On Differentiability of Metric Projections onto Moving Convex Sets *

Darinka Dentcheva
Humboldt-Universität Berlin
10099 Berlin, Germany
dary@mathematik.hu-berlin.de

Abstract

We consider properties of the metric projections onto moving convex sets in normed linear spaces. Under certain conditions about the norm, directional differentiability (of higher order) of the metric projections at boundary points is characterized. The characterization is formulated in terms differentiability of multifunctions and properties of the set-derivatives are shown.

Keywords: Ben-Tal-Zowe differentiability, metric projection, Wijsman topology, E-space

1 Introduction

This paper deals with differentiability properties of the metric projection in normed spaces. Let X be a linear normed space. We denote the distance between points in X by $d(x,y) = \|x-y\|$. Given a nonempty closed subset $C \subseteq X$, the distance function $d : X \to R_+$ associated with this set is defined by:

$$d(x,C) := \inf\{\|x-y\| : y \in C\}. \tag{1}$$

Furthermore, let $\Pr : X \rightrightarrows C$ be the mapping given by

$$\Pr(x,C) = \operatorname{argmin}\{\|x-y\| : y \in C\}. \tag{2}$$

The problem (1) is called a best approximation problem and the mapping (2) is called a metric projection map or the operator of best approximation. There are very many papers devoted to the existence and properties of the metric projections. Many of those investigations are motivated by the fact that approximation and optimization

*Supported by Program HSP III of the Humboldt University Berlin

problems can be casted in the above setting. We refer the interested reader to the monograph [23] for an overview of the theory of best approximation in normed linear spaces by elements of linear subspaces. In general, it is clear that the set $\Pr(x, C)$ may be empty. A set C is called *proximinal* provided $\Pr(x, C)$ is nonempty for each x. If moreover $\Pr(x, C)$ is singleton for each $x \in X$ C is called *Chebyshev*.

In case X is a Hilbert space and C a convex set, it is known that for all $x \in X$ the metric projection exists and is unique. Moreover, the mapping $\Pr(\cdot, C)$ is Lipschitz continuous and for every $x \in C$, it is directionally differentiable (see [27]). Let us emphasize that the metric projection to nonempty convex sets in reflexive Banach spaces always exists. As a matter of fact, its existence characterizes reflexivity (see [16], Sec.19).
The uniqness of the projection is provided by a strongly convex norm of the space X. The continuity of $\Pr(\cdot, C)$ requires additional assumptions. There are examples of a discontinuous $\Pr(\cdot, C)$ with C a linear subspace of a reflexive Banach space (cf. [11]). Let us recall that a Banach space is called an *E-space* if the following is true:

- X is reflexive;
- the norm of X is strongly convex;
- X has the Kadec-Klee property, i.e.,

$$x_n \to x \text{ weakly },\ \|x_n\| \to \|x\| \text{ imply } \|x_n - x\| \to 0.$$

Under the assumption that X is an E-space, it is known that $\Pr(\cdot, C)$ is Differentiability properties of $\Pr(\cdot, C)$ for X being finite dimensional or a Hilbert space are studied in a number of publications (see, e.g., [12, 13, 14, 15, 24]) and the references therein. The continuity in the situation of moving sets is studied with respect to different convergences of the sets (see, e.g., [3, 11, 23]). Differentiability in situation of moving sets is studied in finite dimensions in [2, 18, 25, 26]. Also second order directional differentiability is considered for the finite dimensional case in [2].

The goal of this paper is to generalize some of the results of the latter papers to the situation of E-space. Also, we present two extension of those results: necessary and sufficient conditions for the Hadamard-directional differentiability of the metric projections are given, and, moreover, conditions for higher order differentiability are identified.
We adopt the idea of characterizing these property by some tangential approximations of a multifunction. An appropriate approximation for this purpose is described in section 2 of this paper. Several concepts for approximating multifunctions have been introduced in the literature. We shall briefly state the relations of this concept to those considered in [2, 19, 25, 26] and other frequently used in the literature concepts. Since the considered approximation has a value in its own and maybe interesting for other applications, some of its properties are shown in section 3. Finally, the concepts

are applied to characterize the higher order differentiability of the mapping $\Pr(\cdot,\cdot)$ in section 4.

2 Tangential Approximation of Multifunctions

Let $F : X \rightrightarrows Y$ be defined on a linear normed space X and its images be nonempty closed convex subsets of a linear normed space Y. We denote the graph of F by graph F.

Let $\mathcal{F}(Y)$ denote the hyperspace of nonempty closed sets in Y, and let $\mathcal{F}_c(Y)$ stand for the corresponding hyperspace of nonempty closed convex sets. We endow these spaces with the Wijsman topology determined by the norm and denoted by τ. Then τ is the weak topology determined by the family $\{d(y,\cdot) : y \in Y\}$. We recall some basic facts concerning the topological space $(\mathcal{F}(Y), \tau)$ and refer to [3, 4] for further details. A subbase of τ consists of all sets of the form:

$$\{A \in \mathcal{F}(Y) : d(y, A) < \varepsilon\} \text{ and } \{A \in \mathcal{F}(Y) : d(y, A) > \varepsilon\} \quad (y \in Y,\ \varepsilon > 0).$$

A sequence of closed sets $\{A_n\}$, $A_n \subseteq Y$ converges to some $A \subseteq Y$ in $(\mathcal{F}(Y), \tau)$ (written $A_n \overset{W}{\to} A$) if and only if the sequence of the distance functions converges pointwise, i.e.,

$$A_n \overset{W}{\to} A \quad \text{if and only if} \quad d(y, A_n) \to d(y, A) \text{ for all } y \in Y,$$

The space $(\mathcal{F}(Y), \tau)$ is metrizable if and only if Y is separable. For $\{y_i\}$ being a countable dense subset of the separable metric space (Y, d) the following metric is compatible with the Wijsman topology:

$$\rho_d(A, B) = \sum_{i=1}^{\infty} 2^{-i} \min\{1, |d(y_i, A) - d(y_i, B)|\}.$$

Moreover, if (Y, d) is a complete separable metric space, then $(\mathcal{F}(Y), \tau)$ is a Polish space, i.e., it is separable and there is a complete metric compatible with the topology (cf. [5]).
Differentiability of a multifunction is usually introduced by choosing a class of approximating multifunctions. Several concepts of differentiability have been investigated in the literature. They typically represent some cone-approximation of the graph of the multifunction. Furthermore, corresponding second-order tangent sets have been developed. Some geometrical concepts are presented in [1], which correspond to looking at certain tangent cone to the graph of the multifunction. A stronger differentiability notion related to interesting phenomena in optimization, called proto-differentiability, has been introduced by Rockafellar in [20]. Another attractive concept of differentiability, called semi-differentiability, goes back to Penot [19] and corresponds to the concept of tangential approximation due to Shapiro ([25, 26].
The notion of tangential approximation is modified in this paper by using a metric convergence on the hyperspace of nonempty closed subsets of Y, which is better suited

to the infinite-dimensional case. Using this notion we obtain results on the first- and higher-order differentiability of the metric projection onto moving closed convex sets in E-spaces generalizing those established in [2, 25, 26] in the finite-dimensional setting.

Definition 2.1 *A mapping $F : X \rightrightarrows Y$ is called directionally differentiable at a point $(x, y) \in \operatorname{graph} F$ in direction $h \in X$, if for any sequence $t_n \downarrow 0$ the limit*

$$F'(x, y; h) = \lim_{t_n \downarrow 0} \frac{F(x + t_n h) - y}{t_n}$$

exists in $(\mathcal{F}(Y), \tau)$.

The original definition of semi-differentiability (cf.[19]) is modified in the following way:

Definition 2.2 *A mapping $F : X \rightrightarrows Y$ is called semi-differentiable at a point $(x, y) \in \operatorname{graph} F$, if for all sequences $t_n \downarrow 0$ and $h_n \to h$ the limit*

$$DF(x, y)(h) = \lim_{t_n \downarrow 0, h_n \to h} \frac{F(x + t_n h_n) - y}{t_n}$$

exists in $(\mathcal{F}(Y), \tau)$ for all $h \in X$.

The convergence $h_n \to h$ and all convergences on X will be understood as the norm-convergence. Therefore, we can omit the indices in what follows and write simply $t \downarrow 0, h \to \bar{h}$.

We recall that, originally, these two notions are associated with the Kuratowski-Painlevé convergence of the corresponding differential quotients in the literature. For Y being infinite dimensional, there is no topologization of this convergence in general (cf.[3]). This means that, in general, there is no topology on $\mathcal{F}(Y)$ such that the convergence of elements of $\mathcal{F}(Y)$ in the sense of Kuratowski-Painlevé will force convergence with respect to the topology. We require some stronger convergence in the definitions above. The Wijsman topology is a natural counterpart of the Kuratowski-Painlevé convergence for infinite dimensional spaces. Both convergences coincide if and only if Y is finite dimensional.

The various notions of first order differentiability are extended to higher order properties e.g. in [1, 2, 22]. The approach used to introduce a higher order derivative is to define sets that are higher order “tangent” to the graph of the multifunction. This way of introducing second-order derivatives typically leads to a set-valued counterpart of second order directional derivative in the sense of Ben-Tal and Zowe ([7]. Let us illustrate this by discussing the notions considered in [2]. A set $A \subseteq I\!R^n$ is called pseudo-differentiable in [2] if

$$\liminf_{t \downarrow 0} \frac{A - x}{t} = \limsup_{t \downarrow 0} \frac{A - x}{t}.$$

The limit set builds a first-order derivative. The set A is second-order pseudo-differentiable with respect to y taken from the first-order derivative if

$$\liminf_{t\downarrow 0} \frac{A-x-ty}{t^2/2} = \limsup_{t\downarrow 0} \frac{A-x-ty}{t^2/2}.$$

Having in mind that the Kuratowski-Painlevé convergence of closed sets in $I\!R^n$ is characterized by the convergence of the distance function, it follows that A is pseudo-differentiable at x if the distance function $d(\cdot, A)$ is directionally differentiable, and A is twice pseudo-differentiable if $d(\cdot, A)$ admits a second order directional derivative in the sense of Ben-Tal and Zowe, i.e., setting $d_A(\cdot) = d(\cdot, A)$

$$d''_A(x;h,v) = \lim_{t\downarrow 0} \frac{d_A(x+th+t^2v/2) - d_A(x) - td'_A(x;h)}{t^2/2}.$$

The second order semi-differentiability is considered there in a similar way, i.e., the multifunction $F : I\!R^m \rightrightarrows I\!R^n$ is called twice semi-differentiable at (x,y) into a direction (h,z) if the limit

$$\lim_{t\downarrow 0, v\to v_0} \frac{F(x+th+t^2v/2) - y - tz}{t^2/2}$$

exists for all $v_0 \in I\!R^m$. Here $z \in DF(x,y)(h)$.

The definition is suitable for our consideration. We only change the topology with respect to which the limit is taken, and extend the notion to higher order differentiability inductively. To simplify notation we shall use

$$\mathbf{h}^i = (h^1, \dots, h^i);\ \text{respectively}\ \mathbf{z}^i = (z^1, \dots, z^i) \tag{3}$$

for a proper direction $\mathbf{h}^i$ and for an element $\mathbf{z}^i$, respectively, for which the following holds:

$$z^1 \in DF(x,y)(h^1), z^2 \in D^2F(x,y;\mathbf{h}^1,\mathbf{z}^1)(h^2), \dots, z^i \in D^{k-1}F(x,y;\mathbf{h}^{i-2},\mathbf{z}^{i-2})(h^{i-1}). \tag{4}$$

Let us suppose, that $F : X \rightrightarrows Y$ is directionally differentiable of order $k-1$ at (x,y) into the directions $\mathbf{h}^{k-1}$) with respect to $\mathbf{z}^{k-2}$ which satisfy the condition (4). Furthermore, we set:

$$\Delta^k F(x,y;\mathbf{h}^{k-1},\mathbf{z}^{k-1})(v) = \frac{F(x+th^1+\ldots+t^{k-1}h^{k-1}+t^kv) - y - tz^1 - \ldots - t^{k-1}z^{k-1}}{t^k} \tag{5}$$

for the k-th differential quotient into the direction $\mathbf{h}^k := (\mathbf{h}^{k-1}, v)$ with respect to $\mathbf{z}^{k-1}$.

Definition 2.3 *The mapping $F : X \rightrightarrows Y$ is called directionally differentiable of order k at (x,y) into the direction $\mathbf{h}^k$ with respect to $\mathbf{z}^{k-1}$ as in (3-4), if for all sequences $t \downarrow 0$, $v \to h^k$ the limit*

$$D^kF(x,y;\mathbf{h}^{k-1},\mathbf{z}^{k-1})(h^k) = \lim_{t\downarrow 0, v\to h^k} \Delta^k F(x,y;\mathbf{h}^{k-1},\mathbf{z}^{k-1})(v)$$

exists in $(\mathcal{F}(Y), \tau)$.

Let us emphasize that a multifunction is directionally differentiable (of order k) at a certain point if the corresponding limit exists and is, in addition, nonempty. This requirement is implicitly included in the definitions above since only nonempty subsets are included in $\mathcal{F}(Y)$ and we consider the convergence there. The existence of the limit with respect to the Kuratowski-Painlevé convergence does not guarantee nonempty derivatives.
Various differentiability concepts for finite dimensions are compared in [2, 20]. In [2], it is shown for mappings in finite-dimensional spaces that the tangential approximations considered in [25, 26], if they exist, coincide with the semi-derivatives of [19].
Higher order derivatives of multifunctions are considered also in [9]. The concept there uses the same topology on $\mathcal{F}(Y)$ but differs from this one. The differentiability concept considered in [9] corresponds to the existence of a kind of a Taylor's expansion into some direction. It is suitable for applications to the so-called delta method.

3 The Process Nature of the Derivatives

Recall that a multifunction is called:

- *closed* if graph F is closed;
- *a process* if graph F is a cone;

It is not difficult to see that the semi-derivatives are closed sets. We shall show that that $DF(x, y)(\cdot)$ is a closed continuous process, and that the higher order derivative are closed, continuous "process-like" multifunctions of the directions. Moreover, the derivatives have convex images, if F has convex images.

A multifunction $F : X \rightrightarrows Y$ is *τ-continuous* at a point $\bar{x} \in X$ if for all sequences $x \to \bar{x}$

$$\tau - \lim_{x \to \bar{x}} F(x) = F(\bar{x}).$$

Proposition 3.1 *The multifunctions $D^i F(x, y; h^1, \ldots, h^{i-1})(\cdot)$, $i = 1, 2, \ldots$ have the following properties:*

(a) If $F(x)$ are convex sets for all x in a neighbourhoods of $\bar{x}$, then the semi-derivatives at $(\bar{x}, y)$ have convex images, too.

(b) $z \in D^i F(x, y; \mathbf{h}^{i-1}, \mathbf{z}^{i-1})(h^i)$ if and only if $\lambda^i z \in D^i F(x, y; \lambda h^1, \ldots, \lambda^{i-1} h^{i-1}, \lambda z^1, \ldots, \lambda^{i-1} z^{i-1}))(\lambda^i h^i)$, i.e., $DF(x, y)$ is a process.

Proof: Let $u, w \in D^iF(x,y;\mathbf{h}^{i-1},\mathbf{z}^{i-1})$ and $\lambda \in (0,1)$. According to the definition, given sequences $t_n \downarrow 0$ and $v_n \to v$, $v_n, v \in X$ we obtain that both:

$$d\Big(u, \Delta^iF(x,y;\mathbf{h}^{i-1},\mathbf{z}^{i-1})(v_n)\Big) \text{ and } d\Big(w, \Delta^iF(x,y;\mathbf{h}^{i-1},\mathbf{z}^{i-1})(v_n)\Big)$$

converge to 0. Consequently, there are sequences u_n and w_n such that $u_n, w_n \in \Delta^iF(x,y;\mathbf{h}^{i-1},\mathbf{z}^{i-1})(v_n)$ and $u_n \to u$, $w_n \to w$. Since $F(x+th^1+\ldots+t^{k-1}h^{k-1}+t^kv)$ is convex, the points $\lambda(y+t_nz^1+\ldots+t_n^{i-1}z^{i-1}+t_n^iu_n)+(1-\lambda)(y+t_nz^1+\ldots+t_n^{i-1}z^{i-1}+t_n^iw_n) \in F(x+th^1+\ldots+t^{k-1}h^{k-1}+t^kv)$. Consequently, $\lambda u_n+(1-\lambda)w_n \in \Delta^iF(x,y;\mathbf{h}^{i-1},\mathbf{z}^{i-1})(v_n)$. Hence, for all $y \in Y$

$$\begin{aligned}
&d\Big(\lambda u+(1-\lambda)w, \Delta^iF(x,y;\mathbf{h}^{i-1},\mathbf{z}^{i-1})(v)\Big) = \\
&\lim_{n\to\infty} d\Big(\lambda u+(1-\lambda)w, \Delta^iF(x,y;\mathbf{h}^{i-1},\mathbf{z}^{i-1})(v_n)\Big) \leq \\
&\lim_{n\to\infty} d\Big(\lambda u+(1-\lambda)w, \lambda u_n+(1-\lambda)w_n\Big)+ \\
&\lim_{n\to\infty} d\Big(\lambda u_n+(1-\lambda)w_n, \Delta^iF(x,y;\mathbf{h}^{i-1},\mathbf{z}^{i-1})(v_n)\Big) = 0
\end{aligned}$$

This proves (a).
The second assertion will be proved by induction. Let $z \in DF(x,y)(h)$. By definition $\lambda z \in DF(x,y)(\lambda h)$, if $d(\lambda z, t_n^{-1}(F(x+t_n\lambda h_n)-y))$ converges to 0 while $n \to \infty$. Denoting $\hat{t}_n = \lambda t_n$, we have:

$$d\Big(\lambda z, t_n^{-1}(F(x+t_n\lambda h_n)-y)\Big) = \lambda d\Big(z, \hat{t}_n^{-1}(F(x+\hat{t}_nh_n)-y)\Big)$$

Since $\hat{t}_n \downarrow 0$ it holds $d\Big(z, \hat{t}_n^{-1}(F(x+\hat{t}_nh_n)-y)\Big) \to 0$ and this proves the assertion for the case $i=1$. Let us assume that it is true for $i \leq k$. Given $z \in D^{k+1}F(x,y;\mathbf{h}^k,\mathbf{z}^k)(h^{k+1})$, we have to prove that $\lambda^{k+1}z \in D^{k+1}F(x,y;\lambda h^1,\ldots,\lambda^kh^k,\lambda z^1,\ldots,\lambda^kz^k)(\lambda^{k+1}h^{k+1})$ By the induction assumption $\lambda^iz_i \in D^iF(x,y;\lambda h^1,\ldots,\lambda^{i-1}h^{i-1},\lambda z^1,\ldots,\lambda^iz^i)(\lambda^ih^i)$ for $i \leq k$. Denoting $\hat{t}_n = \lambda t_n$, we have:

$$\begin{aligned}
&d\Big(\lambda^{k+1}z, t_n^{-k-1}(F(x+t_n\lambda h^1+\ldots+t_n^{k+1}\lambda^{k+1}h_n^{k+1})-y-t_n\lambda z^1-\ldots-t_n^k\lambda^kz^k)\Big) = \\
&\lambda^{k+1}d\Big(z, \hat{t}_n^{-k-1}(F(x+\hat{t}_nh^1+\ldots+\hat{t}_n^kh_n^{k+1})-y-\hat{t}_nz^1-\ldots-\hat{t}_n^kz^k)\Big)
\end{aligned}$$

The right-hand-side converges to 0 by definition, and by the assumption of the proposition. This completes the proof. $\square$

The statement (b) is a counterpart of the properties of the higher order contingent sets, considered in [1], Section 4.7.

Proposition 3.2 *Let $F : X \rightrightarrows Y$ be directionally differentiable of order m, ($m \geq 1$) at $(\bar{x},\bar{y})$ into the direction $(h^1,\ldots,h^{k-1},h_n^k)$ with respect to $u = (z^1,\ldots,z^{k-1})$, ($n = 0,1,2,\ldots$), according to the notation (3-4). Assume that $h_n^k \to h_0^k$. Then the derivative of order $k \leq m$ is a closed and τ-continuous multifunction of the direction v, i.e.,*

$$D^kF(\bar{x},\bar{y};\mathbf{h}^{k-1},u)(h_0^k) = \tau - \lim_{h_n^k\to h_0^k} D^kF(\bar{x},\bar{y};\mathbf{h}^{k-1},u)(h_n^k).$$

Proof: Let $\mathbf{h}_n^k$ be a direction as in the assumption, and $u = (z^1, \dots, z^{k-1})$ be elements of the corresponding directional derivative. We consider $d(y, D^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, u)(h_n^k))$ for arbitrary but fixed $y \in Y$. Given an $\varepsilon > 0$, then there exists a positive number t_n (that can be taken smaller than t_{n-1}) for any $n = 1, 2, \dots$ such that

$$|d\Big(y, D^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, u)(h_n^k)\Big) - d\Big(y, \Delta^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, u)(h_n^k)\Big)| \leq \varepsilon$$

by the definition of a semi-derivative and τ-convergence. For the same reason the following inequality holds for n large enough:

$$|d\Big(y, D^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, u)(h_0^k)\Big) - d\Big(y, \Delta^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, u)(h_n^k)\Big)| \leq \varepsilon$$

Consequently, we obtain the estimation

$$|d\Big(y, D^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, u)(h_0^k)\Big) - d\Big(y, D^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, u)(h_n^k)\Big)| \leq 2\varepsilon$$

Letting $\varepsilon \to 0$ we obtain the convergence of the distance functions to the k-th order semi-derivative. We conclude the desired assertion since y was arbitrary. $\square$

Let us now observe the following useful property:

(P) Given an element $z^k \in D^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, \mathbf{z}^{k-1})(h_0^k)$, there exists a sequence $z_n^k \in D^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, \mathbf{z}^{k-1})(h_n^k)$ that converges to z^k. Indeed,

$$0 = d\Big(z^k, D^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, \mathbf{z}^{k-1})(h_0^k)\Big) = \lim_{n\to\infty} d\Big(z^k, D^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, \mathbf{z}^{k-1})(h_n^k)\Big),$$

which imply the existence of an element $z_n^k \in D^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, \mathbf{z}^{k-1})(h_n^k)$ such that $0 = \lim_{n\to\infty} \|z_n^k - z^k\|$.

Proposition 3.3 *Let $F : X \rightrightarrows Y$ be directionally differentiable of order k ($k \geq 1$) at certain point $(\bar{x}, \bar{y})$, then semi-derivative is a lower semi-continuous multifunction of the direction, i.e. whenever $V \cap D^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, \mathbf{z}^{k-1})(\bar{h}) \neq \emptyset$ for some open set $V \subseteq Y$, there is a neighborhood $U \subseteq X$ of $\bar{h}$, such that $V \cap D^k F(\bar{x}, \bar{y}; \mathbf{h}^{k-1}, \mathbf{z}^{k-1})(h) \neq \emptyset$ for all $h \in U$.*

Proof: We show the statement for the first-order semi-derivative; the property for the higher-order derivative follows in the same way. Assume the opposite. This means that there is an open set $V \subseteq Y$ such that $V \cap DF(\bar{x}, \bar{y})(\bar{h}) \neq \emptyset$ and for any neighborhood $U_\varepsilon \subseteq X$ of $\bar{h}$, there is a point $h_\varepsilon \in U_\varepsilon$ such that $V \cap DF(\bar{x}, \bar{y})(h_\varepsilon) = \emptyset$. Let $z \in V \cap DF(\bar{x}, \bar{y})(\bar{h})$. There exists a $\delta > 0$ such that $\tilde{z} \in V$ for all $\tilde{z} : \|\tilde{z} - z\| \leq \delta$. Taking $\varepsilon = 1/n$, we obtain a sequence $h_n \to \bar{h}$. At the same time, it holds

$$d(z, DF(\bar{x}, \bar{y})(h_n)) \geq \delta \text{ and } d(z, DF(\bar{x}, \bar{y})(\bar{h})) = 0.$$

This is a contradiction to the τ-continuity proved in the previous proposition. $\square$

It is shown in [2, 20] that proto-differentiability, resp. directional differentiability for pseudo-Lipschitzian multifunctions is equivalent to semi-differentiability. The latter relation was established also in our setting in [9]. Moreover, the pseudo-Lipschitz behavior of the multifunction plays a crucial role in the investigations in [26].

Recall that a multifunction $F : X \rightrightarrows Y$ is said to be *pseudo-Lipschitzian* at a point $(\bar{x}, \bar{y}) \in \operatorname{graph} F$ if there exist neighbourhoods $U \subseteq X$ of $\bar{x}$ and $V \subseteq Y$ of $\bar{y}$, and a positive constant L such that

$$F(x_1) \cap V \subseteq F(x_2) + L\|x_1 - x_2\|B, \tag{6}$$

whenever $x_1,\ x_2 \in U$. Here B stands for the closed unit ball in Y.
Let us recall that F is pseudo-Lipschitzian at $(\bar{x}, \bar{y}) \in \operatorname{graph} F$ if and only if the distance function $d(\cdot, F(\cdot))$ is locally Lipschitz at $(\bar{y}, \bar{x})$. This result is formulated in [21] for finite-dimensional spaces but the proof works in the infinite-dimensional case, too (see also [26]), and is omitted here.

Proposition 3.4 ([9]) *If $F : X \rightrightarrows Y$ is pseudo-Lipschitzian at a point $(\bar{x}, \bar{y}) \in \operatorname{graph} F$ and directionally differentiable in every direction at that point, then it is semi-differentiable at $(\bar{x}, \bar{y})$ and the semi-derivative coincide with the directional derivative.*

Proposition 3.5 ([9]) *Suppose that $F : X \rightrightarrows Y$ is pseudo-Lipschitzian around some point $(x, y) \in \operatorname{graph} F$ with a constant L and semi-differentiable at that point. Then the semiderivative at (x, y) is Lipschitzian multifunction of the direction with the same Lipschitz constant.*

4 Metric projections onto moving convex sets

First we cite some results concerning the continuity properties of the metric projection. This will play a key role for showing differentiability of the metric projection.
At this place, another topology on the hyperspace $\mathcal{F}(Y)$ becomes useful. Let m denotes the Mosco-topology on $\mathcal{F}(Y)$.
A sequence $\{A_n\}$ of weakly closed subsets of Y converges to a set $A \subseteq Y$ in $(\mathcal{F}(Y), m)$ if both of the following conditions are met:

(i) for each $y \in A$ there are points $y_n \in A_n$ such that $\{y_n\}$ converges in norm to y;

(ii) whenever points $y_k \in A_{n_k},\ k = 1, 2, \ldots$, are given, then the weak convergence of $\{y_n\}$ to $y \in Y$ implies $y \in A$.

This convergence is introduced by Mosco ([17]) as a modification of the sequential characterization of Kuratowski-Painlevé convergence (valid in a first countable space) in order to take into consideration the weak topology.

The following relation between the Mosco- and the Wijsman-topology is known:

Lemma 4.1 **([3])** *Let Y be a normed linear space. Then $\tau = m$ on $\mathcal{F}_c(Y)$ if and only if Y is reflexive and the dual space Y^* has the weak* Kadec-Klee property, i.e., weak* convergence in Y^* and convergence of the norms imply norm convergence in Y^*.*

We will use the following result:

Lemma 4.2 **([11])** *Let Y and its dual Y^* be E-spaces and A, $A_n \in \mathcal{F}_c(Y)$, $n = 1, 2, \ldots$. Then the following three statements are equivalent:*

(i) $m - \lim_{n\to\infty} A_n = A$

(ii) $\lim\limits_{n\to\infty} d(y, A_n) = d(y, A)$ *for all* $y \in Y$

(iii) $\lim\limits_{n\to\infty} \mathrm{Pr}(y, A_n) = \mathrm{Pr}(y, A)$ *with respect to the norm-convergence in Y.*

Moreover, it holds
$\lim\limits_{n\to\infty} \mathrm{Pr}(y_n, A_n) = \mathrm{Pr}(y, A)$*, whenever* $m - \lim_{n\to\infty} A_n = A$ *and* $\lim_{n\to\infty} y_n = y$.

It is also shown in [11] that the Mosco-topology is the weakest one that guarantees strong continuity of the metric projection onto moving convex sets.
Examples of E-spaces are Euclidean spaces, Hilbert spaces, uniformly convex Banach spaces (see e.g.[11]). Each reflexive space is potentially an E-space, since it can be renormed with an equivalent locally uniformly convex norm by a theorem of Troyanski (cf.[10], Theorem 1, p.164).
Let a multifunction F with nonempty closed convex images be given. We assume that Y is an E-space and consider the metric projection as a mapping $f_F : X \times Y \to Y$ defined by setting

$$f_F(x, y) = \mathrm{Pr}(y, F(x)).$$

The mapping f_F is well-defined because Y is an E-space and, thus, the projections exist and are unique. We shall use the following

Lemma 4.3 **([3],p.35)** *A sequence of closed sets $\{A_n\}, A_n \subseteq Y$ converges to some $A \subseteq Y$ in $(\mathcal{F}(Y), \tau)$ if and only if both of the following conditions are met:*

(1) *Whenever A meets a nonempty open subset V of X, then $A_n \cap V \neq \emptyset$ eventually;*

(2) *Denoting $U_r(y) = \{x : d(x, y) < r\}$ for r a positive number, whenever $0 < \varepsilon < \varrho$ and $U_\varrho(y) \cap A = \emptyset$, then $U_\varepsilon(y) \cap A_n = \emptyset$ eventually.*

Theorem 4.4 *Let both Y and Y^* be E-spaces, and F be a multifunction with nonempty closed convex images.*
$F : X \rightrightarrows Y$ is semi-differentiable at $(\bar{x}, \bar{y})$ if and only if the metric projection f_F is Hadamard-directionally differentiable at $(\bar{x}, \bar{y})$ in any direction $(h, z) \in X \times Y$ with derivative

$$f'_F(\bar{x}, \bar{y}; h, z) = \mathrm{Pr}(z, DF(\bar{x}, \bar{y})(h)).$$

Proof: Suppose that F is semi-differentiable. Let $(h,z) \in X \times Y$ be a given direction and $t_n \downarrow 0$, $(h_n, z_n) \to (h,z)$. We consider

$$\bar{z}_n = \Pr(z_n, \frac{F(\bar{x}+t_n h_n)-\bar{y}}{t_n})$$

and $\bar{z} = \Pr(z, DF(\bar{x},\bar{y})(h))$. By definition and invariance of the metric:

$$d(z_n,\bar{z}_n) = \inf_{y\in F(\bar{x}+t_n h_n)} d(z_n, \frac{y-\bar{y}}{t_n}) = \frac{1}{t_n} \inf_{y\in F(\bar{x}+t_n h_n)} d(\bar{y}+t_n z_n, y).$$

Therefore, denoting $y_n = \Pr(\bar{y}+t_n z_n, F(\bar{x}+t_n h_n))$, we have $y_n = \bar{y}+t_n\bar{z}_n$. Recall that the differential quotients τ-converge to the semiderivative and τ-convergence coincide with the Mosco-convergence in this case, by the convexity of the images and Lemma 4.1. Applying Lemma 4.2, we have $\bar{z}_n \to \bar{z}$ and, consequently,

$$\frac{y_n-\bar{y}}{t_n} \to \bar{z},$$

which proves the differentiability of f_F.

Now let us suppose, that f_F is Hadamard-directionally differentiable at (x,y), $y \in F(x)$ into any direction (z,h). We define the sets $C(h) := \{f'_F(y,x;z,h) : z \in Y\}$ for all $h \in X$, and intend to show that $C(h) = DF(x,y;h)$. We can write:

$$f_F(y+tz, x+th) = f_F(y,x) + tf'_F(y,x;z,h) + o(t),$$

which is equivalent to $\Pr(y+tz, F(x+th)) = y + tf'_F(y,x;z,h) + o(t)$. On the other hand $\Pr(y+tz, F(x+th)) = y + t\Pr(z, t^{-1}(F(x+th)-y))$. Consequently, $\Pr(z, t^{-1}(F(x+th)-y)) = f'_F(y,x;z,h) + o(t)/t$, which implies the convergence $\Pr(z,t^{-1}(F(x+th)-y)) \to f'_F(y,x;z,h)$ while $t \downarrow 0$. Assume that $t^{-1}(F(x+th)-y) \xrightarrow{W} C(h)$ does not hold. We shall use the characterization by Lemma 4.3. According to it, one of the conditions (1) or (2) is violated.

Case 1: The condition (1) is violated. Thus, there is a open subset V of Y such that $V \cap C(h) \neq \emptyset$ and $V \cap t_n^{-1}(F(x+t_n h_n)-y) = \emptyset$ for some subsequence $t_n \downarrow 0, h_n \to h$. Let $f'_F(y,x;\hat{z},h) \in V \cap C(h)$ for some $\hat{z} \in Y$. Since $\Pr(\hat{z}, t^{-1}(F(x+th)-y)) \to f'_F(y,x;\hat{z},h)$, and V is open, we obtain that $V \cap t_n^{-1}(F(x+th)-y) \neq \emptyset$.

Case 2: The condition (2) is violated. Thus, there are a positive number ϱ, and a sequence $\alpha_n \downarrow 0$ such that $U_\varrho(u) \cap C(y) = \emptyset$, and $U_{\alpha_n}(u) \cap t_n^{-1}(F(x+t_n h_n)-y) \neq \emptyset$. Therefore, there exists a sequence $u_n \in t_n^{-1}(F(x+t_n h_n)-y), u_n \to u$. Let $\bar{u_n} = \Pr(u, t^{-1}(F(x+th)-y))$. Obviously, $\|u-\bar{u_n}\| \leq \|u-u_n\|$, and the latter term converges to 0. By definition of $C(h)$, we have $u = f'_F(x,y;u,h) \in C(y)$, which is a contradiction to $U_\varrho(y) \cap C(y) = \emptyset$.

This proves that $t^{-1}(F(x+th)-y) \xrightarrow{W} C(h)$, i.e., $C(h) = DF(x,y;h)$, which proves the semi-differentiability of F because h was arbitrary. □

Theorem 4.5 *Let F be directionally differentiable of order k at $(\bar{x},\bar{y})$ into a direction $\mathbf{h}^k := (h^1,\dots,h^k)$ with respect to $\mathbf{z}^{k-1} := (z^1,\dots,z^{k-1})$, where $\mathbf{z}^{k-1})$ satisfies (4). Then the metric projection has directional derivative in sense of Ben-Tal and Zowe of order k at $(\bar{x},\bar{y})$ into the direction $(h^1,z^1,\dots,h^{k-1},z^{k-1},h^k,z)$, which is given by the formula:*

$$f_F^{(k)}(\bar{x},\bar{y};\mathbf{h}^{k-1},\mathbf{z}^{k-1})(h^k,z) = \Pr(z, D^kF(\bar{x},\bar{y};\mathbf{h}^{k-1},\mathbf{z}^{k-1})(h^k)) \text{ for all } z\in Y.$$

Moreover, $f_F^{(k)}(\bar{x},\bar{y};\mathbf{h}^{k-1},\mathbf{z}^{k-1})(\cdot,\cdot)$ is continuous.

Proof: We shall show the assertion for the case of second-order differentiability. The generalization is straightforward. Let $t_n \downarrow 0$, $(h_n^2, z_n) \to (h^2, z)$. We have to show that

$$\frac{\Pr(\bar{y}+t_nz^1+t_n^2z_n, F(\bar{x}+t_nh^1+t_n^2h_n^2)) - t_nf_F'(\bar{x},\bar{y};h^1,z^1)}{t_n^2}$$

converges to $\Pr(z, D^kF(\bar{x},\bar{y};\mathbf{h}^{k-1},\mathbf{z}^{k-1})(h^k))$. By the previous theorem $f_F'(\bar{x},\bar{y};h^1,z^1) = \Pr(z^1, DF(\bar{x},\bar{y})(h^1))$. According to (4), we have $z^1 \in DF(\bar{x},\bar{y})(h^1)$, and, consequently, $f_F'(\bar{x},\bar{y};h^1,z^1) = z^1$. We denote

$$w = \Pr(z, D^2F(\bar{x},\bar{y};h^1,z^1)(h^2)),$$

and

$$\bar{v}_n = \Pr(z_n, \frac{F(\bar{x}+t_nh^1+t_n^2h_n^2) - \bar{y} - t_nz^1}{t_n^2}).$$

Furthermore, we consider $p_n = \Pr(\bar{y}+t_nz^1+t_n^2z_n, F(\bar{x}+t_nh^1+t_n^2h_n^2))$. It holds by the same reasoning as in the previous theorem $p_n = \bar{y}+t_nz^1+t_n^2\bar{v}_n$. Using the Definition 2.3, Lemma 4.1 and the uniform continuity of the metric projection according to Lemma 4.2, we obtain

$$\frac{p_n - \bar{y} - t_nz^1}{t_n^2} = \bar{v}_n \to w.$$

This completes the proof of the differentiability. The continuity of the derivative follows by virtue of Lemma 4.2, Proposition 3.2, and Lemma 4.1. □

This result represents a generalization of the results on directional differentiability of metric projections established in [2, 25, 26] for finite-dimensional spaces. Moreover, it shows that the existence of a tangential approximation of F is not only sufficient (as in the above papers) but also a necessary condition. In some sense it is an extension of the results in [24] because on the one hand more general spaces are considered and on the other hand also higher-order differentiability is shown. However, no direct comparison can be made because [24] deals with projection onto non-convex sets.
Let us now consider the case of a fixed set, i.e., $F(x) = A$. In this case the directional derivative coincide with the tangent sets to A according to the definition in [2, 8]. Then the directional derivatives of the metric projection is given by the projection onto the tangent sets. For the case of $k = 1$, this fact is observed in [14, 25, 26] in

finite dimensional setting, and in [23, 27, 12] in Hilbert spaces. The existence of a k-order tangent set to A at a point y with respect to $z^1 \in T_A(y), z^2 \in T_A^2(y; z^1)$, *etc.* is equivalent to the existence of paths $y + tz^1 + t^2z^2 + \ldots + t^kz^k$ at the boundary of A. This corresponds to the main result in [15], which says that if a closed and convex set A in a Hilbert space has C^K-smooth boundary ($k \geq 2$), then the metric projection $\mathrm{Pr}(\cdot, A)$ is C^{k-1}. In [12] also the opposite (in some sense) is stated. The latter paper contains also interesting examples of closed convex sets such that the metric projection is nowhere Fréchet, or Gateaux differentiable.

References

[1] J.-P. Aubin and H. Frankowska, Set-Valued Analysis, Birkhäuser, Boston 1990.

[2] A. Auslender and R. Cominetti, A comparative study of multifunction differentiability with applications in mathematical programming, *Mathematics of Operations Research*, 16 (1991), 240-258.

[3] G. Beer, Topologies on Closed and Closed Convex Sets, Kluwer, Mathematics and its Applications, Vol.268, 1993

[4] G. Beer, Wijsman convergence: a survey, *Set-Valued Analysis* 2 (1994) 77-94 .

[5] G. Beer, A Polish topology for the closed subsets of a Polish space, *Proc. Amer. Math. Soc.* 113(1991), 1123-1133.

[6] G. Beer, Wijsman convergence of convex sets under renorming, *Nonlinear Analysis, Theory, Methods& Applications*, 22 (1994) 2, 207-216.

[7] A. Ben-Tal and J. Zowe, A unified theory for first- and second-order conditions for extremum problems in topological vector spaces, *Mathematical Programming Study* 19 (1982) 39-76

[8] R. Cominetti, Metric regularity, tangent sets and second-order optimality conditions, *Applied Mathematics and Optimization* 21 (1990), 265-287.

[9] D. Dentcheva, Approximations, extensions and univalued representations of multifunctions, Rutcor Research Report

[10] J. Diestel, Geometry of Banach spaces, Lecture Notes in Mathematics Vol.485, Springer-Verlag, Berlin, 1975.

[11] A. Dontchev and T. Zollezzi, Well-posed optimization problems, Lecture Notes in Mathematics Vol.1543, Springer-Verlag, Berlin, 1993.

[12] S. Fitzpatrick and R.R.Phelps, Differentiability of the metric projection in Hilbert space, *Trans. Amer. Math. Soc.* 270 (1982) 483-501.

[13] A. Haraux, How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities, *J. Math. Soc. Japan* 29 (1977) 615-631.

[14] J.B. Hiriart-Urruty, At what points is the projection mapping differentiable? *Amer. Math. Monthly* 89 (1982) 456-460.

[15] R.B. Holmes, Smoothness of certain metric projections on Hilbert space, *Trans. Amer. Math. Soc.* 184 (1973) 87-100.

[16] R.B. Holmes, Geometric functional analysis, Springer-Verlag, New York, 1975.

[17] U. Mosco, Convergence of convex sets and solutions of variational inequalities, *Adv. in Math.* 3 (1969), 510-585.

[18] K. Malanowski, Differentiability with respect to parameters of solutions to convex programming problems, *Math. Programming* 33 (1985) 352-361.

[19] J.-P. Penot, Differentiability of relations and differential stability of perturbed optimization problems, *SIAM J. Control and Optimization* 22 (1984) 4, 529-551

[20] R. T. Rockafellar, Proto-differentiability of set-valued mappings and its application in optimization, *in:* H. Attouch et al. (eds.), Analyse Non Lineaire, Gauthier-Villars, Paris, 1989, 449-482.

[21] R. T. Rockafellar, Lipschitzian properties of multifunctions, *J. Nonlinear Analysis: Theory, Mathods, Applications* 9 (1985) 867-885.

[22] A. Seeger, Second-order directional derivatives in parametric optimization, *Mathematics of Operations Research* 13 (1988) 124-139.

[23] I. Singer, Best approximation in normed linear spaces by elements of linear subspaces, Die Grundlehren der mathematischen Wissenschaften, Bd. 171, Springer-Verlag, 1970.

[24] A. Shapiro, Existence and differentiability of metric projections in Hilbert spaces, *SIAM J. Optimization* 4 (1994) 1, 130-141.

[25] A. Shapiro, On differentiability of metric projections in $\mathbb{R}^n$: Boundary Case, *Proc.Amer. Math. Sic.* 99, 123-128.

[26] A. Shapiro, Directional differentiability of metric projections onto moving sets at boundary points, *Journal of Mathematical Analysis and Applications* 13 (1988) 392-403.

[27] E.H.Zarantonello, Projections on convex sets in Hilbert space and spectral theory, *in* Contributions to Nonlinear Functional Analysis, 237-424, Academic Press, New York, 1971.