SIAM J. OPTIM.
Vol. 19, No. 2, pp. 863–882

SUBDIFFERENTIAL CALCULUS RULES IN CONVEX ANALYSIS: A UNIFYING APPROACH VIA POINTWISE SUPREMUM FUNCTIONS*

A. HANTOUTE†, M. A. LÓPEZ‡, AND C. ZĂLINESCU§

Abstract. We provide a rule to calculate the subdifferential set of the pointwise supremum of an arbitrary family of convex functions defined on a real locally convex topological vector space. Our formula is given exclusively in terms of the data functions and does not require any assumption either on the index set on which the supremum is taken or on the involved functions. Some other calculus rules, namely chain rule formulas of standard type, are obtained from our main result via new and direct proofs.

Key words. convex analysis, convex subdifferential, pointwise supremum function, calculus rules

AMS subject classifications. 52A41, 90C25, 15A39

DOI. 10.1137/070700413

1. Introduction.

Many operations with convex functions preserve convexity, and so it is natural to ask if the subdifferential of the resulting function can be written in terms of the data functions. Specific to convex analysis is the classical operation of taking the pointwise supremum of an arbitrarily indexed family of convex functions. It has no equivalence in the classical theory of differentiable analysis and constitutes a largely used tool in convex optimization, in theory as well as in practice (see, for instance, [1], [10], and the references therein). In [5] and [8] certain specific techniques relying on the supremum function were applied in the framework of semi-infinite linear optimization.

In this paper, we provide explicit characterizations for the subdifferential mapping of the supremum function of an arbitrarily indexed family of convex functions, exclusively in terms of the data functions. The main virtue of our approach is that the index set over which the supremum is taken is arbitrary, without any algebraic or topological structure. Also the convex functions we consider in this paper are general, defined on a separated locally convex space, and not necessarily lower semicontinuous (lsc) or even proper. Further, we do not assume regularity conditions such as the continuity of the supremum function, the continuity of the data functions, conditions on their domains, and the like.

Since many convex functions can be written as the supremum of continuous affine mappings, numerous operations dealing with such (convex) functions can be formulated as a pointwise supremum of other functions whose subdifferentials can easily be characterized. Specifically, we have proved that our formulas also lead to other

*Received by the editors August 17, 2007; accepted for publication (in revised form) March 26, 2008; published electronically August 13, 2008. Research supported by grants MTM2005-08572-C03 (01) from MEC (Spain) and FEDER (E.U.), ACOMP06/117 and ACOMP/2007/247-292 from Generalitat Valenciana (Spain), and ID-PCE-379 (Romania).
http://www.siam.org/journals/siopt/19-2/70041.html

†Operations Research Center, Miguel Hernández University of Elche, 03202 Elche (Alicante), Spain (hantoute@ua.es).

‡Department of Statistics and Operations Research, University of Alicante, 03071 Alicante, Spain (marco.antonio@ua.es).

§University "Al. I. Cuza" Iaşi, Faculty of Mathematics, Bd. Carol I, Nr. 11, 700506 Iaşi, Romania and Institute of Mathematics Octav Mayer, Iaşi, Romania (zalinesc@uaic.ro).

calculus rules for the subdifferentials of certain operations with convex functions, such as the sum and the composition with affine applications. In this way, our approach gives rise to a unifying view of many well-known calculus rules in convex analysis.

Deriving calculus rules for subdifferentials is one of the first issues raised in convex analysis. Consequently, many earlier contributions dealing with pointwise supremum functions can be found in the literature. See, for instance, [26] to trace out the historical origins of the issue, as well as [2], [3], [4], [12], [13], [15], [20], [21], and [27]. This is why we make a short historical review of some of these results.

Consider the pointwise supremum $f := \sup_{t\in T} f_t$ of a collection of convex functions $f_t : X \to \mathbb{R} \cup \{+\infty\}$, $t \in T \neq \emptyset$, defined on a separated locally convex space X, and let $z \in \operatorname{dom} f$. When T is finite and each f_t is continuous at z, a basic result due to Dubovitskij and Milyutin asserts that (see, e.g., [13])

$$\partial f(z) = \operatorname{co}\left(\bigcup\nolimits_{t\in T(z)} \partial f_t(z)\right),$$

where

$$T(z) := \{t \in T \mid f_t(z) = f(z)\},$$

and co stands for the convex hull. When T is a separated compact topological space and the function $(t, x) \to f_t(x)$ is upper semicontinuous with respect to t for each x, then assuming that each f_t is continuous at z, the following formula can be found, for instance, in [32, Thm. 2.4.18]:

$$\partial f(z) = \operatorname{cl}\left(\operatorname{co} \bigcup\nolimits_{t\in T(z)} \partial f_t(z)\right),$$

where the closure, cl, is taken in the topological dual space X^* with respect to the weak* topology $w^* = \sigma(X^*, X)$.

According to [26], the last result was first established by Levin [15] for a finite-valued convex function defined on $\mathbb{R}^n$. The continuity assumption on the data functions is weakened in [29] and [21, Thm. 4].

Even in simple situations dealing with finitely many functions, the problem is involved so that simple examples in the Euclidean space show that these nice formulae above do not hold in general. Nevertheless, in order to overcome this difficulty, Brøndsted [2] used the concept of ε-subdifferential to establish the following formula, which is valid when $T = \{1, 2, \ldots, k\}$ and all of the functions f_i, $i = 1, 2, \ldots, k$ agree at z :

$$\partial f(z) = \bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left(\operatorname{co} \bigcup\nolimits_{i=1}^{k} \partial_\varepsilon f_i(z)\right).$$

In the case of an infinite collection of convex functions (T infinite), and following [10, p. 405], the most elaborated results are due to Valadier in [27] where, in the context of normed vector spaces and assuming that the supremum function f is continuous at z, the subdifferential $\partial f(z)$ is expressed by considering not only z but all nearby points around it. More precisely, denoting by $\|\cdot\|$ the corresponding norm in X, the following formula is given in [27]:

$$\partial f(z) = \bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left[\operatorname{co}\left(\bigcup\{\partial f_t(y) \mid y \in X,\ t \in T : \|y - z\| \leq \varepsilon,\ f_t(z) \geq f(z) - \varepsilon\}\right)\right].$$

By using the concept of ε-subdifferential, Volle [28] obtained another characterization of $\partial f(z)$ where only the nominal point z appears but in terms of approximate subgradients:

$$\partial f(z) = \bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\left[\operatorname{co}\left(\bigcup\{\partial_\varepsilon f_t(z) \mid t \in T : f_t(z) \geq f(z) - \varepsilon\}\right)\right].$$

It is worth noting that if either all of the functions f_t are affine or if the space X is Banach, then the last two formulas above are equivalent. The equivalence for affine functions is clear while in the Banach spaces setting this observation is partly due to Brøndsted–Rockafellar's theorem, expressing the ε-subdifferential by means of exact subdifferentials at nearby points. As it can be seen, the advantage of using such an enlargement of the subdifferential, namely, the ε-subdifferential, is to avoid qualifications type conditions. Such an idea is exploited in the survey paper [11] (see also references therein) to provide many calculus rules without requiring any regularity condition.

Recently, in [7], the following characterization for the subdifferential ∂f is given when $f_t : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, $t \in T$, are proper convex functions and T is arbitrary:

$$\partial f(z) = \bigcap_{\varepsilon>0} \mathrm{cl}\left[\mathrm{co}\left(\bigcup\{\partial_\varepsilon f_t(z) \mid t \in T : f_t(z) \geq f(z) - \varepsilon\}\right) + \mathrm{N}_{\mathrm{dom}\, f}(z)\right],$$

where $\mathrm{N}_{\mathrm{dom}\, f}(z)$ stands for the normal cone to the domain of f ($\mathrm{dom}\, f$) at z, provided that either the f_t's are lsc or that the relative interiors of their (effective) domains have a common point. In this setting, the formula above implies the one given by Volle [28], since $\mathrm{N}_{\mathrm{dom}\, f}(z) = \{\theta\}$ whenever z is a continuity point of the supremum function f. Further, when dealing with a finite number of functions the term $\mathrm{N}_{\mathrm{dom}\, f}(x)$ can be removed from the formula above which, consequently, entails the one of Brøndsted [2].

At this step, the purpose of the present paper is twofold. First, we extend the last formula from [7] to the setting of convex functions defined on locally convex spaces and which are not necessarily proper or lsc. To this aim, we consider those collections of functions satisfying the following closedness criterion, which holds for a broad class of convex functions and obviously covers the case of lsc functions:

$$\mathrm{cl}\, f = \sup_{t \in T} \mathrm{cl}\, f_t, \tag{1}$$

where $\mathrm{cl}\, f$ and $\mathrm{cl}\, f_t$ stand for the lsc hull of the convex functions f and f_t, respectively. Second, we give a unified approach for the framework of calculus rules in convex analysis. In fact, our characterization of ∂f also allows us to obtain formulas for the subdifferential of the resulting function in many operations as the sum of convex functions and the composition of an affine continuous mapping with a convex function. In this way, we provide direct and easier proofs for the basic chain rules when some supplementary qualification conditions are assumed.

The summary of the paper is as follows. In section 2 we introduce the main tools and basic results used in the paper. In section 3 we give the aimed formula for the subdifferential of the supremum of an arbitrary family of convex functions. After a series of auxiliary lemmas the main result is stated in Theorem 4. In it we use a closedness criterion which is studied in Corollary 9. We close this section by deriving some other formulae in Corollaries 7 (for affine functions), 8 (for finite-dimensional spaces or, more generally, when the relative interior of the domain of the supremum function f is not empty), 10 (Volle's formula), and 12 (Brøndsted's formula). In section 4 we introduce a unifying framework for deriving subdifferential calculus rules. Namely, in Theorem 13 we give a formula for the subdifferential of the sum of a convex function and another convex function precomposed with a continuous affine mapping. Theorem 13 constitutes a slight extension of Hiriart-Urruty–Phelps formula (Corollary 14). It also yields an easy derivation of the basic chain rule (Corollary 16) when some supplementary conditions are assumed, namely, the Moreau–Rockafellar constraint qualification.

2. Notations and basic tools. In this paper X and Y stand for (real) separated locally convex spaces. Their topological dual spaces are respectively denoted by X^* and Y^*. The spaces X and X^* (Y and Y^*) are paired in duality by the bilinear form $(x^*,x) \in X^* \times X \mapsto \langle x^*,x\rangle := \langle x,x^*\rangle := x^*(x)$ ($(y^*,y) \in Y^* \times Y \mapsto \langle y^*,y\rangle$, respectively). Throughout the paper, the sole topology defined on the dual spaces is the w^*–topology. The zero vectors in the involved spaces are all denoted by θ, and the neighborhoods of θ are called θ–neighborhoods. We use the notation $\overline{\mathbb{R}} := \mathbb{R}\cup\{-\infty,+\infty\}$.

We first recall some basic results of convex analysis which can be found, e.g., in the books [17] and [32] and the references therein (see also [10] and [22]). Given two nonempty sets A and B in X (or in X^*, Y, Y^*), we define the algebraic (or Minkowski) sum by

$$A+B := \{a+b \mid a\in A,\ b\in B\}, \quad A+\emptyset := \emptyset + A := \emptyset; \tag{2}$$

moreover, if $\emptyset \neq \Lambda \subset \mathbb{R}$ we set

$$\Lambda A := \{\lambda a \mid \lambda\in\Lambda,\ a\in A\}, \quad \Lambda\emptyset := \emptyset.$$

Furthermore, $\Lambda x := \Lambda\{x\}$, $\lambda A := \{\lambda\}A$, and $x + A := \{x\}+A$.

By $\operatorname{co} A$, $\operatorname{cone} A$, and $\operatorname{aff} A$, we denote the *convex hull,* the *conic hull,* and the *affine hull* of the set A, respectively. Moreover, $\operatorname{int} A$ is the *interior* of A, and $\operatorname{cl} A$ and $\overline{A}$ are indistinctly used for denoting the *closure* of A (w^**–closure* if $A\subset X^*$ or $A\subset Y^*$). In this way, we set $\overline{\operatorname{co}}A := \operatorname{cl}(\operatorname{co} A)$ and $\overline{\operatorname{cone}}A := \operatorname{cl}(\operatorname{cone} A)$. We use $\operatorname{ri} A$ to denote the (topological) *relative interior* of A (i.e., the interior of A in the topology relative to $\operatorname{aff} A$ if $\operatorname{aff} A$ is closed, and the empty set otherwise). We shall use Greek letters for denoting real numbers.

The following properties are applied very often:

$$\operatorname{cl}(A+B) = \operatorname{cl}(A+\operatorname{cl} B), \tag{3}$$

and if A is convex,

$$\lambda \operatorname{ri} A + (1-\lambda)\operatorname{cl} A \subset \operatorname{ri} A \text{ for every } \lambda \in\ ]0,1]. \tag{4}$$

Associated with $A\neq\emptyset$ we consider the sets

$$\begin{aligned} A^\circ &:= \{x^*\in X^* \mid \langle x^*,x\rangle \geq -1 \text{ for all } x\in A\},\\ A^- &:= -(\operatorname{cone} A)^\circ = \{x^*\in X^* \mid \langle x^*,x\rangle \leq 0 \text{ for all } x\in A\}, \text{ and}\\ A^\perp &:= (-A^-)\cap A^- = \{x^*\in X^* \mid \langle x^*,x\rangle = 0 \text{ for all } x\in A\}, \end{aligned}$$

i.e., the (one-sided) *polar*, the *negative dual cone,* and the *orthogonal subspace* (or *annihilator*) of A, respectively. Observe that A° is a closed convex set containing θ, A^- is a closed convex cone, and $A^\perp$ is a closed linear subspace. Further, by the *bipolar theorem,* we have

$$A^{\circ\circ} = \overline{\operatorname{co}}(A\cup\{\theta\}) \text{ and } A^{--} = \overline{\operatorname{cone}}(\operatorname{co} A). \tag{5}$$

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

$$\mathrm{N}_A(x) := \begin{cases} (A-x)^- & \text{if } x\in A,\\ \emptyset & \text{if } x\in X\setminus A. \end{cases}$$

As a consequence of this definition $\mathrm{N}_\emptyset(x) = \emptyset$ for every $x \in X$. If $A \neq \emptyset$ is convex and closed, A_∞ represents its *recession cone* defined as

$$A_\infty := \{y \in X \mid x + \lambda y \in X \text{ for some } x \in X \text{ and all } \lambda \geq 0\}.$$

Given a function $f : X \longrightarrow \overline{\mathbb{R}}$, its *(effective) domain* and *epigraph* are defined by

$$\operatorname{dom} f := \{x \in X \mid f(x) < +\infty\},$$
$$\operatorname{epi} f := \{(x, \alpha) \in X \times \mathbb{R} \mid f(x) \leq \alpha\};$$

moreover, when f is *proper*, that is, $\operatorname{dom} f \neq \emptyset$ and $f(x) > -\infty$ for all $x \in X$, we consider the *graph* of f as being defined by

$$\operatorname{gph} f := \{(x, f(x)) \in X \times \mathbb{R} \mid x \in \operatorname{dom} f\}.$$

So, for f proper one has $\operatorname{epi} f = \operatorname{gph} f + \mathbb{R}_+(\theta, 1)$. We say that f is *convex* if $\operatorname{epi} f$ is convex. In what follows we shall use the convention $+\infty - \infty := +\infty + (-\infty) := +\infty$. Assume that f is convex. The *lower closure* of f is the function $\operatorname{cl} f : X \longrightarrow \overline{\mathbb{R}}$ defined by

$$(\operatorname{cl} f)(x) := \inf\{t \mid (x, t) \in \operatorname{cl}(\operatorname{epi} f)\}.$$

Clearly we have $\operatorname{epi}(\operatorname{cl} f) = \operatorname{cl}(\operatorname{epi} f)$, which implies that $\operatorname{cl} f$ is a lsc convex function dominated by f; i.e., $\operatorname{cl} f \leq f$. Equivalently, we have

$$(\operatorname{cl} f)(x) = \liminf_{y \to x} f(y) \quad \forall x \in X.$$

Further, it can be checked that $\operatorname{cl}(\operatorname{dom}(\operatorname{cl} f)) = \operatorname{cl}(\operatorname{dom} f)$. If $(\operatorname{cl} f)(x) = f(x)$, then f is lsc at x. If there exists $x_0 \in X$ such that $(\operatorname{cl} f)(x_0) = -\infty$, then $(\operatorname{cl} f)(x) = -\infty$ for all $x \in \operatorname{dom}(\operatorname{cl} f)$. We shall denote by $\Lambda(X)$ the set of all the proper convex functions on X, and by $\Gamma(X)$ the subset of $\Lambda(X)$ consisting of the lsc functions; the sets $\Lambda(X^*)$ and $\Gamma(X^*)$ are defined in a similar way.

The *Fenchel conjugate* of f is the function $f^* : X^* \longrightarrow \overline{\mathbb{R}}$ given by

$$f^*(x^*) := \sup\{\langle x^*, x\rangle - f(x) \mid x \in X\}.$$

The functions f and $\operatorname{cl} f$ have the same conjugate; i.e., $f^* = (\operatorname{cl} f)^*$. The *biconjugate* of f is the function $f^{**} : X \longrightarrow \overline{\mathbb{R}}$ given by

$$f^{**}(x) := \sup\{\langle x^*, x\rangle - f^*(x^*) \mid x^* \in X^*\}.$$

Let us recall here that $f^* \in \Gamma(X^*)$ if and only if $\operatorname{dom} f \neq \emptyset$ and there exist $x^* \in X^*$ and $\alpha \in \mathbb{R}$ such that $f(x) \geq \langle x^*, x\rangle + \alpha$ for all $x \in X$; this happens, for instance, when $f \in \Gamma(X)$ in which case we have $f^{**} = f$.

The *support* and the *indicator* functions of $A \neq \emptyset$ are, respectively, defined as

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

and

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

The function σ_A is sublinear, lsc, and satisfies

$$\sigma_A = \sigma_{\overline{\text{co}}A} = \mathrm{I}^*_{\overline{\text{co}}A}. \tag{6}$$

Moreover, it is known that $(\operatorname{dom}\sigma_A)^- = (\overline{\text{co}}A)_\infty$ (e.g., [29, p. 142]) or equivalently, by using (5),

$$\operatorname{cl}(\operatorname{dom}\sigma_A) = \left[(\overline{\text{co}}A)_\infty\right]^- . \tag{7}$$

If $A_1, \ldots, A_m \subset X$ are nonempty sets ($m \geq 2$), then clearly $\sigma_{A_1} + \cdots + \sigma_{A_m} = \sigma_{A_1+\cdots+A_m}$ and $\max\{\sigma_{A_1}, \ldots, \sigma_{A_m}\} = \sigma_{A_1\cup\cdots\cup A_m}$; moreover, if $1 \leq k < m$, then

$$\sigma_{A_1} + \cdots + \sigma_{A_k} + \max\left\{\sigma_{A_{k+1}}, \ldots, \sigma_{A_m}\right\} = \sigma_{A_1+\cdots+A_k+(A_{k+1}\cup\cdots\cup A_m)}.$$

Hence

$$\operatorname{dom}\sigma_{A_1+\cdots+A_m} = \operatorname{dom}\sigma_{A_1\cup\cdots\cup A_m} = \operatorname{dom}\sigma_{A_1+\cdots+A_k+(A_{k+1}\cup\cdots\cup A_m)}.$$

Using (6) and (7) we get

$$\begin{aligned} \left[\overline{\text{co}}(A_1 + \cdots + A_m)\right]_\infty &= \left[\overline{\text{co}}(A_1 \cup \cdots \cup A_m)\right]_\infty \\ &= \left[\overline{\text{co}}\left(A_1 + \cdots + A_k + (A_{k+1} \cup \cdots \cup A_m)\right)\right]_\infty . \end{aligned} \tag{8}$$

If f is convex and $\varepsilon \geq 0$, the *ε-subdifferential* of f at a point $x \in X$ such that $f(x) \in \mathbb{R}$ is the w^*–closed convex set

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

If $f(x) \notin \mathbb{R}$, then we set $\partial_\varepsilon f(x) := \emptyset$. In particular, for $\varepsilon = 0$ we get $\partial f(x) := \partial_0 f(x)$, the *subdifferential* of f at x. Given $x \in X$ and $\varepsilon \geq 0$ we recall the following properties: $\partial f(x) = \cap_{\varepsilon>0}\partial_\varepsilon f(x)$ and $\partial_\varepsilon f(x) = \partial_\varepsilon f(x) + \mathrm{N}_{\operatorname{dom} f}(x)$; moreover, as a simple computation shows (see also [32, Exer. 2.23]),

$$\left[\partial_\varepsilon f(x)\right]_\infty = \mathrm{N}_{\operatorname{dom} f}(x) \text{ for all } x \in \operatorname{dom} f \text{ and } \varepsilon \geq 0 \text{ with } \partial_\varepsilon f(x) \neq \emptyset. \tag{9}$$

If f is not proper, then $\partial_\varepsilon f(x) = \emptyset$ for all $x \in X$. If f is lsc at x and $f(x) \in \mathbb{R}$, then

$$\partial_\varepsilon\left(\operatorname{cl} f\right)(x) = \partial_\varepsilon f(x). \tag{10}$$

If $\partial f(x) \neq \emptyset$, then we have

$$(\operatorname{cl} f)(x) = f(x) \text{ and } \partial\left(\operatorname{cl} f\right)(x) = \partial f(x). \tag{11}$$

If $f \in \Lambda(X)$ and $f(x) \in \mathbb{R}$, then we have $\partial_\varepsilon f(x) \neq \emptyset$ for all $\varepsilon > 0$ if and only if f is lsc at x. Moreover, we have

$$\partial_\varepsilon f(x) = \{x^* \in X^* \mid f(x) + f^*(x^*) \leq \langle x^*, x\rangle + \varepsilon\} \text{ for all } \varepsilon \geq 0. \tag{12}$$

If A is convex and $x \in A$,

$$\partial \mathrm{I}_A(x) = (\operatorname{cone}(A - x))^- = \mathrm{N}_A(x).$$

Finally, if $f \in \Gamma(X)$, then for every $x \in \operatorname{dom} f$, $u \in X$ and $\varepsilon > 0$, we have (see [32, Thm. 2.4.11])

$$f'_\varepsilon(x, u) := \inf_{\lambda>0} \frac{f(x + \lambda u) - f(x) + \varepsilon}{\lambda} = \sigma_{\partial_\varepsilon f(x)}(u). \tag{13}$$

3. Calculus rules for the subdifferential of the supremum function. In this section we consider a nonempty family $\{f_t \mid t \in T\}$ of convex functions $f_t : X \to \overline{\mathbb{R}}$ defined on a (separated) real locally convex space X. The corresponding *pointwise supremum function* $f : X \to \overline{\mathbb{R}}$, given by

$$f(x) := \sup\{f_t(x) \mid t \in T\}, \tag{14}$$

is also convex; our main purpose in this section is to provide a formula for the subdifferential ∂f of f in terms exclusively of the data functions f_t, $t \in T$. The following simple example draws aside, in general, the possibility of writing ∂f in terms of ∂f_t, $t \in T$.

Example 1. [11, Ex. 2.1] Let $f_1, f_2 : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be defined by

$$f_1(x) = \begin{cases} -2\sqrt{x} & \text{if } x \geq 0, \\ +\infty & \text{if } x < 0, \end{cases} \quad \text{and } f_2(x) = f_1(-x)$$

so that $f := \max\{f_1, f_2\} = \mathrm{I}_{\{0\}}$. Then, we easily check that $\partial f(0) = \mathbb{R}$ while both $\partial f_1(0)$ and $\partial f_2(0)$ are empty.

Nevertheless, Theorem 4 below provides a characterization of ∂f, which involves the approximate subdifferentials of the data functions. To start with, we first establish two elementary lemmas.

LEMMA 1. *Let $h \in \Lambda(X)$ and $A \subset \operatorname{dom} h$ be a convex set. If $\operatorname{ri} A \neq \emptyset$, then $\inf_A h = \inf_{\operatorname{cl} A} h$.*

Proof. Set $\mu := \inf_A h$. Fix some $x_0 \in \operatorname{ri} A$ and consider $x \in \operatorname{cl} A$. Take $x_n := (1 - \frac{1}{n})x + \frac{1}{n}x_0$ for $n \geq 1$; then

$$\mu \leq h(x_n) \leq (1 - \tfrac{1}{n})h(x) + \tfrac{1}{n}h(x_0).$$

Taking the limit we get $\mu \leq h(x)$; hence $\mu \leq \inf_{\operatorname{cl} A} h$. ☐

The following simple result is an immediate consequence of (10) and (11).

LEMMA 2. *Let $h \in \Lambda(X)$ and $x \in \operatorname{dom} h$. If $\operatorname{cl} h \in \Lambda(X)$, then*

$$\partial_\varepsilon h(x) = \partial_{\operatorname{cl} h(x) - h(x) + \varepsilon} \operatorname{cl} h(x) \quad \textit{for all} \quad \varepsilon \in \mathbb{R}.$$

Hence $\partial_\varepsilon h(x) \neq \emptyset$ for $\varepsilon > h(x) - \operatorname{cl} h(x)$, and $\partial_\varepsilon h(x) = \emptyset$ for $\varepsilon < h(x) - \operatorname{cl} h(x)$.

From now on, we fix the following notations. Given $z \in X$ and $\varepsilon > 0$ we set

$$\mathcal{F}_z := \{L \subset X \mid L \text{ is a finite-dimensional linear subspace, with } z \in L\},$$

and

$$T_\varepsilon(z) := \{t \in T \mid f_t(z) \geq f(z) - \varepsilon\},$$

where f_t and f are defined as in (14).

The following lemma provides the first extension of Proposition 3 in [7] to general locally convex spaces; [7, Prop. 3] is established in $\mathbb{R}^n$ using subdifferential calculus for support functions. Here we give a direct proof, which, in particular, does not appeal to the Fenchel linearization of the functions f_t.

LEMMA 3. *Let $f_t \in \Gamma(X)$ for $t \in T \neq \emptyset$ and set $f := \sup_{t \in T} f_t$. Assume that $z \in \operatorname{dom} f$ and that $\operatorname{ri}(\operatorname{dom} f) \neq \emptyset$, then*

$$\partial f(z) = \bigcap_{\varepsilon > 0} \operatorname{cl}\left(\operatorname{co}\left(\bigcup_{t \in T_\varepsilon(z)} \partial_{\alpha\varepsilon} f_t(z)\right) + \mathrm{N}_{\operatorname{dom} f}(z)\right) \quad \forall \alpha > 0.$$

Proof. Fix $\alpha > 0$. Denote by A the set in the right-hand side of the above equality. Without loss of generality (w.l.o.g.) we assume that $z = \theta$ and $f(\theta) = 0$. Set

$$T_\varepsilon := T_\varepsilon(\theta), \quad A_\varepsilon := \operatorname{co}\left(\bigcup_{t\in T_\varepsilon(z)} \partial_{\alpha\varepsilon} f_t(z)\right).$$

Note first that, $A_\varepsilon \subset \partial_{(1+\alpha)\varepsilon} f(\theta)$, which together with $\mathrm{N}_{\operatorname{dom} f}(\theta) = (\partial_{(1+\alpha)\varepsilon} f_t(\theta))_\infty$, implies that $\operatorname{cl}(A_\varepsilon + \mathrm{N}_{\operatorname{dom} f}(\theta)) \subset \partial_{(1+\alpha)\varepsilon} f(\theta)$. Indeed,

$$\begin{aligned}(15)\qquad \langle x, x^*\rangle &\le f_t(x) - f_t(\theta) + \alpha\varepsilon \\ &\le f(x) + (1+\alpha)\varepsilon \quad \forall t \in T_\varepsilon,\ \forall x^* \in \partial_{\alpha\varepsilon} f_t(\theta),\ \forall x \in X,\end{aligned}$$

whence $x^* \in \partial_{(1+\alpha)\varepsilon} f(\theta)$. Hence $A \subset \cap_{\varepsilon>0} \partial_{(1+\alpha)\varepsilon} f(\theta) = \partial f(\theta)$.

Let us prove now that $\partial f(\theta) \subset A$. Notice that $f = h^*$, where $h := \inf_{t\in T} f_t^* \ge f^* \ge 0$. Moreover, for $x^* \notin A_\varepsilon$ and $\varepsilon > 0$, we have that $h(x^*) \ge (1\wedge\alpha)\varepsilon := \min\{1,\alpha\}\varepsilon$. Indeed, if $t \in T_\varepsilon$, then $x^* \notin \partial_{\alpha\varepsilon} f_t(\theta)$, and so $f_t^*(x^*) \ge f_t(\theta) + f_t^*(x^*) > \langle \theta, x^*\rangle + \alpha\varepsilon = \alpha\varepsilon$, while, for $t \in T \setminus T_\varepsilon$ we have that $f_t^*(x^*) \ge \langle \theta, x^*\rangle - f_t(\theta) > -f(\theta) + \varepsilon = \varepsilon$. Hence $f_t^*(x^*) \ge (1\wedge\alpha)\varepsilon$ for every $t \in T$, and so $h(x^*) \ge (1\wedge\alpha)\varepsilon$. Take now $\bar{x}^* \in X^*$, which is not in $\operatorname{cl}(A_\varepsilon + \mathrm{N}_{\operatorname{dom} f}(\theta))$ for (some) $\varepsilon > 0$. Using a separation theorem, there exist $\bar{x} \in X$ and $\gamma > 0$ such that

$$(16)\qquad \langle \bar{x}, \bar{x}^*\rangle > \gamma + \langle \bar{x}, x^*\rangle + \langle \bar{x}, u^*\rangle \text{ for all } x^* \in A_\varepsilon \text{ and all } u^* \in \mathrm{N}_{\operatorname{dom} f}(\theta).$$

It follows that $\bar{x} \in (\mathrm{N}_{\operatorname{dom} f}(\theta))^- = \operatorname{cl}(\mathbb{R}_+ \operatorname{dom} f)$. Furthermore, note that from (15) we get $\operatorname{dom} f \subset \operatorname{dom} \sigma_{A_\varepsilon}$, and so $C := \mathbb{R}_+(\operatorname{dom} f) \subset \operatorname{dom} \sigma_{A_\varepsilon} = \operatorname{dom}(\sigma_{A_\varepsilon} - \bar{x}^*)$. Since $\operatorname{aff} C = \operatorname{aff}(\operatorname{dom} f)$ and $\operatorname{ri}(\operatorname{dom} f) \ne \emptyset$, we have that $\operatorname{ri} C \ne \emptyset$. Using Lemma 1 for $\sigma_{A_\varepsilon} - \bar{x}^*$ and C we obtain that one can take $\bar{x} \in \operatorname{dom} f$.

For $\lambda \in]0,1[$ we have

$$\begin{aligned} f(\lambda\bar{x}) &= \sup\{\langle \lambda\bar{x}, x^*\rangle - h(x^*) \mid x^* \in X^*\} \\ &= \max\left\{ \sup_{x^*\in A_\varepsilon} [\langle \lambda\bar{x}, x^*\rangle - h(x^*)],\ \sup_{x^*\in X^*\setminus A_\varepsilon} [\langle \lambda\bar{x}, x^*\rangle - h(x^*)] \right\}. \end{aligned}$$

But, $h \ge 0$ and $\langle \bar{x}, \bar{x}^*\rangle \ge \gamma + \sigma_{A_\varepsilon}(\bar{x})$ (being a consequence of (16)) allow us to write

$$\begin{aligned} \sup_{x^*\in A_\varepsilon} [\langle \lambda\bar{x}, x^*\rangle - h(x^*)] &\le \sup_{x^*\in A_\varepsilon} \langle \lambda\bar{x}, x^*\rangle = \lambda\sigma_{A_\varepsilon}(\bar{x}) \\ &\le \lambda(-\gamma + \langle \bar{x}, \bar{x}^*\rangle) < \langle \lambda\bar{x}, \bar{x}^*\rangle, \end{aligned}$$

while the fact that $h \ge (1\wedge\alpha)\varepsilon$ on $X^* \setminus A_\varepsilon$ implies that

$$\begin{aligned} \sup_{x^*\in X^*\setminus A_\varepsilon} [\langle \lambda\bar{x}, x^*\rangle - h(x^*)] &\le \sup_{x^*\in X^*\setminus A_\varepsilon} \lambda[\langle \bar{x}, x^*\rangle - h(x^*)] + \sup_{x^*\in X^*\setminus A_\varepsilon} (1-\lambda)[-h(x^*)] \\ &\le \lambda h^*(\bar{x}) - (1-\lambda)(1\wedge\alpha)\varepsilon = \lambda f(\bar{x}) - (1-\lambda)(1\wedge\alpha)\varepsilon. \end{aligned}$$

Thus, since

$$\lambda f(\bar{x}) - (1-\lambda)(1\wedge\alpha)\varepsilon < \langle \lambda\bar{x}, x^*\rangle$$

for $\lambda \in]0,1[$ sufficiently small, for such λ we have $f(\lambda\bar{x}) < \langle \lambda\bar{x}, x^*\rangle$, whence $x^* \notin \partial f(\theta)$ because $f(\theta) = 0$. The proof is complete. ∎

Now we are ready to give the main result of the paper in which we establish the formula of the subdifferential of the supremum function f defined in (14).

THEOREM 4. *Let $\{f_t \mid t \in T\}$ be a nonempty family of convex functions $f_t : X \to \overline{\mathbb{R}}$ and set $f := \sup_{t\in T} f_t$. Assume that*

$$\operatorname{cl} f = \sup\{\operatorname{cl} f_t \mid t \in T\}.$$

Then, for every $z \in X$, we have

$$\partial f(z) = \bigcap_{L\in\mathcal{F}_z,\varepsilon>0} \operatorname{cl}\left(\operatorname{co}\left(\bigcup_{t\in T_\varepsilon(z)} \partial_{\alpha\varepsilon} f_t(z)\right) + \mathrm{N}_{L\cap \operatorname{dom} f}(z)\right) \quad \textit{for all } \alpha > 0.$$

Proof. Fix $\alpha > 0$ and denote by A the set in the right-hand side of the preceding equality.

Note first that the conclusion holds if $f(z) \notin \mathbb{R}$. Indeed, if $f(z) = +\infty$, then $\partial f(z) = \emptyset = \mathrm{N}_{L\cap \operatorname{dom} f}(z)$ for every $L \in \mathcal{F}_z$, and the conclusion holds trivially (taking into account (2)). If $f(z) = -\infty$, then $f_t(z) = -\infty$ for all $t \in T$, and so $\partial f(z) = \partial_{\alpha\varepsilon} f_t(z) = \emptyset$ for all $t \in T$ and all $\varepsilon > 0$, and again the conclusion holds trivially.

In the rest of the proof we assume that $f(z) \in \mathbb{R}$ and so, w.l.o.g., we take $z = \theta$ and $f(\theta) = 0$. To simplify the writing we use the notation

$$T_\varepsilon := T_\varepsilon(\theta), \qquad A_\varepsilon := \operatorname{co}\left(\bigcup_{t\in T_\varepsilon} \partial_{\alpha\varepsilon} f_t(\theta)\right), \qquad \mathcal{F} := \mathcal{F}_\theta.$$

The inclusion $A \subset \partial f(\theta)$ easily follows by the definition of A_ε. Indeed, fix $x \in \operatorname{dom} f$, and let $L \in \mathcal{F}$. Then, by setting $E := L + \mathbb{R}x$ we get

$$\langle x, x^* + u^*\rangle \le \langle x, x^*\rangle \le f_t(x) - f_t(\theta) + \alpha\varepsilon \le f(x) + (1+\alpha)\varepsilon$$

for all $t \in T_\varepsilon$, $x^* \in \partial_{\alpha\varepsilon} f_t(\theta)$, and $u^* \in \mathrm{N}_{E\cap \operatorname{dom} f}(\theta)$, whence

$$\langle x, v^*\rangle \le f(x) + (1+\alpha)\varepsilon \text{ for all } v^* \in \operatorname{cl}\left(A_\varepsilon + \mathrm{N}_{E\cap \operatorname{dom} f}(\theta)\right).$$

Because $E \in \mathcal{F}$ and $\mathrm{N}_{E\cap \operatorname{dom} f}(\theta) \subset \mathrm{N}_{L\cap \operatorname{dom} f}(\theta)$, we deduce that

$$\langle x, v^*\rangle \le f(x) + (1+\alpha)\varepsilon \text{ for all } \varepsilon > 0 \text{ and } v^* \in A.$$

Hence $\langle x, v^*\rangle \le f(x) - f(\theta)$ for all $x \in \operatorname{dom} f$ and $v^* \in A$. Therefore, $A \subset \partial f(\theta)$. To prove the inclusion $\partial f(\theta) \subset A$ it suffices to assume that $\partial f(\theta) \ne \emptyset$ in which case, by (11),

$$\partial f(\theta) = \partial\left(\operatorname{cl} f\right)(\theta) \text{ and } (\operatorname{cl} f)(\theta) = f(\theta) = 0. \tag{17}$$

For this aim we shall introduce a family of functions satisfying the assumptions of Lemma 3.

Let us set $S := \{t \in T \mid \operatorname{cl} f_t \text{ is not proper}\}$. Then $\operatorname{cl} f_t$ takes its values in $\{-\infty, +\infty\}$ for $t \in S$ and so, because $(\operatorname{cl} f_t)(\theta) \le (\operatorname{cl} f)(\theta) = 0$ for $t \in T$, we obtain that $(\operatorname{cl} f_t)(\theta) = -\infty$ for $t \in S$; using our hypothesis we get $T \setminus S \ne \emptyset$.

Fix $L \in \mathcal{F}$ and define the family of functions $\{g_t \mid t \in T\} \subset \Gamma(X)$ by

$$g_t(x) := \begin{cases} \max\{(\operatorname{cl} f_t)(x), -1\} & \text{for } t \in S, \\ (\operatorname{cl} f_t)(x) & \text{for } t \in T \setminus S \end{cases}$$

and set

$$g(x) := \sup\{g_t(x) + \langle x, x^*\rangle \mid x^* \in L^\perp,\ t \in T\}.$$

(Observe that $g = \sup_{t\in T} g_t + \mathrm{I}_L$.) Then, since $g_t \geq \mathrm{cl}\, f_t$ for every $t \in T$, the current assumption yields

$$g = \sup_{t\in T} g_t + \mathrm{I}_L \geq \sup_{t\in T} \mathrm{cl}\, f_t + \mathrm{I}_L = \mathrm{cl}\, f + \mathrm{I}_L.$$

Furthermore, thanks to (17), there exists a convex neighborhood U of θ such that $(\mathrm{cl}\, f)(x) > -1$ for every $x \in U$. Hence for $x \in U \cap L$ we have either $(\mathrm{cl}\, f)(x) = +\infty \geq g(x)$ or $(\mathrm{cl}\, f)(x) < +\infty$; in this case for $t \in S$ one has $(\mathrm{cl}\, f_t)(x) = -\infty$, and so $g_t(x) = -1 \leq (\mathrm{cl}\, f)(x)$, while for $t \in T \setminus S$ one has $g_t(x) = (\mathrm{cl}\, f_t)(x) \leq (\mathrm{cl}\, f)(x)$. We deduce that $g(x) \leq (\mathrm{cl}\, f)(x)$ for $x \in U \cap L$. Therefore,

$$g(x) = (\mathrm{cl}\, f)(x) + \mathrm{I}_L(x) \quad \text{for every } x \in U. \tag{18}$$

Moreover, because $L \cap U \cap \mathrm{dom}\, f \subset L \cap U \cap \mathrm{dom}(\mathrm{cl}\, f) = U \cap \mathrm{dom}\, g$, we get

$$\mathrm{N}_{\mathrm{dom}\, g}(\theta) \subset \mathrm{N}_{L\cap \mathrm{dom}\, f}(\theta). \tag{19}$$

Now set

$$T'_\varepsilon := \{t \in T \mid g_t(\theta) \geq -\varepsilon\}.$$

Then $T'_\varepsilon \subset T_\varepsilon \setminus S$ for $\varepsilon \in]0, 1[$. In fact, since $g_t(\theta) = -1$ for $t \in S$, we have that $T'_\varepsilon \subset T \setminus S$. Hence, for $t \in T'_\varepsilon$ we write $0 \geq f_t(\theta) \geq (\mathrm{cl}\, f_t)(\theta) = g_t(\theta) \geq -\varepsilon$, and so $t \in T_\varepsilon$. Moreover, for $t \in T'_\varepsilon$ we have that $\partial_{\alpha\varepsilon}(\mathrm{cl}\, f_t)(\theta) \subset \partial_{(1+\alpha)\varepsilon} f_t(\theta)$. Indeed, since we have $f_t(\theta) - (\mathrm{cl}\, f_t)(\theta) \leq f(\theta) - g_t(\theta) = g(\theta) - g_t(\theta) \leq \varepsilon$, Lemma 2 yields $\partial_{\alpha\varepsilon}(\mathrm{cl}\, f_t)(\theta) = \partial_{\alpha\varepsilon+f_t(\theta)-(\mathrm{cl}\, f_t)(\theta)} f_t(\theta) \subset \partial_{(1+\alpha)\varepsilon} f_t(\theta)$. In view of these observations we get

$$\mathrm{co}\left(\bigcup_{t\in T'_\varepsilon} \partial_{\alpha\varepsilon} g_t(\theta)\right) \subset \mathrm{co}\left(\bigcup_{t\in T'_\varepsilon} \partial_{(1+\alpha)\varepsilon} f_t(\theta)\right) \quad \text{for all } \varepsilon \in]0, 1[. \tag{20}$$

Now we go back to the proof of the inclusion $\partial f(\theta) \subset A$. We apply Lemma 3 for the family $\{g_{(t,x^*)} \mid (t, x^*) \in T \times L^\perp\} \subset \Gamma(X)$ with $g_{(t,x^*)} := g_t + x^*$ and α (this is possible because $g = \sup\{g_{(t,x^*)} \mid (t, x^*) \in T \times L^\perp\}$ and $\mathrm{dom}\, g \subset L$, and so $\mathrm{ri}(\mathrm{dom}\, g) \neq \emptyset$, L being a finite-dimensional space). We obtain

$$\begin{aligned}\partial g(\theta) &= \bigcap_{\varepsilon>0} \mathrm{cl}\left(\mathrm{co}\left(\bigcup_{t\in T'_\varepsilon, x^*\in L^\perp} \partial_{\alpha\varepsilon}(g_t + x^*)(\theta)\right) + \mathrm{N}_{\mathrm{dom}\, g}(\theta)\right)\\ &= \bigcap_{\varepsilon>0} \mathrm{cl}\left(\mathrm{co}\left(\bigcup_{t\in T'_\varepsilon} \partial_{\alpha\varepsilon} g_t(\theta)\right) + L^\perp + \mathrm{N}_{\mathrm{dom}\, g}(\theta)\right).\end{aligned}$$

Then in view of the evident fact that $L^\perp + \mathrm{N}_{L\cap\mathrm{dom}\, f}(\theta) \subset \mathrm{N}_{L\cap\mathrm{dom}\, f}(\theta)$, and using (19) and (20), we get

$$\begin{aligned}\partial g(\theta) &\subset \bigcap_{\varepsilon\in]0,1[} \mathrm{cl}\left(\mathrm{co}\left(\bigcup_{t\in T'_\varepsilon} \partial_{\alpha\varepsilon} g_t(\theta)\right) + \mathrm{N}_{L\cap\mathrm{dom}\, f}(\theta)\right)\\ &\subset \bigcap_{\varepsilon\in]0,1[} \mathrm{cl}\left(\mathrm{co}\left(\bigcup_{t\in T_\varepsilon} \partial_{(1+\alpha)\varepsilon} f_t(\theta)\right) + \mathrm{N}_{L\cap\mathrm{dom}\, f}(\theta)\right).\end{aligned}$$

Hence, for each $\varepsilon \in]0,1[$ we obtain that, taking into account (17) and (18),

$$\begin{aligned}\partial f(\theta) &\subset \partial(\operatorname{cl} f)(\theta) + L^{\perp} = \partial(\operatorname{cl} f)(\theta) + \partial \mathrm{I}_L(\theta) \subset \partial\left(\operatorname{cl} f + \mathrm{I}_L\right)(\theta)\\ &= \partial g(\theta) \subset \operatorname{cl}\left(\operatorname{co}\left(\bigcup_{t\in T_\varepsilon} \partial_{(1+\alpha)\varepsilon} f_t(\theta)\right) + \mathrm{N}_{L\cap \operatorname{dom} f}(\theta)\right)\end{aligned}$$

for all $\varepsilon \in]0,1[$. Since $\delta := \frac{\alpha}{1+\alpha}\varepsilon \in]0,\varepsilon[\subset]0,1[$ (for $\varepsilon \in]0,1[$), we also have that

$$\partial f(\theta) \subset \operatorname{cl}\left(\operatorname{co}\left(\bigcup_{t\in T_\delta} \partial_{\alpha\varepsilon} f_t(\theta)\right) + \mathrm{N}_{L\cap \operatorname{dom} f}(\theta)\right) \subset \operatorname{cl}\left(A_\varepsilon + \mathrm{N}_{L\cap \operatorname{dom} f}(\theta)\right)$$

for all $\varepsilon \in]0,1[$. Since $\varepsilon \in]0,1[$ and $L \in \mathcal{F}$ were arbitrarily chosen, we obtain that

$$\partial f(\theta) \subset \bigcap_{L\in\mathcal{F},\varepsilon\in]0,1[} \operatorname{cl}\left(A_\varepsilon + \mathrm{N}_{L\cap \operatorname{dom} f}(\theta)\right) = \bigcap_{L\in\mathcal{F},\varepsilon>0} \operatorname{cl}\left(A_\varepsilon + \mathrm{N}_{L\cap \operatorname{dom} f}(\theta)\right) = A.$$

The proof is complete. ∎

Theorem 4 provides a complete description for ∂f only in terms of the data functions f_t, $t \in T$. Other descriptions will be provided in Theorem 6 below. We first establish the following lemma, which provides a straightforward infinite-dimensional extension of the corresponding statements in [7, Prop. 4].

LEMMA 5. *Let $T \neq \emptyset$ and $\{f_t \mid t \in T\} \subset \Gamma(X)$, and set $f := \sup\{f_t \mid t \in T\}$. Then, for every $z \in \operatorname{dom} f$, we have that*

$$\begin{aligned}
\mathrm{N}_{\operatorname{dom} f}(z) &= \{v^* \in X^* \mid (v^*, \langle v^*, z\rangle) \in [\overline{\operatorname{co}}\,(\cup_{t\in T} \operatorname{gph} f_t^*)]_\infty\} && (21)\\
&= \{v^* \in X^* \mid (v^*, \langle v^*, z\rangle) \in [\overline{\operatorname{co}}\,(\cup_{t\in T} \operatorname{epi} f_t^*)]_\infty\} && (22)\\
&= \{v^* \in X^* \mid (v^*, \langle v^*, z\rangle) \in (\operatorname{epi} f^*)_\infty\} && (23)\\
&= \{v^* \in X^* \mid (v^*, \langle v^*, z\rangle) \in \operatorname{epi}(\sigma_{\operatorname{dom} f})\}\,. && (24)
\end{aligned}$$

Proof. We assume that f is proper. Statement (24) is just the definition of $\mathrm{N}_{\operatorname{dom} f}(z)$. As seen in Lemma 3, we have that

$$(\inf_{t\in T} f_t^*)^* = \sup_{t\in T} f_t^{**} = \sup_{t\in T} f_t = f.$$

Since f is proper we obtain that

$$f^* = (\inf_{t\in T} f_t^*)^{**} = \overline{\operatorname{co}}\left(\inf_{t\in T} f_t^*\right),$$

that is, $\operatorname{epi} f^* = \overline{\operatorname{co}}(\cup_{t\in T} \operatorname{epi} f_t^*)$; moreover, by [32, Exer. 2.23] one has $(\operatorname{epi} f^*)_\infty = \operatorname{epi}(\sigma_{\operatorname{dom} f})$. Using these two relations we get statements (22) and (23). To finish the proof, it suffices to establish the equality between the sets appearing in the right-hand sides of (21) and (22), say, $E_1(z)$ and $E_2(z)$, respectively, or simply the inclusion $E_2(z) \subset E_1(z)$, since the opposite inclusion is trivial. Indeed, because for any proper function $g : X \to \overline{\mathbb{R}}$ one has $\operatorname{gph} g + \mathbb{R}_+(\theta,1) = \operatorname{epi} g$, we obtain that

$$\overline{\operatorname{co}}\,(\cup_{t\in T} \operatorname{gph} f_t^*) \subset \operatorname{cl}\left[\overline{\operatorname{co}}\,(\cup_{t\in T} \operatorname{gph} f_t^*) + \mathbb{R}_+(\theta,1)\right] = \overline{\operatorname{co}}\,(\cup_{t\in T} \operatorname{epi} f_t^*) = \operatorname{epi} f^*.$$

Since f^* is proper, we have $[\overline{\operatorname{co}}\,(\cup_{t\in T} \operatorname{gph} f_t^*)]_\infty \cap -[\mathbb{R}_+(\theta,1)]_\infty = \{(\theta,0)\}$, and so by [30, Cor. 3.12] (see also [16, Thm. 1.1]), we obtain that $\overline{\operatorname{co}}\,(\cup_{t\in T} \operatorname{gph} f_t^*) + \mathbb{R}_+(\theta,1)$ is closed, whence $\overline{\operatorname{co}}\,(\cup_{t\in T} \operatorname{gph} f_t^*) + \mathbb{R}_+(\theta,1) = \overline{\operatorname{co}}\,(\cup_{t\in T} \operatorname{epi} f_t^*)$, and

$$\begin{aligned}[\overline{\operatorname{co}}\,(\cup_{t\in T} \operatorname{epi} f_t^*)]_\infty &= [\overline{\operatorname{co}}\,(\cup_{t\in T} \operatorname{gph} f_t^*) + \mathbb{R}_+(\theta,1)]_\infty\\ &= [\overline{\operatorname{co}}\,(\cup_{t\in T} \operatorname{gph} f_t^*)]_\infty + \mathbb{R}_+(\theta,1).\end{aligned}$$

Take $v^* \in E_2(z)$; using the preceding relation, $(v^*, \langle v^*, z\rangle) = (x^*, \eta + \lambda)$ for some $(x^*, \eta) \in [\overline{\mathrm{co}}\,(\cup_{t\in T} \operatorname{gph} f_t^*)]_\infty$, and $\lambda \geq 0$. Moreover, since $\operatorname{dom} f \times \{-1\} \subset \operatorname{dom}(\sigma_{\operatorname{epi} f^*}) \subset [(\operatorname{epi} f^*)_\infty]^-$, we obtain that

$$\operatorname{dom} f \times \{-1\} \subset [(\overline{\mathrm{co}}\,(\cup_{t\in T} \operatorname{epi} f_t^*))_\infty]^- \subset [(\overline{\mathrm{co}}\,(\cup_{t\in T} \operatorname{gph} f_t^*))_\infty]^-,$$

and so $\langle (x^*, \eta), (z, -1)\rangle \leq 0$. Since $v^* = x^*$, it follows that

$$\lambda = \langle (v^*, \eta), (z, -1)\rangle = \langle (x^*, \eta), (z, -1)\rangle \leq 0;$$

hence $\lambda = 0$, and so $(v^*, \langle v^*, z\rangle) = (x^*, \eta) \in [\overline{\mathrm{co}}\,(\cup_{t\in T} \operatorname{gph} f_t^*)]_\infty$. This shows that $v^* \in E_1(z)$. ∎

We have the following theorem in which, for simplicity, we suppose that $f_t \in \Gamma(X)$ for all $t \in T$.

THEOREM 6. *Let $T \neq \emptyset$ and $\{f_t \mid t \in T\} \subset \Gamma(X)$, and set $f := \sup_{t\in T} f_t$. Then, for every $z \in X$ and every $\alpha > 0$, we have that*

$$\partial f(z) = \bigcap_{L\in\mathcal{F}_z,\varepsilon>0} \overline{\mathrm{co}}\Big(A_L + \bigcup_{t\in T_\varepsilon(z)} \partial_{\alpha\varepsilon} f_t(z)\Big) = \bigcap_{L\in\mathcal{F}_z,\varepsilon>0} \overline{\mathrm{co}}\Big(B_L + \bigcup_{t\in T_\varepsilon(z)} \partial_{\alpha\varepsilon} f_t(z)\Big),$$

where

$$A_L := \left\{v^* \in X^* \,\middle|\, (v^*, \langle v^*, z\rangle) \in \left[\overline{\mathrm{co}}\Big((L^\perp \times \mathbb{R}_+) \cup \Big(\bigcup_{t\in T} \operatorname{epi} f_t^*\Big)\Big)\right]_\infty\right\},$$
$$B_L := \left\{v^* \in X^* \,\middle|\, (v^*, \langle v^*, z\rangle) \in \left[\overline{\mathrm{co}}\Big((L^\perp \times \{0\}) \cup \Big(\bigcup_{t\in T} \operatorname{gph} f_t^*\Big)\Big)\right]_\infty\right\}.$$

Proof. According to Theorem 4 it suffices to write $\mathrm{N}_{L\cap\operatorname{dom} f}(z)$ in terms of the data functions f_t for each $L \in \mathcal{F}_z$. Indeed, by Lemma 5 applied to the family $\{f_t \mid t \in T\} \cup \{\mathrm{I}_L\} \subset \Gamma(X)$, we have $\mathrm{N}_{L\cap\operatorname{dom} f}(z) = A_L = B_L$; we used the fact that $(\mathrm{I}_L)^* = \mathrm{I}_{L^\perp}$, and so $\operatorname{epi}(\mathrm{I}_L)^* = \operatorname{epi}(\mathrm{I}_{L^\perp}) = L^\perp \times \mathbb{R}_+$ and $\operatorname{gph}(\mathrm{I}_L)^* = \operatorname{gph}(\mathrm{I}_{L^\perp}) = L^\perp \times \{0\}$. ∎

In the affine case (f_t affine) our formula takes a simpler form.

COROLLARY 7. *Assume that $T \neq \emptyset$ and $f := \sup\{\langle a_t^*, \cdot\rangle - \beta_t \mid t \in T\}$, with $a_t^* \in X^*$ and $\beta_t \in \mathbb{R}$. Then, for every $z \in X$, we have that*

$$\partial f(z) = \bigcap_{L\in\mathcal{F}_z,\varepsilon>0} \operatorname{cl}\left(\operatorname{co}\{a_t^* \mid t \in T_\varepsilon(z)\} + B_L\right),$$

where $T_\varepsilon(z) := \{t \in T \mid \langle a_t^*, z\rangle - \beta_t \geq f(z) - \varepsilon\}$ and

$$B_L := \left\{v^* \in X^* \mid (v^*, \langle v^*, z\rangle) \in \left[\overline{\mathrm{co}}\left((L^\perp \times \{0\}) \cup \{(a_t^*, \beta_t) \mid t \in T\}\right)\right]_\infty\right\}.$$

In particular, for a given nonempty set $A \subset X^*$, we have that

$$\partial\sigma_A(z) = \bigcap_{L\in\mathcal{F}_z,\varepsilon>0} \operatorname{cl}\left(\operatorname{co}(A_\varepsilon) + \left[\overline{\mathrm{co}}\left(L^\perp \cup A\right)\right]_\infty \cap \{z\}^\perp\right),$$

where $A_\varepsilon := \{a^* \in A \mid \langle z, a^*\rangle \geq \sigma_A(z) - \varepsilon\}$.

Proof. These formulae easily follow by Theorem 6, similarly as in [7, Prop. 1]. ∎

The following corollary gives us a simplified representation for the subdifferential set of f when $\operatorname{ri}(\operatorname{dom} f) \neq \emptyset$. This is also an extension of Lemma 3 when the functions f_t are not necessarily lsc.

COROLLARY 8. *Let $\{f_t \mid t \in T\}$ be a nonempty family of convex functions $f_t : X \to \overline{\mathbb{R}}$, and set $f := \sup_{t\in T} f_t$. Assume that* $\mathrm{ri}(\mathrm{dom}\, f) \neq \emptyset$. *Then, for every $z \in X$ and $\alpha > 0$, we have that*

$$\partial f(z) = \bigcap_{\varepsilon>0} \mathrm{cl}\left(\mathrm{co}\left(\bigcup_{t\in T_\varepsilon(z)} \partial_{\alpha\varepsilon} f_t(z)\right) + \mathrm{N}_{\mathrm{dom}\, f}(z)\right).$$

Proof. The inclusion "$\supset$" follows immediately by Theorem 4, since we have $\mathrm{N}_{\mathrm{dom}\, f}(z) \subset \mathrm{N}_{L\cap \mathrm{dom}\, f}(z)$ for every $L \in \mathcal{F}_z$. To prove the inclusion "$\subset$", let $\alpha > 0$ be fixed, and let $\partial f(z) \neq \emptyset$ (otherwise the inclusion is obvious). We (may) assume that $z = \theta$ and $f(\theta) = 0$. Then it suffices to show that $\partial f(\theta) \subset \mathrm{cl}\left(\mathrm{co}\left(\bigcup_{t\in T_\varepsilon(\theta)} \partial_{\alpha\varepsilon} f_t(\theta)\right) + \mathrm{N}_{\mathrm{dom}\, f}(\theta)\right)$ for any given $\varepsilon > 0$. Let $V \in \mathcal{V}$, that is, V is a θ–neighborhood in X^*, and $L \in \mathcal{F}_\theta$ be such that $L^\perp \subset V$. We may suppose w.l.o.g. that $L \cap \mathrm{ri}(\mathrm{dom}\, f) \neq \emptyset$, which in particular, implies that $L \cap \mathrm{ri}(\mathbb{R}_+ \mathrm{dom}\, f) \neq \emptyset$. Using (4) we obtain that $\mathrm{cl}\,(L \cap \mathbb{R}_+ \mathrm{dom}\, f) = L \cap \mathrm{cl}(\mathbb{R}_+ \mathrm{dom}\, f)$; this implies that (see [32, p. 7])

$$\mathrm{N}_{L\cap \mathrm{dom}\, f}(\theta) = (L \cap \mathrm{cl}(\mathbb{R}_+ \mathrm{dom}\, f))^- = \mathrm{cl}(L^- + (\mathbb{R}_+ \mathrm{dom}\, f)^-) = \mathrm{cl}\left(L^\perp + \mathrm{N}_{\mathrm{dom}\, f}(\theta)\right).$$

So, by using once again Theorem 4 and (3), we obtain that

$$\begin{aligned} \partial f(\theta) &\subset \mathrm{cl}\left[\mathrm{co}\left(\textstyle\bigcup_{t\in T_\varepsilon(\theta)} \partial_{\alpha\varepsilon} f_t(\theta)\right) + \mathrm{N}_{L\cap \mathrm{dom}\, f}(\theta)\right] \\ &= \mathrm{cl}\left[\mathrm{co}\left(\textstyle\bigcup_{t\in T_\varepsilon(\theta)} \partial_{\alpha\varepsilon} f_t(\theta)\right) + L^\perp + \mathrm{N}_{\mathrm{dom}\, f}(\theta)\right] \\ &\subset \mathrm{co}\left(\textstyle\bigcup_{t\in T_\varepsilon(\theta)} \partial_{\alpha\varepsilon} f_t(\theta)\right) + \mathrm{N}_{\mathrm{dom}\, f}(\theta) + V. \end{aligned}$$

As V is an arbitrary θ–neighborhood, we get that

$$\begin{aligned} \partial f(\theta) &\subset \bigcap_{V\in\mathcal{V}} \left(\mathrm{co}\left(\bigcup_{t\in T_\varepsilon(\theta)} \partial_{\alpha\varepsilon} f_t(\theta)\right) + \mathrm{N}_{\mathrm{dom}\, f}(\theta) + V\right) \\ &= \mathrm{cl}\left(\mathrm{co}\left(\textstyle\bigcup_{t\in T_\varepsilon(\theta)} \partial_{\alpha\varepsilon} f_t(\theta)\right) + \mathrm{N}_{\mathrm{dom}\, f}(\theta)\right), \end{aligned}$$

which finishes the proof. □

From a geometric point of view the closedness criterion given in Theorem 4 is equivalent to

$$\mathrm{cl}\left(\bigcap_{t\in T} \mathrm{epi}\, f_t\right) = \bigcap_{t\in T} \mathrm{cl}\,(\mathrm{epi}\, f_t)\,, \tag{25}$$

which is itself satisfied by a wide variety of convex functions as the following result shows.

COROLLARY 9. *Let $\{f_t \mid t \in T\}$ be a nonempty family of convex functions $f_t : X \to \overline{\mathbb{R}}$, and set $f := \sup_{t\in T} f_t$. Assume that one of the following conditions holds*:

(i) *All of the functions f_t, with $t \in T$ are lsc.*

(ii) *There exists $x_0 \in \mathrm{dom}\, f$ such that f_t is continuous at x_0 for every $t \in T$.*

(iii) *$T := \{1, \ldots, k, k + 1\}$, and there exists $x_0 \in \cap_{i=1}^{k+1} \mathrm{dom}\, f_i$ such that $f_1, \ldots, f_k$ are continuous at x_0.*

(iv) *$X = \mathbb{R}^n$ and $\mathrm{dom}\, f \cap (\cap_{t\in T} \mathrm{ri}(\mathrm{dom}\, f_t))$ is nonempty.*

Then, we have that

$$\mathrm{cl}\, f = \sup\{\mathrm{cl}\, f_t \mid t \in T\},$$

and, consequently, for every $z \in X$ and $\alpha > 0$, it holds that

$$\partial f(z) = \bigcap_{L \in \mathcal{F}_z, \varepsilon > 0} \operatorname{cl}\left(\operatorname{co}\left(\bigcup_{t \in T_\varepsilon(z)} \partial_{\alpha\varepsilon} f_t(z) \right) + \mathrm{N}_{L \cap \operatorname{dom} f}(z) \right).$$

Proof. Setting $A_t := \operatorname{epi} f_t$ for $t \in T$ and $A := \operatorname{epi} f$, one has always $A = \cap_{t \in T} A_t$, and we have to show that $\operatorname{cl} A = \cap_{t \in T} \operatorname{cl}(A_t)$. The inclusion $\operatorname{cl} A \subset \cap_{t \in T} \operatorname{cl}(A_t)$ being obvious, it remains to prove that $\operatorname{cl} A \supset \cap_{t \in T} \operatorname{cl}(A_t)$ in each of the following cases.

(i) It is immediate.

(ii) First observe that [31, Lem. 13] is valid even if f is not proper. Consider $\mu > f(x_0)$. Applying this result we obtain that $y_0 := (x_0, \mu) \in \bigcap_{t \in T} \operatorname{int} A_t$. Now if $x \in \bigcap_{t \in T} \operatorname{cl} A_t$, then $(1 - \lambda)x + \lambda y_0 \in \bigcap_{t \in T} \operatorname{int} A_t \subset A$ for every $\lambda \in]0, 1[$, whence $x \in \operatorname{cl} A$.

(iii) Set $B := \bigcap_{t=1}^{k} A_t$. Then, similarly as in (ii), we can show that $y_0 := (x_0, \mu) \in A_{k+1} \cap \operatorname{int} B$. Hence

$$\operatorname{cl}\left(\bigcap_{t \in T} A_t\right) = \operatorname{cl}(A_{k+1} \cap B) = \operatorname{cl} A_{k+1} \cap \operatorname{cl} B = \operatorname{cl} A_{k+1} \cap \left(\bigcap_{t=1}^{k} \operatorname{cl} A_t\right) = \bigcap_{t \in T} \operatorname{cl} A_t.$$

(iv) This is practically [22, Thm. 9.4].

Taking into account Theorem 4, the final conclusion follows. ∎

The following result (for $\alpha = 1$) is due to Volle (see, e.g., [28, Thm. A]) and is originally established in the context of normed spaces.

COROLLARY 10. *Let $\{f_t \mid t \in T\}$ be a nonempty family of convex functions $f_t : X \to \overline{\mathbb{R}}$, and set $f := \sup_{t \in T} f_t$. Assume that f is finite and continuous at $z \in X$. Then, we have*

$$\partial f(z) = \bigcap_{\varepsilon > 0} \overline{\operatorname{co}}\left(\bigcup_{t \in T_\varepsilon(z)} \partial_{\alpha\varepsilon} f_t(z)\right) \quad \textit{for all } \alpha > 0.$$

Proof. Because f is finite and continuous at z, we have that $z \in \operatorname{int}(\operatorname{dom} f)$, and so $\mathrm{N}_{\operatorname{dom} f}(z) = \{\theta\}$. Further, as $z \in \cap_{t \in T} \operatorname{int}(\operatorname{dom} f)$ Condition (ii) of Corollary 9 yields $\operatorname{cl} f = \sup\{\operatorname{cl} f_t \mid t \in T\}$. Of course, $\operatorname{ri}(\operatorname{dom} f) = \operatorname{int}(\operatorname{dom} f) \neq \emptyset$, and so the conclusion follows from Corollary 8. ∎

In order to derive Brøndsted's formula (Corollary 12 below) we shall need the following result on normal cones.

LEMMA 11. (i) *Let $g_1, \ldots, g_k \in \Gamma(X)$, $f \in \Gamma(Y)$, and consider a continuous affine mapping $A : X \to Y$, where X and Y are (separated) locally convex spaces. Then, for every $z \in \operatorname{dom}(g_1 + \cdots + g_k + f \circ A)$ and all $\varepsilon, \varepsilon_1, \ldots, \varepsilon_k > 0$, we have that*

$$\mathrm{N}_{\operatorname{dom}(g_1 + \cdots + g_k + f \circ A)}(z) = \left[\operatorname{cl}\left(\partial_{\varepsilon_1} g_1(z) + \cdots + \partial_{\varepsilon_k} g_k(z) + A_0^* \partial_\varepsilon f(Az)\right)\right]_\infty,$$

where A_0 is the linear part of A, and A_0^* is the adjoint of A_0.

(ii) *Let $\{f_1, \ldots, f_m\} \subset \Gamma(X)$, with $m \geq 2$ and $0 \leq k \leq m$. Then, for all $z \in \bigcap_{t=1}^{m} \operatorname{dom} f_t$ and all $\varepsilon_1, \ldots, \varepsilon_m > 0$, we have that*

$$\mathrm{N}_{\cap_{t=1}^{m} \operatorname{dom} f_t}(z) = \left[\operatorname{cl}\left(\partial_{\varepsilon_1} f_1(z) + \cdots + \partial_{\varepsilon_k} f_k(z) + \operatorname{co}\left(\partial_{\varepsilon_{k+1}} f_{k+1}(z) \cup \cdots \cup \partial_{\varepsilon_m} f_m(z)\right)\right)\right]_\infty,$$

where $C_1 + \cdots + C_k := \emptyset$ if $k = 0$ and $C_{k+1} \cup \cdots \cup C_m := \emptyset$ if $k = m$.

Proof. (i) Using (7) and (13), as well as the fact that $\mathbb{R}_+(B\cap C)=\mathbb{R}_+B\cap\mathbb{R}_+C$ when B and C are convex sets containing θ, we get that

$$\begin{aligned}&[(\mathrm{cl}\,(\partial_{\varepsilon_1}g_1(z)+\cdots+\partial_{\varepsilon_k}g_k(z)+A_0^*\partial_\varepsilon f(Az)))_\infty]^-\\&\quad=\mathrm{cl}(\mathrm{dom}(\sigma_{\partial_{\varepsilon_1}g_1(z)}+\cdots+\sigma_{\partial_{\varepsilon_k}g_k(z)}+\sigma_{\partial_\varepsilon f(Az)}\circ A_0))\\&\quad=\mathrm{cl}\left(\mathrm{dom}((g_1)'_{\varepsilon_1}(z,\cdot))\cap\cdots\cap\mathrm{dom}((g_k)'_{\varepsilon_k}(z,\cdot))\cap A_0^{-1}\,\mathrm{dom}(f'_\varepsilon(Az,\cdot))\right)\\&\quad=\mathrm{cl}\left(\mathbb{R}_+(\mathrm{dom}\,g_1-z)\cap\cdots\cap\mathbb{R}_+(\mathrm{dom}\,g_k-z)\cap A_0^{-1}(\mathbb{R}_+(\mathrm{dom}\,f-Az))\right)\\&\quad=\mathrm{cl}\left(\mathbb{R}_+(\mathrm{dom}(g_1+\cdots+g_k+f\circ A)-z)\right),\end{aligned}$$

whence the conclusion follows using (5).

(ii) Taking $f=0$ in (i) and observing that $\mathrm{dom}(g_1+\cdots+g_k)=\bigcap_{t=1}^k\mathrm{dom}\,g_t$, we get that $\mathrm{N}_{\cap_{t=1}^k\mathrm{dom}\,g_t}(z)=[\mathrm{cl}\,(\partial_{\varepsilon_1}g_1(z)+\cdots+\partial_{\varepsilon_k}g_k(z))]_\infty$. The conclusion follows now using (8). ∎

The following result is due to Brøndsted (e.g., [2]); see also [7, Prop. 7] where such a formula is extended to families of infinitely many convex functions defined on $\mathbb{R}^n$.

Corollary 12. *Consider the convex functions $f_i:X\to\overline{\mathbb{R}}$ for $i=1,\ldots,k$, and set $f:=\max\{f_1,\ldots,f_k\}$. Assume that*

$$\mathrm{cl}\,f=\max\{\mathrm{cl}\,f_1,\ldots,\mathrm{cl}\,f_k\}.$$

Given $z\in X$ such that $(\mathrm{cl}\,f)(z)=(\mathrm{cl}\,f_i)(z)$ for $i=1,\ldots,k$, we have that

$$\partial f(z)=\bigcap_{\varepsilon>0}\overline{\mathrm{co}}\left(\bigcup_{i=1}^k\partial_\varepsilon f_i(z)\right).$$

Proof. It suffices to establish the inclusion "⊂" in the nontrivial case $\partial f(z)\neq\emptyset$. According to (11), the function f is proper and satisfies $f(z)=(\mathrm{cl}\,f)(z)\in\mathbb{R}$ and $\partial f(z)=\partial\,(\mathrm{cl}\,f)\,(z)$. Because

$$(\mathrm{cl}\,f_i)(z)\le f_i(z)\le f(z)=(\mathrm{cl}\,f)(z)=(\mathrm{cl}\,f_i)(z),$$

we obtain that $(\mathrm{cl}\,f_i)(z)=f_i(z)=f(z)\in\mathbb{R}$ for all $i\in T:=\{1,\ldots,k\}$; hence the functions $\mathrm{cl}\,f_i$, with $i\in T$, are proper. Furthermore, using (10) we get

$$\partial_\varepsilon(\mathrm{cl}\,f_i)(z)=\partial_\varepsilon f_i(z)\text{ for all }\varepsilon>0\text{ and }i\in T.\tag{26}$$

Fix $\varepsilon>0$; it is clear that $T_\varepsilon(z)=T$. Let $V\in\mathcal{V}$, that is, V is a convex θ–neighborhood in X^*, and take $L\in\mathcal{F}_z$ such that $L^\perp\subset V$ ($\Leftrightarrow L^\perp\subset\frac{1}{2}V$). Applying Theorem 4 for $\{\mathrm{cl}\,f_1,\ldots,\mathrm{cl}\,f_k\}$ and $\alpha=1$, we have that

$$\partial\,(\mathrm{cl}\,f)\,(z)\subset\mathrm{cl}\left(\mathrm{co}\left(\bigcup_{i\in T}\partial_\varepsilon(\mathrm{cl}\,f_i)(z)\right)+\mathrm{N}_{L\cap\mathrm{dom}(\mathrm{cl}\,f)}(z)\right).$$

But Lemma 11(ii) applied to $\{\mathrm{cl}\,f_1,\ldots,\mathrm{cl}\,f_k,\mathrm{I}_L\}$ implies that

$$\mathrm{N}_{L\cap\mathrm{dom}(\mathrm{cl}\,f)}(z)=\left[\overline{\mathrm{co}}\left(L^\perp+\left(\bigcup_{i\in T}\partial_\varepsilon(\mathrm{cl}\,f_i)(z)\right)\right)\right]_\infty,$$

where we used the property $\partial_\varepsilon\mathrm{I}_L(z)=L^\perp$. Thus, taking into account (3) and (26), we get that

$$\begin{aligned}\partial f(z)=\partial\,(\mathrm{cl}\,f)\,(z)&\subset\mathrm{cl}\left(\overline{\mathrm{co}}\left(\bigcup_{i\in T}\partial_\varepsilon(\mathrm{cl}\,f_i)(z)\right)+\left[\overline{\mathrm{co}}\left(L^\perp+\left(\bigcup_{i\in T}\partial_\varepsilon(\mathrm{cl}\,f_i)(z)\right)\right)\right]_\infty\right)\\&\subset\mathrm{cl}\left(\overline{\mathrm{co}}\left(L^\perp+\left(\bigcup_{i\in T}\partial_\varepsilon f_i(z)\right)\right)+\left[\overline{\mathrm{co}}\left(L^\perp+\left(\bigcup_{i\in T}\partial_\varepsilon f_i(z)\right)\right)\right]_\infty\right)\\&=\overline{\mathrm{co}}\left(L^\perp+\left(\bigcup_{i\in T}\partial_\varepsilon f_i(z)\right)\right)=\mathrm{cl}\left(L^\perp+\mathrm{co}\left(\bigcup_{i\in T}\partial_\varepsilon f_i(z)\right)\right)\\&\subset L^\perp+\mathrm{co}\left(\bigcup_{i\in T}\partial_\varepsilon f_i(z)\right)+\tfrac{1}{2}V\subset\mathrm{co}\left(\bigcup_{i\in T}\partial_\varepsilon f_i(z)\right)+V.\end{aligned}$$

Consequently,

$$\partial f(z) \subset \bigcap_{V\in\mathcal{V}} \left(\operatorname{co}\left(\bigcup_{i\in T}\partial_\varepsilon f_i(z)\right) + V\right) = \overline{\operatorname{co}}\left(\bigcup_{i\in T}\partial_\varepsilon f_i(z)\right).$$

Finally, the conclusion follows by taking the intersection over $\varepsilon > 0$. ∎

4. Other calculus rules. Throughout this section, we consider two convex functions $f : Y \to \overline{\mathbb{R}}$ and $g : X \to \overline{\mathbb{R}}$, where X and Y are (separated) real locally convex spaces, and a continuous affine mapping $A : X \to Y$ defined by

$$Ax = A_0x + b,$$

where A_0 is the linear part of A and $b \in Y$. We denote by A_0^* the adjoint operator of A_0.

We show that our rule given in Theorem 4, providing formulas for the subdifferential of the supremum function, also gives calculus rules for other operations expressed by means of the convex function $g+f\circ A$. The resulting formulas are not new, but our aim here is to highlight the unifying character of Theorem 4, which also yields alternative proofs that do not rely on the commonly used approach based on conjugation theory [23].

At the first stage, we derive in the following theorem a slight extension of the Hiriart-Urruty–Phelps formula [11]. This allows us to express the subdifferential of $g + f \circ A$ in terms of the approximate subdifferentials of f and g. For comparative purposes, when the involved spaces X and Y are Banach, this is equivalent to writing $\partial(g + f \circ A)$ in terms of the subdifferentials of the data functions at nearby points (e.g., [14], [18], and [25]).

THEOREM 13. *Let us consider two convex functions $f : Y \to \overline{\mathbb{R}}$ and $g : X \to \overline{\mathbb{R}}$, where X and Y are (separated) real locally convex spaces, and a continuous affine mapping $A : X \to Y$, i.e., $Ax = A_0x + b$, where A_0 is the linear part of A and $b \in Y$. Assume that the following holds (when it makes sense):*

$$\operatorname{cl}(g + f \circ A) = (\operatorname{cl} g) + (\operatorname{cl} f) \circ A.$$

Then, for every $z \in X$, we have that

$$\partial(g + f \circ A)(z) = \bigcap_{\varepsilon>0} \operatorname{cl}\left(\partial_\varepsilon g(z) + A_0^*\partial_\varepsilon f(Az)\right),$$

where A_0^* is the adjoint operator of A_0.

Proof. Let us set $\varphi := g + f \circ A$, and $\psi := (\operatorname{cl} g) + (\operatorname{cl} f) \circ A$. The inclusion "$\supset$" always holds, and consequently, it suffices to establish the opposite one when $\partial\varphi(z) \neq \emptyset$. In such a case, by (11) and the current assumption, we have

$$(\operatorname{cl} g)(z) + (\operatorname{cl} f)(Az) = (\operatorname{cl} \varphi)(z) = \varphi(z) = g(z) + f(Az) \in \mathbb{R},$$

and

$$\partial\varphi(z) = \partial(\operatorname{cl}\varphi)(z) = \partial((\operatorname{cl} g) + (\operatorname{cl} f) \circ A)(z) = \partial\psi(z). \tag{27}$$

Hence, $(\operatorname{cl} g)(z) = g(z) \in \mathbb{R}$ and $(\operatorname{cl} f)(Az) = f(Az) \in \mathbb{R}$, and so $\operatorname{cl} f \in \Gamma(Y)$ and $\operatorname{cl} g \in \Gamma(X)$. Furthermore, according to (10), for every $\varepsilon \geq 0$, one has $\partial_\varepsilon(\operatorname{cl} g)(z) = \partial_\varepsilon g(z)$ and $\partial_\varepsilon(\operatorname{cl} f)(Az) = \partial_\varepsilon f(Az)$.

Now, by the Legendre–Fenchel linearization of $\mathrm{cl}\, f$, we write that for every $x \in X$,

$$\begin{aligned}\psi(x) &= (\mathrm{cl}\, g)(x) + (\mathrm{cl}\, f)(Ax) \\ &= (\mathrm{cl}\, g)(x) + \sup\{\langle y^*, Ax\rangle - f^*(y^*) \mid y^* \in \mathrm{dom}\, f^*\} \\ &= \sup\{(\mathrm{cl}\, g)(x) + \langle A_0^* y^*, x\rangle + \langle y^*, b\rangle - f^*(y^*) \mid y^* \in \mathrm{dom}\, f^*\}.\end{aligned}$$

So, applying Theorem 4 (with $\alpha = 1$) together with Corollary 9(i),

$$\partial\psi(z) = \bigcap_{L\in\mathcal{F}_z, \varepsilon>0} \mathrm{cl}\left(\mathrm{co}\left(\bigcup_{y^*\in T_\varepsilon(z)} \left(\partial_\varepsilon(\mathrm{cl}\, g)(z) + A_0^* y^*\right)\right) + \mathrm{N}_{L\cap\mathrm{dom}\,\psi}(z)\right),$$

where, by (12),

$$\begin{aligned}T_\varepsilon(z) &= \{y^* \in Y^* \mid (\mathrm{cl}\, g)(z) + \langle A_0^* y^*, z\rangle + \langle y^*, b\rangle - f^*(y^*) \ge \psi(z) - \varepsilon\} \\ &= \{y^* \in Y^* \mid (\mathrm{cl}\, f)(Az) + f^*(y^*) \le \langle y^*, Az\rangle + \varepsilon\} = \partial_\varepsilon\,(\mathrm{cl}\, f)(Az).\end{aligned}$$

Hence

$$\partial\psi(z) = \bigcap_{L\in\mathcal{F}_z, \varepsilon>0} \mathrm{cl}\left(\partial_\varepsilon(\mathrm{cl}\, g)(z) + A_0^*\partial_\varepsilon f(Az) + \mathrm{N}_{L\cap\mathrm{dom}\,\psi}(z)\right).$$

Now let $V \in \mathcal{V}$ (that is, V is a convex θ–neighborhood in X^*), and let $L \in \mathcal{F}_z$ be such that $L^\perp \subset V$. Then, for every $\varepsilon > 0$, from Lemma 11(i) we get

$$\mathrm{N}_{L\cap\mathrm{dom}\,\psi}(z) = \left[\mathrm{cl}\left(\partial_\varepsilon(\mathrm{cl}\, g)(z) + A_0^*\partial_\varepsilon(\mathrm{cl}\, f)(Az) + L^\perp\right)\right]_\infty,$$

so that, by taking into account (3), (27) leads us to

$$\begin{aligned}\partial\varphi(z) = \partial\psi(z) &\subset \mathrm{cl}\Big(\mathrm{cl}\left(\partial_\varepsilon(\mathrm{cl}\, g)(z) + A_0^*\partial_\varepsilon\,(\mathrm{cl}\, f)(Az) + L^\perp\right) \\ &\qquad + \left[\mathrm{cl}\left(\partial_\varepsilon(\mathrm{cl}\, g)(z) + A_0^*\partial_\varepsilon\,(\mathrm{cl}\, f)(Az) + L^\perp\right)\right]_\infty\Big) \\ &= \mathrm{cl}\left(\partial_\varepsilon(\mathrm{cl}\, g)(z) + A_0^*\partial_\varepsilon\,(\mathrm{cl}\, f)(Az) + L^\perp\right) \\ &\subset \partial_\varepsilon(\mathrm{cl}\, g)(z) + A_0^*\partial_\varepsilon\,(\mathrm{cl}\, f)(Az) + V \\ &= \partial_\varepsilon g(z) + A_0^*\partial_\varepsilon f(Az) + V,\end{aligned}$$

and consequently,

$$\partial\varphi(z) \subset \textstyle\bigcap_{\varepsilon>0}\bigcap_{V\in\mathcal{V}} \left(\partial_\varepsilon g(z) + A_0^*\partial_\varepsilon f(Az) + V\right) = \bigcap_{\varepsilon>0} \mathrm{cl}\left(\partial_\varepsilon g(z) + A_0^*\partial_\varepsilon f(Az)\right).$$

The proof is complete. ☐

Taking f and g to be lsc in Theorem 13 we obtain the following result of Hiriart-Urruty–Phelps [9].

COROLLARY 14. *Let f, g, and A be as in Theorem* 13. *If f and g are, in addition, lsc, then for every $z \in X$, we have that*

$$\partial(g + f\circ A)(z) = \bigcap_{\varepsilon>0} \mathrm{cl}\left(\partial_\varepsilon g(z) + A_0^*\partial_\varepsilon f(Az)\right).$$

In Corollary 16 below we derive the well-known Moreau–Rockafellar's formula on the sum (e.g., [19], p. 47). But, first, we need the following lemma, which gives us information about the closure of convex functions. Its proof does not appeal to the framework of Fenchel duality.

LEMMA 15. *Let $f : Y \to \overline{\mathbb{R}}$ and $g : X \to \overline{\mathbb{R}}$ be convex functions, and $A : X \to Y$ be a continuous affine mapping. Assume that f is finite and continuous at Ax_0 for some $x_0 \in (\operatorname{dom} g) \cap A^{-1}(\operatorname{dom} f)$. Then*

$$\operatorname{cl}(f \circ A + g) = (\operatorname{cl} f) \circ A + (\operatorname{cl} g).$$

Proof. Because $\operatorname{cl} f \le f$, $\operatorname{cl} g \le g$, and $(\operatorname{cl} f) \circ A + (\operatorname{cl} g)$ is lsc, one has $(\operatorname{cl} f) \circ A + (\operatorname{cl} g) \le \operatorname{cl}(f \circ A + g)$. Moreover, in our hypothesis f and $\operatorname{cl} f$ are proper. To establish the converse inequality it suffices to take

$$x \in (\operatorname{dom}(\operatorname{cl} g)) \cap A^{-1}(\operatorname{dom}(\operatorname{cl} f)) \subset (\operatorname{dom}(\operatorname{cl} g)) \cap A^{-1}(\operatorname{cl}(\operatorname{dom} f))$$

such that $(\operatorname{cl}(f \circ A + g))(x) > -\infty$.

Let us fix $\lambda \in]0, 1[$ and set $x_\lambda := \lambda x_0 + (1-\lambda)x$ $[\in (\operatorname{dom}(\operatorname{cl} g)) \cap A^{-1}(\operatorname{cl}(\operatorname{dom} f))]$. Since $Ax_0 \in \operatorname{int}(\operatorname{dom} f)$ and $Ax \in \operatorname{cl}(\operatorname{dom} f)$, (4) yields

$$Ax_\lambda = A(\lambda x_0 + (1-\lambda)x) = \lambda Ax_0 + (1-\lambda)Ax \in \operatorname{int}(\operatorname{dom} f),$$

and so f is continuous at Ax_λ. Now let $(x_i)_{i\in I} \subset X$ be a net which converges to x and satisfies $(\operatorname{cl} g)(x_\lambda) = \lim_i g(\lambda x_0 + (1-\lambda)x_i)$. Since $\lim_i f(\lambda Ax_0 + (1-\lambda)Ax_i) = f(Ax_\lambda) = (\operatorname{cl} f)(Ax_\lambda)$, we obtain that

$$\begin{aligned}(\operatorname{cl}(f \circ A + g))(x_\lambda) &\le \liminf_i \big(f(\lambda Ax_0 + (1-\lambda)Ax_i) + g(\lambda x_0 + (1-\lambda)x_i)\big) \\ &= (\operatorname{cl} f)(\lambda Ax_0 + (1-\lambda)Ax) + (\operatorname{cl} g)(x_\lambda) \\ &\le \lambda\big((\operatorname{cl} f)(Ax_0) + (\operatorname{cl} g)(x_0)\big) + (1-\lambda)\big((\operatorname{cl} f)(Ax) + (\operatorname{cl} g)(x)\big).\end{aligned}$$

Whence, as $\lambda \downarrow 0$ we get

$$\liminf_{\lambda\to 0}(\operatorname{cl}(f \circ A + g))(x_\lambda) \le (\operatorname{cl} f)(Ax) + (\operatorname{cl} g)(x),$$

and so $(\operatorname{cl}(f \circ A + g))(x) \le (\operatorname{cl} f)(Ax) + (\operatorname{cl} g)(x)$. The proof is complete. □

COROLLARY 16. *Let $f : Y \to \overline{\mathbb{R}}$ and $g : X \to \overline{\mathbb{R}}$ be convex functions, and $A : X \to Y$ be a continuous affine mapping with linear part A_0. Assume that f is finite and continuous at Ax_0 for some $x_0 \in (\operatorname{dom} g) \cap A^{-1}(\operatorname{dom} f)$. Then, for every $z \in X$, we have that*

$$\partial(f \circ A + g)(z) = A_0^* \partial f(Az) + \partial g(z).$$

Proof. It is enough to show that $\partial(f \circ A + g)(z) \subset A_0^* \partial f(Az) + \partial g(z)$. Taking into account Theorem 13 and Lemma 15, it suffices to prove that

$$\bigcap_{\varepsilon>0} \operatorname{cl}\big(A_0^* \partial_\varepsilon f(Az) + \partial_\varepsilon g(z)\big) \subset A_0^* \partial f(Az) + \partial g(z) \tag{28}$$

for the nontrivial case $\partial(g + f \circ A)(z) \ne \emptyset$; hence $z \in (\operatorname{dom} g) \cap A^{-1}(\operatorname{dom} f)$ and $g(z), f(Az) \in \mathbb{R}$.

Indeed, for x^* in the set from the left-hand side of (28) and for each $r = 1, 2, \ldots,$ there are nets $(v_i^*)_{i\in I} \subset \partial_{1/r} f(Az)$ and $(u_i^*)_{i\in I} \subset \partial_{1/r} g(z)$ such that $u_i^* + A_0^* v_i^* \to x^*$; thus we may assume that, for every $i \in I$,

$$\langle u_i^* + A_0^* v_i^*, z - x_0\rangle \le \langle x^*, z - x_0\rangle + 1.$$

Since $u_i^* \in \partial_{1/r} g(z)$ and $r \geq 1$, this implies that

$$\langle v_i^*, Az - Ax_0 \rangle \leq \langle u_i^*, x_0 - z \rangle + \langle x^*, z - x_0 \rangle + 1 \leq g(x_0) - g(z) + \langle x^*, z - x_0 \rangle + 2.$$

Because f is continuous at Ax_0, there exists a symmetric θ–neighborhood $U \subset Y$ such that $\sup_{y \in U} f(y + Ax_0) \leq f(Ax_0) + 1$. Hence, for all $y \in U$,

$$\begin{aligned} \langle v_i^*, y \rangle &= \langle v_i^*, Az - Ax_0 \rangle + \langle v_i^*, y + Ax_0 - Az \rangle \\ &\leq \langle v_i^*, Az - Ax_0 \rangle + f(y + Ax_0) - f(Az) + 1 \\ &\leq g(x_0) - g(z) + \langle x^*, z - x_0 \rangle + f(Ax_0) - f(Az) + 4 \leq \mu \end{aligned}$$

for some $\mu > 0$. This shows that $\inf\{\langle v_i^*, y \rangle \mid y \in U\} \geq -\mu$, and so $(v_i^*)_{i \in I} \subset (\mu^{-1}U)^\circ$. Hence, by Alaoglu–Bourbaki's Theorem we may suppose w.l.o.g. that $(v_i^*)_{i \in I}$ and $(u_i^*)_{i \in I}$ w^*–converge to some $v_r^* \in \partial_{1/r} f(Az) \cap (\mu^{-1}U)^\circ$ and $u_r^* \in \partial_{1/r} g(z)$, respectively, and so $x^* = u_r^* + A_0^* v_r^*$. By the same argument we may suppose that $(v_r^*)_r$ and $(u_r^*)_r$ also w^*–converge to some $v^* \in \partial f(Az)$ and $u^* \in \partial g(z)$ and $x^* = u^* + A_0^* v^* \in \partial g(z) + A_0^* \partial f(Az)$. The proof is complete. □

Concluding remarks. (1) The preceding proof still works under more general regularity conditions, as those studied in Theorem 2.8.3 of [32].

(2) It should be noted that Lemma 5 can be easily deduced from Corollary 2.6.3 of [32], which is itself an extension of Corollary 14.

(3) Our main result in section 3 gives the formula for the subdifferential of the pointwise supremum $f := \sup_{t \in T} f_t$ of an arbitrary family of convex functions $f_t : X \to \overline{\mathbb{R}}$, $t \in T$. An important special case, which commonly appears in applications, corresponds to the so-called continuous model (e.g., [13], [24], and [32, Thm. 2.4.18]); see also [6]. There, the index set T is a (separated) compact space, and the parametrized mappings $t \to f_t(x)$ are upper semicontinuous for every $x \in X$. Such a situation is intermediate between the finite ([29]) and the general cases, and it is approached in a forthcoming paper.

(4) For further examples (in $\mathbb{R}^n$) in relation with our formula given in Theorem 4, the reader is addressed to references [6] and [7].

REFERENCES

[1] A. Auslender and M. Teboulle, *Asymptotic Cones and Functions in Optimization and Variational Inequalities*, Springer, New York, 2003.

[2] A. Brøndsted, *On the subdifferential of the supremum of two convex functions*, Math. Scand., 31 (1972), pp. 225–230.

[3] J. M. Danskin, *The Theory of Max-Min and its Applications to Weapons Allocations Problems*, Springer, New York, 1967.

[4] M. A. Goberna and M. A. López, *Optimal value function in semi-infinite programming*, J. Optim. Theory Appl., 59 (1988), pp. 261–279.

[5] M. A. Goberna and M. A. López, *Linear Semi-Infinite Optimization*, John Wiley and Sons, Chichester, UK, 1998.

[6] A. Hantoute, *Subdifferential set of the supremum of lower semicontinuous convex functions and the conical hull property*, Top, 14 (2006), pp. 355–374.

[7] A. Hantoute and M. A. López, *A complete characterization of the subdifferential set of the supremum of an arbitrary family of convex functions*, J. Convex Anal., 15 (2008), to appear.

[8] A. Hantoute and M. A. López, *Characterization of total ill-posedness in semi-infinite linear optimization*, J. Comput. Appl. Math., 217 (2008), pp. 350–364.

[9] J.-B. Hiriart-Urruty and R. R. Phelps, *Subdifferential calculus using ε-subdifferentials*, J. Funct. Anal., 118 (1993), pp. 154–166.

[10] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms* I, II, Springer, Berlin, 1993.

[11] J.-B. Hiriart-Urruty, M. Moussaoui, A. Seeger, and M. Volle, *Subdifferential calculus without qualification conditions, using approximate subdifferentials: A survey*, Nonlinear Anal., 24 (1995), pp. 1727–1754.

[12] A. D. Ioffe and V. L. Levin, *Subdifferentials of convex functions*, Trudy Mos. Mat. Obs., 26 (1972), pp. 3–73 (Russian).

[13] A. D. Ioffe and V. H. Tikhomirov, *Theory of Extremal Problems*, in Stud. Math. Appl. 6, North-Holland, Amsterdam, 1979.

[14] F. Jules and M. Lassonde, *Formulas for subdifferentials of sums of convex functions*, J. Convex Anal., 9 (2002), pp. 519–533.

[15] V. L. Levin, *An application of Helly's theorem in convex programming, problems of best approximation and related questions*, Mat. Sb., 79(121) (1969), pp. 250–263. Math. USSR, Sb. 8 (1969), pp. 235–247 (in English).

[16] D. T. Luc, *Recession cones and the domination property in vector optimization*, Math. Program., 49 (1990), pp. 113–122.

[17] J.-J. Moreau, *Fonctionnelles Convexes*, Instituto Poligrafico e Zecca dello Stato, Rome, Italy, 2003.

[18] J.-P. Penot, *Subdifferential calculus without qualification assumptions*, J. Convex Anal., 3 (1996), pp. 207–219.

[19] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, 2nd ed., Lecture Notes in Math. 1364, Springer, Berlin, 1993.

[20] B. N. Pschenichnyi, *Convex programming in a normalized space*, Kibern., 5 (1965), pp. 46–54 (in Russian); translated as Cybern., 1 (1965), pp. 46–57.

[21] R. T. Rockafellar, *Directionally Lipschitzian functions and subdifferential calculus*, Proc. London Math. Soc., 39 (1979), pp. 331–355.

[22] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, N.J., 1970.

[23] R. T. Rockafellar, *Conjugate Duality and Optimization*, in CBMS Reg. Conf. Ser. Appl. Math. 16, SIAM, Philadelphia, 1974.

[24] V. N. Solov'ev, *The subdifferential and the directional derivatives of the maximum of a family of convex functions*, Izv. Ross Akad. Nauk Ser. Mat., 65 (2001), pp. 107–132.

[25] L. Thibault, *Sequential convex subdifferential calculus and sequential Lagrange multipliers*, SIAM J. Control Optim., 35 (1997), pp. 1434–1444.

[26] V. M. Tikhomirov, *Analysis* II*, Convex Analysis and Approximation Theory*, Encyclopedia Math. Sci. 14, R. V. Gamkrelidze, ed., Springer, New York, 1987.

[27] M. Valadier, *Sous-différentiels d'une borne supérieure et d'une somme continue de fonctions convexes*, C. R. Acad. Sci. Paris Sér. A-B Math., 268 (1969), pp. 39–42.

[28] M. Volle, *Sous-différentiel d'une enveloppe supérieure de fonctions convexes*, C. R. Acad. Sci. Paris Sér. I Math., 317 (1993), pp. 845–849.

[29] M. Volle, *On the subdifferential of an upper envelope of convex functions*, Acta Math. Vietnam., 19 (1994), pp. 137–148.

[30] C. Zălinescu, *Stability for a class of nonlinear optimization problems and applications*, in Nonsmooth Optimization and Related Topics, Ettore Majorana Internat. Sci. Ser. Phys. Sci., 43, Plenum, New York, 1989, pp. 437–458.

[31] C. Zălinescu, *On several results about convex set functions*, J. Math. Anal. Appl., 328 (2007), pp. 1451–1470.

[32] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, Singapore, 2002.