Journal of Convex Analysis
Volume 2 (1995), No.1/2, 211–227

On Nonconvex Subdifferential Calculus in Banach Spaces[1]

Boris S. Mordukhovich, Yongheng Shao
Department of Mathematics, Wayne State University,
Detroit, Michigan 48202, U.S.A.
e-mail: boris@math.wayne.edu

Received July 1994
Revised manuscript received March 1995

Dedicated to R. T. Rockafellar on his 60th Birthday

We study a concept of subdifferential for general extended-real-valued functions defined on arbitrary Banach spaces. At any point considered this basic subdifferential is a nonconvex set of subgradients formed by sequential weak-star limits of the so-called Fréchet ε-subgradients of the function at points nearby. It is known that such and related constructions possess full calculus under special geometric assumptions on Banach spaces where they are defined. In this paper we establish some useful calculus rules in the general Banach space setting.

Keywords : Nonsmooth analysis, generalized differentiation, extended-real-valued functions, nonconvex calculus, Banach spaces, sequential limits

1991 Mathematics Subject Classification: Primary 49J52; Secondary 58C20

1. Introduction

It is well known that for studying local behavior of nonsmooth functions one can successfully use an appropriate concept of *subdifferential* (set of subgradients) which replaces the classical gradient at points of nondifferentiability in the usual two-sided sense. Such a concept first appeared for convex functions in the context of *convex analysis* that has had many significant applications to the range of problems in optimization, economics, mechanics, etc. One of the most important branches of convex analysis is *subdifferential calculus* for convex extended-real-valued functions that was developed in the 1960's, mainly by Moreau and Rockafellar; see [23] and [25, 27] for expositions and references.
The first concept of subdifferential for general nonconvex functions was introduced by Clarke (1973) who performed pioneering work in the area of *nonsmooth analysis* spread far beyond the scope of convexity. In particular, Clarke developed a comprehensive subdifferential calculus for his generalized gradients of locally Lipschitzian functions defined

[1] This research was partially supported by the National Science Foundation under grants DMS–9206989 and DMS-9404128.

ISSN 0944-6532 / \$ 2.50 ⓒ Heldermann Verlag

on Banach spaces; see [2] and references therein. In [26, 28], Rockafellar established and sharpened a number of calculus rules for Clarke's generalized gradients of nonconvex lower semicontinuous (l.s.c.) functions that are not necessarily locally Lipschitzian.

In this paper we study another subdifferential concept for extended-real-valued functions on Banach spaces that appeared in Kruger-Mordukhovich [10, 11] as a generalization of Mordukhovich's finite dimensional construction in [16, 17].This subdifferential corresponds to the collection of *sequential* limiting points of the so-called Fréchet ε-subgradients under small perturbations. In contrast to Clarke's subdifferential, such a set of limiting subgradients usually turns out to be *nonconvex* (may be even not closed) that makes its analysis more complicated. Nevertheless, a number of principal calculus rules and applications have been established for this nonconvex subdifferential and related constructions in finite and infinite dimensions under some geometric assumptions on the structure of (infinite dimensional) Banach spaces. We refer the reader to [3, 4, 7, 9, 13–22, 29, 30] for more details and further information.

To the best of our knowledge, the most general calculus results are obtained in the recent paper [22] for the case of l.s.c. functions defined on Asplund spaces. This class of Banach spaces (see [24]) includes, in particular, every space with a Fréchet differentiable renorm (hence every reflexive space) and appears to be convenient for many applications. On the other hand, some standard Banach spaces (like $C,\ L^1,\ L^\infty$) are not Asplund that reduces the range of possible applications of the subdifferential theory in [22], e.g., in optimal control.

This paper is concerned with developing subdifferential calculus for the limiting nonconvex constructions in arbitrary Banach spaces. In this general framework one cannot hope to obtain full calculus for such sequential constructions based on Fréchet subgradients; our limiting subdifferential may be even empty for certain locally Lipschitzian functions outside of Asplund spaces. Nevertheless, we are able to prove a number of useful calculus results in general Banach spaces including sum rules, chain rules, product and quotient rules, subdifferentiation of marginal functions, etc. Most (but not all) of these results involve assumptions about strict differentiability of some components in compositions. Moreover, we establish parallel calculus rules for the basic subdifferential and its singular counterpart important for characterizing non-Lipschitzian functions. Note that the principal calculus results obtained in this paper are expressed in the form of *equalities* (not just inclusions) without subdifferential regularity assumptions on all the components.

Let us emphasize that in proving the main results we essentially use the original representations of the basic and singular subdifferentials in terms of sequential limits of the *ε-Fréchet* (not exact Fréchet) counterparts. For the case of Asplund spaces we can always take $\varepsilon = 0$ in these representations (see [22]), but in the general Banach space setting one cannot dismiss $\varepsilon > 0$ from the original limiting constructions without loss of the principal calculus rules.

We also mention another line of infinite dimensional generalizations of nonconvex subdifferential constructions in [16, 17] that was developed by Ioffe under the name of "approximate subdifferentials"; see [5] and references therein. Those constructions, based on *topological* limits of the so-called Dini subdifferentials and ε-subdifferentials, are more complicated and may be broader than our basic sequential constructions even for locally Lipschitzian functions on Banach spaces with Fréchet differentiable renorms; see [22] for details and discussions. On the other hand, the best of such topological constructions, called the G-subdifferential, enjoy full calculus in general Banach spaces. We refer the reader to [5] and the recent paper of Jourani-Thibault [8] for more information.

The rest of the paper is organized as follows. Section 2 deals with definitions and preliminary material. In Section 3 we prove sum rules for the basic and singular subdifferentials. Section 4 contains the main results of the paper about subdifferentiation of marginal functions and chain rules for compositions of functions and mappings. In Section 5 we provide some other calculus formulas for subdifferentials of products, quotients, and minimum functions as well as a nonsmooth version of the mean value theorem.

Throughout the paper we use standard notation except special symbols introduced when they are defined. All spaces considered are Banach whose norms are always denoted by $\|\cdot\|$. For any space X we consider its dual space X^* equipped with the weak-star topology. Recall that $\mathrm{cl}\,\Omega$ means the *closure* of a nonempty set $\Omega \subset X$ while notation cl^* is used for the *weak-star topological closure* in X^*. The adjoint (dual) operator to a linear continuous operator A is denoted by A^*.

In contrast to the case of single-valued mappings $\Phi : X \to Y$, the symbol $\Phi : X \rightrightarrows Y$ stands for a *multifunction* from X into Y with *graph*

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

In this paper we often consider multifunctions Φ from X into the dual space X^*. For such objects the expression

$$\limsup_{x \to \bar{x}} \Phi(x)$$

always means the *sequential Kuratowski-Painlevé upper limit* with respect to the norm topology in X and the weak-star topology in X^*, i.e.,

$$\limsup_{x \to \bar{x}} \Phi(x) := \{x^* \in X^* |\ \exists \text{ sequences } \ x_k \to \bar{x} \ \text{ and } \ x_k^* \xrightarrow{w^*} x^* \\ \text{with } \ x_k^* \in \Phi(x_k) \ \text{ for all } \ k = 1, 2, \ldots\}.$$

If $\varphi :\to \bar{\mathbb{R}} := [-\infty, \infty]$ is an *extended-real-valued function* then, as usual,

$$\mathrm{dom}\,\varphi := \{x \in X \ \text{ with } \ |\varphi(x)| < \infty\}, \quad \mathrm{epi}\,\varphi := \{(x, \mu) \in X \times \mathbb{R} |\ \mu \geq \varphi(x)\}.$$

In this case $\limsup \varphi(x)$ and $\liminf \varphi(x)$ denote the *upper and lower limits* of such (scalar) functions in the classical sense. Depending on context, the symbols $x \xrightarrow{\varphi} \bar{x}$ and $x \xrightarrow{\Omega} \bar{x}$ mean, respectively, that $x \to \bar{x}$ with $\varphi(x) \to \varphi(\bar{x})$ and $x \to \bar{x}$ with $x \in \Omega$.
Throughout the paper we use the convention that $a + \emptyset = \emptyset + b = \emptyset$ for any elements a and b.

2. Basic definitions and properties

This section is devoted to presenting preliminary material on the basic generalized differentiability concepts studied in the paper. Let us start with the definitions of normal elements to arbitrary sets in Banach spaces as appeared in Kruger-Mordukhovich [10, 11].

Definition 2.1. Let Ω be a nonempty subset of the Banach space X and let $\varepsilon \geq 0$.

(i) Given $x \in \text{cl}\,\Omega$, the nonempty set

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

is called the *set of (Fréchet) ε-normals* to Ω at x. When $\varepsilon = 0$ the set (2.1) is a cone which is called the *prenormal cone or Fréchet normal cone* to Ω at x and is denoted by $\hat{N}(x;\Omega)$. If $x \notin \text{cl}\,\Omega$ we set $\hat{N}_\varepsilon(x;\Omega) = \emptyset$ for all $\varepsilon \geq 0$.

(ii) Let $\bar{x} \in \text{cl}\,\Omega$. The nonempty cone

$$N(\bar{x};\Omega) := \limsup_{x \to \bar{x},\ \varepsilon \downarrow 0} \hat{N}_\varepsilon(x;\Omega) \tag{2.2}$$

is called the *normal cone* to Ω at $\bar{x}$. We set $N(\bar{x};\Omega) = \emptyset$ for $\bar{x} \notin \text{cl}\,\Omega$.

Note that in the finite dimensional case $X = \mathbb{R}^n$ the normal cone (2.2) coincides with the one in Mordukhovich [16]:

$$N(\bar{x};\Omega) = \limsup_{x \to \bar{x}} [\text{cone}(x - \Pi(x,\Omega))]$$

where "cone" stands for the conic hull of a set and $\Pi(x,\Omega)$ means the Euclidean projection of x on the closure of Ω.

One can observe that the sets (2.1) are convex for any $\varepsilon \geq 0$ while the normal cone (2.2) is *nonconvex* even in simple finite dimensional situations (e.g., for $\Omega = \text{gph}\,|x|$ at $\bar{x} = 0 \in \mathbb{R}^2$). Moreover, if the space X is infinite dimensional, the weak-star topology in X^* is not necessarily sequential and the sequential upper limit in (2.2) does not ensure either the weak-star closedness or the weak-star sequential closedness of the normal cone. Nevertheless, this limiting construction possesses a broader spectrum of useful properties in comparison with the prenormal (Fréchet) cone and its ε-perturbations; see [1, 3, 4, 6, 9–11, 13–22, 29–31] for more details and related material.

If Ω is convex then

$$\hat{N}_\varepsilon(\bar{x};\Omega) = \{x^* \in X^* |\ \langle x^*, \omega - \bar{x}\rangle \leq \varepsilon\|\omega - \bar{x}\| \text{ for any } \omega \in \Omega\} \quad \forall \varepsilon \geq 0$$

and both normal and prenormal cones at $\bar{x} \in \text{cl}\,\Omega$ coincide with the normal cone in the sense of convex analysis [25].

Now we define the basic subdifferential constructions in this paper geometrically using the normal cone (2.2).

Definition 2.2. Let $\varphi : X \to \bar{\mathbb{R}} := [-\infty, \infty]$ be an extended-real-valued function and $\bar{x} \in \text{dom}\,\varphi$. The sets

$$\partial\varphi(\bar{x}) := \{x^* \in X^* |\ (x^*, -1) \in N((\bar{x}, \varphi(\bar{x}));\ \text{epi}\,\varphi)\}, \tag{2.3}$$

$$\partial^\infty\varphi(\bar{x}) := \{x^* \in X^* |\ (x^*, 0) \in N((\bar{x}, \varphi(\bar{x}));\ \text{epi}\,\varphi)\} \tag{2.4}$$

are called, respectively, the *subdifferential* and the *singular subdifferential* of φ at $\bar{x}$. We let $\partial\varphi(\bar{x}) = \partial^\infty\varphi(\bar{x}) = \emptyset$ if $\bar{x} \notin \text{dom}\varphi$.

The subdifferential (2.3) generalizes the concept of *strict* derivative to the case of nonsmooth functions and is reduced to the subdifferential of convex analysis if φ is convex. For any function $\varphi : X \to \bar{\mathbb{R}}$ on Banach space X the subdifferential (2.3) may be smaller (never bigger) than Clarke's generalized gradient $\partial_C\varphi(\bar{x})$ [2]. This was proved in Kruger [9] by using the representation of $\partial_C\varphi(\bar{x})$ in terms of Rockafellar's subderivative [26]. Furthermore, it follows from [9, Theorem 1] that for functions φ l.s.c. around $\bar{x} \in \operatorname{dom}\varphi$ the subdifferential (2.3) can be represented in the analytic form

$$\partial\varphi(\bar{x}) = \limsup_{x \xrightarrow{\varphi} \bar{x},\ \varepsilon \downarrow 0} \hat{\partial}_\varepsilon \varphi(x) \tag{2.5}$$

based on the ε-subdifferential constructions

$$\hat{\partial}_\varepsilon\varphi(x) := \{x^* \in X^* |\ \liminf_{u \to x} \frac{\varphi(u) - \varphi(x) - \langle x^*, u - x\rangle}{\|u - x\|} \geq -\varepsilon\},\ \ \varepsilon \geq 0. \tag{2.6}$$

When $\varepsilon = 0$ the set (2.6) is called the *presubdifferential or Fréchet subdifferential* of φ at x and is denoted by $\hat{\partial}\varphi(x)$.

One can easily check that the normal cone (2.2) to any set $\Omega \subset X$ at $\bar{x} \in \Omega$ is expressed in the subdifferential forms

$$N(\bar{x}; \Omega) = \partial\delta(\bar{x}, \Omega) = \partial^\infty\delta(\bar{x}, \Omega) \tag{2.7}$$

where $\delta(\cdot, \Omega)$ is the *indicator function* of Ω, i.e., $\delta(x, \Omega) = 0$ if $x \in \Omega$ and $\delta(x, \Omega) = \infty$ if $x \notin \Omega$.

In this paper we also use the following derivative-like concept for multifunctions related to the normal cone (2.2).

Definition 2.3. Let $\Phi : X \rightrightarrows Y$ be a multifunction between Banach spaces X and Y, and let $(\bar{x}, \bar{y}) \in \operatorname{cl} \operatorname{gph} \Phi$. The multifunction $D^*\Phi(\bar{x}, \bar{y})$ from Y^* into X^* defined by

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

is called the *coderivative* of Φ at $(\bar{x}, \bar{y})$. The symbol $D^*\Phi(\bar{x})(y^*)$ is used in (2.8) when Φ is single-valued at $\bar{x}$ and $\bar{y} = \Phi(\bar{x})$. We let $D^*\Phi(\bar{x}, \bar{y})(y^*) = \emptyset$ if $(\bar{x}, \bar{y}) \notin \operatorname{cl} \operatorname{gph} \Phi$.

The coderivative (2.8) turns out to be a generalization of the classical strict differentiability concept to the case of nonsmooth mappings and multifunctions. Recall that a mapping $\Phi : X \to Y$ single-valued around $\bar{x}$ is called *strictly differentiable* at $\bar{x}$ with the derivative $\Phi'(\bar{x})$ if

$$\lim_{x \to \bar{x}, u \to \bar{x}} \frac{\Phi(x) - \Phi(u) - \Phi'(\bar{x})(x - u)}{\|x - u\|} = 0. \tag{2.9}$$

It is well known that any mapping Φ continuously Fréchet differentiable around $\bar{x}$ is strictly differentiable at $\bar{x}$ but not vice versa. Based on the definitions, one can derive that

$$D^*\Phi(\bar{x})(y^*) = (\Phi'(\bar{x}))^* y^* \quad \forall y^* \in Y^*$$

if Φ is strictly differentiable at $\bar{x}$. Therefore, the coderivative (2.8) is consistent with the adjoint linear operator to the classical strict derivative.

3. Sum rules

In this section we obtain important calculus results for nonconvex subdifferentials in Section 2 related to representation of subgradients for sums of functions.

Theorem 3.1.

(i) *Let $\varphi : X \to \mathbb{R}$ be strictly differentiable at $\bar{x}$ and let $\psi : X \to \bar{\mathbb{R}}$ be l.s.c. around this point. Then one has*

$$\partial(\varphi + \psi)(\bar{x}) = \varphi'(\bar{x}) + \partial\psi(\bar{x}). \tag{3.1}$$

(ii) *Let $\varphi : X \to \mathbb{R}$ be Lipschitz continuous around $\bar{x}$ and let $\psi : X \to \bar{\mathbb{R}}$ be l.s.c. around this point. Then*

$$\partial^\infty(\varphi + \psi)(\bar{x}) = \partial^\infty\psi(\bar{x}). \tag{3.2}$$

In particular, one has $\partial^\infty\varphi(\bar{x}) = \{0\}$ for any locally Lipschitzian function.

Proof. Let us establish (i). First we verify that

$$\partial(\varphi + \psi)(\bar{x}) \subset \varphi'(\bar{x}) + \partial\psi(\bar{x}) \tag{3.3}$$

for $\bar{x} \in \operatorname{dom} \psi$. By definition (2.9) of the strict derivative, for any sequence $\gamma_\nu \downarrow 0$ there exists a sequence $\delta_\nu \downarrow 0$ such that

$$|\varphi(z) - \varphi(x) - \langle \varphi'(\bar{x}), z - x\rangle| \le \gamma_\nu \|x - z\| \quad \forall x, z \in B_{\delta_\nu}(\bar{x}), \;\; \nu = 1, 2, \ldots. \tag{3.4}$$

Now let us consider $x^* \in \partial(\varphi + \psi)(\bar{x})$. Using representation (2.5), we find sequences $x_k \to \bar{x}$, $(\varphi + \psi)(x_k) \to (\varphi + \psi)(\bar{x})$, $x_k^* \xrightarrow{w^*} x^*$, and $\varepsilon_k \downarrow 0$ as $k \to \infty$ such that

$$x_k^* \in \hat{\partial}_{\varepsilon_k}(\varphi + \psi)(x_k) \quad \forall k = 1, 2, \ldots. \tag{3.5}$$

Due to $x_k \to \bar{x}$ as $k \to \infty$ one can choose a sequence $k_1 < k_2 < \ldots < k_\nu < \ldots$ of positive integers satisfying $\|x_{k_\nu} - \bar{x}\| \le \delta_\nu/2$ for all $\nu = 1, 2, \ldots$. By virtue of (3.5) we pick $0 < \eta_\nu \le \delta_\nu/2$ such that

$$\begin{aligned} \psi(x) - \psi(x_{k_\nu}) + \varphi(x) - \varphi(x_{k_\nu}) - \langle x_{k_\nu}^*, x - x_{k_\nu}\rangle \;\ge\; -2\varepsilon_{k_\nu}\|x - x_{k_\nu}\| \\ \forall x \in B_{\eta_\nu}(x_{k_\nu}), \;\; \nu = 1, 2, \ldots. \end{aligned} \tag{3.6}$$

Observe that $x \in B_{\delta_\nu}(\bar{x})$ whenever $x \in B_{\eta_\nu}(x_{k_\nu})$ for all $\nu = 1, 2, \ldots$ It follows from (3.4) and (3.6) that

$$\begin{aligned} \psi(x) - \psi(x_{k_\nu}) - \langle x_{k_\nu}^* - \varphi'(\bar{x}), x - x_{k_\nu}\rangle \;\ge\; -(2\varepsilon_{k_\nu} + \gamma_\nu)\|x - x_{k_\nu}\| \\ \forall x \in B_{\eta_\nu}(x_{k_\nu}), \;\; \nu = 1, 2, \ldots. \end{aligned}$$

This implies that

$$x_{k_\nu}^* - \varphi'(\bar{x}) \in \hat{\partial}_{\bar{\varepsilon}_\nu}\psi(x_{k_\nu}) \;\text{ with }\; \bar{\varepsilon}_\nu := 2\varepsilon_{k_\nu} + \gamma_\nu, \;\; \nu = 1, 2, \ldots. \tag{3.7}$$

Taking into account that $\psi(x_{k_\nu}) \to \psi(\bar{x})$ as $\nu \to \infty$, we derive from (2.5) and (3.7) that $x^* - \varphi'(\bar{x}) \in \partial\psi(\bar{x})$ and, therefore, get (3.3). Then applying (3.3) to the sum of functions

$(\psi+\varphi)+(-\varphi)$, one arrives at the inclusion opposite to (3.3) that establishes equality (3.1) for the case of $\bar{x}\in\operatorname{dom}\psi$. The case of $\bar{x}\notin\operatorname{dom}\psi$ is trivial. This ends the proof of assertion (i).

To prove (ii) we just check the inclusion

$$\partial^\infty(\varphi+\psi)(\bar{x})\subset\partial^\infty\psi(\bar{x}) \tag{3.8}$$

under the assumptions made. This implies equality (3.2) similarly to (i).
Let us establish (3.8) for $\bar{x}\in\operatorname{dom}\psi$. To furnish this we consider $x^*\in\partial^\infty(\varphi+\psi)(\bar{x})$ and use definition (2.4). In this way one can find sequences $x_k^*\xrightarrow{w^*}x^*,\ \alpha_k\to 0,\ x_k\to\bar{x},\ r_k\to(\varphi+\psi)(\bar{x}),\ \varepsilon_k\downarrow 0$, and $\delta_k\downarrow 0$ such that $r_k\geq\varphi(x_k)+\psi(x_k)$ and

$$\langle x_k^*,x-x_k\rangle+\alpha_k(r-r_k)\leq 2\varepsilon_k(\|x-x_k\|+|r-r_k|) \tag{3.9}$$

for all $(x,r)\in\operatorname{epi}(\varphi+\psi)$ with $x\in B_{\delta_k}(x_k)$ and $|r-r_k|\leq\delta_k,\ k=1,2,\ldots$.
Let l be a Lipschitz modulus of φ around $\bar{x}$. We denote $\bar{\delta}_k:=\delta_k/2(l+1)$ and $\bar{r}_k:=r_k-\varphi(x_k),\ k=1,2,\ldots$. It is easy to see that $\bar{r}_k\geq\psi(x_k)$ for all k and $\bar{r}_k\to\psi(\bar{x})$ as $k\to\infty$. Observe that for any fixed $k=1,2,\ldots$ and any $(x,\bar{r})\in\operatorname{epi}\psi$ satisfying $x\in B_{\bar{\delta}_k}(x_k)$ and $|\bar{r}-\bar{r}_k|\leq\bar{\delta}_k$ one has

$$(x,\bar{r}+\varphi(x))\in\operatorname{epi}(\varphi+\psi)\ \text{ and }\ |(\bar{r}+\varphi(x))-r_k|\leq\delta_k.$$

By virtue of (3.9) we get

$$\langle x_k^*,x-x_k\rangle+\alpha_k(\bar{r}-\bar{r}_k)\leq\bar{\varepsilon}_k(\|x-x_k\|+|\bar{r}-\bar{r}_k|)\ \text{ with }\ \bar{\varepsilon}_k:=2\varepsilon_k(1+l)+|\alpha_k|l$$

for any $(x,\bar{r})\in\operatorname{epi}\psi$ with $x\in B_{\bar{\delta}_k}(x_k)$ and $|\bar{r}-\bar{r}_k|\leq\bar{\delta}_k,\ k=1,2,\ldots$. This implies

$$(x_k^*,\alpha_k)\in\hat{N}_{\bar{\varepsilon}_k}((x_k,\bar{r}_k);\operatorname{epi}\psi).$$

Now using definition (2.4), we conclude that $x^*\in\partial\psi(\bar{x})$ which proves (3.8) and, therefore, equality (3.2). Letting $\psi\equiv 0$ in (3.2), one has $\partial^\infty\varphi(\bar{x})=\{0\}$ for any function φ Lipschitz continuous around $\bar{x}$. This completes the proof of the theorem. □

Remark 3.2. In [30], Thibault proved an analogue of assertion (ii) in Theorem 3.1 using another definition of singular subgradients

$$\partial_1^\infty\varphi(\bar{x}):=\limsup_{x\xrightarrow{\varphi}\bar{x};\ \varepsilon,\lambda\downarrow 0}\lambda\hat{\partial}_\varepsilon\varphi(x). \tag{3.10}$$

In the case of Asplund spaces X both constructions (2.4) and (3.10) are equivalent (see [22, Theorem 2.9]) but in the general Banach space setting they may be different.

4. Subdifferentiation of marginal functions and chain rules

In this section we obtain the main results of the paper related to subdifferentiation of the so-called *marginal functions* of the form

$$m(x):=\inf\{\varphi(x,y)|\ y\in\Phi(x)\} \tag{4.1}$$

where $\varphi : X \times Y \to \bar{\mathbb{R}}$ is an extended-real-valued function and $\Phi : X \rightrightarrows Y$ is a multifunction between Banach spaces. This class includes *value functions* in parametric optimization problems that play an important role in nonsmooth analysis, optimization theory, and various applications. It is well known that marginal functions (4.1) are usually *nonsmooth* even for smooth φ and simple constant sets Φ. Therefore, to compute generalized (in some sense) derivatives for (4.1) is a challenging issue which has many significant applications to optimization, sensitivity analysis, etc. We refer the reader to [2, 3, 19, 22, 29, 30] for various results in this direction and further bibliographies. Our goal is to provide formulas that express the subdifferential (2.3) and singular subdifferential (2.4) of (4.1) in terms of differential constructions for φ and Φ in general Banach spaces.

Note that when $\Phi : X \to Y$ happens to be a single-valued mapping the marginal function (4.1) is just a *composition* of φ and Φ which we denote by

$$(\varphi \circ \Phi)(x) := \varphi(x, \Phi(x)). \tag{4.2}$$

In this case subdifferential formulas for (4.1) are related to *chain rules* for subdifferentiation of compositions.

Let us consider the *minimum set*

$$M(x) := \{y \in \Phi(x) |\ \varphi(x, y) = m(x)\} \tag{4.3}$$

associated with (4.1). In the sequel we need the following *lower semicompactness* property for the multifunction $M : X \rightrightarrows Y$ around the reference point $\bar{x}$: there exists a neighborhood U of $\bar{x}$ such that for any $x \in U$ and any sequence $x_k \to x$ as $k \to \infty$ there is a sequence $y_k \in M(x_k), k = 1, 2, \ldots$, which contains a subsequence convergent in the norm topology of Y. Obviously, any multifunction lower semicontinuous around $\bar{x}$ is lower semicompact around this point. It always happens, in particular, when Φ is single-valued and continuous around $\bar{x}$. If $\dim Y < \infty$ the lower semicompactness property is inherent in every multifunction whose values are nonempty and uniformly bounded near the reference point.

First we present results on subdifferentiation of (4.1) for the case of general multifunctions Φ between Banach spaces.

Theorem 4.1. *Let $\Phi : X \rightrightarrows Y$ have closed graph, let M in* (4.3) *be lower semicompact around $\bar{x} \in \operatorname{dom} m$, and let φ be l.s.c. on* $\operatorname{gph} \Phi$ *and strictly differentiable at $(\bar{x}, \bar{y})$ for any $\bar{y} \in M(\bar{x})$. Then one has*

$$\partial m(\bar{x}) \subset \bigcup_{\bar{y} \in M(\bar{x})} [\varphi'_x(\bar{x}, \bar{y}) + D^*\Phi(\bar{x}, \bar{y})(\varphi'_y(\bar{x}, \bar{y}))], \tag{4.4}$$

$$\partial^\infty m(\bar{x}) \subset \bigcup [D^*\Phi(\bar{x}, \bar{y})(0) |\ \bar{y} \in M(\bar{x})]. \tag{4.5}$$

Proof. First we check that the marginal function (4.1) is l.s.c. around $\bar{x}$ under the assumptions made. Indeed, let U be a neighborhood of $\bar{x}$ from the local semicompactness condition for M. Taking any $x \in U$ and sequence $x_k \to x$, we find a sequence $y_k \in M(x_k)$ that contains a subsequence convergent to some point $y \in Y$ with $(x, y) \in \operatorname{gph} \Phi$. Since φ is l.s.c. on $\operatorname{gph} \Phi$ one has

$$m(x) \le \varphi(x, y) \le \liminf_{k\to\infty} \varphi(x_k, y_k) = \liminf_{k\to\infty} m(x_k)$$

that ensures the lower semicontinuity of $m(x)$ around $\bar{x}$.

Now let us consider a function $f : X \times Y \to \bar{\mathbb{R}}$ defined by

$$f(x,y) := \varphi(x,y) + \delta((x,y), \mathrm{gph}\, \Phi) \tag{4.6}$$

and let us prove that

$$\partial m(\bar{x}) \subset \{x^* \in X^* |\ (x^*, 0) \in \partial f(\bar{x}, \bar{y}), \bar{y} \in M(\bar{x})\}. \tag{4.7}$$

Taking $x^* \in \partial m(\bar{x})$ and using representation (2.5), we find sequences $x_k \to \bar{x}$, $x_k^* \xrightarrow{w^*} x^*$, and $\varepsilon_k \downarrow 0$ such that $m(x_k) \to m(\bar{x})$ as $k \to \infty$ and $x_k^* \in \hat{\partial}_{\varepsilon_k} m(x_k)$ for all $k = 1, 2, \ldots$. Therefore, there exists a sequence $\rho_k \downarrow 0$ as $k \to \infty$ with

$$\langle x_k^*, x - x_k \rangle \le m(x) - m(x_k) + 2\varepsilon_k \|x - x_k\| \quad \forall x \in B_{\rho_k}(x_k),\ k = 1, 2 \ldots.$$

By definitions (4.1) and (4.2) of m and M, for any $y_k \in M(x_k)$ one has

$$\langle (x_k^*, 0), (x, y) - (x_k, y_k) \rangle \le f(x,y) - f(x_k, y_k) + 2\varepsilon_k (\|x - x_k\| + \|y - y_k\|)$$

for all $(x, y) \in B_{\rho_k}((x_k, y_k))$, $k = 1, 2, \ldots$. Due to (2.5) and (4.6) this yields

$$(x_k^*, 0) \in \hat{\partial}_{2\varepsilon_k} f(x_k, y_k) \quad \forall y_k \in M(x_k),\ k = 1, 2 \ldots. \tag{4.8}$$

Now using the lower semicompactness of M around $\bar{x}$, one can select a sequence $y_k \in M(x_k)$ which contains a subsequence convergent to some point $\bar{y} \in \Phi(\bar{x})$. Since $m(x_k) \to m(\bar{x})$ one has $\bar{y} \in M(\bar{x})$ and $f(x_k, y_k) \to f(\bar{x}, \bar{y})$ as $k \to \infty$. By virtue of (4.8) and (2.5) we conclude that $(x^*, 0) \in \partial f(\bar{x}, \bar{y})$ and finish the proof of inclusion (4.7).

Next we note that the function f in (4.6) is represented as the sum of two functions satisfying the assumptions of Theorem 3.1(i) at the point $(\bar{x}, \bar{y}) \in \mathrm{gph}\, M$. Employing the sum rule (3.1) and taking into account the first equality in (2.7) as well as Definition 2.3 of the coderivative, one can easily arrive at inclusion (4.4).

It remains to verify (4.5). To furnish this we first prove that

$$\partial^\infty m(\bar{x}) \subset \{x^* \in X^* |\ (x^*, 0) \in \partial^\infty f(\bar{x}, \bar{y}),\ \bar{y} \in M(\bar{x})\}. \tag{4.9}$$

Picking any $x^* \in \partial^\infty m(\bar{x})$ and using definition (2.4), one gets sequences $x_k^* \xrightarrow{w^*} x^*$, $\lambda_k \to 0$, $x_k \to \bar{x}$, $r_k \to m(\bar{x})$, $\varepsilon_k \downarrow 0$, and $\rho_k \downarrow 0$ as $k \to \infty$ such that $r_k \ge m(x_k)$ and

$$\langle x_k^*, x - x_k \rangle + \lambda_k (r - r_k) \le 2\varepsilon_k (\|x - x_k\| + |r - r_k|) \tag{4.10}$$

for all $(x, r) \in \mathrm{epi}\ m$ with $x \in B_{\rho_k}(x_k)$ and $|r - r_k| \le \rho_k$, $k = 1, 2, \ldots$. Taking into account representation (4.6) and choosing any $y_k \in M(x_k)$, we note that $r_k \ge f(x_k, y_k)$ and (4.10) implies the estimate

$$\begin{aligned} &\langle (x_k^*, 0), (x, y) - (x_k, y_k) \rangle + \lambda_k (r - r_k) \\ &\le 2\varepsilon_k (\|x - x_k\| + \|y - y_k\| + |r - r_k|) \end{aligned}$$

for all $((x,y),r) \in \mathrm{epi}\, f$ with $(x,y) \in B_{\rho_k}((x_k,y_k))$ and $|r - r_k| \le \rho_k,\ k = 1,2,\ldots$. This yields

$$((x_k^*,0),\lambda_k) \in \hat{N}_{2\varepsilon_k}(((x_k,y_k),r_k),\mathrm{epi}\, f), \quad k = 1,2,\ldots. \tag{4.11}$$

Now using the lower semicompactness property of M around $\bar{x}$, we select a sequence $y_k \in M(x_k)$ that contains a subsequence convergent to some point $\bar{y} \in \Phi(\bar{x})$. Since $r_k \to m(\bar{x})$, we conclude that $\bar{y} \in M(\bar{x})$ and $r_k \to f(\bar{x},\bar{y})$ as $k \to \infty$. Passing to the limit in (4.11) as $k \to \infty$, one gets $(x^*,0) \in \partial^\infty f(\bar{x},\bar{y})$ and arrives at the inclusion (4.9).
To obtain the final inclusion (4.5) from (4.9) we apply Theorem 3.1(ii) to the sum of functions in (4.6) taking into account the second representation of the normal cone in (2.7) and construction (2.8). This ends the proof of the theorem. □

Now we consider the case when Φ is single-valued and locally Lipschitzian around $\bar{x}$. Then inclusion (4.5) is trivial but (4.4) carries some chain rule information about the subdifferential (2.3) of composition (4.2). Let us provide more precise and convenient chain rule in the form of equality using instead of the coderivative of Φ the subdifferential of its *Lagrange scalarization*

$$\langle y^*,\Phi\rangle(x) := \langle y^*,\Phi(x)\rangle \quad \forall y^* \in Y^*,\ x \in X.$$

It is proved in [22] that one always has

$$\partial\langle y^*,\Phi\rangle(\bar{x}) \subset D^*\Phi(\bar{x})(y^*) \quad \forall y^* \in Y^*$$

for any mapping $\Phi : X \to Y$ between Banach spaces which is single-valued and continuous around $\bar{x}$.

Theorem 4.2. *Let $\Phi : X \to Y$ be single-valued and Lipschitz continuous around $\bar{x}$ and let $\varphi : X \times Y \to \mathbb{R}$ be strictly differentiable at $(\bar{x},\Phi(\bar{x}))$. Then one has*

$$\partial(\varphi \circ \Phi)(\bar{x}) = \varphi'_x(\bar{x},\bar{y}) + \partial\langle \varphi'_y(\bar{x},\bar{y}),\Phi\rangle(\bar{x}) \ \text{ with } \ \bar{y} = \Phi(\bar{x}). \tag{4.12}$$

Proof. Due to the strict differentiability of φ at $(\bar{x},\bar{y})$ with $\bar{y} = \Phi(\bar{x})$, for any sequence $\gamma_\nu \downarrow 0$ there exists a sequence $\rho_\nu \downarrow 0$ such that

$$\begin{aligned} &|\varphi(u,\Phi(u)) - \varphi(x,\Phi(x)) - \langle \varphi'_x(\bar{x},\bar{y}),u - x\rangle - \langle \varphi'_y(\bar{x},\bar{y}),\Phi(u) - \Phi(x)\rangle| \\ &\le \gamma_\nu(\|u - x\| + \|\Phi(u) - \Phi(x)\|) \quad \forall x,u \in B_{\rho_\nu}(\bar{x}),\ \nu = 1,2,\ldots. \end{aligned} \tag{4.13}$$

Let us pick any $x^* \in \partial(\varphi\circ\Phi)(\bar{x})$. By virtue of (2.5) one gets sequences $x_k \to \bar{x}$, $x_k^* \xrightarrow{w^*} x^*$, and $\varepsilon_k \downarrow 0$ such that $(\varphi\circ\Phi)(x_k) \to (\varphi\circ\Phi)(\bar{x})$ and

$$x_k^* \in \hat{\partial}_{\varepsilon_k}(\varphi\circ\Phi)(x_k) \quad \forall k = 1,2,\ldots. \tag{4.14}$$

Since $x_k \to \bar{x}$ as $k \to \infty$ we select a sequence $k_1 < k_2 < \ldots < k_\nu < \ldots$ of positive integers satisfying $\|x_{k_\nu} - \bar{x}\| \le \rho_\nu/2$ for all ν. By virtue of (4.14) one can choose $0 < \eta_\nu \le \rho_\nu/2$ such that

$$\begin{aligned} &\varphi(x,\Phi(x)) - \varphi(x_{k_\nu},\Phi(x_{k_\nu})) - \langle x_{k_\nu}^*,x - x_{k_\nu}\rangle \\ &\ge -2\varepsilon_{k_\nu}\|x - x_{k_\nu}\| \quad \forall x \in B_{\eta_\nu}(x_{k_\nu}),\ \ \nu = 1,2,\ldots. \end{aligned} \tag{4.15}$$

Let l be a Lipschitz modulus of Φ in some neighborhood of $\bar{x}$ that contains all x_k for k sufficiently large. Noting that $x \in B_{\rho_\nu}(\bar{x})$ if $x \in B_{\eta_\nu}(x_{k_\nu})$ for all ν, we derive from (4.13) and (4.15) that

$$\begin{aligned}&\langle \varphi'_y(\bar{x},\bar{y}), \Phi(x)\rangle - \langle \varphi'_y(\bar{x},\bar{y}), \Phi(x_{k_\nu})\rangle - \langle x^*_{k_\nu} - \varphi'_x(\bar{x},\bar{y}), x - x_{k_\nu}\rangle \\ &\geq -[2\varepsilon_{k_\nu} + \gamma_\nu(l+1)]\|x - x_{k_\nu}\| \quad \forall x \in B_{\eta_\nu}(x_{k_\nu}),\ \nu = 1, 2, \ldots .\end{aligned}$$

From here and (2.6) one has

$$x^*_{k_\nu} - \varphi'_x(\bar{x},\bar{y}) \in \hat{\partial}_{\bar{\varepsilon}_\nu}\langle \varphi'_y(\bar{x},\bar{y}), \Phi(x_{k_\nu})\rangle \ \text{ with } \ \bar{\varepsilon}_\nu := 2\varepsilon_{k_\nu} + \gamma_\nu(l+1) \tag{4.16}$$

for all $\nu = 1, 2, \ldots$. Passing to the limit in (4.16) as $\nu \to \infty$ and using (2.5), we obtain

$$x^* - \varphi'_x(\bar{x},\bar{y}) \in \partial\langle \varphi'_y(\bar{x},\bar{y}), \Phi(\bar{x})\rangle$$

that proves the inclusion "$\subset$" in (4.12). To verify the opposite inclusion in (4.12) we employ the similar arguments starting with a point $x^* \in \langle \varphi'_y(\bar{x},\bar{y}), \Phi\rangle(\bar{x})$. This ends the proof of the theorem. □

Now let us obtain chain rules for the basic and singular subdifferentials of compositions (4.2) in the case of nonsmooth functions φ and strictly differentiable mappings Φ between arbitrary Banach spaces.

Theorem 4.3.

(i) *Let $\Phi : X \to Y$ be strictly differentiable at $\bar{x}$ with $\Phi'(\bar{x})$ invertible (i.e., surjective and one-to-one) and let $\varphi : X \times Y \to \bar{\mathbb{R}}$ be represented as*

$$\varphi(x,y) = \varphi_1(x) + \varphi_2(y)$$

where φ_1 is strictly differentiable at $\bar{x}$ and φ_2 is l.s.c. around $\bar{y} = \Phi(\bar{x})$. Then one has

$$\partial(\varphi \circ \Phi)(\bar{x}) = \varphi'_1(\bar{x}) + (\Phi'(\bar{x}))^*\partial\varphi_2(\bar{y}).$$

(ii) *Let φ and Φ satisfy the assumptions in (i) except that φ_1 is Lipschitz continuous around $\bar{x}$. Then*

$$\partial^\infty(\varphi \circ \Phi)(\bar{x}) = (\Phi'(\bar{x}))^*\partial^\infty\varphi_2(\bar{y}).$$

Proof. In the case considered one has

$$(\varphi \circ \Phi)(x) = \varphi_1(x) + \varphi_2(\Phi(x)). \tag{4.17}$$

To prove (i) we use Theorem 3.1(i) that yields

$$\partial(\varphi \circ \Phi)(\bar{x}) = \varphi'_1(\bar{x}) + \partial(\varphi_2 \circ \Phi)(\bar{x}).$$

Thus we should show that

$$\partial(\varphi_2 \circ \Phi)(\bar{x}) = (\Phi'(\bar{x}))^*\partial\varphi_2(\bar{y}) \ \text{ with } \ \bar{y} = \Phi(\bar{x}) \tag{4.18}$$

First we verify the inclusion "$\supset$" in (4.18). Let us consider any $y^* \in \partial\varphi_2(\bar{y})$ and find sequences $y_k \to \bar{y}$, $y_k^* \xrightarrow{w^*} y^*$, and $\varepsilon_k \downarrow 0$ such that $\varphi_2(y_k) \to \varphi_2(\bar{y})$ as $k \to \infty$ and $y_k^* \in \hat{\partial}_{\varepsilon_k}\varphi_2(y_k)$ for all $k = 1, 2, \ldots$ Employing the inverse mapping theorem for strictly differentiable mappings [12], we conclude that Φ^{-1} is locally single-valued and strictly differentiable at $\bar{y}$ with the strict derivative $(\Phi'(\bar{x}))^{-1}$ at this point. Therefore, Φ is a local homeomorphism around $\bar{x}$. Now using the procedure similar to the proof of Theorem 4.2, for any sequence $\gamma_\nu \downarrow 0$ one gets $\rho_\nu \downarrow 0$, $x_\nu \to \bar{x}$, and a subsequence $k_\nu \to \infty$ as $\nu \to \infty$ such that

$$\varphi_2(\Phi(x)) - \varphi_2(\Phi(x_\nu)) - \langle(\Phi'(\bar{x}))^* y_{k_\nu}^*, x - x_\nu\rangle \geq -(2l\varepsilon_{k_\nu} + \gamma_\nu\|y_{k_\nu}^*\|)\|x - x_\nu\|$$

for any $x \in B_{\rho_\nu}(x_\nu)$, where l is a Lipschitz modulus of Φ in some neighborhood of $\bar{x}$ containing all $B_{\rho_\nu}(x_\nu)$ as $\nu = 1, 2, \ldots$. Due to (2.5) the latter implies the inclusion "$\supset$" in (4.18).

To verify the opposite inclusion we represent φ_2 in the form

$$\varphi_2(y) = (\psi \circ \Phi^{-1})(y) \ \text{ with } \ \psi(x) := (\varphi_2 \circ \Phi)(x).$$

Now applying the inclusion "$\supset$" in (4.18) to the composition $\psi \circ \Phi^{-1}$ and taking into account that $(\Phi^{-1})'(\bar{y}) = (\Phi'(\bar{x}))^{-1}$, we obtain the inclusion "$\subset$" in (4.18). This ends the proof of assertion (i) in the theorem.

It remains to establish (ii). Under the assumptions made one has

$$\partial^\infty(\varphi \circ \Phi)(\bar{x}) = \partial^\infty(\varphi_2 \circ \Phi)(\bar{x})$$

due to Theorem 3.1(ii). The equality

$$\partial^\infty(\varphi_2 \circ \Phi)(\bar{x}) = (\Phi'(\bar{x}))^* \partial^\infty\varphi_2(\Phi(\bar{x}))$$

can be proved similarly to (4.18) using definitions (2.4) and (2.2); cf. the proof of (4.5) in Theorem 4.1. $\square$

In conclusion of this section we present a useful corollary of Theorem 4.3 providing a representation of the normal cone (2.2) to sets of the form

$$\Phi^{-1}(\Lambda) := \{x \in X |\ \Phi(x) \in \Lambda)\}$$

where $\Phi : X \to Y$ and $\Lambda \subset Y$.

Corollary 4.4. *Let Φ be strictly differentiable at $\bar{x}$ with $\Phi'(\bar{x})$ invertible, and let Λ be closed around the point $\bar{y} = \Phi(\bar{x}) \in \Lambda$. Then one has*

$$N(\bar{x}; \Phi^{-1}(\Lambda)) = (\Phi'(\bar{x}))^* N(\bar{y}; \Lambda).$$

Proof. This follows directly from Theorem 4.3(i) when $\varphi_1 = 0$ and $\varphi_2 = \delta(\cdot, \Lambda)$. $\square$

5. Other calculus rules

In this concluding section of the paper we obtain some additional calculus formulas for the subdifferentials (2.3) and (2.4). Let us start with the following *product rules* involving locally Lipschitzian functions.

Theorem 5.1. *Let $\varphi_i : X \to \mathbb{R}$, $i = 1, 2$, be Lipschitz continuous around $\bar{x}$. Then one has*

$$\partial(\varphi_1 \cdot \varphi_2)(\bar{x}) = \partial(\varphi_2(\bar{x})\varphi_1 + \varphi_1(\bar{x})\varphi_2)(\bar{x}). \tag{5.1}$$

If, in addition, one of the functions (say φ_1) is strictly differentiable at $\bar{x}$, then

$$\partial(\varphi_1 \cdot \varphi_2)(\bar{x}) = \varphi_1'(\bar{x})\varphi_2(\bar{x}) + \partial(\varphi_1(\bar{x})\varphi_2)(\bar{x}). \tag{5.2}$$

Proof. To obtain (5.1) we apply the chain rule in Theorem 4.2 with a smooth function $\varphi : \mathbb{R}^2 \to \mathbb{R}$ and a Lipschitzian mapping $\Phi : X \to \mathbb{R}^2$ defined by

$$\Phi(x) := (\varphi_1(x), \varphi_2(x)) \ \text{ and } \ \varphi(y_1, y_2) := y_1 \cdot y_2.$$

When φ_1 is strictly differentiable at $\bar{x}$, equality (5.2) follows from (5.1) by virtue of Theorem 3.1(i). □

In the same way we obtain the following *quotient rules* for subdifferentials of Lipschitzian functions.

Theorem 5.2. *Let $\varphi_i : X \to \mathbb{R}$, $i = 1, 2$, be Lipschitz continuous around $\bar{x}$ and let $\varphi_2(\bar{x}) \neq 0$. Then one has*

$$\partial(\varphi_1/\varphi_2)(\bar{x}) = \frac{\partial(\varphi_2(\bar{x})\varphi_1 - \varphi_1(\bar{x})\varphi_2)(\bar{x})}{[\varphi_2(\bar{x})]^2}. \tag{5.3}$$

If, in addition, φ_1 is strictly differentiable at $\bar{x}$, then

$$\partial(\varphi_1/\varphi_2)(\bar{x}) = \frac{\varphi_1'(\bar{x})\varphi_2(\bar{x}) + \partial(-\varphi_1(\bar{x})\varphi_2)(\bar{x})}{[\varphi_2(\bar{x})]^2}. \tag{5.4}$$

Proof. Apply Theorem 4.2 to the composition $(\varphi_1/\varphi_2)(x) = (\varphi \circ \Phi)(x)$ where $\Phi : X \to \mathbb{R}^2$ and $\varphi : \mathbb{R}^2 \to \mathbb{R}$ are defined by

$$\Phi(x) := (\varphi_1(x), \varphi_2(x)) \ \text{ and } \ \varphi(y_1, y_2) := y_1/y_2.$$

□

Remark 5.3. (a) One can obtain some product and quotient rules for the basic and singular subdifferentials of l.s.c. functions using Theorem 4.1 instead of Theorem 4.2 in the arguments above.
(b) The right-hand side of the product rule (5.2) is equal to

$$\varphi_1'(\bar{x})\varphi_2(\bar{x}) + \varphi_1(\bar{x})\partial\varphi_2(\bar{x})$$

when $\varphi_1(\bar{x}) \geq 0$. (A similar conclusion holds for the numerator in the right-hand side of the quotient rule (5.4)). This follows from the obvious equality $\partial(\alpha\varphi)(\bar{x}) = \alpha\partial\varphi(\bar{x})$ valid for any φ and $\alpha \geq 0$. On the other hand, one has

$$\partial(\alpha\varphi)(\bar{x}) = \alpha\partial^+\varphi(\bar{x}) \text{ for } \alpha < 0$$

where $\partial^+\varphi(\bar{x}) := -\partial(-\varphi)(\bar{x})$ is the so-called *superdifferential* of φ at $\bar{x}$ that can be equivalently defined in terms of the normal cone (2.2) to the hypograph of φ or similarly to (2.5); see [18]. Note that the subdifferential and superdifferential may be considerably different for nonsmooth functions (e.g., for $\varphi(x) = |x|$ where $\partial\varphi(0) = [-1, 1]$ but $\partial^+\varphi(0) = \{-1, 1\}$).

Corollary 5.4. *Let $\varphi : X \to \mathbb{R}$ be Lipschitz continuous around $\bar{x}$ with $\varphi(\bar{x}) \neq 0$. Then one has*

$$\partial(1/\varphi)(\bar{x}) = -\frac{\partial^+\varphi(\bar{x})}{\varphi^2(\bar{x})}.$$

Proof. This follows from Theorem 5.2 with $\varphi_1 = 1$ and $\varphi_2 = \varphi$. □

Now we consider a collection of extended-real-valued functions φ_i, $i = 1, \ldots, n$ $(n \geq 2)$, and define the *minimum function*

$$(\min\varphi_i)(x) := \min\{\varphi_i(x)|\ i = 1, \ldots, n\}.$$

Denoting by

$$I(x) := \{i \in \{1, \ldots, n\}|\ \varphi_i(x) = (\min\varphi_i)(\bar{x})\},$$

we have the following subdifferentiation formula for the mimimum function in any Banach space.

Theorem 5.5. *Let functions $\varphi_i : X \to \bar{\mathbb{R}}$ be l.s.c. at $\bar{x}$ for $i \notin I(\bar{x})$ and be l.s.c. around this point for $i \in I(\bar{x})$ together with* $\min\varphi_i$. *Then one has*

$$\partial(\min\varphi_i)(\bar{x}) \subset \bigcup\{\partial\varphi_i(\bar{x})|\ i \in I(\bar{x})\}. \tag{5.5}$$

Proof. To justify (5.5) let us consider any sequence $\{x_k\} \subset X$ such that $x_k \to \bar{x}$ and $\varphi_i(x_k) \to (\min\varphi_i)(\bar{x})$ for $i \in I(\bar{x})$ as $k \to \infty$. Using the l.s.c. of φ_i at $\bar{x}$ for $i \notin I(\bar{x})$, one can conclude that $I(x_k) \subset I(\bar{x})$. By virtue of (2.6) this yields

$$\hat{\partial}_\varepsilon(\min\varphi_i)(x_k) \subset \bigcup\{\hat{\partial}_\varepsilon\varphi_i(x_k)|\ i \in I(\bar{x})\}$$

for any $\varepsilon \geq 0$ and $k = 1, 2, \ldots$ The latter implies (5.5) due to representation (2.5) for functions l.s.c. around $\bar{x}$. □

In conclusion let us present a nonsmooth version of the *mean value theorem* in Banach spaces in terms of the subdifferential constructions under consideration. To this end we define the *symmetric subdifferential*

$$\partial^0\varphi(\bar{x}) := \partial\varphi(\bar{x}) \cup \partial^+\varphi(\bar{x}) = \partial\varphi(\bar{x}) \cup [-\partial(-\varphi)(\bar{x})] \tag{5.6}$$

for any function $\varphi : X \to \bar{\mathbb{R}}$ at $\bar{x}$. One can see that (5.6) always possesses the classical symmetry property $\partial^0(-\varphi)(\bar{x}) = -\partial^0\varphi(\bar{x})$ and may be nonconvex in simple situations. First we observe the following nonsmooth analogue of the Fermat *stationary principle*.

Proposition 5.6. *Let $\varphi : X \to \bar{\mathbb{R}}$ and let $\bar{x} \in \operatorname{dom} \varphi$. Then $0 \in \partial\varphi(\bar{x})$ if φ has a local mimimum at $\bar{x}$ and $0 \in \partial^+\varphi(\bar{x})$ if φ has a local maximum at $\bar{x}$. So $0 \in \partial^0\varphi(\bar{x})$ if $\bar{x}$ is either minimum or maximum point of φ.*

Proof. It follows directly from (2.6) as $\varepsilon = 0$ that $0 \in \hat{\partial}\varphi(\bar{x})$ if $\bar{x}$ provides a local minimum to φ. Due to (2.5) one always has $\hat{\partial}\varphi(\bar{x}) \subset \partial\varphi(\bar{x})$. This implies the stationary principle for the case of local minima. The other statements in the proposition follow from here by virtue of the definitions. □

Theorem 5.7. *Let $a,\ b \in X$ and let $\varphi : X \to \bar{\mathbb{R}}$ be continuous in $[a,b] := \{a + t(b - a)|\ 0 \le t \le 1\}$. Then there is a number $\theta \in (0,1)$ such that*

$$\varphi(b) - \varphi(a) \in \partial^0_t \varphi(a + \theta(b - a)) \tag{5.7}$$

where the right-hand side is the symmetric subdifferential (5.6) *of the function $t \to \varphi(a + t(b - a))$ at $t = \theta$.*

Proof. Following the line in standard calculus, we consider a function $f : [0,1] \to \mathbb{R}$ defined by

$$f(t) := \varphi(a + t(b - a)) + t(\varphi(a) - \varphi(b)), \quad 0 \le t \le 1. \tag{5.8}$$

Obviously, f is continuous in $[0,1]$ and $f(0) = f(1) = \varphi(a)$. According to the classical Weierstrass theorem the function f attains both global minimum and maximum on $[0,1]$. Excluding the trivial case when f is constant in $[0,1]$, we conclude that there is an interior point $\theta \in (0,1)$ where f attains either its minimal or maximal value over the interval $[0,1]$. Using Proposition 5.6, one has $0 \in \partial^0 f(\theta)$. To obtain (5.7) we now apply Theorem 3.1(i) to the sum of functions in (5.8). This ends the proof of the theorem that is similar to standard calculus. □

Remark 5.8.

(a) Theorem 5.7 implies the classical mean value theorem in Banach spaces when φ is strictly differentiable at every point $x \in (a,b)$. One gets this applying the chain rule to the composition

$$\varphi(a + t(b - a)) = (\varphi \circ \Phi)(t),\ 0 \le t \le 1,$$

in (5.7) with $\Phi(t) := a + t(b - a)$; cf. Theorem 4.2. Note also that we cannot change $\partial^0\varphi$ for $\partial\varphi$ in (5.7) as follows from the example $\varphi(x) = -|x|$ on $[0,1]$.

(b) The arguments used above allow to establish analogues of the results obtained in this paper for the case of modified singular subdifferentials (3.10).

(c) In the case of Asplund spaces the subdifferential constructions of this paper satisfy more developed calculus for l.s.c. functions that is mainly based on *extremal principles* valid in such spaces; see [22].

Acknowledgment. The authors gratefully acknowledge helpful remarks made by the referees.

References

[1] J. M. Borwein, H. M. Strojwas: Proximal analysis and boundaries of closed sets in Banach spaces. I: Theory, Can. J. Math. 38 (1986), 431–452.

[2] F. H. Clarke: Optimizatiom and Nonsmooth Analysis, Wiley, 1983.

[3] F. H. Clarke: Methods of Dynamic and Nonsmooth Optimization, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 57, Soc. Indust. Appl. Math., Philadelphia, Pa., 1989.

[4] A. D. Ioffe: Approximate subdifferentials and applications. I: The finite dimensional theory, Trans. Amer. Math. Soc. 281 (1984), 389–416.

[5] A. D. Ioffe: Approximate subdifferential and applications. III: The metric theory, Mathematika 36 (1989), 1–38.

[6] A. D. Ioffe: Proximal analysis and approximate subdifferentials, J. London Math. Soc. 41 (1990), 175–192.

[7] A. Jourani, L. Thibault: A note on Fréchet and approximate subdifferentials of composite functions, Bull. Austral. Math. Soc. 49 (1994), 111–116.

[8] A. Jourani, L. Thibault: Extensions of subdifferential calculus rules in Banach spaces and applications, Can. J. Math. (to appear)

[9] A. Y. Kruger: Properties of generalized differentials, Siberian Math. J. 26 (1985), 822–832.

[10] A. Y. Kruger, B. S. Mordukhovich: Extremal points and Euler equations in nonsmooth optimization, Dokl. Akad. Nauk BSSR 24 (1980), 684–687.

[11] A. Y. Kruger, B. S. Mordukhovich: Generalized normals and derivatives, and necessary optimality conditions in nondifferentiable programming, Depon. VINITI # 408-80, 494-80, Moscow, 1980.

[12] E. B. Leach: A note on inverse function theorem, Proc. Amer. Math. Soc. 12 (1961), 694–697.

[13] P. D. Loewen: Limits of Fréchet normals in nonsmooth analysis, Optimization and Nonlinear Analysis (A. D. Ioffe, L. Marcus, and S. Reich, eds.), Pitman Research Notes Math. Ser. 244, 1992, pp. 178–188.

[14] P. D. Loewen: Optimal Control via Nonsmooth Analysis, CRM Proceedings and Lecture Notes, Vol. 2, Amer. Math. Soc., Providence, RI, 1993.

[15] P. D. Loewen, R. T. Rockafellar: Optimal control for unbounded differential inclusions, SIAM J. Control Optim. 32 (1994), 442–470.

[16] B. S. Mordukhovich: Maximum principle in problems of time optimal control with nonsmooth constraints, J. Appl. Math. Mech. 40 (1976), 960–969.

[17] B. S. Mordukhovich: Metric approximations and necessary optimality conditions for general classes of nonsmooth extremal problems, Soviet Math. Dokl. 22 (1980), 526–530.

[18] B. S. Mordukhovich: Approximation Methods in Problems of Optimization and Control, "Nauka", Moscow, 1988.

[19] B. S. Mordukhovich: Sensitivity analysis in nonsmooth optimization, Theoretical Aspects of Industrial Design (D. A. Field and V. Komkov, eds.), SIAM Proc. Appl. Math., vol. 58, Soc. Indust. Appl. Math., Philadelphia, Pa., 1992, pp. 32–46.

[20] B. S. Mordukhovich: Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis, Trans. Amer. Math. Soc. 343 (1994), 609–658.

[21] B. S. Mordukhovich: Discrete approximations and refined Euler-Lagrange conditions for nonconvex differential inclusions, SIAM J. Control Optim. 33 (1995), 882–915.

[22] B. S. Mordukhovich, Y. Shao: Nonsmooth sequential analysis in Asplund spaces, Trans. Amer. Math. Soc. (to appear).

[23] J. J. Moreau: Fonctionelles Convexes, Collége de France, Paris, 1966.

[24] R. R. Phelps: Convex Functions, Monotone Operators and Differentiability, 2nd edition, Lecture Notes in Mathematics, Vol. 1364, Springer, 1993.

[25] R. T. Rockafellar: Convex Analysis, Princeton Univ. Press, Princeton, N.J., 1970.

[26] R. T. Rockafellar: Directionally Lipschitzian functions and subdifferential calculus, Proc. London Math. Soc. 39 (1979), 331-355.

[27] R. T. Rockafellar: The Theory of Subgradients and its Applications to Problems of Optimization. Convex and Nonconvex Functions, Heldermann Verlag, Berlin, 1981.

[28] R. T. Rockafellar: Extensions of subgradient calculus with applications to optimization, Nonlinear Anal. 9 (1985), 665–698.

[29] R. T. Rockafellar, R. J-B Wets: Variational Analysis, Springer (to appear).

[30] L. Thibault: On subdifferentials of optimal value functions, SIAM J. Control Optim. 29 (1991), 1019–1036.

[31] J. S. Treiman: Clarke's gradients and epsilon-subgradients in Banach spaces, Trans. Amer. Math. Soc. 294 (1986), 65–78.

HIER :

Leere Seite
228