



**HAL**  
open science

## Enlargements of the Moreau-Rockafellar Subdifferential

Malek Abbasi, Alexander y Kruger, Michel Théra

► **To cite this version:**

Malek Abbasi, Alexander y Kruger, Michel Théra. Enlargements of the Moreau-Rockafellar Subdifferential. 2021. hal-02961001v2

**HAL Id: hal-02961001**

**<https://hal.science/hal-02961001v2>**

Preprint submitted on 24 Jan 2021 (v2), last revised 12 Mar 2021 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Enlargements of the Moreau–Rockafellar Subdifferential

Malek Abbasi · Alexander Y. Kruger · Michel Théra

*Dedicated to Terry Rockafellar, the ruler of Convex Analysis and Endolandia on the occasion of his 85th birthday.*

Received: date / Accepted: date

**Abstract** This paper proposes three enlargements of the conventional Moreau–Rockafellar subdifferential: the sup-, sup\*- and symmetric subdifferentials. They satisfy the most fundamental properties of the conventional subdifferential: convexity, weak\*-closedness and, if the function is bounded, weak\*-compactness. Moreover, they are nonempty for the Rainwater function, possess certain calculus rules, and provide necessary and sufficient optimality conditions for not necessarily convex functions.

**Keywords** Moreau–Rockafellar subdifferential · Convex function · Normal cone · Directional derivative · Optimality conditions.

**Mathematics Subject Classification (2000)** 49J52 · 49J53 · 90C30

## 1 Introduction

Throughout the paper,  $X$  is a real topological vector space. The topological (continuous) dual of  $X$  is denoted by  $X^*$ .

Various problems coming from different areas can be formulated as

$$\boxed{\text{Find } x \in \text{dom } T \text{ such that } 0 \in T(x),} \quad (\mathbf{OP})$$

where  $T : X \rightrightarrows X^*$  is a set-valued mapping, and the set  $\text{dom } T := \{x \in X : T(x) \neq \emptyset\}$  is the domain of  $T$ ; cf. [1]. We refer to this problem as *Original Problem (OP)*. When  $T := \partial f$  is the subdifferential mapping of some extended-real-valued convex function  $f : X \rightarrow \mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$ , then (OP) becomes the Fermat rule:

$$0 \in T(x) \iff x \text{ minimizes } f, \quad (1)$$

which is one of the central facts in optimization theory.

---

Malek Abbasi

Department of Mathematics, University of Isfahan, Isfahan, Iran

E-mail: malek.abbasi@sci.ui.ac.ir, malekabbasi72@gmail.com

Alexander Y. Kruger ORCID 0000-0002-7861-7380

Centre for Informatics and Applied Optimization, School of Engineering, IT and Physical Sciences, Federation University, Ballarat, Vic, Australia

E-mail: a.kruger@federation.edu.au

Michel Théra ORCID 0000-0001-9022-6406

Mathematics and Computer Science Department, University of Limoges

Federation University, Ballarat, Vic, Australia

E-mail: michel.thera@unilim.fr, m.thera@federation.edu.au

It is often convenient to study an auxiliary problem:

$$\boxed{\text{Find } x \in \text{dom } T' \text{ such that } 0 \in T'(x),}$$

where a set-valued mapping  $T' : X \rightrightarrows X^*$  is an *enlargement* of  $T$ , i.e.  $T(x) \subset T'(x)$  and  $T'(x)$  is convex for all  $x \in X$ , and  $T'$  is “close” to  $T$  in a certain sense. Solutions to the latter problem can serve as approximate solutions to **(OP)**. For a general theory of enlargements of monotone operators and their numerous applications we refer the readers to the book by Burachik and Iusem [1] and papers [?, ?].

The  $\varepsilon$ -subdifferential mapping  $\partial_\varepsilon f$  of a proper lower semicontinuous convex function  $f : X \rightarrow \mathbb{R}_\infty$ , defined for all  $x \in \text{dom } f$  and  $\varepsilon \geq 0$  by

$$\partial_\varepsilon f(x) := \{x^* \in X^* : \langle x^*, y - x \rangle \leq f(y) - f(x) + \varepsilon \quad \forall y \in X\},$$

represents a typical enlargement of the conventional subdifferential. It plays a key role in the theory of extremal problems and has been successfully used to construct numerical methods. Several other enlargements of the subdifferential mapping have been proposed in the framework of the *abstract convexity* theory; see, e.g., [2, 3]. For a great account of how much one can extend convex analysis to a wider framework in which some convexity features remain, we refer the readers to [4].

A convenient enlargement of an arbitrary monotone operator  $T : X \rightarrow X^*$ :

$$T^\varepsilon x := \{x^* \in X^* : \langle x^* - y^*, y - x \rangle \geq -\varepsilon \quad \forall (y, y^*) \in \text{gph } T\}$$

was proposed by Revalski and Théra [5]. It has convex and weak\*-closed values for all  $\varepsilon \geq 0$ , and  $Tx \subset T^\varepsilon x$  for all  $\varepsilon \geq 0$  and  $x \in X$ . Moreover,  $T^\varepsilon$  possesses other nice properties: local boundedness, demi-closed graph, Lipschitz continuity and the Brøndsted–Rockafellar property. The particular case  $T = \partial f$  was studied in [6]. Note that  $\partial_\varepsilon f(x) \subset (\partial f)^\varepsilon(x)$ . Martínez–Legaz and Théra [7] provided an example showing that the latter inclusion can be strict.

Finding an enlargement with the right properties is of major importance for applications. Motivated by (1), we consider the following *Enlargement Problem (EP)*:

$$\boxed{\text{Find an enlargement } T' \text{ of } T \text{ such that } 0 \in T'(x) \text{ for all } x \in \text{dom } T'.} \quad \text{(EP)}$$

Suppose that  $T'$  is a solution to **(EP)**. Then  $\text{dom } T \subset \text{dom } T'$ . For any  $\varepsilon \in [0, 1]$ , one can define another enlargement  $T_\varepsilon : X \rightrightarrows X^*$  by setting  $T_\varepsilon(x) := \varepsilon T'(x) + (1 - \varepsilon)T(x)$  for  $x \in \text{dom } T$  and  $T_\varepsilon(x) := T'(x)$  for  $x \in \text{dom } T' \setminus \text{dom } T$ . Obviously  $T(x) \subset T_\varepsilon(x) \subset T'(x)$ ,  $T_0(x) = T(x)$  for all  $x \in \text{dom } T$ , and  $T_1(x) = T'(x)$  for all  $x \in X$ . The mappings  $T_\varepsilon$  can be used to define perturbations of the problem **(OP)**.

Given a nonempty subset  $A \subset X$  with  $A \cap \text{dom } T' \neq \emptyset$  and a number  $\varepsilon \in (0, 1]$ , we can formulate the *Auxiliary Problem (AP)*:

$$\boxed{\text{Find } x \in A \text{ such that } 0 \in T_\varepsilon(x),} \quad \text{(AP)}$$

which may be more tractable and easier to handle.

Finding an appropriate enlargement  $T'$  (as a solution to **(EP)**) plays a key role in this procedure. In this paper, we consider the case  $T := \partial f$ , and aim at finding solutions (enlargements) of the problem **(EP)** such that  $T'$  satisfies some fundamental properties of  $\partial f$  including convexity, weak\*-closedness, weak\*-compactness and certain calculus rules.

It is well known that, even if a convex function  $f$  is continuous at  $\bar{x}$ , the subdifferential  $\partial f(\bar{x})$  can be empty. The next example is due to Rainwater [8]. (We refer the readers to [9, Example 4.2.10] for another example of this kind.) Recall that  $\ell^2(\mathbb{N})$  is the linear space of all real sequences  $x := (x_n)$  such that  $\|x\|^2 := \sum_{n=1}^{\infty} |x_n|^2 < \infty$ .

*Example 1.1* Define  $f : \ell^2(\mathbb{N}) \rightarrow \mathbb{R}$  by

$$f(x) := \begin{cases} -\sum_{n=1}^{\infty} (2^{-n} + x_n)^{\frac{1}{2}} & \text{if } x \in C, \\ +\infty & \text{otherwise,} \end{cases} \quad (2)$$

where  $C := \{x \in \ell^2(\mathbb{N}) : |x_n| \leq 2^{-n}, n = 1, 2, \dots\}$ . Observe that  $f$  is convex and continuous on  $C$ , and  $\partial f(x) = \emptyset$  for all  $x = (x_n) \in \ell^2(\mathbb{N})$  such that  $|x_n| < 2^{-n}$  for infinitely many  $n \in \mathbb{N}$ .

This paper proposes three enlargements of the conventional subdifferential: the *sup-subdifferential*, *sup\*-subdifferential* (section 2) and *symmetric subdifferential* (section 4), which provide additional tools for analyzing extended-real-valued functions when the conventional convex analysis is not applicable. Each of these enlargements is nonempty for the Rainwater function in Example 1.1 and satisfies the fundamental properties of the Moreau–Rockafellar subdifferential: convexity, weak\*-closedness and, if the function is bounded, weak\*-compactness. They also possess some calculus rules and provide some optimality conditions for convex and nonconvex nonsmooth functions.

The sup\*-subdifferential coincides with the Moreau–Rockafellar subdifferential at every point at which the minimum of the function is attained. The sup-subdifferential contains the sup\*-subdifferential and coincides with  $\partial f$  at  $x$  if and only if  $x$  minimizes  $f$ . For other points there are some connections with the conventional subdifferential if the involved function is upper semi-continuous. The symmetric subdifferential contains a nonzero element at each point  $\bar{x}$ , where  $f$  is directionally differentiable, and there exists a direction  $\bar{d}$  such that the maximum of  $f'(\bar{x}; \bar{d})$  and  $f'(\bar{x}; -\bar{d})$  is finite and positive. Note that the function (2) satisfies this condition.

Our basic notation is standard. Let  $f : X \rightarrow \mathbb{R}_{\infty}$  be an extended-real-valued function, and  $\bar{x} \in \text{dom } f := \{x \in X : f(x) < +\infty\}$ . The *directional derivative* of  $f$  at  $\bar{x}$  in direction  $d \in X$  is defined by the following limit:

$$f'(\bar{x}; d) := \lim_{t \rightarrow 0^+} \frac{f(\bar{x} + td) - f(\bar{x})}{t}.$$

We say that  $f$  is *directionally differentiable* at  $\bar{x}$  if the above limit exists in  $\mathbb{R} \cup \{\pm\infty\}$  for all  $d \in X$ . If  $f$  is convex, its subdifferential at  $\bar{x}$  is the set:

$$\partial f(\bar{x}) := \{x^* \in X^* : \langle x^*, x \rangle \leq f(\bar{x} + x) - f(\bar{x}) \quad \forall x \in X\}.$$

The paper is organized as follows. In section 2, we define the sup- and sup\*-subdifferentials, verify their basic properties and establish some necessary and sufficient optimality conditions. In section 3, we study sup- and sup\*-subdifferentials of upper semi-continuous functions and derive some representations in the convex and nonconvex cases. In section 4, we define the symmetric subdifferential and state its fundamental properties. In particular, we establish conditions under which the symmetric subdifferential is nonempty. In section 5, we prove some calculus rules for the sup-, sup\*- and symmetric subdifferentials.

## 2 Sup- and Sup\*-Subdifferentials

In this section,  $X$  is assumed to be a normed vector space, and  $\mathbb{B}_X$  and  $\mathbb{B}_{X^*}$  denote the closed unit balls in  $X$  and  $X^*$ , respectively.

A subset  $\mathcal{E} \subset \mathbb{B}_{X^*}$  is said to be *norm-generating* if, for any  $x \in X$ , there exists  $e^* \in \mathcal{E}$  such that  $|\langle e^*, x \rangle| = \|x\|$ . The collection of all weak\*-closed norm-generating subsets is denoted by  $\mathcal{F}$ . By the Hahn–Banach theorem,  $\mathbb{B}_{X^*} \in \mathcal{F}$ .

The following example demonstrates that the canonical basis of  $\mathbb{R}^n$  is a closed norm-generating subset.

*Example 2.1* Equip  $\mathbb{R}^n$  with the max norm  $\|x\|_{\max} := \max\{|x_k| : 1 \leq k \leq n\}$  for  $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ . Let  $\{e_1, e_2, \dots, e_n\}$  be the canonical basis of  $\mathbb{R}^n$ . For any  $x \in \mathbb{R}^n$ , we have  $\|x\|_{\max} = |x_k|$  for some  $1 \leq k \leq n$ , and  $|\langle e_k, x \rangle| = |x_k| = \|x\|_{\max}$ . Hence,  $\{e_1, e_2, \dots, e_n\}$  is a closed norm-generating subset of  $\mathbb{R}^n$ .

Moreover, any norm-generating subset  $\mathcal{E}$  of  $\mathbb{R}^n$  contains either  $e_k$  or  $-e_k$  for all  $k = 1, \dots, n$ . Indeed, by the definition of  $\mathcal{E}$ , for any  $k = 1, \dots, n$ , there exists some  $u = (u_1, \dots, u_n) \in \mathcal{E}$  such that  $|\langle u, e_k \rangle| = \|e_k\|_{\max} = 1$ . Hence,  $u_k = \pm 1$ . On the other hand, by the definition of the dual norm,  $\|u\|_{(\mathbb{R}^n)^*} = \sum_{i=1}^n |u_i| \leq 1$ . Hence,  $u_i = 0$  for all  $i \neq k$ , and therefore  $u$  equals either  $e_k$  or  $-e_k$ .

Using the same arguments, the above example can be easily extended to the case of an  $\ell^p(\mathbb{N})$  space ( $1 \leq p \leq +\infty$ ). Recall that  $\ell^p(\mathbb{N})$  is the linear space of all real sequences  $x := (x_k)$  such that  $\|x\|_p^p := \sum_{k=1}^{\infty} |x_k|^p < \infty$  if  $p < \infty$ , and  $\|x\|_{\infty} := \max_{k \in \mathbb{N}} |x_k| < \infty$ .  $\square$

*Example 2.2* The canonical basis  $\{e_1, e_2, \dots\}$  (i.e.  $e_k$  is a sequence whose only non-zero entry is a “1” in the  $k$ th coordinate) is a norm-generating subset of  $\ell^{\infty}(\mathbb{N})$ . Any norm-generating subset of  $\ell^p(\mathbb{N})$  ( $1 \leq p \leq +\infty$ ) contains either  $e_k$  or  $-e_k$  for all  $k = 1, 2, \dots$

Given  $x \in X$  and  $u^* \in X^*$ , denote

$$\tau_{u^*}(x) := \begin{cases} \left| \left\langle u^*, \frac{x}{\|x\|} \right\rangle \right| & \text{if } x \neq 0, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

Observe that  $0 \leq \tau_{u^*}(x) \leq \|u^*\|$ .

In what follows, we consider a function  $f : X \rightarrow \mathbb{R}_{\infty}$  and assume that  $\bar{x} \in \text{dom } f$ . The sets

$$\partial_{\text{sup}} f(\bar{x}) := \left\{ x^* \in X^* : \langle x^*, x \rangle \leq \sup_{0 \leq t \leq 1} f(\bar{x} + tx) - f(\bar{x}) \quad \forall x \in X \right\}, \quad (4)$$

$$\partial_{\text{sup}}^{\mathcal{E}} f(\bar{x}) := \left\{ x^* \in X^* : \langle x^*, x \rangle \leq \sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}) \quad \forall x \in X \right\}, \quad (5)$$

$$\partial_{\text{sup}}^* f(\bar{x}) := \bigcap_{\mathcal{E} \in \mathcal{F}} \partial_{\text{sup}}^{\mathcal{E}} f(\bar{x}) \quad (6)$$

are called, respectively, the *sup-subdifferential*, *sup $_{\mathcal{E}}$ -subdifferential* and *sup\*-subdifferential* of  $f$  at  $\bar{x}$ . The set (5) is determined by a given norm-generating set  $\mathcal{E} \in \mathcal{F}$ .

**Proposition 2.1**  $\partial_{\text{sup}} f(\bar{x}) = \partial_{\text{sup}}^{\mathbb{B}_{X^*}} f(\bar{x})$ .

*Proof* If  $u^* \in \mathbb{B}_{X^*}$ , then  $0 \leq \tau_{u^*}(x) \leq 1$ . Hence,  $\partial_{\text{sup}}^{\mathbb{B}_{X^*}} f(\bar{x}) \subset \partial_{\text{sup}} f(\bar{x})$ . Let  $x \in X$ . Then there exists  $u^* \in \mathbb{B}_{X^*}$  such that  $\langle u^*, x \rangle = \|x\|$ . Hence, for any  $t \in [0, 1]$ , we have  $tu^* \in \mathbb{B}_{X^*}$ ,  $\langle tu^*, x \rangle = t\|x\|$  and  $\tau_{tu^*}(x) = t$ , and consequently,  $\partial_{\text{sup}} f(\bar{x}) \subset \partial_{\text{sup}}^{\mathbb{B}_{X^*}} f(\bar{x})$ .  $\square$

The next example is an extension of Example 1.1 to the case of an  $\ell^p(\mathbb{N})$  space ( $1 \leq p < \infty$ ).

*Example 2.3* Define  $f : \ell^p(\mathbb{N}) \rightarrow \mathbb{R}$  by

$$f(x) := \begin{cases} -\sum_{n=1}^{\infty} (2^{-\frac{2n}{p}} + x_n)^{\frac{1}{2}} & \text{if } x \in C, \\ +\infty & \text{otherwise,} \end{cases}$$

where  $C := \{x \in \ell^p(\mathbb{N}) : |x_n| \leq 2^{-\frac{2n}{p}}, n = 1, 2, \dots\}$ . Let  $\bar{x} \in C$ . We show that  $0 \in \partial_{\text{sup}}^* f(\bar{x})$ . The set  $C$  is convex. Each summand in the first part of the definition of  $f$  is continuous and convex, and its absolute value is bounded from above by  $2^{-\frac{n}{p} + \frac{1}{2}}$ .

Hence, the series is uniformly convergent; this shows that  $f$  is continuous on  $C$  and convex. Let  $(e_k)$  denote the canonical basis of  $\ell^p(\mathbb{N})$ . Let  $\mathcal{E} \in \mathcal{F}$ . In view of Example 2.2,  $e_k \in \mathcal{E}$  for all  $k \in \mathbb{N}$ . Now let  $x = (x_n) \in C$  and  $x \neq 0$ . For all  $k \in \mathbb{N}$ , we have

$$\left\langle \pm e_k, \frac{x}{\|x\|_p} \right\rangle = \frac{\pm x_k}{\|x\|_p}.$$

Hence,

$$f\left(\bar{x} + \frac{|x_k|}{\|x\|_p} x\right) = f\left(\bar{x} + \left\langle \pm e_k, \frac{x}{\|x\|_p} \right\rangle x\right) \leq \sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x).$$

Since  $f$  is continuous, by letting  $k \rightarrow \infty$ , we obtain  $\sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}) \geq 0$ . This implies that  $0 \in \partial_{sup}^{\mathcal{E}} f(\bar{x})$ , and consequently,  $0 \in \partial_{sup}^* f(\bar{x})$ .  $\square$

The next two propositions collect some basic properties of the subdifferentials defined above.

**Proposition 2.2** *The following assertions hold true.*

- (i)  $\partial_{sup}^{\mathcal{E}} f(\bar{x})$  is convex and weak\*-closed for all  $\mathcal{E} \in \mathcal{F}$ . As a consequence,  $\partial_{sup} f(\bar{x})$  and  $\partial_{sup}^* f(\bar{x})$  are convex and weak\*-closed.
- (ii)  $0 \in \partial_{sup} f(\bar{x})$ . If  $\bar{x}$  is a local maximizer of  $f$ , then  $\partial_{sup} f(\bar{x}) = \{0\}$ .
- (iii) Suppose that the function  $x \mapsto f(\bar{x} + x)$  is bounded on  $\mathbb{B}_X$ . Then  $\partial_{sup}^{\mathcal{E}} f(\bar{x})$  is weak\*-compact for all  $\mathcal{E} \in \mathcal{F}$ . As a consequence,  $\partial_{sup} f(\bar{x})$  and  $\partial_{sup}^* f(\bar{x})$  are weak\*-compact.
- (iv) Suppose that  $X$  is finite dimensional and  $f$  is continuous. Then  $\partial_{sup}^{\mathcal{E}} f(\bar{x})$  is compact for all  $\mathcal{E} \in \mathcal{F}$ . As a consequence,  $\partial_{sup} f(\bar{x})$  and  $\partial_{sup}^* f(\bar{x})$  are compact.

*Proof* (i) Let  $\mathcal{E} \in \mathcal{F}$ . For any  $x_1^*, x_2^* \in \partial_{sup}^{\mathcal{E}} f(\bar{x})$ ,  $\alpha, \beta \in [0, 1]$  with  $\alpha + \beta = 1$ , and  $x \in X$ , we have

$$\langle \alpha x_1^* + \beta x_2^*, x \rangle \leq \sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}), \quad \langle x_2^*, x \rangle \leq \sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}),$$

and consequently,  $\langle \alpha x_1^* + \beta x_2^*, x \rangle \leq \sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x})$ . Hence,  $\partial_{sup}^{\mathcal{E}} f(\bar{x})$  is convex. Let  $(x_\gamma^*)_{\gamma \in \Gamma}$  be a net in  $\partial_{sup}^{\mathcal{E}} f(\bar{x})$  converging to some  $x^* \in X^*$  in the weak\* topology of  $X^*$ . Let  $x \in X$ . For all  $\gamma \in \Gamma$ , we have  $\langle x_\gamma^*, x \rangle \leq \sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x})$ , and consequently,  $\langle x^*, x \rangle \leq \sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x})$ . Hence,  $\partial_{sup}^{\mathcal{E}} f(\bar{x})$  is weak\*-closed. In view of Proposition 2.1, and thanks to the fact that the intersection of convex and weak\*-closed sets is convex and weak\*-closed, the other two subdifferentials are convex and weak\*-closed too.

(ii) We have  $\sup_{0 \leq t \leq 1} f(\bar{x} + tx) \geq f(\bar{x})$  for all  $x \in X$ . It follows from definition (4) that  $0 \in \partial_{sup} f(\bar{x})$ . Let  $\bar{x}$  be a local maximizer of  $f$ . Then there is a number  $\delta > 0$  such that  $f(\bar{x} + x) \leq f(\bar{x})$  for all  $x \in \delta \mathbb{B}_X$ . If  $x^* \in \partial_{sup} f(\bar{x})$ , then by definition (4), for all  $x \in \delta \mathbb{B}_X$ , we have  $\langle x^*, x \rangle \leq \sup_{0 \leq t \leq 1} f(\bar{x} + tx) - f(\bar{x}) = 0$ , and consequently,  $x^* = 0$ . Hence,  $\partial_{sup} f(\bar{x}) = \{0\}$ .

(iii) Suppose that  $|f(\bar{x} + x)| \leq M < +\infty$  for all  $x \in \mathbb{B}_X$ . Let  $\mathcal{E} \in \mathcal{F}$  and  $x^* \in \partial_{sup}^{\mathcal{E}} f(\bar{x})$ . Then

$$\|x^*\| = \sup_{x \in \mathbb{B}_X} \langle x^*, x \rangle \leq \sup_{x \in \mathbb{B}_X, u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}) \leq M - f(\bar{x}).$$

Thus,  $\partial_{sup}^{\mathcal{E}} f(\bar{x})$  is bounded and therefore weak\*-compact by the Banach–Alaoglu–Bourbaki theorem. The second assertion follows since the intersection of weak\*-compact sets is weak\*-compact.

(iv) Recall that the closed unit ball in a finite dimensional space is compact, and therefore the continuity of  $f$  implies that the function  $x \mapsto f(\bar{x} + x)$  is bounded on  $\mathbb{B}_X$ . The assertion follows from (vi).  $\square$

**Proposition 2.3** *Let  $f$  be convex. The following assertions hold true.*

- (i)  $\partial f(\bar{x}) \subset \partial_{sup}^{\mathcal{E}} f(\bar{x}) \subset \partial_{sup} f(\bar{x})$  for all  $\mathcal{E} \in \mathcal{F}$ . As a consequence,  $\partial f(\bar{x}) \subset \partial_{sup}^* f(\bar{x}) \subset \partial_{sup} f(\bar{x})$ .
- (ii) If  $X = \mathbb{R}$ , then  $\partial_{sup}^* f(\bar{x}) = \partial f(\bar{x})$ .
- (iii)  $\bar{x}$  minimizes  $f$  if and only if  $\partial f(\bar{x}) = \partial_{sup}^* f(\bar{x}) = \partial_{sup} f(\bar{x})$ .
- (iv) If  $\partial_{sup} f(\bar{x})$  is a singleton, then either  $\bar{x}$  minimizes  $f$  or  $\partial f(\bar{x}) = \emptyset$ .

*Proof* (i) Let  $\mathcal{E} \in \mathcal{F}$ ,  $x^* \in \partial f(\bar{x})$  and  $x \in X$ . Then  $\langle x^*, x \rangle \leq f(\bar{x} + x) - f(\bar{x})$ , and there exists  $\hat{u}^* \in \mathcal{E}$  such that  $|\langle \hat{u}^*, x \rangle| = \|x\|$ , i.e.  $\tau_{\hat{u}^*}(x) = 1$ . Therefore

$$f(\bar{x} + x) = f(\bar{x} + \tau_{\hat{u}^*}(x)x) \leq \sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x),$$

and consequently,  $\langle x^*, x \rangle \leq \sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x})$ . It follows that  $x^* \in \partial_{sup}^{\mathcal{E}} f(\bar{x})$ , and consequently,  $\partial f(\bar{x}) \subset \partial_{sup}^{\mathcal{E}} f(\bar{x})$ . The opposite inclusion  $\partial_{sup}^{\mathcal{E}} f(\bar{x}) \subset \partial_{sup} f(\bar{x})$  is straightforward from definitions (4) and (5). The second claim is a consequence of the first one.

(ii)  $\partial f(\bar{x}) \subset \partial_{sup}^* f(\bar{x})$  by (i). The set  $\mathcal{E} := \{1\}$  is norm-generating in  $\mathbb{R}$ .  $\tau_1(x) = 1$  if  $x \neq 0$ . Hence,  $f(\bar{x} + \tau_1(x)x) = f(\bar{x} + x)$  for all  $x \in \mathbb{R}$ , and consequently,  $\partial_{sup}^* f(\bar{x}) \subset \partial_{sup}^{\mathcal{E}} f(\bar{x}) \subset \partial f(\bar{x})$ .

(iii) If  $\partial f(\bar{x}) = \partial_{sup} f(\bar{x})$ , then by Proposition 2.2(ii),  $0 \in \partial f(\bar{x})$ , and consequently,  $\bar{x}$  minimizes  $f$ . Conversely, suppose that  $\bar{x} \in X$  is a minimizer of  $f$ . Let  $x^* \in \partial_{sup} f(\bar{x})$  and  $x \in X$ . Then

$$\langle x^*, x \rangle \leq \sup_{0 \leq t \leq 1} f(\bar{x} + tx) - f(\bar{x}) \leq \sup_{0 \leq t \leq 1} t(f(\bar{x} + x) - f(\bar{x})) = f(\bar{x} + x) - f(\bar{x}),$$

It follows that  $x^* \in \partial(\bar{x})$ , and consequently,  $\partial_{sup} f(\bar{x}) \subset \partial f(\bar{x})$ . In view of (i), we have  $\partial f(\bar{x}) = \partial_{sup}^* f(\bar{x}) = \partial_{sup} f(\bar{x})$ .

(iv) Let  $\partial_{sup} f(\bar{x})$  be a singleton. By Proposition 2.2(ii),  $\partial_{sup} f(\bar{x}) = \{0\}$ . Hence, by (i), either  $\partial f(\bar{x}) = \emptyset$  or  $\partial f(\bar{x}) = \{0\}$ . In the latter case,  $\bar{x}$  minimizes  $f$ .  $\square$

As a by-product of Proposition 2.3(ii), we see that the  $\text{sup}^*$ -subdifferential can be empty at some points. Thanks to Proposition 2.2(ii), the  $\text{sup}$ -subdifferential is always nonempty.

Proposition 2.2(ii) also gives a simple necessary condition for local optimality. Now we discuss some more local conditions. For every  $\delta \in (0, +\infty]$ , we define a localization of  $\partial_{sup} f(\bar{x})$ :

$$\partial_{sup}^{(\delta)} f(\bar{x}) := \{x^* \in X^* : \langle x^*, x \rangle \leq \sup_{0 \leq t \leq 1} f(\bar{x} + tx) - f(\bar{x}), \|x\| < \delta\}. \quad (7)$$

Observe that  $0 \in \partial_{sup}^{(\delta)} f(\bar{x})$  for all  $\delta \in (0, +\infty]$ , and

$$\partial_{sup} f(\bar{x}) = \partial_{sup}^{(+\infty)} f(\bar{x}) = \bigcap_{\delta > 0} \partial_{sup}^{(\delta)} f(\bar{x}).$$

Analyzing the proof of Proposition 2.2(ii), we can see that, using (7), we can formulate a slightly sharper necessary condition.

**Proposition 2.4** *If  $\bar{x}$  is a local maximizer of  $f$ , then  $\partial_{sup}^{(\delta)} f(\bar{x}) = \{0\}$  for all  $\delta \in (0, +\infty]$ .*

*Example 2.4* (See Fig. 2.) Define  $f : \mathbb{R} \rightarrow \mathbb{R}$  as

$$f(x) := \begin{cases} x^2 - 1 & \text{if } |x| \geq 1, \\ 2x^2 - 2 & \text{if } 0 < x < 1, \\ -x - 1 & \text{if } -1 < x \leq 0. \end{cases}$$

Note that  $f$  is not convex and fails to be continuous at 0 (although it is upper semi-continuous). For all  $x \in \mathbb{R}$ , one has

$$\sup_{0 \leq t \leq 1} (-f)(tx) = - \inf_{0 \leq t \leq 1} f(tx) = \begin{cases} 2 & \text{if } x \geq 0, \\ 1 & \text{if } x < 0, \end{cases}$$

and consequently, for each  $\delta > 0$ ,  $\partial_{sup}^{(\delta)}(-f)(0)$  consists of all numbers  $a$  such that  $ax \leq 1$  for all  $x \in (0, \delta)$  and  $ax \leq 0$  for all  $x \in (-\delta, 0)$ . Hence,  $\partial_{sup}^{(\delta)}(-f)(0) = [0, \frac{1}{\delta}]$  and, by Proposition 2.4, 0 fails to be a local minimizer of  $f$ . Observe that  $\partial_{sup}(-f)(0) = \bigcap_{\delta > 0} \partial_{sup}^{(\delta)}(-f)(0) = \{0\}$ , and Proposition 2.2(ii) does not allow us to make this conclusion. One can also check that  $\partial_{sup}(-f)(x) \neq \{0\}$  for all  $x \neq 0$ . Thus,  $f$  has no local minimizers. Similarly,  $\partial_{sup}f(x) \neq \{0\}$  for all  $x \in X$ , and consequently,  $f$  has no local maximizers.  $\square$

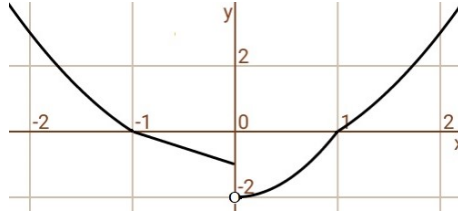


Fig. 1 The graph of  $f$  (Example 2.4)

We now show that under some additional assumptions, the necessary optimality condition in Proposition 2.4 becomes sufficient. We first consider the case  $X := \mathbb{R}$ . Given a function  $f : \mathbb{R} \rightarrow \mathbb{R}_\infty$  and a point  $\bar{x} \in \text{dom } f$ , consider the one-sided limits:

$$f'_+(\bar{x}) := \liminf_{x \downarrow 0} \frac{f(\bar{x} + x) - f(\bar{x})}{x}, \quad f'_-(\bar{x}) := \limsup_{x \uparrow 0} \frac{f(\bar{x} + x) - f(\bar{x})}{x}. \quad (8)$$

Observe that the above limits can be infinite, and

$$f'_-(\bar{x}) = - \liminf_{x \downarrow 0} \frac{f(\bar{x} - x) - f(\bar{x})}{x}.$$

Given an  $\alpha \in [-\infty, +\infty]$ , we use the standard notations  $\alpha_+ := \max\{\alpha, 0\}$  and  $\alpha_- := \min\{\alpha, 0\}$ .

**Proposition 2.5** *Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  and  $\bar{x} \in \text{dom } f$ .*

(i) *For any  $\varepsilon > 0$ , there exists a number  $\delta > 0$  such that*

$$((f'_-(\bar{x}) + \varepsilon)_-, (f'_+(\bar{x}) - \varepsilon)_+) \subset \partial_{sup}^{(\delta)}f(\bar{x}).$$

(ii) *Suppose that  $\partial_{sup}^{(\delta)}f(\bar{x}) = \{0\}$  for all  $\delta > 0$ . Then  $f'_-(\bar{x}) \geq 0$  and  $f'_+(\bar{x}) \leq 0$ . As a consequence,  $f$  is upper semi-continuous at  $\bar{x}$ . If both  $f'_+(\bar{x})$  and  $f'_-(\bar{x})$  are nonzero, then  $\bar{x}$  is a local maximizer of  $f$ .*

*Proof* (i) Let  $\varepsilon > 0$  and  $a \in ((f'_-(\bar{x}) + \varepsilon)_-, (f'_+(\bar{x}) - \varepsilon)_+)$ . If  $f'_+(\bar{x}) < +\infty$ , set  $M_1 := f'_+(\bar{x})$ ; otherwise set  $M_1 := a + \varepsilon$ . If  $f'_-(\bar{x}) > -\infty$ , set  $M_2 := f'_-(\bar{x})$ ; otherwise set  $M_2 := a - \varepsilon$ . Choose a number  $\delta > 0$  such that

$$\begin{aligned} \frac{f(\bar{x} + x) - f(\bar{x})}{x} &> M_1 - \varepsilon \quad \text{for all } x \in (0, \delta), \\ \frac{f(\bar{x} + x) - f(\bar{x})}{x} &< M_2 + \varepsilon \quad \text{for all } x \in (-\delta, 0). \end{aligned}$$



Then, for any  $x \in (-\delta, \delta)$ ,

$$\begin{aligned} \sup_{0 \leq t \leq 1} f(\bar{x} + tx) - f(\bar{x}) &\geq \max_{t \in \{0,1\}} f(\bar{x} + tx) - f(\bar{x}) = (f(\bar{x} + x) - f(\bar{x}))_+ \\ &\geq \begin{cases} (M_1 - \varepsilon)_+ x \geq ax & \text{if } x \geq 0, \\ (M_2 + \varepsilon)_- x \geq ax & \text{if } x \leq 0. \end{cases} \end{aligned}$$

Hence,  $a \in \partial_{sup}^{(\delta)} f(\bar{x})$ .

(ii) By (i),  $f'_-(\bar{x}) + \varepsilon \geq 0$  and  $f'_+(\bar{x}) - \varepsilon \leq 0$  for all  $\varepsilon > 0$ , and consequently,  $f'_-(\bar{x}) \geq 0$  and  $f'_+(\bar{x}) \leq 0$ . In view of the definitions (8), these inequalities imply that  $f$  is upper semi-continuous at  $\bar{x}$ . If  $f'_+(\bar{x})$  and  $f'_-(\bar{x})$  are nonzero, then  $f'_-(\bar{x}) > 0$  and  $f'_+(\bar{x}) < 0$ ; hence,  $\bar{x}$  is a local maximizer of  $f$ .  $\square$

*Remark 2.1* Condition  $\partial_{sup}^{(\delta)} f(\bar{x}) = \{0\}$  in Proposition 2.5(ii) is essential. For example, for the function  $x \mapsto f(x) := x$ , we have  $f'_-(0) = f'_+(0) = 1$ , but 0 is obviously not a maximizer of  $f$ . Notice that  $\partial_{sup}^{(\delta)} f(0) = [0, 1]$  for all  $\delta > 0$ .

The following example demonstrates that condition  $f'_+(\bar{x}) \neq 0$  in Proposition 2.5(ii) is not necessary.

*Example 2.5* (See Fig. 2.) Define  $f : \mathbb{R} \rightarrow \mathbb{R}$  as

$$f(x) := \begin{cases} x^2 - 1 & \text{if } |x| \geq 1, \\ 2 - 2x^2 & \text{if } 0 \leq x < 1, \\ x + 1 & \text{if } -1 < x < 0. \end{cases}$$

The point  $x = 0$  is obviously a local maximizer. Moreover, it is easy to check (as in Example 2.4) that  $\partial_{sup} f(0) = \{0\}$ . At the same time,  $f'_+(0) = 0$ , i.e. the sufficient conditions in Proposition 2.5(ii) are not satisfied. One can also check that  $\partial_{sup} f(x) \neq \{0\}$  for all  $x \neq 0$ . Thus,  $f$  has no other local maximizers.

Observe that  $\partial_{sup}^{(\delta)}(-f)(\pm 1) = \{0\}$  for all  $\delta > 0$ , while  $f'_+(\pm 1) > 0$  and  $f'_-(\pm 1) < 0$ . It follows from Proposition 2.5(ii) applied to the function  $-f$  that  $x = -1$  and  $x = 1$  are local minimizers of  $f$ .  $\square$

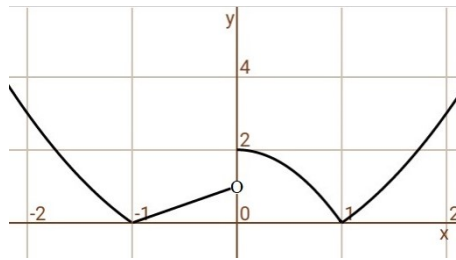


Fig. 2 The graph of  $f$  (Example 2.5)

*Remark 2.2* Flores Bazán [10] used a derivative notion defined geometrically via the epigraph:

$$\text{epi} D_e^R f(\bar{x}; \cdot) := \overline{\bigcup_{t \geq 0} t(\text{epi} f - (\bar{x}, f(\bar{x})))}.$$

It is easy to see that  $\bar{x}$  is a (global) minimizer of  $f$  if and only if  $D_e^R f(\bar{x}; v) \geq 0$  for all  $v \in X$ . This criterion obviously allows one to recover some of the conclusions in Examples 2.4 and 2.5.

Next, we proceed to extending Proposition 2.5(ii) to the setting of a general normed vector space. We are going to use a weaker ‘one-dimensional’ notion of optimality.

**Definition 2.1** A point  $\bar{x}$  is an algebraic local maximizer (minimizer) of  $f$  if, for every  $x \in X$ , there exists a number  $\varepsilon > 0$  such that  $f(\bar{x} + tx) \leq f(\bar{x})$  ( $f(\bar{x}) \leq f(\bar{x} + tx)$ ) for all  $t \in (0, \varepsilon)$ .

*Remark 2.3* In the above definition, it suffices to consider only unit vectors  $x \in X$ . Condition  $t \in (0, \varepsilon)$  can be replaced by a two-sided constraint  $t \in (-\varepsilon, \varepsilon)$ . Thus,  $\bar{x}$  is an algebraic local maximizer (minimizer) of  $f$  if and only if, for every  $x \in X$ , the number 0 is an algebraic local maximizer (minimizer) of the function  $t \mapsto f(\bar{x} + tx)$  on  $\mathbb{R}$ .

A local maximizer (minimizer) of  $f$  is an algebraic local maximizer (minimizer) of  $f$ . The opposite assertion is not true in general, unless  $X = \mathbb{R}$ .

The lower directional derivative of  $f$  at  $\bar{x}$  in the direction  $x \in X$  is defined as follows:

$$f'_+(\bar{x}; x) := \liminf_{t \downarrow 0} \frac{f(\bar{x} + tx) - f(\bar{x})}{t}.$$

For each  $x \in X$ , the function  $f_x : \mathbb{R} \rightarrow \mathbb{R}_\infty$  is defined by  $f_x(t) := f(\bar{x} + tx)$  for all  $t \in \mathbb{R}$ . Obviously,  $(f_x)'_+(0) = f'_+(\bar{x}; x)$ . The next proposition is a consequence of Proposition 2.5(ii).

**Proposition 2.6** If  $\partial_{sup}^{(\delta)} f_x(0) = \{0\}$  for some  $x \in X$  and all  $\delta > 0$ , then  $f'_+(\bar{x}; x) \leq 0$ . As a consequence, if for all  $x \neq 0$  and all  $\delta > 0$ , we have  $\partial_{sup}^{(\delta)} f_x(0) = \{0\}$  and  $f'_+(\bar{x}; x) \neq 0$ , then  $\bar{x}$  is an algebraic local maximizer of  $f$ .

The next example demonstrates that condition  $\partial_{sup}^{(\delta)} f_x(0) = \{0\}$  in the second part of Proposition 2.6 is essential.

*Example 2.6* Consider the continuous function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  given by

$$f(x, y) := \begin{cases} \sqrt{x^2 + y^2} \sin\left(\frac{1}{\sqrt{x^2 + y^2}}\right) & \text{if } x^2 + y^2 \neq 0, \\ 0 & \text{if } x = y = 0. \end{cases}$$

Obviously  $(0, 0)$  fails to be an algebraic local maximizer of  $f$ . We are going to show that  $\partial_{sup}^{(\delta)} f_{(1,0)}(0)$  contains nonzero elements, where  $\delta := \frac{2}{\pi}$ ; more precisely,  $[-\frac{1}{5}, \frac{1}{5}] \subset \partial_{sup}^{(\delta)} f_{(1,0)}(0)$ . One can easily check that

$$f_{(1,0)}(x) := \begin{cases} |x| \sin\left(\frac{1}{|x|}\right) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Now let  $a \in [-\frac{1}{5}, \frac{1}{5}]$ . For each  $x \in (-\delta, 0) \cup (0, \delta)$ , set  $k_x := \left\lfloor (2\pi)^{-1} \left(\frac{1}{|x|} - \frac{\pi}{2}\right) \right\rfloor + 1$  and  $t_x := \left[ (2k_x\pi + \frac{\pi}{2}) |x| \right]^{-1}$ . Observe that  $k_x \in \mathbb{N}$  and  $0 < t_x < 1$ , and consequently,

$$\sup_{0 \leq t \leq 1} f_{(1,0)}(tx) \geq f_{(1,0)}(t_x x) = t_x |x| = \frac{1}{2k_x\pi + \frac{\pi}{2}} \geq \frac{|x|}{1 + 2\pi|x|} > \frac{|x|}{5} \geq ax.$$

Hence,  $a \in \partial_{sup}^{(\delta)} f_{(1,0)}(0)$ . Notice that for every  $(x, y) \neq (0, 0)$  one has

$$f'_+((0, 0); (x, y)) = \liminf_{t \downarrow 0} \frac{|t| \sin\left(\frac{1}{|t| \|(x, y)\|_2}\right)}{t} \|(x, y)\|_2 = -\|(x, y)\|_2 \neq 0,$$

where  $\|(x, y)\|_2 = \sqrt{x^2 + y^2}$ . □

### 3 Sup- and Sup\*-Subdifferentials of Upper Semi-Continuous Functions

In this section, we assume that  $X$  is a normed vector space, a function  $f : X \rightarrow \mathbb{R}_\infty$  is upper semi-continuous, and  $\bar{x} \in \text{dom } f$ . We derive some representations for the sup- and sup\*-subdifferentials in the convex and nonconvex cases. Recall that  $f$  is upper semi-continuous at  $\bar{x}$  if  $\limsup_{x \rightarrow \bar{x}} f(x) \leq f(\bar{x})$ . A function is said to be upper semi-continuous if it is upper semi-continuous at every  $x \in X$ .

Given a subset  $\mathcal{E} \in \mathcal{F}$ , a function  $\tau_{u^*} : X \rightarrow \mathbb{R}$  defined for  $u^* \in \mathcal{E}$  by (3), and a point  $x \in X$ , set

$$\tau_{\mathcal{E}}(x) := \min \left\{ \tau_{x^*}(x) : f(\bar{x} + \tau_{x^*}(x)x) = \sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x), x^* \in \mathcal{E} \right\}, \quad (9)$$

$$\tau(x) := \min \left\{ \lambda \in [0, 1] : f(\bar{x} + \lambda x) = \sup_{0 \leq t \leq 1} f(\bar{x} + tx) \right\}. \quad (10)$$

**Lemma 3.1** *Let  $\mathcal{E} \in \mathcal{F}$ . The functions  $\tau_{\mathcal{E}}, \tau : X \rightarrow [0, 1]$ , given by (9) and (10) are well-defined, and  $\tau \leq \tau_{\mathcal{E}}$ . If, furthermore,  $f$  is convex, then, for all  $x \in X$ ,*

$$\begin{aligned} \sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}) &\leq \tau_{\mathcal{E}}(x)(f(\bar{x} + x) - f(\bar{x})), \\ \sup_{0 \leq t \leq 1} f(\bar{x} + tx) - f(\bar{x}) &\leq \tau(x)(f(\bar{x} + x) - f(\bar{x})). \end{aligned}$$

*Proof* Let  $x \in X$ . Define a function  $Q_x : X^* \rightarrow \mathbb{R}_\infty$ :

$$Q_x(x^*) := f(\bar{x} + \tau_{x^*}(x)x), \quad x^* \in X^*.$$

We claim that  $Q_x$  is weak\*-upper semi-continuous. Indeed, suppose that  $(x_v^*)$  is a net in  $X^*$  which converges to some  $x^*$  in the weak\* topology. It follows that  $\tau_{x_v^*}(x) \rightarrow \tau_{x^*}(x)$  and, since  $f$  is upper semi-continuous,  $\limsup_{x_v^* \rightarrow x^*} Q_x(x_v^*) \leq Q_x(x^*)$ , i.e.  $Q_x$  is weak\*-upper semi-continuous. By the Banach–Alaoglu theorem,  $\mathcal{E}$  is weak\*-compact, and therefore, there exists  $x^* \in \mathcal{E}$  such that  $Q_x(x^*) = \sup_{u^* \in \mathcal{E}} Q_x(u^*)$ , i.e.

$$f(\bar{x} + \tau_{x^*}(x)x) = \sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x).$$

The point  $x^* \in \mathcal{E}$  defined above and the corresponding number  $\tau_{x^*}(x)$  are in general not unique. Nevertheless, one can easily check that the set of all such numbers is compact in  $[0, 1]$ , and consequently, the function (9) is well-defined. Let  $f$  be convex. Since  $\tau_{\mathcal{E}}(x) \in [0, 1]$ , we have

$$\sup_{u^* \in \mathcal{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}) = f(\bar{x} + \tau_{\mathcal{E}}(x)x) - f(\bar{x}) \leq \tau_{\mathcal{E}}(x)(f(\bar{x} + x) - f(\bar{x})).$$

In the particular case  $\mathcal{E} := \mathbb{B}_{X^*}$ , it follows that the function (10) is well-defined, and the second inequality holds in the convex case. The inequality  $\tau \leq \tau_{\mathcal{E}}$  is a direct consequence of the definitions (9) and (10).  $\square$

Recall that, for a nonempty subset  $A \subset X$ , the negative polar cone to  $A$  is defined as  $A^\circ := \{x^* \in X^* : \langle x^*, x \rangle \leq 0 \quad \forall x \in A\}$ . If  $A = \emptyset$ , we set  $A^\circ := X^*$ . The normal cone to a convex subset  $A \subset X$  at  $\bar{x} \in A$  is defined as

$$N_A(\bar{x}) := \{x^* \in X^* : \langle x^*, x - \bar{x} \rangle \leq 0 \quad \forall x \in A\}.$$

Thus, if  $A$  is convex and  $0 \in A$ , then  $A^\circ$  is just the normal cone to  $A$  at zero.

Given a subset  $C \subset X$ , we denote

$$\partial_C f(\bar{x}) := \{x^* \in X^* : \langle x^*, x \rangle \leq f(\bar{x} + x) - f(\bar{x}) \quad \forall x \in C\},$$

with the convention  $\partial_C f(\bar{x}) := X^*$  if  $C = \emptyset$ .

**Theorem 3.1** *Let the functions  $\tau$  and  $\tau_{\mathcal{E}}$  be defined by (10) and (9), respectively. Then*

$$\partial_{\text{sup}} f(\bar{x}) = \bigcap_{0 < \lambda \leq 1} \{ \lambda \partial_{C_{\lambda}} f(\bar{x}) \} \cap (\tau^{-1}(0))^{\circ}, \quad (11)$$

$$\partial_{\text{sup}}^{\mathcal{E}} f(\bar{x}) = \bigcap_{0 < \lambda \leq 1} \{ \lambda \partial_{C_{\lambda}^{\mathcal{E}}} f(\bar{x}) \} \cap (\tau_{\mathcal{E}}^{-1}(0))^{\circ} \quad \text{for all } \mathcal{E} \in \mathcal{F}, \quad (12)$$

$$\partial_{\text{sup}}^* f(\bar{x}) = \bigcap_{\mathcal{E} \in \mathcal{F}, 0 < \lambda \leq 1} \{ \lambda \partial_{C_{\lambda}^{\mathcal{E}}} f(\bar{x}) \} \cap (\tau_{\mathcal{E}}^{-1}(0))^{\circ},$$

where  $C_{\lambda} := \lambda (\tau^{-1}(\lambda) \setminus \tau^{-1}(0))$  and  $C_{\lambda}^{\mathcal{E}} := \lambda (\tau_{\mathcal{E}}^{-1}(\lambda) \setminus \tau_{\mathcal{E}}^{-1}(0))$ .

*Proof* Let  $\mathcal{E} \in \mathcal{F}$ . By definition (5) and Lemma 3.1, we have

$$\partial_{\text{sup}}^{\mathcal{E}} f(\bar{x}) = \{x^* \in X^* : \langle x^*, x \rangle \leq f(\bar{x} + \tau_{\mathcal{E}}(x)x) - f(\bar{x}) \quad \forall x \in X\}.$$

One can easily check that  $\partial_{\text{sup}}^{\mathcal{E}} f(\bar{x}) = B_{\mathcal{E}} \cap (\tau_{\mathcal{E}}^{-1}(0))^{\circ}$ , where

$$B_{\mathcal{E}} := \{x^* \in X^* : \langle x^*, x \rangle \leq f(\bar{x} + \tau_{\mathcal{E}}(x)x) - f(\bar{x}) \quad \forall x \in X \setminus \{\tau_{\mathcal{E}}^{-1}(0)\}\}.$$

Next we check that

$$\begin{aligned} B_{\mathcal{E}} &= \bigcap_{0 < \lambda \leq 1} \{x^* \in X^* : \langle x^*, x \rangle \leq f(\bar{x} + \lambda x) - f(\bar{x}) \quad \forall x \in \tau_{\mathcal{E}}^{-1}(\lambda) \setminus \tau_{\mathcal{E}}^{-1}(0)\} \\ &= \bigcap_{0 < \lambda \leq 1} \{\lambda x^* \in X^* : \langle \lambda x^*, x \rangle \leq f(\bar{x} + x) - f(\bar{x}) \quad \forall x \in \lambda (\tau_{\mathcal{E}}^{-1}(\lambda) \setminus \tau_{\mathcal{E}}^{-1}(0))\} \\ &= \bigcap_{0 < \lambda \leq 1} \{\lambda \partial_{C_{\lambda}^{\mathcal{E}}} f(\bar{x})\}. \end{aligned}$$

This proves (12). The other two representations are consequences of (12).  $\square$

Set  $L_f^>(\bar{x}) := \{x \in X : f(\bar{x} + x) > f(\bar{x})\}$ . The sets  $L_f^<(\bar{x})$ ,  $L_f^=(\bar{x})$  and  $L_f^{\leq}(\bar{x})$  are defined in a similar way. The next lemma provides explicit representations of the functions  $\tau_{\mathcal{E}}$  and  $\tau$  defined by (9) and (10) for an upper semi-continuous convex function.

**Lemma 3.2** *Let  $f$  be convex and  $\mathcal{E} \in \mathcal{F}$ . Then  $\tau_{\mathcal{E}}(x) = \tau(x) = 1$  for all  $x \in L_f^>(\bar{x})$ , and  $\tau(x) = 0$  for all  $x \in L_f^<(\bar{x})$ . If  $0 \in \mathcal{E}$ , then  $\tau_{\mathcal{E}}(x) = 0$  for all  $x \in L_f^{\leq}(\bar{x})$ .*

*Proof* By Lemma 3.1, we have

$$f(\bar{x} + tx) - f(\bar{x}) \leq \tau_{\mathcal{E}}(x)(f(\bar{x} + x) - f(\bar{x})) \quad (13)$$

for all  $t \in T_{\mathcal{E}} := \{\tau_{u^*}(x) : u^* \in \mathcal{E}\}$  and all  $x \in X$ . By the definition of  $\mathcal{F}$ , we always have  $1 \in T_{\mathcal{E}}$ , and  $0 \in T_{\mathcal{E}}$  if  $0 \in \mathcal{E}$ , particularly if  $\mathcal{E} = \mathbb{B}_{X^*}$ .

If  $x \in L_f^>(\bar{x})$ , then, by letting  $t = 1$  in (13), we obtain  $\tau_{\mathcal{E}}(x) \geq 1$ , and therefore,  $\tau_{\mathcal{E}}(x) = 1$ ; in particular,  $\tau(x) = 1$ . Let  $0 \in \mathcal{E}$ . If  $x \in L_f^<(\bar{x})$ , then, by letting  $t = 0$  in (13), we get  $\tau_{\mathcal{E}}(x) \leq 0$ , and therefore,  $\tau_{\mathcal{E}}(x) = 0$ ; in particular,  $\tau(x) = 0$ . If  $x \in L_f^=(\bar{x})$ , then, for all  $t \in [0, 1]$ , we have

$$f(\bar{x} + tx) = f((1-t)\bar{x} + t(\bar{x} + x)) \leq (1-t)f(\bar{x}) + tf(\bar{x} + x) = f(\bar{x}),$$

and consequently,  $\max_{t \in T_{\mathcal{E}}} f(\bar{x} + tx)$  is attained at  $t = 0$ . It follows from definition (9) that  $\tau_{\mathcal{E}}(x) = 0$ ; in particular,  $\tau(x) = 0$ .  $\square$

Using Lemma 3.2, we can simplify the conclusions of Theorem 3.1 for upper semi-continuous convex functions.

**Corollary 3.1** *Let  $f$  be convex and  $0 \in \mathcal{E} \in \mathcal{F}$ . Then*

$$\partial_{sup}^{\mathcal{E}} f(\bar{x}) = \partial_{sup} f(\bar{x}) = \partial_{L_f^>(\bar{x})} f(\bar{x}) \cap N_{L_f^{\leq}(\bar{x})}(0). \quad (14)$$

*As a consequence,  $\partial f(\bar{x}) = \partial_{L_f^>(\bar{x})} f(\bar{x}) \cap N_{L_f^{\leq}(\bar{x})}(0)$  if and only if  $\bar{x}$  minimizes  $f$ .*

*Proof* By Lemma 3.2, for all  $0 < \lambda < 1$ , we have  $\tau_{\mathcal{E}}^{-1}(\lambda) = \tau^{-1}(\lambda) = \emptyset$ , and consequently, using the notations in Theorem 3.1,  $\partial_{C_{\lambda}^{\mathcal{E}}} f(\bar{x}) = \partial_{C_{\lambda}} f(\bar{x}) = X^*$ . We also have  $C_1^{\mathcal{E}} = C_1 = L_f^>(\bar{x})$  and  $\tau_{\mathcal{E}}^{-1}(0) = \tau^{-1}(0) = L_f^{\leq}(\bar{x})$ . Hence, representations (11) and (12) reduce to (14). The last assertion follows thanks to Proposition 2.3(ii).  $\square$

**Corollary 3.2** *Let  $X := \ell^p(\mathbb{N})$  with  $p \geq 1$  and  $f$  be convex. Then  $\partial_{sup}^* f(\bar{x}) = \partial_{sup} f(\bar{x})$ . As a consequence,  $\partial_{sup}^* f(\bar{x}) = \partial f(\bar{x})$  if and only if  $\bar{x}$  minimizes  $f$ .*

*Proof* Let  $(e_k)$  denote the canonical basis of  $\ell^p(\mathbb{N})$  and  $\mathcal{E} \in \mathcal{F}$ . In view of Example 2.2,  $e_k \in \mathcal{E}$  for all  $k \in \mathbb{N}$ . The sequence  $(e_k)$  converges to 0 in the weak\* topology of  $\ell^q(\mathbb{N})$  where  $q$  and  $p$  are convex conjugates. Since  $\mathcal{E}$  is weak\*-closed,  $0 \in \mathcal{E}$ . The assertion follows from definition (6) and Corollary 3.1.  $\square$

**Remark 3.1** In general real topological vector spaces, the equality  $\partial_{sup}^* f(\bar{x}) = \partial f(\bar{x})$  may hold even if  $\bar{x}$  does not minimize  $f$ ; cf. Proposition 2.3(ii).

The sup-subdifferential can be connected with certain directional derivatives. Indeed, if  $f$  is convex, then for all  $x \in X$ , the function

$$h \mapsto \frac{\sup_{0 \leq t \leq h} f(\bar{x} + tx) - f(\bar{x})}{h}$$

is nondecreasing, and the function

$$x \mapsto f'_{sup}(\bar{x}; x) := \lim_{h \downarrow 0} \frac{\sup_{0 \leq t \leq h} f(\bar{x} + tx) - f(\bar{x})}{h}$$

is positively homogeneous (note that the limit exists in  $\mathbb{R} \cup \{\pm\infty\}$ ). It follows that

$$\partial_{sup} f(\bar{x}) = \{x^* \in X^* : \langle x^*, x \rangle \leq f'_{sup}(\bar{x}; x) \quad \forall x \in X\}.$$

#### 4 Symmetric Subdifferential

In this section, we assume that  $X$  is a topological vector space,  $f : X \rightarrow \mathbb{R}_{\infty}$ , and  $\bar{x} \in \text{dom } f$ . The symmetric subdifferential of  $f$  at  $\bar{x}$  is defined as

$$\partial_{sym} f(\bar{x}) := \{x^* \in X^* : \langle x^*, d \rangle \leq f'_{sym}(\bar{x}; d) \quad \forall d \in X\},$$

where

$$f'_{sym}(\bar{x}; d) := \lim_{t \downarrow 0} \frac{\max\{f(\bar{x} + td), f(\bar{x} - td)\} - f(\bar{x})}{t}$$

is the symmetric directional derivative of  $f$  at  $\bar{x}$  in direction  $d$  (if the limit exists in  $\mathbb{R} \cup \{\pm\infty\}$ ). If  $f$  is convex, then  $f'_{sym}(\bar{x}; d)$  exists, and is finite if  $\bar{x} \in \text{int dom } f$ . Indeed,

$$f'_{sym}(\bar{x}; d) = \max\{f'(\bar{x}; d), f'(\bar{x}; -d)\}, \quad (15)$$

where  $f'(\bar{x}; d)$  denotes the conventional directional derivative of  $f$  at  $\bar{x}$  in direction  $d \in X$ . Note that, if the (double-sided) limit  $\lim_{t \rightarrow 0} (f(\bar{x} + td) - f(\bar{x}))/t$  exists, then

$$f'_{sym}(\bar{x}; d) = |f'(\bar{x}; d)|. \quad (16)$$

When  $f$  is convex and continuous at  $\bar{x}$ , then

$$f'(\bar{x}; d) = \max\{\langle x^*, d \rangle : x^* \in \partial f(\bar{x})\}. \quad (17)$$

The following proposition, which is a direct consequence of (15) and (17), states a similar property for the symmetric directional derivative (see [11, formula (4)]).

**Proposition 4.1** *Suppose that  $f$  is convex and continuous at  $\bar{x}$ . Then*

$$f'_{\text{sym}}(\bar{x}; d) = \max \{ \langle x^*, d \rangle : x^* \in \partial f(\bar{x}) \cup \{-\partial f(\bar{x})\} \}.$$

The set  $\partial_{\text{sym}} f(\bar{x})$  is convex, weak\*-closed and symmetric. If the function  $d \mapsto f(\bar{x} + d)$  is bounded on a neighborhood of the origin, then  $\partial_{\text{sym}} f(\bar{x})$  is also weak\*-compact.  $\partial_{\text{sym}} f(\bar{x})$  contains  $\partial f(\bar{x})$ , since  $f'(\bar{x}; d) \leq f'_{\text{sym}}(\bar{x}; d)$  for all  $d \in X$ . Hence, we have the following sufficient condition of minimality.

**Proposition 4.2** *Suppose that  $f$  is convex. If  $\partial_{\text{sym}} f(\bar{x}) = \partial f(\bar{x}) \neq \emptyset$ , then  $\bar{x}$  minimizes  $f$ .*

*Proof* Under the assumptions,  $\partial f(\bar{x})$  is symmetric, and therefore  $0 \in \partial f(\bar{x})$ .  $\square$

The following theorem provides a sufficient condition under which the symmetric subdifferential is nonempty.

**Theorem 4.1** *Let  $f$  be convex. If there exists  $\bar{d} \in X$  such that*

$$0 < \max \{ f'(\bar{x}; \bar{d}), f'(\bar{x}; -\bar{d}) \} < +\infty,$$

*then  $\partial_{\text{sym}} f(\bar{x})$  contains a nonzero element.*

*Proof* One can easily check that  $f'_{\text{sym}}(\bar{x}; \alpha d) = |\alpha| f'_{\text{sym}}(\bar{x}; d)$  for all  $\alpha \in \mathbb{R}$  and  $d \in X$ . The function  $d \mapsto f'_{\text{sym}}(\bar{x}; d)$  is sub-additive. Indeed, for all  $d_1, d_2 \in X$ , we have

$$\begin{aligned} f'_{\text{sym}}(\bar{x}; d_1 + d_2) &= \max \{ f'(\bar{x}; d_1 + d_2), f'(\bar{x}; -d_1 - d_2) \} \\ &\leq \max \{ f'(\bar{x}; d_1) + f'(\bar{x}; d_2), f'(\bar{x}; -d_1) + f'(\bar{x}; -d_2) \} \\ &\leq \max \{ f'(\bar{x}; d_1), f'(\bar{x}; -d_1) \} + \max \{ f'(\bar{x}; d_2), f'(\bar{x}; -d_2) \} \\ &= f'_{\text{sym}}(\bar{x}; d_1) + f'_{\text{sym}}(\bar{x}; d_2). \end{aligned}$$

Thus, the function  $d \mapsto f'_{\text{sym}}(\bar{x}; d)$  is sublinear on  $X$ . Now let  $H := \mathbb{R}\{\bar{d}\}$  be the subspace generated by the nontrivial singleton  $\{\bar{d}\}$ . Define the functional  $l^* \in H^*$  as  $\langle l^*, h \rangle := \alpha f'_{\text{sym}}(\bar{x}; \bar{d})$ , where  $h = \alpha \bar{d}$ . Note that  $l^*$  is well-defined since  $\bar{d} \neq 0$  and  $f'_{\text{sym}}(\bar{x}; \bar{d})$  is finite. It follows that, for all  $h \in H$ ,

$$\langle l^*, h \rangle \leq |\alpha| f'_{\text{sym}}(\bar{x}; \bar{d}) = f'_{\text{sym}}(\bar{x}; \alpha \bar{d}) = f'_{\text{sym}}(\bar{x}; h).$$

By the Hahn–Banach theorem,  $l^*$  can be extended to a functional  $x^* \in X^*$  satisfying  $\langle x^*, d \rangle \leq f'_{\text{sym}}(\bar{x}; d)$  for all  $d \in X$ . Thus,  $0 \neq x^* \in \partial_{\text{sym}} f(\bar{x})$ .  $\square$

*Example 4.1* We consider the function  $f : \ell^2(\mathbb{N}) \rightarrow \mathbb{R}$  in Example 1.1. Let  $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots) \in C$  and  $\bar{d} = e_k$ , the  $k$ th basis vector in  $\ell^2(\mathbb{N})$  for some  $k \in \mathbb{N}$ . One can easily check that

$$f'(\bar{x}; \bar{d}) = -\frac{1}{2}(\bar{x}_k + 2^{-k})^{-\frac{1}{2}}, \quad f'(\bar{x}; -\bar{d}) = \frac{1}{2}(\bar{x}_k + 2^{-k})^{-\frac{1}{2}}.$$

Hence,  $\partial_{\text{sym}} f(\bar{x})$  contains a nonzero element.

*Example 4.2* Define  $f : \mathbb{R} \rightarrow \mathbb{R}$  as

$$f(x) := \begin{cases} x & \text{if } x > 0, \\ 1 - x & \text{if } x \leq 0. \end{cases}$$

It is discontinuous at 0 (though upper semi-continuous) and fails to be convex. One can easily check that  $\partial_{\text{sym}} f(0) = [-1, 1]$ . Indeed,  $\max \{ f'(0; d), f'(0; -d) \} = |d|$  for all  $d \in X$ .

## 5 Calculus Rules in Banach Spaces

In this section, we establish certain calculus rules for  $\partial_{sup}$ ,  $\partial_{sup}^*$  and  $\partial_{sym}$  for proper continuous convex functions on Banach spaces. Let  $f : X \rightarrow \mathbb{R}_\infty$  and  $\bar{x} \in \text{dom } f$ .

The next three rules are immediate from the definitions:

$$\partial_{sup}(\lambda f)(\bar{x}) = \lambda \partial_{sup}f(\bar{x}), \quad \partial_{sup}^*(\lambda f)(\bar{x}) = \lambda \partial_{sup}^*f(\bar{x}), \quad \partial_{sym}(\lambda f)(\bar{x}) = \lambda \partial_{sym}f(\bar{x})$$

for all  $\lambda > 0$ . If  $\partial_{sym}f(\bar{x}) \neq \emptyset$  and  $\partial_{sup}^*f(\bar{x}) \neq \emptyset$ , then these equalities also hold for  $\lambda = 0$ . We now proceed to sum rules.

We start with an auxiliary lemma for the symmetric subdifferential. Recall that a Banach space is *Asplund* if every continuous convex function on an open convex set is Fréchet differentiable on some its dense  $G_\delta$  subset, or equivalently, if the dual of each its separable subspace is separable [12, 13].

**Lemma 5.1** *Let  $f : X \rightarrow \mathbb{R}_\infty$  be a convex function on a Banach space, continuous at  $\bar{x}$ . Assume that the function  $x \mapsto f(\bar{x} + x)$  is bounded on  $\mathbb{B}_X$ . Then*

$$\partial_{sym}f(\bar{x}) = cl^{w^*} \text{co}(\partial f(\bar{x}) \cup (-\partial f(\bar{x}))), \quad (18)$$

$$\partial_{sup}f(\bar{x}) = cl^{w^*} \text{co}(\partial f(\bar{x}) \cup \{0\}), \quad (19)$$

where  $cl^{w^*}$  represents the closure with respect to the weak\* topology.

If  $X$  is *Asplund*, then  $cl^{w^*}$  in (18) can be replaced by the closure with respect to the norm topology.

*Proof* By the assumptions,  $\partial_{sym}f(\bar{x})$  is nonempty and weak\*-compact. Hence, by the Krein–Milman theorem [14],  $\partial_{sym}f(\bar{x})$  contains extreme points. Moreover,

$$\partial_{sym}f(\bar{x}) = cl^{w^*} \text{co} \text{ext}(\partial_{sym}f(\bar{x})),$$

where  $\text{ext } \partial_{sym}f(\bar{x})$  denotes the set of all extreme points of  $\partial_{sym}f(\bar{x})$ . By Proposition 4.1,  $\text{ext } \partial_{sym}f(\bar{x}) \subset \partial f(\bar{x}) \cup (-\partial f(\bar{x}))$ . Hence,

$$\partial_{sym}f(\bar{x}) \subset cl^{w^*} \text{co}(\partial f(\bar{x}) \cup (-\partial f(\bar{x}))).$$

On the other hand,  $\partial f(\bar{x}) \subset \partial_{sym}f(\bar{x})$  and, since  $\partial_{sym}f(\bar{x})$  is symmetric, convex and weak\*-closed,  $cl^{w^*} \text{co}(\partial f(\bar{x}) \cup (-\partial f(\bar{x}))) \subset \partial_{sym}f(\bar{x})$ . This proves (18). The proof of the second equality goes along the same lines as that of (18); we therefore give only a sketch of it. Since  $f$  is continuous,  $f'_{sup}(\bar{x}; \cdot) = \max\{f'(\bar{x}; \cdot), 0\}$  ( $\tau(x)$  equals either 1 or 0 for all  $x \in X$ ), and therefore

$$\partial_{sup}f(\bar{x}) = \{x^* \in X^* : \langle x^*, x \rangle \leq \max\{f'(\bar{x}; d), 0\} \quad \forall x \in X\}.$$

One can easily check that  $f'_{sup}(\bar{x}; d) = \max\{\langle x^*, d \rangle : x^* \in \partial f(\bar{x}) \cup \{0\}\}$ . By the Krein–Milman theorem,  $\partial_{sup}f(\bar{x}) = cl^{w^*} \text{co}(\partial f(\bar{x}) \cup \{0\})$ . This proves (19). If  $X$  is *Asplund*, its dual  $X^*$  has the Radon–Nikodým property [12], and it follows from the Edgar–Lindenstrauss theorem [15, 16] that the weak\*-closure can be replaced by the norm closure.  $\square$

*Remark 5.1* The above proof uses the fact that the dual of an *Asplund* space has the Radon–Nikodým property. In fact, a Banach space is *Asplund* if and only if its dual has the Radon–Nikodým property [12, Theorem 5.7], [13, Theorem 6].

**Theorem 5.1** *Let  $A : X \rightarrow Y$  be a bounded linear map between Banach spaces,  $f : X \rightarrow \mathbb{R}_\infty$  and  $g : Y \rightarrow \mathbb{R}_\infty$  be proper convex functions such that  $f$  and  $g \circ A$  are finite and continuous at  $\bar{x}$ . Suppose that  $0 \in \text{core}(\text{dom } g - A \text{ dom } f)$ , and the functions  $x \mapsto f(\bar{x} + x)$  and  $y \mapsto g(A\bar{x} + y)$  are bounded on  $\mathbb{B}_X$  and  $\mathbb{B}_Y$ , respectively. The following assertions hold true.*

- (i)  $\partial_{\text{sym}}(f + g \circ A)(\bar{x}) \subset \partial_{\text{sym}}f(\bar{x}) + A^* \partial_{\text{sym}}g(A\bar{x})$ .  
 Furthermore, if for any  $d \in X$ ,  $f'(\bar{x}; d) \geq f'(\bar{x}; -d)$  implies  $g'(A\bar{x}; Ad) \geq g'(A\bar{x}; -Ad)$ , then  $\partial_{\text{sym}}(f + g \circ A)(\bar{x}) = \partial_{\text{sym}}f(\bar{x}) + A^* \partial_{\text{sym}}g(A\bar{x})$ .
- (ii)  $\partial_{\text{sup}}(f + g \circ A)(\bar{x}) \subset \partial_{\text{sup}}f(\bar{x}) + A^* \partial_{\text{sup}}g(A\bar{x})$ .  
 Furthermore, if for any  $d \in X$  we have  $f'(\bar{x}; d) \geq 0 \iff g'(A\bar{x}; Ad) \geq 0$ , then  $\partial_{\text{sup}}(f + g \circ A)(\bar{x}) = \partial_{\text{sup}}f(\bar{x}) + A^* \partial_{\text{sup}}g(A\bar{x})$ .

*Proof* (i). The function  $x \mapsto (f + g \circ A)(\bar{x} + x)$  is bounded on  $\mathbb{B}_X$ , and therefore satisfies the conditions of Lemma 5.1. The adjoint operator  $A^* : Y^* \rightarrow X^*$  is weak\*-continuous, and therefore, maps a weak\*-compact set in  $Y^*$  to a weak\*-compact set in  $X^*$ . From these observations, the convex subdifferential sum and chain rules [9], and Lemma 5.1, we have

$$\begin{aligned}
 & \partial_{\text{sym}}(f + g \circ A)(\bar{x}) \\
 &= cl^{w^*} \text{co} (\partial(f + g \circ A)(\bar{x}) \cup (-\partial(f + g \circ A)(\bar{x}))) \\
 &= cl^{w^*} \text{co} ((\partial f(\bar{x}) + A^* \partial g(A\bar{x})) \cup (-\partial f(\bar{x}) - A^* \partial g(A\bar{x}))) \\
 &\subset cl^{w^*} \text{co} ((\partial f(\bar{x}) \cup (-\partial f(\bar{x}))) + (A^* \partial g(A\bar{x}) \cup (-A^* \partial g(A\bar{x})))) \\
 &\subset cl^{w^*} \left( cl^{w^*} \text{co} (\partial f(\bar{x}) \cup (-\partial f(\bar{x}))) + cl^{w^*} A^* \text{co} (\partial g(A\bar{x}) \cup (-\partial g(A\bar{x}))) \right) \\
 &\subset cl^{w^*} \left( cl^{w^*} \text{co} (\partial f(\bar{x}) \cup (-\partial f(\bar{x}))) + A^* cl^{w^*} \text{co} (\partial g(A\bar{x}) \cup (-\partial g(A\bar{x}))) \right) \\
 &= cl^{w^*} (\partial_{\text{sym}}f(\bar{x}) + A^* \partial_{\text{sym}}g(A\bar{x})) = \partial_{\text{sym}}f(\bar{x}) + A^* \partial_{\text{sym}}g(A\bar{x}),
 \end{aligned}$$

since the sum of two weak\*-compact sets is weak\*-closed.

Now suppose that  $x^* \in \partial_{\text{sym}}f(\bar{x})$ ,  $y^* \in \partial_{\text{sym}}g(A\bar{x})$  and  $u^* = A^*y^*$ . Let  $d \in X$ . By the assumptions, we have

$$\begin{aligned}
 \langle x^* + u^*, d \rangle &= \langle x^*, d \rangle + \langle y^*, Ad \rangle \\
 &\leq \max\{f'(\bar{x}; d), f'(\bar{x}; -d)\} + \max\{g'(A\bar{x}; Ad), g'(A\bar{x}; -Ad)\} \\
 &= \max\{f'(\bar{x}; d) + g'(A\bar{x}; Ad), f'(\bar{x}; -d) + g'(A\bar{x}; -Ad)\} \\
 &= \max\{(f + g \circ A)'(\bar{x}; d), (f + g \circ A)'(\bar{x}; -d)\}.
 \end{aligned}$$

It follows that  $x^* + u^* \in \partial_{\text{sym}}(f + g \circ A)(\bar{x})$ , and therefore,

$$\partial_{\text{sym}}f(\bar{x}) + A^* \partial_{\text{sym}}g(A\bar{x}) \subset \partial_{\text{sym}}(f + g \circ A)(\bar{x}).$$

- (ii) The proof goes along the same lines as that of (i). We therefore omit it.  $\square$

The next statement is a straightforward consequence of Theorem 5.1 and Corollary 3.2.

**Theorem 5.2** *Let  $p \geq 1$  and  $f, g : \ell^p(\mathbb{N}) \rightarrow \mathbb{R}_\infty$  be proper convex functions, continuous at  $\bar{x} \in \text{dom } f \cap \text{dom } g$ . Suppose that  $0 \in \text{core}(\text{dom } f - \text{dom } g)$ , and the functions  $x \mapsto f(\bar{x} + x)$  and  $x \mapsto g(\bar{x} + x)$  are bounded on  $\mathbb{B}_X$ . Then*

$$\partial_{\text{sup}}^*(f + g)(\bar{x}) \subset \partial_{\text{sup}}^*f(\bar{x}) + \partial_{\text{sup}}^*g(\bar{x}).$$

*If for any  $d \in X$  we have  $f'(\bar{x}; d) \geq 0 \iff g'(\bar{x}; d) \geq 0$ , then*

$$\partial_{\text{sup}}^*(f + g)(\bar{x}) = \partial_{\text{sup}}^*f(\bar{x}) + \partial_{\text{sup}}^*g(\bar{x}).$$

## Acknowledgments

The authors thank the anonymous referees for their valuable comments that have led to improvements in the presentation. References [4, 6, 10] have been provided by one of the referees.

The research was supported by the Australian Research Council, project DP160100854.



## References

1. Burachik, R.S., Iusem, A.N.: Set-Valued Mappings and Enlargements of Monotone Operators, *Springer Optimization and Its Applications*, vol. 8. Springer, New York (2008)
2. Kutateladze, S.S., Rubinov, A.M.: Minkowski duality, and its applications. *Uspehi Mat. Nauk* **27**(3(165)), 127–176 (1972)
3. Dolecki, S., Kurcyusz, S.: On  $\Phi$ -convexity in extremal problems. *SIAM J. Control Optim.* **16**(2), 277–300 (1978). URL <https://doi.org/10.1137/0316018>
4. Penot, J.P.: What is quasiconvex analysis? *Optimization* **47**(1-2), 35–110 (2000). URL <https://doi.org/10.1080/02331930008844469>
5. Revalski, J.P., Théra, M.: Enlargements and sums of monotone operators. *Nonlinear Anal., Ser. A: Theory Methods* **48**(4), 505–519 (2002). URL [https://doi.org/10.1016/S0362-546X\(00\)00201-7](https://doi.org/10.1016/S0362-546X(00)00201-7)
6. Carrasco-Olivera, D., Flores-Bazán, F.: On the representation of approximate subdifferentials for a class of generalized convex functions. *Set-Valued Anal.* **13**(2), 151–166 (2005). URL <https://doi.org/10.1007/s11228-004-4316-4>
7. Martínez-Legaz, J., Théra, M.:  $\varepsilon$ -subdifferentials in terms of subdifferentials. *Set-Valued Anal.* **4**(4), 327–332 (1996). URL <https://doi.org/10.1007/BF00436109>
8. Rainwater, J.: Yet more on the differentiability of convex functions. *Proc. Amer. Math. Soc.* **103**(3), 773–778 (1988). URL <https://doi.org/10.2307/2046851>
9. Borwein, J., Zhu, Q.J.: Techniques of Variational Analysis, *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC*, vol. 20. Springer-Verlag, New York (2005)
10. Flores-Bazán, F.: Radial epiderivatives and asymptotic functions in nonconvex vector optimization. *SIAM J. Optim.* **14**(1), 284–305 (2003). URL <https://doi.org/10.1137/S1052623401392111>
11. Kruger, A.Y.: Generalized differentials of nonsmooth functions and necessary conditions for an extremum. *Sibirsk. Mat. Zh.* **26**(3), 78–90, 224 (1985). (In Russian; English transl.: *Siberian Math. J.* **26** (1985), 370–379)
12. Phelps, R.R.: Convex Functions, Monotone Operators and Differentiability, *Lecture Notes in Mathematics*, vol. 1364, 2nd edn. Springer-Verlag, Berlin (1993)
13. Yost, D.: Asplund spaces for beginners. *Acta Univ. Carolin. Math. Phys.* **34**(2), 159–177 (1993)
14. Bourbaki, N.: *Espaces Vectoriels Topologiques. Chapitres 1 à 5*, new edn. Masson, Paris (1981). *Éléments de mathématique. [Elements of mathematics]*
15. Stegall, C.: The Radon-Nikodym property in conjugate Banach spaces. *Trans. Amer. Math. Soc.* **206**, 213–223 (1975). URL <https://doi.org/10.2307/1997154>
16. Phelps, R.R.: Dentability and extreme points in Banach spaces. *J. Functional Analysis* **17**, 78–90 (1974). URL [https://doi.org/10.1016/0022-1236\(74\)90005-6](https://doi.org/10.1016/0022-1236(74)90005-6)