MATH. SCAND. 31 (1972), 225–230

ON THE SUBDIFFERENTIAL OF THE SUPREMUM OF TWO CONVEX FUNCTIONS

ARNE BRØNDSTED

1. Introduction.

Let (X,Y) be a pair of real vector spaces in duality under a non-degenerate bilinear function $(x,y) \to \langle x,y \rangle$, and let f be a convex function on X with values in $]-\infty, +\infty]$. Suppose that $x_0 \in X$ is a point where f has a finite value. Then a point $y_0 \in Y$ is called a *subgradient* of f at x_0, provided that

$$\forall x \in X: \; f(x) \geqq f(x_0) + \langle x - x_0, y_0 \rangle .$$

The (possibly empty) set of subgradients of f at x_0 is called the *subdifferential* of f at x_0 and is denoted $\partial f(x_0)$. If $\partial f(x_0)$ is non-empty, then f is said to be *subdifferentiable* at x_0.

A locally convex, necessarily separated, vector space topology on X is said to be *compatible* (with the duality), if the continuous linear forms on X are the functions $x \to \langle x,y \rangle$, where $y \in Y$. Similarly we may speak of compatible topologies on Y. All compatible topologies on one of the spaces X and Y determine the same closed convex subsets of the space, and, as a consequence, the same lower semi-continuous convex functions on the space. In particular, one may speak of such sets and such functions without specifying the topology as long as it is compatible,—which we shall always assume.

With the preceding in mind we may state that the subdifferential $\partial f(x_0)$ of a function f at a point x_0 is a closed convex set. Furthermore, note that if f and g are functions such that $f \leqq g$ and $f(x_0) = g(x_0) < +\infty$, then $\partial g(x_0) \supseteqq \partial f(x_0)$. These two facts clearly imply that for any two functions f_1 and f_2 such that $f_1(x_0) = f_2(x_0) < +\infty$, we have

$$\partial(\sup(f_1,f_2))(x_0) \supseteqq \operatorname{cl}\operatorname{conv}(\partial f_1(x_0) \cup \partial f_2(x_0)) ,$$

where "conv" denotes convex hull and "cl" denotes closure (under any compatible topology). This was noted by C. Lescarret in [1], and it was asked whether equality holds in general. The question was motivated by

Received October 20, 1971.

a theorem of M. Valadier, see [1], stating that if f_1 and f_2 are continuous at x_0 under some compatible topology, then in fact

$$\partial(\sup(f_1,f_2))(x_0) = \operatorname{conv}(\partial f_1(x_0) \cup \partial f_2(x_0)) .$$

(Taking closure of the convex hull of $\partial f_1(x_0)$ and $\partial f_2(x_0)$ is superfluous. For by a theorem of J. J. Moreau, the continuity implies $\sigma(Y,X)$-compactness of $\partial f_1(x_0)$ and $\partial f_2(x_0)$, whence the convex hull is $\sigma(Y,X)$-compact and therefore closed.) The answer to the question is negative, however. In fact, one can merely take $X = Y = \mathsf{R}$ with $\langle x,y\rangle = xy$, take $f_1(x) = -x^{\frac{1}{2}}$ for $x \geqq 0$ and $f_1(x) = +\infty$ for $x < 0$, and take $f_2(x) = f_1(-x)$. Then $\partial f_1(0) = \partial f_2(0) = \emptyset$, whereas $\partial(\sup(f_1,f_2))(0) = \mathsf{R}$.

By this example, the possibility of describing $\partial(\sup(f_1,f_2))(x_0)$ in terms of $\partial f_1(x_0)$ and $\partial f_2(x_0)$ for lower semi-continuous convex functions f_1 and f_2 is excluded. To compensate, we shall show that the *approximate subgradients* are of use. For $\varepsilon > 0$, a point $y_0 \in Y$ is called an *ε-subgradient* of a convex function f on X at a point $x_0 \in X$ where f has a finite value, provided that

$$\forall x \in X:\ f(x) \geqq f(x_0) - \varepsilon + \langle x - x_0, y_0\rangle .$$

The set of ε-subgradients of f at x_0 is denoted $\partial_\varepsilon f(x_0)$; it is a closed convex set, non-empty if f is lower semi-continuous. For functions f_1 and f_2 with $f_1(x_0) = f_2(x_0) < +\infty$ we clearly have

$$\partial_\varepsilon(\sup(f_1,f_2))(x_0) \supseteqq \operatorname{cl}\operatorname{conv}(\partial_\varepsilon f_1(x_0) \cup \partial_\varepsilon f_2(x_0))$$

for any $\varepsilon > 0$. Since for any function f

$$\bigcap\nolimits_{\varepsilon>0} \partial_\varepsilon f(x_0) = \partial f(x_0) ,$$

it follows that

$$\partial(\sup(f_1,f_2))(x_0) \supseteqq \bigcap\nolimits_{\varepsilon>0} \operatorname{cl}\operatorname{conv}(\partial_\varepsilon f_1(x_0) \cup \partial_\varepsilon f_2(x_0)) .$$

Our main result states that if f_1 and f_2 are lower semi-continuous, then in fact we have identity.

Concepts and results treated in the notes of J. J. Moreau [2] are used without reference to the original papers; such references can be found in [2]. The papers in question are due to J. J. Moreau, R. T. Rockafellar and the author.

The author is indebted to R. T. Rockafellar for some valuable comments, allowing improvements of an earlier version of this paper.

2. Main result.

An inspection of the proof of the theorem quoted above due to M. Valadier shows that if one is satisfied with a weaker conclusion, then a weaker assumption works. In fact, one has:

LEMMA (M. Valadier). *Let f_1 and f_2 be convex functions on X with values in $]-\infty, +\infty]$. Let x_0 be a point in X such that $f_1(x_0)=f_2(x_0)<+\infty$ and such that the directional derivatives $f_1'(x_0;\cdot)$ and $f_2'(x_0;\cdot)$ of f_1 and f_2 at x_0 are lower semi-continuous on X. Then*

$$\partial(\sup(f_1,f_2))(x_0) = \operatorname{cl}\operatorname{conv}(\partial f_1(x_0) \cup \partial f_2(x_0)) .$$

For completeness, we include the proof. Let $f_0=\sup(f_1,f_2)$. Then clearly

$$f_0'(x_0;\cdot) = \sup(f_1'(x_0;\cdot), f_2'(x_0;\cdot)) .$$

Denoting "greatest lower semi-continuous convex minorant" by the symbol $\wedge$, and using the lower semi-continuity of $f_1'(x_0;\cdot)$ and $f_2'(x_0;\cdot)$, we get by conjugation

$$\begin{aligned} f_0'(x_0;\cdot)^* &= \big(\sup(f_1'(x_0;\cdot), f_2'(x_0;\cdot))\big)^* \\ &= \wedge\big(f_1'(x_0;\cdot)^*, f_2'(x_0;\cdot)^*\big) . \end{aligned}$$

Now, for any convex function f we have

$$f'(x_0;\cdot)^* = \psi(\partial f(x_0);\cdot) ,$$

where $\psi(\partial f(x_0);\cdot)$ is the indicator function of $\partial f(x_0)$. We therefore get

$$\begin{aligned} \psi(\partial f_0(x_0);\cdot) &= \wedge\big(\psi(\partial f_1(x_0);\cdot), \psi(\partial f_2(x_0);\cdot)\big) \\ &= \psi\big(\operatorname{cl}\operatorname{conv}(\partial f_1(x_0) \cup \partial f_2(x_0));\cdot\big) , \end{aligned}$$

which gives the desired conclusion. (To obtain M. Valadier's theorem from the lemma, use the facts that continuity of f_i at x_0 implies continuity of $f_i'(x_0;\cdot)$ as well as $\sigma(Y,X)$-compactness of $\partial f_i(x_0)$; cf. a remark in section 1.)

Having established this preliminary result, we shall then pass to the main result:

THEOREM. *Let f_1 and f_2 be lower semi-continuous convex functions on X with values in $]-\infty, +\infty]$, and let x_0 be a point in X such that $f_1(x_0)= f_2(x_0)<+\infty$. Then*

$$\partial(\sup(f_1,f_2))(x_0) = \textstyle\bigcap_{\varepsilon>0} \operatorname{cl}\operatorname{conv}(\partial_\varepsilon f_1(x_0) \cup \partial_\varepsilon f_2(x_0)) .$$

Proof. By a remark in section 1, we shall prove that $\partial(\sup(f_1,f_2))(x_0)$ is contained in

$$\operatorname{cl}\operatorname{conv}(\partial_\varepsilon f_1(x_0) \cup \partial_\varepsilon f_2(x_0))$$

for each $\varepsilon > 0$. So, let $\varepsilon > 0$ be fixed. Clearly, there is no loss of generality in assuming $x_0 = o$ and $f_1(o) = f_2(o) = \varepsilon$. For $i = 1,2$ we put

$$g_i(x) = \sup\{\langle x,y\rangle \mid y \in \partial_\varepsilon f_i(o)\}$$

for $x \in X$, i.e. g_i is the support function of $\partial_\varepsilon f_i(o)$. So g_i is lower semi-continuous, convex and positively homogeneous. Being convex and positively homogeneous, g_i equals its directional derivative $g_i'(o;\,\cdot)$ at o. Invoking also the lower semi-continuity of g_i, this has two consequences. Firstly, we may apply the lemma to g_1 and g_2, thereby obtaining

$$\partial(\sup(g_1,g_2))(o) = \operatorname{cl}\operatorname{conv}(\partial g_1(o) \cup \partial g_2(o)) .$$

Secondly, it follows that g_i is the support function of $\partial g_i(o)$, whence $\partial g_i(o) = \partial_\varepsilon f_i(o)$. Consequently

$$\partial(\sup(g_1,g_2))(o) = \operatorname{cl}\operatorname{conv}(\partial_\varepsilon f_1(o) \cup \partial_\varepsilon f_2(o)) .$$

We shall complete the proof by proving the inclusion

$$(1) \qquad \partial(\sup(f_1,f_2))(o) \subseteqq \partial(\sup(g_1,g_2))(o) .$$

In other words, we shall prove that for any $y \in Y$,

$$(2) \qquad \forall x \in X\colon\ \sup(f_1(x),f_2(x)) \geqq \varepsilon + \langle x,y\rangle ,$$

implies

$$(3) \qquad \forall x \in X\colon\ \sup(g_1(x),g_2(x)) \geqq \langle x,y\rangle .$$

Suppose for a moment that (2) $\Rightarrow$ (3) has been proved when $y = o$, and let $y_0 \in Y$ be arbitrary. Let

$$\bar f_i(x) = f_i(x) - \langle x,y_0\rangle, \quad \bar g_i(x) = g_i(x) - \langle x,y_0\rangle ,$$

for $x \in X$, $i = 1,2$. Then $\bar f_i$ fulfils the same assumptions as f_i, and an easy calculation shows that we also have

$$\bar g_i(x) = \sup\{\langle x,y\rangle \mid y \in \partial_\varepsilon \bar f_i(o)\} .$$

Therefore, it follows that (2) $\Rightarrow$ (3) is valid for $\bar f_i$ and $\bar g_i$ when $y = o$. But this means that (2) $\Rightarrow$ (3) is valid for f_i and g_i when $y = y_0$. In conclusion, it suffices to prove (2) $\Rightarrow$ (3) when $y = o$.

So, assume that (2) holds for $y = o$, and consider an arbitrary $x_1 \in X$. Since $g_1(o) = g_2(o) = 0$, there is nothing to prove for $x_1 = o$. Therefore, let $x_1 \neq o$. Suppose that $f_i(\lambda_i x_1) < \varepsilon$ for some $\lambda_i \in]0,1]$, $i = 1,2$. Then, since $f_i(o)=\varepsilon$, we have $f_i(x)<\varepsilon$ for all $x \in]o,\lambda_i x_i]$ by the convexity of f_i. But this clearly contradicts the assumption. Hence, we have either $f_1(x) \geqq \varepsilon$ for all $x \in [o,x_1]$, or $f_2(x) \geqq \varepsilon$ for all $x \in [o,x_1]$. Assuming that $f_1(x) \geqq \varepsilon$ for all $x \in [o,x_1]$, it then follows that the epigraph of f_1, i.e. the set

$$\operatorname{epi} f_1 = \{(x,r) \in X \times \mathsf{R} \mid f_1(x) \leqq r\},$$

has no points in common with any of the segments

$$I_n = [(o,0),(x_1,\varepsilon - n^{-1})],$$

where $n \in \mathsf{N}$. Since f_1 is lower semi-continuous and convex, $\operatorname{epi} f_1$ is closed and convex. Hence, by a standard separation theorem, there exists for each $n \in \mathsf{N}$ a closed hyperplane in $X \times \mathsf{R}$ strictly separating $\operatorname{epi} f_1$ and I_n. In other words, for each $n \in \mathsf{N}$ there exists $y_n \in Y$ and $c_n \in \mathsf{R}$, such that

$$\forall t \in [0,1]:\ t(\varepsilon - n^{-1}) < \langle t x_1, y_n\rangle + c_n\,, \tag{4}$$

$$\forall x \in X:\ \langle x, y_n\rangle + c_n < f_1(x)\,. \tag{5}$$

By setting $t=0$ in (4), it follows that $0<c_n$. Since $f_1(o)=\varepsilon$, (5) then shows that $y_n \in \partial_\varepsilon f_1(o)$. From the definition of g_1 we thereby get

$$\sup_{n\in\mathsf{N}} \langle x_1, y_n\rangle \leqq g_1(x_1)\,. \tag{6}$$

On the other hand, by setting $t=1$ in (4) and $x=o$ in (5), it follows that $-n^{-1}<\langle x_1,y_n\rangle$ which implies

$$0 \leqq \sup_{n\in\mathsf{N}} \langle x_1, y_n\rangle\,. \tag{7}$$

Combining (6) and (7) we get $0 \leqq g_1(x_1)$, whence $0 \leqq \sup(g_1(x_1), g_2(x_1))$ as desired.

We would like to point out that it is possible to give a (slightly more direct) proof which avoids use of the notion of directional derivative. This becomes clear when one notes that the lemma is only used for positively homogeneous functions. The lemma seems, however, to be of independent interest, and for that reason we have preferred to put it in the foreground.

3. Remarks.

We note that the lemma and the theorem are easily extended to any finite family of functions f_i.

Regarding the possibility of extending the results to infinite families, note that for an infinite family $(f_i)_{i\in I}$ we need not have

$$f_0'(x_0;\ \cdot) = \sup_{i\in I} f_i'(x_0;\ \cdot)\,, \tag{8}$$

where f_0 denotes the supremum of the function f_i. With the additional condition (8), the proof of the lemma still works. On the other hand, if (8) is not fulfilled, then the proof shows that the lemma does not hold. The condition (8) is not sufficient to make the proof of the theorem work for an infinite family.

ADDENDUM. The author's attention has been called to M. Valadier's paper [3], in which the theorem of M. Valadier quoted in section 1 above is obtained as a special case of a more general result.

REFERENCES

1. C. Lescarret, *Sous-différentiabilité de fonctions composées,* (Travaux du Séminaire d'Analyse Unilatérale I), Université de Montpellier, 1968–1969.
2. J. J. Moreau, *Fonctionelles convexes,* (Séminaire sur les Equations aux Dérivées Partielles II), Collège de France, Paris, 1966–1967.
3. M. Valadier, *Sous-différentiels d'une borne supérieure et d'une somme continue de fonctions convexes,* C. R. Acad. Sci. Paris Sér. A 268 (1969), 39–42.

UNIVERSITY OF COPENHAGEN
DENMARK