Vietnam Journal of Mathematics (2020) 48:569–588
https://doi.org/10.1007/s10013-020-00403-5

ORIGINAL ARTICLE

Subdifferential of the Supremum via Compactification of the Index Set

R. Correa[1,2] · A. Hantoute[3,4] · M. A. López[4,5]

Received: 21 September 2019 / Accepted: 30 January 2020 / Published online: 30 April 2020

Abstract

We give new characterizations for the subdifferential of the supremum of an arbitrary family of convex functions, dropping out the standard assumptions of compactness of the index set and upper semi-continuity of the functions with respect to the index (J. Convex Anal. 26, 299–324, 2019). We develop an approach based on the compactification of the index set, giving rise to an appropriate enlargement of the original family. Moreover, in contrast to the previous results in the literature, our characterizations are formulated exclusively in terms of exact subdifferentials at the nominal point. Fritz–John and KKT conditions are derived for convex semi-infinite programming.

Keywords Supremum of convex functions · Subdifferentials · Stone–Čech compactification · Convex semi-infinite programming · Optimality conditions

Mathematics Subject Classification (2010) 46N10 · 52A41 · 90C25

Dedicated by his coauthors to Prof. Marco A. López on his 70th birthday.

✉ A. Hantoute
ahantoute@dim.uchile.cl

R. Correa
rcorrea@dim.uchile.cl

M. A. López
marco.antonio@ua.es

1 Universidad de O'Higgins, Rancagua, Chile
2 DIM-CMM of Universidad de Chile, Santiago, Chile
3 Center for Mathematical Modeling (CMM), Universidad de Chile, Santiago, Chile
4 Universidad de Alicante, Alicante, Spain
5 CIAO, Federation University, Ballarat, Australia

1 Introduction

If we consider the pointwise supremum $f := \sup_{t \in T} f_t$ of a collection of convex functions $f_t : X \to \mathbb{R} \cup \{\pm\infty\}$, $t \in T \neq \emptyset$, T arbitrary, defined on a separated locally convex space X, a challenging problem along the recent history of optimization (specially, in the decades of the 60s and 70s of the 20th century) has been to obtain formulas for the subdifferential of the supremum, $\partial f(x)$, at any point x of the effective domain of f, in terms of the subdifferentials of the data functions, $\partial f_t(x),\ t \in T$.

Since many convex functions, such as the Fenchel conjugate, the sum, the composition with affine applications, etc., can be expressed as the supremum of affine or convex functions, formulas characterizing the subdifferential of the supremum were expected to play a crucial role in convex and variational analysis, leading to a variety of calculus rules and allowing a deeper analysis for some relevant problems in this area. For instance, any formula for the subdifferential of the supremum function can be seen as a useful tool in deriving KKT-type optimality conditions for a convex optimization problem. This is due to the fact that any set of convex constraints, even an infinite set, can be replaced by a unique convex constraint involving the supremum function. An alternative approach consists of replacing the constraints by the indicator function of the feasible set. It turns out that, under certain constraint qualifications, its subdifferential (i.e., the normal cone to the feasible set) appears in the so-called Fermat optimality principle, and its relation with the subdifferential of the supremum function can be then conveniently exploited.

Let us quote the following paragraph extracted from [15]: "One of the most specific constructions in convex or nonsmooth analysis is certainly taking the supremum of a (possibly infinite) collection of functions. In the years 1965–1970, various calculus rules concerning the subdifferential of sup-functions started to emerge; working in that direction and using various assumptions, several authors contributed to this calculus rule: B.N. Pshenichnyi, A.D. Ioffe, V.L. Levin, R.T. Rockafellar, A. Sotskov, etc.; however, the most elaborated results of that time were due to M. Valadier (1969); he made use of ε-active indices in taking the supremum of the collection of functions."

Therefore, it is clear that the mathematical interest of this topic was widely recognized since the very beginning of the convex and variational analysis history. A sample of remarkable contributions to this topic are: Brøndsted [1], Ekeland and Temam [10], Ioffe [17], Ioffe and Levin [18], Ioffe and Tikhomirov [19], Levin [20], Pschenichnyi [23], Rockafellar [26], Valadier [30], etc. See, for instance, Tikhomirov [29] to trace out the historical origins of the issue.

In a series of papers ([3–6, 12–14], etc.) we provided alternative characterizations of the subdifferential supremum in various settings, and applied them to derive calculus rules in convex analysis.

In [7] we addressed the problem of characterizing the subdifferential of the supremum of a compactly-indexed family of extended real-valued convex functions. These assumptions, which are standard in the literature of convex analysis and non-differentiable semi-infinite programming, are the compactness of the index set T and the upper semi-continuity of the constraint functions with respect to the index t. A couple of questions arise in a natural way. The first basic one is the following: Is it possible to remove these assumptions? A second more precise question is: By using a compactification of the index set and an appropriate enlargement of the original family of data functions, is there any chance for getting rid of these assumptions, but keeping alive the possibility of still applying the theory developed under them?

In this framework, we propose in the current paper an approach based on the Stone–Čech compactification of the index set T, as well as a natural procedure for building an appropriate enlargement of the original family ensuring the fulfillment of the minimal requirements of continuity of the functions with respect to the index. Moreover, in contrast to previous approaches, our characterizations are formulated exclusively in terms of exact subdifferentials at the nominal point.

Formula (10) constitutes the main result of the paper. It provides an explicit expression of the subdifferential of the supremum function for any family of convex functions, dropping the usual standard assumptions in the literature (upper semi-continuity and compactness conditions; see, e.g. [1, 7, 19, 27, 30]). Namely, compared with the formula

$$\partial f(x) = \bigcap\nolimits_{L\in\mathcal{F}(x),\varepsilon>0} \overline{\mathrm{co}} \left\{\bigcup\nolimits_{t\in T_\varepsilon(x)} \partial_\varepsilon (f_t + \mathrm{I}_{L\cap \mathrm{dom}\, f})(x)\right\}$$

(see (2) and (3) for the definition of $\mathcal{F}(x)$ and $T_\varepsilon(x)$, respectively), which can be easily derived from the main result in [14, Theorem 4], formula (10) involves the convex hull of the union of the exact subdifferentials of exclusively the active functions, up to an appropriate enlargement of the original family of functions.

The paper is structured as follows. After a short section introducing the notation, the main result in the section devoted to preliminaries is formula (4) in Proposition 1, which slightly improves Proposition 2 in [6] as it uses the convex hull instead of the closed convex hull. In Section 3 the compactification process is described in detail, and an appropriate enlargement of the original family $\{f_t, t \in T\}$ is built, through formula (6), in order to guarantee the (upper-semi) continuity requirements with respect to the index t which allow to apply the results in [7]. Our main result in Section 3, Theorem 1, provides the aimed characterization of the subdifferential of f in non-compact frameworks. It comes after some needed technical lemmas, and some corollaries are also established under certain specific assumptions. An example illustrates the compactification approach, and the last section provides Fritz–John and KKT-type optimality conditions for the convex semi-infinite optimization problem such that the compact/continuity assumptions in [6, Theorem 5 and Corollary 6] are again dropped.

2 Notation

Let X be a (real) separated locally convex space, whose topological dual space is X^*, which is endowed with the w^*-topology. The spaces X and X^* are paired in duality by the bilinear form $(x^*, x) \in X^* \times X \mapsto \langle x^*, x\rangle := \langle x, x^*\rangle := x^*(x)$. The zero vectors in X and X^* are denoted by θ. Closed, convex and balanced neighborhoods of θ are called θ-neighborhoods. We use the notation $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ and $\mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$, and adopt the convention $(+\infty) + (-\infty) = (-\infty) + (+\infty) = +\infty$.

Given two nonempty sets A and B in X (or in X^*), we define the *algebraic* (or *Minkowski*) *sum* by

$$A + B := \{a + b : a \in A, b \in B\}, \quad A + \emptyset = \emptyset + A = \emptyset.$$

By $\mathrm{co}(A)$, $\mathrm{cone}(A)$, and $\mathrm{aff}(A)$, we denote the *convex*, the *conical convex* (i.e., $\mathrm{cone}\, A := \mathbb{R}_+(\mathrm{co}\, A)$), and the *affine hulls* of the set A, respectively. Moreover, $\mathrm{int}(A)$ is the *interior* of A and $\mathrm{cl}\, A$ and $\overline{A}$ are indistinctly used for denoting the *closure* of A. We use $\mathrm{ri}(A)$ to denote the (topological) *relative interior* of A (i.e., the interior of A in the topology relative to $\mathrm{aff}(A)$ if $\mathrm{aff}(A)$ is closed, and the empty set otherwise).

Associated with $A \neq \emptyset$ we consider the *orthogonal subspace* given by

$$A^{\perp} := \{x^* \in X^* : \langle x^*, x\rangle = 0 \text{ for all } x \in A\}.$$

The following relation is fulfilled

$$\bigcap\nolimits_{L \in \mathcal{F}} (A + L^{\perp}) \subset \operatorname{cl} A, \tag{1}$$

where $\mathcal{F}$ is the family of finite-dimensional linear subspaces of X.

If $A \subset X$ is convex and $x \in X$, we define the *normal cone* to A at x as

$$\mathrm{N}_A(x) := \{x^* \in X^* : \langle x^*, y - x\rangle \leq 0 \text{ for all } y \in A\},$$

if $x \in A$, and the empty set otherwise.

Given a function $f : X \longrightarrow \overline{\mathbb{R}}$, its *(effective) domain* is

$$\operatorname{dom} f := \{x \in X : f(x) < +\infty\}.$$

We say that f is *proper* when $\operatorname{dom} f \neq \emptyset$ and $f(x) > -\infty$ for all $x \in X$.

Given $x \in X$ and $\varepsilon \geq 0$, the ε-*subdifferential* of f at x is

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

when $x \in \operatorname{dom} f$, and $\partial_\varepsilon f(x) := \emptyset$ when $f(x) \notin \mathbb{R}$. The elements of $\partial_\varepsilon f(x)$ are called ε-*subgradients* of f at x. The *subdifferential* of f at x is $\partial f(x) := \partial_0 f(x)$, whose elements are called *subgradients* of f at x.

The support and the indicator functions of $A \subset X$ are respectively defined as

$$\sigma_A(x^*) := \sup\{\langle x^*, x\rangle : x \in A\} \quad \text{for } x^* \in X^*,$$

and

$$\mathrm{I}_A(x) := \begin{cases} 0 & \text{if } x \in A, \\ +\infty & \text{if } x \in X \setminus A. \end{cases}$$

3 Preliminary Results

We give a first characterization of the subdifferential of the supremum

$$f := \sup_{t \in T} f_t,$$

of a family of extended real-valued convex functions $\{f_t,\ t \in T\}$, defined on a (separated) real locally convex space X, and indexed by an arbitrary (possibly, infinite) set T.

We shall need the following result which slightly improves Proposition 2 in [6], as it uses the convex hull instead of the closed convex hull. Our main result, given in Theorem 1, provides the general characterization of the subdifferential of f in non-necessarily compact frameworks.

Given $x \in X$ and $\varepsilon \geq 0$, we shall denote

$$\mathcal{F}(x) := \{L \text{ is a finite-dimensional linear subspace of } X \text{ containing } x\}, \tag{2}$$

$$T_\varepsilon(x) := \{t \in T : f_t(x) \geq f(x) - \varepsilon\} \quad \text{and} \quad T(x) := T_0(x). \tag{3}$$

Proposition 1 *Fix $x \in X$. We assume there is some $\varepsilon_0 > 0$ such that (i) $T_{\varepsilon_0}(x)$ is compact and (ii) for each net $(t_i)_i \subset T_{\varepsilon_0}(x)$ converging to $t \in T_{\varepsilon_0}(x)$ we have that*

$$\limsup_i f_{t_i}(z) \leq f_t(z) \quad \textit{for all } z \in \operatorname{dom} f.$$

Then

$$\partial f(x) = \bigcap\nolimits_{L \in \mathcal{F}(x)} \operatorname{co}\left\{\bigcup\nolimits_{t \in T(x)} \partial(f_t + \mathrm{I}_{L \cap \operatorname{dom} f})(x)\right\}. \tag{4}$$

Proof According to [6, Proposition 2], where (4) was established with $\overline{\operatorname{co}}$ instead of co, we only need to prove that the sets

$$E_L := \operatorname{co}\left\{\bigcup\nolimits_{t \in T(x)} \partial(f_t + \mathrm{I}_{L \cap \operatorname{dom} f})(x)\right\}, \quad L \in \mathcal{F}(x),$$

are closed under the current hypothesis. Let us denote, for $t \in T(x)$ and $L \in \mathcal{F}(x)$,

$$\tilde{g}_t := f_t + \mathrm{I}_{L \cap \operatorname{dom} f}, \; g_t := \tilde{g}_{t|_L},$$

so that

$$\operatorname{dom} \tilde{g}_t = \operatorname{dom} f_t \cap (L \cap \operatorname{dom} f) = L \cap \operatorname{dom} f.$$

Take a net $(u_i^*)_i \subset E_L$ such that $u_i^* \to u^* \in X^*$. We denote by z_i^* the restriction of u_i^* to the finite-dimensional subspace L, so that

$$(z_i^*)_i \subset \operatorname{co}\left\{\bigcup\nolimits_{t \in T(x)} \partial g_t(x)\right\} \subset L^*,$$

where L^* is the dual of L. Then, by applying Charathéodory's Theorem in L^*, for each i there are some $\lambda_{i,1}, \ldots, \lambda_{i,n+1} \geq 0$ with $\lambda_{i,1} + \cdots + \lambda_{i,n+1} = 1$, and elements $z_{i,k}^* \in \partial g_{t_{i,k}}(x)$ with $t_{i,k} \in T(x)$ and $k \in K := \{1, \ldots, n+1\}$ (hence, the functions $g_{t_{i,k}}$, $k \in K$, are all proper), such that

$$z_i^* = \lambda_{i,1} z_{i,1}^* + \ldots + \lambda_{i,n+1} z_{i,n+1}^*,$$

where n is the dimension of L.

We may assume that each $(\lambda_{i,k})_i$, $k \in K$, converges to some $\lambda_k \geq 0$ such that $\lambda_1 + \cdots + \lambda_{n+1} = 1$. Also, since $(t_{i,k})_i \subset T(x) \subset T_{\varepsilon_0}(x)$ and this last set is compact by assumption, we may assume that $t_{i,k} \to t_k \in T_{\varepsilon_0}(x)$, $k \in K$. Moreover, using again the assumption, we have

$$\limsup_i g_{t_{i,k}}(z) \leq g_{t_k}(z) \quad \text{for all } z \in L \cap \operatorname{dom} f, \; k \in K;$$

in particular, $g_{t_k}(z) > -\infty$ for all $z \in L \cap \operatorname{dom} f$, $k \in K$, and (recall that $t_{i,k} \in T(x)$)

$$f(x) = \limsup_i g_{t_{i,k}}(x) \leq g_{t_k}(x) = f_{t_k}(x) \leq f(x) \quad \text{for all } k \in K,$$

showing that $t_k \in T(x)$ for all $k \in K$. Consequently, taking into account that $(t_{i,k})_i \subset T(x)$ for all $k \in K$, for every $z \in L \cap \operatorname{dom} f$ $(= \operatorname{dom} g_{t_k}, k \in K)$ we obtain

$$\begin{aligned}
\langle z^*, z - x \rangle &= \lim_i \langle \lambda_{i,1} z_{i,1}^* + \cdots + \lambda_{i,n+1} z_{i,n+1}^*, z - x \rangle \\
&\leq \lambda_1 \limsup_i g_{t_{i,1}}(z) + \cdots + \lambda_{n+1} \limsup_i g_{t_{i,n+1}}(z) \\
&\quad + \limsup_i (-\lambda_1 g_{t_{i,1}}(x) - \cdots - \lambda_{n+1} g_{t_{i,n+1}}(x)) \\
&\leq \lambda_1 g_{t_1}(z) + \cdots + \lambda_{n+1} g_{t_{n+1}}(z) + \limsup_i (-\lambda_1 g_{t_{i,1}}(x) - \cdots - \lambda_{n+1} g_{t_{i,n+1}}(x)) \\
&= \lambda_1 g_{t_1}(z) + \cdots + \lambda_{n+1} g_{t_{n+1}}(z) - f(x) \\
&= \sum_{k \in K_+} \lambda_k g_{t_k}(z) - \sum_{k \in K_+} \lambda_k g_{t_k}(x),
\end{aligned}$$

where $K_+ := \{k \in K : \lambda_k > 0\}$. Hence, using Rockafellar's subdifferential sum rule [25], as $g_{t_k}(z) + \mathrm{I}_{L \cap \mathrm{dom}\, f}(z) = g_{t_k}(z)$ and $\mathrm{ri}(\mathrm{dom}\, g_{t_k}) = \mathrm{ri}(L \cap \mathrm{dom}\, f) \neq \emptyset$ for all $z \in L$ and all $k \in K$, we obtain

$$z^* \in \partial \left(\sum_{k \in K_+} \lambda_k g_{t_k} \right)(x) = \sum_{k \in K_+} \lambda_k \partial g_{t_k}(x).$$

Then, using the extension theorem, we can take an extension v^* of z^* to X^* such that

$$v^* \in \sum_{k \in K_+} \lambda_k \partial \tilde{g}_{t_k}(x),$$

satisfying $u^* - v^* \in L^\perp$. Therefore

$$\begin{aligned} u^* \in v^* + L^\perp \subset & \sum_{k \in K_+} \lambda_k \partial \tilde{g}_{t_k}(x) + L^\perp \\ & \subset \sum_{k \in K_+} \lambda_k \partial (\tilde{g}_{t_k} + \mathrm{I}_L)(x) = \sum_{k \in K_+} \lambda_k \partial \tilde{g}_{t_k}(x) \in E_L. \end{aligned}$$

□

4 Compactification Approach

Given a non-empty family of extended real-valued convex functions

$$f_t : X \to \overline{\mathbb{R}}, \quad t \in T,$$

defined on a (separated) real locally convex space X, and indexed by an arbitrary (possibly, infinite) set T, we consider the corresponding supremum function

$$f := \sup_{t \in T} f_t.$$

Here, in order to apply the methodology proposed in [7], we endow the index set T with some topology. When no topology is known on T we frequently use the discrete one. We denote by $\mathcal{C}(T, [0, 1])$ the set of continuous functions from T to $[0, 1]$, and consider the product space $[0, 1]^{\mathcal{C}(T,[0,1])}$, which is compact for the product topology (by Tychonoff theorem). We shall regard the index set T as a subset of $[0, 1]^{\mathcal{C}(T,[0,1])}$, and write $T \subset [0, 1]^{\mathcal{C}(T,[0,1])}$, by using the mapping $\mathfrak{d}: T \to [0, 1]^{\mathcal{C}(T,[0,1])}$, which assigns to each $t \in T$ the evaluation function $\mathfrak{d}(t) \equiv \gamma_t \in [0, 1]^{\mathcal{C}(T,[0,1])}$, defined as

$$\gamma_t(\varphi) := \varphi(t), \quad \varphi \in \mathcal{C}(T, [0, 1]).$$

The closure of T in $[0, 1]^{\mathcal{C}(T,[0,1])}$ for the product topology is the compact set

$$\widehat{T} := \mathrm{cl}(\mathfrak{d}(T)), \tag{5}$$

and is referred to as the Stone–Čech compactification of T, usually denoted by βT. Remember that for $\gamma \in \widehat{T}$ and a net $(\gamma_i)_i \subset \widehat{T}$, we have $\gamma_i \to \gamma$ when

$$\gamma_i(\varphi) \to \gamma(\varphi) \quad \text{for all } \varphi \in \mathcal{C}(T, [0, 1]).$$

When T is completely regular; i.e., compact Hausdorff, $\widehat{T}$ is Hausdorff (see, i.e., [22, §38]), and the convergences in $\mathfrak{d}(T)$ and T are the same.

Next, we enlarge the original family $\{f_t,\ t \in T\}$ by introducing the functions $f_\gamma : X \to \overline{\mathbb{R}}$, $\gamma \in \widehat{T}$, defined by

$$f_\gamma(z) := \limsup_{\gamma_t \to \gamma,\ t \in T} f_t(z); \tag{6}$$

that is,

$$f_\gamma(z) = \sup \left\{ \limsup_i f_{t_i}(z) \,\middle|\, \begin{array}{l} (t_i)_i \subset T,\ \varphi(t_i) \to \gamma(\varphi), \\ \forall \varphi \in \mathcal{C}(T, [0, 1]) \end{array} \right\}.$$

Observe that the family $\{f_\gamma, \gamma \in \widehat{T}\}$ includes the elements of the form f_{γ_t}, $t \in T$, given by

$$f_{\gamma_t}(z) = \limsup_{\gamma_s \to \gamma_t,\ s \in T} f_s(z),$$

which may not belong to the original family $\{f_t,\ t \in T\}$, as well as the functions f_γ with $\gamma \in \widehat{T} \backslash \mathfrak{d}(T)$.

Remark 1 Observe that, for all $t \in T$ and $z \in X$,

$$f_{\gamma_t}(z) \geq \limsup_{s \to t,\ s \in T} f_s(z) \geq f_t(z), \tag{7}$$

and that the first inequality may be strict. Indeed, one may have that $f_{\gamma_t}(z) = \lim_i f_{t_i}(z)$ for some $\gamma_{t_i} \to \gamma_t$ such that $(t_i)_i$ does not converge to t. This may happen, for instance, when T is compact but not Hausdorff. On the other side, if T is completely regular, for example compact Hausdorff, then

$$f_{\gamma_t}(z) = \limsup_{s \to t,\ s \in T} f_s(z).$$

The new functions f_γ, $\gamma \in \widehat{T}$, provide the same supremum f as the original ones f_t, $t \in T$:

Lemma 1 *The functions f_γ, $\gamma \in \widehat{T}$, are convex, and we have*

$$\sup_{\gamma \in \widehat{T}} f_\gamma = \sup_{t \in T} f_t = f.$$

Proof The convexity of the f_γ's follows easily from the convexity of the f_t's. Next, for each $\gamma \in \widehat{T}$ and $z \in X$, we have

$$f_\gamma(z) = \limsup_{\gamma_s \to \gamma,\ s \in T} f_s(z) \leq f(z),$$

entailing that $\sup_{\gamma \in \widehat{T}} f_\gamma \leq f$. In addition, if the sequence $(t_n)_n \subset T$ is such that $f(z) = \lim_n f_{t_n}(z)$, with $z \in X$, then there exist a subnet $(t_i)_i$ of $(t_n)_n$ and $\gamma \in \widehat{T}$ such that $\gamma_{t_i} \to \gamma$, and we get

$$f_\gamma(z) \geq \limsup_i f_{t_i}(z) = \lim_n f_{t_n}(z) = f(z),$$

showing that $\sup_{\gamma \in \widehat{T}} f_\gamma \geq f$. □

Now, given $x \in X$, with $f(x) \in \mathbb{R}$, and $\varepsilon \geq 0$, we introduce the extended ε-active index set of f at x by

$$\widehat{T}_\varepsilon(x) := \left\{ \gamma \in \widehat{T} :\ f_\gamma(x) \geq f(x) - \varepsilon \right\}; \tag{8}$$

and the extended active index set of f at x

$$\widehat{T}(x) := \widehat{T}_0(x) = \left\{ \gamma \in \widehat{T} :\ f_\gamma(x) = f(x) \right\}. \tag{9}$$

Moreover, taking into account (7), for each $t \in T(x)$ we have that

$$f(x) \geq f_{\gamma_t}(x) \geq f_t(x) = f(x);$$

that is,

$$\mathfrak{d}(T(x)) \subset \widehat{T}(x).$$

The set $\widehat{T}(x)$ is a nonempty set in spite of the possible emptiness of $T(x)$. More generally, we have:

Lemma 2 *The sets* $\widehat{T}_\varepsilon(x)$, $\varepsilon \geq 0$ *and* $x \in \operatorname{dom} f$, *are nonempty and compact.*

Proof It is enough to prove that $\widehat{T}(x)$ is nonempty and closed; the general case when $\varepsilon > 0$ is similar. Fix $x \in \operatorname{dom} f$. For a sequence $(t_n)_n \subset T$ such that $\lim_n f_{t_n}(x) = f(x)$ there will exist, due to the compactness of $\widehat{T}$, a subnet $(t_i)_i \subset T$ such that $\gamma_{t_i} \to \gamma \in \widehat{T}$, and then (7) ensures that

$$f(x) = \lim_i f_{t_i}(x) \leq \lim_i f_{\gamma_{t_i}}(x) \leq \limsup_{\gamma_t \to \gamma,\, t \in T} f_t(x) = f_\gamma(x) \leq f(x);$$

that is, $\gamma \in \widehat{T}(x)$ and this set is nonempty.

Next, we show that $\widehat{T}(x)$ is closed. We take a net $(\gamma_i)_i \subset \widehat{T}(x)$ that converges to γ $(\in \widehat{T})$. Then, by the definition of the f_γ's, for each i we find a net $(t_{ij})_j \subset T$ such that $\gamma_{t_{ij}} \to_j \gamma_i$ and

$$f(x) = f_{\gamma_i}(x) = \lim_j f_{t_{ij}}(x).$$

Thus, there exists a diagonal net $\left(\gamma_{t_{ij_i}}, f_{t_{ij_i}}(x)\right)_i \subset \widehat{T} \times \mathbb{R}$ such that $\gamma_{t_{ij_i}} \to_i \gamma$ and $f_{t_{ij_i}}(x) \to_i f(x)$; that is,

$$f_\gamma(x) \geq \limsup_i f_{t_{ij_i}}(x) = \lim_i f_{t_{ij_i}}(x) = f(x),$$

and so $\gamma \in \widehat{T}(x)$. □

Lemma 3 *If* $x \in \operatorname{dom} f$, *then*

$$\widehat{T}(x) = \bigcap\nolimits_{\varepsilon > 0} \operatorname{cl}\big(\mathfrak{d}(T_\varepsilon(x))\big).$$

Proof Take $\gamma \in \widehat{T}(x)$. Then there exists a net $(t_i)_i \subset T$ such that $\gamma_{t_i} \to \gamma$ and

$$f_\gamma(x) = \lim_i f_{t_i}(x) = f(x).$$

Hence, for each $\varepsilon > 0$ there exists an i_0 such that

$$t_i \in T_\varepsilon(x) \quad \text{for all } i \succeq i_0,$$

where $\succeq$ defines the order in the directed set. In other words, $\gamma_{t_i} \in \mathfrak{d}(T_\varepsilon(x))$ for all $i \succeq i_0$. This entails that $\gamma \in \operatorname{cl}(\mathfrak{d}(T_\varepsilon(x)))$, and we get $\gamma \in \bigcap_{\varepsilon > 0} \operatorname{cl}(\mathfrak{d}(T_\varepsilon(x)))$, by the arbitrariness of $\varepsilon > 0$.

Conversely, take $\gamma \in \bigcap_{\varepsilon > 0} \operatorname{cl}(\mathfrak{d}(T_\varepsilon(x)))$. Then, for each integer number k and each neighborhood U of γ, there exists some $\gamma_{t_{(k,U)}} \in U$ with $t_{(k,U)} \in T_{\frac{1}{k}}(x)$; that is (by (7)),

$$f(x) - \frac{1}{k} \leq f_{t_{(k,U)}}(x) \leq f_{\gamma_{t_{(k,U)}}}(x) \leq f(x) \leq 0.$$

Since $\widehat{T}$ is compact Hausdorff (coming form the complete regularity of T), the net $(\gamma_{t_{(k,U)}})_{(k,U)}$ converges and its limit must be equal to γ. Then

$$0 \geq f(x) \geq f_\gamma(x) \geq \limsup_{(k,U)} f_{t_{(k,U)}}(x) = 0,$$

and so $\gamma \in \widehat{T}(x)$. □

Let us examine the concepts introduced above in a compact (possibly, non-Hausdorff) framework. We denote by $\sim$ the equivalence relation on T given by

$$t_1 \sim t_2 \quad \Longleftrightarrow \quad \varphi(t_1) = \varphi(t_2) \quad \text{for all } \varphi \in \mathcal{C}(T, [0, 1]),$$

and by $\tilde{t}$ the equivalence class of $t \in T$. It is known that $\widehat{T}$ and T are homeomorphic when T is compact Hausdorff.

Lemma 4 *Assume that T is compact (possibly, non-Hausdorff). Then, provided that the mapping $t \longmapsto f_t(x)$ is continuous on T, the following assertions hold true for each $x \in$* dom f :

(i) $f_t(x) = f_s(x)$ *for all $s, t \in T$ such that $s \sim t$.*
(ii)

$$\widehat{T}(x) = \left\{ \tilde{t} \in T/\sim \ : \ f_{\tilde{t}}(x) = \limsup_{\tilde{s} \to \tilde{t}} f_s(x) = f(x) \right\} = \left\{ \tilde{t} \in T/\sim \ : \ t \in T(x) \right\},$$

where $\tilde{s} \to \tilde{t}$ means that $\varphi(s) \to \varphi(t)$ for all $\varphi \in \mathcal{C}(T, [0, 1])$.

Proof Under the current hypothesis it can be proved that $\widehat{T}$ and the quotient space $T/\sim$ are homeomorphic, by means of the mapping $\tilde{t} \in T/\sim \ \longmapsto \gamma_t \in \widehat{T}$.

(i) Since $f_{(\cdot)}(x)$ is continuous and T is compact we can easily prove the existence of $m > 0$ such that

$$|f_t(x)| \leq m \quad \text{for all } t \in T.$$

Thus, using the positive and the negative parts of $f_{(\cdot)}(x)$, $f_{(\cdot)}^+(x)$ and $f_{(\cdot)}^-(x)$, we have $m^{-1} f_{(\cdot)}^+(x), m^{-1} f_{(\cdot)}^-(x) \in \mathcal{C}(T, [0, 1])$ and so, for all $s, t \in T$ such that $s \sim t$,

$$f_t(x) = m\left(m^{-1} f_t^+(x) - m^{-1} f_t^-(x)\right) = m\left(m^{-1} f_s^+(x) - m^{-1} f_s^-(x)\right) = f_s(x).$$

(ii) If $\tilde{t} \in \widehat{T}(x)$, then there exists a net $(t_i)_i \subset T$ such that

$$\tilde{t_i} \to \tilde{t} \quad \text{and} \quad f(x) = f_{\tilde{t}}(x) = \lim_i f_{t_i}(x).$$

Since T is compact we may assume that $t_i \to s \in T$, and the continuity of $f_{(\cdot)}(x)$ entails

$$f(x) = \lim_i f_{t_i}(x) = f_s(x);$$

that is, $s \in T(x)$. Now, fix $\varphi \in \mathcal{C}(T, [0, 1])$. From the one hand, since $\tilde{t_i} \to \tilde{t}$, we have that

$$\varphi(t_i) \to \varphi(t).$$

On the other hand, the continuity of φ yields

$$\varphi(t_i) \to \varphi(s),$$

and we get $\varphi(t) = \varphi(s)$; that is, $\tilde{t} = \tilde{s}$.

Conversely, if $\tilde{t} \in T/\sim$ is such that $t \in T(x)$, then by (7) we get

$$f(x) \geq \limsup_{\tilde{s} \to \tilde{t}} f_s(x) = f_{\tilde{t}}(x) \geq f_t(x) = f(x);$$

and we are done. □

Now we give the main result of the paper, for general index sets and dropping both the upper semi-continuity-type condition and the compactness assumption assumed in Proposition 1.

Theorem 1 *Let $\{f_t, t \in T\}$ be a nonempty family of extended real-valued convex functions, and consider $f = \sup_{t \in T} f_t$. Then, for every $x \in \operatorname{dom} f$, we have*

$$\partial f(x) = \bigcap\nolimits_{L \in \mathcal{F}(x)} \operatorname{co} \left\{ \bigcup\nolimits_{\gamma \in \widehat{T}(x)} \partial(f_\gamma + \mathrm{I}_{L \cap \operatorname{dom} f})(x) \right\}, \tag{10}$$

where f_γ, $\widehat{T}(x)$ and $\mathcal{F}(x)$ are defined in (6), (9), and (2), respectively, and T is equipped with a completely regular topology.

In order to prove Theorem 1, we first establish the following key lemma, which constitutes the bridge with the compact framework.

Lemma 5 *Assume that f is proper and take $x \in \operatorname{dom} f$, with $f(x) \in \mathbb{R}$, and $\varepsilon \geq 0$.*

(i) *Every net $(\gamma_i)_i \subset \widehat{T}_\varepsilon(x)$ has an accumulation point $\gamma \in \widehat{T}_\varepsilon(x)$ such that*

$$\limsup_i f_{\gamma_i}(z) \leq f_\gamma(z) \quad \text{for all } z \in \operatorname{dom} f. \tag{11}$$

(ii) *If T is completely regular, then (11) holds for every net $(\gamma_i)_i \subset \widehat{T}_\varepsilon(x)$ converging to $\gamma \in \widehat{T}_\varepsilon(x)$.*

Proof (i) Fix a net $(\gamma_i)_i \subset \widehat{T}_\varepsilon(x)$ and, due to the compactness of $\widehat{T}_\varepsilon(x)$ established in Lemma 2, let $\gamma \in \widehat{T}_\varepsilon(x)$ be such that $\gamma_i \to \gamma$ (without loss of generality). Take $z \in \operatorname{dom} f$, so that $f_\gamma(z) \leq f(z) < +\infty$. Next, for each i there will exist a net $(t_{ij})_j \subset T$ such that

$$\gamma_{t_{ij}} \to_j \gamma_i, \quad f_{\gamma_i}(z) = \lim_j f_{t_{ij}}(z).$$

For every fixed $\delta > 0$ we may suppose, without loss of generality, that for all i

$$f_{t_{ij}}(z) \geq f_{\gamma_i}(z) - \delta \quad \text{eventually on } j.$$

Then there exists a diagonal net $(t_{ij_i})_i \subset T$ such that $\gamma_{t_{ij_i}} \to_i \gamma$ and

$$f_{t_{ij_i}}(z) \geq f_{\gamma_i}(z) - \delta \quad \text{for all } i.$$

Consequently,

$$f_\gamma(z) \geq \limsup_i f_{t_{ij_i}}(z) \geq \limsup_i f_{\gamma_i}(z) - \delta,$$

and we get, as $\delta \downarrow 0$,

$$f_\gamma(z) \geq \limsup_i f_{\gamma_i}(z).$$

(ii) Fix a net $(\gamma_i)_{i \in I} \subset \widehat{T}_\varepsilon(x)$ such that $\gamma_i \to \gamma \in \widehat{T}_\varepsilon(x)$, and take $z \in \operatorname{dom} f$ with $f_\gamma(z) < +\infty$. By assertion (i) the inequality (11) holds for some accumulation point of $(\gamma_i)_i$, which must be γ (because $\widehat{T}$ is Hausdorff).

□

Proof (of Theorem 1) By Lemma 1, the functions $\{f_\gamma, \gamma \in \widehat{T}\}$ are convex and satisfy

$$f = \sup_{\gamma \in \widehat{T}} f_\gamma.$$

According to Lemma 2, the sets $\widehat{T}_\varepsilon(x)$ are compact for every $\varepsilon \geq 0$, while Lemma 5(ii) entails the upper semi-continuity of the mappings $\gamma \longmapsto f_\gamma(z)$, $z \in \operatorname{dom} f$. Consequently, Proposition 1 applies and yields the desired formula. □

Corollary 1 *If* $f_{|\operatorname{aff}(\operatorname{dom} f)}$ *is continuous on* $\operatorname{ri}(\operatorname{dom} f)$ *(assumed to be nonempty), then for every* $x \in X$

$$\partial f(x) = \overline{\operatorname{co}} \left\{ \bigcup\nolimits_{\gamma \in \widehat{T}(x)} \partial (f_\gamma + \mathrm{I}_{\operatorname{dom} f})(x) \right\}. \tag{12}$$

Proof Under the current assumption, for every $L \in \mathcal{F}(x)$ and $\gamma \in \widehat{T}(x)$ such that $L \cap \operatorname{ri}(\operatorname{dom} f) \neq \emptyset$ we have that (see, e.g., [3, Theorem 15(iii)])

$$\partial(f_\gamma + \mathrm{I}_{L \cap \operatorname{dom} f})(x) = \partial\left((f_\gamma + \mathrm{I}_{\operatorname{dom} f}) + \mathrm{I}_L\right)(x) = \operatorname{cl}(\partial(f_\gamma + \mathrm{I}_{\operatorname{dom} f})(x) + L^\perp).$$

Now, given a convex neighborhood $U \subset X^*$ of the origin, we choose $L \in \mathcal{F}(x)$ such that $L \cap \operatorname{ri}(\operatorname{dom} f) \neq \emptyset$ and $L^\perp \subset U$. Then Theorem 1 yields

$$\begin{aligned} \partial f(x) &\subset \operatorname{co}\left\{ \bigcup\nolimits_{\gamma \in \widehat{T}(x)} \partial(f_\gamma + \mathrm{I}_{L \cap \operatorname{dom} f})(x) \right\} \\ &= \operatorname{co}\left\{ \bigcup\nolimits_{\gamma \in \widehat{T}(x)} \operatorname{cl}\left(\partial(f_\gamma + \mathrm{I}_{\operatorname{dom} f})(x) + L^\perp\right) \right\} \\ &\subset \operatorname{co}\left\{ \bigcup\nolimits_{\gamma \in \widehat{T}(x)} \partial(f_\gamma + \mathrm{I}_{\operatorname{dom} f})(x) \right\} + U + U, \end{aligned}$$

and we get, by intersecting over the U's,

$$\partial f(x) \subset \overline{\operatorname{co}} \left\{ \bigcup\nolimits_{\gamma \in \widehat{T}(x)} \partial (f_\gamma + \mathrm{I}_{\operatorname{dom} f})(x) \right\}.$$

The conclusion follows as the opposite inclusion is straightforward. □

Theorem 2 *If* f *is finite and continuous at some point, then for every* $x \in X$

$$\partial f(x) = \overline{\operatorname{co}} \left\{ \bigcup\nolimits_{\gamma \in \widehat{T}(x)} \partial f_\gamma(x) \right\} + \mathrm{N}_{\operatorname{dom} f}(x). \tag{13}$$

Proof Fix $x \in \operatorname{dom} f$. By taking into account that $f_\gamma \leq f$, Corollary 1 yields

$$\begin{aligned} \partial f(x) &= \overline{\operatorname{co}} \left\{ \bigcup\nolimits_{\gamma \in \widehat{T}(x)} \partial (f_\gamma + \mathrm{I}_{\operatorname{dom} f})(x) \right\} \\ &= \overline{\operatorname{co}} \left\{ \bigcup\nolimits_{\gamma \in \widehat{T}(x)} \partial f_\gamma(x) + \mathrm{N}_{\operatorname{dom} f}(x) \right\} \\ &= \partial \sigma_{\bigcup_{\gamma \in \widehat{T}(x)} \partial f_\gamma(x) + \mathrm{N}_{\operatorname{dom} f}(x)}(\theta) \\ &= \partial (\sigma_{\bigcup_{\gamma \in \widehat{T}(x)} \partial f_\gamma(x)} + \sigma_{\mathrm{N}_{\operatorname{dom} f}(x)})(\theta). \end{aligned} \tag{14}$$

Additionally, for a neighborhood U_{x_0} of $x_0 \in \operatorname{int}(\operatorname{dom} f)$ such that $U_{x_0} \subset \operatorname{dom} f$, we have

$$\sigma_{\mathrm{N}_{\operatorname{dom} f}(x)}(U_{x_0} - x) \leq 0,$$

showing that $\sigma_{\mathrm{N}_{\operatorname{dom} f}(x)}$ is continuous at $x_0 - x$. At the same time, we have

$$\sigma_{\bigcup_{\gamma \in \widehat{T}(x)} \partial f_\gamma(x)}(x_0 - x) \leq \sup_{\gamma \in \widehat{T}(x)} (f_\gamma(x_0) - f_\gamma(x)) \leq f(x_0) - f(x) < +\infty,$$

and so, thanks to the Moreau–Rockafellar sum rule, (14) implies that

$$\begin{aligned}\partial f(x) &= \partial\left(\sigma_{\bigcup_{\gamma\in\widehat{T}(x)}\partial f_\gamma(x)} + \sigma_{\mathrm{N}_{\mathrm{dom}\, f}(x)}\right)(\theta)\\ &= \partial\sigma_{\bigcup_{\gamma\in\widehat{T}(x)}\partial f_\gamma(x)}(\theta) + \partial\sigma_{\mathrm{N}_{\mathrm{dom}\, f}(x)}(\theta)\\ &= \overline{\mathrm{co}}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}(x)}\partial f_\gamma(x)\right\} + \mathrm{N}_{\mathrm{dom}\, f}(x).\end{aligned}$$

□

The following corollary is a straightforward consequence of Theorems 1 and 2. We introduce the functions $\tilde{f}_t : X \to \overline{\mathbb{R}}$, $t \in T$, given by

$$\tilde{f}_t(z) := \limsup_{s\to t,\, s\in T} f_s(z), \tag{15}$$

and denote

$$\widetilde{T}(x) := \{t \in T : \ \tilde{f}_t(x) = f(x)\}. \tag{16}$$

Corollary 2 *Assume that T is compact Hausdorff. Then, for every $x \in \mathrm{dom}\, f$, we have*

$$\partial f(x) = \bigcap\nolimits_{L\in\mathcal{F}(x)} \mathrm{co}\left\{\bigcup\nolimits_{t\in\widetilde{T}(x)} \partial(\tilde{f}_t + \mathrm{I}_{L\cap\mathrm{dom}\, f})(x)\right\}.$$

If, in addition, f is finite and continuous at some point, then for every $x \in \mathrm{dom}\, f$

$$\partial f(x) = \mathrm{N}_{\mathrm{dom}\, f}(x) + \overline{\mathrm{co}}\left\{\bigcup\nolimits_{t\in\widetilde{T}(x)} \partial\tilde{f}_t(x)\right\}.$$

Proof Since T is compact Hausdorff; hence, completely regular, we have that $T \equiv \widehat{T}$ and, for all $t \in T$ and $z \in X$,

$$f_{\gamma_t}(z) = \lim_{\gamma_s\to\gamma_t,\, s\in T} f_s(z) = \lim_{s\to t,\, s\in T} f_s(z) = \tilde{f}_t(z).$$

In other words, the first formula is a consequence of Theorem 1. Similarly, the second statement of the theorem follows from Theorem 2. □

The following corollary shows how to deduce Valadier's formula ([30]), given in the compact setting (see [19, Theorem 3, p. 201] and [31, Theorem 2.4.18]).

Corollary 3 *Assume that T is compact Hausdorff. Let $U \subset X$ be an open set such that:*

(i) $f_t(x) \in \mathbb{R}$ *for all* $t \in T$ *and* $x \in U$,
(ii) $t \in T \longmapsto f_t(x)$ *is upper semi-continuous for each* $x \in U$,
(iii) $x \in U \longmapsto f_t(x)$ *is continuous for each* $t \in T$.

Then for every $x \in U$ we have

$$\partial f(x) = \overline{\mathrm{co}}\left\{\bigcup\nolimits_{t\in T(x)} \partial f_t(x)\right\}.$$

Proof Assume first that X is a Banach space. Then, using classical arguments (see, e.g., [19, 31]), it is shown that the supremum function $f = \sup_{t\in T} f_t$ is finite and, so, continuous on U. Thus, by Corollary 2, for each $x \in U$ we have

$$\partial f(x) = \mathrm{N}_{\mathrm{dom}\, f}(x) + \overline{\mathrm{co}}\left\{\bigcup\nolimits_{t\in\widetilde{T}(x)} \partial\tilde{f}_t(x)\right\} = \overline{\mathrm{co}}\left\{\bigcup\nolimits_{t\in\widetilde{T}(x)} \partial\tilde{f}_t(x)\right\}, \tag{17}$$

where $\tilde{f}_t$ and $\widetilde{T}(x)$ are defined in (15) and (16), respectively.

Take $t \in \widetilde{T}(x)$. On the one hand, using the compactness assumption, there exist some net $(t_i)_i \subset T$ and $t \in T$ such that $f(x) = \lim_i f_{t_i}(x)$ and $(t_i)_i$ converges to $t \in T$. But we have, due to assumption (ii),

$$f(x) = \limsup_i f_{t_i}(x) \leq f_t(x),$$

and so $t \in T(x)$.

On the other hand, also by assumption (ii), for all $z \in U$ we have

$$\tilde{f}_t(z) = \limsup_{s \to t} f_s(z) \leq f_t(z),$$

and both functions $\tilde{f}_t$ and f_t coincide at x. Consequently, $\partial \tilde{f}_t(x) \subset \partial f_t(x)$ and (17) yields

$$\partial f(x) = \overline{\text{co}} \left\{ \bigcup\nolimits_{t \in \widetilde{T}(x)} \partial \tilde{f}_t(x) \right\} \subset \overline{\text{co}} \left\{ \bigcup\nolimits_{t \in T(x)} \partial f_t(x) \right\}.$$

Thus, we are done since the opposite inclusion is straightforward.

We consider now the case when X is any locally convex space. We fix $x \in U$ and $x^* \in \partial f(x)$. Given an $L \in \mathcal{F}(x)$, we introduce the convex functions $g_t : L \to \overline{\mathbb{R}}$, $t \in T$, defined as

$$g_t := (f_t + \mathrm{I}_L)_{|L};$$

that is, g_t is the restriction of $f_t + \mathrm{I}_L$ to L, and consider the associated supremum

$$g := \sup_{t \in T} g_t = (f + \mathrm{I}_L)_{|L}.$$

Therefore, since the family $\{g_t, t \in T\}$ satisfies the requirements of the paragraph above, we obtain

$$\partial g(x) = \overline{\text{co}} \left\{ \bigcup\nolimits_{t \in T(x)} \partial g_t(x) \right\}.$$

Now, take $x^* \in \partial f(x)$, so that $\hat{x}^* := x^*_{|L} \in \partial g(x) = \overline{\text{co}} \left\{ \bigcup_{t \in T(x)} \partial g_t(x) \right\}$. Then, thanks to the fact that L^* is isomorphic to the quotient space $X^*/L^\perp$, for every θ-neighborhood $V \subset X^*$ we have that

$$\hat{x}^* \in \text{co} \left\{ \bigcup\nolimits_{t \in T(x)} \partial g_t(x) \right\} + V_{|L},$$

where $V_{|L} := \{u^*_{|L} : u^* \in V\}$ is a θ-neighborhood in $X^*/L^\perp$. In other words, there are $u^* \in V, \lambda_1, \ldots, \lambda_k \geq 0, t_1, \ldots, t_k \in T(x)$ and $\hat{x}^*_1, \cdots, \hat{x}^*_k \in L^*$ such that $\lambda_1 + \cdots + \lambda_k = 1$, $\hat{x}^*_j \in \partial g_{t_j}(x)$, $j = 1, \ldots, k$, $k \geq 1$, and

$$\hat{x}^* = \lambda_1 \hat{x}^*_1 + \cdots + \lambda_k \hat{x}^*_k + u^*_{|L}.$$

Moreover, by the Hahn–Banach theorem, we extend $\hat{x}^*_1, \ldots, \hat{x}^*_k$ to $x^*_1, \ldots, x^*_k \in X^*$, which satisfy

$$\langle x^*, u \rangle = \lambda_1 \langle x^*_1, u \rangle + \cdots + \lambda_k \langle x^*_k, u \rangle + \langle u^*, u \rangle \quad \text{for all } u \in L;$$

that is, $x^* \in \lambda_1 x^*_1 + \cdots + \lambda_k x^*_k + u^* + L^\perp$. But $x^*_j \in \partial (f_{t_j} + \mathrm{I}_L)(x)$, $j = 1, \ldots, k$, and so

$$\begin{aligned} x^* &\in \text{co} \left\{ \bigcup\nolimits_{t \in T(x)} \partial (f_t + \mathrm{I}_L)(x) \right\} + V + L^\perp \\ &= \text{co} \left\{ \bigcup\nolimits_{t \in T(x)} \partial f_t(x) \right\} + V + L^\perp, \end{aligned}$$

where the last equality follows by applying the Moreau–Rockafellar sum rule (thanks to assumption (iii)). Finally, because L and V were arbitrarily chosen, we deduce that $x^* \in \overline{\text{co}} \left\{ \bigcup_{t \in T(x)} \partial f_t(x) \right\}$ (see (1)), and the inclusion "$\subset$" follows. □

Corollary 4 *Assume that* $X = \mathbb{R}^n$. *Then* (12) *and* (13) *hold with* co *instead of* $\overline{\text{co}}$.

Proof Similarly as in the proof Proposition 1, we can prove that the set

$$\text{co}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}(x)} \partial(f_\gamma + \text{I}_{\text{dom}\, f})(x)\right\}$$

is closed and (12) holds with co instead of $\overline{\text{co}}$; that is,

$$\partial f(x) = \text{co}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}(x)} \partial(f_\gamma + \text{I}_{\text{dom}\, f})(x)\right\}. \tag{18}$$

In addition, if f is finite and continuous at some point in dom f, then each function f_γ ($\leq f$), $\gamma \in \widehat{T}(x)$, is finite and continuous at the same point, and (18) yields (13) with co instead of $\overline{\text{co}}$,

$$\partial f(x) = \text{co}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}(x)} \partial(f_\gamma + \text{I}_{\text{dom}\, f})(x)\right\} = \text{co}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}(x)} \partial f_\gamma(x)\right\} + \text{N}_{\text{dom}\, f}(x).$$

□

Example 1 Consider the family of convex functions $g_{2n+1}, h_{2n}, n \in \mathbb{N}$, defined on $\mathbb{R}$ as

$$g_{2n+1}(z) := \max\left\{\frac{nz}{n+1}, 0\right\}, \quad h_{2n}(z) := \max\left\{\frac{-nz}{n+1}, 0\right\}.$$

We introduce the family $\{f_n, n \in \mathbb{N}\}$ such that $f_{2n+1} := g_{2n+1}$ and $f_{2n} := h_{2n}$, together with the supremum function

$$f = \sup_{n\in\mathbb{N}} f_n = \sup_{n\in\mathbb{N}} \{g_{2n+1}, h_{2n}\}.$$

Obviously,

$$f(x) = |x| \quad \text{and} \quad \partial f(x) = \begin{cases} [-1, 1] & \text{if } x = 0, \\ \{-1\} & \text{if } x > 0, \\ \{1\} & \text{if } x > 0, \end{cases}$$

and

$$T(x) = \begin{cases} \mathbb{N} & \text{if } x = 0, \\ \emptyset & \text{if } x \neq 0. \end{cases}$$

Thus, if we apply (4) in Proposition 1, we reach a false conclusion as the assumption there is not satisfied in this case:

$$\partial f(x) = \begin{cases}]-1, 1[& \text{if } x = 0, \\ \emptyset & \text{if } x \neq 0. \end{cases}$$

The Stone–Čech compactification of $\mathbb{N}$ is given by

$$\begin{aligned} \widehat{\mathbb{N}} &= \mathbb{N} \cup \left\{\lim_i \gamma_{n_i} : (n_i)_i \subset \mathbb{N}, n_i \to +\infty\right\} \\ &= \mathbb{N} \cup \left\{\lim_i \gamma_{2n_i}, \ \lim_i \gamma_{2n_i+1} : (n_i)_i \subset \mathbb{N}, n_i \to +\infty\right\}, \end{aligned}$$

whereas the f_γ's, $\gamma \in \widehat{\mathbb{N}}$, take the form

$$f_\gamma = \begin{cases} g_{2n+1} & \text{if } \gamma = \gamma_{2n+1} \equiv 2n+1, \\ h_{2n} & \text{if } \gamma = \gamma_{2n} \equiv 2n, \end{cases}$$

for $\gamma \in \mathbb{N}$, and

$$f_\gamma = \limsup_{\gamma_n \to \gamma} f_n$$

for $\gamma \in \widehat{\mathbb{N}} \setminus \mathbb{N}$. Equivalently, we consider the family

$$\left\{g_{2n+1}, h_{2n}, n \in \mathbb{N};\ g_{\bar{\gamma}}, h_{\bar{\gamma}}\right\},$$

where $g_{\bar{\gamma}}, h_{\bar{\gamma}} : \mathbb{R} \to \mathbb{R}$ are defined as

$$\begin{aligned} g_{\bar{\gamma}}(z) &= \limsup_{n\to\infty} g_{2n+1}(z) = \max\{z, 0\}, \\ h_{\bar{\gamma}}(z) &= \limsup_{n\to\infty} h_{2n}(z) = \max\{-z, 0\}. \end{aligned}$$

It is easily checked that this new family has the same properties as the original one, $\{f_\gamma, \gamma \in \widehat{\mathbb{N}}\}$. In other words, we have enlarged the original family of functions by adding $g_{\bar{\gamma}}$ and $h_{\bar{\gamma}}$. Therefore, applying (10), we get

$$\begin{aligned} \partial f(0) &= \operatorname{co}\left\{\bigcup\nolimits_{n\in T(0)} \partial g_{2n+1}(0) \bigcup \partial g_{\bar{\gamma}}(0) \bigcup\nolimits_{n\in T(0)} \partial h_{2n}(0) \bigcup \partial h_{\bar{\gamma}}(0)\right\} \\ &= \operatorname{co}\left\{\bigcup\nolimits_{n\geq 1}\left[\frac{n}{n+1}, 0\right]\bigcup [0,1] \bigcup\nolimits_{n\geq 1}\left[\frac{-n}{n+1}, 0\right]\bigcup[-1,0]\right\} = [-1,1], \end{aligned}$$

and, for $x \neq 0$, say $x = 1$,

$$\partial f(1) = \partial g_{\bar{\gamma}}(1) = \{1\}.$$

Observe that the presence of the new functions $g_{\bar{\gamma}}$ and $h_{\bar{\gamma}}$ is necessary, since the subdifferentials at 0 of the data functions g_{2n+1} and h_{2n} do not lead us to the whole subdifferential of the supremum function f, as they do not include the subgradients -1 and 1.

In order to decompose the subdifferential term involved in formula (10) we need to impose some additional continuity or lower semi-continuity conditions on the initial functions. The assumption in Theorem 2 gives the first example, where the continuity of the supremum function allows to characterize $\partial f(x)$ by means only of the sets $\partial f_\gamma(x)$. We give next an alternative representation of $\partial f(x)$ by means of the ε-subdifferentials of the f_γ's, under the condition

$$\operatorname{cl} f = \sup_{t\in T}(\operatorname{cl} f_t), \tag{19}$$

where $\operatorname{cl} f$ and $\operatorname{cl} f_t$ are the *closed hulls* (*lower semi-continuous regularizations*) of the respective functions.

Proposition 2 *If* (19) *holds, then for every* $x \in \operatorname{dom} f$

$$\partial f(x) = \bigcap\nolimits_{\varepsilon>0, L\in\mathcal{F}(x)} \overline{\operatorname{co}}\left\{\bigcup\nolimits_{\gamma\in\widehat{T}(x)} \partial_\varepsilon f_\gamma(x) + \mathrm{N}_{L\cap \operatorname{dom} f}(x)\right\},$$

where f_γ, $\widehat{T}(x)$ *and* $\mathcal{F}(x)$ *are defined in* (6), (9), *and* (2), *respectively, and* T *is a completely regular topological space.*

Proof It suffices to apply [7, Theorem 3.8] to the family $\{f_\gamma, \gamma \in \widehat{T}\}$. □

We discuss next a nonconvex counterpart of formula (10), under the following condition introduced in [21],

$$f^{**} = \sup_{t\in T} f_t^{**}, \tag{20}$$

where f^{**} and f_t^{**} are the *biconjugates* of the respective functions. In the convex case, and assuming that the *conjugates* f^* and f_t^* are proper, (19) is equivalent to the last relation.

Proposition 3 *Let $\{f_t, t \in T\}$ be a nonempty family of extended real-valued non-necessarily convex functions, and consider $f = \sup_{t\in T} f_t$. If condition (20) holds, then for every $x \in$* dom f

$$\partial f(x) = \bigcap_{L\in\mathcal{F}(x)} \text{co}\left\{\bigcup_{\gamma\in\widehat{T}(x)} \partial(f_\gamma + \text{I}_{L\cap\text{dom}\, f})(x)\right\}$$
$$= \bigcap_{\varepsilon>0,\, L\in\mathcal{F}(x)} \overline{\text{co}}\left\{\bigcup_{\gamma\in\widehat{T}(x)} \partial_\varepsilon f_\gamma(x) + \text{N}_{L\cap\text{dom}\, f}(x)\right\},$$

where f_γ, $\widehat{T}(x)$, and $\mathcal{F}(x)$ are defined in (6), (9), and (2), respectively, and T is equipped with a completely regular topology.

Proof Assume that $\partial f(x) \neq \emptyset$, so that $f(x) = f^{**}(x)$ and $\partial f(x) = \partial f^{**}(x)$. Then, by applying Theorem 1 to the family $\{f_t^{**}, t \in T\}$, we obtain

$$\partial f(x) = \partial f^{**}(x) = \bigcap_{L\in\mathcal{F}(x)} \text{co}\left\{\bigcup_{\gamma\in\widehat{T}^1(x)} \partial(g_\gamma + \text{I}_{L\cap\text{dom}\, f})(x)\right\}, \tag{21}$$

where $g_\gamma : X \to \overline{\mathbb{R}}$, $\gamma \in \widehat{T}$, are defined by

$$g_\gamma(z) := \limsup_{\gamma_t\to\gamma,\, t\in T} f_t^{**}(z),$$

and

$$\widehat{T}^1(x) := \left\{\gamma \in \widehat{T} : g_\gamma(x) = f(x)\right\}.$$

Observe that for every $\gamma \in \widehat{T}^1(x)$ we have that

$$f(x) = \limsup_{\gamma_t\to\gamma,\, t\in T} f_t^{**}(x) \leq \limsup_{\gamma_t\to\gamma,\, t\in T} f_t(x) = f_\gamma(x) \leq f(x),$$

and so $\gamma \in \widehat{T}(x) = \{\gamma \in \widehat{T} : f_\gamma(x) = f(x)\}$. Moreover, since

$$g_\gamma(z) = \limsup_{\gamma_t\to\gamma,\, t\in T} f_t^{**}(z) \leq \limsup_{\gamma_t\to\gamma,\, t\in T} f_t(z) = f_\gamma(z) \quad \text{for all } z \in X,$$

we deduce that for all $L \in \mathcal{F}(x)$

$$\partial\left(g_\gamma + \text{I}_{L\cap\text{dom}\, f}\right)(x) \subset \partial\left(f_\gamma + \text{I}_{L\cap\text{dom}\, f}\right)(x).$$

Thus, the inclusion "⊂" in the first statement follows from (21), and we are done since the opposite inclusion is easily verified.

The second statement follows similarly by using Proposition 2 instead of Theorem 1. □

5 An Application to Optimality Conditions

In this section, we revise the optimality conditions for convex semi-infinite programming established in [6], by removing the compactness of the set indexing the constraints.

Aside [6], a significant precedent of the results in this section can be found in [11, Chapter 7], where KKT conditions are established for convex semi-infinite optimization with finite-valued functions, using a closedness condition which is implied by some version of Slater's qualification. Many KKT conditions exist in the literature which are obtained via different approaches: approximate subdifferentials of the data functions ([3, 16]), the exact subdifferentials at close points [28], Farkas–Minkowski-type closedness criteria [8] in convex semi-infinite optimization, strong CHIP-like qualifications for convex optimization with non necessarily convex $\mathcal{C}^1$-constraints [2] (see, also, [9] for locally Lipschitz constraints), among others.

Here we consider the following optimization problem

$$(\mathcal{P}): \qquad \inf_{f_t(x)\leq 0,\, t\in T} f_0(x),$$

where T is a completely regular topological space, and $f_t : \mathbb{R}^n \to \mathbb{R}_\infty$, for $t \in T \cup \{0\}$ (we assume, without loss of generality, that $0 \notin T$), are proper and convex. Problem $(\mathcal{P})$ is equivalent to

$$\inf_{f(x)\leq 0} f_0(x),$$

where

$$f := \sup_{t\in T} f_t.$$

Let the set $\widehat{T}$ and the convex functions $f_\gamma : \mathbb{R}^n \to \mathbb{R}_\infty$, $\gamma \in \widehat{T}$, be as defined in (5) and (6), respectively. We also denote

$$\widehat{A}(x) := \{\gamma \in \widehat{T} : f_\gamma(x) = 0\},$$

so that, by Lemma 3, for every feasible point $x \in \mathbb{R}^n$ for $(\mathcal{P})$ we have

$$\widehat{A}(x) = \bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\mathfrak{d}(A_\varepsilon(x))\right),$$

where

$$A_\varepsilon(x) := \{t \in T :\ f_t(x) \geq -\varepsilon\}, \quad \varepsilon > 0.$$

The following theorem establishes Fritz–John-type necessary optimality conditions for problem $(\mathcal{P})$. The main feature of this result and the subsequent corollary is the absence of any compactness and continuity assumptions on the index set and the mappings $t \longmapsto f_t(z)$, as they were required in [6, Theorem 5].

Theorem 3 *Assume that $\bar{x}$ is an optimal solution of $(\mathcal{P})$. Then we have*

(a)

$$0_n \in \operatorname{co}\left\{\partial(f_0 + \mathrm{I}_{\operatorname{dom} f})(\bar{x}) \cup \bigcup\nolimits_{\gamma\in\widehat{A}(\bar{x})} \partial(f_\gamma + \mathrm{I}_{\operatorname{dom} f_0 \cap \operatorname{dom} f})(\bar{x})\right\}.$$

(b) *Moreover, under the condition*

$$\operatorname{ri}(\operatorname{dom} f_\gamma) \cap \operatorname{ri}(\operatorname{dom} f) \neq \emptyset \quad \textit{for all } \gamma \in \widehat{A}(\bar{x}) \cup \{0\},$$

we have

$$0_n \in \operatorname{co}\left\{\partial f_0(\bar{x}) \cup \bigcup\nolimits_{\gamma\in\widehat{A}(\bar{x})} \partial f_\gamma(\bar{x})\right\} + \mathrm{N}_{\operatorname{dom} f}(\bar{x}) + \mathrm{N}_{\operatorname{dom} f_0}(\bar{x}),$$

Proof We consider the supremum function $g : \mathbb{R}^n \to \mathbb{R}_\infty$, defined as

$$g(x) := \sup\{f_0(x) - f_0(\bar{x}),\ f_t(x),\ t \in T\} = \max\{f_0(x) - f_0(\bar{x}),\ f(x)\},$$

so that $\operatorname{dom} g = \operatorname{dom} f_0 \cap \operatorname{dom} f$. It is easily verified that $\bar{x}$ is a global minimum of g; that is, $0_n \in \partial g(\bar{x})$.

We endow the set $T \cup \{0\}$ with the topology generated by the open sets of T and $\{0\}$, which makes it completely regular. Then the compactification of $T \cup \{0\}$ can be identified with $\widehat{T} \cup \{0\}$. Consequently, and according to Corollary 4, $\bar{x}$ satisfies

$$0_n \in \partial g(\bar{x}) = \operatorname{co}\left\{\partial(f_0 + \mathrm{I}_{\operatorname{dom} f})(\bar{x}) \cup \bigcup\nolimits_{\gamma\in\widehat{A}(\bar{x})} \partial(f_\gamma + \mathrm{I}_{\operatorname{dom} f_0 \cap \operatorname{dom} f})(\bar{x})\right\},$$

which is condition (a).

(b) Under the current assumptions, by using the classical sum rule ([25]), we get from the one hand

$$\partial\left(f_0 + \mathrm{I}_{\mathrm{dom}\, f}\right)(\bar{x}) = \partial f_0(\bar{x}) + \mathrm{N}_{\mathrm{dom}\, f}(\bar{x}),$$

and from the other hand, since

$$f_\gamma + \mathrm{I}_{\mathrm{dom}\, f_0 \cap \mathrm{dom}\, f} = f_\gamma + \mathrm{I}_{\mathrm{dom}\, f} + \mathrm{I}_{\mathrm{dom}\, f_0} \quad \text{and} \quad \mathrm{dom}\,(f_\gamma + \mathrm{I}_{\mathrm{dom}\, f}) = \mathrm{dom}\, f,$$

we obtain

$$\partial\left(f_\gamma + \mathrm{I}_{\mathrm{dom}\, f_0 \cap \mathrm{dom}\, f}\right)(\bar{x}) = \partial f_\gamma(\bar{x}) + \mathrm{N}_{\mathrm{dom}\, f_0}(\bar{x}) + \mathrm{N}_{\mathrm{dom}\, f}(\bar{x}).$$

Thus, the conclusion follows from (a). □

Remark 2 In particular, if $f(\bar{x}) < 0$, then the last condition reads

$$0_n \in \partial\left(f_0 + \mathrm{I}_{\mathrm{dom}\, f}\right)(\bar{x}),$$

as $f_\gamma(\bar{x}) \leq f(\bar{x}) < 0$ for all $\gamma \in \widehat{A}$, and so $\widehat{A}(\bar{x}) = \emptyset$.

Remark 3 Observe that the strong Slater condition; i.e., the existence of some $x_0 \in \mathrm{dom}\, f_0$ such that $f(x_0) < 0$, does not imply that x_0 is an interior point of the feasible set. This is what happens in the following example. Take $T := [0, +\infty[$, $f_0 \equiv 0$ and let $f_t : \mathbb{R} \to \mathbb{R}$, $t \in T$, be defined as

$$f_t(x) := \max\{tx - 1, -tx - 1\}.$$

The point 0 is a strong Slater point, but $0 \notin \mathrm{int}(\{x \in \mathbb{R} : \; f_t(x) \leq 0, \; t \in T\})$.

We derive next the KKT conditions for problem $(\mathcal{P})$ under the Slater qualification.

Corollary 5 *Under the strong Slater condition; that is,*

$$f(x_0) < 0 \quad \textit{for some } x_0 \in \mathrm{dom}\, f_0,$$

the point $\bar{x}$ is optimal for $(\mathcal{P})$ if and only if

$$0_n \in \partial\left(f_0 + \mathrm{I}_{\mathrm{dom}\, f}\right)(\bar{x}) + \mathrm{cone}\left\{\bigcup\nolimits_{\gamma \in \widehat{A}(\bar{x})} \partial\left(f_\gamma + \mathrm{I}_{\mathrm{dom}\, f_0 \cap \mathrm{dom}\, f}\right)(\bar{x})\right\}. \tag{22}$$

Proof Assume first that $f(\bar{x}) = 0$. By Theorem 3(a), $\bar{x}$ is optimal if and only if either

$$0_n \in \mathrm{co}\left\{\bigcup\nolimits_{\gamma \in \widehat{A}(\bar{x})} \partial\left(f_\gamma + \mathrm{I}_{\mathrm{dom}\, f_0 \cap \mathrm{dom}\, f}\right)(\bar{x})\right\} \tag{23}$$

or (22) holds.

Moreover, by Theorem 1 we have that

$$\begin{aligned}\mathrm{co}\left\{\bigcup\nolimits_{\gamma \in \widehat{A}(\bar{x})} \partial\left(f_\gamma + \mathrm{I}_{\mathrm{dom}\, f_0 \cap \mathrm{dom}\, f}\right)(\bar{x})\right\} &= \partial\left(\sup_{t \in T}\left(f_t + \mathrm{I}_{\mathrm{dom}\, f_0 \cap \mathrm{dom}\, f}\right)\right)(\bar{x}) \\ &= \partial\left(f + \mathrm{I}_{\mathrm{dom}\, f_0}\right)(\bar{x}),\end{aligned}$$

and so relation (23) is equivalent to

$$0_n \in \mathrm{co}\left\{\bigcup\nolimits_{\gamma \in \widehat{A}(\bar{x})} \partial\left(f_\gamma + \mathrm{I}_{\mathrm{dom}\, f_0 \cap \mathrm{dom}\, f}\right)(\bar{x})\right\} = \partial\left(f + \mathrm{I}_{\mathrm{dom}\, f_0}\right)(\bar{x}),$$

equivalently, $f(x) \geq f(\bar{x}) = 0$ for all $x \in \mathrm{dom}\, f_0$; and this contradicts the strong Slater condition.

Finally, if $f(\bar{x}) < 0$, then (22) follows by Theorem 3(a). □

Acknowledgements Research supported by CONICYT (Fondecyt 1190012 and 1190110), Proyecto/Grant PIA AFB-170001, MICIU of Spain and Universidad de Alicante (Grant Beatriz Galindo BEA- GAL 18/00205), and Research Project PGC2018-097960-B-C21 from MICINN, Spain. The research of the third author is also supported by the Australian ARC - Discovery Projects DP 180100602.

The authors wish to thank the referee for the valuable comments and suggestions which have contributed to improve the first version of this paper.

References

1. Brøndsted, A.: On the subdifferential of the supremum of two convex functions. Math. Scand. **31**, 225–230 (1972)
2. Chieu, N.H., Jeyakumar, V., Li, G., Mohebi, H.: Constraint qualifications for convex optimization without convexity of constraints: new connections and applications to best approximation. Eur. J. Oper. Res. **265**, 19–25 (2018)
3. Correa, R., Hantoute, A., López, M.A.: Weaker conditions for subdifferential calculus of convex functions. J. Funct. Anal. **271**, 1177–1212 (2016)
4. Correa, R., Hantoute, A., López, M.A.: Towards supremum-sum subdifferential calculus free of qualification conditions. SIAM J. Optim. **26**, 2219–2234 (2016)
5. Correa, R., Hantoute, A., López, M.A.: Valadier-like formulas for the supremum function I. J. Convex Anal. **25**, 1253–1278 (2018)
6. Correa, R., Hantoute, A., López, M.A.: Moreau–rockafellar type formulas for the subdifferential of the supremum function. SIAM J. Optim. **29**, 1106–1130 (2019)
7. Correa, R., Hantoute, A., López, M.A.: Valadier-like formulas for the supremum function II: the compactly indexed case. J. Convex Anal. **26**, 299–324 (2019)
8. Dinh, N., Goberna, M.A., López, M.A.: From linear to convex systems: consistency, Farkas' lemma and applications. J. Convex Anal. **13**, 113–133 (2006)
9. Dutta, J., Lalitha, C.S.: Optimality conditions in convex optimization revisited. Optim. Lett. **7**, 221–229 (2013)
10. Ekeland, I., Témam, R.: Convex Analysis and Variational Problems. North-holland & American Elsevier, Amsterdam (1976)
11. Goberna, M.A., López, M.A.: Linear Semi-infinite Optimization. J. Wiley, Chichester (1998)
12. Hantoute, A.: Subdifferential set of the supremum of lower semi-continuous convex functions and the conical hull property. Top **14**, 355–374 (2006)
13. Hantoute, A., López, M.A.: A complete characterization of the subdifferential set of the supremum of an arbitrary family of convex functions. J. Convex Anal. **15**, 831–858 (2008)
14. Hantoute, A., López, M.A., Zălinescu, C.: Subdifferential calculus rules in convex analysis: a unifying approach via pointwise supremum functions. SIAM J. Optim. **19**, 863–882 (2008)
15. Hiriart-Urruty, J.-B.: Convex analysis and optimization in the past 50 years: some snapshots. In: Demyanov, V.F., Pardalos, P.M., Batsyn, M. (eds.) Constructive Nonsmooth Analysis and Related Topics. Springer Optimization and Its Applications, vol. 87, pp. 245–253. Springer, New York (2014)
16. Hiriart-Urruty, J.-B., Phelps, R.R.: Subdifferential calculus using ε-subdifferentials. J. Funct. Anal. **118**, 154–166 (1993)
17. Ioffe, A.D.: A note on subdifferentials of pointwise suprema. Top **20**, 456–466 (2012)
18. Ioffe, A.D., Levin, V.L.: Subdifferentials of convex functions. Tr. Moskov Mat. Obshch **26**, 3–73 (1972). (In Russian)
19. Ioffe, A.D., Tikhomirov, V.H.: Theory of Extremal Problems. Studies in Mathematics and Its Applications, vol. 6. North-Holland, Amsterdam (1979)
20. Levin, V.L.: An application of Helly's theorem in convex programming, problems of best approximation and related questions. Mat. Sb., Nov. Ser. **79**, 250–263 (1969). Engl. Trans.: Math. USSR, Sb. 8, 235–247 (1969)
21. López, M.A., Volle, M.: A formula for the set of optimal solutions of relaxed minimization problems. Applications to subdifferential calculus. J. Convex Anal. **17**, 1057–1075 (2010)
22. Munkres, J.: Topology, 2nd edn. Prentice Hall, Upper Saddle River (2000)
23. Pschenichnyi, B.N.: Convex programming in a normalized space. Kibernetika **5**, 46–54 (1965). (Russian); Engl. Trans.: Cybernetics 1, 46–57 (1965)
24. Rockafellar, R.T., Brøndsted, A.: On the subdifferentiability of convex functions. Proc. Am. Math. Soc. **16**, 605–611 (1965)

25. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
26. Rockafellar, R.T.: Directionally Lipschitzian functions and subdifferential calculus. Proc. Lond. Math. Soc. **39**, 331–355 (1979)
27. Solov'ev, V.N.: The subdifferential and the directional derivatives of the maximum of a family of convex functions. Izvestiya RAN: Ser. Mat. **65**, 107–132 (2001)
28. Thibault, L.: Sequential convex subdifferential calculus and sequential Lagrange multipliers. SIAM J. Control Optim. **35**, 1434–1444 (1997)
29. Tikhomirov, V.M.: Analysis II: convex analysis and approximation theory. In: Gamkrelidze, R.V. (ed.) Encyclopaedia of Mathematical Sciences, vol. 14. Springer, Berlin (1990)
30. Valadier, M.: Sous-différentiels d'une borne supérieure et d'une somme continue de fonctions convexes. C. R. Acad. Sci. Paris Sé,r. A-B **268**, A39–A42 (1969)
31. Zălinescu, C.: Convex Analysis in General Vector Spaces. World Scientific, River Edge (2002)