Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Approximation Theory 134 (2005) 80–101

Journal of Approximation Theory

www.elsevier.com/locate/jat

A comparative study of various notions of approximation of sets

Ali Agadi, Jean-Paul Penot*

Laboratoire de Mathématiques Appliquées, Université de Pau, UMR CNRS 5142, France

Received 1 April 1996; received in revised form 19 September 2004; accepted 6 October 2004

Communicated by P.B. Borwein
Available online 10 March 2005

Abstract

We compare various notions of approximations of sets. Several of them are one-sided versions of existing notions. We devote a particular attention to the case where the approximating set is a translated cone. We point out some consequences for nonsmooth analysis and optimization.

Keywords: Approximation; Distance function; Normal cone; Optimality condition; Tangent cone

1. Introduction

Numerous papers have been devoted to the approximation of functions; a comparatively small number of papers deal with approximation of sets. Still several notions of approximation of sets have been introduced in the literature for various aims, either in a one-sided way [22,23,28] or in a symmetric way [4,39,41,46,47]. The first appearance of a notion of local approximation of a set goes back at least to the work of Lusternik (see [20]) who gave conditions in order that the set of solutions of the nonlinear equation

$$f(x) = 0, \qquad (1)$$

* Corresponding author. Département de Mathématiques Appliquées, Faculté des Sciences, Av. de l'Université, 64000 Pau Cedex, France. Fax: +33 5 59 40 75 55.

E-mail address: jean-paul.penot@univ-pau.fr (J.-P. Penot).

doi:10.1016/j.jat.2004.10.015

given by a mapping $f : X \to Y$ between two Banach spaces, is approximated around a solution x_0 at which f is differentiable by the set of solutions of the linearized equation

$$f'(x_0)(x - x_0) = 0. \tag{2}$$

As one can imagine, such a topic is of interest in mathematics and outside mathematics. As a matter of fact, a number of practitioners linearize nonlinear phenomena without checking the assumptions which would guarantee the validity of the approximation. Still, simple examples show that without surjectivity of the derivative of f at x_0 or some condition replacing this assumption, the local behavior of the solution set S of (1) may be very different from the local behavior of the solution set $S'(x_0)$ of (2).

In [22,23,28], a notion of approximation is used to give necessary and sufficient optimality conditions for some mathematical programming problems in infinite-dimensional spaces. In [4,11,46,47], the authors have used the notions of approximation cone and of proto-differentiability of sets to study the differentiability of the metric projection. In [13,15–17], the Fréchet differentiability of a mapping between two normed vector spaces (n.v.s.'s) is characterized with the help of a notion of conic convergence and a concept of Fréchet approximating cone. In [39,41], the Hausdorff–Pompeiu distance is used to define a notion of tangency for sets which can serve to study the stability of systems of inequalities. In a companion paper [34], the differentiability and subdifferentiability of the distance function to a closed subset is studied with the help of the notion of approximation.

Our aim in this paper is to introduce unilateral (or one-sided) approximation notions of some known symmetric tangency concepts and to present a comparative study of the various notions one can find in the literature. We show in Section 3 that almost all presented notions are equivalent. Our approach is related to some recent notions of convergence for families of sets, but we do not insist on this aspect, although it is important for applications (see, for instance [5,6,7,32,33,36,37,45,49] for recent contributions and references). In Section 4, we give necessary conditions and sufficient conditions for the existence of an approximation cone, and we point out the links with proto-differentiability and B-differentiability, two notions which have been extensively used for the study of variational inequalities and generalized equations [25,27,42], etc. The remaining part of the paper is devoted to applications. In Section 5, we show that the existence of an approximating cone ensures equality between the normal cone and the normal cone in the sense of Fréchet. Using some transversality (or qualification) conditions, we give a criterion in order that the constraint set of a mathematical programming problem is approximated by its linearizing cone; in fact our study is more general and bears on the preservation of the approximation property under some operations such as intersections and inverse images.

Let us mention briefly some other motivations for the present study. Each of the problems we mention requires a precise definition of the notion of approximation which is involved, what justifies the comparison we undertake here.

(A) Let C be a closed convex subset of a n.v.s. E and let $c_1, c_2 : [0, 1] \to E$ which are tangent at 0; under which conditions can one assert that the projections $P_C \circ c_1$ and $P_C \circ c_2$ are tangent at 0? A similar question can be raised for the f-projections and f-farthest point mappings in the sense of Pai and Govindarajulu [26] and for centers as in Beer and Pai [8].

(B) Let A be an approximation at some point e of some subset S of a n.v.s. E. Let $f : X \to E$ be a mapping from another n.v.s. into E such that for x close to some point $x_0 \in X$ the point $f(x)$ has best approximations $a(x)$ and $s(x)$ in A and S, respectively. Under which conditions are $a(\cdot)$ and $s(\cdot)$ tangent at x_0?

(C) Let F and G be two set-valued mappings from a metric space X into some Euclidean space E. Suppose G is an approximation to F at some point x_0 of X. Assume some continuity conditions (such as in [12,18,38,52] for instance) ensuring the existence of continuous selections of F and G. Can one find continuous selections f and g of F and G, respectively, which are tangent at x_0?

(D) Given a set-valued mapping F from $[0, 1]$ into some Euclidean space E and some subdivision σ of the interval $[0, 1]$, one can define an approximation F_σ of F satisfying natural conditions (see [2,53]). Suppose G is another set-valued mapping from $[0, 1]$ into E which is an approximation to F at 0 in the sense that F and G are tangent at 0. Is G_σ an approximation of F_σ at 0?

Let us complete the indication of these tracks with a general observation. One of the most remarkable advances of nonsmooth analysis and variational analysis during the last decades has been to provide a unified treatment for the study of functions, sets, mappings and multimappings (see for example [3,45]). The passages from one object to another one have been particularly fruitful for what concern convergence questions, conditioning properties, error bounds and infinitesimal analysis. One may expect that such a stream also is of interest for approximation questions.

2. Preliminaries

In the sequel A, B and C are subsets of n.v.s. X, and a is a point in $\mathrm{cl}\, A \cap \mathrm{cl}\, B$, where $\mathrm{cl}\, C$ denotes the closure of the set C. We will compare several notions of tangency at a for A and B. Let us describe our notations.

The open unit ball of X is denoted by U_X i.e. $U_X := \{x \in X : \|x\| < 1\}$. For $x \in X$ and $r \in \mathbb{P} := (0, +\infty)$, we set: $A_r = A \cap rU_X$, and for a subset B of X

$$d_B(x) = d(x, B) = \inf \{d(x, y) : y \in B\}$$

represents the distance from x to the set B. By convention we set $\inf \emptyset = \infty$ and $\sup \emptyset = 0$. The *Hausdorff–Pompeiu excess* of the set A over the set B is given by

$$e(A, B) = \sup \{d(x, B) : x \in A\}.$$

The *Hausdorff–Pompeiu* distance between the sets A and B is given by

$$d(A, B) = \max(e(A, B), e(B, A)).$$

And finally, for $r \in \mathbb{P} := (0, +\infty)$, we set

$$e_r(A, B) = e(A_r, B), \quad d_r(A, B) = \max(e_r(A, B), e_r(B, A)).$$

Given $\varepsilon > 0$ the *ε-conical neighborhood* of the subset B of X is the set

$$V_\varepsilon(B) = \{x \in X : d(x, B) < \varepsilon\|x\|\} \cup \{0\}.$$

If the set C is a cone, then $V_\varepsilon(C)$ is a cone and then our definition coincides with the notion introduced or used in [15–17,21].

Before treating the question of approximations of sets, let us recall a familiar concept of tangency for mappings. Given n.v.s. X and Y, a subset A of X, $a \in \mathrm{cl}\, A$, two mappings $g, h : A \to Y$ such $g(a) = h(a)$ are said to be (Fréchet)-*tangent at a* with respect to (w.r.t.) A if

$$\lim_{x(\in A\setminus\{a\})\to a} \frac{\|g(x) - h(x)\|}{\|x - a\|} = 0.$$

The following definition has been introduced by Mignot [24] in connection with variational inequalities and used by several authors including Dontchev and Hager [14], Pang [27], Robinson [43] for similar purposes or in view of an implicit function theorem (see also [28]).

Definition 1. The mapping $f : A \to Y$ is said to be *B-differentiable* (or boundedly differentiable) at a w.r.t. A if g given by

$$g(x) = f(x + a) - f(a)$$

is Fréchet-tangent at 0 w.r.t. $A - a$ to some positively homogenous mapping denoted by $f'(a)(.)$.

We observe that if $A = X$ and if $f'(a)(.)$ is linear and continuous, then f is Fréchet-differentiable at a. The just quoted papers have shown that a number of useful properties of Fréchet-differentiability are preserved when one drops linearity.

3. Approximations of sets

In the present section we display various known notions of approximation for sets, introduce some others and compare them. In [22,23,28] notions of approximations for sets are introduced which are useful for necessary and sufficient optimality conditions of first and second order, for mathematical programming problems in an infinite-dimensional space. The following definition is an attempt to encompass these notions.

Definition 2 (*Maurer [22], Maurer and Zowe [23] and Penot [28]*). The subset A of X is said to be *M–Z approximated* by the subset B at $a \in \mathrm{cl}(A)$ if there exists a mapping $h : A \to B$ which is tangent to the identity mapping I_X of X at a w.r.t. A:

$$\lim_{x(\in A\setminus\{a\})\to a} \frac{\|h(x) - x\|}{\|x - a\|} = 0.$$

We say that the sets A and B are *M–Z tangent* (or tangent in the sense of Maurer–Zowe) at a, if A is approximated by B at a and B is approximated by A at a.

The following notion has been introduced in [46,47] for studying the differentiability of the metric projection in a normed space and the differentiability of the multifunctions.

We introduce here a one-sided version of this concept which does not suppose that B is a translate of a cone; however, that case is the main case of interest.

Definition 3. The subset A of X is said to be *S-approximated* by B at $a \in \mathrm{cl}(A)$ if

$$\lim_{x(\in A\setminus\{a\})\to a} \frac{d(x,B)}{\|x-a\|} = 0.$$

The sets A and B are *S-tangent* (or tangent in the sense of Shapiro) at a if A is *S-approximated* by B at a and B is *S-approximated* by A at a.

The following definition seems to be closely related to the preceding one. Here for a real number r we set: $r_+ = \max(r, 0)$.

Definition 4. The subset A of X is said to be boundedly approximated (in short *B-approximated)* by B at $a \in \mathrm{cl}(A)$ if

$$\lim_{x(\in X\setminus\{a\})\to a} \frac{(d(x,B) - d(x,A))_+}{\|x-a\|} = 0.$$

The sets A and B are said to be *B-tangent* (or boundedly tangent) at a if A is *B-approximated* by B at a and B is *B-approximated* by A at a.

Clearly, the sets A and B are B-tangent at a if only if A and B are tangent at a in the sense of Auslender and Cominetti [4], or A–C tangent, i.e.

$$\lim_{x(\in X\setminus\{a\})\to a} \frac{|\, d(x,A) - d(x,B)\,|}{\|x-a\|} = 0. \tag{3}$$

The following notion has been introduced by Robinson [39] for the study of the stability of some mathematical programming problems. We give here a one-sided version of his concept.

Definition 5. We say that the set A is *R-approximated* by B at $a \in \mathrm{cl}(A)$ if

$$\lim_{r\to 0_+} r^{-1} e_r(A-a, B-a) = 0.$$

The sets A and B are said to be *R-tangent* at a (or tangent in the sense of Robinson) if

$$\lim_{r\to 0_+} r^{-1} d_r(A-a, B-a) = 0$$

i.e. A is *R-approximated* by B at a and B is *R-approximated* by A at a.

Although the following notion did not explicitly appear in the papers [15–17] it is clearly in the spirit of these papers.

Definition 6. We say that the set A is *F-approximated* by B at $a \in \mathrm{cl}(A)$ if for each $\varepsilon > 0$ there exists $\eta > 0$ such that

$$(A-a)_\eta \subset V_\varepsilon(B-a).$$

The sets A and B are said to be *F-tangent* (or tangent in the sense of Fabian) at a if A (resp., B) is *F-approximated* by B (resp., A) at a.

If $B = a + C$, where the set C is a closed cone, then we obtain the concept of Fréchet cone introduced in [15–17].

The last notion we present is a one-sided version of a concept suggested by Demyanov (personal communication).

Definition 7. We say that the set A is D-approximated by B at $a \in \mathrm{cl}(A)$ if

$$\lim_{r\to 0_+} r^{-1} e((A-a)_r, (B-a)_r) = 0.$$

The sets A and B are said to be D-tangent (or tangent in the sense of Demyanov) at a if

$$\lim_{r\to 0_+} r^{-1} d((A-a)_r, (B-a)_r) = 0,$$

i.e. if A is D-approximated by B at a and B is D-approximated by A at a.

Let us start with a comparison between the notions of R-approximation and D-approximation and consequently between the notions of tangency in the sense of Demyanov and in the sense of Robinson.

Proposition 8. *Let A and B be two nonempty sets and let $a \in \mathrm{cl}\, A \cap \mathrm{cl}\, B$. Then the following implication holds*: (i) $\Longrightarrow$ (ii)

(i) *A is D-approximated by B at a*;
(ii) *A is R-approximated by B at a*.

Proof. Without loss of generality we suppose that $a = 0$. As $B_r := B \cap rU_X \subset B$, for each $x \in A_r$ we have $d(x, B) \leqslant d(x, B_r)$, hence, taking the supremum on x in A_r,

$$r^{-1} e(A_r, B) \leqslant r^{-1} e(A_r, B_r)$$

and we obtain the implication announced above. $\square$

Example. In general the implication (ii) $\Longrightarrow$ (i) does not hold. Indeed, take $X = \mathbb{R}$ and for some decreasing sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ of $(0, \frac{1}{4})$ with $\lim_{n\to\infty} \varepsilon_n = 0$ set

$$A = \left\{2^{-n} : n \in \mathbb{N}\right\} \cup \{0\} \quad \text{and} \quad B = \left\{2^{-n}(1+\varepsilon_n) : n \in \mathbb{N}\right\} \cup \{0\}.$$

Let us show that A is R-approximated by B at 0, but that A is not D-approximated by B at 0.

For $r_n = 2^{-n}$, $r \in (r_n, r_{n-1}]$ we have

$$e_r(A, B) = \sup_{p \geqslant n} d(2^{-p}, B) = 2^{-n}\varepsilon_n = r_n \varepsilon_n$$

so that, for $r \in (r_n, r_{n-1}]$, we get

$$r^{-1} e_r(A, B) \leqslant r_n^{-1} e_r(A, B) \leqslant \varepsilon_n.$$

Hence A is R-approximated by B at 0.

However, for $r \in (r_n, r_n(1+\varepsilon_n))$ we have

$$e(A_r, B_r) = \sup_{p \geqslant n} d(2^{-p}, B_r) = 2^{-n-1}(1-\varepsilon_{n+1}) > \tfrac{1}{2} r(1-\varepsilon_{n+1}),$$

so that

$$r^{-1} e(A_r, B_r) > \tfrac{1}{2}(1-\varepsilon_{n+1}) > \tfrac{3}{8}.$$

Therefore $\liminf_{r \to 0_+} r^{-1} e(A_r, B_r) > 0$, and A is not D-approximated by B at 0.

Let us give conditions ensuring the equivalence between R-approximations and D-approximations. These conditions require the following lemma, in which a set C is said to be *starshaped* at 0 if $tc \in C$ for each $c \in C$ and each $t \in [0, 1]$.

Lemma 9. *Let A and B be two closed subsets of X such that B is starshaped at* $0 \in A \cap B$. *Then, for any real number* $r > 0$, *one has*

$$e(A_r, B) \leqslant e(A_r, B_r) \leqslant 2e(A_r, B).$$

Proof. The first inequality follows from the inclusion $B_r \subset B$. For proving the second one, we will show that for each real number $s > e(A_r, B)$ we have $e(A_r, B_r) \leqslant 2s$.

For each $x \in A_r$, there exists $y \in B$ such that $\|x - y\| < s$, so that $y \in (r+s)U_X$. As B is starshaped at 0, setting $z := r(r+s)^{-1} y$, we have $z \in B_r$, $\|z - y\| < s$ and

$$\|x - z\| \leqslant \|x - y\| + \|z - y\| < 2s,$$

so that we get $d(x, B_r) \leqslant 2s$ for each $x \in A_r$, and consequently $e(A_r, B_r) \leqslant 2s$. □

The following corollary is an immediate consequence of the preceding lemma. Here, a set C is said to be *starshaped at* c_0 if $c_0 + t(c - c_0) \in C$ for each $c \in C$ and each $t \in [0, 1]$.

Corollary 10. *If B is starshaped at a, then the following assertions are equivalent*:

(i) *A is D-approximated by B at a*;
(ii) *A is R-approximated by B at a*.

Corollary 11. *Suppose that A and B are starshaped at a. Then the following assertions are equivalent*:

(i) *A and B are D-tangent at a*;
(ii) *A and B are R-tangent at a*.

Now we can state the main result of this section which completes the relationships disclosed in [4] between assertions (2)–(4) of the statement.

Theorem 12. *Let A and B be two nonempty subsets of X and let* $a \in \operatorname{cl} A \cap \operatorname{cl} B$. *Then the following assertions are equivalent*:

(1) *A is MZ-approximated by B at a*;
(2) *A is S-approximated by B at a*;
(3) *A is B-approximated by B at a*;
(4) *A is R-approximated by B at a*;
(5) *A is F-approximated by B at a*.

Under one of these assumptions we say that the set A is approximated by the set B at a (*or the set B is an approximation to the set A at a*).

Proof. The implication (1) $\Longrightarrow$ (2) is immediate: if $h : A \to B$ is Fréchet-tangent to the identity mapping I_X of X w.r.t. A at a, then for $x \in A$ we have

$$d(x, B) \leqslant d(x, h(x)) = o(\|x - a\|).$$

Conversely, let us consider the multifunction $H : X \rightrightarrows X$ given by

$$H(a) = \{a\}, \quad H(x) = \left\{ y \in B : \|x - y\| \leqslant d(x, B) + \|x - a\|^2 \right\} \quad \text{for } x \neq a.$$

It has nonempty values, so that we can pick a selection h of H. Then, for $x \in A$ we have

$$\|h(x) - x\| \leqslant d(x, B) + \|x - a\|^2 = o(\|x - a\|).$$

In order to deal with the other equivalences, let us introduce the functions σ, β, ρ given for $r > 0$ by

$$\sigma(r) := \sup \left\{ r^{-1} d(x, B) : x \in A, \|x - a\| = r \right\},$$
$$\beta(r) := \sup \left\{ r^{-1} (d(x, B) - d(x, A))_+ : x \in X, \|x - a\| = r \right\},$$
$$\rho(r) := r^{-1} e_r(A - a, B - a)$$

and let us note the following immediate observation about the (upper) nondecreasing hull θ_N of a function $\theta : \mathbb{R} \longrightarrow \overline{\mathbb{R}}_+ := [0, +\infty]$ which is given by

$$\theta_N(t) = \sup \left\{ \theta(s) : s \in [0, t] \right\}.$$

Obviously, θ_N is the least nondecreasing function majorizing θ and we have $\lim_{r \to 0_+} \theta(r) = 0$ if, and only if, $\lim_{r \to 0_+} \theta_N(r) = 0$.

The implications (3) $\Longrightarrow$ (2), (4) $\Longrightarrow$ (2) are consequences of the following obvious inequalities:

$$\beta \geqslant \sigma, \quad \rho \geqslant \sigma.$$

Since for $x \in A$, $x \neq a$ such that $s := \|x - a\| < r$ we have

$$r^{-1} d(x, B) \leqslant s^{-1} d(x, B) \leqslant \sigma(s) \leqslant \sigma_N(s),$$

taking the supremum on $A \cap (a + rU_X)$, we obtain

$$\rho(r) \leqslant \sigma_N(r)$$

which shows that (2) $\Longrightarrow$ (4).

Let us show that (4) $\Longrightarrow$ (3) by proving that $\beta(r)\leqslant 2\rho(2r)$.

We observe that for $x \in X$ with $r := \|x-a\| > 0$ we have

$$d(x, A) = d(x-a, (A-a)_{2r})$$

since otherwise we could find $u \in A$ with $\|u-a\| \geqslant 2r$ such that $\|x-u\| < d(x-a, (A-a)_{2r})$ and as $d(x-a, (A-a)_{2r}) \leqslant \|x-a\| = r$ we would have $\|u-a\| \leqslant \|u-x\| + \|x-a\| < 2r$, a contradiction.

Now, as $d(., B)$ is Lipschitzian with rate 1, for any $x \in X$ with $r := \|x-a\|$ and for $w \in A \cap (a + 2rU_X)$ we have $d(x, B) - \|x-w\| \leqslant d(w, B)$, hence, taking the supremum on w, we obtain

$$d(x, B) - d(x-a, (A-a)_{2r}) \leqslant e_{2r}(A-a, B-a).$$

By what precedes we get

$$d(x, B) - d(x, A) \leqslant e_{2r}(A-a, B-a);$$

it follows that $\beta(r)\leqslant 2\rho(2r)$.

We finish the proof by showing that (2) $\Longleftrightarrow$ (5). Now (2) holds if and only if for each $\varepsilon > 0$ there exists $\eta > 0$ such that for each $x \in A\backslash\{a\}$ with $\|x-a\| < \eta$ one has

$$d(x, B) < \varepsilon \|x-a\|$$

or, for each $z = x - a \in (A-a) \cap \eta U_X = (A-a)_\eta$ with $z \neq 0$, one has

$$d(z, B-a) < \varepsilon \|z\|$$

or

$$(A-a)_\eta \subset V_\varepsilon(B-a);$$

thus, (2) holds if and only if (5) holds. □

This result implies a symmetric version.

Corollary 13. *Given $a \in \operatorname{cl} A \cap \operatorname{cl} B$, the following assertions are equivalent*:

(1) *A and B are* (*M–Z*) *tangent at a*;
(2) *A and B are S-tangent at a*;
(3) *A and B are B-tangent at a*;
(4) *A and B are R-tangent at a*;
(5) *A and B are F-tangent at a*.

When one of these assumptions is satisfied we say that the sets A and B are tangent at a.

Example. Let $f, g : X \to Y$ be two mappings between n.v.s., with graphs F, G, respectively. If f and g are tangent at $\overline{x}$, then F and G are tangent at $a := (\overline{x}, f(\overline{x})) = (\overline{x}, g(\overline{x}))$: the map $h : (x, y) \to (x, y - f(x) + g(x))$ and its inverse $h^{-1} : (u, v) \to (u, v - g(u) + f(u))$ are tangent to the identity mapping at a w.r.t. F and G, respectively. Conversely, if F and G

are tangent at a and if f and g are Lipschitzian, then f and g are tangent at $\overline{x}$. More generally, if $h : F \to G$ is tangent to I_X at a w.r.t. F and if k is its first component, for $(x, y) \in F$ one has $h(x, y) = (k(x), g(k(x)))$ and, if l is the Lipschitz rate of g,

$$\begin{aligned} \|f(x) - g(x)\| &\leqslant \|f(x) - g(k(x))\| + \|g(k(x)) - g(x)\| \\ &\leqslant \|(x, f(x)) - h(x, f(x))\| + l\|k(x) - x\| \\ &\leqslant (1 + l)\|(x, f(x)) - h(x, f(x))\|. \end{aligned}$$

One cannot drop the Lipschitz assumption, as the example of the functions $f, g : \mathbb{R} \to \mathbb{R}$ given by $f(x) := \sqrt{|x|}$, $g(x) = 2\sqrt{|x|}$ show.

Let us record for future use the following characterization of B-differentiability which is a specialization of the preceding example. It encompasses previous results of Durdil [15,16].

Proposition 14. *Let $f : X \to Y$ be Lipschitzian around $\overline{x}$, and let $h : X \to Y$ be a positively homogeneous Lipschitzian mapping. Then the set $A :=$ Graph(f) and the set $B := (\overline{x}, f(\overline{x}))+$Graph$(h)$ are tangent at $a := (\overline{x}, f(\overline{x}))$ if, and only if, f is B-differentiable at $\overline{x}$ with B-derivative $f'(\overline{x}) = h$.*

If h is linear and continuous, then f is Fréchet-differentiable at $\overline{x}$ if A and B are tangent.

To conclude this section, we compare the previous different notions of tangency with the notion of directional tangency.

Definition 15 (*Auslender and Cominetti [4]*). The subsets A and B of X are said to be *tangent at $a \in$ cl $A \cap$ cl B in the direction $v \in X$* if

$$\lim_{t \to 0+} t^{-1} \mid d(a + tv, A) - d(a + tv, B) \mid = 0.$$

They are said to be *directionally tangent* at a if there are tangent in any direction $v \in X$.

One has the following relationship between these two concepts; we present the proof for completeness.

Proposition 16 (*Auslender and Cominetti [4]*). *Assertion* (i) *below implies assertion* (ii). *If the n.v.s. X is finite-dimensional these assertions are equivalent*

(i) *A and B are tangent at a*;
(ii) *A and B are directionally tangent at a*.

Proof. (i) $\Longrightarrow$ (ii) is immediate by considering the notion of B-tangency.

(ii) $\Longrightarrow$ (i) Suppose that relation (3) does not hold. Then there exist $\varepsilon > 0$ and a sequence $(x_n)_{n \in \mathbb{N}}$ converging to a, such that:

$$\mid d(x_n, A) - d(x_n, B) \mid > \varepsilon \|x_n - a\| \quad \text{for any } n \in \mathbb{N}.$$

We set $t_n := \|x_n - a\| > 0$, since if $x_n = a$ the inequality does not hold. Let $v_n := t_n^{-1}(x_n - a)$. Then the preceding relation becomes:

$$| d(a + t_n v_n, A) - d(a + t_n v_n, B) | > \varepsilon t_n \quad \text{for any } n \in \mathbb{N}.$$

Since X is finite dimensional, we may assume that the sequence $(v_n)_{n\in\mathbb{N}}$ converges to some unit vector v. Since the distance function is Lipschitzian with rate 1, we may replace v_n by v in the last relation and ε by a smaller ε'. This is a contradiction with the definition of directionally tangent sets. □

In general, if X is an infinite-dimensional space, the implication (ii) $\Longrightarrow$ (i) does not hold. To see that, let A be the graph of a locally Lipschitzian mapping $f : X \to Y$ which is not B-differentiable at $\overline{x}$ but is directionally differentiable at $\overline{x}$ i.e.

$$f'(\overline{x}, u) := \lim_{t\to 0_+} t^{-1}(f(\overline{x} + tu) - f(\overline{x}))$$

exists for all $u \in X$. Let B be the graph of $g : x \longmapsto f(\overline{x}) + f'(\overline{x})(x - \overline{x})$ which is easily seen to be Lipschitzian. Then Proposition 14 shows that A and B are not B-tangent at $a := (\overline{x}, f(\overline{x}))$. However, it is easy to show that A and $a + B$ are directionally tangent at a. More precisely, one has the following lemma when the Lipschitz rate of f around $\overline{x}$ is not greater than 1.

Lemma 17 (*Agadi [1]*). *Under the preceding assumptions, for any $u \in X$, $v \in Y$, one has*:

$$d'_A((\overline{x}, f(\overline{x})), (u, v)) = d((u, v), B) = \left\| f'(\overline{x}, u) - v \right\|$$

where d'_A is the directional derivative of $d_A := d(., A)$ at $(\overline{x}, f(\overline{x}))$ in the direction (u, v).

4. Approximation cones and tangent cones

In this section, we consider the case in which the subset A of X has an approximation at a which is a translated cone. We extend finite-dimensional results of Shapiro [46,47] and we study some consequences of the existence of an approximation cone. We first recall this notion introduced by Shapiro [46,47] for the study of the directional derivative of the metric projection.

Definition 18. A closed cone C of X is said to be an approximation cone to the set A at a, if A and $a + C$ are tangent at a.

Now let us recall some classical notions of tangent cones. From now on σ denotes the weak topology of X (an arbitrary topology weaker than the norm topology could also be considered).

Definition 19. The tangent cone (or contingent cone or Bouligand cone) to the set A at a is the set $T(A, a)$ of vectors $v \in X$ such that there exist sequences $(t_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ in

$\mathbb{P} := (0, \infty)$ and X, respectively, such that $\lim_{n\to\infty} t_n = 0$, $\lim_{n\to\infty} v_n = v$, and $a + t_n v_n \in A$ for each $n \in \mathbb{N}$.

The weak tangent cone $T^\sigma(A, a)$ is the set of vectors v such that there exist bounded nets $(t_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ in $\mathbb{P}$ and X, respectively, such that $\lim_n t_n = 0$, $\lim_n v_n = v$ for σ and $a + t_n v_n \in A$ for each $n \in N$.

The following variant corresponds to a classical notion too. It has an attractive kinematic interpretation in terms of velocities of trajectories in A.

Definition 20. The incident cone (or adjacent cone or intermediate cone) to the set A at a is the set $T^i(A, a)$ of vectors $v \in X$ such that for any sequence $(t_n)_{n\in\mathbb{N}}$ of $\mathbb{P}$ with limit 0 there exists a sequence $(v_n)_{n\in\mathbb{N}}$ of X with limit v such that $a + t_n v_n \in A$ for each $n \in \mathbb{N}$.

It is well-known that $T^i(A, a)$ and $T(A, a)$ are closed cones with $T^i(A, a) \subset T(A, a)$; furthermore

$$T^i(A, a) = \left\{ v \in X : \lim_{t\to 0_+} t^{-1} d(a + tv, A) = 0 \right\}, \tag{4}$$

$$T(A, a) = \left\{ v \in X : \liminf_{t\to 0_+} t^{-1} d(a + tv, A) = 0 \right\}. \tag{5}$$

Definition 21 *(Aubin and Frankowska [3], Auslender and Cominetti [4], Rockafellar [44])* The set A is said to be proto-differentiable (resp., pseudo-differentiable) at a if the incident cone to A at a and the tangent cone (resp., the weak tangent cone) coincide, i.e.

$$T^i(A, a) = T(A, a) \quad (\text{resp., } T^i(A, a) = T^\sigma(A, a)).$$

In the first case one also says that A is derivable at a and one writes $A'(a)$ for the tangent cone to A at a. If A is a closed convex set, then A is proto-differentiable and in fact pseudo-differentiable at any $a \in A$ and $A'(a)$ is the closure $\mathrm{cl}(\mathbb{R}_+(A - a))$ of $\mathbb{R}_+(A - a)$.

The next proposition describes a consequence of an approximation property for proto-differentiability.

Proposition 21. *Let A and B be two subsets of X. If A is approximated by B at $a \in A \cap B$, then the following inclusions hold*:

(i) $T^i(A, a) \subset T^i(B, a)$;
(ii) $T(A, a) \subset T(B, a)$;
(iii) $T^\sigma(A, a) \subset T^\sigma(B, a)$.

Proof. Immediate by using the Maurer–Zowe approach, for instance. □

The following corollary ensues; in view of relations (4), (5) its conclusions (i) and (ii) are valid when A and B are directionally tangent at a.

Corollary 22. *If the sets A and B are tangent at $a \in A \cap B$, then*:

(i) $T^i(A, a) = T^i(B, a)$;

(ii) $T(A, a) = T(B, a)$;
(iii) $T^{\sigma}(A, a) = T^{\sigma}(B, a)$.

Thus, if the sets A and B are tangent at $a \in A \cap B$, then A is proto-differentiable (resp., pseudo-differentiable) at a if, and only if, B is proto-differentiable (resp., pseudo-differentiable) at a.

A necessary condition for the existence of an approximation cone can be derived from the fact that the tangent cone to a closed cone at the origin is the cone itself.

Corollary 23. *If the set A is approximated by $a + C$, with C a σ-closed cone, then one has $T^{\sigma}(A, a) \subset C$. If furthermore, the cone C is an approximation cone to A at a, then A is pseudo-differentiable at a*:

$$T^{i}(A, a) = T^{\sigma}(A, a) = C.$$

It follows from this corollary that if the set A is approximated by $a + T(A, a)$ at a, then this approximation is optimal in terms of inclusion. It is not always the case that a set has an approximation cone, as follows:

Example. Let A be the graph of Lipschitzian mapping f between two infinite-dimensional spaces X and Y, with rate $k \leqslant 1$. We suppose that the directional derivative $f'(\overline{x}, .)$ exists, with

$$f'(\overline{x}, v) = \lim_{t \to 0_+} t^{-1}(f(\overline{x} + tv) - f(\overline{x})),$$

but that f is not B-differentiable at $\overline{x}$. Let $a := (\overline{x}, f(\overline{x}))$. It is easy to see that (see Lemma 17 and [1]):

$$A'(a) := T^{i}(A, a) = T(A, a) = \text{graph}(f'(\overline{x}, .))$$

i.e. the set A is proto-differentiable at a. However, we have seen that $A'(a)$ is not an approximation cone to the set A at a.

However, in a finite-dimensional space, we have the following positive result.

Proposition 24. *Suppose that X is finite dimensional. Then the set A is approximated by $a + T(A, a)$ at a.*

This easy result is a consequence of a more general fact requiring the following definition which is obviously satisfied by any set contained in a finite-dimensional space and by any finite-dimensional submanifold of an arbitrary n.v.s.

Definition 25 (*Penot [29]*). The set A is said to be tangentially compact at $a \in \text{cl}\, A$, if for any sequence $(a_n)_n$ of A converging to a with $a_n \neq a$ the sequence $(\|a_n - a\|^{-1} (a_n - a))_n$ has a converging subsequence.

The following example shows that a set A may be tangentially compact at a without being locally compact at a.

Example. Let W be a normed space and let $A \subset X := W \times \mathbb{R}$ be the epigraph of a function $f : W \to \mathbb{R}$ such that $f(0) = 0$, $f(w)/\|w\| \to +\infty$ as $w \to 0$, $w \neq 0$. Then A is tangentially compact at $a := (0, 0)$.

Theorem 26. *If the set A is tangentially compact at a, then A is approximated by $a+T(A,a)$ at a and $T(A,a)$ is locally compact at 0. Conversely, if A is approximated by $a+T(A,a)$ at a and $T(A,a)$ is locally compact at 0 then A is tangentially compact at a.*

Proof. Suppose that A is not S-approximated by $a + T(A,a)$ at a. Then there exists $\varepsilon > 0$ and a sequence (x_n) of A such that $(x_n) \to a$ and

$$d(x_n, a + T(A,a)) > \varepsilon \|x_n - a\| \quad \text{for all } n \in \mathbb{N}.$$

We set $t_n := \|x_n - a\|$ (> 0) and $u_n := t_n^{-1}(x_n - a)$. Then $\|u_n\| = 1$ and the preceding relation becomes

$$d(u_n, T(A,a)) > \varepsilon \quad \text{for all } n \in \mathbb{N}.$$

As the set A is tangentially compact at a, taking a subsequence if necessary, one can find $u \in X$ such that $(u_n) \to u$ as $n \to \infty$. Then $u \in T(A,a)$, a contradiction with the relation $d(u, T(A,a)) \geqslant \varepsilon$ which stems from the Lipschitz property of $d(\cdot, T(A,a))$.

In order to prove that $T(A,a)$ is locally compact at 0 it suffices to show that any sequence (u_n) of unit vectors of $T(A,a)$ has a converging subsequence. Given a sequence $(\varepsilon_n) \to 0$ in $(0,1)$, we can find $a_n \in A$ and $t_n \in (0, \varepsilon_n)$ such that $z_n := u_n - t_n^{-1}(a_n - a)$ satisfies $\|z_n\| < \varepsilon_n$ for each $n \in \mathbb{N}$. Then, $t_n^{-1}\|a_n - a\| < 2$ and a subsequence of $\left(t_n^{-1}\|a_n - a\|, \|a_n - a\|^{-1}(a_n - a)\right)$ converges to some limit $(q, u) \in [0,2] \times T(A,a)$. The corresponding subsequence of (u_n) then converges to $qu \in T(A,a)$.

Now let us prove the converse. Let (a_n) be a sequence of $A \backslash \{a\}$ converging to a. Let $h : A \to a + T(A,a)$ be tangent to I_X on A, and let $k(x) := h(x) - a$. Then $k(a_n) = a_n - a + r_n z_n$ with $r_n := \|a_n - a\|$, $(z_n) \to 0$. Then $(r_n^{-1}\|k(a_n)\|) \to 1$ and since $T(A,a) \cap \operatorname{cl} U_X$ is compact, $(\|k(a_n)\|^{-1} k(a_n))$ has a converging subsequence, $\left(r_n^{-1}(a_n - a)\right) = \left(r_n^{-1}k(a_n) - z_n\right)$ has a converging subsequence too. □

A pleasant consequence of the existence of a convex approximation cone is the fact that a number of tangent cones coincide. This result can be seen as a necessary condition for the existence of a convex approximation cone. The tangent cones we consider are defined as follows.

Definition 27 (*Jofre and Penot [19], Penot and Terpolilli [35], Treiman [50,51]*). Given $a \in A$ the b-tangent cone $T^b(A,a)$ is the set of $v \in X$ such that for any sequence $((r_n, a_n))_n$ in $\mathbb{P} \times A$ with limit $(0, a)$ such that $(r_n^{-1}(a_n - a))_n$ is bounded, there exists a sequence $(v_n)_n \longrightarrow v$ such that $a_n + r_n v_n \in A$ for all n. The set of v for which this property holds whenever $(r_n^{-1}(a_n - a))_n$ converges (resp., converges to some element of $T^i(A,a)$) is denoted by $T^p(A,a)$ (resp., $T^q(A,a)$).

It is known that $T^b(A,a)$, $T^p(A,a)$, $T^q(A,a)$ are closed convex cones satisfying the following obvious inclusions:

$$T^b(A,a) \subset T^p(A,a) \subset T^q(A,a) \subset T^i(A,a) \subset T(A,a). \tag{6}$$

Theorem 28. *If A has a convex approximation cone C at $a \in A$ then*

$$T^b(A,a) = T^p(A,a) = T^q(A,a) = T^i(A,a) = T(A,a) = C.$$

Proof. If C is an approximation cone of the set A at a, then one has $T(A,a) \subset C$. In view of the inclusions (6) it suffices to show that $C \subset T^b(A,a)$. Without loss of generality we suppose that $a = 0$.

Let $v \in C$ and let $((r_n, a_n))_n$ be a sequence of $\mathbb{P} \times A$ with limit $(0,0)$ such that the sequence $(r_n^{-1}a_n)_n$ is bounded. By assumption, we can find a sequence (w_n) with limit 0 in X such that $c_n := a_n + r_n w_n \in C$ for each n and a sequence (v_n) with limit v in X such that $a'_n := c_n + r_n v_n \in A$ for each n. Then

$$r_n^{-1}(a'_n - a_n) = v_n + w_n \to v$$

and we have shown that $v \in T^b(A,a)$. □

Let us observe that the existence of a convex approximation cone does not suffice to ensure that the circa-tangent cone (or Clarke tangent cone) $T^{\uparrow}(A,a) := \{v \in X : \lim_{t\to 0_+, a'(\in A)\to a}(1/t)d(a'+tv, A) = 0\}$ coincides with the previous tangent cones.

Example. Let f be the even function of one real variable such that for a decreasing sequence (r_n) with limit 0 and such that $r_n^{-2}(r_n - r_{n+1}) \to 0$ one has $f(r_n) = r_n^2$ for even n and 0 for n odd, f being affine on each interval $[r_{n+1}, r_n]$. Then f is differentiable at 0 and the epigraph A of f is approximated by the upper plane C at $(0,0)$ but the Clarke tangent cone to A at $(0,0)$ is $\{0\} \times \mathbb{R}_+$. Similar assertions hold for the even function f given by $f(0) = 0$, $f(x) = x^2 \sin(1/x)^2$ for $x \neq 0$.

Proposition 29. *For a set A, a closed cone C and $a \in A$, among the following assertions one has the implications* (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (iv):

(i) *C is an approximation cone to A at a*;
(ii) *$d(.,A)$ is B-differentiable at a*;
(iii) *$d(.,A)$ is directionally differentiable at a*;
(iv) *A is proto-differentiable at a.*

Furthermore, when (i) *holds one has*

$$A'(a) = C = \left\{v \in X : d'_A(a,v) = 0\right\} \quad \text{and} \quad d'_A(a,.) = d(.,C).$$

Proof. (i) $\Longrightarrow$ (ii) If C is an approximation cone to A at a then A and $a + C$ are B-tangent at a, so that we have

$$\lim_{x(\neq 0)\to 0} \frac{|\, d(a+x, A) - d(x, C)\,|}{\|x\|} = 0.$$

As C is a cone, the distance function $d(., C)$ is positively homogeneous, so that this relation shows that d_A is B-differentiable at a with B-derivative $d'_A(a, \cdot)$ given by

$$d'_A(a, v) = d(v, C), \quad v \in X.$$

(ii) $\Longrightarrow$ (iii) If a Lipschitzian map is B-differentiable, then it has a directional derivative in any direction [48].

(iii) $\Longrightarrow$ (iv) Let $v \in T(A, a)$, then

$$0 = \liminf_{t \to 0_+} t^{-1} d_A(a + tv) = d'_A(a, v) = \lim_{t \to 0_+} t^{-1} d(a + tv, A)$$

so we have $v \in T^i(A, a)$. $\square$

Corollary 30. *Let X be a Banach space such that for any subset A of X and any $a \in A$ the set A has an approximation cone C. Then X is finite dimensional.*

Proof. Under the assumption, for any subset A of X and any $a \in A$ the set A has an approximation cone $C = A'(a)$ and $d'_A(a, \cdot) = d_{A'(a)}(\cdot)$. Applying [9] Theorem 2, we get that X is finite dimensional. In fact, $X = \{0\}$ since in any one-dimensional space one can find a subset which is not proto-differentiable. $\square$

We have seen that if an approximation cone to the set A at a exists, then it is unique and it is the cone $C := T^i(A, a) = T(A, a)$. As a step to a converse, let us present some conditions ensuring that $a + T^i(A, a)$ and $a + T(A, a)$ are approximated by A at a.

Proposition 31. *Let $a \in A$. If the distance function $d(., A)$ is B-differentiable at a, then $a + T(A, a)$ (and a fortiori $a + T^i(A, a)$) is approximated by A at a (A is an approximation to the set $a + T(A, a)$ at a).*

Proof. When $d(., A)$ is B-differentiable at a, the set A is proto-differentiable at a and

$$T^i(A, a) = T(A, a) = \left\{ v \in X : d'_A(a, v) = 0 \right\}.$$

As $d(x, A) = d'_A(a, x - a) + r(\|x - a\|)$, where $\lim_{x (\in A \setminus \{a\}) \to a} \frac{r(\|x - a\|)}{\|x - a\|} = 0$, for $x \in a + T(A, a)$, we have $d(x, A) = r(\|x - a\|)$, i.e. $a + T(A, a)$ is S-approximated by A at a. $\square$

Proposition 32. *Suppose that $T^i(A, a)$ is locally compact at 0. Then $a + T^i(A, a)$ is approximated by A at a (A is an approximation to the set $a + T^i(A, a)$ at a).*

Proof. As $C := T^i(A, a)$ is closed cone, C is locally compact at 0 if and only if there exists $r > 0$ such that $C \cap \text{cl}\,(rU_X)$ is compact, if and only if $C \cap S_X$ is compact, where $S_X = \{x \in X : \|x\| = 1\}$. Suppose that $a + C$ is not S-approximated by A at a: there exists $\varepsilon > 0$ and $u_n \in C$ such that $(u_n) \to 0$ and

$$d(a + u_n, A) > \varepsilon \|u_n\| \quad \text{for all } n \in \mathbb{N}.$$

Let $t_n = \|u_n\|$ and $v_n = t_n^{-1}u_n$. We have $v_n \in C \cap S_X$. As $C \cap S_X$ is compact, there is a subsequence $(v_k)_k$ of the sequence $(v_n)_n$ and a vector $v \in C$, such that $(v_k)_k \longrightarrow v$. Then we get a contradiction with the relation

$$t_n^{-1}d(a + t_n v, A) \geqslant \varepsilon - \|v_n - v\| \quad \text{for all } n \in \mathbb{N}. \qquad \square$$

The following statement follows from the local compactness of the cone $T(A, a)$ at 0 (Theorem 26) and from the fact that $T^i(A, a)$ is a closed subset of $T(A, a)$.

Lemma 33. *If the set A is tangentially compact at a, then $T^i(A, a)$ is locally compact at* 0.

The following corollaries are immediate consequences of the preceding results.

Corollary 34. *If the set A is tangentially compact at a and if A is proto-differentiable at a, then $C := T^i(A, a) = T(A, a)$ is an approximation cone to the set A at a.*

Corollary 35. *If the set A is tangentially compact at a and if the distance function $d_A(.)$ is B-differentiable at a, then the set given by*

$$C := \left\{v \in X : d'_A(a, v) = 0\right\}$$

is an approximation cone to A at a.

The following proposition generalizes results in [4,46] in which X is a finite-dimensional space. It is obtained by combining previous assertions.

Proposition 36. *If the set A is tangentially compact at $a \in A$, then the following assertions are equivalent*:

(i) *$C := T(A, a)$ is an approximation cone to A at a*;
(ii) *$d(., A)$ is B-differentiable at a*;
(iii) *$d(., A)$ is directionally differentiable at a*;
(iv) *A is proto-differentiable at a.*

Furthermore, one has

$$T^i(A, a) = T(A, a) = C = \left\{v \in X : d'_A(a, v) = 0\right\},$$
$$d'_A(a, \cdot) = d(\cdot, C).$$

The following example shows that in any infinite-dimensional space the preceding equivalences fail if A is not tangentially compact at a, even when $T(A, a)$ is locally compact.

Example. Let (u_n) be a sequence of the unit sphere of the infinite-dimensional space X which does not have any cluster point and let $A := \left\{2^{-n}u_n : n \in \mathbb{N}\right\} \cup \{0\}$. Then $T(A, 0) = \{0\}$ and A is not approximated by $T(A, a)$ at $a = 0$.

5. Approximations and normal cones

It is the purpose of this section to show that an approximation property implies some consequences on the normal cone, in particular that the normal cone coincides with the normal cone in the sense of Fréchet.

We denote by X^* the topological dual of the n.v.s. X. The *normal cone* to the set A at $a \in A$ is the cone $N(A, a)$ given by:

$$N(A,a) = (T(A,a))^0 = \left\{x^* \in X^* : \left\langle x^*, v\right\rangle \leqslant 0 \ \forall v \in T(A,a)\right\}.$$

The *Fréchet normal cone* $N^-(A, a)$ to the set A at a is given by:

$$N^-(A,a) = \left\{ x^* \in X^* : \limsup_{x(\in A\setminus\{a\})\to a} \left\langle x^*, \frac{x-a}{\|x-a\|}\right\rangle \leqslant 0 \right\}.$$

Clearly $N^-(A, a) \subset N(A, a)$: given $x^* \in N^-(A, a)$, for each $v \in T(A, a)$ with norm 1 one can find a sequence $((r_n, v_n))_n \subset \mathbb{P} \times X$ with limit $(0, v)$ such that $a_n := a + r_n v_n \in A$ for each n, so that $(r_n^{-1} \|a_n - a\|) \to 1$ and $\langle x^*, v\rangle = \lim_{n\to\infty} \left\langle x^*, r_n^{-1}(a_n - a)\right\rangle \leqslant 0$, i.e. $x^* \in N(A, a)$.

The following consequence of the approximation property is noteworthy.

Theorem 37. *If the set A is approximated by $a + T(A, a)$ at a, then $N(A, a) = N^-(A, a)$.*

Proof. Let $x^* \in N(A, a)$. As the set A is approximated by $a + T(A, a)$ at a it is S-approximated by $a + T(A, a)$ at a

$$\lim_{x(\in A\setminus\{a\})\to a} \frac{d(x, a+T(A,a))}{\|x-a\|} = 0.$$

So, for each $\varepsilon > 0$, there exists $\eta > 0$ such that for all $x \in (A \setminus \{a\}) \cap (a + \eta U_X)$, there exists $v \in T(A, a)$ with $\|v - (x - a)\| \leqslant \varepsilon \|x - a\|$. Thus, one has

$$\left\langle x^*, x-a\right\rangle \leqslant \left\langle x^*, v\right\rangle + \varepsilon \left\|x^*\right\| \|x-a\| \leqslant \varepsilon \left\|x^*\right\| \|x-a\|,$$

i.e. $\langle x^*, \|x - a\|^{-1} (x - a)\rangle \leqslant \varepsilon \|x^*\|$ and, as ε is arbitrary, one gets $x^* \in N^-(A, a)$. □

6. Approximations and operations

In this section, we present a slight extension of an approximation result of Maurer–Zowe [23] about an intersection and an inverse image. Here $A := B \cap g^{-1}(C)$, where $g : X \longrightarrow Y$ is B-differentiable at $a \in A$, with a Lipschitzian derivative $g'(a)$, X and Y are Banach spaces and B and C are arbitrary subsets of X and Y, respectively. In [23] $B = X$, C is a closed convex cone and C' is the tangent cone to C at $g(a)$; there it is shown that A is approximated by $a + L(A, a)$, where $L(A, a)$ is the linearized cone of A at a, given by:

$$L(A,a) := T(B,a) \cap (g'(a))^{-1}(T(C, g(a))).$$

Moreover, the qualification condition used here is a tangential condition, hence is more general than the condition used in [23].

Theorem 38. *Suppose B and C are approximated by* $a + B'$ *and* $g(a) + C'$, *respectively, and suppose that the following metric regularity condition* (M) *is satisfied*

$$(\mathrm{M}) \quad \begin{cases} \exists\, k > 0, r > 0 : \forall u \in B' \cap rU_X, \\ d(u, B' \cap (g'(a))^{-1}(C')) \leqslant kd(g'(a)(u), C'). \end{cases}$$

Then the set A is approximated by $A' := a + B' \cap g'(a)^{-1}(C')$ *at a.*

Proof. Let us show that A is S-approximated by A' at a. Let $b : B \to a + B'$ be such that b is tangent to I_X w.r.t. B at a; let $c : C \to g(a) + C'$ be tangent to I_Y w.r.t. C at $g(a)$. Let $r(x) := g(x) - g(a) - g'(a)(b(x) - a)$, so that $\lim_{x(\in B\setminus\{a\})\to a} \|x - a\|^{-1} r(x) = 0$. For $x \in A$ we have

$$g'(a)(b(x) - a) = g(x) - g(a) - r(x) \in C - g(a) - r(x),$$

so that, for $x \in A$ close enough to a, using (M) with $u := b(x) - a \in B'$, we can find $h(x) \in A'$ such that

$$\begin{aligned} d(b(x) - a, h(x) - a) &\leqslant 2kd(g'(a)(b(x) - a) + r(x), C') + 2k\|r(x)\| \\ &\leqslant 2kd(g(x) - g(a), c(g(x)) - g(a)) + 2k\|r(x)\| \\ &= o(\|x - a\|), \end{aligned}$$

so that $\|x - h(x)\| \leqslant \|x - b(x)\| + o(\|x - a\|)$ and h is tangent to I_A at a. □

Corollary 39. *Suppose X and Y are Banach spaces and g is Fréchet-differentiable at a. Suppose B and C are approximated by* $a + B'$ *and* $g(a) + C'$, *respectively, where* B' *and* C' *are closed convex cones of X and Y, respectively. Suppose that the following condition is satisfied*:

$$(\mathrm{L}) \quad g'(a)B' - C' = Y.$$

Then the set A is approximated by $A' := a + B' \cap g'(a)^{-1}(C')$ *at a.*

Proof. Let us show that condition (M) is satisfied. Thanks to the Robinson and Ursescu [40] open mapping theorem, condition (L) is satisfied if, and only if, there is some $\alpha > 0$ such that:

$$U_Y \subset g'(a)(B' \cap \alpha U_X) - C' \cap \alpha U_Y. \tag{7}$$

Now, given $u \in B'$, we can pick $c'(u) \in C'$ such that $d(g'(a)(u), c'(u)) \leqslant 2d(g'(a)(u), C')$.

Using (7) and an homogeneity argument we can find some $b'(u) \in B'$, $c''(u) \in C'$ with

$$\left\|b'(u)\right\| \leqslant \alpha \left\|c'(u) - g'(a)(u)\right\|, \quad \left\|c''(u)\right\| \leqslant \alpha \left\|c'(u) - g'(a)(u)\right\|,$$
$$c'(u) - g'(a)(u) = g'(a)b'(u) - c''(u).$$

Then we get, as $u + b'(u) \in B' \cap g'(a)^{-1}(c'(u) + c''(u))$ with $c'(u) + c''(u) \in C'$,

$$d(u, B' \cap g'(a)^{-1}(C')) \leqslant \|u - (u + b'(u))\| \leqslant 2\alpha d(g'(a)(u), C').$$

It remains to apply the preceding theorem. □

Taking $B' = \mathrm{co}(T(B, a))$ and $C' = \mathrm{co}(T(C, g(a)))$ we deduce the following consequence.

Corollary 40. *Suppose that the sets B and C are pseudo-convex at a and $g(a)$, respectively, in the sense that $B \subset a+\mathrm{co}(T(B, a))$ and $C \subset g(a)+\mathrm{co}(T(C, g(a)))$. Under condition* (L), *with $B' = \mathrm{co}(T(B, a))$ and $C' = \mathrm{co}(T(C, g(a)))$, the set $A = B \cap g^{-1}(C)$ is approximated by $A' = a + B' \cap g'(a)^{-1}(C')$ at a.*

The last corollary has been proved in [22,23], in the case where $B = X$ and C is a closed convex cone, and under the regularity condition

$$\text{(R)} \quad 0 \in \mathrm{int}(g(a) + g'(a)(B - a) - C)$$

which is more classical (see [41,28] for instance), but more exacting than condition (L).

The following proposition, close to results in [30], points out the interest for optimization theory of the notions of approximation we considered. As there and elsewhere, we say that a is a *minimizer of order one* of a function f on A if there exist some $\alpha, \rho > 0$ such that

$$f(x) \geqslant f(a) + \alpha \|x - a\| \quad \forall x \in A \cap (a + \rho U_X).$$

Proposition 41. *Suppose the admissible set A is approximated at a by a set B and let $f, g : X \to \mathbb{R}$ be tangent at a w.r.t. B, with f locally Lipschitzian around a. If a is a minimizer of order one of g on B, then it is also a minimizer of order one of f on A.*

Proof. Without loss of generality we may suppose $a = 0$, $f(a) = g(a) = 0$. Let $\gamma > 0$ be such that $g(v) \geqslant \gamma\|v\|$ for $v \in B$, $\|v\|$ small enough. Let k be the Lipschitz rate of f on some neighborhood of 0 and let $h : A \to B$ be tangent to I_X on A at 0. For any $\beta > 0$, $\delta > 0$ one can find a neighborhood V of 0 such that for $x \in A \cap V$ one has

$$\begin{aligned} \|h(x) - x\| &\leqslant \beta\|x\|, \\ f(x) &\geqslant f(h(x)) - k\beta\|x\|, \\ f(h(x)) &\geqslant g(h(x)) - \delta\|h(x)\|, \\ g(h(x)) &\geqslant \gamma\|h(x)\|, \end{aligned}$$

so that, combining these inequalities one gets

$$f(x) \geqslant (\gamma - \delta)\|h(x)\| - k\beta\|x\| \geqslant ((\gamma - \delta)(1 - \beta) - k\beta)\|x\|,$$

with $\alpha := (\gamma - \delta)(1 - \beta) - k\beta > 0$, provided β, δ have been chosen small enough. □

It follows that the notion of minimizer of order one is invariant under approximations of the functions and the sets.

References

[1] A. Agadi, Approximation des ensembles en un point et à l'infini, Thesis, University of Pau, June 1995.
[2] Z. Artstein, Piecewise linear approximations of set-valued maps, J. Approx. Theory 56 (1) (1989) 41–47.
[3] J.-P. Aubin, H. Frankowska, Set-Valued Analysis, Birkhäuser, Basel, 1990.
[4] A. Auslender, R. Cominetti, A comparative study of multifunction differentiability with application in mathematical programming, Math. Oper. Res. 16 (1991) 240–258.
[5] D. Aze, J.-P. Penot, Operations on convergent families of sets and functions, Optimization 21 (1990) 521–534.
[6] D. Aze, A. Rahmouni, Lipschitz behavior of the Legendre–Fenchel transform, Set-Valued Anal. 2 (1994) 35–48.
[7] D. Aze, A. Rahmouni, Intrinsic bounds for Kuhn–Tucker points of perturbed convex programs, in: R. Durier, C. Michelot (Eds.), Seventh French–German Conference on Optimization, Dijon, June 1994, Lecture Notes in Economics and Mathematics Systems, vol. 429, Springer, Berlin, 1995, pp. 17–35.
[8] G. Beer, D. Pai, On convergence of convex sets and relative Chebyshev centers, J. Approx. Theory 62 (2) (1990) 147–169.
[9] J. Borwein, M. Fabian, A note on regularity of sets and of distance functions in Banach space, J. Math. Anal. Appl. 182 (1994) (566–570) 17–35.
[10] B. Brosowski, R. Wegmann, On the lower semicontinuity of the set-valued metric projection, J. Approx. Theory 8 (1) (1973) 84–100.
[11] A.L. Brown, Set valued mappings, continuous selections, and metric projections, J. Approx. Theory 57 (1) (1989) 48–68.
[12] A.L. Brown, F. Deutsch, V. Indumathi, P.S. Kenderov, Lower semicontinuity concepts, continuous selections, and set valued metric projections, J. Approx. Theory 115 (1) (2002) 120–143.
[13] J. Danes, J. Durdil, A note on geometric characterization of differentiability, Comm. Math. Univ. Carolinae 17 (1976) 195–204.
[14] A. Dontchev, A. Hager, Implicit functions, Lipschitz maps and stability in optimization, Math. Oper. Res. 19 (1994) 753–768.
[15] J. Durdil, On the geometric characterization of the differentiability I, Comment. Math. Univ. Carolinae 15 (1974) 521–540.
[16] J. Durdil, On the geometric characterization of the differentiability II, Comment. Math. Univ. Carolinae 15 (1974) 727–744.
[17] M. Fabian, Theory of Fréchet cones, Casopis pro pestovani Mat. 107 (1982) 37–58.
[18] T. Fischer, A continuity condition for the existence of a continuous selection for a set-valued mapping, J. Approx. Theory 49 (1987) 340–345.
[19] A. Jofre, J.-P. Penot, Comparing new notions of tangent cones, J. London Math. Soc. (2) 40 (1989) 280–290.
[20] L.A. Liusternik, Conditional extrema of functionals, Mat. Sb. 41 (1934) 390–401.
[21] D.T. Luc, J.-P. Penot, Convergence of asymptotic directions, Trans. Amer. Math. Soc. 353 (2001) 4095–4121.
[22] H. Maurer, First and second-order sufficient conditions in mathematical programming and optimal control, Math. Prog. Study 14 (1981) 163–177.
[23] H. Maurer, J. Zowe, First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems, Math. Programming 16 (1979) 98–110.
[24] F. Mignot, Contrôle dans les inégalités variationnelles elliptiques, J. Funct. Anal. 22 (1976) 130–185.
[25] J. Outrata, M. Kocvara, J. Zowe, Nonsmooth approach to optimization problems with equilibrium constraints. Theory, applications and numerical results, Nonconvex Optimization and its Applications, vol. 28, Kluwer, Dordrecht, 1998.
[26] D.V. Pai, P. Govindarajulu, On set-valued f-projections and f-farthest point mappings, J. Approx. Theory 42 (1984) 4–13.
[27] J.S. Pang, Newton's method for B-differentiable equations, Math. Oper. Res. 15 (1990) 311–341.
[28] J.-P. Penot, On regularity conditions in mathematical programming, Math. Programming Study 19 (1982) 167–199.
[29] J.-P. Penot, Compact nets, filters, and relations, J. Math. Anal. Appl. 93 (2) (1983) 400–417.
[30] J.-P. Penot, Generalized higher order derivatives and higher order optimality conditions, preprint, University of Santiago, November 1984.

[31] J.-P. Penot, Differentiability of relations and differential stability of perturbed optimization problems, SIAM J. Control Optim. 22 (4) (1984) 529–551.
[32] J.-P. Penot, Miscellaneous incidences of convergence theories in optimization and nonsmooth analysis, I: behavior of solutions, Set-Valued Anal. 2 (1–2) (1994) 259–274.
[33] J.-P. Penot, Conditioning convex and nonconvex problems, J. Optim. Theory Appl. 90 (3) (1996) 535–554.
[34] J.-P. Penot, R. Ratsimahalo, Subdifferentials of distance functions approximations and enlargements, Acta Math. Sinica, in press.
[35] J.-P. Penot, P. Terpolilli, Cônes tangents et singularités, C. R. Acad. Sci. Paris 296 (1983) 721–724.
[36] J.-P. Penot, C. Zalinescu, Continuity of usual operations and variational convergences, Set-Valued Anal. 11 (2003) 225–256.
[37] J.-P. Penot, C. Zalinescu, Bounded (Hausdorff) convergence: basic facts and applications, in: F. Giannessi and A. Maugeri (Eds.), Variational Analysis and Applications, Proceedings of Erice Conference 2003, Kluwer, Dordrecht, 2005.
[38] K. Przeslawski, L.E. Rybinski, Concepts of lower semicontinuity and continuous selections for convex valued multifunctions, J. Approx. Theory 68 (3) (1992) 262–282.
[39] S.M. Robinson, Stability theory for systems of inequalities, I: linear systems, SIAM J. Numer. Anal. 12 (1975) 754–769.
[40] S.M. Robinson, Regularity and stability for convex multivalued functions, Math. Oper. Res. 1 (1976) 130–143.
[41] S.M. Robinson, Stability theory for systems of inequalities, II: differentiable non linear systems, SIAM J. Numer. Anal. 12 (1976) 497–512.
[42] S.M. Robinson, Strongly regular generalized equations, Math. Oper. Res. 5 (1980) 43–62.
[43] S.M. Robinson, An implicit function theorem for a class of nonsmooth functions, Math. Oper. Res. 16 (2) (1991) 292–309.
[44] R.T. Rockafellar, Proto-differentiability of set-valued mappings and its applications in optimization, Ann. Inst. H. Poincaré, Anal. Non-Linéaire 6 (1989) 449–482.
[45] R.T. Rockafellar, R.J.-B. Wets, Variational Analysis, Springer, Berlin, 1997.
[46] A. Shapiro, On differentiability of metric projection in $\mathbb{R}^n$: boundary case, Proc. Amer. Math. Soc. 99 (1987) 123–128.
[47] A. Shapiro, Directional differentiability of metric projections onto moving sets at boundary points, J. Math. Anal. Appl. 131 (2) (1988) 392–403.
[48] A. Shapiro, On concepts of directional differentiability, J. Optim. Theory Appl. 66 (1990) 477–487.
[49] Y. Sonntag, C. Zalinescu, Set convergence, an attempt of classification, Trans. Amer. Math. Soc. 340 (1) (1995) 199–226.
[50] J. Treiman, Shrinking generalized gradients, Nonlinear Anal. Theory Methods Appl. 12 (12) (1988) 1429–1450.
[51] J. Treiman, Finite dimensional optimality conditions: B-gradients, J. Optim. Theory Appl. 62 (1) (1989) 139–150.
[52] V.A. Ubhaya, Lipschitzian selections in approximation from nonconvex sets of bounded functions, J. Approx. Theory 56 (2) (1989) 217–224.
[53] R.A. Vitale, Approximation of convex set-valued functions, J. Approx. Theory 26 (4) (1979) 301–316.