Set-Valued and Variational Analysis (2024) 32:8
https://doi.org/10.1007/s11228-024-00712-8

RECENT FRANCO-CHILEAN COLLABORATIONS IN VARIATIONAL ANALYSIS

Conjugation-Based Approach to the ε-Subdifferential of Convex Suprema

Rafael Correa[1] · **Abderrahim Hantoute**[2] · **Marco A. López**[2,3]

Received: 12 July 2023 / Accepted: 21 February 2024 / Published online: 12 March 2024
© The Author(s) 2024

Abstract
We provide new characterizations of the ε-subdifferential of the supremum of an arbitrary family of convex functions. The resulting formulas only involve approximate subdifferentials of adequate convex combinations of the data functions. Families of convex functions with a concavity-like property are introduced and their relationship with affine models is studied. The role of the lower semi-continuity of the data functions is also analyzed.

Keywords Fenchel conjugation · Concavity-like · ε-Subdifferential · Supremum function

Mathematics Subject Classification 26B05 · 26J25 · 49H05

1 Introduction

There have been many works that analyze the subdifferential and ε-subdifferential of the supremum of convex functions in the setting of locally convex spaces. The first ones were established in the late 1960's, during the emergence of convex analysis ([1, 15, 21], etc.). The subject has also known a growing interest during the last decades, where many general results have been established (see, for instance, [7, 9, 11, 14, 16, 17, 19, 22], and references therein).

However, there is one useful result in ([2]), but not very well-known, that falls outside the scope of the previous results and, to our knowledge, has not been adequately studied. This is our purpose in this paper as well as to extend it. More precisely, in the case of finitely many proper convex functions $f_1, \ldots, f_m : X \to \mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$, $m \geq 1$, defined on a

✉ A. Hantoute
hantoute@ua.es

R. Correa
rcorrea@dim.uchile.cl

M.A. López
marco.antonio@ua.es

1 DIM-CMM, Universidad de Chile, Santiago de Chile, Chile

2 Department of Mathematics, Universidad de Alicante, Alicante, Spain

3 Federation University, Ballarat, Australia

locally convex space X, for every x within the effective domain of the maximum function $f := \max_{1\leq k\leq m} f_k$, and for every $\varepsilon \geq 0$, the ε-subdifferential of f at x verifies

$$\partial_\varepsilon f(x) = \bigcup_{\lambda\in\Delta_m} \partial_{(\varepsilon+g_\lambda(x)-f(x))} g_\lambda(x), \tag{1}$$

where $g_\lambda := \sum_{1\leq k\leq m} \lambda_k f_k$ and Δ_m is the canonical simplex in $\mathbb{R}^m$. When $\varepsilon = 0$ this formula reduces to

$$\partial f(x) = \bigcup \{\partial g_\lambda(x) : \lambda \in \Delta_m,\ g_\lambda(x) = f(x)\}.$$

This result is obtained in Corollary 15 as a consequence of our general approach. These two formulas cannot be obtained from the previous works on the supremum because, on the one hand, the convex functions involved in (1) are not necessarily lower semi-continuous and, on the other hand, the right-hand side of (1) involves approximate subdifferentials with parameters that do not exceed ε. These two features are not covered by the works cited before, for which the lower semi-continuity (or a close lower semi-continuity-like property) is critical and cannot be removed in general. Moreover, most of the previous characterizations of $\partial_\varepsilon f(x)$ involve approximate subdifferentials of data functions with parameters greater than ε.

The main objective of this paper is to develop a theory that extends formula (1) to families with infinitely many convex functions $f_t : X \to \mathbb{R}_\infty$, $t \in T$. For this purpose, we first provide new formulas for the approximate subdifferential of the supremum of a family of affine functions (Theorem 4). These formulas are crucial when dealing with families of convex functions that are concave-like (see Definition 1 and Theorem 7).

The extension from affine to families of convex functions satisfying the proposed concavity-like property is performed by means of Moreau's theorem on the representation of the biconjugate function (e.g., [7]). General counterparts of these formulas for families not satisfying the concave-like property are derived in Theorem 12 by adding to the original family those functions which are finite convex combinations of the form $\sum_{t\in T} \lambda_t f_t$.

All these results are obtained under a lower semi-continuity-like condition (20), which has been extensively used in convex subdifferential calculus (e.g., [7, 9, 11], etc). However, we show in Corollary 15 that such a closedness condition is not necessary in the case of finite families; in fact, in this case, condition (20) is automatically satisfied by the augmented family $\left\{g_\lambda := \sum_{t\in T} \lambda_t f_t, \lambda \in \Delta_m,\ \sum_{t\in T} \lambda_t = 1\right\}$, where $m = |T|$, which has the same supremum as the original one.

Some of the results presented here, notably in Theorem 12, improve similar formulas in [18], by removing lower semi-continuity conditions on data functions and dropping out the normal cone to the effective domain of the supremum function (or replacing it with smaller sets). Finally, to illustrate the scope of our results, a general approximate KKT condition for a convex optimization problem is provided in Corollary 16 without resorting to the Slater condition or any other standard constraint qualification.

The paper begins with Section 2 which provides the necessary notation and preliminary results. Section 3 is devoted to affine families, where new characterizations of the ε-subdifferential of the supremum of these families are given. The results of Section 4, dealing with concave-like families, are at an intermediate level of generality between the affine framework and the general scenario, which is developed in Section 5 where the main result is Theorem 12.

2 Preliminaries

This section contains the background material, including the notation, that will be used throughout the paper. Given a real locally convex space (lcs, for short) X, its topological dual X^* is endowed with the w^*-topology. The zero-vector in any linear space is θ, and the family of closed convex balanced neighborhoods of θ is denoted $\mathcal{N}_X$.

We represente $\mathbb{R}_+ := [0, +\infty[$, $\mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$, $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$, and adopt the convention $(+\infty) + (-\infty) = (-\infty) + (+\infty) = +\infty$ and $0 \cdot (+\infty) = +\infty$. Given a nonempty set T, we denote the cardinal of T by $|T|$. The set $\mathbb{R}_+^T$ is the cone of functions from T to $\mathbb{R}_+$. The support of a function $\lambda \in \mathbb{R}_+^T$ is $\operatorname{supp} \lambda := \{t \in T : \ \lambda_t \neq 0\}$, $\mathbb{R}_+^{(T)} := \{\lambda \in \mathbb{R}_+^T : \operatorname{supp} \lambda \text{ is finite}\}$, and

$$\Delta(T) := \left\{ \lambda \in \mathbb{R}_+^{(T)} : \sum_{t \in T} \lambda_t = 1 \right\}. \tag{2}$$

In particular, for $m \geq 1$ we define

$$\Delta_m := \left\{ \lambda \in \mathbb{R}_+^m : \sum_{1 \leq i \leq m} \lambda_i = 1 \right\} \text{ and } \Delta_m^* := \{\lambda \in \Delta_m : \lambda_i > 0 \ \forall i\}.$$

The algebraic sum of two sets A and B in X (or in X^*) is

$$A + B := \{a + b : \ a \in A, \ b \in B\}, \quad A + \emptyset = \emptyset + A = \emptyset. \tag{3}$$

The sets $\operatorname{co} A$ and $\overline{\operatorname{co}} A$ refer to the convex and the closed convex hulls of A, respectively, while $\operatorname{cl} A$ and $\overline{A}$ are indistinctly representing the closure of A (w^*-closure if $A \subset X^*$). Given a function $f : X \longrightarrow \overline{\mathbb{R}}$, $\operatorname{dom} f := \{x \in X : \ f(x) < +\infty\}$ and $\operatorname{epi} f := \{(x, \lambda) \in X \times \mathbb{R} : \ f(x) \leq \lambda\}$ are, respectively, its (effective) domain and epigraph. Given $\alpha \in \mathbb{R}$, we denote $[f \leq \alpha] := \{x \in X : f(x) \leq \alpha\}$; the set $[f < \alpha]$ is defined similarly. We write $f \in \Gamma_0(X)$ if f is proper, convex and lsc; that is, $\operatorname{dom} f \neq \emptyset$ and $f > -\infty$, $\operatorname{epi} f$ is convex and closed. Equivalently, f is lsc at x if it matches its closed hull $\bar{f}$ at x, where $\operatorname{epi}(\bar{f}) = \operatorname{cl}(\operatorname{epi} f)$. Given $\varepsilon \in \mathbb{R}$, the ε-subdifferential of f at $x \in X$ is the set

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

with $\partial_\varepsilon f(x) := \emptyset$ if $x \notin \operatorname{dom} f$ or $\varepsilon < 0$. The elements of $\partial_\varepsilon f(x)$ are called ε-subgradients of f at x, and the subdifferential of f at x is $\partial f(x) := \partial_0 f(x)$. The ε-normal set to a set $A \subset X$ at x is $\mathrm{N}_A^\varepsilon(x) := \partial_\varepsilon \mathrm{I}_A(x)$, where I_A is the indicator function of A. If $f : X \to \mathbb{R}_\infty$ is convex, then $\partial_\varepsilon f(x) \neq \emptyset$ provided that $\varepsilon > 0$ and f is lsc at $x \in \operatorname{dom} f$. However, if f is not lsc at x, then

$$\partial_\alpha f(x) = \emptyset, \text{ for all } \alpha \in [0, f(x) - \overline{f}(x)[; \tag{4}$$

otherwise, there would exist $x_\alpha^* \in X^*$ such that

$$\left\langle x_\alpha^*, y - x \right\rangle \leq f(y) - f(x) + \alpha, \text{ for all } y \in X;$$

hence, $\overline{f}(x) \in \mathbb{R}$. By taking the closure on both sides we arrive at the contradiction $f(x) - \overline{f}(x) \leq \alpha$. The relation in (4) is also true if $\overline{f}(x) = -\infty$.

The following relations will also be used in the paper: for every $x \in X$ and $\varepsilon \geq 0$, we have

$$\partial_{\varepsilon} f(x) = \bigcap_{\delta > \varepsilon} \partial_{\delta} f(x), \tag{5}$$

$$\partial_{\varepsilon} f(x) \subset \partial_{\varepsilon + \bar{f}(x) - f(x)} \overline{f}(x) \subset \partial_{\varepsilon} \bar{f}(x), \tag{6}$$

and, if $g : X \to \overline{\mathbb{R}}$ is another function such that $f \leq g$ and $x \in \operatorname{dom} g$, then

$$\partial_{\varepsilon} f(x) \subset \partial_{\varepsilon + g(x) - f(x)} g(x). \tag{7}$$

The following property gives an inner approximation of the ε-subdifferential ([7, Proposition 4.1.9] and [12]).

Lemma 1 *Given a convex function $f : X \to \mathbb{R}_{\infty}$, $x \in \operatorname{dom} f$ and $\varepsilon > 0$, we have that*

$$\partial_{\varepsilon} f(x) = \operatorname{cl}\left(\bigcup_{0 < \delta < \varepsilon} \partial_{\delta} f(x) \right),$$

provided that the last set is nonempty.

The Fenchel conjugate of the function $f : X \to \overline{\mathbb{R}}$ is the function $f^* : X^* \to \overline{\mathbb{R}}$ defined by $f^*(x^*) := \sup_{x \in X} \{\langle x^*, x \rangle - f(x)\}$. Hence, $x^* \in \partial_{\varepsilon} f(x)$ if and only if $f(x) + f^*(x^*) \leq \langle x^*, x \rangle + \varepsilon$. An example of conjugate functions is the support function of a set $C \subset X^*$ which is $\sigma_C = (\mathrm{I}_C)^*$, where I_C is the indicator function; i.e., $\mathrm{I}_C(x) = 0$ if $x \in C$, and $\mathrm{I}_C(x) = +\infty$ otherwise. Then, the Moreau theorem that establishes that the biconjugate of a function having a continuous affine minorant coincides with its closed convex hull, implies that

$$(\sigma_C)^* = ((\mathrm{I}_C)^*)^* = \mathrm{I}_{\overline{\mathrm{co}}(C)}.$$

Consequently, we obtain the expression of the ε-subdifferential of σ_C.

Proposition 2 *If $C \subset X^*$ is a non-empty set, then*

$$\partial_{\varepsilon} \sigma_C(x) = \left\{ x^* \in \overline{\mathrm{co}}(C) : \left\langle x^*, x \right\rangle \geq \sigma_C(x) - \varepsilon \right\}, \text{ for all } x \in X \text{ and } \varepsilon \geq 0.$$

3 Supremum of Affine Functions

In this section, we extend Proposition 2 to the larger family of suprema of affine functions. Given a non-empty set $C \subset X^* \times \mathbb{R}$, Theorem 4 below provides different characterizations of the ε-subdifferential of the supremum function $\phi : X \to \mathbb{R}_{\infty}$ given by

$$\phi(\cdot) := \sup_{(x^*, \beta) \in C} (\left\langle x^*, \cdot \right\rangle - \beta) = \sigma_C(\cdot, -1) = \sigma_{\overline{\mathrm{co}} C}(\cdot, -1). \tag{8}$$

A finite-dimensional version of Theorem 4 has been given in [9] when $\varepsilon = 0$. We point out that the following result cannot be derived, at least directly, from subdifferential calculus rules of the supremum such us those established for example in [5, 6, 11, 16].

Given $x \in \operatorname{dom}\phi$ and $\varepsilon \geq 0$, we introduce the set

$$C_\varepsilon(x) := \{(x^*, \beta) \in \overline{\text{co}}(C) : \langle x^*, x\rangle - \beta \geq \phi(x) - \varepsilon\}, \tag{9}$$

which corresponds to the set of points in $\overline{\text{co}}(C)$ where the ε-supremum in ϕ is attained. Similarly, we have

Lemma 3 *Given $x \in \operatorname{dom}\phi$ and $\varepsilon \geq 0$, we have*

$$C_\varepsilon(x) = \{(x^*, \beta) \in \overline{\text{co}}(C) : \phi(x) \geq \langle x^*, x\rangle - \beta \geq \phi(x) - \varepsilon\}. \tag{10}$$

Proof The inclusion "$\supset$" is clear, whereas the inclusion "$\subset$" is a straightforward consequence of the third equality in (8). $\square$

Theorem 4 *Let ϕ be the function in (8). Then, for every $x \in \operatorname{dom}\phi$ and $\varepsilon \geq 0$, we have*

$$\partial_\varepsilon \phi(x) = \left\{x^* \in X^* : (x^*, \beta) \in C_\varepsilon(x) \text{ for some } \beta \in \mathbb{R}\right\} \tag{11}$$

and, whenever $\varepsilon > 0$,

$$\partial_\varepsilon \phi(x) = \text{cl}\left\{x^* \in X^* : (x^*, \beta) \in \text{co}(C),\ \phi(x) \geq \langle x^*, x\rangle - \beta \geq \phi(x) - \varepsilon\right\}. \tag{12}$$

Proof Fix $x \in \operatorname{dom}\phi$ and $\varepsilon > 0$. We introduce the continuous linear mapping $A_0 : X \to X \times \mathbb{R}$ defined by $A_0 z := (z, 0)$ whose adjoint is the continuous linear mapping $A_0^* : X^* \times \mathbb{R} \to X^*$ given by $A_0^*(z^*, \beta) = z^*$. Then, since $\phi = \sigma_C \circ (A_0 + (\theta, -1))$, we get (e.g., [13])

$$\partial_\varepsilon \phi(x) = \text{cl}(A_0^*(\partial_\varepsilon \sigma_C(x, -1))),$$

and Proposition 2 implies that

$$\partial_\varepsilon \phi(x) = \text{cl}\{x^* \in X^* : (x^*, \beta) \in C_\varepsilon(x) \text{ for some } \beta \in \mathbb{R}\}.$$

Hence, to establish (11), we only need to show that the set

$$B_\varepsilon := \left\{x^* \in X^* : (x^*, \beta) \in C_\varepsilon(x) \text{ for some } \beta \in \mathbb{R}\right\}$$

is closed. To this aim, we pick a net $(x_i^*)_i \subset B_\varepsilon$ that converges to some $x^* \in X^*$, and choose another net $(\beta_i)_i \subset \mathbb{R}$ such that $(x_i^*, \beta_i)_i \subset C_\varepsilon(x)$; that is, taking into account Lemma 3,

$$(x_i^*, \beta_i) \in \overline{\text{co}}(C) \text{ and } \phi(x) \geq \langle x_i^*, x\rangle - \beta_i \geq \phi(x) - \varepsilon, \text{ for all } i. \tag{13}$$

Hence, as $(x_i^*)_i$ converges, we may assume without loss of generality that $(\beta_i)_i$ converges to some β such that $(x^*, \beta) \in \overline{\text{co}}(C)$ and $\langle x^*, x\rangle - \beta \geq \phi(x) - \varepsilon$; in other words, $x^* \in B_\varepsilon$ and this set is closed. Then, (11) has been proved for $\varepsilon > 0$.

We consider now $\varepsilon = 0$. From (5) and the paragraph above we obtain

$$\partial \phi(x) = \bigcap_{\delta > 0} \partial_\delta \phi(x) = \bigcap_{\delta > 0} B_\delta.$$

Thus, given any $x^* \in \partial\phi(x)$, for every $\delta > 0$ there exists some $\beta_\delta \in \mathbb{R}$ such that $(x^*, \beta_\delta) \in \overline{\text{co}}(C)$ and $\phi(x) \geq \langle x^*, x\rangle - \beta_\delta \geq \phi(x) - \delta$, according to Lemma 3. In particular, as $\delta \downarrow 0$, we get $\beta_\delta \to \beta_0 := \langle x^*, x\rangle - \phi(x)$ and $(x^*, \beta_0) \in \overline{\text{co}}(C)$; that is, $x^* \in \{x^* \in X^* : (x^*, \beta) \in C_0(x) \text{ for some } \beta \in \mathbb{R}\}$ and the inclusion "$\subset$" in (11) follows.

To prove the converse inclusion we take

$$x^* \in \{x^* \in X^* : \ (x^*, \beta) \in C_0(x) \text{ for some } \beta \in \mathbb{R}.$$

Then, there exists some $\beta \in \mathbb{R}$ such that $(x^*, \beta) \in \overline{\text{co}}(C)$ and $\langle x^*, x\rangle - \beta = \langle (x, -1), (x^*, \beta)\rangle = \phi(x)$. Hence, using Proposition 2, $(x^*, \beta) \in \partial\sigma_C(x, -1) = \partial\sigma_C(A_0 x + (\theta, -1))$ and we deduce that

$$\begin{aligned} x^* = A_0^*(x^*, \beta) &\in A_0^*\partial\sigma_C(A_0 x + (\theta, -1)) \\ &\subset \partial\left(\sigma_C \circ (A_0 + (\theta, -1))\right)(x) = \partial\phi(x). \end{aligned}$$

To establish (12) we first claim that, for all $x \in \text{dom}\,\phi$ and $\varepsilon > 0$,

$$\partial_\varepsilon\phi(x) = \bigcap_{\delta > \varepsilon} \text{cl}\,(E_\delta)\,, \tag{14}$$

where

$$E_\delta := \left\{x^* \in X^* : (x^*, \beta) \in \text{co}(C),\ \phi(x) \geq \left\langle x^*, x\right\rangle - \beta \geq \phi(x) - \delta,\ \beta \in \mathbb{R}\right\}.$$

Indeed, given $\delta > \delta_1 > \varepsilon > 0$ and $x^* \in \partial_\varepsilon\phi(x) \subset \partial_{\delta_1}\phi(x)$, (11) gives rise to some $\beta \in \mathbb{R}$ such that $(x^*, \beta) \in \overline{\text{co}}(C)$ and $\langle x^*, x\rangle - \beta \geq \phi(x) - \delta_1$; that is,

$$\left\langle x^*, x\right\rangle - \beta > \phi(x) - \delta.$$

This yields the existence of a net $(x_i^*, \beta_i) \in \text{co}(C)$ converging to (x^*, β) such that $\left\langle x_i^*, x\right\rangle - \beta_i > \phi(x) - \delta$; that is, $x_i^* \in E_\delta$. In other words, $x^* \in \text{cl}\,(E_\delta)$ and the arbitrariness of $\delta > \varepsilon$ entails the inclusion "$\subset$" in (14). The claim is proved because the opposite inclusion "$\supset$" in (14) easily follows from (11) as a consequence of the following inclusions,

$$\bigcap_{\delta > \varepsilon} \text{cl}\,(E_\delta) \subset \bigcap_{\delta > \varepsilon} \text{cl}\,(B_\delta) = \bigcap_{\delta > \varepsilon} \text{cl}\,(\partial_\delta\phi(x)) = \bigcap_{\delta > \varepsilon} \partial_\delta\phi(x) = \partial_\varepsilon\phi(x).$$

Now, taking into account Lemma 1 and the fact that $\phi \in \Gamma_0(X)$, (14) and (11) give rise to (12):

$$\begin{aligned} \partial_\varepsilon\phi(x) &= \text{cl}\left(\bigcup_{0<\gamma<\varepsilon} \partial_\gamma\phi(x)\right) = \text{cl}\left(\bigcup_{0<\gamma<\varepsilon}\ \bigcap_{\delta>\gamma} \text{cl}\,(E_\delta)\right) \\ &\subset \text{cl}\,(E_\varepsilon) \subset \text{cl}\,(B_\varepsilon) = \partial_\varepsilon\phi(x). \end{aligned}$$

□

The following example illustrates Theorem 4.

Example 1 Consider the function $\phi : \mathbb{R} \to \mathbb{R}_\infty$ defined by

$$\phi(x) := \sup_{t>0}\,(tx - 1/t)\,,$$

which is represented as in (8) with $C = \{(t, 1/t) : \ t > 0\}$.

Then, $\phi(0) = 0$ and (11) gives rise to

$$\partial\phi(0) = \{\alpha \in \mathbb{R} : (\alpha, \beta) \in \overline{\text{co}}(C),\ \beta = 0\} = \{\alpha \in \mathbb{R} : (\alpha, 0) \in \overline{\text{co}}(C)\}\,,$$

where $C = \{(t, 1/t) : t > 0\}$. This entails that $\partial\phi(0) = \emptyset$. Now, when $\varepsilon > 0$, (11) reads

$$\partial_\varepsilon \phi(0) = \{\alpha > 0 : 1/\alpha \leq \beta \leq \varepsilon\} = [1/\varepsilon, +\infty[.$$

Observe that straightforward calculus give $\phi(x) = -2\sqrt{-x} + \mathrm{I}_{\mathbb{R}_-}(x)$.

The second example shows that (12) cannot be extended to the case $\varepsilon = 0$.

Example 2 Consider the function $\phi : \mathbb{R} \to \mathbb{R}_\infty$ defined by

$$\phi(x) := \max\left\{\sup_{t>0}(tx - 1/t), 0\right\}.$$

So, $\phi(0) = 0$ and ϕ is written in the form $\phi(x) := \sigma_C(x, -1)$, where

$$C := \{(t, 1/t), t > 0; (0, 0)\}.$$

In this case, we have

$$\mathrm{co}(C) = \{(0, 0)\} \cup \{(\alpha, \beta) : \alpha, \beta > 0\} \text{ and } \overline{\mathrm{co}}(C) = \{(\alpha, \beta) : \alpha, \beta \geq 0\}.$$

Consequently, (12) reads

$$\partial_\varepsilon \phi(0) = \mathrm{cl}\,\{\alpha \in \mathbb{R} : (\alpha, \beta) \in \mathrm{co}(C),\ 0 \leq \beta \leq \varepsilon\} = \mathbb{R}_+, \text{ for all } \varepsilon > 0.$$

However, this last relation does not hold when $\varepsilon = 0$ because, thanks to formula (11), we have

$$\partial\phi(0) = \{\alpha \in \mathbb{R} : (\alpha, \beta) \in \overline{\mathrm{co}}(C),\ \beta = 0\} = \mathbb{R}_+,$$

while $\mathrm{cl}\,\{\alpha \in \mathbb{R} : (\alpha, \beta) \in \mathrm{co}(C),\ \beta = 0\} = \{0\} \subsetneqq \partial\phi(0)$.

The following corollary, which constitutes an improvement of [10, Lemma 1] (see, also, [20]), is given for the purpose of illustrating the scope of Theorem 4. Then formula (15) below can be seen as an extension of the well-known result for functions f in $\Gamma_0(X)$:

$$\partial_\varepsilon f^* = (\partial_\varepsilon f)^{-1}, \text{ for all } \varepsilon \geq 0.$$

Corollary 5 *Given a function $f : X \to \mathbb{R}_\infty$, for every $x^* \in X^*$ and $\varepsilon > 0$ we have*

$$\partial_\varepsilon f^*(x^*) = \mathrm{cl}\left\{\sum_{x\in\mathrm{dom}\, f} \lambda_x (\partial_{\varepsilon_x} f)^{-1}(x^*) :\ \lambda \in \Delta(\mathrm{dom}\, f),\ \sum_{x\in\mathrm{dom}\, f} \lambda_x \varepsilon_x \leq \varepsilon,\ \varepsilon_x \geq 0\right\}. \quad (15)$$

In particular, if f is convex, then

$$\partial_\varepsilon f^*(x^*) = \mathrm{cl}((\partial_\varepsilon f)^{-1}(x^*)). \quad (16)$$

Proof First, we may additionally assume that f^* is proper (and so will be f); otherwise, (15) would hold trivially since all the sets involved will be empty. Next, we apply Theorem 4 with the set $C := \{(x, f(x)) \colon x \in \mathrm{dom}\, f\}$ and the supremum function

$$\phi(\cdot) := \sup_{(x,\beta)\in C} (\langle \cdot, x\rangle - \beta) = f^*(\cdot).$$

Then, for every given $x^* \in X^*$ and $\varepsilon > 0$, formula (12) yields

$$\partial_\varepsilon f^*(x^*) = \mathrm{cl}\left\{ \sum_{x\in \mathrm{dom}\, f} \lambda_x x : \lambda \in \Delta(\mathrm{dom}\, f),\ \sum_{x\in \mathrm{supp}\,\lambda} \lambda_x(\langle x^*, x\rangle - f(x)) \geq f^*(x^*) - \varepsilon \right\}.$$

Next, taking $\varepsilon_x := f(x) + f^*(x^*) - \langle x^*, x\rangle\ (\geq 0)$, we get $\sum_{x\in \mathrm{dom}\, f} \lambda_x \varepsilon_x \leq \varepsilon$ and $x \in (\partial_\varepsilon f)^{-1}(x^*)$, and the last expression is equivalent to (15).

If f is convex, then every $z := \sum_{x\in \mathrm{dom}\, f} \lambda_x x$, $\lambda \in \Delta(\mathrm{dom}\, f)$ and $\sum_{x\in \mathrm{supp}\,\lambda} \lambda_x \varepsilon_x \leq \varepsilon$, satisfies: for every $y \in \mathrm{dom}\, f$,

$$\begin{aligned} \langle x^*, y - z\rangle &= \sum_{x\in \mathrm{dom}\, f} \lambda_x \langle x^*, y - x\rangle \\ &\leq \sum_{x\in \mathrm{dom}\, f} \lambda_x (f(y) - f(x) + \varepsilon_x) \leq f(y) - f(z) + \varepsilon; \end{aligned}$$

that is, $z \in (\partial_\varepsilon f)^{-1}(x^*)$. Therefore $\sum_{x\in \mathrm{dom}\, f} \lambda_x (\partial_{\varepsilon_x} f)^{-1}(x^*) \subset (\partial_\varepsilon f)^{-1}(x^*)$, and (15) implies the inclusion "$\subset$" in (16). Finally, (16) follows because the opposite inclusion "$\supset$" is straightforward. □

4 Concave-Like Setting

The concept of concave-like has been recognized as a relaxation of concavity of mappings, namely in relation to minimax theorems ([8]). Here, abusing the language, we adopt it to families of functions.

Definition 1 A family of functions $\{f_t : X \to \overline{\mathbb{R}},\ t \in T\}$ is said to be concave-like if each convex combination of its elements is dominated by a member of the family; that is, for each $\lambda \in \Delta(T)$, there exists some $s \in T$ such that

$$\sum_{t\in \mathrm{supp}\,\lambda} \lambda_t f_t \leq f_s. \tag{17}$$

Typical examples of concave-like families are:

(i) $\{\langle a, \cdot\rangle - b,\ (a, b) \in C\}$, where C is convex.

(ii) $\{f_t : X \to \overline{\mathbb{R}}, t \in T\}$, where T is convex and the mappings $t \mapsto f_t(x)$, $x \in X$, are concave.

(iii) $\{f_t : X \to \overline{\mathbb{R}}, t \in T\}$, where $(T, \preccurlyeq)$ is an ordered set and the net $(f_t)_{t\in T}$ is non-decreasing.

(iv) Also, if $\{f_t : X \to \overline{\mathbb{R}}, t \in T\}$ is concave-like, then the families $\{\bar{f}_t, t \in T\}$ and $\{\bar{f}_t : t \in S\}$, $S := \{t \in T : \bar{f}_t > -\infty\}$, are also concave-like. Actually, given any $\lambda \in \Delta(T)$, there exists some $s \in T$ such that

$$f_\lambda := \sum_{t\in \mathrm{supp}\,\lambda} \lambda_t f_t \leq f_s.$$

Then

$$\sum_{t\in \mathrm{supp}\,\lambda} \lambda_t \bar{f}_t \leq \sum_{t\in \mathrm{supp}\,\lambda} \lambda_t f_t \leq f_s, \tag{18}$$

and we conclude that $\sum_{t\in\operatorname{supp}\lambda}\lambda_t \bar{f}_t \leq \bar{f}_s$. If $\lambda \in \Delta(S)$ $(\subset \Delta(T))$ and s is as above, then (18) implies that $s \in S$, and the family $\{\bar{f}_t : t \in S\}$ is concave-like too.

Actually, every family of functions can be enlarged to a concave-like family without changing the associated supremum:

Lemma 6 *Given functions* $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, *the family*

$$\left\{ f_\lambda := \sum_{t\in\operatorname{supp}\lambda} \lambda_t f_t,\ \lambda \in \Delta(T) \right\}$$

is always concave-like, and we have that $\sup_{t\in T} f_t = \sup_{\lambda\in\Delta(T)} f_\lambda$.

Proof Given any $\mu \in \Delta(\Delta(T))$, it can be easily checked that the function

$$\sum_{\lambda\in\operatorname{supp}\mu} \mu_\lambda f_\lambda = \sum_{\lambda\in\operatorname{supp}\mu,\ t\in\operatorname{supp}\lambda} \mu_\lambda \lambda_t f_t$$

has the form f_{λ_0} for some $\lambda_0 \in \Delta(T)$. Finally, the equality $\sup_{t\in T} f_t = \sup_{\lambda\in\Delta(T)} f_\lambda$ is straightforward. □

Theorem 7 below characterizes the ε-subdifferential of the supremum $f := \sup_{t\in T} f_t$ of concave-like families of convex functions. The formula given there distinguishes between the role played by proper and improper closures (of the f_t's). While $\partial_\varepsilon f(x)$ strongly depends on the functions indexed in

$$T^p := \{t \in T : \bar{f}_t \in \Gamma_0(X)\}, \tag{19}$$

the rest of functions having improper closures, indexed in $T^i := T \setminus T^p$, only intervene through their effective domains.

Theorem 7 *Given convex functions* $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, *and* $f := \sup_{t\in T} f_t$, *we suppose that the family* $\{f_t, t \in T^p\}$ *is concave-like and*

$$\bar{f} = \sup_{t\in T} \bar{f}_t. \tag{20}$$

Then, for all $x \in \operatorname{dom} f$ *and* $\varepsilon > 0$, *we have*

$$\partial_\varepsilon f(x) = \operatorname{cl}\left\{ \bigcup_{t\in T^p} \partial_{\varepsilon+f_t(x)-f(x)}(f_t + \mathrm{I}_D)(x) \right\}, \tag{21}$$

where

$$D := \bigcap_{t\in T^i} \operatorname{cl}(\operatorname{dom} f_t)\ (= X \text{ if } T^i = \emptyset).$$

Conversely, (21) *also implies* (20) *provided that* f^* *is proper.*

Proof We proceed by steps.

Step 1: The inclusion "$\supset$" in (21) always holds because, for every $x \in \operatorname{dom} f$ and $\varepsilon > 0$,

$$\bigcup_{t\in T^p} \partial_{\varepsilon+f_t(x)-f(x)}(f_t + \mathrm{I}_D)(x) \subset \partial_\varepsilon(f + \mathrm{I}_{\operatorname{dom} f})(x) = \partial_\varepsilon f(x).$$

Step 2: In this step, we prove (21) under the additional assumption that $\{f_t, t \in T\} \subset \Gamma_0(X)$, so that $T^p = T$ and (20) automatically holds. Then, according to Moreau Theorem, we have

$$f = \sup_{t\in T} f_t = \sup_{t\in T}(f_t^*)^* = \sup_{t\in T,\ x^*\in \mathrm{dom}\, f_t^*} \{\langle x^*, \cdot\rangle - f_t^*(x^*)\}$$

and Theorem 4 entails, for all $x \in \mathrm{dom}\, f$ and $\varepsilon > 0$,

$$\partial_\varepsilon f(x) = \mathrm{cl}\left\{x^* \in X^* : (x^*, \beta) \in \mathrm{co}\,(C)\,,\ \langle x^*, x\rangle - \beta \geq f(x) - \varepsilon\right\}, \tag{22}$$

where

$$C := \{(x^*, f_t^*(x^*)) : x^* \in \mathrm{dom}\, f_t^*,\ t \in T\}.$$

Take $(x^*, \beta) \in \mathrm{co}\,(C)$ such that $\langle x^*, x\rangle - \beta \geq f(x) - \varepsilon$. Then there are $x_i^* \in \mathrm{dom}\, f_{t_i}^*$ and $t_i \in T (i = 1, \ldots, k, k \geq 1)$ together with $\lambda \in \Delta_k^*$ such that

$$(x^*, \beta) = \sum_{1\leq i\leq k} \lambda_i (x_i^*, f_{t_i}^*(x_i^*))$$

and, using the Fenchel inequality,

$$\sum_{1\leq i\leq k} \lambda_i f_{t_i}(x) - \varepsilon \leq f(x) - \varepsilon \leq \sum_{1\leq i\leq k} \lambda_i \left(\langle x_i^*, x\rangle - f_{t_i}^*(x_i^*)\right) \leq \sum_{1\leq i\leq k} \lambda_i f_{t_i}(x) \leq f_{t_0}(x), \tag{23}$$

where $t_0 \in T$ comes from the concave-like assumption. In other words, for $\varepsilon_i := f_{t_i}(x) + f_{t_i}^*(x_i^*) - \langle x_i^*, x\rangle (\geq 0)$ the last inequalities yield

$$\begin{aligned}\sum_{1\leq i\leq k} \lambda_i\varepsilon_i + f_{t_0}(x) - \sum_{1\leq i\leq k} \lambda_i f_{t_i}(x) &= f_{t_0}(x) - \sum_{1\leq i\leq k} \lambda_i \left(\langle x_i^*, x\rangle - f_{t_i}^*(x_i^*)\right) \\ &\leq f_{t_0}(x) - \sum_{1\leq i\leq k} \lambda_i f_{t_i}(x) + \varepsilon \leq f_{t_0}(x) - f(x) + 2\varepsilon.\end{aligned} \tag{24}$$

Thus, since $x \in \mathrm{dom}\, f \subset \mathrm{dom}\, f_{t_0}$ and $\lambda_i > 0$ for $i = 1, \ldots, k$, we conclude from the definition of the ε_i's, that

$$x^* = \sum_{1\leq i\leq k} \lambda_i x_i^* \in \sum_{1\leq i\leq k} \lambda_i \partial_{\varepsilon_i} f_{t_i}(x) = \sum_{1\leq i\leq k} \partial_{\lambda_i\varepsilon_i}(\lambda_i f_{t_i})(x);$$

that is, taking into account (7), (24), and denoting $\varepsilon_\lambda := \Sigma_{1\leq i\leq k}\, \lambda_i\varepsilon_i$, $f_\lambda := \sum_{1\leq i\leq k}\lambda_i f_{t_i}$,

$$x^* \in \partial_{\varepsilon_\lambda} f_\lambda(x) \subset \partial_{\varepsilon_\lambda + f_{t_0}(x) - f_\lambda(x)} f_{t_0}(x) \subset \partial_{f_{t_0}(x) - f(x) + \varepsilon} f_{t_0}(x).$$

Therefore the inclusion "$\subset$" in (21) follows from (22), and we are done because the other inclusion "$\supset$" is easily checked.

Step 3: We prove in this step that (21) also holds under condition (20). In fact, we only need to verify the inclusion "$\subset$" in (21) in the case when $\partial_\varepsilon f(x) \neq \emptyset$. This implies $f(x) \leq \overline{f}(x) + \varepsilon \leq f(x) + \varepsilon < +\infty$ and we may assume, for simplicity, that $x = \theta$, $f(\theta) = 0$ and, *a fortiori*, that f and $\bar{f}$ are proper. More precisely, due to (20), we have that $\sup_{t\in T} \bar{f}_t(\theta) = \bar{f}(\theta) \in \mathbb{R}$ and some of the $\bar{f}_t$'s must be proper ($T^p \neq \emptyset$). Moreover, using (20), we have that

$$\bar{f} = \sup_{t\in T} \bar{f}_t = \sup_{t\in T^p} \left(\bar{f}_t + \mathrm{I}_D\right). \tag{25}$$

Thus, since the family $\{\bar{f}_t + \mathrm{I}_D,\ t \in T^p\}$ is concave-like (see (18)), by taking into account (6) and (7) the second step yields

$$\partial_\varepsilon f(\theta) \subset \partial_{(\varepsilon+\bar{f}(\theta))} \bar{f}(\theta) = \mathrm{cl}\left\{ \bigcup_{t\in T^p} \partial_{\varepsilon+\bar{f}(\theta)+\bar{f}_t(\theta)-\bar{f}(\theta)} \left(\bar{f}_t + \mathrm{I}_D\right)(\theta) \right\}$$

$$= \mathrm{cl}\left\{ \bigcup_{t\in T^p} \partial_{\varepsilon+\bar{f}_t(\theta)} (\bar{f}_t + \mathrm{I}_D)(\theta) \right\} \tag{26}$$

$$\subset \mathrm{cl}\left\{ \bigcup_{t\in T^p} \partial_{\varepsilon+f_t(\theta)} (f_t + \mathrm{I}_D)(\theta) \right\}, \tag{27}$$

and the desired inclusion follows. Hence, in view of the first step, (21) also follows in the current case.

Step 4: We prove in this step that (21) implies (20), under the additional assumption that f^* is proper. If (21) holds, then for all $x \in \operatorname{dom} f$ and $\varepsilon > 0$ we have

$$\partial_\varepsilon f(x) = \mathrm{cl}\left\{ \bigcup_{t\in T^p} \partial_{\varepsilon+f_t(x)-f(x)} (f_t + \mathrm{I}_D)(x) \right\} \subset \mathrm{cl}\left\{ \bigcup_{t\in T^p} \partial_{\varepsilon+\bar{f}_t(x)-f(x)} \left(\bar{f}_t + \mathrm{I}_D\right)(x) \right\}.$$

Thus, since

$$g := \sup_{t\in T} \bar{f}_t = \sup_{t\in T^p} \left(\bar{f}_t + \mathrm{I}_D\right) \le \bar{f} \le f,$$

we conclude, from (7), that

$$\partial_\varepsilon f(x) \subset \partial_{\varepsilon+g(x)-f(x)} g(x) \subset \partial_\varepsilon g(x), \text{ for all } x \in \operatorname{dom} f \text{ and } \varepsilon > 0. \tag{28}$$

Therefore there exists some constant c such that $\bar{f} - g = c$ (see, i.e., [3, Theorem 5.3]); that is, $c \ge 0$. Moreover, since f^* is proper, for each $\varepsilon > 0$ there exists some $x_\varepsilon \in \operatorname{dom} f$ such that $\partial_\varepsilon f(x_\varepsilon) \neq \emptyset$. Hence, (28) implies that $\partial_{\varepsilon+g(x_\varepsilon)-f(x_\varepsilon)} g(x_\varepsilon) \neq \emptyset$, which in turn yields

$$0 \le \varepsilon + g(x_\varepsilon) - f(x_\varepsilon) \le \varepsilon + g(x_\varepsilon) - \bar{f}(x_\varepsilon) = \varepsilon - c;$$

that is, $0 \le c \le \varepsilon$ and we deduce that $c = 0$ as $\varepsilon \downarrow 0$. □

The closure condition (20) not only holds for lsc functions but also in other natural situations as it is shown in [11]. Let us show now that every family of the form $\left\{\sum_{1\le k\le m} \lambda_k f_k,\ \lambda \in \Delta(T)\right\}$, where $f_1, \ldots, f_k : X \to \mathbb{R}_\infty$ are proper convex functions, satisfies condition (20).

Lemma 8 *Given proper convex functions $f_k : X \to \mathbb{R}_\infty$, $1 \le k \le m$, the family $\{g_\lambda := \sum_{1\le k\le m} \lambda_k f_k,\ \lambda \in \Delta_m\}$ satisfies condition (20).*

Proof We know by Lemma 6 that $\max_{1\le k\le m} f_k = \max_{\lambda\in\Delta_m} g_\lambda$, and so for every $z \in X$ we obtain

$$\bar{f}(z) = \sup_{U\in\mathcal{N}_X} \inf_{y\in U} f(z+y) = \sup_{U\in\mathcal{N}_X} \inf_{y\in X} \max_{\lambda\in\Delta_m} \left(g_\lambda(z+y) + \mathrm{I}_U(y)\right).$$

Thus, by the minimax theorem ([23, Theorem 2.10.2])

$$\bar{f}(z) = \sup_{U\in\mathcal{N}_X} \max_{\lambda\in\Delta_m} \inf_{y\in U} g_\lambda(z+y) = \max_{\lambda\in\Delta_m} \sup_{U\in\mathcal{N}} \inf_{y\in U} g_\lambda(z+y) = \max_{\lambda\in\Delta_m} \overline{g_\lambda}(z),$$

and condition (20) is satisfied by $\{g_\lambda, \lambda \in \Delta_m\}$. □

Some consequences of Theorem 7 come next. First, we give another useful representation of $\partial_\varepsilon f(x)$ which explicitly involves the approximate subdifferential of the data functions $f_t, t \in T^p$.

Corollary 9 *With the assumptions of Theorem 7 we have, for all $x \in \operatorname{dom} f$ and $\varepsilon > f(x) - \bar{f}(x)$,*

$$\partial_\varepsilon f(x) = \operatorname{cl}\left\{ \bigcup_{\alpha \geq 0,\ t \in T^p} \partial_{\varepsilon-\alpha+f_t(x)-f(x)} f_t(x) + \mathrm{N}_D^\alpha(x) \right\}, \tag{29}$$

where $D = \cap_{t\in T^i} \operatorname{cl}(\operatorname{dom} f_t)$ $(= X$ if $T^i = \emptyset)$.

Proof Fix $x \in \operatorname{dom} f$ and $\varepsilon > f(x) - \bar{f}(x)$ such that $\partial_\varepsilon f(x) \neq \emptyset$; hence, both f and $\bar{f}$ belong to $\Gamma_0(X)$. The combination of (26), (27) and (21) leads us to the following characterization of $\partial_\delta f(x)$, for all $\delta > 0$,

$$\partial_\delta f(x) = \operatorname{cl}\left\{ \bigcup_{t\in T^p} \partial_{\delta+\bar{f}_t(x)-f(x)} (\bar{f}_t + \mathrm{I}_D)(x) \right\}. \tag{30}$$

Now the condition $\varepsilon > f(x) - \bar{f}(x)$ implies that $\partial_\delta f(x) \neq \emptyset$ for some $0 < \delta < \varepsilon$. Indeed, choosing any $\delta > 0$ such that $\varepsilon - 2\delta > f(x) - \bar{f}(x) \geq 0$, any element $z^* \in \partial_\delta \bar{f}(x)$ (this set is nonempty because $\bar{f} \in \Gamma_0(X)$, as we observed above) satisfies

$$f(x) + f^*(z^*) + 2\delta - \varepsilon \leq \bar{f}(x) + f^*(z^*) = \bar{f}(x) + (\bar{f})^*(z^*) \leq \langle z^*, x\rangle + \delta.$$

Thus,

$$f(x) + f^*(z^*) \leq \langle z^*, x\rangle + \varepsilon - \delta,$$

and so $z^* \in \partial_{\varepsilon-\delta} f(x)$. Therefore, using Lemma 1, (30) reads

$$\partial_\varepsilon f(x) = \operatorname{cl}\left(\bigcup_{0<\delta<\varepsilon} \partial_\delta f(x) \right) = \operatorname{cl}\left(\bigcup_{0<\delta<\varepsilon} \operatorname{cl}\left\{ \bigcup_{t\in T^p} \partial_{\delta+\bar{f}_t(x)-f(x)} (\bar{f}_t + \mathrm{I}_D)(x) \right\} \right).$$

Furthermore, by adjusting the δ involved in the last union (we make it a little larger), we also write

$$\partial_\varepsilon f(x) \subset \operatorname{cl}\left(\bigcup_{t\in T^p,\ f(x)-\bar{f}_t(x)<\delta<\varepsilon} \partial_{\delta+\bar{f}_t(x)-f(x)} (\bar{f}_t + \mathrm{I}_D)(x) \right).$$

At this moment, applying the subdifferential sum rule from [13] (see, also, [4]), we get

$$\begin{aligned} \partial_\varepsilon f(x) &\subset \operatorname{cl}\left(\bigcup_{t\in T^p,\ f(x)-\bar{f}_t(x)<\delta<\varepsilon} \operatorname{cl}\left(\bigcup_{\alpha\geq 0} \partial_{\delta-\alpha+\bar{f}_t(x)-f(x)} \bar{f}_t(x) + \mathrm{N}_D^\alpha(x) \right) \right) \\ &\subset \operatorname{cl}\left(\bigcup_{t\in T^p,\ \alpha\geq 0} \partial_{\varepsilon-\alpha+\bar{f}_t(x)-f(x)} \bar{f}_t(x) + \mathrm{N}_D^\alpha(x) \right). \end{aligned}$$

Hence, the nontrivial inclusion in (29) follows as (see (7))

$$\partial_{\varepsilon-\alpha+\bar{f}_t(x)-f(x)}\bar{f}_t(x) \subset \partial_{\varepsilon-\alpha+f_t(x)-f(x)}f_t(x). \qquad \square$$

The following results gives an equivalent reformulation of $\partial_\varepsilon f(x)$ that highlights the role played by the ε-active indices at x,

$$T_\varepsilon(x) := \{s \in T^p : f_s(x) \geq f(x) - \varepsilon\},\ \varepsilon \geq 0.$$

Corollary 10 *With the assumptions of Theorem 7 we have, for all $x \in \operatorname{dom} f$ and $\varepsilon > f(x) - \bar{f}(x)$,*

$$\partial_\varepsilon f(x) = \operatorname{cl}\left\{\bigcup_{\substack{0\leq\gamma\leq\varepsilon \\ t\in T_{\varepsilon-\gamma}(x)}} \partial_\gamma(f_t + \mathrm{I}_D)(x)\right\} \tag{31}$$

$$= \operatorname{cl}\left\{\bigcup_{\substack{0\leq\alpha\leq\gamma\leq\varepsilon \\ t\in T_{\varepsilon-\gamma}(x)}} \partial_{\gamma-\alpha} f_t(x) + \mathrm{N}_D^\alpha(x)\right\}, \tag{32}$$

where $D = \cap_{t\in T^i} \operatorname{cl}(\operatorname{dom} f_t)$ $(= X$ if $T^i = \emptyset)$. In particular, we have

$$\partial f(x) = \bigcap_{\delta>0} \operatorname{cl}\left\{\bigcup_{t\in T_\delta(x)} \partial_\delta(f_t + \mathrm{I}_D)(x)\right\} \tag{33}$$

$$= \bigcap_{\delta>0} \operatorname{cl}\left\{\bigcup_{t\in T_{\delta-\alpha}(x),\ \alpha\geq 0} \partial_{\delta-\alpha} f_t(x) + \mathrm{N}_D^\alpha(x)\right\}. \tag{34}$$

Proof Relation (31) follows easily from Theorem 7, defining $\gamma := \varepsilon + f_t(x) - f(x)$ in (21). Similarly, (32) comes from (29). In addition, due to (5), (33) is a direct consequence of (31). Thus, only (34) remains to be checked. More specifically, since it suffices to verify the nontrivial inclusion "$\subset$", we may suppose that $\partial f(x) \neq \emptyset$, so that $\partial f(x) = \partial \bar{f}(x)$ and $f(x) = \bar{f}(x) \in \mathbb{R}$. Therefore, by (29),

$$\partial f(x) = \bigcap_{\varepsilon>0} \partial_\varepsilon f(x) = \bigcap_{\varepsilon > f(x)-\bar{f}(x)} \partial_\varepsilon f(x)$$

$$= \bigcap_{\varepsilon>0} \operatorname{cl}\left\{\bigcup_{\alpha\geq 0,\ t\in T^p} \partial_{\varepsilon-\alpha+f_t(x)-f(x)} f_t(x) + \mathrm{N}_D^\alpha(x)\right\},$$

and the desired formula follows by setting $\gamma := \varepsilon + f_t(x) - f(x)$ in the last expression. $\square$

The closure can be removed from (21) and subsequent formulas in the compact-continuous setting.

Corollary 11 *In addition to the assumptions of Theorem 7, suppose that T is compact and $\limsup_{t\to s} f_t(z) \leq f_s(z)$ for all $s \in T$ and all $z \in D := \cap_{t\in T^i} \operatorname{cl}(\operatorname{dom} f_t)$ $(= X$ if $T^i = \emptyset)$. Then, for all $x \in \operatorname{dom} f$ and $\varepsilon > 0$, we have*

$$\partial_\varepsilon f(x) = \bigcup_{t\in T} \partial_{\varepsilon+f_t(x)-f(x)}(f_t + \mathrm{I}_D)(x), \tag{35}$$

and, provided that $\varepsilon > f(x) - \bar{f}(x)$,

$$\partial_\varepsilon f(x) = \bigcup_{\alpha \geq 0,\ t \in T} \partial_{\varepsilon - \alpha + f_t(x) - f(x)} f_t(x) + \mathrm{N}_D^\alpha(x). \tag{36}$$

Proof We give the proof of the first equality, the second follows as in the previous corollaries. Fix $x \in \operatorname{dom} f$ and $\varepsilon > 0$, and take $x^* \in \partial_\varepsilon f(x)$. Then, according to Theorem 7, there are nets $x_j^* \to x^*$ and $(t_j)_j \subset T^p$ such that

$$x_j^* \in \partial_{\varepsilon + f_{t_j}(x) - f(x)}(f_{t_j} + \mathrm{I}_D)(x), \text{ for all } j;$$

hence, $\varepsilon + f_{t_j}(x) - f(x) \geq 0$. By the current compactness condition, we may assume without loss of generality that $t_j \to t$ for some $t \in T$. Thus, the upper semicontinuity assumption implies that $\varepsilon + f_t(x) - f(x) \geq 0$ (as $x \in \operatorname{dom} f \subset D$) and $x^* \in \partial_{\varepsilon + f_t(x) - f(x)}(f_t + \mathrm{I}_D)(x)$. This yields the nontrivial inclusion "$\subset$" in (35). □

We close this section by applying Theorem 7 to derive a characterization of the ε-normal set to sublevel sets. See [10, Corollary 7] for the case of lsc convex functions.

Example 3 Consider a convex function $f : X \to \overline{\mathbb{R}}$, having a proper conjugate and satisfying

$$\operatorname{cl}([f \leq 0]) = \left[\bar{f} \leq 0\right]; \tag{37}$$

this last property occurs if, for instance, $[f < 0] \neq \emptyset$. Then, for every $x \in [f \leq 0]$ and $\varepsilon > 0$, we have

$$\mathrm{N}_{[f \leq 0]}^\varepsilon(x) = \operatorname{cl}\left(\bigcup_{t > 0} t \partial_{\frac{\varepsilon}{t} + f(x)} f(x)\right). \tag{38}$$

In order to prove (38) we define the functions

$$f_t := tf,\ t \in T := \,]0, +\infty[,\ \text{and } g := \sup_{t > 0}(tf).$$

Obviously, $T^p = T$ and $\{f_t, t \in T\}$ is concave-like. Therefore, because $g = \mathrm{I}_{[f \leq 0]}$ and (20) holds (as a consequence of (37)), Theorem 7 (with $D = X$) entails

$$\mathrm{N}_{[f \leq 0]}^\varepsilon(x) = \partial_\varepsilon g(x) = \operatorname{cl}\left(\bigcup_{t > 0} \partial_{\varepsilon + tf(x)}(tf)(x)\right) = \operatorname{cl}\left(\bigcup_{t > 0} t \partial_{\frac{\varepsilon}{t} + f(x)} f(x)\right).$$

5 Epsilon-Subdifferential of Suprema

The results of the previous section are applied here to the extended family of funtions

$$f_\lambda := \sum_{t \in \operatorname{supp} \lambda} \lambda_t f_t,\ \lambda \in \Delta(T^p), \tag{39}$$

where $T^p = \{t \in T : \bar{f}_t \in \Gamma_0(X)\}$. This family, by construction, satisfies the concave-like property and has the same supremum function. We give next the main theorem.

Theorem 12 *Given convex functions $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, and $f := \sup_{t\in T} f_t$, we suppose that (20) holds. Then, for any $x \in \operatorname{dom} f$ and $\varepsilon > 0$, we have*

$$\partial_\varepsilon f(x) = \operatorname{cl}\left\{ \bigcup_{\lambda\in\Delta(T^p)} \partial_{\varepsilon+f_\lambda(x)-f(x)} \left(f_\lambda + \mathrm{I}_D\right)(x) \right\}, \tag{40}$$

where $D := \bigcap_{t\in T\setminus T^p} \operatorname{cl}(\operatorname{dom} f_t)$.

Proof We fix $x \in \operatorname{dom} f$ and $\varepsilon > 0$ and suppose, without loss of generality, that $\partial_\varepsilon f(x) \neq \emptyset$. Since (see (25) and Lemma 6)

$$\bar{f} = \sup_{t\in T^p} \left(\bar{f}_t + \mathrm{I}_D\right) = \sup_{\lambda\in\Delta(T^p)} \sum_{t\in\operatorname{supp}\lambda} \lambda_t \left(\bar{f}_t + \mathrm{I}_D\right)$$

and the family, in $\Gamma_0(X)$,

$$\left\{ g_\lambda := \sum_{t\in\operatorname{supp}\lambda} \lambda_t \left(\bar{f}_t + \mathrm{I}_D\right),\ \lambda \in \Delta(T^p) \right\},$$

is concave-like, Theorem 7 gives rise to

$$\begin{aligned}\partial_\varepsilon f(x) \subset \partial_{\varepsilon+\bar{f}(x)-f(x)} \bar{f}(x) &= \operatorname{cl}\left\{ \bigcup_{\lambda\in\Delta(T^p)} \partial_{\varepsilon+\bar{f}(x)-f(x)+g_\lambda(x)-\bar{f}(x)} g_\lambda(x) \right\} \\ &= \operatorname{cl}\left\{ \bigcup_{\lambda\in\Delta(T^p)} \partial_{\varepsilon+g_\lambda(x)-f(x)} g_\lambda(x) \right\}.\end{aligned}$$

Thus, denoting $\tilde{f}_\lambda := \sum_{t\in\operatorname{supp}\lambda} \lambda_t \bar{f}_t$ and using (7), we deduce

$$\begin{aligned}\partial_\varepsilon f(x) &\subset \operatorname{cl}\left\{ \bigcup_{\lambda\in\Delta(T^p)} \partial_{\varepsilon+\tilde{f}_\lambda(x)-f(x)} \left(\tilde{f}_\lambda + \mathrm{I}_D\right)(x) \right\} \\ &\subset \operatorname{cl}\left\{ \bigcup_{\lambda\in\Delta(T^p)} \partial_{\varepsilon+f_\lambda(x)-f(x)} \left(f_\lambda + \mathrm{I}_D\right)(x) \right\},\end{aligned}$$

and the nontrivial inclusion "$\subset$" in (40) holds. □

In order to give a finite-dimensional counterpart of our main theorem in the finite-dimensional setting we use the following lemma.

Lemma 13 *Given convex functions $f_1, \cdots, f_m : \mathbb{R}^n \longrightarrow \mathbb{R}_\infty$, $m \geq 1$, we denote $g_\lambda := \sum_{1\leq i\leq m} \lambda_i f_i$, $\lambda \in \Delta_m$. Then we have*

$$\inf_{\mathbb{R}^n} \max_{1\leq i\leq m} f_i = \max_{\lambda\in\Delta_m,\ |\operatorname{supp}\lambda|\leq n+1} \inf_{\mathbb{R}^n} g_\lambda,$$

and consequently, for every $x \in \mathbb{R}^n$, $\alpha \in \Delta_m$ and $\varepsilon \in \mathbb{R}$,

$$\partial_{\varepsilon+g_\alpha(x)} g_\alpha(x) \subset \bigcup_{\lambda\in\Delta_m,\ |\operatorname{supp}\lambda|\leq n+1} \partial_{\varepsilon+g_\lambda(x)} g_\lambda(x). \tag{41}$$

Proof The first conclusion is a consequence of Carathéodory lemma and the minimax theorem. To show the second statement, we fix $x \in \mathbb{R}^n$, $\alpha \in \Delta_m$ and $\varepsilon \in \mathbb{R}$. If $\theta \in \partial_{\varepsilon+g_\alpha(x)} g_\alpha(x)$, then by the first part of the lemma there exists some $\lambda \in \Delta_m$ such that $|\operatorname{supp}\lambda| \leq n+1$ and

$$-\varepsilon \leq \inf_{\mathbb{R}^n} g_\alpha \leq \inf_{\mathbb{R}^n} \max_{1\leq i\leq m} f_i = \inf_X g_\lambda;$$

that is, $\theta \in \partial_{\varepsilon+g_\lambda(x)} g_\lambda(x)$.

More generally, if $x^* \in \partial_{\varepsilon+g_\alpha(x)} g_\alpha(x)$ and $\tilde{g}_\alpha(\cdot) := \sum_{1\leq i\leq m} \alpha_i (f_i - \langle x^*, \cdot\rangle)$, then $\theta \in \partial_{\varepsilon+g_\alpha(x)} \tilde{g}_\alpha(x)$ and the paragraph above yields some $\lambda \in \Delta_m$ such that $|\operatorname{supp}\lambda| \leq n+1$ and

$$\theta \in \partial_{\varepsilon+g_\alpha(x)-\tilde{g}_\alpha(x)+\tilde{g}_\lambda(x)} \tilde{g}_\lambda(x) = \partial_{\varepsilon+\langle x^*,x\rangle+\tilde{g}_\lambda(x)} \tilde{g}_\lambda(x) = \partial_{\varepsilon+g_\lambda(x)} g_\lambda(x) - x^*,$$

where we also denoted $\tilde{g}_\lambda(\cdot) := \sum_{1\leq i\leq m} \lambda_i (f_i - \langle x^*, \cdot\rangle)$. Hence, $x^* \in \partial_{\varepsilon+g_\lambda(x)} g_\lambda(x)$ and the desired inclusion follows. □

Theorem 14 *Under the assumptions of Theorem 12, if $X = \mathbb{R}^n$ and $\varepsilon > 0$, then*

$$\partial_\varepsilon f(x) = \operatorname{cl}\left\{ \bigcup_{\lambda\in\Delta(T^p),\ |\operatorname{supp}\lambda|\leq n+1} \partial_{\varepsilon+f_\lambda(x)-f(x)} \left(f_\lambda + \mathrm{I}_D\right)(x) \right\}. \tag{42}$$

If additionally, T is compact and $\limsup_{s\to t} f_s(z) \leq f_t(z)$ for all $z \in D$, for $\varepsilon \geq 0$ we have

$$\partial_\varepsilon f(x) = \bigcup_{\lambda\in\Delta(T),\ |\operatorname{supp}\lambda|\leq n+1} \partial_{\varepsilon+f_\lambda(x)-f(x)} \left(f_\lambda + \mathrm{I}_{\cap_{t\in T} \operatorname{dom} f_t}\right)(x). \tag{43}$$

Proof Formula (42) follows by combining (40) and (41). In order to prove formula (43), we fix $x \in \operatorname{dom} f$, $\varepsilon \geq 0$ and take $x^* \in \partial_\varepsilon f(x)$. Then formula (42) yields some nets $\varepsilon_i \downarrow \varepsilon$ ($\varepsilon_i \equiv \varepsilon$ if $\varepsilon > 0$), $x_i^* \to x^*$ and $(\lambda_i)_i \subset \Delta(T^p)$ such that $|\operatorname{supp}\lambda_i| \leq n+1$ and

$$x_i^* \in \partial_{\varepsilon_i+f_{\lambda_i}(x)-f(x)} \left(f_{\lambda_i} + \mathrm{I}_D\right)(x), \text{ for all } i,$$

where $f_{\lambda_i} = \sum_{t\in\operatorname{supp}\lambda_i} \lambda_i(t) f_t$ (see (39)); that is,

$$\left\langle x_i^*, z-x\right\rangle \leq f_{\lambda_i}(z) - f_{\lambda_i}(x) + (\varepsilon_i + f_{\lambda_i}(x) - f(x)) \tag{44}$$

$$= f_{\lambda_i}(z) - f(x) + \varepsilon_i, \text{ for all } z \in D \text{ and all } i. \tag{45}$$

We consider the net $(t_{i,1}, \ldots, t_{i,n+1})_i$ such that $\operatorname{supp}\lambda_i \subset (t_{i,1}, \ldots, t_{i,n+1})$. Then, taking into account the current continuity-compactness assumptions, we may assume without loss of generality that for all $1 \leq j \leq n+1$

$$\lambda_i(t_{i,j}) \to \lambda_j \geq 0 \text{ and } t_{i,j} \to t_j \in T;$$

hence, $(\lambda_1, \ldots, \lambda_{n+1}) \in \Delta_{n+1}$. Let us define the function $\lambda \in \Delta(T)$ as $\lambda(t_j) := \lambda_j$ if $1 \leq j \leq n+1$, and $\lambda(t) := 0$ if $t \neq t_j$. Then, by passing to the limit on i in (44) we obtain

$$\left\langle x^*, z-x\right\rangle \leq \sum_{1\leq j\leq n+1} \lambda(t_j) f_t(z) - f(x) + \varepsilon, \text{ for all } z \in D,$$

which in turn implies, taking into account the convention $0 \cdot (+\infty) = +\infty$, that for all $z \in \cap_{t\in T} \operatorname{dom} f_t (\subset D)$

$$\left\langle x^*, z-x\right\rangle \leq f_\lambda(z) - f(x) + \varepsilon = f_\lambda(z) - f_\lambda(x) + f_\lambda(x) - f(x) + \varepsilon.$$

Thus, $x^* \in \partial_{\varepsilon+f_\lambda(x)-f(x)}\left(f_\lambda + \mathrm{I}_{\cap_{t\in T}\operatorname{dom} f_t}\right)(x)$ and the non-trivial inclusion "$\subset$" in (43) follows. $\square$

Remark 1 Theorem 12 covers Theorem 7. In fact, if the family $\{f_t, t \in T^p\}$ is concave-like, then for each $\lambda \in \Delta(T^p)$ there exists some $t \in T^p$ such that $f_\lambda \le f_t$, and (40) entails, for all $x \in \operatorname{dom} f$ and $\varepsilon > 0$,

$$\partial_\varepsilon f(x) \subset \operatorname{cl}\left\{\bigcup_{t\in T^p} \partial_{\varepsilon+f_t(x)-f(x)}\left(f_t + \mathrm{I}_D\right)(x)\right\}.$$

Formula in (21) follows because the opposite of the last inclusion always holds.

The case of the maximum of a finite family comes easily from Theorem 12.

Corollary 15 *Given a finite family of proper convex functions $f_k : X \to \mathbb{R}_\infty$, $k \in T := \{1, \cdots, m\}$, and $f = \max_{1\le k\le m} f_k$, for every $x \in \operatorname{dom} f$ and $\varepsilon \ge 0$ we have*

$$\partial_\varepsilon f(x) = \bigcup_{\lambda\in\Delta_m} \partial_{\varepsilon+g_\lambda(x)-f(x)} g_\lambda(x), \tag{46}$$

where $g_\lambda := \sum_{1\le k\le m} \lambda_k f_k$. In particular, we have that

$$\partial f(x) = \bigcup\left\{\partial g_\lambda(x) : \lambda \in \Delta_m,\ g_\lambda(x) = f(x)\right\}.$$

Proof According to Lemma 8, the family $\{g_\lambda, \lambda \in \Delta_m\}$ satisfies condition (20). Moreover, since that

$$\sum_{k\in\operatorname{supp}\mu} \mu_k g_{\lambda_k} \in \{g_\lambda,\ \lambda \in \Delta_m\},\ \text{for all } \mu \in \Delta_m,$$

and $f = \max_{\lambda\in\Delta_m} g_\lambda$, by applying Theorem 12 to the family $\{g_\lambda, \lambda \in \Delta_m\}$ we obtain for all $x \in \operatorname{dom} f$ and $\varepsilon > 0$

$$\partial_\varepsilon f(x) \subset \operatorname{cl}\left\{\bigcup_{\lambda\in\Delta_m} \partial_{\varepsilon+g_\lambda(x)-f(x)}\left(g_\lambda + \mathrm{I}_D\right)(x)\right\} = \operatorname{cl}\left\{\bigcup_{\lambda\in\Delta_m} \partial_{\varepsilon+g_\lambda(x)-f(x)} g_\lambda(x)\right\};$$

the last equality holds because we have (due to convention $0 \cdot (+\infty) = +\infty$, if $x \notin D$ then $g_\lambda(x) = +\infty$)

$$g_\lambda + \mathrm{I}_D = \sum_{1\le k\le m} \lambda_k f_k + \sum_{1\le k\le m,\ \overline{f}_k \text{ improper}} \mathrm{I}_{\operatorname{cl}(\operatorname{dom} f_k)} = g_\lambda.$$

Therefore, for every $x^* \in \partial_\varepsilon f(x)$, there exist nets $x_i^* \to x^*$ and $(\lambda_i) \subset \Delta_m$ such that $x_i^* \in \partial_{\varepsilon+g_{\lambda_i}(x)-f(x)} g_{\lambda_i}(x)$ for all i; that is,

$$\left\langle x_i^*, y - x\right\rangle \le g_{\lambda_i}(y) - f(x) + \varepsilon,\ \text{for all } y \in X.$$

Thus, since may assume without loss of generality that $\lambda_i \to \lambda \in \Delta_m$, by taking limits on i in the last inequality we deduce that $x^* \in \partial_{\varepsilon+g_\lambda(x)-f(x)} g_\lambda(x)$ and the nontrivial inclusion in (46) follows. $\square$

We finish this work by providing a general approximate KKT condition for the convex optimization problem

$$(P) \qquad \inf_{f(x):=\sup\limits_{t\in T} f_t \le 0} g(x),$$

where the functions $f_t, g : X \to \overline{\mathbb{R}}$, $t \in T$, are lsc proper and convex. Next, we establish some KKT ε-optimality conditions without resorting to the Slater condition or any other constraint qualification.

Corollary 16 *Let $\bar{x} \in (\operatorname{dom} g) \cap [f \le 0]$ and $\varepsilon > 0$. Then, in the setting above, $\bar{x}$ is ε-minimum of (P) if and only if*

$$\theta \in \operatorname{cl}\left(\bigcup_{\substack{\alpha\in\mathbb{R},\ \mu>0 \\ \lambda\in\Delta(T)}} \partial_{\varepsilon-\alpha} g(\bar{x}) + \mu \partial_{\frac{\alpha}{\mu}+f_\lambda(\bar{x})} f_\lambda(\bar{x}) \right), \tag{47}$$

where $f_\lambda := \sum_{t\in\operatorname{supp}\lambda} \lambda_t f_t$.

Proof Assume that $\bar{x}$ is an ε-minimum of (P). Then, by [13] and Example 3,

$$\theta \in \operatorname{cl}\left(\bigcup_{\alpha\in\mathbb{R}} \partial_{\varepsilon-\alpha} g(\bar{x}) + \mathrm{N}^{\alpha}_{[f\le 0]}(\bar{x}) \right) \subset \operatorname{cl}\left(\bigcup_{\alpha\in\mathbb{R},\ \mu>0} \left\{ \partial_{\varepsilon-\alpha} g(\bar{x}) + \mu \partial_{\frac{\alpha}{\mu}+f(\bar{x})} f(\bar{x}) \right\} \right).$$

Therefore, appealing to Theorem 12, we obtain (47). The proof is finished as the sufficiency of (47) is straightforward. □

Acknowledgements The authors are very grateful for the referees' thorough reviews, which have definitely contributed to improving the latest version of the article.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. The research of the first author is supported by Centro Basal CMM FB210005 and by Fondecyt Regular 1240335. Second and third authors are supported by the Research Project PID2022-136399NB-C21 from MICINN, Spain. The research of the second author is also supported by MICIU of Spain and Universidad de Alicante (Contract Beatriz Galindo BEA- GAL 18/00205), AICO/2021/165 of Generalitat Valenciana CMM FB210005.

Data Availability Not applicable.

Declarations

Competing Interests On behalf of the authors, the corresponding author states that there are no Competing Interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Brøndsted, A.: On the subdifferential of the supremum of two convex functions. Math. Scand. **31**, 225–230 (1972)
2. Combari, C., Laghdir, M., Thibault, L.: Sous-différentiels de fonctions convexes composées. Ann. Sci. Math. Qué. **18**, 554–561 (1994)
3. Correa, R., Hantoute, A., García, Y.: Nonconvex integration using ε-subdifferentials. Optimization **67**, 2205–2227 (2018)
4. Correa, R., Hantoute, A., López, M.A.: Weaker conditions for subdifferential calculus of convex functions. J. Funct. Anal. **271**, 1177–1212 (2016)
5. 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)
6. 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)
7. Correa, R., Hantoute, A., López, M.A.: Fundamentals of Convex Analysis and Optimization. A Supremum Function Approach. Springer Series in Operations Research and Financial Engineering. Springer, Berlin (2023)
8. Fan, K.: Minimax theorems. Proc. Natl. Acad. Sci. USA **39**, 42–47 (1953)
9. 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)
10. Hantoute, A., Svensson, A.: A general representation of δ-normal sets to sublevels of convex functions. Set-Valued Var. Anal. **25**, 651–678 (2017)
11. 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)
12. Hantoute, A., Jourani, A., Vicente-Pérez, J.: Lebesgue infinite sums of convex functions: subdifferential calculus. J. Convex Anal. **30**, 1053–1072 (2023)
13. Hiriart-Urruty, J.-B., Moussaoui, M., Seeger, A., Volle, M.: Subdifferential calculus without qualifi cation conditions, using approximate subdifferentials: a survey. Nonlinear Anal. TMA **24**, 1727–1754 (1995)
14. Ioffe, A.D.: A note on subdifferentials of pointwise suprema. Top **20**, 456–466 (2012)
15. Ioffe, A.D., Levin, V.L.: Subdifferentials of convex functions. Tr. Mosk. Mat. Obŝ. **26**, 3–73 (1972). (Russian)
16. Li, C., Ng, K.F.: Subdifferential calculus rules for supremum functions in convex analysis. SIAM J. Optim. **21**, 782–797 (2011)
17. López, M.A., Volle, M.: A formula for the set of optimal solutions of a relaxed minimization problem. Applications to subdifferential calculus. J. Convex Anal. **17**, 1057–1075 (2010)
18. López, M.A., Volle, M.: On the subdifferential of the supremum of an arbitrary family of extended real-valued functions. Rev. R. Acad. Cienc. Exactas Fís. Nat., Ser. A Mat. **105**, 3–21 (2011)
19. Mordukhovich, B., Nghia, T.T.A.: Subdifferentials of nonconvex supremum functions and their applications to semi-infinite and infinite programs with Lipschitzian data. SIAM J. Optim. **23**, 406–431 (2013)
20. Pérez-Aros, P.: Formulae for the conjugate and the subdifferential of the supremum function. J. Optim. Theory Appl. **180**, 397–427 (2019)
21. Valadier, V.: Sous-différentiels d'une borne supérieure et d'une somme continue de fonctions convexes. C. R. Acad. Sci., Sér. 1 Math. **268**, 39–42 (1969)
22. Volle, M.: Sous-différentiels d'une enveloppe supérieure de fonctions convexes. C. R. Acad. Sci., Sér. 1 Math. **317**, 845–849 (1993)
23. Zălinescu, C.: Convex Analysis in General Vector Spaces. World Scientific, River Edge (2002)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.