P_U)$. Since T is firmly nonexpansive and $y \in F(T)$, we have

$$\begin{aligned} \|x - y\|^2 &\geq \|T(x) - T(y)\|^2 + \|(x - y) - (T(x) - T(y))\|^2 \\ &= \|T(x) - y\|^2 + \|x - T(x)\|^2, \end{aligned} \quad (5.3.2)$$

By assumption, $\operatorname{Fix}(T, P_U)$ is nonempty, and by Corollary 2.5.1, it is closed and convex. We take now a specific point in $\operatorname{Fix}(T, P_U)$, namely $\bar{y} = P_{\operatorname{Fix}(T,P_U)}(x)$, and rewrite **EB1** as

$$\|x - T(x)\|^2 \geq \omega^2 \|x - \bar{y}\|^2. \quad (5.3.3)$$

Combining (5.3.2) and (5.3.3), we get

$$\begin{aligned} \|x - \bar{y}\|^2 &\geq \|x - T(x)\|^2 + \|T(x) - \bar{y}\|^2 \\ &\geq \omega^2 \|x - \bar{y}\|^2 + \|T(x) - \bar{y}\|^2. \end{aligned} \quad (5.3.4)$$

Rearranging (5.3.4), we conclude that

$$\begin{aligned}(1-\omega^2)\left\|x-\bar{y}\right\|^2 &\geq \left\|T(x)-\bar{y}\right\|^2 \\ &= \left\|T(x)-D(x)\right\|^2+\left\|D(x)-\bar{y}\right\|^2+2\langle T(x)-D(x), D(x)-\bar{y}\rangle \end{aligned} \quad (5.3.5)$$

Note that $\langle T(x)-D(x),\bar{y}-D(x)\rangle = \langle T(x)-P_U(T(x)),\bar{y}-P_U(T(x))\rangle \leq 0$, by Proposition 2.2.1, since $\bar{y}\in U$. Hence,

$$\left\|T(x)-\bar{y}\right\|^2 \geq \left\|T(x)-D(x)\right\|^2+\left\|D(x)-\bar{y}\right\|^2. \quad (5.3.6)$$

Let $\hat{y} = P_{\mathrm{Fix}(T,P_U)}(D(x))$. From (5.3.5) and (5.3.6) we obtain

$$\begin{aligned}(1-\omega^2)\left\|x-\bar{y}\right\|^2 &\geq \left\|T(x)-\bar{y}\right\|^2 \\ &\geq \left\|T(x)-D(x)\right\|^2+\left\|D(x)-\bar{y}\right\|^2 \\ &\geq \left\|D(x)-\bar{y}\right\|^2 \geq \left\|D(x)-\hat{y}\right\|^2,\end{aligned} \quad (5.3.7)$$

where the second inequality holds by (5.3.6) and the last one follows from the definition of orthogonal projection. From (5.3.7) we conclude, recalling the definitions of $\bar{y},\hat{y}$, that

$$\mathrm{dist}(D(x),\mathrm{Fix}(T,P_U))^2 \leq (1-\omega^2)\mathrm{dist}(x,\mathrm{Fix}(T,P_U))^2, \quad (5.3.8)$$

which shows that the second inequality in (5.3.1) holds. Next we look at the first one. Let $\tilde{y} = P_{\mathrm{Fix}(T,P_U)}(C(x))$. We have that

$$\begin{aligned}\left\|C(x)-\tilde{y}\right\|^2 &\leq \left\|C(x)-\hat{y}\right\|^2 \leq \left\|D(x)-\hat{y}\right\|^2 \\ &\leq \left\|D(x)-\bar{y}\right\|^2 \leq (1-\omega^2)\left\|x-\bar{y}\right\|^2,\end{aligned} \quad (5.3.9)$$

where the first and the third inequality hold by the definition of orthogonal projection, the second one follows from Proposition 5.2.6(i) and the last one holds by (5.3.7). Note that the first inequality in (5.3.1) follows immediately from (5.3.9), in view of the definitions of $\tilde{y},\bar{y}$. ■

Corollary 5.3.1. Under the assumptions of Proposition 5.3.1, let $\{z^k\}_{k\in\mathbb{N}}, \{x^k\}_{k\in\mathbb{N}}$ be the sequences generated by MAP and CRM respectively, for solving FPP(T,P_U), *i.e.*, $z^{k+1} = D(z^k)$, and $x^{k+1} = C(x^k)$, starting from some $z^0\in\mathbb{R}^n$ and $x^0\in U$. Then the scalar sequences $\{a^k\},\{b^k\}$, defined as $a^k = \mathrm{dist}(z^k,\mathrm{Fix}(T,P_U))$ and $b^k = \mathrm{dist}(x^k,\mathrm{Fix}(T,P_U))$, converge Q-linearly to zero with asymptotic constants bounded above by $\sqrt{1-\omega^2}$, with ω as in **EB1**.

Proof. It follows from (5.3.1) that, for all $x\in U$,

$$\mathrm{dist}(D(x),\mathrm{Fix}(T,P_U))^2 \leq (1-\omega^2)\mathrm{dist}(x,\mathrm{Fix}(T,P_U))^2, \quad (5.3.10)$$

and that, for all $z\in U$,

$$\mathrm{dist}(C(x),\mathrm{Fix}(T,P_U))^2 \leq (1-\omega^2)\mathrm{dist}(x,\mathrm{Fix}(T,P_U))^2, \quad (5.3.11)$$

In view of the definitions of $\{x^k\}_{k\in\mathbb{N}}, \{z^k\}_{k\in\mathbb{N}}$, and remembering that both sequences are contained in U, by Proposition 5.2.4 in the case of $\{x^k\}_{k\in\mathbb{N}}$ and by definition of D in the case of $\{z^k\}_{k\in\mathbb{N}}$, we get from (5.3.10), (5.3.11),

$$\frac{\operatorname{dist}(z^{k+1}, \operatorname{Fix}(T, P_U))}{\operatorname{dist}(z^k, \operatorname{Fix}(T, P_U))} \leq \sqrt{1-\omega^2}, \tag{5.3.12}$$

$$\frac{\operatorname{dist}(x^{k+1}, \operatorname{Fix}(T, P_U))}{\operatorname{dist}(x^k, \operatorname{Fix}(T, P_U))} \leq \sqrt{1-\omega^2}. \tag{5.3.13}$$

The result follows immediately from (5.3.12), (5.3.13). ■

We show next that the sequences $\{x^k\}_{k\in\mathbb{N}}$ and $\{z^k\}_{k\in\mathbb{N}}$ are R-linearly convergent under Assumption **EB1**, with asymptotic constants bounded by $\sqrt{1-\omega^2}$, where ω is the **EB1** parameter.

Theorem 5.3.1. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a firmly nonexpansive operator and $U \subset \mathbb{R}^n$ is an affine manifold. Assume that $\operatorname{Fix}(T, P_U) \neq \varnothing$ and that condition **EB1** Holds. Consider the sequences $\{z^k\}_{k\in\mathbb{N}}, \{x^k\}_{k\in\mathbb{N}}$ generated by MAP and CRM respectively, for solving $\operatorname{Fix}(T, P_U)$, *i.e.*, $x^{k+1} = D(x^k)$ and $z^{k+1} = C(z^k)$, starting from some $z^0 \in \mathbb{R}^n$ and some $x^0 \in U$. Then both sequences converge R-linearly to points in $\operatorname{Fix}(T, P_U)$, with asymptotic constants bounded above by $\sqrt{1-\omega^2}$, with ω as in assumption **EB1**.

Proof. It follows from Corollary 5.3.1 that the scalar sequences $\{a^k\}_{k\in\mathbb{N}}, \{b^k\}_{k\in\mathbb{N}}$, defined as $a^k = \operatorname{dist}(z^k, \operatorname{Fix}(T, P_U))$ and $b^k = \operatorname{dist}(x^k, \operatorname{Fix}(T, P_U))$ are Q-linearly convergent to 0 with asymptotic constant bounded above by $\sqrt{1-\omega^2} < 1$, and hence R-linearly convergent to 0, with the same asymptotic constant. By Corollary 5.2.1, the sequence $\{x^k\}_{k\in\mathbb{N}}$ is Fejér monotone with respect to $\operatorname{Fix}(T, P_U)$, and the same holds for the sequence $\{z^k\}$, in view of (5.2.7). By Theorem 3.1.1, both sequences converge to points in $\operatorname{Fix}(T, P_U)$. Finally, by Lemma 3.1.1, both sequences converge R-linearly convergent to their limit points in the intersection, with asymptotic constants bounded by $\sqrt{1-\omega^2}$. ■

We mention that in Chapter 3, we showed that for CFP under EB, CRM achieves an asymptotic constant of linear convergence better than MAP. We have not been able to prove such superiority in the case of FPP. However, the numerical results exhibited in Section 5.4 strongly suggest that the asymptotic constant of CRM is indeed better than the MAP one. The task of establishing such theoretical superiority is left as an open problem.

5.4 Numerical comparisons between CRM and PPM for solving FPP

We report here numerical comparisons between CRM and PPM for solving FPP with p firmly nonexpansive operators.

All operators in this section belong to the family studied in Section 5.1, *i.e.,* they are convex combinations of orthogonal projections onto a finite number of closed convex sets with nonempty intersection. In view of Proposition 5.1.2(ii), these operators are ensured to have fixed-points. Hence, in view of Proposition 5.1.3 they are not orthogonal projections themselves.

The construction of the problems is as follows: for each instance, we choose randomly a number $r \in \{3,4,5\}$ (r is the number of convex sets in the convex combination). Then we sample values $\lambda_1,\dots,\lambda_r \in (0,1)$ with uniform distribution. We define $\mu_i = \lambda_i/(\sum_{\ell=1}^r \lambda_l)$. and we take the firmly nonexpansive operator T as $T = \sum_{i=1}^r \mu_i P_{\mathcal{E}_i}$, where $\mathcal{E}_i$ is an ellipsoid and $P_{\mathcal{E}_i}$ is the orthogonal projection onto it. Each ellipsoid $\mathcal{E}_i$ is of the form $\mathcal{E}_i := \{x \in \mathbb{R}^n : g_i(x) \le 0\}$, where $g_i : \mathbb{R}^n \to \mathbb{R}$ is given as $g_i(x) = x^t A_i x + 2(b^i)^t x - \alpha_i$, with $A_i \in \mathbb{R}^{n\times n}$ symmetric positive definite, $b^i \in \mathbb{R}^n$ and $0 < \alpha_i \in \mathbb{R}$.

Each matrix A_i is of the form $A_i = \gamma I + B_i^t B_i$, with $B_i \in \mathbb{R}^{n\times n}$, $\gamma \in \mathbb{R}_{++}$, where I stands for the identity matrix. The matrix B_i is a sparse matrix sampled from the standard normal distribution with sparsity density $p = 2n^{-1}$ and each vector b^i is sampled from the uniform distribution between $[0,1]$. We then choose each α_i so that $\alpha_i > (b^i)^t A b^i$, which ensures that 0 belongs to every $\mathcal{E}_i$, so that the intersection of the ellipsoids is nonempty. As explained above, this ensures that each instance of FPP has solutions.

In order to compute the projection onto the ellipsoids, we use a version of the Alternating Direction Method of Multipliers (ADMM) suited for this purpose, see [44]. A short explanation of how ADMM works is given in subsection 4.5. The stopping criterion for ADMM is as follows: we stop the ADMM iterative process when the norm of the difference between 2 consecutive ADMM iterates is less than 10^{-8}. We also fix a maximum number of 10 000 ADMM iterations.

For CRM, we use Pierra's product space reformulation, as explained in Section 1.4. We implement PPM directly from its definition (see Subsection 1.1.2). The stopping criterion for both CRM and PPM is similar to the one for the ADMM subroutine, but with a different tolerance: the iterative process stops when the norm of the difference between 2 consecutive CRM or PPM iterates is less than 10^{-6}. The maximum number of iterations is fixed at 50 000 for both algorithms.

The experiments consist of solving, with CRM and PPM, 250 instances of FPP selected as follows. We consider the following values for the dimension n: $\{10, 30, 50, 100, 200\}$, and for each n we take p firmly nonexpansive operators with $p \in \{10, 25, 50, 100, 200\}$. For each of these 25 pairs (n,p), we randomly generate 10 instances of FPP with the above explained procedure.

The initial point x^0 is of the form $(\eta,\dots,\eta) \in \mathbb{R}^n$, with $\eta < 0$ and $|\eta|$ sufficiently large so as to guarantee that x^0 is far from all the ellipsoids.

The computational experiments were carried out on an Intel Xeon W-2133 3.60GHz with 32GB of RAM running Ubuntu 20.04. We implemented all experiments in Julia programming language v1.6 (see [22]). The codes of our experiments are fully available at: https://github.com/Mirza-Reza/FPP.

We report in Table 5.1 the following descriptive statistics for CRM and PPM: mean, maximum (max), minimum (min) and standard deviation (std) for iteration count (it) and CPU time in seconds (CPU (s)). In particular, the ratio of the CPU time (in average for all instances) of PPM with respect to CRM is 7.69, meaning that CRM is, on the average, almost eight times faster that PPM.

We report next similar statistics, but separately for each dimension n.

Looking at Table 5.2, we observe that the CPU time for PPM grows linearly with the

Table 5.1: Statistics for all instances, reporting number of iterations and CPU time

Method		mean	max	min	std
CRM	`it`	144.2880	554	23	95.2581
	`CPU(s)`	14.6048	120.3020	0.2729	22.4890
PPM	`it`	5977.3520	25000	209	6385.9388
	`CPU(s)`	112.3315	1085.9685	1.2483	190.3078

dimension n, while the growth of the CRM CPU time is somewhat higher than linear. As a consequence, the superiority of CRM over PPM, measured in terms of the quotient between the PPM CPU time and the CRM CPU time, is slightly decreasing with n: it goes from a ratio of 9.17 for $n = 10$ to a ratio of 7.56 for $n = 200$. This said, it is clear that CRM vastly outperforms PPM in terms of CPU time for all the values of n tested in our experiments.

Table 5.2: Statistics for instances of each dimension n, reporting number of iterations and CPU time

Method		mean	max	min	std
CRM	`it`	141.8400	512	28	99.1028
n=10	`CPU(s)`	2.3247	6.8150	0.2729	1.9756
PPM	`it`	6024.5400	19163	209	6425.5744
n=10	`CPU(s)`	21.3369	92.19569	1.2483	22.4132
CRM	`it`	153.5000	526	46	92.2150
n=30	`CPU(s)`	4.6989	16.6523	0.7607	4.1754
PPM	`it`	5608.4400	18353	500	5956.7588
n=30	`CPU(s)`	42.9296	174.9737	2.9861	46.2969
CRM	`it`	129.5000	469	23	91.7052
n=50	`CPU(s)`	6.8152	17.1668	1.0480	5.0391
PPM	`it`	5288.5200	24680	423	5548.5052
n=50	`CPU(s)`	53.3709	222.7307	3.5744	55.7054
CRM	`it`	152.0400	399	28	84.1924
n=100	`CPU(s)`	15.5937	41.2581	1.9661	12.4246
PPM	`it`	7224.4200	21978	540	7663.8605
n=100	`CPU(s)`	114.4037	428.8247	6.3108	108.4765
CRM	`it`	144.5600	554	42	105.7244
n=200	`CPU(s)`	43.5915	120.3019	5.0157	34.3053
PPM	`it`	5740.8400	22378	370	5948.5707
n=200	`CPU(s)`	329.6167	1085.9685	19.0842	315.8783

Next, we report in the next table similar statistics, but separately for problems involving p firmly nonexpansive operators, for each value of p.
Table 5.3 indicates that both the CRM and the PPM CPU time grow slightly less than linearly in p, the number of firmly nonexpansive operators in each instance of FPP, but the growth in both cases seems to become linear for $p \geq 50$. Consistently with this behavior, the ratio between the PPM CPU time and the CRM CPU time is about 3 for $p = 10, 25$ and about 8, for $p = 50, 100, 200$. Again, for all values of p, CRM turns out to be highly better than PPM in terms of CPU time.

Table 5.3: Statistics for instances of FPP problems with p firmly nonexpansive operators, reporting number of iterations and CPU time

Method		**mean**	**max**	**min**	**std**
fneCRM	`it`	91.0000	263	28	50.1745
p=10	`CPU(s)`	2.8569	13.1619	0.2729	2.8807
PPM	`it`	1316.6800	6765	209	1264.5453
p=10	`CPU(s)`	13.1578	50.8767	1.2483	11.4271
CRM	`it`	113.7000	469	36	83.5955
p=25	`CPU(s)`	6.5062	45.2021	0.6664	8.9416
PPM	`it`	2865.9200	14617	650	2651.0789
p=25	`CPU(s)`	34.6541	242.2805	2.9785	47.5093
CRM	`it`	128.8000	331	23	76.9244
p=50	`CPU(s)`	10.438 80	46.8045	1.3100	11.8677
PPM	`it`	4949.4200	25000	870	5531.4013
p=50	`CPU(s)`	88.4859	602.8599	6.5347	125.0821
CRM	`it`	166.2800	526	49	91.1888
p=100	`CPU(s)`	18.8065	70.6532	2.4265	20.1719
PPM	`it`	7077.4600	25000	1586	4970.7775
p=100	`CPU(s)`	143.0699	729.1966	12.0125	171.2874
CRM	`it`	221.6600	554	88	105.5789
p=200	`CPU(s)`	34.4157	120.3019	4.7277	35.5202
PPM	`it`	13 677.28	25000	4015	6856.3914
p=200	`CPU(s)`	282.2900	1085.9685	31.8832	295.7094

Finally, we exhibit the performance profile, in the sense of [33], for all the instances. Again the superiority of CRM with respect to PPM is fully corroborated.

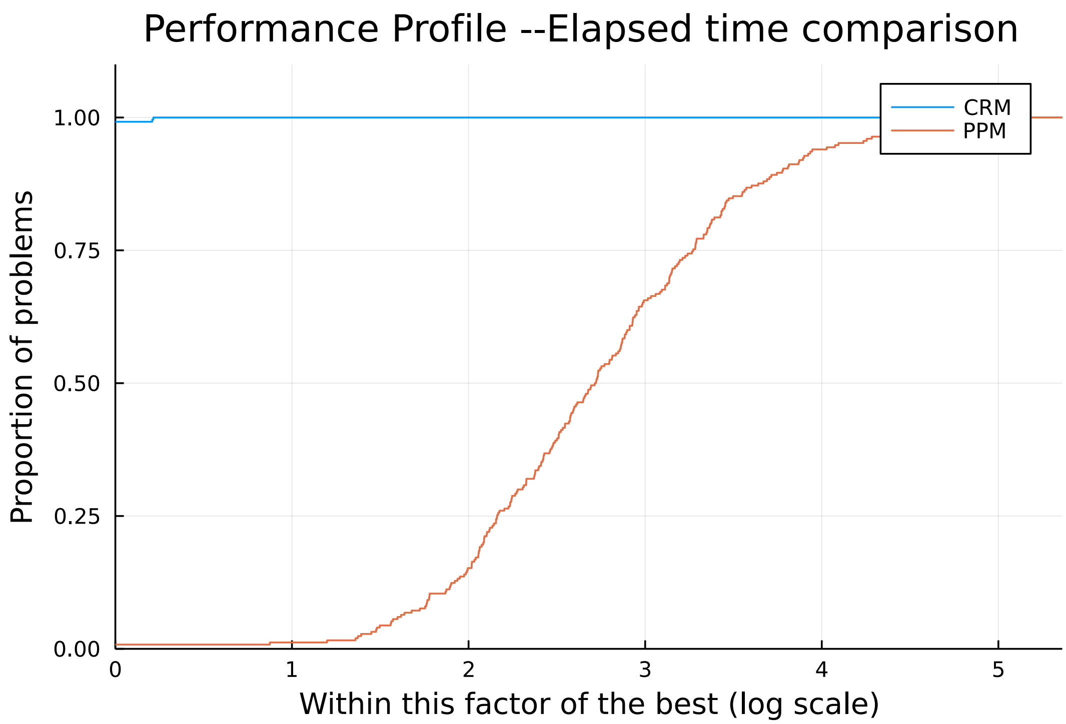

Figure 5.1: Performance profile of experiments with ellipsoidal feasibility – CRM vs PPM

Bibliography

[1] Agmon, S. The relaxation method for linear inequalities. *Canadian Journal of Mathematics 6* (1954), 382–392.

[2] Aragón Artacho, F. J., Campoy, R., and Tam, M. K. The Douglas-Rachford algorithm for convex and nonconvex feasibility problems. *Mathematical Methods of Operations Research 91*, 2 (2020), 201–240.

[3] Araújo, G., Arefidamghani, R., Behling, R., Bello-Cruz, Y., Iusem, A., and Santos, L.-R. Circumcentering approximate reflections for solving the convex feasibility problem. *Fixed Point Theory and Algorithms for Sciences and Engineering*, 1 (2022), 1–30.

[4] Arefidamghani, R., Behling, R., Bello-Cruz, Y., Iusem, A. N., and Santos, L.-R. The circumcentered-reflection method achieves better rates than alternating projections. *Computational Optimization and Applications 79*, 2 (2021), 507–530.

[5] Auslender, A. *Optimisation, Méthodes numériques.* Masson, 1976.

[6] Bauschke, H. H., and Borwein, J. M. On the convergence of von Neumann's alternating projection algorithm for two sets. *Set-Valued Analysis 1*, 2 (1993), 185–212.

[7] Bauschke, H. H., and Borwein, J. M. On projection algorithms for solving convex feasibility problems. *SIAM Review 38*, 3 (1996), 367–426.

[8] Bauschke, H. H., and Combettes, P. L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, vol. 408. Springer, 2011.

[9] Bauschke, H. H., Dao, M. N., Noll, D., and Phan, H. M. Proximal point algorithm, Douglas-Rachford algorithm and alternating projections: a case study. *arXiv preprint arXiv:1501.06603* (2015).

[10] Bauschke, H. H., Ouyang, H., and Wang, X. On circumcenters of finite sets in Hilbert spaces. *arXiv preprint arXiv:1807.02093* (2018).

[11] Bauschke, H. H., Ouyang, H., and Wang, X. Circumcentered methods induced by isometries. *Vietnam Journal of Mathematics 48*, 3 (2020), 471–508.

[12] Bauschke, H. H., Ouyang, H., and Wang, X. Best approximation mappings in Hilbert spaces. *Mathematical Programming* (2021), 1–47.

[13] Behling, R., Bello-Cruz, J. Y., and Santos, L.-R. On the linear convergence of the circumcentered-reflection method. *Operations Research Letters 46*, 2 (2018), 159–162.

[14] Behling, R., Bello-Cruz, J. Y., and Santos, L.-R. The block-wise circumcentered-reflection method. *Computational Optimization and Applications* (2019), 1–25.

[15] Behling, R., Bello-Cruz, Y., Iusem, A. N., and Santos, L.-R. On the centralization of the circumcentered-reflection method. *arXiv preprint arXiv:2111.07022* (2021).

[16] Behling, R., Bello-Cruz, Y., Lara-Urdaneta, H., Oviedo, H., and Santos, L.-R. Circumcentric directions of cones. *arXiv preprint arXiv:2112.08314* (2021).

[17] Behling, R., Bello-Cruz, Y., and Santos, L.-R. On the circumcentered-reflection method for the convex feasibility problem. *Numerical Algorithms 86*, 4 (2021), 1475–1494.

[18] Behling, R., Cruz, J. Y. B., and Santos, L.-R. Circumcentering the Douglas–Rachford method. *Numerical Algorithms 78*, 3 (2018), 759–776.

[19] Behling, R., Fischer, A., Haeser, G., Ramos, A., and Schönefeld, K. On the constrained error bound condition and the projected Levenberg–Marquardt method. *Optimization 66*, 8 (2017), 1397–1411.

[20] Behling, R., and Iusem, A. N. The effect of calmness on the solution set of systems of nonlinear equations. *Mathematical Programming 137*, 1 (2013), 155–165.

[21] Benzi, M. Gianfranco Cimmino's contributions to numerical mathematics. *Atti del Seminario di Analisi Matematica, Dipartimento di Matematica dell'Universita di Bologna. Volume Speciale: Ciclo di Conferenze in Ricordo di Gianfranco Cimmino* (2004), 87–109.

[22] Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. Julia: A fresh approach to numerical computing. *SIAM Review 59*, 1 (2017), 65–98.

[23] Bregman, L. M. The method of successive projection for finding a common point of convex sets. *Sov. Math. Dok. 162*, 3 (1965), 688–692.

[24] Cegielski, A. Generalized relaxations of nonexpansive operators and convex feasibility problems. *Contemporary Mathematics 513* (2010), 111–123.

[25] Cegielski, A. *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, vol. 2057. Springer, 2012.

[26] Censor, Y., and Cegielski, A. Projection methods: an annotated bibliography of books and reviews. *Optimization 64*, 11 (2015), 2343–2358.

[27] Censor, Y., and Segal, A. The split common fixed point problem for directed operators. *Journal of Convex Analysis 16*, 2 (2009), 587–600.

[28] Cheney, W., and Goldstein, A. A. Proximity maps for convex sets. *Proceedings of the American Mathematical Society 10*, 3 (1959), 448–450.

[29] Cimmino, G. Calcolo spprossimato per le soluzioni del sistemi di equazioni lineari. *La Ricerca Scientifica* (Roma) 1 (1938), 326–333.

[30] Dai, Y. H. Fast algorithms for projection on an ellipsoid. *SIAM Journal on Optimization 16*, 4 (2006), 986–1006.

[31] de Pierro, A. R., and Iusem, A. N. A parallel projection method of finding a common point of a family of convex sets. *Pesquisa Operacional 5*, 1 (1985), 1–20.

[32] Dizon, N., Hogan, J., and Lindstrom, S. B. Circumcentering reflection methods for nonconvex feasibility problems. *arXiv preprint arXiv:1910.04384* (2019).

[33] Dolan, E. D., and Moré, J. J. Benchmarking optimization software with performance profiles. *Mathematical Programming 91*, 2 (2002), 201–213.

[34] 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.

[35] Drusvyatskiy, D., Ioffe, A. D., and Lewis, A. S. Transversality and alternating projections for nonconvex sets. *Foundations of Computational Mathematics 15*, 6 (2015), 1637–1651.

[36] Eremin, I. I. Generalization of the relaxation method of Motzkin-Agmon. *Uspekhi Matematicheskikh Nauk 20*, 2 (1965), 183–187.

[37] Euklides, and Heath, T. L. *The Thirteen Books of Euclid's Elements. Vol. 2, Books III-IX*. Dover, 1956.

[38] Fukushima, M. An outer approximation algorithm for solving general convex programs. *Operations Research 31*, 1 (1983), 101–113.

[39] Glowinski, R., and Marrocco, A. Approximation par éléments finis d'ordre un et résolution par pénalisation dualité d'une classe de problèmes non linéaires. *RAIRO, Recherche Opérationelle 9,2* (1975), 41–76.

[40] Gubin, L., Polyak, B. T., and Raik, E. The method of projections for finding the common point of convex sets. *USSR Computational Mathematics and Mathematical Physics 7*, 6 (1967), 1–24.

[41] Hiriart-Urruty, J. B., and Lemaréchal, C. *Convex Analysis and Minimization Algorithms I: Fundamentals*, vol. 305. Springer, 2013.

[42] Hoffman, A. J. On approximate solutions of systems of linear inequalities. In *Selected Papers Of Alan J Hoffman: With Commentary*. World Scientific, 2003, pp. 174–176.

[43] Iusem, A. N., and de Pierro, A. R. On the set of weighted least squares solutions of systems of convex inequalities. *Commentationes Mathematicae Universitatis Carolinae 25*, 4 (1984), 667–678.

[44] Jia, Z., Cai, X., and Han, D. Comparison of several fast algorithms for projection onto an ellipsoid. *Journal of Computational and Applied Mathematics 319* (2017), 320–337.

[45] Kaczmarz, S. Angenäherte Auflösung von Systemen Linearer Gleichungen. *Bulletin de L'Académie Polonaise des Sciences et Lettres A-35* (1937), 355–357.

[46] Kaczmarz, S. Approximate solution of systems of linear equations. *International Journal of Control 57*, 6 (1993), 1269–1271.

[47] Kanzow, C., Yamashita, N., and Fukushima, M. Withdrawn: Levenberg–Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints. *Journal of Computational and Applied Mathematics 173*, 2 (2005), 321–343.

[48] Kayalar, S., and Weinert, H. L. Error bounds for the method of alternating projections. *Mathematics of Control, Signals and Systems 1*, 1 (1988), 43–59.

[49] Kruger, A. Y. About intrinsic transversality of pairs of sets. *Set-Valued and Variational Analysis 26*, 1 (2018), 111–142.

[50] Lin, A., and Han, S.-P. A class of methods for projection on the intersection of several ellipsoids. *SIAM Journal on Optimization 15*, 1 (2004), 129–138.

[51] Lindstrom, S. B., and Sims, B. Survey: sixty years of Douglas–Rachford. *Journal of the Australian Mathematical Society 110*, 3 (2021), 333–370.

[52] 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.

[53] Maligranda, L. Stefan Kaczmarz (1895-1939). *Antiquitates Mathematicae 1*, 1 (2014), 15–62.

[54] Motzkin, T. S., and Schoenberg, I. J. The relaxation method for linear inequalities. *Canadian Journal of Mathematics 6* (1954), 393–404.

[55] Moudafi, A. A note on the split common fixed-point problem for quasi-nonexpansive operators. *Nonlinear Analysis: Theory, Methods* & Applications 74, 12 (2011), 4083–4087.

[56] Parks, P. Stefan Kaczmarz (1895–1939). *International Journal of Control 57*, 6 (1993), 1263–1267.

[57] Pierra, G. Decomposition through formalization in a product space. *Mathematical Programming 28*, 1 (1984), 96–115.

[58] Robinson, S. M. *Generalized Equations and Their Solutions, Part II: Applications to Nonlinear Programming.* Springer, 1982.

[59] Rockafellar, R. T., and Wets, R. J.-B. *Variational Analysis*, vol. 317. Springer, 2009.

[60] Von Neumann, J. *Functional Operators: The Geometry of Orthogonal Spaces*, vol. 2. Princeton University Press, 1951.

[61] Yao, Y., Liou, Y.-C., and Yao, J.-C. Split common fixed point problem for two quasi-pseudo-contractive operators and its algorithm construction. *Fixed Point Theory and Applications 2015*, 1 (2015), 1–19.

[62] Zhao, J., and He, S. Alternating Mann iterative algorithms for the split common fixed-point problem of quasi-nonexpansive mappings. *Fixed Point Theory and Applications 2013*, 1 (2013), 1–12.