The Method of Alternating Relaxed Projections for two nonconvex sets

Heinz H. Bauschke[*], Hung M. Phan[†], and Xianfu Wang[‡]

September 24, 2013

Dedicated to Boris Mordukhovich on the occasion of his 65th Birthday

Abstract

The Method of Alternating Projections (MAP), a classical algorithm for solving feasibility problems, has recently been intensely studied for nonconvex sets. However, intrinsically available are only local convergence results: convergence occurs if the starting point is not too far away from solutions to avoid getting trapped in certain regions. Instead of taking full projection steps, it can be advantageous to underrelax, i.e., to move only part way towards the constraint set, in order to enlarge the regions of convergence.

In this paper, we thus systematically study the Method of Alternating Relaxed Projections (MARP) for two (possibly nonconvex) sets. Complementing our recent work on MAP, we establish local linear convergence results for the MARP. Several examples illustrate our analysis.

2010 Mathematics Subject Classification: Primary 65K05; Secondary 47N10, 49J52, 49M20, 65K10, 90C25, 90C26.

Keywords: Feasibility problem, linear convergence, method of alternating projections, method of alternating relaxed projections, normal cone, projection operator.

1 Introduction

We assume throughout this paper that

$$X \text{ is a Euclidean space} \tag{1}$$

[*]Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: heinz.bauschke@ubc.ca.

[†]Mathematics and Statistics, University of Victoria, PO Box 3060 STN CSC, Victoria, B.C. V8W 3R4, Canada. E-mail: hphan@uvic.ca.

[‡]Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: shawn.wang@ubc.ca.

with inner product $\langle\cdot,\cdot\rangle$ and associated norm $\|\cdot\|$ and that

(2)	A and B are nonempty closed subsets of X.

Our aim is to solve the feasibility problem

$$\text{find } x \in A \cap B. \tag{3}$$

(We do not *a priori* assume that $A \cap B \neq \varnothing$.) We assume that it is possible to evaluate the *projection operators* (nearest point mappings) P_A and P_B associated with the constraints sets A and B respectively. The operators P_A and P_B are generally set-valued; they are single-valued only in the convex case. The celebrated *Method of Alternating Projections (MAP)*, whose origins can be traced back to von Neumann [28] and Wiener [30], with starting point $b_{-1} \in X$ generates sequences according to the update rule[1]

$$(\forall n \in \mathbb{N}) \quad a_n \in P_A b_{n-1} \quad \text{and} \quad b_n \in P_B a_n. \tag{4}$$

If A and B are convex, then this method is well understood; see, e.g., [2, 5, 7, 11, 12, 13, 16, 17, 18, 19] and the references therein for extensions and variants. The convergence theory for the MAP and related methods is much more delicate in the absence of convexity; see, e.g., [8, 9, 15, 23, 24] and the references therein.

Simple examples can be constructed to show that in general one cannot expect global convergence of the MAP when $A \cap B \neq \varnothing$:

Example 1.1 (unrelaxed MAP) Suppose that $X = \mathbb{R}$, that $A = \{-3, 2\}$ and that $B = \{-3, 6\}$. Then $A \cap B = \{-3\} \neq \varnothing$. Now set $b_{-1} := 0$. Then $a_0 := P_A b_{-1} = P_A 0 = 2$ (since $|2-0| = 2 < 3 = |-3-0|$) and $b_0 := P_B a_0 = P_B 2 = 6$ (since $|6-2| = 4 < 5 = |-3-2|$) and clearly $a_1 := P_A b_0 = 2$. It follows that

$$(\forall n \in \mathbb{N}) \; a_n = 2 \text{ and } b_n = 6. \tag{5}$$

Thus, the sequences generated by the MAP do not converge to a point in $A \cap B$ (see Figure 1).

To improve this situation, we study in this paper the *Method of Alternating Relaxed Projections*, where the unrelaxed projection steps are replaced by underrelaxed versions; e.g., the projection operators P_A and P_B may be replaced by $(1-\lambda)\operatorname{Id} + \lambda P_A$ and $(1-\mu)\operatorname{Id} + \mu P_B$, where λ and μ belong to $]0,1]$. In the convex case, there are several pertinent references including [2, 3, 11, 14, 20, 21, 22, 29].

The idea of *regularizing* operators is of course not new; MARP can be seen as regularizing the straight projection operators. To demonstrate the potential of this approach, let us revisit Example 1.1:

Example 1.2 (MARP for Example 1.1) Let X, A, B, and b_{-1} be as in Example 1.1. Rather than iterating P_A and P_B, we now iterate $\frac{1}{2}\operatorname{Id} + \frac{1}{2}P_A$ and $\frac{1}{2}\operatorname{Id} + \frac{1}{2}P_B$. Then $a_0 = (\frac{1}{2}\operatorname{Id} + \frac{1}{2}P_A)b_{-1} = \frac{1}{2}b_{-1} + \frac{1}{2}P_A b_{-1} = \frac{1}{2}0 + \frac{1}{2}2 = 1$, $b_0 = (\frac{1}{2}\operatorname{Id} + \frac{1}{2}P_B)a_0 = \frac{1}{2}1 + \frac{1}{2}(-3) = -2$, $a_1 = (\frac{1}{2}\operatorname{Id} + \frac{1}{2}P_A)b_0 = \frac{1}{2}(-2) + \frac{1}{2}(-3) = -5/2$, $b_1 = \cdots = -11/4$, $a_2 = -23/8$, $b_2 = -47/16, \ldots$ (see Figure 1), and the sequences generated converge[2] to -3, the unique point in $A \cap B$, as desired.

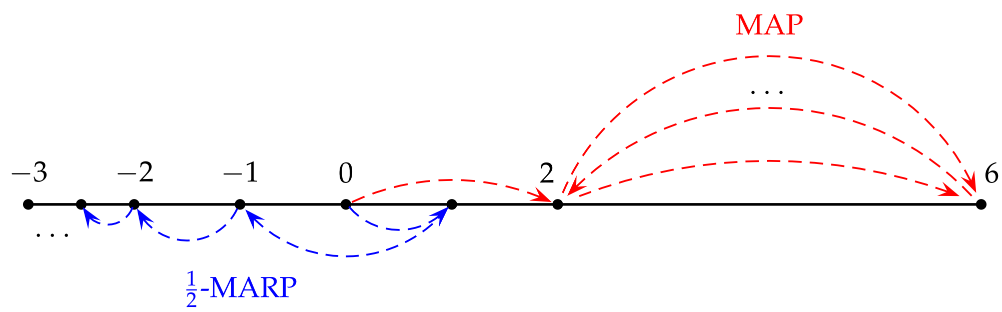

Figure 1: MAP vs MARP

The goal of this paper is to systematically study the MARP and to provide sufficient conditions for convergence.

The tools used are from variational analysis; we extend techniques recently introduced in [8, 9].

Our main results are the following:

- Theorem 4.3 is a powerful abstract linear convergence result that is applicable in particular to the MARP;
- Theorem 5.11 provides a local linear convergence result for the MARP in the presence of a CQ condition;
- Theorem 6.4 guarantees local linear convergence of the MARP under some regularity assumptions.

The paper is organized as follows: After reviewing auxiliary notions in Section 2, we introduce the MARP in Section 3 and obtain some basic properties. Abstract linear convergence results are presented in Section 4. Local linear convergence results based on CQ conditions and on regularity are provided in Sections 5 and 6, respectively. In Section 7, we discuss linearly vanishing relaxation parameters. Various examples illustrating the general theory are constructed in Sections 8 and 9.

We conclude this section with some notational comments. We write $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$, $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$, and $\mathbb{N} = \mathbb{Z} \cap \mathbb{R}_+$. The *distance function* is $d_A \colon x \mapsto \inf_{a \in A} \|x - a\|$ and the (generally set-valued) *projection operator* is $P_A \colon x \mapsto \{a \in A \mid \|x - a\| = d_A(x)\}$. Given a subset S of X, we write $\operatorname{int} S$, $\operatorname{ri} S$, and $\overline{S}$ for the interior, the relative interior and the closure of S, respectively. If u and v are points in X, we write $[u, v] = \{(1-\lambda)u + \lambda v \mid 0 \leq \lambda \leq 1\}$, $]u, v] = \{(1-\lambda)u + \lambda v \mid 0 < \lambda \leq 1\}$, and similarly for $[u, v[$ and $]u, v[$. We also set $\operatorname{ball}(u; r) = \{x \in X \mid \|x - u\| \leq r\}$. For notation not explicitly stated in this paper, and background material in convex and variational analysis, we refer the reader to [7, 10, 25, 26, 27, 31].

2 Auxiliary Notions

In this section, we collect several technical definitions for future use. For further results and comments, see [8, 9] and the references therein. We start with the restricted normal cone, which is not

[1] We follow a common but convenient abuse of notation and write $a_n = P_A b_{n-1}$ etc. if the set of nearest points is a singleton.

[2] This will follow from Example 8.2 below.

only central to our analysis but also a generalized version of the Mordukhovich normal cone, an object known to be of critical importance in modern variational analysis.

Definition 2.1 (restricted normal cones) (See [8, Definition 2.1].) *Let $a \in A$.*

(i) *The B*-restricted proximal normal cone *of A at a is*

$$\widehat{N}_A^B(a) := \operatorname{cone}\Big(\big(B \cap P_A^{-1}a\big) - a\Big) = \operatorname{cone}\Big(\big(B - a\big) \cap \big(P_A^{-1}a - a\big)\Big). \tag{6}$$

(ii) *The B*-restricted normal cone *$N_A^B(a)$ is implicitly defined by $u \in N_A^B(a)$ if and only if there exist sequences $(a_n)_{n\in\mathbb{N}}$ in A and $(u_n)_{n\in\mathbb{N}}$ in $\widehat{N}_A^B(a_n)$ such that $a_n \to a$ and $u_n \to u$.*

Definition 2.2 (regularity of sets) (See [8, Definition 8.1].) *Let $c \in B$, $\varepsilon \geq 0$, and $\delta > 0$. Then B is* (A, ε, δ)-regular at c *if*

$$\left.\begin{array}{c} (y, b) \in B \times B, \\ \|y - c\| \leq \delta, \|b - c\| \leq \delta, \\ u \in \widehat{N}_B^A(b) \end{array}\right\} \quad \Rightarrow \quad \langle u, y - b\rangle \leq \varepsilon \|u\| \cdot \|y - b\|. \tag{7}$$

The set B is called A-superregular *at $c \in B$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that B is (A, ε, δ)-regular at c. When $A = X$, we say "B* is (ε, δ)-regular*" or "B* is superregular*", i.e., the prefix "X-" is omitted.*

Definition 2.3 (linear convergence) *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X, let $c \in X$, let $\alpha \in\,]0, 1[$. Then* $(x_n)_{n\in\mathbb{N}}$ converges to c linearly with rate α *if there exists $M \in \mathbb{R}_+$ such that*[3]

$$(\forall n \in \mathbb{N}) \quad \|x_n - c\| \leq M\alpha^n. \tag{8}$$

Definition 2.4 (CQ-number) (See [8, Definition 6.1].) *Let $\widetilde{A}$ and $\widetilde{B}$ be nonempty subsets of X, let $c \in X$, and let $\delta \in \mathbb{R}_{++}$. The* CQ-number *at c associated with $(A, \widetilde{A}, B, \widetilde{B})$ and δ is*

$$\theta_\delta := \theta_\delta\big(A, \widetilde{A}, B, \widetilde{B}\big) := \sup\left\{\langle u, v\rangle \;\middle|\; \begin{array}{l} u \in \widehat{N}_A^{\widetilde{B}}(a), v \in -\widehat{N}_B^{\widetilde{A}}(b), \|u\| \leq 1, \|v\| \leq 1, \\ \|a - c\| \leq \delta, \|b - c\| \leq \delta. \end{array}\right\}, \tag{9}$$

and the limiting CQ-number *at c associated with $(A, \widetilde{A}, B, \widetilde{B})$ is*

$$\overline{\theta} := \overline{\theta}\big(A, \widetilde{A}, B, \widetilde{B}\big) := \lim_{\delta \downarrow 0} \theta_\delta\big(A, \widetilde{A}, B, \widetilde{B}\big). \tag{10}$$

Definition 2.5 (CQ condition) (See [8, Definition 6.6].) *Let $\widetilde{A}$ and $\widetilde{B}$ be nonempty subsets of X, and let $c \in X$. Then the* $(A, \widetilde{A}, B, \widetilde{B})$-CQ condition *holds at c if*

$$N_A^{\widetilde{B}}(c) \cap \big(-N_B^{\widetilde{A}}(c)\big) \subseteq \{0\}. \tag{11}$$

We recall the following equivalence from [8, Theorem 6.8]:

$$N_A^{\widetilde{B}}(c) \cap \big(-N_B^{\widetilde{A}}(c)\big) \subseteq \{0\} \;\Leftrightarrow\; \overline{\theta}\big(A, \widetilde{A}, B, \widetilde{B}\big) < 1. \tag{12}$$

[3] Note that one may alternatively require that (8) only holds eventually at the expense of possibly enlarging M; see, e.g., [9, Remark 3.7].

3 MARP: Basic Properties

Definition 3.1 *Let $y \in X$ and let $\lambda \in \,]0,1]$. Then the vectors in the set*

$$(1-\lambda)y + \lambda P_A y = \big\{(1-\lambda)y + \lambda a \;\big|\; a \in P_A y\big\} \tag{13}$$

are called λ-relaxed projections of y on A.

Note that the 1-relaxed projections are precisely the original (unrelaxed) projections. From now on, we assume that

$$\boldsymbol{\lambda} = (\lambda_n)_{n\in\mathbb{N}} \text{ and } \boldsymbol{\mu} = (\mu_n)_{n\in\mathbb{N}} \text{ are sequences in }]0,1], \quad \text{and } \alpha_0 := \max\{\lambda_0, \mu_0\}. \tag{14}$$

Definition 3.2 (Method of Alternating Relaxed Projections (MARP)) *Let $y_{-1} \in X$ be the starting point. The method of alternating $(\boldsymbol{\lambda},\boldsymbol{\mu})$-relaxed projections between A and B (the $(\boldsymbol{\lambda},\boldsymbol{\mu})$-MARP or just MARP in short) generates sequences $\mathbf{x} := (x_n)_{n\in\mathbb{N}}$ and $\mathbf{y} := (y_n)_{n\in\mathbb{N}}$ as follows:*

$$\begin{aligned}(\forall n\in\mathbb{N})\quad y_{n-1} &\mapsto a_n \in P_A y_{n-1} \mapsto x_n := (1-\lambda_n)y_{n-1} + \lambda_n a_n \\ &\mapsto b_n \in P_B x_n \mapsto y_n := (1-\mu_n)x_n + \mu_n b_n \mapsto \cdots.\end{aligned} \tag{15}$$

We call $(\mathbf{x},\mathbf{y})$ also $(\boldsymbol{\lambda},\boldsymbol{\mu})$-MARP or simply MARP sequences.

If $(\forall n\in\mathbb{N})$ $\lambda_n = \mu_n = 1$, then $(x_n)_{n\in\mathbb{N}} = (a_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}} = (b_n)_{n\in\mathbb{N}}$, and the MARP reduces to the classic method of alternating projections (MAP).

Unless specified otherwise, we assume for the remainder of this paper that

$$\big(\mathbf{x} = (x_n)_{n\in\mathbb{N}}, \mathbf{y} = (y_n)_{n\in\mathbb{N}}\big) \text{ are } (\boldsymbol{\lambda},\boldsymbol{\mu})\text{-MARP sequences with starting point } y_{-1}. \tag{16}$$

The following simple result turns out to be quite useful.

Proposition 3.3 *Let $y \in X$, $a \in P_A y$, $\lambda \in \,]0,1]$, and set $x := (1-\lambda)y + \lambda a$. Then the following hold:*

(i) *$P_A(x) = a$.*

(ii) *$x - y = \lambda(a-y)$ and thus $\|x-y\| = \lambda\|a-y\| = \lambda d_A(y)$.*

(iii) *$\lambda(x-a) = (1-\lambda)(y-x)$.*

Proof. (i): Suppose that $a' \in A \smallsetminus \{a\}$. *Case 1: $x \in [y,a']$.* Then $\|y-a'\| > \|y-a\|$ because y, a, a' lie on the same ray. So

$$\|x-a'\| = \|y-a'\| - \|y-x\| > \|y-a\| - \|y-x\| = \|x-a\|. \tag{17}$$

Case 2: $x \notin [y,a']$. Then

$$\|x-a'\| > \|y-a'\| - \|y-x\| \geq \|y-a\| - \|y-x\| = \|x-a\|. \tag{18}$$

In either case, $\|x - a'\| > \|x - a\|$ and therefore $a = P_A(x)$.

(ii): Indeed, $x - y = \lambda(a - y) \Leftrightarrow x = (1-\lambda)y + \lambda a$.

(iii): We have: $\lambda(x-a) = (1-\lambda)(y-x) \Leftrightarrow \lambda(x-a) = (1-\lambda)y - (1-\lambda)x \Leftrightarrow -\lambda a = (1-\lambda)y - x \Leftrightarrow x = (1-\lambda)y + \lambda a$. ■

Definition 3.4 (projection absorbing) *Let S be a nonempty subset of X. Then S is A-*projection absorbing *(or* projection absorbing with respect to A*), if*

$$(\forall s \in S)(\forall a \in P_A s) \quad [s, a] \subseteq S. \tag{19}$$

Remark 3.5 Let S be a subset of X that is A-projection absorbing.

(i) Clearly, X is A-projection absorbing.

(ii) If S is B-projection absorbing, then S is also $A \cup B$-projection absorbing because $(\forall s \in S)$ $P_{A\cup B}(s) \subseteq P_A s \cup P_B s$. The opposite implication is not necessarily true: for example, if $X = \mathbb{R}^2$, $S = A = \mathbb{R} \times \{1\}$, and $B = \mathbb{R} \times \{0\}$, then S is $A \cup B$-projection absorbing but not B-projection absorbing.

(iii) If S is convex and $P_A(S) \subseteq S$, then S is A-projection absorbing.

(iv) On the other hand, if A is convex and $S = \overline{X \smallsetminus A}$, then S is (usually not convex but) still A-projection absorbing.

The notion of a projection absorbing set is important because of the following result pertaining to the orbit of the MARP.

Proposition 3.6 *Let S be a subset of X that is both A-projection absorbing and B-projection absorbing. If $y_{-1} \in S$, then $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ lie in S.*

Proof. This follows readily by using mathematical induction. ■

Lemma 3.7 *Set $\beta := \max\{d_A(y_{-1}), d_B(y_{-1})\}$. Then the following hold:*

$$\|x_0 - y_{-1}\| = \lambda_0 d_A(y_{-1}) \leq \lambda_0 \beta, \tag{20a}$$

$$\max\{d_A(x_0), d_B(x_0)\} \leq \|x_0 - y_{-1}\| + \beta \leq (1+\lambda_0)\beta, \tag{20b}$$

$$\|y_0 - x_0\| = \mu_0 d_B(x_0) \leq \mu_0(1+\lambda_0)\beta, \tag{20c}$$

$$\max\{\|y_0 - x_0\|, \|x_0 - y_{-1}\|\} \leq \alpha_0(1+\alpha_0)\beta. \tag{20d}$$

Proof. Using Proposition 3.3(ii), we have $\|x_0 - y_{-1}\| = \lambda_0 d_A(y_{-1}) \leq \lambda_0\beta$. Thus, (20a) holds. The nonexpansiveness of distance functions implies (20b). On the one hand, using Proposition 3.3(ii) again, we see that $\|y_0 - x_0\| = \mu_0 d_B(x_0)$. On the other hand, (20b) yields $d_B(x_0) \leq (1+\lambda_0)\beta$. Altogether, we obtain (20c). Finally, (20d) follows from (20a) and (20c). ■

The following lemma is important for our analysis.

Lemma 3.8 *Let $n \in \mathbb{N}$ and let $\theta \in [0,1]$ be such that*

$$\langle y_n - x_n, y_{n-1} - x_n\rangle \leq \theta\|y_n - x_n\| \cdot \|x_n - y_{n-1}\|. \tag{21}$$

Then

$$\|x_{n+1} - y_n\| \leq \tfrac{\lambda_{n+1}}{\lambda_n}\sqrt{\lambda_n^2 + (1-\lambda_n)^2 + 2\theta\lambda_n(1-\lambda_n)} \cdot \max\left\{\|y_n - x_n\|, \|x_n - y_{n-1}\|\right\}. \tag{22}$$

Proof. Proposition 3.3(iii) yields

$$x_n - a_n = \tfrac{1-\lambda_n}{\lambda_n}(y_{n-1} - x_n). \tag{23}$$

Combining (23) with assumption (21), we have

$$\langle y_n - x_n, x_n - a_n\rangle = \tfrac{1-\lambda_n}{\lambda_n}\langle y_n - x_n, y_{n-1} - x_n\rangle \le \tfrac{\theta(1-\lambda_n)}{\lambda_n}\|y_n - x_n\| \cdot \|x_n - y_{n-1}\|. \tag{24}$$

Substituting (23) and (24) into

$$\|y_n - a_n\|^2 = \|y_n - x_n\|^2 + \|x_n - a_n\|^2 + 2\langle y_n - x_n, x_n - a_n\rangle \tag{25}$$

gives

$$\|y_n - a_n\|^2 \le \|y_n - x_n\|^2 + \tfrac{(1-\lambda_n)^2}{\lambda_n^2}\|x_n - y_{n-1}\|^2 + 2\tfrac{\theta(1-\lambda_n)}{\lambda_n}\|y_n - x_n\| \cdot \|x_n - y_{n-1}\|. \tag{26}$$

Multiplying both sides by λ_{n+1}^2, we have

$$\lambda_{n+1}^2\|y_n - a_n\|^2 \le \tfrac{\lambda_{n+1}^2}{\lambda_n^2}\big(\lambda_n^2 + (1-\lambda_n)^2 + 2\theta\lambda_n(1-\lambda_n)\big)\max{}^2\big\{\|y_n - x_n\|, \|x_n - y_{n-1}\|\big\}. \tag{27}$$

From Proposition 3.3(ii), we have

$$\|x_{n+1} - y_n\| = \lambda_{n+1} d_A(y_n)\| \le \lambda_{n+1}\|y_n - a_n\|. \tag{28}$$

Combining with (27), we obtain the result. ■

A proof analogous to that of Lemma 3.8 (or interchanging the roles of A and B) yields the following result.

Lemma 3.9 *Let $n \in \mathbb{N}$ and let $\theta \in [0,1]$ be such that*

$$\langle x_{n+1} - y_n, x_n - y_n\rangle \le \theta\|x_{n+1} - y_n\| \cdot \|y_n - x_n\|. \tag{29}$$

Then

$$\|y_{n+1} - x_{n+1}\| \le \tfrac{\mu_{n+1}}{\mu_n}\sqrt{\mu_n^2 + (1-\mu_n)^2 + 2\theta\mu_n(1-\mu_n)} \cdot \max\big\{\|x_{n+1} - y_n\|, \|y_n - x_n\|\big\}. \tag{30}$$

4 Abstract Linear Convergence

In this section, we provide convergence results that refine and complement those of [9, Proposition 3.8] and [23][4].

Lemma 4.1 (abstract linear convergence) *Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\ge -1}$ be sequences in X. Assume that there exist constants $M \in \mathbb{R}_+$ and $\rho \in [0,1[$ such that*

$$(\forall n \in \mathbb{N}) \quad \max\big\{d(y_n, x_n), d(x_n, y_{n-1})\big\} \le M\rho^n. \tag{31}$$

Then there exists $\bar{c} \in X$ such that

$$(\forall n \in \mathbb{N}) \quad \max\big\{d(x_n, \bar{c}), d(y_n, \bar{c})\big\} \le \frac{M(1+\rho)}{1-\rho} \cdot \rho^n; \tag{32}$$

consequently, $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge linearly to $\bar{c}$ with rate ρ.

[4] In fact, the results in this section hold true in any complete metric space.

Proof. We have

$$(\forall n \in \mathbb{N}) \quad d(y_n, y_{n-1}) \leq d(y_n, x_n) + d(x_n, y_{n-1}) \leq 2\rho^n M. \tag{33}$$

Hence, for every $k \in \{n+1, n+2, \ldots\}$,

$$d(y_k, y_n) \leq \sum_{i=n+1}^{k} d(y_i, y_{i-1}) \leq 2M \sum_{i=n+1}^{k} \rho^i \leq \frac{2M\rho^{n+1}}{1-\rho}. \tag{34}$$

Thus $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence with, say, limit $\bar{c} \in X$. Letting $k \to +\infty$ in (34), we see that

$$d(y_n, \bar{c}) \leq \frac{2M\rho^{n+1}}{1-\rho}. \tag{35}$$

It also follows that

$$d(x_n, \bar{c}) \leq d(x_n, y_n) + d(y_n, \bar{c}) \leq M\rho^n + \frac{2M\rho^{n+1}}{1-\rho} = \frac{M(1+\rho)\rho^n}{1-\rho}. \tag{36}$$

Therefore, (36) implies that

$$(\forall n \in \mathbb{N}) \quad \max\big\{d(x_n, \bar{c}), d(y_n, \bar{c})\big\} \leq \frac{M(1+\rho)}{1-\rho} \cdot \rho^n, \tag{37}$$

as claimed. ■

Definition 4.2 (alternating contraction property) *Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\geq -1}$ be sequences in X, let $c \in X$, and let $(r, \rho) \in \mathbb{R}_{++} \times [0, 1[$. We say that $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\geq -1}$ have the* alternating contraction property at c with parameters (r, ρ) *if the following implication holds whenever $n \in \mathbb{N}$ and $(u_1, u_2, u_3, u_4) \in \{(y_{n-1}, x_n, y_n, x_{n+1}), (x_n, y_n, x_{n+1}, y_{n+1})\}$:*

$$\max\big\{d(u_2, c), d(u_3, c)\big\} \leq r \quad \Rightarrow \quad d(u_3, u_4) \leq \rho \max\big\{d(u_1, u_2), d(u_2, u_3)\big\}. \tag{38}$$

Theorem 4.3 (abstract linear convergence) *Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\geq -1}$ be sequences in X, and let $(r, \rho) \in \mathbb{R}_{++} \times [0, 1[$. Assume that the sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\geq -1}$ have the alternating contraction property at y_{-1} with parameters (r, ρ). Assume further that*

$$M := \max\big\{d(y_0, x_0), d(x_0, y_{-1})\big\} \leq \tfrac{r(1-\rho)}{2}. \tag{39}$$

Then $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge linearly to a point $\bar{c} \in X$ with rate ρ; more precisely,

$$(\forall n \in \mathbb{N}) \quad \max\big\{d(x_n, \bar{c}), d(y_n, \bar{c})\big\} \leq \tfrac{M(1+\rho)}{1-\rho}\rho^n \leq \tfrac{r(1+\rho)}{2}\rho^n. \tag{40}$$

In addition, $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$, and hence $\bar{c}$, all lie in $\mathrm{ball}(y_{-1}; r)$.

Proof. Using (39), we estimate

$$(\forall n \in \mathbb{N}) \quad 2M \sum_{i=0}^{n} \rho^i \leq \frac{2M}{1-\rho} \leq r. \tag{41}$$

We now show by induction that the following holds for every $n \in \mathbb{N}$:

$$d(x_n, y_{-1}) \leq \Big(\sum_{i=0}^{n} \rho^i + \sum_{i=0}^{n-1} \rho^i\Big) M, \tag{42a}$$

(42b) $$\max\big\{d(y_n, x_n), d(x_n, y_{n-1})\big\} \leq \rho^n M.$$

Clearly, in view of the definition of M, (42) holds when $n = 0$.

Now assume that (42) holds for some $n \in \mathbb{N}$. First, using (41) and (42), we have

(43) $$d(x_n, y_{-1}) \leq r \quad \text{and} \quad d(y_n, y_{-1}) \leq d(y_n, x_n) + d(x_n, y_{-1}) \leq r.$$

So the contraction property applied to the quadruple $(y_{n-1}, x_n, y_n, x_{n+1})$ implies

(44) $$d(x_{n+1}, y_n) \leq \rho \max\{d(y_n, x_n), d(x_n, y_{n-1})\} \leq \rho^{n+1} M.$$

It follows that

(45a) $$d(x_{n+1}, y_{-1}) \leq d(x_{n+1}, y_n) + d(y_n, x_n) + d(x_n, y_{-1})$$

(45b) $$\leq \rho^{n+1} M + \rho^n M + \Big(\sum_{i=0}^{n} \rho^i + \sum_{i=0}^{n-1} \rho^i\Big) M$$

(45c) $$= \Big(\sum_{i=0}^{n+1} \rho^i + \sum_{i=0}^{n} \rho^i\Big) M$$

(45d) $$\leq r.$$

So (42a) holds with n replaced by $n+1$. Next, the contraction property applied to the quadruple $(x_n, y_n, x_{n+1}, y_{n+1})$ yields

(46) $$d(y_{n+1}, x_{n+1}) \leq \rho \max\big\{d(x_{n+1}, y_n), d(y_n, x_n)\big\}.$$

In view of (46), (44), and (42b), we deduce that

(47) $$\max\big\{d(y_{n+1}, x_{n+1}), d(x_{n+1}, y_n)\big\} \leq \rho^{n+1} M,$$

i.e., (42b) holds with n replaced by $n+1$. Thus, by induction, (42) holds for every $n \in \mathbb{N}$.

Combining (42b), Lemma 4.1, and (41), we obtain

(48) $$(\forall n \in \mathbb{N}) \quad \max\{d(x_n, \bar{c}), d(y_n, \bar{c})\} \leq \frac{M(1+\rho)}{1-\rho}\rho^n \leq \frac{r(1+\rho)}{2}\rho^n.$$

Finally, (43) implies that the sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$, and consequently their common limit $\bar{c}$, lie in $\mathrm{ball}(y_{-1}; r)$. ■

5 Linear Convergence of the MARP and the CQ Condition

The sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ produced by the MARP need not lie in the sets A and B, respectively. Therefore, the techniques utilized for the method of alternating projections in [8, 9] and [23] cannot be directly applied. In this section, we present a new technique which relies on the geometry of Euclidean spaces.

In addition to our assumptions on the sets A and B, the relaxation parameter sequences $\boldsymbol{\lambda} = (\lambda_n)_{n\in\mathbb{N}}$ and $\boldsymbol{\mu} = (\mu_n)_{n\in\mathbb{N}}$, and the MARP sequences $\mathbf{x} = (x_n)_{n\in\mathbb{N}}$ and $\mathbf{y} = (y_n)_{n\in\mathbb{N}}$ with starting point $y_{-1} \in X$ (see (2), (14), and (16)), we assume the following in this section:

(49a) S is a subset of X that is projection absorbing with respect to A and B, $y_{-1} \in S$, and

$$(49b)\qquad (\forall n \in \mathbb{N}) \quad \lambda_n \geq \lambda_{n+1} \to \lambda_\infty \ \text{ and } \ \mu_n \geq \mu_{n+1} \to \mu_\infty, \quad \text{and} \quad \alpha_\infty := \min\{\lambda_\infty, \mu_\infty\}.$$

We start with a technical result.

Lemma 5.1 *Let $\mu \in \,]0,1]$ and let $\theta \in [0,1[$. Then*

$$0 < \mu^2 + (1-\mu)^2 + 2\theta\mu(1-\mu) = 1 - 2(1-\theta)\mu(1-\mu) \leq 1, \tag{50}$$

and the last inequality is an equality if and only if $\mu = 1$.

Proof. Clearly,

$$\begin{aligned}
&\text{(51a)} & 0 \leq 2(1+\theta)\mu(1-\mu) &= 2\mu(1-\mu) + 2\theta\mu(1-\mu) \\
&\text{(51b)} & &\leq \mu^2 + (1-\mu)^2 + 2\theta\mu(1-\mu) \\
&\text{(51c)} & &= \mu^2 + (1-\mu)^2 + 2\mu(1-\mu) - 2(1-\theta)\mu(1-\mu) \\
&\text{(51d)} & &= (\mu + (1-\mu))^2 - 2(1-\theta)\mu(1-\mu) \\
&\text{(51e)} & &= 1 - 2(1-\theta)\mu(1-\mu) \leq 1.
\end{aligned}$$

Note that equality in (51a) occurs exactly when $\mu = 1$; in this case, the inequality (51b) is strict. Furthermore, equality in (51e) occurs exactly when $\mu = 1$. ■

The following result will help us later in this section to identify the convergence rate of the MARP.

Lemma 5.2 *Let $\theta \in [0,1[$ and define $\hat{\rho} \in \mathbb{R}_+$ implicitly by*

$$\hat{\rho}^2 := \sup_{n\in\mathbb{N}} \left\{ \begin{array}{l} \frac{\lambda_{n+1}^2}{\lambda_n^2}\big(\lambda_n^2 + (1-\lambda_n)^2 + 2\theta\lambda_n(1-\lambda_n)\big), \\ \frac{\mu_{n+1}^2}{\mu_n^2}\big(\mu_n^2 + (1-\mu_n)^2 + 2\theta\mu_n(1-\mu_n)\big) \end{array} \right\}. \tag{52}$$

Then $0 < \hat{\rho} \leq \sqrt{1 - 2(1-\theta)\min\big\{\alpha_0(1-\alpha_0), \alpha_\infty(1-\alpha_\infty)\big\}} \leq 1$; consequently, if $1 > \alpha_0 \geq \alpha_\infty > 0$, then $\hat{\rho} < 1$.

Proof. Let us first consider

$$\sigma := \sup_{n\in\mathbb{N}} \frac{\mu_{n+1}^2}{\mu_n^2}\big(\mu_n^2 + (1-\mu_n)^2 + 2\theta\mu_n(1-\mu_n)\big), \tag{53}$$

the corresponding supremum involving $\boldsymbol{\lambda}$ is treated similarly. Lemma 5.1 yields $\sigma > 0$. Since $(\forall n \in \mathbb{N})\ 0 < \frac{\mu_\infty}{\mu_0} \leq \frac{\mu_{n+1}}{\mu_n} \leq 1$ and hence $\frac{\mu_{n+1}^2}{\mu_n^2} \leq 1$, we estimate with the help of Lemma 5.1 that

$$\begin{aligned}
&\text{(54a)} & 0 < \sigma &\leq \sup_{n\in\mathbb{N}} \big(\mu_n^2 + (1-\mu_n)^2 + 2\theta\mu_n(1-\mu_n)\big) \\
&\text{(54b)} & &= \sup_{n\in\mathbb{N}} \big(1 - 2(1-\theta)\mu_n(1-\mu_n)\big)
\end{aligned}$$

$$(54c)\qquad = 1-2(1-\theta)\min\{\mu_0(1-\mu_0),\mu_\infty(1-\mu_\infty)\}$$

because any minimizer of the function $\mu \mapsto \mu(1-\mu)$ restricted to the interval $[\mu_\infty, \mu_0]$ must be one of the endpoints of the interval. The conclusion now follows by combining this estimate with its $\boldsymbol{\lambda}$ counterpart. ■

The following result provides information about the location of limits of the MARP.

Proposition 5.3 *Suppose both sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ generated by the MARP converge to $\bar{c} \in X$. Then the following hold:*

(i) *If $\lambda_\infty > 0$, then $\bar{c} \in A$.*

(ii) *If $\mu_\infty > 0$, then $\bar{c} \in B$.*

(iii) *If $\alpha_\infty > 0$, then $\bar{c} \in A \cap B$.*

Proof. Clearly, $x_n - y_n \to 0$ and $y_n - x_{n+1} \to 0$.

(i): Suppose that $\lambda_\infty > 0$. By Proposition 3.3(ii), $0 \leftarrow \|x_{n+1} - y_n\| = \lambda_{n+1} d_A(y_n)$. Since $\lambda_\infty > 0$, it follows that $d_A(y_n) \to 0$. Hence, $\bar{c} \in A$.

(ii): The proof is analogous to that of (i).

(iii): Combine (i) and (ii). ■

The following examples illustrate that no conclusion can be drawn about the location of the limit point when $\lambda_\infty = 0$ or $\mu_\infty = 0$.

Example 5.4 (MARP limit point lies outside $A \cup B$ and $\lambda_\infty = \mu_\infty = 0$) Suppose that $X = S = \mathbb{R}$, that $A = B = \mathbb{R}_-$. Let $\delta \in \mathbb{R}_{++}$, and assume that $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ satisfy

$$(55)\qquad (\forall n \in \mathbb{N}) \quad \lambda_n = \mu_n = 1 - \sqrt{\frac{\delta + 2^{-(n+1)}}{\delta + 2^{-n}}} \in \left]0,1\right[.$$

Then $\lambda_\infty = \mu_\infty = 0$. Suppose that $y_{-1} \in \mathbb{R}_{++} = X \smallsetminus (A \cup B)$. Then the $(\boldsymbol{\lambda}, \boldsymbol{\mu})$-MARP sequences are

$$(56)\qquad (\forall n \in \mathbb{N}) \quad x_n = y_{n-1}\sqrt{\frac{\delta + 2^{-(n+1)}}{\delta + 2^{-n}}} \quad \text{and} \quad y_n = x_n\sqrt{\frac{\delta + 2^{-(n+1)}}{\delta + 2^{-n}}},$$

which inductively leads to

$$(57)\qquad x_n = y_{-1}\sqrt{\frac{\delta + 2^{-(n+1)}}{\delta + 2^{-n}}}\left(\frac{\delta + 2^{-n}}{\delta + 1}\right) \quad \text{and} \quad y_n = y_{-1}\frac{\delta + 2^{-(n+1)}}{\delta + 1}.$$

Note that $\lim_{n\in\mathbb{N}} x_n = \lim_{n\in\mathbb{N}} y_n = \frac{\delta y_{-1}}{\delta+1} \notin A \cup B$.

Example 5.5 (MARP limit point lies in $A \cap B$ and $\lambda_\infty = \mu_\infty = 0$) Suppose that $X = S = \mathbb{R}$, that $A = B = \mathbb{R}_-$, and that $\lambda_\infty = \mu_\infty = 0$ while $\sum_{n\in\mathbb{N}} \lambda_n = \sum_{n\in\mathbb{N}} \mu_n = +\infty$. Furthermore, assume that $y_{-1} = \eta \in \mathbb{R}_{++}$. Then

$$(58)\qquad (\forall n \in \mathbb{N}) \quad x_n = \eta\prod_{i=0}^{n}(1-\lambda_i)\prod_{i=0}^{n-1}(1-\mu_i) \quad \text{and} \quad y_n = \eta\prod_{i=0}^{n}(1-\lambda_i)\prod_{i=0}^{n}(1-\mu_i),$$

and so

$$\ln(y_n/\eta) = \ln\Big(\prod_{i=0}^{n}(1-\lambda_i)\prod_{i=0}^{n}(1-\mu_i)\Big) = \sum_{i=0}^{n}\ln(1-\lambda_i) + \sum_{i=0}^{n}\ln(1-\mu_i) \tag{59a}$$

$$\leq \sum_{i=0}^{n}(-\lambda_i) + \sum_{i=0}^{n}(-\mu_i) \to -\infty. \tag{59b}$$

It follows that $y_n \to 0$ and thus $x_n \to 0$. Hence $\lim_{n\in\mathbb{N}} x_n = \lim_{n\in\mathbb{N}} y_n \in A\cap B$.

Example 5.6 (MARP limit point lies outside $A\cap B$ and $\lambda_\infty > 0 = \mu_\infty$) Suppose that $X = S = \mathbb{R}^2$, that $A = \mathbb{R}\times\{0\}$, that $B = \{0\}\times\mathbb{R}$, that $(\forall n\in\mathbb{N})$ $\lambda_n = \frac{1}{2}$ and $\mu_n = 1 - \frac{1+2^{-(n+1)}}{1+2^{-n}}$, and that $y_{-1} = (\eta,\zeta) \in X \smallsetminus A$. Then

$$(\forall n\in\mathbb{N})\quad x_n = \Big(\tfrac{\eta}{2^{n+1}}, \tfrac{\zeta(1+2^{-n})}{2}\Big) \quad\text{and}\quad y_n = \Big(\tfrac{\eta}{2^{n+1}}, \tfrac{\zeta(1+2^{-(n+1)})}{2}\Big); \tag{60}$$

therefore, $\lim_{n\in\mathbb{N}} x_n = \lim_{n\in\mathbb{N}} y_n = (0,\frac{\zeta}{2}) \in B\smallsetminus A$.

We now present the main convergence result of this section.

Theorem 5.7 (local linear convergence) *Let $r\in\mathbb{R}_{++}$ and let $\theta\in[0,1[$. Assume that the following hold:*

(i) $\alpha_0 < 1$;

(ii) $\hat{\rho} \leq \rho < 1$, *where $\hat{\rho}$ is as in Lemma 5.2;*

(iii) $\max\big\{d_A(y_{-1}), d_B(y_{-1})\big\} \leq \frac{r(1-\rho)}{2\alpha_0(1+\alpha_0)}$;

(iv)

$$\big(\forall x\in S\cap \operatorname{ball}(y_{-1};r)\big)\big(\forall a\in P_A x\big)\big(\forall b\in P_B x\big)\quad \langle a-x, x-b\rangle \leq \theta\|a-x\|\cdot\|x-b\|. \tag{61}$$

Then the MARP sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge linearly with rate ρ to some point $\bar{c} \in$ ball$(y_{-1};r)$ and

$$(\forall n\in\mathbb{N})\quad \max\big\{\|x_n-\bar{c}\|, \|y_n-\bar{c}\|\big\} \leq \tfrac{r(1+\rho)}{2}\rho^n. \tag{62}$$

Furthermore, if $\min\{\lambda_\infty,\mu_\infty\} > 0$, then $\bar{c}\in A\cap B$.

Proof. Combining (49) and Proposition 3.6, we deduce that $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ lie in S.

We now claim that $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\geq -1}$ have the alternating contraction property at y_{-1} with parameters (r,ρ) (see Definition 4.2). Let us fix $n\in\mathbb{N}$ and check (38) for the quadruple $(y_{n-1}, x_n, y_n, x_{n+1})$; the other quadruple $(x_n, y_n, x_{n+1}, y_{n+1})$ is treated similarly. We assume that $\|x_n - y_{-1}\| \leq r$. Since $a_n\in P_A x_n$ and $b_n \in P_B x_n$ (see Definition 3.2), (61) yields

$$\langle a_n - x_n, x_n - b_n\rangle \leq \theta\|a_n - x_n\|\cdot\|x_n - b_n\|. \tag{63}$$

Proposition 3.3(iii) implies that $y_{n-1} - x_n = \frac{\lambda_n}{1-\lambda_n}(x_n - a_n)$ and $y_n - x_n = \frac{\mu_n}{1-\mu_n}(b_n - x_n)$; thus,

$$\langle y_n - x_n, y_{n-1} - x_n\rangle \leq \theta\|y_n - x_n\|\cdot\|x_n - y_{n-1}\|. \tag{64}$$

Hence, by Lemma 3.8 and assumption (ii),

$$\|x_{n+1} - y_n\| \leq \rho \max\big\{\|y_n - x_n\|, \|x_n - y_{n-1}\|\big\}. \tag{65}$$

Thus (38) holds, as claimed.

It now follows from (20d) of Lemma 3.7 and assumption (iii) that

$$M := \max\big\{\|y_0 - x_0\|, \|x_0 - y_{-1}\|\big\} \leq \alpha_0(1+\alpha_0)\max\big\{d_A(y_{-1}), d_B(y_{-1})\big\} \leq \tfrac{r(1-\rho)}{2}. \tag{66}$$

Hence, by Theorem 4.3, $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge linearly to $\bar{c} \in \operatorname{ball}(y_{-1}, r)$ and

$$(\forall n \in \mathbb{N}) \quad \max\big\{\|x_n - \bar{c}\|, \|y_n - \bar{c}\|\big\} \leq \tfrac{r(1+\rho)}{2}\rho^n. \tag{67}$$

Finally, recall Proposition 5.3. ■

Remark 5.8 (best bound for the convergence rate) In Theorem 5.7, the linear rate is tied to the constant $\hat{\rho}$ defined by (52). The computation of $\hat{\rho}$ appears to be hard in general; however, the upper bound provided in Lemma 5.2 is minimized when $\lambda_0 = \lambda_\infty = \mu_0 = \mu_\infty = \frac{1}{2}$, i.e., when $(\forall n \in \mathbb{N})\ \lambda_n = \mu_n = \frac{1}{2}$, in which case

$$0 < \hat{\rho} = \sqrt{\frac{1+\theta}{2}} < 1. \tag{68}$$

The following result concerns global convergence. As a consequence, it somewhat surprisingly guarantees the *nonemptiness* of the intersection.

Corollary 5.9 (global convergence) *Assume that* $1 > \alpha_0 \geq \alpha_\infty > 0$ *and that there exists* $\theta \in [0,1[$ *such that*

$$\big(\forall x \in S\big)\big(\forall a \in P_A x\big)\big(\forall b \in P_B x\big) \quad \langle a - x, x - b\rangle \leq \theta\|a - x\| \cdot \|x - b\|. \tag{69}$$

Then the MARP sequences $(x_n)_{n\in\mathbb{N}}$ *and* $(y_n)_{n\in\mathbb{N}}$ *converge linearly with rate* $\hat{\rho}$ *to some point in* $A \cap B$*, where* $\hat{\rho} \in\,]0,1[$ *is defined in Lemma 5.2.*

Example 5.10 (two subspaces) Suppose that A and B are affine subspaces with $A \cap B \neq \varnothing$, that $S = \operatorname{aff}(A \cup B)$, and that $1 > \alpha_0 \geq \alpha_\infty > 0$. Then there exists $\theta \in [0,1[$ such that $\hat{\rho} \in\,]0,1[$, where $\hat{\rho}$ is defined in Lemma 5.2. Moreover, the MARP sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge linearly with rate $\hat{\rho}$ to some point in $A \cap B$.

Proof. (See also [6, Theorem 5.7] for a closely related result.) After translating if necessary, we assume that A and B are linear subspaces, and that $S = A + B$. Let $x \in S$, let $a \in P_A x$, and let $b \in P_B x$. Using [8, Theorem 3.5], we have $x - a \in \widehat{N}^S_A(a) \subseteq N^S_A(a) = N_A(a) \cap S = A^\perp \cap (A+B)$. Similarly, $x - b \in B^\perp \cap (A+B)$. Since A and B are subspaces and $A^\perp \cap (A+B) \cap B^\perp \cap (A+B) = (A+B) \cap (A+B)^\perp = \{0\}$, we set

$$\theta := \max\Big\langle A^\perp \cap (A^\perp \cap B^\perp)^\perp \cap \operatorname{ball}(0;1), B^\perp \cap (A^\perp \cap B^\perp)^\perp \cap \operatorname{ball}(0;1)\Big\rangle < 1. \tag{70}$$

(Thus, θ is the cosine of the *Friedrichs angle* between $A^\perp$ and $B^\perp$, which is identical to the cosine of the Friedrichs angle between A and B.) Hence (69) holds and the conclusion now follows from Corollary 5.9. ■

In the spirit of [23] and [9], we now guarantee local linear convergence when the CQ condition holds.

Theorem 5.11 (local convergence via CQ condition) *Suppose that* $1 > \alpha_0 \geq \alpha_\infty > 0$, *that* $c \in A \cap B$ *and that the* (A, S, B, S)*-CQ holds at* c, *i.e. (see Definition 2.5),*

$$N_A^S(c) \cap (-N_B^S(c)) = \{0\}. \tag{71}$$

In view of (12), *the limiting CQ number associated with* (A, S, B, S) *(see Definition 2.4) satisfies*

$$\overline{\theta} = \max\Big\{ \langle u, v\rangle \;\Big|\; u \in N_A^S(c), v \in -N_B^S(c), \|u\| \leq 1, \|v\| \leq 1 \Big\} < 1. \tag{72}$$

Let $\theta \in \left]\overline{\theta}, 1\right[$. *Then there exists* $\delta > 0$ *such that whenever the starting point* y_{-1} *lies in* $S \cap \operatorname{ball}(c;\delta)$, *the sequences* $(x_n)_{n\in\mathbb{N}}$ *and* $(y_n)_{n\in\mathbb{N}}$ *generated by the MARP converge linearly to a point in* $A \cap B$ *with rate* $\hat{\rho} \in \left]0, 1\right[$ *(see Lemma 5.2).*

Proof. There exists $\varepsilon > 0$ sufficiently small such that $\theta_{2\varepsilon} \leq \theta$, where $\theta_{2\varepsilon}$ is the CQ number associated with (A, S, B, S) and 2ε (see Definition 2.4). We claim that

$$\delta := \tfrac{\varepsilon(1-\hat{\rho})}{1-\hat{\rho}+2\alpha_0(1+\alpha_0)} \tag{73}$$

does the job.

To this end, assume that $y_{-1} \in S \cap \operatorname{ball}(c;\delta)$ and set

$$r := \tfrac{2\delta\alpha_0(1+\alpha_0)}{1-\hat{\rho}}. \tag{74}$$

Since $c \in A \cap B$, we deduce that

$$\max\big\{d_A(y_{-1}), d_B(y_{-1})\big\} \leq \|y_{-1} - c\| \leq \delta = \tfrac{r(1-\hat{\rho})}{2\alpha_0(1+\alpha_0)}, \tag{75}$$

which is assumption (iii) of Theorem 5.7.

Now let $x \in S \cap \operatorname{ball}(y_{-1}; r)$, let $a \in P_A x$, and let $b \in P_B x$. Using (73) and (74), we estimate

$$\|x - c\| \leq \|x - y_{-1}\| + \|y_{-1} - c\| \leq r + \delta = \varepsilon. \tag{76}$$

Hence, $\|a - c\| \leq \|a - x\| + \|x - c\| = d_A(x) + \|x - c\| \leq 2\|x - c\| \leq 2\varepsilon$. Analogously, $\|b - c\| \leq 2\varepsilon$. On the other hand, $a - x \in -\widehat{N}_A^S(a)$ and $x - b \in \widehat{N}_B^S(b)$. It thus follows from the definition of the CQ-number (see (9)) and our choice of ε that

$$\langle a - x, x - b\rangle \leq \theta_{2\varepsilon}\|a - x\| \cdot \|x - b\| \leq \theta\|a - x\| \cdot \|x - b\|, \tag{77}$$

which is assumption (iv) of Theorem 5.7. Therefore, Theorem 5.7 implies that $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge linearly to a point $\bar{c} \in A \cap B \cap \operatorname{ball}(y_{-1}; r)$ and

$$(\forall n \in \mathbb{N}) \quad \max\big\{\|x_n - \bar{c}\|, \|y_n - \bar{c}\|\big\} \leq \tfrac{r(1+\hat{\rho})}{2}\hat{\rho}^n = \tfrac{\varepsilon\alpha_0(1+\alpha_0)(1+\hat{\rho})}{1-\hat{\rho}+2\alpha_0(1+\alpha_0)}\hat{\rho}^n. \tag{78}$$

We also note that $\bar{c} \in \operatorname{ball}(c;\varepsilon)$ because $\|\bar{c} - c\| \leq \|\bar{c} - y_{-1}\| + \|y_{-1} - c\| \leq r + \delta = \varepsilon$. ■

Finally, we use Aharoni and Censor's [1, Theorem 1] to obtain a linear convergence rate result in the convex case.

Corollary 5.12 (two convex sets) *Suppose that* A *and* B *are convex with* $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$, *that* $S = \operatorname{aff}(A \cup B)$, *and that* $1 > \alpha_0 \geq \alpha_\infty > 0$. *Then the sequences* $(x_n)_{n\in\mathbb{N}}$ *and* $(y_n)_{n\in\mathbb{N}}$ *generated by the MARP converge linearly to a point in* $A \cap B$.

Proof. It is known that the MARP sequences converge to some point $c \in A \cap B$; see, e.g., the aforementioned [1, Theorem 1]. By [8, Proposition 7.5], the (A, S, B, S)-CQ condition holds at c. In view of (12), the limiting CQ number associated with (A, S, B, S) (see Definition 2.4) satisfies

$$\overline{\theta} = \max\left\{ \langle u, v\rangle \;\middle|\; u \in N_A^S(c), v \in -N_B^S(c), \|u\| \le 1, \|v\| \le 1 \right\} < 1. \tag{79}$$

Let $\theta \in \,]\overline{\theta}, 1[$ and obtain $\delta > 0$ as in Theorem 5.11. Since $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge to c, there exists $n_0 \in \mathbb{N}$ such that $y_{n_0} \in B(c; \delta)$. The conclusion therefore follows from Theorem 5.11 (applied to the MARP with starting point $y_{n_0} \in S$). ■

6 Linear Convergence of the MARP and Regularity

We now investigate the MARP in the presence of regularity. We uphold the assumptions (49) of the previous section.

The following result is a counterpart of Lemma 3.8; it refines [23, Theorem 5.2] and [9, Proposition 3.4].

Lemma 6.1 *Let $\theta \in [0, 1[$, let $\delta > 0$, let $\varepsilon \ge 0$ and let $n \in \mathbb{N}$. Suppose that $c \in A$, that A is $(S, \varepsilon, 2\delta)$-regular at c (see Definition 2.2), and that the quadruple $(y_{n-1}, x_n, y_n, x_{n+1})$ generated by the MARP (see Definition 3.2) with starting point $y_{-1} \in S$ satisfies*

$$\{x_n, y_n\} \subseteq \operatorname{ball}(c; \delta) \quad \textit{and} \quad \langle x_{n+1} - y_n, x_n - y_n\rangle \le \theta \|x_{n+1} - y_n\| \cdot \|x_n - y_n\|. \tag{80}$$

Then

$$\|x_{n+1} - y_n\| \le \tfrac{\lambda_{n+1}}{\lambda_n}\big(\theta\lambda_n + 2\varepsilon + 1 - \lambda_n\big) \max\big\{\|y_n - x_n\|, \|x_n - y_{n-1}\|\big\}. \tag{81}$$

Furthermore, the following implication holds:

$$\lambda_n = 1 \;\Rightarrow\; \|x_{n+1} - y_n\| \le \lambda_{n+1}(\theta + 2\varepsilon)\|y_n - x_n\|. \tag{82}$$

Proof. Using Proposition 3.3(i), we have $\|a_n - c\| \le \|x_n - a_n\| + \|x_n - c\| = d_A(x_n) + \|x_n - c\| \le 2\|x_n - c\| \le 2\delta$. Moreover, $\|a_{n+1} - c\| \le \|a_{n+1} - y_n\| + \|y_n - c\| = d_A(y_n) + \|y_n - c\| \le 2\|y_n - c\| \le 2\delta$. Since $y_n - a_{n+1} \in \widehat{N}_A^S(a_{n+1})$ and A is $(S, \varepsilon, 2\delta)$-regular at c, we obtain

$$\langle a_{n+1} - y_n, a_{n+1} - a_n\rangle \le \varepsilon \|a_{n+1} - y_n\| \cdot \|a_{n+1} - a_n\|. \tag{83}$$

Now $a_{n+1} - y_n = \frac{1}{\lambda_{n+1}}(x_{n+1} - y_n)$ (by Proposition 3.3(ii)) and (80) imply

$$\langle a_{n+1} - y_n, x_n - y_n\rangle \le \theta \|a_{n+1} - y_n\| \cdot \|x_n - y_n\|. \tag{84}$$

Adding (83), (84) and $\langle a_{n+1} - y_n, a_n - x_n\rangle \le \|a_{n+1} - y_n\| \cdot \|a_n - x_n\|$, we obtain

$$\begin{aligned} \|a_{n+1} - y_n\|^2 \le{}& \theta\|a_{n+1} - y_n\| \cdot \|x_n - y_n\| \\ &+ \varepsilon\|a_{n+1} - y_n\| \cdot \|a_{n+1} - a_n\| + \|a_{n+1} - y_n\| \cdot \|a_n - x_n\|; \end{aligned} \tag{85}$$

thus,

$$\|a_{n+1} - y_n\| \le \theta\|y_n - x_n\| + \varepsilon\|a_{n+1} - a_n\| + \|a_n - x_n\|. \tag{86}$$

Substituting $\|a_{n+1} - a_n\| \leq \|a_{n+1} - y_n\| + \|a_n - y_n\| \leq 2\|a_n - y_n\| \leq 2\|y_n - x_n\| + 2\|x_n - a_n\|$ into (86) results in

$$\|a_{n+1} - y_n\| \leq (\theta + 2\varepsilon)\|y_n - x_n\| + (1 + 2\varepsilon)\|a_n - x_n\|. \tag{87}$$

Therefore, since $\|a_n - x_n\| = \frac{1-\lambda_n}{\lambda_n}\|x_n - y_{n-1}\|$ and $\|a_{n+1} - y_n\| = \frac{1}{\lambda_{n+1}}\|x_{n+1} - y_n\|$ by Proposition 3.3(iii)&(ii), we obtain

$$\|x_{n+1} - y_n\| \leq \tfrac{\lambda_{n+1}}{\lambda_n}\big((\theta + 2\varepsilon)\lambda_n\|y_n - x_n\| + (1 + 2\varepsilon)(1 - \lambda_n)\|x_n - y_{n-1}\|\big) \tag{88a}$$

$$\leq \tfrac{\lambda_{n+1}}{\lambda_n}\big((\theta + 2\varepsilon)\lambda_n + (1 + 2\varepsilon)(1 - \lambda_n)\big)\max\big\{\|y_n - x_n\|, \|x_n - y_{n-1}\|\big\} \tag{88b}$$

$$= \tfrac{\lambda_{n+1}}{\lambda_n}\big(\theta\lambda_n + 2\varepsilon + 1 - \lambda_n\big)\max\big\{\|y_n - x_n\|, \|x_n - y_{n-1}\|\big\}, \tag{88c}$$

which is (81), as announced. Finally, (82) follows from (88a). ■

Analogously to the proof of Lemma 6.1, we obtain the following result.

Lemma 6.2 *Let $\theta \in [0, 1[$, let $\delta > 0$, let $\varepsilon \geq 0$ and let $n \in \mathbb{N}$. Suppose that $c \in B$, that B is $(S, \varepsilon, 2\delta)$-regular at c (see Definition 2.2), and that the quadruple $(x_n, y_n, x_{n+1}, y_{n+1})$ generated by the MARP (see Definition 3.2) with starting point $y_{-1} \in S$ satisfies*

$$\{y_n, x_{n+1}\} \subseteq \mathrm{ball}(c; \delta) \quad \textit{and} \quad \langle y_{n+1} - x_{n+1}, y_n - x_{n+1}\rangle \leq \theta\|y_{n+1} - x_{n+1}\| \cdot \|y_n - x_{n+1}\|. \tag{89}$$

Then

$$\|y_{n+1} - x_{n+1}\| \leq \tfrac{\mu_{n+1}}{\mu_n}\big(\theta\mu_n + 2\varepsilon + 1 - \mu_n\big)\max\big\{\|x_{n+1} - y_n\|, \|y_n - x_n\|\big\}. \tag{90}$$

Furthermore, the following implication holds:

$$\mu_n = 1 \;\Rightarrow\; \|y_{n+1} - x_{n+1}\| \leq \mu_{n+1}(\theta + 2\varepsilon)\|x_{n+1} - y_n\|. \tag{91}$$

The next result will be useful later in this section.

Lemma 6.3 *Assume that $\alpha_\infty > 0$ and let $\varepsilon \in \mathbb{R}_{++}$ and $\theta \in [0, 1[$ be such that $(1 - \theta)\alpha_\infty > 2\varepsilon$. Then*

$$0 < \hat{\kappa} := \sup_{n \in \mathbb{N}}\Big\{\tfrac{\lambda_{n+1}}{\lambda_n}\big(\theta\lambda_n + 2\varepsilon + 1 - \lambda_n\big), \tfrac{\mu_{n+1}}{\mu_n}\big(\theta\mu_n + 2\varepsilon + 1 - \mu_n\big)\Big\} \tag{92a}$$

$$\leq 1 - \big((1 - \theta)\alpha_\infty - 2\varepsilon\big) < 1. \tag{92b}$$

Proof. Clearly, $0 < \hat{\kappa}$. Since $\theta - 1 < 0$, we obtain

$$(\forall n \in \mathbb{N}) \quad \theta\lambda_n + 2\varepsilon + 1 - \lambda_n = (\theta - 1)\lambda_n + 1 + 2\varepsilon \leq (\theta - 1)\alpha_\infty + 1 + 2\varepsilon \tag{93}$$

and $\theta\mu_n + 2\varepsilon + 1 - \mu_n \leq (\theta - 1)\alpha_\infty + 1 + 2\varepsilon$. Therefore, $\hat{\kappa} \leq (\theta - 1)\alpha_\infty + 1 + 2\varepsilon < 1$. ■

The proof of the following result is partially similar to that of Theorem 5.7; however, the linear rates of convergence obtained are different.

Theorem 6.4 (MARP with regularity of sets) *Let $\varepsilon \geq 0$, $\delta > 0$, and $\theta \in [0, 1 - 2\varepsilon[$. Assume that the following hold:*

(i) *A and B are $(S, \varepsilon, 2\delta)$-regular at $c \in A \cap B$;*

(ii) $\big(\forall x \in S \cap \operatorname{ball}(c;\delta)\big)\big(\forall a \in P_A x\big)\big(\forall b \in P_B x\big) \quad \langle a-x, x-b\rangle \leq \theta\|a-x\|\cdot\|x-b\|;$

(iii) $\alpha_\infty > 2\varepsilon/(1-\theta) \geq 0.$

Assume also that the starting point y_{-1} of the MARP sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ satisfies

$$y_{-1} \in S \quad \textit{and} \quad \|y_{-1}-c\| \leq \frac{\delta(1-\hat{\kappa})}{1-\hat{\kappa}+2\alpha_0(1+\alpha_0)}, \tag{94}$$

where $\hat{\kappa} \in \,]0,1[$ is as in (92). Then $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge linearly to a point $\bar{c} \in A\cap B\cap$ $\operatorname{ball}(c;\delta)$ with rate $\hat{\kappa}$; indeed,

$$(\forall n \in \mathbb{N}) \quad \max\big\{\|x_n-\bar{c}\|, \|y_n-\bar{c}\|\big\} \leq \frac{\delta\alpha_0(1+\alpha_0)(1+\hat{\kappa})}{1-\hat{\kappa}+2\alpha_0(1+\alpha_0)}\hat{\kappa}^n. \tag{95}$$

Furthermore, if $(\forall n \in \mathbb{N})$ $\lambda_n = \mu_n = 1$, then $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge linearly with rate $\hat{\kappa}^2 = (\theta+2\varepsilon)^2$:

$$(\forall n \in \mathbb{N}) \quad \max\big\{\|x_n-\bar{c}\|, \|y_n-\bar{c}\|\big\} \leq \frac{2\delta(1+\hat{\kappa}^2)}{(1+\hat{\kappa})(5-\hat{\kappa})}\hat{\kappa}^{2n}. \tag{96}$$

Proof. Set

$$r := \frac{2\delta\alpha_0(1+\alpha_0)}{1-\hat{\kappa}+2\alpha_0(1+\alpha_0)}. \tag{97}$$

We claim that

$$(x_n)_{n\in\mathbb{N}} \text{ and } (y_n)_{n\in\mathbb{N}} \text{ have the alternating contraction property} \tag{98}$$

at y_{-1} with parameter $(r,\hat{\kappa})$ (recall Definition 4.2). Let $n \in \mathbb{N}$ and consider first the quadruple $(y_{n-1}, x_n, y_n, x_{n+1})$. In order to prove (38), we start by assuming that

$$\max\big\{\|x_n-y_{-1}\|, \|y_n-y_{-1}\|\big\} \leq r. \tag{99}$$

Then, using (94) and (97), we obtain

$$\max\big\{\|x_n-c\|, \|y_n-c\|\big\} \leq r+\|y_{-1}-c\| \leq r+\frac{\delta(1-\hat{\kappa})}{1-\hat{\kappa}+2\alpha_0(1+\alpha_0)} = \delta. \tag{100}$$

Applying (ii) with y_n, $a_{n+1} \in P_A y_n$, and $b_n = P_B y_n$, we see that

$$\langle a_{n+1}-y_n, y_n-b_n\rangle \leq \theta\|a_{n+1}-y_n\|\cdot\|y_n-b_n\|. \tag{101}$$

On the other hand, Proposition 3.3(ii)&(iii) implies $a_{n+1}-y_n = \frac{1}{\lambda_{n+1}}(x_{n+1}-y_n)$ and $y_n - b_n = \frac{1-\mu_n}{\mu_n}(x_n-y_n)$. Altogether,

$$\langle x_{n+1}-y_n, x_n-y_n\rangle \leq \theta\|x_{n+1}-y_n\|\cdot\|x_n-y_n\|. \tag{102}$$

In view of Lemma 6.1, we now deduce

$$\|x_{n+1}-y_n\| \leq \hat{\kappa}\max\big\{\|y_n-x_n\|, \|x_n-y_{n-1}\|\big\}. \tag{103}$$

This verifies (38) for the quadruple $(y_{n-1}, x_n, y_n, x_{n+1})$. The quadruple $(x_n, y_n, x_{n+1}, y_{n+1})$ is treated similarly (invoke Lemma 6.2 instead of Lemma 6.1). Therefore, (98) holds.

Next, using inequality (20d) of Lemma 3.7, the assumption that $c \in A \cap B$ (see (i)), and (94), we obtain

$$\max\{\|y_0 - x_0\|, \|x_0 - y_{-1}\|\} \leq \alpha_0(1+\alpha_0)\max\{d_A(y_{-1}), d_B(y_{-1})\} \tag{104a}$$

$$\leq \alpha_0(1+\alpha_0)\|y_{-1} - c\| \tag{104b}$$

$$\leq \frac{\alpha_0(1+\alpha_0)\delta(1-\hat{\kappa})}{1-\hat{\kappa}+2\alpha_0(1+\alpha_0)} \tag{104c}$$

$$= \frac{r(1-\hat{\kappa})}{2}. \tag{104d}$$

Thus, Theorem 4.3 and (97) yield the existence of $\bar{c} \in \operatorname{ball}(y_{-1}; r)$ such that

$$(\forall n \in \mathbb{N}) \quad \max\{\|x_n - \bar{c}\|, \|y_n - \bar{c}\|\} \leq \frac{r(1+\hat{\kappa})}{2}\hat{\kappa}^n = \frac{\delta\alpha_0(1+\alpha_0)(1+\hat{\kappa})}{1-\hat{\kappa}+2\alpha_0(1+\alpha_0)}\hat{\kappa}^n. \tag{105}$$

Furthermore, Theorem 4.3 also states that (99) holds for every $n \in \mathbb{N}$; consequently, so does its consequence (100). Also, Proposition 5.3, assumption (iii) and (100) imply that $\bar{c} \in A \cap B \cap \operatorname{ball}(c; \delta)$.

Finally, we additionally assume that $(\forall n \in \mathbb{N})\ \lambda_n = \mu_n = 1$. Then $\alpha_0 = \alpha_\infty = 1$, $\hat{\kappa} = \theta + 2\varepsilon$, and $r = \frac{4\delta}{5-\hat{\kappa}}$. Combining (100), (102), (82) and (91) yields

$$(\forall n \in \mathbb{N}) \quad \|x_{n+1} - y_n\| \leq \hat{\kappa}\|y_n - x_n\| \text{ and } \|y_n - x_n\| \leq \hat{\kappa}\|x_n - y_{n-1}\|; \tag{106}$$

consequently, $\|x_{n+1} - y_n\| \leq \hat{\kappa}^2\|x_n - y_{n-1}\| = \hat{\kappa}^2 \max\{\|y_n - x_n\|, \|x_n - y_{n-1}\|\}$ and similarly $\|y_{n+1} - x_{n+1}\|^2 \leq \hat{\kappa}^2 \max\{\|x_{n+1} - y_n\|, \|y_n - x_n\|\}$. Thus, $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ have the contraction property at y_{-1} with parameters $(r, \hat{\kappa}^2)$. Now, (39) holds with $(r, \hat{\kappa}^2)$ because of (104) and

$$M := \max\{\|y_0 - x_0\|, \|x_0 - y_{-1}\|\} \leq \tfrac{r(1-\hat{\kappa})}{2} \leq \tfrac{r(1-\hat{\kappa}^2)}{2}. \tag{107}$$

Hence, Theorem 4.3 implies that $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge linearly with rate $\hat{\kappa}^2 = (\theta + 2\varepsilon)^2$; in fact,

$$(\forall n \in \mathbb{N}) \quad \max\{\|x_n - \bar{c}\|, \|y_n - \bar{c}\|\} \leq \tfrac{M(1+\hat{\kappa}^2)}{1-\hat{\kappa}^2}\hat{\kappa}^{2n} \leq \tfrac{2\delta(1+\hat{\kappa}^2)}{(1+\hat{\kappa})(5-\hat{\kappa})}\hat{\kappa}^{2n}. \tag{108}$$

This completes the proof. ■

Remark 6.5 (comparing rates in the $(S, 0, 2\delta)$-regular case) Consider Theorem 6.4 when A and B are $(S, 0, 2\delta)$-regular at $c \in A \cap B$, and $\alpha_\infty > 0$. This happens, e.g., when A and B are convex. Since $\varepsilon = 0$, we have $\theta \in [0, 1[$. Consider the function

$$f\colon\ ]0,1] \to\ ]0,1[\ : \lambda \mapsto \frac{\theta\lambda + 1 - \lambda}{\sqrt{\lambda^2 + (1-\lambda)^2 + 2\theta\lambda(1-\lambda)}}. \tag{109}$$

Then

$$f'\colon \lambda \mapsto \frac{-\lambda(1-\theta^2)}{\big(\lambda^2 + (1-\lambda)^2 + 2\theta\lambda(1-\lambda)\big)^{3/2}} < 0. \tag{110}$$

Hence f is strictly decreasing and therefore $(\forall n \in \mathbb{N})$ $f(\alpha_\infty) \geq f(\lambda_\infty) \geq f(\lambda_n)$; consequently,

$$(111a) \qquad (\forall n \in \mathbb{N}) \quad \tfrac{\lambda_{n+1}}{\lambda_n}\big(\theta\lambda_n + 1 - \lambda_n\big) \leq f(\alpha_\infty)\tfrac{\lambda_{n+1}}{\lambda_n}\sqrt{\lambda_n^2 + (1-\lambda_n)^2 + 2\theta\lambda_n(1-\lambda_n)}.$$

Similarly,

$$(111b) \qquad (\forall n \in \mathbb{N}) \quad \tfrac{\mu_{n+1}}{\mu_n}\big(\theta\mu_n + 1 - \mu_n\big) \leq f(\alpha_\infty)\tfrac{\mu_{n+1}}{\mu_n}\sqrt{\mu_n^2 + (1-\mu_n)^2 + 2\theta\mu_n(1-\mu_n)}.$$

On the other hand, $\hat{\kappa}$ defined in (92) becomes

$$(112a) \qquad \hat{\kappa} = \sup_{n\in\mathbb{N}} \Big\{ \tfrac{\lambda_{n+1}}{\lambda_n}\big(\theta\lambda_n + 1 - \lambda_n\big), \tfrac{\mu_{n+1}}{\mu_n}\big(\theta\mu_n + 1 - \mu_n\big) \Big\}$$

while $\hat{\rho}$ defined by (52) satisfies

$$(112b) \qquad \hat{\rho} = \sup_{n\in\mathbb{N}} \left\{ \begin{array}{l} \tfrac{\lambda_{n+1}}{\lambda_n}\sqrt{\lambda_n^2 + (1-\lambda_n)^2 + 2\theta\lambda_n(1-\lambda_n)}, \\ \tfrac{\mu_{n+1}}{\mu_n}\sqrt{\mu_n^2 + (1-\mu_n)^2 + 2\theta\mu_n(1-\mu_n)} \end{array} \right\}.$$

Altogether,

$$(113) \qquad \hat{\kappa} \leq f(\alpha_\infty)\hat{\rho} < \hat{\rho}.$$

Therefore, the rate $\hat{\kappa}$ is always better than the rate $\hat{\rho}$.

Remark 6.6 (best bound for the convergence rate) In Theorem 6.4, the linear rate is bounded above by the constant $\hat{\kappa}$ defined in (92). Again, the actual computation of $\hat{\kappa}$ seems to be hard in general; however, the upper bound in Lemma 6.3 is minimized when $(\forall n \in \mathbb{N})$ $\lambda_n = \mu_n = 1$, in which case not only

$$(114) \qquad \hat{\kappa}^2 = (\theta + 2\varepsilon)^2$$

but we also obtain a better rate from Theorem 6.4, namely $\hat{\kappa}^2$! Comparing to the best bound derived in Remark 5.8, we note that for fixed $\theta \in [0,1[$ and for all $\varepsilon > 0$ sufficiently small

$$(115) \qquad (\theta + 2\varepsilon)^2 < \theta + 2\varepsilon < \sqrt{\theta} < \sqrt{\frac{1+\theta}{2}}.$$

Thus, when $\theta \to 1^-$, we expect the linear rate of convergence for the MARP to approach that of the unrelaxed MAP.

7 MARP with linearly vanishing relaxation parameters

In this section, we consider the $(\boldsymbol{\lambda}, \boldsymbol{\mu})$-MARP sequences with linearly vanishing relaxation parameters; specifically, we assume that

$$(116) \qquad \eta := \sup_{n\in\mathbb{N}} \left\{ \frac{\lambda_{n+1}}{\lambda_n}, \frac{\mu_{n+1}}{\mu_n} \right\} < 1.$$

A concrete instance occurs when $(\forall n \in \mathbb{N})$ $\lambda_n = \lambda_0\eta^n$ and $\mu_n = \mu_0\eta^n$.

The following result guarantees that the MARP sequences *always converge linearly and globally without any assumption on regularity or CQ-type conditions whatsoever.*

Theorem 7.1 *The MARP sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge linearly to some point $\bar{c} \in X$ with rate η; moreover,*

$$(\forall n \in \mathbb{N}) \quad \max\big\{\|x_n - \bar{c}\|, \|y_n - \bar{c})\|\big\} \leq \tfrac{M(1+\eta)}{1-\eta} \cdot \eta^n, \tag{117}$$

where $M := \max\{\|y_0 - x_0\|, \|x_0 - y_{-1}\|\}$, and

$$\|\bar{c} - y_{-1}\| \leq \tfrac{2\alpha_0(1+\alpha_0)}{1-\eta} \max\big\{d_A(y_{-1}), d_B(y_{-1})\big\}. \tag{118}$$

Proof. Let $n \in \mathbb{N}$. Clearly, $\langle y_n - x_n, y_{n-1} - x_n\rangle \leq 1 \cdot \|y_n - x_n\| \cdot \|x_n - y_{n-1}\|$ because of Cauchy-Schwarz. Lemma 3.8 (applied with $\theta = 1$) yields

$$\|x_{n+1} - y_n\| \leq \tfrac{\lambda_{n+1}}{\lambda_n} \max\big\{\|y_n - x_n\|, \|x_n - y_{n-1}\|\big\} \tag{119a}$$

$$\leq \eta \max\big\{\|y_n - x_n\|, \|x_n - y_{n-1}\|\big\}. \tag{119b}$$

On the other hand, by using Lemma 3.9, we similarly obtain

$$\|y_{n+1} - x_{n+1}\| \leq \eta \max\big\{\|x_{n+1} - y_n\|, \|y_n - x_n\|\big\}. \tag{120}$$

Altogether,

$$\max\big\{\|y_{n+1} - x_{n+1}\|, \|x_{n+1} - y_n\|\big\} \leq \eta \max\big\{\|y_n - x_n\|, \|x_n - y_{n-1}\|\big\} \tag{121a}$$

$$\vdots \tag{121b}$$

$$\leq \eta^{n+1} \max\big\{\|y_0 - x_0\|, \|x_0 - y_{-1}\|\big\} \tag{121c}$$

$$= M\eta^{n+1}. \tag{121d}$$

Thus

$$(\forall n \in \mathbb{N}) \quad \max\big\{\|y_n - x_n\|, \|x_n - y_{n-1}\|\big\} \leq M\eta^n \tag{122}$$

because (122) holds for $n = 0$ by the definition of M. Therefore, by Lemma 4.1, there exists a point $\bar{c} \in X$ such that

$$(\forall n \in \mathbb{N}) \quad \max\big\{\|x_n - \bar{c}\|, \|y_n - \bar{c})\|\big\} \leq \tfrac{M(1+\eta)}{1-\eta} \cdot \eta^n, \tag{123}$$

i.e., (117) holds. In particular, $\|x_0 - \bar{c}\| \leq \frac{M(1+\eta)}{1-\eta}$ and thus

$$\|\bar{c} - y_{-1}\| \leq \|x_0 - y_{-1}\| + \|x_0 - \bar{c}\| \leq M + \tfrac{M(1+\eta)}{1-\eta} = \tfrac{2M}{1-\eta}. \tag{124}$$

On the other hand, Lemma 3.7 yields $M \leq \alpha_0(1 + \alpha_0)\max\{d_A(y_{-1}), d_B(y_{-1})\}$. Altogether, we obtain (118). ■

Remark 7.2 It is interesting to compare the results of this section to some of the results of previous sections. On the one hand, Theorem 7.1 yields universal and global linear convergence; however, the location of the limit is not known to be in the intersection $A \cap B$. On the other hand, Theorem 5.11 and Theorem 6.4 guarantee linear convergence when a CQ condition or regularity holds, respectively; nevertheless, these results are only local. We appear to witness here an "uncertainty principle" which pits quality of convergence against location of the limit. It would be highly desirable to design hybrid methods that guarantee global convergence to a point in the intersection (or to prove that such an undertaking is hopeless).

8 Further Examples

Proposition 8.1 *Suppose that $X = \mathbb{R}$ and let $(a,b,c) \in \mathbb{R}_{++} \times \mathbb{R}_{++} \times \mathbb{R}_{--}$ satisfy*

$$\max\{a, b-2a\} < |c| = -c < \sqrt{a^2 + (b-a)^2} < b. \tag{125}$$

Suppose that $A = \{a, c\}$, that $B = \{b, c\}$, that

$$(\forall n \in \mathbb{N}) \quad \lambda_n = \mu_n = \lambda \in \left]\frac{a+c+\sqrt{c^2-a^2}}{2a}, \frac{b+c}{2a}\right[\subset \left]0,1\right[\tag{126}$$

and that $y_{-1} = 0$. Then the following hold:

(i) *The MAP cycles between a and b, and thus does not converge to a point in $A \cap B$.*

(ii) *The λ-MARP converges linearly to $c \in A \cap B$ with rate $1 - \lambda$.*

Proof. On the one hand, $c^2 < a^2 + (b-a)^2 \Leftrightarrow (b-a)^2 > c^2 - a^2 \Leftrightarrow b - a = |b-a| > \sqrt{c^2 - a^2} \Leftrightarrow$

$$b + c > a + c + \sqrt{c^2 - a^2} = \sqrt{|c| - a}\Big(\sqrt{|c| + a} - \sqrt{|c| - a}\Big) > 0. \tag{127}$$

On the other hand, $2a > b + c > 0$. Altogether, the interval from which λ is drawn is well defined and we have

$$b + c > 2a\lambda > a + c + \sqrt{c^2 - a^2}. \tag{128}$$

Since $a = |a| < |c|$, it follows that $P_A y_{-1} = P_A 0 = a$. Hence

$$x_0 = \lambda a. \tag{129}$$

Now $b - 2a < |c| \Leftrightarrow b - a < a + |c| \Leftrightarrow |b - a| < a - c \Leftrightarrow |b - a| < |c - a|$, so $P_B a = b$ and obviously $P_A b = a$. This proves (i).

Next, $P_B x_0 = c \Leftrightarrow |c - \lambda a| < |b - \lambda a| \Leftrightarrow \lambda a + |c| < b - \lambda a \Leftrightarrow 2\lambda a < b - |c| = b + c$. By the first inequality in (128), $P_B x_0 = c$ and thus $y_0 = (1-\lambda)x_0 + \lambda P_B x_0 = (1-\lambda)\lambda a + \lambda c$, i.e.,

$$y_0 = (1-\lambda)\lambda a + \lambda c. \tag{130}$$

We have $P_A y_0 = c$ if and only if $y_0 < (c + a)/2$, which is equivalent to $2(1-\lambda)\lambda a + 2\lambda c < a + c$. Viewed in terms of λ, this is a quadratic inequality which holds because of the second inequality in (128). It follows that $x_1 = (1-\lambda)y_0 + \lambda P_A y_0$, i.e.,

$$x_1 = (1-\lambda)\big((1-\lambda)\lambda a + \lambda c\big) + \lambda c. \tag{131}$$

Furthermore, $x_0 - x_1 = \lambda a(2-\lambda)(\lambda - c/a) > 0$ and so $x_1 < x_0$. Since already $P_B x_0 = c$, it follows that $P_B x_1 = c$ and therefore

$$y_1 = (1-\lambda)x_1 + \lambda c. \tag{132}$$

Thus, $(\forall n \in \{2, 3, \ldots\})$ $x_n = T y_{n-1}$ and $y_n = T x_n$, where $T\colon x \mapsto (1-\lambda)x + \lambda c$ is $(1-\lambda)$-Lipschitz continuous with unique fixed point c. This completes the proof of (ii). ■

Example 8.2 (Examples 1.1 and 1.2 revisited) Consider Proposition 8.1 with $c = -3 < a = 2 < b = 6$. Then $\max\{a, b-2a\} = \max\{2, 6-2\cdot 2\} = \max\{2,2\} = 2 < 3 = -c < 4.47 \approx \sqrt{20} = \sqrt{2^2+(6-2)^2} = \sqrt{a^2+(b-a)^2} < 6 = b$, and therefore (125) holds. We have $b+c = 3$, $2a = 4$, and $a + c + \sqrt{c^2-a^2} = -1 + \sqrt{9-4} = -1+\sqrt{5} \approx 1.23$. Hence (126) turns into

$$\frac{\sqrt{5}-1}{4} < \lambda < \frac{3}{4}; \tag{133}$$

note that since $(-1+\sqrt{5})/4 \approx 0.31$, we can choose in particular $\lambda = \frac{1}{2}$. By Proposition 8.1, the MAP with starting point 0 fails while the $\frac{1}{2}$-MARP converges linearly with rate $\frac{1}{2}$.

We have seen in the last example that MAP can fail to find a solution while MARP is able to solve the the problem. On the other hand, MAP can be faster than MARP:

Example 8.3 (MARP and nonsummable relaxation parameters) Suppose that $X = \mathbb{R}^2$, that $A = \mathbb{R}\times\{0\}$, and that $B = \{0\}\times\mathbb{R}$. Then $A\cap B = \{(0,0)\}$. On the one hand, regardless of the location of y_{-1}, the sequences for the $(1,1)$-MARP, i.e., MAP, satisfy $y_0 = x_1 = y_1 = \cdots = 0 \in A\cap B$ and thus convergence occurs in finitely many steps. On the other hand, let us now consider the MARP. Writing $y_{-1} = (\eta_1, \eta_2)$, one checks that for every $n \in \mathbb{N}$,

$$x_n = \big(\eta_1 \textstyle\prod_{i=0}^{n-1}(1-\mu_i), \eta_2 \prod_{i=0}^{n}(1-\lambda_i)\big) \tag{134a}$$

and

$$y_n = \big(\eta_1 \textstyle\prod_{i=0}^{n}(1-\mu_i), \eta_2 \prod_{i=0}^{n}(1-\lambda_i)\big). \tag{134b}$$

Thus if one of the relaxation parameters encountered is one, then we obtain finite convergence in the corresponding coordinate. So assume that $(\forall n\in\mathbb{N})\ \max\{\lambda_n,\mu_n\} < 1$, that $\eta_1 \neq 0$, and that $\eta_2 \neq 0$. If $\lambda_n \to 0$ and $\mu_n \to 0$, then (similarly to the discussion of Example 5.5 or see [4, Proposition 2.1]), we have the following characterizations:

(i) $(\lim_{n\in\mathbb{N}} x_n, \lim_{n\in\mathbb{N}} y_n) = (0,0) \Leftrightarrow \sum_{n\in\mathbb{N}}\lambda_n = \sum_{n\in\mathbb{N}}\mu_n = +\infty$.

(ii) $(\lim_{n\in\mathbb{N}} x_n, \lim_{n\in\mathbb{N}} y_n) \in A\smallsetminus\{(0,0)\} \Leftrightarrow \sum_{n\in\mathbb{N}}\lambda_n = +\infty, \sum_{n\in\mathbb{N}}\mu_n < +\infty$.

(iii) $(\lim_{n\in\mathbb{N}} x_n, \lim_{n\in\mathbb{N}} y_n) \in B\smallsetminus\{(0,0)\} \Leftrightarrow \sum_{n\in\mathbb{N}}\lambda_n < +\infty, \sum_{n\in\mathbb{N}}\mu_n = +\infty$.

(iv) $(\lim_{n\in\mathbb{N}} x_n, \lim_{n\in\mathbb{N}} y_n) \notin (A\cup B) \Leftrightarrow \sum_{n\in\mathbb{N}}\lambda_n < +\infty, \sum_{n\in\mathbb{N}}\mu_n < +\infty$.

This shows that when $\lambda_\infty = \mu_\infty = 0$, all possibilities for $\lim_{n\in\mathbb{N}}(x_n,y_n)$ occur. See also Examples 5.4, 5.5, and 5.6.

Remark 8.4 (convergence rates: actual vs upper bounds) In the previous sections, we have established upper bounds for the linear convergence rates. Let us now make some comments on the tightness of these estimates.

Consider the set up in Example 8.3 with $(\forall n\in\mathbb{N})\ \lambda_n = \mu_n = \lambda \in\ ]0,1]$. Then (134) yields the *actual rate*

$$\hat{\rho}_{\text{actual}} := 1-\lambda \in [0,1[\,. \tag{135}$$

Let us now turn to the estimates established earlier. On the one hand, since Corollary 5.9 holds with $\theta = 0$, we obtain from (52) that

$$\hat{\rho} = \sqrt{\lambda^2 + (1-\lambda)^2} > 1 - \lambda = \hat{\rho}_{\text{actual}}. \tag{136}$$

On the other hand, the assumptions of Theorem 6.4 are satisfied with $\varepsilon = 0$, $\delta = +\infty$, $S = X = \mathbb{R}^2$, and $\theta = 0$. Thus, the upper bound computed using (92) satisfies

$$\hat{\kappa} = 1 - \lambda = \hat{\rho}_{\text{actual}}. \tag{137}$$

9 A Doubly Non-Superregular Example

In this final section, we assume that $X = \mathbb{R}^2$. We shall construct A and B exhibiting various intriguing properties. We shall use tools from Euclidean geometry. Given $(s, u, v) \in X^3$, we denote the signed angle from the ray $\mathbb{R}_+ \times \{0\}$ to u by $\widehat{u}$; furthermore, $\widehat{usv} = \widehat{vsu}$ stands for the usual (nonsigned) angle at the point s.

9.1 The Set Up

We assume that

$$4w \in \left[0, \tfrac{\pi}{2}\right] \quad \text{and} \quad \cos(4w) = \tfrac{3}{4}, \tag{138}$$

so $w \approx 0.18$. Define

$$f\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto \begin{cases} 0, & \text{if } x \in \left]-\infty, 0\right] \cup \left]1, +\infty\right[; \\ (\tan w)\big(x - \frac{1}{2^k}\big), & \text{if } x \in \left]\frac{3}{2\cdot 2^{k+1}}, \frac{1}{2^k}\right] \text{ and } k \in \mathbb{N}; \\ -(\tan w)\big(x - \frac{1}{2^{k+1}}\big), & \text{if } x \in \left]\frac{1}{2^{k+1}}, \frac{3}{2\cdot 2^{k+1}}\right] \text{ and } k \in \mathbb{N}. \end{cases} \tag{139}$$

Moreover, denote by $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ the reflector with respect to the line $y = (\tan 2w)x$. Now we assume that (see Figure 2)

$$A = \big\{(x, y) \in \mathbb{R}^2 \bigm| y \le f(x)\big\}, \;\; B = \Phi(A), \;\; \text{and} \;\; c = (0, 0) \in A \cap B, \tag{140}$$

and we also set

$$(\forall k \in \mathbb{N}) \quad s_k := \big(\tfrac{1}{2^k}, 0\big) \;\; \text{and} \;\; z_k := \Phi(s_k). \tag{141}$$

9.2 The Normal Cones

Note that

$$(\forall a \in A)(\forall k \in \{1, 2, \ldots\}) \quad N_A(a) \subseteq N_A(s_k) := \big\{u \in \mathbb{R}^2 \bigm| \tfrac{\pi}{2} - w \le \widehat{u} \le \tfrac{\pi}{2} + w\big\}. \tag{142}$$

Let A' be the reflection of A about $\mathbb{R} \times \{0\}$. Then, since $\widehat{z_0 c s_0} = 4w$ and B is obtained by rotating A' by the angle $4w$ about the origin c, it follows that

$$(\forall a' \in A') \quad N_{A'}(a') \subseteq N_{A'}(s_k) = -N_A(s_k) = \big\{u \in \mathbb{R}^2 \bigm| -\tfrac{\pi}{2} - w \le \widehat{u} \le -\tfrac{\pi}{2} + w\big\} \tag{143}$$

and

$$(\forall b \in B) \quad N_B(b) \subseteq \big\{u \in \mathbb{R}^2 \bigm| -\tfrac{\pi}{2} + 3w \le \widehat{u} \le -\tfrac{\pi}{2} + 5w\big\}. \tag{144}$$

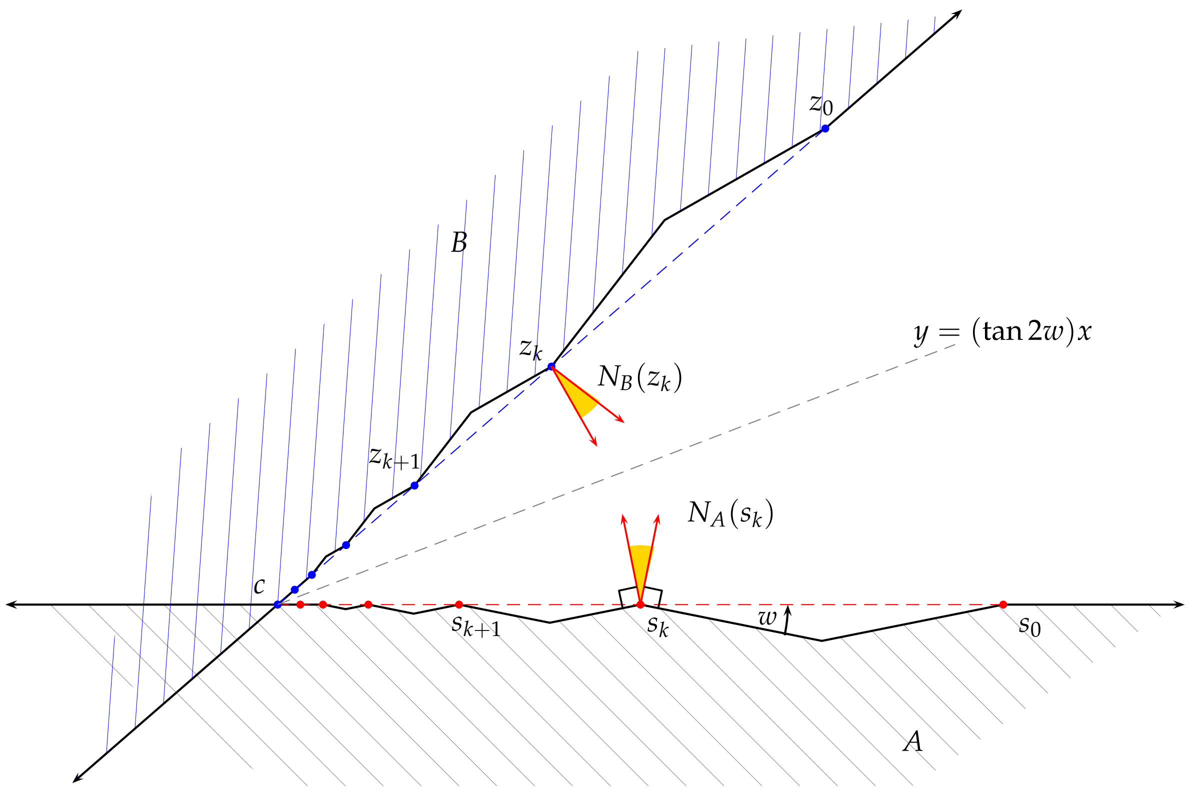

Figure 2: Non-superregular sets in $\mathbb{R}^2$

9.3 The CQ-number at c associated with $(A, \operatorname{bdry} A, B, \operatorname{bdry} B)$

Let $\delta > 0$. Then for every $k \in \{1, 2, \ldots\}$, the closed region W (see Figure 3) is a subset of $P_A^{-1}(s_k)$; thus,

$$\widehat{N}_A^{\operatorname{bdry} B}(s_k) = \operatorname{cone}\big((P_A^{-1}(s_k) \cap \operatorname{bdry} B) - s_k\big) = \operatorname{cone}(W - s_k) = N_A(s_k) \tag{145}$$

and

$$\widehat{N}_B^{\operatorname{bdry} A}(z_k) = N_B(z_k). \tag{146}$$

We now compute the CQ-number at c associated with $(A, \operatorname{bdry} A, B, \operatorname{bdry} B)$ and δ. Since for every $k \in \{1, 2, \ldots\}$, the normal cones $\widehat{N}_A^{\operatorname{bdry} B}(s_k)$ and $\widehat{N}_B^{\operatorname{bdry} A}(z_k)$ are the largest possible, it suffices in (9) to take the supremum over the points $B(c; \delta) \cap \{s_k \mid k \in \{1, 2, \ldots\}\}$ and $B(c; \delta) \cap \{z_k \mid k \in \{1, 2, \ldots\}\}$, respectively:

$$\theta_\delta = \sup\left\{ \langle u, v\rangle \;\middle|\; \begin{array}{l} u \in -\widehat{N}_A^{\operatorname{bdry} B}(s_k), v \in \widehat{N}_B^{\operatorname{bdry} A}(z_l), \|u\| \le 1, \|v\| \le 1, \\ \|s_k\| \le \delta, \|z_l\| \le \delta, (k, l) \in \{1, 2, \ldots\}^2 \end{array} \right\}. \tag{147}$$

It thus follows from (143) and (144) that

(148)

$$\begin{aligned} \theta_\delta &= \sup\left\{ \langle u, v\rangle \;\middle|\; \widehat{u} \in \left[-\tfrac{\pi}{2} - w, -\tfrac{\pi}{2} + w\right] \text{ and } \widehat{v} \in \left[-\tfrac{\pi}{2} + 3w, -\tfrac{\pi}{2} + 5w\right], \|u\| \le 1, \|v\| \le 1 \right\} \\ &= \cos 2w = \sqrt{\tfrac{1+\cos 4w}{2}} = \sqrt{\tfrac{7}{8}}. \end{aligned}$$

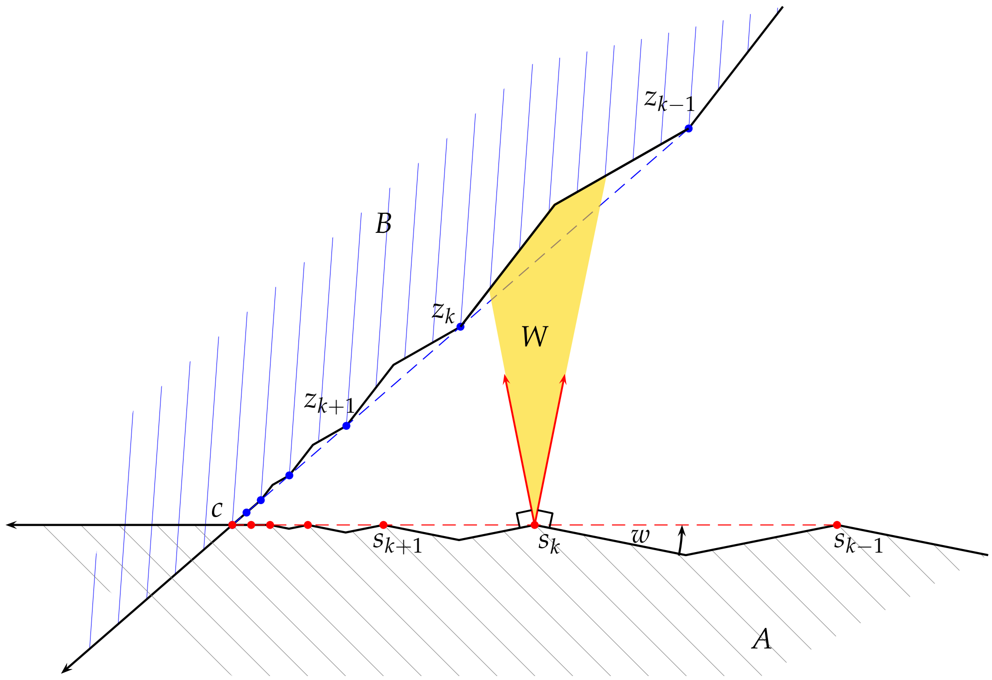

Figure 3: Inverse projections of s_k

Therefore,

$$(\forall \delta > 0) \quad 0.93 < \overline{\theta} = \theta_\delta = \sqrt{\tfrac{7}{8}} < 0.94. \tag{149}$$

9.4 A lower bound for ε in the (ε, δ)-regular case

Let $k \in \{1, 2, \ldots\}$, and set $d := \|s_{k+1} - c\|$. Then $\|z_k - c\| = \|s_k - c\| = 2d$. Now set $\beta_1 := \|z_k - s_{k+1}\|$ and $\beta_2 := \|z_k - s_k\|$. Noticing that $\widehat{z_k c s_k} = 4w$ and using the cosine theorem for the two triangles $\triangle c z_k s_{k+1}$ and $\triangle c z_k s_k$, we have

$$\beta_1^2 = d^2 + (2d)^2 - 2d(2d)(\cos 4w) = d^2 + 4d^2 - 4d^2(\tfrac{3}{4}) = 2d^2, \tag{150a}$$

$$\beta_2^2 = (2d)^2 + (2d)^2 - 2(2d)(2d)(\cos 4w) = 4d^2 + 4d^2 - 8d^2(\tfrac{3}{4}) = 2d^2. \tag{150b}$$

Hence, $\beta_1 = \beta_2 = d\sqrt{2}$. This also implies $P_A z_k \subseteq [s, s_k] \cup [s, s_{k+1}]$ (see Figure 4). The cosine theorem for the triangle $\triangle z_k s_k s_{k+1}$ gives

$$\cos \widehat{s_{k+1} z_k s_k} = \frac{\beta_1^2 + \beta_2^2 - d^2}{2\beta_1\beta_2} = \frac{2d^2 + 2d^2 - d^2}{2d\sqrt{2}d\sqrt{2}} = \frac{3}{4} > 0. \tag{151}$$

So we conclude that $\widehat{s_{k+1} z_k s_k} = 4w$. Next, since $\widehat{s s_{k+1} s_k} = \widehat{s_{k+1} s_k s} = w$, we have $\widehat{s_{k+1} s s_k} = \pi - 2w$. On the other hand,

$$\widehat{s s_k z_k} = w + \widehat{s_{k+1} s_k z_k} = w + \frac{\pi - 4w}{2} = \frac{\pi}{2} - w. \tag{152}$$

Altogether,

$$\widehat{s s_k z_k} = \widehat{s s_{k+1} z_k} = \widehat{s_k s z_k} = \widehat{s_{k+1} s z_k} = \frac{\pi}{2} - w, \tag{153}$$

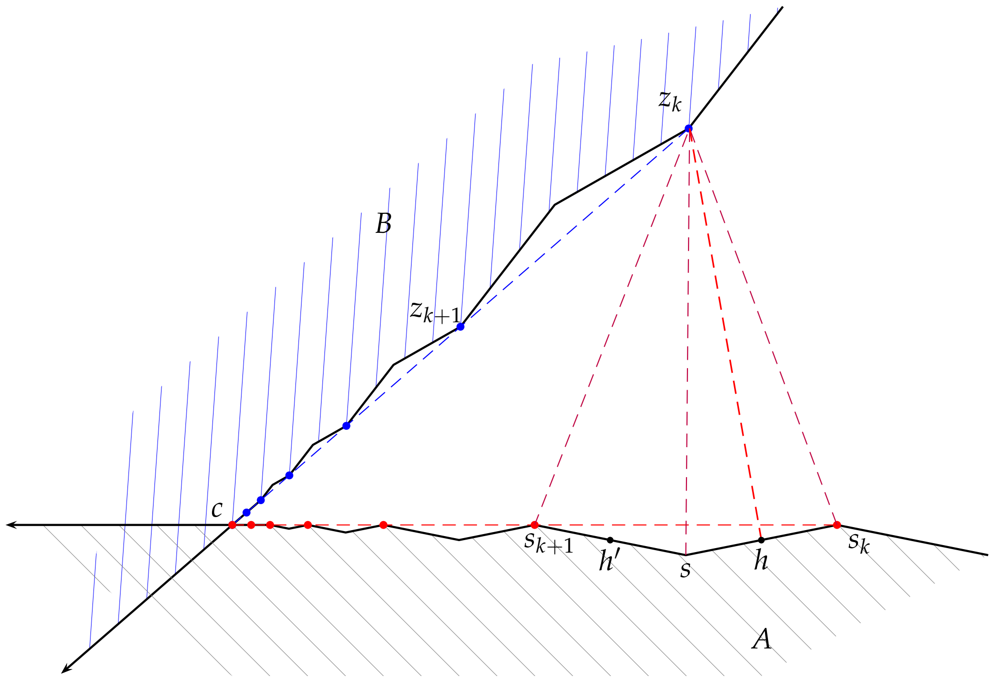

Figure 4: Projections of z_k on A

i.e., we have two isosceles triangles $\triangle z_k s s_{k+1}$ and $\triangle z_k s s_k$. Let h and h' be the two mid-points of $[s, s_k]$ and $[s, s_{k+1}]$. Then, $P_A z_k = \{h, h'\}$. Clearly $u := z_k - h \in \widehat{N}_A^{\mathrm{bdry}\, B}(h)$. Noticing that $\widehat{h'hz_k} = \widehat{shz_k} - \widehat{shh'} = \frac{\pi}{2} - w$, we have

$$\left\langle \frac{u}{\|u\|}, \frac{s_{k+1}-h}{\|s_{k+1}-h\|} \right\rangle = \cos \widehat{s_{k+1}hz_k} > \cos \widehat{h'hz_k} = \cos(\tfrac{\pi}{2} - w) = \sin w > 0.17; \tag{154}$$

consequently,

$$\langle u, s_{k+1} - h \rangle > (0.17) \cdot \|u\| \cdot \|s_{k+1} - h\|. \tag{155}$$

Now we assume that A is $(\mathrm{bdry}\, B, \varepsilon, \delta)$-regular at c for some $\varepsilon \geq 0$ and $\delta > 0$. Since $s_n \to c$ and $z_n \to c$, eventually all the points s_{k+1}, s_k, z_k, h', h lie in $\mathrm{ball}(c; \delta)$. From the above argument, we have $u \in \widehat{N}_A^{\mathrm{bdry}\, B}(h)$ and

$$(0.17) \cdot \|u\| \cdot \|s_{k+1} - h\| < \langle u, s_{k+1} - h \rangle \leq \varepsilon \cdot \|u\| \cdot \|s_{k+1} - h\|. \tag{156}$$

Thus

$$\varepsilon > 0.17. \tag{157}$$

Similarly, if B is $(\mathrm{bdry}\, A, \varepsilon, \delta)$-regular, then $\varepsilon > 0.17$.

9.5 For the MAP, [9, Proposition 3.12] is never applicable

Consider [9, Proposition 3.12] with $(\widetilde{A}, \widetilde{B}) = (\mathrm{bdry}\, A, \mathrm{bdry}\, B)$ and $I = J$ singletons. Clearly [9, (51)] holds. The two assumptions of [9, Proposition 3.12] are the following:

(i) A is $(\operatorname{bdry} B, \varepsilon, 3\delta)$-regular.

(ii) The CQ-number $\theta_{3\delta}$ at c associated with $(A, \operatorname{bdry} A, B, \operatorname{bdry} B)$ and 3δ satisfies $\theta_{3\delta} < 1 - 2\varepsilon$.

Assume that (i) holds. On the one hand, $\theta_{3\delta} > 0.93$ by (149). On the other hand, $\varepsilon > 0.17$ by (157). If (ii) holds, then we obtain the absurdity $0.93 < \theta_{3\delta} < 1 - 2\varepsilon < 0.66$. We conclude that (i) and (ii) cannot hold concurrently, which implies that [9, Proposition 3.12] is not applicable.

9.6 For the MAP, [9, Theorems 3.14 and 3.17] are never applicable

In view of (157), we note that A is not $(\operatorname{bdry} B)$-superregular at c, and that B is not $(\operatorname{bdry} A)$-superregular at c. Therefore, the results in [23] are not applicable, and neither are [9, Theorems 3.14 and 3.17] with $(\widetilde{A}, \widetilde{B}) = (\operatorname{bdry} A, \operatorname{bdry} B)$ and I and J singletons.

9.7 For the MARP, we deduce convergence with a linear rate

Indeed, suppose that $S = X$. The (A, X, B, X)-CQ condition holds, and so does the $(A, \operatorname{bdry} A, B, \operatorname{bdry} B)$-CQ condition. Hence, Theorem 5.11 applies and yields local convergence for the MARP sequences. Moreover, by (149), we can make ε in (73), and hence $\delta = +\infty$, arbitrarily large. Thus the MARP converges with a linear rate regardless of the starting point. Note that Corollary 5.9 also yields the global convergence result.

The figures suggest that the sequences generated by the MAP also converge with a linear rate. It would be interesting either to find theorems that allow for this conclusion or to at least obtain a partition of A and B so that the results of [9] are applicable to the induced collections $(\widetilde{A}, \widetilde{B})$.

Acknowledgments

The authors are grateful to two referees for their pertinent and constructive comments. HHB was partially supported by a Discovery Grant and an Accelerator Supplement of the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chair Program. HMP was supported by a PIMS postdoctoral fellowship, University of British Columbia Okanagan internal grant, and University of Victoria internal grant. XW was partially supported by a Discovery Grant of the Natural Sciences and Engineering Research Council of Canada.

References

[1] R. AHARONI AND Y. CENSOR, Block-iterative projection methods for parallel computation of solutions to convex feasibility problems, *Linear Algebra and its Applications* **120** (1989), 165–175.

[2] J.B. BAILLON, R.E. BRUCK, AND S. REICH, On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces, *Houston Journal of Mathematics* **4** (1978), 1–9.

[3] J.-B. BAILLON, P.L. COMBETTES, AND R. COMINETTI, Asymptotic behaviour of compositions of under-relaxed nonexpansive operators, arXiv preprint 1304.7078 (April 2013) URL: `http://arxiv.org/abs/1304.7078`

[4] H.H. BAUSCHKE, The approximation of fixed points of compositions of nonexpansive mappings in Hilbert space, *Journal of Mathematical Analysis and Applications* **202** (1996), 150–159.

[5] H.H. Bauschke and J.M. Borwein, On the convergence of von Neumann's alternating projection algorithm for two sets, *Set-Valued Analysis* **1(2)** (1993), 185–212.

[6] H.H. Bauschke and J.M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Review* **38(3)** (1996), 367–426.

[7] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer (2011).

[8] H.H. Bauschke, D.R. Luke, H.M. Phan, and X. Wang, Restricted normal cones and the method of alternating projections: theory, *Set-Valued and Variational Analysis* **21** (2013), 431–473.

[9] H.H. Bauschke, D.R. Luke, H.M. Phan, and X. Wang, Restricted normal cones and the method of alternating projections: applications, *Set-Valued and Variational Analysis* **21** (2013), 475–501.

[10] J.M. Borwein and J.D. Vanderwerff, *Convex Functions*, Cambridge University Press (2010).

[11] R.E. Bruck and S. Reich, Nonexpansive projections and resolvents of accretive operators in Banach spaces, *Houston Journal of Mathematics* **3** (1977), 459–470.

[12] A. Cegielski, *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, Springer (2012).

[13] Y. Censor and S.A. Zenios, *Parallel Optimization*, Oxford University Press (1997).

[14] P.L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization* **53** (2004), 475–504.

[15] P.L. Combettes and H.J. Trussell, Method of successive projections for finding a common point of sets in metric spaces, *Journal of Optimization Theory and Applications* **67** (1990), 487–507.

[16] F. Deutsch, *Best Approximation in Inner Product Spaces*, Springer (2001).

[17] F. Deutsch and H. Hundal, The rate of convergence for the cyclic projections algorithm I: angles between convex sets, *Journal of Approximation Theory* **142** (2006), 36–55.

[18] F. Deutsch and H. Hundal, The rate of convergence for the cyclic projections algorithm II: norms of nonlinear operators, *Journal of Approximation Theory* **142** (2006), 56–82.

[19] F. Deutsch and H. Hundal, The rate of convergence for the cyclic projections algorithm III: regularity of convex sets, *Journal of Approximation Theory* **155** (2008), 155–184.

[20] K. Goebel and W.A. Kirk, *Topics in Metric Fixed Point Theory*, Cambridge University Press (1990).

[21] K. Goebel and S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, Marcel Dekker (1984).

[22] L.G. Gubin, B.T. Polyak, and E.V. Raik, The method of projections for finding the common point of convex sets, *USSR Computational Mathematics and Mathematical Physics* **7** (1967), 1–24.

[23] A.S. Lewis, D.R. Luke, and J. Malick, Local linear convergence for alternating and averaged nonconvex projections, *Foundations of Computational Mathematics* **9** (2009), 485–513.

[24] D.R. Luke, Finding best approximation pairs relative to a convex and prox-regular set in a Hilbert space, *SIAM Journal on Optimization* **19(2)** (2008), 714–739.

[25] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation I*, Springer-Verlag (2006).

[26] R.T. Rockafellar, *Convex Analysis*, Princeton University Press (1970).

[27] R.T. Rockafellar and R. J-B Wets, *Variational Analysis (corrected 3rd printing)*, Springer (2009).

[28] J. von Neumann, *Functional Operators Vol. II. The Geometry of Orthogonal Spaces*, Annals of Mathematical Studies #22, Princeton University Press (1950).

[29] X. Wang and H.H. Bauschke, Compositions and averages of two resolvents: relative geometry of fixed point sets and a partial answer to a question by C. Byrne, *Nonlinear Analysis* **20** (2012), 131–153.

[30] N. Wiener, On the factorization of matrices, *Commentarii Mathematici Helvetici* **29** (1955), 97–111.

[31] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific Publishing (2002).