TAIWANESE JOURNAL OF MATHEMATICS
Vol. 10, No. 2, pp. 299-326, February 2006
This paper is available online at http://www.math.nthu.edu.tw/tjm/

SUBGRADIENTS OF DISTANCE FUNCTIONS AT OUT-OF-SET POINTS

Boris S. Mordukhovich and Nguyen Mau Nam

Abstract. This paper deals with the classical *distance function* to closed sets and its extension to the case of set-valued mappings. It has been well recognized that the distance functions play a crucial role in many aspects of variational analysis, optimization, and their applications. One of the most remarkable properties of even the classical distance function is its *intrinsic nonsmoothness*, which requires the usage of generalized differential constructions for its study and applications. In this paper we present new results in theser directions using mostly the generalized differential constructions introduced earlier by the first author, as well as their recent modifications. We pay the main attention to studying subgradients of the distance functions in *out-of-set* points, which is essentially more involved in comparison with the in-set case. Most of the results obtained are new in both finite-dimensional and infinite-dimensional settings; some of them of provide essential improvements of known results even for convex sets.

1. Introduction

This paper is devoted to the study of generalized differential properties of *distance functions*, which play a remarkable role in variational analysis, optimization, and their applications; see, e.g., the books [3, 15, 19] for more discussions and references. Since the *standard/classical distance function*

(1.1) $$d(x;\Omega) := \inf_{y\in\Omega} \|x-y\|, \quad x \in X,$$

is *not differentiable* (while always Lipschitz continuous on X) even for the simplest sets $\Omega \subset X$, tools of generalized differentiation are heavily needed for its study and

Received November 1, 2004; Accepted April 7, 2005.
2000 *Mathematics Subject Classification*: 49J52, 90C30.
Key words and phrases: Variational analysis and optimization, Distance functions, Generalized differentiation, Lipschitzian stability.
Research was partially supported by the National Science Foundation under grant DMS-0304989 and by the Australian Research Council under grant DP-0451168.

applications. We refer the reader to [2, 6-9, 11, 13, 15, 16, 19, 21] among other publications devoted to computing and estimating various subgradient sets for the classical distance function (1.1) in finite and infinite dimensions. Note that there are two principal and essentially different cases for generalized differentiation of (1.1): the *in-set* case of $\bar{x} \in \Omega$ and the *out-of-set* case of $\bar{x} \notin \Omega$. The latter case is much more involved and less investigated.

In this paper we consider, along with the standard distance function (1.1), it extension

$$\rho(z,x) := \inf_{y \in F(z)} \|x - y\| = d(F(z);x) \tag{1.2}$$

built upon the *generating* set-valued mapping $F\colon Z \rightrightarrows X$. The latter function, called the *general distance function* in what follows, may be essentially more complicated than (1.1). In particular, it is not generally Lipschitz continuous and even lower semicontinuous (l.s.c.) around given/reference points. Some generalized differential properties of (1.2) were studied in [4, 5, 16, 20]. Again, there are two principal settings for studying local properties of (1.2): the *in-set* case of $(\bar{z},\bar{x}) \in \operatorname{gph} F$ and the *out-of-set* case of $(\bar{z},\bar{x}) \notin \operatorname{gph} F$. To the best of our knowledge, the latter case has been investigated only in the papers [4, 16].

The present paper can be considered as a continuation and development of our previous one [16] being, in contrast to [16], entirely devoted to the out-of-set case for both the general and standard distance functions. New developments concern, first of all, the involvement of *intermediate points* between the reference and projection ones into upper subgradient estimates; see below. This brings us to new results even for *convex* sets in finite dimensions. We establish also new relationships between *singular subgradients* and *mixed coderivatives* of *marginal/value functions* that are directly applied to the general distance function (1.2) in the non-Lipschitzian case. Moreover, we extend the class of subdifferentials under consideration in comparison with [16] and obtain new applications to the projection nonemptiness and Lipschitz stability.

The rest of the paper is organized as follows. In Section 2 we briefly discuss some preliminary material needed in what follows. Section 3 collects new *upper estimates* for various subgradients of the distance functions (1.1) and (1.2) involving *intermediate* points. In Section 4 we present upper estimates for the new type of *right-sided limiting subgradients* for both distance functions under consideration. Section 5 is devoted to establishing relationships between singular subgradients of *marginal functions*, including the general distance function (1.2), and mixed coderivatives of the generating set-valued mappings. Finally, Section 6 contains some new applications of the main results obtained in the paper.

Unless otherwise stated, all the spaces under consideration are Banach, with X^* denoting the dual space of X. As usual, $I\!B$ and $I\!B^*$ stand for the closed unit balls of the space in question and its dual, while S and S^* denote the corresponding unit spheres. The notation $B(\bar{x};\delta) := \bar{x} + \delta I\!B$ stands for the closed ball centered at $\bar{x}$ with radius δ. Note also that $I\!N := \{1, 2, \ldots\}$ and that the convention $0 \cdot \emptyset = 0$ is used in what follows.

2. Preliminaries

This section contains some preliminary material, which is widely used in the main body of the paper. The reader can find more details and references in the books by Rockafellar and Wets [19] in finite dimensions and by Mordukhovich [15] in both finite-dimensional and infinite-dimensional spaces.

Given $\Omega \subset X$ and $\varepsilon \geq 0$, define the (Fréchet-like) *ε-normals* to Ω at $\bar{x} \in \Omega$ by

$$\widehat{N}_\varepsilon(\bar{x};\Omega) := \Big\{ x^* \in X^* \Big| \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \leq \varepsilon \Big\}, \tag{2.1}$$

where $x \xrightarrow{\Omega} \bar{x}$ means that $x \to \bar{x}$ with $x \in \Omega$. When $\varepsilon = 0$, the set $\widehat{N}_0(\bar{x};\Omega)$ in (2.1) is a cone called the *Fréchet normal cone* and denoted by $\widehat{N}(\bar{x};\Omega)$.

The *basic/limiting normal cone* $N(\bar{x};\Omega)$ is obtained from $\widehat{N}_\varepsilon(x;\Omega)$ by taking the *sequential* Painlevé-Kuratowski upper (or outer) limit in the weak* topology w^* of X^* as

$$N(\bar{x};\Omega) := \mathop{\mathrm{Lim\,sup}}_{\substack{x \xrightarrow{\Omega} \bar{x} \\ \varepsilon \downarrow 0}} \widehat{N}_\varepsilon(x;\Omega). \tag{2.2}$$

One can equivalently put $\varepsilon = 0$ in (2.2) when Ω is closed around $\bar{x}$ and when the space X is *Asplund*, i.e., a Banach space whose separable subspaces have separable duals. This class of spaces is sufficiently large including, in particular, every reflexive space; see, e.g., [18] for more information. The cone of *proximal normals* is defined by

$$N_P(\bar{x};\Omega) := \Big\{ x^* \in X^* \Big| \exists \delta > 0,\ \eta > 0 \ \text{ such that } \ \langle x^*, x - \bar{x}\rangle \leq \eta\|x - \bar{x}\|^2 \ \text{for all } \ x \in B(\bar{x};\delta) \Big\}.$$

If the space X is *Hilbert*, then the basic normal cone (2.2) can be equivalently obtained as the the sequential Painleveé-Kuratowski limit of proximal normals instead of Fréchet ones in (2.1) with $\varepsilon = 0$:

$$N(\bar{x};\Omega) = \mathop{\mathrm{Lim\,sup}}_{x \xrightarrow{\Omega} \bar{x}} N_P(x;\Omega),$$

which reduces to the normal cone introduced by Mordukhovich [12] in finite dimensions.

Let $f\colon X \to \overline{I\!R} := [-\infty, \infty]$ be an extended-real-valued function finite at $\bar{x}$. The set

$$\widehat{\partial}_\varepsilon f(\bar{x}) := \Big\{x^* \in X^* \Big|\ \liminf_{x\to\bar{x}} \frac{f(x)-f(\bar{x})-\langle x^*, x-\bar{x}\rangle}{\|x-\bar{x}\|} \ge -\varepsilon\Big\} \tag{2.3}$$

is called the *ε-subdifferential* of f at $\bar{x}$. If $\varepsilon = 0$, then $\widehat{\partial}_0 f(\bar{x})$ is said to be the *Fréchet subdifferential* of f at $\bar{x}$ and is denoted by $\widehat{\partial} f(\bar{x})$. Similarly to the case of normals, the *proximal subdifferential* of f at $\bar{x}$ is defined by

$$\begin{aligned}\partial_P f(\bar{x}) : = \big\{x^* \in X^* \big|\ &\exists\delta > 0,\ \eta > 0 \ \text{ such that } \ \langle x^*, x-\bar{x}\rangle \\ &\le f(x) - f(\bar{x}) + \eta\|x-\bar{x}\|^2 \ \text{ for all } \ x \in B(\bar{x};\delta)\Big\}.\end{aligned}$$

The *basic/limiting subgradient* of f at $\bar{x}$ is defined by

$$\partial f(\bar{x}) := \mathop{\mathrm{Lim\,sup}}_{\substack{x \xrightarrow{f} \bar{x} \\ \varepsilon\downarrow 0}} \widehat{\partial}_\varepsilon f(x), \tag{2.4}$$

where $x \xrightarrow{f} \bar{x}$ means that $x \to \bar{x}$ and $f(x) \to f(\bar{x})$. Note that $\widehat{\partial}_\varepsilon f(x)$ can be replaced by $\widehat{\partial}\varphi(x) := \widehat{\partial}_0 f(x)$ in (2.4) when X is Asplund while f is lower semicontinuous around $\bar{x}$. Moreover, one can equivalently use the proximal subdifferential under the "$\mathrm{Lim\,sup}$" in (2.4) if X is Hilbert. Let us mention the geometric representation of the basic subdifferential:

$$\partial f(\bar{x}) = \big\{x^* \in X^* \big|\ (x^*, -1) \in N((\bar{x}, f(\bar{x})); \mathrm{epi}\, f)\big\}$$

via the epigraph $\mathrm{epi}\, f := \{(x,\mu) \in X \times I\!R |\ \mu \ge f(x)\}$ of f. It follows from (2.4) that

$$\widehat{\partial} f(\bar{x}) \subset \partial f(\bar{x}), \tag{2.5}$$

while the equality in (2.5) defines the class of *lower regular* functions [13, 15], which particularly includes the case of Clarke regularity as defined in [19].

Recall the *singular subdifferential* construction for $f\colon X \to \overline{I\!R}$ at $\bar{x}$ defined by

$$\partial^\infty f(\bar{x}) = \mathop{\mathrm{Lim\,sup}}_{\substack{x \xrightarrow{f} \bar{x} \\ \varepsilon,\lambda\downarrow 0}} \lambda\widehat{\partial}_\varepsilon f(x). \tag{2.6}$$

This construction makes sense only for non-Lipschitzian functions, since $\partial^\infty f(\bar{x}) = \{0\}$ if f is Lipschitz continuous around $\bar{x}$. Note that $\varepsilon > 0$ can be equivalently omitted in (2.6) if X is Asplund and f is l.s.c. around $\bar{x}$. Observe that in the latter case the singular subdifferential (2.6) admits the equivalent geometric representation

$$\partial^{\infty} f(\bar{x}) := \big\{x^* \in X^* \big| \; (x^*, 0) \in N((\bar{x}, f(\bar{x})); \mathrm{epi}\, f)\big\}. \tag{2.7}$$

We also need to recall some directional derivative/subderivative constructions used in what follows. The *Rockafellar subderivative* of $f\colon X \to \overline{I\!R}$ at $\bar{x}$ is defined by

$$f^{\uparrow}(\bar{x}; h) := \sup_{\delta>0} \Big[\limsup_{(x,\alpha)\downarrow^f \bar{x}, t\downarrow 0} \Big(\inf_{h' \in B(\bar{x};\delta)} \frac{f(x+th') - \alpha}{t} \Big) \Big],$$

where $(x,\alpha) \downarrow^f \bar{x}$ means that $(x,\alpha) \in \mathrm{epi} f, (x,\alpha) \to (\bar{x}, f(\bar{x}))$. If f is l.s.c. around $\bar{x}$, then

$$f^{\uparrow}(\bar{x}; h) := \sup_{\delta>0} \Big[\limsup_{x \xrightarrow{f} \bar{x}, t\downarrow 0} \Big(\inf_{h' \in B(\bar{x};\delta)} \frac{f(x+th') - f(x)}{t} \Big) \Big],$$

where $x \xrightarrow{f} \bar{x}$ stands for $f(x) \to f(\bar{x})$ with $x \to \bar{x}$. Moreover, when f is locally Lipschitzian around $\bar{x}$, the subderivative $f^{\uparrow}(\bar{x}; h)$ agrees with the *Clarke directional derivative*

$$f^{\circ}(\bar{x}; h) := \limsup_{x\to\bar{x}, t\downarrow 0} \frac{f(x+th) - f(x)}{t}.$$

Finally, the *Dini-Hadamard directional derivative* of f at $\bar{x}$ is given by

$$f^{-}(\bar{x}; h) := \liminf_{h\to h, t\downarrow 0} \frac{f(\bar{x}+th') - f(\bar{x})}{t},$$

which is simplified by

$$f^{-}(\bar{x}; h) := \liminf_{t\downarrow 0} \frac{f(\bar{x}+th) - f(\bar{x})}{t}$$

when f is locally Lipschitz around $\bar{x}$. The corresponding *Clarke and Dini-Hadamard subdifferentials* of f at $\bar{x}$ are defined by

$$\partial_C f(\bar{x}) := \Big\{x^* \in X^* \Big| \langle x^*, h\rangle \le f^{\uparrow}(\bar{x}; h), \text{for all } h \in X\Big\},$$

$$\partial^- f(\bar{x}) := \Big\{x^* \in X^* \Big| \langle x^*, h\rangle \le f^{-}(\bar{x}; h), \text{for all } h \in X\Big\}.$$

We say that f is *directionally regular* at $\bar{x}$ if $f^{-}(\bar{x}; h) = f^{\uparrow}(\bar{x}; h)$ for all $h \in X$, which implies that $\partial_C f(\bar{x}) = \partial^- f(\bar{x})$.

Considering a set-valued mapping $F\colon X \rightrightarrows Y$ between Banach spaces with the graph

$$\mathrm{gph}\, F := \big\{(x,y) \in X \times Y \big|\ y \in F(x)\big\},$$

we recall its *normal coderivative* $D^*_N F(\bar{x},\bar{y})\colon X^* \rightrightarrows Y^*$ and *mixed coderivative* $D^*_M F(\bar{x},\bar{y})$ at $(\bar{x},\bar{y}) \in \mathrm{gph}\, F$ defined respectively by

$$D^*_N F(\bar{x},\bar{y})(y^*) := \big\{x^* \in X^* \big|\ (x^*,-y^*) \in N((\bar{x},\bar{y}); \mathrm{gph}\, F)\big\}, \tag{2.8}$$

$$D^*_M F(\bar{x},\bar{y})(y^*) := \Big\{x^* \in X^* \Big| \exists \varepsilon_k \downarrow 0,\ (x_k,y_k) \stackrel{\mathrm{gph}\, F}{\to} (\bar{x},\bar{y}),\ x^*_k \stackrel{w^*}{\longrightarrow} x^*, \\ y^*_k \stackrel{\|\cdot\|}{\longrightarrow} y^* \ \text{ with } \ (x^*_k,-y^*_k) \in \widehat{N}_{\varepsilon_k}((x_k,y_k); \mathrm{gph}\, F)\Big\}, \tag{2.9}$$

where $\stackrel{w^*}{\to}$ signifies the weak* *sequential* convergence in X^*, while $\stackrel{\|\cdot\|}{\to}$ stands for the norm convergence in the dual space; we omit $\|\cdot\|$ in the latter. We can equivalently put $\varepsilon_k = 0$ in (2.9) if X and Y are Asplund and if the graph of F is closed around $(\bar{x},\bar{y})$. Clearly $D^*_M F(\bar{x},\bar{y})(y^*) \subset D^*_N F(\bar{x},\bar{y})(y^*)$, where the equality holds if $\dim Y < \infty$ and in more general settings of "strong coderivative normality" listed in [14, Proposition 3.2] and [15, Proposition 4.9]. Observe that the basic and singular subdifferentials in (2.4) and (2.7) can be described as

$$\partial f(\bar{x}) = D^* E_f(\bar{x}, f(\bar{x}))(1) \ \text{ and } \ \partial^\infty f(\bar{x}) = D^* E_f(\bar{x}, f(\bar{x}))(0)$$

via the coderivative of the epigraphical multifunction $E_f\colon X \rightrightarrows I\!R$ associated with f by $E_f(x) := \{\mu \in I\!R |\ \mu \ge f(x)\}$.

One of the most fundamental differences between variational analysis in finite and infinite dimensions, crucial for many aspects of generalized differentiation and optimization, is the necessary of imposing additional compactness requirements in infinite-dimensional spaces that ensure the *nontriviality* while passing to the limit in the weak* topology. In this paper we use the following general properties that are automatic in finite dimensions, hold for "reasonably good" sets and mappings, and are preserved under various operations; see [15] for the comprehensive theory and applications.

A set Ω is *sequentially normally compact* (SNC) at $\bar{x}$ if for any sequences $\varepsilon_k \downarrow 0$, $x_k \stackrel{\Omega}{\to} \bar{x}$ and $x^*_k \in \widehat{N}_{\varepsilon_k}(x_k;\Omega)$ one has

$$\big[x^*_k \stackrel{w^*}{\to} 0\big] \Longrightarrow \big[\|x^*_k\| \to 0\big] \ \text{ as } \ k \to \infty, \tag{2.10}$$

where ε_k can be omitted if X is Asplund and if Ω is locally closed around $\bar{x}$. The SNC condition is automatic when Ω satisfies the so-called "compactly epi-Lipschitzian" property in the sense of Borwein and Strojwas, particularly when it is convex and finite-codimensional with nonempty relative interior; see [15] for more details. We say that a set $\Omega \subset X \times Y$ is *SNC with respect to* X at $(\bar{x}, \bar{y}) \in \Omega$ if (2.10) holds for any sequences $\varepsilon_k \downarrow 0$, $(x_k, y_k) \stackrel{\Omega}{\to} (\bar{x}, \bar{y})$, and $(x_k^*, y_k^*) \in \widehat{N}_{\varepsilon_k}((x_k, y_k); \Omega)$ as $k \in I\!N$.

A set-valued mapping $F\colon X \rightrightarrows Y$ is *SNC* at $(\bar{x}, \bar{y}) \in \text{gph}\, F$ if its graph enjoys this property. For the case of mappings a more subtle *partial SNC* (i.e., *PSNC*) property can be defined. We say that F is *PSNC* at $(\bar{x}, \bar{y})$ if for any sequences $\varepsilon_k \downarrow 0$, $(x_k, y_k) \stackrel{\text{gph}\, F}{\to} (\bar{x}, \bar{y})$ and $(x_k^*, y_k^*) \in \widehat{N}_{\varepsilon_k}((x_k, y_k); \text{gph}\, F)$ one has

$$\big[x_k^* \stackrel{w^*}{\to} 0,\ \|y_k^*\| \to 0\big] \Longrightarrow \big[\|x_k^*\| \to 0\big] \ \text{ as } \ k \to \infty,$$

where $\varepsilon_k = 0$ in the Asplund space and closed graph setting. The PSNC property always holds when F is *Lipschitz-like* around $(\bar{z}, \bar{x})$ in the following sense of Aubin [1]: there exist neighborhoods V of $\bar{z}$ and W of $\bar{x}$ as well as modulus $\ell \geq 0$ such that

$$F(u) \cap W \subset F(v) + \ell \|u - v\| I\!B \ \text{ whenever } \ u, v \in V. \tag{2.11}$$

This reduces to the classical (Hausdorff) *local Lipschitzian* behavior of F around $\bar{z}$ for $W = X$ in (2.11). The Lipschitz-like property of F is known to be *equivalent* to the *metric regularity* and *linear openness* properties of the *inverse* mapping F^{-1}; these three equivalent properties play a fundamental role in many aspects of nonlinear analysis especially those related to optimization; see [3, 15, 19] and the references therein.

Finally in these preliminaries, let us mention a version of the SNC property for extended-real-valued functions $f\colon X \to \overline{I\!R}$ finite at $\bar{x}$. Namely, f is *sequentially normally epi-compact* (SNEC) at $\bar{x}$ if its epigraph is SNC at $(\bar{x}, f(\bar{x}))$. This property always holds for locally Lipschitzian functions and their appropriate extensions.

3. Upper Estimates for Various Subdifferentials of Distance Functions

In this section we derive some *upper estimates* of all the subdifferentials defined in Section 2 for the *general distance function* $\rho(z, x) = d(F(z); x)$ and its *standard* specification $d(x; \Omega)$ at *out-of-set* points. The main new feature of the results obtained is that they involve *all intermediate points* between a given out-of set point and its projections on the set. This allows us to essentially improve known results even for convex subsets of finite-dimensions.

We start with the following statement, which can be easily derived from the result by Bounkhel [4, Proposition 3.2].

Lemma 3.1. (Projections in Banach space) *Let $F\colon Z \rightrightarrows X$ be a set-valued mapping between Banach spaces with $(\bar z,\bar x) \notin \mathrm{gph}\,F$ satisfying $\Pi(\bar x;F(\bar z)) \neq \emptyset$. Then for any $t \in (0,1]$ and $\bar\omega \in \Pi(\bar x;F(\bar z))$ one has*

$$\bar\omega \in \Pi(t\bar\omega + (1-t)\bar x; F(\bar z)).$$

This gives, in particular, that

$$\bar\omega \in \Pi(t\bar\omega + (1-t)\bar x;\Omega) \ \text{ whenever } \ t \in (0,1]$$

for any $\Omega \subset X$, any $\bar x \notin \Omega$ with $\Pi(\bar x;\Omega) \neq \emptyset$, and any $\bar\omega \in \Pi(\bar x;\Omega)$.

The next proposition establishes useful upper estimates for ε-subgradients of the distance functions at the reference points via those at *intermediate* points. Note that the upper estimate in this proposition and the subsequent results involve intermediate points corresponding to *every* $t \in (0,1]$ in what follows.

Proposition 3.2. (Upper estimates for ε-subgradients of distance functions via intermediate points) *Let $F\colon Z \rightrightarrows X$ be a set-valued mapping between Banach spaces. Assume that $(\bar z,\bar x) \notin \mathrm{gph}\,F$ and that $\Pi(\bar x;F(\bar z)) \neq \emptyset$. Then for any $t \in (0,1]$ we have the inclusion*

$$\widehat\partial_\varepsilon \rho(\bar z,\bar x) \subset \bigcap_{\bar\omega \in \Pi(\bar x;F(\bar z))} \widehat\partial_\varepsilon \rho(\bar z, t\bar\omega + (1-t)\bar x) \cap C^*_\varepsilon, \quad \varepsilon \geq 0, \tag{3.1}$$

where $C^*_\varepsilon := \{(z^*,x^*) \in Z^* \times X^* \big| \ 1-\varepsilon \leq ||x^*|| \leq 1+\varepsilon\}$. In particular,

$$\widehat\partial_\varepsilon d(\bar x;\Omega) \subset \bigcap_{\bar\omega \in \Pi(\bar x;\Omega)} \widehat\partial_\varepsilon d(t\bar\omega + (1-t)\bar x;\Omega) \cap [1-\varepsilon, 1+\varepsilon]S^* \tag{3.2}$$

for any $\Omega \subset X$ and $\bar x \notin \Omega$ satisfying $\Pi(\bar x;\Omega) \neq \emptyset$.

Proof. To justify (3.1), take an arbitrary $(z^*,x^*) \in \widehat\partial_\varepsilon \rho(\bar z,\bar x)$ and $\bar\omega \in \Pi(\bar x;F(\bar z))$ with any $t \in (0,1]$. Putting $\widetilde x := \bar x + t(\bar\omega - \bar x)$, one gets by Lemma 3.1 that

$$\bar\omega \in \Pi(\widetilde x;F(\bar z)) \ \text{ and } \ \rho(\bar z;\widetilde x) = (1-t)\|\bar x - \bar\omega\| = (1-t)\rho(\bar z,\bar x).$$

Given $\eta > 0$ find $\delta > 0$ by Definition (2.4) of ε-subgradients such that

$$\langle z^*, z-\bar z\rangle + \langle x^*, x-\bar x\rangle \leq \rho(z,x) - \rho(\bar z,\bar x) + (\varepsilon+\eta)(\|z-\bar z\| + \|x-\bar x\|)$$

whenever $\|z-\bar z\| < \delta$ and $\|x-\bar x\| < \delta$. Then for any $(z,x) \in Z \times X$ such that $||z-\bar z|| < \delta$ and $||x-\widetilde x|| = ||(x-\widetilde x+\bar x)-\bar x|| < \delta$, one has

$$\begin{aligned}\langle z^*, z-\bar z\rangle+\langle x^*, x-\widetilde x\rangle &\le \rho(z;x-\widetilde x+\bar x)-\rho(\bar z,\bar x)+(\varepsilon+\eta)(\|z-\bar z\|+\|x-\widetilde x\|)\\ &\le \rho(z,x)+\|\widetilde x-\bar x\|-\|\bar x-\bar\omega\|+(\varepsilon+\eta)(\|z-\bar z\|+\|x-\widetilde x\|)\\ &\le \rho(z,x)+t\|\bar\omega-\bar x\|-\|\bar x-\bar\omega\|+(\varepsilon+\eta)(\|z-\bar z\|+\|x-\widetilde x\|)\\ &= \rho(z,x)-(1-t)\|\bar\omega-\bar x\|+(\varepsilon+\eta)(\|z-\bar z\|+\|x-\widetilde x\|)\\ &= \rho(z,x)-\rho(\bar z,\widetilde x)+(\varepsilon+\eta)(\|z-\bar z\|+\|x-\widetilde x\|)\end{aligned}$$

This gives $(z^*,x^*)\in\widehat\partial_\varepsilon\rho(\bar z,\widetilde x)$. Since $x^*\in\widehat\partial_\varepsilon d(\bar x;\Omega)$ for $\Omega:=F(\bar z)$ with $\bar x\notin\Omega$, we get by [10, Proposition 1.5] that

$$1-\varepsilon\le\|x^*\|\le 1+\varepsilon,$$

which completes the proof of (3.1). Inclusion (3.2) is an obvious specification of (3.1) for $F(\cdot)=\Omega\subset X$. ■

It happens that counterparts of the upper estimates (3.1) and (3.2) from Proposition 3.2 hold not only for Fréchet subgradients but also for proximal and Dini-Hadamard subgradients from Section 2.

Theorem 3.3. (Upper estimates for Fréchet, proximal, and Dini-Hadamard subgradients of distance functions via intermediate points) *Let in the setting of Proposition* 3.2 *the symbol* $\partial^\bullet$ *stand for one of the following subdifferentials: Fréchet, proximal, and Dini-Hadamard. Then one has*

$$\partial\rho^\bullet(\bar z,\bar x)\subset\bigcap_{\bar\omega\in\Pi(\bar x;F(\bar z))}\partial^\bullet\rho(\bar z,t\bar\omega+(1-t)\bar x)\cap C^*, \tag{3.3}$$

where $C^*:=\{(z^*,x^*)\in Z^*\times X^*\big|\ ||x^*||=1\}$. *In particular,*

$$\partial^\bullet d(\bar x;\Omega)\subset\bigcap_{\bar\omega\in\Pi(\bar x;\Omega)}\partial^\bullet d(t\bar\omega+(1-t)\bar x;\Omega)\cap S^*. \tag{3.4}$$

Proof. We need to justify (3.3) only for $\partial^\bullet=\partial^-$, since it has been proved in Proposition 3.2 for Fréchet subgradients as $\varepsilon=0$ and can be derived by the same arguments for the case of proximal subgradients.

To proceed for $\partial^\bullet\rho=\partial^-\rho$, fix any $\bar\omega\in\Pi(\bar x;F(\bar z))$ and $t\in(0,1]$. First consider the Dini-Hadamard directional derivative and show that

$$\begin{gathered}\rho^-((\bar z,\bar x);(\bar h,\bar k))\le\rho^-((\bar z,t\bar\omega+(1-t)\bar x);(\bar h,\bar k))\\ \text{whenever }\ (\bar h,\bar k)\in Z\times X.\end{gathered} \tag{3.5}$$

Suppose without loss of generality that $\rho^-((\bar z,\bar x);(\bar h,\bar k))>-\infty$, since otherwise one obviously has $\partial^-\rho(\bar z,\bar x)=\emptyset$ and (3.3) holds trivially. Putting $\alpha:=$

$\rho^-((\bar{z},\bar{x});(\bar{h},\bar{k}))$ and taking into account that $\rho(\bar{z},\cdot)$ is Lipschitz continuous, we get

$$\rho^-((\bar{z},\bar{x});(\bar{h},\bar{k})) = \liminf_{h\to\bar{h},\lambda\downarrow 0} \frac{\rho(\bar{z}+\lambda h,\bar{x}+\lambda\bar{k})-\rho(\bar{z},\bar{x})}{\lambda}.$$

Then for any $\epsilon>0$, find $\delta>0$ such that

$$\rho(\bar{z}+\lambda h,\bar{x}+\lambda\bar{k})-\rho(\bar{z},\bar{x})\ge\lambda(\alpha-\epsilon) \ \text{ whenever } \ ||h-\bar{h}||<\delta,\ 0<\lambda<\delta.$$

Furthermore, for such h and λ one has

$$\begin{aligned}\rho(\bar{z}+\lambda h,\widetilde{x}+\lambda\bar{k})-\rho(\bar{z},\widetilde{x}) &= \rho(\bar{z}+\lambda h,\bar{x}+t(\bar{\omega}-\bar{x})+\lambda\bar{k})-(1-t)\rho(\bar{z},\bar{x})\\ &\ge \rho(\bar{z}+\lambda h,\bar{x}+\lambda\bar{k})-t||\bar{\omega}-\bar{x}||-(1-t)\rho(\bar{z},\bar{x})\\ &= \rho(\bar{z}+\lambda h,\bar{x}+\lambda\bar{k})-\rho(\bar{z},\bar{x})\\ &\ge \lambda(\alpha-\epsilon).\end{aligned}$$

Since $\epsilon>0$ is arbitrary, this gives (3.5) while dividing by $\lambda>0$ and taking "$\liminf$" in both sides of the latter inequality. To complete the proof of (3.3), pick any $(z^*,x^*)\in\partial^-\rho(\bar{z},\bar{x})$ and get the estimates

$$\langle(z^*,x^*),(\bar{h},\bar{k})\rangle\le\rho^-((\bar{z},\bar{x});(\bar{h},\bar{k}))\le\rho^-((\bar{z},t\bar{\omega}+(1-t)\bar{x});(\bar{h},\bar{k}))$$
$$\text{for all } \ (\bar{h},\bar{k})\in Z\times X.$$

Thus, by definition of Dini-Hadamard subgradients, we arrive at the inclusion

$$\partial^-\rho(\bar{z},\bar{x})\subset\partial^-\rho(\bar{z},t\bar{\omega}+(1-t)\bar{x}).$$

The last part $(z^*,x^*)\in C^*$ follows from [4, Proposition 3.2]. ■

It is easy to observe the following consequence of Theorem 3,3 involving subgradient estimates for distance functions via corresponding *normals* at intermediate points.

Corollary 1.1. (upper estimates for Fréchet, proximal, and Dini-Hadamard subgradients of distance functions via normals at intermediate points) *Let in the setting of Proposition* 3.2 *the symbols* $\partial^\bullet$ *and* $N^\bullet$ *stand for the Fréchet subdifferential and normal cone as well as for proximal subdifferential and normal cone, respectively. Given any* $t\in(0,1]$ *and* $\bar{\omega}\in\Pi(\bar{x};F(\bar{z}))$, *consider the set-valued mapping*

$$F_{t,\bar{\omega}}(z):=\big\{x\in X\big|\ d(F(z);x)\le t_{\bar{\omega}}\big\} \ \textit{ with } \ t_{\bar{\omega}}:=d(F(\bar{z});t\bar{\omega}+(1-t)\bar{x}).$$

Then one has the inclusion

$$\text{(3.6)} \qquad \partial^\bullet\rho(\bar{z},\bar{x})\subset\bigcap_{\bar{\omega}\in\Pi(\bar{x};F(\bar{z}))}N^\bullet((\bar{z},t\bar{\omega}+(1-t)\bar{x});\operatorname{gph}F_{t,\bar{\omega}})\cap C^*,$$

where $C^* := \{(z^*, x^*) \in Z^* \times X^* \big| \, ||x^*|| = 1\}$. *In particular,*

$$\partial^\bullet d(\bar{x}; \Omega) \subset \bigcap_{\bar{\omega} \in \Pi(\bar{x}; \Omega)} N^\bullet(t\bar{\omega} + (1-t)\bar{x}; \Omega_{t,\bar{\omega}}) \cap S^*,$$

where $\Omega_{t,\bar{\omega}} := \{x \in X \big| \, d(x; \Omega) \le t_{\bar{\omega}}\}$ *with* $t_{\bar{\omega}} := d(t\bar{\omega} + (1-t)\bar{x}; \Omega)$.

Proof. We need only proving (3.6). Fix arbitrary $\bar{\omega} \in \Pi(\bar{x}; F(\bar{z}))$ and $t \in (0, 1]$. Put $\widetilde{x} := t\bar{\omega} + (1-t)\bar{x}$. It follows from Theorem 3.3 that $\partial^\bullet \rho(\bar{z}, \bar{x}) \subset \partial^\bullet \rho(\bar{z}, \widetilde{x}) \cap C^*$. Using the definition of Fréchet (resp. proximal) subdifferential, we have

$$\widehat{\partial}\rho(\bar{z}, \widetilde{x}) \subset \widehat{N}((\bar{z}, \widetilde{x}); \mathrm{gph}\, F_{t,\bar{\omega}}).$$

This directly implies (3.6) due to (3.3). ∎

Observe that for $t = 1$ one obviously has

$$F_{t,\bar{\omega}} \equiv F \quad \text{and} \quad \Omega_{t,\bar{\omega}} \equiv \Omega$$

provided that F is closed-graph and that Ω is closed. Thus

$$\partial^\bullet \rho(\bar{z}, \bar{x}) \subset \textstyle\bigcap_{\bar{\omega} \in \Pi(\bar{x}; F(\bar{z}))} N^\bullet((\bar{z}, \bar{\omega}); \mathrm{gph}\, F) \cap C^* \quad \text{and} \quad \partial^\bullet d(\bar{x}; \Omega) \subset \textstyle\bigcap_{\bar{\omega} \in \Pi(\bar{x}; \Omega)} N^\bullet(\bar{\omega}; \Omega) \cap S^*.$$

It immediately follows from Theorem 3.3 that estimates (3.3) and (3.4) therein hold for the basic/limiting and Clarke subdifferentials provided that the corresponding *lower regularity* and *directional regularity* assumptions from Section 2 are fulfilled. However, such regularity assumptions for distance functions are very restrictive at *out-of-set* points. In particular, for the standard distance function in finite dimensions they are equivalent to its *smoothness*; see [15, Subsection 1.3.3].

The following natural question arises. Would it be possible to keep inclusions (3.3) and (3.4) with $\partial^\bullet$ stands for either the limiting subdifferential or for the Clarke subdifferential with no regularity assumptions? The answer happens to be *no* even for the *standard* distance functions in *finite-dimensional* spaces.

Example 3.5. (Failure of the intermediate subdifferential estimates for limiting and Clarke subgradients). There is a closed subset of $I\!R^2$ such that the inclusion

$$\partial^\bullet d(\bar{x}; \Omega) \subset \bigcap_{\bar{\omega} \in \Pi(\bar{x}; \Omega)} \partial^\bullet d(t\bar{\omega} + (1-t)\bar{x}; \Omega) \cap S^*$$

does not hold for some $\bar{x} \notin \Omega$ and $t \in (0, 1]$, where $\partial^\bullet$ stands for either the limiting subdifferential or for the Clarke subdifferential of the distance function.

Proof. Consider the set

$$\Omega := \big\{(x,y) \in I\!R^2 \big|\ x^2 + y^2 \ge 1\big\}$$

and the point $\bar{x} = (0,0) \notin \Omega$. Then

$$\partial d(\bar{x};\Omega) = S \ \text{ and } \ \partial_C d(\bar{x};\Omega) = I\!B,$$

while for $\bar{\omega} = (1,0)$ and $t = 1/2$ we have

$$\partial d(t\bar{\omega} + (1-t)\bar{x};\Omega) = \partial_C d(t\bar{\omega} + (1-t)\bar{x};\Omega) = \{(1,0)\},$$

which justifies the conclusions of this example. ■

Observe also that the estimates of Theorem 3.3 *essentially improve* known results even for the case of *convex* sets in *finite dimensions* when all the subdifferentials considered in Section 2 reduce to the subdifferential of convex analysis.

Example 3.6. (Improvement of known results for convex sets). There is a closed convex set $\Omega \subset I\!R^2$ and a point $\bar{x} \notin \Omega$ such that, for some $\bar{\omega} \in \Pi(\bar{x};\Omega)$ and $t \in (0,1)$, the subgradient sets $\partial^\bullet d(t\bar{\omega} + (1-t)\bar{x};\Omega)$ reduce to the same singleton for all the subdifferentials $\partial^\bullet$ under consideration being strictly smaller than

$$\bigcap_{\bar{\omega}\in\Pi(\bar{x};\Omega)} N(\bar{\omega};\Omega) \cap S^* \subset \bigcap_{\bar{\omega}\in\Pi(\bar{x};\Omega)} N(\bar{\omega};\Omega) \cap B^*,$$

where N stands for the normal cone of convex analysis.

Proof. It follows from Theorem 12 in Burke et al. [6] that

$$\partial d(\bar{x};\Omega) \subset \bigcap_{\bar{\omega}\in\Pi(\bar{x};\Omega)} N(\bar{\omega};\Omega) \cap I\!B^* \tag{3.7}$$

for convex sets in Banach spaces. Consider the case of Ω and $\bar{x}$ given by

$$\Omega := \text{epi}\,|\cdot| \subset I\!R^2 \quad \text{and} \quad \bar{x} := (0,-1) \notin \Omega.$$

It is easy to check that $\partial^\bullet d(\bar{x};\Omega) = \{(0,-1)\}$ for all the mentioned subdifferentials $\partial^\bullet$ reduced of course to the classical subdifferential ∂ of convex analysis. One also easily gets $\{\bar{\omega}\} = \{(0,0)\} = \Pi(\bar{x};\Omega)$ and

$$N(\bar{\omega};\Omega) \cap S^* = \big\{(u,v) \in I\!R^2 \big|\ u^2 + v^2 = 1,\ v \le -|u|\big\}.$$

On the other hand, for $t = 1/2$ and $t\bar{\omega} + (1-t)\bar{x} = (0,-1/2)$ we directly compute that $\partial d(t\bar{\omega} + (1-t)\bar{x};\Omega) = \{(0,-1)\}$, and the upper estimate of Theorem 3.3 is *exact* in this case being much sharper than (3.7) in general. ■

Observe also that the counterpart of the inclusion (3.7) formulated in [6, Theorem 12] for the Clarke subdifferential and normal cone requires the *directional regularity* assumption that was missing therein. Otherwise, one has

$$\bigcap_{\bar\omega\in\Pi(\bar x;\Omega)} N_C(\bar\omega;\Omega)\cap B^*=\big\{(0,0)\big\}$$

in the setting of Example 3.6, and the Clarke counterpart of inclusion (3.7) *is violated*.

Nevertheless, the next two interrelated theorems show that certain natural analogs of (3.7) for the standard distance function $d(\cdot;\Omega)$ and its extensions ρ hold in terms of our basic *limiting subgradients* with *no regularity* assumptions but under some *well-posedness* requirements, which are *automatic* in many important settings (e.g., in reflexive spaces with equivalent Kadec norms); see [16] for more details. First let us derive upper estimates for limiting subgradients of the distance functions at out-of-set points via those at intermediate points built as above.

Theorem 3.7. (Upper estimates for limiting subgradients of distance functions via limiting subgradients at intermediate points). *Let $F\colon Z\rightrightarrows X$ be a closed-graph multifunction between Banach spaces. Assume that $\Pi(\bar x;F(\bar z))\neq\emptyset$ and that the following well-posed conditions hold: for any sequences $\varepsilon_k\downarrow 0$ and $(z_k,x_k)\xrightarrow{\rho}(\bar z,\bar x)$ with $\widehat\partial_{\varepsilon_k}\rho(z_k,x_k)\neq\emptyset$ there is a sequence of $\omega_k\in\Pi(x_k;F(z_k))$ containing a convergent subsequence. Then for any $t\in(0,1]$ one has*

$$\partial\rho(\bar z,\bar x)\subset\bigcup_{\bar\omega\in\Pi(\bar x;F(\bar z))}\partial\rho(\bar z,t\bar\omega+(1-t)\bar x)\cap D^*,\tag{3.8}$$

where $D^*:=\big\{(z^*,x^*)\in Z^*\times X^*\big|\ ||x^*||\le 1\big\}$. In particular,

$$\partial d(\bar x;\Omega)\subset\bigcup_{\bar\omega\in\Pi(\bar x;\Omega)}\partial d(t\bar\omega+(1-t)\bar x;\Omega)\cap I\!\!B^*\ \text{ as }\ t\in(0,1]\tag{3.9}$$

for every closed set $\Omega\subset X$ and $\bar x\notin\Omega$ with $\Pi(\bar x;\Omega)\neq\emptyset$ provided the following well-posedness: for any sequences $\varepsilon_k\downarrow 0$, $x_k\to\bar x$ with $\widehat\partial_{\varepsilon_k}d(x_k;\Omega)\neq\emptyset$ there is a sequence $\omega_k\in\Pi(x_k;\Omega)$ containing a convergent subsequence.

Proof. Fix $(z^*,x^*)\in\partial\rho(\bar z,\bar x)$ and find by definition sequences

$$\varepsilon_k\downarrow 0,\ \ (z_k,x_k)\xrightarrow{\rho}(\bar z,\bar x),\ \ \text{and}\ \ (z_k^*,x_k^*)\xrightarrow{w^*}(z^*,x^*)\ \ \text{as}\ \ k\to\infty$$

satisfying $(x_k^*,z_k^*)\in\widehat\partial_{\varepsilon_k}\rho(z_k,x_k)$ for all $k\in I\!\!N$. Employing the well-posedness and closed-graph assumptions, we have a subsequence of $\omega_k\in\Pi(x_k;F(z_k))$ converging to some point $\bar\omega\in F(\bar z)$. By the relations

$$||x_k-\omega_k||=\rho(z_k,x_k)\to\rho(\bar z,\bar x)=||\bar x-\bar\omega||$$

we have $\bar\omega \in \Pi(\bar x; F(\bar z))$. Given any $t \in (0,1]$, it follows from Proposition 3.2 that

$$(z_k^*, x_k^*) \in \widehat\partial_{\varepsilon_k}\rho(z_k, t\omega_k + (1-t)x_k) \ \text{ and } \ ||x_k^*|| \le 1 + \varepsilon_k, \quad k \in I\!N.$$

Taking into account that

$$(z_k, t\omega_k + (1-t)x_k) \to (\bar z, t\bar\omega + (1-t)\bar x) \ \text{ and } \ ||x^*|| \le \liminf ||x_k^*|| \le 1 \ \text{ as } \ k \to \infty,$$

we get by Lemma 3.1 that

$$\rho(z_k, t\omega_k + (1-t)x_k) = (1-t)\rho(z_k, x_k) \to (1-t)\rho(\bar z, \bar x) = \rho(\bar z, t\bar\omega + (1-t)\bar x).$$

Thus $(z_k, t\omega_k + (1-t)x_k) \xrightarrow{\rho} (\bar z, t\bar\omega + (1-t)\bar x)$, which implies $(z^*, x^*) \in \partial\rho(\bar z, t\bar\omega + (1-t)\bar x)$ and hence justifies (3.8). As usual, inclusion (3.9) follows from (3.8) by considering the constant mapping $F(\cdot) \equiv \Omega$. ■

Note that one can equivalently put $\varepsilon_k = 0$ in the well-posedness requirements of Theorem 3.7 when the spaces X and Z are Asplund and when the general distance function ρ is lower semicontinuous. Assuming in addition that the graph of $F_{t,\bar\omega}$ is closed that obviously holds when ρ is l.s.c. (of course, it is redundant for the standard one $d(\cdot;\Omega)$), we arrive at the following estimates involving *limiting normals* at intermediate points. We refer the reader to the formulation of Corollary 3.4 for the symbols $F_{t,\bar\omega}$ and $\Omega_{t,\bar\omega}$ and to [16, Section 5] for general conditions ensuring the lower semicontinuity of ρ. Observe that Theorem 3.8 essentially improves our previous results [16] of the *projection type* corresponding to the case of $t = 1$.

Theorem 3.8. (Upper estimates for limiting subgradients of distance functions via limiting normals at intermediate points). *Suppose that all the assumptions of Theorem* 3.7 *hold. For any fixed* $t \in (0,1]$ *assume in addition that* gph $F_{t\bar\omega}$ *is closed whenever* $\bar\omega \in \Pi(\bar x; F(\bar z))$. *Then one has*

$$\partial\rho(\bar z, \bar x) \subset \bigcup_{\bar\omega \in \Pi(\bar x; F(\bar z))} N((\bar z, t\bar\omega + (1-t)\bar x); \text{gph}\, F_{t,\bar\omega}) \cap D^* \tag{3.10}$$

with $D^* = \{(z^*, x^*) \in Z^* \times X^* \big|\ ||x^*|| \le 1\}$. *If, in particular,* $\Omega \subset X$ *is a closed set and* $\bar x \notin \Omega$ *with* $\Pi(\bar x; \Omega) \ne \emptyset$, *then*

$$\partial d(\bar x; \Omega) \subset \bigcup_{\bar\omega \in \Pi(\bar x; \Omega)} N(t\bar\omega + (1-t)\bar x; \Omega_{t,\bar\omega}) \cap I\!B^* \ \textit{ whenever } \ 0 < t \le 1.$$

Proof. To justify (3.10), fix any $t \in (0,1]$ and $(z^*, x^*) \in \partial\rho(\bar z, \bar x)$. As in the proof of Theorem 3.7, find sequences

$$\varepsilon_k \downarrow 0, \ (z_k, x_k) \xrightarrow{\rho} (\bar z, \bar x), \ (z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*), \ \text{ and } \ \omega_k \to \bar\omega \ \text{ as } \ k \to \infty$$

satisfying $(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho(z_k, x_k)$, $\omega_k \in \Pi(x_k; F(z_k))$, and $\bar{\omega} \in \Pi(\bar{x}; F(\bar{z}))$. Since the graph of $F_{t,\bar{\omega}}$ is closed while $(\bar{z}, \bar{x}) \notin \text{gph}\, F_{t,\bar{\omega}}$, we use result of [5, Lemma 3.1] and find a neighborhood V of $(\bar{z}, \bar{x})$ such that for any $(z, x) \in V$ one has

$$\begin{aligned} &\rho(z, x) = \rho_{t,\bar{\omega}}(z, x) + t_{\bar{\omega}} \quad \text{with} \quad \rho_{t,\bar{\omega}}(z, x) := d(F_{t,\bar{\omega}}(z); x) \\ &\text{and} \quad t_{\bar{\omega}} := d(F(\bar{z}); \widetilde{x}), \end{aligned} \tag{3.11}$$

where $\widetilde{x} := t\bar{\omega} + (1-t)\bar{x}$. Hence

$$t_{\bar{\omega}} = (1-t)||\bar{x} - \bar{\omega}||, \quad d(F_{t,\bar{\omega}}(\bar{z}); \bar{x}) = t||\bar{x} - \bar{\omega}|| = ||\bar{x} - \widetilde{x}||, \quad \text{and} \quad \widetilde{x} \in \Pi(\bar{x}; F_{t,\bar{\omega}}(\bar{z})).$$

Now for any fixed $k \in I\!N$ consider a continuous function $\varphi\colon [0, 1] \to I\!R$ defined by

$$\varphi(\lambda) := d(F(z_k); \lambda\omega_k + (1-\lambda)x_k) - (1-t)||\bar{x} - \bar{\omega}||.$$

Since $\varphi(1) = -(1-t)||\bar{x} - \bar{\omega}|| \le 0$ and $\varphi(0) = ||x_k - \omega_k|| - (1-t)||\bar{x} - \bar{\omega}|| > 0$ for large k, we find, by the classical intermediate value theorem, such $\lambda_k \in (0, 1]$ that

$$d(F(z_k); \lambda_k\omega_k + (1-\lambda_k)x_k) = t_{\bar{\omega}} = (1-t)||\bar{x} - \bar{\omega}||.$$

Suppose without loss of generality that $\lambda_k \to \lambda \in [0, 1]$ as $k \to \infty$. Then Lemma 3.1 gives

$$d(F(z_k); \lambda_k\omega_k + (1-\lambda_k)x_k) = (1-\lambda_k)||x_k - \omega_k|| \to (1-\lambda)||\bar{x} - \bar{\omega}|| \quad \text{as} \quad k \to \infty.$$

The latter implies that $\lambda = t$ and $\lambda_k\omega_k + (1-\lambda_k)x_k \in F_{t,\bar{\omega}}(z_k)$ converges to $t\bar{\omega} + (1-t)\bar{x}$. Then employing (3.11) with large k, we get

$$\begin{aligned} \rho_{t,\bar{\omega}}(z_k, x_k) &= \rho(z_k, x_k) - t_{\bar{\omega}} = ||x_k - \omega_k|| - t_{\bar{\omega}} \\ &= ||x_k - \omega_k|| - d(F(z_k); \lambda_k\omega_k + (1 - \lambda_k x_k)) \\ &= ||x_k - \omega_k|| - (1-\lambda_k)||x_k - \omega_k|| \\ &= \lambda_k||x_k - \omega_k|| \\ &= ||x_k - (\lambda_k\omega_k + (1-\lambda_k)x_k)||, \end{aligned}$$

which implies that $\lambda_k\omega_k + (1-\lambda_k)x_k \in \Pi(x_k; F_{t,\bar{\omega}}(z_k))$. Using again (3.11) together with Proposition 3.2, we justify the inclusions

$$\begin{aligned} (z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho(z_k, x_k) &= \widehat{\partial}_{\varepsilon_k}\rho_{t,\bar{\omega}}(z_k, x_k) \subset \widehat{\partial}_{\varepsilon_k}\rho_{t,\bar{\omega}}(z_k, \lambda_k\omega_k + (1-\lambda_k)x_k) \cap C^* \\ &\subset \widehat{N}_{\varepsilon_k}((z_k, \lambda_k\omega_k + (1-\lambda_k)x_k); \text{gph}\, F_{t,\bar{\omega}}) \cap C^*. \end{aligned}$$

for large k. Hence $(z^*, x^*) \in N((\bar{z}, t\bar{\omega} + (1-t)\bar{x}); \text{gph}\, F_{t,\bar{\omega}})$ and $||x^*|| \le \liminf ||x_k^*|| \le 1$, which completes the proof of the theorem. ■

4. Right-sided Limiting Subgradients of Distance Functions

As mentioned, some of the developments in Section 3 can be treated as extensions of our previous results of the *projection type* obtained in [16]. Observe that it is *very essential* that $t > 0$ in all the "intermediate" results of Section 2. Actually the main theorems obtained above simply *are not valid* when $t = 0$; see the example below. The passage to the limit as $t \downarrow 0$ requires involving new constructions and arguments that are actually equivalent to those presented in this section that follows the corresponding developments in [16].

In this section we always assume that $F\colon Z \rightrightarrows X$ is a *closed-graph* mapping between Banach spaces. Fix any point $(\bar{z}, \bar{x}) \in Z \times X$ and put $r := d(F(\bar{z}); \bar{x})$. Recall that the *enlargement mapping* $F_r\colon Z \rightrightarrows X$ is defined by

$$F_r(z) := \{x \in X \,|\, d(F(z); x) \le r\} \tag{4.1}$$

and observe that $F_r \equiv F$ if only if $r = 0$, which corresponds to the case of $(\bar{z}, \bar{x}) \in \text{gph}\, F$. We have the following relationship

$$N((\bar{z}, \bar{x}); \text{gph}\, F) = \bigcup_{\lambda \ge 0} \lambda \partial \rho(\bar{z}, \bar{x}), \quad (\bar{z}, \bar{x}) \in \text{gph}\, F, \tag{4.2}$$

between our basic/limiting subdifferential (2.4) of the general distance function at *in-set points* and the basic normal cone (2.2) to $\text{gph}\, F$ for an arbitrary closed-graph mapping $F\colon Z \rightrightarrows X$ established by Thibault [20]. However, we *cannot* keep such a relationship between $\partial\rho(\bar{z}, \bar{x})$ at *out-of-set points* and the basic normal to the graph of the enlargement, even for the case of standard distance functions in finite-dimensional spaces. Indeed, consider the set

$$\Omega := \left\{(x, y) \in I\!R^2 \,\middle|\, x^2 + y^2 \ge 1\right\}$$

and the point $(0,0) \notin \Omega$. Then $\Omega_r = I\!R^2$ with $r = 1$, $N(\bar{x}; \Omega_r) = \{0\}$ while $\partial d(\bar{x}; \Omega)$ is the whole unit sphere of $I\!R^2$.

To establish a counterpart of (4.2) in out-of-set points, we need the *new limiting modification* of the basic subdifferentials, which gives a smaller set of subgradients; namely, those which are obtained as limits of ε-subgradients at point x_k, where the function values are on the *right side* of $f(\bar{x})$, i.e., $f(x_k) \ge f(\bar{x})$.

Definition 4.1. (Right-sided limiting subgradients). Let $f\colon X \to \overline{I\!R}$ be finite at $\bar{x}$. Then the (limiting) right-sided subdifferential of f at $\bar{x}$ is

$$\partial_{\ge} f(\bar{x}) := \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{f+} \bar{x} \\ \varepsilon \downarrow 0}} \widehat{\partial}_\varepsilon f(x),$$

where $x \stackrel{f+}{\to} \bar{x}$ means that $x \to \bar{x}$ with $f(x) \to f(\bar{x})$ and $f(x) \geq f(\bar{x})$.

The right-sided limiting subdifferential first appeared in our previous paper [16] devoted to the study and applications of distance functions. While reading it, Lionel Thibault drew our attention that a different but somehow related sided subdifferential of the standard distance function, involving limits of Clarke normals, was introduced by Cornet and Czarnecki [8] in finite dimensions to establish existence theorems for generalized equilibria.

Observe that it is possible to put equivalently $\varepsilon = 0$ in the above limiting constructions if X is *Asplund,* if f is *l.s.c.* around $\bar{x}$, and if $\widehat{\partial} f(x) = \emptyset$ whenever $f(x) = f(\bar{x})$ and x near $\bar{x}$. One obviously has

$$\widehat{\partial} f(\bar{x}) \subset \partial_{\geq} f(\bar{x}) \subset \partial f(\bar{x}). \tag{4.3}$$

Observe that the equalities hold in both inclusions of (4.3) when φ is *lower regular* at $\bar{x}$, in particular, when f is convex. In general both inclusions in (4.3) may be *strict* even for the standard distance function in finite dimensions; see [16].

Using the Ekeland variational principle, we can prove the following auxiliary result establishing the relationship between Fréchet ε-subgradients of distance functions in term of (nonempty) *perturbed projections*; cf. [16, Theorem 3.6].

Lemma 4.2. (Estimates of ε-subgradients for distance functions via normal at perturbed projections.) *Let* $F\colon Z \rightrightarrows X$ *with* $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$. *Then for any* $\varepsilon \geq 0$, *for any* $(z^*, x^*) \in \widehat{\partial}_\varepsilon \rho(\bar{z}, \bar{x})$, *and for any* $\eta > 0$ *there exists* $(v, u) \in \operatorname{gph} F$ *such that*

$$||v - \bar{z}|| \leq \eta, \ ||u - \bar{x}|| \leq \rho(\bar{z}, \bar{x}) + \eta, \ \text{ and } \ (z^*, x^*) \in \widehat{N}_{\varepsilon+\eta}((v, u); \operatorname{gph} F).$$

The next theorem, presented here for completeness (cf. [16]), gives appropriate analogs of representation (4.2) at *out-of-set* points with using of enlargement mapping (4.1) and replacing the limiting subdifferential $\partial \rho(\bar{z}, \bar{x})$ by its *right-sided* counterpart.

Theorem 4.3. (relationships between right-sided subgradients of distance functions and limiting normals to enlargements). *Let* $F\colon Z \rightrightarrows X$ *with* $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$, *and let* $r := \rho(\bar{z}, \bar{x})$. *Assume that* $\operatorname{gph} F_r$ *is closed. Then one has the inclusion*

$$\partial_{\geq} \rho(\bar{z}, \bar{x}) \subset N((\bar{z}, \bar{x}); \operatorname{gph} F_r) \cap (Z^* \times I\!B^*), \tag{4.4}$$

which admits the stronger form

$$\partial_{\geq} \rho(\bar{z}, \bar{x}) \subset N((\bar{z}, \bar{x}); \operatorname{gph} F_r) \cap \left[Z^* \times (I\!B^* \setminus \{0\})\right] \tag{4.5}$$

if the set $\operatorname{gph} F_r \subset Z \times X$ *is SNC at* $(\bar{z}, \bar{x})$ *with respect to* X. *Moreover,*

$$\partial_{\geq} \rho(\bar{z}, \bar{x}) \subset N((\bar{z}, \bar{x}); \operatorname{gph} F_r) \cap (Z^* \times S^*) \tag{4.6}$$

if X is finite-dimensional.

Proof. To justify (4.4), pick $(z^*,x^*)\in\partial_{\ge}\rho(\bar z,\bar x)$ and Definition 4.1 find sequences $\varepsilon_k\downarrow 0$, $(z_k,x_k)\xrightarrow{\rho}(\bar z,\bar x)$, and $(z_k^*,x_k^*)\xrightarrow{w^*}(z^*,x^*)$ satisfying

$$\rho(z_k,x_k)\ge\rho(\bar z,\bar x)>0\ \text{ and }\ (z_k^*,x_k^*)\in\widehat\partial\rho_{\varepsilon_k}(z_k,x_k)\ \text{ for all }\ k\in I\!\!N.$$

Since $(\bar z,\bar x)\notin\operatorname{gph}F$, we have $(z_k,x_k)\notin\operatorname{gph}F$ for all large $k\in I\!\!N$. It is not difficult to check that

$$(z_k^*,x_k^*)\in\widehat N_{\varepsilon_k}((z_k,x_k);\operatorname{gph}F_r),\ 1-\varepsilon_k\le\|x_k^*\|\le 1+\varepsilon_k \tag{4.7}$$

provided that there is a subsequence (z_k,x_k) such that $\rho(z_k,x_k)=r=\rho(\bar x,\bar z)$. If the opposite holds, we use the result by Bounkhel and Thibault [5, Lemma 3.1] to ensure the representation

$$\rho(z,x)=r+\rho_r(z;x)\ \text{ for all }\ (z,x)\notin\operatorname{gph}F_r\ \text{ with }\ \rho_r(z,x):=d(x;F_r(z)).$$

This directly implies that

$$(z_k^*,x_k^*)\in\widehat\partial_{\varepsilon_k}\rho(z_k,x_k)=\widehat\partial_{\varepsilon_k}(r+\rho_r)(z_k,x_k)=\widehat\partial_{\varepsilon_k}\rho_r(z_k,x_k),\ \ k\in I\!\!N.$$

Denote $\eta_k:=\rho_r(z_k,x_k)=\rho(z_k,x_k)-r\downarrow 0$ as $k\to\infty$ and, by Lemma 4.2, find $(v_k,u_k)\in\operatorname{gph}F_r$ satisfying

$$\begin{aligned}&\|z_k-v_k\|\le\eta_k,\ \|x_k-u_k\|\le\rho_r(z_k,x_k)+\eta_k\le 2\eta_k,\quad\text{and}\\&(z_k^*,x_k^*)\in\widehat N_{\varepsilon_k+\eta_k}((v_k,u_k);\operatorname{gph}F_r),\ 1-\varepsilon_k\le\|x_k^*\|\le 1+\varepsilon_k.\end{aligned}\tag{4.8}$$

Passing to the limit as $k\to\infty$ in both relations (4.7) and (4.8), we arrive at

$$(z^*,x^*)\in N((\bar z,\bar x);\operatorname{gph}F_r)\ \text{ with }\ \|x^*\|\le 1,$$

which justifies (4.4). Moreover, $\|x^*\|\ne 0$ under the SNC requirement on the graph of F_r with respect to X by the above constructions. This gives (4.5). When X is finite dimensional, we get $\|x^*\|\ge 1$ by passing to the limit in the lower estimate $\|x_k^*\|\ge 1-\varepsilon_k$ of (4.7) and (4.8). This gives (4.6) and completes the proof of the theorem. ■

The next result deals with the *standard distance function* $d(\cdot;\Omega)$. Its first part is a direct consequence of Theorem 4.3, with the notation

$$\Omega_r:=\big\{x\in X\big|\ d(x;\Omega)\le r\big\}\ \text{ as }\ r:=d(\bar x;\Omega)$$

standing for the corresponding *enlargement* of Ω at $\bar x\notin\Omega$. The second part is certainly of independent interest.

Theorem 4.4. (Relationships between right-sided subgradients of the standard distance function and basic normals to the set enlargement). *Let Ω be a closed subset in a Banach space X, and let $\bar x\notin\Omega$ with $r=d(\bar x;\Omega)$. Then the following hold:*

(i) *One has the inclusion*

$$\partial_{\geq} d(\bar{x};\Omega) \subset N(\bar{x};\Omega_r) \cap I\!\!B^*,$$

where the stronger inclusion

$$\partial_{\geq} d(\bar{x};\Omega) \subset \big[N(\bar{x};\Omega_r) \cap I\!\!B^*\big] \setminus \{0\}$$

is fulfilled if Ω_r is SNC at $\bar{x}$. Moreover,

$$\partial_{\geq} d(\bar{x};\Omega) \subset N(\bar{x};\Omega_r) \cap S^*$$

if the space X is finite-dimensional.

(ii) *One always has the equality*

(4.9) $$N(\bar{x};\Omega_r) = \bigcup_{\lambda \geq 0} \lambda \partial_{\geq} d(\bar{x};\Omega).$$

Proof. We need to justify assertion (ii). The inclusion "$\supset$" in (4.9) follows from (4.4). To proof the opposite inclusion, we first verify that

(4.10) $$N(\bar{x};\Omega_r) \setminus \{0\} \subset \bigcup_{\lambda > 0} \lambda \partial_{\geq} d(\bar{x};\Omega).$$

To proceed, pick any normal $x^* \in N(\bar{x};\Omega_r) \setminus \{0\}$ and find sequences $\varepsilon_k \downarrow 0$, $x_k \stackrel{\Omega_r}{\to} \bar{x}$, and $x_k^* \in \widehat{N}_{\varepsilon_k}(x_k;\Omega_r)$ with $x_k^* \stackrel{w^*}{\to} x^*$ as $k \to \infty$. Since $0 < \|x^*\| \leq \liminf_{k\to\infty} \|x_k^*\|$, there is $m > 0$ and a subsequence of $\{x_k^*\}$ (with no relabeling) such that $\|x_k^*\| \geq m$ for all $k \in I\!\!N$. One has $d(x_k;\Omega) \leq r$ by $x_k \in \Omega_r$ and the closedness of Ω. Observe that $d(x_k;\Omega) = r$ for all $k \in I\!\!N$, since $x_k^* = 0$ otherwise. Thus

$$\frac{x^*}{\|x_k^*\|} \in \Big\{x_k^* \in \widehat{N}_{\varepsilon_k/m}(x_k;\Omega_r) \Big| \; \|x^*\| = 1\Big\}.$$

Then modifying slightly the proof of Theorem 3.6 in Bounkhel and Thibault 3.6, we find a bounded sequence of positive numbers α_k such that

$$\frac{x_k^*}{\|x_k^*\|} \in \widehat{\partial}_{\alpha_k \varepsilon_k/m} d(x_k;\Omega), \quad k \in I\!\!N.$$

Due to the boundedness of $\{\|x_k^*\|\}$, we assume with no loss of generality that $\|x_k^*\| \to \widetilde{\lambda}$ as $k \to \infty$ for some number $\widetilde{\lambda} > 0$. Therefore

$$x^* \in \widetilde{\lambda} \partial_{\geq} d(\bar{x};\Omega) \subset \bigcup_{\lambda \geq 0} \partial_{\geq} d(\bar{x};\Omega)$$

by definition of the right-sided subdifferential. Adding $\lambda = 0$ to the union on the right-hand side of (4.10), we see that $x^* = 0$ belongs to this set due to our convention $0 \cdot \emptyset = 0$. Thus we arrive at (4.9) and complete the proof of the theorem. ■

5. Relationship Between Singular Subgradients of Distance Functions and Coderivatives of Generating Mappings

The primary goal of this section is to establish relationships between the *singular subdifferential* (2.6) of the distance function ρ defined in (1.2) and the *mixed coderivative* (2.9) of the generating mapping $F\colon Z \rightrightarrows X$ in (1.2). Note that this question does not make sense for the standard distance function (1.1), which is always globally Lipschitz continuous with therefore $\partial^\infty d(\bar{x};\Omega) = \{0\}$.

Observe that the distance function (1.2) belongs to the class of the so-called *marginal functions* given generally by

$$\mu(x) := \inf\big\{\varphi(x,y)\big|\ y \in G(x)\big\}, \tag{5.1}$$

where $\varphi\colon X \times Y \to \overline{I\!R}$ is a l.s.c. function and where $G\colon X \rightrightarrows Y$ is a closed-graph multifunction around the reference points. Marginal functions of type (5.1) play indeed a crucial role in variational analysis and optimization; see, e.g., [15, 19] and the references therein. In particular, they describe *optimal values* in minimization problems being often called for this reason by *value functions*.

We associate with the marginal function (5.1) the *solution map*

$$S(x) := \big\{y \in G(x)\big|\ \varphi(x,y) = \mu(x)\big\}, \tag{5.2}$$

which is assumed to be nonempty in what follows. We say that S is *μ-inner semicontinuous* at $(\bar{x},\bar{y}) \in \operatorname{gph} S$ if for any sequences $\varepsilon_k \downarrow 0$ and $x_k \xrightarrow{\mu} \bar{x}$ such that $\widehat{\partial}_{\varepsilon_k}\mu(x_k) \neq \emptyset$, there is a sequence $y_k \in S(x_k)$ containing a subsequence that converges to $\bar{y}$. The mapping S is said to be *μ-inner semicompact* at $\bar{x}$ if for any sequences $\varepsilon_k \downarrow 0$ and $x_k \xrightarrow{\mu} \bar{x}$ with $\widehat{\partial}_{\varepsilon_k}\mu(x_k) \neq \emptyset$ there is a sequence $y_k \in S(x_k)$ containing a subsequence that converges to some $\bar{y} \in S(\bar{x})$. Observe as usual that we can equivalently put $\varepsilon_k = 0$ if both spaces X and Y are *Asplund* and if μ is lower semicontinuous.

The following theorem establishes important relationships between the singular subdifferential of the marginal function (5.1) and the *mixed* coderivative of the generating mapping G that essentially improve the previously known ones [17] obtained in terms of the bigger *normal* coderivative (2.8).

Theorem 5.1. (Singular subgradients of marginal functions). *Let X and Y be Asplund. The following assertions hold:*

(i) *Assume that φ is locally Lipschitzian around $(\bar{x}, \bar{y})$ and the solution map S is μ-inner semicontinuous at $(\bar{x}, \bar{y})$. Then*

$$\partial^\infty \mu(\bar{x}) \subset D^*_M G(\bar{x}, \bar{y})(0). \tag{5.3}$$

(ii) *Assume that S is μ-inner semicompact at $\bar{x}$ and φ is locally Lipschitzian around $(\bar{x}, \bar{y})$ for all $\bar{y} \in S(\bar{x})$. Then*

$$\partial^\infty \mu(\bar{x}) \subset \bigcup_{\bar{y} \in S(\bar{x})} D^*_M G(\bar{x}, \bar{y})(0).$$

Proof. To justify (5.3), fix any $x^* \in \partial^\infty \mu(\bar{x})$ and have by definition that

$$x^* \in \mathop{\mathrm{Lim\,sup}}_{\substack{x \xrightarrow{\mu} \bar{x} \\ \varepsilon \downarrow 0,\, \lambda \downarrow 0}} \lambda \widehat{\partial}_\varepsilon \mu(x),$$

i.e., there are sequences $\varepsilon_k \downarrow 0$, $\lambda_k \downarrow 0$, $x_k \xrightarrow{\mu} \bar{x}$, and $x^*_k \in \widehat{\partial}_{\varepsilon_k} \mu(x_k)$ such that $\lambda_k x^*_k \xrightarrow{w^*} x^*$ as $k \to \infty$. Since S is μ-inner semicontinuous at $(\bar{x}, \bar{y})$, we can find $y_k \in S(x_k)$ whose subsequence, with no relabeling, converges to $\bar{y}$. It follows by definition from $x^*_k \in \widehat{\partial}_{\varepsilon_k} \mu(x_k)$ that for any $\eta > 0$ there is $\gamma > 0$ such that

$$\langle x^*_k, x - x_k \rangle \le \mu(x) - \mu(x_k) + (\varepsilon_k + \eta)||x - x_k|| \quad \text{whenever} \quad x \in x_k + \gamma I\!B.$$

Considering the function

$$\phi(x, y) := \varphi(x, y) + \delta((x, y); \mathrm{gph}\, G),$$

we easily conclude that

$$\langle (x^*_k, 0), (x - x_k, y - y_k) \rangle \le \phi(x, y) - \phi(x_k, y_k) + (\varepsilon_k + \eta)(||x - x_k|| + ||y - y_k||)$$

whenever $(x, y) \in (x_k, y_k) + \gamma I\!B$, which gives $(x^*_k, 0) \in \widehat{\partial}_{\varepsilon_k} \phi(x_k, y_k)$.

Fix now any sequence $\eta_k \downarrow 0$. Since φ is locally Lipschitzian around $(\bar{x}, \bar{y})$ while X and Y are Asplund, we apply the "fuzzy" sum rule for ε-subgradients of ϕ (see, e.g., [15, 17]) and find sequences

$$(x_{1k}, y_{1k}) \xrightarrow{\varphi} (\bar{x}, \bar{y}),\ (x_{2k}, y_{2k}) \xrightarrow{\mathrm{gph}\, G} (\bar{x}, \bar{y}),$$
$$(x^*_{1k}, y^*_{1k}) \in \widehat{\partial} \varphi(x_{1k}, y_{1k}), \quad \text{and} \quad (x^*_{2k}, y^*_{2k}) \in \widehat{N}((x_{2k}, y_{2k}); \mathrm{gph}\, G)$$

satisfying the estimate

$$||(x^*_k, 0) - (x^*_{1k}, y^*_{1k}) - (x^*_{2k}, y^*_{2k})|| \le \varepsilon_k + \eta_k$$

or, equivalently, the following ones:

$$||x_k^* - x_{1k}^* - x_{2k}^*|| \le \varepsilon_k + \eta_k \ \text{ and } \ ||y_{1k}^* + y_{2k}^*|| \le \varepsilon_k + \eta_k. \tag{5.4}$$

It follows from the Lipschitz continuity of φ with some modulus $\ell > 0$ that $||(x_{1k}^*, y_{1k}^*)|| \le \ell$, which implies that $\lambda_k \|(x_{1k}^*, y_{1k}^*)\| \to 0$ as $k \to \infty$. By (5.4) we therefore have

$$\lambda_k \|y_{2k}^*\| \to 0 \ \text{ and } \ \lambda_k x_{2k}^* \xrightarrow{w^*} x^* \ \text{ as } \ k \to \infty. \tag{5.5}$$

Taking into account that

$$\lambda_k(x_{2k}^*, y_{2k}^*) \in \widehat{N}((x_{2k}, y_{2k}); \text{gph}\, G) \ \text{ for all } \ k \in I\!N$$

and using the definition of the mixed coderivative (2.9), we derive from the convergence relations (5.5) that $x^* \in D_M^* G(\bar{x}, \bar{y})(0)$, which gives (5.3) and completes the proof of (i).

The proof of assertion (ii) is similar with using the μ-inner semicontinuity condition for S instead of the μ-inner semicompactness one in (i). ■

To include the distance function (1.2) into framework of Theorem 5.1, we need to consider a slightly more general class of marginal functions related to minimization problems with the so-called *moving sets* of feasible solutions. Namely, consider marginal functions in the form

$$\mu(x, y) := \inf \big\{ \varphi(y, z) \big|\, z \in G(x) \big\}, \tag{5.6}$$

where $\varphi\colon Y \times Z \to \overline{I\!R}$ and $G\colon X \rightrightarrows Z$. The corresponding solution sets are given by

$$S(x, y) := \big\{ z \in G(x) \big|\, \varphi(y, z) = \mu(x, y) \big\}. \tag{5.7}$$

We impose the standing assumptions on φ and G as for the case of (5.1) and (5.2). In fact, the following results for the more general class of marginal functions (5.6) are easily derived from Theorem 5.1.

Corollary 5.2. (singular subgradients of marginal functions over moving sets). *Let μ and S be given in* (5.6) *and* (5.7)*, and let the spaces X, Y, Z be Asplund. The following assertions hold.*

(i) *Assume that S is μ-inner semicontinuous at $((\bar{x}, \bar{y}), \bar{z})$ and that φ is locally Lipschitzian around $(\bar{y}, \bar{z})$. Then*

$$\partial^\infty \mu(\bar{x}, \bar{y}) \subset \big\{ (x^*, 0) \big|\, x^* \in D_M^* G(\bar{x}, \bar{z})(0) \big\}.$$

(ii) *If S is μ-inner semicompact at $(\bar{x},\bar{y})$ and φ is locally Lipschitzian around $(\bar{y},\bar{z})$ for all $\bar{z}\in S(\bar{x},\bar{y})$, then*

$$\partial^\infty\mu(\bar{x},\bar{y})\subset\bigcup_{\bar{z}\in S(\bar{x},\bar{y})}\big\{(x^*,0)\big|\,x^*\in D_M^*G(\bar{x},\bar{z})(0)\big\}.$$

Proof. To proof (i), put $u=(x,y)$ and define

$$\widetilde{G}(u)=\widetilde{G}(x,y):=G(x),\quad \widetilde{\varphi}(u,z):=\varphi(y,z).$$

Then we have the representation

$$\mu(x,y)=\mu(u)=\inf\big\{\widetilde{\varphi}(u,z)\big|\,z\in\widetilde{G}(u)\big\}.$$

Applying Theorem 5.1 with $(\bar{x},\bar{y})$ replaced by $(\bar{u},\bar{z})=(\bar{x},\bar{y},\bar{z})$, we get

$$\partial^\infty\mu(\bar{x},\bar{y})=\partial^\infty\mu(\bar{u})\subset\big\{(x^*,y^*)\big|\,(x^*,y^*)\in D_M^*\widetilde{G}((\bar{x},\bar{y}),\bar{z})(0)\big\}.$$

It is easy to observe the inclusion

$$D_M^*\widetilde{G}((\bar{x},\bar{y}),\bar{z})(0)\subset\big\{(x^*,0)\big|\,x^*\in D_M^*G(\bar{x},\bar{z})(0)\big\}.$$

Summarizing all the above, we arrive at

$$\partial^\infty\mu(\bar{x},\bar{y})\subset\big\{(x^*,0)\big|\,x^*\in D_M^*G(\bar{x},\bar{z})(0)\big\}$$

and complete the proof of (i). The proof of (ii) is similar. ■

Finally in this section, we establish relationships between the singular subdifferential of the distance function (1.2) and the mixed coderivative of the mapping $F_{t,\bar{\omega}}$ from Corollary 3.4 of Section 3 that depends on *intermediate projection* points.

Theorem 5.3. (Singular subgradients of distance functions at out-of-set points via intermediate projections). *Let $F\colon Z\rightrightarrows X$ be a closed-graph mapping between Asplund spaces, and let $(\bar{x},\bar{y})\notin$ gph F with $\Pi(\bar{x};F(\bar{z}))\neq\emptyset$. Assume that the well-posed condition of Theorem 3.7 holds. For any fixed $t\in(0,1]$ suppose in addition that* gph $F_{t,\bar{\omega}}$ *is closed whenever $\bar{\omega}\in\Pi(\bar{x};F(\bar{z}))$. Then we have the inclusion*

$$\partial^\infty\rho(\bar{z},\bar{x})\subset\bigcup_{\bar{\omega}\in\Pi(\bar{x};F(\bar{z}))}\Big\{(x^*,0)\Big|\,x^*\in D_M^*F_{t,\bar{\omega}}(\bar{x},t\bar{w}+(1-t)\bar{x})(0)\Big\}.\tag{5.8}$$

In particular, one has (as $t=1$) that

$$\partial^\infty\rho(\bar{z},\bar{x})\subset\bigcup_{\bar{w}\in\Pi(\bar{x};F(\bar{z}))}\Big\{(x^*,0)\Big|\,x^*\in D_M^*F(\bar{z},\bar{w})(0)\Big\}.\tag{5.9}$$

Proof. Clearly the inclusion (5.9) follows directly from Corollary 5.2. It remains to justify (5.8). Fix $\bar{\omega} \in \Pi(\bar{x}; F(\bar{z}))$, $t \in (0,1]$, and $(z^*, x^*) \in \partial^\infty \rho(\bar{z}, \bar{x})$. Then find sequences $\varepsilon_k \downarrow 0$, $\lambda_k \downarrow 0$, $(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x})$, and $(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k} \rho(z_k, x_k)$ satisfying

$$\lambda_k (z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*) \ \text{ as } \ k \to \infty.$$

Repeating the proof of Theorem 3.8, we get sequences $\omega_k \in \Pi(x_k; F(z_k))$ and $\bar{\omega} \in \Pi(\bar{x}; F(\bar{z}))$ satisfying the relations

$$t_k \omega_k + (1 - t_k) x_k \to t\bar{\omega} + (1 - t)\bar{x} \ \text{ with } \ t_k \omega_k + (1 - t_k) x_k \in F_{t,\bar{\omega}}(z_k), \ \text{ and}$$

$$(z_k^*, x_k^*) \in \widehat{N}_{\varepsilon_k}((z_k, t_k \omega_k + (1 - t_k) x_k); \operatorname{gph} F_{t,\bar{\omega}}) \ \text{ with } \ 1 - \varepsilon_k \le ||x_k^*|| \le 1 + \varepsilon_k.$$

Hence $\lambda_k (z_k^*, x_k^*) \in \widehat{N}_{\lambda_k \varepsilon_k}((z_k, t_k \omega_k + (1 - t_k) x_k); \operatorname{gph} F_{t,\bar{\omega}})$ and

$$\lambda_k \|x_k^*\| \to 0, \quad \lambda_k z_k^* \xrightarrow{w^*} z^* \ \text{ as } \ k \to \infty.$$

Thus $x^* = 0$ and $z^* \in D_M^* F_{t,\bar{\omega}}(\bar{z}, t\bar{\omega} + (1 - t)\bar{x})(0)$, which completes the proof.■

6. Some Applications

There are a great many of possible applications of generalized differentiation results for both distance functions (1.1) and (1.2) under consideration; see, e.g., [2, 3-8, 15, 16, 19] and the references therein for some applications of previously known results in this direction. In this section we choose to present two new applications of the results obtained above. The first one gives new conditions for the *projection nonemptiness* in infinite dimensions; the second applications ensures the *Lipschitz continuity* of the general distance function ρ from (1.2), which strongly relates to *Lipschitzian stability* of constraint and variational systems; cf. [16]. The next theorem provides, besides efficient conditions for the projection nonemptiness, refined *upper estimates* for the limiting subdifferential of distance functions (1.1) and (1.2) in the (range) Hilbert space setting.

Theorem 6.1. (Sufficient conditions for the projection nonemptiness via limiting subgradients).*Let $F\colon Z \rightrightarrows X$ be a closed-graph mapping from Asplund space Z to a Hilbert space X, and let $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$. Assume that ρ is l.s.c. around $(\bar{z}, \bar{x})$ and that*

$$\partial \rho(\bar{z}, \bar{x}) \cap \big\{ (z^*, x^*) \in Z^* \times X^* \big| \ ||x^*|| = 1 \big\} \ne \emptyset.$$

Then $\Pi(\bar{x}; F(\bar{z}))$ *is nonempty. Moreover,*

$$\partial\rho(\bar{z},\bar{x}) \bigcap \Big\{(z^*,x^*) \in Z^* \times X^* \Big|\, ||x^*|| = 1\Big\}$$
$$\subset \bigcup_{\bar{\omega}\in\Pi(\bar{x};F(\bar{z}))} \Big\{(z^*,x^*) \in N((\bar{z},\bar{\omega}); \operatorname{gph} F) \Big|\, x^* = \frac{\bar{x} - \Pi(\bar{x}; F(\bar{z}))}{\rho(\bar{z},\bar{x})}\Big\}.$$

In particular, $\Pi(\bar{x};\Omega) \neq \emptyset$ *for any closed subset* $\Omega \subset X$ *of a Hilbert space with* $\bar{x} \notin \Omega$ *and* $\partial d(\bar{x};\Omega) \cap S^* \neq \emptyset$. *Furthermore, in the latter case one has*

$$\partial d(\bar{x};\Omega) \cap S^* \subset \frac{\bar{x} - \Pi(\bar{x};\Omega)}{d(\bar{x};\Omega)}.$$

Proof. Fix any $(z^*,x^*) \in \partial\rho(\bar{z},\bar{x})$ with $||x^*|| = 1$. By definition there are sequences

$$(z_k,x_k) \xrightarrow{\rho} (\bar{z},\bar{x}) \text{ and } (z_k^*,x_k^*) \xrightarrow{w^*} (z^*,x^*) \text{ such that } (z_k^*,x_k^*) \in \widehat{\partial}\rho(z_k,x_k).$$

Then $x_k^* \in \widehat{\partial} d(\cdot;F(z_k))$ at $x = x_k$ with $x_k \notin F(z_k)$ when k is sufficiently large. It follows from [21, Theorem 5.3] that $\Pi(x_k;F(z_k))$ is a singleton $\{\omega_k\}$ and that

$$x_k^* = \frac{x_k - \omega_k}{\rho(z_k,x_k)}.$$

Using now Corollary 3.4 for the case of Fréchet normals and subgradients, we conclude that $(z_k^*,x_k^*) \in \widehat{N}((z_k,\omega_k); \operatorname{gph} F)$. Since $||x_k^*|| = 1 \to 1 = ||x^*||$, $x_k^* \xrightarrow{w} x^*$, and the norm in Hilbert space is *Kadec*, one has $x_k^* \xrightarrow{||\cdot||} x^*$. This implies that

$$\omega_k \to \bar{x} - x^*\rho(\bar{z},\bar{x}) \text{ as } k \to \infty.$$

Putting $\bar{\omega} := \bar{x} - x^*\rho(\bar{z},\bar{x})$, we obtain

$$||\bar{x} - \bar{\omega}|| = ||x^*\rho(\bar{z},\bar{x})|| = \rho(\bar{z},\bar{x}).$$

Hence $\Pi(\bar{x};F(\bar{z})) \neq \emptyset$, $(z^*,x^*) \in N((\bar{z},\bar{\omega}); \operatorname{gph} F)$, and

$$x^* = \frac{\bar{x} - \bar{\omega}}{\rho(\bar{z},\bar{x})} \in \frac{\bar{x} - \Pi(\bar{x};F(\bar{z}))}{\rho(\bar{z},\bar{x})},$$

which complete the proof of the theorem. ∎

The last theorem gives efficient conditions ensuring the Lipschitz continuity of the general distance function (1.2) at *out-of-set* points.

Theorem 6.2. (Sufficient conditions for Lipschitzian continuity of the general distance function at out-of-set points).*Let $F\colon Z \rightrightarrows X$ be a closed-graph mapping between Asplund spaces, and let $(\bar z,\bar x)\notin \mathrm{gph}\,F$. Assume that $\Pi(\bar x;F(\bar z))\neq\emptyset$, that ρ is l.s.c., and that the well-posedness condition Theorem 3.7 holds. Suppose also that there is $t\in(0,1]$ such that the mapping $F_{t,\bar\omega}$ defined in Corollary 3.4 is Lipschitz-like around $(\bar z;t\bar\omega+(1-t)\bar x)$ for all $\bar\omega\in\Pi(\bar x;F(\bar z))$. Then ρ is locally Lipschitzian around $(\bar z,\bar x)$.*

In particular, if the well-posedness condition Theorem 3.7 *holds and the original mapping F is Lipschitz-like around $(\bar x,\bar\omega)$ for any $\bar\omega\in\Pi(\bar x;F(\bar z))$, then ρ is locally Lipschitzian around $(\bar z,\bar x)$.*

Proof. Using [15, Lemma 2.36], it is not hard to prove that a *necessary and sufficient* condition for a l.s.c. function $f\colon X\to\overline{I\!R}$ defined on an Asplund space X to be *SNEC* is as follows: for any sequences $\lambda_k\downarrow 0$, $x_k\xrightarrow{f}\bar x$, and $x_k^*\in\lambda_k\widehat\partial f(x_k)$ one has

$$\big[x_k^*\xrightarrow{w^*}0\big]\Longrightarrow\big[\|x_k^*\|\to 0\big]\ \text{ as }\ k\to\infty. \tag{6.2}$$

Now fix any $\bar\omega\in\Pi(\bar x;F(\bar z))$ and also fix $t\in(0,1]$ from in the assumptions of the theorem. Since $F_{t,\bar\omega}$ is assumed to be Lipschitz-like around $(\bar z;t\bar\omega+(1-t)\bar x)$, we have

$$D_M^*F_{t,\bar\omega}(\bar z,t\bar\omega+(1-t)\bar x)(0)=\{0\}\ \text{ for all }\ \bar\omega\in\Pi(\bar x;F(\bar z))$$

due to the coderivative criterion from [14, Theorem 4.3]. Then the singular subdifferential inclusion from Theorem 5.3 valid under the well-posedness condition implies that $\partial^\infty\rho(\bar z,\bar x)=\{0\}$. To ensure the Lipschitz continuity of ρ around $(\bar z,\bar x)$, it remains, by [15, Theorem 3.49], to show that ρ is SNEC at $(\bar z,\bar x)$.

To proceed, let us employ the above characterization of the SNEC property and consider arbitrary sequences

$$\lambda_k\downarrow 0,\ \ (z_k,x_k)\xrightarrow{\rho}(\bar z,\bar x),\ \text{ and }\ (z_k^*,x_k^*)\in\lambda_k\widehat\partial\rho(z_k,x_k)$$

with $(z_k^*,x_k^*)\xrightarrow{w^*}0$ as $k\to\infty$. We need to prove that $\|(z_k^*,x_k^*)\|\to 0$. Take a sequence $(\widetilde z_k^*,\widetilde x_k^*)\in\widehat\partial\rho(z_k,x_k)$ satisfying $(z_k^*,x_k^*)=\lambda_k(\widetilde z_k^*,\widetilde x_k^*)$. Then, by the assumed well-posedness condition, find a sequence $\{\omega_k\}\in\Pi(x_k;F(z_k))$ which has a subsequence (without relabeling) converging to some $\bar\omega\in\Pi(\bar x;F(\bar z))$. Using the argument in the proof of Theorem 3.8, one can find a sequence $\widetilde\omega_k\in\Pi(x_k;F_{t,\bar\omega}(z_k))$ such that $\widetilde\omega_k\to\widetilde\omega\in\Pi(\bar x;F_{t,\bar\omega}(\bar z))$ and

$$(\widetilde z_k^*,\widetilde x_k^*)\in\widehat N((z_k,\widetilde\omega_k);\mathrm{gph}\,F_{t,\bar\omega})\ \text{ with }\ \|\widetilde x_k^*\|=1,$$

which implies that $\|x_k^*\|=\lambda_k\|\widetilde x_k^*\|\to 0$ as $k\to\infty$. Taking into account that

$$(z_k^*,x_k^*)=\lambda_k(\widetilde z_k^*,\widetilde x_k^*)\in\widehat N((z_k,\widetilde\omega_k);\mathrm{gph}\,F_{t,\bar\omega})\ \text{ with }\ \|x_k^*\|\to 0$$

and using again [14, Theorem 3.3], we ensure that $F_{t,\bar{\omega}}$ is PSNC at $(\bar{z}, \widetilde{\omega})$, and hence $\|z_k^*\| \to 0$ as $k \to \infty$. This completes the proof of the theorem. ■

REFERENCES

1. J.-P. Aubin, Lipschitz behavior of solutions to convex minimization problems, *Math. Oper. Res.*, **9** (1984), 87-111.
2. J. M. Borwein and J. R. Giles, The proximal normal formula in Banach space, *Trans. Amer. Math. Soc.*, **302** (1987), 371-381.
3. J. M. Borwein and Q. J. Zhu, *Techniques of Variational Analysis. An Introduction*, Springer, New York, to appear.
4. Bounkhel, M.: On the distance function associated with a set-valued mapping. *J. Nonlinear Convex Anal.*, **2** (2001), 265-283.
5. M. Bounkhel and L. Thibault, On various notions of regularity of sets in nonsmooth analysis. *Nonlinear Anal.*, **48** (2002), 223-246.
6. J. V. Burke, M. C. Ferris and M. Quian, On the Clarke subdifferential of the distance function of a closed set, *J. Math. Anal. Appl.*, **166** (1992), 199-213.
7. F. H. Clarke, Y. S. Ledyaev, R. J. Stern and P. R Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, New York, 1998.
8. B. Cornet and M.-O. Czarnecki, Existence of generalized equilibria: Necessary and sufficient conditions, *Nonlinear Anal.*, **44** (2001), 555-574.
9. Ioffe, A. D.: Proximal analysis and approximate subdifferentials, *J. London Math. Soc.*, **2** (1990), 175-192.
10. A. Jourani and L. Thibault, Metric regularity and subdifferential calculus in Banach space, *Set-Valued Anal.*, **3** (1995), 87-100.
11. A. Kruger, Depon. VINITI **1331**, Moscow, 1981.
12. B. S. Mordukhovich, Maximum principle in problems of time optimal control with nonsmooth constraints, *J. Appl. Math. Mech.*, **40** (1976), 960-969.
13. B. S.Mordukhovich, *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow, 1988.
14. B. S. Mordukhovich, Coderivative analysis of variational systems, *J. Global Optim.*, **28** (2004), 347-362.
15. B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, Vol. I: Basic Theory, Vol. II: Applications*, Springer, Berlin, to appear.
16. B. S. Mordukhovich and N. M. Nam, Subgradients of distance functions with applications to Lipschitzian stability, *Math. Progr. Ser. B*, to appear.
17. B. S. Mordukhovich and Y. Shao, Nonsmooth sequential analysis in Asplund spaces. *Trans. Amer. Math. Soc.*, **348** (1996), 1235-1280.

18. R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, 2nd ed., Springer, Berlin, 1993.

19. R. T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

20. L. Thibault, On subdifferentials of optimal value functions, *SIAM J. Control Optim.*, **29** (1991), 1019-1036.

21. Z. Wu and J. J. Ye, Equivalence among various derivatives and subdifferentialas of distance function, *J. Math. Anal. Appl.*, **282** (2003), 629-647.

Boris S. Mordukhovich and Nguyen Mau Nam
Department of Mathematics,
Wayne State University,
Detroit, MI 48202,
U.S.A.