Noname manuscript No.
(will be inserted by the editor)

Enlargements of the Moreau–Rockafellar Subdifferential

Malek Abbasi · **Alexander Y. Kruger** · **Michel Théra**

Received: date / Accepted: date

Abstract The Moreau–Rockafellar subdifferential is a highly important notion in convex analysis and optimization theory. But there are many functions which fail to be subdifferentiable at certain points. In particular, there is a continuous convex function defined on $\ell^2(\mathbb{N})$, whose Moreau–Rockafellar subdifferential is empty at every point of its domain. This paper proposes some enlargements of the Moreau–Rockafellar subdifferential: the sup*-subdifferential, sup-subdifferential and symmetric subdifferential, all of them being nonempty for the mentioned function. These enlargements satisfy the most fundamental properties of the Moreau–Rockafellar subdifferential: convexity, weak*-closedness, weak*-compactness and, under some additional assumptions, possess certain calculus rules. The sup* and sup subdifferentials coincide with the Moreau–Rockafellar subdifferential at every point at which the function attains its minimum, and if the function is upper semi-continuous, then there are some relationships for the other points. They can be used to detect minima and maxima of arbitrary functions.

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

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

1 Motivation

Throughout the paper, X is a real topological vector space. The topological (continuous) dual of X is denoted by X^*. If $T : X \rightrightarrows X^*$ is a set-valued mapping, the set of all $x \in X$ such that $T(x)$ is nonempty is the domain of T and is denoted by $\operatorname{dom} T$.

Various problems coming from different areas can be formulated as

$$\boxed{\text{Find } x \in \operatorname{dom} T \text{ such that } 0 \in T(x)} \tag{OP}$$

Malek Abbasi
Department of Mathematics, University of Isfahan, Isfahan, Iran
E-mail: malek.abbasi@sci.ui.ac.ir
Alexander Y. Kruger ORCID 0000-0002-7861-7380
Centre for Informatics and Applied Optimization, School of Science, Engineering and Information Technology, Federation University Australia, 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
E-mail: michel.thera@icloud.com

(see [8–10] and the references therein). We denote this problem by (OP) (Original Problem) for easy reference. When T is the subdifferential mapping of some extended-real-valued convex function $f : X \to \mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$, then (OP) becomes the Fermat rule:

$$0 \in T(x) \quad \Longleftrightarrow \quad x \text{ minimizes } f$$

which is one of the central facts in optimization theory.

In some situations, it can be convenient to consider an *enlargement* of T: a set-valued mapping $T' : X \rightrightarrows X^*$ such that $T(x) \subset T'(x)$ and $T'(x)$ is convex for all $x \in X$, and being "close" to T in a sense which will be specified later, and study the auxiliary problem

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

Solutions of the latter problem can serve as approximate solutions of the original problem (OP).

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

$$\partial_\varepsilon f(x) = \{x^* \in X^* : f(y) - f(x) \geq \langle x^*, y - x\rangle - \varepsilon \quad \text{for all } y \in X\}$$

is an enlargement of the conventional Moreau–Rockafellar subdifferential. It plays a key role in the theory of extremal problems and has been successfully used to construct numerical methods for minimizing convex functions.

Following this idea, given a monotone operator A acting between X and X^*, and $\varepsilon \geq 0$, Revalski and Théra [23] defined an enlargement $A_\varepsilon : X \to X^*$ of A by

$$A^\varepsilon x := \{x^* \in X^* : \langle y - x, x^* - y^*\rangle \geq -\varepsilon \quad \text{for all} \quad (y, y^*) \in \operatorname{gph} A\},$$

where $x \in X$ and $\varepsilon \geq 0$. A^ε has convex and weak*-closed values for all $\varepsilon \geq 0$ and

$$Ax \subset A^\varepsilon x \quad \text{for all } \varepsilon \geq 0 \text{ and } x \in X.$$

When $A = \partial f$, it holds $\partial_\varepsilon f(x) \subset (\partial f)^\varepsilon(x)$, and the inclusion can be strict, as the next example by Martínez-Legaz and Théra [16] shows:

$$f(x) = x^2, \quad 0 \notin \partial_{\frac{1}{2}} f(1) \quad \text{but} \quad 0 \in (\partial f)^{\frac{1}{2}}(1).$$

Now consider the following problem:

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

For the sake of convenience, we denote this problem by (EP) (Enlargement Problem).

Suppose that T' is a solution of (EP). For $\varepsilon \in [0,1]$, define $T_\varepsilon(x) := \varepsilon T'(x) + (1-\varepsilon)T(x)$ for $x \in \operatorname{dom} T$ and $T_\varepsilon(x) := T'(x)$ for $x \in \operatorname{dom} T' \setminus \operatorname{dom} T$. Obviously $T(x) \subset T_\varepsilon(x) \subset T'(x)$ and $T_1(x) = T'(x)$ for all $x \in X$. Let $A \subset X$ be a given nonempty subset of X with $A \cap \operatorname{dom} T' \neq \emptyset$. Set

$$\varepsilon_0 := \inf\{\varepsilon \in [0,1] : \ \exists x \in A \text{ such that } 0 \in T_\varepsilon(x)\}.$$

Obviously $0 \leq \varepsilon_0 \leq 1$. Hence, there exist a sequence $(x_n) \in A$ and a decreasing sequence $(\varepsilon_n) \in]0,1]$ converging to ε_0 such that $0 \in T_{\varepsilon_n}(x_n)$. If A is compact and T' satisfies some continuity properties, then one can show that $0 \in T_{\varepsilon_0}(\bar{x})$ for some $\bar{x} \in A$. Therefore, such a T_{ε_n} (and T_{ε_0}) allows us to define perturbations of the problem (OP). If $\varepsilon_0 = 0$ (and therefore $T_{\varepsilon_0} = T$), then we can formulate the following Auxiliary Problem:

$$\boxed{\text{Find } x \in A \text{ such that } 0 \in T_{\varepsilon_n}(x),} \tag{AP1}$$

which may be more tractable and somewhat easier to handle. In this case, x_n is a solution of (AP1) and, under the compactness and continuity assumptions, the sequence (x_n) has a subsequence which converges to some solution $\bar{x} \in A$ of (OP). If $\varepsilon_0 > 0$, then (OP) fails to have a solution in A. In this case, instead of A, we can consider an increasing sequence of closed subsets $A_n \subset X$ ($A_1 \subset A_2 \subset \cdots$) with $X = \cup_{n=1}^{\infty} A_n$ such that $A_1 \cap \mathrm{dom}\, T' \neq \emptyset$ (implying $A_n \cap \mathrm{dom}\, T' \neq \emptyset$ for all $n \in \mathbb{N}$).

Then $\varepsilon_n := \inf\{\varepsilon \in [0,1] : \ \exists x \in A_n \text{ such that } T_\varepsilon(x)\}$ is a decreasing sequence in $[0,1]$ (converging to zero), and we can formulate another Auxiliary Problem:

$$\boxed{\text{Find } x_n \in \mathrm{dom}\, T_{\varepsilon_n} \text{ such that } 0 \in T_{\varepsilon_n}(x_n).} \tag{AP2}$$

Finding an appropriate enlargement T' (as a solution of (EP)) plays a key role in this procedure. Notice that $T' := X^*$ is a solution of (EP) but it is not appropriate for our purposes. Indeed, by letting $T' := X^*$, we have $T_\varepsilon(x) = X^*$ for all $0 < \varepsilon \leq 1$ and $T_\varepsilon(x) = T$ for $\varepsilon = 0$. Such an enlargement is useless.

The main scope of this paper is to find solutions (enlargements) of problem (EP), close to T in a certain sense, when T is the subdifferential operator: $T := \partial f$. In this case, finding a close solution of problem (EP) means to find an enlargement T' of ∂f such that T' satisfies the fundamental properties of ∂f such as convexity, weak*-closedness, weak*-compactness and certain calculus rules.

The enlargements $\partial_{sup} f$ and $\partial_{sym} f$ defined below and the corresponding to them set-valued mapping T_ε satisfy the mentioned fundamental properties. Moreover, the subdifferential equations $\partial_{sup} f(x) = \{0\}$ and $\partial_{sup}(-f)(x) = \{0\}$ can be used for detecting minima and maxima of an arbitrary function f; cf. Corollary 3.1, and Examples 3.4 and 3.5.

The paper is dedicated to Terry Rockafellar, the ruler of Convex Analysis and Endolandia on the occasion of his 85th birthday.

2 Introduction

Let us start with recalling some basic concepts and terminology. Let $f : X \to \mathbb{R}_\infty$ be an extended-real-valued function. The directional derivative of f at $\bar{x} \in \mathrm{dom}\, f := \{x \in X : f(x) < +\infty\}$ in direction $d \in X$, denoted by $f'(\bar{x}; d)$, is defined by the following limit:

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

We say that f is directionally differentiable at $\bar{x}$ if the above limit exists in $\mathbb{R} \cup \{\pm\infty\}$ for all $d \in X$. In this case, the subdifferential of f at $\bar{x}$ is the set (cf. [12, 14, 21])

$$\partial f(\bar{x}) := \{x^* \in X^* : \langle x^*, d\rangle \leq f'(\bar{x}; d) \quad \forall d \in X\}. \tag{1}$$

If f is convex, then f is directionally differentiable at every $\bar{x} \in \mathrm{dom}\, f$, and the set (1) coincides with the Moreau–Rockafellar subdifferential of f at $\bar{x}$:

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

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

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

If $C = \emptyset$, we set $\partial_C f(\bar{x}) := X^*$.

The Moreau–Rockafellar subdifferential has proven to be a powerful tool in convex analysis and optimization theory (see [1–6, 11, 15, 17, 18, 25, 28, 30] and the references therein). It is well known that, even if f is continuous at $\bar{x}$, the subdifferential $\partial f(\bar{x})$ can be empty. The next example is due to Rainwater [22]. (We refer the readers to [5, 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 $\|x\| := \sum_{n=1}^{\infty} |x_n|^2 < \infty$.

Example 2.1 Define $f : \ell^2(\mathbb{N}) \to \mathbb{R}$ by

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

where

$$C := \{x \in \ell^2(\mathbb{N}) : |x_n| \le 2^{-n}, \quad n = 1, 2, \cdots\}.$$

It is easy to check that f is convex, continuous on C, and $\partial f(x) = \emptyset$ for all $x \in \mathscr{N}(C)$, where

$$\mathscr{N}(C) := \{x \in \ell^2(\mathbb{N}) : |x_n| < 2^{-n} \quad \text{for infinitely many } n\}.$$

This paper proposes three enlargements of the subdifferential of a directionally differentiable function: the sup*-subdifferential, sup-subdifferential (section 3) and symmetric subdifferential (section 4). Each of these enlargements is nonempty for the Rainwater function in Example 2.1 and satisfies the fundamental properties of the Moreau–Rockafellar subdifferential: convexity, weak*-closedness and weak*-compactness. 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 ∂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 sup*- and sup-subdifferentials also provide some optimality conditions for convex and nonconvex nonsmooth functions (Proposition 3.2, Examples 3.4 and 3.5, and Corollaries 3.1 and 3.3). The symmetric subdifferential under a mild condition contains a nonzero element. More precisely, if there exists a direction $\bar{d}$ such that the maximum of $f'(\bar{x};\bar{d})$ and $f'(\bar{x};-\bar{d})$ is positive and finite, then the symmetric subdifferential of f at $\bar{x}$ contains a nonzero continuous linear functional. Note that the function (2) satisfies this condition. Thus, the symmetric subdifferential of the function (2) contains a nonzero element at every point at which the Moreau–Rockafellar subdifferential of this function is empty (Example 4.1). The mentioned enlargements also possess some calculus rules, and therefore are close to the subdifferential operator ∂f in the sense that they behave very much like ∂f. This is why we call each of these enlargements a "subdifferential".

The paper is organized as follows. In section 3, we define the sup- and sup*-subdifferentials and verify their properties. In section 4, we define the symmetric subdifferential and state its fundamental properties. In section 5, we prove some calculus rules for these subdifferentials.

In what follows, we consider a directionally differentiable function $f : X \to \mathbb{R}_\infty$ defined on a real topological vector space.

3 The 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 ball in X and X^*, respectively.

The subset $\mathscr{E} \subset \mathbb{B}_{X^*}$ is said to be norm-generating, if for any $x \in X$ there exists $e^* \in \mathscr{E}$ such that $|\langle e^*, x\rangle| = \|x\|$. The collection of all weak* closed norm-generating subsets is denoted by $\mathscr{F}$. By the Hahn–Banach theorem, $\mathbb{B}_{X^*} \in \mathscr{F}$ (see [24, 26]).

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

Example 3.1 Equip $\mathbb{R}^n$ with the max norm $\|x\|_{\max} := \max\{|x_k| : 1 \le k \le n\}$ for $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$. Let $\{e_1, e_2, \ldots, e_n\}$ be the canonical basis of $\mathbb{R}^n$. For any $x \in \mathbb{R}^n$, we have $\|x\|_{\max} = |x_k|$ for some $1 \le k \le n$, and $|\langle e_k, x\rangle| = |x_k| = \|x\|_{\max}$. Hence, $\{e_1, e_2, \ldots, e_n\}$ is a closed norm-generating subset of $\mathbb{R}^n$.

Moreover, any norm-generating subset $\mathscr{E}$ of $\mathbb{R}^n$ contains either e_k or $-e_k$ for all $k = 1, \ldots, n$. Indeed, by the definition of $\mathscr{E}$, for any $k = 1, \ldots, n$, there exists some $u = (u_1, \ldots, u_n) \in \mathscr{E}$ 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, $\|u\|_{(\mathbb{R}^n)^*} = \sum_{i=1}^n |u_i| \le 1$. Hence, $u_i = 0$ for all $i \ne k$, 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 $(1 \le p \le +\infty)$. Recall that $\ell^p(\mathbb{N})$ is the linear space of all real sequences $x := (x_k)$ such that $\|x\|_p := \sum_{k=1}^\infty |x_k|^p < \infty$ if $p < \infty$, and $\|x\|_\infty := \max_{k\in\mathbb{N}} |x_k| < \infty$.

Example 3.2 The canonical basis $\{e_1, e_2, \ldots, \}$ (i.e., e_k is a sequence whose only non-zero entry is a "1" in the kth coordinate) is a norm-generating subset of $\ell^\infty(\mathbb{N})$. Any norm-generating subset of $\ell^p(\mathbb{N})$ $(1 \le p \le +\infty)$ contains either e_k or $-e_k$ for all $k = 1, 2, \ldots$.

3.1 Definitions and Fundamental Properties

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

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

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

Let $\bar{x} \in \operatorname{dom} f$. The sets

$$\partial^{\mathscr{E}}_{sup} f(\bar{x}) := \left\{ x^* \in X^* : \langle x^*, x\rangle \le \sup_{u^*\in\mathscr{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}) \quad \forall x \in X \right\}, \tag{3}$$

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

$$\partial^{\star}_{sup} f(\bar{x}) := \bigcap_{\mathscr{E}\in\mathscr{F}} \partial^{\mathscr{E}}_{sup} f(\bar{x}) \tag{5}$$

are called, respectively, the $\sup_{\mathscr{E}}$-subdifferential, sup-subdifferential and $\sup^\star$-sub-differential of f at $\bar{x}$. The first one determined by a given norm-generating set $\mathscr{E} \in \mathscr{F}$. Note that $\partial_{sup} f(\bar{x})$ is a particular case of $\partial^{\mathscr{E}}_{sup} f(\bar{x})$ with $\mathscr{E} := \mathbb{B}_{X^*}$.

Proposition 3.1 $\partial_{sup} f(\bar{x}) = \partial^{\mathbb{B}_{X^*}}_{sup} f(\bar{x})$.

Proof If $u^* \in \mathbb{B}_{X^*}$, then $0 \le \tau_{u^*}(x) \le 1$. Hence, $\partial^{\mathbb{B}_{X^*}}_{sup} f(\bar{x}) \subset \partial_{sup} f(\bar{x})$. 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, $\partial_{sup} f(\bar{x}) \subset \partial^{\mathbb{B}_{X^*}}_{sup} f(\bar{x})$. □

The next example is an extension of Example 2.1 to the case of an $\ell^p(\mathbb{N})$ space $(1 \le p < \infty)$.

Example 3.3 Define $f : \ell^p(\mathbb{N}) \to \mathbb{R}$ 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

$$C := \{x \in \ell^p(\mathbb{N}) : |x_n| \le 2^{-\frac{2n}{p}}, \quad n = 1, 2, \dots\}.$$

Let $\bar{x} \in C$. We show that $0 \in \partial^{\star}_{sup} f(\bar{x})$. 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 $\mathscr{E} \in \mathscr{F}$. In view of Example 3.2, $e_k \in \mathscr{E}$, for all $k \in \mathbb{N}$. Now let $x = (x_n) \in C$ and $x \ne 0$. 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,

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

Since f is continuous, by letting $k \to \infty$, we obtain

$$\sup_{u^* \in \mathscr{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}) \ge 0.$$

This implies that $0 \in \partial^{\mathscr{E}}_{sup} f(\bar{x})$, and consequently, $0 \in \partial^{\star}_{sup} f(\bar{x})$. □

Proposition 3.2 *Let $\bar{x} \in \operatorname{dom} f$. The following assertions hold true:*

(i) $\partial^{\mathscr{E}}_{sup} f(\bar{x})$ is convex and weak*-closed for all $\mathscr{E} \in \mathscr{F}$. As a consequence, $\partial_{sup} f(\bar{x})$ and $\partial^{\star}_{sup} f(\bar{x})$ are convex and weak*-closed.
(ii) $0 \in \partial_{sup} f(\bar{x})$. If $\bar{x}$ maximizes f, then $\partial_{sup} f(\bar{x}) = \{0\}$.
(iii) If f is convex, then $\partial f(\bar{x}) \subset \partial^{\mathscr{E}}_{sup} f(\bar{x}) \subset \partial_{sup} f(\bar{x})$ for all $\mathscr{E} \in \mathscr{F}$. As a consequence, $\partial f(\bar{x}) \subset \partial^{\star}_{sup} f(\bar{x}) \subset \partial_{sup} f(\bar{x})$.
(iv) If f is convex, then $\bar{x}$ minimizes f if and only if $\partial_{sup} f(\bar{x}) = \partial f(\bar{x})$. As a consequence, $\bar{x}$ minimizes f if and only if $\partial f(\bar{x}) = \partial^{\star}_{sup} f(\bar{x}) = \partial_{sup} f(\bar{x})$.
(v) If f is convex and $\partial_{sup} f(\bar{x})$ is a singleton, then either $\bar{x}$ minimizes f or $\partial f(\bar{x}) = \emptyset$.
(vi) Suppose that the function $x \mapsto f(\bar{x} + x)$ is bounded on $\mathbb{B}_X$. Then $\partial^{\mathscr{E}}_{sup} f(\bar{x})$ is weak*-compact for all $\mathscr{E} \in \mathscr{F}$. As a consequence, $\partial_{sup} f(\bar{x})$ and $\partial^{\star}_{sup} f(\bar{x})$ are weak*-compact.
(vii) Suppose that X is finite dimensional and f is continuous. Then $\partial^{\mathscr{E}}_{sup} f(\bar{x})$ is compact for all $\mathscr{E} \in \mathscr{F}$. As a consequence, $\partial_{sup} f(\bar{x})$ and $\partial^{\star}_{sup} f(\bar{x})$ are compact.

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

$$\langle x_1^*, x \rangle \le \sup_{u^* \in \mathscr{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}),$$
$$\langle x_2^*, x \rangle \le \sup_{u^* \in \mathscr{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}),$$

and consequently,

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

Hence, $\partial^{\mathscr{E}}_{sup} f(\bar{x})$ is convex. Let $(x_\gamma^*)_{\gamma \in \Gamma}$ be a net in $\partial^{\mathscr{E}}_{sup} f(\bar{x})$ converging to some $x^* \in X^*$ in weak*-topology of X^*. Let $x \in X$. For all $\gamma \in \Gamma$, we have

$$\langle x_\gamma^*, x \rangle \le \sup_{u^* \in \mathscr{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}),$$

and consequently,

$$\langle x^*, x \rangle \le \sup_{u^* \in \mathscr{E}} f(\bar{x} + \tau_{u^*}(x)x) - f(\bar{x}).$$

Hence, $\partial_{sup}^{\mathscr{E}} f(\bar{x})$ is weak*-closed. In view of Proposition 3.1, 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 $\sup_{0\le t\le 1} f(\bar{x}+tx) \ge f(\bar{x})$ for all $x \in X$. It follows from definition (4) that $0 \in \partial_{sup} f(\bar{x})$. If $\bar{x}$ maximizes f, then $f(\bar{x}+tx) - f(\bar{x}) \le 0$ for all $x \in X$ and all $0 \le t \le 1$, and consequently, $\sup_{0\le t\le 1} f(\bar{x}+tx) = f(\bar{x})$ for all $x \in X$. Hence, $\partial_{sup} f(\bar{x}) = \{0\}$.

(iii). Let f be convex, $\mathscr{E} \in \mathscr{F}$, $x^* \in \partial f(\bar{x})$ and $x \in X$. Then

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

and there exists $\hat{u}^* \in \mathscr{E}$ such that $|\langle \hat{u}^*, x\rangle| = \|x\|$, i.e. $\tau_{\hat{u}^*}(x) = 1$. Therefore

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

and consequently,

$$\langle x^*, x\rangle \le \sup_{u^*\in\mathscr{E}} f(\bar{x}+\tau_{u^*}(x)x) - f(\bar{x}).$$

It follows that $x^* \in \partial_{sup}^{\mathscr{E}} f(\bar{x})$, and consequently, $\partial f(\bar{x}) \subset \partial_{sup}^{\mathscr{E}} f(\bar{x})$. The opposite inclusion $\partial_{sup}^{\mathscr{E}} f(\bar{x}) \subset \partial_{sup} f(\bar{x})$ is straightforward from definitions (3) and (4). The second claim is a consequence of the first one.

(iv). Let f be convex. If $\partial f(\bar{x}) = \partial_{sup} f(\bar{x})$, then by (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

$$\begin{aligned}\langle x^*, x\rangle &\le \sup_{0\le t\le 1} f(\bar{x}+tx) - f(\bar{x})\\ &\le \sup_{0\le t\le 1} t(f(\bar{x}+x) - f(\bar{x})) = f(\bar{x}+x) - f(\bar{x}),\end{aligned}$$

It follows that $x^* \in \partial(\bar{x})$, and consequently, $\partial_{sup} f(\bar{x}) \subset \partial f(\bar{x})$. In view of (iii), we have $\partial_{sup} f(\bar{x}) = \partial f(\bar{x})$. The second claim is a consequence of the first one and (iii).

(v). Let f be convex and $\partial_{sup} f(\bar{x})$ be a singleton. By (ii), $\partial_{sup} f(\bar{x}) = \{0\}$. Hence by (iii), either $\partial f(\bar{x}) = \emptyset$ or $\partial f(\bar{x}) = \{0\}$. In the latter case, $\bar{x}$ minimizes f.

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

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

Thus, $\partial_{sup}^{\mathscr{E}} f(\bar{x})$ 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.

(vii). 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). □

Proposition 3.2(ii) yields necessary conditions of optimality.

Corollary 3.1 *Let $\bar{x} \in \operatorname{dom} f$. If $\bar{x}$ maximizes f, then $\partial_{sup} f(\bar{x}) = \{0\}$.*
If $\bar{x}$ minimizes f, then $\partial_{sup}(-f)(\bar{x}) = \{0\}$.

Example 3.4 Define $f:\mathbb{R}\to\mathbb{R}$ as

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

By Proposition 3.2(ii), $0\in\partial_{sup}f(0)$. Moreover, for any $x\in[-1,1]$, one has $\sup_{0\leq t\leq 1} f(tx) = f(0)$, and consequently, if $a\in\partial_{sup}f(0)$, then $ax\leq 0$ for all $x\in[-1,1]$, which yields $a=0$. Hence, $\partial_{sup}f(0)=\{0\}$. One can also check that $\partial_{sup}(-f)(\pm 1)=\{0\}$, while at all other points the sup-subdifferential of both f and $-f$ is not equal to $\{0\}$. By Corollary 3.1, the points ± 1 and 0 are the only candidates for the function f to attain its local minima and maxima, respectively (Fig. 1 shows that this is actually the case). Note that f is not convex and fails to be continuous at zero (although it is upper semi-continuous at 0).

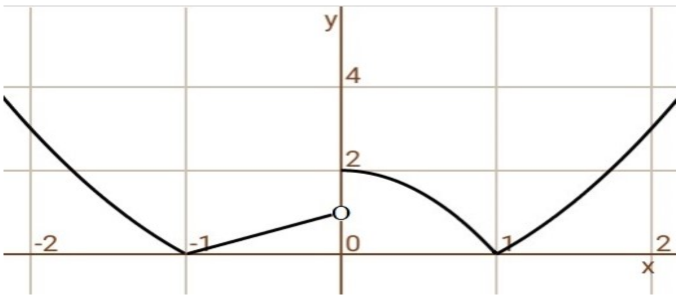

Fig. 1 The graph of f (Example 3.4)

Example 3.5 Define $f:\mathbb{R}\to\mathbb{R}$ as

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

By Proposition 3.2(ii), $0\in\partial_{sup}f(0)$. Moreover, for any $x\in\mathbb{R}$, 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, if $a\in\partial_{sup}(-f)(0)$, then $ax\leq 1$ for all $x>0$ and $ax\leq 0$ for all $x<0$, which yields $a=0$. Hence, $\partial_{sup}(-f)(0)=\{0\}$. One can also check that at all other points the sup-subdifferential of both f and $-f$ is not equal to $\{0\}$. By Corollary 3.1, the point 0 is the only candidate for the function f to attain its local minimum (see Fig. 2 below). However, 0 fails to be a minimizer of f.

The next example shows that for a convex function of a single real variable the sup*-subdifferential reduces to the conventional one.

Example 3.6 Let $f:\mathbb{R}\to\mathbb{R}_\infty$ be convex and $\bar{x}\in\operatorname{dom} f$. Then $\partial^\star_{sup}f(\bar{x})=\partial f(\bar{x})$. Indeed, $\partial f(\bar{x})\subset\partial^\star_{sup}f(\bar{x})$ by Proposition 3.2(iii). Note that the set $\mathscr{E}:=\{1\}$ is norm-generating in $\mathbb{R}$, and $\tau_1(x)=1$ if $x\neq 0$. Hence, $f(\bar{x}+\tau_1(x)x)=f(\bar{x}+x)$ for all $x\in\mathbb{R}$, and consequently, $\partial^\star_{sup}f(\bar{x})\subset\partial^{\mathscr{E}}_{sup}f(\bar{x})\subset\partial f(\bar{x})$.

As a byproduct of Example 3.6, we see that the sup*-subdifferential can be empty at some points. Recall that, in view of Proposition 3.2(ii), the sup-subdifferential is always nonempty.

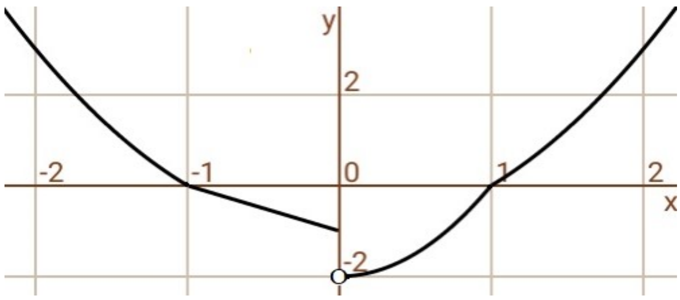

Fig. 2 The graph of f (Example 3.5)

3.2 Sup- and Sup*-Subdifferentials of Upper Semi-Continuous Functions

In this section, we derive relationships between the sup-subdifferential, sup*-subdifferential and ∂_C subdifferential of an upper semi-continuous function. Recall that a function $f : X \to \mathbb{R}_\infty$ is upper semi-continuous at $\bar{x}$ if

$$\limsup_{x\to\bar{x}} f(x) \le f(\bar{x}).$$

A function is said to be upper semi-continuous if it is upper semi-continuous at every $x \in X$. We begin with the following result about upper semi-continuous functions.

Proposition 3.3 *Let f be upper semi-continuous, $\bar{x} \in \operatorname{dom} f$ and $\mathscr{E} \in \mathscr{F}$. Then the function $\tau_{\mathscr{E}} : X \to [0,1]$, defined for all $x \in X$ by*

$$\tau_{\mathscr{E}}(x) := \min\left\{\tau_{x^*}(x) : \ f(\bar{x}+\tau_{x^*}(x)x) = \sup_{u^*\in\mathscr{E}} f(\bar{x}+\tau_{u^*}(x)x),\ x^* \in \mathscr{E}\right\}, \qquad (6)$$

is well-defined. If, furthermore, f is convex, then

$$\sup_{u^*\in\mathscr{E}} f(\bar{x}+\tau_{u^*}(x)x) - f(\bar{x}) \le \tau_{\mathscr{E}}(x)(f(\bar{x}+x)-f(\bar{x})) \quad \text{for all} \quad x \in X.$$

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

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

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

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

The point $x^* \in \mathscr{E}$ 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 (6) is well-defined. Let f be convex. Since $\tau_{\mathscr{E}}(x) \in [0,1]$, we have

$$\begin{aligned}\sup_{u^*\in\mathscr{E}} f(\bar{x}+\tau_{u^*}(x)x) - f(\bar{x}) &= f(\bar{x}+\tau_{\mathscr{E}}(x)x) - f(\bar{x}) \\ &\le \tau_{\mathscr{E}}(x)(f(\bar{x}+x)-f(\bar{x})).\end{aligned}$$

This completes the proof. ☐

With $\mathscr{E} := \mathbb{B}_{X^*}$, Proposition 3.3 yields the following corollary.

Corollary 3.2 *Let f be upper semi-continuous and $\bar{x} \in \operatorname{dom} f$. Then the function $\tau : X \to [0,1]$:*

$$\tau(x) := \min\left\{\lambda \in [0,1] : f(\bar{x}+\lambda x) = \sup_{0\le t\le 1} f(\bar{x}+tx)\right\}, \quad x \in X \tag{7}$$

is well-defined. If, furthermore, f is convex, then

$$\sup_{0\le t\le 1} f(\bar{x}+tx) - f(\bar{x}) \le \tau(x)(f(\bar{x}+x) - f(\bar{x})) \quad \textit{for all} \quad x \in X.$$

Remark 3.1 Comparing (6) and (7), one can notice that, under the conditions of Proposition 3.3, it holds $0 \le \tau(x) \le \tau_{\mathscr{E}}(x) \le 1$ for all $\mathscr{E} \in \mathscr{F}$ and $x \in X$.

Recall that, for a nonempty subset $A \subset X$, the negative polar cone to A is defined as

$$A^\circ := \{x^* \in X^* : \langle x^*, x\rangle \le 0 \quad \forall x \in A\}.$$

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

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

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

Theorem 3.1 *Let f be upper semi-continuous and $\bar{x} \in \operatorname{dom} f$. Then*

$$\partial^{\mathscr{E}}_{sup} f(\bar{x}) = \bigcap_{0<\lambda\le 1} \left\{\lambda \partial_{C^{\mathscr{E}}_\lambda} f(\bar{x})\right\} \bigcap \left(\tau^{-1}_{\mathscr{E}}(0)\right)^\circ \quad \textit{for all } \mathscr{E} \in \mathscr{F}, \tag{8}$$

$$\partial^{\star}_{sup} f(\bar{x}) = \bigcap_{\mathscr{E}\in\mathscr{F},\, 0<\lambda\le 1} \left\{\lambda \partial_{C^{\mathscr{E}}_\lambda} f(\bar{x})\right\} \bigcap \left(\tau^{-1}_{\mathscr{E}}(0)\right)^\circ,$$

$$\partial_{sup} f(\bar{x}) = \bigcap_{0<\lambda\le 1} \{\lambda \partial_{C_\lambda} f(\bar{x})\} \bigcap \left(\tau^{-1}(0)\right)^\circ, \tag{9}$$

where $C^{\mathscr{E}}_\lambda := \lambda\left(\tau^{-1}_{\mathscr{E}}(\lambda) \setminus \tau^{-1}_{\mathscr{E}}(0)\right)$, $C_\lambda := \lambda\left(\tau^{-1}(\lambda) \setminus \tau^{-1}(0)\right)$, and the functions $\tau_{\mathscr{E}}$ and τ are defined by (6) and (7), respectively.

Proof Let $\mathscr{E} \in \mathscr{F}$. By definition (3) and Proposition 3.3, we have

$$\partial^{\mathscr{E}}_{sup} f(\bar{x}) = \{x^* \in X^* : \langle x^*, x\rangle \le f(\bar{x} + \tau_{\mathscr{E}}(x)x) - f(\bar{x}) \quad \forall x \in X\}.$$

One can easily check that $\partial^{\mathscr{E}}_{sup} f(\bar{x}) = B_{\mathscr{E}} \cap \left(\tau^{-1}_{\mathscr{E}}(0)\right)^\circ$, where

$$B_{\mathscr{E}} := \left\{x^* \in X^* : \langle x^*, x\rangle \le f(\bar{x} + \tau_{\mathscr{E}}(x)x) - f(\bar{x}) \quad \forall x \in X \setminus \{\tau^{-1}_{\mathscr{E}}(0)\}\right\}.$$

Next we check that

$$\begin{aligned} B_{\mathscr{E}} &= \bigcap_{0<\lambda\le 1} \left\{x^* \in X^* : \langle x^*, x\rangle \le f(\bar{x}+\lambda x) - f(\bar{x}) \quad \forall x \in \tau^{-1}_{\mathscr{E}}(\lambda) \setminus \tau^{-1}_{\mathscr{E}}(0)\right\} \\ &= \bigcap_{0<\lambda\le 1} \left\{\lambda x^* \in X^* : \langle x^*, x\rangle \le f(\bar{x}+x) - f(\bar{x}) \quad \forall x \in \lambda\left(\tau^{-1}_{\mathscr{E}}(\lambda) \setminus \tau^{-1}_{\mathscr{E}}(0)\right)\right\} \\ &= \bigcap_{0<\lambda\le 1} \left\{\lambda \partial_{C^{\mathscr{E}}_\lambda} f(\bar{x})\right\}. \end{aligned}$$

This proves (8). The other two representations are consequences of (8). $\square$

Set

$$\begin{aligned} L_f^{>}(\bar{x}) &:= \{x\in X : f(\bar{x}+x) > f(\bar{x})\};\\ L_f^{<}(\bar{x}) &:= \{x\in X : f(\bar{x}+x) < f(\bar{x})\};\\ L_f^{=}(\bar{x}) &:= \{x\in X : f(\bar{x}+x) = f(\bar{x})\};\\ L_f^{\leq}(\bar{x}) &:= \{x\in X : f(\bar{x}+x) \leq f(\bar{x})\}. \end{aligned}$$

The following proposition provides explicit representations of the functions $\tau_{\mathscr{E}}$ and τ defined by (6) and (7) for an upper semi-continuous convex function.

Proposition 3.4 *Let f be convex upper semi-continuous, $\mathscr{E}\in\mathscr{F}$ and $\bar{x}\in \operatorname{dom} f$. Then $\tau_{\mathscr{E}}(x)=\tau(x)=1$ for all $x\in L_f^{>}(\bar{x})$, and $\tau(x)=0$ for all $x\in L_f^{\leq}(\bar{x})$.*

If $0\in\mathscr{E}$, then $\tau_{\mathscr{E}}(x)=0$ for all $x\in L_f^{\leq}(\bar{x})$.

Proof By Proposition 3.3, we have

$$f(\bar{x}+tx)-f(\bar{x}) \leq \tau_{\mathscr{E}}(x)(f(\bar{x}+x)-f(\bar{x})) \tag{10}$$

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

If $x\in L_f^{>}(\bar{x})$, then, by letting $t=1$ in (10), we obtain $\tau_{\mathscr{E}}(x)\geq 1$, and therefore, $\tau_{\mathscr{E}}(x)=1$; in particular, $\tau(x)=1$. Let $0\in\mathscr{E}$. If $x\in L_f^{<}(\bar{x})$, then, by letting $t=0$ in (10), we get $\tau_{\mathscr{E}}(x)\leq 0$, and therefore, $\tau_{\mathscr{E}}(x)=0$; in particular, $\tau(x)=0$. If $x\in L_f^{=}(\bar{x})$, then, for all $t\in[0,1]$, we have

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

and consequently, $\max_{t\in T_{\mathscr{E}}} f(\bar{x}+tx)$ is attained at $t=0$. It follows from definition (6) that $\tau_{\mathscr{E}}(x)=0$; in particular, $\tau(x)=0$. □

Using Proposition 3.4, we can simplify the conclusions of Theorem 3.1 for upper semi-continuous convex functions.

Corollary 3.3 *Let f be convex upper semi-continuous, $0\in\mathscr{E}\in\mathscr{F}$ and $\bar{x}\in\operatorname{dom} f$. Then*

$$\partial^{\mathscr{E}}_{sup} f(\bar{x}) = \partial_{sup} f(\bar{x}) = \partial_{L_f^{>}(\bar{x})} f(\bar{x}) \cap N_{L_f^{\leq}(\bar{x})}(0). \tag{11}$$

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 Proposition 3.4, for all $0<\lambda<1$, we have $\tau_{\mathscr{E}}^{-1}(\lambda)=\tau^{-1}(\lambda)=\emptyset$, and consequently, using the notations in Theorem 3.1, $\partial_{C_\lambda^{\mathscr{E}}} f(\bar{x}) = \partial_{C_\lambda} f(\bar{x}) = X^*$. We also have $C_1^{\mathscr{E}}=C_1=L_f^{>}(\bar{x})$ and $\tau_{\mathscr{E}}^{-1}(0)=\tau^{-1}(0)=L_f^{\leq}(\bar{x})$. Hence, representations (8) and (9) reduce to (11). The last assertion follows thanks to Proposition 3.2(iv). □

Corollary 3.4 *Let $X:=\ell^p(\mathbb{N})$ with $p\geq 1$, f be convex upper semi-continuous, and $\bar{x}\in\operatorname{dom} f$. Then $\partial^{\star}_{sup} f(\bar{x}) = \partial_{sup} f(\bar{x})$. As a consequence, $\partial^{\star}_{sup} f(\bar{x}) = \partial f(\bar{x})$ if and only if $\bar{x}$ minimizes f.*

Proof Let (e_k) denote the canonical basis of $\ell^p(\mathbb{N})$ and $\mathscr{E}\in\mathscr{F}$. In view of Example 3.2, $e_k\in\mathscr{E}$, 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 $\mathscr{E}$ is weak*-closed, $0\in\mathscr{E}$. The assertion follows from definition (5) and Corollary 3.3. □

Remark 3.2 In general real topological vector spaces, the equality $\partial^{\star}_{sup} f(\bar{x}) = \partial f(\bar{x})$ can hold even if $\bar{x}$ does not minimize f; cf. Example 3.6.

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

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

is nondecreasing, and the function

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

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

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

4 The Symmetric Subdifferential

Let X be a linear topological space and $\bar{x} \in \operatorname{dom} f$. The symmetric subdifferential of f at $\bar{x}$ is defined as

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

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 $\bar{x} \in \operatorname{int} \operatorname{dom} f$. Indeed,

$$f'_{sym}(\bar{x};d) = \max\left\{f'(\bar{x};d), f'(\bar{x};-d)\right\}, \tag{12}$$

where $f'(\bar{x};d)$ denotes the conventional directional derivative of f at $\bar{x}$ in direction $d \in X$. Note that, if the limit

$$\lim_{t \to 0} \frac{f(\bar{x}+td) - f(\bar{x})}{t}$$

exists, then

$$f'_{sym}(\bar{x};d) = |f'(\bar{x};d)|. \tag{13}$$

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

$$f'(\bar{x};d) = \max\{\langle x^*, d \rangle : \quad x^* \in \partial f(\bar{x})\}. \tag{14}$$

The following proposition, which is a direct consequence of (12) and (14), states a similar property for the symmetric directional derivative (see [13, formula (4)]).

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

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

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

Proposition 4.2 *Suppose that f is convex and $\bar{x} \in \operatorname{dom} f$. If $\partial_{sym} f(\bar{x}) = \partial f(\bar{x}) \neq \emptyset$, 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 and $\bar{x} \in \operatorname{dom} f$. If there exists $\bar{d} \in X$ such that*

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

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

Proof One can easily check that

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

for all $\alpha \in \mathbb{R}$ and $d \in X$. The function $d \mapsto f'_{sym}(\bar{x};d)$ is sub-additive. Indeed,

$$\begin{aligned} f'_{sym}(\bar{x};d_1+d_2) &= \max\left\{f'(\bar{x};d_1+d_2), f'(\bar{x};-d_1-d_2)\right\} \\ &\leq \max\left\{f'(\bar{x};d_1)+f'(\bar{x};d_2), f'(\bar{x};-d_1)+f'(\bar{x};-d_2)\right\} \\ &\leq \max\left\{f'(\bar{x};d_1), f'(\bar{x};-d_1)\right\} + \max\left\{f'(\bar{x};d_2), f'(\bar{x};-d_2)\right\} \\ &= f'_{sym}(\bar{x};d_1) + f'_{sym}(\bar{x};d_2) \quad \text{for all } d_1, d_2 \in X. \end{aligned}$$

Thus, the function $d \mapsto f'_{sym}(\bar{x};d)$ 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 $\langle l^*, h\rangle := \alpha f'_{sym}(\bar{x};\bar{d})$, where $h = \alpha\bar{d}$. Note that l^* is well-defined since $\bar{d} \neq 0$ and $f'_{sym}(\bar{x};\bar{d})$ is finite. It follows that, for all $h \in H$,

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

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

Example 4.1 We consider the function $f : \ell^2(\mathbb{N}) \to \mathbb{R}$ in Example 2.1. Let $\bar{x} = (\bar{x}_1, \bar{x}_2, \ldots) \in C$ and $\bar{d} = e_k$, the kth 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, $\partial_{sym} f(\bar{x})$ contains a nonzero element.

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

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

It is discontinuous (though upper semi-continuous) at 0 and fails to be convex. One can easily check that $\partial_{sym} f(0) = [-1, 1]$. Indeed,

$$\max\left\{f'(0;d), f'(0;-d)\right\} = |d| \quad \text{for all } d \in X.$$

5 Calculus Rules in Banach Spaces

In this section, we establish certain calculus rules for ∂_{sup}, $\partial^{\star}_{sup}$ and ∂_{sym} for proper continuous convex functions on Banach spaces.

The next three rules are immediate from the definitions (as long as $\bar{x} \in \operatorname{dom} f$):

$$\partial_{sup}(\lambda f)(\bar{x}) = \lambda \partial_{sup} f(\bar{x}), \;\; \partial^{\star}_{sup}(\lambda f)(\bar{x}) = \lambda \partial^{\star}_{sup} f(\bar{x}), \;\; \partial_{sym}(\lambda f)(\bar{x}) = \lambda \partial_{sym} f(\bar{x})$$

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

5.1 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_δ subset, or equivalently, if the dual of each its separable subspace is separable [20, 29].

Lemma 5.1 *Let $f : X \to \mathbb{R}_\infty$ 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*

$$\partial_{sym} f(\bar{x}) = cl^{w^*} \operatorname{co} \left(\partial f(\bar{x}) \cup (-\partial f(\bar{x}))\right), \tag{15}$$

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

If X is Asplund, then cl^{w^*} in (15) can be replaced by the closure with respect to the norm topology.

Proof By assumptions, $\partial_{sym} f(\bar{x})$ is nonempty and weak*-compact. Hence, by the Krein–Milman theorem [7], $\partial_{sym} f(\bar{x})$ contains extreme points. Moreover,

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

where $\operatorname{ext} \partial_{sym} f(\bar{x})$ denotes the set of all extreme points of $\partial_{sym} f(\bar{x})$. By Proposition 4.1,

$$\operatorname{ext} \partial_{sym} f(\bar{x}) \subset \partial f(\bar{x}) \cup (-\partial f(\bar{x})).$$

Hence,

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

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,

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

This proves (15).

If X is Asplund, its dual $X^\star$ has the Radon–Nikodým property [20], and it follows from the Edgar–Lindenstrauss theorem [19, 27] that the weak*-closure can be replaced by the norm closure. □

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 [20, Theorem 5.7], [29, Theorem 6].

Theorem 5.1 *Let $A : X \to Y$ be a bounded linear map between Banach spaces, $f : X \to \mathbb{R}_\infty$ and $g : Y \to \mathbb{R}_\infty$ be proper convex functions such that f and $g \circ A$ are finite and continuous at $\bar{x}$. Suppose that $0 \in \operatorname{core}(\operatorname{dom} g - A \operatorname{dom} f)$, 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. Then the following assertions hold true.*

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

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

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 $A^*: Y^* \to X^*$ 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 [5], and Lemma 5.1, we have

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

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

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

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

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

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

(ii) The proof goes along the same lines as that of (i). 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

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

One can easily check that

$$f'_{sup}(\bar{x};d)=\max\left\{\langle x^*,d\rangle:\quad x^*\in\partial f(\bar{x})\cup\{0\}\right\}.$$

By the Krein–Milman theorem, $\partial_{sup}f(\bar{x})=cl^{w^*}\,\mathrm{co}\,(\partial f(\bar{x})\cup\{0\})$. □

5.2 Sup*-Subdifferential Sum Rule in $\ell^p(\mathbb{N})$

The next statement is a straightforward consequence of Theorem 5.1 and Corollary 3.4.

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

$$\partial^\star_{sup}(f+g)(\bar{x})\subset\partial^\star_{sup}f(\bar{x})+\partial^\star_{sup}g(\bar{x}).$$

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

$$\partial^\star_{sup}(f+g)(\bar{x})=\partial^\star_{sup}f(\bar{x})+\partial^\star_{sup}g(\bar{x}).$$

6 Acknowledgments

The research was supported by the Australian Research Council, project DP160100854 and the FMJH program PGMO.

References

1. Adly, S.: A Variational Approach to Nonsmooth Dynamics. SpringerBriefs in Mathematics. Springer, Cham (2017). DOI 10.1007/978-3-319-68658-5
2. Aragón, F.J., Goberna, M., López, M.A., Rodríguez, M.L.: Nonlinear Optimization. Springer Undergraduate Texts in Mathematics and Technology. Springer, Cham (2019). DOI 10.1007/978-3-030-11184-7
3. Aragón Artacho, F.J., Borwein, J.M., Martín-Márquez, V., Yao, L.: Applications of convex analysis within mathematics. Math. Program., Ser. B **148**(1-2), 49–88 (2014). DOI 10.1007/s10107-013-0707-3
4. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, second edn. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, Cham (2017). DOI 10.1007/978-3-319-48311-5
5. 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)
6. Borwein, J.M., Vanderwerff, J.: Convex Functions: Constructions, Characterizations and Counterexamples, *Encyclopedia of Mathematics and its Applications*, vol. 109. Cambridge University Press, Cambridge (2010). DOI 10.1017/CBO9781139087322
7. Bourbaki, N.: Espaces Vectoriels Topologiques. Chapitres 1 à 5, new edn. Masson, Paris (1981). Éléments de mathématique. [Elements of mathematics]
8. 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)
9. Burachik, R.S., Iusem, A.N., Svaiter, B.F.: Enlargement of monotone operators with applications to variational inequalities. Set-Valued Anal. **5**(2), 159–180 (1997). DOI 10.1023/A:1008615624787
10. Burachik, R.S., Martínez-Legaz, J., Rezaie, M., Théra, M.: Erratum to: An additive subfamily of enlargements of a maximally monotone operator [MR3428692]. Set-Valued Var. Anal. **23**(4), 667–668 (2015). DOI 10.1007/s11228-015-0344-5
11. Giles, J.R.: Convex Analysis with Application in the Differentiation of Convex Functions, *Research Notes in Mathematics*, vol. 58. Pitman (Advanced Publishing Program), Boston, Mass.-London (1982)
12. Ioffe, A.D., Tihomirov, V.M.: Theory of Extremal Problems, *Studies in Mathematics and its Applications*, vol. 6. North-Holland Publishing Co., Amsterdam-New York (1979)
13. 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)
14. Lemaréchal, C.: An introduction to the theory of nonsmooth optimization. Optimization **17**(6), 827–858 (1986). DOI 10.1080/02331938608843204
15. Lucchetti, R.: Convexity and Well-Posed Problems, *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC*, vol. 22. Springer, New York (2006). DOI 10.1007/0-387-31082-7
16. Martinez-Legaz, J., Théra, M.: ε-subdifferentials in terms of subdifferentials. Set-Valued Anal. **4**(4), 327–332 (1996). DOI 10.1007/BF00436109
17. Mordukhovich, B.S.: Variational Analysis and Applications. Springer Monographs in Mathematics. Springer, Cham (2018). DOI 10.1007/978-3-319-92775-6
18. Peypouquet, J.: Convex Optimization in Normed Spaces. SpringerBriefs in Optimization. Springer, Cham (2015). DOI 10.1007/978-3-319-13710-0
19. Phelps, R.R.: Dentability and extreme points in Banach spaces. J. Functional Analysis **17**, 78–90 (1974). DOI 10.1016/0022-1236(74)90005-6
20. Phelps, R.R.: Convex Functions, Monotone Operators and Differentiability, *Lecture Notes in Mathematics*, vol. 1364, 2nd edn. Springer-Verlag, Berlin (1993)
21. Pshenichnyi, B.N.: Necessary Conditions for an Extremum, *Pure and Applied Mathematics*, vol. 4. Marcel Dekker, Inc., New York (1971)
22. Rainwater, J.: Yet more on the differentiability of convex functions. Proc. Amer. Math. Soc. **103**(3), 773–778 (1988). DOI 10.2307/2046851
23. Revalski, J.P., Théra, M.: Enlargements and sums of monotone operators. Nonlinear Anal., Ser. A: Theory Methods **48**(4), 505–519 (2002). DOI 10.1016/S0362-546X(00)00201-7
24. Riesz, F., Sz.-Nagy, B.: Functional Analysis. Dover Books on Advanced Mathematics. Dover Publications, Inc., New York (1990)
25. Rockafellar, R.T.: Convex Analysis. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ (1997)
26. Rudin, W.: Real and Complex Analysis, third edn. McGraw-Hill Book Co., New York (1987)
27. Stegall, C.: The Radon-Nikodym property in conjugate Banach spaces. Trans. Amer. Math. Soc. **206**, 213–223 (1975). DOI 10.2307/1997154

28. Théra, M.: Subdifferential calculus for convex operators. J. Math. Anal. Appl. **80**(1), 78–91 (1981). DOI 10.1016/0022-247X(81)90093-7
29. Yost, D.: Asplund spaces for beginners. Acta Univ. Carolin. Math. Phys. **34**(2), 159–177 (1993)
30. Zălinescu, C.: Convex Analysis in General Vector Spaces. World Scientific Publishing Co., Inc., River Edge, NJ (2002). DOI 10.1142/9789812777096