Subdifferentials of Distance Functions in Banach Spaces

NG, Kwong Wing

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Master of Philosophy
in
Mathematics

The Chinese University of Hong Kong
September 2010

Thesis / Assessment Committee

Professor LEUNG Chi Wai (Chair)

Professor NG Kung Fu (Thesis Supervisor)

Professor FENG De Jun (Committee Member)

Professor LI Chong (External Examiner)

Abstract

This thesis discusses the generalized differential properties of several distance functions defined on Banach spaces, which are of paramount importance in variational analysis, optimization and many other areas. It is well recognized that the standard distance function, which measures the distance from a moving point to a fixed subset, is intrinsically nonsmooth, rendering the machinery of classical differential calculus insufficient for a comprehensive study. Among the various generalized differential devices invented to study such nonsmooth functions are the Fréchet subdifferential, the proximal subdifferential, a family of subdifferentials due to Mordukhovich, and their dual normal constructions. With a wealth of new tools, the generalized differential properties of the standard distance function have been thoroughly studied in the literature. However, there are a number of less acquainted generalizations of the standard distance function, including the generalized distance function, which denotes the distance from a moving point to a moving subset, and the perturbed distance function, which signifies the perturbed distance from a moving point to a fixed subset. Mainly based on the publications of Mordukhovich and Nam, together with the work of Wang, Li and Xu, estimates and alternative characterizations of various subdifferentials and normal objects related to these generalizations of the standard distance function are delineated and studied systematically in the thesis. A slight improvement of a theorem established by Wang, Li and Xu is also included.

摘要

本文旨在討論定義於巴拿赫空間 (Banach space) 上的若干距離函數之廣義微分性質。在變分分析、最優化理論等範疇中，距離函數 (distance function) 素來扮演著重要的角色。為人熟悉的標準距離函數 (standard distance function) 表示空間裡一移動點至一固定集合之距離；其固有的非光滑性使得傳統微分學的工具不適用於其分析上。為了解此等非光滑函數，廣義微分工具繼而興起，如 Fréchet 次微分， proximal 次微分， Mordukhovich 所研發之一系列次微分，還有與各種次微分對應的法向構作等。在林林總總的工具協助下，已有大量文獻致力研究標準距離函數的廣義微分性質，但從推廣標準距離函數而得到的新種距離函數卻仍有待廣泛研究。這些新種距離函數其中包括表示空間裡一移動點至一移動集合之距離的廣義距離函數 (generalized distance function)，以及量度空間裡一移動點至一固定集合之擾動距離的擾動距離函數 (perturbed distance function)。以 Mordukhovich-Nam 和 Wang-Li-Xu 的論文為基礎，本文有條不紊地描繪及估量與這些新種距離函數有關的各種次微分和法向構作，並對 Wang-Li-Xu 的一定理稍作推廣。

Acknowledgments

I wish to express my most heartfelt gratitude to my supervisor, Prof. K. F. Ng, for his kind understanding of the difficulties I encountered as a complete novice and an unhealthy student in the course of writing this thesis, and his unfailing patience explaining abstruse technicalities to me and proofreading the manuscript. I am also thankful to Dr. L. F. Cheung, who not only inspired me to work on this fruitful area but was also willing to share with me his thoughts on the subject.

In addition, grateful acknowledgment is due to the help of many friends, who spared their invaluable time to discuss with me the content of the manuscript in spite of their busy schedules. In particular, discussions with Mr. F. M. Lau and Mr. H. L. Lee helped me work out a number of details in the thesis.

Last but not least, I would like to thank my family, to whom I am deeply indebted. They endured my extended preoccupation with generosity and grace. Above all, writing the thesis caused my caring mother a great deal of pain and worry. It was with the indispensable support of my family that I managed to complete the thesis.

Contents

Introduction

Modern variational analysis may be regarded as an outgrowth of the traditional subjects of calculus of variations and mathematical programming. Nonsmooth functions, sets with nonsmooth boundaries and set-valued mappings, which arise naturally and ubiquitously in mathematics, are predominant in the framework of variational analysis. Conforming to the historical approach to optimization, which relies heavily on the theory of classical differential calculus, generalized differentiation lies at the heart of variational analysis.

The primary goal of this thesis is to explore the generalized differential properties of several distance functions defined on arbitrary Banach spaces. Distance functions are vital in optimization and variational analysis. They often appear in nonlinear programming and constrained optimization problems even with smooth initial data. For instance, distance functions were used to establish notable multiplier existence theorems in constrained optimization in [10] and to devise efficient algorithms for solving systems of nonlinear equations in [11] and [13]. Such results were mostly obtained via perturbation, penalization and approximation techniques.

Over the years, tremendous effort has been continually devoted to investigating the generalized differential properties of the *standard distance function* $d(\cdot, \Omega) : X \to \mathbb{R}$ defined by

$$d(x, \Omega) := \inf\{\|x - w\| : w \in \Omega\},$$

which measures the distance from a moving point in a Banach space X to a fixed nonempty subset $\Omega \subset X$ (see, for example, [4, 9, 12, 15, 25, 33, 38]). In spite of its intrinsic nonsmoothness, its global Lipschitz continuity has proved to be helpful in its study. Estimates and representations of its many subdifferentials frequently employ corresponding normal objects, enlargements and projections.

One possible extension of the standard distance function is the *generalized distance* *function* $\rho : \mathrm{dom}\, F \times X \to \mathbb{R}$ defined by

$$\rho(z, x) := d(x, F(z)) = \inf\{\|x - w\| : w \in F(z)\},$$

where Z and X are both Banach spaces, and the set-valued mapping $F : Z \rightrightarrows X$ serves to produce different subsets of X. The generalized distance function signifies the distance from a moving point in X to a moving subset of X. It was Rockafellar who first considered the generalized distance function at points belonging to $\mathrm{gph}\, F$ and proved in [32] that the local Lipschitz continuity of the generalized distance function is equivalent to the local Lipschitz-like property of F. In stark contrast to the standard distance function, the generalized distance function is in general neither locally Lipschitz nor locally lower semicontinuous, which has led to a lot of difficulties in its study. Estimates and representations of its subdifferentials do not only involve dual normal constructions, enlargements and projections, but often also perturbed projections and coderivatives.

Whether the point of interest belongs to $\mathrm{gph}\, F$ affects the generalized differential properties of the generalized distance function significantly. Some results pertaining to the case in which the point of interest belongs to $\mathrm{gph}\, F$ were proved by Thibault in [36], while an inconsiderable collection of formulae concerning the case in which the point of interest lies out of $\mathrm{gph}\, F$ are available in [6] and [7]. Proceeding further, an instructive observation is that the generalized distance function indeed belongs to a more general class of functions known as generalized marginal functions, which are in many instances drawn on to develop central theorems in duality theory of minimization problems (see [34]). Descriptions of subdifferentials of the generalized distance function and the generalized marginal function ascertained in [26] and [27] are surveyed in this thesis, covering whenever possible both the case in which the point of interest lies in $\mathrm{gph}\, F$ and that in which the point lies out of $\mathrm{gph}\, F$.

Besides the generalized distance function, another popular extension of the standard distance function is the *perturbed distance function* $d^J(\cdot, \Omega) : X \to \mathbb{R}$ defined by

$$d^J(x, \Omega) := \inf\{\|x - w\| + J(w) : w \in \Omega\},$$

which indicates the perturbed distance from a moving point in a Banach space X to a fixed nonempty subset $\Omega \subset X$, with the perturbation generated by a lower semicontinuous function $J : \Omega \to \mathbb{R}$. The perturbed distance function was first analyzed in [1] by Baranger, who proved that the set of points in a uniformly convex Banach space for which the perturbed minimization problem has a solution is a dense G_δ-subset, provided that J is bounded below. Since then, a multitude of existence results have been discovered (see [16, 17]) and applied to tackle optimal control problems governed by partial differential equations (see [2, 20, 28]). It should be noted that the convexity of Ω plays a principal role in the study of the perturbed distance function. Conclusions about subdifferentials of the perturbed distance function communicated in [37], which embrace both the case in which Ω is convex and that in which Ω is nonconvex, are examined systematically in this thesis.

The rest of the thesis is comprised of five chapters. Chapter 1 gives a brief overview of the preliminary materials to prepare for subsequent chapters. Chapter 2 gathers some fundamental estimates and alternative representations of Fréchet-like, limiting and singular subdifferentials of the generalized distance function. A major motivation in Chapter 2 is to characterize subdifferentials of the generalized distance function by means of dual normal constructions. Estimates of Fréchet-like and limiting subdifferentials are developed via their dual normal objects, enlargements, projections and perturbed projections while those of singular subdifferentials are acquired via coderivatives. Special assumptions utilized in this chapter are the criteria for well-posedness of the best approximation problem, and a simple sufficient condition for fulfilling one of the criteria is supplied. With the use of intermediate points, Chapter 3 continues to investigate other estimates of various subdifferentials of the generalized distance function, a number of which may be viewed as extensions of the analogous results obtained via projections in Chapter 2. A prominent establishment in Chapter 3 pertains to limiting subdifferentials of the generalized distance function in a Hilbert space setting and provides efficient conditions to guarantee the nonemptiness of projection sets as a by-product. Chapter 4 turns to study singular subdifferentials of the marginal function and the generalized marginal function, with an emphasis on reducing results to the corresponding ones for the standard distance function and the generalized distance function. As in Chapter 2, mixed coderivatives are employed in the derivation of upper estimates. Chapter 5 deals with the perturbed distance function. While reasonable estimates may be given generally for a few subdifferentials of the perturbed distance function, exact formulae are available at points which are self-solutions to the perturbed minimization problem, provided that some mild assumptions are satisfied. As in Chapter 4, reduction to the analogous results for the standard distance function is highlighted. This concludes the outline of the thesis.

Chapter 1

Preliminaries

In this chapter, basic definitions and notations to be used throughout the thesis are introduced. Most of these are standard in nonsmooth analysis and variational analysis.

1.1 Basic Notations and Conventions

Unless otherwise stated, X is always a real Banach space with dual space X^*. The norm on X and that on X^* are denoted by $\|\cdot\|_X$ and $\|\cdot\|_{X^*}$ respectively. When the meaning is clear from the context, both norms are conveniently denoted by $\|\cdot\|$. The canonical pairing on $X^* \times X$ is represented by $\langle \cdot, \cdot \rangle$ and the evaluation $x^*(x)$ is represented by $\langle x^*, x \rangle$. Adopting the usual notations, $\boldsymbol{B}_X$ and $\boldsymbol{B}_{X^*}$ stand for the *closed* unit balls, while $\boldsymbol{S}_X$ and $\boldsymbol{S}_{X^*}$ stand for the unit spheres, in X and X^* respectively. In general, the closed balls in X and X^* with radius $r > 0$ centered at $\bar{x}$ are denoted by $\boldsymbol{B}_X(\bar{x}, r)$ and $\boldsymbol{B}_{X^*}(\bar{x}, r)$ respectively. The symbols $\boldsymbol{S}_X(\bar{x}, r)$ and $\boldsymbol{S}_{X^*}(\bar{x}, r)$ are defined similarly. When two or three Banach spaces are involved, unless otherwise specified, Y and Z also denote real Banach spaces.

Let $X_1, X_2, \cdots, X_n$ be Banach spaces. The product space $\prod X := X_1 \times X_2 \times \cdots \times X_n$ is equipped with the ℓ_1-norm defined by

$$\|(x_1, x_2, \cdots, x_n)\|_{\prod X} := \|x_1\|_{X_1} + \|x_2\|_{X_2} + \cdots + \|x_n\|_{X_n}.$$

Note that $\prod X$ is also a Banach space with respect to the ℓ_1-norm. For convenience, the closed ball and the sphere in $\prod X$ with "radius" $r > 0$ centered at $(\bar{x}_1, \cdots, \bar{x}_n)$ are defined slightly differently from the above, namely

$$\boldsymbol{B}_{\prod X}((\bar{x}_1, \cdots, \bar{x}_n), r) := \{(x_1, \cdots, x_n) \in X_1 \times \cdots X_n : \|x_i - \bar{x}_i\| \leq r, \quad i = 1, \cdots, n\},$$

$$\boldsymbol{S}_{\prod X}((\bar{x}_1, \cdots, \bar{x}_n), r) := \{(x_1, \cdots, x_n) \in X_1 \times \cdots X_n : \|x_i - \bar{x}_i\| = r, \quad i = 1, \cdots, n\},$$

$\boldsymbol{B}_{\prod X} := \boldsymbol{B}_{\prod X}(0, 1)$, and $\boldsymbol{S}_{\prod X} := \boldsymbol{S}_{\prod X}(0, 1)$.

Let $\mathbb{R}$ denote the set of all real numbers, $\mathbb{R}_+$ denote the set of all *positive* real numbers and $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ denote the extended real line. Moreover, $\mathbb{N} := \{1, 2, 3, \cdots\}$ stands for the set of all natural numbers.

At this point, it is convenient to introduce the basic topological and geometrical notations that are needed later. For any subset $\Omega \subset X$, the notations $\operatorname{cl}\Omega$, $\operatorname{int}\Omega$, $\operatorname{co}\Omega$ and $\operatorname{bd}\Omega$ respectively stand for the *closure*, the *interior*, the *convex hull* and the *boundary* of Ω with respect to the norm topology of X. Likewise, $w\text{-}\operatorname{cl}\Omega$ indicates the *weak closure* of Ω, the closure of Ω with respect to the weak topology of X. The *conical hull* of Ω is defined by $\operatorname{cone}\Omega := \{ax \in X : a \geq 0 \text{ and } x \in X\}$. In particular, the *apex* $0 \in \operatorname{cone}\Omega$ and $\operatorname{cone}\Omega$ is nonempty. Conforming to the practice in convex analysis, Ω is said to be a *cone* if $\Omega = \operatorname{cone}\Omega$. Furthermore, for any subset $\Lambda \subset X^*$, the symbol $\operatorname{cl}^* \Lambda$ signifies the *weak* closure* of Λ, the closure of Λ with respect to the weak* topology of X^*.

As for convergence, there are several notations indicating different types of convergence. While "$\xrightarrow{w^*}$ " and "w^*-lim " denote *weak* convergence*, "$\xrightarrow{w}$ " and "w-lim " mean *weak convergence*. In addition, "$\rightarrow$ " and "lim " stand for the ordinary *norm convergence*, which is sometimes emphasized by the notation "$\xrightarrow{\|\cdot\|}$". If $\Omega \subset X$ and $\bar{x} \in \operatorname{cl}\Omega$, then $x \xrightarrow{\Omega} \bar{x}$ means $x \rightarrow \bar{x}$ with $x \in \Omega$. Let $f : X \rightarrow \overline{\mathbb{R}}$ and $\bar{x} \in X$. The notation $x \xrightarrow{f} \bar{x}$ means $x \rightarrow \bar{x}$ with $f(x) \rightarrow f(\bar{x})$; the strengthened version $x \xrightarrow{f+} \bar{x}$ means $x \rightarrow \bar{x}$ with $f(x) \geq f(\bar{x})$ and $f(x) \rightarrow f(\bar{x})$.

Let $f : X \rightarrow \overline{\mathbb{R}}$ be an extended real-valued function. The *effective domain* of f is given by $\operatorname{dom} f := \{x \in X : f(x) < \infty\}$. f is called a *proper* function if $\operatorname{dom} f \neq \varnothing$ and $f(x) \neq -\infty$ for all $x \in X$. f is described as *improper* if it is not proper.

Throughout the disquisition, arithmetic involving the empty set $\varnothing$ and the extended real numbers $\pm\infty$ is inevitable. Regarding the empty set, below are some of the most customary conventions:

$$\Omega + \varnothing = \varnothing; \qquad 0 \cdot \varnothing = \{0\}; \qquad a \cdot \varnothing = \varnothing \text{ for all } a \in \mathbb{R}\backslash\{0\};$$

$$\inf \varnothing = \infty; \qquad \sup \varnothing = -\infty; \qquad \|\varnothing\| = \infty.$$

Regarding infinity, the following common conventions are adopted:

$$0 \cdot \infty = \infty \cdot 0 = 0; \qquad 0 \cdot (-\infty) = (-\infty) \cdot 0 = 0; \qquad -(-\infty) = \infty;$$

$$x + \infty = \infty + x = \infty \quad \text{and} \quad x - \infty = (-\infty) + x = -\infty \qquad \text{for all } x \in \mathbb{R};$$

$$x \cdot \infty = \infty \cdot x = \infty \quad \text{and} \quad x \cdot (-\infty) = (-\infty) \cdot x = -\infty \qquad \text{for all } x > 0;$$

$$x \cdot \infty = \infty \cdot x = -\infty \quad \text{and} \quad x \cdot (-\infty) = (-\infty) \cdot x = \infty \qquad \text{for all } x < 0.$$

The expressions $\infty - \infty$, $(-\infty) + \infty$ and $\dfrac{\infty}{\infty}$ are undefined.

1.2 Fundamental Results in Banach Space Theory and Variational Analysis

This section presents a few standard theorems in Banach space theory and variational analysis.

Ekeland's Variational Principle is commonly regarded as the first published general variational principle. It turns out to be a characterization of complete metric spaces. A proof may be found in any standard text on nonsmooth analysis or variational analysis, such as [15] and [25].

Theorem 1.2.1 (Ekeland's Variational Principle). *Let (X, d) be a metric space.*

(a) *Assume that X is complete and that $f : X \to \overline{\mathbb{R}}$ is a proper lower semicontinuous function bounded below. Suppose there exist $x_0 \in X$ and $\varepsilon > 0$ satisfying*

$$f(x_0) \leq \inf_{x \in X} f(x) + \varepsilon.$$

Then for any $\lambda > 0$, there exists $\bar{x} \in X$ such that

(i) $f(\bar{x}) \leq f(x_0)$,

(ii) $d(\bar{x}, x_0) \leq \lambda$, *and*

(iii) $f(x) + \frac{\varepsilon}{\lambda} d(x, \bar{x}) > f(\bar{x})$ *for all* $x \neq \bar{x}$.

(b) *Conversely, X is complete if for every Lipschitz continuous function $f : X \to \mathbb{R}$ bounded below and every $\varepsilon > 0$, there exists $\bar{x} \in X$ such that*

(i') $f(\bar{x}) \leq \inf_{x \in X} f(x) + \varepsilon$, *and*

(iii') $f(x) + \varepsilon d(x, \bar{x}) > f(\bar{x})$ *for all* $x \neq \bar{x}$.

Entailed below are two basic results in the theory of Banach spaces. The first one points to the lower semicontinuity of the norms $\|\cdot\|_X$ and $\|\cdot\|_{X^*}$ with respect to the weak topology of X and the weak* topology of X^* respectively; the second one is a useful characterization of reflexive spaces. Both results, together with their proofs, may be found in [22].

Theorem 1.2.2. **(a)** *Let* $\{x_\alpha\}_{\alpha\in I}$ *be a net in* X *such that* $x_\alpha \xrightarrow{w} \bar{x}$ *for some* $\bar{x} \in X$. *Then* $\liminf_{\alpha} \|x_\alpha\| \geq \|\bar{x}\|$. *In other words,* $\|\cdot\|_X$ *is lower semicontinuous with respect to the weak topology of* X.

(b) *Let* $\{x_\alpha^*\}_{\alpha\in I}$ *be a net in* X^* *such that* $x_\alpha^* \xrightarrow{w^*} \bar{x}^*$ *for some* $\bar{x}^* \in X^*$. *Then* $\liminf_{\alpha} \|x_\alpha^*\| \geq \|\bar{x}^*\|$. *In other words,* $\|\cdot\|_{X^*}$ *is lower semicontinuous with respect to the weak* topology of* X^*.

Remarks 1.2.3. (i) If X is *finite dimensional*, then its weak topology and its norm topology coincide. It follows that $\|\cdot\|_X$ is *continuous* with respect to the weak topology of X.

(ii) If X is *finite dimensional*, then the weak* topology and the norm topology of X^* coincide. It follows that $\|\cdot\|_{X^*}$ is *continuous* with respect to the weak* topology of X^*.

Theorem 1.2.4. *A normed space is reflexive if and only if each of its bounded sequences has a weakly convergent subsequence.*

1.3 Set-Valued Mappings

This section introduces *set-valued mappings*, which are in stark contrast to the usual single-valued functions. Since the investigation in this section does not draw on any norm structure, X and Y may be taken as mere *topological spaces*.

As suggested by the terminology, a *set-valued mapping* or *multifunction* F between X and Y, denoted by $F : X \rightrightarrows Y$ or $F : X \to \mathcal{P}(Y)$, is a mapping from X into the power set $\mathcal{P}(Y)$ of Y.

Just as to single-valued functions, the following specifications are fundamental to set-valued mappings.

Definition 1.3.1. Let $F : X \rightrightarrows Y$ be a set-valued mapping, $\Omega \subset X$ and $\Theta \subset Y$.

(a) The ***domain of*** F is $\operatorname{dom} F := \{x \in X : F(x) \neq \varnothing\}$.

(b) The ***range of*** F is $\operatorname{range} F := \{y \in Y : y \in F(x) \text{ for some } x \in X\}$.

(c) The ***image of*** Ω ***under*** F is $F(\Omega) := \{y \in Y : y \in F(x) \text{ for some } x \in \Omega\}$.

(d) The ***inverse image of*** Θ ***under*** F is $F^{-1}(\Theta) := \{x \in X : F(x) \cap \Theta \neq \varnothing\}$.

(e) The ***graph of*** F is $\operatorname{gph} F := \{(x, y) \in X \times Y : y \in F(x)\}$.

Definition 1.3.2. Let $F : X \rightrightarrows Y$ be a set-valued mapping.

(a) F is said to be ***closed-valued*** (respectively ***convex-valued***) if $F(x)$ is closed (respectively convex) for all $x \in X$.

(b) F is said to be ***closed-graph*** if $\operatorname{gph} F$ is closed.

As the basic building blocks in the development of a full calculus, limit concepts form an integral part of the theory of set-valued mappings. However, limit concepts for set-valued mappings are much more complicated than their counterparts for single-valued functions. Although only upper limits are needed in the forthcoming disquisition, other related limits are also covered below for the sake of completeness.

Definition 1.3.3. Let $\bar{x} \in X$ and $F : X \rightrightarrows Y$ be a set-valued mapping.

(a) The ***sequential Painlevé-Kuratowski upper*** or ***outer limit of*** F ***as*** $x \to \bar{x}$ is defined by

$$\operatorname*{Lim\,sup}_{x\to\bar{x}} F(x) := \big\{y \in Y : \text{there exist sequences } \{x_k\}_{k=1}^{\infty} \subset X \text{ and } \{y_k\}_{k=1}^{\infty} \subset Y \text{ with } x_k \to \bar{x} \text{ and } y_k \to y \text{ such that } y_k \in F(x_k) \text{ for all } k \in \mathbb{N}\big\}.$$

(b) The ***sequential Painlevé-Kuratowski lower*** or ***inner limit of*** F ***as*** $x \to \bar{x}$ is defined by

$$\operatorname*{Lim\,inf}_{x\to\bar{x}} F(x) := \big\{y \in Y : \text{for any sequence } \{x_k\}_{k=1}^{\infty} \subset X \text{ with } x_k \to \bar{x}, \text{ there exists a sequence } \{y_k\}_{k=1}^{\infty} \subset Y \text{ with } y_k \to y \text{ such that } y_k \in F(x_k) \text{ for all } k \in \mathbb{N}\big\}.$$

(c) Suppose $\operatorname*{Lim\,sup}_{x\to\bar{x}} F(x) = \operatorname*{Lim\,inf}_{x\to\bar{x}} F(x)$. The ***sequential Painlevé-Kuratowski limit of*** F ***as*** $x \to \bar{x}$ is defined by

$$\operatorname*{Lim}_{x\to\bar{x}} F(x) := \operatorname*{Lim\,sup}_{x\to\bar{x}} F(x) = \operatorname*{Lim\,inf}_{x\to\bar{x}} F(x).$$

Remark 1.3.4. Analogous to the familiar inequality involving the usual upper limit and lower limit for single-valued functions, a conspicuous relation between the upper limit and the lower limit for set-valued mappings defined above is

$$\operatorname*{Lim\,inf}_{x\to\bar{x}} F(x) \subset \operatorname*{Lim\,sup}_{x\to\bar{x}} F(x).$$

1.4 Enlargements and Projections

The investigation in subsequent chapters employs extensively the devices of *enlargements* and *projections*, which are the subjects of this section.

Definition 1.4.1. Let $\Omega \subset X$ be a nonempty subset. The ***standard distance function*** $d(\cdot, \Omega) : X \to \mathbb{R}$ ***associated with*** Ω is defined by

$$d(x, \Omega) := \inf\{\|x - w\| : w \in \Omega\}.$$

Remark 1.4.2. An immediate consequence of the above definition is that $d(\cdot, \Omega) \equiv d(\cdot, \operatorname{cl}\Omega)$.

Definition 1.4.3. Let $\Omega \subset X$ be a nonempty subset, $\bar{x} \in X$ and $r \geq 0$.

(a) The r***-enlargement of*** Ω is defined by

$$\Omega_r := \{x \in X : d(x, \Omega) \leq r\}.$$

(b) The r***-thickening of*** Ω is defined by

$$\Omega_r^t := \Omega + r\boldsymbol{B}_X.$$

(c) The ***projection set of*** $\bar{x}$ ***onto*** Ω is defined by

$$\Pi(\bar{x}, \Omega) := \{w \in \Omega : \|w - \bar{x}\| = d(\bar{x}, \Omega)\}.$$

(d) The r***-perturbed projection set of*** $\bar{x}$ ***onto*** Ω is defined by

$$\Pi_r(\bar{x}, \Omega) := \{w \in \Omega : \|w - \bar{x}\| \leq d(\bar{x}, \Omega) + r\}.$$

Remark 1.4.4. While Ω_r, $\Pi(\bar{x}, \Omega)$ and $\Pi_r(\bar{x}, \Omega)$ are necessarily closed, Ω_r^t is *not* closed in general.

In many applications, it is important to have *nonempty* projection sets. In this light, a simple sufficient condition to ensure nonvoid projection sets is hereby included.

Proposition 1.4.5. *Let X be reflexive and $\Omega \subset X$ be a nonempty weakly closed subset. Then for any $x \in X$, $\Pi(x, \Omega) \neq \varnothing$.*

Proof. Let $x \in X$. For each $k \in \mathbb{N}$, there exists $w_k \in \Omega$ such that

$$d(x, \Omega) + \frac{1}{k} > \|x - w_k\| \geq \|w_k\| - \|x\|. \tag{1.1}$$

Then $\|w_k\| \leq \|x\| + d(x, \Omega) + 1$ for all $k \in \mathbb{N}$ and $\{w_k\}_{k=1}^{\infty}$ is a bounded sequence in Ω. In view of the reflexivity of X, Theorem 1.2.4 implies that $\{w_k\}_{k=1}^{\infty}$ has a weakly convergent subsequence. By passing to this subsequence if necessary, assume that $w_k \xrightarrow{w} \bar{w}$ for some $\bar{w} \in X$. Since Ω is weakly closed, $\bar{w} \in \Omega$ and hence $d(x, \Omega) \leq \|x - \bar{w}\|$. In light of the lower semicontinuity of $\|\cdot\|$ with respect to the weak topology of X, it follows from (1.1) that

$$d(x, \Omega) = \liminf_{k \to \infty} \left(d(x, \Omega) + \frac{1}{k} \right) \geq \liminf_{k \to \infty} \|x - w_k\| \geq \|x - \bar{w}\|.$$

By definition, $\bar{w} \in \Pi(x, \Omega) \neq \varnothing$. □

It is evident from Definition 1.4.3 that enlargements and thickenings are closely related concepts. Their precise relationship is stated in the next result.

Proposition 1.4.6. (cf. [29, Lemma 27]) *Let $\Omega \subset X$ be a nonempty subset and $r \geq 0$. Then*

(a) $\Omega_r = \operatorname{cl} \Omega_r^t$;

(b) $\Omega_r = \Omega_r^t$ *if and only if* $\Pi(x, \Omega) \neq \varnothing$ *for all* $x \in X$ *with* $d(x, \Omega) = r$.

Proof. **(a)** Let $x \in \Omega_r^t$. By definition, $x = w + ru$ for some $w \in \Omega$ and $u \in \boldsymbol{B}_X$. Then $d(x, \Omega) \leq \|x - w\| = \|ru\| \leq r$ and $x \in \Omega_r$. Hence $\Omega_r \supset \Omega_r^t$ and $\Omega_r \supset \operatorname{cl} \Omega_r^t$ follows from noting that Ω_r is closed.

Consider the opposite inclusion. Let $x \in \Omega_r$. By definition, $d(x, \Omega) \leq r$. For any $\varepsilon > 0$, there exists $w_\varepsilon \in \Omega$ such that $r + \varepsilon \geq d(x, \Omega) + \varepsilon > \|x - w_\varepsilon\|$.

It follows that $\frac{\|x-w_\varepsilon\|}{r+\varepsilon} \le 1$ and $\frac{x-w_\varepsilon}{r+\varepsilon} \in \boldsymbol{B}_X$. Take $y_\varepsilon = w_\varepsilon + r\left(\frac{x-w_\varepsilon}{r+\varepsilon}\right)$. Then $y_\varepsilon \in \Omega + r\boldsymbol{B}_X = \Omega_r^t$. Moreover, observe that

$$\begin{aligned}\|y_\varepsilon - x\| &= \left\|w_\varepsilon + r\left(\frac{x-w_\varepsilon}{r+\varepsilon}\right) - x\right\| = \left|\frac{r}{r+\varepsilon} - 1\right| \|x - w_\varepsilon\| \\ &\le \left(\frac{\varepsilon}{r+\varepsilon}\right)(r+\varepsilon) = \varepsilon\end{aligned}$$

and hence $y_\varepsilon \in \boldsymbol{B}_X(x,\varepsilon)$. Consequently, $y_\varepsilon \in \boldsymbol{B}_X(x,\varepsilon) \cap \Omega_r^t \neq \varnothing$ for any $\varepsilon > 0$ and $x \in \operatorname{cl}\Omega_r^t$. Thus $\Omega_r \subset \operatorname{cl}\Omega_r^t$. The desired equality holds.

(b) Suppose $\Omega_r = \Omega_r^t$. Let $x \in X$ with $d(x,\Omega) = r$. Then $x \in \Omega_r = \Omega_r^t$. By definition, $x = w + ru$ for some $w \in \Omega$ and $u \in \boldsymbol{B}_X$. It follows that $\|x - w\| = \|ru\| \le r = d(x,\Omega)$. On the other hand, $w \in \Omega$ implies that $\|x - w\| \ge d(x,\Omega)$. As a result, $w \in \Pi(x,\Omega) \neq \varnothing$.

Suppose $\Pi(x,\Omega) \neq \varnothing$ for all $x \in X$ with $d(x,\Omega) = r$. Let $x \in \Omega_r^t$. There exist $w \in \Omega$ and $u \in \boldsymbol{B}_X$ such that $x = w + ru$. Thus $d(x,\Omega) \le \|x - w\| = \|ru\| \le r$ and $x \in \Omega_r$, implying $\Omega_r^t \subset \Omega_r$. Conversely, let $x \in \Omega_r$. By definition, $d(x,\Omega) \le r$. If $d(x,\Omega) = r$, then $\Pi(x,\Omega) \neq \varnothing$ by assumption and $\|x - w_1\| = d(x,\Omega) = r$ for some $w_1 \in \Omega$. If $d(x,\Omega) < r$, then $\|x - w_2\| < d(x,\Omega) + \delta < r$ for some $\delta > 0$ and $w_2 \in \Omega$. In both cases, there exist $w \in \Omega$ and $u \in \boldsymbol{B}_X$ such that $x - w = ru$ or $x = w + ru$. Therefore $x \in \Omega_r^t$ and $\Omega_r \subset \Omega_r^t$. As a result, $\Omega_r = \Omega_r^t$. □

Corollary 1.4.7. ([29, Lemma 27]) *Let X be reflexive, $\Omega \subset X$ be a nonempty weakly closed subset and $r \ge 0$. Then $\Omega_r = \Omega_r^t$.*

Proof. Since X is reflexive and $\Omega \subset X$ is a nonempty weakly closed subset, Proposition 1.4.5 implies that $\Pi(x,\Omega) \neq \varnothing$ for all $x \in X$. The conclusion then follows from Proposition 1.4.6(b) immediately. □

1.5 Subdifferentials

This section focuses on a number of popular derivative-like constructions in variational analysis devised for the study of nonsmooth functions.

Recall the *subdifferential* in convex analysis, which was originally introduced for convex functions, and the *proximal subdifferential* in nonsmooth analysis, which was first intended for proper lower semicontinuous functions.

Definition 1.5.1. Let $f : X \to \overline{\mathbb{R}}$ be finite at $\bar{x} \in X$.

(a) The ***subdifferential*** (*in the sense of convex analysis*) ***of*** f ***at*** $\bar{x}$ is defined by

$$\partial^c f(\bar{x}) := \{x^* \in X^* : f(x) \geq f(\bar{x}) + \langle x^*, x - \bar{x}\rangle \text{ for all } x \in X\}.$$

The elements of this set are known as ***subgradients of*** f ***at*** $\bar{x}$.

(b) The ***proximal subdifferential of*** f ***at*** $\bar{x}$ is defined by

$$\begin{aligned}\partial^p f(\bar{x}) := \big\{x^* \in X^* : \;& \text{there exist } \delta > 0 \text{ and } \eta > 0 \text{ such that} \\ & \langle x^*, x - \bar{x}\rangle \leq f(x) - f(\bar{x}) + \eta\|x - \bar{x}\|^2 \\ & \text{for all } x \in \boldsymbol{B}_X(\bar{x}, \delta)\big\}.\end{aligned}$$

The elements of this set are known as ***proximal subgradients of*** f ***at*** $\bar{x}$.

Remark 1.5.2. Observe that $\partial^c f(\bar{x})$ is *closed* and *convex.* In particular, if X is *reflexive,* then $\partial^c f(\bar{x})$ is *weakly*-closed.* On the other hand, $\partial^p f(\bar{x})$ is *convex* but *not necessarily closed.*

While the aforementioned subdifferentials have been extensively studied in the literature, another class of subdifferentials has been more recently developed by Mordukhovich and his collaborators to provide alternative approximating instruments. See the comprehensive two-volume monograph [25] for further discussion.

Definition 1.5.3. Let $f : X \to \overline{\mathbb{R}}$ be finite at $\bar{x} \in X$ and $\varepsilon \geq 0$. The (***Fréchet-like***) ***ε-subdifferential of f at $\bar{x}$*** is defined by

$$\widehat{\partial}_\varepsilon f(\bar{x}) := \left\{ x^* \in X^* : \liminf_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \geq -\varepsilon \right\}.$$

The elements of this set are known as (***Fréchet-like***) ***ε-subgradients of f at $\bar{x}$***. In particular, $\widehat{\partial} f(\bar{x}) := \widehat{\partial}_0 f(\bar{x})$ is called the ***Fréchet subdifferential of f at $\bar{x}$*** and its elements are known as ***Fréchet subgradients of f at $\bar{x}$.***

Remarks 1.5.4. (i) Observe that $\widehat{\partial}_\varepsilon f(\bar{x})$ is *closed* and *convex*. In particular, if X is *reflexive*, then $\widehat{\partial}_\varepsilon f(\bar{x})$ is *weakly*-closed.*

(ii) Note the *monotone* property of $\widehat{\partial}_\varepsilon f(\bar{x})$ with respect to ε: if $0 \leq \varepsilon_1 \leq \varepsilon_2$, then $\widehat{\partial}_{\varepsilon_1} f(\bar{x}) \subset \widehat{\partial}_{\varepsilon_2} f(\bar{x})$.

The above definition gives a paramount characterization of $\widehat{\partial}_\varepsilon f(\bar{x})$.

Proposition 1.5.5. *Let $f : X \to \overline{\mathbb{R}}$ be finite at $\bar{x} \in X$ and $\varepsilon \geq 0$. Then $x^* \in \widehat{\partial}_\varepsilon f(\bar{x})$ if and only if for any $\gamma > 0$, there exists $\delta > 0$ such that for all $x \in X$ with $\|x - \bar{x}\| \leq \delta$,*

$$\langle x^*, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) + (\varepsilon + \gamma)\|x - \bar{x}\|.$$

In other words, $x^* \in \widehat{\partial}_\varepsilon f(\bar{x})$ if and only if for any $\gamma > 0$, the function $\psi : X \to \overline{\mathbb{R}}$ defined by $\psi(x) = f(x) - f(\bar{x}) - \langle x^*, x - \bar{x} \rangle + (\varepsilon + \gamma)\|x - \bar{x}\|$ attains a local minimum at $\bar{x}$.

Proof. Consider the first assertion of the proposition. Assume $x^* \in \widehat{\partial}_\varepsilon f(\bar{x})$. By definition,

$$\ell := \sup_{\delta > 0} \inf_{0 < \|x - \bar{x}\| \leq \delta} \frac{f(x) - f(\bar{x}) - \langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} = \liminf_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \geq -\varepsilon.$$

Let $\gamma > 0$. Suppose ℓ is finite. There exists $\delta_1 > 0$ such that for all $x \in X$ with $0 < \|x - \bar{x}\| \leq \delta_1$,

$$\frac{f(x) - f(\bar{x}) - \langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \geq \ell - \gamma \geq -\varepsilon - \gamma,$$

that is,

$$\langle x^*, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) + (\varepsilon + \gamma)\|x - \bar{x}\|. \tag{1.2}$$

Otherwise $\ell = \infty$. There exists $\delta_2 > 0$ such that for all $x \in X$ with $0 < \|x - \bar{x}\| \leq \delta_2$,

$$\frac{f(x) - f(\bar{x}) - \langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \geq 0,$$

which, upon rearrangement, produces

$$\langle x^*, x - \bar{x}\rangle \leq f(x) - f(\bar{x}) \leq f(x) - f(\bar{x}) + (\varepsilon + \gamma)\|x - \bar{x}\|. \tag{1.3}$$

In both cases, in view of (1.2) and (1.3), there exists $\delta > 0$ such that the same inequality holds for all $x \in X$ with $0 < \|x - \bar{x}\| \leq \delta$. Moreover, equality trivially holds for $x = \bar{x}$. This proves one implication.

For the opposite implication, let $\gamma > 0$. By assumption, there exists $\delta > 0$ such that for all $x \in X$ with $\|x - \bar{x}\| \leq \delta$,

$$\langle x^*, x - \bar{x}\rangle \leq f(x) - f(\bar{x}) + (\varepsilon + \gamma)\|x - \bar{x}\|. \tag{1.4}$$

In particular, for all $x \in X$ with $0 < \|x - \bar{x}\| \leq \delta$, rearranging (1.4) shows

$$\frac{f(x) - f(\bar{x}) - \langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \geq -\varepsilon - \gamma.$$

Passing to the limit, one sees that $\liminf_{x \to \bar{x}} \dfrac{f(x) - f(\bar{x}) - \langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \geq -\varepsilon - \gamma$, which, since $\gamma > 0$ is arbitrary, reduces to

$$\liminf_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \geq -\varepsilon.$$

By definition, $x^* \in \widehat{\partial}_\varepsilon f(\bar{x})$ and the other implication holds.

The second assertion of the proposition follows from the first by noting that $\psi(\bar{x}) = 0$.

□

For the purpose of the subsequent exposition, ε-subgradients of the standard distance function are of special interest.

Proposition 1.5.6. *Let $f : X \to \overline{\mathbb{R}}$ be finite at $\bar{x} \in X$ and $\varepsilon \geq 0$. Suppose f is locally Lipschitz at $\bar{x}$ with rank $\ell \geq 0$. Then $\|x^*\| \leq \ell + \varepsilon$ for all $x^* \in \widehat{\partial}_\varepsilon f(\bar{x})$.*

Proof. Let $x^* \in \widehat{\partial}_\varepsilon f(\bar{x})$. Since f is locally Lipschitz at $\bar{x}$ with rank ℓ, there exists $\delta_1 > 0$ such that for all $x \in X$ with $\|x - \bar{x}\| \leq \delta_1$,

$$|f(x) - f(\bar{x})| \leq \ell\|x - \bar{x}\|. \tag{1.5}$$

Let $\eta > 0$. In light of Proposition 1.5.5, there exists $\delta_1 \geq \delta_2 > 0$ such that for all $x \in X$ with $\|x - \bar{x}\| \leq \delta_2$,

$$\begin{aligned}\langle x^*, x - \bar{x}\rangle &\leq f(x) - f(\bar{x}) + (\varepsilon + \eta)\|x - \bar{x}\| \\ &\leq \ell\|x - \bar{x}\| + (\varepsilon + \eta)\|x - \bar{x}\| \\ &= (\ell + \varepsilon + \eta)\|x - \bar{x}\|,\end{aligned}$$

where the second inequality follows from (1.5). Using the linearity of x^*, one has

$$\|x^*\| = \sup_{x \neq \bar{x}} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} = \sup_{0 < \|x - \bar{x}\| \leq \delta_2} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \leq \ell + \varepsilon + \eta.$$

Since $\eta > 0$ is arbitrary, $\|x^*\| \leq \ell + \varepsilon$. This completes the proof of the assertion. □

Proposition 1.5.7. (cf. [18, Proposition 1.5]) *Let $\Omega \subset X$ be a nonempty subset, $\bar{x} \in X$ and $\varepsilon \geq 0$. Then for any $x^* \in \widehat{\partial}_\varepsilon d(\bar{x}, \Omega)$,*

(a) $\|x^*\| \leq 1 + \varepsilon$;

(b) $1 - \varepsilon \leq \|x^*\| \leq 1 + \varepsilon$ *if it is further supposed that $\bar{x} \notin \operatorname{cl}\Omega$.*

Proof. **(a)** The conclusion follows from Proposition 1.5.6 readily by noting that $d(\cdot, \Omega)$ is Lipschitz with rank 1.

(b) Let $x^* \in \widehat{\partial}_\varepsilon d(\bar{x}, \Omega)$ and $\eta > 0$. By **(a)**, $\|x^*\| \leq 1 + \varepsilon$. It suffices to show that $\|x^*\| \geq 1 - \varepsilon$. Using Proposition 1.5.5, there exists $\delta > 0$ such that for all $x \in X$ with $\|x - \bar{x}\| \leq \delta$,

$$\langle x^*, x - \bar{x}\rangle \leq d(x, \Omega) - d(\bar{x}, \Omega) + (\varepsilon + \eta)\|x - \bar{x}\|. \tag{1.6}$$

Note that $\bar{x} \notin \operatorname{cl}\Omega$ implies $d(\bar{x}, \Omega) > 0$. Let $0 < t < \min\left\{1, \frac{\delta}{2d(\bar{x},\Omega)}\right\}$. Then $(1 + t^2)d(\bar{x}, \Omega) > d(\bar{x}, \Omega)$ implies $(1 + t^2)d(\bar{x}, \Omega) > \|\bar{x} - w_t\| > 0$ for some $w_t \in \Omega$, or equivalently, $d(\bar{x}, \Omega) > \frac{\|\bar{x}-w_t\|}{1+t^2}$. Let $y_t = (1-t)\bar{x} + tw_t$. Check that

$$\bar{x} - y_t = t(\bar{x} - w_t), \quad y_t - w_t = (1-t)(\bar{x} - w_t), \text{ and}$$
$$\|y_t - \bar{x}\| = t\|\bar{x} - w_t\| < \frac{\delta\|\bar{x} - w_t\|}{2d(\bar{x}, \Omega)} < \frac{\delta\|\bar{x} - w_t\|}{(1+t^2)d(\bar{x}, \Omega)} < \frac{\delta\|\bar{x} - w_t\|}{\|\bar{x} - w_t\|} = \delta.$$

Putting $x = y_t$ in (1.6) yields

$$\begin{aligned}
&\langle x^*, t(w_t - \bar{x})\rangle = \langle x^*, y_t - \bar{x}\rangle \\
&\le d(y_t, \Omega) - d(\bar{x}, \Omega) + (\varepsilon + \eta)\|y_t - \bar{x}\| \\
&\le \|y_t - w_t\| - d(\bar{x}, \Omega) + (\varepsilon + \eta)\|y_t - \bar{x}\| \\
&\le (1-t)\|\bar{x} - w_t\| - \frac{\|\bar{x} - w_t\|}{1+t^2} + (\varepsilon + \eta)t\|\bar{x} - w_t\| \\
&= \left(\frac{-t^3 + t^2 - t}{1+t^2}\right)\|\bar{x} - w_t\| + (\varepsilon + \eta)t\|\bar{x} - w_t\|,
\end{aligned}$$

which can be rearranged as

$$\frac{\langle x^*, \bar{x} - w_t\rangle}{\|\bar{x} - w_t\|} \ge \frac{t^2 - t + 1}{1+t^2} - (\varepsilon + \eta).$$

In view of of this inequality,

$$\|x^*\| = \sup_{u \ne 0} \frac{\langle x^*, u\rangle}{\|u\|} \ge \frac{\langle x^*, \bar{x} - w_t\rangle}{\|\bar{x} - w_t\|} \ge \frac{t^2 - t + 1}{1+t^2} - (\varepsilon + \eta). \tag{1.7}$$

Since $\eta > 0$ is arbitrary, letting $t \to 0$ in (1.7) shows that $\|x^*\| \ge 1 - \varepsilon$. The result is verified. □

The next proposition states one of the most significant relationships of the preceding subdifferentials. The reader may refer to [34, Proposition 9.1.9] for more details.

Proposition 1.5.8. *Let $f : X \to \overline{\mathbb{R}}$ be finite at $\bar{x} \in X$. Then $\partial^c f(\bar{x}) \subset \partial^p f(\bar{x}) \subset \widehat{\partial} f(\bar{x})$. If f is further supposed to be locally Lipschitz at $\bar{x}$ with rank $\ell \ge 0$, then $\partial^c f(\bar{x}) \subset \partial^p f(\bar{x}) \subset \widehat{\partial} f(\bar{x}) \subset \ell \boldsymbol{B}_{X^*}$.*

At this point, it is worthwhile to digress from the introduction of subdifferential constructions to consider a special class of Banach spaces and an essential topological property.

Definition 1.5.9. A Banach space X is said to be an ***Asplund space*** if every continuous convex function defined on a nonempty open convex subset $D \subset X$ is Fréchet differentiable at each point of some dense G_δ subset of D.

An equivalent characterization is that *an Asplund space is a Banach space whose separable subspaces have separable duals.* Asplund spaces are not rare. Indeed, the class of all Asplund spaces is large enough to include all *reflexive spaces* and in particular, all *Hilbert spaces.*

Definition 1.5.10. Let $\Omega \subset X$ be a nonempty subset and $\bar{x} \in X$. Ω is said to be ***locally closed at*** $\bar{x}$ if there exists a neighbourhood U of $\bar{x}$ such that $U \cap \Omega$ is closed.

One of the most important calculus rules for ε-subdifferentials is the so-called *fuzzy sum rule*, a version of which is catered for Asplund spaces. A proof may be found in [25, Theorem 2.33].

Theorem 1.5.11 (Semi-Lipschitz Fuzzy Sum Rule for ε-subdifferentials). *Let X be an Asplund space, $\varphi_i : X \to \overline{\mathbb{R}}$ be proper functions, where $i = 1, 2$, and $\bar{x} \in \operatorname{dom} \varphi_1 \cap \operatorname{dom} \varphi_2$. Suppose φ_1 is locally Lipschitz at $\bar{x}$ and φ_2 is lower semicontinuous on a neighbourhood of $\bar{x}$. Then for any $\varepsilon \geq 0$ and $\eta > 0$,*

$$\widehat{\partial}_\varepsilon(\varphi_1 + \varphi_2)(\bar{x}) \subset \bigcup \big\{ \widehat{\partial}\varphi_1(x_1) + \widehat{\partial}\varphi_2(x_2) + (\varepsilon + \eta)\boldsymbol{B}_{X^*} : x_i \in \boldsymbol{B}_X(\bar{x}, \eta), \\ |\varphi_i(x_i) - \varphi_i(\bar{x})| \leq \eta,\ i = 1, 2 \big\}.$$

The last two derivative-like constructions considered in this section are limiting ones built upon ε-subdifferentials.

Definition 1.5.12. Let $f : X \to \overline{\mathbb{R}}$ be finite at $\bar{x} \in X$.

(a) The ***limiting subdifferential*** or ***basic subdifferential of*** f ***at*** $\bar{x}$ is defined by

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

The elements of this set are known as ***limiting subgradients*** or ***basic subgradients of f at $\bar{x}$.***

(b) The ***singular subdifferential of f at $\bar{x}$*** is defined by

$$\partial^\infty f(\bar{x}) := \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{f} \bar{x} \\ \varepsilon, \lambda \downarrow 0}} \lambda \widehat{\partial}_\varepsilon f(x).$$

The elements of this set are known as ***singular subgradients of f at $\bar{x}$.***

Remark 1.5.13. If X is an *Asplund space* and f is *lower semicontinuous on a neighbourhood of* $\bar{x}$, then the *limiting subdifferential* and the *singular subdifferential* of f at $\bar{x}$ admit the simpler representations

$$\partial f(\bar{x}) = \operatorname*{Lim\,sup}_{x \xrightarrow{f} \bar{x}} \widehat{\partial} f(x) \quad \text{and} \quad \partial^\infty f(\bar{x}) = \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{f} \bar{x} \\ \lambda \downarrow 0}} \lambda \widehat{\partial} f(x).$$

1.6 Sets of Normals

In this section, normal objects dual to the derivative-like constructions in the previous section are considered.

Definition 1.6.1. Let $\Omega \subset X$ and $\bar{x} \in \Omega$.

(a) The ***normal cone*** (*in the sense of convex analysis*) ***to*** Ω ***at*** $\bar{x}$ is defined by

$$N^c(\bar{x};\Omega) := \{x^* \in X^* : \langle x^*, x - \bar{x}\rangle \leq 0 \text{ for all } x \in \Omega\}.$$

The elements of this set are known as ***normals to*** Ω ***at*** $\bar{x}$.

(b) The ***proximal normal cone to*** Ω ***at*** $\bar{x}$ is defined by

$$N^p(\bar{x};\Omega) := \big\{x^* \in X^* : \text{there exist } \delta > 0 \text{ and } \eta > 0 \text{ such that}$$
$$\langle x^*, x - \bar{x}\rangle \leq \eta\|x - \bar{x}\|^2 \text{ for all } x \in \boldsymbol{B}_X(\bar{x},\delta) \cap \Omega\big\}.$$

The elements of this set are known as ***proximal normals to*** Ω ***at*** $\bar{x}$.

Remarks 1.6.2. (i) If $\bar{x} \in \operatorname{cl}\Omega$, it is a conspicuous consequence of the above definitions that $N^c(\bar{x};\Omega) = N^c(\bar{x};\operatorname{cl}\Omega)$ and $N^p(\bar{x};\Omega) = N^p(\bar{x};\operatorname{cl}\Omega)$.

(ii) Observe that $N^c(\bar{x};\Omega)$ is *closed* and *convex*. In particular, if X is *reflexive*, then $N^c(\bar{x};\Omega)$ is *weakly*-closed*. On the other hand, $N^p(\bar{x};\Omega)$ is *convex* but *not necessarily closed*.

(iii) Note the *monotone* property of $N^c(\bar{x};\Omega)$ and $N^p(\bar{x};\Omega)$ with respect to set inclusion: if $\bar{x} \in \Omega_1 \subset \Omega_2$, then $N^c(\bar{x};\Omega_2) \subset N^c(\bar{x};\Omega_1)$ and $N^p(\bar{x};\Omega_2) \subset N^p(\bar{x};\Omega_1)$.

Definition 1.6.3. Let $\Omega \subset X$ with $\bar{x} \in \Omega$ and $\varepsilon \geq 0$. The ***set of*** (***Fréchet-like***) ***ε-normals to*** Ω ***at*** $\bar{x}$ is defined by

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

The elements of this set are known as (***Fréchet-like***) ***ε-normals to*** Ω ***at*** $\bar{x}$. In particular, $\widehat{N}(\bar{x};\Omega) := \widehat{N}_0(\bar{x};\Omega)$ is called the ***Fréchet normal cone to*** Ω ***at*** $\bar{x}$ and its elements are known as ***Fréchet normals to*** Ω ***at*** $\bar{x}$.

Remarks 1.6.4. (i) If $\bar{x} \in \operatorname{cl}\Omega$, it is a conspicuous consequence of the above definition that $\widehat{N}_\varepsilon(\bar{x};\Omega) = \widehat{N}_\varepsilon(\bar{x};\operatorname{cl}\Omega)$.

(ii) Observe that $\widehat{N}_\varepsilon(\bar{x};\Omega)$ is *closed* and *convex*. In particular, if X is *reflexive*, then $\widehat{N}_\varepsilon(\bar{x};\Omega)$ is *weakly*-closed.*

(iii) While $\widehat{N}(\bar{x};\Omega)$ is a cone, $\widehat{N}_\varepsilon(\bar{x};\Omega)$ is *not* a cone for any $\varepsilon > 0$.

(iv) Note the *monotone* properties of $\widehat{N}_\varepsilon(\bar{x};\Omega)$ with respect to ε and with respect to set inclusion:

- If $0 \le \varepsilon_1 \le \varepsilon_2$, then $\widehat{N}_{\varepsilon_1}(\bar{x};\Omega) \subset \widehat{N}_{\varepsilon_2}(\bar{x};\Omega)$.
- If $\bar{x} \in \Omega_1 \subset \Omega_2$, then $\widehat{N}_\varepsilon(\bar{x};\Omega_2) \subset \widehat{N}_\varepsilon(\bar{x};\Omega_1)$.

(v) The counterpart of the subdifferential inclusion relation in Proposition 1.5.8 holds for the dual normal objects: $N^c(\bar{x};\Omega) \subset N^p(\bar{x};\Omega) \subset \widehat{N}(\bar{x};\Omega)$.

One also has a principal description of $\widehat{N}_\varepsilon(\bar{x};\Omega)$ analogous to that of ε-subdifferentials as a direct consequence of the preceding definition.

Proposition 1.6.5. *Let $\Omega \subset X$ with $\bar{x} \in \Omega$ and $\varepsilon \ge 0$. Then $x^* \in \widehat{N}_\varepsilon(\bar{x};\Omega)$ if and only if for any $\gamma > 0$, there exists $\delta > 0$ such that for all $x \in \Omega$ with $\|x - \bar{x}\| \le \delta$,*

$$\langle x^*, x - \bar{x}\rangle \le (\varepsilon + \gamma)\|x - \bar{x}\|.$$

In other words, $x^* \in \widehat{N}_\varepsilon(\bar{x};\Omega)$ if and only if for any $\gamma > 0$, the function $\zeta : \Omega \to \overline{\mathbb{R}}$ defined by $\zeta(x) = -\langle x^*, x - \bar{x}\rangle + (\varepsilon + \gamma)\|x - \bar{x}\|$ attains a local minimum at $\bar{x}$.

Proof. Consider the first assertion of the proposition. Assume $x^* \in \widehat{N}_\varepsilon(\bar{x};\Omega)$. By definition,

$$\ell := \inf_{\delta>0} \sup_{\substack{x\in\Omega \\ 0<\|x-\bar{x}\|\le\delta}} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} = \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \le \varepsilon.$$

Let $\gamma > 0$. Suppose ℓ is finite. There exists $\delta_1 > 0$ such that for all $x \in \Omega$ with $0 < \|x - \bar{x}\| \leq \delta_1$,

$$\frac{\langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq \ell + \gamma \leq \varepsilon + \gamma,$$

which implies

$$\langle x^*, x - \bar{x} \rangle \leq (\varepsilon + \gamma)\|x - \bar{x}\|. \tag{1.8}$$

Otherwise $\ell = -\infty$. There exists $\delta_2 > 0$ such that for all $x \in \Omega$ with $0 < \|x - \bar{x}\| \leq \delta_2$,

$$\frac{\langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0,$$

which gives

$$\langle x^*, x - \bar{x} \rangle \leq 0 \leq (\varepsilon + \gamma)\|x - \bar{x}\|. \tag{1.9}$$

In both cases, in view of (1.8) and (1.9), there exists $\delta > 0$ such that the same inequality holds for all $x \in \Omega$ with $0 < \|x - \bar{x}\| \leq \delta$. Moreover, equality trivially holds for $x = \bar{x}$. This proves one implication.

For the opposite implication, let $\gamma > 0$. By assumption, there exists $\delta > 0$ such that for all $x \in \Omega$ with $\|x - \bar{x}\| \leq \delta$,

$$\langle x^*, x - \bar{x} \rangle \leq (\varepsilon + \gamma)\|x - \bar{x}\|. \tag{1.10}$$

In particular, for all $x \in \Omega$ with $0 < \|x - \bar{x}\| \leq \delta$, rearranging (1.10) shows

$$\frac{\langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq \varepsilon + \gamma.$$

Passing to the limit, one sees that $\limsup\limits_{x \xrightarrow{\Omega} \bar{x}} \dfrac{\langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq \varepsilon + \gamma$, which, since $\gamma > 0$ is arbitrary, reduces to

$$\limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq \varepsilon.$$

By definition, $x^* \in \widehat{N}_\varepsilon(\bar{x}; \Omega)$ and the other implication holds.

The second assertion of the proposition follows from the first by noting that $\zeta(\bar{x}) = 0$.

□

Proposition 1.6.6. *Let $\Omega \subset X$ with $\bar{x} \in \Omega$ and $\varepsilon \geq 0$. Then for any $\alpha > 0$, $x^* \in \widehat{N}_\varepsilon(\bar{x}; \Omega)$ if and only if $\alpha x^* \in \widehat{N}_{\alpha\varepsilon}(\bar{x}; \Omega)$.*

Proof. Let $\gamma > 0$ and $\alpha > 0$. Suppose $x^* \in \widehat{N}_\varepsilon(\bar{x};\Omega)$. Owing to Proposition 1.6.5, there exists $\delta > 0$ such that for all $x \in \Omega$ with $\|x - \bar{x}\| \le \delta$,

$$\langle x^*, x - \bar{x}\rangle \le \left(\varepsilon + \frac{\gamma}{\alpha}\right)\|x - \bar{x}\|,$$

and hence
$$\langle \alpha x^*, x - \bar{x}\rangle \le \alpha\left(\varepsilon + \frac{\gamma}{\alpha}\right)\|x - \bar{x}\| = (\alpha\varepsilon + \gamma)\|x - \bar{x}\|.$$

Using Proposition 1.6.5 again, $\alpha x^* \in \widehat{N}_{\alpha\varepsilon}(\bar{x};\Omega)$.

Conversely, suppose $\alpha x^* \in \widehat{N}_{\alpha\varepsilon}(\bar{x};\Omega)$. The above implies that

$$x^* = \frac{1}{\alpha}(\alpha x^*) \in \widehat{N}_{\frac{1}{\alpha}(\alpha\varepsilon)}(\bar{x};\Omega) = \widehat{N}_\varepsilon(\bar{x};\Omega).$$

The result is verified. □

Definition 1.6.7. Let $\Omega \subset X$ and $\bar{x} \in \Omega$. The ***limiting normal cone*** or ***basic normal cone to Ω at $\bar{x}$*** is defined by

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

The elements of this set are known as ***limiting normals*** or ***basic normals to Ω at $\bar{x}$***.

Remarks 1.6.8. (i) Clearly, $\widehat{N}(\bar{x};\Omega) \subset N(\bar{x};\Omega)$.

(ii) If $\bar{x} \in \operatorname{cl}\Omega$, it is a conspicuous consequence of the above definition that $N(\bar{x};\Omega) \subset N(\bar{x};\operatorname{cl}\Omega)$.

(iii) If X is an *Asplund space* and Ω is *locally closed at* $\bar{x}$, then the *limiting normal cone* to Ω at $\bar{x}$ admits the simpler representation

$$N(\bar{x};\Omega) = \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}(x;\Omega).$$

On the other hand, if X is a *Hilbert space*, then the *limiting normal cone* to Ω at $\bar{x}$ admits the simpler representation

$$N(\bar{x};\Omega) = \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} N^p(x;\Omega).$$

One of the most recognized relationships between these normal objects and their dual subdifferentials is provided by *indicator functions.*

Definition 1.6.9. Let $\Omega \subset X$. The ***indicator function*** $\delta_\Omega : X \to \overline{\mathbb{R}}$ ***of*** Ω is defined by

$$\delta_\Omega(x) := \begin{cases} 0 & \text{if } x \in \Omega, \\ \infty & \text{if } x \notin \Omega. \end{cases}$$

Proposition 1.6.10. *Let* $\Omega \subset X$ *and* $\bar{x} \in \Omega$. *Then* $N^\bullet(\bar{x};\Omega) = \partial^\bullet \delta_\Omega(\bar{x})$, *where* $(N^\bullet, \partial^\bullet)$ *stands for* (N^c, ∂^c), (N^p, ∂^p) *or* $(\widehat{N}, \widehat{\partial})$.

Proof. **(a)** Let $x^* \in N^c(\bar{x};\Omega)$. By definition, $\langle x^*, x - \bar{x}\rangle \leq 0$ for all $x \in \Omega$. This implies for all $x \in \Omega$, in view of $\delta_\Omega(x) = \delta_\Omega(\bar{x}) = 0$, that $\delta_\Omega(x) \geq \delta_\Omega(\bar{x}) + \langle x^*, x - \bar{x}\rangle$. On the other hand, for all $x \notin \Omega$, $\delta_\Omega(x) = \infty$ and the inequality $\delta_\Omega(x) \geq \delta_\Omega(\bar{x}) + \langle x^*, x - \bar{x}\rangle$ trivially holds. With the inequality valid for all $x \in X$, one sees that $x^* \in \partial^c \delta_\Omega(\bar{x})$ and $N^c(\bar{x};\Omega) \subset \partial^c \delta_\Omega(\bar{x})$.

Consider the reverse inclusion. Let $x^* \in \partial^c \delta_\Omega(\bar{x})$. For all $x \in X$, the inequality $\delta_\Omega(x) \geq \delta_\Omega(\bar{x}) + \langle x^*, x - \bar{x}\rangle$ holds. In particular, for all $x \in \Omega$, with $\delta_\Omega(x) = \delta_\Omega(\bar{x}) = 0$, the inequality simplifies to $\langle x^*, x - \bar{x}\rangle \leq 0$. Hence $x^* \in N^c(\bar{x};\Omega)$ and $N^c(\bar{x};\Omega) \supset \partial^c \delta_\Omega(\bar{x})$. This proves the first equality.

(b) Let $x^* \in N^p(\bar{x};\Omega)$. By definition, there exist $\delta > 0$ and $\eta > 0$ such that for all $x \in \boldsymbol{B}_X(\bar{x},\delta) \cap \Omega$, $\langle x^*, x - \bar{x}\rangle \leq \eta\|x - \bar{x}\|^2 = \delta_\Omega(x) - \delta_\Omega(\bar{x}) + \eta\|x - \bar{x}\|^2$, since $\delta_\Omega(x) = \delta_\Omega(\bar{x}) = 0$. On the other hand, for all $x \in \boldsymbol{B}_X(\bar{x},\delta)\backslash\Omega$, $\delta_\Omega(x) = \infty$ and the inequality $\langle x^*, x - \bar{x}\rangle \leq \delta_\Omega(x) - \delta_\Omega(\bar{x}) + \eta\|x - \bar{x}\|^2$ trivially holds. With the inequality valid for all $x \in \boldsymbol{B}_X(\bar{x},\delta)$, one has $x^* \in \partial^p \delta_\Omega(\bar{x})$ and $N^p(\bar{x};\Omega) \subset \partial^p \delta_\Omega(\bar{x})$.

Consider the reverse inclusion. Let $x^* \in \partial^p \delta_\Omega(\bar{x})$. Then for all $x \in \boldsymbol{B}_X(\bar{x},\delta)$, there holds $\langle x^*, x - \bar{x}\rangle \leq \delta_\Omega(x) - \delta_\Omega(\bar{x}) + \eta\|x - \bar{x}\|^2$. In particular, for all $x \in \boldsymbol{B}_X(\bar{x},\delta) \cap \Omega$, since $\delta_\Omega(x) = \delta_\Omega(\bar{x}) = 0$, the inequality reduces to $\langle x^*, x - \bar{x}\rangle \leq$

$\eta\|x-\bar{x}\|^2$. Thus $x^* \in N^p(\bar{x};\Omega)$ and $N^p(\bar{x};\Omega) \supset \partial^p\delta_\Omega(\bar{x})$. This establishes the second equality.

(c) Let $x^* \in \widehat{N}(\bar{x};\Omega)$ and $\gamma > 0$. Using Proposition 1.6.5, there exists $\lambda > 0$ such that for all $x \in \Omega$ with $\|x-\bar{x}\| \le \lambda$,

$$\langle x^*, x-\bar{x}\rangle \le \gamma\|x-\bar{x}\|. \tag{1.11}$$

Fix any $x \in X$ with $\|x-\bar{x}\| \le \lambda$. Suppose $x \in \Omega$. Then $\delta_\Omega(x) = \delta_\Omega(\bar{x}) = 0$. It follows from (1.11) that

$$\langle x^*, x-\bar{x}\rangle \le \delta_\Omega(x) - \delta_\Omega(\bar{x}) + \gamma\|x-\bar{x}\|. \tag{1.12}$$

Otherwise $x \notin \Omega$ and hence $\delta_\Omega(x) = \infty$. Thus inequality (1.12) trivially holds. In both cases, Proposition 1.5.5 implies that $x^* \in \widehat{\partial}\delta_\Omega(\bar{x})$ and $\widehat{N}(\bar{x};\Omega) \subset \widehat{\partial}\delta_\Omega(\bar{x})$.

Consider the reverse inclusion. Let $x^* \in \widehat{\partial}\delta_\Omega(\bar{x})$ and $\gamma > 0$. Employing Proposition 1.5.5 again, there exists $\eta > 0$ such that for all $x \in X$ with $\|x-\bar{x}\| \le \eta$,

$$\langle x^*, x-\bar{x}\rangle \le \delta_\Omega(x) - \delta_\Omega(\bar{x}) + \gamma\|x-\bar{x}\|. \tag{1.13}$$

Fix any $x \in \Omega$ with $\|x-\bar{x}\| \le \eta$. Note that $\delta_\Omega(x) = \delta_\Omega(\bar{x}) = 0$, reducing (1.13) to

$$\langle x^*, x-\bar{x}\rangle \le \gamma\|x-\bar{x}\|.$$

By virtue of Proposition 1.6.5, $x^* \in \widehat{N}(\bar{x};\Omega)$ and $\widehat{N}(\bar{x};\Omega) \supset \widehat{\partial}\delta_\Omega(\bar{x})$. This justifies the third equality. □

Indeed, the same duality relation also holds for sets of ε-normals and limiting normal cones. The reader may consult Section 1.3.1 and Section 1.3.2 of [25] for details.

Proposition 1.6.11. *Let $\Omega \subset X$ and $\bar{x} \in \Omega$. Then $\widehat{N}_\varepsilon(\bar{x};\Omega) = \widehat{\partial}_\varepsilon\delta_\Omega(\bar{x})$ for any $\varepsilon \ge 0$, and $N(\bar{x};\Omega) = \partial\delta_\Omega(\bar{x})$.*

1.7 Coderivatives

While a number of derivative-like constructions have been introduced for single-valued functions, this section describes several less acquainted derivative-like constructions for set-valued mappings. Collectively known as *coderivatives*, these constructions are natural extensions of the classical *adjoint derivative operators* of smooth single-valued functions and allow pointwise approximation of set-valued mappings using elements of dual spaces.

Definition 1.7.1. Let $F : X \rightrightarrows Y$ be a set-valued mapping with $\operatorname{dom} F \neq \varnothing$ and $(\bar{x}, \bar{y}) \in \operatorname{gph} F$.

(a) Let $\varepsilon \geq 0$. The ***ε-coderivative of F at*** $(\bar{x}, \bar{y})$ is the multifunction $\widehat{D}^*_\varepsilon F(\bar{x}, \bar{y}) : Y^* \rightrightarrows X^*$ defined by

$$\widehat{D}^*_\varepsilon F(\bar{x}, \bar{y})(y^*) := \{x^* \in X^* : (x^*, -y^*) \in \widehat{N}_\varepsilon((\bar{x}, \bar{y}); \operatorname{gph} F)\}.$$

In particular, $\widehat{D}^* F(\bar{x}, \bar{y}) := \widehat{D}^*_0 F(\bar{x}, \bar{y})$ is called the ***Fréchet coderivative of F at*** $(\bar{x}, \bar{y})$.

(b) The ***normal coderivative of F at*** $(\bar{x}, \bar{y})$ is the multifunction $D^*_N F(\bar{x}, \bar{y}) : Y^* \rightrightarrows X^*$ defined by

$$D^*_N F(\bar{x}, \bar{y})(y^*) := \operatorname*{Lim\,sup}_{\substack{(x,y)\to(\bar{x},\bar{y}) \\ \bar{y}^* \xrightarrow{w^*} y^* \\ \varepsilon \downarrow 0}} \widehat{D}^*_\varepsilon F(x, y)(\bar{y}^*).$$

That is, $x^* \in D^*_N F(\bar{x}, \bar{y})(y^*)$ if and only if there exist sequences $\{\varepsilon_k\}_{k=1}^\infty \subset \mathbb{R}_+$, $\{(x_k, y_k)\}_{k=1}^\infty \subset X \times Y$ and $\{(x^*_k, y^*_k)\}_{k=1}^\infty \subset X^* \times Y^*$ such that $\varepsilon_k \downarrow 0$, $(x_k, y_k) \xrightarrow{\operatorname{gph} F} (\bar{x}, \bar{y})$, $(x^*_k, y^*_k) \xrightarrow{w^*} (x^*, y^*)$ and $(x^*_k, -y^*_k) \in \widehat{N}_{\varepsilon_k}((x_k, y_k); \operatorname{gph} F)$ for all $k \in \mathbb{N}$.

(c) The ***mixed coderivative of F at*** $(\bar{x}, \bar{y})$ is the multifunction $D^*_M F(\bar{x}, \bar{y}) : Y^* \rightrightarrows X^*$ defined by

$$D^*_M F(\bar{x}, \bar{y})(y^*) := \operatorname*{Lim\,sup}_{\substack{(x,y)\to(\bar{x},\bar{y}) \\ \bar{y}^* \xrightarrow{\|\cdot\|} y^* \\ \varepsilon \downarrow 0}} \widehat{D}^*_\varepsilon F(x, y)(\bar{y}^*).$$

That is, $x^* \in D^*_M F(\bar{x},\bar{y})(y^*)$ if and only if there exist sequences $\{\varepsilon_k\}_{k=1}^\infty \subset \mathbb{R}_+$, $\{(x_k,y_k)\}_{k=1}^\infty \subset X \times Y$ and $\{(x_k^*,y_k^*)\}_{k=1}^\infty \subset X^* \times Y^*$ such that $\varepsilon_k \downarrow 0$, $(x_k,y_k) \xrightarrow{\operatorname{gph} F} (\bar{x},\bar{y})$, $x_k^* \xrightarrow{w^*} x^*$, $y_k^* \xrightarrow{\|\cdot\|} y^*$ and $(x_k^*,-y_k^*) \in \widehat{N}_{\varepsilon_k}((x_k,y_k);\operatorname{gph} F)$ for all $k \in \mathbb{N}$.

Remarks 1.7.2. (i) Clearly, $\widehat{D}^* F(\bar{x},\bar{y})(y^*) \subset D^*_M F(\bar{x},\bar{y})(y^*) \subset D^*_N F(\bar{x},\bar{y})(y^*)$ for any $y^* \in Y^*$.

(ii) Observe that the *normal coderivative* may be alternatively characterized by the serviceable description

$$D^*_N F(\bar{x},\bar{y})(y^*) = \{x^* \in X^* : (x^*,-y^*) \in N((\bar{x},\bar{y});\operatorname{gph} F)\}.$$

This shows that the normal coderivative is uniquely determined by the limiting normal cone to $\operatorname{gph} F$ and explains the name of the coderivative.

(iii) The primary difference between the definition of the normal coderivative and that of the mixed coderivative is that *weak* convergence* is used for both X^* and Y^* in the definition of the *normal coderivative*, while *weak* convergence* is used for X^* and *norm convergence* is used for Y^* in the definition of the *mixed coderivative.* This justifies the choice of the terminology *mixed coderivative.*

(iv) If X and Y are *Asplund spaces* and $\operatorname{gph} F$ is *locally closed at* $(\bar{x},\bar{y})$, then the *mixed coderivative* of F at $(\bar{x},\bar{y})$ admits the simpler representation

$$D^*_M F(\bar{x},\bar{y})(y^*) = \operatorname*{Lim\,sup}_{\substack{(x,y)\to(\bar{x},\bar{y}) \\ \bar{y}^* \xrightarrow{\|\cdot\|} y^*}} \widehat{D}^* F(x,y)(\bar{y}^*).$$

That is, $x^* \in D^*_M F(\bar{x},\bar{y})(y^*)$ if and only if there exist sequences $\{(x_k,y_k)\}_{k=1}^\infty \subset X \times Y$ and $\{(x_k^*,y_k^*)\}_{k=1}^\infty \subset X^* \times Y^*$ such that $(x_k,y_k) \xrightarrow{\operatorname{gph} F} (\bar{x},\bar{y})$, $x_k^* \xrightarrow{w^*} x^*$, $y_k^* \xrightarrow{\|\cdot\|} y^*$ and $(x_k^*,-y_k^*) \in \widehat{N}((x_k,y_k);\operatorname{gph} F)$ for all $k \in \mathbb{N}$.

A number of further properties of coderivatives are needed in later chapters and are included here without proof. Details are available in [25, Theorem 1.41] and [25, Theorem 1.43 & Theorem 1.44] respectively.

Definition 1.7.3. Let $F : X \rightrightarrows Y$ be a set-valued mapping with $\operatorname{dom} F \neq \varnothing$.

(a) Let $U \subset X$ and $V \subset Y$ be nonempty subsets. F is said to be ***Lipschitz-like on U relative to V*** with rank $\ell \geq 0$ if for all $x, u \in U$,

$$F(x) \cap V \subset F(u) + \ell\|x - u\|\boldsymbol{B}_Y.$$

(b) Let $(\bar{x}, \bar{y}) \in \operatorname{gph} F$. F is said to be ***locally Lipschitz-like*** or ***pseudo-Lipschitz*** or ***Aubin at*** $(\bar{x}, \bar{y})$ ***with rank*** $\ell \geq 0$ if there exist neighbourhoods U of $\bar{x}$ and V of $\bar{y}$ such that F is Lipschitz-like on U relative to V with rank ℓ.

As pointed out in [25], the local Lipschitz-like property can be regarded as a localization of Lipschitz behaviour not only relative to a point of the domain but also relative to a particular point of the *image set* $\bar{y} \in F(\bar{x})$, and admits an efficient characterization in terms of the local Lipschitz continuity of the *generalized distance function*, which is the focus of the next chapter.

Theorem 1.7.4. *Let $F : X \rightrightarrows Y$ be a set-valued mapping and $(\bar{x}, \bar{y}) \in \operatorname{gph} F$. Then F is locally Lipschitz-like at $(\bar{x}, \bar{y})$ if and only if $\rho : \operatorname{dom} F \times Y \to \mathbb{R}$ defined by*

$$\rho(x, y) := d(y, F(x)) = \inf\{\|y - w\| : w \in F(x)\}$$

is locally Lipschitz at $(\bar{x}, \bar{y})$.

Theorem 1.7.5. *Let $F : X \rightrightarrows Y$ be a set-valued mapping and $(\bar{x}, \bar{y}) \in \operatorname{gph} F$. Suppose F is locally Lipschitz-like at $(\bar{x}, \bar{y})$ with rank $\ell \geq 0$. The following statements hold:*

(a) *Let $\varepsilon \geq 0$. Then there exists $\eta > 0$ such that for all $x \in \boldsymbol{B}_X(\bar{x}, \eta)$, $y \in F(x) \cap \boldsymbol{B}_Y(\bar{y}, \eta)$, and $y^* \in Y^*$,*

$$\sup\left\{\|x^*\| : x^* \in \widehat{D}^*_\varepsilon F(x, y)(y^*)\right\} \leq \ell\|y^*\| + \varepsilon(1 + \ell).$$

(b) $D^*_M F(\bar{x}, \bar{y})(0) = \{0\}$.

Chapter 2

The Generalized Distance Function - Basic Estimates

While the standard distance function measures the distance from a moving point to a fixed destination set, it is natural to consider the distance from a moving point to a *moving destination set* as a generalization. This gives rise to a function of two variables, the *generalized distance function*, which is the subject of this chapter. Most of the results covered in this chapter first appeared in [26].

2.1 Elementary Properties of the Generalized Distance Function

Definition 2.1.1. Let $F : Z \rightrightarrows X$ be a set-valued mapping with $\operatorname{dom} F \neq \varnothing$. The ***generalized distance function*** $\rho : \operatorname{dom} F \times X \to \mathbb{R}$ ***associated with*** F is defined by

$$\rho(z, x) := d(x, F(z)) = \inf\{\|x - w\| : w \in F(z)\}.$$

The generalized distance function allows the destination set to vary by employing a set-valued mapping F. As a generalization of the standard distance function, it may be reduced easily to the latter, which concerns a fixed nonempty destination set $\Omega \subset X$, by taking $F \equiv \Omega$.

An elementary property of the generalized distance function is used repeatedly in the subsequent exposition.

Proposition 2.1.2. *Let $F : Z \rightrightarrows X$ be a set-valued mapping and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$ with $\bar{z} \in \operatorname{dom} F$. Suppose $\operatorname{gph} F$ is locally closed at $(\bar{z}, \bar{x})$. Then $\rho(\bar{z}, \bar{x}) > 0$.*

Proof. Suppose $\rho(\bar{z}, \bar{x}) = 0$. For each $k \in \mathbb{N}$, there exists $w_k \in F(\bar{z})$ such that

$$\frac{1}{k} = \frac{1}{k} + \rho(\bar{z}, \bar{x}) = \frac{1}{k} + d(\bar{x}, F(\bar{z})) > \|\bar{x} - w_k\|. \tag{2.1}$$

Since $\operatorname{gph} F$ is locally closed at $(\bar{z}, \bar{x})$, there exists $\delta > 0$ such that $\boldsymbol{B}_{Z\times X}((\bar{z}, \bar{x}), \delta) \cap \operatorname{gph} F$ is closed. By considering the tail of $\{w_k\}_{k=1}^{\infty}$ if necessary, assume that $\|w_k - \bar{x}\| \leq \delta$ for all $k \in \mathbb{N}$. Then $\{(\bar{z}, w_k)\}_{k=1}^{\infty}$ is a sequence in the closed set $\boldsymbol{B}_{Z\times X}((\bar{z}, \bar{x}), \delta) \cap \operatorname{gph} F$. On the other hand, letting $k \to \infty$ in inequality (2.1) yields $w_k \to \bar{x}$. It follows from $(\bar{z}, w_k) \to (\bar{z}, \bar{x})$ that $(\bar{z}, \bar{x}) \in \boldsymbol{B}_{Z\times X}((\bar{z}, \bar{x}), \delta) \cap \operatorname{gph} F$. In particular, $(\bar{z}, \bar{x}) \in \operatorname{gph} F$, which contradicts the initial assumption. Thus $\rho(\bar{z}, \bar{x}) > 0$. □

Building upon the generalized distance function, some more definitions are made.

Definition 2.1.3. Let $F : Z \rightrightarrows X$ be a set-valued mapping with $\operatorname{dom} F \neq \varnothing$, $(\bar{z}, \bar{x}) \in Z \times X$ and $r \geq 0$.

(a) The r***-enlargement of*** F is the set-valued mapping $F_r : \operatorname{dom} F \rightrightarrows X$ defined by

$$F_r(z) := \{x \in X : d(x, F(z)) \leq r\}.$$

(b) The r***-generalized distance function*** $\rho_r : \operatorname{dom} F \times X \to \mathbb{R}$ ***associated with*** F is defined by

$$\rho_r(z, x) := d(x, F_r(z)) = \inf\{\|x - w\| : w \in F_r(z)\}.$$

(c) The r***-perturbed projection set of*** $(\bar{z}, \bar{x})$ ***onto*** $\operatorname{gph} F$ is defined by

$$\Theta_r^F(\bar{z}, \bar{x}) := \{(v, u) \in \operatorname{gph} F : \|v - \bar{z}\| \leq r \quad \text{and} \quad \|u - \bar{x}\| \leq \rho(\bar{z}, \bar{x}) + r\}.$$

Remark 2.1.4. A moment's reflection on the definition of the r-enlargement of F reveals that $\operatorname{dom} F = \operatorname{dom} F_r$.

In order to avoid trivial statements, ***all set-valued mappings*** $F : Z \rightrightarrows X$ ***considered in the rest of this chapter are conveniently assumed to satisfy*** $\operatorname{dom} F = Z \neq \varnothing$. In the same light, ***all subsets*** $\Omega \subset X$ ***in this chapter are presumed to be nonempty***.

The r-generalized distance function ρ_r has some obvious relationships with the generalized distance function ρ. Two of these are included here.

Proposition 2.1.5. (cf. [9, Lemma 3.1]) *Let* $F : Z \rightrightarrows X$ *be a set-valued mapping,* $r \geq 0$ *and* $(\bar{z}, \bar{x}) \notin \operatorname{gph} F_r$. *Then* $\rho_r(\bar{z}, \bar{x}) = \rho(\bar{z}, \bar{x}) - r$.

Proof. Let $v \in F_r(\bar{z})$ and $\varepsilon > 0$. By definition, $d(v, F(\bar{z})) \leq r$. There exists $w_\varepsilon \in F(\bar{z})$ such that $\|v - w_\varepsilon\| < d(r, F(\bar{z})) + \varepsilon \leq r + \varepsilon$. It follows that

$$\begin{aligned}\|v - \bar{x}\| &\geq \|\bar{x} - w_\varepsilon\| - \|w_\varepsilon - v\| \\ &> d(\bar{x}, F(\bar{z})) - (r + \varepsilon) \\ &= \rho(\bar{z}, \bar{x}) - r - \varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ and $v \in F_r(\bar{z})$ are arbitrary, $\|v - \bar{x}\| \geq \rho(\bar{z}, \bar{x}) - r$ for all $v \in F_r(\bar{z})$. Consequently, $\rho_r(\bar{z}, \bar{x}) = d(\bar{x}, F_r(\bar{z})) \geq \rho(\bar{z}, \bar{x}) - r$.

Conversely, let $y \in F(\bar{z})$. Since $(\bar{z}, \bar{x}) \notin \operatorname{gph} F_r$, $\bar{x} \notin F_r(\bar{z})$ and $d(\bar{x}, F(\bar{z})) > r$. Define $h : [0, \infty) \to [0, \infty)$ by

$$h(s) = d(s\bar{x} + (1 - s)y, F(\bar{z})).$$

It follows that h is continuous, $h(0) = 0$ and $h(1) > r$. By the intermediate value theorem, there exists $s_0 \in [0, 1)$ such that $h(s_0) = r$. Take $w = s_0\bar{x} + (1 - s_0)y$ so that w, $\bar{x}$ and y are collinear. Then $h(s_0) = d(w, F(\bar{z})) = r$ and $w \in F_r(\bar{z})$. Note that

$$\|\bar{x} - y\| = \|\bar{x} - w\| + \|w - y\| \geq d(\bar{x}, F_r(\bar{z})) + d(w, F_r(\bar{z})) = \rho_r(\bar{z}, \bar{x}) + r.$$

Since $y \in F(\bar{z})$ is arbitrary, $\|\bar{x} - y\| \geq \rho_r(\bar{z}, \bar{x}) + r$ holds for all $y \in F(\bar{z})$, which implies that $\rho(\bar{z}, \bar{x}) = d(\bar{x}, F(\bar{z})) \geq \rho_r(\bar{z}, \bar{x}) + r$. Equivalently, $\rho_r(\bar{z}, \bar{x}) \leq \rho(\bar{z}, \bar{x}) - r$. The assertion is ascertained. □

Proposition 2.1.6. *Let $F : Z \rightrightarrows X$ be a set-valued mapping, $r \geq 0$ and $(\bar{z}, \bar{x}) \in Z \times X$. If ρ is locally Lipschitz at $(\bar{z}, \bar{x})$ with rank $\ell \geq 0$, then ρ_r is also locally Lipschitz at $(\bar{z}, \bar{x})$ with rank ℓ.*

Proof. Since ρ is locally Lipschitz at $(\bar{z}, \bar{x})$ with rank ℓ, there exists $\delta > 0$ such that for all $(z_i, x_i) \in Z \times X$ with $\|x_i - \bar{x}\| \leq \delta$ and $\|z_i - \bar{z}\| \leq \delta$, where $i = 1, 2$,

$$|\rho(z_1, x_1) - \rho(z_2, x_2)| \leq \ell(\|x_1 - x_2\| + \|z_1 - z_2\|). \tag{2.2}$$

There are three different cases:

Case 1: $x_1 \in F_r(z_1)$ and $x_2 \in F_r(z_2)$.

Note that $\rho_r(z_1, x_1) = \rho_r(z_2, x_2) = 0$. It follows that

$$|\rho_r(z_1, x_1) - \rho_r(z_2, x_2)| = 0 \leq \ell(\|x_1 - x_2\| + \|z_1 - z_2\|).$$

Case 2: $x_1 \notin F_r(z_1)$ and $x_2 \notin F_r(z_2)$.

For $k = 1, 2$, noting that $(z_k, x_k) \notin \operatorname{gph} F_r$ and applying Proposition 2.1.5, one sees that $\rho_r(z_k, x_k) = \rho(z_k, x_k) - r$. Employing (2.2),

$$\begin{aligned} |\rho_r(z_1, x_1) - \rho_r(z_2, x_2)| &= |(\rho(z_1, x_1) - r) - (\rho(z_2, x_2) - r)| \\ &= |\rho(z_1, x_1) - \rho(z_2, x_2)| \\ &\leq \ell(\|x_1 - x_2\| + \|z_1 - z_2\|). \end{aligned}$$

Case 3: $x_i \notin F_r(z_i)$ and $x_j \in F_r(z_j)$, where $i \neq j$ and $1 \leq i, j \leq 2$.

Observe that $\rho_r(z_j, x_j) = 0$ and $\rho(z_j, x_j) = d(x_j, F(z_j)) \leq r$. Moreover, $(z_i, x_i) \notin \operatorname{gph} F_r$. By Proposition 2.1.5 again, $\rho_r(z_i, x_i) = \rho(z_i, x_i) - r$. Using (2.2),

$$\begin{aligned} |\rho_r(z_1, x_1) - \rho_r(z_2, x_2)| &= \rho_r(z_1, x_1) = \rho(z_1, x_1) - r \\ &\leq \rho(z_1, x_1) - \rho(z_2, x_2) \leq \ell(\|x_1 - x_2\| + \|z_1 - z_2\|). \end{aligned}$$

In all three cases, ρ_r is locally Lipschitz at $(\bar{z}, \bar{x})$ with rank ℓ. □

As suggested heuristically by celebrated theorems in convex analysis, normal cones and sets of ε-normals to enlargements are integral ingredients of estimates and characterizations of their dual subdifferentials of the generalized distance function. That these normal cones and sets of ε-normals to enlargements are all well-defined is a rather obvious fact.

Proposition 2.1.7. *Let $F : Z \rightrightarrows X$ be a set-valued mapping, $(\bar{z}, \bar{x}) \in Z \times X$ and $r = \rho(\bar{z}, \bar{x})$. Then $(\bar{z}, \bar{x}) \in \operatorname{gph} F_r$.*

Proof. Note that $d(\bar{x}, F(\bar{z})) = \rho(\bar{z}, \bar{x}) = r$ implies $\bar{x} \in F_r(\bar{z})$ and hence $(\bar{z}, \bar{x}) \in \operatorname{gph} F_r$. □

Remark 2.1.8. This proposition ensures that the normal objects $N^c((\bar{z}, \bar{x}); \operatorname{gph} F_r)$, $N^p((\bar{z}, \bar{x}); \operatorname{gph} F_r)$, $\widehat{N}_\varepsilon((\bar{z}, \bar{x}); \operatorname{gph} F_r)$, $N((\bar{z}, \bar{x}); \operatorname{gph} F_r)$ and the coderivatives $\widehat{D}^*_\varepsilon F_r(\bar{z}, \bar{x})$, $D^*_N F_r(\bar{z}, \bar{x})$, $D^*_M F_r(\bar{z}, \bar{x})$ are all well-defined.

The rest of the chapter presents a collection of estimates of Fréchet-like and limiting subdifferentials of the generalized distance function. These estimates are not only fundamental in the theory of the generalized distance function but are also readily reducible as special cases to the analogous results pertaining to the standard distance function, which are often of independent interest and are hence entailed in this thesis separately as corollaries. Most of these corollaries follow immediately from their preceding results concerning the generalized distance function by taking $Z = \{\bar{z}\}$ and $F \equiv \Omega$. Only corollaries which involve other technicalities are proved individually.

2.2 Fréchet-Like Subdifferentials of the Generalized Distance Function

In the influential paper *On the Clarke subdifferential of the distance function of a closed set* [12], Burke, Ferris and Qian exhibited a collection of elegant estimates of Clarke subdifferentials of the standard distance function by means of thickenings and projections defined in Section 1.4. This has motivated endeavours to produce fundamental estimates of Fréchet-like subdifferentials of the generalized distance function via the comparable tools of enlargements and projections. As emphasized in [12], results regarding the standard distance function at points situated in the underlying set are notably different from those at points lying out of the set; similar distinction is also relevant in the analysis of the generalized distance function. While subdifferentiation of the generalized distance function at points lying in $\operatorname{gph} F$ has been investigated to a certain extent by Thibault in [36], little has been known about that at points lying out of $\operatorname{gph} F$. In this section, both cases are dealt with whenever possible.

The first proposition provides upper estimates of ε-subdifferentials of the generalized distance function via enlargements.

Proposition 2.2.1. ([26, Proposition 3.1]) *Let $F : Z \rightrightarrows X$ be a set-valued mapping, $(\bar{z}, \bar{x}) \in Z \times X$ and $r = \rho(\bar{z}, \bar{x})$. For any $\varepsilon \geq 0$, the following statements hold:*

(a) $\widehat{\partial}_\varepsilon \rho(\bar{z}, \bar{x}) \subset \left\{(z^*, x^*) \in \widehat{N}_\varepsilon((\bar{z}, \bar{x}); \operatorname{gph} F_r) : \|x^*\| \leq 1 + \varepsilon\right\}$.

(b) *If $r > 0$, then*

$$\widehat{\partial}_\varepsilon \rho(\bar{z}, \bar{x}) \subset \left\{(z^*, x^*) \in \widehat{N}_\varepsilon((\bar{z}, \bar{x}); \operatorname{gph} F_r) : 1 - \varepsilon \leq \|x^*\| \leq 1 + \varepsilon\right\}.$$

Proof. (a) Let $\varepsilon \geq 0$, $(z^*, x^*) \in \widehat{\partial}_\varepsilon \rho(\bar{z}, \bar{x})$ and $\gamma > 0$. By Proposition 1.5.5, there exists $\delta > 0$ such that for all $(z, x) \in Z \times X$ with $\|z - \bar{z}\| \leq \delta$ and $\|x - \bar{x}\| \leq \delta$,

$$\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \leq \rho(z, x) - \rho(\bar{z}, \bar{x}) + (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x - \bar{x}\|). \quad (2.3)$$

Fix any $(z,x) \in \operatorname{gph} F_r$ with $\|z-\bar{z}\| \le \delta$ and $\|x-\bar{x}\| \le \delta$. Then $x \in F_r(z)$ and hence $\rho(z,x) = d(x, F(z)) \le r = \rho(\bar{z},\bar{x})$, reducing (2.3) to

$$\langle z^*, z-\bar{z}\rangle + \langle x^*, x-\bar{x}\rangle \le (\varepsilon+\gamma)(\|z-\bar{z}\| + \|x-\bar{x}\|).$$

Applying Proposition 1.6.5, one sees that $(z^*,x^*) \in \widehat{N}_\varepsilon((\bar{z},\bar{x}); \operatorname{gph} F_r)$.

Moreover, by taking $z=\bar{z}$ in (2.3), one has for all $x \in X$ with $\|x-\bar{x}\| \le \delta$,

$$\begin{aligned}\langle x^*, x-\bar{x}\rangle &\le \rho(\bar{z},x) - \rho(\bar{z},\bar{x}) + (\varepsilon+\gamma)\|x-\bar{x}\| \\ &= d(x, F(\bar{z})) - d(\bar{x}, F(\bar{z})) + (\varepsilon+\gamma)\|x-\bar{x}\|.\end{aligned}$$

Proposition 1.5.5 implies $x^* \in \widehat{\partial}_\varepsilon d(\bar{x}, F(\bar{z}))$. In view of Proposition 1.5.7(a), $\|x^*\| \le 1+\varepsilon$. The assertion is verified.

(b) Let $\varepsilon \ge 0$ and $(z^*,x^*) \in \widehat{\partial}_\varepsilon \rho(\bar{z},\bar{x})$. By **(a)**, $(z^*,x^*) \in \widehat{N}_\varepsilon((\bar{z},\bar{x}); \operatorname{gph} F_r)$ and it has also been shown that $x^* \in \widehat{\partial}_\varepsilon d(\bar{x}, F(\bar{z}))$. It follows from $d(\bar{x}, F(\bar{z})) = \rho(\bar{z},\bar{x}) = r > 0$ that $\bar{x} \notin \operatorname{cl} F(\bar{z})$. Employing Proposition 1.5.7(b) gives $1-\varepsilon \le \|x^*\| \le 1+\varepsilon$. This completes the proof of the proposition. □

Corollary 2.2.2. *Let $\Omega \subset X$, $\bar{x} \in X$ and $r = d(\bar{x}, \Omega)$. For any $\varepsilon \ge 0$, the following statements hold:*

(a) $\widehat{\partial}_\varepsilon d(\bar{x}, \Omega) \subset \widehat{N}_\varepsilon(\bar{x}; \Omega_r) \cap (1+\varepsilon)\boldsymbol{B}_{X^*}$.

(b) *If $r > 0$, then* $\widehat{\partial}_\varepsilon d(\bar{x}, \Omega) \subset \widehat{N}_\varepsilon(\bar{x}; \Omega_r) \cap [1-\varepsilon, 1+\varepsilon]\boldsymbol{S}_{X^*}$.

If ρ is locally Lipschitz at $(\bar{z},\bar{x})$, it is also possible to obtain lower estimates. Note that an extra constant which depends on the local Lipschitz rank of ρ is involved.

Theorem 2.2.3. (cf. [26, Theorem 3.2]) *Let $F : Z \rightrightarrows X$ be a set-valued mapping, $(\bar{z},\bar{x}) \in Z \times X$ and $r = \rho(\bar{z},\bar{x})$. Suppose ρ is locally Lipschitz at $(\bar{z},\bar{x})$ with rank $\ell \ge 0$. The following statements hold:*

(a) *If $r = 0$, then for any $\varepsilon \geq 0$,*

$$\left\{(z^*, x^*) \in \widehat{N}_\varepsilon((\bar{z}, \bar{x}); \operatorname{gph} F_0) : \|x^*\| \leq 1 + \varepsilon\right\} \subset \widehat{\partial}_{(2\ell+1)\varepsilon}\rho(\bar{z}, \bar{x}).$$

(b) *If $r > 0$, then for any $\varepsilon \geq 0$,*

$$\left\{(z^*, x^*) \in \widehat{N}_\varepsilon((\bar{z}, \bar{x}); \operatorname{gph} F_r) : 1 - \varepsilon \leq \|x^*\| \leq 1 + \varepsilon\right\} \subset \widehat{\partial}_{(2\ell+1)\varepsilon}\rho(\bar{z}, \bar{x}).$$

Proof. Since ρ is locally Lipschitz at $(\bar{z}, \bar{x})$ with rank ℓ, by Proposition 2.1.6, ρ_r is also locally Lipschitz at $(\bar{z}, \bar{x})$ with rank ℓ for any $r \geq 0$. Hence there exists $\delta_1 > 0$ such that for any $(z_i, x_i) \in Z \times X$ with $\|z_i - \bar{z}\| \leq \delta_1$ and $\|x_i - \bar{x}\| \leq \delta_1$, where $i = 1, 2$,

$$|\rho(z_1, x_1) - \rho(z_2, x_2)| \leq \ell(\|x_1 - x_2\| + \|z_1 - z_2\|), \text{ and} \tag{2.4}$$

$$|\rho_r(z_1, x_1) - \rho_r(z_2, x_2)| \leq \ell(\|x_1 - x_2\| + \|z_1 - z_2\|). \tag{2.5}$$

(a) Let $\varepsilon \geq 0$, $\gamma > 0$ and $(z^*, x^*) \in \widehat{N}_\varepsilon((\bar{z}, \bar{x}); \operatorname{gph} F_0)$ with $\|x^*\| \leq 1 + \varepsilon$. Owing to Proposition 1.6.5, there exists $\delta_1 \geq \delta_2 > 0$ such that for all $(z, x) \in \operatorname{gph} F_0$ with $\|z - \bar{z}\| \leq \delta_2$ and $\|x - \bar{x}\| \leq \delta_2$,

$$\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \leq (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x - \bar{x}\|). \tag{2.6}$$

Take $\delta_3 = \min\left\{\frac{\delta_2}{4(\ell+1)}, \frac{1}{2}\right\} > 0$. Fix any $(z, x) \neq (\bar{z}, \bar{x})$ with $\|z - \bar{z}\| \leq \delta_3$ and $\|x - \bar{x}\| \leq \delta_3$. If $(z, x) \in \operatorname{gph} F_0$, then $x \in F_0(z)$ and $\rho(z, x) = d(x, F(z)) = 0 = \rho(\bar{z}, \bar{x})$. Thus (2.6) is equivalent to

$$\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \leq \rho(z, x) - \rho(\bar{z}, \bar{x}) + (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x - \bar{x}\|). \tag{2.7}$$

Otherwise $(z, x) \notin \operatorname{gph} F_0$. Note that $(\|z - \bar{z}\| + \|x - \bar{x}\|)^2 > 0$. Choose $x_1 \in F(z)$ such that

$$\begin{aligned} \|x - x_1\| &< \rho(z, x) + (\|z - \bar{z}\| + \|x - \bar{x}\|)^2 \\ &= (\rho(z, x) - \rho(\bar{z}, \bar{x})) + (\|z - \bar{z}\| + \|x - \bar{x}\|)^2 \\ &\leq \ell(\|x - \bar{x}\| + \|z - \bar{z}\|) + \|z - \bar{z}\| + \|x - \bar{x}\| \\ &\leq 2(\ell + 1)\delta_3 \leq (\ell + 1)\frac{2\delta_2}{4(\ell + 1)} = \frac{\delta_2}{2}. \end{aligned} \tag{2.8}$$

Check that $\|x_1 - \bar{x}\| \leq \|x_1 - x\| + \|x - \bar{x}\| \leq \frac{\delta_2}{2} + \delta_3 \leq \frac{\delta_2}{2} + \frac{\delta_2}{4(\ell+1)} \leq \delta_2$ and $\|z - \bar{z}\| \leq \delta_3 \leq \delta_2$. Moreover, since $x_1 \in F(z)$, one has $d(x_1, F(z)) = 0$. This implies $x_1 \in F_0(z)$ and $(z, x_1) \in \operatorname{gph} F_0$. It follows from (2.6) that

$$\langle z^*, z - \bar{z}\rangle + \langle x^*, x_1 - \bar{x}\rangle \leq (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x_1 - \bar{x}\|). \tag{2.9}$$

Using estimates (2.4), (2.8) and (2.9), one sees that

$$\begin{aligned}
&\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \\
&= \langle z^*, z - \bar{z}\rangle + \langle x^*, x_1 - \bar{x}\rangle + \langle x^*, x - x_1\rangle \\
&\leq (\varepsilon + \gamma)(\|x_1 - \bar{x}\| + \|z - \bar{z}\|) + \langle x^*, x - x_1\rangle \\
&\leq (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x - \bar{x}\| + \|x - x_1\|) + \|x^*\|\|x - x_1\| \\
&\leq (\varepsilon + \gamma + \|x^*\|)\|x - x_1\| + (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x - \bar{x}\|) \\
&\leq (2\varepsilon + \gamma + 1)(\rho(z, x) + (\|x - \bar{x}\| + \|z - \bar{z}\|)^2) \\
&\quad + (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x - \bar{x}\|) \\
&\leq (2\varepsilon + \gamma + 1)(\|x - \bar{x}\| + \|z - \bar{z}\|)^2 + (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x - \bar{x}\|) \\
&\quad + \rho(z, x) - \rho(\bar{z}, \bar{x}) + (2\varepsilon + \gamma)(\rho(z, x) - \rho(\bar{z}, \bar{x})) \\
&\leq \rho(z, x) - \rho(\bar{z}, \bar{x}) + (2\varepsilon + \gamma + 1)(\|x - \bar{x}\| + \|z - \bar{z}\|)^2 \\
&\quad + (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x - \bar{x}\|) + \ell(2\varepsilon + \gamma)(\|x - \bar{x}\| + \|z - \bar{z}\|) \\
&= \rho(z, x) - \rho(\bar{z}, \bar{x}) + (2\varepsilon + \gamma + 1)(\|x - \bar{x}\| + \|z - \bar{z}\|)^2 \\
&\quad + ((2\ell + 1)\varepsilon + (\ell + 1)\gamma)(\|z - \bar{z}\| + \|x - \bar{x}\|).
\end{aligned} \tag{2.10}$$

Rearranging inequalities (2.7) and (2.10), there holds

$$\begin{aligned}
&\frac{\rho(z, x) - \rho(\bar{z}, \bar{x}) - \langle (z^*, x^*), (z, x) - (\bar{z}, \bar{x})\rangle}{\|(z, x) - (\bar{z}, \bar{x})\|} \\
&\geq \begin{cases} -(\varepsilon + \gamma) & \text{if } (z, x) \in \operatorname{gph} F_0, \\ -(2\ell + 1)\varepsilon - (\ell + 1)\gamma - (2\varepsilon + \gamma + 1)(\|z - \bar{z}\| + \|x - \bar{x}\|) & \text{if } (z, x) \notin \operatorname{gph} F_0. \end{cases}
\end{aligned}$$

Since $\gamma > 0$ is arbitrary, passing to the limit, one has

$$\liminf_{(z,x)\to(\bar{z},\bar{x})} \frac{\rho(z, x) - \rho(\bar{z}, \bar{x}) - \langle (z^*, x^*), (z, x) - (\bar{z}, \bar{x})\rangle}{\|(z, x) - (\bar{z}, \bar{x})\|} \geq -(2\ell + 1)\varepsilon.$$

By definition, $(z^*, x^*) \in \widehat{\partial}_{(2\ell+1)\varepsilon}\rho(\bar{z}, \bar{x})$. The assertion holds.

(b) Let $\varepsilon \geq 0$, $\eta > 0$ and $(z^*, x^*) \in \widehat{N}_\varepsilon((\bar{z}, \bar{x}); \operatorname{gph} F_r)$ with $1 - \varepsilon \leq \|x^*\| \leq 1 + \varepsilon$. Due to Proposition 1.6.5, there exists $\delta_1 \geq \delta_4 > 0$ such that for all $(z, x) \in \operatorname{gph} F_r$ with $\|z - \bar{z}\| \leq \delta_4$ and $\|x - \bar{x}\| \leq \delta_4$,

$$\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \leq (\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{x}\|). \tag{2.11}$$

On the other hand, $d(\bar{x}, F(\bar{z})) = \rho(\bar{z}, \bar{x}) = r$ implies $\bar{x} \in F_r(\bar{z})$ and $\rho_r(\bar{z}, \bar{x}) = d(\bar{x}, F_r(\bar{z})) = 0$. Note also that $F_r \equiv (F_r)_0$. Applying the result of **(a)** to F_r and ρ_r in place of F and ρ respectively reveals that $(z^*, x^*) \in \widehat{\partial}_{(2\ell+1)\varepsilon}\rho_r(\bar{z}, \bar{x})$. Employing Proposition 1.5.5, there exists $\delta_4 \geq \delta_5 > 0$ such that for all $(z, x) \in Z \times X$ with $\|z - \bar{z}\| \leq \delta_5$ and $\|x - \bar{x}\| \leq \delta_5$,

$$\begin{aligned} &\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \\ &\leq \rho_r(z, x) - \rho_r(\bar{z}, \bar{x}) + ((2\ell + 1)\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{x}\|) \\ &= \rho_r(z, x) + ((2\ell + 1)\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{x}\|). \end{aligned} \tag{2.12}$$

Take $\delta_6 = \frac{\delta_5}{2\ell+1} > 0$. Fix any $(z, x) \neq (\bar{z}, \bar{x})$ with $\|z - \bar{z}\| \leq \delta_6$ and $\|x - \bar{x}\| \leq \delta_6$. If $(z, x) \notin \operatorname{gph} F_r$, then Proposition 2.1.5 implies $\rho_r(z, x) = \rho(z, x) - r$. In light of (2.12), one has

$$\begin{aligned} &\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \\ &\leq \rho(z, x) - r + ((2\ell + 1)\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{x}\|) \\ &= \rho(z, x) - \rho(\bar{z}, \bar{x}) + ((2\ell + 1)\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{x}\|). \end{aligned} \tag{2.13}$$

If $(z, x) \in \operatorname{gph} F_r$, then $x \in F_r(z)$ and $\rho(z, x) = d(x, F(z)) \leq r = \rho(\bar{z}, \bar{x})$. Choose $x_0 \in X$ with $\|x_0\| = 1$ such that $1 - \varepsilon - \eta \leq \|x^*\| - \eta < \langle x^*, x_0\rangle$. Take

$\widetilde{x} = x + (\rho(\bar{z},\bar{x}) - \rho(z,x))x_0$. Note that

$$\begin{aligned} d(\widetilde{x}, F(z)) &= \inf_{y\in F(z)} \|\widetilde{x} - y\| \\ &= \inf_{y\in F(z)} \|x + (\rho(\bar{z},\bar{x}) - \rho(z,x))x_0 - y\| \\ &\le \inf_{y\in F(z)} \|x - y\| + \|(\rho(\bar{z},\bar{x}) - \rho(z,x))x_0\| \\ &= d(x, F(z)) + \rho(\bar{z},\bar{x}) - \rho(z,x) \\ &= \rho(z,x) + r - \rho(z,x) = r, \end{aligned}$$

which implies $\widetilde{x} \in F_r(z)$ and $(z,\widetilde{x}) \in \operatorname{gph} F_r$. Moreover,

$$\begin{aligned} \|\widetilde{x} - \bar{x}\| &\le \|\widetilde{x} - x\| + \|x - \bar{x}\| \\ &= \|(\rho(\bar{z},\bar{x}) - \rho(z,x))x_0\| + \|x - \bar{x}\| \\ &= |\rho(\bar{z},\bar{x}) - \rho(z,x)| + \|x - \bar{x}\| \\ &\le \ell(\|z - \bar{z}\| + \|x - \bar{x}\|) + \|x - \bar{x}\| \\ &\le (2\ell + 1)\delta_6 = \delta_5 \le \delta_4 \end{aligned}$$

and $\|z - \bar{z}\| \le \delta_6 \le \delta_4$. It follows from (2.11) and the above estimates that

$$\begin{aligned} &\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \\ &= \langle z^*, z - \bar{z}\rangle + \langle x^*, \widetilde{x} - \bar{x}\rangle + \langle x^*, x - \widetilde{x}\rangle \\ &\le (\varepsilon + \eta)(\|z - \bar{z}\| + \|\widetilde{x} - \bar{x}\|) + \langle x^*, (\rho(z,x) - \rho(\bar{z},\bar{x}))x_0\rangle \\ &\le (\varepsilon + \eta)(\|z - \bar{z}\| + \ell(\|z - \bar{z}\| + \|x - \bar{x}\|) + \|x - \bar{x}\|) \\ &\quad + (\rho(z,x) - \rho(\bar{z},\bar{x}))(1 - \varepsilon - \eta) \\ &\le (\varepsilon + \eta)(\ell + 1)(\|z - \bar{z}\| + \|x - \bar{x}\|) + \rho(z,x) - \rho(\bar{z},\bar{x}) \\ &\quad + (\varepsilon + \eta)(\rho(\bar{z},\bar{x}) - \rho(z,x)) \\ &\le (\varepsilon + \eta)(\ell + 1)(\|z - \bar{z}\| + \|x - \bar{x}\|) + \rho(z,x) - \rho(\bar{z},\bar{x}) \\ &\quad + \ell(\varepsilon + \eta)(\|x - \bar{x}\| + \|z - \bar{z}\|) \\ &= (2\ell + 1)(\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{x}\|) + \rho(z,x) - \rho(\bar{z},\bar{x}). \end{aligned} \tag{2.14}$$

Rearranging inequalities (2.13) and (2.14) yields

$$\frac{\rho(z,x)-\rho(\bar{z},\bar{x})-\langle(z^*,x^*),(z,x)-(\bar{z},\bar{x})\rangle}{\|(z,x)-(\bar{z},\bar{x})\|}\geq\begin{cases}-(2\ell+1)(\varepsilon+\eta) & \text{if } (z,x)\in\operatorname{gph}F_r,\\ -((2\ell+1)\varepsilon+\eta) & \text{if } (z,x)\notin\operatorname{gph}F_r.\end{cases}$$

Since $\eta>0$ is arbitrary, passing to the limit, one sees that

$$\liminf_{(z,x)\to(\bar{z},\bar{x})}\frac{\rho(z,x)-\rho(\bar{z},\bar{x})-\langle(z^*,x^*),(z,x)-(\bar{z},\bar{x})\rangle}{\|(z,x)-(\bar{z},\bar{x})\|}\geq-(2\ell+1)\varepsilon.$$

By definition, $(z^*,x^*)\in\widehat{\partial}_{(2\ell+1)\varepsilon}\rho(\bar{z},\bar{x})$. This proves the desired inclusion. □

Corollary 2.2.4. *Let $\Omega\subset X$, $\bar{x}\in X$ and $r=d(\bar{x},\Omega)$. The following statements hold:*

(a) *If $r=0$, then for any $\varepsilon\geq 0$, $\widehat{N}_\varepsilon(\bar{x};\Omega_0)\cap(1+\varepsilon)\boldsymbol{B}_{X^*}\subset\widehat{\partial}_{3\varepsilon}d(\bar{x},\Omega)$.*

(b) *If $r>0$, then for any $\varepsilon\geq 0$, $\widehat{N}_\varepsilon(\bar{x};\Omega_r)\cap[1-\varepsilon,1+\varepsilon]\boldsymbol{S}_{X^*}\subset\widehat{\partial}_{3\varepsilon}d(\bar{x},\Omega)$.*

Corollary 2.2.5. *Let $F:Z\rightrightarrows X$ be a closed-graph mapping, $(\bar{z},\bar{x})\in Z\times X$ and $r=\rho(\bar{z},\bar{x})$. Suppose ρ is locally Lipschitz at $(\bar{z},\bar{x})$. The following statements hold:*

(a) *If $(\bar{z},\bar{x})\in\operatorname{gph}F$, then $\widehat{\partial}\rho(\bar{z},\bar{x})=\left\{(z^*,x^*)\in\widehat{N}((\bar{z},\bar{x});\operatorname{gph}F):\|x^*\|\leq 1\right\}$.*

(b) *If $(\bar{z},\bar{x})\notin\operatorname{gph}F$, then $\widehat{\partial}\rho(\bar{z},\bar{x})=\left\{(z^*,x^*)\in\widehat{N}((\bar{z},\bar{x});\operatorname{gph}F_r):\|x^*\|=1\right\}$.*

Proof. By assumption, F is closed-graph. If $(\bar{z},\bar{x})\in\operatorname{gph}F$, then $r=d(\bar{x},F(\bar{z}))=0$ and $\operatorname{gph}F_r=\operatorname{gph}F$. On the other hand, if $(\bar{z},\bar{x})\notin\operatorname{gph}F$, then $r=d(\bar{x},F(\bar{z}))>0$. The first equality follows by taking $\varepsilon=0$ in Proposition 2.2.1(a) and Theorem 2.2.3(a). Similarly, the second equality follows by taking $\varepsilon=0$ in Proposition 2.2.1(b) and Theorem 2.2.3(b). □

Corollary 2.2.6. *Let $\Omega\subset X$ be closed, $\bar{x}\notin\Omega$ and $r=d(\bar{x},\Omega)$. Then*

$$\widehat{\partial}d(\bar{x},\Omega)=\widehat{N}(\bar{x};\Omega_r)\cap\boldsymbol{S}_{X^*}.$$

Another equality connecting Fréchet subdifferentials of the generalized distance function and Fréchet normal cones to $\operatorname{gph}F_r$ further sheds light on the duality between these two families.

Proposition 2.2.7. (cf. [26, Proposition 3.4]) *Let $F : Z \rightrightarrows X$ be a set-valued mapping, $(\bar{z}, \bar{x}) \in Z \times X$ and $r = \rho(\bar{z}, \bar{x})$. Suppose ρ is locally Lipschitz at $(\bar{z}, \bar{x})$. Then*

$$\widehat{N}((\bar{z}, \bar{x}); \operatorname{gph} F_r) = \bigcup_{\lambda \geq 0} \lambda \widehat{\partial} \rho(\bar{z}, \bar{x}).$$

Proof. Let $(z^*, x^*) \in \widehat{N}((\bar{z}, \bar{x}); \operatorname{gph} F_r)$. Suppose $x^* \neq 0$. Then $\lambda := \|x^*\| > 0$. Noting that $\widehat{N}((\bar{z}, \bar{x}); \operatorname{gph} F_r)$ is a cone, one has $\frac{1}{\lambda}(z^*, x^*) \in \widehat{N}((\bar{z}, \bar{x}); \operatorname{gph} F_r)$. Invoking Theorem 2.2.3(a) and (b) with $\varepsilon = 0$, one sees that $\frac{1}{\lambda}(z^*, x^*) \in \widehat{\partial} \rho(\bar{z}, \bar{x})$ and hence $(z^*, x^*) \in \lambda \widehat{\partial} \rho(\bar{z}, \bar{x})$. Otherwise $x^* = 0$. Observe that $d(\bar{x}, F(\bar{z})) = \rho(\bar{z}, \bar{x}) = r$ implies $\bar{x} \in F_r(\bar{z})$ and $(\bar{z}, \bar{x}) \in \operatorname{gph} F_r$. Since ρ is locally Lipschitz at $(\bar{z}, \bar{x})$, it follows from Proposition 2.1.6 that ρ_r is also locally Lipschitz at $(\bar{z}, \bar{x})$. By virtue of Theorem 1.7.4, F_r is Lipschitz-like at $(\bar{z}, \bar{x})$ with some rank $\ell \geq 0$. Applying Theorem 1.7.5 yields

$$\sup \left\{ \|z^*\| : (z^*, 0) \in \widehat{N}((\bar{z}, \bar{x}); \operatorname{gph} F_r) \right\} = \sup \left\{ \|z^*\| : z^* \in \widehat{D}^* F_r(\bar{z}, \bar{x})(0) \right\} \leq 0.$$

Hence $\|z^*\| = 0$ and $z^* = 0$. Consequently, $(z^*, x^*) = (0, 0) \in 0 \cdot \widehat{\partial} \rho(\bar{z}, \bar{x})$. In both cases, the inclusion $\widehat{N}((\bar{z}, \bar{x}); \operatorname{gph} F_r) \subset \bigcup_{\lambda \geq 0} \lambda \widehat{\partial} \rho(\bar{z}, \bar{x})$ is valid.

Consider the opposite inclusion. By Proposition 2.2.1, $\widehat{\partial} \rho(\bar{z}, \bar{x}) \subset \widehat{N}((\bar{z}, \bar{x}); \operatorname{gph} F_r)$. Since $\widehat{N}((\bar{z}, \bar{x}); \operatorname{gph} F_r)$ is a cone, $\lambda \widehat{\partial} \rho(\bar{z}, \bar{x}) \subset \widehat{N}((\bar{z}, \bar{x}); \operatorname{gph} F_r)$ for all $\lambda \geq 0$ and hence $\bigcup_{\lambda \geq 0} \lambda \widehat{\partial} \rho(\bar{z}, \bar{x}) \subset \widehat{N}((\bar{z}, \bar{x}); \operatorname{gph} F_r)$. Thus the desired equality holds. □

Corollary 2.2.8. *Let $\Omega \subset X$, $\bar{x} \in X$ and $r = d(\bar{x}, \Omega)$. Then*

$$\widehat{N}(\bar{x}; \Omega_r) = \bigcup_{\lambda \geq 0} \lambda \widehat{\partial} d(\bar{x}, \Omega).$$

The next proposition provides upper estimates of ε-subdifferentials via projections introduced in Definition 1.4.3.

Proposition 2.2.9. (cf. [26, Proposition 3.5]) *Let $F : Z \rightrightarrows X$ be a closed-graph mapping, $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$ with $\Pi(\bar{x}, F(\bar{z})) \neq \varnothing$, and $\varepsilon \geq 0$. Then for any $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$,*

$$\widehat{\partial}_\varepsilon \rho(\bar{z}, \bar{x}) \subset \left\{ (z^*, x^*) \in \widehat{N}_\varepsilon((\bar{z}, \bar{y}); \operatorname{gph} F) : 1 - \varepsilon \leq \|x^*\| \leq 1 + \varepsilon \right\}.$$

Proof. Let $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$, $(z^*, x^*) \in \widehat{\partial}_\varepsilon \rho(\bar{z}, \bar{x})$ and $\eta > 0$. Thus $\bar{y} \in F(\bar{z})$, which implies $(\bar{z}, \bar{y}) \in \operatorname{gph} F$ and $\widehat{N}_\varepsilon((\bar{z}, \bar{y}); \operatorname{gph} F)$ is well-defined. Using Proposition 1.5.5, there exists $\delta > 0$ such that for all $(z, x) \in Z \times X$ with $\|z - \bar{z}\| \le \delta$ and $\|x - \bar{x}\| \le \delta$,

$$\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \le \rho(z, x) - \rho(\bar{z}, \bar{x}) + (\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{x}\|). \tag{2.15}$$

Fix any $(z, x) \in \operatorname{gph} F$ with $\|z - \bar{z}\| \le \delta$ and $\|x - \bar{y}\| \le \delta$. Then $x \in F(z)$ and hence $\rho(z, x) = d(x, F(z)) = 0$. Note that $\|\bar{x} - \bar{y}\| = d(\bar{x}, F(\bar{z})) = \rho(\bar{z}, \bar{x})$ and $\|(x - \bar{y} + \bar{x}) - \bar{x}\| = \|x - \bar{y}\| \le \delta$. Employing (2.15), one has

$$\begin{aligned}
&\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{y}\rangle \\
&= \langle z^*, z - \bar{z}\rangle + \langle x^*, (x - \bar{y} + \bar{x}) - \bar{x}\rangle \\
&\le \rho(z, x - \bar{y} + \bar{x}) - \rho(\bar{z}, \bar{x}) + (\varepsilon + \eta)(\|z - \bar{z}\| + \|(x - \bar{y} + \bar{x}) - \bar{x}\|) \\
&\le \rho(z, x) + \|\bar{x} - \bar{y}\| - \|\bar{x} - \bar{y}\| + (\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{y}\|) \\
&= (\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{y}\|).
\end{aligned}$$

In light of Proposition 1.6.5, $(z^*, x^*) \in \widehat{N}_\varepsilon((\bar{z}, \bar{y}); \operatorname{gph} F)$.

Moreover, by taking $z = \bar{z}$ in (2.15), one sees that for all $x \in X$ with $\|x - \bar{x}\| \le \delta$,

$$\begin{aligned}
\langle x^*, x - \bar{x}\rangle &\le \rho(\bar{z}, x) - \rho(\bar{z}, \bar{x}) + (\varepsilon + \eta)\|x - \bar{x}\| \\
&= d(x, F(\bar{z})) - d(\bar{x}, F(\bar{z})) + (\varepsilon + \eta)\|x - \bar{x}\|.
\end{aligned}$$

By Proposition 1.5.5 again, $x^* \in \widehat{\partial}_\varepsilon d(\bar{x}, F(\bar{z}))$. Since F is closed-graph, $F(\bar{z})$ is closed; $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$ implies that $\bar{x} \notin F(\bar{z}) = \operatorname{cl} F(\bar{z})$. Applying Proposition 1.5.7(b) yields $1 - \varepsilon \le \|x^*\| \le 1 + \varepsilon$. This completes the proof of the proposition. □

Corollary 2.2.10. *Let $\Omega \subset X$ be closed, $\bar{x} \notin \Omega$ with $\Pi(\bar{x}, \Omega) \neq \varnothing$, and $\varepsilon \ge 0$. Then for any $\bar{y} \in \Pi(\bar{x}, \Omega)$,*

$$\widehat{\partial}_\varepsilon d(\bar{x}, \Omega) \subset \widehat{N}_\varepsilon(\bar{y}; \Omega) \cap [1 - \varepsilon, 1 + \varepsilon]\boldsymbol{S}_{X^*}.$$

However, the requirement that the projection set be nonempty is often too stringent for application. It is desirable to obtain an analogue of the previous proposition without using projections. Indeed, similar but more involved estimates may be obtained by means of *Ekeland's Variational Principle* and *perturbed projections* stated respectively in Theorem 1.2.1 and Definition 2.1.3.

Theorem 2.2.11. ([26, Theorem 3.6]) *Let $F : Z \rightrightarrows X$ be a closed-graph mapping, $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$ and $\varepsilon \geq 0$. Then for any $\eta > 0$,*

$$\widehat{\partial}_\varepsilon \rho(\bar{z}, \bar{x}) \subset \bigcup_{(v,u)\in \Theta_\eta^F(\bar{z},\bar{x})} \left\{(z^*, x^*) \in \widehat{N}_{\varepsilon+\eta}((v,u); \operatorname{gph} F) : 1-\varepsilon \leq \|x^*\| \leq 1+\varepsilon\right\}.$$

Proof. Let $\eta > 0$, $(z^*, x^*) \in \widehat{\partial}_\varepsilon \rho(\bar{z}, \bar{x})$ and $0 < \gamma < \frac{\eta}{2}$. Using Proposition 1.5.5, there exists $\delta > 0$ such that for all $(z, x) \in Z \times X$ with $\|z - \bar{z}\| \leq \delta$ and $\|x - \bar{x}\| \leq \delta$,

$$\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \leq \rho(z, x) - \rho(\bar{z}, \bar{x}) + (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x - \bar{x}\|). \tag{2.16}$$

Let $0 < \widetilde{\eta} < \min\left\{\gamma, \frac{\delta}{2}, 1\right\}$. There exists $\bar{y} \in F(\bar{z})$ such that

$$\|\bar{x} - \bar{y}\| < d(\bar{x}, F(\bar{z})) + \widetilde{\eta}^2 = \rho(\bar{z}, \bar{x}) + \widetilde{\eta}^2. \tag{2.17}$$

Let $W = \operatorname{gph} F \cap \boldsymbol{B}_{Z\times X}((\bar{z}, \bar{y}), \delta)$, which is a closed and hence complete metric space. Define $\varphi : W \to \mathbb{R}$ by

$$\varphi(z, x) = -\langle z^*, z - \bar{z}\rangle - \langle z^*, x - \bar{y}\rangle + (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x - \bar{y}\|) + \widetilde{\eta}^2.$$

Then φ is continuous and in particular lower semicontinuous on W. For any $(z, x) \in W$, one has $x \in F(z)$ and hence $\rho(z, x) = d(x, F(z)) = 0$. Moreover, $\|z - \bar{z}\| \leq \delta$ and $\|x - \bar{y}\| \leq \delta$, from which $\|(x - \bar{y} + \bar{x}) - \bar{x}\| = \|x - \bar{y}\| \leq \delta$ follows. In view of (2.16) and (2.17), one sees that

$$\begin{aligned}
&\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{y}\rangle \\
&= \langle z^*, z - \bar{z}\rangle + \langle x^*, (x - \bar{y} + \bar{x}) - \bar{x}\rangle \\
&\leq \rho(z, x - \bar{y} + \bar{x}) - \rho(\bar{z}, \bar{x}) + (\varepsilon + \gamma)(\|z - \bar{z}\| + \|(x - \bar{y} + \bar{x}) - \bar{x}\|) \\
&< (\rho(z, x) + \|\bar{x} - \bar{y}\|) - (\|\bar{x} - \bar{y}\| - \widetilde{\eta}^2) + (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x - \bar{y}\|) \\
&= \widetilde{\eta}^2 + (\varepsilon + \gamma)(\|z - \bar{z}\| + \|x - \bar{y}\|).
\end{aligned} \tag{2.18}$$

Rearranging (2.18) yields $\varphi(z,x) > 0 = \varphi(\bar{z},\bar{y}) - \widetilde{\eta}^2$ for all $(z,x) \in W$, which implies

$$\inf_{(z,x)\in W} \varphi(z,x) \geq \varphi(\bar{z},\bar{y}) - \widetilde{\eta}^2,$$

and φ is bounded below. Employing *Ekeland's Variational Principle* (Theorem 1.2.1), there exists $(v,u) \in W$ with $\|v-\bar{z}\| \leq \widetilde{\eta}$ and $\|u-\bar{y}\| \leq \widetilde{\eta}$ such that for all $(z,x) \in W$, $\varphi(z,x) + \widetilde{\eta}(\|v-z\| + \|u-x\|) \geq \varphi(v,u)$, which is equivalent to

$$\begin{aligned}
&\langle z^*, z-v\rangle + \langle x^*, x-u\rangle \\
&\leq (\varepsilon+\gamma)(\|z-\bar{z}\| - \|v-\bar{z}\| + \|x-\bar{y}\| - \|u-\bar{y}\|) + \widetilde{\eta}(\|v-z\| + \|u-x\|) \\
&\leq (\varepsilon+\gamma)(\|z-v\| + \|x-u\|) + \widetilde{\eta}(\|z-v\| + \|x-u\|) \\
&\leq (\varepsilon+2\gamma)(\|z-v\| + \|x-u\|) \\
&\leq (\varepsilon+\eta)(\|z-v\| + \|x-u\|). \qquad (2.19)
\end{aligned}$$

Fix any $(z,x) \in \operatorname{gph} F$ with $\|z-v\| \leq \widetilde{\eta}$ and $\|x-u\| \leq \widetilde{\eta}$. Note that $\|z-\bar{z}\| \leq \|z-v\| + \|v-\bar{z}\| \leq 2\widetilde{\eta} \leq \delta$ and $\|x-\bar{y}\| \leq \|x-u\| + \|u-\bar{y}\| \leq 2\widetilde{\eta} \leq \delta$. Hence $(z,x) \in \operatorname{gph} F \cap \boldsymbol{B}_{Z\times X}((v,u),\delta) = W$. It follows from (2.19) that for any $\lambda > 0$,

$$\begin{aligned}
\langle z^*, z-v\rangle + \langle x^*, x-u\rangle &\leq (\varepsilon+\eta)(\|z-v\| + \|x-u\|) \\
&\leq ((\varepsilon+\eta)+\lambda)(\|z-v\| + \|x-u\|).
\end{aligned}$$

In light of Proposition 1.6.5, $(z^*,x^*) \in \widehat{N}_{\varepsilon+\eta}((v,u);\operatorname{gph} F)$, which is well-defined since $(v,u) \in \operatorname{gph} F$.

Moreover, in view of (2.17),

$$\begin{aligned}
\|u-\bar{x}\| \leq \|u-\bar{y}\| + \|\bar{y}-\bar{x}\| &\leq \widetilde{\eta} + \rho(\bar{z},\bar{x}) + \widetilde{\eta}^2 \\
&< \rho(\bar{z},\bar{x}) + 2\widetilde{\eta} < \rho(\bar{z},\bar{x}) + 2\gamma \\
&< \rho(\bar{z},\bar{x}) + \eta
\end{aligned}$$

and $\|v-\bar{z}\| \leq \widetilde{\eta} < \eta$, which imply $(v,u) \in \Theta_\eta^F(\bar{z},\bar{x})$. Owing to the assumptions that F is closed-graph and $(\bar{z},\bar{x}) \notin \operatorname{gph} F$, $F(\bar{z})$ is closed and $\bar{x} \notin F(\bar{z}) = \operatorname{cl} F(\bar{z})$. Invoking Proposition 1.5.7(b) again, one has $1-\varepsilon \leq \|x^*\| \leq 1+\varepsilon$. Consequently, the assertion is substantiated. □

Remark 2.2.12. Unlike in Proposition 2.2.9, the projection set $\Pi(\bar{x}, F(\bar{z}))$ is *not* assumed to be nonempty in this theorem. While $\Pi(\bar{x}, F(\bar{z}))$ may be empty, the perturbed projection set $\Theta_\eta^F(\bar{z}, \bar{x})$ is guaranteed by *Ekeland's Variational Principle* to be nonempty.

Corollary 2.2.13. *Let $\Omega \subset X$ be closed, $\bar{x} \notin \Omega$ and $\varepsilon \geq 0$. Then for any $\eta > 0$,*

$$\widehat{\partial}_\varepsilon d(\bar{x}, \Omega) \subset \bigcup_{x \in \Pi_\eta(\bar{x}, \Omega)} \left(\widehat{N}_{\varepsilon+\eta}(x; \Omega) \cap [1-\varepsilon, 1+\varepsilon] \boldsymbol{S}_{X^*} \right).$$

2.3 Limiting and Singular Subdifferentials of the Generalized Distance Function

This section is devoted to developing estimates of limiting and singular subdifferentials of the generalized distance function via limiting normal cones and mixed coderivatives first discussed in Definition 1.6.7 and Definition 1.7.1 respectively. Results of the previous section prove to be essential tools in this section.

Recall from Corollary 2.2.8 that Fréchet normal cones to $\operatorname{gph} F_r$ can be delineated as unions of nonnegative multiples of Fréchet subdifferentials of the generalized distance function. Indeed, it was proved by Thibault in [36] using *Ekeland's Variational Principle* that the same relation holds for limiting normal cones to $\operatorname{gph} F$ and limiting subdifferentials of the generalized distance function at points *belonging to* $\operatorname{gph} F$.

Theorem 2.3.1. ([36, Proposition 2.7]) *Let $F : Z \rightrightarrows X$ be a closed-graph mapping and $(\bar{z}, \bar{x}) \in \operatorname{gph} F$. Then*

$$N((\bar{z}, \bar{x}); \operatorname{gph} F) = \bigcup_{\lambda \geq 0} \lambda \partial \rho(\bar{z}, \bar{x}).$$

A primary motivation for the study in this section is to attempt to extend the preceding equality to points *not belonging to* $\operatorname{gph} F$. However, it turns out that limiting subdifferentials of the generalized distance function at such points are too large for the equality to hold. For this reason, smaller limiting constructions are needed. The reader is suggested to refer to [25] for further development.

Definition 2.3.2. Let $f : X \to \overline{\mathbb{R}}$ be finite at $\bar{x} \in X$.

(a) The ***right-sided limiting subdifferential of*** f ***at*** $\bar{x}$ is defined by

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

The elements of this set are known as ***right-sided limiting subgradients of*** f ***at*** $\bar{x}$.

(b) The ***right-sided singular subdifferential of*** f ***at*** $\bar{x}$ is defined by

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

The elements of this set are known as ***right-sided singular subgradients of*** f ***at*** $\bar{x}$.

Remarks 2.3.3. (i) It follows immediately that $\widehat{\partial} f(\bar{x}) \subset \partial_{\geq} f(\bar{x}) \subset \partial f(\bar{x})$ and $\partial_{\geq}^{\infty} f(\bar{x}) \subset \partial^{\infty} f(\bar{x})$.

(ii) An important observation is that $\partial_{\geq} f(\bar{x}) = \partial f(\bar{x})$ if f attains a *local minimum* at $\bar{x}$. In particular, $\partial_{\geq} d(\bar{x}, \Omega) = \partial d(\bar{x}, \Omega)$ for any $\Omega \subset X$ with $\bar{x} \in \Omega$.

In nonsmooth calculus problems, it is often necessary to impose *additional compactness requirements* in order to arrive at interesting results. One such requirement which ensures equivalence between weak* convergence and norm convergence of sequences in sets of ε-normals to zero is especially relevant to the subsequent exposition. A more elaborate explanation of this property is available in [25].

Definition 2.3.4. Let $X, X_1, \cdots, X_n$ be Banach spaces.

(a) A subset $\Omega \subset X$ is said to be ***sequentially normally compact at*** $\bar{x} \in \Omega$ if for any sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{x_k\}_{k=1}^{\infty} \subset \Omega$ and $\{x_k^*\}_{k=1}^{\infty} \subset X^*$ such that $\varepsilon_k \downarrow 0$, $x_k \to \bar{x}$, $x_k^* \in \widehat{N}_{\varepsilon_k}(x_k; \Omega)$ for all $k \in \mathbb{N}$ and $x_k^* \xrightarrow{w^*} 0$, one has $\|x_k^*\| \to 0$.

(b) A subset $\Omega \subset X_1 \times \cdots \times X_n$ is said to be ***sequentially normally compact with respect to*** X_i, where $1 \leq i \leq n$, ***at*** $(\bar{x}_1, \cdots, \bar{x}_n) \in \Omega$ if for any sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{(x_k^1, \cdots, x_k^n)\}_{k=1}^{\infty} \subset \Omega$ and $\{(x_k^{1*}, \cdots, x_k^{n*})\}_{k=1}^{\infty} \subset X_1^* \times \cdots \times X_n^*$ such that $\varepsilon_k \downarrow 0$, $(x_k^1, \cdots, x_k^n) \to (\bar{x}_1, \cdots, \bar{x}_n)$, $(x_k^{1*}, \cdots, x_k^{n*}) \in \widehat{N}_{\varepsilon_k}((x_k^1, \cdots, x_k^n); \Omega)$ for all $k \in \mathbb{N}$ and $x_k^{i*} \xrightarrow{w^*} 0$, one has $\|x_k^{i*}\| \to 0$.

Remarks 2.3.5. (i) If $\bar{x} \in \operatorname{cl} \Omega \subset X$, then the sequential normal compactness of $\operatorname{cl} \Omega$ at $\bar{x}$ implies that of Ω at $\bar{x}$.

(ii) If X is finite dimensional, then $\Omega \subset X$ is automatically sequentially normally compact at any $\bar{x} \in \Omega$.

(iii) If X_i is finite dimensional, then $\Omega \subset X_1 \times \cdots \times X_n$ is automatically sequentially normally compact with respect to X_i at any $(\bar{x}_1, \cdots, \bar{x}_n) \in \Omega$.

Drawing on the tool of enlargements defined in Definition 2.1.3 as in the last section, the next theorem provides important upper estimates of right-sided limiting subdifferentials of the generalized distance function under various assumptions.

Theorem 2.3.6. (cf. [26, Theorem 4.3]) *Let $F : Z \rightrightarrows X$ be a set-valued mapping, $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$ and $r = \rho(\bar{z}, \bar{x})$. Suppose $\operatorname{gph} F$ is locally closed at $(\bar{z}, \bar{x})$ and $\operatorname{gph} F_r$ is closed. The following statements hold:*

(a) $\partial_{\geq}\rho(\bar{z}, \bar{x}) \subset \{(z^*, x^*) \in N((\bar{z}, \bar{x}); \operatorname{gph} F_r) : \|x^*\| \leq 1\}$.

(b) *If $\operatorname{gph} F_r \subset Z \times X$ is sequentially normally compact with respect to X at $(\bar{z}, \bar{x})$, then*

$$\partial_{\geq}\rho(\bar{z}, \bar{x}) \subset \{(z^*, x^*) \in N((\bar{z}, \bar{x}); \operatorname{gph} F_r) : 0 < \|x^*\| \leq 1\}.$$

(c) *If X is finite dimensional, then*

$$\partial_{\geq}\rho(\bar{z}, \bar{x}) \subset \{(z^*, x^*) \in N((\bar{z}, \bar{x}); \operatorname{gph} F_r) : \|x^*\| = 1\}.$$

(d) *If ρ is locally Lipschitz at $(\bar{z}, \bar{x})$, then*

$$\{(z^*, x^*) \in N((\bar{z}, \bar{x}); \operatorname{gph} F_r) : 0 < \|x^*\| \leq 1\} \subset \bigcup_{\lambda > 0} \lambda \partial_{\geq}\rho(\bar{z}, \bar{x}).$$

Proof. (a) Let $(z^*, x^*) \in \partial_{\geq}\rho(\bar{z}, \bar{x})$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^{\infty} \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty} \subset Z^* \times X^*$ such that

$$\varepsilon_k \downarrow 0, \tag{2.20}$$

$$(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x}), \tag{2.21}$$

$$(z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*), \tag{2.22}$$

$$\rho(z_k, x_k) \geq \rho(\bar{z}, \bar{x}) \text{ for all } k \in \mathbb{N}, \text{ and} \tag{2.23}$$

$$(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho(z_k, x_k) \text{ for all } k \in \mathbb{N}. \tag{2.24}$$

Since gph F is locally closed at $(\bar{z}, \bar{x})$ and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$, by Proposition 2.1.2, $r = \rho(\bar{z}, \bar{z}) > 0$. In view of (2.23), for all $k \in \mathbb{N}$, $d(x_k, F(z_k)) = \rho(z_k, x_k) > 0$, which implies $x_k \notin F(z_k)$ and $(z_k, x_k) \notin \operatorname{gph} F$.

Suppose there are infinitely many (z_k, x_k) such that $\rho(z_k, x_k) = r$. By passing to this subsequence of $\{(z_k, x_k)\}_{k=1}^{\infty}$ together with the corresponding subsequences of $\{\varepsilon_k\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, assume that $d(x_k, F(z_k)) = \rho(z_k, x_k) = r$ and hence $(z_k, x_k) \in \operatorname{gph} F_r$ for all $k \in \mathbb{N}$. By virtue of (2.21), $(z_k, x_k) \xrightarrow{\operatorname{gph} F_r} (\bar{z}, \bar{x})$. For each $k \in \mathbb{N}$, employing (2.24) and Proposition 2.2.1(b), one sees that

$$(z_k^*, x_k^*) \in \widehat{N}_{\varepsilon_k}((z_k, x_k); \operatorname{gph} F_r) \text{ and } 1 - \varepsilon_k \leq \|x_k^*\| \leq 1 + \varepsilon_k. \tag{2.25}$$

Otherwise, by considering the tail of $\{(z_k, x_k)\}_{k=1}^{\infty}$ together with the corresponding terms of $\{\varepsilon_k\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, assume that $d(x_k, F(z_k)) = \rho(z_k, x_k) > r$ and hence $(z_k, x_k) \notin \operatorname{gph} F_r$ for all $k \in \mathbb{N}$. Let $\gamma > 0$. Employing (2.24) and Proposition 1.5.5, for each $k \in \mathbb{N}$, there exists $\delta_k > 0$ such that for all $(z, x) \in Z \times X$ with $\|z - z_k\| \leq \delta_k$ and $\|x - x_k\| \leq \delta_k$,

$$\langle z_k^*, z - z_k\rangle + \langle x_k^*, x - x_k\rangle \leq \rho(z, x) - \rho(z_k, x_k) + (\varepsilon_k + \gamma)(\|z - z_k\| + \|x - x_k\|). \tag{2.26}$$

Using Proposition 2.1.5, $\rho(z_k, x_k) = \rho_r(z_k, x_k) + r$. Fix any $k \in \mathbb{N}$ and $(z, x) \in Z \times X$ with $\|z - z_k\| \leq \delta_k$ and $\|x - x_k\| \leq \delta_k$. If $(z, x) \in \operatorname{gph} F_r$, then $x \in F_r(z)$, which implies $\rho(z, x) = d(x, F(z)) \leq r$ and $\rho_r(z, x) = d(x, F_r(z)) = 0$. Hence $\rho(z, x) \leq \rho_r(z, x) + r$. If $(z, x) \notin \operatorname{gph} F_r$, due to Proposition 2.1.5 again, $\rho(z, x) = \rho_r(z, x) + r$. In both cases, it follows from (2.26) that

$$\begin{aligned}
&\langle z_k^*, z - z_k\rangle + \langle x_k^*, x - x_k\rangle \\
&\leq \rho(z, x) - \rho(z_k, x_k) + (\varepsilon_k + \gamma)(\|z - z_k\| + \|x - x_k\|) \\
&\leq (\rho_r(z, x) + r) - (\rho_r(z_k, x_k) + r) + (\varepsilon_k + \gamma)(\|z - z_k\| + \|x - x_k\|) \\
&\leq \rho_r(z, x) - \rho_r(z_k, x_k) + (\varepsilon_k + \gamma)(\|z - z_k\| + \|x - x_k\|).
\end{aligned}$$

Owing to Proposition 1.5.5, $(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho_r(z_k, x_k)$. For all $k \in \mathbb{N}$, let $\eta_k = \rho_r(z_k, x_k) = \rho(z_k, x_k) - r > 0$. In view of (2.21), $\eta_k \downarrow 0$. For each $k \in \mathbb{N}$, applying Theorem 2.2.11 to F_r and ρ_r in place of F and ρ respectively, there exists $(v_k, u_k) \in \Theta_{\eta_k}^{F_r}(z_k, x_k)$ such that

$$(z_k^*, x_k^*) \in \widehat{N}_{\varepsilon_k+\eta_k}((v_k, u_k); \operatorname{gph} F_r) \text{ and } 1 - \varepsilon_k \le \|x_k^*\| \le 1 + \varepsilon_k. \tag{2.27}$$

Then $(v_k, u_k) \in \operatorname{gph} F_r$, $\|z_k - v_k\| \le \eta_k$ and $\|x_k - u_k\| \le \rho_r(z_k, x_k) + \eta_k = 2\eta_k$. Note that $\|v_k - \bar{z}\| \le \|v_k - z_k\| + \|z_k - \bar{z}\|$ and $\|u_k - \bar{x}\| \le \|u_k - x_k\| + \|x_k - \bar{x}\|$. It follows from $\eta_k, \varepsilon_k \downarrow 0$ and $(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x})$ that $\varepsilon_k + \eta_k \downarrow 0$ and $(v_k, u_k) \xrightarrow{\operatorname{gph} F_r} (\bar{z}, \bar{x})$.

In both cases, $(z^*, x^*) \in N((\bar{z}, \bar{x}); \operatorname{gph} F_r)$. Using the lower semicontinuity of $\|\cdot\|$ with respect to the weak* topology of X^* in connection with the second relation of both (2.25) and (2.27) yields

$$\|x^*\| \le \liminf_{k\to\infty} \|x_k^*\| \le \liminf_{k\to\infty}(1 + \varepsilon_k) = 1.$$

Hence $\partial_{\ge}\rho(\bar{z}, \bar{x}) \subset \{(z^*, x^*) \in N((\bar{z}, \bar{x}); \operatorname{gph} F_r) : \|x^*\| \le 1\}$ holds.

(b) Let $(z^*, x^*) \in \partial_{\ge}\rho(\bar{z}, \bar{x})$. By **(a)**, $(z^*, x^*) \in N((\bar{z}, \bar{x}); \operatorname{gph} F_r)$ and $\|x^*\| \le 1$. It suffices to show that $\|x^*\| > 0$. Suppose $\|x^*\| = 0$. Then $x^* = 0$ and (2.22) implies $x_k^* \xrightarrow{w^*} 0$. Since $\operatorname{gph} F_r$ is sequentially normally compact with respect to X at $(\bar{z}, \bar{x})$, one has $\|x_k^*\| \to 0$. On the other hand, owing to the second relation of both (2.25) and (2.27) and $\varepsilon_k \downarrow 0$, $\|x_k^*\| \to 1$, which is a contradiction. This shows that $\|x^*\| > 0$ and completes the proof.

(c) Let $(z^*, x^*) \in \partial_{\ge}\rho(\bar{z}, \bar{x})$. By **(a)**, $(z^*, x^*) \in N((\bar{z}, \bar{x}); \operatorname{gph} F_r)$. It suffices to show that $\|x^*\| = 1$. Since X is finite dimensional, $\|\cdot\|$ is continuous with respect to the weak* topology of X^*. Letting $k \to \infty$ in the second relation of both (2.25) and (2.27) gives $\|x^*\| = \lim_{k\to\infty} \|x_k^*\| = 1$. The result follows.

(d) Let $(z^*, x^*) \in N((\bar{z}, \bar{x}); \operatorname{gph} F_r)$ with $0 < \|x^*\| \le 1$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^{\infty} \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty} \subset Z^* \times X^*$

such that

$$\varepsilon_k \downarrow 0, \tag{2.28}$$

$$(z_k, x_k) \xrightarrow{\text{gph} F_r} (\bar{z}, \bar{x}), \tag{2.29}$$

$$(z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*), \text{ and} \tag{2.30}$$

$$(z_k^*, x_k^*) \in \widehat{N}_{\varepsilon_k}((z_k, x_k); \text{gph} F_r) \text{ for all } k \in \mathbb{N}. \tag{2.31}$$

Suppose ρ is locally Lipschitz at $(\bar{z}, \bar{x})$ with some rank $\ell \geq 0$. There exists $\delta' > 0$ such that ρ is locally Lipschitz on $\boldsymbol{B}_{Z\times X}((\bar{z}, \bar{x}), \delta')$ with rank ℓ. Let $0 < \delta < \delta'$. Owing to (2.29), by considering the tail of $\{(z_k, x_k)\}_{k=1}^{\infty}$ together with the corresponding terms of $\{\varepsilon_k\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, assume that for all $k \in \mathbb{N}$, $(z_k, x_k) \in \boldsymbol{B}_{Z\times X}((\bar{z}, \bar{x}), \delta)$, so that ρ is locally Lipschitz at (z_k, x_k) with rank ℓ.

Since $\|\cdot\|$ is lower semicontinuous with respect to the weak* topology of X^*, it follows from (2.30) that $\beta_1 := \liminf_{k\to\infty} \|x_k^*\| \geq \|x^*\| > 0$. There exists $K \in \mathbb{N}$ such that $\|x_k^*\| \geq \beta_1 - \frac{1}{2}\|x^*\| \geq \frac{1}{2}\|x^*\| := \beta_2 > 0$ for all $k \geq K$. Again, by considering the tail of $\{(z_k, x_k)\}_{k=1}^{\infty}$ together with the corresponding terms of $\{\varepsilon_k\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, assume that $\|x_k^*\| \geq \beta_2 > 0$ for all $k \in \mathbb{N}$. Then for each $k \in \mathbb{N}$, it follows from (2.31) and Proposition 1.6.6 that

$$\begin{aligned}\frac{1}{\|x_k^*\|}(z_k^*, x_k^*) &\in \left\{(z^*, x^*) \in \widehat{N}_{\frac{\varepsilon_k}{\|x_k^*\|}}((z_k, x_k); \text{gph} F_r) : \|x^*\| = 1\right\} \\ &\subset \left\{(z^*, x^*) \in \widehat{N}_{\frac{\varepsilon_k}{\beta_2}}((z_k, x_k); \text{gph} F_r) : \|x^*\| = 1\right\},\end{aligned}$$

which upon applying Theorem 2.2.3(b) yields

$$\frac{1}{\|x_k^*\|}(z_k^*, x_k^*) \in \widehat{\partial}_{\frac{(2\ell+1)\varepsilon_k}{\beta_2}} \rho(z_k, x_k).$$

Note that $\frac{(2\ell+1)\varepsilon_k}{\beta_2} \downarrow 0$ due to (2.28).

In view of (2.29), $(z_k, x_k) \in \text{gph} F_r$, which implies $x_k \in F_r(z_k)$ and $\rho(z_k, x_k) = d(x_k, F(z_k)) \leq r = \rho(\bar{z}, \bar{x})$. Suppose $\rho(z_k, x_k) < r$ for infinitely many $k \in \mathbb{N}$. With regard to (2.28), there exists $M \in \mathbb{N}$ such that $\varepsilon_M < \beta_2$ and $\rho(z_M, x_M) < r$.

Let $\gamma > 0$. Invoking (2.31) and Proposition 1.6.5, there exists $\widehat{\delta} > 0$ such that for all $(z, x) \in \operatorname{gph} F_r$ with $\|z - z_M\| \le \widehat{\delta}$ and $\|x - x_M\| \le \widehat{\delta}$,

$$\langle z_M^*, z - z_M \rangle + \langle x_M^*, x - x_M \rangle \le (\varepsilon_M + \gamma)(\|z - z_M\| + \|x - x_M\|). \tag{2.32}$$

Let $\widehat{\eta} = \min\{\widehat{\delta}, r - \rho(z_M, x_M)\} > 0$. For all $x \in X$ with $\|x - x_M\| \le \widehat{\eta}$,

$$\begin{aligned} d(x, F(z_M)) &\le \|x - x_M\| + d(x_M, F(z_M)) \\ &\le \widehat{\eta} + d(x_M, F(z_M)) \le r - \rho(z_M, x_M) + \rho(z_M, x_M) = r \end{aligned}$$

and $(z_M, x) \in \operatorname{gph} F_r$, implying by (2.32) that

$$\langle x_M^*, x - x_M \rangle \le (\varepsilon_M + \gamma)\|x - x_M\|.$$

In light of the linearity of x_M^*, one sees that

$$\|x_M^*\| = \sup_{x \ne x_M} \frac{\langle x_M^*, x - x_M \rangle}{\|x - x_M\|} = \sup_{0 < \|x - x_M\| \le \widehat{\eta}} \frac{\langle x_M^*, x - x_M \rangle}{\|x - x_M\|} \le \varepsilon_M + \gamma.$$

Since $\gamma > 0$ is arbitrary, $\|x_M^*\| \le \varepsilon_M < \beta_2$, which contradicts the earlier statement that $\|x_k^*\| \ge \beta_2$ for all $k \in \mathbb{N}$. Hence $\rho(z_k, x_k) = r$ for sufficiently large $k \in \mathbb{N}$. Once again, by considering the tail of $\{(z_k, x_k)\}_{k=1}^{\infty}$ together with the corresponding terms of $\{\varepsilon_k\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, assume that $\rho(z_k, x_k) = r$ for all $k \in \mathbb{N}$. In particular, it follows from (2.29) that $(z_k, x_k) \xrightarrow{\rho+} (\bar{z}, \bar{x})$.

Moreover, the convergence of $\{x_k\}_{k=1}^{\infty}$ implies that $\{\|x_k\|\}_{k=1}^{\infty}$ is a bounded sequence in $\mathbb{R}$ and has a convergent subsequence by Bolzano-Weierstrass theorem. By further passing to this subsequence of $\{(z_k, x_k)\}_{k=1}^{\infty}$ together with the corresponding subsequences of $\{\varepsilon_k\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, assume that $\|x_k^*\| \to \mu$ for some $\mu \ge \beta_1 > 0$. Owing to (2.30), $\frac{1}{\|x_k^*\|}(z_k^*, x_k^*) \xrightarrow{w^*} \frac{1}{\mu}(z^*, x^*)$. Therefore $\frac{1}{\mu}(z^*, x^*) \in \partial_{\ge}\rho(\bar{z}, \bar{x})$ and hence

$$(z^*, x^*) \in \mu\partial_{\ge}\rho(\bar{z}, \bar{x}) \subset \bigcup_{\lambda > 0} \lambda\partial_{\ge}\rho(\bar{z}, \bar{x}).$$

The conclusion is established. □

Indeed, the main argument used in the proof of Theorem 2.3.6(a) is also crucial in the proof of a forthcoming key result. The conclusion of the argument is hereby advantageously restated as a separate lemma.

Lemma 2.3.7. *Let $F : Z \rightrightarrows X$ be a set-valued mapping, $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$ and $r = \rho(\bar{z}, \bar{x})$. Suppose $\operatorname{gph} F$ is locally closed at $(\bar{z}, \bar{x})$ and $\operatorname{gph} F_r$ is closed. Assume further that $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^{\infty} \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty} \subset Z^* \times X^*$ are sequences satisfying $\varepsilon_k \downarrow 0$, $(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x})$, $\rho(z_k, x_k) \geq \rho(\bar{z}, \bar{x})$, and $(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k} \rho(z_k, x_k)$ for all $k \in \mathbb{N}$. Then there exist two sequences $\{\gamma_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{(v_k, u_k)\}_{k=1}^{\infty} \subset Z \times X$ and a subsequence $\{(\widetilde{z}_k^*, \widetilde{x}_k^*)\}_{k=1}^{\infty}$ of $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ such that for all $k \in \mathbb{N}$,*

$$(\widetilde{z}_k^*, \widetilde{x}_k^*) \in \widehat{N}_{\gamma_k}((v_k, u_k); \operatorname{gph} F_r), \quad 1 - \gamma_k \leq \|\widetilde{x}_k^*\| \leq 1 + \gamma_k;$$

and

$$\gamma_k \downarrow 0, \qquad (v_k, u_k) \xrightarrow{\operatorname{gph} F_r} (\bar{z}, \bar{x}).$$

Having established Theorem 2.3.6, the desired alternative description of limiting normal cones to $\operatorname{gph} F_r$ may be derived without much difficulty.

Corollary 2.3.8. (cf. [26, Corollary 4.4]) *Let $F : Z \rightrightarrows X$ be a set-valued mapping, $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$ and $r = \rho(\bar{z}, \bar{x})$. Suppose $\operatorname{gph} F$ is locally closed at $(\bar{z}, \bar{x})$ and $\operatorname{gph} F_r$ is closed. Assume further that $\operatorname{gph} F_r \subset Z \times X$ is sequentially normally compact with respect to X at $(\bar{z}, \bar{x})$ and ρ is locally Lipschitz at $(\bar{z}, \bar{x})$. Then*

$$N((\bar{z}, \bar{x}); \operatorname{gph} F_r) = \bigcup_{\lambda \geq 0} \lambda \partial_{\geq} \rho(\bar{z}, \bar{x}).$$

Proof. By Theorem 2.3.6(a), $N((\bar{z}, \bar{x}); \operatorname{gph} F_r) \supset \partial_{\geq} \rho(\bar{z}, \bar{x})$. Since $N((\bar{z}, \bar{x}); \operatorname{gph} F_r)$ is a cone, $N((\bar{z}, \bar{x}); \operatorname{gph} F_r) \supset \lambda \partial_{\geq} \rho(\bar{z}, \bar{x})$ for all $\lambda \geq 0$, which justifies the inclusion

$$N((\bar{z}, \bar{x}); \operatorname{gph} F_r) \supset \bigcup_{\lambda \geq 0} \lambda \partial_{\geq} \rho(\bar{z}, \bar{x}).$$

Consider the reverse inclusion. Let $(z^*, x^*) \in N((\bar{z}, \bar{x}); \operatorname{gph} F_r)$. Suppose $x^* \neq 0$. Then $\|x^*\| > 0$. Noting that $N((\bar{z}, \bar{x}); \operatorname{gph} F_r)$ is a cone, one has

$$\frac{1}{\|x^*\|}(z^*, x^*) \in \{(z^*, x^*) \in N((\bar{z}, \bar{x}); \operatorname{gph} F_r) : 0 < \|x^*\| \leq 1\}.$$

An application of Theorem 2.3.6(d) yields $\frac{1}{\|x^*\|}(z^*,x^*) \in \bigcup_{\lambda>0} \lambda\partial_{\geq}\rho(\bar{z},\bar{x})$ and hence $(z^*,x^*) \in \bigcup_{\lambda>0} \lambda\partial_{\geq}\rho(\bar{z},\bar{x})$.

Otherwise $x^* = 0$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{(z_k,x_k)\}_{k=1}^{\infty} \subset Z \times X$ and $\{(z_k^*,x_k^*)\}_{k=1}^{\infty} \subset Z^* \times X^*$ such that

$$\varepsilon_k \downarrow 0, \tag{2.33}$$

$$(z_k,x_k) \xrightarrow{\text{gph}\, F_r} (\bar{z},\bar{x}), \tag{2.34}$$

$$(z_k^*,x_k^*) \xrightarrow{w^*} (z^*,0), \text{ and} \tag{2.35}$$

$$(z_k^*,x_k^*) \in \widehat{N}_{\varepsilon_k}((z_k,x_k);\text{gph}\, F_r) \text{ for all } k \in \mathbb{N}. \tag{2.36}$$

Since ρ is locally Lipschitz at $(\bar{z},\bar{x})$, by Proposition 2.1.6, ρ_r is also locally Lipschitz at $(\bar{z},\bar{x})$. Moreover, $d(\bar{x},F(\bar{z})) = \rho(\bar{z},\bar{x}) = r$ implies that $\bar{x} \in F_r(\bar{z})$ and $(\bar{z},\bar{x}) \in \text{gph}\, F_r$. In light of Theorem 1.7.4, F_r is locally Lipschitz-like at $(\bar{z},\bar{x})$ with some rank $\ell \geq 0$. For each $k \in \mathbb{N}$, employing Theorem 1.7.5(a), there exists $\eta_k > 0$ independent of x_k^* such that for all $(z,x) \in \text{gph}\, F_r$ with $\|z-\bar{z}\| \leq \eta_k$ and $\|x-\bar{x}\| \leq \eta_k$,

$$\sup\left\{\|z^*\| : z^* \in \widehat{D}_{\varepsilon_k}F_r(z,x)(-x_k^*)\right\} \leq \ell\|-x_k^*\| + \varepsilon_k(1+\ell),$$

that is,
$$\sup\left\{\|z^*\| : (z^*,x_k^*) \in \widehat{N}_{\varepsilon_k}((z,x);\text{gph}\, F_r)\right\} \leq \ell\|x_k^*\| + \varepsilon_k(1+\ell). \tag{2.37}$$

In view of (2.34), $(z_k,x_k) \in \text{gph}\, F_r$ for all $k \in \mathbb{N}$, and by passing to a subsequence of $\{(z_k,x_k)\}_{k=1}^{\infty}$ together with the corresponding subsequences of $\{\varepsilon_k\}_{k=1}^{\infty}$ and $\{(z_k^*,x_k^*)\}_{k=1}^{\infty}$ if necessary, assume that $\|z_k-\bar{z}\| \leq \eta_k$ and $\|x_k-\bar{x}\| \leq \eta_k$ for all $k \in \mathbb{N}$. It follows from (2.36) and (2.37) that for each $k \in \mathbb{N}$,

$$\|z_k^*\| \leq \sup\left\{\|z^*\| : (z^*,x_k^*) \in \widehat{N}_{\varepsilon_k}((z_k,x_k);\text{gph}\, F_r)\right\} \leq \ell\|x_k^*\| + \varepsilon_k(1+\ell). \tag{2.38}$$

Note that (2.35) and the sequential normal compactness of $\text{gph}\, F_r$ with respect to X at $(\bar{z},\bar{x})$ imply that $\|x_k^*\| \to 0$. Together with (2.33) and (2.38), one sees that $\|z_k^*\| \to 0$. Invoking the lower semicontinuity of $\|\cdot\|$ with respect to the weak* topology of Z^* gives

$$\|z^*\| \leq \liminf_{k\to\infty} \|z_k^*\| = 0.$$

Therefore $\|z^*\| = 0$ and $z^* = 0$. Hence $(z^*, x^*) = (0,0) \in 0 \cdot \partial_{\geq}\rho(\bar{z}, \bar{x}) \subset \bigcup_{\lambda \geq 0} \lambda\partial_{\geq}\rho(\bar{z}, \bar{x})$.

In both cases, the reverse inclusion

$$N((\bar{z}, \bar{x}); \operatorname{gph} F_r) \subset \bigcup_{\lambda \geq 0} \lambda\partial_{\geq}\rho(\bar{z}, \bar{x})$$

is valid. The proof of the corollary is complete. □

Corollary 2.3.9. *Let $\Omega \subset X$ be closed, $\bar{x} \notin \Omega$ and $r = d(\bar{x}, \Omega)$. Suppose Ω_r is also closed. The following statements hold:*

(a) $\partial_{\geq} d(\bar{x}, \Omega) \subset N(\bar{x}; \Omega_r) \cap \boldsymbol{B}_{X^*}$.

(b) *If Ω_r is sequentially normally compact at $\bar{x}$, then*

$$\partial_{\geq} d(\bar{x}, \Omega) \subset N(\bar{x}; \Omega_r) \cap (\boldsymbol{B}_{X^*} \backslash \{0\}).$$

(c) *If X is finite dimensional, then*

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

(d) $N(\bar{x}; \Omega_r) \cap (\boldsymbol{B}_{X^*} \backslash \{0\}) \subset \bigcup_{\lambda > 0} \lambda\partial_{\geq} d(\bar{x}, \Omega)$.

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

Proof. As in most corollaries, **(a)**, **(b)**, **(c)** and **(d)** follow from Theorem 2.3.6(a), (b), (c) and (d) respectively by taking $Z = \{\bar{z}\}$ and $F \equiv \Omega$. It suffices to prove **(e)**. Note that **(a)** implies $\partial_{\geq} d(\bar{x}, \Omega) \subset N(\bar{x}; \Omega_r)$. Due to the fact that $N(\bar{x}; \Omega_r)$ is a cone, $\lambda\partial_{\geq} d(\bar{x}, \Omega) \subset N(\bar{x}; \Omega_r)$ for all $\lambda \geq 0$ and hence

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

Consider the opposite inclusion. Let $x^* \in N(\bar{x}; \Omega_r) \backslash \{0\}$. Then $\|x^*\| > 0$. Using the fact that $N(\bar{x}; \Omega_r)$ is a cone again, $\frac{x^*}{\|x^*\|} \in N(\bar{x}; \Omega_r) \cap (\boldsymbol{B}_{X^*} \backslash \{0\})$. In light of **(d)**, one has

$\frac{x^*}{\|x^*\|} \in \bigcup_{\lambda>0} \lambda\partial_{\geq} d(\bar{x}, \Omega)$ and thus $x^* \in \bigcup_{\lambda>0} \lambda\partial_{\geq} d(\bar{x}, \Omega)$. Moreover, $\{0\} = 0 \cdot \partial_{\geq} d(\bar{x}, \Omega) \subset \bigcup_{\lambda\geq 0} \lambda\partial_{\geq} d(\bar{x}, \Omega)$. Consequently, there holds

$$N(\bar{x}; \Omega_r) \subset \bigcup_{\lambda\geq 0} \lambda\partial_{\geq} d(\bar{x}, \Omega). \qquad \square$$

Remark 2.3.10. Unlike Corollary 2.3.8, Corollary 2.3.9(e) does not impose any sequential normal compactness assumption.

The next key theorem contends that singular subdifferentials of the generalized distance function at points *belonging to* gph F may be described in terms of mixed coderivatives introduced in Section 1.7. Since the proof is rather involved, the theorem is established using the following lemma, which is also of independent interest.

Lemma 2.3.11. (cf. [26, Lemma 4.6]) *Let* $F : Z \rightrightarrows X$ *be a set-valued mapping and* $(\bar{z}, \bar{x}) \in \operatorname{gph} F$. *Suppose* ρ *is upper semicontinuous at* $(\bar{z}, \bar{x})$. *Then for any* $\varepsilon \geq 0$, $\gamma > 0$ *and* $(z^*, x^*) \in \widehat{N}_\varepsilon((\bar{z}, \bar{x}); \operatorname{gph} F)$,

$$(z^*, x^*) \in (\|x^*\| + \varepsilon + \gamma)\widehat{\partial}_{\frac{\varepsilon}{\|x^*\|+\varepsilon+\gamma}} \rho(\bar{z}, \bar{x}).$$

Proof. Let $\varepsilon \geq 0$, $\gamma > 0$ and $(z^*, x^*) \in \widehat{N}_\varepsilon((\bar{z}, \bar{x}); \operatorname{gph} F)$. Fix $0 < \eta < \gamma$. By Proposition 1.6.5, there exists $0 < \delta_1 < 1$ such that for all $(z, x) \in \operatorname{gph} F$ with $\|z - \bar{z}\| \leq \delta_1$ and $\|x - \bar{x}\| \leq \delta_1$,

$$\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \leq (\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{x}\|). \tag{2.39}$$

Since $(\bar{z}, \bar{x}) \in \operatorname{gph} F$, $\bar{x} \in F(\bar{z})$ and $\rho(\bar{z}, \bar{x}) = d(\bar{x}, F(\bar{z})) = 0$. In view of the upper semicontinuity of ρ at $(\bar{z}, \bar{x})$, there exists $\delta_2 > 0$ such that for all $(z, x) \in Z \times X$ with $\|z - \bar{z}\| \leq \delta_2$ and $\|x - \bar{x}\| \leq \delta_2$,

$$\rho(z, x) = \rho(z, x) - \rho(\bar{z}, \bar{x}) \leq \frac{\delta_1}{4}. \tag{2.40}$$

Take $\delta = \min\left\{\frac{\delta_1}{4}, \delta_2\right\} > 0$. Fix any $(z, x) \neq (\bar{z}, \bar{x})$ with $\|z - \bar{z}\| \leq \delta$ and $\|x - \bar{x}\| \leq \delta$. If $(z, x) \in \operatorname{gph} F$, then $x \in F(z)$ and $\rho(z, x) = d(x, F(z)) = 0$. Thus (2.39) becomes

$$\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \leq (\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{x}\|) + (\|x^*\| + \varepsilon + \gamma)(\rho(z, x) - \rho(\bar{z}, \bar{x})). \tag{2.41}$$

Otherwise $(z,x) \notin \operatorname{gph} F$. Note that $(\|z-\bar{z}\|+\|x-\bar{x}\|)^2 > 0$. In light of (2.40), it possible to choose $x_1 \in F(z)$ such that

$$\begin{aligned}\|x-x_1\| &< d(x,F(z)) + (\|z-\bar{z}\|+\|x-\bar{x}\|)^2 \\ &= \rho(z,x) + (\|z-\bar{z}\|+\|x-\bar{x}\|)^2 \\ &\le \frac{\delta_1}{4} + 4\delta^2 \le \frac{\delta_1}{4} + \frac{\delta_1^2}{4} \\ &\le \frac{\delta_1}{4} + \frac{\delta_1}{4} = \frac{\delta_1}{2}.\end{aligned} \tag{2.42}$$

Moreover, $\|x_1-\bar{x}\| \le \|x_1-x\| + \|x-\bar{x}\| \le \frac{\delta_1}{2} + \delta \le \frac{\delta_1}{2} + \frac{\delta_1}{4} \le \delta_1$. Using estimates (2.39) and (2.42),

$$\begin{aligned}&\langle z^*, z-\bar{z}\rangle + \langle x^*, x-\bar{x}\rangle \\ =&\langle z^*, z-\bar{z}\rangle + \langle x^*, x_1-\bar{x}\rangle + \langle x^*, x-x_1\rangle \\ \le&(\varepsilon+\eta)(\|z-\bar{z}\|+\|x_1-\bar{x}\|) + \langle x^*, x-x_1\rangle \\ \le&(\varepsilon+\eta)(\|z-\bar{z}\|+\|x-\bar{x}\|+\|x-x_1\|) + \|x^*\|\|x-x_1\| \\ \le&(\|x^*\|+\varepsilon+\eta)\,\|x-x_1\| + (\varepsilon+\eta)(\|z-\bar{z}\|+\|x-\bar{x}\|) \\ \le&(\|x^*\|+\varepsilon+\gamma)(\rho(z,x)+(\|z-\bar{z}\|+\|x-\bar{x}\|)^2) \\ &\quad + (\varepsilon+\eta)(\|z-\bar{z}\|+\|x-\bar{x}\|) \\ \le&(\|x^*\|+\varepsilon+\gamma)(\|z-\bar{z}\|+\|x-\bar{x}\|)^2 + (\varepsilon+\eta)(\|z-\bar{z}\|+\|x-\bar{x}\|) \\ &\quad + (\|x^*\|+\varepsilon+\gamma)(\rho(z,x)-\rho(\bar{z},\bar{x})).\end{aligned} \tag{2.43}$$

Rearranging inequalities (2.41) and (2.43), there holds

$$\begin{aligned}&\frac{\rho(z,x)-\rho(\bar{z},\bar{x}) - \left\langle \frac{1}{\|x^*\|+\varepsilon+\gamma}(z^*,x^*), (z,x)-(\bar{z},\bar{x})\right\rangle}{\|(z,x)-(\bar{z},\bar{x})\|} \\ &\ge \begin{cases} -\frac{\varepsilon+\eta}{\|x^*\|+\varepsilon+\gamma} & \text{if } (z,x)\in \operatorname{gph} F, \\ -\frac{\varepsilon+\eta}{\|x^*\|+\varepsilon+\gamma} - (\|z-\bar{z}\|+\|x-\bar{x}\|) & \text{if } (z,x)\notin \operatorname{gph} F.\end{cases}\end{aligned}$$

Since $0<\eta<\gamma$ is arbitrary, passing to the limit, one has

$$\liminf_{(z,x)\to(\bar{z},\bar{x})} \frac{\rho(z,x)-\rho(\bar{z},\bar{x}) - \left\langle \frac{1}{\|x^*\|+\varepsilon+\gamma}(z^*,x^*), (z,x)-(\bar{z},\bar{x})\right\rangle}{\|(z,x)-(\bar{z},\bar{x})\|} \ge -\frac{\varepsilon}{\|x^*\|+\varepsilon+\gamma}.$$

By definition, $\frac{1}{\|x^*\|+\varepsilon+\gamma}(z^*, x^*) \in \widehat{\partial}_{\frac{\varepsilon}{\|x^*\|+\varepsilon+\gamma}}\rho(\bar{z}, \bar{x})$ and hence

$$(z^*, x^*) \in (\|x^*\| + \varepsilon + \gamma)\widehat{\partial}_{\frac{\varepsilon}{\|x^*\|+\varepsilon+\gamma}}\rho(\bar{z}, \bar{x}). \qquad \square$$

Theorem 2.3.12. ([26, Theorem 4.7]) *Let $F : Z \rightrightarrows X$ be a closed-graph mapping and $(\bar{z}, \bar{x}) \in \operatorname{gph} F$. Suppose ρ is upper semicontinuous at $(\bar{z}, \bar{x})$. Then*

$$\partial^\infty \rho(\bar{z}, \bar{x}) = \{(z^*, 0) \in Z^* \times X^* : z^* \in D_M^* F(\bar{z}, \bar{x})(0)\}.$$

Proof. Let $(z^*, x^*) \in \partial^\infty \rho(\bar{z}, \bar{x})$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^\infty \subset \mathbb{R}_+$, $\{\lambda_k\}_{k=1}^\infty \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^\infty \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^\infty \subset Z^* \times X^*$ such that

$$\varepsilon_k \downarrow 0, \quad \lambda_k \downarrow 0, \tag{2.44}$$

$$(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x}), \tag{2.45}$$

$$\lambda_k(z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*), \text{ and} \tag{2.46}$$

$$(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho(z_k, x_k) \text{ for all } k \in \mathbb{N}. \tag{2.47}$$

Suppose there are infinitely many (z_k, x_k) such that $\rho(z_k, x_k) = 0$. By passing to this subsequence of $\{(z_k, x_k)\}_{k=1}^\infty$ together with the corresponding subsequences of $\{\varepsilon_k\}_{k=1}^\infty$, $\{\lambda_k\}_{k=1}^\infty$ and $\{(z_k^*, x_k^*)\}_{k=1}^\infty$ if necessary, assume that for all $k \in \mathbb{N}$, $d(x_k, F(z_k)) = \rho(z_k, x_k) = 0$, which implies $x_k \in F(z_k)$ and $(z_k, x_k) \in \operatorname{gph} F$, since $\operatorname{gph} F$ is closed. Note also that $F_0 \equiv F$. For each $k \in \mathbb{N}$, using (2.47) and Proposition 2.2.1(a),

$$(z_k^*, x_k^*) \in \widehat{N}_{\varepsilon_k}((z_k, x_k); \operatorname{gph} F) \text{ and } \|x_k^*\| \leq 1 + \varepsilon_k. \tag{2.48}$$

Applying Proposition 1.6.6 to the first relation of (2.48), one sees that for all $k \in \mathbb{N}$,

$$\lambda_k(z_k^*, x_k^*) \in \widehat{N}_{\lambda_k\varepsilon_k}((z_k, x_k); \operatorname{gph} F). \tag{2.49}$$

In light of (2.44) and (2.45), $\lambda_k\varepsilon_k \downarrow 0$ and $(z_k, x_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{x})$.

Otherwise, by considering the tail of $\{(z_k, x_k)\}_{k=1}^\infty$ together with the corresponding terms of $\{\varepsilon_k\}_{k=1}^\infty$, $\{\lambda_k\}_{k=1}^\infty$ and $\{(z_k^*, x_k^*)\}_{k=1}^\infty$ if necessary, assume that for all $k \in \mathbb{N}$, $d(x_k, F(z_k)) = \rho(z_k, x_k) > 0$, which implies $x_k \notin F(z_k)$ and $(z_k, x_k) \notin \operatorname{gph} F$. Let $\eta_k = \rho(z_k, x_k) > 0$ for all $k \in \mathbb{N}$. Since $(\bar{z}, \bar{x}) \in \operatorname{gph} F$, $\bar{x} \in F(\bar{z})$ and $\rho(\bar{z}, \bar{x}) = d(\bar{x}, F(\bar{z})) = 0$.

In view of (2.45), $\eta_k \downarrow 0$. For each $k \in \mathbb{N}$, by virtue of (2.47) and Theorem 2.2.11, there exists $(v_k, u_k) \in \Theta^F_{\eta_k}(z_k, x_k)$ such that

$$(z_k^*, x_k^*) \in \widehat{N}_{\varepsilon_k+\eta_k}((v_k, u_k); \operatorname{gph} F) \text{ and } 1 - \varepsilon_k \le \|x_k^*\| \le 1 + \varepsilon_k. \tag{2.50}$$

Then $(v_k, u_k) \in \operatorname{gph} F$, $\|z_k - v_k\| \le \eta_k$ and $\|x_k - u_k\| \le \rho(z_k, x_k) + \eta_k = 2\eta_k$. Applying Proposition 1.6.6 to the first relation of (2.50), one sees that for all $k \in \mathbb{N}$,

$$\lambda_k(z_k^*, x_k^*) \in \widehat{N}_{\lambda_k(\varepsilon_k+\eta_k)}((v_k, u_k); \operatorname{gph} F). \tag{2.51}$$

Note that $\|v_k - \bar{z}\| \le \|v_k - z_k\| + \|z_k - \bar{z}\|$ and $\|u_k - \bar{x}\| \le \|u_k - x_k\| + \|x_k - \bar{x}\|$. It follows from (2.44) and (2.45) that $\lambda_k(\varepsilon_k + \eta_k) \downarrow 0$ and $(v_k, u_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{x})$.

Owing to the second relation of both (2.48) and (2.50) and $\varepsilon_k \downarrow 0$, $\{\|x_k^*\|\}_{k=1}^\infty$ is a bounded sequence in $\mathbb{R}$. With $\lambda_k \downarrow 0$, one has $\|\lambda_k x_k^*\| = \lambda_k \|x_k^*\| \to 0$, hence $\lambda_k x_k^* \to 0$ and in turn $-\lambda_k x_k^* \to 0$. Thus in both cases, $z^* \in D_M^* F(\bar{z}, \bar{x})(0)$.

Using the lower semicontinuity of $\|\cdot\|$ with respect to the weak* topology of X^* yields

$$\|x^*\| \le \liminf_{k\to\infty} \|\lambda_k x_k^*\| = 0.$$

Therefore $\|x^*\| = 0$ and $x^* = 0$. The inclusion

$$\partial^\infty \rho(\bar{z}, \bar{x}) \subset \{(z^*, 0) \in Z^* \times X^* : z^* \in D_M^* F(\bar{z}, \bar{x})(0)\}$$

is established.

Consider the reverse inclusion. Let $z^* \in D_M^* F(\bar{z}, \bar{x})(0)$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^\infty \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^\infty \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^\infty \subset Z^* \times X^*$ such that

$$\varepsilon_k \downarrow 0, \tag{2.52}$$

$$(z_k, x_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{x}), \tag{2.53}$$

$$z_k^* \xrightarrow{w^*} z^*, \qquad x_k^* \to 0, \text{ and} \tag{2.54}$$

$$(z_k^*, x_k^*) \in \widehat{N}_{\varepsilon_k}((z_k, x_k); \operatorname{gph} F) \text{ for all } k \in \mathbb{N}. \tag{2.55}$$

In light of (2.53), for all $k \in \mathbb{N}$, $(z_k, x_k) \in \operatorname{gph} F$, which implies $x_k \in F(z_k)$ and thus $\rho(z_k, x_k) = d(x_k, F(z_k)) = 0$. It follows that $(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x})$. Observing that $\rho \geq 0$ and $\rho(\bar{z}, \bar{x}) = 0$, the upper semicontinuity of ρ at $(\bar{z}, \bar{x})$ implies the existence of $\delta > 0$ such that ρ is upper semicontinuous on $\boldsymbol{B}_{Z\times X}((\bar{z}, \bar{x}), \delta)$. Due to (2.53) again, by considering the tail of $\{(z_k, x_k)\}_{k=1}^{\infty}$ together with the corresponding terms of $\{\varepsilon_k\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, assume that for all $k \in \mathbb{N}$, $(z_k, x_k) \in \boldsymbol{B}_{Z\times X}((\bar{z}, \bar{x}), \delta)$ and hence ρ is upper semicontinuous at (z_k, x_k). For each $k \in \mathbb{N}$, employing (2.55) and Lemma 2.3.11 yields

$$(z_k^*, x_k^*) \in (\|x_k^*\| + \varepsilon_k + \sqrt{\varepsilon_k})\, \widehat{\partial}_{\frac{\varepsilon_k}{\|x_k^*\|+\varepsilon_k+\sqrt{\varepsilon_k}}} \rho(z_k, x_k). \tag{2.56}$$

For all $k \in \mathbb{N}$, let $\lambda_k = \|x_k^*\|+\varepsilon_k+\sqrt{\varepsilon_k}$, $\widehat{\varepsilon}_k = \frac{\varepsilon_k}{\|x_k^*\|+\varepsilon_k+\sqrt{\varepsilon_k}}$, and $(\widehat{z}_k^*, \widehat{x}_k^*) = \frac{1}{\lambda_k}(z_k^*, x_k^*)$. It follows from $x_k^* \to 0$ that $\|x_k^*\| \to 0$. In view of (2.52), (2.54) and (2.56),

$$\begin{aligned}
&\lambda_k = \|x_k^*\| + \varepsilon_k + \sqrt{\varepsilon_k} \downarrow 0, \quad \lambda_k(\widehat{z}_k^*, \widehat{x}_k^*) = (z_k^*, x_k^*) \xrightarrow{w^*} (z^*, 0),\\
&\widehat{\varepsilon}_k = \frac{\varepsilon_k}{\|x_k^*\| + \varepsilon_k + \sqrt{\varepsilon_k}} \leq \frac{\varepsilon_k}{\varepsilon_k + \sqrt{\varepsilon_k}} = \frac{\sqrt{\varepsilon_k}}{\sqrt{\varepsilon_k}+1} \downarrow 0, \text{ and}\\
&(\widehat{z}_k^*, \widehat{x}_k^*) = \frac{1}{\lambda_k}(z_k^*, x_k^*) = \frac{(z_k^*, x_k^*)}{\|x_k^*\| + \varepsilon_k + \sqrt{\varepsilon_k}} \in \widehat{\partial}_{\widehat{\varepsilon}_k}\rho(z_k, x_k) \text{ for all } k \in \mathbb{N}.
\end{aligned}$$

By definition, $(z^*, 0) \in \partial^{\infty}\rho(\bar{z}, \bar{x})$. This establishes

$$\partial^{\infty}\rho(\bar{z}, \bar{x}) \supset \{(z^*, 0) \in Z^* \times X^* : z^* \in D_M^*F(\bar{z}, \bar{x})(0)\}$$

and completes the proof of the equality. □

At points *not belonging to* $\operatorname{gph} F$, instead of an equality, there is only an analogous *inclusion* involving *enlargements*. Moreover, right-sided singular subdifferentials are used in place of singular subdifferentials.

Theorem 2.3.13. (cf. [26, Theorem 4.8]) *Let $F : Z \rightrightarrows X$ be a set-valued mapping, $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$ and $r = \rho(\bar{z}, \bar{x})$. Suppose* $\operatorname{gph} F$ *is locally closed at $(\bar{z}, \bar{x})$ and* $\operatorname{gph} F_r$ *is closed. Then*

$$\partial_{\geq}^{\infty}\rho(\bar{z}, \bar{x}) \subset \{(z^*, 0) \in Z^* \times X^* : z^* \in D_M^*F_r(\bar{z}, \bar{x})(0)\}.$$

Proof. Let $(z^*, x^*) \in \partial^{\infty}_{\geq}\rho(\bar{z}, \bar{x})$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{\lambda_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^{\infty} \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty} \subset Z^* \times X^*$ such that

$$\varepsilon_k \downarrow 0, \quad \lambda_k \downarrow 0, \tag{2.57}$$

$$(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x}), \tag{2.58}$$

$$\lambda_k(z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*), \tag{2.59}$$

$$\rho(z_k, x_k) \geq \rho(\bar{z}, \bar{x}) \text{ for all } k \in \mathbb{N}, \text{ and} \tag{2.60}$$

$$(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho(z_k, x_k) \text{ for all } k \in \mathbb{N}. \tag{2.61}$$

In view of the assumptions and using Lemma 2.3.7, there exist two sequences $\{\gamma_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$ and $\{(v_k, u_k)\}_{k=1}^{\infty} \subset Z \times X$ which, by passing to appropriate subsequences of $\{\lambda_k\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, may be assumed to satisfy the aforementioned conditions $\lambda_k \downarrow 0$ and $\lambda_k(z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*)$, as well as the additional conditions $\gamma_k \downarrow 0$, $(v_k, u_k) \xrightarrow{\operatorname{gph} F_r} (\bar{z}, \bar{x})$,

$$(z_k^*, x_k^*) \in \widehat{N}_{\gamma_k}((v_k, u_k); \operatorname{gph} F_r) \text{ and } 1 - \gamma_k \leq \|x_k^*\| \leq 1 + \gamma_k \tag{2.62}$$

for all $k \in \mathbb{N}$. In light of the first relation of (2.62) and Proposition 1.6.6, one sees that for each $k \in \mathbb{N}$,

$$\lambda_k(z_k^*, x_k^*) \in \widehat{N}_{\lambda_k\gamma_k}((v_k, u_k); \operatorname{gph} F_r).$$

Note that, due to (2.57), $\lambda_k\gamma_k \downarrow 0$. Moreover, the second relation of (2.62) and $\gamma_k \downarrow 0$ reveal that $\{\|x_k^*\|\}_{k=1}^{\infty}$ is a bounded sequence in $\mathbb{R}$. With $\lambda_k \downarrow 0$, one has $\|\lambda_k x_k^*\| = \lambda_k\|x_k^*\| \to 0$, hence $\lambda_k x_k^* \to 0$ and in turn $-\lambda_k x_k^* \to 0$. Thus $z^* \in D_M^*F_r(\bar{z}, \bar{x})(0)$.

Using the lower semicontinuity of $\|\cdot\|$ with respect to the weak* topology of X^* yields

$$\|x^*\| \leq \liminf_{k\to\infty} \|\lambda_k x_k^*\| = 0.$$

Therefore $\|x^*\| = 0$ and $x^* = 0$. This justifies

$$\partial^{\infty}_{\geq}\rho(\bar{z}, \bar{x}) \subset \{(z^*, 0) \in Z^* \times X^* : z^* \in D_M^*F_r(\bar{z}, \bar{x})(0)\}.$$

□

The rest of this section collects a few upper estimates of limiting and singular subdifferentials of the generalized distance function via projections. At this point, it is necessary to state certain *criteria for well-posedness of the best approximation problem.*

Definition 2.3.14. **(a)** Let $F : Z \rightrightarrows X$ be a set-valued mapping and $(\bar{z}, \bar{x}) \in Z \times X$. ***The first criterion for well-posedness of the best approximation problem from*** $(\bar{z}, \bar{x})$ ***to*** $\operatorname{gph} F$ ***via the generalized distance function*** (or simply the first criterion for well-posedness via the generalized distance function) is that, for any sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$ and $\{(z_k, x_k)\}_{k=1}^{\infty} \subset Z \times X$ with $\varepsilon_k \downarrow 0$, $(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x})$ and $\widehat{\partial}_{\varepsilon_k}\rho(z_k, x_k) \neq \varnothing$ for all $k \in \mathbb{N}$, there exists a sequence $\{y_k\}_{k=1}^{\infty} \subset X$ with $y_k \in \Pi(x_k, F(z_k))$ for all $k \in \mathbb{N}$ which has a convergent subsequence.

(b) Let $\Omega \subset X$ and $\bar{x} \in X$. ***The first criterion for well-posedness of the best approximation problem from*** $\bar{x}$ ***to*** Ω ***via the standard distance function*** (or simply the first criterion for well-posedness via the standard distance function) is that, for any sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$ and $\{x_k\}_{k=1}^{\infty} \subset X$ with $\varepsilon_k \downarrow 0$, $x_k \to \bar{x}$ and $\widehat{\partial}_{\varepsilon_k} d(x_k, \Omega) \neq \varnothing$ for all $k \in \mathbb{N}$, there exists a sequence $\{y_k\}_{k=1}^{\infty} \subset X$ with $y_k \in \Pi(x_k, \Omega)$ for all $k \in \mathbb{N}$ which has a convergent subsequence.

Remarks 2.3.15. (i) If X and Z are *Asplund spaces* and ρ is *lower semicontinuous at* $(\bar{z}, \bar{x})$, the first criterion for well-posedness via the generalized distance function can be simplified by taking $\varepsilon_k = 0$ for all $k \in \mathbb{N}$.

(ii) If X is an *Asplund space* and Ω is *locally closed at* $\bar{x}$, the first criterion for well-posedness via the standard distance function can be simplified by taking $\varepsilon_k = 0$ for all $k \in \mathbb{N}$.

Definition 2.3.16. **(a)** Let $F : Z \rightrightarrows X$ be a closed-graph mapping, $(\bar{z}, \bar{x}) \in Z \times X$ and ρ be lower semicontinuous at $(\bar{z}, \bar{x})$. ***The second criterion for well-posedness of the best approximation problem from*** $(\bar{z}, \bar{x})$ ***to*** $\operatorname{gph} F$ ***via***

the generalized distance function (or simply the second criterion for well-posedness via the generalized distance function) is that, for any sequences $\{(v_k, x_k)\}_{k=1}^{\infty} \subset Z \times X$ and $\{u_k\}_{k=1}^{\infty} \subset X$ such that $\|u_k - x_k\| - \rho(v_k, x_k) \to 0$, $(v_k, x_k) \to (\bar{z}, \bar{x})$ and $u_k \in F(v_k)$ for all $k \in \mathbb{N}$, $\{u_k\}_{k=1}^{\infty}$ has a convergent subsequence.

(b) Let $\Omega \subset X$ be closed and $\bar{x} \in X$. ***The second criterion for well-posedness of the best approximation problem from $\bar{x}$ to Ω via the standard distance function*** (or simply the second criterion for well-posedness via the standard distance function) is that, for any sequences $\{x_k\}_{k=1}^{\infty} \subset X$ and $\{u_k\}_{k=1}^{\infty} \subset \Omega$ such that $\|u_k - \bar{x}\| \to d(\bar{x}, \Omega)$ and $x_k \to \bar{x}$, $\{u_k\}_{k=1}^{\infty}$ has a convergent subsequence.

Remarks 2.3.17. (i) The main difference between the two criteria for well-posedness via the generalized distance function is that, instead of imposing sequential compactness on the *projection* sequence $\{y_k\}_{k=1}^{\infty}$ with $y_k \in \Pi(x_k, F(z_k))$ for all $k \in \mathbb{N}$ in the first criterion, sequential compactness is imposed on the *in-graph* sequence $\{u_k\}_{k=1}^{\infty}$ with $(v_k, u_k) \in \operatorname{gph} F$ for all $k \in \mathbb{N}$ in the second criterion.

(ii) The main difference between the two criteria for well-posedness via the standard distance function is that, instead of imposing sequential compactness on the *projection* sequence $\{y_k\}_{k=1}^{\infty}$ with $y_k \in \Pi(x_k, \Omega)$ for all $k \in \mathbb{N}$ in the first criterion, sequential compactness is imposed on the *in-set* sequence $\{u_k\}_{k=1}^{\infty}$ with $u_k \in \Omega$ for all $k \in \mathbb{N}$ in the second criterion.

Prior to stating the main theorems, it is beneficial to examine two simple consequences of the criteria for well-posedness via the generalized distance function.

Lemma 2.3.18. *Let $F : Z \rightrightarrows X$ be a closed-graph mapping and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$. Suppose $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^{\infty} \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty} \subset Z^* \times X^*$ are sequences satisfying $\varepsilon_k \downarrow 0$, $(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x})$, and $(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k} \rho(z_k, x_k)$ for all $k \in \mathbb{N}$. Assume further that the first criterion for well-posedness of the best approximation* *problem from* $(\bar{z}, \bar{x})$ *to* $\operatorname{gph} F$ *via the generalized distance function is satisfied. Then there exist* $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$*, a sequence* $\{\widetilde{y}_k\}_{k=1}^{\infty} \subset X$ *and corresponding subsequences* $\{\widetilde{\varepsilon}_k\}_{k=1}^{\infty}$ *of* $\{\varepsilon_k\}_{k=1}^{\infty}$*,* $\{(\widetilde{z}_k, \widetilde{x}_k)\}_{k=1}^{\infty}$ *of* $\{(z_k, x_k)\}_{k=1}^{\infty}$*,* $\{(\widetilde{z}_k^*, \widetilde{x}_k^*)\}_{k=1}^{\infty}$ *of* $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ *such that for all* $k \in \mathbb{N}$*,*

$$(\widetilde{z}_k, \widetilde{x}_k) \notin \operatorname{gph} F, \qquad \widetilde{y}_k \in \Pi(\widetilde{x}_k, F(\widetilde{z}_k)), \qquad 1 - \widetilde{\varepsilon}_k \leq \|\widetilde{x}_k^*\| \leq 1 + \widetilde{\varepsilon}_k,$$

$$(\widetilde{z}_k^*, \widetilde{x}_k^*) \in \widehat{\partial}_{\widetilde{\varepsilon}_k} \rho(\widetilde{z}_k, \widetilde{x}_k) \cap \widehat{N}_{\widetilde{\varepsilon}_k}((\widetilde{z}_k, \widetilde{y}_k); \operatorname{gph} F);$$

and $$(\widetilde{z}_k, \widetilde{y}_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{y}), \qquad \widetilde{\varepsilon}_k \downarrow 0, \qquad (\widetilde{z}_k, \widetilde{x}_k) \xrightarrow{\rho} (\bar{z}, \bar{x}).$$

Proof. Since the first criterion for well-posedness of the best approximation problem from $(\bar{z}, \bar{x})$ to $\operatorname{gph} F$ via the generalized distance function is satisfied, there exists a sequence $\{y_k\}_{k=1}^{\infty} \subset X$ with $y_k \in \Pi(x_k, F(z_k))$ for all $k \in \mathbb{N}$ which has a convergent subsequence $\{\widetilde{y}_k\}_{k=1}^{\infty}$ with $\widetilde{y}_k \to \bar{y}$ for some $\bar{y} \in X$. Due to the assumptions about the sequences, by passing to the corresponding subsequences $\{\widetilde{\varepsilon}_k\}_{k=1}^{\infty}$ of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{(\widetilde{z}_k, \widetilde{x}_k)\}_{k=1}^{\infty}$ of $\{(z_k, x_k)\}_{k=1}^{\infty}$ and $\{(\widetilde{z}_k^*, \widetilde{x}_k^*)\}_{k=1}^{\infty}$ of $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, one has

$$\widetilde{\varepsilon}_k \downarrow 0, \tag{2.63}$$

$$(\widetilde{z}_k, \widetilde{x}_k) \xrightarrow{\rho} (\bar{z}, \bar{x}), \tag{2.64}$$

$$\widetilde{y}_k \in \Pi(\widetilde{x}_k, F(\widetilde{z}_k)) \text{ for all } k \in \mathbb{N}, \text{ and} \tag{2.65}$$

$$(\widetilde{z}_k^*, \widetilde{x}_k^*) \in \widehat{\partial}_{\widetilde{\varepsilon}_k} \rho(\widetilde{z}_k, \widetilde{x}_k) \text{ for all } k \in \mathbb{N}. \tag{2.66}$$

For all $k \in \mathbb{N}$, it follows from (2.65) that $\widetilde{y}_k \in F(\widetilde{z}_k)$, which implies $(\widetilde{z}_k, \widetilde{y}_k) \in \operatorname{gph} F$, and

$$\|\widetilde{x}_k - \widetilde{y}_k\| = d(\widetilde{x}_k, F(\widetilde{z}_k)) = \rho(\widetilde{z}_k, \widetilde{x}_k). \tag{2.67}$$

Owing to (2.64), letting $k \to \infty$ in (2.67) yields $\|\bar{x} - \bar{y}\| = \rho(\bar{z}, \bar{x}) = d(\bar{x}, F(\bar{z}))$. Moreover, $(\widetilde{z}_k, \widetilde{y}_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{y})$. Since $\{(\widetilde{z}_k, \widetilde{y}_k)\}_{k=1}^{\infty}$ is a sequence in $\operatorname{gph} F$, which is closed, $(\bar{z}, \bar{y}) \in \operatorname{gph} F$ and hence $\bar{y} \in F(\bar{z})$. Thus $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$.

Since $\operatorname{gph} F$ is closed and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$, $\bar{x} \notin F(\bar{z})$ and $\rho(\bar{z}, \bar{x}) = d(\bar{x}, F(\bar{z})) > 0$. In view of (2.64), by considering the tail of $\{(\widetilde{z}_k, \widetilde{x}_k)\}_{k=1}^{\infty}$ together with the corresponding terms of $\{\widetilde{\varepsilon}_k\}_{k=1}^{\infty}$, $\{\widetilde{y}_k\}_{k=1}^{\infty}$ and $\{(\widetilde{z}_k^*, \widetilde{x}_k^*)\}_{k=1}^{\infty}$ if necessary, assume that for all $k \in \mathbb{N}$,

$d(\widetilde{x}_k, F(\widetilde{z}_k)) = \rho(\widetilde{z}_k, \widetilde{x}_k) > 0$, which implies $\widetilde{x}_k \notin F(\widetilde{z}_k)$ and $(\widetilde{z}_k, \widetilde{x}_k) \notin \operatorname{gph} F$. In light of (2.66) and Theorem 2.2.9, for each $k \in \mathbb{N}$,

$$(\widetilde{z}_k^*, \widetilde{x}_k^*) \in \widehat{N}_{\widetilde{\varepsilon}_k}((\widetilde{z}_k, \widetilde{y}_k); \operatorname{gph} F) \text{ and } 1 - \widetilde{\varepsilon}_k \leq \|\widetilde{x}_k^*\| \leq 1 + \widetilde{\varepsilon}_k.$$

The conclusion is ascertained. □

Lemma 2.3.19. *Let $F : Z \rightrightarrows X$ be a closed-graph mapping and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$. Suppose $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^{\infty} \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty} \subset Z^* \times X^*$ are sequences satisfying $\varepsilon_k \downarrow 0$, $(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x})$, and $(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k} \rho(z_k, x_k)$ for all $k \in \mathbb{N}$. Assume further that ρ is lower semicontinuous at $(\bar{z}, \bar{x})$ and the second criterion for well-posedness of the best approximation problem from $(\bar{z}, \bar{x})$ to $\operatorname{gph} F$ via the generalized distance function is satisfied. Then there exist $\bar{u} \in \Pi(\bar{x}, F(\bar{z}))$, a sequence $\{(\widetilde{v}_k, \widetilde{u}_k)\}_{k=1}^{\infty} \subset Z \times X$ and corresponding subsequences $\{\widetilde{\varepsilon}_k\}_{k=1}^{\infty}$ of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{(\widetilde{z}_k, \widetilde{x}_k)\}_{k=1}^{\infty}$ of $\{(z_k, x_k)\}_{k=1}^{\infty}$, $\{(\widetilde{z}_k^*, \widetilde{x}_k^*)\}_{k=1}^{\infty}$ of $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ such that for all $k \in \mathbb{N}$,*

$$(\widetilde{z}_k, \widetilde{x}_k) \notin \operatorname{gph} F, \qquad (\widetilde{v}_k, \widetilde{u}_k) \in \Theta_{\widetilde{\varepsilon}_k}^F(\widetilde{z}_k, \widetilde{x}_k), \qquad 1 - \widetilde{\varepsilon}_k \leq \|\widetilde{x}_k^*\| \leq 1 + \widetilde{\varepsilon}_k,$$

$$(\widetilde{z}_k^*, \widetilde{x}_k^*) \in \widehat{\partial}_{\widetilde{\varepsilon}_k} \rho(\widetilde{z}_k, \widetilde{x}_k) \cap \widehat{N}_{2\widetilde{\varepsilon}_k}((\widetilde{v}_k, \widetilde{u}_k); \operatorname{gph} F);$$

and $$(\widetilde{v}_k, \widetilde{u}_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{u}), \qquad \widetilde{\varepsilon}_k \downarrow 0, \qquad (\widetilde{z}_k, \widetilde{x}_k) \xrightarrow{\rho} (\bar{z}, \bar{x}).$$

Proof. Since $\operatorname{gph} F$ is closed and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$, $\bar{x} \notin F(\bar{z})$ and $\rho(\bar{z}, \bar{x}) = d(\bar{x}, F(\bar{z})) > 0$. In view of the condition $(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x})$, by considering the tail of $\{(z_k, x_k)\}_{k=1}^{\infty}$ together with the corresponding terms of $\{\varepsilon_k\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, assume that for all $k \in \mathbb{N}$, $d(x_k, F(z_k)) = \rho(z_k, x_k) > 0$, which implies $x_k \notin F(z_k)$ and $(z_k, x_k) \notin \operatorname{gph} F$. In light of the assumption that $(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k} \rho(z_k, x_k)$ for all $k \in \mathbb{N}$ and applying Theorem 2.2.11, for each $k \in \mathbb{N}$, there exists $(v_k, u_k) \in \Theta_{\varepsilon_k}^F(z_k, x_k)$ such that

$$(z_k^*, x_k^*) \in \widehat{N}_{2\varepsilon_k}((v_k, u_k); \operatorname{gph} F) \text{ and } 1 - \varepsilon_k \leq \|x_k^*\| \leq 1 + \varepsilon_k. \tag{2.68}$$

Then for all $k \in \mathbb{N}$, $(v_k, u_k) \in \operatorname{gph} F$, $\|z_k - v_k\| \leq \varepsilon_k$, and

$$\|x_k - u_k\| \leq \rho(z_k, x_k) + \varepsilon_k. \tag{2.69}$$

Note that $u_k \in F(v_k)$ implies $\|x_k - u_k\| \geq d(x_k, F(v_k)) = \rho(v_k, x_k)$. Moreover, there holds $\|v_k - \bar{z}\| \leq \|v_k - z_k\| + \|z_k - \bar{z}\|$. Thus the conditions $\varepsilon_k \downarrow 0$ and $(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x})$ guarantee the convergence relations $(v_k, x_k) \to (\bar{z}, \bar{x})$ and $2\varepsilon_k \downarrow 0$. Invoking the lower semicontinuity of ρ at $(\bar{z}, \bar{x})$ yields

$$\rho(\bar{z}, \bar{x}) \leq \liminf_{k\to\infty} \rho(v_k, x_k). \tag{2.70}$$

Owing to (2.69) and (2.70),

$$\begin{aligned}
0 &\leq \liminf_{k\to\infty}(\|x_k - u_k\| - \rho(v_k, x_k)) \leq \limsup_{k\to\infty}(\|x_k - u_k\| - \rho(v_k, x_k)) \\
&\leq \limsup_{k\to\infty}(\rho(z_k, x_k) + \varepsilon_k - \rho(v_k, x_k)) \\
&\leq \lim_{k\to\infty} \rho(z_k, x_k) + \lim_{k\to\infty} \varepsilon_k - \liminf_{k\to\infty} \rho(v_k, x_k) \\
&= \rho(\bar{z}, \bar{x}) - \liminf_{k\to\infty} \rho(v_k, x_k) \leq 0.
\end{aligned}$$

It follows that

$$\limsup_{k\to\infty}(\|x_k - u_k\| - \rho(v_k, x_k)) = \liminf_{k\to\infty}(\|x_k - u_k\| - \rho(v_k, x_k)) = 0,$$

which implies $\lim_{k\to\infty}(\|x_k - u_k\| - \rho(v_k, x_k)) = 0$. Since the second criterion for well-posedness of the best approximation problem from $(\bar{z}, \bar{x})$ to $\operatorname{gph} F$ via the generalized distance function is satisfied, $\{u_k\}_{k=1}^{\infty}$ has a convergent subsequence $\{\widetilde{u}_k\}_{k=1}^{\infty}$ with $\widetilde{u}_k \to \bar{u}$ for some $\bar{u} \in X$. Due to the assumptions about the sequences, by passing to the corresponding subsequences $\{\widetilde{\varepsilon}_k\}_{k=1}^{\infty}$ of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{\widetilde{v}_k\}_{k=1}^{\infty}$ of $\{v_k\}_{k=1}^{\infty}$, $\{(\widetilde{z}_k, \widetilde{x}_k)\}_{k=1}^{\infty}$ of $\{(z_k, x_k)\}_{k=1}^{\infty}$ and $\{(\widetilde{z}_k^*, \widetilde{x}_k^*)\}_{k=1}^{\infty}$ of $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, one has for all $k \in \mathbb{N}$,

$$\begin{aligned}
&(\widetilde{z}_k, \widetilde{x}_k) \notin \operatorname{gph} F, \qquad (\widetilde{v}_k, \widetilde{u}_k) \in \Theta^F_{\widetilde{\varepsilon}_k}(\widetilde{z}_k, \widetilde{x}_k), \qquad 1 - \widetilde{\varepsilon}_k \leq \|\widetilde{x}_k^*\| \leq 1 + \widetilde{\varepsilon}_k, \\
&(\widetilde{z}_k^*, \widetilde{x}_k^*) \in \widehat{\partial}_{\widetilde{\varepsilon}_k} \rho(\widetilde{z}_k, \widetilde{x}_k) \cap \widehat{N}_{2\widetilde{\varepsilon}_k}((\widetilde{v}_k, \widetilde{u}_k); \operatorname{gph} F);
\end{aligned}$$

and $\quad \widetilde{\varepsilon}_k \downarrow 0, \qquad (\widetilde{z}_k, \widetilde{x}_k) \xrightarrow{\rho} (\bar{z}, \bar{x}).$

Observe that $\{(\widetilde{v}_k, \widetilde{u}_k)\}_{k=1}^{\infty}$ is a sequence in $\operatorname{gph} F$ and $(\widetilde{v}_k, \widetilde{u}_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{u})$. Since $\operatorname{gph} F$ is closed, $(\bar{z}, \bar{u}) \in \operatorname{gph} F$. Note that $\bar{u} \in F(\bar{z})$ gives $\|\bar{x} - \bar{u}\| \geq d(\bar{x}, F(\bar{z}))$. On the other hand, employing (2.69) again, one obtains

$$\|\bar{x} - \bar{u}\| = \lim_{k\to\infty} \|\widetilde{x}_k - \widetilde{u}_k\| \le \lim_{k\to\infty} (\rho(\widetilde{z}_k, \widetilde{x}_k) + \widetilde{\varepsilon}_k) = \rho(\bar{z}, \bar{x}) = d(\bar{x}, F(\bar{z})).$$

As a result, $\bar{u} \in \Pi(\bar{x}, F(\bar{z}))$. This establishes the conclusion. □

The criteria for well-posedness serve as the principal assumptions in the next two theorems and their corollaries.

Theorem 2.3.20. ([26, Theorem 4.9]) *Let $F : Z \rightrightarrows X$ be a closed-graph mapping and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$. Suppose the first criterion for well-posedness of the best approximation problem from $(\bar{z}, \bar{x})$ to $\operatorname{gph} F$ via the generalized distance function is satisfied. The following statements hold:*

(a) $\partial\rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{y}\in\Pi(\bar{x},F(\bar{z}))} \{(z^*, x^*) \in N((\bar{z}, \bar{y}); \operatorname{gph} F) : \|x^*\| \le 1\}.$

(b) *If $\operatorname{gph} F \subset Z \times X$ is sequentially normally compact with respect to X at any $(\bar{z}, \bar{y}) \in Z \times X$ with $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$, then*

$$\partial\rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{y}\in\Pi(\bar{x},F(\bar{z}))} \{(z^*, x^*) \in N((\bar{z}, \bar{y}); \operatorname{gph} F) : 0 < \|x^*\| \le 1\}.$$

(c) *If X is finite dimensional, then*

$$\partial\rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{y}\in\Pi(\bar{x},F(\bar{z}))} \{(z^*, x^*) \in N((\bar{z}, \bar{y}); \operatorname{gph} F) : \|x^*\| = 1\}.$$

(d) $\partial^\infty\rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{y}\in\Pi(\bar{x},F(\bar{z}))} \{(z^*, 0) \in Z^* \times X^* : z^* \in D^*_M F(\bar{z}, \bar{y})(0)\}.$

Proof. **(a)** Let $(z^*, x^*) \in \partial\rho(\bar{z}, \bar{x})$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^\infty \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^\infty \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^\infty \subset Z^* \times X^*$ such that

$$\varepsilon_k \downarrow 0, \tag{2.71}$$

$$(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x}), \tag{2.72}$$

$$(z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*), \text{ and} \tag{2.73}$$

$$(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho(z_k, x_k) \text{ for all } k \in \mathbb{N}. \tag{2.74}$$

In view of the assumptions and employing Lemma 2.3.18, there exist $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$ and a sequence $\{y_k\}_{k=1}^{\infty} \subset X$ which, by passing to appropriate subsequences of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{(z_k, x_k)\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, may be assumed to satisfy the aforementioned conditions (2.71), (2.72), (2.73) and (2.74), as well as the additional conditions $(z_k, y_k) \xrightarrow{\text{gph}\, F} (\bar{z}, \bar{y})$,

$$(z_k^*, x_k^*) \in \widehat{N}_{\varepsilon_k}((z_k, y_k); \text{gph}\, F) \text{ and } 1 - \varepsilon_k \leq \|x_k^*\| \leq 1 + \varepsilon_k \tag{2.75}$$

for all $k \in \mathbb{N}$. By definition, $(z^*, x^*) \in N((\bar{z}, \bar{y}); \text{gph}\, F)$, which is well-defined since $\bar{y} \in F(\bar{z})$ and hence $(\bar{z}, \bar{y}) \in \text{gph}\, F$.

Using the lower semicontinuity of $\|\cdot\|$ with respect to the weak* topology of X^* in connection with the second relation of (2.75) yields

$$\|x^*\| \leq \liminf_{k \to \infty} \|x_k^*\| \leq \liminf_{k \to \infty} (1 + \varepsilon_k) = 1.$$

This ascertains

$$\partial \rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{y} \in \Pi(\bar{x}, F(\bar{z}))} \{(z^*, x^*) \in N((\bar{z}, \bar{y}); \text{gph}\, F) : \|x^*\| \leq 1\}.$$

(b) Let $(z^*, x^*) \in \partial \rho(\bar{z}, \bar{x})$. By **(a)**, $(z^*, x^*) \in N((\bar{z}, \bar{y}); \text{gph}\, F)$ for some $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$ and $\|x^*\| \leq 1$. It suffices to show that $\|x^*\| > 0$. Suppose $\|x^*\| = 0$. Then $x^* = 0$ and (2.73) implies $x_k^* \xrightarrow{w^*} 0$. Since $\text{gph}\, F$ is sequentially normally compact with respect to X at $(\bar{z}, \bar{y})$, one has $\|x_k^*\| \to 0$. On the other hand, owing to the second relation of (2.75) and $\varepsilon_k \downarrow 0$, $\|x_k^*\| \to 1$, which is a contradiction. This shows that $\|x^*\| > 0$ and completes the proof.

(c) Let $(z^*, x^*) \in \partial \rho(\bar{z}, \bar{x})$. By **(a)**, $(z^*, x^*) \in \bigcup_{\bar{y} \in \Pi(\bar{x}, F(\bar{z}))} N((\bar{z}, \bar{y}); \text{gph}\, F)$. It suffices to show that $\|x^*\| = 1$. Since X is finite dimensional, $\|\cdot\|$ is continuous with respect to the weak* topology of X^*. Letting $k \to \infty$ in the second relation of (2.75) gives $\|x^*\| = \lim_{k \to \infty} \|x_k^*\| = 1$. The result follows.

(d) Let $(z^*, x^*) \in \partial^\infty \rho(\bar{z}, \bar{x})$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^\infty \subset \mathbb{R}_+$, $\{\lambda_k\}_{k=1}^\infty \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^\infty \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^\infty \subset Z^* \times X^*$ such that

$$\varepsilon_k \downarrow 0, \quad \lambda_k \downarrow 0, \tag{2.76}$$

$$(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x}), \tag{2.77}$$

$$\lambda_k(z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*), \text{ and} \tag{2.78}$$

$$(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k} \rho(z_k, x_k) \text{ for all } k \in \mathbb{N}. \tag{2.79}$$

As in the proof of **(a)**, in view of the assumptions and employing Lemma 2.3.18, there exist $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$ and a sequence $\{y_k\}_{k=1}^\infty \subset X$ which, by passing to appropriate subsequences of $\{\varepsilon_k\}_{k=1}^\infty$, $\{\lambda_k\}_{k=1}^\infty$, $\{(z_k, x_k)\}_{k=1}^\infty$ and $\{(z_k^*, x_k^*)\}_{k=1}^\infty$ if necessary, may be assumed to satisfy the aforementioned conditions (2.76), (2.77), (2.78) and (2.79), as well as the additional conditions $(z_k, y_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{y})$,

$$(z_k^*, x_k^*) \in \widehat{N}_{\varepsilon_k}((z_k, y_k); \operatorname{gph} F) \text{ and } 1 - \varepsilon_k \le \|x_k^*\| \le 1 + \varepsilon_k \tag{2.80}$$

for all $k \in \mathbb{N}$. Using the first relation of (2.80) and Proposition 1.6.6, one sees that for all $k \in \mathbb{N}$,

$$\lambda_k(z_k^*, x_k^*) \in \widehat{N}_{\lambda_k \varepsilon_k}((z_k, y_k); \operatorname{gph} F). \tag{2.81}$$

Note that, due to (2.76), $\lambda_k \varepsilon_k \downarrow 0$. Moreover, the second relation of (2.80) and $\varepsilon_k \downarrow 0$ reveal that $\{\|x_k^*\|\}_{k=1}^\infty$ is a bounded sequence in $\mathbb{R}$. With $\lambda_k \downarrow 0$, one has $\|\lambda_k x_k^*\| = \lambda_k \|x_k^*\| \to 0$, hence $\lambda_k x_k^* \to 0$ and in turn $-\lambda_k x_k^* \to 0$. Thus $z^* \in D_M^* F(\bar{z}, \bar{y})(0)$, which is well-defined since $\bar{y} \in F(\bar{z})$ and hence $(\bar{z}, \bar{y}) \in \operatorname{gph} F$.

Owing to the lower semicontinuity of $\|\cdot\|$ with respect to the weak* topology of X^*,

$$\|x^*\| \le \liminf_{k \to \infty} \|\lambda_k x_k^*\| = 0.$$

Therefore $\|x^*\| = 0$ and $x^* = 0$. This substantiates

$$\partial^\infty \rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{y} \in \Pi(\bar{x}, F(\bar{z}))} \{(z^*, 0) \in Z^* \times X^* : z^* \in D_M^* F(\bar{z}, \bar{y})(0)\}. \qquad \square$$

Theorem 2.3.21. (cf. [26, Remark 4.12]) *Let $F : Z \rightrightarrows X$ be a closed-graph mapping and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$. Suppose ρ is lower semicontinuous at $(\bar{z}, \bar{x})$ and the second criterion for well-posedness of the best approximation problem from $(\bar{z}, \bar{x})$ to $\operatorname{gph} F$ via the generalized distance function is satisfied. Then the conclusions of Theorem 2.3.20 hold.*

Proof. (a) Let $(z^*, x^*) \in \partial\rho(\bar{z}, \bar{x})$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^{\infty} \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty} \subset Z^* \times X^*$ such that

$$\varepsilon_k \downarrow 0, \tag{2.82}$$

$$(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x}), \tag{2.83}$$

$$(z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*), \text{ and} \tag{2.84}$$

$$(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho(z_k, x_k) \text{ for all } k \in \mathbb{N}. \tag{2.85}$$

In view of the assumptions and employing Lemma 2.3.19, there exist $\bar{u} \in \Pi(\bar{x}, F(\bar{z}))$ and a sequence $\{(v_k, u_k)\}_{k=1}^{\infty} \subset Z \times X$ which, by passing to appropriate subsequences of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{(z_k, x_k)\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, may be assumed to satisfy the aforementioned conditions (2.82), (2.83), (2.84) and (2.85), as well as the additional conditions $(v_k, u_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{u})$,

$$(z_k^*, x_k^*) \in \widehat{N}_{2\varepsilon_k}((v_k, u_k); \operatorname{gph} F) \text{ and } 1 - \varepsilon_k \leq \|x_k^*\| \leq 1 + \varepsilon_k \tag{2.86}$$

for all $k \in \mathbb{N}$. Note that $2\varepsilon_k \downarrow 0$ owing to (2.82). By definition, $(z^*, x^*) \in N((\bar{z}, \bar{u}); \operatorname{gph} F)$, which is well-defined since $\bar{u} \in F(\bar{z})$ and hence $(\bar{z}, \bar{u}) \in \operatorname{gph} F$.

Using the lower semicontinuity of $\|\cdot\|$ with respect to the weak* topology of X^* in connection with the second relation of (2.86) yields

$$\|x^*\| \leq \liminf_{k\to\infty} \|x_k^*\| \leq \liminf_{k\to\infty}(1 + \varepsilon_k) = 1.$$

This ascertains

$$\partial\rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{u}\in\Pi(\bar{x}, F(\bar{z}))} \{(z^*, x^*) \in N((\bar{z}, \bar{u}); \operatorname{gph} F) : \|x^*\| \leq 1\}.$$

(b) Let $(z^*, x^*) \in \partial\rho(\bar{z}, \bar{x})$. By **(a)**, $(z^*, x^*) \in N((\bar{z}, \bar{u}); \operatorname{gph} F)$ for some $\bar{u} \in \Pi(\bar{x}, F(\bar{z}))$ and $\|x^*\| \leq 1$. It suffices to show that $\|x^*\| > 0$. Suppose $\|x^*\| = 0$. Then $x^* = 0$ and (2.84) implies $x_k^* \xrightarrow{w^*} 0$. Since $\operatorname{gph} F$ is sequentially normally compact with respect to X at $(\bar{z}, \bar{u})$, one has $\|x_k^*\| \to 0$. On the other hand, owing to the second relation of (2.86) and $\varepsilon_k \downarrow 0$, $\|x_k^*\| \to 1$, which is a contradiction. This shows that $\|x^*\| > 0$ and completes the proof.

(c) Let $(z^*, x^*) \in \partial\rho(\bar{z}, \bar{x})$. By **(a)**, $(z^*, x^*) \in \bigcup_{\bar{u} \in \Pi(\bar{x}, F(\bar{z}))} N((\bar{z}, \bar{u}); \operatorname{gph} F)$. It suffices to show that $\|x^*\| = 1$. Since X is finite dimensional, $\|\cdot\|$ is continuous with respect to the weak* topology of X^*. Letting $k \to \infty$ in the second relation of (2.86) gives $\|x^*\| = \lim_{k\to\infty} \|x_k^*\| = 1$. The result follows.

(d) Let $(z^*, x^*) \in \partial^\infty\rho(\bar{z}, \bar{x})$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^\infty \subset \mathbb{R}_+$, $\{\lambda_k\}_{k=1}^\infty \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^\infty \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^\infty \subset Z^* \times X^*$ such that

$$\varepsilon_k \downarrow 0, \quad \lambda_k \downarrow 0, \tag{2.87}$$

$$(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x}), \tag{2.88}$$

$$\lambda_k(z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*), \text{ and} \tag{2.89}$$

$$(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho(z_k, x_k) \text{ for all } k \in \mathbb{N}. \tag{2.90}$$

As in the proof of **(a)**, in view of the assumptions and employing Lemma 2.3.19, there exist $\bar{u} \in \Pi(\bar{x}, F(\bar{z}))$ and a sequence $\{(v_k, u_k)\}_{k=1}^\infty \subset Z \times X$ which, by passing to appropriate subsequences of $\{\varepsilon_k\}_{k=1}^\infty$, $\{\lambda_k\}_{k=1}^\infty$, $\{(z_k, x_k)\}_{k=1}^\infty$ and $\{(z_k^*, x_k^*)\}_{k=1}^\infty$ if necessary, may be assumed to satisfy the aforementioned conditions (2.87), (2.88), (2.89) and (2.90), as well as the additional conditions $(v_k, u_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{u})$,

$$(z_k^*, x_k^*) \in \widehat{N}_{2\varepsilon_k}((v_k, u_k); \operatorname{gph} F) \text{ and } 1 - \varepsilon_k \leq \|x_k^*\| \leq 1 + \varepsilon_k \tag{2.91}$$

for all $k \in \mathbb{N}$. Using the first relation of (2.91) and Proposition 1.6.6, one sees that for all $k \in \mathbb{N}$,

$$\lambda_k(z_k^*, x_k^*) \in \widehat{N}_{2\lambda_k\varepsilon_k}((v_k, u_k); \operatorname{gph} F). \tag{2.92}$$

Note that, due to (2.87), $2\lambda_k\varepsilon_k \downarrow 0$. Moreover, the second relation of (2.91) and $\varepsilon_k \downarrow 0$ reveal that $\{\|x_k^*\|\}_{k=1}^{\infty}$ is a bounded sequence in $\mathbb{R}$. With $\lambda_k \downarrow 0$, one has $\|\lambda_k x_k^*\| = \lambda_k\|x_k^*\| \to 0$, hence $\lambda_k x_k^* \to 0$ and in turn $-\lambda_k x_k^* \to 0$. Thus $z^* \in D_M^* F(\bar{z}, \bar{u})(0)$, which is well-defined since $\bar{u} \in F(\bar{z})$ and hence $(\bar{z}, \bar{u}) \in \operatorname{gph} F$.

Owing to the lower semicontinuity of $\|\cdot\|$ with respect to the weak* topology of X^*,

$$\|x^*\| \leq \liminf_{k\to\infty} \|\lambda_k x_k^*\| = 0.$$

Therefore $\|x^*\| = 0$ and $x^* = 0$. This substantiates

$$\partial^{\infty}\rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{u}\in\Pi(\bar{x}, F(\bar{z}))} \{(z^*, 0) \in Z^* \times X^* : z^* \in D_M^* F(\bar{z}, \bar{u})(0)\}. \qquad \square$$

Corollary 2.3.22. *Let $\Omega \subset X$ be closed and $\bar{x} \notin \Omega$. Suppose the first or the second criterion for well-posedness of the best approximation problem from $\bar{x}$ to Ω via the standard distance function is satisfied. The following statements hold:*

(a) $\partial d(\bar{x}, \Omega) \subset \bigcup_{\bar{y}\in\Pi(\bar{x},\Omega)} N(\bar{y}; \Omega) \cap \boldsymbol{B}_{X^*}$.

(b) *If Ω is sequentially normally compact at any $\bar{y} \in \Pi(\bar{x}, \Omega)$, then*

$$\partial d(\bar{x}, \Omega) \subset \bigcup_{\bar{y}\in\Pi(\bar{x},\Omega)} N(\bar{y}; \Omega) \cap (\boldsymbol{B}_{X^*}\backslash\{0\}).$$

(c) *If X is finite dimensional, then*

$$\partial d(\bar{x}, \Omega) \subset \bigcup_{\bar{y}\in\Pi(\bar{x},\Omega)} N(\bar{y}; \Omega) \cap \boldsymbol{S}_{X^*}.$$

Theorem 2.3.20 and Theorem 2.3.21 demonstrate perceptibly the significance of the criteria for well-posedness. However, it is in general not easy to check whether these criteria are satisfied. The concluding result of this section exhibits a simple condition which guarantees the fulfillment of the first criterion for well-posedness.

Definition 2.3.23. A normed space X is said to have the ***Kadets-Klee property*** or the ***Radon-Riesz property*** or ***property (H)*** if for any sequence $\{x_k\}_{k=1}^{\infty} \subset X$ and any $\bar{x} \in X$ such that $x_k \xrightarrow{w} \bar{x}$ and $\|x_k\| \to \|\bar{x}\|$, one has $x_k \to \bar{x}$.

This is equivalent to saying that X *has the Kadets-Klee property if norm convergence and weak convergence agree on* $\boldsymbol{S}_X$. It is well-known that every *locally uniformly convex* space, and in particular every *reflexive* space, admits an equivalent Kadets-Klee norm.

The first criterion for well-posedness is fulfilled under mild assumptions in a space possessing the Kadets-Klee property.

Theorem 2.3.24. (cf. [26, Corollary 4.10]) *Let X be a reflexive Banach space having the Kadets-Klee property and $(\bar{z},\bar{x}) \in Z \times X$. Suppose $F : Z \rightrightarrows X$ is a closed-graph mapping with respect to the norm×weak topology of $Z \times X$. Then the first criterion for well-posedness of the best approximation problem from $(\bar{z},\bar{x})$ to* gph F *via the generalized distance function is satisfied.*

Proof. Let $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$ and $\{(z_k,x_k)\}_{k=1}^{\infty} \subset Z \times X$ be sequences such that $\varepsilon_k \downarrow 0$, $(z_k,x_k) \xrightarrow{\rho} (\bar{z},\bar{x})$ and $\widehat{\partial}_{\varepsilon_k}\rho(z_k,x_k) \neq \varnothing$ for all $k \in \mathbb{N}$. Since X is reflexive and F is closed-graph with respect to the norm×weak topology of $Z \times X$, for all $k \in \mathbb{N}$, $F(z_k)$ is weakly closed and hence $\Pi(x_k, F(z_k)) \neq \varnothing$ in view of Proposition 1.4.5. For each $k \in \mathbb{N}$, let $y_k \in \Pi(x_k, F(z_k))$. Then $y_k \in F(z_k)$, which implies $(z_k, y_k) \in \operatorname{gph} F$, and

$$\|x_k - y_k\| = d(x_k, F(z_k)) = \rho(z_k, x_k). \tag{2.93}$$

It follows from the assumptions $\rho(z_k,x_k) \to \rho(\bar{z},\bar{x})$ and $x_k \to \bar{x}$ that $\{y_k - x_k\}_{k=1}^{\infty}$ and $\{x_k\}_{k=1}^{\infty}$ are both bounded sequences in X and in turn $\{y_k\}_{k=1}^{\infty}$ is also a bounded sequence in X. Invoking Theorem 1.2.4, $\{y_k\}_{k=1}^{\infty}$ has a weakly convergent subsequence $\{y_{k_l}\}_{l=1}^{\infty}$ such that $y_{k_l} \xrightarrow{w} \bar{y}$ for some $\bar{y} \in X$. It suffices to prove that $y_{k_l} \to \bar{y}$.

By passing to the corresponding subsequences $\{\varepsilon_{k_l}\}_{l=1}^{\infty}$ of $\{\varepsilon_k\}_{k=1}^{\infty}$ and $\{(z_{k_l},x_{k_l})\}_{l=1}^{\infty}$ of $\{(z_k,x_k)\}_{k=1}^{\infty}$ if necessary, one sees that $\varepsilon_{k_l} \downarrow 0$, $(z_{k_l},x_{k_l}) \xrightarrow{\rho} (\bar{z},\bar{x})$, $\widehat{\partial}_{\varepsilon_{k_l}}\rho(z_{k_l},x_{k_l}) \neq \varnothing$ for all $l \in \mathbb{N}$. Note that $\{(z_{k_l},y_{k_l})\}_{l=1}^{\infty}$ is a sequence in gph F, which is closed with respect to the norm×weak topology of $Z \times X$, and $(z_{k_l},y_{k_l}) \to (\bar{z},\bar{y})$ with respect to the same topology. Consequently, $(\bar{z},\bar{y}) \in \operatorname{gph} F$ and hence $\bar{y} \in F(\bar{z})$. Thus

$$\|\bar{x} - \bar{y}\| \geq d(\bar{x}, F(\bar{z})) = \rho(\bar{z},\bar{x}). \tag{2.94}$$

On the other hand, since $x_{k_l} - y_{k_l} \xrightarrow{w} \bar{x} - \bar{y}$, using the lower semicontinuity of $\|\cdot\|$ with respect to the weak topology of X in connection with (2.93) reveals that

$$\|\bar{x} - \bar{y}\| \le \liminf_{l\to\infty} \|x_{k_l} - y_{k_l}\| = \liminf_{l\to\infty} \rho(z_{k_l}, x_{k_l}) = \rho(\bar{z}, \bar{x}). \tag{2.95}$$

Inequalities (2.94) and (2.95) together yield $\|\bar{x} - \bar{y}\| = \rho(\bar{z}, \bar{x})$.

Check that, by virtue of (2.93) again,

$$\lim_{l\to\infty} \|x_{k_l} - y_{k_l}\| = \lim_{l\to\infty} \rho(z_{k_l}, x_{k_l}) = \rho(\bar{z}, \bar{x}) = \|\bar{x} - \bar{y}\|.$$

In light of the Kadets-Klee property of X, one has $x_{k_l} - y_{k_l} \to \bar{x} - \bar{y}$. As a result,

$$\lim_{l\to\infty} y_{k_l} = \lim_{l\to\infty} (x_{k_l} - (x_{k_l} - y_{k_l})) = \lim_{l\to\infty} x_{k_l} - \lim_{l\to\infty} (x_{k_l} - y_{k_l}) = \bar{x} - (\bar{x} - \bar{y}) = \bar{y}.$$

By definition, the first criterion for well-posedness of the best approximation problem from $(\bar{z}, \bar{x})$ to $\operatorname{gph} F$ via the generalized distance function is fulfilled. □

Remark 2.3.25. The conditions $\varepsilon_k \downarrow 0$ and $\widehat{\partial}_{\varepsilon_k}\rho(z_k, x_k) \neq \varnothing$ for all $k \in \mathbb{N}$ are indeed not needed in the above proof.

Corollary 2.3.26. *Let X be a reflexive Banach space having the Kadets-Klee property and $\bar{x} \in X$. Suppose $\Omega \subset X$ is weakly closed. Then the first criterion for well-posedness of the best approximation problem from $\bar{x}$ to Ω via the standard distance function is satisfied.*

Chapter 3

The Generalized Distance Function - Estimates via Intermediate Points

This chapter continues to survey various subdifferentials of the generalized distance function. Having established some basic estimates, further results can be derived via *intermediate points* situated on line segments with endpoints being given points *not belonging to* $\operatorname{gph} F$ and their projections. The new approach adopted in this chapter, together with other mild assumptions, leads to more refined estimates. Many of these may be regarded as improved versions of the estimates via projections communicated in the previous chapter, since projections are merely special intermediate points.

The majority of the theorems presented in this chapter were again first ascertained by Mordukhovich and Nam in [26] and [27]. As in the previous chapter, ***all set-valued mappings*** $F : Z \rightrightarrows X$ ***in this chapter are presumed to enjoy the serviceable property*** $\operatorname{dom} F = Z \neq \varnothing$ ***and all subsets*** $\Omega \subset X$ ***are assumed to be nonempty.*** Moreover, most corollaries are results about the standard distance function which follow immediately from their counterparts about the generalized distance function by taking $Z = \{\bar{z}\}$ and $F \equiv \Omega$. The proofs of such corollaries are omitted.

3.1 Fréchet-Like and Limiting Subdifferentials of the Generalized Distance Function via Intermediate Points

This section collects a number of estimates of Fréchet-like and limiting subdifferentials of the generalized distance function via intermediate points.

The discussion commences with an elementary but essential lemma about intermediate points.

Lemma 3.1.1. (cf. [27, Lemma 3.1]) *Let $F : Z \rightrightarrows X$ be a set-valued mapping and $(\bar{z}, \bar{x}) \in Z \times X$ with $\Pi(\bar{x}, F(\bar{z})) \neq \varnothing$. For any $t \in [0,1]$ and $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$, the following statements hold:*

(a) $d(t\bar{y} + (1-t)\bar{x}, F(\bar{z})) = (1-t)\|\bar{y} - \bar{x}\| = \|(t\bar{y} + (1-t)\bar{x}) - \bar{y}\|$.

(b) $\bar{y} \in \Pi(t\bar{y} + (1-t)\bar{x}, F(\bar{z}))$.

Proof. Let $t \in [0,1]$ and $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$. By definition, $\|\bar{x} - \bar{y}\| = d(\bar{x}, F(\bar{z}))$ and $\bar{y} \in F(\bar{z})$. It follows that

$$\begin{aligned} d(t\bar{y} + (1-t)\bar{x}, F(\bar{z})) &= d(t(\bar{y} - \bar{x}) + \bar{x}, F(\bar{z})) \geq d(\bar{x}, F(\bar{z})) - t\|\bar{y} - \bar{x}\| \\ &= \|\bar{x} - \bar{y}\| - t\|\bar{y} - \bar{x}\| = (1-t)\|\bar{y} - \bar{x}\|. \end{aligned}$$

On the other hand, since $\bar{y} \in F(\bar{z})$,

$$d(t\bar{y} + (1-t)\bar{x}, F(\bar{z})) \leq \|(t\bar{y} + (1-t)\bar{x}) - \bar{y}\| = (1-t)\|\bar{y} - \bar{x}\|.$$

Combining the inequalities, $d(t\bar{y}+(1-t)\bar{x}, F(\bar{z})) = (1-t)\|\bar{y}-\bar{x}\| = \|(t\bar{y}+(1-t)\bar{x})-\bar{y}\|$. Therefore $\bar{y} \in \Pi(t\bar{y} + (1-t)\bar{x}, F(\bar{z}))$. The assertions are justified. □

Corollary 3.1.2. *Let $\Omega \subset X$ and $\bar{x} \in X$ with $\Pi(\bar{x}, \Omega) \neq \varnothing$. For any $t \in [0,1]$ and $\bar{y} \in \Pi(\bar{x}, \Omega)$, the following statements hold:*

(a) $d(t\bar{y} + (1-t)\bar{x}, \Omega) = (1-t)\|\bar{y} - \bar{x}\| = \|(t\bar{y} + (1-t)\bar{x}) - \bar{y}\|$.

(b) $\bar{y} \in \Pi(t\bar{y} + (1-t)\bar{x}, \Omega)$.

With intermediate points, it is possible to arrive at estimates of ε-subdifferentials of the generalized distance function without using sets of ε-normals.

Proposition 3.1.3. (cf. [27, Proposition 3.2]) *Let $F : Z \rightrightarrows X$ be a closed-graph mapping and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$ with $\Pi(\bar{x}, F(\bar{z})) \neq \varnothing$. Then for any $\varepsilon \geq 0$, $t \in [0,1]$ and $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$,*

$$\widehat{\partial}_\varepsilon \rho(\bar{z}, \bar{x}) \subset \{(z^*, x^*) \in \widehat{\partial}_\varepsilon \rho(\bar{z}, t\bar{y} + (1-t)\bar{x}) : 1 - \varepsilon \leq \|x^*\| \leq 1 + \varepsilon\}.$$

Proof. Let $\varepsilon \geq 0$, $t \in [0,1]$ and $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$. Then $\|\bar{x} - \bar{y}\| = d(\bar{x}, F(\bar{z})) = \rho(\bar{z}, \bar{x})$. Let $(z^*, x^*) \in \widehat{\partial}_\varepsilon \rho(\bar{z}, \bar{x})$ and $\eta > 0$. Employing Proposition 1.5.5, there exists $\delta > 0$ such that for all $(z, x) \in Z \times X$ with $\|z - \bar{z}\| \leq \delta$ and $\|x - \bar{x}\| \leq \delta$,

$$\langle z^*, z - \bar{z} \rangle + \langle x^*, x - \bar{x} \rangle \leq \rho(z, x) - \rho(\bar{z}, \bar{x}) + (\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{x}\|). \tag{3.1}$$

Define $\bar{v} = t\bar{y} + (1-t)\bar{x}$. Using Lemma 3.1.1(a), $\rho(\bar{z}, \bar{v}) = d(\bar{v}, F(\bar{z})) = (1-t)\|\bar{y} - \bar{x}\|$. Fix any $(z, x) \in Z \times X$ with $\|z - \bar{z}\| \leq \delta$ and $\|x - \bar{v}\| \leq \delta$. Note that $\|(x - \bar{v} + \bar{x}) - \bar{x}\| = \|x - \bar{v}\| \leq \delta$ and $\|\bar{v} - \bar{x}\| = t\|\bar{x} - \bar{y}\|$. It follows from (3.1) that

$$\begin{aligned}
&\langle z^*, z - \bar{z} \rangle + \langle x^*, x - \bar{v} \rangle \\
&= \langle z^*, z - \bar{z} \rangle + \langle x^*, (x - \bar{v} + \bar{x}) - \bar{x} \rangle \\
&\leq \rho(z, x - \bar{v} + \bar{x}) - \rho(\bar{z}, \bar{x}) + (\varepsilon + \eta)(\|z - \bar{z}\| + \|(x - \bar{v} + \bar{x}) - \bar{x}\|) \\
&\leq \rho(z, x) + \|\bar{v} - \bar{x}\| - \|\bar{x} - \bar{y}\| + (\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{v}\|) \\
&= \rho(z, x) + t\|\bar{x} - \bar{y}\| - \|\bar{x} - \bar{y}\| + (\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{v}\|) \\
&= \rho(z, x) - (1-t)\|\bar{x} - \bar{y}\| + (\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{v}\|) \\
&= \rho(z, x) - \rho(\bar{z}, \bar{v}) + (\varepsilon + \eta)(\|z - \bar{z}\| + \|x - \bar{v}\|).
\end{aligned}$$

In view of Proposition 1.5.5, $(z^*, x^*) \in \widehat{\partial}_\varepsilon \rho(\bar{z}, \bar{v})$. On the other hand, since F is closed-graph and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$, $\bar{x} \notin F(\bar{z})$ and $\rho(\bar{z}, \bar{x}) = d(\bar{x}, F(\bar{z})) > 0$. Applying Proposition 2.2.1(b), one has $1 - \varepsilon \leq \|x^*\| \leq 1 + \varepsilon$. The assertion holds. □

Corollary 3.1.4. *Let $\Omega \subset X$ be closed and $\bar{x} \notin \Omega$ with $\Pi(\bar{x}, \Omega) \neq \varnothing$. Then for any $\varepsilon \geq 0$, $t \in [0,1]$ and $\bar{y} \in \Pi(\bar{x}, \Omega)$,*

$$\widehat{\partial}_\varepsilon d(\bar{x}, \Omega) \subset \widehat{\partial}_\varepsilon d(t\bar{y} + (1-t)\bar{x}, \Omega) \cap [1-\varepsilon, 1+\varepsilon]\boldsymbol{S}_{X^*}.$$

Likewise, estimates of limiting subdifferentials of the generalized distance function not involving limiting normal cones can be derived via intermediate points. However, for limiting subdifferentials, it is necessary to further ensure that a criterion for well-posedness of the best approximation problem is fulfilled.

Theorem 3.1.5. (cf. [27, Theorem 3.7]) *Let $F : Z \rightrightarrows X$ be a closed-graph mapping and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$ with $\Pi(\bar{x}, F(\bar{z})) \neq \varnothing$. Suppose the first criterion for well-posedness of the best approximation problem from $(\bar{z}, \bar{x})$ to $\operatorname{gph} F$ via the generalized distance function is satisfied. Then for any $t \in [0,1]$,*

$$\partial\rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{y} \in \Pi(\bar{x}, F(\bar{z}))} \{(z^*, x^*) \in \partial\rho(\bar{z}, t\bar{y} + (1-t)\bar{x}) : \|x^*\| \leq 1\}.$$

Proof. Let $t \in [0,1]$ and $(z^*, x^*) \in \partial\rho(\bar{z}, \bar{x})$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^\infty \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^\infty \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^\infty \subset Z^* \times X^*$ such that

$$\varepsilon_k \downarrow 0, \tag{3.2}$$

$$(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x}), \tag{3.3}$$

$$(z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*), \text{ and} \tag{3.4}$$

$$(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho(z_k, x_k) \text{ for all } k \in \mathbb{N}. \tag{3.5}$$

In view of the assumptions and employing Lemma 2.3.18, there exist $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$ and a sequence $\{y_k\}_{k=1}^\infty \subset X$ which, by passing to appropriate subsequences of $\{\varepsilon_k\}_{k=1}^\infty$, $\{(z_k, x_k)\}_{k=1}^\infty$ and $\{(z_k^*, x_k^*)\}_{k=1}^\infty$ if necessary, may be assumed to satisfy the aforementioned conditions (3.2), (3.3), (3.4) and (3.5), as well as the additional conditions $(z_k, y_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{y})$, $(z_k, x_k) \notin \operatorname{gph} F$ and $y_k \in \Pi(x_k, F(z_k))$ for all $k \in \mathbb{N}$. Note that $\|\bar{y} - \bar{x}\| = d(\bar{x}, F(\bar{z})) = \rho(\bar{z}, \bar{x})$ and $\|y_k - x_k\| = d(x_k, F(z_k)) = \rho(z_k, x_k)$ for all

$k \in \mathbb{N}$. In light of (3.5) and Proposition 3.1.3, for each $k \in \mathbb{N}$, one has

$$(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k} \rho(z_k, ty_k + (1-t)x_k) \text{ and } 1 - \varepsilon_k \le \|x_k^*\| \le 1 + \varepsilon_k. \tag{3.6}$$

By virtue of Lemma 3.1.1(a), one sees that

$$\begin{aligned} \rho(z_k, ty_k + (1-t)x_k) &= d(ty_k + (1-t)x_k, F(z_k)) \\ &= (1-t)\|y_k - x_k\| = (1-t)\rho(z_k, x_k) \quad \text{and} \\ \rho(\bar{z}, t\bar{y} + (1-t)\bar{x}) &= d(t\bar{y} + (1-t)\bar{x}, F(\bar{z})) \\ &= (1-t)\|\bar{y} - \bar{x}\| = (1-t)\rho(\bar{z}, \bar{x}). \end{aligned}$$

Then the convergence relations $(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x})$ and $y_k \to \bar{y}$ together imply that $(z_k, ty_k + (1-t)x_k) \xrightarrow{\rho} (\bar{z}, t\bar{y} + (1-t)\bar{x})$. By definition, $(z^*, x^*) \in \partial\rho(\bar{z}, t\bar{y} + (1-t)\bar{x})$.

Using the lower semicontinuity of $\|\cdot\|$ with respect to the weak* topology of X^* in connection with the second relation of (3.6) yields

$$\|x^*\| \le \liminf_{k\to\infty} \|x_k^*\| \le \liminf_{k\to\infty} (1 + \varepsilon_k) = 1.$$

This establishes

$$\partial\rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{y} \in \Pi(\bar{x}, F(\bar{z}))} \{(z^*, x^*) \in \partial\rho(\bar{z}, t\bar{y} + (1-t)\bar{x}) : \|x^*\| \le 1\}. \qquad \square$$

Corollary 3.1.6. *Let $\Omega \subset X$ be closed and $\bar{x} \notin \Omega$ with $\Pi(\bar{x}, \Omega) \neq \varnothing$. Suppose the first criterion for well-posedness of the best approximation problem from $\bar{x}$ to Ω via the standard distance function is satisfied. Then for any $t \in [0,1]$,*

$$\partial d(\bar{x}, \Omega) \subset \bigcup_{\bar{y} \in \Pi(\bar{x}, \Omega)} \partial d(t\bar{y} + (1-t)\bar{x}, \Omega) \cap \boldsymbol{B}_{X^*}.$$

In addition to intermediate points, the next couple of theorems also utilize *enlargements.* The proofs of these theorems are largely similar and rely on a crucial argument which essentially forms the skeletons of the proofs. In order to avoid reproducing the tedious reasoning, it is desirable to establish this critical part independently.

Proposition 3.1.7. *Let $F : Z \rightrightarrows X$ be a closed-graph mapping, $(\bar{z},\bar{x}) \notin \operatorname{gph} F$ with $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$ and $t \in (0,1]$. Suppose $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{y_k\}_{k=1}^{\infty} \subset X$, $\{(z_k,x_k)\}_{k=1}^{\infty} \subset Z \times X$ and $\{(z_k^*,x_k^*)\}_{k=1}^{\infty} \subset Z^* \times X^*$ are sequences satisfying $\varepsilon_k \downarrow 0$, $(z_k,x_k) \xrightarrow{\rho} (\bar{z},\bar{x})$, $(z_k,y_k) \xrightarrow{\operatorname{gph} F} (\bar{z},\bar{y})$, $y_k \in \Pi(x_k, F(z_k))$ and $(z_k^*,x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho(z_k,x_k)$ for all $k \in \mathbb{N}$. Assume further that $\operatorname{gph} F_{t_{\bar{y}}}$ is closed, where $t_{\bar{y}} = \rho(\bar{z}, t\bar{y} + (1-t)\bar{x})$. Then there exist a sequence $\{\widetilde{v}_k\}_{k=1}^{\infty} \subset X$ and corresponding subsequences $\{\widetilde{\varepsilon}_k\}_{k=1}^{\infty}$ of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{\widetilde{y}_k\}_{k=1}^{\infty}$ of $\{y_k\}_{k=1}^{\infty}$, $\{(\widetilde{z}_k,\widetilde{x}_k)\}_{k=1}^{\infty}$ of $\{(z_k,x_k)\}_{k=1}^{\infty}$, $\{(\widetilde{z}_k^*,\widetilde{x}_k^*)\}_{k=1}^{\infty}$ of $\{(z_k^*,x_k^*)\}_{k=1}^{\infty}$ such that for all $k \in \mathbb{N}$,*

$$\widetilde{y}_k \in \Pi(\widetilde{x}_k, F(\widetilde{z}_k)), \qquad 1-\widetilde{\varepsilon}_k \le \|\widetilde{x}_k^*\| \le 1+\widetilde{\varepsilon}_k,$$

$$(\widetilde{z}_k^*,\widetilde{x}_k^*) \in \widehat{\partial}_{\widetilde{\varepsilon}_k}\rho(\widetilde{z}_k,\widetilde{x}_k) \cap \widehat{\partial}_{\widetilde{\varepsilon}_k}\rho_{t_{\bar{y}}}(\widetilde{z}_k,\widetilde{v}_k) \cap \widehat{N}_{\widetilde{\varepsilon}_k}((\widetilde{z}_k,\widetilde{v}_k); \operatorname{gph} F_{t_{\bar{y}}});$$

and

$$\widetilde{\varepsilon}_k \downarrow 0, \quad (\widetilde{z}_k,\widetilde{x}_k) \xrightarrow{\rho} (\bar{z},\bar{x}), \quad (\widetilde{z}_k,\widetilde{y}_k) \xrightarrow{\operatorname{gph} F} (\bar{z},\bar{y}),$$

$$(\widetilde{z}_k,\widetilde{v}_k) \xrightarrow{\operatorname{gph} F_{t_{\bar{y}}}} (\bar{z}, t\bar{y}+(1-t)\bar{x}).$$

Proof. Since $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$, $\|\bar{y}-\bar{x}\| = d(\bar{x}, F(\bar{z})) = \rho(\bar{z},\bar{x})$. Let $\bar{v} = t\bar{y} + (1-t)\bar{x}$. Using Lemma 3.1.1(a),

$$t_{\bar{y}} = \rho(\bar{z},\bar{v}) = d(\bar{v}, F(\bar{z})) = (1-t)\|\bar{y}-\bar{x}\| = (1-t)\rho(\bar{z},\bar{x}).$$

Since $\operatorname{gph} F$ is closed and $(\bar{z},\bar{x}) \notin \operatorname{gph} F$, $\bar{x} \notin F(\bar{z})$ and $\rho(\bar{z},\bar{x}) = d(\bar{x}, F(\bar{z})) > 0$. Then $t \in (0,1]$ guarantees that $d(\bar{x}, F(\bar{z})) = \rho(\bar{z},\bar{x}) > (1-t)\rho(\bar{z},\bar{x}) = t_{\bar{y}}$. Thus $\bar{x} \notin F_{t_{\bar{y}}}(\bar{z})$ and $(\bar{z},\bar{x}) \notin \operatorname{gph} F_{t_{\bar{y}}}$. In view of $(z_k,x_k) \xrightarrow{\rho} (\bar{z},\bar{x})$, by considering the tail of $\{(z_k,x_k)\}_{k=1}^{\infty}$ together with the corresponding terms of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{y_k\}_{k=1}^{\infty}$ and $\{(z_k^*,x_k^*)\}_{k=1}^{\infty}$ if necessary, assume that for all $k \in \mathbb{N}$, $d(x_k, F(z_k)) = \rho(z_k,x_k) > t_{\bar{y}}$, which implies $x_k \notin F_{t_{\bar{y}}}(z_k)$, hence $(z_k,x_k) \notin \operatorname{gph} F_{t_{\bar{y}}}$, and in turn

$$\rho(z_k,x_k) = \rho_{t_{\bar{y}}}(z_k,x_k) + t_{\bar{y}}$$

by virtue of Proposition 2.1.5.

For each $k \in \mathbb{N}$, define $\varphi_k : [0,1] \to \mathbb{R}$ by

$$\varphi_k(\lambda) = d(\lambda y_k + (1-\lambda)x_k, F(z_k)) - t_{\bar{y}},$$

which is obviously continuous. Moreover, $y_k \in F(z_k)$ since $y_k \in \Pi(x_k, F(z_k))$. Check that

$$\varphi_k(1) = d(y_k, F(z_k)) - t_{\bar{y}} = -t_{\bar{y}} \leq 0, \text{ and}$$
$$\varphi_k(0) = d(x_k, F(z_k)) - t_{\bar{y}} = \rho(z_k, x_k) - t_{\bar{y}} > 0.$$

Then the intermediate value theorem guarantees the existence of $\lambda_k \in (0, 1]$ such that

$$d(\lambda_k y_k + (1 - \lambda_k)x_k, F(z_k)) - t_{\bar{y}} = \varphi_k(\lambda_k) = 0,$$

that is, $$d(\lambda_k y_k + (1 - \lambda_k)x_k, F(z_k)) = t_{\bar{y}} = (1 - t)\rho(\bar{z}, \bar{x}).$$

Since $\{\lambda_k\}_{k=1}^{\infty}$ is a bounded sequence in $\mathbb{R}$, due to Bolzano-Weierstrass theorem, it has a convergent subsequence $\{\widetilde{\lambda}_k\}_{k=1}^{\infty}$ such that $\widetilde{\lambda}_k \to \bar{\lambda}$ for some $\bar{\lambda} \in [0, 1]$. By passing to the corresponding subsequences $\{\widetilde{\varepsilon}_k\}_{k=1}^{\infty}$ of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{\widetilde{y}_k\}_{k=1}^{\infty}$ of $\{y_k\}_{k=1}^{\infty}$, $\{(\widetilde{z}_k, \widetilde{x}_k)\}_{k=1}^{\infty}$ of $\{(z_k, x_k)\}_{k=1}^{\infty}$ and $\{(\widetilde{z}_k^*, \widetilde{x}_k^*)\}_{k=1}^{\infty}$ of $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, one has for all $k \in \mathbb{N}$,

$$(\widetilde{z}_k, \widetilde{x}_k) \notin \operatorname{gph} F_{t_{\bar{y}}}, \quad \widetilde{y}_k \in \Pi(\widetilde{x}_k, F(\widetilde{z}_k)), \quad (\widetilde{z}_k^*, \widetilde{x}_k^*) \in \widehat{\partial}_{\widetilde{\varepsilon}_k}\rho(\widetilde{z}_k, \widetilde{x}_k),$$
$$\rho(\widetilde{z}_k, \widetilde{x}_k) = \rho_{t_{\bar{y}}}(\widetilde{z}_k, \widetilde{x}_k) + t_{\bar{y}}, \tag{3.7}$$
$$d(\widetilde{\lambda}_k\widetilde{y}_k + (1 - \widetilde{\lambda}_k)\widetilde{x}_k, F(\widetilde{z}_k)) = t_{\bar{y}} = (1 - t)\rho(\bar{z}, \bar{x}); \tag{3.8}$$

and $$\widetilde{\varepsilon}_k \downarrow 0, \quad (\widetilde{z}_k, \widetilde{x}_k) \xrightarrow{\rho} (\bar{z}, \bar{x}), \quad (\widetilde{z}_k, \widetilde{y}_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{y}).$$

Let $\widetilde{v}_k = \widetilde{\lambda}_k\widetilde{y}_k + (1 - \widetilde{\lambda}_k)\widetilde{x}_k$ for all $k \in \mathbb{N}$. Since $\widetilde{y}_k \in \Pi(\widetilde{x}_k, F(\widetilde{z}_k))$, $\|\widetilde{y}_k - \widetilde{x}_k\| = d(\widetilde{x}_k, F(\widetilde{z}_k)) = \rho(\widetilde{z}_k, \widetilde{x}_k)$. Owing to Lemma 3.1.1(a) and $(\widetilde{z}_k, \widetilde{x}_k) \xrightarrow{\rho} (\bar{z}, \bar{x})$, one obtains for all $k \in \mathbb{N}$,

$$d(\widetilde{v}_k, F(\widetilde{z}_k)) = (1 - \widetilde{\lambda}_k)\|\widetilde{y}_k - \widetilde{x}_k\| = (1 - \widetilde{\lambda}_k)\rho(\widetilde{z}_k, \widetilde{x}_k) \to (1 - \bar{\lambda})\rho(\bar{z}, \bar{x}). \tag{3.9}$$

Comparing (3.8) and (3.9), one has $(1-\bar{\lambda})\rho(\bar{z}, \bar{x}) = (1-t)\rho(\bar{z}, \bar{x})$, and in particular, $t = \bar{\lambda}$. Thus $\widetilde{v}_k = \widetilde{\lambda}_k\widetilde{y}_k + (1-\widetilde{\lambda}_k)\widetilde{x}_k \to \bar{\lambda}\bar{y} + (1-\bar{\lambda})\bar{x} = t\bar{y} + (1-t)\bar{x} = \bar{v}$. In light of (3.8), $\widetilde{v}_k \in F_{t_{\bar{y}}}(\widetilde{z}_k)$ and $(\widetilde{z}_k, \widetilde{v}_k) \in \operatorname{gph} F_{t_{\bar{y}}}$ for all $k \in \mathbb{N}$. It follows that $(\widetilde{z}_k, \widetilde{v}_k) \xrightarrow{\operatorname{gph} F_{t_{\bar{y}}}} (\bar{z}, \bar{v})$.

Due to (3.7), (3.8) and (3.9), one sees that

$$\begin{aligned}d(\widetilde{x}_k, F_{t_{\bar{y}}}(\widetilde{z}_k)) &= \rho_{t_{\bar{y}}}(\widetilde{z}_k, \widetilde{x}_k) = \rho(\widetilde{z}_k, \widetilde{x}_k) - t_{\bar{y}} = \rho(\widetilde{z}_k, \widetilde{x}_k) - d(\widetilde{v}_k, F(\widetilde{z}_k))\\ &= \|\widetilde{y}_k - \widetilde{x}_k\| - (1-\widetilde{\lambda}_k)\|\widetilde{y}_k - \widetilde{x}_k\| = \widetilde{\lambda}_k\|\widetilde{y}_k - \widetilde{x}_k\|\\ &= \|\widetilde{x}_k - (\widetilde{\lambda}_k\widetilde{y}_k + (1-\widetilde{\lambda}_k)\widetilde{x}_k)\| = \|\widetilde{x}_k - \widetilde{v}_k\|.\end{aligned}$$

Therefore $\widetilde{v}_k \in \Pi(\widetilde{x}_k, F_{t_{\bar{y}}}(\widetilde{z}_k))$ for all $k \in \mathbb{N}$.

Let $\gamma > 0$. For each $k \in \mathbb{N}$, since $(\widetilde{z}_k^*, \widetilde{x}_k^*) \in \widehat{\partial}_{\widetilde{\varepsilon}_k}\rho(\widetilde{z}_k, \widetilde{x}_k)$, an application of Proposition 1.5.5 shows that there exists $\delta_1^k > 0$ such that for all $(z,x) \in Z \times X$ with $\|z - \widetilde{z}_k\| \le \delta_1^k$ and $\|x - \widetilde{x}_k\| \le \delta_1^k$,

$$\langle \widetilde{z}_k^*, z - \widetilde{z}_k\rangle + \langle \widetilde{x}_k^*, x - \widetilde{x}_k\rangle \le \rho(z,x) - \rho(\widetilde{z}_k, \widetilde{x}_k) + (\widetilde{\varepsilon}_k + \gamma)(\|z - \widetilde{z}_k\| + \|x - \widetilde{x}_k\|). \quad (3.10)$$

Fix any $(z,x) \in Z\times X$ with $\|z-\widetilde{z}_k\| \le \delta_1^k$ and $\|x-\widetilde{x}_k\| \le \delta_1^k$. If $(z,x) \in \operatorname{gph} F_{t_{\bar{y}}}$, then $x \in F_{t_{\bar{y}}}(z)$, which implies $\rho_{t_{\bar{y}}}(z,x) = d(x, F_{t_{\bar{y}}}(z)) = 0$ and $\rho(z,x) = d(x, F(z)) \le t_{\bar{y}} = \rho_{t_{\bar{y}}}(z,x) + t_{\bar{y}}$. If $(z,x) \notin \operatorname{gph} F_{t_{\bar{y}}}$, then $\rho(z,x) = \rho_{t_{\bar{y}}}(z,x) + t_{\bar{y}}$ by virtue of Proposition 2.1.5. In both cases, there holds

$$\rho(z,x) \le \rho_{t_{\bar{y}}}(z,x) + t_{\bar{y}}. \quad (3.11)$$

It follows from (3.7), (3.10) and (3.11) that

$$\begin{aligned}&\langle \widetilde{z}_k^*, z - \widetilde{z}_k\rangle + \langle \widetilde{x}_k^*, x - \widetilde{x}_k\rangle\\ &\le (\rho_{t_{\bar{y}}}(z,x) + t_{\bar{y}}) - (\rho_{t_{\bar{y}}}(\widetilde{z}_k, \widetilde{x}_k) + t_{\bar{y}}) + (\widetilde{\varepsilon}_k + \gamma)(\|z - \widetilde{z}_k\| + \|x - \widetilde{x}_k\|)\\ &\le \rho_{t_{\bar{y}}}(z,x) - \rho_{t_{\bar{y}}}(\widetilde{z}_k, \widetilde{x}_k) + (\widetilde{\varepsilon}_k + \gamma)(\|z - \widetilde{z}_k\| + \|x - \widetilde{x}_k\|).\end{aligned}$$

Invoking Proposition 1.5.5, $(\widetilde{z}_k^*, \widetilde{x}_k^*) \in \widehat{\partial}_{\widetilde{\varepsilon}_k}\rho_{t_{\bar{y}}}(\widetilde{z}_k, \widetilde{x}_k)$ for all $k \in \mathbb{N}$. Since $\operatorname{gph} F_{t_{\bar{y}}}$ is closed, applying Proposition 3.1.3, one obtains for each $k \in \mathbb{N}$,

$$(\widetilde{z}_k^*, \widetilde{x}_k^*) \in \widehat{\partial}_{\widetilde{\varepsilon}_k}\rho_{t_{\bar{y}}}(\widetilde{z}_k, \widetilde{v}_k) \text{ and } 1 - \widetilde{\varepsilon}_k \le \|\widetilde{x}_k^*\| \le 1 + \widetilde{\varepsilon}_k. \quad (3.12)$$

In view of the first relation of (3.12) and Proposition 1.5.5, there exists $\delta_2^k > 0$ such that for all $(z,x) \in Z \times X$ with $\|z - \widetilde{z}_k\| \le \delta_2^k$ and $\|x - \widetilde{v}_k\| \le \delta_2^k$,

$$\langle \widetilde{z}_k^*, z - \widetilde{z}_k\rangle + \langle \widetilde{x}_k^*, x - \widetilde{v}_k\rangle \le \rho_{t_{\bar{y}}}(z,x) - \rho_{t_{\bar{y}}}(\widetilde{z}_k, \widetilde{v}_k) + (\widetilde{\varepsilon}_k + \gamma)(\|z - \widetilde{z}_k\| + \|x - \widetilde{v}_k\|). \quad (3.13)$$

Fix any $(z,x) \in \operatorname{gph} F_{t_{\bar{y}}}$ with $\|z - \widetilde{z}_k\| \le \delta_2^k$ and $\|x - \widetilde{v}_k\| \le \delta_2^k$. Then $x \in F_{t_{\bar{y}}}(z)$ and $\rho_{t_{\bar{y}}}(z,x) = d(x, F_{t_{\bar{y}}}(z)) = 0 \le \rho_{t_{\bar{y}}}(\widetilde{z}_k, \widetilde{x}_k)$, reducing (3.13) to

$$\langle \widetilde{z}_k^*, z - \widetilde{z}_k\rangle + \langle \widetilde{x}_k^*, x - \widetilde{v}_k\rangle \le (\widetilde{\varepsilon}_k + \gamma)(\|z - \widetilde{z}_k\| + \|x - \widetilde{v}_k\|).$$

Using Proposition 1.6.5, one obtains for all $k \in \mathbb{N}$, $(\widetilde{z}_k^*, \widetilde{x}_k^*) \in \widehat{N}_{\widetilde{\varepsilon}_k}((\widetilde{z}_k, \widetilde{v}_k); \operatorname{gph} F_{t_{\bar{y}}})$, which is well-defined since $(\widetilde{z}_k, \widetilde{v}_k) \in \operatorname{gph} F_{t_{\bar{y}}}$. This completes the proof of the proposition. □

Utilizing intermediate points, the next theorem enhances the results of Theorem 2.3.20(a) and Theorem 2.3.20(d).

Theorem 3.1.8. (cf. [27, Theorem 3.8 & Theorem 5.3]) *Let $F : Z \rightrightarrows X$ be a closed-graph mapping and $(\bar{z},\bar{x}) \notin \operatorname{gph} F$. For any $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$ and $t \in (0,1]$, assume that $\operatorname{gph} F_{t_{\bar{y}}}$ is closed, where $t_{\bar{y}} = \rho(\bar{z}, t\bar{y} + (1-t)\bar{x})$. Suppose further that the first criterion for well-posedness of the best approximation problem from $(\bar{z},\bar{x})$ to $\operatorname{gph} F$ via the generalized distance function is satisfied. Then for any $t \in (0,1]$, the following statements hold:*

(a) $\partial\rho(\bar{z},\bar{x}) \subset \bigcup_{\bar{y}\in\Pi(\bar{x},F(\bar{z}))} \{(z^*,x^*) \in N((\bar{z}, t\bar{y} + (1-t)\bar{x}); \operatorname{gph} F_{t_{\bar{y}}}) : \|x^*\| \le 1\}$.

(b) $\partial^\infty\rho(\bar{z},\bar{x}) \subset \bigcup_{\bar{y}\in\Pi(\bar{x},F(\bar{z}))} \{(z^*,0) \in Z^* \times X^* : z^* \in D_M^* F_{t_{\bar{y}}}(\bar{z}, t\bar{y} + (1-t)\bar{x})(0)\}$.

Proof. **(a)** Let $t \in (0,1]$ and $(z^*,x^*) \in \partial\rho(\bar{z},\bar{x})$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^\infty \subset \mathbb{R}_+$, $\{(z_k,x_k)\}_{k=1}^\infty \subset Z \times X$ and $\{(z_k^*,x_k^*)\}_{k=1}^\infty \subset Z^* \times X^*$ such that

$$\varepsilon_k \downarrow 0, \tag{3.14}$$

$$(z_k,x_k) \xrightarrow{\rho} (\bar{z},\bar{x}), \tag{3.15}$$

$$(z_k^*,x_k^*) \xrightarrow{w^*} (z^*,x^*), \text{ and} \tag{3.16}$$

$$(z_k^*,x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho(z_k,x_k) \text{ for all } k \in \mathbb{N}. \tag{3.17}$$

In view of the assumptions and using Lemma 2.3.18, there exist $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$ and a sequence $\{y_k\}_{k=1}^\infty \subset X$ which, by passing to appropriate subsequences of

$\{\varepsilon_k\}_{k=1}^{\infty}$, $\{(z_k, x_k)\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ if necessary, may be assumed to satisfy the aforementioned conditions (3.14), (3.15), (3.16) and (3.17), as well as the additional conditions $(z_k, y_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{y})$ and $y_k \in \Pi(x_k, F(z_k))$ for all $k \in \mathbb{N}$. Further employing Proposition 3.1.7, there exists a sequence $\{v_k\}_{k=1}^{\infty} \subset X$ which, by passing to appropriate subsequences of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{y_k\}_{k=1}^{\infty}$, $\{(z_k, x_k)\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ again if necessary, may be assumed to satisfy all preceding conditions, as well as the new conditions $(z_k, v_k) \xrightarrow{\operatorname{gph} F_{t_{\bar{y}}}} (\bar{z}, t\bar{y} + (1-t)\bar{x})$,

$$(z_k^*, x_k^*) \in \widehat{N}_{\varepsilon_k}((z_k, v_k); \operatorname{gph} F_{t_{\bar{y}}}) \text{ and } 1 - \varepsilon_k \leq \|x_k^*\| \leq 1 + \varepsilon_k \tag{3.18}$$

for all $k \in \mathbb{N}$. Consequently, $(z^*, x^*) \in N((\bar{z}, t\bar{y} + (1-t)\bar{x}); \operatorname{gph} F_{t_{\bar{y}}})$.

Invoking the lower semicontinuity of $\|\cdot\|$ with respect to the weak* topology of X^* in connection with the second relation of (3.18) yields

$$\|x^*\| \leq \liminf_{k\to\infty} \|x_k^*\| \leq \liminf_{k\to\infty}(1 + \varepsilon_k) = 1.$$

This proves

$$\partial\rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{y}\in\Pi(\bar{x}, F(\bar{z}))} \{(z^*, x^*) \in N((\bar{z}, t\bar{y} + (1-t)\bar{x}); \operatorname{gph} F_{t_{\bar{y}}}) : \|x^*\| \leq 1\}.$$

(b) Let $t \in (0, 1]$ and $(z^*, x^*) \in \partial^{\infty}\rho(\bar{z}, \bar{x})$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{\lambda_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{(z_k, x_k)\}_{k=1}^{\infty} \subset Z \times X$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty} \subset Z^* \times X^*$ such that

$$\varepsilon_k \downarrow 0, \quad \lambda_k \downarrow 0, \tag{3.19}$$

$$(z_k, x_k) \xrightarrow{\rho} (\bar{z}, \bar{x}), \tag{3.20}$$

$$\lambda_k(z_k^*, x_k^*) \xrightarrow{w^*} (z^*, x^*), \text{ and} \tag{3.21}$$

$$(z_k^*, x_k^*) \in \widehat{\partial}_{\varepsilon_k}\rho(z_k, x_k) \text{ for all } k \in \mathbb{N}. \tag{3.22}$$

As in the proof of **(a)**, in view of the assumptions and using Lemma 2.3.18, there exist $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$ and a sequence $\{y_k\}_{k=1}^{\infty} \subset X$ which, by passing to appropriate subsequences of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{\lambda_k\}_{k=1}^{\infty}$, $\{(z_k, x_k)\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$

if necessary, may be assumed to satisfy the aforementioned conditions (3.19), (3.20), (3.21) and (3.22), as well as the additional conditions $(z_k, y_k) \xrightarrow{\text{gph}\, F} (\bar{z}, \bar{y})$ and $y_k \in \Pi(x_k, F(z_k))$ for all $k \in \mathbb{N}$. Further employing Proposition 3.1.7, there exists a sequence $\{v_k\}_{k=1}^{\infty} \subset X$ which, by passing to appropriate subsequences of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{\lambda_k\}_{k=1}^{\infty}$, $\{y_k\}_{k=1}^{\infty}$, $\{(z_k, x_k)\}_{k=1}^{\infty}$ and $\{(z_k^*, x_k^*)\}_{k=1}^{\infty}$ again if necessary, may be assumed to satisfy all preceding conditions, as well as the new conditions $(z_k, v_k) \xrightarrow{\text{gph}\, F_{t_{\bar{y}}}} (\bar{z}, t\bar{y} + (1-t)\bar{x})$,

$$(z_k^*, x_k^*) \in \widehat{N}_{\varepsilon_k}((z_k, v_k); \text{gph}\, F_{t_{\bar{y}}}) \text{ and } 1 - \varepsilon_k \leq \|x_k^*\| \leq 1 + \varepsilon_k \tag{3.23}$$

for all $k \in \mathbb{N}$. Using the first relation of (3.23) and Proposition 1.6.6, one sees that for all $k \in \mathbb{N}$,

$$\lambda_k(z_k^*, x_k^*) \in \widehat{N}_{\lambda_k \varepsilon_k}((z_k, y_k); \text{gph}\, F_{t_{\bar{y}}}). \tag{3.24}$$

Note that, due to (3.19), $\lambda_k \varepsilon_k \downarrow 0$. Moreover, the second relation of (3.23) and $\varepsilon_k \downarrow 0$ reveal that $\{\|x_k^*\|\}_{k=1}^{\infty}$ is a bounded sequence in $\mathbb{R}$. With $\lambda_k \downarrow 0$, one has $\|\lambda_k x_k^*\| = \lambda_k \|x_k^*\| \to 0$, hence $\lambda_k x_k^* \to 0$ and in turn $-\lambda_k x_k^* \to 0$. It follows that $z^* \in D_M^* F(\bar{z}, t\bar{y} + (1-t)\bar{x})(0)$.

Owing to the lower semicontinuity of $\|\cdot\|$ with respect to the weak* topology of X^*,

$$\|x^*\| \leq \liminf_{k \to \infty} \|\lambda_k x_k^*\| = 0.$$

Therefore $\|x^*\| = 0$ and $x^* = 0$. This verifies

$$\partial^{\infty} \rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{y} \in \Pi(\bar{x}, F(\bar{z}))} \{(z^*, 0) \in Z^* \times X^* : z^* \in D_M^* F_{t_{\bar{y}}}(\bar{z}, t\bar{y} + (1-t)\bar{x})(0)\}. \quad \square$$

Remark 3.1.9. Taking $t = 1$ in the theorem, one has $t_{\bar{y}} = 0$, which implies $F_{t_{\bar{y}}} = F$ since F is closed-graph. Thus the conclusions reduce to the estimates via projections established in Theorem 2.3.20(a) and Theorem 2.3.20(d) respectively:

$$\partial \rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{y} \in \Pi(\bar{x}, F(\bar{z}))} \{(z^*, x^*) \in N((\bar{z}, \bar{y}); \text{gph}\, F) : \|x^*\| \leq 1\}, \text{ and}$$

$$\partial^{\infty} \rho(\bar{z}, \bar{x}) \subset \bigcup_{\bar{y} \in \Pi(\bar{x}, F(\bar{z}))} \{(z^*, 0) \in Z^* \times X^* : z^* \in D_M^* F(\bar{z}, \bar{y})(0)\}.$$

Corollary 3.1.10. *Let $\Omega \subset X$ be closed and $\bar{x} \notin \Omega$. For any $\bar{y} \in \Pi(\bar{x}, \Omega)$ and $t \in (0, 1]$, assume that $\Omega_{t_{\bar{y}}}$ is closed, where $t_{\bar{y}} = d(t\bar{y} + (1-t)\bar{x}, \Omega)$. Suppose further that the first criterion for well-posedness of the best approximation problem from $\bar{x}$ to Ω via the standard distance function is satisfied. Then for any $t \in (0, 1]$,*

$$\partial d(\bar{x}, \Omega) \subset \bigcup_{\bar{y} \in \Pi(\bar{x}, \Omega)} N(t\bar{y} + (1-t)\bar{x}; \Omega_{t_{\bar{y}}}) \cap \boldsymbol{B}_{X^*}.$$

Remark 3.1.11. Taking $t = 1$ in the corollary, one has $t_{\bar{y}} = 0$, which implies $\Omega_{t_{\bar{y}}} = \Omega$ since Ω is closed. Thus the conclusion reduces to the estimate via projections in Corollary 2.3.22(a) established for the case in which the first criterion for well-posedness of the best approximation problem is assumed to be satisfied:

$$\partial d(\bar{x}, \Omega) \subset \bigcup_{\bar{y} \in \Pi(\bar{x}, \Omega)} N(\bar{y}; \Omega) \cap \boldsymbol{B}_{X^*}.$$

The remaining part of this section focuses on a *Hilbert space* setting.

Lemma 3.1.12. *Let X be a Hilbert space, $\Omega \subset X$ and $\bar{x} \in X$ with $\Pi(\bar{x}, \Omega) \neq \varnothing$. The following statements hold:*

(a) *For any $t \in (0, 1]$ and $\bar{y} \in \Pi(\bar{x}, \Omega)$,*

$$\Pi(t\bar{y} + (1-t)\bar{x}, \Omega) = \{\bar{y}\}.$$

(b) *If Ω is closed, $\bar{x} \notin \Omega$ and $\widehat{\partial} d(\bar{x}, \Omega) \neq \varnothing$, then $\Pi(\bar{x}, \Omega)$ is a singleton and*

$$\widehat{\partial} d(\bar{x}, \Omega) = \frac{\bar{x} - \Pi(\bar{x}, \Omega)}{d(\bar{x}, \Omega)}.$$

Proof. (a) Let $t \in (0, 1]$ and $\bar{y} \in \Pi(\bar{x}, \Omega)$. By definition, $\|\bar{x} - \bar{y}\| = d(\bar{x}, \Omega)$ and $\bar{y} \in \Omega$. Define $v = t\bar{y} + (1-t)\bar{x}$. If $v = \bar{x}$, then $t \neq 0$ implies $\bar{y} = \bar{x} = v$ and $\Pi(v, \Omega) = \Pi(\bar{y}, \Omega) = \{\bar{y}\}$. Otherwise $v \neq \bar{x}$ and hence $\|v - \bar{x}\| > 0$. In view of Corollary 3.1.2, one has $d(v, \Omega) = (1-t)\|\bar{y} - \bar{x}\| = \|v - \bar{y}\|$ and $\bar{y} \in \Pi(v, \Omega)$. Assume there exists $u \neq \bar{y}$ such that $u \in \Pi(v, \Omega)$. Then $\|v - u\| = d(v, \Omega) = (1-t)\|\bar{y} - \bar{x}\|$ and $u \in \Omega$. There are two possible cases:

Case 1: $\bar{x}$, v and u are non-collinear.

Since X is Hilbert and hence strictly convex, the strict triangle inequality holds for $\bar{x}$, v and u. It follows that

$$\begin{aligned}\|\bar{x} - u\| &< \|\bar{x} - v\| + \|v - u\| \\ &= t\|\bar{y} - \bar{x}\| + (1-t)\|\bar{y} - \bar{x}\| \\ &= \|\bar{y} - \bar{x}\| = d(\bar{x}, \Omega),\end{aligned}$$

which contradicts $u \in \Omega$.

Case 2: $\bar{x}$, v and u are collinear.

There exists $\beta \in \mathbb{R}$ such that $u = \beta v + (1-\beta)\bar{x}$. Using $\bar{y} = \frac{1}{t}v + \left(1 - \frac{1}{t}\right)\bar{x}$, one sees that

$$|1-\beta|\|v - \bar{x}\| = \|v - u\| = d(v, \Omega) = \|v - \bar{y}\| = \left(\frac{1}{t} - 1\right)\|v - \bar{x}\|.$$

Therefore $\frac{1}{t} - 1 = |1-\beta|$, or equivalently, $\frac{1}{t} - 1 = \pm(1-\beta)$. Suppose $\frac{1}{t} - 1 = \beta - 1$. Then $\frac{1}{t} = \beta$ and thus $u = \bar{y}$, which contradicts the assumption $u \neq \bar{y}$. Otherwise $\frac{1}{t} - 1 = 1 - \beta$, which simplifies to $\beta = 2 - \frac{1}{t}$. Observe that

$$|\beta|\|v - \bar{x}\| = \|u - \bar{x}\| \geq d(\bar{x}, \Omega) = \|\bar{y} - \bar{x}\| = \frac{1}{t}\|v - \bar{x}\|,$$

which gives $|\beta| \geq \frac{1}{t}$. If $\beta \leq -\frac{1}{t}$, then $2 \leq 0$, which is a contradiction. If $\beta \geq \frac{1}{t}$, then $t \geq 1$. Since $t \in (0, 1]$, it is only possible that $t = 1$ and in turn $\beta = 1$. Therefore $u = \bar{y}$, which again contradicts the assumption $u \neq \bar{y}$. Consequently, $\Pi(v, \Omega) = \{\bar{y}\}$.

(b) Let $t \in (0, 1)$ and $\bar{y} \in \Pi(\bar{x}, \Omega)$. By definition, $\|\bar{x} - \bar{y}\| = d(\bar{x}, \Omega)$ and $\bar{y} \in \Omega$. Define $v = t\bar{y} + (1-t)\bar{x}$. By virtue of **(a)**, $\Pi(v, \Omega) = \{\bar{y}\}$. Moreover, $\bar{x} \notin \Omega$ implies $\bar{x} \neq \bar{y}$ and hence $v \notin \Omega$. By assumption, $\widehat{\partial} d(\bar{x}, \Omega) \neq \varnothing$. Taking $\varepsilon = 0$ in Corollary 3.1.4 yields

$$\varnothing \neq \widehat{\partial} d(\bar{x}, \Omega) \subset \widehat{\partial} d(v, \Omega). \tag{3.25}$$

Using [38, Theorem 5.3], one has

$$\widehat{\partial} d(v, \Omega) = \left\{\frac{v - \bar{y}}{\|v - \bar{y}\|}\right\} = \left\{\frac{(1-t)(\bar{x} - \bar{y})}{\|(1-t)(\bar{x} - \bar{y})\|}\right\} = \left\{\frac{\bar{x} - \bar{y}}{d(\bar{x}, \Omega)}\right\}. \tag{3.26}$$

In view of (3.25) and (3.26), $\widehat{\partial}d(\bar{x},\Omega) = \widehat{\partial}d(v,\Omega) = \left\{\frac{\bar{x}-\bar{y}}{d(\bar{x},\Omega)}\right\}$ is a singleton. Since $\bar{y} \in \Pi(\bar{x},\Omega)$ is arbitrary, $\widehat{\partial}d(\bar{x},\Omega) = \left\{\frac{\bar{x}-y}{d(\bar{x},\Omega)}\right\} = \left\{\frac{\bar{x}-\bar{y}}{d(\bar{x},\Omega)}\right\}$ for all $y \in \Pi(\bar{x},\Omega)$ and in turn $\Pi(\bar{x},\Omega) = \{\bar{y}\}$ is also a singleton. It follows that

$$\widehat{\partial}d(\bar{x},\Omega) = \frac{\bar{x} - \Pi(\bar{x},\Omega)}{d(\bar{x},\Omega)}.$$ □

With this lemma, efficient conditions which guarantee the nonemptiness of projection sets and refined upper estimates of limiting subdifferentials of the generalized distance function may be supplied.

Theorem 3.1.13. ([27, Theorem 6.1]) *Let $F : Z \rightrightarrows X$ be a closed-graph mapping from an Asplund space Z to a Hilbert space X and $(\bar{z},\bar{x}) \notin \operatorname{gph} F$. Suppose ρ is lower semicontinuous on a neighbourhood of $(\bar{z},\bar{x})$. The following statements hold:*

(a) *If $\{(z^*,x^*) \in \partial\rho(\bar{z},\bar{x}) : \|x^*\| = 1\} \neq \varnothing$, then $\Pi(\bar{x},F(\bar{z})) \neq \varnothing$.*

(b) $\{(z^*,x^*) \in \partial\rho(\bar{z},\bar{x}) : \|x^*\| = 1\}$
$$\subset \bigcup_{\bar{y}\in\Pi(\bar{x},F(\bar{z}))} \left\{(z^*,x^*) \in N((\bar{z},\bar{y});\operatorname{gph} F) : x^* = \frac{\bar{x}-\bar{y}}{\rho(\bar{z},\bar{x})}\right\}.$$

Proof. Let $(z^*,x^*) \in \partial\rho(\bar{z},\bar{x})$ with $\|x^*\| = 1$. By virtue of the assumptions on Z and X, the product space $Z \times X$ is Asplund. In light of Remark 1.5.13, there exist sequences $\{(z_k,x_k)\}_{k=1}^{\infty} \subset Z \times X$ and $\{(z_k^*,x_k^*)\}_{k=1}^{\infty} \subset Z^* \times X^*$ such that

$$(z_k,x_k) \xrightarrow{\rho} (\bar{z},\bar{x}), \tag{3.27}$$

$$(z_k^*,x_k^*) \xrightarrow{w^*} (z^*,x^*), \text{ and} \tag{3.28}$$

$$(z_k^*,x_k^*) \in \widehat{\partial}\rho(z_k,x_k) \text{ for all } k \in \mathbb{N}. \tag{3.29}$$

Since $\operatorname{gph} F$ is closed and $(\bar{z},\bar{x}) \notin \operatorname{gph} F$, $\rho(\bar{z},\bar{x}) = d(\bar{x},F(\bar{z})) > 0$. In view of (3.27), by considering the tail of $\{(z_k,x_k)\}_{k=1}^{\infty}$ together with the corresponding terms of $\{(z_k^*,x_k^*)\}_{k=1}^{\infty}$ if necessary, assume that for all $k \in \mathbb{N}$, $d(x_k,F(z_k)) = \rho(z_k,x_k) > 0$, which implies $x_k \notin F(z_k)$ and $(z_k,x_k) \notin \operatorname{gph} F$.

Let $\gamma > 0$. For each $k \in \mathbb{N}$, using (3.29) and Proposition 1.5.5, there exists $\delta_k > 0$ such that for all $(z, x) \in Z \times X$ with $\|z - z_k\| \leq \delta_k$ and $\|x - x_k\| \leq \delta_k$,

$$\langle z^*, z - z_k \rangle + \langle x^*, x - x_k \rangle \leq \rho(z, x) - \rho(z_k, x_k) + \gamma(\|z - z_k\| + \|x - x_k\|). \tag{3.30}$$

Taking $z = z_k$ in (3.30), one sees that for all $x \in X$ with $\|x - x_k\| \leq \delta_k$,

$$\begin{aligned}\langle x^*, x - x_k \rangle &\leq \rho(z_k, x) - \rho(z_k, x_k) + \gamma\|x - x_k\| \\ &= d(x, F(z_k)) - d(x_k, F(z_k)) + \gamma\|x - x_k\|.\end{aligned}$$

Invoking Proposition 1.5.5, $x_k^* \in \widehat{\partial} d(x_k, F(z_k)) \neq \varnothing$ for all $k \in \mathbb{N}$. Furthermore, $F(z_k)$ is closed. Lemma 3.1.12(b) reveals that $\Pi(x_k, F(z_k))$ is a singleton and

$$\widehat{\partial} d(x_k, F(z_k)) = \frac{x_k - \Pi(x_k, F(z_k))}{d(x_k, F(z_k))} = \{x_k^*\}. \tag{3.31}$$

Let $\Pi(x_k, F(z_k)) = \{y_k\}$ for all $k \in \mathbb{N}$. Then $\|x_k - y_k\| = d(x_k, F(z_k))$. It follows from (3.31) that

$$x_k^* = \frac{x_k - y_k}{d(x_k, F(z_k))} = \frac{x_k - y_k}{\rho(z_k, x_k)}, \tag{3.32}$$

which implies $\|x_k^*\| = 1$ and $y_k = x_k - \rho(z_k, x_k)x_k^*$ for all $k \in \mathbb{N}$. Since X is Hilbert, X^* is also Hilbert and hence possesses the Kadets-Klee property. Moreover, weak convergence and weak* convergence in X^* are equivalent, so that (3.28) implies $x_k^* \xrightarrow{w} x^*$. Owing to the observation $\|x_k^*\| = \|x^*\| = 1$ for all $k \in \mathbb{N}$, the Kadets-Klee property of X^* guarantees that $x_k^* \to x^*$. Together with (3.27), this shows $\bar{y} := \lim_{k\to\infty} y_k = \bar{x} - \rho(\bar{z}, \bar{x})x^*$, which can be rearranged as $x^* = \frac{\bar{x} - \bar{y}}{\rho(\bar{z}, \bar{x})}$. Then $\|\bar{x} - \bar{y}\| = \|\rho(\bar{z}, \bar{x})x^*\| = \rho(\bar{z}, \bar{x}) = d(\bar{x}, F(\bar{z}))$. Note that $y_k \in F(z_k)$ and hence $(z_k, y_k) \in \operatorname{gph} F$ for all $k \in \mathbb{N}$. It follows that $\{(z_k, y_k)\}_{k=1}^{\infty}$ is a sequence in $\operatorname{gph} F$ with $(z_k, y_k) \xrightarrow{\operatorname{gph} F} (\bar{z}, \bar{y})$. Since $\operatorname{gph} F$ is closed, $(\bar{z}, \bar{y}) \in \operatorname{gph} F$ and $\bar{y} \in F(\bar{z})$. Therefore $\bar{y} \in \Pi(\bar{x}, F(\bar{z})) \neq \varnothing$ and $N((\bar{z}, \bar{y}); \operatorname{gph} F)$ is well-defined.

From (3.29) and Proposition 2.2.9, one sees that $(z_k^*, x_k^*) \in \widehat{N}((z_k, y_k); \operatorname{gph} F)$ for all $k \in \mathbb{N}$. Owing to Remarks 1.6.8(iii), $(z^*, x^*) \in N((\bar{z}, \bar{y}); \operatorname{gph} F)$. The inclusion

$$\{(z^*, x^*) \in \partial\rho(\bar{z}, \bar{x}) : \|x^*\| = 1\} \subset \bigcup_{\bar{y} \in \Pi(\bar{x}, F(\bar{z}))} \left\{(z^*, x^*) \in N((\bar{z}, \bar{y}); \operatorname{gph} F) : x^* = \frac{\bar{x} - \bar{y}}{\rho(\bar{z}, \bar{x})}\right\}$$

is established. □

Several consequences of the theorem should be highlighted.

Corollary 3.1.14. *Let X be a Hilbert space, $\Omega \subset X$ be closed and $\bar{x} \notin \Omega$. The following statements hold:*

(a) *If $\partial d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*} \neq \varnothing$, then $\Pi(\bar{x}, \Omega) \neq \varnothing$.*

(b) $\partial_{\geq} d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*} \subset \partial d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*} \subset \dfrac{\bar{x} - \Pi(\bar{x}, \Omega)}{d(\bar{x}, \Omega)}$.

(c) *If $\widehat{\partial} d(\bar{x}, \Omega) \neq \varnothing$, then*

$$\partial_{\geq} d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*} = \partial d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*} = \frac{\bar{x} - \Pi(\bar{x}, \Omega)}{d(\bar{x}, \Omega)}$$

is a singleton.

(d) *If $X = \mathbb{R}^n$, then $\partial d(\bar{x}, \Omega) = \dfrac{\bar{x} - \Pi(\bar{x}, \Omega)}{d(\bar{x}, \Omega)}$.*

Proof. **(a)** The result follows from Theorem 3.1.13(a) by taking $Z = \{\bar{z}\}$ and $F \equiv \Omega$.

(b) Taking $Z = \{\bar{z}\}$ and $F \equiv \Omega$ in Theorem 3.1.13(b) gives $\partial d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*} \subset \frac{\bar{x} - \Pi(\bar{x}, \Omega)}{d(\bar{x}, \Omega)}$. On the other hand, Remarks 2.3.3(i) implies $\partial_{\geq} d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*} \subset \partial d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*}$. The assertion holds.

(c) Since $\widehat{\partial} d(\bar{x}, \Omega) \neq \varnothing$, in light of Lemma 3.1.12(b), $\Pi(\bar{x}, \Omega)$ is a singleton and $\widehat{\partial} d(\bar{x}, \Omega) = \frac{\bar{x} - \Pi(\bar{x}, \Omega)}{d(\bar{x}, \Omega)}$. Let $\Pi(\bar{x}, \Omega) = \{\bar{y}\}$. Then $\|\bar{x} - \bar{y}\| = d(\bar{x}, \Omega)$ and $\left\| \frac{\bar{x} - \bar{y}}{d(\bar{x}, \Omega)} \right\| = 1$. This implies $\widehat{\partial} d(\bar{x}, \Omega) = \left\{ \frac{\bar{x} - \bar{y}}{d(\bar{x}, \Omega)} \right\} = \widehat{\partial} d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*}$. By **(b)** and Remarks 2.3.3(i), one has

$$\left\{ \frac{\bar{x} - \bar{y}}{d(\bar{x}, \Omega)} \right\} = \widehat{\partial} d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*} \subset \partial_{\geq} d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*} \subset \frac{\bar{x} - \Pi(\bar{x}, \Omega)}{d(\bar{x}, \Omega)} = \left\{ \frac{\bar{x} - \bar{y}}{d(\bar{x}, \Omega)} \right\}.$$

It follows that

$$\widehat{\partial} d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*} = \partial_{\geq} d(\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*} = \frac{\bar{x} - \Pi(\bar{x}, \Omega)}{d(\bar{x}, \Omega)} = \left\{ \frac{\bar{x} - \bar{y}}{d(\bar{x}, \Omega)} \right\},$$

which is a singleton.

(d) This is stated in [33, Example 8.53]. □

Remark 3.1.15. For $X = \mathbb{R}^n$, compared to the upper estimate $\bigcup\limits_{\bar{y} \in \Pi(\bar{x}, \Omega)} N(\bar{y}; \Omega) \cap \boldsymbol{S}_{X^*}$ of limiting subdifferentials of the standard distance function given in Corollary 2.3.22(c), the exact representation $\frac{\bar{x} - \Pi(\bar{x}, \Omega)}{d(\bar{x}, \Omega)}$ in **(d)** is a remarkable improvement. In general, there holds

$$\frac{\bar{x} - \Pi(\bar{x}, \Omega)}{d(\bar{x}, \Omega)} \subset \bigcup_{\bar{y} \in \Pi(\bar{x}, \Omega)} N(\bar{y}; \Omega) \cap \boldsymbol{S}_{X^*};$$

the inclusion may be strict even for convex Ω.

3.2 Fréchet and Proximal Subdifferentials of the Generalized Distance Function via Intermediate Points

The use of intermediate points does not only produce new estimates of Fréchet-like and limiting subdifferentials, but also of the classical Fréchet and proximal subdifferentials of the generalized distance function. This section explores refined estimates of these classical subdifferentials via intermediate points.

The starting point is some elementary properties of proximal subdifferentials.

Lemma 3.2.1. *Let $F : Z \rightrightarrows X$ be a set-valued mapping, $(\bar{z}, \bar{x}) \in Z \times X$ and $r = \rho(\bar{z}, \bar{x})$. Then $\partial^p \rho(\bar{z}, \bar{x}) \subset N^p((\bar{z}, \bar{x}); \operatorname{gph} F_r)$.*

Proof. Let $(z^*, x^*) \in \partial^p \rho(\bar{z}, \bar{x})$. By definition, there exist $\delta > 0$ and $\eta > 0$ such that for all $(z, x) \in Z \times X$ with $\|z - \bar{z}\| \leq \delta$ and $\|x - \bar{x}\| \leq \delta$,

$$\langle (z^*, x^*), (z, x) - (\bar{z}, \bar{x}) \rangle \leq \rho(z, x) - \rho(\bar{z}, \bar{x}) + \eta \|(z, x) - (\bar{z}, \bar{x})\|^2. \tag{3.33}$$

Fix any $(z, x) \in \operatorname{gph} F_r$ with $\|z - \bar{z}\| \leq \delta$ and $\|x - \bar{x}\| \leq \delta$. Then $x \in F_r(z)$ and hence $\rho(z, x) = d(x, F(z)) \leq r = \rho(\bar{z}, \bar{x})$, reducing (3.33) to

$$\langle (z^*, x^*), (z, x) - (\bar{z}, \bar{x}) \rangle \leq \eta \|(z, x) - (\bar{z}, \bar{x})\|^2.$$

Consequently, $(z^*, x^*) \in N^p((\bar{z}, \bar{x}); \operatorname{gph} F_r)$, which verifies the assertion. □

Lemma 3.2.2. (cf. [18, Proposition 1.5]) *Let $F : Z \rightrightarrows X$ be a set-valued mapping and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$. Suppose $\operatorname{gph} F$ is locally closed at $(\bar{z}, \bar{x})$. Then for any $(z^*, x^*) \in \partial^p \rho(\bar{z}, \bar{x})$, $\|x^*\| = 1$.*

Proof. Let $(z^*, x^*) \in \partial^p \rho(\bar{z}, \bar{x})$. By definition, there exist $\delta > 0$ and $\eta > 0$ such that for all $(z, x) \in Z \times X$ with $\|z - \bar{z}\| \leq \delta$ and $\|x - \bar{x}\| \leq \delta$,

$$\langle (z^*, x^*), (z, x) - (\bar{z}, \bar{x}) \rangle \leq \rho(z, x) - \rho(\bar{z}, \bar{x}) + \eta \|(z, x) - (\bar{z}, \bar{x})\|^2. \tag{3.34}$$

Since $d(\cdot, F(\bar{z}))$ is Lipschitz with rank 1, by putting $z = \bar{z}$ in (3.34), one sees that for all $x \neq \bar{x}$ with $\|x - \bar{x}\| \leq \delta$,

$$\begin{aligned}\langle x^*, x - \bar{x}\rangle &\leq \rho(\bar{z}, x) - \rho(\bar{z}, \bar{x}) + \eta\|x - \bar{x}\|^2 \\ &\leq d(x, F(\bar{z})) - d(\bar{x}, F(\bar{z})) + \eta\|x - \bar{x}\|^2 \\ &\leq \|x - \bar{x}\| + \eta\|x - \bar{x}\|^2,\end{aligned} \tag{3.35}$$

which, upon rearrangement, takes the form

$$\frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \leq 1 + \eta\|x - \bar{x}\|. \tag{3.36}$$

On the other hand, by linearity of x^*, for any $\gamma > 0$,

$$\|x^*\| = \sup_{x \neq \bar{x}} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} = \sup_{0<\|x-\bar{x}\|\leq\gamma} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|},$$

which implies $\quad \|x^*\| = \inf_{\gamma>0} \sup_{0<\|x-\bar{x}\|\leq\gamma} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} = \limsup_{x\to\bar{x}} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|}.$

Hence it follows from (3.36) that

$$\|x^*\| = \limsup_{x\to\bar{x}} \frac{\langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \leq \limsup_{x\to\bar{x}}(1 + \eta\|x - \bar{x}\|) = 1.$$

Consider the opposite inequality. Since gph F is locally closed at $(\bar{z}, \bar{x})$ and $(\bar{z}, \bar{x}) \notin$ gph F, applying Proposition 2.1.2, $\rho(\bar{z}, \bar{x}) > 0$. Let $0 < t < \min\left\{1, \frac{\delta}{2\rho(\bar{z},\bar{x})}\right\}$. Then $(1 + t^2)\rho(\bar{z}, \bar{x}) > \rho(\bar{z}, \bar{x}) = d(\bar{x}, F(\bar{z}))$ implies $(1 + t^2)\rho(\bar{z}, \bar{x}) > \|\bar{x} - w_t\|$ for some $w_t \in F(\bar{z})$, or equivalently, $\rho(\bar{z}, \bar{x}) > \frac{\|\bar{x}-w_t\|}{1+t^2}$. Note that $\bar{x} \notin F(\bar{z})$ and $\bar{x} \neq w_t$. Let $y_t = (1 - t)\bar{x} + tw_t$. Check that

$$\bar{x} - y_t = t(\bar{x} - w_t), \quad y_t - w_t = (1 - t)(\bar{x} - w_t), \quad y_t \neq \bar{x}, \text{ and}$$
$$\|y_t - \bar{x}\| = t\|\bar{x} - w_t\| < \frac{\delta\|\bar{x} - w_t\|}{2\rho(\bar{z}, \bar{x})} < \frac{\delta\|\bar{x} - w_t\|}{(1 + t^2)\rho(\bar{z}, \bar{x})} < \frac{\delta\|\bar{x} - w_t\|}{\|\bar{x} - w_t\|} = \delta.$$

Putting $x = y_t$ in (3.35) yields

$$\begin{aligned}&\langle x^*, t(w_t - \bar{x})\rangle = \langle x^*, y_t - \bar{x}\rangle \\ &\leq d(y_t, F(\bar{z})) - d(\bar{x}, F(\bar{z})) + \eta\|y_t - \bar{x}\|^2 \\ &\leq \|y_t - w_t\| - \rho(\bar{z}, \bar{x}) + \eta\|y_t - \bar{x}\|^2 \\ &\leq (1 - t)\|\bar{x} - w_t\| - \frac{\|\bar{x} - w_t\|}{1 + t^2} + \eta t^2\|\bar{x} - w_t\|^2,\end{aligned}$$

which can be rearranged to give

$$\begin{aligned}\frac{\langle x^*, \bar{x} - w_t\rangle}{\|\bar{x} - w_t\|} &\geq \frac{1}{t}\left(t - 1 + \frac{1}{1+t^2}\right) - \eta t\|\bar{x} - w_t\| \\ &\geq \frac{t^2 - t + 1}{1+t^2} - \eta t\|\bar{x} - w_t\|.\end{aligned}$$

In view of of this inequality,

$$\|x^*\| = \sup_{u\neq 0} \frac{\langle x^*, u\rangle}{\|u\|} \geq \frac{\langle x^*, \bar{x} - w_t\rangle}{\|\bar{x} - w_t\|} \geq \frac{t^2 - t + 1}{1+t^2} - \eta t\|\bar{x} - w_t\|. \tag{3.37}$$

Letting $t \to 0$ in (3.37) shows that $\|x^*\| \geq 1$. This substantiates the assertion. □

The counterpart of Proposition 3.1.3 for Fréchet and proximal subdifferentials may be easily established by means of the preceding lemmas.

Theorem 3.2.3. (cf. [27, Theorem 3.3]) *Let $F : Z \rightrightarrows X$ be a closed-graph mapping and $(\bar{z}, \bar{x}) \notin \operatorname{gph} F$ with $\Pi(\bar{x}, F(\bar{z})) \neq \varnothing$. Then for any $t \in [0, 1]$ and $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$,*

$$\partial^\bullet \rho(\bar{z}, \bar{x}) \subset \{(z^*, x^*) \in \partial^\bullet \rho(\bar{z}, t\bar{y} + (1-t)\bar{x}) : \|x^*\| = 1\},$$

where $\partial^\bullet$ stands for $\widehat{\partial}$ or ∂^p.

Proof. **(a)** Consider the inclusion for $\widehat{\partial}$. Taking $\varepsilon = 0$ in Proposition 3.1.3, the conclusion follows immediately.

(b) Consider the inclusion for ∂^p. Let $t \in [0, 1]$, $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$ and $(z^*, x^*) \in \partial^p \rho(\bar{z}, \bar{x})$. By definition, there exist $\eta > 0$ and $\delta > 0$ such that for all $(z, x) \in Z \times X$ with $\|z - \bar{z}\| \leq \delta$ and $\|x - \bar{x}\| \leq \delta$,

$$\langle z^*, z - \bar{z}\rangle + \langle x^*, x - \bar{x}\rangle \leq \rho(z, x) - \rho(\bar{z}, \bar{x}) + \eta(\|z - \bar{z}\| + \|x - \bar{x}\|)^2. \tag{3.38}$$

Define $\bar{v} = t\bar{y} + (1-t)\bar{x}$. Employing Lemma 3.1.1(a), $\rho(\bar{z}, \bar{v}) = d(\bar{v}, F(\bar{z})) = (1-t)\|\bar{x} - \bar{y}\|$. Fix any $(z, x) \in Z \times X$ with $\|z - \bar{z}\| \leq \delta$ and $\|x - \bar{v}\| \leq \delta$. Note that $\|(x - \bar{v} + \bar{x}) - \bar{x}\| = \|x - \bar{v}\| \leq \delta$. Moreover, $\|\bar{x} - \bar{y}\| = d(\bar{x}, F(\bar{z})) = \rho(\bar{z}, \bar{x})$

and $\|\bar{v}-\bar{x}\| = t\|\bar{x}-\bar{y}\|$. By virtue of (3.38),

$$\begin{aligned}
&\langle z^*, z-\bar{z}\rangle + \langle x^*, x-\bar{v}\rangle \\
&= \langle z^*, z-\bar{z}\rangle + \langle x^*, (x-\bar{v}+\bar{x})-\bar{x}\rangle \\
&\leq \rho(z, x-\bar{v}+\bar{x}) - \rho(\bar{z},\bar{x}) + \eta(\|z-\bar{z}\| + \|(x-\bar{v}+\bar{x})-\bar{x}\|)^2 \\
&\leq \rho(z,x) + \|\bar{v}-\bar{x}\| - \|\bar{x}-\bar{y}\| + \eta(\|z-\bar{z}\| + \|x-\bar{v}\|)^2 \\
&= \rho(z,x) + t\|\bar{x}-\bar{y}\| - \|\bar{x}-\bar{y}\| + \eta(\|z-\bar{z}\| + \|x-\bar{v}\|)^2 \\
&= \rho(z,x) - (1-t)\|\bar{x}-\bar{y}\| + \eta(\|z-\bar{z}\| + \|x-\bar{v}\|)^2 \\
&= \rho(z,x) - \rho(\bar{z},\bar{v}) + \eta(\|z-\bar{z}\| + \|x-\bar{v}\|)^2.
\end{aligned}$$

By definition, $(z^*, x^*) \in \partial^p \rho(\bar{z},\bar{v})$. Moreover, it follows from Lemma 3.2.2 that $\|x^*\| = 1$. This completes the proof of

$$\partial^p \rho(\bar{z},\bar{x}) \subset \{(z^*,x^*) \in \partial^p \rho(\bar{z}, t\bar{y} + (1-t)\bar{x}) : \|x^*\| = 1\}. \qquad \square$$

Corollary 3.2.4. *Let $\Omega \subset X$ be closed and $\bar{x} \notin \Omega$ with $\Pi(\bar{x},\Omega) \neq \varnothing$. Then for any $t \in [0,1]$ and $\bar{y} \in \Pi(\bar{x},\Omega)$,*

$$\partial^\bullet d(\bar{x},\Omega) \subset \partial^\bullet d(t\bar{y} + (1-t)\bar{x}, \Omega) \cap \boldsymbol{S}_{X^*},$$

where $\partial^\bullet$ stands for $\widehat{\partial}$ or ∂^p.

Theorem 3.2.3 also relates Fréchet and proximal subdifferentials of the generalized distance function to their respective normal objects.

Proposition 3.2.5. (cf. [27, Corollary 1.1]) *Let $F : Z \rightrightarrows X$ be a closed-graph mapping and $(\bar{z},\bar{x}) \notin \operatorname{gph} F$ with $\Pi(\bar{x}, F(\bar{z})) \neq \varnothing$. Then for any $t \in [0,1]$ and $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$,*

$$\partial^\bullet \rho(\bar{z},\bar{x}) \subset \left\{(z^*,x^*) \in N^\bullet((\bar{z}, t\bar{y} + (1-t)\bar{x}); \operatorname{gph} F_{t_{\bar{y}}}) : \|x^*\| = 1\right\},$$

where $t_{\bar{y}} = \rho(\bar{z}, t\bar{y} + (1-t)\bar{x})$ and $(\partial^\bullet, N^\bullet)$ stands for $(\widehat{\partial}, \widehat{N})$ or (∂^p, N^p).

Proof. **(a)** Consider the inclusion for $(\widehat{\partial}, \widehat{N})$. Let $t \in [0,1]$, $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$ and $(z^*, x^*) \in \widehat{\partial}\rho(\bar{z},\bar{x})$. Using Proposition 3.1.3 with $\varepsilon = 0$, one sees that $(z^*, x^*) \in$

$\widehat{\partial}\rho(\bar{z}, t\bar{y} + (1-t)\bar{x})$ and $\|x^*\| = 1$. Then, applying Proposition 2.2.1 with $\varepsilon = 0$ and $r = \rho(\bar{z}, t\bar{y} + (1-t)\bar{x}) = t_{\bar{y}}$ yields $(z^*, x^*) \in \widehat{N}((\bar{z}, t\bar{y} + (1-t)\bar{x}); \text{gph}\, F_{t_{\bar{y}}})$. This justifies

$$\widehat{\partial}\rho(\bar{z}, \bar{x}) \subset \left\{(z^*, x^*) \in \widehat{N}((\bar{z}, t\bar{y} + (1-t)\bar{x}); \text{gph}\, F_{t_{\bar{y}}}) : \|x^*\| = 1\right\}.$$

(b) Consider the inclusion for (∂^p, N^p). Let $t \in [0,1]$, $\bar{y} \in \Pi(\bar{x}, F(\bar{z}))$ and $(z^*, x^*) \in \partial^p\rho(\bar{z}, \bar{x})$. Theorem 3.2.3 implies $(z^*, x^*) \in \partial^p\rho(\bar{z}, t\bar{y} + (1-t)\bar{x})$ and $\|x^*\| = 1$. Then, using Lemma 3.2.1 with $r = \rho(\bar{z}, t\bar{y} + (1-t)\bar{x}) = t_{\bar{y}}$, one has $(z^*, x^*) \in N^p((\bar{z}, t\bar{y} + (1-t)\bar{x}); \text{gph}\, F_{t_{\bar{y}}})$. It follows that

$$\partial^p\rho(\bar{z}, \bar{x}) \subset \left\{(z^*, x^*) \in N^p((\bar{z}, t\bar{y} + (1-t)\bar{x}); \text{gph}\, F_{t_{\bar{y}}}) : \|x^*\| = 1\right\}. \qquad \square$$

Corollary 3.2.6. *Let $\Omega \subset X$ be closed and $\bar{x} \notin \Omega$ with $\Pi(\bar{x}, \Omega) \neq \varnothing$. Then for any $t \in [0,1]$ and $\bar{y} \in \Pi(\bar{x}, \Omega)$,*

$$\partial^\bullet d(\bar{x}, \Omega) \subset N^\bullet(t\bar{y} + (1-t)\bar{x}; \Omega_{t_{\bar{y}}}) \cap \boldsymbol{S}_{X^*},$$

where $t_{\bar{y}} = d(t\bar{y} + (1-t)\bar{x}, \Omega)$ and $(\partial^\bullet, N^\bullet)$ stands for $(\widehat{\partial}, \widehat{N})$ or (∂^p, N^p).

Remark 3.2.7. Taking $t = 1$ in Proposition 3.2.5 and Corollary 3.2.6, one has $t_{\bar{y}} = 0$, which implies $\text{gph}\, F_{t_{\bar{y}}} = \text{gph}\, F$ and $\Omega_{t_{\bar{y}}} = \Omega$, since F is closed-graph and Ω is closed. Thus the conclusions reduce to the following estimates via projections, with the ones for $(\widehat{\partial}, \widehat{N})$ already established in Proposition 2.2.9 and Corollary 2.2.10 respectively:

$$\partial^\bullet\rho(\bar{z}, \bar{x}) \subset \{(z^*, x^*) \in N^\bullet((\bar{z}, \bar{y}); \text{gph}\, F) : \|x^*\| = 1\}, \text{ and}$$

$$\partial^\bullet d(\bar{x}, \Omega) \subset N^\bullet(\bar{y}; \Omega) \cap \boldsymbol{S}_{X^*}.$$

Chapter 4

The Marginal Function

This chapter endeavours to extend earlier results regarding singular subdifferentials of the standard distance function and the generalized distance function to more general classes of functions, which include the standard distance function and the generalized distance function as illustrative examples. These extended results published by Mordukhovich and Nam in [27] may be used to derive efficient subdifferential chain rules for compositions involving nonsmooth mappings. See [25] for more development.

4.1 Singular Subdifferentials of the Marginal Function

Indeed, the standard distance function belongs to a more general class of functions known as *marginal functions*, which are prominent in variational analysis, optimization and control theory. In particular, they are intimately related to the study of Lagrange multipliers and sensitivity analysis.

Definition 4.1.1. Let $\varphi : X \times Y \to \overline{\mathbb{R}}$ be a lower semicontinuous function and $G : X \rightrightarrows Y$ be a closed-graph mapping. Then $\mu : X \to \overline{\mathbb{R}}$ defined by

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

is called the ***marginal function*** (or ***value function***) ***generated by*** φ ***and*** G, and

$S : X \rightrightarrows Y$ defined by

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

is called the ***solution mapping associated with*** μ.

In other words, the marginal function describes the *optimal value* in a parametric minimization problem of the form

$$\text{minimize} \quad \varphi(x, y) \quad \text{subject to} \quad y \in G(x).$$

It is for this reason that the marginal function is also known as the *value function*. Akin to the standard distance function, the marginal function is nonsmooth and does not admit any classical derivative, even for smooth initial data.

For any nonempty closed subset $\Omega \subset X$, by considering the continuous function $\| \bullet_1 - \bullet_2 \|_X$ and the closed-graph mapping $F \equiv \Omega$, it is clear that the standard distance function $d(\cdot, \Omega) : X \to \mathbb{R}$ defined by

$$d(x, \Omega) := \inf\{\|w - x\| : w \in \Omega\} = \inf\{\|w - x\| : w \in F(x)\}$$

is the marginal function generated by $\| \bullet_1 - \bullet_2 \|_X$ and F, and the projection mapping $\Pi(\cdot, \Omega) : X \rightrightarrows X$ defined by

$$\Pi(x, \Omega) := \{w \in \Omega : \|w - x\| = d(x, \Omega)\} = \{w \in F(x) : \|w - x\| = d(x, \Omega)\}$$

is the solution mapping associated with $d(\cdot, \Omega)$.

Definition 4.1.2. Let $\varphi : X \times Y \to \overline{\mathbb{R}}$ be a lower semicontinuous function and $G : X \rightrightarrows Y$ be a closed-graph mapping. Define $\mu : X \to \overline{\mathbb{R}}$ to be the marginal function generated by φ and G, and $S : X \rightrightarrows Y$ to be the solution mapping associated with μ. S is said to be

(a) ***μ-inner semicontinuous at*** $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ if for any sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$ and $\{x_k\}_{k=1}^{\infty} \subset X$ with $\varepsilon_k \downarrow 0$, $x_k \xrightarrow{\mu} \bar{x}$ and $\widehat{\partial}_{\varepsilon_k} \mu(x_k) \neq \varnothing$ for all $k \in \mathbb{N}$, there exists a sequence $\{y_k\}_{k=1}^{\infty} \subset Y$ with $y_k \in S(x_k)$ for all $k \in \mathbb{N}$ which has a subsequence converging to $\bar{y}$;

(b) ***μ-inner semicompact at*** $\bar{x} \in X$ if for any sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$ and $\{x_k\}_{k=1}^{\infty} \subset X$ with $\varepsilon_k \downarrow 0$, $x_k \xrightarrow{\mu} \bar{x}$ and $\widehat{\partial}_{\varepsilon_k}\mu(x_k) \neq \varnothing$ for all $k \in \mathbb{N}$, there exists a sequence $\{y_k\}_{k=1}^{\infty} \subset Y$ with $y_k \in S(x_k)$ for all $k \in \mathbb{N}$ which has a subsequence converging to some $\bar{y} \in S(\bar{x})$.

Remark 4.1.3. Obviously, the μ-inner semicontinuity of S at $(\bar{x}, \bar{y})$ implies the μ-inner semicompactness of S at $\bar{x}$.

In the context of the standard distance function $d(\cdot, \Omega)$ associated with a nonempty closed subset $\Omega \subset X$, the $d(\cdot, \Omega)$-inner semicompactness of the projection mapping $\Pi(\cdot, \Omega)$ at $\bar{x} \in X$ is precisely the first criterion for well-posedness of the best approximation problem from $\bar{x}$ to Ω via the standard distance function. Indeed, both μ-inner semicontinuity and μ-inner semicompactness impose certain sequential compactness on the sequence $\{y_k\}_{k=1}^{\infty}$ with $(x_k, y_k) \in \operatorname{gph} S$ for all $k \in \mathbb{N}$. This central idea behind the two notions produces an interesting result about mixed coderivatives of the generating set-valued mapping of the marginal function.

Proposition 4.1.4. *Let X and Y be Asplund spaces, $\varphi : X \times Y \to \overline{\mathbb{R}}$ be a lower semicontinuous function and $G : X \rightrightarrows Y$ be a closed-graph mapping. Define $\mu : X \to \overline{\mathbb{R}}$ to be the marginal function generated by φ and G, and $S : X \rightrightarrows Y$ to be the solution mapping associated with μ. Suppose $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{\lambda_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{x_k\}_{k=1}^{\infty} \subset X$ and $\{x_k^*\}_{k=1}^{\infty} \subset X^*$ are sequences satisfying $\varepsilon_k \downarrow 0$, $\lambda_k \downarrow 0$, $x_k \xrightarrow{\mu} \bar{x}$ for some $\bar{x} \in X$, $\lambda_k x_k^* \xrightarrow{w^*} x^*$ for some $x^* \in X^*$, and $x_k^* \in \widehat{\partial}_{\varepsilon_k}\mu(x_k)$ for all $k \in \mathbb{N}$. Assume further that there exists a sequence $\{y_k\}_{k=1}^{\infty} \subset Y$ with $y_k \in S(x_k)$ for all $k \in \mathbb{N}$ which has a subsequence converging to some $\bar{y} \in S(\bar{x})$, and φ is locally Lipschitz at $(\bar{x}, \bar{y})$. Then $x^* \in D_M^* G(\bar{x}, \bar{y})(0)$.*

Proof. By passing to the convergent subsequence of $\{y_k\}_{k=1}^{\infty}$ together with the corresponding subsequences of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{\lambda_k\}_{k=1}^{\infty}$, $\{x_k\}_{k=1}^{\infty}$ and $\{x_k^*\}_{k=1}^{\infty}$ if necessary, assume that $y_k \to \bar{y}$. Suppose φ is locally Lipschitz at $(\bar{x}, \bar{y})$ with some rank $\ell \geq 0$. There exists $\alpha' > 0$ such that φ is finite and locally Lipschitz on $\boldsymbol{B}_{X\times Y}((\bar{x}, \bar{y}), \alpha')$ with rank ℓ. Let $0 < \alpha < \alpha'$. Since $x_k \to \bar{x}$ and $y_k \to \bar{y}$, by considering the tail of $\{x_k\}_{k=1}^{\infty}$ and $\{y_k\}_{k=1}^{\infty}$ together with the corresponding terms of $\{\varepsilon_k\}_{k=1}^{\infty}$, $\{\lambda_k\}_{k=1}^{\infty}$ and $\{x_k^*\}_{k=1}^{\infty}$ if necessary, assume that for all $k \in \mathbb{N}$, $(x_k, y_k) \in \boldsymbol{B}_{X\times Y}((\bar{x}, \bar{y}), \alpha)$, so that φ is finite and locally Lipschitz at (x_k, y_k) with rank ℓ. Accordingly, for all $k \in \mathbb{N}$, there exists $\alpha_k > 0$ such that φ is finite and locally Lipschitz on $\boldsymbol{B}_{X\times Y}((x_k, y_k), \alpha_k)$ with rank ℓ.

Let $\eta > 0$. For each $k \in \mathbb{N}$, in light of $x_k^* \in \widehat{\partial}_{\varepsilon_k}\mu(x_k)$ and Proposition 1.5.5, there exists $\gamma_k > 0$ such that for all $x \in X$ with $\|x - x_k\| \leq \gamma_k$,

$$\langle x_k^*, x - x_k\rangle \leq \mu(x) - \mu(x_k) + (\varepsilon_k + \eta)\|x - x_k\|. \tag{4.1}$$

Let $\beta_k = \min\{\alpha_k, \gamma_k\} > 0$ for all $k \in \mathbb{N}$. Fix any $k \in \mathbb{N}$ and $(x, y) \in X \times Y$ with $\|x - x_k\| \leq \beta_k$ and $\|y - y_k\| \leq \beta_k$. Note that $\varphi(x, y)$ is finite as $\|x - x_k\| \leq \alpha_k$ and $\|y - y_k\| \leq \alpha_k$. Suppose $(x, y) \in \operatorname{gph} G$. Then $y \in G(x)$ and $\delta_{\operatorname{gph} G}(x, y) = 0$. Thus

$$\mu(x) \leq \varphi(x, y) = \varphi(x, y) + \delta_{\operatorname{gph} G}(x, y) = (\varphi + \delta_{\operatorname{gph} G})(x, y).$$

Otherwise $(x, y) \notin \operatorname{gph} G$. Then $\delta_{\operatorname{gph} G}(x, y) = \infty$ and there trivially holds

$$\mu(x) \leq \varphi(x, y) + \delta_{\operatorname{gph} G}(x, y) = (\varphi + \delta_{\operatorname{gph} G})(x, y).$$

In both cases, $\mu(x) \leq (\varphi + \delta_{\operatorname{gph} G})(x, y)$. Since $y_k \in S(x_k) \subset G(x_k)$, $(x_k, y_k) \in \operatorname{gph} G$ and hence $\delta_{\operatorname{gph} G}(x_k, y_k) = 0$, which implies $\mu(x_k) = \varphi(x_k, y_k) = \varphi(x_k, y_k) + \delta_{\operatorname{gph} G}(x_k, y_k) = (\varphi + \delta_{\operatorname{gph} G})(x_k, y_k)$. Using $\|x - x_k\| \leq \gamma_k$, it follows from (4.1) that

$$\begin{aligned}
&\langle x_k^*, x - x_k\rangle + \langle 0, y - y_k\rangle \\
&= \langle x_k^*, x - x_k\rangle \\
&\leq \mu(x) - \mu(x_k) + (\varepsilon_k + \eta)\|x - x_k\| \\
&\leq (\varphi + \delta_{\operatorname{gph} G})(x, y) - (\varphi + \delta_{\operatorname{gph} G})(x_k, y_k) + (\varepsilon_k + \eta)(\|x - x_k\| + \|y - y_k\|).
\end{aligned}$$

Applying Proposition 1.5.5 again, $(x_k^*, 0) \in \widehat{\partial}_{\varepsilon_k}(\varphi + \delta_{\operatorname{gph} G})(x_k, y_k)$ for all $k \in \mathbb{N}$.

Let $\{\eta_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$ be a sequence such that $\eta_k < \alpha_k$ for all $k \in \mathbb{N}$ and $\eta_k \downarrow 0$. By virtue of the assumptions on X and Y, $X \times Y$ is an Asplund space. Since $\operatorname{gph} G$ is closed, $\delta_{\operatorname{gph} G}$ is lower semicontinuous. Moreover, φ and $\delta_{\operatorname{gph} G}$ are both proper functions. Invoking the *semi-Lipschitz fuzzy sum rule for ε-subdifferentials* (Theorem 1.5.11), for each $k \in \mathbb{N}$, there exist $(x_{1k}, y_{1k}), (x_{2k}, y_{2k}) \in \boldsymbol{B}_{X\times Y}((x_k, y_k), \eta_k)$, $(x_{1k}^*, y_{1k}^*) \in \widehat{\partial}\varphi(x_{1k}, y_{1k})$, $(x_{2k}^*, y_{2k}^*) \in \widehat{\partial}\delta_{\operatorname{gph} G}(x_{2k}, y_{2k})$ and $(x_{3k}^*, y_{3k}^*) \in \boldsymbol{B}_{X^*\times Y^*}$ such that

$$(x_k^*, 0) = (x_{1k}^*, y_{1k}^*) + (x_{2k}^*, y_{2k}^*) + (\varepsilon_k + \eta_k)(x_{3k}^*, y_{3k}^*), \tag{4.2}$$

$$|\varphi(x_{1k}, y_{1k}) - \varphi(x_k, y_k)| \leq \eta_k, \text{ and} \tag{4.3}$$

$$|\delta_{\operatorname{gph} G}(x_{2k}, y_{2k}) - \delta_{\operatorname{gph} G}(x_k, y_k)| \leq \eta_k. \tag{4.4}$$

With regard to $(x_{3k}^*, y_{3k}^*) \in \boldsymbol{B}_{X^*\times Y^*}$, there hold $\|x_{3k}^*\| \leq 1$ and $\|y_{3k}^*\| \leq 1$. Rearranging (4.2), one sees that for all $k \in \mathbb{N}$,

$$x_k^* - x_{1k}^* - x_{2k}^* = (\varepsilon_k + \eta_k)x_{3k}^* \quad \text{and} \quad -y_{1k}^* - y_{2k}^* = (\varepsilon_k + \eta_k)y_{3k}^*,$$

which imply $\|x_k^* - x_{1k}^* - x_{2k}^*\| \leq \varepsilon_k + \eta_k$ and $\|y_{1k}^* + y_{2k}^*\| \leq \varepsilon_k + \eta_k$. On the other hand, observe that $(x_{1k}, y_{1k}) \in \boldsymbol{B}_{X\times Y}((x_k, y_k), \eta_k) \subset \operatorname{int} \boldsymbol{B}_{X\times Y}((x_k, y_k), \alpha_k)$. Thus φ is locally Lipschitz at (x_{1k}, y_{1k}) with rank ℓ for all $k \in \mathbb{N}$. Since $(x_{1k}^*, y_{1k}^*) \in \widehat{\partial}\varphi(x_{1k}, y_{1k})$, employing Proposition 1.5.8 gives $\|(x_{1k}^*, y_{1k}^*)\| \leq \ell$ and hence $\{\|(x_{1k}^*, y_{1k}^*)\|\}_{k=1}^{\infty}$ is a bounded sequence in $\mathbb{R}$. These estimates yield for all $k \in \mathbb{N}$,

$$\begin{aligned}\|y_{2k}^*\| &\leq \|y_{1k}^* + y_{2k}^*\| + \|y_{1k}^*\| \\ &\leq \|y_{1k}^* + y_{2k}^*\| + \|(x_{1k}^*, y_{1k}^*)\| \\ &\leq \varepsilon_k + \eta_k + \ell.\end{aligned}$$

Due to $\varepsilon_k \downarrow 0$ and $\eta_k \downarrow 0$, $\{\|x_k^* - x_{1k}^* - x_{2k}^*\|\}_{k=1}^{\infty}$ and $\{\|y_{2k}^*\|\}_{k=1}^{\infty}$ are both bounded sequences in $\mathbb{R}$. With $\lambda_k \downarrow 0$, one has $\|\lambda_k y_{2k}^*\| = \lambda_k \|y_{2k}^*\| \to 0$, hence $\lambda_k y_{2k}^* \to 0$ and in turn

$$-\lambda_k y_{2k}^* \to 0. \tag{4.5}$$

Similarly, $\|\lambda_k(x_{1k}^*, y_{1k}^*)\| = \lambda_k\|(x_{1k}^*, y_{1k}^*)\| \to 0$, so that $\lambda_k(x_{1k}^*, y_{1k}^*) \to (0, 0)$, and in particular $\lambda_k x_{1k}^* \to 0$. Applying the same argument gives $\|\lambda_k(x_k^* - x_{1k}^* - x_{2k}^*)\| =$

$\lambda_k\|x_k^* - x_{1k}^* - x_{2k}^*\| \to 0$, so that $\lambda_k(x_k^* - x_{1k}^* - x_{2k}^*) \to 0$. These convergence relations, together with $\lambda_k x_k^* \xrightarrow{w^*} x^*$, produce

$$\lambda_k x_{2k}^* = \lambda_k x_k^* - \lambda_k x_{1k}^* - \lambda_k(x_k^* - x_{1k}^* - x_{2k}^*) \xrightarrow{w^*} x^*. \tag{4.6}$$

For all $k \in \mathbb{N}$, owing to (4.4),

$$\delta_{\operatorname{gph} G}(x_{2k}, y_{2k}) = |\delta_{\operatorname{gph} G}(x_{2k}, y_{2k}) - \delta_{\operatorname{gph} G}(x_k, y_k)| \le \eta_k < \infty,$$

so that $\delta_{\operatorname{gph} G}(x_{2k}, y_{2k}) = 0$ and $(x_{2k}, y_{2k}) \in \operatorname{gph} G$. Then $(x_{2k}, y_{2k}) \in \boldsymbol{B}_{X\times Y}((x_k, y_k), \eta_k)$ implies

$$\|x_{2k} - \bar{x}\| \le \|x_{2k} - x_k\| + \|x_k - \bar{x}\| \le \eta_k + \|x_k - \bar{x}\|, \text{ and}$$
$$\|y_{2k} - \bar{y}\| \le \|y_{2k} - y_k\| + \|y_k - \bar{y}\| \le \eta_k + \|y_k - \bar{y}\|.$$

It follows from $x_k \to \bar{x}$, $y_k \to \bar{y}$ and $\eta_k \downarrow 0$ that

$$(x_{2k}, y_{2k}) \xrightarrow{\operatorname{gph} G} (\bar{x}, \bar{y}). \tag{4.7}$$

Furthermore, for all $k \in \mathbb{N}$, since $(x_{2k}^*, y_{2k}^*) \in \widehat{\partial}\delta_{\operatorname{gph} G}(x_{2k}, y_{2k}) = \widehat{N}((x_{2k}, y_{2k}); \operatorname{gph} G)$, which is a cone, one arrives at

$$(\lambda_k x_{2k}^*, \lambda_k y_{2k}^*) = \lambda_k(x_{2k}^*, y_{2k}^*) \in \widehat{N}((x_{2k}, y_{2k}); \operatorname{gph} G) \subset \widehat{N}_{\varepsilon_k}((x_{2k}, y_{2k}); \operatorname{gph} G). \tag{4.8}$$

With (4.5), (4.6), (4.7), (4.8) and $\varepsilon_k \downarrow 0$, one obtains $x^* \in D_M^* G(\bar{x}, \bar{y})(0)$, which is well-defined since $\bar{y} \in S(\bar{x}) \subset G(\bar{x})$ and thus $(\bar{x}, \bar{y}) \in \operatorname{gph} G$. □

The theorem below, the validity of which relies heavily on Proposition 4.1.4, establishes significant relationships between singular subdifferentials of the marginal function and mixed coderivatives of its generating mapping in an Asplund space setting.

Theorem 4.1.5. (cf. [27, Theorem 5.1])) *Let X and Y be Asplund spaces, $\varphi : X \times Y \to \overline{\mathbb{R}}$ be a lower semicontinuous function and $G : X \rightrightarrows Y$ be a closed-graph mapping. Define $\mu : X \to \overline{\mathbb{R}}$ to be the marginal function generated by φ and G, and $S : X \rightrightarrows Y$ to be the solution mapping associated with μ. The following statements hold:*

(a) *If S is μ-inner semicontinuous at $(\bar{x},\bar{y}) \in \operatorname{gph} S$ and φ is locally Lipschitz at $(\bar{x},\bar{y})$, then*

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

(b) *If S is μ-inner semicompact at $\bar{x} \in X$ and φ is locally Lipschitz at $(\bar{x},\bar{y})$ 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. **(a)** Let $x^* \in \partial^{\infty}\mu(\bar{x})$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{\lambda_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{x_k\}_{k=1}^{\infty} \subset X$ and $\{x_k^*\}_{k=1}^{\infty} \subset X^*$ such that

$$\varepsilon_k \downarrow 0, \quad \lambda_k \downarrow 0, \tag{4.9}$$

$$x_k \xrightarrow{\mu} \bar{x}, \tag{4.10}$$

$$\lambda_k x_k^* \xrightarrow{w^*} x^*, \text{ and} \tag{4.11}$$

$$x_k^* \in \widehat{\partial}_{\varepsilon_k}\mu(x_k) \text{ for all } k \in \mathbb{N}. \tag{4.12}$$

Note that $(\bar{x},\bar{y}) \in \operatorname{gph} S$ is equivalent to $\bar{y} \in S(\bar{x})$. Since S is μ-inner semicontinuous at $(\bar{x},\bar{y})$, there exists a sequence $\{y_k\}_{k=1}^{\infty} \subset Y$ with $y_k \in S(x_k)$ for all $k \in \mathbb{N}$ which has a subsequence converging to $\bar{y}$. Applying Proposition 4.1.4 yields $x^* \in D_M^* G(\bar{x},\bar{y})(0)$. This substantiates

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

(b) Let $x^* \in \partial^{\infty}\mu(\bar{x})$. By definition, there exist sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{\lambda_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$, $\{x_k\}_{k=1}^{\infty} \subset X$ and $\{x_k^*\}_{k=1}^{\infty} \subset X^*$ satisfying (4.9) to (4.12). Since S is μ-inner semicompact at $\bar{x}$, there exists a sequence $\{y_k\}_{k=1}^{\infty} \subset Y$ with $y_k \in S(x_k)$ for all $k \in \mathbb{N}$ which has a subsequence converging to some $\bar{y} \in S(\bar{x})$. Applying Proposition 4.1.4 yields $x^* \in D_M^* G(\bar{x},\bar{y})(0)$. This verifies

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

4.2 Singular Subdifferentials of the Generalized Marginal Function

In order to include the generalized distance function in the framework, it is necessary to consider a slightly larger class of marginal functions which are related to minimization problems with *moving sets of feasible solutions.*

Definition 4.2.1. Let $\varphi : Y \times Z \to \overline{\mathbb{R}}$ be a lower semicontinuous function and $G : X \rightrightarrows Z$ be a closed-graph mapping. Then $\mu : X \times Y \to \overline{\mathbb{R}}$ defined by

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

is called the ***generalized marginal function*** (or ***generalized value function***) ***generated by*** φ and G, and $S : X \times Y \rightrightarrows Z$ defined by

$$S(x, y) := \{z \in G(x) : \varphi(y, z) = \mu(x, y)\}$$

is called the ***generalized solution mapping associated with*** μ.

Akin to the marginal function, the generalized marginal function describes the *optimal value* in a parametric minimization problem of the form

$$\text{minimize} \quad \varphi(y, z) \quad \text{subject to} \quad z \in G(x).$$

This justifies the use of the alternative terminology *generalized value function.*

By considering the continuous function $\| \bullet_1 - \bullet_2 \|_X : X \times X \to \mathbb{R}$ and any closed-graph mapping $F : Z \rightrightarrows X$, it is clear that the generalized distance function $\rho : \operatorname{dom} F \times X \to \mathbb{R}$ associated with F defined by

$$\rho(z, x) := d(x, F(z)) = \inf\{\|w - x\| : w \in F(z)\}$$

is the generalized marginal function generated by $\| \bullet_1 - \bullet_2 \|_X$ and F, and the projection mapping $\widetilde{\Pi} : \operatorname{dom} F \times X \rightrightarrows X$ defined by

$$\widetilde{\Pi}(z, x) := \Pi(x, F(z)) = \{w \in F(z) : \|w - x\| = d(x, F(z)) = \rho(z, x)\}$$

is the generalized solution mapping associated with ρ.

Definition 4.2.2. Let $\varphi : Y \times Z \to \overline{\mathbb{R}}$ be a lower semicontinuous function and $G : X \rightrightarrows Z$ be a closed-graph mapping. Define $\mu : X \times Y \to \overline{\mathbb{R}}$ to be the generalized marginal function generated by φ and G, and $S : X \times Y \rightrightarrows Z$ to be the generalized solution mapping associated with μ. S is said to be

(a) ***μ-inner semicontinuous at*** $((\bar{x}, \bar{y}), \bar{z}) \in \operatorname{gph} S$ if for any sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$ and $\{(x_k, y_k)\}_{k=1}^{\infty} \subset X \times Y$ with $\varepsilon_k \downarrow 0$, $(x_k, y_k) \xrightarrow{\mu} (\bar{x}, \bar{y})$ and $\widehat{\partial}_{\varepsilon_k} \mu(x_k, y_k) \neq \varnothing$ for all $k \in \mathbb{N}$, there exists a sequence $\{z_k\}_{k=1}^{\infty} \subset Z$ with $z_k \in S(x_k, y_k)$ for all $k \in \mathbb{N}$ which has a subsequence converging to $\bar{z}$;

(b) ***μ-inner semicompact at*** $(\bar{x}, \bar{y}) \in X \times Y$ if for any sequences $\{\varepsilon_k\}_{k=1}^{\infty} \subset \mathbb{R}_+$ and $\{(x_k, y_k)\}_{k=1}^{\infty} \subset X \times Y$ with $\varepsilon_k \downarrow 0$, $(x_k, y_k) \xrightarrow{\mu} (\bar{x}, \bar{y})$ and $\widehat{\partial}_{\varepsilon_k} \mu(x_k, y_k) \neq \varnothing$ for all $k \in \mathbb{N}$, there exists a sequence $\{z_k\}_{k=1}^{\infty} \subset Z$ with $z_k \in S(x_k, y_k)$ for all $k \in \mathbb{N}$ which has a subsequence converging to some $\bar{z} \in S(\bar{x}, \bar{y})$.

Remark 4.2.3. Clearly, the μ-inner semicontinuity of S at $((\bar{x}, \bar{y}), \bar{z})$ implies the μ-inner semicompactness of S at $(\bar{x}, \bar{y})$.

Again, both conditions impose some kind of sequential compactness on the sequence $\{z_k\}_{k=1}^{\infty}$ with $((x_k, y_k), z_k) \in \operatorname{gph} S$ for all $k \in \mathbb{N}$. In the context of the generalized distance function ρ associated with a closed-graph mapping $F : Z \rightrightarrows X$, the ρ-inner semicompactness of the projection mapping $\widetilde{\Pi}$ at $(\bar{z}, \bar{x}) \in Z \times X$ is precisely the first criterion for well-posedness of the best approximation problem from $(\bar{z}, \bar{x})$ to $\operatorname{gph} F$ via the generalized distance function.

Theorem 4.1.5 can be easily extended to hold for the generalized marginal function.

Theorem 4.2.4. (cf. [27, Corollary 5.2]) *Let X, Y and Z be Asplund spaces, $\varphi : Y \times Z \to \overline{\mathbb{R}}$ be a lower semicontinuous function and $G : X \rightrightarrows Z$ be a closed-graph mapping. Define $\mu : X \times Y \to \overline{\mathbb{R}}$ to be the generalized marginal function generated by*

φ and G, and $S : X \times Y \rightrightarrows Z$ to be the generalized solution mapping associated with μ. The following statements hold:

(a) *If S is μ-inner semicontinuous at $((\bar{x},\bar{y}),\bar{z}) \in \operatorname{gph} S$ and φ is locally Lipschitz at $(\bar{y},\bar{z})$, then*

$$\partial^{\infty}\mu(\bar{x},\bar{y}) \subset \{(x^*,0) \in X^* \times Y^* : x^* \in D_M^* G(\bar{x},\bar{z})(0)\}.$$

(b) *If S is μ-inner semicompact at $(\bar{x},\bar{y}) \in X \times Y$ and φ is locally Lipschitz at $(\bar{y},\bar{z})$ for all $\bar{z} \in S(\bar{x},\bar{y})$, then*

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

Proof. By virtue of the assumptions on X and Y, $X \times Y$ is an Asplund space. Let $\widetilde{G} : X \times Y \rightrightarrows Z$ be defined by $\widetilde{G}(x,y) = G(x)$ and $\widetilde{\varphi} : (X \times Y) \times Z \to \overline{\mathbb{R}}$ be defined by $\widetilde{\varphi}((x,y),z) = \varphi(y,z)$. Since G is closed-graph and φ is lower semicontinuous, $\widetilde{G}$ is also closed-graph and $\widetilde{\varphi}$ is also lower semicontinuous. Then the generalized marginal function $\mu : X \times Y \to \overline{\mathbb{R}}$ given by

$$\mu(x,y) = \inf\{\varphi(y,z) : z \in G(x)\} = \inf\{\widetilde{\varphi}((x,y),z) : z \in \widetilde{G}(x,y)\}$$

may be regarded as the marginal function generated by $\widetilde{\varphi}$ and $\widetilde{G}$, and the generalized solution mapping $S : X \times Y \rightrightarrows Z$ given by

$$\begin{aligned} S(x,y) &= \{z \in G(x) : \varphi(y,z) = \mu(x,y)\} \\ &= \{z \in \widetilde{G}(x,y) : \widetilde{\varphi}((x,y),z) = \mu(x,y)\} \end{aligned}$$

may be regarded as the solution mapping associated with μ.

In this extended setting, it is instructive to first demonstrate that for any $(\bar{x},\bar{y}) \in X \times Y$ and $\bar{z} \in G(\bar{x})$,

$$D_M^* \widetilde{G}((\bar{x},\bar{y}),\bar{z})(0) \subset \{(x^*,0) \in X^* \times Y^* : x^* \in D_M^* G(\bar{x},\bar{z})(0)\}. \tag{4.13}$$

Let $(\bar{x},\bar{y}) \in X \times Y$ and $\bar{z} \in G(\bar{x})$. Since $\bar{z} \in G(\bar{x}) = \widetilde{G}(\bar{x},\bar{y})$, $(\bar{x},\bar{z}) \in \operatorname{gph} G$ and $((\bar{x},\bar{y}),\bar{z}) \in \operatorname{gph}\widetilde{G}$, which guarantee that $D_M^*G(\bar{x},\bar{z})(0)$ and $D_M^*\widetilde{G}((\bar{x},\bar{y}),\bar{z})(0)$ are both well-defined.

Let $(x^*,y^*) \in D_M^*\widetilde{G}((\bar{x},\bar{y}),\bar{z})(0)$. Owing to Remarks 1.7.2(iv), there exist sequences $\{((x_k,y_k),z_k)\}_{k=1}^{\infty} \subset (X \times Y) \times Z$ and $\{((x_k^*,y_k^*),z_k^*)\}_{k=1}^{\infty} \subset (X^* \times Y^*) \times Z^*$ such that

$$((x_k,y_k),z_k) \xrightarrow{\operatorname{gph}\widetilde{G}} ((\bar{x},\bar{y}),\bar{z}), \tag{4.14}$$

$$(x_k^*,y_k^*) \xrightarrow{w^*} (x^*,y^*), \quad -z_k^* \to 0, \text{ and} \tag{4.15}$$

$$((x_k^*,y_k^*),z_k^*) \in \widehat{N}(((x_k,y_k),z_k);\operatorname{gph}\widetilde{G}) \text{ for all } k \in \mathbb{N}. \tag{4.16}$$

Let $\gamma > 0$. For all $k \in \mathbb{N}$, using (4.16) and Proposition 1.6.5, there exists $\delta_k > 0$ such that for all $((x,y),z) \in \operatorname{gph}\widetilde{G}$ with $\|x-x_k\| \le \delta_k$, $\|y-y_k\| \le \delta_k$ and $\|z-z_k\| \le \delta_k$,

$$\begin{aligned}\langle x_k^*, x-x_k\rangle + \langle y_k^*, y-y_k\rangle + \langle z_k^*, z-z_k\rangle \\ \le \gamma(\|x-x_k\| + \|y-y_k\| + \|z-z_k\|).\end{aligned} \tag{4.17}$$

Fix any $(x,z) \in \operatorname{gph} G$ with $\|x-x_k\| \le \delta_k$ and $\|z-z_k\| \le \delta_k$. Then $z \in G(x) = \widetilde{G}(x,y_k)$ and hence $((x,y_k),z) \in \operatorname{gph}\widetilde{G}$. Putting $y = y_k$ in (4.17), one has

$$\langle x_k^*, x-x_k\rangle + \langle z_k^*, z-z_k\rangle \le \gamma(\|x-x_k\| + \|z-z_k\|).$$

Employing Proposition 1.6.5 again yields $(x_k^*,z_k^*) \in \widehat{N}((x_k,z_k);\operatorname{gph} G)$ for all $k \in \mathbb{N}$. In view of (4.14), there hold $(x_k,z_k) \to (\bar{x},\bar{z})$ and $((x_k,y_k),z_k) \in \operatorname{gph}\widetilde{G}$ for all $k \in \mathbb{N}$. The latter implies $z_k \in \widetilde{G}(x_k,y_k) = G(x_k)$ and $(x_k,z_k) \in \operatorname{gph} G$. Thus $(x_k,z_k) \xrightarrow{\operatorname{gph} G} (\bar{x},\bar{z})$. In light of Remarks 1.7.2(iv), $x^* \in D_M^*G(\bar{x},\bar{z})(0)$.

Note that $z_k \in G(x_k) = \widetilde{G}(x_k,y)$ is equivalent to $((x_k,y),z_k) \in \operatorname{gph}\widetilde{G}$ for all $y \in Y$. Putting $x = x_k$ and $z = z_k$ in (4.17), for all $y \in Y$ with $\|y-y_k\| \le \delta_k$, one arrives at

$$\langle y_k^*, y-y_k\rangle \le \gamma\|y-y_k\|.$$

It follows from the linearity of y_k^* that

$$\|y_k^*\| = \sup_{y \ne y_k} \frac{\langle y^*, y-y_k\rangle}{\|y-y_k\|} = \sup_{0<\|y-y_k\|\le\delta_k} \frac{\langle y^*, y-y_k\rangle}{\|y-y_k\|} \le \gamma.$$

Since $\gamma > 0$ is arbitrary, $\|y_k^*\| \leq 0$ and hence $\|y_k^*\| = 0$ for all $k \in \mathbb{N}$. Invoking the lower semicontinuity of $\|\cdot\|$ with respect to the weak* topology of Y^* yields

$$\|y^*\| \leq \liminf_{k\to\infty} \|y_k^*\| = 0.$$

Therefore $\|y^*\| = 0$ and $y^* = 0$. This proves (4.13).

(a) Let $(x^*, y^*) \in \partial^\infty \mu(\bar{x}, \bar{y})$. The local Lipschitz property of φ at $(\bar{y}, \bar{z})$ implies that $\widetilde{\varphi}$ is locally Lipschitz at $((\bar{x}, \bar{y}), \bar{z})$. In view of the μ-inner semicontinuity of S at $((\bar{x}, \bar{y}), \bar{z})$, applying Theorem 4.1.5(a) gives $(x^*, y^*) \in D_M^*\widetilde{G}((\bar{x}, \bar{y}), \bar{z})(0)$. It follows from (4.13) that $(x^*, y^*) \in \{(x^*, 0) \in X^* \times Y^* : x^* \in D_M^*G(\bar{x}, \bar{z})(0)\}$. Thus

$$\partial^\infty \mu(\bar{x}, \bar{y}) \subset \{(x^*, 0) \in X^* \times Y^* : x^* \in D_M^*G(\bar{x}, \bar{z})(0)\}.$$

(b) Let $(x^*, y^*) \in \partial^\infty \mu(\bar{x}, \bar{y})$. The local Lipschitz property of φ at $(\bar{y}, \bar{z})$ for all $\bar{z} \in S(\bar{x}, \bar{y})$ implies that $\widetilde{\varphi}$ is locally Lipschitz at $((\bar{x}, \bar{y}), \bar{z})$ for all $\bar{z} \in S(\bar{x}, \bar{y})$. In view of the μ-inner semicompactness of S at $(\bar{x}, \bar{y})$, applying Theorem 4.1.5(b) gives $(x^*, y^*) \in D_M^*\widetilde{G}((\bar{x}, \bar{y}), \bar{z})(0)$ for some $\bar{z} \in S(\bar{x}, \bar{y})$. It follows from (4.13) that $(x^*, y^*) \in \{(x^*, 0) \in X^* \times Y^* : x^* \in D_M^*G(\bar{x}, \bar{z})(0)\}$. Hence

$$\partial^\infty \mu(\bar{x}, \bar{y}) \subset \bigcup_{\bar{z} \in S(\bar{x}, \bar{y})} \{(x^*, 0) \in X^* \times Y^* : x^* \in D_M^*G(\bar{x}, \bar{z})(0)\}. \qquad \square$$

Remark 4.2.5. In the context of the generalized distance function ρ associated with a closed-graph mapping F, **(b)** reduces specifically to Theorem 2.3.20(d).

Chapter 5

The Perturbed Distance Function

In previous chapters, the generalized distance function, which generalizes the standard distance function, and its extension, the generalized marginal function, have been studied. This chapter concerns yet another generalization of the standard distance function, the *perturbed distance function*.

5.1 Elementary Properties of the Perturbed Distance Function

Definition 5.1.1. Let $\Omega \subset X$ be a nonempty subset and $J : \Omega \to \mathbb{R}$ be a lower semicontinuous function. The ***J-perturbed distance function*** $d^J(\cdot, \Omega) : X \to \mathbb{R}$ ***associated with*** Ω is defined by

$$d^J(x, \Omega) := \inf\{\|x - w\| + J(w) : w \in \Omega\}.$$

Obviously, the perturbed distance function generalizes the standard distance function by including a perturbation generated by J. If $J \equiv 0$, the perturbed distance function $d^J(\cdot, \Omega)$ reduces immediately to the standard distance function $d(\cdot, \Omega)$.

The following are two elementary properties of the perturbed distance function.

Proposition 5.1.2. *Let $\Omega \subset X$ be a nonempty subset and $J : \Omega \to \mathbb{R}$ be a lower semicontinuous function. Then $d^J(\cdot, \Omega)$ is a Lipschitz function with rank 1.*

Proof. Let $x, y \in X$. Note that

$$\begin{aligned} d^J(x,\Omega) &= \inf\{\|w-x\| + J(w) : w \in \Omega\} \\ &\leq \inf\{\|w-y\| + \|y-x\| + J(w) : w \in \Omega\} \\ &= \inf\{\|w-y\| + J(w) : w \in \Omega\} + \|y-x\| \\ &= d^J(y,\Omega) + \|y-x\|. \end{aligned}$$

Rearranging the inequality gives $d^J(x,\Omega) - d^J(y,\Omega) \leq \|y-x\|$. Interchanging the roles of x and y, one sees that $d^J(y,\Omega) - d^J(x,\Omega) \leq \|y-x\|$. Hence $|d^J(y,\Omega) - d^J(x,\Omega)| \leq \|y-x\|$ for all $x, y \in X$. By definition, $d^J(\cdot,\Omega)$ is Lipschitz with rank 1. □

Proposition 5.1.3. *Let $\Omega \subset X$ be a nonempty convex subset and $J : \Omega \to \mathbb{R}$ be a lower semicontinuous convex function. Then $d^J(\cdot,\Omega)$ is a convex function.*

Proof. Let $x, y \in X$, $t \in (0,1)$ and $\varepsilon > 0$. Then there exist $w_1, w_2 \in \Omega$ such that $d^J(x,\Omega) + \varepsilon > \|x - w_1\| + J(w_1)$ and $d^J(y,\Omega) + \varepsilon > \|y - w_2\| + J(w_2)$. Since J is convex, $J(tw_1 + (1-t)w_2) \leq tJ(w_1) + (1-t)J(w_2)$. Moreover, the convexity of Ω guarantees that $tw_1 + (1-t)w_2 \in \Omega$. Hence

$$\begin{aligned} & td^J(x,\Omega) + (1-t)d^J(y,\Omega) + \varepsilon \\ &= t(d^J(x,\Omega) + \varepsilon) + (1-t)(d^J(y,\Omega) + \varepsilon) \\ &> t(\|x - w_1\| + J(w_1)) + (1-t)(\|y - w_2\| + J(w_2)) \\ &= \|tx - tw_1\| + \|(1-t)y - (1-t)w_2\| + tJ(w_1) + (1-t)J(w_2) \\ &\geq \|(tx + (1-t)y) - (tw_1 + (1-t)w_2)\| + J(tw_1 + (1-t)w_2) \\ &\geq d^J(tx + (1-t)y, \Omega). \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, $td^J(x,\Omega) + (1-t)d^J(y,\Omega) \geq d^J(tx + (1-t)y, \Omega)$. By definition, $d^J(\cdot,\Omega)$ is convex. □

Closely associated with the perturbed distance function is the perturbed minimization problem. For any nonempty $\Omega \subset X$ and $\bar{x} \in X$, the *perturbed minimization* *problem at* $\bar{x}$ *on* Ω is to find $x_0 \in \Omega$ which attains the infimum specified by the perturbed distance function, that is, to find $x_0 \in \Omega$ such that

$$d^J(\bar{x}, \Omega) = \|\bar{x} - x_0\| + J(x_0).$$

The perturbed minimization problem is not always solvable. It is heuristically clear that its solvability depends on a combination of factors, including in particular the choice of Ω, $\bar{x}$ and J. Points that solve the perturbed minimization problem at themselves are not only intriguing but also helpful in the analysis of the perturbed distance function.

Definition 5.1.4. Let $\Omega \subset X$ be a nonempty subset and $J : \Omega \to \mathbb{R}$ be a lower semicontinuous function. The ***self-solution set of the perturbed minimization problem on*** Ω is defined by

$$S(\Omega) := \{x \in \Omega : d^J(x, \Omega) = J(x)\}.$$

Remark 5.1.5. While the inequality $d^J(x, \Omega) \leq J(x)$ evidently holds for all $x \in \Omega$, the opposite inequality may fail to hold at every point in Ω and hence $S(\Omega)$ may be empty.

As in most minimization problems, the perturbed minimization problem may be conveniently tackled using *minimizing sequences.*

Definition 5.1.6. Let $\bar{x} \in X$, $\Omega \subset X$ be a nonempty subset and $J : \Omega \to \mathbb{R}$ be a lower semicontinuous function. A sequence $\{x_k\}_{k=1}^{\infty} \subset \Omega$ is said to be a ***minimizing sequence of the perturbed minimization problem at*** $\bar{x}$ ***on*** Ω if

$$\lim_{k \to \infty} (\|\bar{x} - x_k\| + J(x_k)) = d^J(\bar{x}, \Omega).$$

Remark 5.1.7. Trivially, if $x_0 \in \Omega$ is a solution to the perturbed minimization problem at $\bar{x}$ on Ω, then the constant sequence $\{x_k\}_{k=1}^{\infty} \subset \Omega$ with $x_k = x_0$ for all $k \in \mathbb{N}$ is a minimizing sequence which converges to x_0. Thus every solution to the perturbed minimization problem at $\bar{x}$ on Ω has a minimizing sequence which converges to it.

The notion of well-posedness was first formulated for the perturbed minimization problem in [21].

Definition 5.1.8. Let $\bar{x} \in X$, $\Omega \subset X$ be a nonempty subset and $J : \Omega \to \mathbb{R}$ be a lower semicontinuous function. ***The criterion for well-posedness of the perturbed minimization problem at*** $\bar{x}$ ***on*** Ω is that there is a unique solution $x_0 \in \Omega$, to which every minimizing sequence converges. The perturbed minimization problem is said to be ***well-posed*** if such criterion for well-posedness is satisfied.

With regard to the decisive role played by the convexity of Ω in the study of the generalized differential properties of the perturbed distance function, the case in which Ω is convex and that in which Ω is nonconvex are examined separately. The next two sections exhibit a number of results about Fréchet-like and proximal subdifferentials of the perturbed distance function originally communicated by Wang, Li and Xu in [37], which generalize the respective ones in [9, 12, 25] concerning the standard distance function. Most of these extended results utilize the function $J + \delta_\Omega$, which is *formally not globally defined* since J is only defined on Ω. This moderate hindrance may be easily surmounted by a *compromising* definition.

Definition 5.1.9. Let $\Omega \subset X$ be a nonempty subset and $J : \Omega \to \mathbb{R}$ be a lower semicontinuous function. The function $J + \delta_\Omega : X \to \overline{\mathbb{R}}$ is defined by

$$(J + \delta_\Omega)(x) := \begin{cases} J(x) & \text{if } x \in \Omega, \\ \infty & \text{if } x \notin \Omega. \end{cases}$$

5.2 The Convex Case - Subdifferentials of the Perturbed Distance Function

This section is devoted to the study of subdifferentials of the perturbed distance function for convex Ω. Among the many subdifferentials available, it is natural to use the classical subdifferential in convex analysis in view of the convexity of Ω.

Theorem 5.2.1. (cf. [37, Theorem 3.1]) *Let $\Omega \subset X$ be a nonempty convex subset and $\bar{x} \in X$. Suppose $J : \Omega \to \mathbb{R}$ is lower semicontinuous and convex. The following statements hold:*

(a) *If $\bar{x} \in \Omega$, then $\partial^c d^J(\bar{x}, \Omega) \supset \partial^c(J + \delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}$.*

(b) *If $\bar{x} \in S(\Omega)$, then $\partial^c d^J(\bar{x}, \Omega) \subset \partial^c(J + \delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}$.*

Proof. (a) Let $x^* \in \partial^c(J + \delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}$. Since $x^* \in \boldsymbol{B}_{X^*}$, $\|x^*\| \leq 1$ and hence for all $x, y \in X$,

$$\|y - x\| \geq \|x^*\|\|y - x\| \geq \langle x^*, y - x \rangle = \langle x^*, y - \bar{x} \rangle + \langle x^*, \bar{x} - x \rangle. \tag{5.1}$$

By assumption, $\bar{x} \in \Omega$. Using $x^* \in \partial^c(J + \delta_\Omega)(\bar{x})$, for all $x \in \Omega$,

$$J(x) = (J + \delta_\Omega)(x) \geq (J + \delta_\Omega)(\bar{x}) + \langle x^*, x - \bar{x} \rangle = J(\bar{x}) + \langle x^*, x - \bar{x} \rangle. \tag{5.2}$$

Thus for all $x \in \Omega$ and $y \in X$, by adding (5.1) and (5.2), one has

$$\begin{aligned} \|y - x\| + J(x) &\geq \langle x^*, y - \bar{x} \rangle + \langle x^*, \bar{x} - x \rangle + J(\bar{x}) + \langle x^*, x - \bar{x} \rangle \\ &= \langle x^*, y - \bar{x} \rangle + J(\bar{x}). \end{aligned} \tag{5.3}$$

On the other hand, $d^J(\bar{x}, \Omega) \leq J(\bar{x})$ owing to $\bar{x} \in \Omega$. It follows from (5.3) that for all $y \in X$,

$$\begin{aligned} d^J(y, \Omega) &\geq d^J(y, \Omega) + d^J(\bar{x}, \Omega) - J(\bar{x}) \\ &= \inf\{\|y - x\| + J(x) : x \in \Omega\} + d^J(\bar{x}, \Omega) - J(\bar{x}) \\ &= (\langle x^*, y - \bar{x} \rangle + J(\bar{x})) - J(\bar{x}) + d^J(\bar{x}, \Omega) \\ &= d^J(\bar{x}, \Omega) + \langle x^*, y - \bar{x} \rangle. \end{aligned}$$

By definition, $x^* \in \partial^c d^J(\bar{x}, \Omega)$. This completes the proof.

(b) Let $x^* \in \partial^c d^J(\bar{x}, \Omega)$. Since $\bar{x} \in S(\Omega) \subset \Omega$, $d^J(\bar{x}, \Omega) = J(\bar{x}) = (J + \delta_\Omega)(\bar{x})$. Then for all $x \in X$,

$$d^J(x, \Omega) \geq d^J(\bar{x}, \Omega) + \langle x^*, x - \bar{x} \rangle = (J + \delta_\Omega)(\bar{x}) + \langle x^*, x - \bar{x} \rangle. \qquad (5.4)$$

In particular, for all $x \in \Omega$, it follows from (5.4) that

$$(J + \delta_\Omega)(x) = J(x) \geq d^J(x, \Omega) \geq (J + \delta_\Omega)(\bar{x}) + \langle x^*, x - \bar{x} \rangle.$$

On the other hand, for all $x \notin \Omega$, $(J + \delta_\Omega)(x) = \infty$ and

$$(J + \delta_\Omega)(x) \geq (J + \delta_\Omega)(\bar{x}) + \langle x^*, x - \bar{x} \rangle$$

trivially holds. With the inequality valid for all $x \in X$, $x^* \in \partial^c(J + \delta_\Omega)(\bar{x})$. Moreover, since $d^J(\cdot, \Omega)$ is Lipschitz with rank 1, invoking Proposition 1.5.8 yields $x^* \in \boldsymbol{B}_{X^*}$. This justifies the conclusion. □

The well-known description (see [12, 15, 25]) of subdifferentials of the standard distance function may be regarded as a noticeable consequence of the preceding theorem.

Corollary 5.2.2. *Let $\Omega \subset X$ be a convex subset and $\bar{x} \in \Omega$. Then*

$$\partial^c d(\bar{x}, \Omega) = N^c(\bar{x}; \Omega) \cap \boldsymbol{B}_{X^*}.$$

Proof. Let $J \equiv 0$. Then J is trivially convex and lower semicontinuous. Moreover, $d^J(\cdot, \Omega)$ reduces to $d(\cdot, \Omega)$ and $S(\Omega) = \{x \in \Omega : d(x, \Omega) = 0\} = \Omega$, which implies $\bar{x} \in S(\Omega)$. Applying Theorem 5.2.1(a) and (b), one sees that

$$\partial^c d(\bar{x}, \Omega) = \partial^c \delta_\Omega(\bar{x}) \cap \boldsymbol{B}_{X^*} = N^c(\bar{x}; \Omega) \cap \boldsymbol{B}_{X^*}.$$ □

5.3 The Nonconvex Case - Fréchet-Like and Proximal Subdifferentials of the Perturbed Distance Function

This section proceeds to explore the generalized differential properties of the perturbed distance function for nonconvex Ω. In absence of convexity, Fréchet-like and proximal subdifferentials are more appropriate tools than the subdifferential in convex analysis.

The results of this section make use of a couple of *weakened* versions of the usual Lipschitz property extensively.

Definition 5.3.1. Let $\Omega \subset X$ and $f : X \to \overline{\mathbb{R}}$ be finite at $\bar{x} \in \Omega$.

(a) f is said to be ***centrally Lipschitz at $\bar{x}$ on Ω with rank $\ell \geq 0$*** if for all $x \in \Omega$,

$$|f(x) - f(\bar{x})| \leq \ell \|x - \bar{x}\|.$$

(b) f is said to be ***locally centrally Lipschitz at $\bar{x}$ on Ω with rank $\ell \geq 0$*** if there exists $\rho > 0$ such that for all $x \in \boldsymbol{B}_X(\bar{x}, \rho) \cap \Omega$,

$$|f(x) - f(\bar{x})| \leq \ell \|x - \bar{x}\|.$$

(c) The ***sharp local central Lipschitz rank of f at $\bar{x}$ on Ω*** is defined by

$$\ell_{\bar{x}} := \inf_{\rho > 0} \sup_{x \in (\boldsymbol{B}_X(\bar{x}, \rho) \setminus \{\bar{x}\}) \cap \Omega} \frac{|f(x) - f(\bar{x})|}{\|x - \bar{x}\|}.$$

Remarks 5.3.2. (i) Obviously, if f is centrally Lipschitz at $\bar{x}$ on Ω with some rank ℓ, then f is locally centrally Lipschitz at $\bar{x}$ on Ω with rank ℓ and $\ell_{\bar{x}} \leq \ell$.

(ii) Note that f is locally centrally Lipschitz at $\bar{x}$ on Ω if and only if $\ell_{\bar{x}} < \infty$.

The main results of this section rely on a lemma.

Lemma 5.3.3. ([37, Lemma 3.1]) *Let $\Omega \subset X$, $\bar{x} \in S(\Omega)$ and $J : \Omega \to \mathbb{R}$ be a lower semicontinuous function. Suppose the perturbed minimization problem at $\bar{x}$ on Ω is well-posed. Then for any $\gamma > 0$ and $\rho > 0$, there exists $0 < r < 1$ such that $\|y - \bar{x}\| < \rho$ holds whenever $x \in \boldsymbol{B}_X(\bar{x}, r)$ and $y \in \Omega$ satisfy $\|x - y\| + J(y) \leq d^J(x, \Omega) + \gamma \|x - \bar{x}\|$.*

Proof. Suppose, on the contrary, that there exist $\gamma_0 > 0$ and $\rho_0 > 0$ such that for all $0 < r < 1$,

$$\|x_r - y_r\| + J(y_r) \leq d^J(x_r, \Omega) + \gamma_0\|x_r - \bar{x}\| \quad \text{and} \quad \|y_r - \bar{x}\| \geq \rho_0$$

for some $x_r \in \boldsymbol{B}_X(\bar{x}, r)$ and $y_r \in \Omega$. Then for each $k \in \mathbb{N}$, there exist $x_k \in \boldsymbol{B}_X\left(\bar{x}, \frac{1}{k+1}\right)$ and $y_k \in \Omega$ such that

$$d^J(x_k, \Omega) \leq \|x_k - y_k\| + J(y_k) \leq d^J(x_k, \Omega) + \gamma_0\|x_k - \bar{x}\| \quad \text{and} \quad \|y_k - \bar{x}\| \geq \rho_0. \tag{5.5}$$

Note that $x_k \in \boldsymbol{B}_X\left(\bar{x}, \frac{1}{k+1}\right)$ implies $\|x_k - \bar{x}\| \leq \frac{1}{k+1}$ for all $k \in \mathbb{N}$ and hence $x_k \to \bar{x}$. This together with the continuity of $d^J(\cdot, \Omega)$ guarantees that

$$\lim_{k\to\infty} d^J(x_k, \Omega) = d^J(\bar{x}, \Omega) = \lim_{k\to\infty} (d^J(x_k, \Omega) + \gamma_0\|x_k - \bar{x}\|).$$

Thus letting $k \to \infty$ in the first inequality of (5.5) yields $\lim_{k\to\infty} (\|x_k - y_k\| + J(y_k)) = d^J(\bar{x}, \Omega)$. On the other hand, observe that for all $k \in \mathbb{N}$,

$$0 \leq |\,\|\bar{x} - y_k\| - \|x_k - y_k\|\,| \leq \|(\bar{x} - y_k) - (x_k - y_k)\| = \|\bar{x} - x_k\|. \tag{5.6}$$

Using $x_k \to \bar{x}$ again and letting $k \to \infty$ in (5.6), one has $\lim_{k\to\infty} (\|\bar{x} - y_k\| - \|x_k - y_k\|) = 0$. It follows that

$$\begin{aligned} &\lim_{k\to\infty} (\|\bar{x} - y_k\| + J(y_k)) = \lim_{k\to\infty} (\|\bar{x} - y_k\| - \|x_k - y_k\| + \|x_k - y_k\| + J(y_k)) \\ &= \lim_{k\to\infty} (\|\bar{x} - y_k\| - \|x_k - y_k\|) + \lim_{k\to\infty} (\|x_k - y_k\| + J(y_k)) = d^J(\bar{x}, \Omega). \end{aligned}$$

By definition, $\{y_k\}_{k=1}^{\infty} \subset \Omega$ is a minimizing sequence of the perturbed minimization problem at $\bar{x}$ on Ω, which is assumed to be well-posed. Then $\bar{x} \in S(\Omega)$ must be the unique solution and hence $y_k \to \bar{x}$, which contradicts the second inequality of (5.5). Thus the assertion holds. □

Theorem 5.3.4 below is the analogue of Theorem 5.2.1 for *Fréchet-like subdifferentials*. This result was first stated for *Fréchet subdifferentials* by Wang, Li and Xu in [37, Theorem 3.2]. Theorem 5.3.4, which reduces to their result as a special case for $\varepsilon = 0$, is established by adapting their proof.

Theorem 5.3.4. *Let $\Omega \subset X$, $\bar{x} \in S(\Omega)$ and $J : \Omega \to \mathbb{R}$ be a lower semicontinuous function. For any $\varepsilon \geq 0$, the following statements hold:*

(a) $\widehat{\partial}_\varepsilon d^J(\bar{x}, \Omega) \subset \widehat{\partial}_\varepsilon (J + \delta_\Omega)(\bar{x}) \cap (1 + \varepsilon)\boldsymbol{B}_{X^*}$.

(b) *Suppose the perturbed minimization problem at $\bar{x}$ on Ω is well-posed. If J is locally centrally Lipschitz at $\bar{x}$ on Ω with $\ell_{\bar{x}} < 1$, then*

$$\widehat{\partial}_{\frac{4\varepsilon}{1-\ell_{\bar{x}}}} d^J(\bar{x}, \Omega) \supset \widehat{\partial}_\varepsilon (J + \delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}.$$

Proof. (a) Let $\varepsilon \geq 0$, $x^* \in \widehat{\partial}_\varepsilon d^J(\bar{x}, \Omega)$ and $\gamma > 0$. In view of Proposition 1.5.5, there exists $\alpha > 0$ such that for all $x \in X$ with $\|x - \bar{x}\| \leq \alpha$,

$$\langle x^*, x - \bar{x} \rangle \leq d^J(x, \Omega) - d^J(\bar{x}, \Omega) + (\varepsilon + \gamma)\|x - \bar{x}\|. \tag{5.7}$$

Fix any $x \in X$ with $\|x - \bar{x}\| \leq \alpha$. Note that $d^J(\bar{x}, \Omega) = J(\bar{x}) = (J + \delta_\Omega)(\bar{x})$ is finite as $\bar{x} \in S(\Omega) \subset \Omega$. Suppose $x \in \Omega$. Then $d^J(x, \Omega) \leq J(x) = (J + \delta_\Omega)(x)$. It follows from (5.7) that

$$\langle x^*, x - \bar{x} \rangle \leq (J + \delta_\Omega)(x) - (J + \delta_\Omega)(\bar{x}) + (\varepsilon + \gamma)\|x - \bar{x}\|.$$

Otherwise $x \notin \Omega$. Then $(J + \delta_\Omega)(x) = \infty$ and the same inequality trivially holds. With the inequality valid for all $x \in X$ with $\|x - \bar{x}\| \leq \alpha$, by virtue of Proposition 1.5.5 again, $x^* \in \widehat{\partial}_\varepsilon (J + \delta_\Omega)(\bar{x})$. Moreover, since $d^J(\cdot, \Omega)$ is Lipschitz with rank 1, Proposition 1.5.6 implies $\|x^*\| \leq 1 + \varepsilon$. This ascertains

$$\widehat{\partial}_\varepsilon d^J(\bar{x}, \Omega) \subset \widehat{\partial}_\varepsilon (J + \delta_\Omega)(\bar{x}) \cap (1 + \varepsilon)\boldsymbol{B}_{X^*}.$$

(b) Let $\varepsilon \geq 0$, $x^* \in \widehat{\partial}_\varepsilon (J + \delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}$ and $\gamma > 0$. In light of $x^* \in \widehat{\partial}_\varepsilon (J + \delta_\Omega)(\bar{x})$ and Proposition 1.5.5, there exists $\rho_1 > 0$ such that for all $x \in X$ with $\|x - \bar{x}\| \leq \rho_1$,

$$\langle x^*, x - \bar{x} \rangle \leq (J + \delta_\Omega)(x) - (J + \delta_\Omega)(\bar{x}) + (\varepsilon + \gamma)\|x - \bar{x}\|. \tag{5.8}$$

Since $\ell_{\bar{x}} < 1$, one sees that $\ell_{\bar{x}} < \frac{\ell_{\bar{x}}+1}{2} < 1$. Then there exists $\rho_2 > 0$ such that for all $x \in \boldsymbol{B}_X(\bar{x}, \rho_2) \cap \Omega$,

$$|J(x) - J(\bar{x})| \leq \left(\frac{\ell_{\bar{x}} + 1}{2}\right) \|x - \bar{x}\|. \tag{5.9}$$

Let $\rho = \min\{\rho_1, \rho_2\} > 0$. Since the perturbed minimization problem at $\bar{x}$ on Ω is well-posed, invoking Lemma 5.3.3, there exists $0 < r < 1$ such that for all $x \in \boldsymbol{B}_X(\bar{x}, r)$ and $y \in \Omega$ satisfying $\|x - y\| + J(y) \leq d^J(x, \Omega) + \gamma\|x - \bar{x}\|$, there holds

$$\|y - \bar{x}\| < \rho. \tag{5.10}$$

Fix any $x \in \boldsymbol{B}_X(\bar{x}, r)\backslash\{\bar{x}\}$. Then $\|x - \bar{x}\| > 0$ and there exists $y_x \in \Omega$ such that

$$\|x - y_x\| + J(y_x) < d^J(x, \Omega) + \gamma\|x - \bar{x}\| \leq J(\bar{x}) + (1 + \gamma)\|x - \bar{x}\|, \tag{5.11}$$

which implies by (5.10) that $\|y_x - \bar{x}\| < \rho \leq \rho_1$. In view of (5.8), one has

$$\begin{aligned} \langle x^*, y_x - \bar{x}\rangle &\leq (J + \delta_\Omega)(y_x) - (J + \delta_\Omega)(\bar{x}) + (\varepsilon + \gamma)\|y_x - \bar{x}\| \\ &= J(y_x) - J(\bar{x}) + (\varepsilon + \gamma)\|y_x - \bar{x}\|. \end{aligned} \tag{5.12}$$

Moreover, $\|y_x - \bar{x}\| < \rho \leq \rho_2$ and $y_x \in \Omega$ together imply $y_x \in \boldsymbol{B}_X(\bar{x}, \rho_2) \cap \Omega$. It follows from (5.9) that

$$|J(y_x) - J(\bar{x})| \leq \left(\frac{\ell_{\bar{x}} + 1}{2}\right)\|y_x - \bar{x}\|. \tag{5.13}$$

Employing (5.11) and (5.13), one obtains

$$\begin{aligned} \|y_x - \bar{x}\| &\leq \|y_x - x\| + \|x - \bar{x}\| \\ &\leq J(\bar{x}) + (1 + \gamma)\|x - \bar{x}\| - J(y_x) + \|x - \bar{x}\| \\ &\leq |J(y_x) - J(\bar{x})| + (2 + \gamma)\|x - \bar{x}\| \\ &\leq \left(\frac{\ell_{\bar{x}} + 1}{2}\right)\|y_x - \bar{x}\| + (2 + \gamma)\|x - \bar{x}\|, \end{aligned}$$

which can be rearranged to give

$$\|y_x - \bar{x}\| \leq \left(\frac{2 + \gamma}{1 - \frac{\ell_{\bar{x}} + 1}{2}}\right)\|x - \bar{x}\| = \left(\frac{4 + 2\gamma}{1 - \ell_{\bar{x}}}\right)\|x - \bar{x}\|. \tag{5.14}$$

Since $\bar{x} \in S(\Omega)$ and $x^* \in \boldsymbol{B}_{X^*}$, $d^J(\bar{x},\Omega) = J(\bar{x})$ and $\|x^*\| \leq 1$. Owing to estimates (5.11), (5.12) and (5.14), one arrives at

$$\begin{aligned}
&\langle x^*, x-\bar{x}\rangle \\
&= \langle x^*, x-y_x\rangle + \langle x^*, y_x-\bar{x}\rangle \\
&\leq \|x^*\|\|x-y_x\| + (J(y_x) - J(\bar{x}) + (\varepsilon+\gamma)\|y_x-\bar{x}\|) \\
&\leq \|x-y_x\| + \left(J(y_x) - J(\bar{x}) + \left(\frac{(\varepsilon+\gamma)(4+2\gamma)}{1-\ell_{\bar{x}}}\right)\|x-\bar{x}\|\right) \\
&\leq \left(d^J(x,\Omega) + \gamma\|x-\bar{x}\| - J(y_x)\right) \\
&\quad + \left(J(y_x) - d^J(\bar{x},\Omega) + \left(\frac{(\varepsilon+\gamma)(4+2\gamma)}{1-\ell_{\bar{x}}}\right)\|x-\bar{x}\|\right) \\
&= d^J(x,\Omega) - d^J(\bar{x},\Omega) + \left(\gamma + \frac{(\varepsilon+\gamma)(4+2\gamma)}{1-\ell_{\bar{x}}}\right)\|x-\bar{x}\| \\
&= d^J(x,\Omega) - d^J(\bar{x},\Omega) + \left(\frac{4\varepsilon}{1-\ell_{\bar{x}}} + \gamma\left(1+\frac{2\gamma+2\varepsilon+4}{1-\ell_{\bar{x}}}\right)\right)\|x-\bar{x}\|.
\end{aligned}$$

On the other hand, for $x = \bar{x}$, equality trivially holds in

$$\langle x^*, x-\bar{x}\rangle \leq d^J(x,\Omega) - d^J(\bar{x},\Omega) + \left(\frac{4\varepsilon}{1-\ell_{\bar{x}}} + \gamma\left(1+\frac{2\gamma+2\varepsilon+4}{1-\ell_{\bar{x}}}\right)\right)\|x-\bar{x}\|.$$

With the inequality valid for all $x \in \boldsymbol{B}_X(\bar{x},r)$, by virtue of Proposition 1.5.5 again, $x^* \in \widehat{\partial}_{\frac{4\varepsilon}{1-\ell_{\bar{x}}}} d^J(\bar{x},\Omega)$. This justifies

$$\widehat{\partial}_{\frac{4\varepsilon}{1-\ell_{\bar{x}}}} d^J(\bar{x},\Omega) \supset \widehat{\partial}_{\varepsilon}(J+\delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}. \qquad \square$$

Indeed, the counterpart of Theorem 5.2.1 for proximal subdifferentials also holds. As in Theorem 5.3.4, the perturbed minimization problem at $\bar{x}$ on Ω is required to be well-posed and J is required to be locally centrally Lipschitz at $\bar{x}$ on Ω in order to compensate for the nonconvexity of Ω.

Theorem 5.3.5. (cf. [37, Theorem 3.3]) *Let $\Omega \subset X$, $\bar{x} \in S(\Omega)$ and $J : \Omega \to \mathbb{R}$ be a lower semicontinuous function. The following statements hold:*

(a) $\partial^p d^J(\bar{x},\Omega) \subset \partial^p(J+\delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}$.

(b) *Suppose the perturbed minimization problem at $\bar{x}$ on Ω is well-posed. If J is locally centrally Lipschitz at $\bar{x}$ on Ω with $\ell_{\bar{x}} < 1$, then*

$$\partial^p d^J(\bar{x}, \Omega) \supset \partial^p (J + \delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}.$$

Proof. **(a)** Let $x^* \in \partial^p d^J(\bar{x}, \Omega)$. By definition, there exist $\alpha > 0$ and $\eta > 0$ such that for all $x \in \boldsymbol{B}_X(\bar{x}, \alpha)$,

$$\langle x^*, x - \bar{x} \rangle \leq d^J(x, \Omega) - d^J(\bar{x}, \Omega) + \eta \|x - \bar{x}\|^2. \tag{5.15}$$

Fix any $x \in \boldsymbol{B}_X(\bar{x}, \alpha)$. Note that $d^J(\bar{x}, \Omega) = J(\bar{x}) = (J + \delta_\Omega)(\bar{x})$ is finite as $\bar{x} \in S(\Omega) \subset \Omega$. Suppose $x \in \Omega$. Then $d^J(x, \Omega) \leq J(x) = (J + \delta_\Omega)(x)$. It follows from (5.15) that

$$\langle x^*, x - \bar{x} \rangle \leq (J + \delta_\Omega)(x) - (J + \delta_\Omega)(\bar{x}) + \eta \|x - \bar{x}\|^2.$$

Otherwise $x \notin \Omega$. Then $(J + \delta_\Omega)(x) = \infty$ and the same inequality trivially holds. With the inequality valid for all $x \in \boldsymbol{B}_X(\bar{x}, \alpha)$, by definition, $x^* \in \partial^p (J + \delta_\Omega)(\bar{x})$. Moreover, since $d(\cdot, \Omega)$ is Lipschitz with rank 1, Proposition 1.5.8 implies $x^* \in \boldsymbol{B}_{X^*}$. This ascertains

$$\partial^p d^J(\bar{x}, \Omega) \subset \partial^p (J + \delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}.$$

(b) Let $x^* \in \partial^p (J + \delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}$. In light of $x^* \in \partial^p (J + \delta_\Omega)(\bar{x})$, there exist $\rho_1 > 0$ and $\eta > 0$ such that for all $x \in \boldsymbol{B}_X(\bar{x}, \rho_1)$,

$$\langle x^*, x - \bar{x} \rangle \leq (J + \delta_\Omega)(x) - (J + \delta_\Omega)(\bar{x}) + \eta \|x - \bar{x}\|^2. \tag{5.16}$$

Let $\ell_{\bar{x}} < \ell < 1$. Then there exists $\rho_2 > 0$ such that for all $x \in \boldsymbol{B}_X(\bar{x}, \rho_2) \cap \Omega$,

$$|J(x) - J(\bar{x})| \leq \ell \|x - \bar{x}\|. \tag{5.17}$$

Let $\rho = \min\{\rho_1, \rho_2\} > 0$. Since the perturbed minimization problem at $\bar{x}$ on Ω is well-posed, invoking Lemma 5.3.3, there exists $0 < r < 1$ such that for all

$x \in \boldsymbol{B}_X(\bar{x}, r)$ and $y \in \Omega$ satisfying $\|x - y\| + J(y) \leq d^J(x, \Omega) + \|x - \bar{x}\|$, there holds

$$\|y - \bar{x}\| < \rho. \tag{5.18}$$

Fix any $x \in \boldsymbol{B}_X(\bar{x}, r)\backslash\{\bar{x}\}$. Then $0 < \|x - \bar{x}\| \leq r < 1$ and there exists $y_x \in \Omega$ such that

$$\|x - y_x\| + J(y_x) < d^J(x, \Omega) + \|x - \bar{x}\|^2 \tag{5.19}$$

$$< d^J(x, \Omega) + \|x - \bar{x}\|, \tag{5.20}$$

which implies by (5.18) that $\|y_x - \bar{x}\| < \rho \leq \rho_1$ and $y_x \in \boldsymbol{B}_X(\bar{x}, \rho_1)$. In view of (5.16), one sees that

$$\begin{aligned}\langle x^*, y_x - \bar{x}\rangle &\leq (J + \delta_\Omega)(y_x) - (J + \delta_\Omega)(\bar{x}) + \eta\|y_x - \bar{x}\|^2 \\ &= J(y_x) - J(\bar{x}) + \eta\|y_x - \bar{x}\|^2. \end{aligned} \tag{5.21}$$

Moreover, $\|y_x - \bar{x}\| < \rho \leq \rho_2$ and $y_x \in \Omega$ together imply $y_x \in \boldsymbol{B}_X(\bar{x}, \rho_2) \cap \Omega$. It follows from (5.17) that

$$|J(y_x) - J(\bar{x})| \leq \ell\|y_x - \bar{x}\|. \tag{5.22}$$

Employing (5.20) and (5.22), one obtains

$$\begin{aligned}\|y_x - \bar{x}\| &\leq \|y_x - x\| + \|x - \bar{x}\| \\ &\leq d^J(x, \Omega) + \|x - \bar{x}\| - J(y_x) + \|x - \bar{x}\| \\ &\leq \|x - \bar{x}\| + J(\bar{x}) + \|x - \bar{x}\| - J(y_x) + \|x - \bar{x}\| \\ &\leq |J(y_x) - J(\bar{x})| + 3\|x - \bar{x}\| \\ &\leq \ell\|y_x - \bar{x}\| + 3\|x - \bar{x}\|,\end{aligned}$$

which can be rearranged to give

$$\|y_x - \bar{x}\| \leq \left(\frac{3}{1 - \ell}\right)\|x - \bar{x}\|. \tag{5.23}$$

Since $\bar{x} \in S(\Omega)$ and $x^* \in \boldsymbol{B}_{X^*}$, $d^J(\bar{x},\Omega) = J(\bar{x})$ and $\|x^*\| \leq 1$. Owing to estimates (5.19), (5.21) and (5.23), one arrives at

$$\begin{aligned}
&\langle x^*, x - \bar{x}\rangle \\
&= \langle x^*, x - y_x\rangle + \langle x^*, y_x - \bar{x}\rangle \\
&\leq \|x^*\|\|x - y_x\| + (J(y_x) - J(\bar{x}) + \eta\|y_x - \bar{x}\|^2) \\
&\leq \|x - y_x\| + \left(J(y_x) - J(\bar{x}) + \left(\frac{9\eta}{(1-\ell)^2}\right)\|x - \bar{x}\|^2\right) \\
&\leq \left(d^J(x,\Omega) + \|x - \bar{x}\|^2 - J(y_x)\right) + \left(J(y_x) - d^J(\bar{x},\Omega) + \left(\frac{9\eta}{(1-\ell)^2}\right)\|x - \bar{x}\|^2\right) \\
&= d^J(x,\Omega) - d^J(\bar{x},\Omega) + \left(1 + \frac{9\eta}{(1-\ell)^2}\right)\|x - \bar{x}\|^2.
\end{aligned}$$

On the other hand, for $x = \bar{x}$, equality trivially holds in

$$\langle x^*, x - \bar{x}\rangle \leq d^J(x,\Omega) - d^J(\bar{x},\Omega) + \left(1 + \frac{9\eta}{(1-\ell)^2}\right)\|x - \bar{x}\|^2.$$

With the inequality valid for all $x \in \boldsymbol{B}_X(\bar{x}, r)$, by definition, $x^* \in \partial^p d^J(\bar{x},\Omega)$. This justifies

$$\partial^p d^J(\bar{x},\Omega) \supset \partial^p(J + \delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}. \qquad \square$$

While the perturbed minimization problem at $\bar{x}$ on Ω is required to be well-posed in both Theorem 5.3.4 and Theorem 5.3.5, verifying whether this requisite is fulfilled is in general not an easy task. In this light, it is instructive to have a simple sufficient condition to guarantee the well-posedness of the perturbed minimization problem.

Lemma 5.3.6. (cf. [37, Lemma 3.4]) *Let $\Omega \subset X$, $\bar{x} \in S(\Omega)$ and $J : \Omega \to \mathbb{R}$ be a lower semicontinuous function. Suppose J is centrally Lipschitz at $\bar{x}$ on Ω with rank $0 \leq \ell < 1$. Then the perturbed minimization problem at $\bar{x}$ on Ω is well-posed.*

Proof. Since $\bar{x} \in S(\Omega)$, it is a self-solution to the perturbed minimization problem and there holds $J(\bar{x}) = d^J(\bar{x},\Omega)$. Let $\{x_k\}_{k=1}^{\infty} \subset \Omega$ be any minimizing sequence of the perturbed minimization problem at $\bar{x}$ on Ω. Since J is centrally Lipschitz at $\bar{x}$ on Ω

with rank ℓ, $|J(x_k) - J(\bar{x})| \leq \ell \|x_k - \bar{x}\|$ for all $k \in \mathbb{N}$. Observe that

$$\begin{aligned}\|x_k - \bar{x}\| + J(\bar{x}) &= (\|x_k - \bar{x}\| + J(x_k)) + (J(\bar{x}) - J(x_k)) \\ &\leq (\|x_k - \bar{x}\| + J(x_k)) + \ell\|x_k - \bar{x}\|,\end{aligned}$$

which, upon rearrangement and in view of the assumption $0 \leq \ell < 1$, produces

$$J(\bar{x}) \leq (1-\ell)\|x_k - \bar{x}\| + J(\bar{x}) \leq \|x_k - \bar{x}\| + J(x_k). \tag{5.24}$$

As a minimizing sequence, $\{x_k\}_{k=1}^{\infty}$ satisfies $\lim_{k\to\infty}(\|x_k - \bar{x}\| + J(x_k)) = d^J(\bar{x}, \Omega) = J(\bar{x})$. Letting $k \to \infty$ in (5.24), one has

$$(1-\ell)\lim_{k\to\infty}\|x_k - \bar{x}\| + J(\bar{x}) = \lim_{k\to\infty}((1-\ell)\|x_k - \bar{x}\| + J(\bar{x})) = J(\bar{x}).$$

It follows that $\lim_{k\to\infty}\|x_k - \bar{x}\| = 0$, or equivalently, $x_k \to \bar{x}$. Consequently, any minimizing sequence of the perturbed minimization problem converges to $\bar{x}$, which, by Remark 5.1.7, implies that $\bar{x}$ is the only solution to the perturbed minimization problem. Hence the criterion for well-posedness of the perturbed minimization problem at $\bar{x}$ on Ω is satisfied. □

Below is an immediate consequence of Theorem 5.3.4 and Theorem 5.3.5, which is readily reducible as a special case to a familiar result (see [9]) concerning the standard distance function.

Corollary 5.3.7. (cf. [37, Corollary 3.2]) *Let $\Omega \subset X$, $\bar{x} \in S(\Omega)$ and $J : \Omega \to \mathbb{R}$ be a lower semicontinuous function. Suppose J is centrally Lipschitz at $\bar{x}$ on Ω with rank $0 \leq \ell < 1$. Then*

$$\partial^{\bullet} d^J(\bar{x}, \Omega) = \partial^{\bullet}(J + \delta_{\Omega})(\bar{x}) \cap \boldsymbol{B}_{X^*},$$

where $\partial^{\bullet}$ stands for $\widehat{\partial}$ or ∂^p.

Proof. Since $\bar{x} \in S(\Omega)$ and J is centrally Lipschitz at $\bar{x}$ on Ω with rank $0 \leq \ell < 1$, Lemma 5.3.6 implies the well-posedness of the perturbed minimization problem at $\bar{x}$ on Ω. Moreover, in view of the central Lipschitz assumption, J is automatically locally centrally Lipschitz at $\bar{x}$ on Ω with $\ell_{\bar{x}} \leq \ell < 1$. Taking $\varepsilon = 0$ in Theorem 5.3.4(a) and (b), one obtains $\widehat{\partial} d^J(\bar{x}, \Omega) = \widehat{\partial}(J + \delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}$. Moreover, the inclusions in Theorem 5.3.5(a) and (b) give $\partial^p d^J(\bar{x}, \Omega) = \partial^p(J + \delta_\Omega)(\bar{x}) \cap \boldsymbol{B}_{X^*}$ immediately. □

Corollary 5.3.8. *Let $\Omega \subset X$ and $\bar{x} \in \Omega$. Then*

$$\partial^\bullet d(\bar{x}, \Omega) = N^\bullet(\bar{x}; \Omega) \cap \boldsymbol{B}_{X^*},$$

where $(\partial^\bullet, N^\bullet)$ stands for $(\widehat{\partial}, \widehat{N})$ or (∂^p, N^p).

Proof. Let $J \equiv 0$. Then J is trivially lower semicontinuous and centrally Lipschitz at $\bar{x}$ on Ω with arbitrary rank $\ell \geq 0$. Moreover, $d^J(\cdot, \Omega)$ reduces to $d(\cdot, \Omega)$. Note that $S(\Omega) = \{x \in \Omega : d(x, \Omega) = 0\} = \Omega$ and hence $\bar{x} \in S(\Omega)$. Applying Corollary 5.3.7 gives

$$\widehat{\partial} d(\bar{x}, \Omega) = \widehat{\partial}\delta_\Omega(\bar{x}) \cap \boldsymbol{B}_{X^*} = \widehat{N}(\bar{x}; \Omega) \cap \boldsymbol{B}_{X^*}, \text{ and}$$

$$\partial^p d(\bar{x}, \Omega) = \partial^p \delta_\Omega(\bar{x}) \cap \boldsymbol{B}_{X^*} = N^p(\bar{x}; \Omega) \cap \boldsymbol{B}_{X^*}.$$

□

Bibliography

[1] Baranger, J. Existence de solution pour des problemes d'optimisation nonconvexe. *C. R. Acad. Sci. Paris*, **274**, 307-309, 1972.

[2] Bidaut, M. F. Existence theorems for usual and approximate solutions of optimal control problems. *J. Optim. Theory Appl.*, **15**, 393-411, 1975.

[3] Borwein, J. M., Fitzpatrick, S. P., Giles, J. R. The differentiability of real functions on normed linear spaces using generalized subgradients. *J. Math. Anal. Appl.*, **128**, 512-534, 1987.

[4] Borwein, J. M., Giles, J. R. The proximal normal formula in Banach space. *Trans. Amer. Math. Soc.*, **302**, 371-381, 1987.

[5] Borwein, J. M., Lewis, A. S. *Convex Analysis and Nonlinear Optimization: Theory and Examples.* Springer, New York, 2006.

[6] Bounkhel, M. On the distance function associated with a set-valued mapping. *J. Nonlinear Convex Anal.*, **2**, 265-278, 2001.

[7] Bounkhel, M. Scalarization of the normal Fréchet regularity of set-valued mappings. *New Zealand J. Math.*, **33**, 129-146, 2004.

[8] Bounkhel, M., Jofré, A. Subdifferential stability of the distance function and its applications to nonconvex economies and equilibrium. *J. Nonlinear Convex Anal.*, **5**, 331-347, 2004.

[9] Bounkhel, M., Thibault, L. On various notions of regularity of sets in nonsmooth analysis. *Nonlinear Anal.*, **48**, 223-246, 2002.

[10] Burke, J. V. An exact penalization viewpoint of constrained optimization. Technical Report ANL/MCS-TM-95, Mathematics and Computer Science Division, Argonne National Laboratories, Argonne, IL 60439, 1987.

[11] Burke, J. V. A sequential quadratic programming method for potentially infeasible mathematical programs. *J. Math. Anal. Appl.*, **139**, 319-351, 1989.

[12] Burke, J. V., Ferris, M. C., Qian, M. On the Clarke subdifferential of the distance function of a closed set. *J. Math. Anal. Appl.*, **166**, 199-213, 1992.

[13] Burke, J. V., Han, S. P. A Gauss-Newton approach to solving generalized inequalities. *Math. Oper. Res.*, **11**, 632-643, 1986.

[14] Burke, J. V., Poliquin, R. A. Optimality conditions for non-finite valued convex composite functions. *Math. Progr.*, **57**, 103-120, 1992.

[15] Clarke, F. H. *Optimization and Nonsmooth Analysis.* Wiley, New York, 1983.

[16] Cobzas, S. Nonconvex optimization problems on weakly compact subsets of Banach spaces. *Anal. Numér. Théor. Approx.*, **9**, 19-25, 1980.

[17] Dontchev, A. L., Zolezzi, T. *Well-Posed Optimization Problems.* Springer, Berlin, 1993.

[18] Jourani, A., Thibault, L. Metric regularity and subdifferential calculus in Banach spaces. *Set-Valued Anal.*, **3**, 87-100, 1995.

[19] Kruger, A. Y., Mordukhovich, B. S. Extremal points and the Euler equation in nonsmooth optimization. *Dokl. Akad. Nauk BSSR*, **24**, 684-687, 1980.

[20] Lebourg, G. Perturbed optimization problems in Banach spaces. *Bull. Soc. Math. Fr.*, **60**, 95-111, 1979.

[21] Li, C., Peng, L. H. Porosity of perturbed optimization problems in Banach spaces. *J. Math. Anal. Appl.*, **324**, 751-761, 2006.

[22] Megginson, R. E. *An Introduction to Banach Space Theory.* Springer, New York, 1988.

[23] Mordukhovich, B. S. Maximum principle in problems of time optimal control with nonsmooth constraints. *J. Appl. Math. Mech.*, **40**, 960-969, 1976.

[24] Mordukhovich, B. S. Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions. *Trans. Amer. Math. Soc.*, **340**, 1-35, 1993.

[25] Mordukhovich, B. S. *Variational Analysis and Generalized Differentiation, Vol. I: Basic Theory; Vol. II: Applications.* Springer, Berlin, 2006.

[26] Mordukhovich, B. S., Nam, N. M. Subgradients of distance functions with applications to Lipschitzian stability. *Math. Progr. Ser. B*, **104**, 635-668, 2005.

[27] Mordukhovich, B. S., Nam, N. M. Subgradients of distance functions at out-of-set points. *Taiwan. J. Math.*, **10**, 299-326, 2006.

[28] Peng, L. H., Li, C., Yao, J. C. Well-posedness of a class of perturbed optimization problems in Banach spaces. *J. Math. Anal. Appl.*, **346**, 384-394, 2008.

[29] Penot, J.-P., Ratsimahalo, R. Subdifferentials of distance functions, approximations and enlargements. *Acta Math. Sinica*, **23**, 507-520, 2007.

[30] Phelps, R. R. *Convex Functions, Monotone Operators and Differentiability.* Springer, Berlin, 1993.

[31] Rockafellar, R. T. *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1970.

[32] Rockafellar, R. T. Lipschitzian properties of multifunctions. *Nonlinear Anal.*, **9**, 867-885, 1985.

[33] Rockafellar, R. T., Wets, R. J.-B. *Variational Analysis.* Springer, Berlin, 1998.

[34] Schirotzek, W. *Nonsmooth Analysis.* Springer, Berlin, 2007.

[35] Van Tiel, J. *Convex Analysis: An Introductory Text.* Wiley, New York, 1984.

[36] Thibault, L. On subdifferentials of optimal value functions. *SIAM J. Control Optim.*, **29**, 1019-1036, 1991.

[37] Wang, J. H., Li, C., Xu, H. K. Subdifferentials of perturbed distance functions in Banach spaces. *J. Glob. Optim.*, **46**, 489-501, 2010.

[38] Wu, Z., Ye, J. J. Equivalence among various derivatives and subdifferentials of the distance function. *J. Math. Anal. Appl.*, **282**, 629-647, 2003.

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