# Restricted normal cones and the method of alternating projections

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

May 2, 2012

**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 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. We thoroughly study restricted normal cones from the viewpoint of constraint qualifications and regularity. Numerous examples are provided to illustrate the theory.

**2010 Mathematics Subject Classification:** Primary 49J52, 49M20; Secondary 47H09, 65K05, 65K10, 90C26.

**Keywords:** Constraint qualification, convex set, Friedrichs angle, linear convergence, method of alternating projections, normal cone, projection operator, restricted normal cone, superregularity.

## 1 Introduction

Throughout this paper, we assume that

[*]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.

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

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

| | |
|---|---|
| (1) | $X$ is a Euclidean space |

(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 operator $P_A$ 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 and Norbert Wiener 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., [4], [7], [11], 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 [9]. 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., [25]). 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. We unify their theory with classical convex convergence results. Our principal tool is a new normal cone which we term the* restricted normal cone*. A careful study of restricted normal cones and their applications is carried out. We also allow for constraint sets that are* unions *of superregular* *(or even convex) sets. We shall recover the known optimal convergence rate for the MAP when studying two linear subspaces.* In a parallel paper [5] we apply the tools developed here to the important problem of sparsity optimization with affine constraints.

The remainder of the paper is organized as follows. In Section 2, we collect various auxiliary results that are useful later and to make the later analysis less cluttered. The restricted normal cones are introduced in Section 3. Section 4 focuses on normal cones that are restricted by affine subspaces; the results achieved are critical in the inclusion of convex settings to the linear convergence framework. Further examples and results are provided in Section 5 and Section 6, where we illustrate that the restricted normal cone cannot be obtained by intersections with various natural conical supersets. Section 7 and Section 8 are devoted to constraint qualifications which describe how well the sets $A$ and $B$ relate to each other. In Section 9, we discuss regularity and superregularity, notions that extend the idea of convexity, for sets and collections of sets. We are then in a position to provide in Section 10 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 [6], [22], [23], and [24]; 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 $\rho$ while $\operatorname{sphere}(z;\rho) := \{x \in X \mid d(z,x) = \rho\}$ is the (closed) sphere centered at $z$ with radius $\rho$. 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 results

In this section, we fix some basic notation used throughout this article. We also collect several auxiliary results that will be useful in the sequel.

## 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*

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

*is the corresponding* projection*.*

**Proposition 2.2 (existence)** *Let $A$ be a nonempty closed subset of $X$. Then $(\forall x \in X)\ P_A(x) \neq \varnothing$.*

*Proof.* Let $z \in X$. The function $f\colon X \to \mathbb{R}\colon x \mapsto \|x - z\|^2$ is continuous and $\lim_{\|x\|\to+\infty} f(x) = +\infty$. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $A$ such that $f(x_n) \to \inf f(A)$. Then $(x_n)_{n\in\mathbb{N}}$ is bounded. Since $A$ is closed and $f$ is continuous, every cluster point of $(x_n)_{n\in\mathbb{N}}$ is a minimizer of $f$ over the set $A$, i.e., an element in $P_A z$. ■

**Example 2.3 (sphere)** Let $z \in X$ and $\rho \in \mathbb{R}_{++}$. Set $S := \operatorname{sphere}(z;\rho)$. Then

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

*Proof.* Let $x \in X$. The formula is clear when $x = z$, so we assume $x \neq z$. Set

$$c := z + \rho\frac{x - z}{\|x - z\|} \in S, \tag{7}$$

and let $s = z + \rho b \in S \smallsetminus \{c\}$, i.e., $\|b\| = 1$ and $b \neq (x - z)/\|x - z\|$. Hence, using that $|\|u\| - \|v\|| < \|u - v\| \Leftrightarrow \langle u, v\rangle < \|u\|\|v\|$ and because of Cauchy-Schwarz, we obtain

$$\|x - c\| = \big|\|x - z\| - \rho\big| = \big|\|x - z\| - \|\rho b\|\big| = \big|\|x - z\| - \|s - z\|\big| \tag{8a}$$

$$< \|x - s\|. \tag{8b}$$

We have thus established (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 following result is well known.

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

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

*Proof.* (i)&(ii): [4, Theorem 3.14]. (iii): [4, Proposition 4.8]. (iv): Clear from (iii). ■

## Miscellany

**Lemma 2.6** *Let $A$ and $B$ be subsets of $X$, and let $K$ be a cone in $X$. Then the following hold:*

(i) $\operatorname{cone}(A \cap B) \subseteq \operatorname{cone} A \cap \operatorname{cone} B$.

(ii) $\operatorname{cone}(K \cap B) = K \cap \operatorname{cone} B$.

*Proof.* (i): Clear. (ii): By (i), $\operatorname{cone}(K \cap B) \subseteq (\operatorname{cone} K) \cap (\operatorname{cone} B) = K \cap \operatorname{cone} B$. Now assume that $x \in (K \cap \operatorname{cone} B) \smallsetminus \{0\}$. Then there exists $\beta > 0$ such that $x/\beta \in B$. Since $K$ is a cone, $x/\beta \in K$. Thus $x/\beta \in K \cap B$ and therefore $x \in \operatorname{cone}(K \cap B)$. ■

Note that the inclusion in Lemma 2.6(i) may be strict: indeed, consider the case when $X = \mathbb{R}$, $A := \{1\}$, and $B = \{2\}$.

**Lemma 2.7 (a characterization of convexity)** *Let $A$ be a nonempty closed subset of $X$. Then the following are equivalent:*

(i) *$A$ is convex.*

(ii) $P_A^{-1}(a) - a$ *is a cone, for every* $a \in A$.

(iii) $P_A(x)$ *is a singleton, for every* $x \in X$.

*Proof.* "(i)⇒(ii)": Indeed, it is well known in convex analysis (see, e.g., [24, Proposition 6.17]) that for every $a \in A$, $P_A^{-1}(a) - a$ is equal to the normal cone (in the sense of convex analysis) of $A$ at $a$.

"(ii)⇒(iii)": Let $x \in X$. By Proposition 2.2, $P_A x \neq \varnothing$. Take $a_1$ and $a_2$ in $P_A x$. Then $\|x - a_1\| = \|x - a_2\|$ and $x - a_1 \in P_A^{-1} a_1 - a_1$. Since $P_A^{-1} a - a$ is a cone, we have $2(x - a_1) \in P_A^{-1} a_1 - a_1$. Hence $y := 2x - a_1 \in P_A^{-1} a_1$ and $y - x = x - a_1$. Thus,

(9a) $$\langle y - a_2, a_1 - a_2 \rangle = \langle (y - x) + (x - a_2), (a_1 - x) + (x - a_2) \rangle$$

(9b) $$= \langle y-x, a_1-x\rangle + \langle y-x, x-a_2\rangle + \langle x-a_2, a_1-x\rangle + \|x-a_2\|^2$$
(9c) $$= \langle x-a_1, a_1-x\rangle + \langle x-a_1, x-a_2\rangle + \langle x-a_2, a_1-x\rangle + \|x-a_2\|^2$$
(9d) $$= -\|x-a_1\|^2 + \|x-a_2\|^2$$
(9e) $$= 0.$$

Since $a_1 \in P_A y$, it follows that

(10a) $$\|y-a_1\|^2 = \|y-a_2\|^2 + 2\langle y-a_2, a_2-a_1\rangle + \|a_1-a_2\|^2$$
(10b) $$= \|y-a_2\|^2 + \|a_1-a_2\|^2$$
(10c) $$\geq \|y-a_2\|^2$$
(10d) $$\geq \|y-a_1\|^2.$$

Hence equality holds throughout (10). Therefore, $a_1 = a_2$.

"(iii)⇒(i)": This classical result due to Bunt and to Motzkin on the convexity of Chebyshev sets is well known; for proofs, see, e.g., [11, Chapter 12] or [4, Corollary 21.13]. ■

**Proposition 2.8** *Let $S$ be a convex set. Then the following are equivalent.*

(i) $0 \in \operatorname{ri} S$.

(ii) $\operatorname{cone} S = \operatorname{span} S$.

(iii) $\overline{\operatorname{cone}}\, S = \operatorname{span} S$.

*Proof.* Set $Y = \operatorname{span} S$. Then (i) ⇔ 0 belongs to the interior of $S$ relative to $Y$.

"(i)⇒(ii)": There exists $\delta > 0$ such that for every $y \in Y \smallsetminus \{0\}$, $\delta y/\|y\| \in S$. Hence $y \in \operatorname{cone} S$.

"(ii)⇒(i)": For every $y \in Y$, there exists $\delta > 0$ such that $\delta y \in S$. Now [23, Corollary 6.4.1] applies in $Y$.

"(ii)⇔(iii)": Set $K = \operatorname{cone} S$, which is convex. By [23, Corollary 6.3.1], we have $\operatorname{ri} K = \operatorname{ri} Y \Leftrightarrow \overline{K} = \overline{Y} \Leftrightarrow \operatorname{ri} Y \subseteq K \subseteq \overline{Y}$. Since $\operatorname{ri} Y = Y = \overline{Y}$, we obtain the equivalences: $\operatorname{ri} K = Y \Leftrightarrow \overline{K} = Y \Leftrightarrow K = Y$. ■

## 3 Restricted normal cones: basic properties

Normal cones are fundamental objects in variational analysis; they are used to construct subdifferential operators, and they have found many applications in optimization, optimal control, nonlinear analysis, convex analysis, etc.; see, e.g., [4], [6], [8], [19], [22], [23], [24]. One of the key building blocks is the Mordukhovich (or limiting) normal cone $N_A$, which is obtained by limits of proximal normal vectors. In this section, we propose a new, very flexible, normal cone of $A$, denoted by $N_A^B$, by constraining the proximal normal vectors to a set $B$.

**Definition 3.1 (normal cones)** *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*

$$\widehat{N}_A^B(a) := \operatorname{cone}\Big(\big(B \cap P_A^{-1}a\big) - a\Big) = \operatorname{cone}\Big(\big(B - a\big) \cap \big(P_A^{-1}a - a\big)\Big). \tag{11}$$

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

$$N_A^{\mathrm{prox}}(a) := \widehat{N}_A^X(a) = \operatorname{cone}\big(P_A^{-1}a - a\big). \tag{12}$$

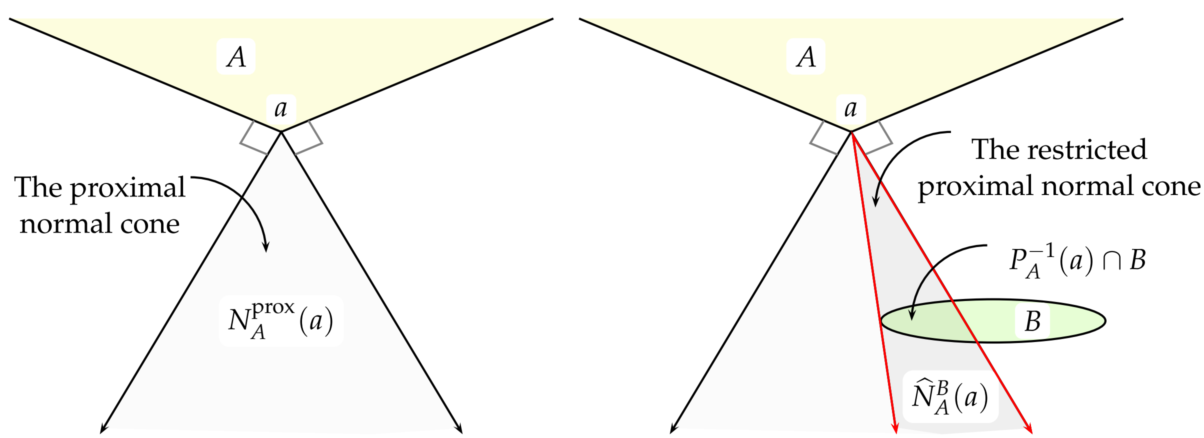
(iii) *The $B$-*restricted normal cone *$N_A^B(a)$ is implicitly defined by $u \in N_A^B(a)$ if and only if there exist sequences $(a_n)_{n\in\mathbb{N}}$ in $A$ and $(u_n)_{n\in\mathbb{N}}$ in $\widehat{N}_A^B(a_n)$ such that $a_n \to a$ and $u_n \to u$.*

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

(v) *The* normal convex from convex analysis *$N_A^{\mathrm{conv}}(a)$ is implicitly defined by $u \in N_A^{\mathrm{conv}}(a)$ 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 $N_A^{\mathrm{prox}}(a_n)$ such that $a_n \to a$ and $u_n \to u$.*

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




**Remark 3.2** Some comments regarding Definition 3.1 are in order.

(i) Clearly, the restricted proximal normal cone generalizes the notion of the classical proximal normal cone. The name "restricted" stems from the fact that the pre-image $P_A^{-1}a$ is restricted to the set $B$.

(ii) See [24, Example 6.16] and [22, Subsection 2.5.2.D on page 240] for further information regarding the classical proximal normal cone, including the fact that

$$u \in N_A^{\text{prox}}(a) \quad \Leftrightarrow \quad a \in A \text{ and } (\exists\, \delta > 0)(\forall x \in A) \ \ \langle u, x - a\rangle \leq \delta \|x - a\|^2. \tag{13}$$

This also implies that: $N_A^{\text{prox}}(a) + (A - a)^{\ominus} \subseteq N_A^{\text{prox}}(a)$.

(iii) Note that $\operatorname{gr} N_A^B = (A \times X) \cap \overline{\operatorname{gr} \widehat{N}_A^B}$. Put differently, $N_A^B(a)$ is the outer (or upper Kuratowski) limit of $\widehat{N}_A^B(x)$ as $x \to a$ in $A$, written

$$N_A^B(a) = \varlimsup_{\substack{x \to a \\ x \in A}} \widehat{N}_A^B(x). \tag{14}$$

See also [24, Chapter 4].

(iv) See [22, Definition 1.1] or [24, Definition 6.3] (where this is called the regular normal cone) for further information regarding $N_A^{\text{Fré}}(a)$.

(v) The Mordukhovich normal cone is also known as the basic or limiting normal cone. Note that $N_A = N_A^X$ and $\operatorname{gr} N_A = (A \times X) \cap \overline{\operatorname{gr} \widehat{N}_A^X} = (A \times X) \cap \overline{\operatorname{gr} N_A^{\text{prox}}}$ and once again $N_A(a)$ is the outer (or upper Kuratowski) limit of $\widehat{N}_A^X(x)$ or $N_A^{\text{prox}}(x)$ as $x \to a$ in $A$. See also [22, page 141] for historical notes.

The next result presents useful characterizations of the Mordukhovich normal cone.

**Proposition 3.3 (characterizations of the Mordukhovich normal cone)** *Let $A$ be a nonempty closed subset of $X$, let $a \in A$, and let $u \in X$. Then the following are equivalent:*

(i) $u \in N_A(a)$.

(ii) *There exist sequences $(\lambda_n)_{n \in \mathbb{N}}$ in $\mathbb{R}_+$, $(b_n)_{n \in \mathbb{N}}$ in $X$, $(a_n)_{n \in \mathbb{N}}$ in $A$ such that $a_n \to a$, $\lambda_n(b_n - a_n) \to u$, and $(\forall n \in \mathbb{N})$ $a_n \in P_A b_n$.*

(iii) *There exist sequences $(\lambda_n)_{n \in \mathbb{N}}$ in $\mathbb{R}_+$, $(x_n)_{n \in \mathbb{N}}$ in $X$, $(a_n)_{n \in \mathbb{N}}$ in $A$ such that $x_n \to a$, $\lambda_n(x_n - a_n) \to u$, and $(\forall n \in \mathbb{N})$ $a_n \in P_A x_n$. (This also implies $a_n \to a$.)*

(iv) *There exist sequences $(a_n)_{n \in \mathbb{N}}$ in $A$ and $(u_n)_{n \in \mathbb{N}}$ in $X$ such that $a_n \to a$, $u_n \to u$, and $(\forall n \in \mathbb{N})$ $u_n \in N_A^{\text{Fré}}(a_n)$.*

*Proof.* "(i)⇔(ii)": Clear from Definition 3.1(vi).

"(iii)⇔(iv)": Noting that the definition of $N_A(a)$ in [22] is the one given in (iv), we see that this equivalence follows from [22, Theorem 1.6].

"(ii)⇒(iii)": Let $(\lambda_n)_{n \in \mathbb{N}}$, $(a_n)_{n \in \mathbb{N}}$, and $(b_n)_{n \in \mathbb{N}}$ be as in (ii). For every $n \in \mathbb{N}$, since $a_n \in P_A b_n$, [24, Example 6.16] implies that $a_n \in P_A[a_n, b_n]$. Now let $(\varepsilon_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1[$ such that $\varepsilon_n a_n \to 0$ and $\varepsilon_n b_n \to 0$. Set

$$(\forall n \in \mathbb{N}) \quad x_n = (1 - \varepsilon_n)a_n + \varepsilon_n b_n = a_n + \varepsilon_n(b_n - a_n) \in [a_n, b_n]. \tag{15}$$

Then $x_n \to a$ and $(\forall n \in \mathbb{N})\ a_n \in P_A x_n$. Furthermore, $(\lambda_n/\varepsilon_n)_{n\in\mathbb{N}}$ lies in $\mathbb{R}_+$ and

$$(\lambda_n/\varepsilon_n)(x_n - a_n) = \lambda_n(b_n - a_n) \to u. \tag{16}$$

"(iii)⇒(ii)": Let $(\lambda_n)_{n\in\mathbb{N}}$, $(x_n)_{n\in\mathbb{N}}$, and $(a_n)_{n\in\mathbb{N}}$ be as in (iii). Since $x_n \to a$ and $a \in A$, we deduce that $0 \leq \|x_n - a_n\| = d_A(x_n) \leq \|x_n - a\| \to 0$. Hence $x_n - a_n \to 0$ which implies that $a_n - a = a_n - x_n + x_n - a \to 0 + 0 = 0$. Therefore, (ii) holds with $(b_n)_{n\in\mathbb{N}} = (x_n)_{n\in\mathbb{N}}$. ■

Here are some basic properties of the restricted normal cone and its relation to various classical cones.

**Lemma 3.4 (basic inclusions among the normal cones)** *Let $A$ and $B$ be nonempty subsets of $X$, and let $a \in A$. Then the following hold:*

(i) $N_A^{\mathrm{conv}}(a) \subseteq N_A^{\mathrm{prox}}(a)$.

(ii) $\widehat{N}_A^B(a) = \operatorname{cone}((B-a) \cap (P_A^{-1}a - a)) \subseteq (\operatorname{cone}(B-a)) \cap N_A^{\mathrm{prox}}(a)$.

(iii) $\widehat{N}_A^B(a) \subseteq \widehat{N}_A^X(a) = N_A^{\mathrm{prox}}(a)$ *and* $N_A^B(a) \subseteq N_A(a)$.

(iv) $\widehat{N}_A^B(a) \subseteq N_A^B(a)$.

(v) *If $A$ is closed, then* $N_A^{\mathrm{prox}}(a) \subseteq N_A^{\mathrm{Fré}}(a)$.

(vi) *If $A$ is closed, then* $N_A^{\mathrm{Fré}}(a) \subseteq N_A(a)$.

(vii) *If $A$ is closed and convex, then* $\widehat{N}_A^X(a) = N_A^{\mathrm{prox}}(a) = N_A^{\mathrm{Fré}}(a) = N_A^{\mathrm{conv}}(a) = N_A(a)$.

(viii) *If $a \in \operatorname{ri}(A)$, then* $\widehat{N}_A^{\operatorname{aff}(A)}(a) = N_A^{\operatorname{aff}(A)}(a) = \{0\}$.

(ix) $(\operatorname{aff}(A) - a)^{\perp} \subseteq (A-a)^{\ominus}$.

(x) $(A-a)^{\ominus} \cap \operatorname{cone}(B-a) \subseteq \widehat{N}_A^B(a) \subseteq \operatorname{cone}(B-a)$.

*Proof.* (i): Take $u \in N_A^{\mathrm{conv}}(a)$ and fix an arbitrary $\delta > 0$. Then $(\forall x \in A)\ \langle u, x-a\rangle \leq 0 \leq \delta\|x-a\|^2$. In view of (13), $u \in N_A^{\mathrm{prox}}(a)$.

(ii): In view of Lemma 2.6, the definitions yield

$$\begin{aligned}
\widehat{N}_A^B(a) &= \operatorname{cone}\big((B \cap P_A^{-1}a) - a\big) = \operatorname{cone}\big((B-a) \cap (P_A^{-1}a - a)\big) && \text{(17a)}\\
&\subseteq \operatorname{cone}\big((B-a) \cap \operatorname{cone}(P_A^{-1}a - a)\big) = \operatorname{cone}\big((B-a) \cap N_A^{\mathrm{prox}}(a)\big) && \text{(17b)}\\
&= \operatorname{cone}(B-a) \cap N_A^{\mathrm{prox}}(a). && \text{(17c)}
\end{aligned}$$

(iii), (iv) and (ix): This is obvious.

(v): Assume that $A$ is closed and take $u \in N_A^{\text{prox}}(a)$. By (13), there exists $\rho > 0$ such that $(\forall x \in A)\ \langle u, x-a\rangle \leq \rho\|x-a\|^2$. Now let $\varepsilon > 0$ and set $\delta = \varepsilon/\rho$. If $x \in A \cap \text{ball}(a;\delta)$, then $\langle u, x-a\rangle \leq \rho\|x-a\|^2 \leq \rho\delta\|x-a\| = \varepsilon\|x-a\|$. Thus, $u \in N_A^{\text{Fré}}(a)$.

(vi): This follows from Proposition 3.3.

(vii): Since $A$ is closed, it follows from (i), (v), and (vi) that

$$N_A^{\text{conv}}(a) \subseteq N_A^{\text{prox}}(a) \subseteq N_A^{\text{Fré}}(a) \subseteq N_A(a). \tag{18}$$

On the other hand, by [22, Proposition 1.5], $N_A(a) \subseteq N_A^{\text{conv}}(a)$ because $A$ is convex.

(viii): By assumption, $(\exists\,\delta > 0)\ \text{ball}(a;\delta) \cap \text{aff}(A) \subseteq A$. Hence $\text{aff}(A) \cap P_A^{-1}a = \{a\}$ and thus $\widehat{N}_A^{\text{aff}(A)}(a) = \{0\}$. Since $a \in \text{ri}(A)$, it follows that $(\forall x \in \text{ball}(a;\delta/2) \cap \text{aff}(A))\ \widehat{N}_A^{\text{aff}(A)}(x) = \{0\}$. Therefore, $N_A^{\text{aff}(A)}(a) = \{0\}$.

(x): Take $u \in ((A-a)^{\ominus} \cap \text{cone}(B-a)) \smallsetminus \{0\}$, say $u = \lambda(b-a)$, where $b \in B$ and $\lambda > 0$. Then $0 \geq \sup\langle A-a, u\rangle = \lambda\sup\langle A-a, b-a\rangle = \sup\lambda\,\langle\overline{\text{conv}}\,A-a, b-a\rangle$. By Fact 2.5(ii), $a = P_{\overline{\text{conv}}\,A}b$ and hence $a = P_A b$. It follows that $u \in \text{cone}((B \cap P_A^{-1}a) - a)$. The left inclusion thus holds. The right inclusion is clear. ■

**Remark 3.5 (on closedness of normal cones)** Let $A$ be a nonempty subset of $X$, let $a \in A$, and let $B$ be a subset of $X$. Then $N_A^B(a)$, $N_A(a)$, and $N_A^{\text{conv}}(a)$ are obviously closed—this is also true for $N_A^{\text{Fré}}(a)$ but requires some work (see [24, Proposition 6.5]). On the other hand, the classical proximal normal cone $N_A^{\text{prox}}(a) = \widehat{N}_A^X(a)$ is not necessarily closed (see, e.g., [24, page 213]), and hence neither is $\widehat{N}_A^B(a)$. For a concrete example, suppose that $X = \mathbb{R}^2$, that $A = \{(0,0)\}$, that $B = \mathbb{R} \times \{1\}$ and that $a = (0,0)$. Then $\widehat{N}_A^B(a) = \big(\mathbb{R} \times \mathbb{R}_{++}\big) \cup \{(0,0)\}$, which is not closed; however, the classical proximal normal cone $N_A^{\text{prox}}(a) = \mathbb{R}^2$ is closed.

The sphere is a nonconvex set for which all classical normal cones coincide:

**Example 3.6 (classical normal cones of the sphere)** Let $z \in X$ and $\rho \in \mathbb{R}_{++}$. Set $S := \text{sphere}(z;\rho)$ and let $s \in S$. Then $N_S^{\text{prox}}(s) = \widehat{N}_S^X(s) = N_S^{\text{Fré}}(s) = N_S(s) = \mathbb{R}(s-z)$.

*Proof*. By Example 2.3, we have $P_S^{-1}(s) = z + \mathbb{R}_+(s-z)$ and so $P_S^{-1}(s) - s = [-1,+\infty[\,\cdot\,(s-z)$. Hence, using Lemma 3.4(v)&(vi), we have

$$N_S^{\text{prox}}(s) = \widehat{N}_S^X(s) = \mathbb{R}(s-z) \subseteq N_S^{\text{Fré}}(s) \subseteq N_S(s) \tag{19a}$$

$$= \varlimsup_{\substack{s'\to s\\ s'\in S}} N_S^{\text{prox}}(s') = \varlimsup_{\substack{s'\to s\\ s'\in S}} \mathbb{R}(s'-z) = \mathbb{R}(s-z) \tag{19b}$$

$$= N_S^{\text{prox}}(s), \tag{19c}$$

as announced. ■

Here are some elementary yet useful calculus rules.

**Proposition 3.7** *Let $A$, $A_1$, $A_2$, $B$, $B_1$, and $B_2$ be nonempty subsets of $X$, let $c \in X$, and suppose that $a \in A \cap A_1 \cap A_2$. Then the following hold:*

(i) *If $A$ and $B$ are convex, then $\widehat{N}_A^B(a)$ is convex.*

(ii) $\widehat{N}_A^{B_1 \cup B_2}(a) = \widehat{N}_A^{B_1}(a) \cup \widehat{N}_A^{B_2}(a)$ *and* $N_A^{B_1 \cup B_2}(a) = N_A^{B_1}(a) \cup N_A^{B_2}(a)$.

(iii) *If $B \subseteq A$, then $\widehat{N}_A^B(a) = N_A^B(a) = \{0\}$.*

(iv) *If $A_1 \subseteq A_2$, then $\widehat{N}_{A_2}^B(a) \subseteq \widehat{N}_{A_1}^B(a)$.*

(v) $-\widehat{N}_A^B(a) = \widehat{N}_{-A}^{-B}(-a)$, $-N_A^B(a) = N_{-A}^{-B}(-a)$, *and* $-N_A(a) = N_{-A}(-a)$.

(vi) $\widehat{N}_A^B(a) = \widehat{N}_{A-c}^{B-c}(a-c)$ *and* $N_A^B(a) = N_{A-c}^{B-c}(a-c)$.

*Proof*. It suffices to establish the conclusions for the restricted proximal normal cones since the restricted normal cone results follows by taking closures (or outer limits). (i): We assume that $B \cap P_A^{-1}a \neq \varnothing$, for otherwise the conclusion is clear. Then $P_A^{-1}(a) = P_{\overline{A}}^{-1}a = (\mathrm{Id} + N_{\overline{A}})a$ is convex (as the image of the maximally monotone operator $\mathrm{Id} + N_{\overline{A}}$ at $a$). Hence $(B \cap P_A^{-1}a) - a$ is convex as well, and so is its conical hull, which is $\widehat{N}_A^B(a)$. (ii): Since $((B_1 \cup B_2) \cap P_A^{-1}a) - a = ((B_1 \cap P_A^{-1}a) - a) \cup ((B_2 \cap P_A^{-1}a) - a)$, the result follows by taking the conical hull. (iii): Clear, because $(B \cap P_A^{-1}a) - a$ is either empty or equal to $\{0\}$. (iv): Suppose $\lambda(b-a) \in \widehat{N}_{A_2}^B(a)$, where $\lambda \geq 0$, $b \in B$, and $a \in P_{A_2}b$. Since $a \in A_1 \subseteq A_2$, we have $a \in P_{A_1}b$. Hence $\lambda(b-a) \in \widehat{N}_{A_1}^B(a)$. (v): This follows by using elementary manipulations and the fact that $P_{-A} = (-\mathrm{Id}) \circ P_A \circ (-\mathrm{Id})$. (vi): This follows readily from the fact that $P_{A-c}^{-1}(a-c) = P_A^{-1}(a) - c$. ■

**Remark 3.8** The restricted normal cone counterparts of items (i) and (iv) are false in general; see Example 5.1 (and also Example 5.4(iv)) below.

The Mordukhovich normal cone (and hence also the Clarke normal cone which contains the Mordukhovich normal cone) strictly contains $\{0\}$ at boundary points (see [22, Corollary 2.24] or [24, Exercise 6.19]); however, the restricted normal cone can be $\{0\}$ at boundary points as we illustrate next.

**Example 3.9 (restricted normal cone at boundary points)** Suppose that $X = \mathbb{R}^2$, set $A := \mathrm{ball}(0;1) = \{x \in \mathbb{R}^2 \mid \|x\| \leq 1\}$ and $B := \mathbb{R} \times \{2\}$, and let $a = (a_1, a_2) \in A$. Then

$$\widehat{N}_A^B(a) = \begin{cases} \mathbb{R}_+ a, & \text{if } \|a\| = 1 \text{ and } a_2 > 0; \\ \{(0,0)\}, & \text{otherwise.} \end{cases} \tag{20}$$

Consequently,

$$N_A^B(a) = \begin{cases} \mathbb{R}_+ a, & \text{if } \|a\| = 1 \text{ and } a_2 \geq 0; \\ \{(0,0)\}, & \text{otherwise.} \end{cases} \tag{21}$$

Thus the restricted normal cone is $\{(0,0)\}$ for all boundary points in the lower half disk that do not "face" the set $B$.

**Remark 3.10** In contrast to Example 3.9, we shall see in Corollary 4.11(ii) below that if $A$ is closed, $B$ is the affine hull of $A$, and $a$ belongs to the relative boundary of $A$, then the restricted normal cone $N_A^B(a)$ strictly contains $\{0\}$.

# 4 Restricted normal cones and affine subspaces

In this section, we consider the case when the restricting set is a suitable affine subspace. This results in further calculus rules and a characterization of interiority notions.

The following four lemmas are useful in the derivation of the main results in this section.

**Lemma 4.1** *Let $A$ and $B$ be nonempty subsets of $X$, and suppose that $c \in A \cap B$. Then*

$$\operatorname{aff}(A \cup B) - c = \operatorname{span}(B - A). \tag{22}$$

*Proof.* Since $c \in A \cap B \subseteq A \cup B$, it is clear that the $\operatorname{aff}(A \cup B) - c$ is a subspace. On the one hand, if $a \in A$ and $b \in B$, then $b - a = 1 \cdot b + (-1) \cdot a + 1 \cdot c - c \in \operatorname{aff}(A \cup B) - c$. Hence $B - A \subseteq \operatorname{aff}(A \cup B) - c$ and thus $\operatorname{span}(B - A) \subseteq \operatorname{aff}(A \cup B) - c$. On the other hand, if $x \in \operatorname{aff}(A \cup B)$, say $x = \sum_{i \in I} \lambda_i a_i + \sum_{j \in J} \mu_j b_j$, where each $a_i$ belongs to $A$, each $b_j$ belongs to $B$, and $\sum_{i \in I} \lambda_i + \sum_{j \in J} \mu_j = 1$, then $x - c = \sum_{i \in I} (-\lambda_i)(c - a_i) + \sum_{j \in J} \mu_j (b_j - c) \in \operatorname{span}(B - A)$. Thus $\operatorname{aff}(A \cup B) - c \subseteq \operatorname{span}(B - A)$. ■

**Lemma 4.2** *Let $A$ be a nonempty subset of $X$, let $a \in A$, and let $u \in (\operatorname{aff}(A) - a)^\perp$. Then*

$$(\forall x \in X) \quad P_A(x + u) = P_A(x). \tag{23}$$

*Proof.* Let $x \in X$. For every $b \in A$, we have

$$\|u + x - b\|^2 = \|u\|^2 + 2\langle u, x - b\rangle + \|x - b\|^2 \tag{24a}$$

$$= \|u\|^2 + 2\langle u, x - a\rangle + 2\langle u, a - b\rangle + \|x - b\|^2 \tag{24b}$$

$$= \|u\|^2 + 2\langle u, x - a\rangle + \|x - b\|^2. \tag{24c}$$

Hence $P_A(x + u) = \operatorname{argmin}_{b \in A} \|u + x - b\|^2 = \operatorname{argmin}_{b \in A} \|x - b\|^2 = P_A x$, as announced. ■

**Lemma 4.3** *Let $A$ be a nonempty subset of $X$, and let $L$ be an affine subspace of $X$ containing $A$. Then*

$$P_A = P_A \circ P_L. \tag{25}$$

*Proof.* Let $a \in A$ and $x \in X$, and set $b = P_L x$. Using [4, Corollary 3.20(i)], we have $x - b \in (L - a)^\perp \subset (\operatorname{aff}(A) - a)^\perp$. In view of Lemma 4.2, we deduce that $(P_A \circ P_L)x = P_A(b) = P_A(b + (x - b)) = P_A x$. ■

**Lemma 4.4** *Let $A$ be a nonempty subset of $X$, let $a \in A$, and let $L$ be an affine subspace of $X$ containing $A$. Then the following hold:*

(i) $\widehat{N}^L_A(a) \perp (L-a)^\perp$.

(ii) $N^L_A(a) \perp (L-a)^\perp$.

*Proof*. Observe that $L - a = \operatorname{par}(A)$ does not depend on the concrete choice of $a \in A$. (i): Using Lemma 3.4(x), we see that $\widehat{N}^L_A(a) \subseteq \operatorname{cone}(L-a) \subseteq \operatorname{span}(L-a) \perp (\operatorname{span}(L-a))^\perp = (L-a)^\perp = (\operatorname{par} A)^\perp$. (ii): By (i), $\operatorname{ran} \widehat{N}^L_A \subseteq \operatorname{par} A$. Since $\operatorname{ran} N^L_A \subseteq \overline{\operatorname{ran} \widehat{N}^L_A}$, it follows that $\operatorname{ran} N^L_A \subseteq \operatorname{par} A = L - a$. ■

For a normal cone restricted to certain affine subspaces, it is possible to derive precise relationships to the Mordukhovich normal cone.

**Theorem 4.5 (restricted vs Mordukhovich normal cone)** *Let $A$ and $B$ be nonempty subsets of $X$, suppose that $a \in A$, and let $L$ be an affine subspace of $X$ containing $A$. Then the following hold:*

$$\widehat{N}^X_A(a) = \widehat{N}^L_A(a) \oplus (L-a)^\perp = \widehat{N}^X_A(a) + (L-a)^\perp, \tag{26a}$$

$$\widehat{N}^L_A(a) = \widehat{N}^X_A(a) \cap (L-a), \tag{26b}$$

$$N_A(a) = N^L_A(a) \oplus (L-a)^\perp = N_A(a) + (L-a)^\perp, \tag{26c}$$

$$N^L_A(a) = N_A(a) \cap (L-a). \tag{26d}$$

*Consequently, the following hold as well:*

$$\widehat{N}^X_A(a) = \widehat{N}^{\operatorname{aff}(A)}_A(a) \oplus (\operatorname{aff}(A)-a)^\perp = \widehat{N}^X_A(a) + (\operatorname{aff}(A)-a)^\perp, \tag{27a}$$

$$\widehat{N}^{\operatorname{aff}(A)}_A(a) = \widehat{N}^X_A(a) \cap (\operatorname{aff}(A)-a), \tag{27b}$$

$$N_A(a) = N^{\operatorname{aff}(A)}_A(a) \oplus \big(\operatorname{aff}(A)-a\big)^\perp = N_A(a) + \big(\operatorname{aff}(A)-a\big)^\perp, \tag{27c}$$

$$N^{\operatorname{aff}(A)}_A(a) = N_A(a) \cap \big(\operatorname{aff}(A)-a\big), \tag{27d}$$

$$a \in A \cap B \;\Rightarrow\; N^{\operatorname{aff}(A\cup B)}_A(a) = N_A(a) \cap \operatorname{span}(A-B). \tag{27e}$$

*Proof*. (26a): Take $u \in \widehat{N}^X_A(a)$. Then there exist $\lambda \geq 0$, $x \in X$, and $a \in P_A x$ such that $\lambda(x-a) = u$. Set $b = P_L x$. By Lemma 4.3, we have $a \in P_A x = (P_A \circ P_L)x = P_A b$. Using [4, Corollary 3.20(i)], we thus deduce that $\lambda(b-a) \in \widehat{N}^L_A(a)$ and $\lambda(x-b) \in (L-b)^\perp = (L-a)^\perp$. Hence $u = \lambda(b-a) + \lambda(x-b) \in \widehat{N}^L_A(a) + (L-a)^\perp = \widehat{N}^L_A(a) \oplus (L-a)^\perp$ by Lemma 4.4(i). We have thus shown that

$$\widehat{N}^X_A(a) \subseteq \widehat{N}^L_A(a) \oplus (L-a)^\perp. \tag{28}$$

On the other hand, Lemma 3.4(iii) implies that $\widehat{N}^L_A(a) \subseteq \widehat{N}^X_A(a)$ and thus

$$\widehat{N}^L_A(a) + (L-a)^\perp \subseteq \widehat{N}^X_A(a) + (L-a)^\perp. \tag{29}$$

Altogether,

$$\widehat{N}_A^X(a) \subseteq \widehat{N}_A^L(a) \oplus (L-a)^\perp \subseteq \widehat{N}_A^X(a) + (L-a)^\perp. \tag{30}$$

To complete the proof of (26a), it thus suffices to show that $\widehat{N}_A^X(a) + (L-a)^\perp \subseteq \widehat{N}_A^X(a)$. To this end, let $u \in \widehat{N}_A^X(a)$ and $v \in (L-a)^\perp \subseteq (\operatorname{aff}(A)-a)^\perp$. Then there exist $\lambda \geq 0$, $b \in X$, and $a \in P_A b$ such that $u = \lambda(b-a)$. If $\lambda = 0$, then $u = 0$ and $u + v = v \in (\operatorname{aff}(A)-a)^\perp \subseteq (A-a)^\ominus = (A-a)^\ominus \cap X = (A-a)^\ominus \cap \operatorname{cone}(X-a) \subseteq \widehat{N}_A^X(a)$ by Lemma 3.4(ix)&(x). Thus, we assume that $\lambda > 0$. By Lemma 4.2, we have $a \in P_A b = P_A(b + \lambda^{-1}v)$. Hence $b + \lambda^{-1}v - a \in \widehat{N}_A^X(a)$ and therefore $\lambda(b + \lambda^{-1}v - a) = \lambda(b-a) + v = u + v \in \widehat{N}_A^X(a)$, as required.

(26b): By Lemma 3.4(iii)&(x), $\widehat{N}_A^L(a) \subseteq \widehat{N}_A^X(a) \cap (L-a)$. Now let $u \in \widehat{N}_A^X(a) \cap (L-a)$. By (26a), we have $u = v + w$, where $v \in \widehat{N}_A^L(a) \subseteq L - a$ and $w \in (L-a)^\perp$. On the other hand, $w = u - v \in (L-a) - (L-a) = L - a$. Altogether $w \in (L-a) \cap (L-a)^\perp = \{0\}$. Hence $u = v \in \widehat{N}_A^L(a)$.

(26c): Let $u \in N_A(a)$. By definition, there exist sequences $(a_n)_{n\in\mathbb{N}}$ in $A$ and $(u_n)_{n\in\mathbb{N}}$ in $X$ such that $a_n \to a$, $u_n \to u$, and $(\forall n \in \mathbb{N})$ $u_n \in \widehat{N}_A^X(a_n)$. By (26a), there exists a sequence $(v_n, w_n)_{n\in\mathbb{N}}$ such that $(a_n, v_n)_{n\in\mathbb{N}}$ lies in $\operatorname{gr} \widehat{N}_A^L$, $(w_n)_{n\in\mathbb{N}}$ lies in $(L-a)^\perp$, and $(\forall n \in \mathbb{N})$ $u_n = v_n + w_n$ and $v_n \perp w_n$. Since $\|u\|^2 \leftarrow \|u_n\|^2 = \|v_n\|^2 + \|w_n\|^2$, the sequences $(v_n)_{n\in\mathbb{N}}$ and $(w_n)_{n\in\mathbb{N}}$ are bounded. After passing to subsequences and relabeling if necessary, we assume $(v_n)_{n\in\mathbb{N}}$ and $(w_n)_{n\in\mathbb{N}}$ are convergent, with limits $v$ and $w$, respectively. It follows that $v \in N_A^L(a)$ and $w \in (L-a)^\perp$; consequently, $u = v + w \in N_A^L(a) \oplus (L-a)^\perp$ by Lemma 4.4(ii). Thus $N_A(a) \subseteq N_A^L(a) \oplus (L-a)^\perp$. On the other hand, by Lemma 3.4(iii), $N_A^L(a) \oplus (L-a)^\perp \subseteq N_A(a) + (L-a)^\perp$. Altogether,

$$N_A(a) \subseteq N_A^L(a) \oplus (L-a)^\perp \subseteq N_A(a) + (L-a)^\perp. \tag{31}$$

It thus suffices to prove that $N_A(a) + (L-a)^\perp \subseteq N_A(a)$. To this end, take $u \in N_A(a)$ and $v \in (L-a)^\perp$. Then there exist sequences $(a_n)_{n\in\mathbb{N}}$ in $A$ and $(u_n)_{n\in\mathbb{N}}$ in $X$ such that $a_n \to a$, $u_n \to u$, and $(\forall n \in \mathbb{N})$ $u_n \in \widehat{N}_A^X(a_n)$. For every $n \in \mathbb{N}$, we have $L - a = L - a_n$ and hence $u_n + v \in \widehat{N}_A^X(a_n) + (L-a_n)^\perp = \widehat{N}_A^X(a_n)$ by (26a). Passing to the limit, we conclude that $u + v \in N_A(a)$.

(26d): First, take $u \in N_A^L(a)$. On the one hand, by Lemma 3.4(iii), $u \in N_A(a)$. On the other hand, by Lemma 4.4(ii), $u \in (L-a)^{\perp\perp} = L - a$. Altogether, we have shown that

$$N_A^L(a) \subseteq N_A(a) \cap (L-a). \tag{32}$$

Conversely, take $u \in N_A(a) \cap (L-a) \subseteq N_A(a)$. By (26c), there exist $v \in N_A^L(a)$ and $w \in (L-a)^\perp$ such that $u = v + w$ and $v \perp w$. By (32), $v \in L - a$. Hence $w = u - v \in (L-a) - (L-a) = L - a$. Since $w \in (L-a)^\perp$, we deduce that $w = 0$. This implies $u = v \in N_A^L(a)$. Therefore, $N_A(a) \cap (L-a) \subseteq N_A^L(a)$.

"Consequently" part: Consider (26) when $L = \operatorname{aff}(A)$ or $L = \operatorname{aff}(A \cup B)$, and recall Lemma 4.1 in the latter case. ■

An immediate consequence of Theorem 4.5 (or of the definitions) is the following result.

**Corollary 4.6 (the $X$-restricted and the Mordukhovich normal cone coincide)** *Let $A$ be a nonempty subset of $X$, and let $a \in A$. Then*

$$N_A^X(a) = N_A(a). \tag{33}$$

The next two results provide some useful calculus rules.

**Corollary 4.7 (restricted normal cone of a sum)** *Let $C_1$ and $C_2$ be nonempty closed convex subsets of $X$, let $a_1 \in C_1$, let $a_2 \in C_2$, and let $L$ be an affine subspace of $X$ containing $C_1 + C_2$. Then*

$$N_{C_1+C_2}^L(a_1 + a_2) = N_{C_1}^{L-a_2}(a_1) \cap N_{C_2}^{L-a_1}(a_2). \tag{34}$$

*Proof.* Set $C = C_1 + C_2$ and $a = a_1 + a_2$. Then (26d) and [24, Exercise 6.44] yield

$$N_C^L(a) = N_C(a) \cap (L - a) = N_{C_1}(a_1) \cap N_{C_2}(a_2) \cap (L - a) \tag{35a}$$

$$= \big(N_{C_1}(a_1) \cap (L - a)\big) \cap \big(N_{C_2}(a_2) \cap (L - a)\big). \tag{35b}$$

Note that $L - a$ is a linear subspace of $X$ containing $C_1 - a_1$ and $C_2 - a_2$. Thus, $L - a_2 = L - a + a_1$ is an affine subspace of $X$ containing $C_1$, and $L - a_1 = L - a + a_2$ is an affine subspace of $X$ containing $C_2$. By (26d),

$$N_{C_1}^{L-a_2}(a_1) = N_{C_1}(a_1) \cap (L - a) \quad \text{and} \quad N_{C_2}^{L-a_1}(a_2) = N_{C_2}(a_2) \cap (L - a). \tag{36}$$

The conclusion follows by combining (35) and (36). ■

**Corollary 4.8 (an intersection formula)** *Let $A$ and $B$ be nonempty closed convex subsets of $X$, and suppose that $a \in A \cap B$. Let $L$ be an affine subspace of $X$ containing $A \cup B$. Then*

$$N_A^L(a) \cap \big(-N_B^L(a)\big) = N_{A-B}^{L-a}(0). \tag{37}$$

*Proof.* Using (26d), Proposition 3.7(v), [24, Exercise 6.44], and again (26d), we obtain

$$N_A^L(a) \cap \big(-N_B^L(a)\big) = N_A(a) \cap (L - a) \cap \big(-N_B(a)\big) \cap (L - a) \tag{38a}$$

$$= \Big(N_A(a) \cap \big(-N_B(a)\big)\Big) \cap (L - a) \tag{38b}$$

$$= \big(N_A(a) \cap N_{-B}(-a)\big) \cap (L - a) \tag{38c}$$

$$= N_{A-B}(0) \cap (L - a) \tag{38d}$$

$$= N_{A-B}^{L-a}(0), \tag{38e}$$

as required. ■

Let us now work towards relating the restricted normal cone to the (relative and classical) interior and to the boundary of a given set.

**Proposition 4.9** *Let $A$ be a nonempty subset of $X$, let $a \in A$, let $L$ be an affine subspace containing $A$, and suppose that $N_A^L(a) = \{0\}$. Then $L = \operatorname{aff}(A)$.*

*Proof*. Using $0 \in N_A^{\operatorname{aff}(A)}(a) \subseteq N_A^L(a) = \{0\}$ and applying (26c) and (27c), we have

$$N_A(a) = 0 + (L - a)^{\perp} = 0 + (\operatorname{aff}(A) - a)^{\perp}. \tag{39}$$

So $L - a = \operatorname{aff}(A) - a$, i.e., $L = \operatorname{aff}(A)$. ∎

**Theorem 4.10** *Let $A$ and $B$ be nonempty subsets of $X$, and let $a \in A$. Then*

$$N_A^B(a) = \{0\} \quad \Leftrightarrow \quad (\exists\, \delta > 0)\big(\forall x \in A \cap \operatorname{ball}(a;\delta)\big)\ \ P_A^{-1}(x) \cap B \subseteq \{x\}. \tag{40}$$

*Furthermore, if $A$ is closed and $B$ is an affine subspace of $X$ containing $A$, then the following are equivalent:*

(i) $N_A^B(a) = \{0\}$.

(ii) $(\exists\, \rho > 0)$ $\operatorname{ball}(a;\rho) \cap B \subseteq A$.

(iii) $B = \operatorname{aff}(A)$ *and* $a \in \operatorname{ri}(A)$.

*Proof*. Note that $N_A^B(a) = \{0\} \Leftrightarrow (\exists\, \delta > 0)$ $(\forall x \in A \cap \operatorname{ball}(a;\delta))$ $\widehat{N}_A^B(x) = \{0\}$. Hence (40) follows from the definition of $\widehat{N}_A^B(x)$.

Now suppose that $A$ is closed and $B$ is an affine subspace of $X$ containing $A$.

"(i)⇒(ii)": Let $\delta > 0$ be as in (40) and set $\rho := \delta/2$. Let $b \in B(a;\rho) \cap B$, and take $x \in P_A b$, which is possible since $A$ is closed. Then $\|b - x\| = d_A(b) \leq \|b - a\| \leq \rho$ and hence

$$\|x - a\| \leq \|x - b\| + \|b - a\| \leq \rho + \rho = 2\rho = \delta. \tag{41}$$

Using (40), we deduce that $b \in P_A^{-1}(x) \cap B \subseteq \{x\} \subseteq A$.

"(ii)⇒(iii)": It follows that $B = \operatorname{aff}(B) \subseteq \operatorname{aff}(A) \subseteq B$; hence, $B = \operatorname{aff}(A)$. Thus $\operatorname{ball}(a;\rho) \cap \operatorname{aff}(A) \subseteq A$, which means that $a \in \operatorname{ri}(A)$.

"(iii)⇒(i)": Lemma 3.4(viii). ∎

**Corollary 4.11 (interior and boundary characterizations)** *Let $A$ be a nonempty closed subset of $X$, and let $a \in A$. Then the following hold:*

(i) $N_A^{\operatorname{aff}(A)}(a) = \{0\} \Leftrightarrow a \in \operatorname{ri}(A)$.

(ii) $N_A^{\operatorname{aff}(A)}(a) \neq \{0\} \Leftrightarrow a \in A \smallsetminus \operatorname{ri}(A)$.

(iii) $N_A(a) = \{0\} \Leftrightarrow a \in \operatorname{int}(A)$.

(iv) $N_A(a) \neq \{0\} \Leftrightarrow a \in A \smallsetminus \text{int}(A)$.

*Proof.* (i): Apply Theorem 4.10 with $B = \text{aff}(A)$. (ii): Clear from (i). (iii): Apply Theorem 4.10 with $B = X$, and recall Corollary 4.6. (iv): Clear from (iii). ■

A second look at the proof of (i)⇒(ii) in Theorem 4.10 reveals that this implication does actually not require the assumption that $B$ be an affine subspace of $X$ containing $A$. The following example illustrates that the converse implication fails even when $B$ is a superset of $\text{aff}(A)$.

**Example 4.12** Suppose that $X = \mathbb{R}^2$, and set $A := \mathbb{R} \times \{0\}$, $a = (0,0)$, and $B = \mathbb{R} \times \{0,2\}$. Then $A = \text{aff}(A) \subseteq B$ and $\text{ball}(a;1) \cap B \subseteq A$; however, $(\forall x \in A)\ \widehat{N}_A^B(x) = \{0\} \times \mathbb{R}_+$ and therefore $N_A^B(a) = \{0\} \times \mathbb{R}_+ \neq \{(0,0)\}$.

## Two convex sets

It is instructive to interpret the previous results for two convex sets:

**Theorem 4.13 (two convex sets: restricted normal cones and relative interiors)** *Let $A$ and $B$ be nonempty convex subsets of $X$. Then the following are equivalent:*

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

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

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

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

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

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

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

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

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

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

*Proof.* By [23, Corollary 6.6.2], (ii) $\Leftrightarrow$ $\text{ri}\, A \cap \text{ri}\, B \neq \varnothing \Leftrightarrow 0 \in \text{ri}\, A - \text{ri}\, B \Leftrightarrow$ (ii).

Applying Proposition 2.8 to $B - A$, and [3, Proposition 3.1.3] to $\overline{\text{cone}}\,(B - A)$, we obtain

$$\text{(ii)} \Leftrightarrow \text{(iii)} \ \Leftrightarrow\ \overline{\text{cone}}\,(B - A) = \text{span}(B - A) \tag{42a}$$

$$\text{(42b)} \qquad \Leftrightarrow \overline{\operatorname{cone}}\,(B-A) \cap \big(\overline{\operatorname{cone}}\,(B-A)\big)^{\oplus} = \{0\}.$$

Let $c \in A \cap B$. Then Corollary 4.8 (with $L = X$) yields $N_A(c) \cap \big(-N_B(c)\big) = N_{A-B}(0) = (A - B)^{\ominus} = (B-A)^{\oplus} = (\overline{\operatorname{cone}}(B-A))^{\oplus}$. Hence

$$\text{(43)} \qquad (\forall c \in C) \quad N_A(c) \cap \big(-N_B(c)\big) \cap \overline{\operatorname{cone}}\,(B-A) = \big(\overline{\operatorname{cone}}\,(B-A)\big)^{\oplus} \cap \overline{\operatorname{cone}}\,(B-A)$$

and

$$\text{(44)} \qquad (\forall c \in C) \quad N_A(c) \cap \big(-N_B(c)\big) \cap \operatorname{span}(B-A) = \big(\overline{\operatorname{cone}}\,(B-A)\big)^{\oplus} \cap \operatorname{span}(B-A).$$

Combining (42), (43), and (44), we see that (ii)–(vii) are equivalent.

Next, Lemma 4.1 and Corollary 4.8 yield the equivalence of (viii)–(x).

Finally, (x)⇔(ii) by Corollary 4.11(i). ■

**Corollary 4.14 (two convex sets: normal cones and interiors)** *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) $N_{A-B}(0) = \{0\}$.

*Proof*. We start by notating that if $C$ is a convex subset of $X$, then $0 \in \operatorname{int} C \Leftrightarrow 0 \in \operatorname{ri} C$ and $\operatorname{span} C = X$. Consequently,

$$\text{(45)} \qquad \text{(i)} \quad \Leftrightarrow \quad 0 \in \operatorname{ri}(B-A) \text{ and } \operatorname{span}(B-A) = X.$$

Assume that (i) holds. Then (45) and Theorem 4.13 imply that $\operatorname{cone}(B-A) = \overline{\operatorname{cone}}\,(B-A) = \operatorname{span}(B-A) = X$. Hence (ii) holds, and from Theorem 4.13 we obtain that (ii)⇒(iii)⇔(iv)⇔(v). Finally, Corollary 4.11(iii) yields the implication (v)⇒(i). ■

## 5 Further examples

In this section, we provide further examples that illustrate particularities of restricted normal cones.

As announced in Remark 3.8, when $a \in A_2 \subsetneqq A_1$, it is possible that the *nonconvex* restricted normal cones satisfy $N^B_{A_1}(a) \not\subseteq N^B_{A_2}(a)$ even when $A_1$ and $A_2$ are both *convex*. This lack of inclusion is also known for the Mordukhovich normal cone (see [22, page 5], where however one of the sets is not convex). Furthermore, the following example also shows that the restricted normal cone cannot be derived from the Mordukhovich normal cone by the simple relativization procedure of intersecting with naturally associated cones and subspaces.

**Example 5.1 (lack of convexity, inclusion, and relativization)** Suppose that $X = \mathbb{R}^2$, and define two nonempty closed *convex* sets by $A := A_1 := \operatorname{epi}(|\cdot|)$ and $A_2 := \operatorname{epi}(2|\cdot|)$. Then $a := (0,0) \in A_2 \subsetneqq A_1$. Furthermore, set $B := \mathbb{R} \times \{0\}$. Then

$$(\forall x = (x_1, x_2) \in A_1) \quad \widehat{N}^B_{A_1}(x) = \begin{cases} \mathbb{R}_+(1,-1), & \text{if } x_2 = x_1 > 0; \\ \mathbb{R}_+(-1,-1), & \text{if } x_2 = -x_1 > 0; \\ \{(0,0)\}, & \text{otherwise,} \end{cases} \tag{46a}$$

$$(\forall x = (x_1, x_2) \in A_2) \quad \widehat{N}^B_{A_2}(x) = \begin{cases} \mathbb{R}_+(2,-1), & \text{if } x_2 = 2x_1 > 0; \\ \mathbb{R}_+(-2,-1), & \text{if } x_2 = -2x_1 > 0; \\ \{(0,0)\}, & \text{otherwise.} \end{cases} \tag{46b}$$

Consequently,

$$N^B_{A_1}(a) = \operatorname{cone}\big\{(1,-1), (-1,-1)\big\}, \tag{47a}$$

$$N^B_{A_2}(a) = \operatorname{cone}\big\{(2,-1), (-2,-1)\big\}. \tag{47b}$$

Note that $N^B_{A_1}(a) \not\subseteq N^B_{A_2}(a)$ and $N^B_{A_2}(a) \not\subseteq N^B_{A_1}(a)$; in fact, $N^B_{A_1}(a) \cap N^B_{A_2}(a) = \{(0,0)\}$. Furthermore, neither $N^B_{A_1}(a)$ nor $N^B_{A_2}(a)$ is convex even though $A_1$, $A_2$, and $B$ are. Finally, observe that $\operatorname{cone}(B-a) = \operatorname{span}(B-a) = B$, that $\operatorname{cone}(B-A) = \mathbb{R} \times \mathbb{R}_-$, that $\operatorname{span}(B-A) = X$, and that $N_A(a) = \operatorname{cone}[(1,-1),(-1,-1)] \neq N^B_A(a)$. Consequently, $\operatorname{cone}(B-a) \cap N_A(a) = \operatorname{span}(B-a) \cap N_A(a) = \{(0,0)\}$, $\operatorname{cone}(B-A) \cap N_A(a) = N_A(a) = \operatorname{span}(B-A) \cap N_A(a)$. Therefore, *$N^B_A(a)$ cannot be obtained by intersecting the Mordukhovich normal cone* with one of the sets $\operatorname{cone}(B-a)$, $\operatorname{span}(B-a)$, $\operatorname{cone}(B-A)$, and $\operatorname{span}(B-A)$.

We shall present some further examples. The proof of the following result is straight-forward and hence omitted.

**Proposition 5.2** *Let $K$ be a closed cone in $X$, and let $B$ be a nonempty cone of $X$. Then*

$$N^B_K(0) = \overline{\bigcup_{x \in K} \widehat{N}^B_K(x)} = \overline{\bigcup_{x \in \operatorname{bdry} K} \widehat{N}^B_K(x)} = \overline{\bigcup_{x \in K} N^B_K(x)} = \overline{\bigcup_{x \in \operatorname{bdry} K} N^B_K(x)}. \tag{48}$$

**Example 5.3** Let $K$ be a closed convex cone in $X$, suppose that $u_0 \in \operatorname{int}(K)$ and that $K \subseteq \{u_0\}^{\oplus}$, and set $B := \{u_0\}^{\perp}$. Then:

(i) $(\forall x \in K \cap B)\ \widehat{N}^B_K(x) = \{0\}$.

(ii) $(\forall x \in K \smallsetminus B)\ \widehat{N}^B_K(x) = N^B_K(x) = N_K(x) = K^{\ominus} \cap \{x\}^{\perp}$.

(iii) $N_K^B(0) = \overline{\bigcup_{x\in K} \widehat{N}_K^B(x)} = \overline{\bigcup_{x\in K\smallsetminus B}(K^\ominus \cap \{x\}^\perp)} = \overline{K^\ominus \cap \bigcup_{x\in K\smallsetminus B}\{x\}^\perp}$.

If one of these unions is closed, then all closures may be omitted.

*Proof.* (i): Let $x \in K \cap B$. It suffices to show that $B \cap P_K^{-1}(x) = \{x\}$. To this end, take $y \in B \cap P_K^{-1}(x)$. By definition of $B$, we have $\langle u_0, x\rangle = 0$ and $\langle u_0, y\rangle = 0$. Hence

$$\langle u_0, y - x\rangle = 0. \tag{49}$$

Furthermore, $x = P_K y$ and hence, using e.g. [4, Proposition 6.27], we have $y - x \in K^\ominus$. Since $u_0 \in \operatorname{int} K$, there exists $\delta > 0$ such that $\operatorname{ball}(u_0;\delta) \subseteq K$. Thus $y - x \in (\operatorname{ball}(u_0;\delta))^\ominus$. In view of (49), $\delta\|y - x\| \le 0$. Therefore, $y = x$.

(ii): Let $x \in K \smallsetminus B$. Using Lemma 3.4(iii)&(iv), Corollary 4.6, Lemma 3.4(vii), and [4, Example 6.39], we have

$$\widehat{N}_K^B(x) \subseteq \widehat{N}_K^X(x) \subseteq N_K^X(x) = N_K(x) = N_K^{\mathrm{conv}}(x) = K^\ominus \cap \{x\}^\perp. \tag{50}$$

Since $x \in K \subseteq \{u_0\}^\oplus$ and $x \notin B$, we have $\langle u_0, x\rangle > 0$. Now take $u \in (K^\ominus \cap \{x\}^\perp) \smallsetminus \{0\}$. Since $u \in K^\ominus$ and $u_0 \in \operatorname{int}(K)$, we have $\langle u, u_0\rangle < 0$. Now set

$$b := x - \frac{\langle u_0, x\rangle}{\langle u_0, u\rangle} u. \tag{51}$$

Then $b \in B$ and $b - x = -\langle u_0, x\rangle \langle u_0, u\rangle^{-1} u \in \mathbb{R}_{++}u \subseteq K^\ominus \cap \{x\}^\perp = N_K^{\mathrm{conv}}(x)$. By [4, Proposition 6.46], $x = P_K b$. Hence $b - x \in \widehat{N}_K^B(x)$ and thus $u \in \widehat{N}_K^B(x)$. Therefore, $K^\ominus \cap \{x\}^\perp \subseteq \widehat{N}_K^B(x)$. In view of (50), and since $\widehat{N}_K^B(x) \subseteq N_K^B(x) \subseteq N_K(x)$ by Lemma 3.4(iii)&(iv), we have established (ii).

(iii): Combine (i), (ii), and Proposition 5.2. ■

**Example 5.4 (ice cream cone)** Suppose that $X = \mathbb{R}^m = \mathbb{R}^{m-1} \times \mathbb{R}$, where $m \in \{2, 3, 4, \ldots\}$, and let $\beta > 0$. Define the corresponding closed convex *ice cream cone* by

$$K := \left\{ x \in \mathbb{R}^m \;\middle|\; \beta\sqrt{x_1^2 + \cdots + x_{m-1}^2} \le x_m \right\}, \tag{52}$$

and set $B := \mathbb{R}^{m-1} \times \{0\}$. Then the following hold:

(i) $\widehat{N}_K^B(0,0) = \{(0,0)\}$.

(ii) $N_K(0,0) = \left\{ y \in \mathbb{R}^m \mid \beta^{-1}\sqrt{y_1^2 + \cdots + y_{m-1}^2} \le -y_m \right\} = \bigcup_{\substack{z\in\mathbb{R}^{m-1}\\ \|z\|\le 1}} \mathbb{R}_+(\beta z, -1)$.

(iii) $(\forall z \in \mathbb{R}^{m-1} \smallsetminus \{0\})\ \widehat{N}_K^B(z, \beta\|z\|) = N_K^B(z, \beta\|z\|) = N_K(z, \beta\|z\|) = \mathbb{R}_+(\beta z, -\|z\|)$.

(iv) $N_K^B(0,0) = \bigcup_{\substack{z\in\mathbb{R}^{m-1}\\ \|z\|=1}} \mathbb{R}_+(\beta z, -1)$, which is a closed cone that is *not convex*.

*Proof*. Clearly, $K$ is closed and convex. Note that $K$ is the lower level set of height 0 of the continuous convex function

$$f\colon \mathbb{R}^m = \mathbb{R}^{m-1} \times \mathbb{R} \to \mathbb{R}\colon x = (z, x_m) \mapsto \beta\|z\| - x_m; \tag{53}$$

hence , by [26, Exercise 2.5(b) and its solution on page 205],

$$\operatorname{int}(K) = \big\{x = (z, x_m) \in \mathbb{R}^{m-1} \times \mathbb{R} \;\big|\; \beta\|z\| < x_m\big\}. \tag{54}$$

Lemma 3.4(iii)&(iv), Corollary 4.6, and Corollary 4.11(iii) imply that

$$\big(\forall x \in \operatorname{int}(K)\big) \quad \widehat{N}^B_K(x) \subseteq \widehat{N}^X_K(x) \subseteq N^X_K(x) = N_K(x) = \{0\}. \tag{55}$$

Write $x = (z, x_m) \in \mathbb{R}^{m-1} \times \mathbb{R} = X$, and assume that $x \in K$. We thus assume that $x \in \operatorname{bdry}(K)$, i.e., $\beta\|z\| = x_m$ by (54), i.e., $x = (z, \beta\|z\|)$. Combining [4, Proposition 16.8] with [26, Corollary 2.9.5] (or [4, Lemma 26.17]) applied to $f$, we obtain

$$N_K\big(z, \beta\|z\|\big) = \operatorname{cone}\big(\beta\partial\|\cdot\|(z) \times \{-1\}\big), \tag{56}$$

where $\partial\|\cdot\|$ denotes the subdifferential operator from convex analysis applied to the Euclidean norm in $\mathbb{R}^{m-1}$. In view of [4, Example 16.25] we thus have

$$N_K\big(z, \beta\|z\|\big) = \begin{cases} \operatorname{cone}\big(\beta\|z\|^{-1}z \times \{-1\}\big), & \text{if } z \neq 0; \\ \operatorname{cone}\big(\operatorname{ball}(0;\beta) \times \{-1\}\big), & \text{if } z = 0. \end{cases} \tag{57}$$

The case $z = 0$ in (57) readily leads to (ii).

Now set $u_0 := (0, 1) \in \mathbb{R}^{m-1} \times \mathbb{R}$. Then $\{u_0\}^\perp = B$ and $\{u_0\}^\oplus = \mathbb{R}^{m-1} \times \mathbb{R}_+ \supseteq K$. Note that $(0, 0) \in K \cap B$ and thus $\widehat{N}^B_K(0, 0) = \{(0, 0)\}$ by Example 5.3(i). We have thus established (i).

Now assume that $z \neq 0$. Then $N_K(z, \beta\|z\|) = \mathbb{R}_+(\beta z, -\|z\|)$. Note that $\beta z \neq 0$ and so $(z, \beta\|z\|) \notin B$. The formulas announced in (iii) therefore follow from Example 5.3(ii).

Next, combining (54), (55), and Example 5.3(iii) as well as utilizing the compactness of the unit sphere in $\mathbb{R}^{m-1}$, we see that

$$N^B_K(0, 0) = \overline{\bigcup_{z \in \mathbb{R}^{m-1} \smallsetminus \{0\}} \mathbb{R}_+(\beta z, -\|z\|)} = \overline{\bigcup_{\substack{z \in \mathbb{R}^{m-1} \\ \|z\|=1}} \mathbb{R}_+(\beta z, -1)} = \bigcup_{\substack{z \in \mathbb{R}^{m-1} \\ \|z\|=1}} \mathbb{R}_+(\beta z, -1). \tag{58}$$

This establishes (iv). ■

**Remark 5.5** Consider Example 5.4. Note that $N^B_K(0, 0)$ is actually the boundary of $N_K(0, 0)$. Furthermore, since $N_K(0, 0) = N^{\text{conv}}_K(0, 0)$ by Lemma 3.4(vii), the formulas in (ii) also describe $K^\ominus$, which is therefore an ice cream cone as well.

## 6 Cones containing restricted normal cones

In this section, we provide various examples illustrating that the restricted (proximal) normal cone does not naturally arise by considering various natural cones containing it.

Let $A$ and $B$ be nonempty subsets of $X$, and let $a \in A$. We saw in Lemma 3.4(ii) that

$$\widehat{N}_A^B(a) = \operatorname{cone}\big((B-a)\cap(P_A^{-1}a-a)\big) \subseteq \operatorname{cone}(B-a)\cap N_A^{\mathrm{prox}}(a). \tag{59}$$

This raises the question whether or not the inclusion in (59) is strict. It turns out and as we shall now illustrate, both conceivable alternatives (equality and strict inclusion) do occur. Therefore, $\widehat{N}_A^B(a)$ is a new construction.

We start with a condition sufficient for equality in (59),

**Proposition 6.1** *Let $A$ and $B$ be nonempty subsets of $X$. Let $A$ be closed and $a \in A$. Assume that one of the following holds:*

(i) *$P_A^{-1}(a) - a$ is a cone.*

(ii) *$A$ is convex.*

*Then $\widehat{N}_A^B(a) = \operatorname{cone}(B-a)\cap N_A^{\mathrm{prox}}(a)$.*

*Proof.* (i): Lemma 2.6(ii). (ii): Combine (i) with Lemma 2.7. ■

The next examples illustrates that equality in (59) can occur even though $P_A^{-1}(a) - a$ is not a cone. Consequently, the assumption that $P_A^{-1}(a) - a$ be a cone in Proposition 6.1 is sufficient—but not necessary—for equality in (59).

**Example 6.2** Suppose that $X = \mathbb{R}^2$, and let $A := X \smallsetminus \mathbb{R}^2_{++}$, $B := \mathbb{R}_+(1,1)$, and $a := (0,1)$. Then one verifies that

$$P_A^{-1}(a) - a = [0,1]\times\{0\}, \tag{60a}$$

$$N_A^{\mathrm{prox}}(a) = \operatorname{cone}(P_A^{-1}a - a) = \mathbb{R}_+\times\{0\}, \tag{60b}$$

$$\operatorname{cone}(B-a) = \big\{(t_1,t_2)\in\mathbb{R}^2 \;\big|\; t_1 \geq 0, t_2 < t_1\big\} \cup \{(0,0)\}, \tag{60c}$$

$$\widehat{N}_A^B(a) = \mathbb{R}_+\times\{0\}. \tag{60d}$$

Hence $\widehat{N}_A^B(a) = \mathbb{R}_+\times\{0\} = \operatorname{cone}(B-a)\cap N_A^{\mathrm{prox}}(a)$.

We now provide an example where the inclusion in (59) is strict.

**Example 6.3** Suppose that $X = \mathbb{R}^2$, let $A := \operatorname{cone}\{(1,0),(0,1)\} = \operatorname{bdry}\mathbb{R}^2_+$, $B := \mathbb{R}_+(2,1)$, and $a := (0,1) \in A$. Then one verifies that

$$P_A^{-1}(a) - a = \left]-\infty, 1\right] \times \{0\}, \tag{61a}$$

$$N_A^{\text{prox}}(a) = \operatorname{cone}(P_A^{-1}a - a) = \mathbb{R} \times \{0\}, \tag{61b}$$

$$\operatorname{cone}(B - a) = \big\{(x_1, x_2) \in \mathbb{R}^2 \;\big|\; x_1 \geq 0, 2x_2 < x_1\big\} \cup \{(0,0)\}, \tag{61c}$$

$$\widehat{N}_A^B(a) = \{(0,0)\}. \tag{61d}$$

Hence $\widehat{N}_A^B(a) = \{(0,0)\} \subsetneqq \mathbb{R}_+ \times \{0\} = \operatorname{cone}(B - a) \cap N_A^{\text{prox}}(a)$, and therefore the inclusion in (59) is strict. In accordance with Proposition 6.1, neither is $P_A^{-1}(a) - a$ a cone nor is $A$ convex.

Let us now turn to the restricted normal cone $N_A^B(a)$. Taking the outer limit in (59) and recalling (14), we obtain

$$N_A^B(a) = \varlimsup_{\substack{x \to a \\ x \in A}} \widehat{N}_A^B(x) \tag{62a}$$

$$\subseteq \varlimsup_{\substack{x \to a \\ x \in A}} \big(\operatorname{cone}(B - x) \cap N_A^{\text{prox}}(x)\big) \tag{62b}$$

$$\subseteq \big(\varlimsup_{\substack{x \to a \\ x \in A}} \operatorname{cone}(B - x)\big) \cap N_A(a). \tag{62c}$$

The inclusions in (62) are optimal in the sense that all possible combinations (strict inclusion and equality) can occur:

- For results and examples illustrating equality in (62b) and equality in (62c), see Proposition 6.5 and Example 6.6 below.
- For an example illustrating equality in (62b) and strict inequality in (62c), see Example 6.7 below.
- For an example illustrating strict inequality in (62b) and equality in (62c), see Example 6.10 below.
- For examples illustrating strict inequality in (62b) and strict inequality in (62c), see Example 6.8 and Example 6.9 below.

The remainder of this section is devoted to providing these examples.

**Proposition 6.4** *Let $A$ and $B$ be nonempty subsets of $X$. Let $A$ be closed $a \in A$. Assume that one of the following holds:*

(i) *$P_A^{-1}(x) - x$ is a cone for every $x \in A$ sufficiently close to $a$.*

(ii) *$A$ is convex.*

*Then* (62b) *holds with equality, i.e.,* $N_A^B(a) = \overline{\lim}_{\substack{x\to a\\ x\in A}} \big(\operatorname{cone}(B-x)\cap N_A^{\text{prox}}(x)\big)$

*Proof*. Indeed, if $x \in A$ is sufficiently close to $a$, then Proposition 6.1 implies that $\widehat{N}_A^B(x) = \operatorname{cone}(B-x)\cap N_A^{\text{prox}}(x)$. Now take the outer limit as $x \to a$ in $A$. ■

**Proposition 6.5** *Let $A$ be a nonempty closed convex subset of $X$, let $B$ be a nonempty subset of $X$, and let $a \in A$. Assume that $x \mapsto \operatorname{cone}(B-x)$ is outer semicontinuous at $a$ relative to $A$, i.e.,*

$$\overline{\lim_{\substack{x\to a\\ x\in A}}}\operatorname{cone}(B-x) = \operatorname{cone}(B-a), \tag{63}$$

*Then* (62) *holds with equalities, i.e.,*

$$N_A^B(a) = \overline{\lim_{\substack{x\to a\\ x\in A}}}\big(\operatorname{cone}(B-x)\cap N_A^{\text{prox}}(x)\big) = \Big(\overline{\lim_{\substack{x\to a\\ x\in A}}}\operatorname{cone}(B-x)\Big)\cap N_A(a). \tag{64}$$

*Proof*. The convexity of $A$ and Lemma 3.4(vii) yield

$$\operatorname{cone}(B-a)\cap N_A(a) = \operatorname{cone}(B-a)\cap N_A^{\text{prox}}(a). \tag{65}$$

On the other hand, Proposition 6.1(ii) and Lemma 3.4(iv) imply

$$\operatorname{cone}(B-a)\cap N_A^{\text{prox}}(a) = \widehat{N}_A^B(a) \subseteq N_A^B(a). \tag{66}$$

Altogether, $\operatorname{cone}(B-a)\cap N_A(a) \subseteq N_A^B(a)$. In view of (63),

$$\Big(\overline{\lim_{\substack{x\to a\\ x\in A}}}\operatorname{cone}(B-x)\Big)\cap N_A(a) \subseteq N_A^B(a). \tag{67}$$

Recalling (62), we therefore obtain (64). ■

**Example 6.6** Let $A$ be a linear subspace of $X$, set $B := A$, and $a := (0,0)$. Then $N_A^B(a) = \{0\}$ by (26d), $N_A(a) = A^\perp$, and $\operatorname{cone}(B-x) = A$, for every $x \in A$. Hence $(\overline{\lim}_{\substack{x\to a\\ x\in A}}\operatorname{cone}(B-x))\cap N_A(a) = \{0\}$ and (62) holds with equalities.

In Proposition 6.5, the convexity and the outer semicontinuity assumptions are both *essential* in the sense that absence of either assumption may make the inclusion (62c) strict; we shall illustrate this in the next three examples.

**Example 6.7** Suppose that $X = \mathbb{R}^2$, and let $A := \operatorname{epi}(|\cdot|)$, $B := \mathbb{R}\times\{0\}$, and $a := (0,0)$. If $x = (x_1, x_2) \in A \smallsetminus \{a\}$, then $x_2 > 0$, $B - x = \mathbb{R}\times\{-x_2\}$, and so $\operatorname{cone}(B-x) = \mathbb{R}\times\mathbb{R}_{--}\cup\{(0,0)\}$. Hence

$$\overline{\lim_{\substack{x\to a\\ x\in A}}}\operatorname{cone}(B-x) = \mathbb{R}\times\mathbb{R}_- \neq \mathbb{R}\times\{0\} = \operatorname{cone}(B-a), \tag{68}$$

i.e., (63) fails. Since $A$ is closed and convex, Lemma 3.4(vii) implies that $N_A(a) = N_A^{\text{conv}}(a) = -A$. Thus

$$\big(\varlimsup_{\substack{x \to a \\ x \in A}} \operatorname{cone}(B - x)\big) \cap N_A(a) = -A. \tag{69}$$

Proposition 6.4(ii) yields equality in (62b), i.e.,

$$N_A^B(a) = \varlimsup_{\substack{x \to a \\ x \in A}} \big(\operatorname{cone}(B - x) \cap N_A^{\text{prox}}(x)\big). \tag{70}$$

Already in Example 5.1 did we observe that

$$N_A^B(a) = \operatorname{cone}\{(1, -1), (-1, -1)\}. \tag{71}$$

Therefore we have

$$N_A^B(a) = \varlimsup_{\substack{x \to a \\ x \in A}} \big(\operatorname{cone}(B - x) \cap N_A^{\text{prox}}(x)\big) \subsetneqq \big(\varlimsup_{\substack{x \to a \\ x \in A}} \operatorname{cone}(B - x)\big) \cap N_A(a), \tag{72}$$

i.e., the inclusion (62c) is strict.

**Example 6.8** Suppose that $X = \mathbb{R}^2$, and let $A := \operatorname{cone}\{(1, 0), (0, 1)\} = \operatorname{bdry} \mathbb{R}^2_+$, $B := \mathbb{R} \times \{1\} \cup \{(1, 0), (-1, 0)\}$, and $a := (0, 0)$. Clearly, $A$ is not convex. If $x = (x_1, x_2) \in A$ is sufficiently close to $a$, we have

$$\operatorname{cone}(B - x) = \begin{cases} \mathbb{R} \times \mathbb{R}_+, & \text{if } x_1 \geq 0; \\ \mathbb{R} \times \mathbb{R}_{++} \cup \operatorname{cone}\{(1, -x_2), (-1, -x_2)\}, & \text{if } x_2 > 0. \end{cases} \tag{73}$$

This yields

$$\varlimsup_{\substack{x \to a \\ x \in A}} \operatorname{cone}(B - x) = \mathbb{R} \times \mathbb{R}_+ = \operatorname{cone}(B - a), \tag{74}$$

i.e., (63) holds. Next, if $x = (x_1, x_2) \in A$, then

$$P_A^{-1}(x) = \begin{cases} \{x_1\} \times \left]-\infty, x_1\right], & \text{if } x_1 > 0 \text{ and } x_2 = 0; \\ \left]-\infty, x_2\right] \times \{x_2\}, & \text{if } x_1 = 0 \text{ and } x_2 > 0; \\ \mathbb{R}^2_-, & \text{if } x_1 = x_2 = 0, \end{cases} \tag{75}$$

and so

$$N_A^{\text{prox}}(x) = \operatorname{cone}\big(P_A^{-1}(x) - x\big) = \begin{cases} \{0\} \times \mathbb{R}, & \text{if } x_1 > 0 \text{ and } x_2 = 0; \\ \mathbb{R} \times \{0\}, & \text{if } x_1 = 0 \text{ and } x_2 > 0; \\ \mathbb{R}^2_-, & \text{if } x_1 = x_2 = 0. \end{cases} \tag{76}$$

It follows that

$$N_A(a) = \varlimsup_{\substack{x \to a \\ x \in A}} N_A^{\text{prox}}(x) = \mathbb{R}^2_- \cup \big(\{0\} \times \mathbb{R}\big) \cup \big(\mathbb{R} \times \{0\}\big). \tag{77}$$

If $x \in A$ is sufficiently close $a$, then

$$\widehat{N}_A^B(x) = \begin{cases} \{(0,0)\}, & \text{if } x \neq a; \\ \mathbb{R}_- \times \{0\}, & \text{if } x = a. \end{cases} \tag{78}$$

It follows that

$$N_A^B(a) = \mathbb{R}_- \times \{0\}. \tag{79}$$

Combining (73) and (76), we obtain for every $x = (x_1, x_2) \in A$ sufficiently close to $a$ that

$$\operatorname{cone}(B - x) \cap N_A^{\text{prox}}(x) = \begin{cases} \{0\} \times \mathbb{R}_+, & \text{if } x_1 > 0 \text{ and } x_2 = 0; \\ \{(0,0)\}, & \text{if } x_1 = 0 \text{ and } x_2 > 0; \\ \mathbb{R}_- \times \{0\}, & \text{if } x_1 = x_2 = 0. \end{cases} \tag{80}$$

Thus

$$\varlimsup_{\substack{x \to a \\ x \in A}} \big( \operatorname{cone}(B - x) \cap N_A^{\text{prox}}(x) \big) = \big( \{0\} \times \mathbb{R}_+ \big) \cup \big( \mathbb{R}_- \times \{0\} \big). \tag{81}$$

Using (79), (81), (74), and (77), we conclude that

$$\begin{aligned} N_A^B(a) &= \mathbb{R}_- \times \{0\} && \text{(82a)} \\ &\subsetneqq \big( \{0\} \times \mathbb{R}_+ \big) \cup \big( \mathbb{R}_- \times \{0\} \big) = \varlimsup_{\substack{a' \to a \\ a' \in A}} \big( \operatorname{cone}(B - x) \cap N_A^{\text{prox}}(x) \big) && \text{(82b)} \\ &\subsetneqq \big( \{0\} \times \mathbb{R}_+ \big) \cup \big( \mathbb{R} \times \{0\} \big) = \Big( \varlimsup_{\substack{x \to a \\ x \in A}} \operatorname{cone}(B - x) \Big) \cap N_A(a). && \text{(82c)} \end{aligned}$$

Therefore, both inclusions in (62) are strict; however, $A$ is not convex while (63) does hold.

**Example 6.9** Suppose that $X = \mathbb{R}^2$, let $A := \operatorname{cone}\{(1,0),(0,1)\} = \operatorname{bdry} \mathbb{R}_+^2$, $B := \mathbb{R}_+(2,1)$ and $a := (0,0)$. Let $x = (x_1, x_2) \in A$. Then (see Example 6.8)

$$P_A^{-1}(x) - x = \begin{cases} \{0\} \times \left]-\infty, x_1\right], & \text{if } x_1 > 0 \text{ and } x_2 = 0; \\ \left]-\infty, x_2\right] \times \{0\}, & \text{if } x_1 = 0 \text{ and } x_2 > 0; \\ \mathbb{R}_-^2, & \text{if } x_1 = x_2 = 0, \end{cases} \tag{83}$$

$$N_A^{\text{prox}}(x) = \begin{cases} \{0\} \times \mathbb{R}, & \text{if } x_1 > 0 \text{ and } x_2 = 0; \\ \mathbb{R} \times \{0\}, & \text{if } x_1 = 0 \text{ and } x_2 > 0; \\ \mathbb{R}_-^2, & \text{if } x_1 = x_2 = 0, \end{cases} \tag{84}$$

and

$$N_A(a) = \varlimsup_{\substack{x \to a \\ x \in A}} N_A^{\text{prox}}(x) = \mathbb{R}_-^2 \cup \big( \{0\} \times \mathbb{R} \big) \cup \big( \mathbb{R} \times \{0\} \big). \tag{85}$$

Thus

$$\widehat{N}_A^B(x) = \operatorname{cone}\big((P_A^{-1}(x) - x) \cap (B - x)\big) = \begin{cases} \{0\} \times \mathbb{R}_+, & \text{if } x_1 > 0 \text{ and } x_2 = 0; \\ \{(0,0)\}, & \text{if } x_1 = 0 \text{ and } x_2 \geq 0. \end{cases} \tag{86}$$

Hence

$$N_A^B(a) = \varlimsup_{\substack{x \to a \\ x \in A}} \widehat{N}_A^B(x) = \{0\} \times \mathbb{R}_+. \tag{87}$$

On the other hand,

$$\operatorname{cone}(B - x) = \begin{cases} \big\{(y_1, y_2) \;\big|\; y_2 \geq 0,\, y_1 < 2y_2\big\} \cup \{(0,0)\}, & \text{if } x_1 > 0 \text{ and } x_2 = 0; \\ \big\{(y_1, y_2) \;\big|\; y_1 \geq 0,\, 2y_2 < y_1\big\} \cup \{(0,0)\}, & \text{if } x_1 = 0 \text{ and } x_2 > 0; \\ B, & \text{if } x_1 = x_2 = 0. \end{cases} \tag{88}$$

Combining (84) and (88), we deduce that

$$\operatorname{cone}(B - x) \cap N_A^{\text{prox}}(x) = \begin{cases} \{0\} \times \mathbb{R}_+, & \text{if } x_1 > 0 \text{ and } x_2 = 0; \\ \mathbb{R}_+ \times \{0\}, & \text{if } x_1 = 0 \text{ and } x_2 > 0; \\ \{(0,0)\}, & \text{if } x_1 = x_2 = 0. \end{cases} \tag{89}$$

Using (88) and (89), we compute

$$\varlimsup_{\substack{x \to a \\ x \in A}} \operatorname{cone}(B - x) = \big\{(y_1, y_2) \;\big|\; y_1 \geq 0 \;\text{ or }\; y_2 \geq 0\big\} = X \smallsetminus \mathbb{R}^2_{--} \neq B = \operatorname{cone}(B - a) \tag{90}$$

and

$$\varlimsup_{\substack{x \to a \\ x \in A}} \big(\operatorname{cone}(B - x) \cap N_A^{\text{prox}}(x)\big) = \big(\{0\} \times \mathbb{R}_+\big) \cup \big(\mathbb{R}_+ \times \{0\}\big) = \operatorname{cone}\{(0,1), (1,0)\}. \tag{91}$$

Using (87), (91), (90), and (85), we conclude that

$$N_A^B(a) = \{0\} \times \mathbb{R}_+ \tag{92a}$$

$$\subsetneqq \big(\{0\} \times \mathbb{R}_+\big) \cup \big(\mathbb{R}_+ \times \{0\}\big) = \varlimsup_{\substack{x \to a \\ x \in A}} \big(\operatorname{cone}(B - x) \cap N_A^{\text{prox}}(x)\big) \tag{92b}$$

$$\subsetneqq \big(\{0\} \times \mathbb{R}\big) \cup \big(\mathbb{R} \times \{0\}\big) = \big(\varlimsup_{\substack{x \to a \\ x \in A}} \operatorname{cone}(B - x)\big) \cap N_A(a). \tag{92c}$$

Therefore, both inclusions in (62) are strict; however, $A$ is not convex and (63) does not hold (see (90)).

Finally, we provide an example where the inclusion (62b) is strict while the inclusion (62c) is an equality.

**Example 6.10** Suppose that $X = \mathbb{R}^2$, let $A := \operatorname{cone}\{(1,0),(0,1)\}$, $B := \{(y_1,y_2) \mid y_1 + y_2 = 1\}$, and $a := (0,0)$. Let $x = (x_1,x_2) \in A$ be sufficiently close to $a$. We compute

$$\operatorname{cone}(B - x) = \{(y_1,y_2) \mid y_1 + y_2 > 0\} \cup \{(0,0)\}, \tag{93a}$$

$$N_A^{\text{prox}}(x) = \begin{cases} \{0\} \times \mathbb{R}, & \text{if } x_1 > 0 \text{ and } x_2 = 0; \\ \mathbb{R} \times \{0\}, & \text{if } x_1 = 0 \text{ and } x_2 > 0; \\ \mathbb{R}^2_-, & \text{if } x_1 = x_2 = 0, \end{cases} \tag{93b}$$

$$\widehat{N}_A^B(x) = \{(0,0)\}. \tag{93c}$$

Furthermore, Example 6.8 (see (77)) implies that $N_A(a) = \mathbb{R}^2_- \cup \big(\{0\} \times \mathbb{R}\big) \cup \big(\mathbb{R} \times \{0\}\big)$. We thus deduce that

$$N_A^B(a) = \{(0,0)\} \tag{94a}$$

$$\subsetneqq \big(\{0\} \times \mathbb{R}_+\big) \cup \big(\mathbb{R}_+ \times \{0\}\big) = \varlimsup_{\substack{x \to a \\ x \in A}} \big(\operatorname{cone}(B - x) \cap N_A^{\text{prox}}(x)\big) \tag{94b}$$

$$= \big(\{0\} \times \mathbb{R}_+\big) \cup \big(\mathbb{R}_+ \times \{0\}\big) = \big(\varlimsup_{\substack{x \to a \\ x \in A}} \operatorname{cone}(B - x)\big) \cap N_A(a). \tag{94c}$$

Therefore, the inclusion (62b) is strict while the inclusion (62c) is an equality.

# 7 Constraint qualification conditions and numbers

Utilizing restricted normal cones, we introduce in this section the notions of *CQ-number*, *joint-CQ-number*, *CQ condition*, and *joint-CQ condition*, where CQ stands for "constraint qualification".

## CQ and joint-CQ numbers

**Definition 7.1 (CQ-number)** *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 $\delta$ is*

$$\theta_\delta := \theta_\delta\big(A,\widetilde{A},B,\widetilde{B}\big) := \sup\left\{ \langle u,v\rangle \;\middle|\; \begin{array}{l} u \in \widehat{N}_A^{\widetilde{B}}(a), v \in -\widehat{N}_B^{\widetilde{A}}(b), \|u\| \le 1, \|v\| \le 1, \\ \|a - c\| \le \delta, \|b - c\| \le \delta. \end{array} \right\}. \tag{95}$$

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

$$\overline{\theta} := \overline{\theta}\big(A,\widetilde{A},B,\widetilde{B}\big) := \lim_{\delta \downarrow 0} \theta_\delta\big(A,\widetilde{A},B,\widetilde{B}\big). \tag{96}$$

Clearly,

$$\theta_\delta\big(A,\widetilde{A},B,\widetilde{B}\big) = \theta_\delta\big(B,\widetilde{B},A,\widetilde{A}\big) \quad \text{and} \quad \overline{\theta}\big(A,\widetilde{A},B,\widetilde{B}\big) = \overline{\theta}\big(B,\widetilde{B},A,\widetilde{A}\big). \tag{97}$$

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

$$c \in \overline{A} \cap \overline{B} \text{ and } 0 < \delta_1 < \delta_2 \quad \Rightarrow \quad 0 \leq \overline{\theta} \leq \theta_{\delta_1} \leq \theta_{\delta_2} \leq 1, \tag{98}$$

while $\theta_\delta$, and hence $\overline{\theta}$, is equal to $-\infty$ if $c \notin \overline{A} \cap \overline{B}$ and $\delta$ is sufficiently small (using the fact that $\sup \varnothing = -\infty$). Using Proposition 3.7(ii)&(vi), we see that

$$\widetilde{A} \subseteq A' \text{ and } \widetilde{B} \subseteq B' \quad \Rightarrow \quad \theta_\delta(A, \widetilde{A}, B, \widetilde{B}) \leq \theta_\delta(A, A', B, B') \tag{99}$$

and, for every $x \in X$,

$$\theta_\delta\big(A, \widetilde{A}, B, \widetilde{B}\big) \text{ at } c \quad = \quad \theta_\delta\big(A - x, \widetilde{A} - x, B - x, \widetilde{B} - x\big) \text{ at } c - x. \tag{100}$$

To deal with unions, it is convenient to extend this notion as follows.

**Definition 7.2 (joint-CQ-number)** *Let $\mathcal{A} := (A_i)_{i\in I}$, $\widetilde{\mathcal{A}} := (\widetilde{A}_i)_{i\in I}$, $\mathcal{B} := (B_j)_{j\in J}$, $\widetilde{\mathcal{B}} := (\widetilde{B}_j)_{j\in J}$ be nontrivial collections*[1] *of nonempty subsets of $X$, let $c \in X$, and let $\delta \in \mathbb{R}_{++}$. The* joint-CQ-number *at $c$ associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ and $\delta$ is*

$$\theta_\delta = \theta_\delta\big(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}}\big) := \sup_{(i,j)\in I\times J} \theta_\delta\big(A_i, \widetilde{A}_i, B_j, \widetilde{B}_j\big), \tag{101}$$

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

$$\overline{\theta} = \overline{\theta}\big(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}}\big) := \lim_{\delta\downarrow 0} \theta_\delta\big(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}}\big). \tag{102}$$

For convenience, we will simply write $\theta_\delta$, $\overline{\theta}$ and omit the possible arguments $(A, \widetilde{A}, B, \widetilde{B})$ and $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ when there is no cause for confusion. If $I$ and $J$ are singletons, then the notions of CQ-number and joint-CQ-number coincide. Also observe that

$$c \in \bigcup_{i\in I} A_i \cap \bigcup_{j\in J} B_j \quad \Rightarrow \quad (\forall \delta \in \mathbb{R}_{++}) \;\; 0 \leq \overline{\theta} \leq \theta_\delta \leq 1 \tag{103}$$

while $\overline{\theta} = \theta_\delta = -\infty$ when $\delta > 0$ is sufficiently small and $c$ does not belong to both $\overline{\bigcup_{i\in I} A_i}$ and $\overline{\bigcup_{j\in J} B_j}$. Furthermore, the joint-CQ-number (and hence the limiting joint-CQ-number as well) really depends only on those sets $A_i$ and $B_j$ for which $c \in \overline{A_i} \cap \overline{B_j}$.

To illustrate this notion, let us compute the CQ-number of two lines. The formula provided is the cosine of the angle between the two lines — as we shall see in Theorem 8.12 below, this happens actually for all linear subspaces although then the angle must be defined appropriately and the proof is more involved.

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

**Proposition 7.3 (CQ-number of two distinct lines through the origin)** *Suppose that $w_a$ and $w_b$ are two vectors in $X$ such that $\|w_a\| = \|w_b\| = 1$. Let $A := \mathbb{R}w_a$, $B := \mathbb{R}w_b$, and $\delta \in \mathbb{R}_{++}$. Assume that $A \cap B = \{0\}$. Then the CQ-number at 0 is*

$$\theta_\delta(A, A, B, B) = |\langle w_a, w_b\rangle|. \tag{104}$$

*Proof.* Set $s := \langle w_a, w_b\rangle$.

Assume first that $s \neq 0$. Let $a = \alpha w_a \in A$ and $b = \beta w_b \in B$. Then $P_A^{-1}(a) - a = N_A(a) = \{w_a\}^\perp$; considering $(B - a) \cap \{w_a\}^\perp$ leads to $\beta s = \alpha$. Hence $(P_A^{-1}(a) - a) \cap (B - a) = \beta w_b - \alpha w_a$ and

$$\widehat{N}_A^B(a) = \operatorname{cone}\big(\alpha s^{-1} w_b - \alpha w_a\big). \tag{105}$$

Similarly,

$$-\widehat{N}_B^A(b) = \operatorname{cone}\big(\beta w_b - \beta s^{-1} w_a\big). \tag{106}$$

Now set $u := \alpha s^{-1} w_b - \alpha w_a \in \widehat{N}_A^B(a)$ and $v := \beta w_b - \beta s^{-1} w_a \in -\widehat{N}_B^A(b)$. One computes

$$\|u\| = \frac{|\alpha|\sqrt{1-s^2}}{|s|}, \quad \|v\| = \frac{|\beta|\sqrt{1-s^2}}{|s|}, \quad \text{and} \quad \langle u, v\rangle = \frac{\alpha\beta(1-s^2)}{s}. \tag{107}$$

Hence

$$\frac{\langle u, v\rangle}{\|u\| \cdot \|v\|} = \operatorname{sgn}(\alpha)\operatorname{sgn}(\beta)s. \tag{108}$$

Choosing $\alpha$ and $\beta$ in $\{-1, 1\}$ appropriately, we arrange for $\langle u, v\rangle / (\|u\| \cdot \|v\|) = |s|$, as claimed.

Now assume that $s = 0$. Arguing similarly, we see that

$$(\forall a \in A) \quad \widehat{N}_A^B(a) = \begin{cases} \{0\}, & \text{if } a \neq 0; \\ B, & \text{if } a = 0, \end{cases} \quad \text{and} \quad (\forall b \in B) \quad \widehat{N}_B^A(b) = \begin{cases} \{0\}, & \text{if } b \neq 0; \\ A, & \text{if } b = 0. \end{cases} \tag{109}$$

This leads to $\theta_\delta(A, A, B, B) = 0 = |s|$, again as claimed. ■

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 closed subsets of $X$ and let $\delta \in \mathbb{R}_{++}$. Set $A := \bigcup_{i\in I} A_i$, $\widetilde{A} := \bigcup_{i\in I} \widetilde{A}_i$, $B := \bigcup_{j\in J} B_j$, $\widetilde{B} := \bigcup_{j\in J} \widetilde{B}_j$, and suppose that $c \in A \cap B$. It is interesting to compare the joint-CQ-number of collections, i.e., $\theta_\delta\big(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}}\big)$, to the CQ-number of the unions, i.e., $\theta_\delta\big(A, \widetilde{A}, B, \widetilde{B}\big)$. We shall see in the following two examples that *neither of them is smaller than the other*; in fact, one of them can be equal to 1 while the other is strictly less than 1.

**Example 7.4 (joint-CQ-number $<$ CQ-number of the unions)** 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

$$\theta_\delta\big(\mathcal{A}, \mathcal{A}, \mathcal{B}, \mathcal{B}\big) = \tfrac{2}{\sqrt{5}} < 1 = \theta_\delta\big(A, A, B, B\big). \tag{110}$$

*Proof*. Using Proposition 7.3, we compute, for the reference point $c$,

$$\theta_\delta(A_1, A_1, B_1, B_1) = \big|\big\langle (0,1,0), \tfrac{1}{\sqrt{2}}(0,1,1)\big\rangle\big| = \tfrac{1}{\sqrt{2}}, \tag{111a}$$

$$\theta_\delta(A_1, A_1, B_2, B_2) = |\,\langle (0,1,0), (1,0,0)\rangle\,| = 0, \tag{111b}$$

$$\theta_\delta(A_2, A_2, B_1, B_1) = \big|\big\langle \tfrac{1}{\sqrt{5}}(2,0,-1), \tfrac{1}{\sqrt{2}}(0,1,1)\big\rangle\big| = \tfrac{1}{\sqrt{10}}, \tag{111c}$$

$$\theta_\delta(A_2, A_2, B_2, B_2) = \big|\big\langle \tfrac{1}{\sqrt{5}}(2,0,-1), (1,0,0\big\rangle\big| = \tfrac{2}{\sqrt{5}}. \tag{111d}$$

Hence $\theta_\delta(\mathcal{A}, \mathcal{A}, \mathcal{B}, \mathcal{B}) = \max_{(i,j)\in I\times J} \theta_\delta(A_i, A_i, B_j, B_j) = \frac{2}{\sqrt{5}} < 1$.

To estimate the CQ-number of the union, set

$$a := (0, \delta, 0) \in A_1 \subseteq A \text{ and } b := (\delta, 0, 0) \in B_2 \subseteq B. \tag{112}$$

Note that $\|a - c\| = \|a\| = \delta$ and $\|b - c\| = \|b\| = \delta$. Now define

$$\widetilde{a} := (\delta, 0, -\delta/2) \in A_2 \subseteq A \;\text{ and }\; \widetilde{b} := (0, \delta, \delta) \in B_1 \subseteq B. \tag{113}$$

Since $\|\widetilde{a} - P_{B_2}\widetilde{a}\| < \|\widetilde{a} - P_{B_1}\widetilde{a}\|$ and $P_{B_2}\widetilde{a} = b$, we have $b = P_B\widetilde{a}$. Since $\|\widetilde{b} - P_{A_1}\widetilde{b}\| < \|\widetilde{b} - P_{A_2}\widetilde{b}\|$ and $P_{A_1}\widetilde{b} = a$, we have $a = P_A\widetilde{b}$. Therefore, $\widetilde{b} \in B \cap P_A^{-1}(a)$ and $\widetilde{a} \in A \cap P_B^{-1}(b)$. It follows that

$$u := \tfrac{1}{\delta}(\widetilde{b} - a) = (0,0,1) \in \widehat{N}_A^B(a), \tag{114a}$$

$$v := \tfrac{2}{\delta}(b - \widetilde{a}) = (0,0,1) \in -\widehat{N}_B^A(b). \tag{114b}$$

Since $\|u\| = \|v\| = 1$, we obtain $1 = \langle u, v\rangle \leq \theta_\delta(A, A, B, B) \leq 1$. ■

**Example 7.5 (CQ-number of the unions < joint-CQ-number)** Suppose that $X = \mathbb{R}$, let $I := J := \{1, 2\}$, $A_1 := B_1 := \mathbb{R}_-$, $A_2 := B_2 := \mathbb{R}_+$, $c := 0$, and $\delta > 0$. Furthermore, set $\mathcal{A} := (A_i)_{i\in I}$, $\mathcal{B} := (B_j)_{j\in I}$, $A := A_1 \cup A_2 = \mathbb{R}$, and $B := B_1 \cup B_2 = \mathbb{R}$. Then

$$\theta_\delta\big(A, A, B, B\big) = 0 < 1 = \theta_\delta\big(\mathcal{A}, \mathcal{A}, \mathcal{B}, \mathcal{B}\big). \tag{115}$$

*Proof*. Lemma 3.4(viii) implies that $(\forall x \in \mathbb{R})\ \widehat{N}_{\mathbb{R}}^{\mathbb{R}}(x) = \{0\}$. Hence $\theta_\delta(\mathbb{R}, \mathbb{R}, \mathbb{R}, \mathbb{R}) = 0$ as claimed. On the other hand, $\widehat{N}_{\mathbb{R}_+}^{\mathbb{R}_-}(0) = \mathbb{R}_-$ and $\widehat{N}_{\mathbb{R}_-}^{\mathbb{R}_+}(0) = \mathbb{R}_+$. Hence $\theta_\delta(\mathbb{R}_-, \mathbb{R}_-, \mathbb{R}_+, \mathbb{R}_+) = 1$ and therefore $\theta_\delta\big(\mathcal{A}, \mathcal{A}, \mathcal{B}, \mathcal{B}\big) = 1$ as well. ■

The two preceding examples illustrated the independence of the two types of CQ-numbers (for the collection and for the union). In some cases, such as Example 7.4, it is beneficial to work with a suitable partition to obtain a CQ-number that is less than one, which in turn is very desirable in applications (see Section 10).

## CQ and joint-CQ conditions

**Definition 7.6 (CQ and joint-CQ conditions)** *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*

$$N_A^{\widetilde{B}}(c) \cap \big( - N_B^{\widetilde{A}}(c)\big) \subseteq \{0\}. \tag{116}$$

(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.,*

$$\big(\forall (i, j) \in I \times J\big) \quad N_{A_i}^{\widetilde{B}_j}(c) \cap \big( - N_{B_j}^{\widetilde{A}_i}(c)\big) \subseteq \{0\}. \tag{117}$$

In view of the definitions, the key case to consider is when $c \in A \cap B$ (or when $c \in A_i \cap B_j$ in the joint-CQ case). The CQ-number is based on the behavior of the restricted proximal normal cone in a neighborhood of the point under consideration — a related notion is that of the exact CQ-number, where we consider the restricted normal cone at the point instead of nearby restricted proximal normal cones.

**Definition 7.7 (exact CQ-number and exact joint-CQ-number)** *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]

$$\overline{\alpha} := \overline{\alpha}\big(A, \widetilde{A}, B, \widetilde{B}\big) := \sup \left\{ \langle u, v\rangle \;\middle|\; u \in N_A^{\widetilde{B}}(c), v \in -N_B^{\widetilde{A}}(c), \|u\| \leq 1, \|v\| \leq 1 \right\}. \tag{118}$$

(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*

$$\overline{\alpha} := \overline{\alpha}(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}}) := \sup_{(i,j)\in I\times J} \overline{\alpha}(A_i, \widetilde{A}_i, B_j, \widetilde{B}_j). \tag{119}$$

The next result relates the various condition numbers defined above.

**Theorem 7.8** *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* (119)*), the joint-CQ-number at $c$ associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ and $\delta > 0$ by $\theta_\delta$ (see* (101)*), and the limiting joint-CQ-number at $c$ associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ by $\overline{\theta}$ (see* (102)*). 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) $\overline{\alpha} \leq \theta_\delta$.

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

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

*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 $\overline{\alpha} = \overline{\theta} < 1$.*

*Proof.* (i): Suppose that $\overline{\alpha} < 1$. The condition for equality in the Cauchy-Schwarz inequality implies that for all $(i,j) \in I \times J$, the intersection $N_{A_i}^{\widetilde{B}_j}(c) \cap (-N_{B_j}^{\widetilde{A}_i}(c))$ is either empty or $\{0\}$. In view of Definition 7.6, we see that the $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$-joint-CQ holds at $c$.

(ii): Let $(i,j) \in I \times J$. Take $u \in N_{A_i}^{\widetilde{B}_j}(c)$ and $v \in -N_{B_j}^{\widetilde{A}_i}(c)$ such that $\|u\| \leq 1$ and $\|v\| \leq 1$. Then, by definition of the restricted normal cone, there exist sequences $(a_n)_{n\in\mathbb{N}}$ in $A_i$, $(b_n)_{n\in\mathbb{N}}$ in $B_j$, $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ in $X$ such that $a_n \to c$, $b_n \to c$, $u_n \to u$, $v_n \to v$, and $(\forall n \in \mathbb{N})$ $u_n \in \widehat{N}_{A_i}^{\widetilde{B}_j}(a_n)$ and $v_n \in -\widehat{N}_{B_j}^{\widetilde{A}_i}(b_n)$. Note that since $\delta > 0$, eventually $a_n$ and $b_n$ lie in $\operatorname{ball}(c;\delta)$; consequently, $\langle u_n, v_n\rangle \leq \theta_\delta(A_i, \widetilde{A}_i, B_j, \widetilde{B}_j)$. Taking the limit as $n \to +\infty$, we obtain $\langle u, v\rangle \leq \theta_\delta(A_i, \widetilde{A}_i, B_j, \widetilde{B}_j) \leq \theta_\delta$. Now taking the supremum over suitable $u$ and $v$, followed by taking the supremum over $(i,j)$, we conclude that $\overline{\alpha} \leq \theta_\delta$.

(iii): This is clear from (ii) and (102).

(iv): Let $(\delta_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}_{++}$ such that $\delta_n \to 0$. Then for every $n \in \mathbb{N}$, there exist

$$i_n \in I,\ j_n \in J,\ a_n \in A_{i_n},\ b_n \in B_{j_n},\ u_n \in \widehat{N}_{A_{i_n}}^{\widetilde{B}_{j_n}}(a_n),\ v_n \in -\widehat{N}_{B_{j_n}}^{\widetilde{A}_{i_n}}(b_n) \tag{120}$$

such that

$$\|a_n - c\| \leq \delta_n,\ \|b_n - c\| \leq \delta_n,\ \|u_n\| \leq 1,\ \|v_n\| \leq 1,\ \text{ and } \langle u_n, v_n\rangle > \theta_{\delta_n} - \delta_n. \tag{121}$$

Since $I$ and $J$ are finite, and after passing to a subsequence and relabeling if necessary, we can and do assume that there exists $(i,j) \in I \times J$ such that $u_n \to u \in N_{A_i}^{\widetilde{B}_j}(c)$ and $v_n \to v \in -N_{B_j}^{\widetilde{A}_i}(c)$. Hence $\overline{\theta} \leftarrow \theta_{\delta_n} - \delta_n < \langle u_n, v_n\rangle \to \langle u, v\rangle \leq \overline{\alpha}$. Hence $\overline{\theta} \leq \overline{\alpha}$. On the other hand, $\overline{\alpha} \leq \overline{\theta}$ by (iii). Altogether, $\overline{\alpha} = \overline{\theta}$.

(v): "$\Rightarrow$": Let $(i,j) \in I \times J$. If $c \notin A_i \cap B_j$, then $\overline{\alpha}(A_i, \widetilde{A}_i, B_j, \widetilde{B}_j) = -\infty$. Now assume that $c \in A_i \cap B_j$. Since the $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$-joint-CQ condition holds, we have $N_{A_i}^{\widetilde{B}_j}(c) \cap -N_{B_j}^{\widetilde{A}_i}(c) = \{0\}$. By Cauchy-Schwarz,

$$\overline{\alpha}(A_i, \widetilde{A}_i, B_j, \widetilde{B}_j) = \sup\left\{\langle u, v\rangle \;\middle|\; u \in N_{A_i}^{\widetilde{B}_j}(c), v \in -N_{B_j}^{\widetilde{A}_i}(c), \|u\| \leq 1, \|v\| \leq 1\right\} < 1. \tag{122}$$

Since $I$ and $J$ are finite and because of (iv), we deduce that $\overline{\theta} = \overline{\alpha} < 1$.
"$\Leftarrow$": Combine (i) with (iv). ■

## 8 CQ conditions and CQ numbers: examples

In this section, we provide further results and examples illustrating CQ conditions and CQ numbers.

First, let us note that the assumption that the sets of indices be finite in Theorem 7.8(iv) is essential:

**Example 8.1 ($\overline{\alpha} < \overline{\theta}$)** Suppose that $X = \mathbb{R}^2$, let $\Gamma \subseteq \mathbb{R}_{++}$ be such that $\sup \Gamma = +\infty$, set $(\forall \gamma \in \Gamma)$ $A_\gamma := \operatorname{epi}(\frac{1}{2}\gamma|\cdot|^2)$, $B := \mathbb{R}_+ \times \mathbb{R}$, $\mathcal{A} := (A_\gamma)_{\gamma\in\Gamma}$, $\widetilde{\mathcal{A}} := (X)_{\gamma\in\Gamma}$, $\mathcal{B} := (B)$, $\widetilde{\mathcal{B}} := (X)$, and $c := (0,0)$. Denote the exact joint-CQ-number at $c$ associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ by $\overline{\alpha}$ (see (119)), the joint-CQ-number at $c$ associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ and $\delta > 0$ by $\theta_\delta$ (see (101)), and the limiting joint-CQ-number at $c$ associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ by $\overline{\theta}$ (see (102)). Then

$$\overline{\alpha} = 0 < 1 = \theta_\delta = \overline{\theta}. \tag{123}$$

*Proof*. Let $\gamma \in \Gamma$ and pick $x > 0$ such that $a := (x, \frac{1}{2}\gamma x^2) \in A_\gamma$ satisfies $\|a\| = \|a - c\| = \delta$, i.e., $x > 0$ and

$$\gamma^2 x^2 = 2\Big(\sqrt{1+\gamma^2\delta^2} - 1\Big) \to +\infty \quad \text{as } \gamma \to +\infty \text{ in } \Gamma. \tag{124}$$

Hence

$$\gamma x \to +\infty, \quad \text{as } \gamma \to +\infty \text{ in } \Gamma. \tag{125}$$

Since $A_\gamma$ is closed and convex, it follows from Lemma 3.4(vii) that

$$u := \frac{(\gamma x, -1)}{\sqrt{\gamma^2 x^2 + 1}} \in \mathbb{R}_+(\gamma x, -1) = N^{\text{conv}}_{A_\gamma}(a) = \widehat{N}^X_{A_\gamma}(a) = N^X_{A_\gamma}(a) = N_{A_\gamma}(a). \tag{126}$$

Furthermore, $v := (1,0) \in -(\mathbb{R}_- \times \{0\}) = -\widehat{N}^X_B(c) = -N^X_B(c) = -N_B(c)$, $\|u\| = \|v\| = 1$, and, in view of (125),

$$1 \geq \theta_\delta \geq \theta_\delta(A_\gamma, X, B, X) \geq \langle u, v\rangle = \frac{\gamma x}{\sqrt{\gamma^2 x^2 + 1}} \tag{127a}$$

$$\to 1 \quad \text{as } \gamma \to +\infty \text{ in } \Gamma. \tag{127b}$$

Thus $\theta_\delta = 1$, which implies that $\overline{\theta} = 1$. Finally, $N_{A_\gamma}(c) = (\{0\} \times \mathbb{R}_-) \perp (\mathbb{R}_+ \times \{0\}) = -N_B(c)$, which shows that $\overline{\alpha} = 0$. ■

For the eventual application of these results to the method of alternating projections, the condition $\overline{\alpha} = \overline{\theta} < 1$ is critical to ensure linear convergence.

The following example illustrates that the CQ-number can be interpreted as a quantification of the CQ condition.

**Example 8.2 (CQ-number quantifies CQ condition)** 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 (116)).

(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 (96)).

*Proof.* The identity (26d) of Theorem 4.5 yields $N_A^L(c) = N_A(c) \cap (L - c)$ and $N_B^L(c) = N_B(c) \cap (L - c)$. Hence

$$N_A^L(c) \cap \big(-N_B^L(c)\big) = N_A(c) \cap \big(-N_B(c)\big) \cap (L - c), \tag{128}$$

and the equivalence of (i) and (ii) is now clear. Finally, Theorem 7.8(iv)&(v) yields the equivalence of (i) and (iii). ■

Depending on the choice of the restricting sets $\widetilde{A}$ and $\widetilde{B}$, the $(A, \widetilde{A}, B, \widetilde{B})$-CQ condition may either hold or fail:

**Example 8.3 (CQ condition depends on restricting sets)** 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,

$$N_A^X(c) \cap \big(-N_B^X(c)\big) = \{0\} \times \mathbb{R}_- \ \text{ while } \ N_A^B(c) \cap \big(-N_B^A(c)\big) = \{(0,0)\}. \tag{129}$$

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

For two spheres, it is possible to quantify the convergence of $\theta_\delta$ to $\overline{\delta} = \overline{\alpha}$:

**Proposition 8.4 (CQ-numbers of two spheres)** *Let $z_1$ and $z_2$ be in $X$, let $\rho_1$ and $\rho_2$ 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 7.1), and the exact CQ-number at $c$ associated with $(S_1, X, S_2, X)$ by $\overline{\alpha}$ (see Definition 7.7). 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*

$$\overline{\alpha} \leq \theta_\delta \leq \overline{\alpha} + \varepsilon, \tag{130}$$

*where $\theta_\delta$ is the CQ-number at $c$ associated with $(S_1, X, S_2, X)$ (see Definition 7.1).*

*Proof*. (i): This follows from Theorem 7.8(iv) and Example 3.6.

(ii): Combine (i) with the characterization of equality in the Cauchy-Schwarz inequality.

Let us now establish (130). By Theorem 7.8(ii), we have $\overline{\alpha} \leq \theta_\delta$. Let $s_1 \in S_1$ be such that $\|s_1 - c\| \leq \delta$, let $u_1 \in \widehat{N}^X_{S_1}(s_1)$ be such that $\|u_1\| = 1$, let $s_2 \in S_2$ be such that $\|s_2 - c\| \leq \delta$, and let $u_2 \in \widehat{N}^X_{S_2}(s_2)$ be such that $\|u_2\| = 1$. By Example 3.6,

$$u_1 = \pm\frac{s_1 - z_1}{\|s_1 - z_1\|} = \pm\frac{s_1 - z_1}{\rho_1} \quad \text{and} \quad u_2 = \pm\frac{s_2 - z_2}{\|s_2 - z_2\|} = \pm\frac{s_2 - z_2}{\rho_2}. \tag{131}$$

Hence

$$\begin{aligned}
\rho_1\rho_2 \langle u_1, u_2\rangle &\leq |\langle s_1 - z_1, s_2 - z_2\rangle| && \text{(132a)}\\
&= |\langle (s_1 - c) + (c - z_1), (s_2 - c) + (c - z_2)\rangle| && \text{(132b)}\\
&\leq |\langle s_1 - c, s_2 - c\rangle| + |\langle s_1 - c, c - z_2\rangle| && \text{(132c)}\\
&\quad + |\langle c - z_1, s_2 - c\rangle| + |\langle c - z_1, c - z_2\rangle| && \text{(132d)}\\
&\leq \delta^2 + \delta(\rho_1 + \rho_2) + \rho_1\rho_2\overline{\alpha} && \text{(132e)}
\end{aligned}$$

and thus, using the definition of $\delta$,

$$\langle u_1, u_2\rangle \leq \overline{\alpha} + \frac{\delta^2 + \delta(\rho_1 + \rho_2)}{\rho_1\rho_2} = \overline{\alpha} + \varepsilon. \tag{133}$$

Therefore, by the definition of $\theta_\delta$, we have $\theta_\delta \leq \overline{\alpha} + \varepsilon$. ■

## Two convex sets

Let us turn to the classical convex setting. We start by noting that well known constraint qualifications are conveniently characterized using our CQ conditions.

**Proposition 8.5** *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 $A \cap B$.*

*Proof*. This is clear from Theorem 4.13. ■

**Proposition 8.6** *Let $A$ and $B$ be nonempty convex subsets of $X$ such that $A \cap B \neq \varnothing$. Then the following are equivalent:*

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

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

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

*Proof.* This is clear from Corollary 4.14. ■

In stark contrast to Proposition 8.5 and 8.6, if the restricting sets are not both equal to $L$ or to $X$, then the CQ-condition may actually depend on the reference point as we shall illustrate now:

**Example 8.7 (CQ condition depends on the reference point)** Suppose that $X = \mathbb{R}^2$, and let $f\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto (\max\{0, x\})^2$, which is a continuous convex function. Set $A := \operatorname{epi} f$ and $B := \mathbb{R} \times \{0\}$, which are closed convex subsets of $X$. Consider first the point $c := (-1, 0) \in A \cap B$. Then $N_A^B(c) = \{(0, 0)\}$ and $N_B^A(c) = \{0\} \times \mathbb{R}_+$; hence,

$$N_A^B(c) \cap \big(-N_B^A(c)\big) = \{(0, 0)\}, \tag{134}$$

i.e., the $(A, A, B, B)$-CQ condition holds at $c$. On the other hand, consider now $d := (0, 0) \in A \cap B$. Then $N_A^B(d) = \{0\} \times \mathbb{R}_-$ and $N_B^A(d) = \{0\} \times \mathbb{R}_+$; thus,

$$N_A^B(d) \cap \big(-N_B^A(d)\big) = \{0\} \times \mathbb{R}_-, \tag{135}$$

i.e., the $(A, A, B, B)$-CQ condition fails at $d$.

## Two linear (or intersecting affine) subspaces

We specialize further to two linear subspaces of $X$. A pleasing connection between CQ-number and the angle between two linear subspaces will be revealed. But first we provide some auxiliary results.

**Proposition 8.8** *Let $A$ and $B$ be linear subspaces of $X$, and let $\delta \in \mathbb{R}_{++}$. Then*

$$\bigcup_{a \in A \cap (B + A^\perp) \cap \operatorname{ball}(0;\delta)} \widehat{N}_A^B(a) = \bigcup_{a \in A \cap \operatorname{ball}(0;\delta)} \widehat{N}_A^B(a) = \bigcup_{a \in A} \widehat{N}_A^B(a) = A^\perp \cap (A + B). \tag{136}$$

*Proof.* Let $a \in A$. Then $P_A^{-1}(a) = a + A^\perp$ and hence $P_A^{-1}(a) - a = A^\perp$. If $B \cap (a + A^\perp) = \varnothing$, then $\widehat{N}_A^B(a) = \{0\}$. Thus we assume that $B \cap (a + A^\perp) \neq \varnothing$, which is equivalent to $a \in A \cap (B + A^\perp)$. Next, by Lemma 3.4(ii), $\widehat{N}_A^B(a) = A^\perp \cap \operatorname{cone}(B - a)$. This implies $(\forall \lambda \in \mathbb{R}_{++})$ $\operatorname{cone}(B - \lambda a) = \operatorname{cone}(\lambda(B - a)) = \operatorname{cone}(B - a)$. Thus,

$$(\forall \lambda \in \mathbb{R}_{++}) \quad \widehat{N}_A^B(\lambda a) = A^\perp \cap \operatorname{cone}(B - \lambda a) = A^\perp \cap \operatorname{cone}(B - a) = \widehat{N}_A^B(a). \tag{137}$$

This establishes not only the first two equalities in (136) but also the third because

$$\bigcup_{a \in A} \widehat{N}_A^B(a) = \bigcup_{a \in A} \big(A^\perp \cap \operatorname{cone}(B - a)\big) = A^\perp \cap \bigcup_{a \in A} \operatorname{cone}(B - a) \tag{138a}$$

$$= A^{\perp} \cap \operatorname{cone}\Big( \bigcup_{a\in A} (B-a) \Big) = A^{\perp} \cap \operatorname{cone}(B-A) = A^{\perp} \cap (B-A) \tag{138b}$$

$$= A^{\perp} \cap (B+A). \tag{138c}$$

The proof is complete. ■

We now introduce two notions of angles between subspaces; for further information, we highly recommend [10] and [11].

**Definition 8.9** *Let $A$ and $B$ be linear subspaces of $X$.*

(i) **(Dixmier angle)** [15] *The* Dixmier angle *between $A$ and $B$ is the number in $[0, \frac{\pi}{2}]$ whose cosine is given by*

$$c_0(A, B) := \sup \big\{ |\langle a, b\rangle| \;\big|\; a \in A, b \in B, \|a\| \leq 1, \|b\| \leq 1 \big\}. \tag{139}$$

(ii) **(Friedrichs angle)** [16] *The* Friedrichs angle *(or simply the* angle*) between $A$ and $B$ is the number in $[0, \frac{\pi}{2}]$ whose cosine is given by*

$$c(A, B) := c_0(A \cap (A \cap B)^{\perp}, B \cap (A \cap B)^{\perp}) \tag{140a}$$

$$= \sup \left\{ |\langle a, b\rangle| \;\middle|\; \begin{array}{l} a \in A \cap (A\cap B)^{\perp}, \|a\| \leq 1, \\ b \in B \cap (A \cap B)^{\perp}, \|b\| \leq 1 \end{array} \right\}. \tag{140b}$$

Let us gather some properties of angles.

**Fact 8.10** *Let $A$ and $B$ be linear subspaces of $X$. Then the following hold:*

(i) *If $A \cap B = \{0\}$, then $c(A, B) = c_0(A, B)$.*

(ii) *If $A \cap B \neq \{0\}$, then $c_0(A, B) = 1$.*

(iii) $c(A, B) < 1$.

(iv) $c(A, B) = c_0(A, B \cap (A \cap B)^{\perp}) = c_0(A \cap (A \cap B)^{\perp}, B)$.

(v) **(Solmon)** $c(A, B) = c(A^{\perp}, B^{\perp})$.

*Proof*. (i)–(iii): Clear from the definitions. (iv): See, e.g., [10, Lemma 2.10(1)] or [11, Lemma 9.5]. (v): See, e.g., [10, Theorem 2.16]. ■

**Proposition 8.11 (CQ-number of two linear subspaces and Dixmier angle)** *Let $A$ and $B$ be linear subspaces of $X$, and let $\delta > 0$. Then*

$$\theta_\delta(A, A, B, B) = c_0\big(A^{\perp} \cap (A+B), B^{\perp} \cap (A+B)\big), \tag{141a}$$

$$\theta_\delta(A, X, B, B) = c_0\big(A^{\perp} \cap (A+B), B^{\perp}\big), \tag{141b}$$

$$\theta_\delta(A, A, B, X) = c_0\big(A^{\perp}, B^{\perp} \cap (A+B)\big), \tag{141c}$$

*where the CQ-numbers at 0 are defined as in* (95).

*Proof.* This follows from Proposition 8.8. ■

We are now in a position to derive a striking connection between the CQ-number and the Friedrichs angle, which underlines a possible interpretation of the CQ-number as a generalized Friedrichs angle between two sets.

**Theorem 8.12 (CQ-number of two linear subspaces and Friedrichs angle)** *Let $A$ and $B$ be linear subspaces of $X$, and let $\delta > 0$. Then*

$$\theta_\delta(A, A, B, B) = \theta_\delta(A, X, B, B) = \theta_\delta(A, A, B, X) = c(A, B) < 1, \tag{142}$$

*where the CQ-number at 0 is defined as in* (95).

*Proof.* On the one hand, using Fact 8.10(v), we have

$$\begin{aligned} c(A, B) &= c(A^\perp, B^\perp) && \text{(143a)} \\ &= c_0\big(A^\perp \cap (A^\perp \cap B^\perp)^\perp, B^\perp \cap (A^\perp \cap B^\perp)^\perp\big) && \text{(143b)} \\ &= c_0\big(A^\perp \cap (A + B), B^\perp \cap (A + B)\big). && \text{(143c)} \end{aligned}$$

On the other hand, Fact 8.10(iv) yields

$$\begin{aligned} c_0\big(A^\perp \cap (A + B), B^\perp\big) &= c_0\big(A^\perp \cap (A^\perp \cap B^\perp)^\perp, B^\perp\big) && \text{(144a)} \\ &= c(A^\perp, B^\perp) && \text{(144b)} \\ &= c_0\big(A^\perp, B^\perp \cap (A^\perp \cap B^\perp)^\perp\big) && \text{(144c)} \\ &= c_0\big(A^\perp, B^\perp \cap (A + B)\big). && \text{(144d)} \end{aligned}$$

Altogether, recalling Proposition 8.11, we obtain the result. ■

The results in this subsection have a simple generalization to intersecting affine subspaces. Indeed, if $A$ and $B$ are *intersecting* affine subspaces, then the corresponding Friedrichs angle is

$$c(A, B) := c(\operatorname{par} A, \operatorname{par} B). \tag{145}$$

Combining (100) with Theorem 8.12, we immediately obtain the following result.

**Corollary 8.13 (CQ-number of two intersecting affine subspaces and Friedrichs angle)** *Let $A$ and $B$ be affine subspaces of $X$, suppose that $c \in A \cap B$, and let $\delta > 0$. Then*

$$\theta_\delta(A, A, B, B) = \theta_\delta(A, X, B, B) = \theta_\delta(A, A, B, X) = c(A, B) < 1, \tag{146}$$

*where the CQ-number at $c$ is defined as in* (95).

## 9 Regularities

In this section, we study a notion of set regularity that is based on restricted normal cones.

**Definition 9.1 (regularity and superregularity)** *Let $A$ and $B$ be nonempty subsets of $X$, and let $c \in X$.*

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

$$\left.\begin{array}{c}(y,b) \in B \times B,\\ \|y-c\| \leq \delta, \|b-c\| \leq \delta,\\ u \in \widehat{N}_B^A(b)\end{array}\right\} \quad \Rightarrow \quad \langle u, y-b\rangle \leq \varepsilon\|u\| \cdot \|y-b\|. \tag{147}$$

*If $B$ is $(X, \varepsilon, \delta)$-regular at $c$, then we also simply speak of $(\varepsilon, \delta)$-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 $(A, \varepsilon, \delta)$-regular at $c$. Again, if $B$ is $X$-superregular at $c$, then we also say that $B$ is superregular at $c$.*

**Remark 9.2** Several comments on Definition 9.1 are in order.

(i) Superregularity with $A = X$ was introduced by Lewis, Luke and Malick in [17, 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 $\delta \in \,]0, d_B(c)[$, $B$ is $(0, \delta)$-regular at $c$; consequently, $B$ is superregular at $c$.

(iii) If $\varepsilon_1 > \varepsilon_2$ and $B$ is $(A, \varepsilon_2, \delta)$-regular at $c$ then $B$ is also $(A, \varepsilon_1, \delta)$-regular at $c$.

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

(v) It follows from Proposition 3.7(ii) 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 $c$.

(vi) If $B$ is convex, then it follows with Lemma 3.4(vii) that $B$ is $(A, 0, +\infty)$-regular at $c$; consequently, $B$ is superregular.

(vii) 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 $c$.

(viii) 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 9.3 below).

As a first example, let us consider the sphere.

**Example 9.3 (sphere)** 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 9.1). However, $S$ is not $(0, \delta)$-regular at $s$.

*Proof.* Let $b \in S$ and $y \in S$. Then $\rho^2 = \|z-y\|^2 = \|z-b\|^2 + \|y-b\|^2 - 2\langle z-b, y-b\rangle = \rho^2 + \|y-b\|^2 - 2\langle z-b, y-b\rangle$, which implies

$$2\langle z-b, y-b\rangle = \|y-b\|^2. \tag{148}$$

On the other hand, by Example 3.6, we have

$$\widehat{N}_S^X(b) \cap \operatorname{sphere}(0;1) = \Big\{ \pm \frac{z-b}{\|z-b\|} \Big\} = \Big\{ \pm \frac{z-b}{\rho} \Big\}. \tag{149}$$

Suppose that $u \in \widehat{N}_S^X(b) \cap \operatorname{sphere}(0;1)$. Combining (148) and (149), we obtain

$$\Big\langle \widehat{N}_S^X(b) \cap \operatorname{sphere}(0;1), y-b \Big\rangle = \Big\{ \pm \frac{1}{2\rho}\|y-b\|^2 \Big\}. \tag{150}$$

Thus if $\|y-s\| \leq \rho\varepsilon$, $\|b-s\| \leq \rho\varepsilon$, and $u \in \widehat{N}_S^X(b) \cap \operatorname{sphere}(0;1)$, then

$$\langle u, y-b\rangle \leq \frac{1}{2\rho}\|y-b\|^2 \leq \frac{1}{2\rho}\big(\|y-s\| + \|s-b\|\big)\|y-b\| \leq \frac{\rho\varepsilon+\rho\varepsilon}{2\rho}\|y-b\| \tag{151}$$

$$= \varepsilon\|u\| \cdot \|y-b\|, \tag{152}$$

which verifies the $(\varepsilon, \rho\varepsilon)$-regularity of $S$ at $s$. Finally, by (150),

$$\max\big\{ \langle \widehat{N}_S^X(b) \cap \operatorname{sphere}(0;1), y-b\rangle \big\} = \frac{1}{2\rho}\|y-b\|^2 > 0 \tag{153}$$

and therefore $S$ is not $(0, \delta)$-regular at $s$. ■

We now characterizes $A$-superregularity using restricted normal cones.

**Theorem 9.4 (characterization of $A$-superregularity)** *Let $A$ and $B$ be nonempty subsets of $X$, and let $c \in X$. Then $B$ is $A$-superregular at $c$ if and only if for every $\varepsilon \in \mathbb{R}_{++}$, there exists $\delta \in \mathbb{R}_{++}$ such that*

$$\left.\begin{array}{c} (y,b) \in B \times B \\ \|y-c\| \leq \delta, \|b-c\| \leq \delta \\ u \in N_B^A(b) \end{array}\right\} \quad \Rightarrow \quad \langle u, y-b\rangle \leq \varepsilon\|u\| \cdot \|y-b\|. \tag{154}$$

*Proof.* "$\Leftarrow$": Clear from Lemma 3.4(iv). "$\Rightarrow$": We argue by contradiction; thus, we assume there exists $\varepsilon \in \mathbb{R}_{++}$ and sequences $(y_n, b_n, u_n)_{n\in\mathbb{N}}$ in $B \times B \times X$ such that $(y_n, b_n) \to (c, c)$ and for every $n \in \mathbb{N}$,

$$u_n \in N_B^A(b_n) \quad \text{and} \quad \langle u_n, y_n - b_n\rangle > \varepsilon\|u_n\| \cdot \|y_n - b_n\|. \tag{155}$$

By the definition of the restricted normal cone, for every $n \in \mathbb{N}$, there exists a sequence $(b_{n,k}, u_{n,k})_{k\in\mathbb{N}}$ in $B \times X$ such that $\lim_{k\in\mathbb{N}} b_{n,k} = b_n$, $\lim_{k\in\mathbb{N}} u_{n,k} = u_n$, and $(\forall k \in \mathbb{N})\ u_{n,k} \in \widehat{N}_B^A(b_{n,k})$. Hence there exists a subsequence $(k_n)_{n\in\mathbb{N}}$ of $(n)_{n\in\mathbb{N}}$ such that $b_{n,k_n} \to c$ and

$$(\forall n \in \mathbb{N}) \quad \langle u_{n,k_n}, y_n - b_{n,k_n}\rangle > \frac{\varepsilon}{2}\|u_{n,k_n}\| \cdot \|y_n - b_{n,k_n}\|. \tag{156}$$

However, this contradicts the $A$-superregularity of $B$ at $c$. ■

When $B = X$, then Theorem 9.4 turns into [17, Proposition 4.4]:

**Corollary 9.5 (Lewis-Luke-Malick)** *Let $B$ be a nonempty subset of $X$ and let $c \in B$. Then $B$ is superregular at $c$ if and only if for every $\varepsilon \in \mathbb{R}_{++}$ there exists $\delta \in \mathbb{R}_{++}$ such that*

$$\left.\begin{array}{c} (y,b) \in B \times B \\ \|y-c\| \le \delta, \|b-c\| \le \delta \\ u \in N_B(b) \end{array}\right\} \quad \Rightarrow \quad \langle u, y-b\rangle \le \varepsilon \|u\| \cdot \|y-b\|. \tag{157}$$

We now introduce the notion of joint-regularity, which is tailored for collections of sets and which turns into Definition 9.1 when the index set is a singleton.

**Definition 9.6 (joint-regularity)** *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 $(A,\varepsilon,\delta)$-joint-regular at $c$ if $\varepsilon \ge 0$, $\delta > 0$, and for every $j \in J$, $B_j$ is $(A,\varepsilon,\delta)$-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 9.1, we may omit the prefix $A$ if $A = X$.*

Here are some verifiable conditions that guarantee joint-(super)regularity.

**Proposition 9.7** *Let $\mathcal{A} := (A_j)_{j\in J}$ and $\mathcal{B} := (B_j)_{j\in J}$ be nontrivial collections of nonempty subsets of $X$, let $c \in X$, let $(\varepsilon_j)_{j\in J}$ be a collection in $\mathbb{R}_+$, and let $(\delta_j)_{j\in J}$ be a collection in $]0,+\infty]$. Set $A := \bigcap_{j\in J} A_j$, $\varepsilon := \sup_{j\in J} \varepsilon_j$, and $\delta := \inf_{j\in J} \delta_j$. Then the following hold:*

(i) *If $\delta > 0$ and $(\forall j \in J)$ $B_j$ is $(A_j,\varepsilon_j,\delta_j)$-regular at $c$, then $\mathcal{B}$ is $(A,\varepsilon,\delta)$-joint-regular at $c$.*

(ii) *If $J$ is finite and $(\forall j \in J)$ $B_j$ is $(A_j,\varepsilon_j,\delta_j)$-regular at $c$, then $\mathcal{B}$ is $(A,\varepsilon,\delta)$-joint-regular at $c$.*

(iii) *If $J$ is finite and $(\forall j \in J)$ $B_j$ is $A_j$-superregular at $c$, then $\mathcal{B}$ is $A$-joint-superregular at $c$.*

*Proof.* (i): Indeed, by Remark 9.2(v), $B_j$ is $(A,\varepsilon,\delta)$-regular at $c$ for every $j \in J$.

(ii): Since $J$ is finite, we have $\delta > 0$ and so the conclusion follows from (i).

(iii): This follows from (ii) and the definitions. ■

**Corollary 9.8 (convexity and regularity)** *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 $c$.*

*Proof.* By Remark 9.2(vi), $B_j$ is $(0,+\infty)$-regular, superregular, and $A$-superregular at $c$, for every $j \in J$. Now apply Proposition 9.7(i)&(iii). ■

The following example illustrates the flexibility gained through the notion of joint-regularity.

**Example 9.9 (two lines: joint-superregularity $\not\Rightarrow$ superregularity of the union)** 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 9.8, $(B_1, B_2)$ 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 $0 < \|y - b\| = \delta\|d_2 - d_1\|$. Using Proposition 3.3(iii), we see that $N_B(b) = \{d_1\}^\perp$. 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.

Let us provide an example of an $A$-superregular set that is not superregular. To do so, we require the following elementary result.

**Lemma 9.10** *Consider in $\mathbb{R}^2$ the sets $C := [(0,1), (m, 1+m^2)] = \big\{(x, 1+mx) \mid x \in [0,m]\big\}$ and $D := [(m,1), (m, 1+m^2)]$, where $m \in \mathbb{R}_{++}$. Let $z \in \mathbb{R}$. Then*

$$P_{C\cup D}(z,0) = \begin{cases} (0,1), & \text{if } z < m/2;\\ \{(0,1),(m,1)\}, & \text{if } z = m/2;\\ (m,1), & \text{if } z > m/2. \end{cases} \tag{158}$$

*Proof.* It is clear that $P_D(z,0) = (m,1)$. We assume that $0 < z < m$ for otherwise (158) is clearly true. We claim that $P_C(z,0) = (0,1)$. Indeed, $f\colon x \mapsto \|(x, 1+mx) - (z,0)\|^2$ is a convex quadratic with minimizer $x_z := (z-m)/(1+m^2)$. The requirement $x_z \geq 0$ from the definition of $C$ forces $z \geq m$, which is a contradiction. Hence $P_C(z,0)$ is a subset of the relative boundary of $C$, i.e., of $\{(0,1),(m,1+m^2)\}$. Clearly, $(0,1)$ is the closer to $(z,0)$ than $(m,1+m^2)$. This verifies the claim. Since $P_{C\cup D}(z,0)$ is the subset of points in $P_C(z,0) \cup P_D(z,0)$ closest to $(z,0)$, the result follows. ■

**Example 9.11 ($A$-superregularity $\not\Rightarrow$ superregularity)** Suppose that $X = \mathbb{R}^2$. As in [17, Example 4.6], we consider $c := (0,0) \in X$ and $B := \operatorname{epi} f$, where

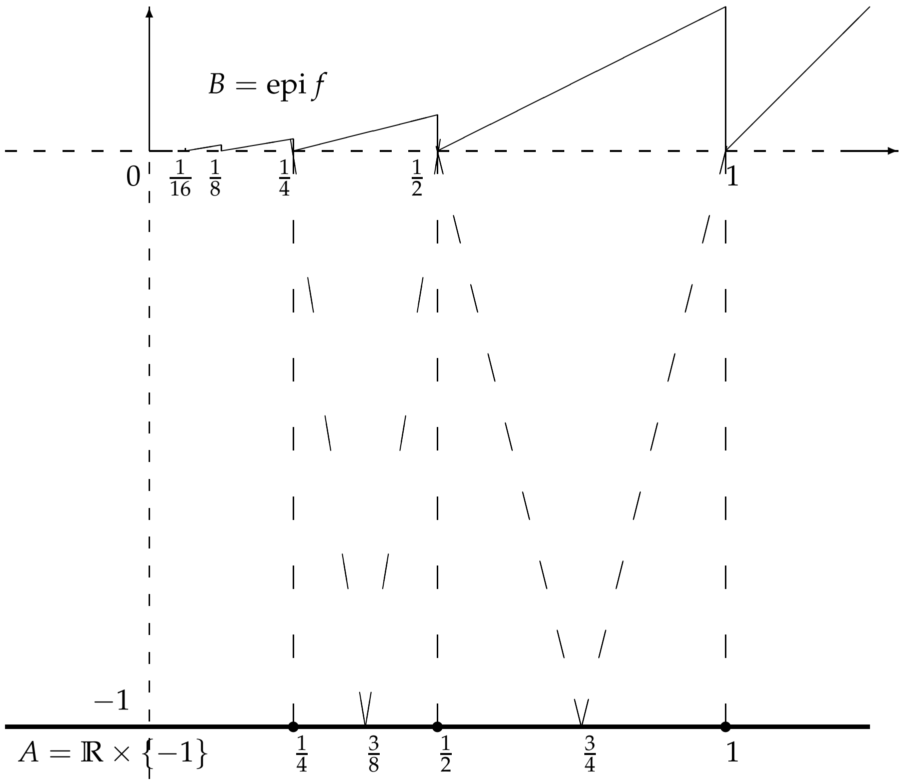
$$f\colon \mathbb{R} \to \left]-\infty,+\infty\right] : x \mapsto \begin{cases} 2^k(x - 2^k), & \text{if } 2^k \leq x < 2^{k+1} \text{ and } k \in \mathbb{Z};\\ 0, & \text{if } x = 0;\\ +\infty, & \text{if } x < 0. \end{cases} \tag{159}$$

Then $B$ is not superregular at $c$; however, $B$ is $A$-superregular at $c$, where $A := \mathbb{R} \times \{-1\}$.

*Proof.* It is stated in [17, Example 4.6] that $B$ is not superregular at $c$ (and that $B$ is Clarke regular at $c$).

To tackle $A$-superregularity, let us determine $P_B(A)$. Let us consider the point $a = (\alpha, -1)$, where $\alpha \in \left[2^{-1}, 1\right[$. Then Lemma 9.10 (see also the picture below) implies that

$$P_B(\alpha,-1) = \begin{cases} (\tfrac{1}{2},0), & \text{if } \tfrac{1}{2} \leq \alpha < \tfrac{3}{4};\\ \big\{(\tfrac{1}{2},0),(1,0)\big\}, & \text{if } \alpha = \tfrac{3}{4};\\ (1,0), & \text{if } \tfrac{3}{4} < \alpha < 1; \end{cases} \tag{160}$$




and more generally,

$$2^k \le \alpha < 2^{k+1} \;\Rightarrow\; P_B(\alpha,-1) = \begin{cases} (2^k,0), & \text{if } 2^k \le \alpha < 2^k + 2^{k-1};\\ \big\{(2^k,0),(2^{k+1},0)\big\}, & \text{if } \alpha = 2^k + 2^{k-1};\\ (2^{k+1},0), & \text{if } 2^k + 2^{k-1} < \alpha < 2^{k+1}. \end{cases} \tag{161}$$

Clearly, if $a \in \mathbb{R}_- \times \{-1\}$, then $P_B(a) = (0,0)$. Let $b \in B$. Then

$$A \cap P_B^{-1}(b) = \begin{cases} \big[2^{k-2}+2^{k-1}, 2^{k-1}+2^k\big] \times \{-1\}, & \text{if } b = (2^k,0) \text{ and } k \in \mathbb{Z};\\ \mathbb{R}_- \times \{-1\}, & \text{if } b = (0,0);\\ \varnothing, & \text{otherwise.} \end{cases} \tag{162}$$

Thus

$$\widehat{N}_B^A(b) = \begin{cases} \operatorname{cone}\Big(\big[-2^{k-2}, 2^{k-1}\big] \times \{-1\}\Big), & \text{if } b = (2^k,0) \text{ and } k \in \mathbb{Z};\\ \{(0,0)\} \cup \big(\mathbb{R}_- \times \mathbb{R}_{--}\big), & \text{if } b = (0,0);\\ \{(0,0)\}, & \text{otherwise.} \end{cases} \tag{163}$$

Let $\varepsilon \in \mathbb{R}_{++}$. Let $K \in \mathbb{Z}$ be such that $2^{K-1} \le \varepsilon$, and let $\delta \in \left]0, 2^K\right]$. Furthermore, let $y = (y_1, y_2) \in B$, let $b = (b_1, b_2) \in B$, let $u \in \widehat{N}_B^A(b)$, and assume that $\|y - c\| \le \delta$ and that $\|b - c\| \le \delta$. We consider three cases.

*Case 1:* $b = (0,0)$. Then $u \in \mathbb{R}^2_-$ and $y \in \mathbb{R}^2_+$; consequently, $\langle u, y-b\rangle = \langle u, y\rangle \le 0 \le \varepsilon\|u\| \cdot \|y - b\|$.

*Case 2:* $b \notin (\{0\} \cup 2^{\mathbb{Z}}) \times \{0\}$. Then $\widehat{N}^A_B(b) = \{(0,0\}$; hence $u = 0$ and so $\langle u, y-b\rangle = 0 \le \varepsilon\|u\| \cdot \|y-b\|$.

*Case 3:* $b \in 2^{\mathbb{Z}} \times \{0\}$, say $b = (2^k, 0)$, where $k \in \mathbb{Z}$. Since $2^k = \|b - 0\| = \|b - c\| \le \delta \le 2^K$, we have $k \le K$. Furthermore, $y_2 \ge 0$, $\max\{|y_1 - b_1|, |y_2 - b_2|\} \le \|y - b\|$, and $u = \lambda(t, -1) = (\lambda t, -\lambda)$ where $t \in [-2^{k-2}, 2^{k-1}]$ and $\lambda \ge 0$. Hence $\lambda \le \|u\|$ and

$$\begin{aligned}
&(164a) && \langle u, y-b\rangle = \lambda t(y_1 - b_1) - \lambda(y_2 - b_2) = \lambda t(y_1 - b_1) - \lambda(y_2 - 0)\\
&(164b) && \le \lambda t(y_1 - b_1) \le \lambda|t| \cdot |y_1 - b|\\
&(164c) && \le \|u\| \cdot 2^{k-1} \cdot \|y-b\| \le 2^{K-1}\|u\| \cdot \|y-b\| \le \varepsilon \cdot \|u\| \cdot \|y-b\|.
\end{aligned}$$

Therefore, in all three cases, we have shown that $\langle u, y-b\rangle \le \varepsilon\|u\| \cdot \|y-b\|$. ■

We now use Example 9.11 to construct an example complementary to Example 9.9.

**Example 9.12 (superregularity of the union $\not\Rightarrow$ joint-superregularity)** Suppose that $X = \mathbb{R}^2$, set $B_1 := \operatorname{epi} f$, where $f$ is as in Example 9.11, $B_2 := X \smallsetminus B_1$, and $c := (0,0)$. Since $B_1 \cup B_2 = X$ is convex, it is clear from Remark 9.2(vi) that $B_1 \cup B_2$ is superregular at $c$. On the other hand, since $B_1$ is not superregular at $c$ (see Example 9.11), it is obvious that $(B_1, B_2)$ is not joint-superregular at $c$.

# 10 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 10.1** *Let $A$ and $B$ be nonempty closed subsets of $X$. Then the following hold*[3]*:*

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

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

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

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

$$a \in (\operatorname{bdry} A) \smallsetminus B \;\Leftrightarrow\; a \in A \smallsetminus B \;\Rightarrow\; b \in B \smallsetminus A \;\Rightarrow\; a \in \operatorname{bdry} A. \tag{165}$$

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

$$b \in (\operatorname{bdry} B) \smallsetminus A \;\Leftrightarrow\; b \in B \smallsetminus A \;\Rightarrow\; a \in A \smallsetminus B \;\Rightarrow\; b \in \operatorname{bdry} B. \tag{166}$$

*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 $\operatorname{aff}(A \cup B) \cap \operatorname{ball}(a;\delta) \subseteq A$. Hence $\widetilde{a} := a + \delta(b-a)/\|b-a\| \in A$ 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).

(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 10.2** *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

$$d(a,b) = d_B(a) \leq d(a,y) = d_A(y) \leq \delta. \tag{167}$$

Thus,

$$d(a,c) \leq d(a,y) + d(y,c) \leq \delta + \delta = 2\delta, \tag{168}$$

which establishes (i). Using (167), 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 (167) and (168), we obtain (iii) via $d(b,c) \leq d(b,a) + d(a,c) \leq \delta + 2\delta = 3\delta$. ■

**Corollary 10.3** *Let $A$ and $B$ be nonempty closed subsets of $X$, let $\rho \in \mathbb{R}_{++}$, and suppose that $c \in A \cap B$. Then*

$$P_A P_B P_A \operatorname{ball}(c;\rho) \subseteq \operatorname{ball}(c; 6\rho). \tag{169}$$

*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 10.2(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 10.4 (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 9.1). 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*

$$\|a - b\| \leq (\theta_{3\delta} + 2\varepsilon)\|a - y\|, \tag{170}$$

*where $\theta_{3\delta}$ the CQ-number at $c$ associated with $(A, \widetilde{A}, B, \widetilde{B})$ (see (95)).*

*Proof.* Lemma 10.2(i)&(iii) yields $\|a - c\| \leq 2\delta$ and $\|b - c\| \leq 3\delta$. On the other hand, $y - a \in \widehat{N}_A^{\widetilde{B}}(a)$ and $b - a \in -\widehat{N}_B^{\widetilde{A}}(b)$. Therefore,

$$\langle b - a, y - a\rangle \leq \theta_{3\delta}\|b - a\| \cdot \|y - a\|. \tag{171}$$

Since $a - b \in \widehat{N}_B^{\widetilde{A}}(b)$, $\|y - c\| \leq \delta$, and $\|b - c\| \leq 3\delta$, we obtain, using the $(\widetilde{A}, \varepsilon, 3\delta)$-regularity of $B$, that $\langle a - b, y - b\rangle \leq \varepsilon\|a - b\| \cdot \|y - b\|$. Moreover, Lemma 10.2(ii) states that $\|y - b\| \leq 2\|a - y\|$. It follows that

$$\langle a - b, y - b\rangle \leq 2\varepsilon\|a - b\| \cdot \|a - y\|. \tag{172}$$

Adding (171) and (172) yields $\|a - b\|^2 \leq (\theta_{3\delta} + 2\varepsilon)\|a - b\| \cdot \|a - y\|$. The result follows. ■

We now provide a result for collections of sets similar to—and relying upon—Proposition 10.4.

**Proposition 10.5 (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 $\varepsilon \geq 0$ 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:*

$$b \in \big((\operatorname{bdry} B) \smallsetminus A\big) \cap \bigcup_{j\in J}(B_j \cap \widetilde{B}_j) \ \Leftrightarrow\ b \in B \smallsetminus A \ \Rightarrow\ a \in A \smallsetminus B. \tag{173}$$

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

$$a \in \big((\operatorname{bdry} A) \smallsetminus B\big) \cap \bigcup_{i\in I}(A_i \cap \widetilde{A}_i) \;\Leftrightarrow\; a \in A \smallsetminus B \;\Rightarrow\; b \in B \smallsetminus A. \tag{174}$$

(v) *Suppose that* $\mathcal{B}$ *is* $(\widetilde{A}, \varepsilon, 3\delta)$*-joint-regular at* $c$ *(see Definition 9.6), 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*

$$\|b - a\| \le (\theta_{3\delta} + 2\varepsilon)\|a - y\|, \tag{175}$$

*where* $\theta_{3\delta}$ *is the joint-CQ-number at* $c$ *associated with* $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ *(see* (101)*).*

(vi) *Suppose that* $\mathcal{A}$ *is* $(\widetilde{B}, \varepsilon, 3\delta)$*-joint-regular at* $c$ *(see Definition 9.6), 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*

$$\|a - b\| \le (\theta_{3\delta} + 2\varepsilon)\|b - x\|, \tag{176}$$

*where* $\theta_{3\delta}$ *is the joint-CQ-number at* $c$ *associated with* $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ *(see* (101)*).*

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

(iii): Note that Lemma 10.1(iv)&(iii) and (ii) yield the implications

$$b \in B \smallsetminus A \;\Leftrightarrow\; b \in (\operatorname{bdry} B) \smallsetminus A \;\Rightarrow\; a \in A \smallsetminus B \;\Leftrightarrow\; a \in (\operatorname{bdry} A) \smallsetminus B \;\Rightarrow\; b \in \bigcup_{j\in J}(B_j \cap \widetilde{B}_j), \tag{177}$$

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

$$\|a - b\| = d_B(a) \le d_{B_j}(a) = \|a - b'\|. \tag{178}$$

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 9.2(v)) $B_j$ 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 $d_{A_i}(y) = d_A(y) = \|y - a\| \le \|y - c\| \le \delta$, we obtain from Proposition 10.4 that

$$\|a - b'\| \le \big(\theta_{3\delta}(A_i, \widetilde{A}_i, B_j, \widetilde{B}_j) + 2\varepsilon\big)\|a - y\|. \tag{179}$$

Combining with (178), we deduce that $\|a - b\| \le \|a - b'\| \le (\theta_{3\delta} + 2\varepsilon)\|a - y\|$.

(vi): This follows from (v) and (97). ■

## An abstract linear convergence result

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

**Definition 10.6 (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 $\gamma$ *if there exists $\mu \in \mathbb{R}_+$ such that*

$$(\forall n \in \mathbb{N}) \quad d(x_n, \bar{x}) \leq \mu\gamma^n. \tag{180}$$

**Remark 10.7 (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

$$(\forall n \in \{n_0, n_0+1, \ldots\}) \quad d(x_n, \bar{x}) \leq \mu_0\gamma^n. \tag{181}$$

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

$$(\forall n \in \mathbb{N}) \quad d(x_n, \bar{x}) \leq \max\{\mu_0, \mu_1\}\gamma^n, \tag{182}$$

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

**Proposition 10.8 (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*

$$\gamma := \alpha\beta < 1 \tag{183a}$$

*and*

$$(\forall n \in \mathbb{N}) \quad d(a_{n+1}, b_n) \leq \alpha d(a_n, b_n) \text{ and } d(a_{n+1}, b_{n+1}) \leq \beta d(a_{n+1}, b_n). \tag{183b}$$

*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*

$$(\forall n \in \mathbb{N}) \quad \max\{d(a_n, c), d(b_n, c)\} \leq \frac{1+\alpha}{1-\gamma} d(a_0, b_0) \cdot \gamma^n; \tag{184}$$

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

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

$$d(a_n, b_n) \leq \beta d(a_n, b_{n-1}) \leq \alpha\beta d(a_{n-1}, b_{n-1}) = \gamma d(a_{n-1}, b_{n-1}) \leq \cdots \leq \gamma^n\delta; \tag{185}$$

hence,

$$d(b_n, b_{n+1}) \leq d(b_n, a_{n+1}) + d(a_{n+1}, b_{n+1}) \leq \alpha d(b_n, a_n) + \gamma d(a_n, b_n) \tag{186a}$$
$$= (\alpha+\gamma) d(a_n, b_n) \leq (\alpha+\gamma)\delta\gamma^n. \tag{186b}$$

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 (185), $d(a_n, b_n) \to 0$ 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 (186),

$$d(b_n, b_m) \leq \sum_{k=n}^{m-1} d(b_k, b_{k+1}) \leq \sum_{k\geq n} d(b_k, b_{k+1}) \leq \sum_{k\geq n} (\alpha+\gamma)\delta\gamma^k = \frac{(\alpha+\gamma)\delta\gamma^n}{1-\gamma}. \tag{187}$$

Hence, using (185) and (187), we estimate that

$$d(a_n, b_m) \leq d(a_n, b_n) + d(b_n, b_m) \leq \delta\gamma^n + \frac{(\alpha+\gamma)\delta\gamma^n}{1-\gamma} = \frac{(1+\alpha)\delta\gamma^n}{1-\gamma}. \tag{188}$$

Letting $m \to +\infty$ in (187) and (188), we obtain (184). ■

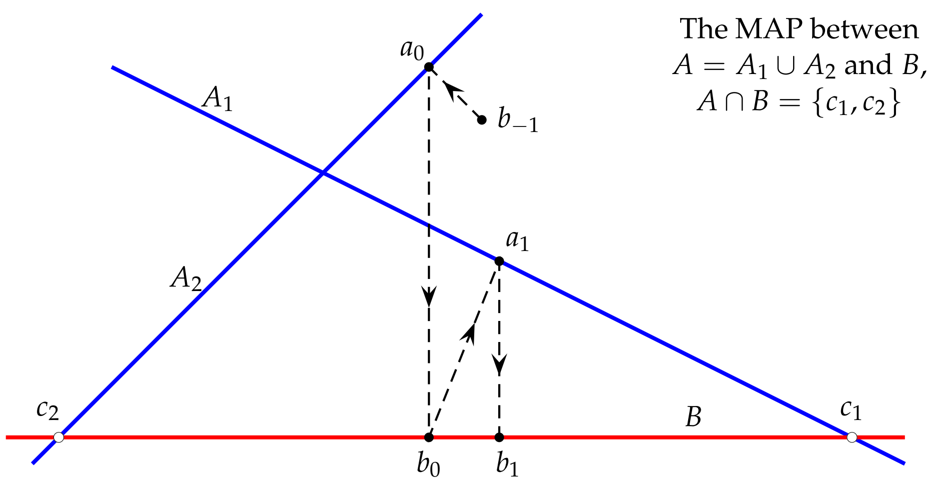
### The sequence generated by the MAP

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

**Definition 10.9 (MAP)** *Let $A$ and $B$ be nonempty closed subsets of $X$, let $b_{-1} \in X$, and let*

$$(\forall n \in \mathbb{N}) \quad a_n \in P_A(b_{n-1}) \text{ and } b_n \in P_B(a_n). \tag{189}$$

*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 10.10** *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:

$$\begin{cases} \mathcal{A} := (A_i)_{i\in I} \text{ and } \mathcal{B} := (B_j)_{j\in J} \text{ are nontrivial collections} \\ \quad \text{of nonempty closed subsets of } X; \\ A := \bigcup_{i\in I} A_i \text{ and } B := \bigcup_{j\in J} B_j \text{ are closed;} \\ c \in A \cap B; \\ \widetilde{\mathcal{A}} := (\widetilde{A}_i)_{i\in I} \text{ and } \widetilde{\mathcal{B}} := (\widetilde{B}_j)_{j\in J} \text{ are collections} \\ \quad \text{of nonempty subsets of } X \text{ such that} \\ \qquad (\forall i \in I)\ \ P_{A_i}\big((\operatorname{bdry} B) \smallsetminus A\big) \subseteq \widetilde{A}_i, \\ \qquad (\forall j \in J)\ \ P_{B_j}\big((\operatorname{bdry} A) \smallsetminus B\big) \subseteq \widetilde{B}_j; \\ \widetilde{A} := \bigcup_{i\in I} \widetilde{A}_i \text{ and } \widetilde{B} := \bigcup_{j\in J} \widetilde{B}_j. \end{cases} \tag{190}$$

**Lemma 10.11 (backtracking MAP)** *Assume that* (190) *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* $b_{n-1} \in ((\mathrm{bdry}\, B) \smallsetminus A) \cap \bigcup_{j \in J}(B_j \cap \widetilde{B}_j)$.

*Proof.* (i): Applying Proposition 10.5(iii) to $b_{n-1} \in B$, $a_n \in P_A b_{n-1}$, $b_n \in P_B a_n$, we obtain

$$b_n \in B \smallsetminus A \;\Leftrightarrow\; b_n \in \big((\mathrm{bdry}\, B) \smallsetminus A\big) \cap \bigcup_{j \in J}(B_j \cap \widetilde{B}_j) \;\Rightarrow\; a_n \in A \smallsetminus B. \tag{191}$$

On the other hand, applying Proposition 10.5(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

$$a_n \in A \smallsetminus B \;\Leftrightarrow\; a_n \in \big((\mathrm{bdry}\, A) \smallsetminus B\big) \cap \bigcup_{i \in I}(A_i \cap \widetilde{A}_i). \tag{192}$$

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 10.12** *Assume that* (190) *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 9.6) and set*

$$\sigma := \begin{cases} 1, & \text{if } \mathcal{B} \text{ is not known to be } (\widetilde{A}, \varepsilon, 3\delta)\text{-joint-regular at } c; \\ 2, & \text{if } \mathcal{B} \text{ is also } (\widetilde{A}, \varepsilon, 3\delta)\text{-joint-regular at } c. \end{cases} \tag{193}$$

(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 7.2).*

*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*

$$\|b_{-1} - c\| \leq \frac{(1 - \theta^\sigma)\delta}{6(2 + \theta - \theta^\sigma)}. \tag{194}$$

*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* $\theta^\sigma$*; in fact,*

$$\|\bar{c} - c\| \leq \delta \quad \text{and} \quad (\forall n \geq 1) \;\; \max\big\{\|a_n - \bar{c}\|, \|b_n - \bar{c}\|\big\} \leq \frac{\delta(1+\theta)}{2 + \theta - \theta^\sigma}\theta^{\sigma(n-1)}. \tag{195}$$

*Proof.* In view of $a_1 \in P_A P_B P_A b_{-1}$ and (194), Corollary 10.3 yields

$$\beta := \|a_1 - c\| \leq \frac{(1 - \theta^\sigma)\delta}{(2 + \theta - \theta^\sigma)} \leq \frac{\delta}{2}. \tag{196}$$

Since $c \in A \cap B$, we have $\theta_{3\delta} \geq 0$ by (98) and hence $\theta > 0$. Using (196), we estimate

$$(\forall n \geq 1) \quad \beta\theta^{\sigma(n-1)} + \beta + \beta(1+\theta)\sum_{k=0}^{n-2}\theta^{\sigma k} \leq \beta + \beta(1+\theta)\sum_{k=0}^{n-1}\theta^{\sigma k} \tag{197a}$$
$$= \beta + \beta(1+\theta)\frac{1-\theta^{\sigma n}}{1-\theta^{\sigma}} \tag{197b}$$
$$\leq \beta + \beta\frac{1+\theta}{1-\theta^{\sigma}} \tag{197c}$$
$$= \beta\Big(\frac{2+\theta-\theta^{\sigma}}{1-\theta^{\sigma}}\Big) \tag{197d}$$
$$\leq \delta. \tag{197e}$$

We now claim that if

$$n \geq 1, \quad \|a_n - b_n\| \leq \beta\theta^{\sigma(n-1)} \quad \text{and} \quad \|a_n - c\| \leq \beta + \beta(1+\theta)\sum_{k=0}^{n-2}\theta^{\sigma k}, \tag{198}$$

then

$$\|a_{n+1} - b_{n+1}\| \leq \theta^{\sigma-1}\|a_{n+1} - b_n\| \leq \theta^{\sigma}\|a_n - b_n\| \leq \beta\theta^{\sigma n}, \tag{199a}$$
$$\|a_{n+1} - c\| \leq \beta + \beta(1+\theta)\sum_{k=0}^{n-1}\theta^{\sigma k}. \tag{199b}$$

To prove this claim, assume that (198) holds. Using (198) and (197), we first observe that

$$\max\big\{\|a_n - c\|, \|b_n - c\|\big\} \leq \|b_n - a_n\| + \|a_n - c\| \tag{200a}$$
$$\leq \beta\theta^{\sigma(n-1)} + \beta + \beta(1+\theta)\sum_{k=0}^{n-2}\theta^{\sigma k} \leq \delta. \tag{200b}$$

We now consider two cases:

*Case 1*: $b_n \in A \cap B$. Then $b_n = a_{n+1} = b_{n+1}$ and thus (199a) holds. Moreover, $\|a_{n+1} - c\| = \|b_n - c\|$ and (199b) follows from (200a).

*Case 2*: $b_n \notin A \cap B$. Then $b_n \in B \smallsetminus A$. Lemma 10.11(i) implies $a_n \in ((\mathrm{bdry}\, A) \smallsetminus B) \cap \bigcup_{i\in I}(A_i \cap \widetilde{A}_i)$ and $b_n \in ((\mathrm{bdry}\, B) \smallsetminus A) \cap \bigcup_{j\in J}(B_j \cap \widetilde{B}_j)$. Note that $\|a_n - c\| \leq \delta$ by (200a), and recall that $\mathcal{A}$ is $(\widetilde{B}, \varepsilon, 3\delta)$-joint-regular at $c$ by (i). It thus follows from Proposition 10.5(vi) (applied to $a_n, b_n, a_{n+1}$) that

$$\|a_{n+1} - b_n\| \leq \theta\|a_n - b_n\|. \tag{201}$$

On the one hand, if $\sigma = 1$, then Proposition 10.10(ii) yields $\|a_{n+1} - b_{n+1}\| \leq \|a_{n+1} - b_n\| = \theta^{\sigma-1}\|a_{n+1} - b_n\|$. On the other hand, if $\sigma = 2$, then $\mathcal{B}$ is $(\widetilde{A}, \varepsilon, 3\delta)$-joint-regular at $c$ by (i); hence, Proposition 10.5(v) (applied to $b_n, a_{n+1}, b_{n+1}$) yields $\|a_{n+1} - b_{n+1}\| \leq \theta\|a_{n+1} - b_n\| = \theta^{\sigma-1}\|a_{n+1} - b_n\|$. Altogether, in either case,

$$\|a_{n+1} - b_{n+1}\| \leq \theta^{\sigma-1}\|a_{n+1} - b_n\|. \tag{202}$$

Combining (202) with (201) and (198) gives

$$\|a_{n+1} - b_{n+1}\| \leq \theta^{\sigma-1}\|a_{n+1} - b_n\| \leq \theta^{\sigma}\|a_n - b_n\| \leq \beta\theta^{\sigma n}, \tag{203}$$

which is (199a). Furthermore, (201), (198) and (200a) yield

$$\begin{aligned}
\|a_{n+1} - c\| &\leq \|a_{n+1} - b_n\| + \|b_n - c\| && \text{(204a)}\\
&\leq \theta\|a_n - b_n\| + \|b_n - c\| && \text{(204b)}\\
&\leq \theta\beta\theta^{\sigma(n-1)} + \beta\theta^{\sigma(n-1)} + \beta + \beta(1+\theta)\sum_{k=0}^{n-2}\theta^{\sigma k} && \text{(204c)}\\
&= \beta + \beta(1+\theta)\sum_{k=0}^{n-1}\theta^{\sigma k}, && \text{(204d)}
\end{aligned}$$

which establishes (199b). Therefore, in all cases, (199) holds.

Since $\|a_1 - b_1\| = d_B(a_1) \leq \|a_1 - c\| = \beta$, we see that (198) holds for $n = 1$. Thus, the above claim and the principle of mathematical induction principle imply that (199) holds for every $n \geq 1$.

Next, (199a) implies

$$(\forall n \geq 1) \quad \|a_{n+1} - b_n\| \leq \theta\|a_n - b_n\| \quad \text{and} \quad \|a_{n+1} - b_{n+1}\| \leq \theta^{\sigma-1}\|a_{n+1} - b_n\|. \tag{205}$$

In view of (205) and $\|a_1 - b_1\| \leq \beta$, Proposition 10.8 yields $\bar{c} \in A \cap B$ such that

$$\begin{aligned}
(\forall n \geq 1) \quad \max\{\|a_n - \bar{c}\|, \|b_n - \bar{c}\|\} &\leq \frac{1+\theta}{1-\theta^{\sigma}}\|a_1 - b_1\| \cdot \theta^{\sigma(n-1)} && \text{(206)}\\
&\leq \frac{1+\theta}{1-\theta^{\sigma}}\beta \cdot \theta^{\sigma(n-1)} && \text{(207)}\\
&\leq \frac{\delta(1+\theta)}{2+\theta-\theta^{\sigma}}\theta^{\sigma(n-1)}. && \text{(208)}
\end{aligned}$$

On the other hand, (199b) and (197) 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 (195). ■

**Remark 10.13** In view of Lemma 10.1(i)&(ii), an aggressive choice for use in (190) is $(\forall i \in I)$ $\widetilde{A}_i = \text{bdry}\, A_i$ and $(\forall j \in J)$ $\widetilde{B}_j = \text{bdry}\, B_j$.

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

**Theorem 10.14 (linear convergence of the MAP and superregularity)** *Assume that* (190) *holds and that $\mathcal{A}$ is $\widetilde{B}$-joint-superregular at $c$ (see Definition 9.6). Denote the limiting joint-CQ-number at $c$ associated with $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ (see Definition 7.2) 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 7.7) 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* $\theta \in \,]\overline{\theta}, 1[$ *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 9.6).*

(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 7.2).*

*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* $\theta$:

$$(\forall n \geq 1)\ \max\{\|a_n - \bar{c}\|, \|b_n - \bar{c}\|\} \leq \frac{\delta(1+\theta)}{2}\theta^{n-1}. \tag{209}$$

*Proof*. Observe that (ii) implies (i) by Theorem 7.8(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 10.12 with $\sigma = 1$. ■

**Corollary 10.15** *Assume that* (190) *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 7.2) by* $\overline{\theta}$, *and assume that* $\overline{\theta} < 1$. *Let* $\theta \in \,]\overline{\theta}, 1[$, *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* $\theta$.

*Proof*. Combine Theorem 10.14 with Corollary 9.8. ■

**Example 10.16 (working with collections and joint notions is useful)** Consider the setting of Example 7.4, 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 10.15 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 9.9! This illustrates that the main result of Lewis-Luke-Malick (see Corollary 10.24 below) is not applicable because two of its hypotheses fail.

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

**Theorem 10.17 (linear convergence of the MAP and double superregularity)** *Assume that* (190) *holds, that* $\mathcal{A}$ *is* $\widetilde{B}$*-joint-superregular at c and that* $\mathcal{B}$ *is* $\widetilde{A}$*-joint-superregular at c (see Definition 9.6). Denote the limiting joint-CQ-number at c associated with* $(\mathcal{A}, \widetilde{\mathcal{A}}, \mathcal{B}, \widetilde{\mathcal{B}})$ *(see Definition 7.2) 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 7.7) 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* $\theta \in \,]\overline{\theta}, 1[$ *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 7.2).*

*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* $\theta^2$*; in fact,*

$$(\forall n \geq 1) \quad \max\left\{\|a_n - \bar{c}\|, \|b_n - \bar{c}\|\right\} \leq \frac{\delta}{2-\theta}\left(\theta^2\right)^{n-1}. \tag{210}$$

*Proof.* The existence of $\delta > 0$ such that (i)–(iii) hold is clear. Then apply Proposition 10.12 with $\sigma = 2$. ■

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

**Corollary 10.18 (refined convergence rate)** *Assume that* (190) *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 7.2).*

*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* $\theta^2$*; in fact,*

$$(\forall n \geq 1) \quad \max\left\{\|a_n - \bar{c}\|, \|b_n - \bar{c}\|\right\} \leq \frac{\delta}{2-\theta}\left(\theta^2\right)^{n-1}. \tag{211}$$

*Proof.* Apply Proposition 10.12 with $\sigma = 2$. ■

Let us illustrate a situation where it is possible to make $\delta$ in Theorem 10.17 precise.

**Example 10.19 (the MAP for two spheres)** Let $z_1$ and $z_2$ be in $X$, let $\rho_1$ and $\rho_2$ 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 $\theta \in \left]\overline{\alpha}, 1\right[$. Then the conclusion of Theorem 10.17 holds with

$$\delta := \min\left\{\frac{\sqrt{(\rho_1+\rho_2)^2 + \rho_1\rho_2(\theta - \overline{\alpha})} - (\rho_1+\rho_2)}{6}, \frac{\varepsilon\rho_1}{3}, \frac{\varepsilon\rho_2}{3}\right\} \tag{212}$$

*Proof*. Combine Example 9.3 (applied with $\varepsilon = (\theta - \overline{\alpha})/4$ there), Proposition 8.4, and Theorem 10.17. ■

Here is a useful special case of Theorem 10.17:

**Theorem 10.20** *Assume that $A$ and $B$ are $L$-superregular, and that*

$$N_A(c) \cap \big( - N_B(c)\big) \cap \big(L - c\big) = \{0\}, \tag{213}$$

*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 8.2 with Theorem 10.17 (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, which does not require the starting point to be sufficiently close to $A \cap B$:

**Theorem 10.21 (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

$$(\forall c \in A \cap B) \quad \|a_0 - c\| \geq \|b_0 - c\| \geq \|a_1 - c\| \geq \|b_1 - c\| \geq \cdots \tag{214}$$

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

$$\|a_{k_n} - c\|^2 \geq \|a_{k_n} - b_{k_n}\|^2 + \|b_{k_n} - c\|^2 \geq \varepsilon^2 + \|a_{k_n+1} - c\|^2 \geq \varepsilon^2 + \|a_{k_{n+1}} - c\|^2 \tag{215}$$

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 (214) (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 8.5, the $(A, L, B, L)$-CQ conditions holds at $\bar{c}$. Thus, by Example 8.2, $N_A(\bar{c}) \cap (-N_B(\bar{c})) \cap (L - \bar{c}) = \{0\}$. Furthermore, Corollary 9.8 and Remark 9.2(vi)&(viii) imply that $A$ and $B$ are $L$-superregular at $\bar{c}$. The conclusion now follows from Theorem 10.20, applied to suitably chosen tails of the sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_{n\in\mathbb{N}})$. ■

**Example 10.22 (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 10.21 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 9.2(vi)). 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 8.12), we obtain from Corollary 10.18 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 [10, Section 3] and [11, Chapter 9].

**Remark 10.23** For further linear convergence results for the MAP in the convex setting we refer the reader to [1], [2], [3], [12], [13], [14], and the references therein. See also [20] and [21] 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 10.14:

**Corollary 10.24 (Lewis-Luke-Malick)** (See [17, 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 10.14(i) with $\widetilde{\mathcal{A}} := \widetilde{\mathcal{B}} := (X)$, $\mathcal{A} := (A)$ and $\mathcal{B} := (B)$. ■

However, even in simple situations, Corollary 10.24 is not powerful enough to recover known convergence results.

**Example 10.25 (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

$$N_A(c) \cap (-N_B(c)) = A^{\perp} \cap B^{\perp} = (A+B)^{\perp} = L^{\perp}. \tag{216}$$

Therefore, the Lewis-Luke-Malick CQ (see [17, Theorem 5.16] and also Corollary 10.24) holds for $(A, B)$ at $c$ if and only if

$$N_A(c) \cap (-N_B(c)) = \{0\} \;\Leftrightarrow\; A + B = X. \tag{217}$$

On the other hand, the CQ provided in Theorem 10.20 (see also Example 10.22) *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 10.24) is unable to achieve this. (It was this example that originally motivated us to pursue the present work.)

**Example 10.26 (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 (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 4.13).

## 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 10.20. 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 10.27** **($(\widetilde{A},\widetilde{B}) = (A,B)$)** Let $A$, $B$, and $c$ be as in Example 8.3, 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 8.5), 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 7.8(v). Thus linear convergence of the MAP is guaranteed by Theorem 10.17.

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 10.28** **($(\widetilde{A},\widetilde{B}) = (\operatorname{bdry} A, \operatorname{bdry} B)$)** 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

$$N_A^B(c) = N_A^X(c) = N_A(c) = \big\{(u_1,u_2) \in \mathbb{R}^2 \;\big|\; u_2 + 2|u_1| \leq 0\big\}, \tag{218a}$$

$$N_B^A(c) = N_B^X(c) = N_B(c) = \big\{(u_1,u_2) \in \mathbb{R}^2 \;\big|\; -u_2 + 3|u_1| \leq 0\big\}, \tag{218b}$$

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

$$N_A^B(c) \cap (-N_B^A(c)) = \big\{(u_1,u_2) \in \mathbb{R}^2 \;\big|\; u_2 + 3|u_1| \leq 0\big\} \neq \{0\}. \tag{219}$$

Consequently, for either $(\widetilde{A},\widetilde{B}) = (A,B)$ or $(\widetilde{A},\widetilde{B}) = (X,X)$, Theorem 10.17 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 10.13. Then

$$N_A^{\widetilde{B}}(c) = \big\{(u_1,u_2) \in \mathbb{R}^2 \;\big|\; u_2 + 2|u_1| = 0\big\}, \tag{220a}$$

$$N_B^{\widetilde{A}}(c) = \big\{(u_1,u_2) \in \mathbb{R}^2 \;\big|\; -u_2 + 3|u_1| = 0\big\}, \tag{220b}$$

$N_A^{\widetilde{B}}(c) \cap (-N_B^{\widetilde{A}}(c)) = \{0\}$ and the $(A,\widetilde{A},B,\widetilde{B})$-CQ condition holds. Hence, using also Theorem 7.8(v), Theorem 10.21 and Theorem 10.17, 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 10.29** **($(\widetilde{A},\widetilde{B}) = (P_A((\operatorname{bdry} B) \smallsetminus A), P_B((\operatorname{bdry} A) \smallsetminus B))$)** 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 $N_A^{\operatorname{bdry} B}(c) \cap (-N_B^{\operatorname{bdry} A}(c)) = \operatorname{bdry} B \neq \{0\}$. On the other hand, if $(\widetilde{A},\widetilde{B}) = (P_A((\operatorname{bdry} B) \smallsetminus A), P_B((\operatorname{bdry} A) \smallsetminus B))$, 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 have introduced restricted normal cones which generalize classical normal cones. We have presented some of their basic properties and shown their usefulness in describing interiority conditions, constraint qualifications, and regularities. The corresponding results were employed to yield new powerful sufficient conditions 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 least aggressive choice, $(\widetilde{A}, \widetilde{B}) = (X, X)$, recovers the framework by Lewis, Luke, and Malick. The choice $(\widetilde{A}, \widetilde{B}) = (\operatorname{aff}(A \cup B), \operatorname{aff}(A \cup B))$ 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 $(\widetilde{A}, \widetilde{B}) = (\operatorname{aff}(A \cup B), \operatorname{aff}(A \cup B))$ fails, one may also try more aggressive choices such as $(\widetilde{A}, \widetilde{B}) = (A, B)$ or $(\widetilde{A}, \widetilde{B}) = (\operatorname{bdry} A, \operatorname{bdry} B)$ to guarantee linear convergence. In a follow-up work [5] 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 rates and radii of convergence, where all other sufficient conditions, particularly those of [18] and [17], fail.

### Acknowledgments

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

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

[3] 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.

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

[5] 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`

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

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

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

[9] 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.

[10] 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.

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

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

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

[19] P.D. Loewen, *Optimal Control via Nonsmooth Analysis*, CRM Proceedings & Lecture Notes, AMS, Providence, RI, 1993.

[20] 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.

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

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

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

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

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

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