Restricted normal cones and the method of alternating projections: applications

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

March 1, 2013

Abstract

The method of alternating projections (MAP) is a common method for solving feasibility problems. While employed traditionally to subspaces or to convex sets, little was known about the behavior of the MAP in the nonconvex case until 2009, when Lewis, Luke, and Malick derived local linear convergence results provided that a condition involving normal cones holds and at least one of the sets is superregular (a property less restrictive than convexity). However, their results failed to capture very simple classical convex instances such as two lines in a three-dimensional space.

In this paper, we extend and develop the Lewis-Luke-Malick framework so that not only any two linear subspaces but also any two closed convex sets whose relative interiors meet are covered. We also allow for sets that are more structured such as unions of convex sets. The key tool required is the restricted normal cone, which is a generalization of the classical Mordukhovich normal cone. Numerous examples are provided to illustrate the theory.

2010 Mathematics Subject Classification: Primary 65K10; Secondary 47H04, 49J52, 49M20, 49M37, 65K05, 90C26, 90C30.

Keywords: Convex set, Friedrichs angle, linear convergence, method of alternating projections, nonconvex set, normal cone, projection operator, restricted normal cone, superregularity.

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

[†]Institut für Numerische und Angewandte Mathematik, Universität Göttingen, Lotzestr. 16–18, 37083 Göttingen, Germany. E-mail: r.luke@math.uni-goettingen.de.

[‡]Department of Mathematics & 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.

1 Introduction

Throughout this paper, we assume that

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

(i.e., finite-dimensional real Hilbert space) with inner product $\langle\cdot,\cdot\rangle$, induced norm $\|\cdot\|$, and induced metric d.

Let A and B be nonempty closed subsets of X. We assume first that A and B are additionally *convex* and that $A \cap B \neq \varnothing$. In this case, the *projection operators* P_A and P_B (a.k.a. projectors or nearest point mappings) corresponding to A and B, respectively, are single-valued with full domain. In order to find a point in the intersection A and B, it is very natural to simply alternate the operators P_A and P_B resulting in the famous *method of alternating projections (MAP)*. Thus, given a starting point $b_{-1} \in X$, sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are generated as follows:

$$(\forall n \in \mathbb{N}) \qquad a_n := P_A b_{n-1}, \quad b_n := P_B a_n. \tag{2}$$

In the present consistent convex setting, both sequences have a common limit in $A \cap B$. Not surprisingly, because of its elegance and usefulness, the MAP has attracted many famous mathematicians, including John von Neumann [28] and Norbert Wiener [29] and it has been independently rediscovered repeatedly. It is out of scope of this article to review the history of the MAP, its many extensions, and its rich and convergence theory; the interested reader is referred to, e.g., [5], [9], [13], and the references therein.

Since X is finite-dimensional and A and B are closed, the convexity of A and B is actually not needed in order to guarantee existence of nearest points. This gives rise to *set-valued* projection operators which for convenience we also denote by P_A and P_B. Dropping the convexity assumption, the MAP now generates sequences via

$$(\forall n \in \mathbb{N}) \qquad a_n \in P_A b_{n-1}, \quad b_n \in P_B a_n. \tag{3}$$

This iteration is much less understood than its much older convex cousin. For instance, global convergence to a point in $A \cap B$ cannot be guaranteed anymore [11]. Nonetheless, the MAP is widely applied to applications in engineering and the physical sciences for finding a point in $A \cap B$ (see, e.g., [27]). Lewis, Luke, and Malick achieved a break-through result in 2009, when there are no normal vectors that are opposite and at least one of the sets is superregular (a property less restrictive than convexity). Their proof techniques were quite different from the well known convex approaches; in fact, the Mordukhovich normal cone was a central tool in their analysis. However, their results were not strong enough to handle well known convex and linear scenarios. For instance, the linear convergence of the MAP for two lines in $\mathbb{R}^3$ cannot be obtained in their framework.

The goal of this paper is to extend the results by Lewis, Luke and Malick to make them applicable in more general settings. Their theory is unified with classical convex convergence results. We even allow for sets *that are* unions *of superregular (or even convex) sets. The known optimal convergence rate for the MAP for two linear subspaces is also recovered.*

Our principal tool is the new *restricted normal cone*, which we carefully investigated in the companion paper [6]. In a parallel paper [7], we apply our results to the important problem of sparsity optimization with affine constraints.

The remainder of the paper is organized as follows. The theoretical machinery from variational analysis underlying our main results is reviewed in Section 2. We are then in a position to provide in Section 3 our main results dealing with the local linear convergence of the MAP.

Notation

The notation employed in this article is quite standard and follows largely [8], [24], [25], and [26]; these books also provide exhaustive information on variational analysis. The real numbers are $\mathbb{R}$, the integers are $\mathbb{Z}$, and $\mathbb{N} := \{z \in \mathbb{Z} \mid z \geq 0\}$. Further, $\mathbb{R}_+ := \{x \in \mathbb{R} \mid x \geq 0\}$, $\mathbb{R}_{++} := \{x \in \mathbb{R} \mid x > 0\}$ and $\mathbb{R}_-$ and $\mathbb{R}_{--}$ are defined analogously. Let R and S be subsets of X. Then the closure of S is $\overline{S}$, the interior of S is $\operatorname{int}(S)$, the boundary of S is $\operatorname{bdry}(S)$, and the smallest affine and linear subspaces containing S are $\operatorname{aff} S$ and $\operatorname{span} S$, respectively. The linear subspace parallel to $\operatorname{aff} S$ is $\operatorname{par} S := (\operatorname{aff} S) - S = (\operatorname{aff} S) - s$, for every $s \in S$. The relative interior of S, $\operatorname{ri}(S)$, is the interior of S relative to $\operatorname{aff}(S)$. The negative polar cone of S is $S^{\ominus} = \{u \in X \mid \sup \langle u, S\rangle \leq 0\}$. We also set $S^{\oplus} := -S^{\ominus}$ and $S^{\perp} := S^{\oplus} \cap S^{\ominus}$. We also write $R \oplus S$ for $R + S := \{r + s \mid (r,s) \in R \times S\}$ provided that $R \perp S$, i.e., $(\forall (r,s) \in R \times S)\ \langle r,s\rangle = 0$. We write $F\colon X \rightrightarrows X$, if F is a mapping from X to its power set, i.e., $\operatorname{gr} F$, the graph of F, lies in $X \times X$. Abusing notation slightly, we will write $F(x) = y$ if $F(x) = \{y\}$. A nonempty subset K of X is a cone if $(\forall \lambda \in \mathbb{R}_+)\ \lambda K := \{\lambda k \mid k \in K\} \subseteq K$. The smallest cone containing S is denoted $\operatorname{cone}(S)$; thus, $\operatorname{cone}(S) := \mathbb{R}_+ \cdot S := \{\rho s \mid \rho \in \mathbb{R}_+, s \in S\}$ if $S \neq \varnothing$ and $\operatorname{cone}(\varnothing) := \{0\}$. The smallest convex and closed and convex subset containing S are $\operatorname{conv}(S)$ and $\overline{\operatorname{conv}}\,(S)$, respectively. If $z \in X$ and $\rho \in \mathbb{R}_{++}$, then $\operatorname{ball}(z;\rho) := \{x \in X \mid d(z,x) \leq \rho\}$ is the closed ball centered at z with radius ρ while $\operatorname{sphere}(z;\rho) := \{x \in X \mid d(z,x) = \rho\}$ is the (closed) sphere centered at z with radius ρ. If u and v are in X, then $[u,v] := \{(1-\lambda)u + \lambda v \mid \lambda \in [0,1]\}$ is the line segment connecting u and v.

2 Auxiliary theoretical results

In this section, we fix some basic notation used throughout this article. We also collect several auxiliary results from [6] that will be useful in the proof of the main results on the MAP.

Projections

Definition 2.1 (distance and projection) *Let A be a nonempty subset of X. Then*

$$d_A\colon X \to \mathbb{R}\colon x \mapsto \inf_{a\in A} d(x,a) \tag{4}$$

is the distance function *of the set A and*

$$P_A\colon X \rightrightarrows X\colon x \mapsto \big\{a \in A \;\big|\; d_A(x) = d(x,a)\big\} \tag{5}$$

is the corresponding projection.

The following result is well known.

Proposition 2.2 (existence) (See, e.g., [6, Proposition 1.2].) *Let A be a nonempty closed subset of X. Then $(\forall x \in X)$ $P_A(x) \neq \varnothing$.*

Example 2.3 (sphere) (See, e.g., [6, Example 1.4].) Let $z \in X$ and $\rho \in \mathbb{R}_{++}$. Set $S := \operatorname{sphere}(z;\rho)$. Then

$$(\forall x \in X) \quad P_S(x) = \begin{cases} z + \rho\frac{x-z}{\|x-z\|}, & \text{if } x \neq z;\\ S, & \text{otherwise.}\end{cases} \tag{6}$$

In view of Proposition 2.2, the next result is in particular applicable to the union of finitely many nonempty closed subsets of X.

Lemma 2.4 (union) *Let $(A_i)_{i\in I}$ be a collection of nonempty subsets of X, set $A := \bigcup_{i\in I} A_i$, let $x \in X$, and suppose that $a \in P_A(x)$. Then there exists $i \in I$ such that $a \in P_{A_i}(x)$.*

Proof. Indeed, since $a \in A$, there exists $i \in I$ such that $a \in A_i$. Then $d(x,a) = d_A(x) \leq d_{A_i}(x) \leq d(x,a)$. Hence $d(x,a) = d_{A_i}(x)$, as claimed. ■

The projection onto a nonempty closed convex set has very nice properties as we point out next.

Fact 2.5 (projection onto closed convex set) *Let C be a nonempty closed convex subset of X, and let x, y and p be in X. Then the following hold:*

(i) *$P_C(x)$ is a singleton.*

(ii) *$P_C(x) = p$ if and only if $p \in C$ and $\sup\langle C - p, x - p\rangle \leq 0$.*

(iii) $\|P_C(x) - P_C(y)\|^2 + \|(\operatorname{Id} - P_C)(x) - (\operatorname{Id} - P_C)(y)\|^2 \leq \|x - y\|^2$.

(iv) $\|P_C(x) - P_C(y)\| \leq \|x - y\|$.

Proof. (i)&(ii): [5, Theorem 3.14]. (iii): [5, Proposition 4.8]. (iv): Clear from (iii). ■

Restricted normal cones

Let us start by reviewing the definitions of various normal cones from variational analysis (see, e.g., [8], [10], [24], [25], and [26] for further information and applications).

Definition 2.6 (normal cones) (See also [6, Definition 2.1].) *Let A and B be nonempty subsets of X, and let a and u be in X. If $a \in A$, then various normal cones of A at a are defined as follows:*

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

$$\widehat{N}_A^B(a) := \operatorname{cone}\Big(\big(B \cap P_A^{-1}a\big) - a\Big) = \operatorname{cone}\Big(\big(B - a\big) \cap \big(P_A^{-1}a - a\big)\Big). \tag{7}$$

(ii) *The (classical)* proximal normal cone *of A at a is*

$$N_A^{\text{prox}}(a) := \widehat{N}_A^X(a) = \operatorname{cone}\big(P_A^{-1}a - a\big). \tag{8}$$

(iii) *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 $u_n \to u$.*

(iv) *The* Fréchet normal cone $N_A^{\text{Fré}}(a)$ *is implicitly defined by $u \in N_A^{\text{Fré}}(a)$ if and only if $(\forall \varepsilon > 0)$ $(\exists\, \delta > 0)$ $(\forall x \in A \cap \operatorname{ball}(a;\delta))$ $\langle u, x - a\rangle \le \varepsilon\|x - a\|$.*

(v) *The* normal convex from convex analysis $N_A^{\text{conv}}(a)$ *is implicitly defined by $u \in N_A^{\text{conv}}(a)$ if and only if* $\sup \langle u, A - a\rangle \le 0$.

(vi) *The* Mordukhovich normal cone $N_A(a)$ *of A at a is implicitly defined by $u \in N_A(a)$ if and only if there exist sequences $(a_n)_{n\in\mathbb{N}}$ in A and $(u_n)_{n\in\mathbb{N}}$ in $N_A^{\text{prox}}(a_n)$ such that $a_n \to a$ and $u_n \to u$.*

If $a \notin A$, then all normal cones are defined to be empty.

In the convex case, all unrestricted normal cones coincide:

Lemma 2.7 (convex case) (See, e.g., [6, Lemma 2.4(vii)].) *Let A be nonempty closed convex subset of X, and let $a \in A$. Then $\widehat{N}_A^X(a) = N_A^{\text{prox}}(a) = N_A^{\text{Fré}}(a) = N_A^{\text{conv}}(a) = N_A(a)$.*

In the following two results, we revisit classical constraint qualifications and provide characterizations in terms of normal cones.

Theorem 2.8 (two convex sets: restricted normal cones and relative interiors) (See [6, Theorem 3.13].) *Let A and B be nonempty convex subsets of X. Then the following are equivalent:*

(i) $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$.

(ii) $0 \in \operatorname{ri}(B - A)$.

(iii) $\operatorname{cone}(B-A) = \operatorname{span}(B-A)$.

(iv) $N_A(c) \cap (-N_B(c)) \cap \overline{\operatorname{cone}}(B-A) = \{0\}$ *for some* $c \in A \cap B$.

(v) $N_A(c) \cap (-N_B(c)) \cap \overline{\operatorname{cone}}(B-A) = \{0\}$ *for every* $c \in A \cap B$.

(vi) $N_A(c) \cap (-N_B(c)) \cap \operatorname{span}(B-A) = \{0\}$ *for some* $c \in A \cap B$.

(vii) $N_A(c) \cap (-N_B(c)) \cap \operatorname{span}(B-A) = \{0\}$ *for every* $c \in A \cap B$.

(viii) $N_A^{\operatorname{aff}(A\cup B)}(c) \cap (-N_B^{\operatorname{aff}(A\cup B)}(c)) = \{0\}$ *for some* $c \in A \cap B$.

(ix) $N_A^{\operatorname{aff}(A\cup B)}(c) \cap (-N_B^{\operatorname{aff}(A\cup B)}(c)) = \{0\}$ *for every* $c \in A \cap B$.

(x) $N_{A-B}^{\operatorname{span}(B-A)}(0) = \{0\}$.

Corollary 2.9 (two convex sets: normal cones and interiors) (See [6, Corollary 3.14].) *Let A and B be nonempty convex subsets of X. Then the following are equivalent:*

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

(ii) $\operatorname{cone}(B-A) = X$.

(iii) $N_A(c) \cap (-N_B(c)) = \{0\}$ *for some* $c \in A \cap B$.

(iv) $N_A(c) \cap (-N_B(c)) = \{0\}$ *for every* $c \in A \cap B$.

(v) $N_{A-B}(0) = \{0\}$.

CQ and joint-CQ numbers

The notions of CQ and joint-CQ numbers can be viewed as quantifications of constraint qualifications.

Definition 2.10 ((joint) CQ-number) (See [6, Definition 6.1 and Definition 6.2].) *Let A, $\widetilde{A}$, B, $\widetilde{B}$, 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*

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

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

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

For nontrivial collections[1] $\mathcal{A} := (A_i)_{i\in I}$, $\widetilde{\mathcal{A}} := (\widetilde{A}_i)_{i\in I}$, $\mathcal{B} := (B_j)_{j\in J}$, $\widetilde{\mathcal{B}} := (\widetilde{B}_j)_{j\in J}$ *of nonempty subsets of* X*, the* joint-CQ-number *at* $c \in X$ *associated with* $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ *and* $\delta > 0$ *is*

$$\theta_\delta = \theta_\delta\big(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}}\big) := \sup_{(i,j)\in I\times J} \theta_\delta\big(A_i, \widetilde{A}_i, B_j, \widetilde{B}_j\big), \tag{11}$$

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

$$\overline{\theta} = \overline{\theta}\big(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}}\big) := \lim_{\delta\downarrow 0} \theta_\delta\big(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}}\big). \tag{12}$$

The CQ-number is obviously an instance of the joint-CQ-number when I and J are singletons. When the arguments are clear from the context we will simply write θ_δ and $\overline{\theta}$.

Clearly,

$$\theta_\delta\big(A, \widetilde{A}, B, \widetilde{B}\big) = \theta_\delta\big(B, \widetilde{B}, A, \widetilde{A}\big) \quad \text{and} \quad \overline{\theta}\big(A, \widetilde{A}, B, \widetilde{B}\big) = \overline{\theta}\big(B, \widetilde{B}, A, \widetilde{A}\big). \tag{13}$$

Note that, $\delta \mapsto \theta_\delta$ is increasing; this makes $\overline{\theta}$ well defined. Furthermore, since 0 belongs to nonempty B-restricted proximal normal cones and because of the Cauchy-Schwarz inequality, we have

$$c \in \overline{A} \cap \overline{B} \text{ and } 0 < \delta_1 < \delta_2 \quad \Rightarrow \quad 0 \le \overline{\theta} \le \theta_{\delta_1} \le \theta_{\delta_2} \le 1, \tag{14}$$

while θ_δ, and hence $\overline{\theta}$, is equal to $-\infty$ if $c \notin \overline{A} \cap \overline{B}$ and δ is sufficiently small (using the fact that $\sup \varnothing = -\infty$).

Example 2.11 (joint-CQ-number < CQ-number of the unions) (See [6, Example 6.4].) Suppose that $X = \mathbb{R}^3$, let $I := J := \{1, 2\}$, $A_1 := \mathbb{R}(0, 1, 0)$, $A_2 := \mathbb{R}(2, 0, -1)$, $B_1 := \mathbb{R}(0, 1, 1)$, $B_2 := \mathbb{R}(1, 0, 0)$, $c := (0, 0, 0)$, and let $\delta > 0$. Furthermore, set $\mathcal{A} := (A_i)_{i\in I}$, $\mathcal{B} := (B_j)_{j\in J}$, $A := A_1 \cup A_2$, and $B := B_1 \cup B_2$. Then

$$\theta_\delta\big(\mathcal{A}, \mathcal{A}, \mathcal{B}, \mathcal{B}\big) = \tfrac{2}{\sqrt{5}} < 1 = \theta_\delta\big(A, A, B, B\big). \tag{15}$$

CQ and joint-CQ conditions

The notions of CQ and joint-CQ conditions are complementary to those of CQ and joint-CQ numbers — while the former build on restricted proximal normals in a neighbourhood of a point of interest, the latter rest on the restricted normal cone at a point.

Definition 2.12 (CQ and joint-CQ conditions) (See [6, Definition 6.6].) *Let* $c \in X$*.*

(i) *Let* A, $\widetilde{A}$, B *and* $\widetilde{B}$ *be nonempty subsets of* X*. Then the* $(A, \widetilde{A}, B, \widetilde{B})$-CQ condition *holds at* c *if*

$$N_A^{\widetilde{B}}(c) \cap \big(-N_B^{\widetilde{A}}(c)\big) \subseteq \{0\}. \tag{16}$$

[1] The collection $(A_i)_{i\in I}$ is said to be *nontrivial* if $I \neq \varnothing$.

(ii) *Let* $\mathcal{A} := (A_i)_{i\in I}$, $\widetilde{\mathcal{A}} := (\widetilde{A}_i)_{i\in I}$, $\mathcal{B} := (B_j)_{j\in J}$ *and* $\widetilde{\mathcal{B}} := (\widetilde{B}_j)_{j\in J}$ *be nontrivial collections of nonempty subsets of X. Then the* $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$-joint-CQ condition *holds at c if for every* $(i,j) \in I \times J$, *the* $(A_i, \widetilde{A}_i, B_j, \widetilde{B}_j)$*-CQ condition holds at c, i.e.,*

$$\big(\forall (i,j) \in I \times J\big) \quad N_{A_i}^{\widetilde{B}_j}(c) \cap \big(-N_{B_j}^{\widetilde{A}_i}(c)\big) \subseteq \{0\}. \tag{17}$$

Definition 2.13 (exact CQ-number and exact joint-CQ-number) (See [6, Definition 6.7].) *Let* $c \in X$.

(i) *Let* A, $\widetilde{A}$, B *and* $\widetilde{B}$ *be nonempty subsets of X. The* exact CQ-number *at c associated with* $(A, \widetilde{A}, B, \widetilde{B})$ *is* [2]

$$\overline{\alpha} := \overline{\alpha}\big(A, \widetilde{A}, B, \widetilde{B}\big) := \sup\left\{ \langle u, v\rangle \;\middle|\; u \in N_A^{\widetilde{B}}(c), v \in -N_B^{\widetilde{A}}(c), \|u\| \le 1, \|v\| \le 1 \right\}. \tag{18}$$

(ii) *Let* $\mathcal{A} := (A_i)_{i\in I}$, $\widetilde{\mathcal{A}} := (\widetilde{A}_i)_{i\in I}$, $\mathcal{B} := (B_j)_{j\in J}$ *and* $\widetilde{\mathcal{B}} := (\widetilde{B}_j)_{j\in J}$ *be nontrivial collections of nonempty subsets of X. The* exact joint-CQ-number *at c associated with* $(\mathcal{A}, \mathcal{B}, \widetilde{\mathcal{A}}, \widetilde{\mathcal{B}})$ *is*

$$\overline{\alpha} := \overline{\alpha}(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}}) := \sup_{(i,j)\in I\times J} \overline{\alpha}(A_i, \widetilde{A}_i, B_j, \widetilde{B}_j). \tag{19}$$

The next result relates the various condition numbers defined above.

Theorem 2.14 (See [6, Theorem 6.8].) *Let* $\mathcal{A} := (A_i)_{i\in I}$, $\widetilde{\mathcal{A}} := (\widetilde{A}_i)_{i\in I}$, $\mathcal{B} := (B_j)_{j\in J}$ *and* $\widetilde{\mathcal{B}} := (\widetilde{B}_j)_{j\in J}$ *be nontrivial collections of nonempty subsets of X. Set* $A := \bigcup_{i\in I} A_i$ *and* $B := \bigcup_{j\in J} B_j$, *and suppose that* $c \in A \cap B$. *Denote the exact joint-CQ-number at c associated with* $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ *by* $\overline{\alpha}$ *(see* (19)*), the joint-CQ-number at c associated with* $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ *and* $\delta > 0$ *by* θ_δ *(see* (11)*), and the limiting joint-CQ-number at c associated with* $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ *by* $\overline{\theta}$ *(see* (12)*). Then the following hold:*

(i) *If* $\overline{\alpha} < 1$, *then the* $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$*-CQ condition holds at c.*

(ii) $\overline{\alpha} \le \theta_\delta$.

(iii) $\overline{\alpha} \le \overline{\theta}$.

Now assume in addition that I and J are finite. Then the following hold:

(iv) $\overline{\alpha} = \overline{\theta}$.

(v) *The* $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$*-joint-CQ condition holds at c if and only if* $\overline{\alpha} = \overline{\theta} < 1$.

[2]Note that if $c \notin A \cap B$, then $\overline{\alpha} = \sup \varnothing = -\infty$.

Examples

Example 2.15 (CQ-number quantifies CQ condition) (See [6, Example 7.2].) Let A and B be subsets of X, and suppose that $c \in A \cap B$. Let L be an affine subspace of X containing $A \cup B$. Then the following are equivalent:

(i) $N_A^L(c) \cap (-N_B^L(c)) = \{0\}$, i.e., the (A, L, B, L)-CQ condition holds at c (see (16)).

(ii) $N_A(c) \cap (-N_B(c)) \cap (L - c) = \{0\}$.

(iii) $\overline{\theta} < 1$, where $\overline{\theta}$ is the limiting CQ-number at c associated with (A, L, B, L) (see (10)).

Example 2.16 (CQ condition depends on restricting sets) (See [6, Example 7.3].) Suppose that $X = \mathbb{R}^2$, and set $A := \operatorname{epi}(|\cdot|)$, $B := \mathbb{R} \times \{0\}$, and $c := (0, 0)$. Then we readily verify that $N_A(c) = N_A^X(c) = -A$, $N_A^B(c) = -\operatorname{bdry} A$, $N_B(c) = N_B^X(c) = \{0\} \times \mathbb{R}$, and $N_B^A(c) = \{0\} \times \mathbb{R}_+$. Consequently,

$$N_A^X(c) \cap \big(-N_B^X(c)\big) = \{0\} \times \mathbb{R}_- \ \text{ while } \ N_A^B(c) \cap \big(-N_B^A(c)\big) = \{(0,0)\}. \tag{20}$$

Therefore, the (A, A, B, B)-CQ condition holds, yet the (A, X, B, X)-CQ condition fails.

The case of two spheres is very pleasant because the quantities can be computed explicitly:

Proposition 2.17 (CQ-numbers of two spheres) (See [6, Example 7.4].) *Let z_1 and z_2 be in X, let ρ_1 and ρ_2 be in $\mathbb{R}_{++}$, set $S_1 := \operatorname{sphere}(z_1; \rho_1)$ and $S_2 := \operatorname{sphere}(z_2; \rho_2)$ and assume that $c \in S_1 \cap S_2$. Denote the limiting CQ-number at c associated with (S_1, X, S_2, X) by $\overline{\theta}$ (see Definition 2.10), and the exact CQ-number at c associated with (S_1, X, S_2, X) by $\overline{\alpha}$ (see Definition 2.13). Then the following hold:*

(i) $\overline{\theta} = \overline{\alpha} = \dfrac{|\langle z_1 - c, z_2 - c\rangle|}{\rho_1\rho_2}$.

(ii) $\overline{\alpha} < 1$ *unless the spheres are identical or intersect only at c.*

Now assume that $\overline{\alpha} < 1$, let $\varepsilon \in \mathbb{R}_{++}$, and set $\delta := \big(\sqrt{(\rho_1 + \rho_2)^2 + 4\rho_1\rho_2\varepsilon} - (\rho_1 + \rho_2)\big)/2 > 0$. Then

$$\overline{\alpha} \le \theta_\delta \le \overline{\alpha} + \varepsilon, \tag{21}$$

where θ_δ is the CQ-number at c associated with (S_1, X, S_2, X) (see Definition 2.10).

Let us revisit the classical constraint qualification for two convex sets.

Proposition 2.18 (See [6, Proposition 7.5].) *Let A and B be nonempty convex subsets of X such that $A \cap B \neq \varnothing$, and set $L = \operatorname{aff}(A \cup B)$. Then the following are equivalent:*

(i) $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$.

(ii) *The* (A, L, B, L)*-CQ condition holds at some point in* $A \cap B$*.*

(iii) *The* (A, L, B, L)*-CQ condition holds at every point in* $A \cap B$*.*

We now turn to two linear subspaces.

Definition 2.19 (angles between two subspaces) *Let* A *and* B *be linear subspaces of* X*.*

(i) **(Dixmier angle)** [17] *The* Dixmier angle *between* A *and* B *is the number in* $[0, \frac{\pi}{2}]$ *whose cosine is given by*

$$c_0(A, B) := \sup\big\{ |\langle a, b\rangle| \;\big|\; a \in A, b \in B, \|a\| \le 1, \|b\| \le 1 \big\}. \tag{22}$$

(ii) **(Friedrichs angle)** [18] *The* Friedrichs angle *(or simply the* angle*) between* A *and* B *is the number in* $[0, \frac{\pi}{2}]$ *whose cosine is given by*

$$c(A, B) := c_0(A \cap (A \cap B)^\perp, B \cap (A \cap B)^\perp) \tag{23a}$$

$$= \sup\left\{ |\langle a, b\rangle| \;\middle|\; \begin{array}{l} a \in A \cap (A\cap B)^\perp, \|a\| \le 1, \\ b \in B \cap (A\cap B)^\perp, \|b\| \le 1 \end{array} \right\}. \tag{23b}$$

Let us state a striking connection between the CQ-number and the Friedrichs angle.

Theorem 2.20 (CQ-number of two linear subspaces and Friedrichs angle) (See [6, Theorem 7.12].) *Let* A *and* B *be linear subspaces of* X*, and let* $\delta > 0$*. Then*

$$\theta_\delta(A, A, B, B) = \theta_\delta(A, X, B, B) = \theta_\delta(A, A, B, X) = c(A, B) < 1, \tag{24}$$

where the CQ-number at 0 *is defined as in* (9)*.*

Regularities

Regularity is a notion of a set that generalizes convexity. We shall also use restricted versions involving restricted normal cones.

Definition 2.21 (regularity and superregularity) (See [6, Definition 8.1].) *Let* A *and* B *be nonempty subsets of* X*, and let* $c \in X$*.*

(i) *We say that* B *is* (A, ε, δ)-regular *at* $c \in X$ *if* $\varepsilon \ge 0$*,* $\delta > 0$*, and*

$$\left.\begin{array}{c} (y, b) \in B \times B, \\ \|y - c\| \le \delta, \|b - c\| \le \delta, \\ u \in \widehat{N}_B^A(b) \end{array}\right\} \quad \Rightarrow \quad \langle u, y - b\rangle \le \varepsilon \|u\| \cdot \|y - b\|. \tag{25}$$

If B *is* (X, ε, δ)*-regular at* c*, then we also simply speak of* (ε, δ)*-regularity.*

(ii) *The set B is called A-*superregular *at $c \in X$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that B is (A, ε, δ)-regular at c. Again, if B is X-superregular at c, then we also say that B is superregular at c.*

Remark 2.22 (See [6, Remark 8.2].) Several comments on Definition 2.21 are in order.

(i) Superregularity with $A = X$ was introduced by Lewis, Luke and Malick in [20, Section 4]. Among other things, they point out that amenability and prox regularity are sufficient conditions for superregularity, while Clarke regularity is a necessary condition.

(ii) The reference point c does not have to belong to B. If $c \notin \overline{B}$, then for every $\delta \in \left]0, d_B(c)\right[$, B is $(0, \delta)$-regular at c; consequently, B is superregular at c.

(iii) If $\varepsilon \in \left[1, +\infty\right[$, then Cauchy-Schwarz implies that B is $(\varepsilon, +\infty)$-regular at every point in X.

(iv) Note that B is $(A_1 \cup A_2, \varepsilon, \delta)$-regular at c if and only if B is both $(A_1, \varepsilon, \delta)$-regular and $(A_2, \varepsilon, \delta)$-regular at c.

(v) If B is convex, then it follows with Lemma 2.7 that B is $(A, 0, +\infty)$-regular at c; consequently, B is superregular.

(vi) Similarly, if B is locally convex at c, i.e., there exists $\rho \in \mathbb{R}_{++}$ such that $B \cap \operatorname{ball}(c; \rho)$ is convex, then B is superregular at c.

(vii) If B is $(A, 0, \delta)$-regular at c, then B is A-superregular at c; the converse, however, is not true in general (see Example 2.23 below).

Example 2.23 (sphere) (See [6, Example 8.3].) Let $z \in X$ and $\rho \in \mathbb{R}_{++}$. Set $S := \operatorname{sphere}(z; \rho)$, suppose that $s \in S$, let $\varepsilon \in \mathbb{R}_{++}$, and let $\delta \in \mathbb{R}_{++}$. Then S is $(\varepsilon, \rho\varepsilon)$-regular at s; consequently, S is superregular at s (see Definition 2.21). However, S is not $(0, \delta)$-regular at s.

The notion of joint-regularity is critical in our analysis of the MAP below.

Definition 2.24 (joint-regularity) (See [6, Definition 8.6].) *Let A be a nonempty subset of X, let $\mathcal{B} := (B_j)_{j \in J}$ be a nontrivial collection of nonempty subsets of X, and let $c \in X$.*

(i) *We say that $\mathcal{B}$ is (A, ε, δ)-joint-regular at c if $\varepsilon \geq 0$, $\delta > 0$, and for every $j \in J$, B_j is (A, ε, δ)-regular at c.*

(ii) *The collection $\mathcal{B}$ is A-joint-superregular at c if for every $j \in J$, B_j is A-superregular at c.*

As in Definition 2.21, we may omit the prefix A if $A = X$.

In the convex case, we note that all regularity notions hold.

Corollary 2.25 (convexity and regularity) (See [6, Corollary 8.8].) *Let $\mathcal{B} := (B_j)_{j \in J}$ be a nontrivial collection of nonempty convex subsets of X, let $A \subseteq X$, and let $c \in X$. Then $\mathcal{B}$ is $(0, +\infty)$-joint-regular, $(A, 0, +\infty)$-joint-regular, joint-superregular, and A-joint-superregular at c.*

Let us explicitly point out that these notions are about collections of sets rather than their unions.

Example 2.26 (two lines: joint-superregularity $\not\Rightarrow$ superregularity of the union) (See [6, Example 8.9].) Suppose that d_1 and d_2 are in sphere$(0;1)$. Set $B_1 := \mathbb{R}d_1$, $B_2 := \mathbb{R}d_2$, and $B := B_1 \cup B_2$, and assume that $B_1 \cap B_2 = \{0\}$. By Corollary 2.25, (B_1, B_2) is joint-superregular at 0. Let $\delta \in \mathbb{R}_{++}$, and set $b := \delta d_1$ and $y := \delta d_2$. Then $\|y - 0\| = \delta$, $\|b - 0\| = \delta$, and $0 < \|y - b\| = \delta\|d_2 - d_1\|$. Furthermore, $N_B(b) = \{d_1\}^\perp$. Note that there exists $v \in \{d_1\}^\perp$ such that $\langle v, d_2\rangle \neq 0$ (for otherwise $\{d_1\}^\perp \subseteq \{d_2\}^\perp \Rightarrow B_2 \subseteq B_1$, which is absurd). Hence there exists $u \in \{d_1\}^\perp = \{b\}^\perp = N_B(b)$ such that $\|u\| = 1$ and $\langle u, d_2\rangle > 0$. It follows that $\langle u, y - b\rangle = \langle u, y\rangle = \delta\langle u, d_2\rangle = \langle u, d_2\rangle \|u\|\|y - b\|/\|d_2 - d_1\|$. Therefore, B is not superregular at 0.

3 The method of alternating projections (MAP)

We now apply the machinery of restricted normal cones and associated results to derive linear convergence results.

On the composition of two projection operators

The method of alternating projections iterates projection operators. Thus, in the next few results, we focus on the outcome of a single iteration of the composition.

Lemma 3.1 *Let A and B be nonempty closed subsets of X. Then the following hold*[3]*:*

(i) $P_A(B \smallsetminus A) \subseteq \mathrm{bdry}_{\mathrm{aff}\, A\cup B} A \subseteq \mathrm{bdry}\, A$.

(ii) $P_B(A \smallsetminus B) \subseteq \mathrm{bdry}_{\mathrm{aff}\, A\cup B}(B) \subseteq \mathrm{bdry}\, B$.

(iii) *If $b \in B$ and $a \in P_A b$, then:*

$$a \in (\mathrm{bdry}\, A) \smallsetminus B \;\Leftrightarrow\; a \in A \smallsetminus B \;\Rightarrow\; b \in B \smallsetminus A \;\Rightarrow\; a \in \mathrm{bdry}\, A. \tag{26}$$

(iv) *If $a \in A$ and $b \in P_B a$, then:*

$$b \in (\mathrm{bdry}\, B) \smallsetminus A \;\Leftrightarrow\; b \in B \smallsetminus A \;\Rightarrow\; a \in A \smallsetminus B \;\Rightarrow\; b \in \mathrm{bdry}\, B. \tag{27}$$

Proof. (i): Take $b \in B \smallsetminus A$ and $a \in P_A b$. Assume to the contrary that there exists $\delta \in \mathbb{R}_{++}$ such that $\mathrm{aff}(A \cup B) \cap \mathrm{ball}(a;\delta) \subseteq A$. Without loss of generality, we assume that $\delta < \|b - a\|$. Then $\widetilde{a} := a + \delta(b - a)/\|b - a\| \in A$ and thus $d_A(b) \leq d(\widetilde{a}, b) < d(a, b) = d_A(b)$, which is absurd.

(ii): Interchange the roles of A and B in (i).

[3]We denote by $\mathrm{bdry}_{\mathrm{aff}\, A\cup B}(S)$ the boundary of $S \subseteq X$ with respect to $\mathrm{aff}(A \cup B)$.

(iii): If $a \in (\operatorname{bdry} A) \smallsetminus B$, then clearly $a \in A \smallsetminus B$. Now assume that $a \in A \smallsetminus B$. If $b \in A$, then $a \in P_A b = \{b\} \subseteq B$, which is absurd. Hence $b \in B \smallsetminus A$ and thus (i) implies that $a \in P_A(B \smallsetminus A) \subseteq \operatorname{bdry} A$.

(iv): Interchange the roles of A and B in (iii). ■

Lemma 3.2 *Let A and B be nonempty closed subsets of X, let $c \in X$, let $y \in B$, let $a \in P_A y$, let $b \in P_B a$, and let $\delta \in \mathbb{R}_+$. Assume that $d_A(y) \leq \delta$ and that $d(y, c) \leq \delta$. Then the following hold:*

(i) $d(a, c) \leq 2\delta$.

(ii) $d(b, y) \leq 2d(a, y) \leq 2\delta$.

(iii) $d(b, c) \leq 3\delta$.

Proof. Since $y \in B$, we have

(28) $$d(a, b) = d_B(a) \leq d(a, y) = d_A(y) \leq \delta.$$

Thus,

(29) $$d(a, c) \leq d(a, y) + d(y, c) \leq \delta + \delta = 2\delta,$$

which establishes (i). Using (28), we also conclude that $d(b, y) \leq d(b, a) + d(a, y) \leq 2d(a, y) \leq 2\delta$; hence, (ii) holds. Finally, combining (28) and (29), we obtain (iii) via $d(b, c) \leq d(b, a) + d(a, c) \leq \delta + 2\delta = 3\delta$. ■

Corollary 3.3 *Let A and B be nonempty closed subsets of X, let $\rho \in \mathbb{R}_{++}$, and suppose that $c \in A \cap B$. Then*

(30) $$P_A P_B P_A \operatorname{ball}(c; \rho) \subseteq \operatorname{ball}(c; 6\rho).$$

Proof. Let $b_{-1} \in \operatorname{ball}(c; \rho)$, $a_0 \in P_A b_{-1}$, $b_0 \in P_B a_0$, and $a_1 \in P_A b_0$. We have $d(a_0, b_{-1}) = d_A(b_{-1}) \leq d(b_{-1}, c) \leq \rho$, so $d_B(a_0) \leq d(a_0, c) \leq d(a_0, b_{-1}) + d(b_{-1}, c) \leq 2\rho$. Applying Lemma 3.2(iii) to the sets B and A, the points a_0, b_0, a_1, and $\delta = 2\rho$, we deduce that $d(a_1, c) \leq 3(2\rho) = 6\rho$. ■

The next two results are essential to guarantee a local contractive property of the composition.

Proposition 3.4 (regularity and contractivity) *Let A and B be nonempty closed subsets of X, let $\widetilde{A}$ and $\widetilde{B}$ be nonempty subsets of X, let $c \in X$, let $\varepsilon \geq 0$, and let $\delta > 0$. Assume that B is $(\widetilde{A}, \varepsilon, 3\delta)$-regular at c (see Definition 2.21). Furthermore, assume that $y \in B \cap \widetilde{B}$, that $a \in P_A(y) \cap \widetilde{A}$, that $b \in P_B(a)$, that $\|y - c\| \leq \delta$, and that $d_A(y) \leq \delta$. Then*

(31) $$\|a - b\| \leq (\theta_{3\delta} + 2\varepsilon)\|a - y\|,$$

where $\theta_{3\delta}$ the CQ-number at c associated with $(A, \widetilde{A}, B, \widetilde{B})$ (see (9)).

Proof. Lemma 3.2(i)&(iii) yields $\|a-c\| \le 2\delta$ and $\|b-c\| \le 3\delta$. On the other hand, $y-a \in \widehat{N}_A^{\widetilde{B}}(a)$ and $b-a \in -\widehat{N}_B^{\widetilde{A}}(b)$ by (7). Therefore,

$$\langle b-a, y-a\rangle \le \theta_{3\delta}\|b-a\| \cdot \|y-a\|. \tag{32}$$

Since $a-b \in \widehat{N}_B^{\widetilde{A}}(b)$, $\|y-c\| \le \delta$, and $\|b-c\| \le 3\delta$, we obtain, using the $(\widetilde{A}, \varepsilon, 3\delta)$-regularity of B, that $\langle a-b, y-b\rangle \le \varepsilon\|a-b\| \cdot \|y-b\|$. Moreover, Lemma 3.2(ii) states that $\|y-b\| \le 2\|a-y\|$. It follows that

$$\langle a-b, y-b\rangle \le 2\varepsilon\|a-b\| \cdot \|a-y\|. \tag{33}$$

Adding (32) and (33) yields $\|a-b\|^2 \le (\theta_{3\delta}+2\varepsilon)\|a-b\| \cdot \|a-y\|$. The result follows. ■

We now provide a result for collections of sets similar to—and relying upon—Proposition 3.4.

Proposition 3.5 (joint-regularity and contractivity) *Let $\mathcal{A} := (A_i)_{i\in I}$ and $\mathcal{B} := (B_j)_{j\in J}$ be nontrivial collections of closed subsets of X, Assume that $A := \bigcup_{i\in I} A_i$ and $B := \bigcup_{j\in J} B_j$ are closed, and that $c \in A \cap B$. Let $\widetilde{\mathcal{A}} := (\widetilde{A}_i)_{i\in I}$ and $\widetilde{\mathcal{B}} := (\widetilde{B}_j)_{j\in J}$ be nontrivial collections of nonempty subsets of X such that $(\forall i \in I)$ $P_{A_i}((\operatorname{bdry} B) \smallsetminus A) \subseteq \widetilde{A}_i$ and $(\forall j \in J)$ $P_{B_j}((\operatorname{bdry} A) \smallsetminus B) \subseteq \widetilde{B}_j$. Set $\widetilde{A} := \bigcup_{i\in I} \widetilde{A}_i$ and $\widetilde{B} := \bigcup_{j\in J} \widetilde{B}_j$, let $\varepsilon \ge 0$ and let $\delta > 0$.*

(i) *If $b \in (\operatorname{bdry} B) \smallsetminus A$ and $a \in P_A(b)$, then $(\exists\, i \in I)$ $a \in P_{A_i}(b) \subseteq A_i \cap \widetilde{A}_i$.*

(ii) *If $a \in (\operatorname{bdry} A) \smallsetminus B$ and $b \in P_B(a)$, then $(\exists\, j \in J)$ $b \in P_{B_j}(a) \subseteq B_j \cap \widetilde{B}_j$.*

(iii) *If $y \in B$, $a \in P_A(y)$ and $b \in P_B(a)$, then:*

$$b \in \big((\operatorname{bdry} B) \smallsetminus A\big) \cap \bigcup_{j\in J}(B_j \cap \widetilde{B}_j) \;\Leftrightarrow\; b \in B \smallsetminus A \;\Rightarrow\; a \in A \smallsetminus B. \tag{34}$$

(iv) *If $x \in A$, $b \in P_B(x)$, and $a \in P_A(b)$, then:*

$$a \in \big((\operatorname{bdry} A) \smallsetminus B\big) \cap \bigcup_{i\in I}(A_i \cap \widetilde{A}_i) \;\Leftrightarrow\; a \in A \smallsetminus B \;\Rightarrow\; b \in B \smallsetminus A. \tag{35}$$

(v) *Suppose that $\mathcal{B}$ is $(\widetilde{A}, \varepsilon, 3\delta)$-joint-regular at c (see Definition 2.24), that $y \in ((\operatorname{bdry} B) \smallsetminus A) \cap \bigcup_{j\in J}(B_j \cap \widetilde{B}_j)$, that $a \in P_A(y)$, that $b \in P_B(a)$, and that $\|y-c\| \le \delta$. Then*

$$\|b-a\| \le (\theta_{3\delta}+2\varepsilon)\|a-y\|, \tag{36}$$

where $\theta_{3\delta}$ is the joint-CQ-number at c associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ (see (11)).

(vi) *Suppose that $\mathcal{A}$ is $(\widetilde{B}, \varepsilon, 3\delta)$-joint-regular at c (see Definition 2.24), that $x \in ((\operatorname{bdry} A) \smallsetminus B) \cap \bigcup_{i\in I}(A_i \cap \widetilde{A}_i)$, that $b \in P_B(x)$, that $a \in P_A(b)$, and that $\|x-c\| \le \delta$. Then*

$$\|a-b\| \le (\theta_{3\delta}+2\varepsilon)\|b-x\|, \tag{37}$$

where $\theta_{3\delta}$ is the joint-CQ-number at c associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ (see (11)).

Proof. (i)&(ii): Clear from Lemma 2.4 and the assumptions.

(iii): Note that Lemma 3.1(iv)&(iii) and (ii) yield the implications

$$b \in B \smallsetminus A \;\Leftrightarrow\; b \in (\operatorname{bdry} B) \smallsetminus A \;\Rightarrow\; a \in A \smallsetminus B \;\Leftrightarrow\; a \in (\operatorname{bdry} A) \smallsetminus B \;\Rightarrow\; b \in \bigcup_{j \in J} (B_j \cap \widetilde{B}_j), \tag{38}$$

which give the conclusion.

(iv): Interchange the roles of A and B in (iii).

(v): There exists $j \in J$ such that $y \in B_j \cap \widetilde{B}_j \cap ((\operatorname{bdry} B) \smallsetminus A)$. Let $b' \in P_{B_j} a$. Then

$$\|a - b\| = d_B(a) \leq d_{B_j}(a) = \|a - b'\|. \tag{39}$$

Since $\mathcal{B}$ is $(\widetilde{A}, \varepsilon, 3\delta)$-joint-regular at c, it is clear that B_j is $(\widetilde{A}, \varepsilon, 3\delta)$-regular at c. Since $y \in (\operatorname{bdry} B) \smallsetminus A$ and because of (i), there exists $i \in I$ such that $a \in P_{A_i} y \subseteq \widetilde{A}_i$. Since $\widetilde{A}_i \subseteq \widetilde{A}$, it follows that (see also Remark 2.22(iv)) B_j is $(\widetilde{A}_i, \varepsilon, 3\delta)$-regular at c. Since $y \in B_j \cap \widetilde{B}_j$, $a \in P_{A_i} y \cap \widetilde{A}_i$, $b' \in P_{B_j} a$, and $d_{A_i}(y) = d_A(y) = \|y - a\| \leq \|y - c\| \leq \delta$, we obtain from Proposition 3.4 that

$$\|a - b'\| \leq \big(\theta_{3\delta}(A_i, \widetilde{A}_i, B_j, \widetilde{B}_j) + 2\varepsilon\big) \|a - y\|. \tag{40}$$

Combining with (39), we deduce that $\|a - b\| \leq \|a - b'\| \leq (\theta_{3\delta} + 2\varepsilon)\|a - y\|$.

(vi): This follows from (v) and (13). ■

An abstract linear convergence result

Let us now focus on algorithmic results (which are actually true even in complete metric spaces).

Definition 3.6 (linear convergence) *Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X, let $\bar{x} \in X$, and let $\gamma \in [0, 1[$. Then $(x_n)_{n \in \mathbb{N}}$* converges linearly *to $\bar{x}$ with* rate γ *if there exists $\mu \in \mathbb{R}_+$ such that*

$$(\forall n \in \mathbb{N}) \quad d(x_n, \bar{x}) \leq \mu \gamma^n. \tag{41}$$

Remark 3.7 (rate of convergence depends only on the tail of the sequence) Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X, let $\bar{x} \in X$, and let $\gamma \in \,]0, 1[$. Assume that there exists $n_0 \in \mathbb{N}$ and $\mu_0 \in \mathbb{R}_+$ such that

$$\big(\forall n \in \{n_0, n_0 + 1, \ldots\}\big) \quad d(x_n, \bar{x}) \leq \mu_0 \gamma^n. \tag{42}$$

Set $\mu_1 := \max\big\{d(x_m, \bar{x})/\gamma^m \;\big|\; m \in \{0, 1, \ldots, n_0 - 1\}\big\}$. Then

$$(\forall n \in \mathbb{N}) \quad d(x_n, \bar{x}) \leq \max\{\mu_0, \mu_1\}\gamma^n, \tag{43}$$

and therefore $(x_n)_{n \in \mathbb{N}}$ converges linearly to $\bar{x}$ with rate γ.

Proposition 3.8 (abstract linear convergence) *Let A and B be nonempty closed subsets of X, let $(a_n)_{n\in\mathbb{N}}$ be a sequence in A, and let $(b_n)_{n\in\mathbb{N}}$ be a sequence in B. Assume that there exist constants $\alpha \in \mathbb{R}_+$ and $\beta \in \mathbb{R}_+$ such that*

$$\gamma := \alpha\beta < 1 \tag{44a}$$

and

$$(\forall n \in \mathbb{N}) \quad d(a_{n+1}, b_n) \leq \alpha d(a_n, b_n) \text{ and } d(a_{n+1}, b_{n+1}) \leq \beta d(a_{n+1}, b_n). \tag{44b}$$

Then $(\forall n \in \mathbb{N})$ $d(a_{n+1}, b_{n+1}) \leq \gamma d(a_n, b_n)$ and there exists $c \in A \cap B$ such that

$$(\forall n \in \mathbb{N}) \quad \max\big\{d(a_n, c), d(b_n, c)\big\} \leq \frac{1+\alpha}{1-\gamma} d(a_0, b_0) \cdot \gamma^n; \tag{45}$$

consequently, $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ converge linearly to c with rate γ.

Proof. Set $\delta := d(a_0, b_0)$. Then for every $n \in \mathbb{N}$,

$$d(a_n, b_n) \leq \beta d(a_n, b_{n-1}) \leq \alpha\beta d(a_{n-1}, b_{n-1}) = \gamma d(a_{n-1}, b_{n-1}) \leq \cdots \leq \gamma^n \delta; \tag{46}$$

hence,

$$d(b_n, b_{n+1}) \leq d(b_n, a_{n+1}) + d(a_{n+1}, b_{n+1}) \leq \alpha d(b_n, a_n) + \gamma d(a_n, b_n) \tag{47a}$$

$$= (\alpha + \gamma) d(a_n, b_n) \leq (\alpha + \gamma)\delta\gamma^n. \tag{47b}$$

Thus $(b_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, so there exists $c \in B$ such that $b_n \to c$. On the other hand, by (46), $d(a_n, b_n) \to 0$ and $(a_n)_{n\in\mathbb{N}}$ lies in A. Hence, $a_n \to c$ and $c \in A$. Thus, $c \in A \cap B$. Fix $n \in \mathbb{N}$ and let $m \geq n$. Using (47),

$$d(b_n, b_m) \leq \sum_{k=n}^{m-1} d(b_k, b_{k+1}) \leq \sum_{k\geq n} d(b_k, b_{k+1}) \leq \sum_{k\geq n} (\alpha + \gamma)\delta\gamma^k = \frac{(\alpha+\gamma)\delta\gamma^n}{1-\gamma}. \tag{48}$$

Hence, using (46) and (48), we estimate that

$$d(a_n, b_m) \leq d(a_n, b_n) + d(b_n, b_m) \leq \delta\gamma^n + \frac{(\alpha+\gamma)\delta\gamma^n}{1-\gamma} = \frac{(1+\alpha)\delta\gamma^n}{1-\gamma}. \tag{49}$$

Letting $m \to +\infty$ in (48) and (49), we obtain (45). ■

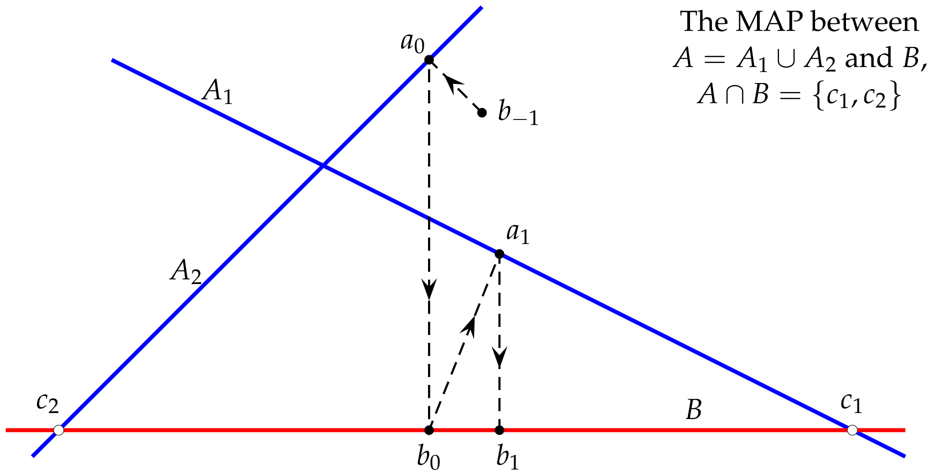
The sequence generated by the MAP

We start with the following definition, which is well defined by Proposition 2.2.

Definition 3.9 (MAP) *Let A and B be nonempty closed subsets of X, let $b_{-1} \in X$, and let*

$$(\forall n \in \mathbb{N}) \quad a_n \in P_A(b_{n-1}) \text{ and } b_n \in P_B(a_n). \tag{50}$$

Then we say that the sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are generated by the method of alternating projections *(with respect to the pair (A, B)) with starting point b_{-1}.*



Our aim is to provide sufficient conditions for linear convergence of the sequences generated by the method of alternating projections. The following two results are simple yet useful.

Proposition 3.10 *Let A and B be nonempty closed subsets of X, and let (a_n) and (b_n) be sequences generated by the method of alternating projections. Then the following hold:*

(i) *The sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ lie in A and B, respectively.*

(ii) $(\forall n \in \mathbb{N})\ \|a_{n+1} - b_{n+1}\| \leq \|a_{n+1} - b_n\| \leq \|a_n - b_n\|$.

(iii) *If $\{a_n\}_{n\in\mathbb{N}} \cap B \neq \varnothing$, or $\{b_n\}_{n\in\mathbb{N}} \cap A \neq \varnothing$, then there exists $c \in A \cap B$ such that for all n sufficiently large, $a_n = b_n = c$.*

Proof. (i): This is clear from the definition.

(ii): Indeed, for every $n \in \mathbb{N}$, $\|a_{n+1} - b_{n+1}\| = d_B(a_{n+1}) \leq \|a_{n+1} - b_n\| = d_A(b_n) \leq \|b_n - a_n\|$ using (i).

(iii): Suppose, say that $a_n \in B$. Then $b_n = P_B a_n = a_n =: c \in A \cap B$ and all subsequent terms of the sequences are equal to c as well. ■

New convergence results for the MAP

We are now in a position to state and derive new linear convergence results. In this section, we shall often assume the following:

$$
(51) \qquad \begin{cases} \mathcal{A} := (A_i)_{i\in I} \text{ and } \mathcal{B} := (B_j)_{j\in J} \text{ are nontrivial collections} \\ \quad \text{of nonempty closed subsets of } X; \\ A := \bigcup_{i\in I} A_i \text{ and } B := \bigcup_{j\in J} B_j \text{ are closed;} \\ c \in A \cap B; \\ \widetilde{\mathcal{A}} := (\widetilde{A}_i)_{i\in I} \text{ and } \widetilde{\mathcal{B}} := (\widetilde{B}_j)_{j\in J} \text{ are collections} \\ \quad \text{of nonempty subsets of } X \text{ such that} \\ \qquad (\forall i \in I) \;\; P_{A_i}\big((\operatorname{bdry} B) \smallsetminus A\big) \subseteq \widetilde{A}_i, \\ \qquad (\forall j \in J) \;\; P_{B_j}\big((\operatorname{bdry} A) \smallsetminus B\big) \subseteq \widetilde{B}_j; \\ \widetilde{A} := \bigcup_{i\in I} \widetilde{A}_i \text{ and } \widetilde{B} := \bigcup_{j\in J} \widetilde{B}_j. \end{cases}
$$

Lemma 3.11 (backtracking MAP) *Assume that* (51) *holds. Let* $(a_n)_{n\in\mathbb{N}}$ *and* $(b_n)_{n\in\mathbb{N}}$ *be generated by the MAP with starting point* b_{-1}. *Let* $n \in \{1, 2, 3, \ldots\}$. *Then the following hold:*

(i) *If* $b_n \notin A$, *then* $a_n \in ((\operatorname{bdry} A) \smallsetminus B) \cap \bigcup_{i\in I}(A_i \cap \widetilde{A}_i)$ *and* $b_n \in ((\operatorname{bdry} B) \smallsetminus A) \cap \bigcup_{j\in J}(B_j \cap \widetilde{B}_j)$.

(ii) *If* $a_n \notin B$, *then* $a_n \in ((\operatorname{bdry} A) \smallsetminus B) \cap \bigcup_{i\in I}(A_i \cap \widetilde{A}_i)$.

(iii) *If* $a_n \notin B$ *and* $n \geq 2$, *then* $b_{n-1} \in ((\operatorname{bdry} B) \smallsetminus A) \cap \bigcup_{j\in J}(B_j \cap \widetilde{B}_j)$.

Proof. (i): Applying Proposition 3.5(iii) to $b_{n-1} \in B$, $a_n \in P_A b_{n-1}$, $b_n \in P_B a_n$, we obtain

$$
(52) \qquad b_n \in B \smallsetminus A \;\Leftrightarrow\; b_n \in \big((\operatorname{bdry} B) \smallsetminus A\big) \cap \bigcup_{j\in J}(B_j \cap \widetilde{B}_j) \;\Rightarrow\; a_n \in A \smallsetminus B.
$$

On the other hand, applying Proposition 3.5(iv) to $a_{n-1} \in A$, $b_{n-1} \in P_B a_{n-1}$, $a_n \in P_A b_{n-1}$, we see that

$$
(53) \qquad a_n \in A \smallsetminus B \;\Leftrightarrow\; a_n \in \big((\operatorname{bdry} A) \smallsetminus B\big) \cap \bigcup_{i\in I}(A_i \cap \widetilde{A}_i).
$$

Altogether, (i) is established.

(ii)&(iii): The proofs are analogous to that of (i). ■

Let us now state and prove a key technical result.

Proposition 3.12 *Assume that* (51) *holds. Suppose that there exist* $\varepsilon \geq 0$ *and* $\delta > 0$ *such that the following hold:*

(i) $\mathcal{A}$ *is* $(\widetilde{B}, \varepsilon, 3\delta)$*-joint-regular at* c *(see Definition 2.24) and set*

$$\sigma := \begin{cases} 1, & \text{if } \mathcal{B} \text{ is not known to be } (\widetilde{A}, \varepsilon, 3\delta)\text{-joint-regular at } c; \\ 2, & \text{if } \mathcal{B} \text{ is also } (\widetilde{A}, \varepsilon, 3\delta)\text{-joint-regular at } c. \end{cases} \tag{54}$$

(ii) $\theta_{3\delta} < 1 - 2\varepsilon$*, where* $\theta_{3\delta}$ *is the joint-CQ-number at* c *associated with* $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ *(see Definition 2.10).*

Set $\theta := \theta_{3\delta} + 2\varepsilon \in \,]0, 1[$*. Let* $(a_n)_{n\in\mathbb{N}}$ *and* $(b_n)_{n\in\mathbb{N}}$ *be sequences generated by the MAP with starting point* b_{-1} *satisfying*

$$\|b_{-1} - c\| \leq \frac{(1-\theta^\sigma)\delta}{6(2+\theta-\theta^\sigma)}. \tag{55}$$

Then $(a_n)_{n\in\mathbb{N}}$ *and* $(b_n)_{n\in\mathbb{N}}$ *converge linearly to some point* $\bar{c} \in A \cap B$ *with rate* θ^σ*; in fact,*

$$\|\bar{c} - c\| \leq \delta \quad \text{and} \quad (\forall n \geq 1) \ \ \max\big\{\|a_n - \bar{c}\|, \|b_n - \bar{c}\|\big\} \leq \frac{\delta(1+\theta)}{2+\theta-\theta^\sigma}\theta^{\sigma(n-1)}. \tag{56}$$

Proof. In view of $a_1 \in P_A P_B P_A b_{-1}$ and (55), Corollary 3.3 yields

$$\beta := \|a_1 - c\| \leq \frac{(1-\theta^\sigma)\delta}{(2+\theta-\theta^\sigma)} \leq \frac{\delta}{2}. \tag{57}$$

Since $c \in A \cap B$, we have $\theta_{3\delta} \geq 0$ by (14) and hence $\theta > 0$. Using (57), we estimate

$$\begin{aligned}
&(\forall n \geq 1) \quad \beta\theta^{\sigma(n-1)} + \beta + \beta(1+\theta)\sum_{k=0}^{n-2}\theta^{\sigma k} \leq \beta + \beta(1+\theta)\sum_{k=0}^{n-1}\theta^{\sigma k} && \text{(58a)}\\
&\qquad = \beta + \beta(1+\theta)\frac{1-\theta^{\sigma n}}{1-\theta^\sigma} && \text{(58b)}\\
&\qquad \leq \beta + \beta\frac{1+\theta}{1-\theta^\sigma} && \text{(58c)}\\
&\qquad = \beta\Big(\frac{2+\theta-\theta^\sigma}{1-\theta^\sigma}\Big) && \text{(58d)}\\
&\qquad \leq \delta. && \text{(58e)}
\end{aligned}$$

We now claim that if

$$n \geq 1, \quad \|a_n - b_n\| \leq \beta\theta^{\sigma(n-1)} \quad \text{and} \quad \|a_n - c\| \leq \beta + \beta(1+\theta)\sum_{k=0}^{n-2}\theta^{\sigma k}, \tag{59}$$

then

$$\|a_{n+1} - b_{n+1}\| \leq \theta^{\sigma-1}\|a_{n+1} - b_n\| \leq \theta^\sigma\|a_n - b_n\| \leq \beta\theta^{\sigma n}, \tag{60a}$$

$$\|a_{n+1} - c\| \le \beta + \beta(1+\theta)\sum_{k=0}^{n-1} \theta^{\sigma k}. \tag{60b}$$

To prove this claim, assume that (59) holds. Using (59) and (58), we first observe that

$$\max\big\{\|a_n - c\|, \|b_n - c\|\big\} \le \|b_n - a_n\| + \|a_n - c\| \tag{61a}$$

$$\le \beta\theta^{\sigma(n-1)} + \beta + \beta(1+\theta)\sum_{k=0}^{n-2} \theta^{\sigma k} \le \delta. \tag{61b}$$

We now consider two cases:

Case 1: $b_n \in A \cap B$. Then $b_n = a_{n+1} = b_{n+1}$ and thus (60a) holds. Moreover, $\|a_{n+1} - c\| = \|b_n - c\|$ and (60b) follows from (61a).

Case 2: $b_n \notin A \cap B$. Then $b_n \in B \smallsetminus A$. Lemma 3.11(i) implies $a_n \in ((\operatorname{bdry} A) \smallsetminus B) \cap \bigcup_{i \in I}(A_i \cap \widetilde{A}_i)$ and $b_n \in ((\operatorname{bdry} B) \smallsetminus A) \cap \bigcup_{j \in J}(B_j \cap \widetilde{B}_j)$. Note that $\|a_n - c\| \le \delta$ by (61a), and recall that $\mathcal{A}$ is $(\widetilde{B}, \varepsilon, 3\delta)$-joint-regular at c by (i). It thus follows from Proposition 3.5(vi) (applied to a_n, b_n, a_{n+1}) that

$$\|a_{n+1} - b_n\| \le \theta\|a_n - b_n\|. \tag{62}$$

On the one hand, if $\sigma = 1$, then Proposition 3.10(ii) yields $\|a_{n+1} - b_{n+1}\| \le \|a_{n+1} - b_n\| = \theta^{\sigma-1}\|a_{n+1} - b_n\|$. On the other hand, if $\sigma = 2$, then $\mathcal{B}$ is $(\widetilde{A}, \varepsilon, 3\delta)$-joint-regular at c by (i); hence, Proposition 3.5(v) (applied to b_n, a_{n+1}, b_{n+1}) yields $\|a_{n+1} - b_{n+1}\| \le \theta\|a_{n+1} - b_n\| = \theta^{\sigma-1}\|a_{n+1} - b_n\|$. Altogether, in either case,

$$\|a_{n+1} - b_{n+1}\| \le \theta^{\sigma-1}\|a_{n+1} - b_n\|. \tag{63}$$

Combining (63) with (62) and (59) gives

$$\|a_{n+1} - b_{n+1}\| \le \theta^{\sigma-1}\|a_{n+1} - b_n\| \le \theta^{\sigma}\|a_n - b_n\| \le \beta\theta^{\sigma n}, \tag{64}$$

which is (60a). Furthermore, (62), (59) and (61a) yield

$$\|a_{n+1} - c\| \le \|a_{n+1} - b_n\| + \|b_n - c\| \tag{65a}$$

$$\le \theta\|a_n - b_n\| + \|b_n - c\| \tag{65b}$$

$$\le \theta\beta\theta^{\sigma(n-1)} + \beta\theta^{\sigma(n-1)} + \beta + \beta(1+\theta)\sum_{k=0}^{n-2} \theta^{\sigma k} \tag{65c}$$

$$= \beta + \beta(1+\theta)\sum_{k=0}^{n-1} \theta^{\sigma k}, \tag{65d}$$

which establishes (60b). Therefore, in all cases, (60) holds.

Since $\|a_1 - b_1\| = d_B(a_1) \le \|a_1 - c\| = \beta$, we see that (59) holds for $n = 1$. Thus, the above claim and the principle of mathematical induction principle imply that (60) holds for every $n \ge 1$.

Next, (60a) implies

$$(\forall n \geq 1) \quad \|a_{n+1} - b_n\| \leq \theta \|a_n - b_n\| \quad \text{and} \quad \|a_{n+1} - b_{n+1}\| \leq \theta^{\sigma-1} \|a_{n+1} - b_n\|. \tag{66}$$

In view of (66) and $\|a_1 - b_1\| \leq \beta$, Proposition 3.8 yields $\bar{c} \in A \cap B$ such that

$$(\forall n \geq 1) \quad \max\left\{\|a_n - \bar{c}\|, \|b_n - \bar{c}\|\right\} \leq \frac{1+\theta}{1-\theta^\sigma} \|a_1 - b_1\| \cdot \theta^{\sigma(n-1)} \tag{67}$$

$$\leq \frac{1+\theta}{1-\theta^\sigma} \beta \cdot \theta^{\sigma(n-1)} \tag{68}$$

$$\leq \frac{\delta(1+\theta)}{2+\theta-\theta^\sigma} \theta^{\sigma(n-1)}. \tag{69}$$

On the other hand, (60b) and (58) imply $(\forall n \geq 1)$ $\|a_{n+1} - c\| \leq \delta$; thus, letting $n \to +\infty$, we obtain $\|\bar{c} - c\| \leq \delta$. This completes the proof of (56). ■

Remark 3.13 In view of Lemma 3.1(i)&(ii), an aggressive choice for use in (51) is $(\forall i \in I)$ $\widetilde{A}_i = \operatorname{bdry} A_i$ and $(\forall j \in J)$ $\widetilde{B}_j = \operatorname{bdry} B_j$.

Our main convergence result on the linear convergence of the MAP is the following:

Theorem 3.14 (linear convergence of the MAP and superregularity) *Assume that* (51) *holds and that $\mathcal{A}$ is $\widetilde{B}$-joint-superregular at c (see Definition 2.24). Denote the limiting joint-CQ-number at c associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ (see Definition 2.10) by $\overline{\theta}$, and the the exact joint-CQ-number at c associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ (see Definition 2.13) by $\overline{\alpha}$. Assume further that one of the following holds:*

(i) $\overline{\theta} < 1$.

(ii) *I and J are finite, and $\overline{\alpha} < 1$.*

Let $\theta \in \left]\overline{\theta}, 1\right[$ and set $\varepsilon := (\theta - \overline{\theta})/3 > 0$. Then there exists $\delta > 0$ such that the following hold:

(iii) *$\mathcal{A}$ is $(\widetilde{B}, \varepsilon, 3\delta)$-joint-regular at c (see Definition 2.24).*

(iv) *$\theta_{3\delta} \leq \overline{\theta} + \varepsilon < 1 - 2\varepsilon$, where $\theta_{3\delta}$ is the joint-CQ-number at c associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ (see Definition 2.10).*

Consequently, suppose the starting point of the MAP b_{-1} satisfies $\|b_{-1} - c\| \leq (1-\theta)\delta/12$. Then $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ converge linearly to some point in $\bar{c} \in A \cap B$ with $\|\bar{c} - c\| \leq \delta$ and rate θ:

$$(\forall n \geq 1) \ \max\{\|a_n - \bar{c}\|, \|b_n - \bar{c}\|\} \leq \frac{\delta(1+\theta)}{2} \theta^{n-1}. \tag{70}$$

Proof. Observe that assumption (ii) is more restrictive than assumption (i) by Theorem 2.14(iv). The definitions of $\widetilde{B}$-joint-superregularity and of $\overline{\theta}$ allow us to find $\delta > 0$ sufficiently small such that both (iii) and (iv) hold. The result thus follows from Proposition 3.12 with $\sigma = 1$. ■

Corollary 3.15 *Assume that* (51) *holds and that, for every $i \in I$, A_i is convex. Denote the limiting joint-CQ-number at c associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ (see Definition 2.10) by $\overline{\theta}$, and assume that $\overline{\theta} < 1$. Let $\theta \in \left]\overline{\theta}, 1\right[$, and let b_{-1}, the starting point of the MAP, be sufficiently close to c. Then $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ converge linearly to some point in $A \cap B$ with rate θ.*

Proof. Combine Theorem 3.14 with Corollary 2.25. ■

Example 3.16 (working with collections and joint notions is useful) Consider the setting of Example 2.11, and suppose that $\widetilde{\mathcal{A}} = \mathcal{A}$ and $\widetilde{\mathcal{B}} = \mathcal{B}$. Note that A_i is convex, for every $i \in I$. Then $\theta_\delta(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}}) < 1 = \theta_\delta(A, A, B, B) = \overline{\theta}(A, X, B, X)$. Hence Corollary 3.15 guarantees linear convergence of the MAP while it is not possible to work directly with the unions A and B due to their condition number being equal to 1 *and* because neither A nor B is superregular by Example 2.26! This illustrates that the main result of Lewis-Luke-Malick (see Corollary 3.25 below) is not applicable because two of its hypotheses fail.

The following result features an improved rate of convergence θ^2 due to the additional presence of superregularity.

Theorem 3.17 (linear convergence of the MAP and double superregularity) *Assume that* (51) *holds, that $\mathcal{A}$ is $\widetilde{B}$-joint-superregular at c and that $\mathcal{B}$ is $\widetilde{A}$-joint-superregular at c (see Definition 2.24). Denote the limiting joint-CQ-number at c associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ (see Definition 2.10) by $\overline{\theta}$, and the the exact joint-CQ-number at c associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ (see Definition 2.13) by $\overline{\alpha}$. Assume further that* (a) *$\overline{\theta} < 1$, or (more restrictively) that* (b) *I and J are finite, and $\overline{\alpha} < 1$ (and hence $\overline{\theta} = \overline{\alpha} < 1$). Let $\theta \in \left]\overline{\theta}, 1\right[$ and $\varepsilon := \frac{\theta - \overline{\theta}}{3}$. Then there exists $\delta > 0$ such that*

(i) *$\mathcal{A}$ is $(\widetilde{B}, \varepsilon, 3\delta)$-joint-regular at c;*

(ii) *$\mathcal{B}$ is $(\widetilde{A}, \varepsilon, 3\delta)$-joint-regular at c; and*

(iii) *$\theta_{3\delta} < \overline{\theta} + \varepsilon = \theta - 2\varepsilon < 1 - 2\varepsilon$, where $\theta_{3\delta}$ is the joint-CQ-number at c associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ (see Definition 2.10).*

Consequently, suppose the starting point of MAP b_{-1} satisfies $\|b_{-1} - c\| \leq \frac{(1-\theta)\delta}{6(2-\theta)}$. Then $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ converge linearly to some point in $\bar{c} \in A \cap B$ with $\|\bar{c} - c\| \leq \delta$ and rate θ^2; in fact,

$$(71) \qquad (\forall n \geq 1) \quad \max\big\{\|a_n - \bar{c}\|, \|b_n - \bar{c}\|\big\} \leq \frac{\delta}{2-\theta}\big(\theta^2\big)^{n-1}.$$

Proof. The existence of $\delta > 0$ such that (i)–(iii) hold is clear. Then apply Proposition 3.12 with $\sigma = 2$. ■

In passing, let us point out a sharper rate of convergence under sufficient conditions stronger than superregularity.

Corollary 3.18 (refined convergence rate) *Assume that* (51) *holds and that there exists* $\delta > 0$ *such that*

(i) $\mathcal{A}$ *is* $(\widetilde{B}, 0, 3\delta)$*-joint-regular at* c*;*

(ii) $\mathcal{B}$ *is* $(\widetilde{A}, 0, 3\delta)$*-joint-regular at* c*; and*

(iii) $\theta < 1$*, where* $\theta := \theta_{3\delta}$ *is the joint-CQ-number at* c *associated with* $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ *(see Definition 2.10).*

Suppose also that the starting point of the MAP b_{-1} *satisfies* $\|b_{-1} - c\| \leq \frac{(1-\theta)\delta}{6(2-\theta)}$*. Then* $(a_n)_{n\in\mathbb{N}}$ *and* $(b_n)_{n\in\mathbb{N}}$ *converge linearly to some point in* $\bar{c} \in A \cap B$ *with* $\|\bar{c} - c\| \leq \delta$ *and rate* θ^2*; in fact,*

$$(\forall n \geq 1) \quad \max\big\{\|a_n - \bar{c}\|, \|b_n - \bar{c}\|\big\} \leq \frac{\delta}{2-\theta}\big(\theta^2\big)^{n-1}. \tag{72}$$

Proof. Apply Proposition 3.12 with $\sigma = 2$. ■

Let us illustrate a situation where it is possible to make δ in Theorem 3.17 precise.

Example 3.19 (the MAP for two spheres) Let z_1 and z_2 be in X, let ρ_1 and ρ_2 be in $\mathbb{R}$, set $A := \operatorname{sphere}(z_1; \rho_1)$ and $B := \operatorname{sphere}(z_2; \rho_2)$, and assume that $\{c\} \subsetneqq A \cap B \subsetneqq A \cup B$. Then $\overline{\alpha} := |\langle z_1 - c, z_2 - c\rangle| / (\rho_1\rho_2) < 1$. Let $\theta \in \,]\overline{\alpha}, 1[$. Then the conclusion of Theorem 3.17 holds with[4]

$$\delta := \min\left\{\frac{\sqrt{(\rho_1+\rho_2)^2 + \rho_1\rho_2(\theta - \overline{\alpha})} - (\rho_1+\rho_2)}{6}, \frac{(\theta-\overline{\alpha})\rho_1}{12}, \frac{(\theta-\overline{\alpha})\rho_2}{12}\right\} \tag{73}$$

Proof. Combine Example 2.23 (applied with $\varepsilon = (\theta - \overline{\alpha})/4$ there), Proposition 2.17, and Theorem 3.17. ■

Here is a useful special case of Theorem 3.17:

Theorem 3.20 *Assume that* A *and* B *are* L*-superregular, and that*

$$N_A(c) \cap \big(-N_B(c)\big) \cap \big(L - c\big) = \{0\}, \tag{74}$$

where $L := \operatorname{aff}(A \cup B)$*. Then the sequences generated by the MAP converge linearly to a point in* $A \cap B$ *provided that the starting point is sufficiently close to* c*.*

Proof. Combine Example 2.15 with Theorem 3.17 (applied with I and J being singletons, and with $\widetilde{A} = \widetilde{B} = L$). ■

We now obtain a well known global linear convergence result for the convex case[5], which does not require the starting point to be sufficiently close to $A \cap B$:

[4] Note that when $\overline{\alpha}$ approaches 1, then δ approaches 0 which is consistent with the lack of linear convergence of the MAP for two spheres intersecting in exactly one point.

[5] This result is part of the folklore and it can be traced back to [19] although it is not stated there explicitly in this form. It also follows by combining [1, Proposition 4.6.1] with [2, Theorem 3.12].

Theorem 3.21 (two convex sets) *Assume that A and B are convex, and $A \cap B \neq \varnothing$. Then for every starting point $b_{-1} \in X$, the sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ generated by the MAP converge to some point in $A \cap B$. The convergence of these sequences is linear provided that* $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$.

Proof. By Fact 2.5(iv), we have

$$(\forall c \in A \cap B) \quad \|a_0 - c\| \geq \|b_0 - c\| \geq \|a_1 - c\| \geq \|b_1 - c\| \geq \cdots \tag{75}$$

After passing to subsequences if needed, we assume that $a_{k_n} \to a \in A$ and $b_{k_n} \to b \in B$. We show that $a = b$ by contradiction, so we assume that $\varepsilon := \|a - b\|/3 > 0$. We have eventually $\max\{\|a_{k_n} - a\|, \|b_{k_n} - b\|\} < \varepsilon$; hence $\|a_{k_n} - b_{k_n}\| \geq \varepsilon$ eventually. By Fact 2.5(iii), we have

$$\|a_{k_n} - c\|^2 \geq \|a_{k_n} - b_{k_n}\|^2 + \|b_{k_n} - c\|^2 \geq \varepsilon^2 + \|a_{k_n+1} - c\|^2 \geq \varepsilon^2 + \|a_{k_{n+1}} - c\|^2 \tag{76}$$

eventually. But this would imply that for all n sufficiently large, and for every $m \in \mathbb{N}$, we have $\|a_{k_n} - c\|^2 \geq m\varepsilon^2 + \|a_{k_{n+m}} - c\|^2 \geq m\varepsilon^2$, which is absurd. Hence $\bar{c} := a = b \in A \cap B$ and now (75) (with $c = \bar{c}$) implies that $a_n \to \bar{c}$ and $b_n \to \bar{c}$.

Next, assume that $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$, and set $L := \operatorname{aff}(A \cup B)$. By Proposition 2.18, the (A, L, B, L)-CQ conditions holds at $\bar{c}$. Thus, by Example 2.15, $N_A(\bar{c}) \cap (-N_B(\bar{c})) \cap (L - \bar{c}) = \{0\}$. Furthermore, Corollary 2.25 and Remark 2.22(v)&(vii) imply that A and B are L-superregular at $\bar{c}$. The conclusion now follows from Theorem 3.20, applied to suitably chosen tails of the sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$. ■

Example 3.22 (the MAP for two linear subspaces) Assume that A and B are linear subspaces of X. Since $0 \in A \cap B = \operatorname{ri} A \cap \operatorname{ri} B$, Theorem 3.21 guarantees the linear convergence of the MAP to some point in $A \cap B$, where $b_{-1} \in X$ is the arbitrary starting point. On the other hand, A and B are $(0, +\infty)$-regular (see Remark 2.22(v)). Since $(\forall \delta \in \mathbb{R}_{++})\ \theta_\delta(A, A, B, B) = c(A, B) < 1$, where $c(A, B)$ is the cosine of the Friedrichs angle between A and B (see Theorem 2.20), we obtain from Corollary 3.18 that the rate of convergence is $c^2(A, B)$. In fact, it is well known that this is the optimal rate, and also that $\lim_n a_n = \lim_n b_n = P_{A\cap B}(b_{-1})$; see [12, Section 3] and [13, Chapter 9].

Remark 3.23 (subspaces vs manifolds) It is tempting to explore the following statement, which is a variant of Example 3.22.

> *Let A and B be C^2 submanifolds of X, and let $c \in A \cap B$ such that the Friedrichs angle between the tangent spaces at c is strictly positive. If the starting point of the MAP is sufficiently close to c, then the sequences generated by the MAP converge linearly to a point in $A \cap B$.*

Interestingly, this statement is *false*. First, we note that the Friedrichs angle is *always* strictly positive by Theorem 2.20. Secondly, consider either (i) two spheres intersecting in precisely one point; or (ii) $A = \mathbb{R} \times \{0\}$ and $\operatorname{epi}(\rho \mapsto \rho^2)$ in $X = \mathbb{R}^2$. In either case, $A \cap B = \{c\}$ is a singleton, and the MAP does not converge linearly to c unless the starting point is c itself.

We conjecture that the statement above is correct if the Friedrichs angle is replaced by the *Dixmier angle*. Unfortunately, this modified statement is of somewhat limited interest because the classical case of two linear subspaces is still not covered (consider two linear subspaces A and B such that $A \cap B \supsetneqq \{0\}$; e.g., two planes in $\mathbb{R}^3$).

Remark 3.24 For further linear convergence results for the MAP in the convex setting we refer the reader to [2], [3], [4], [14], [15], [16], and the references therein. See also [22] and [23] for recent related work for the nonconvex case.

Comparison to Lewis-Luke-Malick results and further examples

The main result of Lewis, Luke, and Malick arises as a special case of Theorem 3.14:

Corollary 3.25 (Lewis-Luke-Malick) (See [20, Theorem 5.16].) *Suppose that $N_A(c) \cap (-N_B(c)) = \{0\}$ and that A is superregular at $c \in A \cap B$. If the starting point of MAP is sufficiently close to c, then the sequences generated by the MAP converge linearly to a point in $A \cap B$.*

Proof. Since $N_A(c) \cap (-N_B(c)) = \{0\}$, we have $\overline{\theta} < 1$. Now apply Theorem 3.14(i) with $\widetilde{\mathcal{A}} := \widetilde{\mathcal{B}} := (X)$, $\mathcal{A} := (A)$ and $\mathcal{B} := (B)$. ■

However, even in simple situations, Corollary 3.25 is not powerful enough to recover known convergence results.

Example 3.26 (Lewis-Luke-Malick CQ may fail even for two subspaces) Suppose that A and B are two linear subspaces of X, and set $L := \operatorname{aff}(A \cup B) = A + B$. For $c \in A \cap B$, we have

$$N_A(c) \cap (-N_B(c)) = A^\perp \cap B^\perp = (A + B)^\perp = L^\perp. \tag{77}$$

Therefore, the Lewis-Luke-Malick CQ (see [20, Theorem 5.16] and also Corollary 3.25) holds for (A, B) at c if and only if

$$N_A(c) \cap (-N_B(c)) = \{0\} \;\Leftrightarrow\; A + B = X. \tag{78}$$

On the other hand, the CQ provided in Theorem 3.20 (see also Example 3.22) *always holds* and we obtain linear convergence of the MAP. However, even for two lines in $\mathbb{R}^3$, the Lewis-Luke-Malick CQ (see Corollary 3.25) is unable to achieve this. (It was this example that originally motivated us to pursue the present work.)

Example 3.27 (Lewis-Luke-Malick CQ is too strong even for convex sets) Assume that A and B are convex (and hence superregular). Then the Lewis-Luke-Malick CQ condition is $0 \in \operatorname{int}(B - A)$ (see Corollary 2.9(i)) while the $(A, \operatorname{aff}(A \cup B), B, \operatorname{aff}(A \cup B))$-CQ is equivalent to the much less restrictive condition $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$ (see Theorem 2.8).

The flexibility of choosing $(\widetilde{A}, \widetilde{B})$

Often, $L = \operatorname{aff}(A \cup B)$ is a convenient choice which yields linear convergence of the MAP as in Theorem 3.20. However, there are situations when this choice for $\widetilde{A}$ and $\widetilde{B}$ is not helpful but when a different, more aggressive, choice does guarantee linear convergence:

Example 3.28 ($(\widetilde{A}, \widetilde{B}) = (A, B)$) Let A, B, and c be as in Example 2.16, and let $L := \operatorname{aff}(A \cup B)$. Since A and B are *convex* and hence *superregular*, the (A, L, B, L)-CQ condition is equivalent to $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$ (see Proposition 2.18), which fails in this case. However, the (A, A, B, B)-CQ condition does hold; hence, the corresponding limiting CQ-number is less than 1 by Theorem 2.14(v). Thus linear convergence of the MAP is guaranteed by Theorem 3.17.

The next example illustrates a situation where the choice $(\widetilde{A}, \widetilde{B}) = (A, B)$ fails while the even tighter choice $(\widetilde{A}, \widetilde{B}) = (\operatorname{bdry} A, \operatorname{bdry} B)$ results in success:

Example 3.29 ($(\widetilde{A}, \widetilde{B}) = (\operatorname{bdry} A, \operatorname{bdry} B)$) Suppose that $X = \mathbb{R}^2$, that $A = \operatorname{epi}(|\cdot|/2)$, that $B = -\operatorname{epi}(|\cdot|/3)$, and that $c = (0, 0)$. Note that $\operatorname{aff}(A \cup B) = X$ and $\operatorname{ri} A \cap \operatorname{ri} B = \varnothing$. Then

$$N_A^B(c) = N_A^X(c) = N_A(c) = \big\{(u_1, u_2) \in \mathbb{R}^2 \;\big|\; u_2 + 2|u_1| \leq 0\big\}, \tag{79a}$$

$$N_B^A(c) = N_B^X(c) = N_B(c) = \big\{(u_1, u_2) \in \mathbb{R}^2 \;\big|\; -u_2 + 3|u_1| \leq 0\big\}, \tag{79b}$$

and so the (A, A, B, B)-CQ condition fails because

$$N_A^B(c) \cap (-N_B^A(c)) = \big\{(u_1, u_2) \in \mathbb{R}^2 \;\big|\; u_2 + 3|u_1| \leq 0\big\} \neq \{0\}. \tag{80}$$

Consequently, for either $(\widetilde{A}, \widetilde{B}) = (A, B)$ or $(\widetilde{A}, \widetilde{B}) = (X, X)$, Theorem 3.17 is not applicable because $\overline{\alpha} = \overline{\theta} = 1$: indeed, $u = (0, -1) \in N_A(c)$ and $v = (0, -1) \in -N_B(c)$, so $1 = \langle u, v\rangle \leq \bar{\alpha} \leq 1$.

On the other hand, let us now choose $(\widetilde{A}, \widetilde{B}) = (\operatorname{bdry} A, \operatorname{bdry} B)$, which is justified by Remark 3.13. Then

$$N_A^{\widetilde{B}}(c) = \big\{(u_1, u_2) \in \mathbb{R}^2 \;\big|\; u_2 + 2|u_1| = 0\big\}, \tag{81a}$$

$$N_B^{\widetilde{A}}(c) = \big\{(u_1, u_2) \in \mathbb{R}^2 \;\big|\; -u_2 + 3|u_1| = 0\big\}, \tag{81b}$$

$N_A^{\widetilde{B}}(c) \cap (-N_B^{\widetilde{A}}(c)) = \{0\}$ and the $(A, \widetilde{A}, B, \widetilde{B})$-CQ condition holds. Hence, using also Theorem 2.14(v), Theorem 3.21 and Theorem 3.17, we deduce linear convergence of the MAP.

However, even the choice $(\widetilde{A}, \widetilde{B}) = (\operatorname{bdry} A, \operatorname{bdry} B)$ may not be applicable to yield the desired linear convergence as the following shows. In this example, we employ the tightest possibility allowed by our framework, namely $(\widetilde{A}, \widetilde{B}) = (P_A((\operatorname{bdry} B) \smallsetminus A), P_B((\operatorname{bdry} A) \smallsetminus B))$.

Example 3.30 ($(\widetilde{A}, \widetilde{B}) = (P_A((\operatorname{bdry} B) \smallsetminus A), P_B((\operatorname{bdry} A) \smallsetminus B))$) Suppose that $X = \mathbb{R}^2$, that $A = \operatorname{epi}(|\cdot|)$, that $B = -A$, and that $c = (0, 0)$. Then $N_A^{\operatorname{bdry} B}(c) = \operatorname{bdry} B = -\operatorname{bdry} A$ and $N_B^{\operatorname{bdry} A}(c) = \operatorname{bdry} A$; hence, the $(A, \operatorname{bdry} A, B, \operatorname{bdry} B)$-CQ condition fails because $N_A^{\operatorname{bdry} B}(c) \cap (-N_B^{\operatorname{bdry} A}(c)) =$

bdry $B \neq \{0\}$. On the other hand, if $(\widetilde{A}, \widetilde{B}) = (P_A((\text{bdry}\, B) \smallsetminus A), P_B((\text{bdry}\, A) \smallsetminus B))$, then $N_A^{\widetilde{B}} = \{0\} = N_B^{\widetilde{A}} = \{0\}$ because $\widetilde{A} = \{c\} = \widetilde{B}$. Thus, the $(A, \widetilde{A}, B, \widetilde{B})$-CQ conditions holds. (Note that the MAP converges in finitely many steps.)

Conclusion

We use the technology of restricted normal cones developed in [6] to develop the least restrictive sufficient conditions to date for linear convergence of the sequences generated by the method of alternating projections applied to two sets A and B. A key ingredient were suitable restricting sets ($\widetilde{A}$ and $\widetilde{B}$). The unrestricted — and hence least aggressive — choice $(\widetilde{A}, \widetilde{B}) = (X, X)$ recovers the framework by Lewis, Luke, and Malick. The choice $(\widetilde{A}, \widetilde{B}) = (\text{aff}(A \cup B), \text{aff}(A \cup B))$ allows us to include basic settings from convex analysis into our framework. Thus, the framework provided here unifies the recent nonconvex results by Lewis, Luke, and Malick with classical convex-analytical settings. When the choice $(\widetilde{A}, \widetilde{B}) = (\text{aff}(A \cup B), \text{aff}(A \cup B))$ fails, one may also try more aggressive choices such as $(\widetilde{A}, \widetilde{B}) = (A, B)$ or $(\widetilde{A}, \widetilde{B}) = (\text{bdry}\, A, \text{bdry}\, B)$ to guarantee linear convergence. In a follow-up work [7] we demonstrate the power of these tools with the important problem of sparsity optimization with affine constraints. Without any assumptions on the regularity of the sets or the intersection we achieve local convergence results, with explicit rates and radii of convergence, where all other sufficient conditions, particularly those of [21] and [20], fail.

Acknowledgments

We would like to thank the referee for her/his helpful comments. HHB was partially supported by the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chair Program. This research was initiated when HHB visited the Institut für Numerische und Angewandte Mathematik, Universität Göttingen because of his study leave in Summer 2011. HHB thanks DRL and the Institut for their hospitality. DRL was supported in part by the German Research Foundation grant SFB755-A4. HMP was partially supported by the Pacific Institute for the Mathematical Sciences and and by a University of British Columbia research grant. XW was partially supported by the Natural Sciences and Engineering Research Council of Canada.

References

[1] H.H. Bauschke, *Projection Algorithms and Monotone Operators*, Ph.D. thesis, Simon Fraser University, Burnaby, B.C., Canada, 1996.

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

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

[4] H.H. Bauschke, J.M. Borwein, and A.S. Lewis, The method of cyclic projections for closed convex sets in Hilbert space, in *Recent Developments in Optimization Theory and Nonlinear Analysis (Jerusalem 1995)*, Y. Censor and S. Reich (editors), Contemporary Mathematics vol. 204, American Mathematical Society, pp. 1–38, 1997.

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

[6] H.H. Bauschke, D.R. Luke, H.M. Phan, and X. Wang, Restricted normal cones and the method of alternating projections: theory, preprint, March 2013.

[7] H.H. Bauschke, D.R. Luke, H.M. Phan, and X. Wang, Restricted normal cones and sparsity optimization with affine constraints, arXiv preprint, May 2012, `http://arxiv.org`

[8] J.M. Borwein and Q.J. Zhu, *Techniques of Variational Analysis*, Springer-Verlag, 2005.

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

[10] F.H. Clarke, Y.S. Ledyaev, R.J. Stern and P.R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer-Verlag, 1998.

[11] 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.

[12] F. Deutsch, The angle between subspaces of a Hilbert space, in *Approximation theory, wavelets and applications (Maratea, 1994)*, S.P. Singh, A. Carbone, and B. Watson (editors), NATO Advanced Science Institutes Series C: Mathematical and Physical Sciences vol. 454, Kluwer, pp. 107–130, 1995.

[13] F. Deutsch, *Best Approximation in Inner Product Spaces*, Springer, 2001.

[14] 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.

[15] 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.

[16] 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.

[17] J. Dixmier, Étude sur les variétés et les opérateurs de Julia, avec quelques applications, *Bulletin de la Société Mathématique de France* 77 (1949), 11–101.

[18] K. Friedrichs, On certain inequalities and characteristic value problems for analytic functions and for functions of two variables, *Transactions of the AMS* 41 (1937), 321–364.

[19] 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.

[20] 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.

[21] A.S. Lewis and J. Malick, Alternating projection on manifolds, *Mathematics of Operations Research* 33 (2008), 216–234.

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

[23] D.R. Luke, Local linear convergence and approximate projections onto regularized sets, *Nonlinear Analysis* 75 (2012), 1531–1546.

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

[25] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

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

[27] H. Stark and Y. Yang, *Vector Space Projections*, Wiley, 1998.

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

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