5-1-2005

Fréchet Subdifferential Calculus and Optimality Conditions in Nondifferentiable Programming

Boris S. Mordukhovich
Wayne State University, boris@math.wayne.edu

Nguyen Mau Nam
Wayne State University

N. D. Yen
Institute of Mathematics, Hanoi, Vietnam

FRÉCHET SUBDIFFERENTIAL CALCULUS AND OPTIMALITY CONDITIONS IN NONDIFFERENTIABLE PROGRAMMING

B. S. MORDUKHOVICH, N. M. NAM, and N. D. YEN

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics
Research Report

2005 Series
#5

This research was partly supported by the National Science Foundation and the Australian Research Council

FRÉCHET SUBDIFFERENTIAL CALCULUS AND OPTIMALITY CONDITIONS IN NONDIFFERENTIABLE PROGRAMMING [1]

B. S. MORDUKHOVICH [2], N. M. NAM [3], N. D. YEN [4]

Dedicated to Diethard Pallaschke in honor of his 65th birthday

We develop various (exact) calculus rules for Fréchet lower and upper subgradients of extended-real-valued functions in general Banach spaces. Then we apply this calculus to derive new necessary optimality conditions for some remarkable classes of problems in constrained optimization including minimization problems for difference-type functions under geometric and operator constraints as well as subdifferential optimality conditions for the so-called weak sharp minima.

Keywords: Variational analysis and constrained optimization; Fréchet normals, subgradients, and coderivatives; Subdifferential calculus; Necessary optimality conditions; Difference-type functions; Geometric and operator constraints; Weak sharp minima

Mathematical Subject Classification 2000: Primary: 49J52, 49J53; Secondary: 90C30

1 Introduction

This paper is devoted to the study of generalized differentiation and its applications to the first-order necessary optimality conditions for constrained optimization problems in infinite-dimensional spaces. Let $\varphi\colon X \to \overline{I\!R} := [-\infty, \infty]$ be an extended-real-valued function on a Banach space X. One of the most simple derivative-like objects, the so-called *Fréchet subdifferential* of φ at a domain point $\bar{x}$ with $|\varphi(\bar{x})| < \infty$, is defined by

$$\widehat{\partial}\varphi(\bar{x}) := \Big\{x^* \in X^* \Big| \liminf_{x\to\bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle x^*, x - \bar{x}\rangle}{||x - \bar{x}||} \geq 0\Big\}. \tag{1.1}$$

When "lim inf" is replaced by "lim" and the inequality "≥ 0" is replaced by the equality "$= 0$" in (1.1), we have the classical definition of the Fréchet derivative/gradient of φ at $\bar{x}$; that is where the name of "Fréchet subgradients" comes from.

It follows directly from definition (1.1) of Fréchet subgradients $x^* \in \widehat{\partial}\varphi(\bar{x})$ that the following *generalized Fermat rule* holds:

$$0 \in \widehat{\partial}\varphi(\bar{x}) \tag{1.2}$$

[1]Research was partially supported by the National Science Foundation under grant DMS-0304989 and by the Australian Research Council under grant DP-0451168.

[2]Department of Mathematics, Wayne State University, Detroit, Michigan 48202, USA; boris@math.wayne.edu

[3]Department of Mathematics, Wayne State University, Detroit, Michigan 48202, USA; nam@math.wayne.edu

[4]Institute of Mathematics, 18 Hoang Quoc Viet Rd., 10307 Hanoi, Vietnam; ndyen@math.ac.vn

whenever $\bar{x}$ is a local minimizer for $\varphi\colon X \to \overline{I\!R}$. Considering now the general *constrained minimization* problem in a Banach space:

$$\text{minimize } \varphi(x) \text{ subject to } x \in \Omega \subset X, \tag{1.3}$$

we equivalently reduce it to the unconstrained problem:

$$\text{minimize } \varphi(x) + \delta(x;\Omega), \quad x \in X, \tag{1.4}$$

involving the *indicator function* $\delta(\cdot;\Omega)$ of the set Ω that equals 0 if $x \in \Omega$ and ∞ otherwise. Applying Fermat's rule (1.2), we get

$$0 \in \widehat{\partial}\big[\varphi + \delta(\cdot;\Omega)\big](\bar{x}),$$

whenever $\bar{x}$ is a local solution to the constrained problem (1.3). To proceed further in constrained optimization and obtain valuable optimality conditions in terms of the initial data, we need to have satisfactory *calculus rules* for Fréchet subgradients, which is generally *not* the case. In particular, the desirable sum rule

$$\widehat{\partial}(\varphi_1 + \varphi_2)(\bar{x}) \subset \widehat{\partial}\varphi_1(\bar{x}) + \widehat{\partial}\varphi_2(\bar{x}) \tag{1.5}$$

does not hold even in the simplest nonsmooth settings, e.g., for the functions $\varphi_1(x) = |x|$ and $\varphi_2(x) = -|x|$ on the real line. On the other hand, Fréchet subgradients satisfy the so-called "fuzzy calculus" under natural conditions involving a broad class of Banach (namely Asplund, particularly reflexive) spaces; see the recent books by Borwein and Zhu [1] and Mordukhovich [4] with the references therein. However, such fuzzy calculus rules, which allow us to approximately represent Fréchet subgradients of sums and other compositions at points of interest via Fréchet subgradients of separate functions at points nearby, are not very useful for a number of applications including necessary optimality conditions in constrained optimization. "Exact" calculus rules dealing only with points of interest (e.g., with minimizers) are essentially more desirable for the majority of applications. Unfortunately, such an exact calculus is not available for Fréchet subgradients in reasonably general settings, which significantly restricts the scope of their applications.

The primary goal of this paper is to develop *exact calculus rules* for Fréchet subgradients and related constructions in arbitrary *Banach spaces*. Surprisingly, it can be done for a variety of calculus rules under certain assumptions mostly related to the nonemptiness of the Fréchet subdifferential or its upper (superdifferential) counterpart for some of the functions involved in compositions. The new calculus results obtained in this paper particularly include chain, difference, product, and quotient rules in a rather surprising generality.

The main tool of our analysis is a *smooth variational description* of Fréchet subgradients in general Banach spaces; see the next section. This description allows us to establish, instead of the sum rule (1.5) that is not true, the following striking *difference rule*

$$\widehat{\partial}(\varphi_1 - \varphi_2)(\bar{x}) \subset \bigcap_{x^* \in \widehat{\partial}\varphi_2(\bar{x})} \big[\widehat{\partial}\varphi_1(\bar{x}) - x^*\big] \subset \widehat{\partial}\varphi_1(\bar{x}) - \widehat{\partial}\varphi_2(\bar{x}), \tag{1.6}$$

provided that $\widehat{\partial}\varphi_2(\bar{x}) \neq \emptyset$. Due to the stationary principle (1.2), the difference rule (1.6) immediately implies that

$$\widehat{\partial}\varphi_2(\bar{x}) \subset \widehat{\partial}\varphi_1(\bar{x}) \tag{1.7}$$

at a local minimum point $\bar{x}$ for the difference $\varphi = \varphi_1 - \varphi_2$ of general extended-real-valued functions in Banach spaces. For the case of convex functions φ_i, their difference $\varphi_1 - \varphi_2$ is called a DC-function (i.e., the difference of convex functions), and the necessary optimality condition (1.7) was first established probably by Hiriart-Urruty [3]. By now the class of DC-functions has become an attractive object useful in many aspects of optimization theory and applications. The reader can find more information on such functions and related topics in, e.g., Pallaschke and Urbański [6], Penot [7], and the references therein. Note that the latter paper contains necessary and sufficient conditions of type (1.7) and their second-order developments in both convex and nonconvex settings.

The difference rule and its consequence (1.7) happen to be useful to derive efficient necessary conditions for *sharp minimizers* in the sense of Polyak [9] and in some other settings. Further results for minimizing of difference functions subject to general geometric and operator/functional constraints are obtained in this paper by combining calculus rules for Fréchet subgradients with a more developed calculus of basic/limiting subgradients by Mordukhovich [4] in the framework of Asplund spaces.

The rest of the paper is organized as follows. In Section 2 we present basic definitions and preliminaries, which are widely used in the sequel. Section 3 contains new calculus rules for Fréchet subgradients of extended-real-valued functions and the corresponding coderivatives of set-valued mappings. Section 4 is devoted to applications of these results and also calculus rules for limiting subgradients to deriving necessary optimality conditions in various nonsmooth problems of constrained *difference programming*, i.e., optimization problems whose cost functions are given as a difference of some nonsmooth functions.

Our notation is basically standard; see the books by Rockafellar and Wets [11] and by Mordukhovich [4]. All the spaces are supposed to be Banach unless otherwise stated.

2 Basic Definitions and Preliminaries

Given an extended-real-valued function $\varphi\colon X \to \overline{I\!R}$, the *Fréchet subdifferential*, i.e., the collection of Fréchet subgradients of φ at $\bar{x} \in \operatorname{dom}\varphi$, was defined in (1.1). Sometimes this construction is called "regular subdifferential" as in Rockafellar and Wets [11]. Furthermore, (1.1) agrees with the Crandall-Lions subdifferential of φ at $\bar{x}$ from the theory of viscosity solutions to nonlinear partial differential equations.

A characteristic feature of Fréchet subgradients is the following *smooth variational description* held in *any Banach* space; see [4, Theorem 1.88(i)].

Proposition 2.1 (smooth variational description of Fréchet subgradients). *Let $\varphi\colon X \to \overline{I\!R}$ be finite at $\bar{x}$. Then $x^* \in \widehat{\partial}\varphi(\bar{x})$ if and only if there are a neighborhood U of $\bar{x}$ and a function $s\colon U \to I\!R$, which is Fréchet differentiable at $\bar{x}$ with the derivative $\nabla s(\bar{x})$ and such that*

$$s(\bar{x}) = \varphi(\bar{x}), \quad \nabla s(\bar{x}) = x^*, \quad \text{and} \quad s(x) \le \varphi(x) \quad \text{for all} \quad x \in U.$$

On the other words, every Fréchet subgradient of an arbitrary function φ at $\bar{x}$ can be equivalently described via the classical Fréchet derivative of another function s such that the difference $\varphi - s$ attains its local minimum at $\bar{x}$. In [4, Theorem 1.88(ii,iii)], the reader can find more involved smooth variational descriptions of Fréchet subgradients of functions defined on "smooth" Banach spaces (i.e., those admitting an equivalent smooth renorm or bump function), but they are not used in this paper.

Along with Fréchet subgradients (which are sometimes called *lower* subgradients, while usually "lower" is taken for granted), we need to employ in what follows an upper counterpart of (1.1) called the *Fréchet upper subdifferential* (sometimes "Fréchet superdifferential") of $\varphi\colon X \to \overline{I\!R}$ at $\bar{x}$ with $|\varphi(\bar{x})| < \infty$ and defined by

$$\widehat{\partial}^+\varphi(\bar{x}) := -\widehat{\partial}(-\varphi)(\bar{x}). \tag{2.1}$$

It is easy to observe from the definitions that if the sets $\widehat{\partial}\varphi(\bar{x})$ and $\widehat{\partial}^+\varphi(\bar{x})$ are nonempty simultaneously, then φ is Fréchet differentiable at $\bar{x}$ with

$$\widehat{\partial}^+\varphi(\bar{x}) = \widehat{\partial}\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}.$$

Due to the symmetry in (2.1), properties of (lower) Fréchet subgradients automatically imply those for the upper counterpart; so it is sufficient to consider only the lower version. Let us emphasize, however, that in some situations we need to use both constructions (1.1) and (2.1), although for different functions involved in calculus rules and optimality conditions—this is one of the *strongest points* made in this paper!

Another useful (while elementary) property of Fréchet subgradients is the following sum rule for two functions on a Banach space when one of these functions is Fréchet differentiable at the point in question; see, e.g., [4, Proposition 1.107(i)].

Proposition 2.2 (sum rule with a differentiable function). *Let $\varphi_1\colon X \to \overline{I\!R}$ be finite at $\bar{x}$, and let $\varphi_2\colon X \to \overline{I\!R}$ be Fréchet differentiable at $\bar{x}$. Then*

$$\widehat{\partial}(\varphi_1 + \varphi_2)(\bar{x}) = \widehat{\partial}\varphi_1(\bar{x}) + \nabla\varphi_2(\bar{x}).$$

As mentioned, a desirable sum rule of the inclusion (not even talking about equality) type (1.5) does not hold for Fréchet subgradients of sums involving both nonsmooth functions. Such a desirable/full calculus is available for the so-called basic (or limiting) subgradients, which can be viewed as a *sequential robust regularization* of Fréchet subgradients and its ε-enlargements. A comprehensive theory and applications of the latter subgradient and associated normal and coderivative constructions are developed in the books by Mordukhovich [4, 5], which are widely used in what follows. We also refer the reader to the book by Rockafellar and Wets [11] for related developments in finite-dimensional spaces.

To describe limiting subgradients for extended-real-valued functions on Banach spaces, we first need to recall an appropriate ε-enlargement of the Fréchet subdifferential known as the *ε-subdifferential* (or the collection of ε-subgradients) of $\varphi\colon X \to \overline{I\!R}$ at $\bar{x}$ with $|\varphi(\bar{x})| < \infty$. It is defined by

$$\widehat{\partial}_\varepsilon\varphi(\bar{x}) := \Big\{x^* \in X^* \Big| \liminf_{x\to\bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \geq -\varepsilon\Big\}, \quad \varepsilon \geq 0, \tag{2.2}$$

where $\widehat{\partial}_0\varphi(\bar{x}) := \widehat{\partial}\varphi(\bar{x})$. Then the *basic/limiting subdifferential* of φ at $\bar{x}$ is defined by

$$\partial\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{\substack{x\xrightarrow{\varphi}\bar{x}\\ \varepsilon\downarrow 0}} \widehat{\partial}_\varepsilon\varphi(x), \tag{2.3}$$

where "Lim sup" stands for the *sequential Painlevé-Kuratowski upper/outer limit* of a set-valued mapping $F\colon X\times[0,\infty)\rightrightarrows X^*$ given by $F(x,\varepsilon) := \widehat{\partial}_\varepsilon\varphi(x)$ with respect to the norm topology of X and weak* topology of X^*, and where $x\xrightarrow{\varphi}\bar{x}$ means that $x\to\bar{x}$ with $\varphi(x)\to\varphi(\bar{x})$; see [4]. Note that, by [4, Theorem 2.34], we can equivalently put $\varepsilon=0$ in (2.3) if φ is lower semicontinuous (l.s.c.) around $\bar{x}$ and if the space X is *Asplund*, i.e., any of its separable subspaces has a separable dual; the reader can find more information on Asplund space in [4] and/or in the book by Phelps [8].

It follows from (2.3) that $\widehat{\partial}\varphi(\bar{x})\subset\partial\varphi(\bar{x})$. A function $\varphi\colon X\to\overline{I\!R}$ is called *lower regular* at $\bar{x}$ with $|\varphi(\bar{x})|<\infty$ if

$$\partial\varphi(\bar{x}) = \widehat{\partial}\varphi(\bar{x}). \tag{2.4}$$

Besides the classical cases of convex functions and those strictly differentiable at $\bar{x}$ (in particular, smooth functions), the latter class includes substantially broader collections of functions encountered in variational analysis and optimization; see [4, 11]. If φ is lower regular at $\bar{x}$ and locally Lipschitzian around this point, then $\widehat{\partial}\varphi(\bar{x})\neq\emptyset$ provided that X is Asplund. This is due to $\partial\varphi(\bar{x})\neq\emptyset$ under the above conditions; see [4, Corollary 2.25].

Given a set $\Omega\subset X$ and a point $\bar{x}\in\Omega$, we define the *Fréchet normal cone* (or prenormal cone) $\widehat{N}(\bar{x};\Omega)$ and the *basic/limiting normal cone* $N(\bar{x};\Omega)$ to Ω at $\bar{x}$ by, respectively,

$$\widehat{N}(\bar{x};\Omega) := \widehat{\partial}\delta(\bar{x};\Omega) \;\text{ and }\; N(\bar{x};\Omega) := \partial\delta(\bar{x};\Omega) \tag{2.5}$$

via the corresponding subdifferential of the indicator function. A set Ω is called *normally regular* at $\bar{x}\in\Omega$ if $N(\bar{x};\Omega)=\widehat{N}(\bar{x};\Omega)$. Besides convex sets and those with smooth boundaries, this class includes also various sets that exhibit locally convex-like and other "nice" behavior; see [4, 11]. For some applications in this paper, we need the construction

$$\partial^\infty\varphi(\bar{x}) := \big\{x^*\in X^*\big|\,(x^*,0)\in N((\bar{x},\varphi(\bar{x}));\operatorname{epi}\varphi)\big\} \tag{2.6}$$

of the *singular subdifferential* of $\varphi\colon X\to\overline{I\!R}$ at $\bar{x}$ with $|\varphi(\bar{x})|<\infty$ defined via the limiting normal cone (2.5) to the epigraph

$$\operatorname{epi}\varphi := \big\{(x,\mu)\in X\times I\!R\big|\,\mu\geq\varphi(x)\big\}.$$

Note that $\partial^\infty\varphi(\bar{x})=\{0\}$ if φ is locally Lipschitzian around $\bar{x}$; see [4, Corollary 1.81].

Given a single-valued mapping $f\colon X\to Y$ between Banach spaces with the graph

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

and a point $(\bar{x},f(\bar{x}))\in\operatorname{gph} f$, we need the following two coderivative constructions

$$\widehat{D}^*f(\bar{x})(y^*) := \big\{x^*\in X^*\big|\,(x^*,-y^*)\in\widehat{N}((\bar{x},f(\bar{x}));\operatorname{gph} f)\big\}, \tag{2.7}$$

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

called, respectively, the *Fréchet coderivative* and the *limiting coderivative* of f at $\bar{x}$. Note that the coderivative (2.8) is called "normal coderivative" and is denoted by D^*_N in [4] to distinguish it from the other limiting coderivatives considered in [4]. Here we do not need to consider any other limiting coderivative but (2.8) and also to use coderivative constructions for set-valued mappings comprehensively studied in [4]. Observe that

$$\widehat{D}^*f(\bar{x})(y^*) = \big\{\nabla f(\bar{x})^*y^*\big\} \text{ and resp. } D^*f(\bar{x})(y^*) = \big\{\nabla f(\bar{x})^*y^*\big\}$$

if f is Fréchet differentiable and strictly differentiable at $\bar{x}$, respectively, where $\nabla f(\bar{x})$ stands for the corresponding derivative of a single-valued mapping.

Finally in this section, we recall two notions of the so-called "normal compactness" for sets and functions, respectively, used in the paper. Both properties are automatically fulfilled in finite dimensions; they concern relationships between weak* and strong convergences to 0 in dual spaces and hold for "reasonably good" sets and functions. Since they are employed in this paper only in the Asplund space setting, we present their simplified definitions equivalent in Asplund spaces to the general ones given in [4].

A set Ω is *sequentially normally compact* (SNC) at $\bar{x}$ if for any sequences $x_k \xrightarrow{\Omega} \bar{x}$ and $x_k^* \in \widehat{N}(x_k; \Omega)$ one has

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

where $x_k \xrightarrow{\Omega} \bar{x}$ stands for $x_k \to \bar{x}$ with $x_k \in \Omega$, and where w^* signifies the weak* topology of X^*. The SNC condition is automatic when Ω satisfies the so-called "compactly epi-Lipschitzian" property in the sense of Borwein and Strojwas, particularly when it is convex and finite-codimensional with nonempty relative interior; see [1, 4] for more details.

A function $\varphi\colon X \to \overline{I\!R}$ is *sequentially normally epi-compact* (SNEC) at a point $\bar{x}$ with $|\varphi(\bar{x})| < \infty$ if its epigraph is SNC at $(\bar{x}, \varphi(\bar{x}))$, which is equivalent (for l.s.c. functions on Asplund spaces) to

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

for any sequences $(x_k, x_k^*, \lambda_k) \in X \times X^* \times [0, \infty)$ with $x_k \xrightarrow{\Omega} \bar{x}$, $\lambda_k \downarrow 0$, and $x_k^* \in \lambda_k \widehat{\partial}\varphi(x_k)$; see [4, Corollary 2.39]. This property always holds for locally Lipschitzian functions and their appropriate extensions.

3 Calculus of Fréchet Subgradients and Coderivatives

In this section we collect new *exact calculus rules* for Fréchet subgradients of extended-real-valued functions on arbitrary Banach spaces. We also present some related calculus results for the coderivative associated with the Fréchet subdifferential.

Our first theorem provides a principal *difference rule* for (lower) Fréchet subgradients, which can be equivalently formulated as a sum rule for lower subgradients with the usage of *upper* subgradients of one the functions involved. As mentioned, the proof of this result is based on the *smooth variational description* of Fréchet subgradients for extended-real-valued functions in the general Banach space setting; see Proposition 2.1.

Theorem 3.1 (difference and sum rules for Fréchet subgradients). *Let $\varphi_i\colon X \to \overline{\mathbb{R}}$ be finite at $\bar{x}$ for $i = 1, 2$. The following assertions hold:*

(i) *Assume that $\widehat{\partial}\varphi_2(\bar{x}) \neq \emptyset$. Then*

$$\widehat{\partial}(\varphi_1 - \varphi_2)(\bar{x}) \subset \bigcap_{x^* \in \widehat{\partial}\varphi_2(\bar{x})} [\widehat{\partial}\varphi_1(\bar{x}) - x^*] \subset \widehat{\partial}\varphi_1(\bar{x}) - \widehat{\partial}\varphi_2(\bar{x}). \tag{3.1}$$

(ii) *Assume that $\widehat{\partial}^+\varphi_1(\bar{x}) \neq \emptyset$. Then*

$$\widehat{\partial}(\varphi_1 + \varphi_2)(\bar{x}) \subset \bigcap_{x^* \in \widehat{\partial}^+\varphi_1(\bar{x})} [x^* + \widehat{\partial}\varphi_2(\bar{x})] \subset \widehat{\partial}^+\varphi_1(\bar{x}) + \widehat{\partial}\varphi_2(\bar{x}).$$

Proof. To prove (3.1), fix any $u^* \in \widehat{\partial}(\varphi_1 - \varphi_2)(\bar{x})$ and $x^* \in \widehat{\partial}\varphi_2(\bar{x})$. Using the smooth variational description of Proposition 2.1 for the Fréchet subgradient $x^* \in \widehat{\partial}\varphi_2(\bar{x})$, we find a real-valued function $s(\cdot)$ defined in a neighborhood U of $\bar{x}$ such that $s(\cdot)$ is Fréchet differentiable at $\bar{x}$ satisfying the relations

$$s(\bar{x}) = \varphi_2(\bar{x}), \quad \nabla s(\bar{x}) = x^*, \quad \text{and} \quad s(x) \le \varphi_2(x) \quad \text{for} \quad x \in U. \tag{3.2}$$

Picking any $\varepsilon > 0$ and applying definition (1.1) to the Fréchet subgradient $u^* \in \widehat{\partial}(\varphi_1 - \varphi_2)(\bar{x})$, we find $\eta > 0$ such that

$$\begin{aligned} \langle u^*, x - \bar{x}\rangle &\le \varphi_1(x) - \varphi_2(x) - (\varphi_1(\bar{x}) - \varphi_2(\bar{x})) + \varepsilon\|x - \bar{x}\| \\ &\le \varphi_1(\bar{x}) - s(x) - (\varphi_1(\bar{x}) - s(\bar{x})) + \varepsilon\|x - \bar{x}\|, \end{aligned}$$

whenever $\|x - \bar{x}\| < \eta$. Using now (3.2) and Proposition 2.2, we get

$$u^* \in \widehat{\partial}(\varphi_1 - s)(\bar{x}) = \widehat{\partial}\varphi_1(\bar{x}) - \nabla s(\bar{x}) = \widehat{\partial}\varphi_1(\bar{x}) - x^*,$$

which justifies the difference rule (3.1).

To prove the sum rule in (ii), it suffices to denote $\psi(x) := -\varphi_1(x)$ and to apply the difference rule (3.1) to the representation $\varphi_1 + \varphi_2 = \varphi_2 - \psi$. $\triangle$

By Proposition 2.2 the inclusions in (i) and (ii) of Theorem 3.1 become equalities if one of the functions is Fréchet differentiable at $\bar{x}$. Equalities hold in other (fully nonsmooth) cases as well. For example, one can check that for

$$\varphi_1(x) := |x| \quad \text{and} \quad \varphi_2(x) := \sqrt{|x|}, \quad x \in \mathbb{R},$$

we have $\widehat{\partial}\varphi_1(0) = [-1, 1]$, $\widehat{\partial}\varphi_2(0) = \mathbb{R}$, and $\widehat{\partial}(\varphi_1 - \varphi_2)(0) = \mathbb{R}$, i.e, the equality holds in (3.1). Observe also that the assumption on $\widehat{\partial}\varphi_2(\bar{x}) \neq \emptyset$ is essential for the fulfillment of (3.1). Example: $\varphi_1(x) := |x|$ and $\varphi_2(x) := -|x|$ with $\bar{x} = 0$.

Let us present two consequences of Theorem 3.1. The first one concerns the so-called *DC-functions*, i.e., differences of convex functions, whose importance has been well recognized in optimization theory and applications; see, e.g., [3, 6, 7]. We keep the notation ∂ for the subdifferential of convex analysis with which both Fréchet and limiting subdifferentials agree in the case of convex functions.

Corollary 3.2 (subgradients of DC-functions). *Let $\varphi := \varphi_1 - \varphi_2$ be a DC-function, where φ_2 is continuous at $\bar{x}$. Then*

$$\widehat{\partial}\varphi(\bar{x}) = \widehat{\partial}(\varphi_1 - \varphi_2)(\bar{x}) \subset \bigcap_{x^* \in \partial\varphi_2(\bar{x})} \big[\partial\varphi_1(\bar{x}) - x^*\big].$$

Proof. It is well known from convex analysis that the subdifferential of every convex function is nonempty at a point of continuity; see, e.g., [8, Proposition 1.11]. $\triangle$

Observe that the continuity requirement on φ_2 in Corollary 3.2 can be relaxed to the significantly weaker one $\bar{x} \in \operatorname{ri}(\operatorname{dom}\varphi_2)$ if $X = \mathbb{R}^n$. This follows from the classical subdifferentiability result of finite-dimensional convex analysis; see [10, Theorem 23.4].

The second corollary of Theorem 3.1 gives a new and useful difference rule for the *limiting subgradients* (2.3) in the the framework of Asplund spaces.

Corollary 3.3 (difference rule for limiting subgradients). *Let X be Asplund, let φ_1 be l.s.c. around $\bar{x}$, and let φ_2 be continuous at $\bar{x}$ and such that $\widehat{\partial}\varphi_2(x)$ is nonempty and uniformly bounded around this point; all the assumptions on φ_2 are automatic when φ_2 is convex and continuous around $\bar{x}$. Then*

$$\partial(\varphi_1 - \varphi_2)(\bar{x}) \subset \partial\varphi_1(\bar{x}) - \partial\varphi_2(\bar{x}). \tag{3.3}$$

Proof. Pick any $x^* \in \partial(\varphi_1 - \varphi_2)(\bar{x})$ and, by the representation of basic/limiting subgradients of l.s.c. functions on Asplund spaces from [4, Theorem 2.34], find sequences $x_k \to \bar{x}$ and $x_k^* \xrightarrow{w^*} x^*$ satisfying $\varphi_1(x_k) - \varphi_2(x_k) \to \varphi_1(\bar{x}) - \varphi_2(\bar{x})$ and

$$x_k^* \in \widehat{\partial}(\varphi_1 - \varphi_2)(x_k) \ \text{ for all } \ k \in \mathbb{N} := \{1, 2, \ldots\}. \tag{3.4}$$

Applying Theorem 3.1(i) in (3.4), we select sequences $x_{ik}^* \in \widehat{\partial}\varphi_i(x_k)$ as $i = 1, 2$ such that

$$x_k^* = x_{1k}^* - x_{2k}^* \ \text{ for all } \ k \in \mathbb{N}. \tag{3.5}$$

Since the sets $\widehat{\partial}\varphi_2(x)$ are uniformly bounded in the dual space to an Asplund space, we assume without loss of generality that the sequence $\{x_{2k}^*\}$ weak* converges to some $x_2^* \in X^*$. The continuity of φ_2 gives $\varphi_2(x_k) \to \varphi_2(\bar{x})$, and hence $\varphi_1(x_k) \to \varphi_1(\bar{x})$ as $k \to \infty$ by the above choice of $\{x_k\}$. By definition of the limiting subdifferential, this immediately implies that $x_2^* \in \partial\varphi_2(\bar{x})$.

Furthermore, it follows from (3.5) that the sequence $\{x_{1k}^*\}$ weak* converges to some $x_1^* \in X^*$, which must belong to $\partial\varphi_1(\bar{x})$ by the discussions above. Passing to the limit in (3.5), we conclude that $x^* \in \partial\varphi_1(\bar{x}) - \partial\varphi_2(\bar{x})$ and complete the proof. $\triangle$

Remark 3.4 (difference rule for proximal subgradients). Similar considerations allow us to establish analogs of Theorem 3.1 and related results presented below for the so-called *proximal subgradients* of $\varphi\colon X \to \overline{\mathbb{R}}$ defined on Banach spaces by: $x^* \in \partial_P\varphi(\bar{x})$ if only if there are positive numbers σ and η such that

$$\langle x^*, x - \bar{x}\rangle \le \varphi(x) - \varphi(\bar{x}) + \sigma\|x - \bar{x}\|^2 \ \text{ whenever } \ \|x - \bar{x}\| \le \eta.$$

The proximal counterpart of the difference rule (3.1) is formulated as:

$$\partial_P(\varphi_1-\varphi_2)(\bar{x})\subset\bigcap_{x^*\in\partial_P\varphi_2(\bar{x})}[\partial_P\varphi_1(\bar{x})-x^*]\subset\partial_P\varphi_1(\bar{x})-\partial_P\varphi_2(\bar{x})$$

provided that $\partial_P\varphi_2(\bar{x})\neq\emptyset$.

Given a single-valued mapping $f\colon X\to Y$ between Banach spaces, we consider its *scalarization* defined by

$$\langle y^*,f\rangle(x):=\langle y^*,f(x)\rangle,\quad x\in X,$$

for any $y^*\in Y^*$. The next proposition gives a useful relationship between the Fréchet coderivative (2.7) of Lipschitzian mappings and Fréchet subgradients of their scalarization.

Proposition 3.5 (scalarization formula). *Let f be Lipschitz continuous around $\bar{x}$. Then one has the equality*

$$\widehat{D}^*f(\bar{x})(y^*)=\widehat{\partial}\langle y^*,f\rangle(\bar{x})\ \text{ for all }\ y^*\in Y^*.$$

Proof. Pick any $x^*\in\widehat{\partial}\langle y^*,f\rangle(\bar{x})$ and employ definition (1.1) of Fréchet subgradients. Given $\varepsilon>0$, find $\eta>0$ such that

$$\langle x^*,x-\bar{x}\rangle\le\langle y^*,f(x)\rangle-\langle y^*,f(\bar{x})\rangle+\varepsilon\|x-\bar{x}\|\ \text{ whenever }\ \|x-\bar{x}\|\le\eta.$$

Furthermore, for all $x\in X$ we obviously have

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

which implies $(x^*,-y^*)\in\widehat{N}((\bar{x},f(\bar{x}));\operatorname{gph}f)$ and hence $x^*\in\widehat{D}^*f(\bar{x})(y^*)$.

To justify the opposite inclusion in the proposition, we assume that f is locally Lipschitzian around $\bar{x}$ with modulus $\ell\ge0$. Taking arbitrary $x^*\in\widehat{D}^*f(\bar{x})(y^*)$ and $\varepsilon>0$, we find $\eta>0$ such that

$$\langle x^*,x-\bar{x}\rangle-\langle y^*,f(x)-f(\bar{x})\rangle\le\varepsilon(\|x-\bar{x}\|+\|f(x)-f(\bar{x})\|)\le(\ell+1)\varepsilon\|x-\bar{x}\|$$

whenever $\|x-\bar{x}\|\le\eta$. Thus one has for such x that

$$\langle y^*,f\rangle(x)-\langle y^*,f\rangle(\bar{x})-\langle x^*,x-\bar{x}\rangle\ge-(\ell+1)\varepsilon\|x-\bar{x}\|,$$

which yields $x^*\in\widehat{\partial}\langle y^*,f\rangle(\bar{x})$ and ends the proof of the scalarization formula. △

Combining now the results of Theorem 3.1 and Proposition 3.5, we arrive at the difference rule for Fréchet coderivatives of Lipschitzian mappings between Banach spaces.

Corollary 3.6 (difference rule for Fréchet coderivatives). *Let $f_i\colon X\to Y$, $i=1,2$, be Lipschitz continuous around $\bar{x}$. Then*

$$\widehat{D}^*(f_1-f_2)(\bar{x})(y^*)\subset\bigcap_{x^*\in\widehat{D}^*f_2(\bar{x})(y^*)}\big[\widehat{D}^*f_1(\bar{x})(y^*)-x^*\big]\ \text{ for all }\ y^*\in Y^*\tag{3.6}$$

provided that $\widehat{D}^*f_2(\bar{x})(y^*)\neq\emptyset$. Furthermore, inclusion (3.6) holds as equality if f_2 is Fréchet differentiable at $\bar{x}$.

Proof. Since $\widehat{\partial}\langle y^*, f_2\rangle(\bar{x}) = \widehat{D}^* f_2(\bar{x})(y^*) \neq \emptyset$ by the scalarization formula, we have from Theorem 3.1 and Proposition 3.5 that

$$\begin{aligned}\widehat{D}^*(f_1 - f_2)(\bar{x})(y^*) = \widehat{\partial}\langle y^*, f_1 - f_2\rangle(\bar{x}) &\subset \bigcap_{x^* \in \widehat{\partial}\langle y^*, f_2\rangle(\bar{x})(y^*)} \big[\widehat{\partial}\langle y^*, f_1\rangle(\bar{x})(y^*) - x^*\big] \\ &= \bigcap_{x^* \in \widehat{D}^* f_2(\bar{x})(y^*)} \big[\widehat{D}^* f_1(\bar{x})(y^*) - x^*\big],\end{aligned}$$

which implies (3.6). The equality in (3.6) when f_2 is Fréchet differentiable at $\bar{x}$ follows immediately from Propositions 2.2 and 3.5. $\triangle$

Next we establish an *exact chain rule* for Fréchet subgradients of *generalized compositions* given by

$$(\varphi \circ f)(x) := \varphi(x, f(x)) \tag{3.7}$$

that involve extended-real-valued functions $\varphi\colon X \times Y \to \overline{I\!R}$ and single-valued mappings $f\colon X \to Y$ between arbitrary Banach spaces. Note that the following theorem and its consequences express Fréchet (lower) subgradients of compositions via Fréchet *upper* subgradients of the outer function φ as in (3.7). The proof of this theorem is based, similarly to Theorem 3.1, on the *smooth variational description* of Fréchet subgradients from Proposition 2.1.

Theorem 3.7 (chain rule with generalized compositions). *Given* (3.7), *suppose that f is Lipschitz continuous around $\bar{x}$ and φ is finite at $(\bar{x}, \bar{y})$, where $\bar{y} := f(\bar{x})$. Furthermore, we assume that $\widehat{\partial}^+\varphi(\bar{x}, \bar{y}) \neq \emptyset$. Then*

$$\widehat{\partial}(\varphi \circ f)(\bar{x}) \subset \bigcap_{(x^*, y^*) \in \widehat{\partial}^+\varphi(\bar{x}, \bar{y})} \big[x^* + \widehat{\partial}\langle y^*, f\rangle(\bar{x})\big]. \tag{3.8}$$

If in addition φ is Fréchet differentiable at $(\bar{x}, \bar{y})$, then

$$\widehat{\partial}(\varphi \circ f)(\bar{x}) = x^* + \widehat{\partial}\langle y^*, f\rangle(\bar{x}) \;\; \textit{with} \;\; (x^*, y^*) := \nabla\varphi(\bar{x}, \bar{y}). \tag{3.9}$$

Proof. Take arbitrary subgradients $u^* \in \widehat{\partial}(\varphi \circ f)(\bar{x})$ and $(x^*, y^*) \in \widehat{\partial}^+\varphi(\bar{x}, \bar{y})$. Since $(-x^*, -y^*) \in \widehat{\partial}(-\varphi)(\bar{x}, \bar{y})$ by definition (2.1), we employ Proposition 2.1 and find a function $s\colon U \to I\!R$, which is Fréchet differentiable at $(\bar{x}, \bar{y})$ satisfying

$$\varphi(\bar{x}, \bar{y}) = s(\bar{x}, \bar{y}), \quad \nabla s(\bar{x}, \bar{y}) = (x^*, y^*), \quad \varphi(x, y) \le s(x, y) \;\text{ for all }\; (x, y) \in U, \tag{3.10}$$

where U is a neighborhood of $(\bar{x}, \bar{y})$ in $X \times Y$. Now employing definition (1.1) for the Fréchet subgradient $u^* \in \widehat{\partial}(\varphi \circ f)(\bar{x})$ and then its smooth variational description in (3.10), for any $\varepsilon > 0$ find $\eta > 0$ such that

$$\begin{aligned}\langle u^*, x - \bar{x}\rangle &\le \varphi(x, f(x)) - \varphi(\bar{x}, f(\bar{x})) + \varepsilon\|x - \bar{x}\| \\ &\le s(x, f(x)) - s(\bar{x}, f(\bar{x})) + \varepsilon\|x - \bar{x}\| \\ &= \langle x^*, x - \bar{x}\rangle + \langle y^*, f(x) - f(\bar{x})\rangle + o(\|x - \bar{x}\| + \|f(x) - f(\bar{x})\|) + \varepsilon\|x - \bar{x}\|\end{aligned}$$

for all $\|x-\bar{x}\|<\eta$. This implies

$$\langle u^*-x^*,x-\bar{x}\rangle-\langle y^*,f(x)-f(\bar{x})\rangle\le o(\|x-\bar{x}\|+\|f(x)-f(\bar{x})\|)+\varepsilon\|x-\bar{x}\|$$

whenever $\|x-\bar{x}\|<\eta$. Hence

$$\limsup_{x\to\bar{x}}\frac{\langle u^*-x^*,x-\bar{x}\rangle-\langle y^*,f(x)-f(\bar{x})\rangle}{\|x-\bar{x}\|+\|f(x)-f(\bar{x})\|}\le\varepsilon\ \text{ whenever }\ \varepsilon\ge 0.$$

Since $\varepsilon>$ is arbitrary, we arrive at the inclusion

$$(u^*-x^*,-y^*)\in\widehat{N}\big((\bar{x},f(\bar{x}));\operatorname{gph}f\big),$$

which yields, by definition (2.7) and Proposition 3.5, that

$$u^*-x^*\in\widehat{D}^*f(\bar{x})(y^*)=\widehat{\partial}\langle y^*,f\rangle(\bar{x}),$$

i.e., $u^*\in x^*+\widehat{\partial}\langle y^*,f\rangle(\bar{x})$. This justifies the inclusion chain rule (3.8).

Let us prove the equality chain rule (3.9) provided that φ is Fréchet differentiable at $(\bar{x},\bar{y})$. Observe that (3.9) readily implies the inclusion

$$\widehat{\partial}(\varphi\circ f)(\bar{x})\subset x^*+\widehat{\partial}\langle y^*,f\rangle(\bar{x}).$$

Thus it remains to justify the opposite inclusion in (3.9). We proceed as follows: suppose that $u^*\notin\widehat{\partial}(\varphi\circ f)(\bar{x})$ and then show that $u^*\notin x^*+\widehat{\partial}\langle y^*,f\rangle(\bar{x})$, which is equivalent to

$$u^*\notin x^*+\widehat{D}^*f(\bar{x})(y^*)\tag{3.11}$$

by the scalarization formula of Proposition 3.5. Since $u^*\notin\widehat{\partial}(\varphi\circ f)(\bar{x})$, we have

$$\liminf_{x\to\bar{x}}\frac{(\varphi\circ f)(x)-(\varphi\circ f)(\bar{x})-\langle u^*,x-\bar{x}\rangle}{\|x-\bar{x}\|}<0.$$

This means that there are $\overline{\varepsilon}>0$ and a sequence $x_k\to\bar{x}$ such that $x_k\ne\bar{x}$ for all k and that

$$\frac{\varphi(x_k,f(x_k))-\varphi(\bar{x},f(\bar{x}))-\langle u^*,x_k-\bar{x}\rangle}{\|x_k-\bar{x}\|}\le-\overline{\varepsilon}.\tag{3.12}$$

Putting $y_k:=f(x_k)$, we conclude from (3.12) that

$$\begin{aligned}\langle u^*,x_k-\bar{x}\rangle&\ge\varphi(x_k,y_k)-\varphi(\bar{x},\bar{y})+\overline{\varepsilon}\|x_k-\bar{x}\|\\&=\langle\nabla\varphi(\bar{x},\bar{y}),(x_k-\bar{x},y_k-\bar{y})\rangle+o(\|x_k-\bar{x}\|+\|y_k-\bar{y}\|)+\overline{\varepsilon}\|x_k-\bar{x}\|\\&=\langle x^*,x_k-\bar{x}\rangle+\langle y^*,y_k-\bar{y}\rangle+o(\|x_k-\bar{x}\|+\|y_k-\bar{y}\|)+\overline{\varepsilon}\|x_k-\bar{x}\|.\end{aligned}$$

Taking into account that

$$\|x_k-\bar{x}\|\ge(1/\ell)\|y_k-\bar{y}\|\ \text{ for large }\ k\in I\!N$$

by the Lipschitz continuity of f around $\bar{x}$ with modulus ℓ, we get

$$\begin{aligned}\langle u^*-x^*,x_k-\bar{x}\rangle-\langle y^*,y_k-\bar{y}\rangle&\ge\frac{\overline{\varepsilon}}{2}\|x_k-\bar{x}\|+\frac{\overline{\varepsilon}}{2\ell}\|y_k-\bar{y}\|+o(\|x_k-\bar{x}\|+\|y_k-\bar{y}\|)\\&\ge\widehat{\varepsilon}(\|x_k-\bar{x}\|+\|y_k-\bar{y}\|)+o(\|x_k-\bar{x}\|+\|y_k-\bar{y}\|),\end{aligned}$$

where $\widehat{\varepsilon} := \min\{\overline{\varepsilon}/2, \overline{\varepsilon}/(2\ell)\}$. This implies

$$\limsup_{(x,y)\xrightarrow{\text{gph}\, f}(\bar{x},\bar{y})} \frac{\langle u^*-x^*, x-\bar{x}\rangle - \langle y^*, y-\bar{y}\rangle}{\|x-\bar{x}\|+\|y-\bar{y}\|} \geq \widehat{\varepsilon},$$

which shows that $(u^*-x^*,-y^*) \notin \widehat{N}((\bar{x},\bar{y});\text{gph}\, f)$ and thus justifies (3.9). $\triangle$

When $\varphi = \varphi(y)$ in (3.7), we immediately get from Theorem 3.7 the following chain rule for usual compositions $(\varphi \circ f)(x) := \varphi(f(x))$.

Corollary 3.8 (chain rules with usual compositions). *Let $f\colon X \to Y$ be Lipschitz continuous around $\bar{x}$, and let $\varphi\colon Y \to \overline{\mathbb{R}}$ be finite at $\bar{y} := f(\bar{x})$. Then*

$$\widehat{\partial}(\varphi \circ f)(\bar{x}) \subset \bigcap_{y^* \in \widehat{\partial}^+\varphi(\bar{y})} \widehat{\partial}\langle y^*, f\rangle(\bar{x}) \tag{3.13}$$

if $\widehat{\partial}^+\varphi(\bar{y}) \neq \emptyset$. Moreover, (3.13) holds as equality when φ is Fréchet differentiable at $\bar{y}$.

Based on the major calculus results of Theorems 3.1 and 3.7, we derive now some other useful calculus rules for Fréchet subgradients in arbitrary Banach spaces. The next theorem gives a general *product rule* involving Lipschitzian functions.

Theorem 3.9 (product rule for Fréchet subgradients). *Let the functions $\varphi_i\colon X \to \mathbb{R}$, $i = 1,2$, be Lipschitz continuous around $\bar{x}$. Assume that $\widehat{\partial}\big(-\varphi_1(\bar{x})\varphi_2\big)(\bar{x}) \neq \emptyset$. Then one has the product rule inclusion*

$$\widehat{\partial}(\varphi_1 \cdot \varphi_2)(\bar{x}) \subset \bigcap_{x^* \in \widehat{\partial}(-\varphi_1(\bar{x})\varphi_2)(\bar{x})} \big[\widehat{\partial}\big(\varphi_2(\bar{x})\varphi_1\big)(\bar{x}) - x^*\big], \tag{3.14}$$

which holds as equality provided that φ_2 is Fréchet differentiable at $\bar{x}$.

Proof. Define $f\colon X \to \mathbb{R}^2$ and $\psi\colon \mathbb{R}^2 \to \mathbb{R}$ by setting

$$f(x) := \big(\varphi_1(x), \varphi_2(x)\big) \;\text{ and }\; \psi(y_1,y_2) := y_1 \cdot y_2.$$

Then $\varphi_1 \cdot \varphi_2 = \psi \circ f$, and we use the chain rule from Corollary 3.8, which is more convenient to write via the coderivative by taking into account the scalarization formula of Proposition 3.5:

$$\widehat{\partial}(\varphi_1 \cdot \varphi_2)(\bar{x}) = \widehat{D}^* f(\bar{x})\big(\nabla\psi(f(\bar{x})\big) = \widehat{D}^* f(\bar{x})\big(\varphi_2(\bar{x}), \varphi_1(\bar{x})\big). \tag{3.15}$$

Since $f(x) = f_1(x) - f_2(x)$ with $f_1(x) := (\varphi_1(x), 0)$ and $f_2(x) := (0, -\varphi_2(x))$, we derive from the difference rule of Corollary 3.6 that

$$\widehat{D}^* f(\bar{x})(u_1,u_2) \subset \bigcap_{x^* \in \widehat{D}^* f_2(\bar{x})(u_1,u_2)} \big[\widehat{D}^* f_1(\bar{x})(u_1,u_2) - x^*\big]. \tag{3.16}$$

Since $\text{gph}\, f_1 = \text{gph}\, \varphi_1 \times \{0\}$, we conclude from the elementary product formula for Fréchet normals that

$$\widehat{N}\big((\bar{x}; f_1(\bar{x}));\text{gph}\, f_1\big) = \widehat{N}\big((\bar{x}, \varphi_1(\bar{x}));\text{gph}\, \varphi_1\big) \times \mathbb{R}$$

and similarly for f_2. Therefore

$$\widehat{D}^* f_1(\bar{x})\big(\varphi_2(\bar{x}),\varphi_1(\bar{x})\big) = \widehat{D}^*\varphi_1(\bar{x})\big(\varphi_2(\bar{x})\big) = \widehat{\partial}\big(\varphi_2(\bar{x})\varphi_1\big)(\bar{x}),$$
$$\widehat{D}^* f_2(\bar{x})\big(\varphi_2(\bar{x}),\varphi_1(\bar{x})\big) = \widehat{D}^*\big(-\varphi_2(\bar{x})\big)\big(\varphi_1(\bar{x})\big) = \widehat{\partial}\big(-\varphi_1(\bar{x})\varphi_2\big)(\bar{x}),$$

and thus the inclusion (3.14) follows from (3.15) and (3.16).

The equality in (3.14) under the Fréchet differentiability of φ_2 at $\bar{x}$ follows from the equality in (3.16) under the equivalent Fréchet differentiability of f_2 at this point due to the equality case of Corollary 3.6. $\triangle$

Before proceeding with the quotient rule for Fréchet subgradients, let us present a simple *reciprocal rule*, which always holds as equality being thus of independent interest.

Proposition 3.10 (reciprocal rule for Fréchet subgradients). *Let $\varphi\colon X \to \mathbb{R}$ be locally Lipschitzian around $\bar{x}$ with $\varphi(\bar{x}) \neq 0$. Then*

$$\widehat{\partial}\big(1/\varphi\big)(\bar{x}) = \frac{\widehat{\partial}(-\varphi)(\bar{x})}{(\varphi(\bar{x}))^2}. \tag{3.17}$$

Proof. Putting $\psi(x) := 1/x$ for all $x \neq 0$, we have $1/\varphi = \psi \circ \varphi$. Thus Corollary 3.8 and Proposition 3.5 ensure the equalities

$$\widehat{\partial}\big(1/\varphi\big)(\bar{x}) = \widehat{\partial}(\psi\circ\varphi)(\bar{x}) = \widehat{\partial}\langle\nabla\psi(\varphi(\bar{x})),\varphi\rangle(\bar{x}) = \widehat{D}^*\varphi(\bar{x})\big(\nabla\psi(\varphi(\bar{x})\big).$$

Since $\nabla\psi(\varphi(\bar{x})) = -1/[(\varphi(\bar{x}))^2]$, we have

$$\widehat{D}^*\varphi(\bar{x})\big(\nabla\psi(\varphi(\bar{x}))\big) = \widehat{D}^*\varphi(\bar{x})\Big(-\frac{1}{(\varphi(\bar{x}))^2}\Big) = \widehat{\partial}\Big(-\frac{1}{(\varphi(\bar{x}))^2}\varphi\Big)(\bar{x}) = \frac{1}{(\varphi(\bar{x}))^2}\widehat{\partial}(-\varphi)(\bar{x}),$$

which justifies the reciprocal rule (3.17). $\triangle$

Next we derive the *quotient rule* for Fréchet subgradients of locally Lipschitzian functions on arbitrary Banach spaces.

Theorem 3.11 (quotient rule for Fréchet subgradients). *Let $\varphi_i\colon X \to \mathbb{R}$, $i = 1,2$, be Lipschitz continuous around $\bar{x}$ with $\varphi_2(\bar{x}) \neq 0$. Assume that $\widehat{\partial}\big(\varphi_1(\bar{x})\varphi_2\big)(\bar{x}) \neq \emptyset$. Then one has the inclusion*

$$\widehat{\partial}\big(\varphi_1/\varphi_2\big)(\bar{x}) \subset \bigcap_{x^*\in\widehat{\partial}(\varphi_1(\bar{x})\varphi_2)(\bar{x})} \frac{\big[\widehat{\partial}\big(\varphi_2(\bar{x})\varphi_1\big)(\bar{x}) - x^*\big]}{(\varphi_2(\bar{x}))^2}, \tag{3.18}$$

which holds as equality if φ_2 is Fréchet differentiable at $\bar{x}$.

Proof. We obviously have the composite representation $\varphi_1/\varphi_2 = \psi\circ f$ with $f\colon X \to \mathbb{R}^2$ and $\psi\colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x) := \big(\varphi_1(x),\varphi_2(x)\big) \;\text{ and }\; \psi(y_1,y_2) := y_1/y_2.$$

Classical calculus gives us the equality

$$\nabla\psi(f(\bar{x})) = \Big(\frac{1}{\varphi_2(\bar{x})}, -\frac{\varphi_1(\bar{x})}{(\varphi_2(\bar{x}))^2}\Big).$$

Observe that $f = f_1 - f_2$ with $f_1(x) := (\varphi_1(x), 0)$ and $f_2(x) := (0, -\varphi_2(x))$. The assumptions made ensure, by the scalarization formula of Proposition 3.5, that

$$\widehat{D}^* f_2\Big(\frac{1}{\varphi_2(\bar{x})}, -\frac{\varphi_1(\bar{x})}{(\varphi_2(x))^2}\Big) = \frac{1}{(\varphi_2(\bar{x}))^2}\widehat{\partial}\big(\varphi_1(\bar{x})\varphi_2\big)(\bar{x}) \neq \emptyset \text{ and}$$

$$\widehat{D}^* f_1(\bar{x})\Big(\frac{1}{\varphi_2(\bar{x})}, -\frac{\varphi_1(\bar{x})}{(\varphi_2(\bar{x}))^2}\Big) = \widehat{\partial}\Big(\frac{\varphi_1}{\varphi_2(\bar{x})}\Big)(\bar{x}).$$

Employing therefore the difference rule from Corollary 3.6 and then the chain rule from Corollary 3.8, we get

$$\begin{aligned}\widehat{\partial}(\varphi_1/\varphi_2)(\bar{x}) &\subset \bigcap_{x^* \in \frac{1}{(\varphi_2(\bar{x}))^2}\widehat{\partial}(\varphi_1(\bar{x})\varphi_2)(\bar{x})} \Big[\widehat{\partial}\Big(\frac{\varphi_1}{\varphi_2(\bar{x})}\Big)(\bar{x}) - x^*\Big] \\ &= \bigcap_{x^* \in \widehat{\partial}\left(\varphi_1(\bar{x})\varphi_2\right)(\bar{x})} \left[\frac{\widehat{\partial}(\varphi_2(\bar{x})\varphi_1)(\bar{x}) - x^*}{(\varphi_2(\bar{x}))^2}\right],\end{aligned}$$

where the inclusion holds as equality if φ_2 is Fréchet differentiable at $\bar{x}$. This completes the proof of the theorem. $\triangle$

Note that the reciprocal rule of Proposition 3.10 is generally independent from the quotient rule of Theorem 3.11, since (3.18) is proved to hold as equality if φ_2 is Fréchet differentiable at $\bar{x}$, while there is no such a restriction in the reciprocal rule (3.17).

To conclude this section, we derive a useful rule for evaluating Fréchet subgradients of the *minimum function*

$$(\wedge\varphi_i)(x) := \min\big\{\varphi_i(x) \big| \, i = 1, ..., n\big\} \text{ with } \varphi_i \colon X \to \overline{\mathbb{R}} \text{ and } n \geq 2$$

important in many applications. Denote $I(x) := \big\{j \in \{1, ..., n\} \big| \, \varphi_j(x) = (\wedge\varphi_i)(x)\big\}$.

Proposition 3.12 (Fréchet subgradients of minimum functions). *We always have the inclusion*

$$\widehat{\partial}(\wedge\varphi_i)(\bar{x}) \subset \bigcap_{j \in I(\bar{x})} \widehat{\partial}\varphi_j(\bar{x}).$$

Proof. Take $x^* \in \widehat{\partial}(\wedge\varphi_i)(\bar{x})$ and for any $\varepsilon > 0$ find $\eta > 0$ such that

$$\langle x^*, x - \bar{x}\rangle \leq (\wedge\varphi_i)(x) - (\wedge\varphi_i)(\bar{x}) + \varepsilon\|x - \bar{x}\|$$

whenever $\|x - \bar{x}\| < \eta$. For such x and for any $j \in I(\bar{x})$ we have

$$\begin{aligned}\langle x^*, x - \bar{x}\rangle \leq \langle x^*, x - \bar{x}\rangle &\leq (\wedge\varphi_i)(x) - (\wedge\varphi_i)(\bar{x}) + \varepsilon\|x - \bar{x}\| \\ &= (\wedge\varphi_i)(x) - \varphi_j(\bar{x}) + \varepsilon\|x - \bar{x}\| \\ &\leq \varphi_j(x) - \varphi_j(\bar{x}) + \varepsilon\|x - \bar{x}\|,\end{aligned}$$

which justifies $x^* \in \widehat{\partial}\varphi_j(\bar{x})$ and completes the proof of the theorem. $\triangle$

4 Necessary Conditions in Nondifferentiable Difference Programming

In this section we apply the calculus results for Fréchet subgradients established in the previous section and also known calculus rules for basic subgradients to derive new necessary optimality conditions for some classes of problems in nondifferentiable programming. The main attention is paid to problems of the so-called *difference programming*, i.e., to optimization problems whose cost functions are given in the *difference form*; see the discussions and references in Section 1.

Let us first present simple conditions for *unconstrained* problems of minimizing *difference functions* $\varphi = \varphi_1 - \varphi_2$ on arbitrary Banach spaces.

Proposition 4.1 (necessary conditions for minimizing difference functions). *Let $\bar{x}$ be a local minimizer of the difference function $\varphi = \varphi_1 - \varphi_2$, where both $\varphi_i \colon X \to \overline{I\!R}$ are finite at $\bar{x}$. Then one has the inclusion*

$$\widehat{\partial}\varphi_2(\bar{x}) \subset \widehat{\partial}\varphi_1(\bar{x}). \tag{4.1}$$

If in addition φ_2 is lower regular at $\bar{x}$, then

$$\partial\varphi_2(\bar{x}) \subset \partial\varphi_1(\bar{x}). \tag{4.2}$$

Proof. Inclusion (4.1) immediately follows from the difference rule of Theorem 3.1 by the Fermat stationary principle (1.1). The second one (4.2) follows from (4.1) due to the lower regularity definition (2.4) and the inclusion $\widehat{\partial}\varphi_1(\bar{x}) \subset \partial\varphi_1(\bar{x})$. $\triangle$

Note that inclusion (4.2) does *not* hold as a necessary optimality condition with no lower regularity assumption. Indeed, minimizing the difference $\varphi_1 - \varphi_2$ with $\varphi_1(x) \equiv 1$ and $\varphi_2(x) = -|x|$, we have $\bar{x} = 0$ as its global minimizer. In the case $\widehat{\partial}\varphi_1(0) = \partial\varphi_1(0) = \{0\}$, $\widehat{\partial}\varphi_2(0) = \emptyset$, and $\partial\varphi_2(0) = \{-1, 1\}$. Thus (4.1) holds, while (4.2) is not satisfied.

Simple necessary optimality conditions of Proposition 4.1 have useful consequences for the study of the so-called "weak sharp minima," for which necessary optimality conditions were previously obtained via Clarke generalized gradients under restrictive regularity assumptions; see, e.g., Burke and Ferris [2] with the references therein. Now we can significantly improve previously known results. Given a proper function $\varphi \colon X \to \overline{I\!R}$ and a nonempty subset $\Omega \subset X$ of a Banach space, recall that $S \subset \Omega$ is a set of *weak sharp minima* for φ relative to $\Omega \subset X$ with modulus $\alpha > 0$ if

$$\varphi(x) \geq \varphi(y) + \alpha \operatorname{dist}(x; S) \text{ for all } x \in \Omega \text{ and } y \in S,$$

where $\operatorname{dist}(x; S)$ stands for the distance function of the set S.

Corollary 4.2 (necessary conditions for unconstrained weak sharp minima). *Let S be the set of weak sharp minima for the function φ relative to the whole space X with modulus α. Then for every $\bar{x} \in S$ we have*

$$\alpha I\!B^* \cap \widehat{N}(\bar{x}; S) \subset \widehat{\partial}\varphi(\bar{x}), \tag{4.3}$$

where $\mathbb{B}^* \subset X^*$ stands for the unit dual ball. If in addition S is normally regular at $\bar{x}$, then

$$\alpha\mathbb{B}^* \cap N(\bar{x}; S) \subset \partial\varphi(\bar{x}). \tag{4.4}$$

Proof. By definition we have

$$\varphi(x) \geq \varphi(y) + \alpha\,\mathrm{dist}(x; S) \;\text{ for all }\; x \in X \;\text{ and }\; y \in S.$$

Thus every $y \in S$ is an optimal solution to the unconstrained problem of minimizing the *difference function* $\psi(x) := \varphi(x) - \alpha\,\mathrm{dist}(x; S)$. Employing Theorem 4.1, we get

$$\alpha\,\widehat{\partial}\mathrm{dist}(y; S) \subset \widehat{\partial}\varphi(y). \tag{4.5}$$

It is well known (see, e.g., [4, Corollary 1.96]) that

$$\widehat{\partial}\mathrm{dist}(\bar{x}; S) = \widehat{N}(\bar{x}; S) \cap \mathbb{B}^*.$$

Substituting it into (4.5), we arrive at (4.3). Inclusion (4.4) in the regularity case follows from (4.2) in Proposition 4.1 and the fact that the normal regularity of a set agrees with the lower regularity of its distance function at any in-set point; see [4, Corollary 1.98]. $\triangle$

Next we consider some *constrained* problems of nondifferentiable difference programming. To derive necessary conditions for constrained problems, we mainly combine the unconstrained results with more developed calculus rules available for basic/limiting normals and subgradients.

Let us start with difference problems under general *geometric* constraints given by

$$\text{minimize }\; \varphi(x) \;\text{ subject to }\; x \in \Omega, \tag{4.6}$$

where the cost function φ is represented as $\varphi = \varphi_1 - \varphi_2$. For convenience we say that φ is *Fréchet decomposable* on Ω at $\bar{x} \in \Omega$ if

$$\widehat{\partial}(\varphi + \delta(\cdot; \Omega))(\bar{x}) \subset \widehat{\partial}\varphi(\bar{x}) + \widehat{N}(\bar{x}; \Omega).$$

It happens, e.g., when φ is Fréchet differentiable at $\bar{x} \in \Omega$, while Ω is an arbitrary nonempty subset of a Banach space. In the first parts of the following theorem and its corollary we impose the decomposition assumption in the general setting of Banach spaces, while the second parts are justified without this assumption via limiting normals and subgradients in the framework of Asplund spaces.

Theorem 4.3 (necessary conditions for difference problems with geometric constraints). *Let $\bar{x}$ be a local solution of problem (4.6) in a Banach space X, where φ is represented as $\varphi_1 - \varphi_2$ with $\varphi_i \colon X \to \overline{\mathbb{R}}$ finite at $\bar{x}$. The following assertions hold:*

(i) *Assume that φ_1 is Fréchet decomposable on Ω. Then one has*

$$\widehat{\partial}\varphi_2(\bar{x}) \subset \widehat{\partial}\varphi_1(\bar{x}) + \widehat{N}(\bar{x}; \Omega). \tag{4.7}$$

(ii) *Suppose that X is Asplund, that φ_1 is l.s.c. around $\bar{x}$, and that Ω is locally closed around this point. Assume in addition that either φ_1 is SNEC at $\bar{x}$ or Ω is SNC at $\bar{x}$, and that the qualification condition*

$$\partial^\infty \varphi_1(\bar{x}) \cap \big(-N(\bar{x};\Omega)\big) = \{0\}. \tag{4.8}$$

is satisfied; all the assumptions on φ_1 hold automatically when φ_1 is Lipschitz continuous around $\bar{x}$. Then one has

$$\widehat{\partial}\varphi_2(\bar{x}) \subset \partial\varphi_1(\bar{x}) + N(\bar{x};\Omega). \tag{4.9}$$

Proof. The problem (4.6) under consideration can obviously be reformulated in the unconstrained difference form:

$$\text{minimize } \big[\varphi_1(x) + \delta(x;\Omega)\big] - \varphi_2(x) \text{ subject to } x \in \Omega. \tag{4.10}$$

By Proposition 4.1 we have

$$\widehat{\partial}\varphi_2(\bar{x}) \subset \widehat{\partial}(\varphi_1 + \delta(\cdot;\Omega))(\bar{x})$$

for a given local minimizer $\bar{x}$ to (4.6). This directly implies the necessary condition (4.7) provided that φ_1 is decomposable on Ω.

To justify (4.9), observe that

$$\widehat{\partial}(\varphi_1 + \delta(\cdot;\Omega))(\bar{x}) \subset \partial(\varphi_1 + \delta(\cdot;\Omega))(\bar{x}) \subset \partial\varphi_1(\bar{x}) + N(\bar{x};\Omega),$$

where the latter inclusion holds due to the sum rule for basic subgradients from [4, Theorem 3.36] under the assumptions made in (ii). If φ_1 is locally Lipschitzian around $\bar{x}$, then it is SNEC by [4, Corollary 1.69] and the qualification condition (4.8) is satisfied due to $\partial^\infty\varphi_1(\bar{x}) = \{0\}$ by [4, Corollary 1.81]. This ends the proof of the theorem. $\triangle$

As a useful corollary of Theorem 4.3, we get the following necessary conditions for weak sharp minima under general geometric constraints.

Corollary 4.4 (necessary conditions for weak sharp minima under geometric constraints). *Let S be the set of weak sharp minima for $\varphi\colon X \to \overline{\mathbb{R}}$ relative to $\Omega \subset X$ with modulus $\alpha > 0$. Then one has:*

(i) *Given $\bar{x} \in S$, assume that φ is Fréchet decomposable on Ω. Then*

$$\alpha\mathbb{B}^* \cap \widehat{N}(\bar{x};S) \subset \widehat{\partial}\varphi(\bar{x}) + \widehat{N}(\bar{x};\Omega).$$

(ii) *Assume that X is Asplund, that φ is l.s.c. around $\bar{x}$, and that Ω is locally closed around this point. Then*

$$\alpha\mathbb{B}^* \cap \widehat{N}(\bar{x};S) \subset \partial\varphi(\bar{x}) + N(\bar{x};\Omega)$$

provided that the qualification condition

$$\partial^\infty\varphi(\bar{x}) \cap \big(-N(\bar{x};\Omega)\big) = \{0\}$$

holds and that either φ is SNEC at $\bar{x}$ or Ω is SNC at this point; all the assumptions on φ are automatic if φ is locally Lipschitzian around $\bar{x}$. Furthermore, we have

$$\alpha I\!B^* \cap N(\bar{x}; S) \subset \partial\varphi(\bar{x}) + N(\bar{x}; \Omega)$$

if in addition S is normally regular at $\bar{x}$.

Proof. This follows from Theorem 4.3 by reducing weak sharp minima to constrained minimization of difference functions; cf. the proof of Corollary 4.2. $\triangle$

Finally in this section, we consider constrained problems of difference programming, where constraints are given in some general *operator form* via single-valued mappings between infinite-dimensional spaces. They particularly include problems with standard *functional* constraints defined by equalities and inequalities with (generally nonsmooth) real-valued functions. Note that the problems under consideration may also include geometric constraints as in Theorem 4.3, while the results obtained therein and in what follows are generally independent for common classes of constrained problems.

The class of difference programming problems with operator constraints considered below is given by:

$$\text{minimize } \varphi_1(x) - \varphi_2(x) \text{ subject to } x \in f^{-1}(\Theta) \cap \Omega, \tag{4.11}$$

where $f\colon X \to Y$, $\Omega \subset X$, $\Theta \subset Y$, and the *inverse image*/preimage of Θ under f is

$$f^{-1}(\Theta) := \big\{x \in X \big|\, f(x) \in \Theta\big\}.$$

Note that the classical case of operator constraints corresponds to the case of $\Theta = \{0\}$ in (4.11) with a mapping/operator f having infinite-dimensional values.

We present two independent theorems concerning necessary optimality conditions for problems (4.11). The first one requires the existence of a *surjective* derivative of f at the point in question, while involves general spaces in model (4.11). The second theorem deals with mappings between Asplund spaces with *no differentiability* assumptions on f.

Theorem 4.5 (necessary conditions for difference programming with surjective operator constraints). *Let $\bar{x}$ be a local solution to problem (4.11) in the Banach space setting, where $\varphi_i\colon X \to \overline{I\!R}$, $i = 1, 2$, are finite at $\bar{x}$, and where $f\colon X \to Y$ is Fréchet differentiable at $\bar{x}$ with the surjective derivative $\nabla f(\bar{x})$. The following assertions hold:*

(i) *Assume that φ_1 is Fréchet differentiable at $\bar{x}$, that $\Omega = X$, and that either f is strictly differentiable at $\bar{x}$ or $\dim Y < \infty$. Then*

$$\widehat{\partial}\varphi_2(\bar{x}) \subset \nabla\varphi_1(\bar{x}) + \nabla f(\bar{x})^* \widehat{N}(f(\bar{x}); \Theta). \tag{4.12}$$

(ii) *Assume that X is Asplund, that φ_1 is l.s.c. around $\bar{x}$, that Ω and Θ are locally closed around $\bar{x}$ and $f(\bar{x})$, respectively, and that f is strictly differentiable at $\bar{x}$. Furthermore, we suppose the the qualification condition*

$$\begin{cases} \big[x_1^* \in \partial^\infty\varphi_1(\bar{x}),\ \ x_2^* \in N(\bar{x}; \Omega),\ \ x_3^* \in \nabla f(\bar{x})^* N(f(\bar{x}); \Theta), \\ x_1^* + x_2^* + x_3^* = 0\big] \Longrightarrow\ x_1^* = x_2^* = x_3^* = 0 \end{cases} \tag{4.13}$$

is satisfied and that either both Ω and Θ are SNC at $\bar{x}$ and $f(\bar{x})$, respectively, or one of these sets has the SNC property while φ_1 is SNEC at $\bar{x}$; the latter assumption on φ_1 is automatic when φ_1 is locally Lipschitzian around $\bar{x}$ in which case the qualification condition (4.13) *reduces to*

$$\nabla f(\bar{x})^* N(f(\bar{x});\Theta) \cap \big(-N(\bar{x};\Omega)\big) = \{0\}. \tag{4.14}$$

Then one has the inclusion

$$\widehat{\partial}\varphi_2(\bar{x}) \subset \partial\varphi_1(\bar{x}) + \nabla f(\bar{x})^* N(f(\bar{x});\Theta) + N(\bar{x};\Omega). \tag{4.15}$$

Proof. To justify (i), we employ assertion (i) of Theorem 4.3 with the geometric constraint given by $x \in f^{-1}(\Theta)$. Using then the calculus rule for computing Fréchet normals to inverse images from [4, Corollary 1.15], we get

$$\widehat{N}(\bar{x}; f^{-1}(\Theta)) = \nabla f(\bar{x})^* \widehat{N}(\bar{y};\Theta)$$

under the assumptions made in (i). Thus (4.12) follows from (4.7) in this setting.

To justify (ii) with $\Omega \neq X$, we assume that X is Asplund (Y may be an arbitrary Banach space) and use further calculus for basic normals and for the sequential normal compactness property available in the Asplund space framework. In this case the geometric constraint in (4.6) is written in the intersection form: $x \in f^{-1}(\Theta) \cap \Omega$. By assertion (ii) of Theorem 4.3 one has the necessary optimality condition in (4.11) given by

$$\widehat{\partial}\varphi_2(\bar{x}) \subset \partial\varphi_1(\bar{x}) + N(\bar{x}; f^{-1}(\Theta) \cap \Omega)$$

when the qualification condition

$$\partial^\infty\varphi_1(\bar{x}) \cap \big[-N(\bar{x}; f^{-1}(\Theta) \cap \Omega)\big] = \{0\}$$

is satisfied and when either φ_1 is SNEC at $\bar{x}$ or $f^{-1}(\Theta) \cap \Omega$ is SNC at this point. Employing the intersection rule from [4, Corollary 3.5], we get

$$N(\bar{x}; f^{-1}(\Theta) \cap \Omega) \subset N(\bar{x}; f^{-1}(\Theta)) + N(\bar{x};\Omega)$$

provided the local closedness of $f^{-1}(\Theta)$ and Ω around $\bar{x}$, the qualification condition

$$N(\bar{x}; f^{-1}(\Theta)) \cap \big(-N(\bar{x};\Omega)\big) = \{0\}, \tag{4.16}$$

and that either $f^{-1}(\Theta)$ or Ω is SNC at $\bar{x}$. Moreover, the intersection $f^{-1}(\Theta) \cap \Omega$ is SNC at $\bar{x}$ if *both* of these sets are SNC at this point under the fulfillment of the qualification condition (4.16); see [4, Corollary 3.81].

Furthermore, by [4, Theorem 1.17] we have the equality representation of basic normals

$$N(\bar{x}; f^{-1}(\Theta)) = \nabla f(\bar{x})^* N(f(\bar{x});\Theta) \tag{4.17}$$

to inverse images of arbitrary sets in Banach spaces under strictly differentiable mappings with surjective derivatives. Observe also that the SNC property of $f^{-1}(\Theta)$ at $\bar{x}$ is equivalent, under the surjectivity of $\nabla f(\bar{x})$, to the SNC property of Θ at $f(\bar{x})$; see [4, Theorem 1.22].

Combining the above calculus results, we arrive at the necessary optimality condition (4.15) under the assumptions made in (ii). If φ_1 is Lipschitz continuous around $\bar{x}$, then it is SNEC at this point with $\partial^\infty\varphi_1(\bar{x}) = \{0\}$ due to the facts mentioned in the proof of Theorem 4.3(ii). Thus the qualification condition (4.13) reduces to (4.14) in this case, which completes the proof of the theorem. $\triangle$

The last result of this paper establishes a general necessary optimality condition for problem (4.11) with operator and geometric constraints without any differentiability (and even more–surjective derivative) assumptions on f. This is done in the setting when *both* spaces X and Y are Asplund allowing us to employ comprehensive generalized differential and SNC calculi from [4, Chapter 3]. In what follow we use the basic/limiting coderivative (2.8) as a proper extension of the adjoint derivative $\nabla f(\bar{x})^* y^*$ for nonsmooth operators. Note the composite notation

$$D^* f(\bar{x}) \circ \Lambda := \big\{ D^* f(\bar{x})(y^*) \big| \, y^* \in \Lambda \big\}.$$

Theorem 4.6 (necessary conditions for difference programming with nondifferentiable operator constraints). *Let $\bar{x}$ be a local solution to problem (4.11), where both spaces X and Y are Asplund, where $\varphi_1 \colon X \to \overline{I\!R}$ is l.s.c. around $\bar{x}$ while $\varphi_2 \colon X \to \overline{I\!R}$ is finite at this point, where $f \colon X \to Y$ is Lipschitz continuous around $\bar{x}$, and where Ω and Θ is locally closed at $\bar{x}$ and $f(\bar{x})$, respectively. Then one has*

$$\widehat{\partial}\varphi_2(\bar{x}) \subset \partial\varphi_1(\bar{x}) + D^* f(\bar{x}) \circ N(f(\bar{x}); \Theta) + N(\bar{x}; \Omega)$$

provided the fulfillment of the qualification conditions

$$\begin{cases} \big[x_1^* \in \partial^\infty\varphi_1(\bar{x}), \;\; x_2^* \in N(\bar{x}; \Omega), \;\; x_3^* \in D^* f(\bar{x}) \circ N(f(\bar{x}); \Theta), \\ x_1^* + x_2^* + x_3^* = 0\big] \Longrightarrow \; x_1^* = x_2^* = x_3^* = 0, \end{cases}$$

$$N(f(\bar{x}); \Theta) \cap \ker D^* f(\bar{x}) = \{0\}, \tag{4.18}$$

and that either both Ω and Θ are SNC at $\bar{x}$ and $f(\bar{x})$, respectively, or one of these sets has the SNC property while φ_1 is SNEC at $\bar{x}$.

Proof. It follows the same pattern as the above proof of Theorem 4.5(ii) with the replacement of the equality (4.17) by the inclusion

$$N(\bar{x}; f^{-1}(\Theta)) \subset D^* f(\bar{x}) \circ N(f(\bar{x}); \Theta),$$

which holds under the fulfillment of the qualification condition (4.18); see [4, Corollary 3.9]. Furthermore, [4, Theorem 3.84] ensures the SNC property of the inverse image $f^{-1}(\Theta)$ at $\bar{x}$ provided that f is locally Lipschitzian around $\bar{x}$, Θ is SNC at $f(\bar{x})$, and the qualification condition (4.18) holds. This completes the proof of the theorem. $\triangle$

Remark 4.7 (necessary conditions for difference programming with equilibrium constraints). Employing generalized differential and SNC calculi developed in [4, Chapter 3], we can derive necessary optimality conditions for problems of difference programming

with other types of constraints similarly to optimization problems with standard cost functions considered by Mordukhovich [5, Chapter 5]. In particular, we can incorporate more general operator constrained of the type

$$x \in F^{-1}(\Theta) := \big\{x \in X \big| \; F(x) \cap \Theta \neq \emptyset\big\}$$

described by set-valued mappings $F\colon X \rightrightarrows Y$. More interesting, the so-called *equilibrium constraints* given by solution maps to "generalized equations"

$$S(x) := \big\{y \in Y \big| \; 0 \in g(x,y) + Q(x,y)\big\} \;\text{ with }\; g\colon X \to Y \;\text{ and }\; Q\colon X \rightrightarrows Y$$

can be taken into account. The latter includes parametric sets of solutions to variational and hemivariational inequalities, complementarity problems, KKT systems, etc.; see [5] for more details, references, and applications.

References

[1] J. M. Borwein and Q. J. Zhu (2005). *Techniques of Variational Analysis.* Springer-Verlag, New York.

[2] J. V. Burke and M. C. Ferris (1993). Weak sharp minima in mathematical programming. *SIAM Journal on Control and Optimization*, **31**, 1340–1359.

[3] J.-B. Hiriart-Urruty (1985). Generalized differentiability, duaslity and optimization for problems dealing with differences of convex functions. *Lecture Notes in Economics and Mathematical Sciences*, **256**, pp. 219–239. Springer-Verlag, Berlin.

[4] B. S. Mordukhovich (2005). *Variational Analysis and Generalized Differentiation, I: Basic Theory.* Springer-Verlag, Berlin.

[5] B. S. Mordukhovich (2005). *Variational Analysis and Generalized Differentiation, II: Applications.* Springer-Verlag, Berlin.

[6] D. Pallaschke and R. Urbański (2002). *Pairs of Compact Convex Sets.* Kluwer Academic Publishers, Dordrecht, The Netherlands.

[7] J.-P. Penot (1998). On the minimization of difference functions. *Journal of Global Optimization*, **12**, 373–382.

[8] R. R. Phelps (1993). *Convex Functions, Monotone Operators and Differentiability*, 2nd edition. Springer-Verlag, Berlin.

[9] B. T. Polyak (1987). *Introduction to Optimization.* Optimization Software, New York.

[10] R. T. Rockafellar (1970). *Convex Analysis.* Princeton University Press, Princeton, New Jersy.

[11] R. T. Rockafellar and R. J-B. Wets (1998). *Variational Analysis.* Springer-Verlag, Berlin.