MATH. SCAND. 34 (1974), 69–83

SUBDIFFERENTIABILITY OF CONVEX FUNCTIONS WITH VALUES IN AN ORDERED VECTOR SPACE

JOCHEM ZOWE[1]

Abstract.

It was shown by Valadier [8] that a convex function defined on a topological vector space X with values in a topological order complete vector lattice Y is subdifferentiable (even regularly subdifferentiable) at each point, where the function is continuous. We will prove that under some assumptions on X and the order cone C this even holds, if Y is an ordered topological vector space. Furthermore we will see that under our assumptions on X and C the Gateaux differentiability of a convex function is equivalent to the existence of only one subgradient. Our result apply e.g. if X is a separable reflexive Banach space and Y is a semireflexive locally convex space ordered by a cone with a weakly compact base.

1. Introduction and notations.

Throughout the following let X and Y be separated locally convex vector spaces over R and let Y be ordered by a closed convex proper cone C. We write $z \leqq y$ for $z,y \in Y$ if $y-z \in C$. With X', Y' we denote the topological duals of X and Y and with $\langle \cdot , \cdot \rangle$ the canonical bilinear forms on the dualities $\langle X,X' \rangle$ and $\langle Y,Y' \rangle$. Furthermore let $\sigma(\cdot,\cdot)$, $\tau(\cdot,\cdot)$, $\beta(\cdot,\cdot)$ stand for the weak, Mackey- and strong topologies with respect to the dual pairs $\langle X,X' \rangle$ and $\langle Y,Y' \rangle$. We write for example X'_σ for X' under $\sigma(X',X)$, $N\sigma(X',X)$ for the neighbourhood filter of 0 in X'_σ, A°_σ ($\bar{A}_\sigma$) for the interior (closure) of a set $A \subset X'_\sigma$ etc. If $A \subset X'$ is convex then $\bar{A}_\sigma = \bar{A}_\tau$ and we omit the subscript.

We will consider a function f mapping a nonvoid convex subset K of X into Y such that for all $x_1,x_2 \in K$ and $\lambda \in \mathsf{R}$, $0 \leq \lambda \leq 1$,

$$f(\lambda x_1 + (1-\lambda)x_2) \leqq \lambda f(x_1) + (1-\lambda) f(x_2) .$$

Received August 8, 1973.

[1] This paper was written during the author's stay in the summer 1973 as a guest at the Department of Mathematics of the University of Dundee, Scotland.

f is called a *convex function*. It is assumed throughout the following without further mentioning that

(1.1) *$x_0 \in K^\circ$ and f is continuous at x_0, when f is regarded as a mapping of X into Y_σ.*

What we are interested in, is the set

$$(1.2) \qquad \partial f(x_0) := \{T \in \mathscr{L}(X, Y) : T(x - x_0) \leqq f(x) - f(x_0) \text{ for all } x \in K\}$$

of *subgradients* of f at x_0 (here $\mathscr{L}(X, Y)$ stands for the set of continuous linear mappings from X into Y). $\partial f(x_0)$ is called the *subdifferential* of f at x_0. In the special case $Y = \mathsf{R}$ the subdifferential is a nonvoid convex compact subset of X'_σ (see [4]). Each y' in the dual cone of C

$$C' := \{y' \in Y' : \langle C, y' \rangle \geq 0\},$$

defines a convex functional

$$(1.3) \qquad y' \circ f(x) := \langle f(x), y' \rangle \quad \text{for all } x \in K.$$

Therefore

(1.4) $\partial(y' \circ f)(x_0)$, $y' \in C'$, *is a nonvoid convex compact subset of* X'_σ.

(Here

$$\partial(y' \circ f)(x_0) = \{x' \in X' : \langle x - x_0, x' \rangle \leq y' \circ f(x) - y' \circ f(x_0) \text{ for all } x \in K\}.)$$

Valadier [8] showed that an analoguous result holds for f itself, if Y is an order complete vector lattice. In his proof this assumption on Y is essentially used to show that $\partial f(x_0)$ is nonvoid. We are going to demonstrate that this result remains valid for an ordered vector space, if $(C')^\circ_\tau \neq \varnothing$. Roughly speaking, our idea is the following: First note that the transpose S of a $T \in \partial f(x_0)$ is a continuous linear mapping from Y'_σ into X'_σ and that $Sy' \in \partial(y' \circ f)(x_0)$ for all $y' \in C'$. Conversely, we will construct a continuous linear mapping S from Y'_σ into X'_σ with

$$Sy' \in \partial(y' \circ f)(x_0), \quad y' \in C',$$

and the transpose will belong to $\partial f(x_0)$. The existence of a $S: C' \to X'$ such that $Sy' \in \partial(y' \circ f)(x_0)$ is an easy consequence of (1.4). The important point is that under the assumption $(C')^\circ_\tau \neq \varnothing$ the mapping S can be chosen to be linear and continuous.

2. Auxiliary propositions.

For the proof of our principal auxiliary result, proposition 2.5, we need some further information about the sets $\partial(y' \circ f)(x_0)$, $y' \in C'$. First let us state some simple well-known facts. The *directional derivative* of $y' \circ f$, $y' \in C'$, at x_0 in the direction h,

$$(y' \circ f)'(x_0; h) = \lim \lambda^{-1}(y' \circ f(x_0 + \lambda h) - y' \circ f(x_0)), \quad \lambda \to 0,\ \lambda > 0\,,$$

is a positively homogeneous, subadditive functional in h, defined for all $h \in X$. We have

$$(2.1) \quad (y' \circ f)'(x_0; h) = \inf\{\lambda^{-1}(y' \circ f(x_0 + \lambda h) - y' \circ f(x_0)) : \lambda > 0, x_0 + \lambda h \in K\}\,,$$

and thus

PROPOSITION 2.1. *Let* $y' \in C'$. *Then* $x' \in \partial(y' \circ f)(x_0)$ *if and only if*

$$\langle h, x' \rangle \le (y' \circ f)'(x_0; h) \quad \textit{for all } h \in X\,.$$

PROPOSITION 2.2. *Let* $y' \in C'$ *and* $h_0 \in X$. *Then*

$$\langle h_0, x_0' \rangle = (y' \circ f)'(x_0; h_0)$$

for some $x_0' \in \partial(y' \circ f)(x_0)$.

PROOF. Consider the convex functional $p(h) := f'(x_0; h)$. We have

$$p(\pm h) \le y' \circ f(x_0 \pm h) - y' \circ f(x_0)$$

for h small enough such that $x_0 \pm h \in K$, and because of $p(h) + p(-h) \ge p(0) = 0$ we get

$$p(h) \le \max\{|y' \circ f(x_0 + h) - y' \circ f(x_0)|, |y' \circ f(x_0) - y' \circ f(x_0 - h)|\}\,.$$

Since $y' \circ f$ is continuous at x_0 by assumption (1.1), p must be continuous at 0. From

$$p(h_0 + h) \le p(h_0) + p(h) \quad \text{and} \quad p(h_0) \le p(h_0 + h) + p(-h)$$

we get

$$-p(-h) \le p(h_0 + h) - p(h_0) \le p(h)\,,$$

that is, p is continuous at h_0 as well. But then p has a subgradient at h_0, that is, for some $x_0' \in X'$

$$\langle h - h_0, x_0' \rangle \le p(h) - p(h_0) \quad \text{for all } h \in X\,.$$

$h = 0$ and $h = 2h_0$ show

$$\langle -h_0, x_0' \rangle \le -p(h_0), \quad \langle h_0, x_0' \rangle \le p(2h_0) - p(h_0) = p(h_0)\,,$$

and thus $\langle h_0,x_0'\rangle = p(h_0) = (y'\circ f)'(x_0;\, h_0)$ and

$$\langle h,x_0'\rangle \leq p(h) = (y'\circ f)'(x_0;\, h) \quad \text{for all } h \in X\,.$$

From proposition 2.1 we get $x_0' \in \partial(y'\circ f)(x_0)$.

Because of proposition 2.1 the hyperplane

$$\{x' \in X' : \langle h_0,x'\rangle = \langle h_0,x_0'\rangle\}$$

supports $\partial(y'\circ f)(x_0)$ at x_0' and consequently x_0' is a boundary point of $\partial(y'\circ f)(x_0)$. It will be a crucial fact for our construction below that under some assumptions on X there exist boundary points x' of $\partial(y'\circ f)(x_0)$ and supporting hyperplanes H with $H \cap \partial(y'\circ f)(x_0) = \{x'\}$. To this let us call a point x_0 of a convex set A in a locally convex vector space E an *exposed point* of A, if there exists $l \in E'$ such that $l(x_0) > l(x)$ for all $x \in A$, $x \neq x_0$ (see [2]). For later use we note a strengthening of the Krein–Milman-Theorem [2]:

LEMMA 2.3. *If A is a convex weakly compact subset of a Banach space E and E is either separable or uniformly convex, then A is the closure of the convex hull of the exposed points of A*

$$A = \operatorname{cl}\operatorname{conv}\exp A\,.$$

In the following we denote by $\exp\partial(y'\circ f)(x_0)$, $y' \in C'$, the set of exposed points of $\partial(y'\circ f)(x_0)$ where X' is endowed with any topology $\mathscr{T}$ consistent with $\langle X',X\rangle$ (that is, $\sigma(X',X) \leq \mathscr{T} \leq \tau(X',X)$). Since $(X'_\sigma)' = (X'_\tau)'$, this set is well-defined.

If $x_0' \in \exp\partial(y'\circ f)(x_0)$ then for some $h_0 \in X$

$$\langle h_0,x_0'\rangle > \langle h_0,x'\rangle \quad \text{for all } x' \in \partial(y'\circ f)(x_0),\ x' \neq x_0'\,,$$

and from proposition 2.1 and 2.2 we obtain $\langle h_0,x_0'\rangle = (y'\circ f)'(x_0;\, h_0)$, that is,

PROPOSITION 2.4. *Let $y' \in C'$ and $x_0' \in \exp\partial(y'\circ f)(x_0)$. Then there exists $h_0 \in X$ such that*

$$\langle h_0,x'\rangle < \langle h_0,x_0'\rangle = (y'\circ f)'(x_0;\, h_0) \quad \textit{for all } x' \in \partial(y'\circ f)(x_0),\ x' \neq x_0'\,.$$

We are now prepared to prove our main auxiliary result:

PROPOSITION 2.5. *Suppose $y_0' \in (C')^\circ_\tau$ and $x_0' \in \operatorname{conv}\exp\partial(y_0'\circ f)(x_0)$. Then there exists a $S \in \mathscr{L}(Y'_\tau,X'_\sigma)$ such that $Sy' \in \partial(y'\circ f)(x_0)$ for all $y' \in C'$. Moreover, $Sy_0' = x_0'$.*

PROOF. It is easily seen that it is sufficient to prove the assertion for $x_0' \in \exp \partial(y_0' \circ f)(x_0)$. Let x_0' be such an element. To y_0' and x_0' fix h_0 as in proposition 2.4 and then choose for every $y' \in C'$ an element x' in X', say Sy', such that

$$(2.2) \qquad Sy' \in \partial(y' \circ f)(x_0), \quad \langle h_0, Sy' \rangle = (y' \circ f)'(x_0; h_0) .$$

That this can be done is the contents of proposition 2.2. Now let

$$y' = \textstyle\sum_{i=1}^k \lambda_i y_i', \quad y_i' \in C', \ \lambda_i \geq 0 \text{ and } k \geq 1 .$$

Going back to the definition of $\partial(y' \circ f)(x_0)$ and $(y' \circ f)'(x_0; h_0)$, it is easily verified that (2.2) holds as well if we replace Sy' by $\sum_{i=1}^k \lambda_i Sy_i'$. Consequently

$$(2.3) \qquad S(\textstyle\sum_{i=1}^k \lambda_i y_i') = \sum_{i=1}^k \lambda_i Sy_i' \quad \text{for all } y_i' \in C', \ \lambda_i \geq 0, \ k \geq 1 ,$$

if we assume that Sy', for $y' \in C'$, is uniquely determined by (2.2). Because of proposition 2.4 this is true for y_0', so that $Sy_0' = x_0'$. Now suppose that for some $y_1' \in C'$, (2.2) is satisfied by x_1', x_1'', $x_1' \neq x_1''$. Since $y_0' \in (C')^\circ_\tau$ we can choose $0 < \lambda < 1$ small enough such that

$$y_2' := \frac{1}{1-\lambda} y_0' - \frac{\lambda}{1-\lambda} y_1' \in C' ;$$

thus Sy_2' is defined. But then (2.2) holds for $y_0' = \lambda y_1' + (1-\lambda) y_2'$, if we replace Sy_0' by

$$x' := \lambda x_1' + (1-\lambda) Sy_2' \quad \text{or} \quad x'' := \lambda x_1'' + (1-\lambda) Sy_2' ,$$

and thus $Sy_0' = x' = x''$ in contradiction to our assumption $x_1' \neq x_1''$. So we see that S maps C' "linearly" in the sense of (2.3) into X'. Since $(C')^\circ_\tau \neq \emptyset$, every $y' \in Y'$ is representable in the form $y' = y_1' - y_2'$ where $y_1', y_2' \in C'$. If $y' = y_3' - y_4'$ is another representation of y' with $y_3', y_4' \in C'$, then $y_1' + y_4' = y_3' + y_2' \in C'$ and from (2.3) we obtain

$$Sy_1' - Sy_2' = Sy_3' - Sy_4' .$$

Thus by

$$Sy' := Sy_1' - Sy_2' \quad \text{where } y' = y'_1 - y_2', \ y_1', y_2' \in C' ,$$

S is uniquely extended to all of Y. Of course, S is linear, $Sy' \in \partial(y' \circ f)(x_0)$ for $y' \in C'$ and $Sy_0' = x_0'$.

It remains to prove the continuity of S. To this end, let

$$U := \{x' \in X' : |\langle h_1, x' \rangle| \leq 1\}, \quad h_1 \in X ,$$

be given. We choose $\lambda > 0$ small enough such that $x_0 \pm \lambda h_1 \in K$ and define $V \in N\tau(Y', Y)$ by

$$V := \{y' \in Y' : |\langle y_i, y' \rangle| \leq \lambda, i = 1,2\}$$

where

$$y_1 := f(x_0 + \lambda h_1) - f(x_0), \quad y_2 := f(x_0) - f(x_0 - \lambda h_1).$$

For $y' \in V \cap C'$ we obtain from proposition 2.1, (2.1) and the definition of V

$$\langle \lambda h_1, Sy' \rangle \leq (y' \circ f)'(x_0; \lambda h_1) \leq \langle y_1, y' \rangle \leq \lambda$$

and similarly $\langle -\lambda h_1, Sy' \rangle \leq \langle -y_2, y' \rangle \leq \lambda$, that is,

$$S(V \cap C') \subset \{x' \in X' : |\langle \lambda h_1, x' \rangle| \leq \lambda\} = U.$$

Since $(C')^\circ_\tau \neq \emptyset$ there exists a convex symmetric $W \in N\tau(Y', Y)$ and a $y' \in C'$ such that $y' + W \subset C'$, $y' + W \subset V$ and thus

$$2W = W - W = (y' + W) - (y' + W) \subset V \cap C' - V \cap C'.$$

We get

$$S(2W) \subset S(V \cap C' - V \cap C') \subset U - U = 2U.$$

This completes the proof.

Taking the adjoint of the above S we obtain

PROPOSITION 2.6. *Suppose X is a Mackey space (that is, X has the topology $\tau(X, X')$), $y_0' \in (C')^\circ_\tau$ and $x_0' \in \operatorname{conv}\exp \partial(y_0' \circ f)(x_0)$. Then there exists a $T \in \partial f(x_0)$ such that $y_0' \circ T = x_0'$.*

PROOF. Let S be the above constructed mapping and define for every $x \in X$ a linear form Tx on Y' by

$$\langle Tx, y' \rangle := \langle x, Sy' \rangle \quad \text{for all } y' \in Y'.$$

Since $S \in \mathscr{L}(Y'_\tau, X'_\sigma)$ we have $Tx \in (Y'_\tau)' = Y$ for $x \in X$; but then $T \in \mathscr{L}(X_\sigma, Y_\sigma)$ and furthermore $T \in \mathscr{L}(X, Y)$, since X is a Mackey space (see [6, chapter IV, 7.4]). By construction

$$(y_0' \circ T)(x) = \langle Tx, y_0' \rangle = \langle x, Sy_0' \rangle = \langle x, x_0' \rangle$$

for all $x \in X$, that is, $y_0' \circ T = x_0'$.

Now assume that $T \notin \partial f(x_0)$, hence $T(\bar{x} - x_0) \not\leq f(\bar{x}) - f(x_0)$ for some $\bar{x} \in K$. Then the compact convex set $\{z\}$,

$$z := f(\bar{x}) - f(x_0) - T(\bar{x} - x_0),$$

and the closed convex set C can be strictly separated by a closed hyperplane, that is, for some $y' \in Y'$ and $\lambda \in \mathsf{R}$

$$\langle y, y' \rangle > \lambda > \langle z, y' \rangle \quad \text{for all } y \in C.$$

Since C is a cone, we see that $y' \in C'$, $\lambda < 0$ and thus $\langle z, y' \rangle < 0$, that is

$$\langle \bar{x} - x_0, Sy' \rangle = \langle T(\bar{x} - x_0), y' \rangle > y' \circ f(\bar{x}) - y' \circ f(x_0)$$

in contradiction to $Sy' \in \partial(y' \circ f)(x_0)$.

3. Main theorems.

In order to be able to apply proposition 2.6 we have to make two assumptions: $(C')^\circ_\tau \neq \emptyset$ and $\exp \partial(y' \circ f)(x_0) \neq \emptyset$ for some $y' \in (C')^\circ_\tau$. Before we give a condition guaranteeing the existence of exposed points, let us note a simple consequence of $(C')^\circ_\tau \neq \emptyset$. To this end, remember that the order cone C is called *normal* with respect to a topology $\mathscr{T}$ on Y, if there exists a base of neighbourhoods V of the origin in $\mathscr{T}$ such that

$$[u, z] := \{y \in Y : u \leqq y \leqq z\} \subset V \quad \text{if } u, z \in V .$$

We have

PROPOSITION 3.1. *If $(C')^\circ_\tau \neq \emptyset$ then C is normal in Y_σ.*

PROOF. Since $(C')^\circ_\tau \neq \emptyset$ each $y_0' \in Y'$ is representable in the form $y_0' = y_1' - y_2'$, $y_1', y_2' \in C'$. Now

$$\{y \in Y : |\langle y, y_i' \rangle| \leq 1, i = 1, 2\} \subset \{y \in Y : |\langle y, y_0' \rangle| \leq 2\};$$

consequently the sets

$$\{y \in Y : |\langle y, y_i' \rangle| \leq 1, i = 1, 2, \ldots, m\}, \quad y_i' \in C' ,$$

form a base for $N\sigma(Y, Y')$. But then the assertion follows easily from the fact that for $y_i' \in C'$ and $u \leqq y \leqq z$

$$\langle u, y_i' \rangle \leq \langle y, y_i' \rangle \leq \langle z, y_i' \rangle .$$

Now let X be semireflexive and normable (hence a reflexive Banach space) and either separable or smoothly convex. Then X'_τ is a Banach space as well and furthermore separable respectively uniformly convex (see [3, § 26,10, (12)]). From (1.4) it follows that $\partial(y' \circ f)(x_0)$, $y' \in C'$, is a nonvoid convex $\sigma(X', X'')$-compact subset of X'_τ; hence by lemma 2.3

LEMMA 3.2. *If X is a reflexive Banach space and either separable or smoothly convex, then*

$$\exp \partial(y' \circ f)(x_0) \neq \emptyset \quad \textit{for all } y' \in C' .$$

Moreover

$$\partial(y' \circ f)(x_0) = \operatorname{cl} \operatorname{conv} \exp \partial(y' \circ f)(x_0) .$$

(Here the closure can be taken in any topology consistent with $\langle X', X \rangle$.)

REMARK. Lemma 3.2 applies for instance to $X = l^p$ and $X = L^p$, $1 < p < \infty$.

Now we can state our first theorem. The central result will be that $\partial f(x_0)$ is nonvoid; the other points are proved similarly as in [8].

THEOREM 3.3. *If*

(a) *X is a reflexive Banach space and is either separable or smoothly convex,*
(b) $(C')^{\circ}_{\tau} \neq \emptyset$,

then $\partial f(x_0)$ is a nonvoid convex equicontinuous subset of $\mathscr{L}(X, Y_\sigma)$.
If, furthermore,

(c) *all order intervals $[u,z]$ are relatively compact in Y_σ,*

then $\partial f(x_0)$ is compact in $\mathscr{L}_s(X, Y_\sigma)$.

(Here $\mathscr{L}_s(X, Y_\sigma)$ is the space $\mathscr{L}(X, Y_\sigma)$ endowed with the topology of simple convergence.)

PROOF. $\partial f(x_0) \neq \emptyset$ is an immediate consequence of assumptions (a), (b), lemma 3.2 and proposition 2.6. The convexity of $\partial f(x_0)$ is obvious.

Now let $V \in N\sigma(Y, Y')$ be given; as shown in the proof of proposition 3.1, we may assume that V is symmetric and that $[u,z] \subset V$ if $u,z \in V$. Because of the continuity of f in x_0 there exists a symmetric neighbourhood U of 0 in X such that $f(x_0 + U) - f(x_0) \subset V$, that is,

$$f(x_0 + h) - f(x_0) \in V, \quad f(x_0) - f(x_0 - h) \in -V = V$$

for all $h \in U$. From $Th \leqq f(x_0 + h) - f(x_0)$ and $T(-h) \leqq f(x_0 - h) - f(x_0)$ for all $T \in \partial f(x_0)$ and $h \in U$ we get

$$Th \in [f(x_0) - f(x_0 - h), \, f(x_0 + h) - f(x_0)] \subset V,$$

showing that $\partial f(x_0)$ is an equicontinuous subset of $\mathscr{L}(X, Y_\sigma)$.

As we have seen, for each $h \in U$ the set $\{Th : T \in \partial f(x_0)\}$ is contained in some order interval and by (c) in a relatively compact subset of Y_σ. Since U is absorbing this holds for all $h \in X$; hence $\partial f(x_0)$ is relatively compact in $\mathscr{L}_s(X, Y_\sigma)$ (by [1, chapitre 3, § 3, n° 5]). The proof will be finished if we can show that $\partial f(x_0)$ is closed in $\mathscr{L}_s(X, Y_\sigma)$. To see this, note that $\mathscr{L}(X, Y) = \mathscr{L}(X, Y_\sigma)$. In fact $\mathscr{L}(X, Y) \subset \mathscr{L}(X, Y_\sigma)$ is trivial. Now for $T \in \mathscr{L}(X, Y_\sigma)$ and $y' \in Y'$, the mapping $x \to \langle Tx, y' \rangle$ is a continuous hence weakly continuous linear form on X, which is equivalent to $T \in \mathscr{L}(X_\sigma, Y_\sigma)$, and, since X is a Mackey space, $T \in \mathscr{L}(X, Y)$ (by [6, chapter IV, 7.4]). Thus

$$\begin{aligned}\partial f(x_0) &= \bigcap_{x \in K} \{T \in \mathscr{L}(X,Y) : T(x-x_0) \leqq f(x) - f(x_0)\} \\ &= \bigcap_{x \in K} \{T \in \mathscr{L}(X,Y_\sigma) : T(x-x_0) \in f(x) - f(x_0) - C\}\end{aligned}$$

and the theorem follows from the fact that for each $x \in X$, the mapping $T \to T(x-x_0)$ from $\mathscr{L}_s(X, Y_\sigma)$ into Y_σ is continuous, and that C is closed in Y_σ.

REMARK. Assumption (a) was only needed to guarantee the existence of at least one exposed point in $\partial(y' \circ f)(x_0)$ for some $y' \in (C')^\circ_\tau$. Our theorem holds, of course, with any hypothesis yielding the existence of such a point.

In section 4 we will give a condition for Y and C under which the assumptions (b) and (c) are satisfied.

The function f is even *regular subdifferentiable* at x_0 in the following sense

THEOREM 3.4. *Under the assumption* (a), (b), (c) *of Theorem* 3.3:

$$y' \circ \partial f(x_0) = \partial(y' \circ f)(x_0) \quad \textit{for all } y' \in C' .$$

PROOF. Note first that $y' \circ \partial f(x_0) \subset \partial(y' \circ f)(x_0)$. Thus lemma 3.2 implies

$$y' \circ \partial f(x_0) \subset \partial(y' \circ f)(x_0) = \operatorname{cl} \operatorname{conv} \exp \partial(y' \circ f)(x_0) \quad \text{for } y' \in C' .$$

From proposition 2.6 we get

$$\operatorname{conv} \exp \partial(y' \circ f)(x_0) \subset y' \circ \partial f(x_0) \quad \text{for } y' \in (C')^\circ_\tau .$$

Now for each $y' \in Y'$, the mapping $T \to y' \circ T$ maps $\mathscr{L}_s(X, Y_\sigma)$ continuously into X'_σ and consequently $y' \circ \partial f(x_0)$, for $y' \in C'$, is compact in X'_σ. The assertion follows for $y' \in (C')^\circ_\tau$ from the above inclusions. Now, suppose $x' \in \partial(y' \circ f)(x_0)$ but $x' \notin y' \circ \partial f(x_0)$ for some $y' \in C'$, $y' \notin (C')^\circ_\tau$. Then by a separation argument

$$(x' + U) \cap (y' \circ \partial f(x_0) + U) = \emptyset \tag{3.1}$$

for some $U \in N\sigma(X', X)$, say $U = \{u \in X' : |\langle \bar{x}, u \rangle| \leq 1\}$. For a fixed $y_1' \in (C')^\circ_\tau$ and $x_1' \in \partial(y_1' \circ f)(x_0)$ we consider the sequences

$$y_j' := j^{-1} y_1' + (1 - j^{-1}) y', \quad x_j' := j^{-1} x_1' + (1 - j^{-1}) x', \quad j = 1, 2, \ldots$$

Then $x_j' \in x' + U$ for j larger than some j_0. Moreover, it is easily verified that $x_j' \in \partial(y_j' \circ f)(x_0)$ for all j, and, since $y_j' \in (C')^\circ_\tau$, that $x_j' \in y_j' \circ \partial f(x_0)$. Hence

$$x_j' \in (x' + U) \cap y_j \circ \partial f(x_0) \quad \text{for } j \geq j_0 . \tag{3.2}$$

Since $\partial f(x_0)$ is compact in $\mathscr{L}_s(X, Y_\sigma)$, there exists $\lambda > 0$ such that $|\langle T\bar{x}, y_1' - y' \rangle| \leq \lambda$ for all $T \in \partial f(x_0)$; hence

$$|\langle \bar{x}, (y_j' - y') \circ T \rangle| = j^{-1} |\langle T\bar{x}, y_1' - y' \rangle| \leq 1$$

for j larger than some j_1, that is,

$$y_j' \circ \partial f(x_0) \subset y' \circ \partial f(x_0) + U ,$$

and because of (3.1),

$$(x' + U) \cap y_j' \circ \partial f(x_0) = \emptyset$$

for $j \geq j_1$. This contradicts (3.2).

Remark. The assumptions of the above theorems are, for example, satisfied in the special case $X = \mathsf{R}^n$, $Y = \mathsf{R}^m$ and C any closed proper convex cone in R^m. Since R^m is order complete if and only if the closed cone C is generated by m linearly independent elements, our theorems are in the finite dimensional case a direct generalization of the results given by Valadier.

In order to give an example, where theorems 3.3 and 3.4 do not hold, let Y be an ordered vector space with topology $\mathscr{T} > \sigma(Y, Y')$. Define $X := Y_\sigma$ and consider any $f \in \mathscr{L}(X, Y_\sigma)$ but $f \notin \mathscr{L}(X, Y)$ (for example $f(x) := x$ for $x \in X$). f is a convex mapping satisfying (1.1) for $x_0 = 0$. It is easily verified that $\partial f(0) = \emptyset$ but $\partial(y' \circ f)(0) \neq \emptyset$.

4. Cones with a compact base.

We will give a condition for Y and C, under which $(C')^\circ_\tau \neq \emptyset$ and all order intervals are relatively compact in Y_σ. In order to do this remember that a nonempty convex subset B of C is called a *base* for C if each $y \in C$, $y \neq 0$, has a unique representation $y = \lambda b$, where $b \in B$, $\lambda \in \mathsf{R}$, $\lambda > 0$ (see [5, chapter I, § 3]). If $(C')^\circ_\tau \neq \emptyset$ and $y_0' \in (C')^\circ_\tau$ then the closed convex set

$$B := \{ y \in C : \langle y, y_0' \rangle = 1 \}$$

is a base for C. Now

$$|\langle y, y' \rangle| \leq 1 \quad \text{for } y \in B \text{ and } y' \in U := (-y_0' + C') \cap (y_0' - C') ,$$

and, since $U \in N\tau(Y', Y)$, the base B is an equicontinuous subset of $(Y'_\tau)' = Y$. By the theorem of Alaoglu–Bourbaki:

PROPOSITION 4.1. *Suppose* $(C')^\circ_\tau \neq \emptyset$ *and* $y_0' \in (C')^\circ_\tau$. *Then*

$$B = \{y \in C : \langle y, y_0' \rangle = 1\}$$

is a weakly compact base for C.

If Y is semireflexive then the converse of 4.1 holds:

PROPOSITION 4.2. *Suppose* Y *is semireflexive and* C *has a weakly compact base* B *lying in a closed hyperplane* H *not containing* 0. *Then* $(C')^\circ_\tau \neq \emptyset$. *Furthermore, all order intervals are relatively compact in* Y_σ.

PROOF. Let $B \subset H = \{y \in Y : \langle y, y_0' \rangle = 1\}$ for some $y_0' \in Y'$. Since B is a base for C we see that $y_0' \in C'$. We will show that

$$[-y_0', y_0'] = (-y_0' + C') \cap (y_0' - C')$$

is a barrel in Y'_τ (that is, a convex circled closed and absorbing set) and thus an element of $N\tau(Y', Y)$, since Y'_τ is barreled (by [6, chapter IV, 5.5]). But then $(C')^\circ_\tau \neq \emptyset$, since

$$y_0' + [-y_0', y_0'] \subset C' .$$

Moreover, we see that C is normal in Y_σ (by proposition 3.1) and thus all order intervals are bounded in Y_σ, hence relatively compact in Y_σ (by [6, chapter IV, 5.5]).

It is easily verified that $[-y_0', y_0']$ is convex, closed and circled since $y_0' \in C'$. In order to see that $[-y_0', y_0']$ is absorbing let $y_1' \in Y'$ be given. We will show that $y_1' \in \lambda[-y_0', y_0']$ where $\lambda > 0$ is such that $B \subset \lambda U$ for

$$U := \{y \in Y : |\langle y, y_1' \rangle| \leq 1\};$$

since B is compact in Y_σ such a λ exists. Now, assume $y_1' \notin \lambda[-y_0', y_0']$, that is,

$$\lambda^{-1} y_1' \notin y_0' - C' \quad \text{or} \quad \lambda^{-1} y_1' \notin -y_0' + C' .$$

Let us consider only the case $\lambda^{-1} y_1' \notin y_0' - C'$ (the other assumption can be dealt with similarly), that is,

$$z' := y_0' - \lambda^{-1} y_1' \notin C' .$$

By a separation argument there is a $y \in Y$, $y \neq 0$, and an $\alpha \in \mathbb{R}$ such that

$$\langle y, y' \rangle > \alpha > \langle y, z' \rangle \quad \text{for all } y' \in C' .$$

Since C is a cone, we get $\alpha < 0$, $y \in C'' := \{u \in Y : \langle u, C' \rangle \geq 0\}$ and, by the bipolar theorem, $y \in C$, that is, $y = \beta b$ where $b \in B$, $\beta > 0$. Thus

$$\beta^{-1}\langle y,z'\rangle = \langle b,z'\rangle = \langle b,y_0'\rangle - \lambda^{-1}\langle b,y_1'\rangle < \alpha\beta^{-1} < 0\,,$$

hence $\lambda^{-1}\langle b,y_1'\rangle > \langle b,y_0'\rangle = 1$, that is, $b \notin \lambda U$ in contradiction to the choice of λ.

From proposition 4.2. we get

THEOREM 4.3. *Let Y be semireflexive. Then theorem* 3.3 *and* 3.4 *hold if the assumptions* (b) *and* (c) *are replaced by*

(b′) *C has a weakly compact base lying in a closed hyperplane not running through* 0.

REMARK. It is easy to construct closed proper convex cones C satisfying (b′). For this purpose let H be a closed hyperplane in Y not containing 0, take a nonempty convex weakly compact subset B in H and define $C := \bigcup_{\lambda \geq 0} \lambda B$.

5. Subdifferentiability and Gateaux differentiability.

Let us note an interesting conclusion from proposition 4.1. Recall that the infimum (if it exists) of the set

$$\{\lambda^{-1}(f(x_0+\lambda h)-f(x_0)) :\ \lambda > 0,\, x_0+\lambda h \in K\}$$

is called the *directional derivative* $f'(x_0;h)$ of f at x_0 in the direction h (cf. [8]).

THEOREM 5.1. *If $(C')^\circ_\tau \neq \emptyset$ then $f'(x_0;h)$ is defined for all $h \in X$. If, in addition, C is normal, then*

$$f'(x_0;h) = \lim \lambda^{-1}(f(x_0+\lambda h)-f(x_0)), \quad \lambda \to 0,\, \lambda > 0\,.$$

PROOF. Let $h \in X$ and assume $x_0 \pm h \in K$ (otherwise replace h by λh, $\lambda > 0$ sufficiently small). For $0 < \mu \leq \nu \leq 1$ we have

$$f(x_0+\mu h) = f\left(\frac{\nu-\mu}{\nu}x_0 + \frac{\mu}{\nu}(x_0+\nu h)\right) \leqq \frac{\nu-\mu}{\nu}f(x_0) + \frac{\mu}{\nu}f(x_0+\nu h)$$

and hereby

$$(5.1) \qquad \mu^{-1}(f(x_0+\mu h)-f(x_0)) \leqq \nu^{-1}(f(x_0+\nu h)-f(x_0)), \quad 0 < \mu \leq \nu \leq 1\,.$$

Similarly one gets

$$f(x_0)-f(x_0-h) \leqq \mu^{-1}(f(x_0+\mu h)-f(x_0)) \quad \text{for } 0 < \mu \leq 1$$

and thus with

$$y_n := n(f(x_0+n^{-1}h)-f(x_0))-f(x_0)+f(x_0-h), \quad n=1,2,\ldots,$$

we have

$$0 \leqq y_n \leqq y_m \quad \text{for } n \geq m,\ n,m \in \mathsf{N}.$$

We will show that the sequence $\{y_n\}_{n\in N}$ converges in Y_σ to some y. This yields for each $n_0 \in \mathsf{N}$ that

$$y \in \mathrm{cl}\{y_n : n \geq n_0\} \subset \mathrm{cl}(y_{n_0}-C) = y_{n_0}-C,$$

that is, y is a lower bound for $\{y_n\}_{n\in\mathsf{N}}$. If z is any other lower bound, that is, $z \leqq y_n$ for all n, then $y_n - z \in C$ and thus $y-z \in \bar{C} = C$, hence $y = \inf\{y_n\}$. The first part of the assertion is an easy consequence from this.

As shown in section 4 each $y_0' \in (C')^\circ_\tau$ determines a representation $y_n = \lambda_n b_n$ where

$$b_n \in B := \{y \in C : \langle y, y_0'\rangle = 1\}$$

and $\lambda_n \geq 0$. From $0 \leqq y_n \leqq y_m$ for $n \geq m$ we obtain $0 \leq \lambda_n \leq \lambda_m$, so that $\{\lambda_n\}$ converges. Moreover, for some $\lambda > 0$ sufficiently large $\{y_n\} \subset \mathrm{conv}(0 \cup \lambda B)$. As $\mathrm{conv}(0 \cup \lambda B)$ is a compact subset of Y_σ, the convergence of $\{y_n\}$ will follow, if we can show that $\{y_n\}$ is a Cauchy sequence in Y_σ. To see this, let $y' \in Y'$ be given. Then $y' = y_1' - y_2'$ where $y_1', y_2' \in (C')^\circ_\tau$. If $y_n = \alpha_n u_n = \beta_n v_n$ are the representations of $y_n \in \{y_n\}_{n\in\mathsf{N}}$ with respect to the bases given by y_1', y_2', then

$$|\langle y_i - y_j, y'\rangle| \leq |\langle y_i - y_j, y_1'\rangle| + |\langle y_i - y_j, y_2'\rangle| = |\alpha_i - \alpha_j| + |\beta_i - \beta_j|$$

and this converges to 0, if $i,j \to \infty$. Thus $\{y_n\}$ is a Cauchy sequence in Y_σ.

Since C is normal in Y_σ (by proposition 3.1), it is an easy consequence of (5.1) that

$$f'(x_0; h) = \lim \lambda^{-1}(f(x_0+\lambda h)-f(x_0)), \quad \lambda \to 0, \lambda > 0,$$

in Y_σ. If in addition C is normal in Y, this even holds in Y (by [5, chapter 2, 3.4]).

Theorem 5.1 shows that the above definition of the directional derivative is in accordance with the definition used in 2.

Recall that f is called Gateaux differentiable at x_0, if

$$\lim \lambda^{-1}(f(x_0+\lambda h)-f(x_0)), \quad \lambda \to 0,$$

exists for all $h \in X$ ([3, § 26,4]). Since for a convex function $f: \mathsf{R}^n \to \mathsf{R}$ the partial derivatives (if they exist) are continuous (see [7, Theorem 4.4.7]), the Gateaux- and Fréchet-differentiability coincide for such an f. Thus 5.2 is a generalization of the well-known theorem that a convex function f: $\mathrm{R}^n \to \mathrm{R}$ is differentiable at x_0 if and only if f has a unique subgradient at x_0 (that is, there exists only one "nonvertical" supporting hyperplane at $(x_0, f(x_0))$ to

$$\{(x,z) \in \mathrm{R}^{n+1} : f(x) \leq z\})$$

(cf. [7, Theorem 4.4.6]).

THEOREM 5.2. *If assumptions* (a), (b), (c) *of theorem* 3.3 *hold and if* C *is normal in* Y, *then* f *is Gateaux differentiable at* x_0 *if and only if* f *has a unique subgradient at* x_0.

PROOF. Suppose

$$d(x_0; h) := \lim \lambda^{-1}(f(x_0+\lambda h) - f(x_0)), \quad \lambda \to 0,$$

exists for all $h \in X$. Let $T \in \partial f(x_0)$. Then for $h \in X$ and all $\lambda > 0$ sufficiently small

$$\lambda T h = T(x_0 + \lambda h - x_0) \leqq f(x_0 + \lambda h) - f(x_0),$$

hence $Th \leqq d(x_0; h)$. For $-h$ we get

$$-Th = T(-h) \leqq d(x_0; -h) = -d(x_0; h)$$

and thus $Th = d(x_0; h)$ for all h, which shows that T is uniquely determined.

Now, suppose that f is not Gateaux differentiable at x_0, and let us show that $\partial f(x_0)$ contains at least two elements. From theorem 5.1 we get $f'(x_0; h) \neq -f'(x_0; -h)$ for some h, hence

$$\langle f'(x_0; h), y' \rangle \neq \langle -f'(x_0; -h), y' \rangle \quad \text{for some } y' \in C',$$

that is,

$$(y' \circ f)'(x_0; h) \neq -(y' \circ f)'(x_0; -h).$$

We choose $x_1', x_2' \in \partial(y' \circ f)(x_0)$ as in proposition 2.2,

$$\langle h, x_1' \rangle = (y' \circ f)'(x_0; h) \quad \text{and} \quad \langle -h, x_2' \rangle = (y' \circ f)'(x_0; -h),$$

and thus $\langle h, x_1' \rangle \neq \langle h, x_2' \rangle$, that is $x_1' \neq x_2'$. The remaining part of the proof follows from theorem 3.4.

NOTE ADDED IN PROOF. Recently M. M. Day pointed out that Lemma 2.3 holds without the assumption that E is either separable or uniformly convex (M. M. Day, *Normed linear spaces*, 3. edition, Springer-Verlag, Heidelberg · New York, 1973, ch. III, 5, 5a). Consequently in Lemma 3.2 and Theorem 3.3 the hypothesis that X is either separable or smoothly convex can be omitted.

REFERENCES

1. N. Bourbaki, *Espaces vectoriels topologiques*, chapitres 3, 4, 5 (Act. Sci. Ind. 1229), 1re éd., Hermann, Paris, 1966.
2. V. L. Klee, *Extremal structure of convex sets* II, Math. Z., 69 (1958), 90–104.
3. G. Köthe, *Topologische lineare Räume* I (Grundlehren Math. Wissensch. 107), Springer-Verlag, Berlin · Göttingen · Heidelberg, 1960.
4. J. J. Moreau, *Fonctionelles convexes* (Séminaire sur les équations aux dérivées partielles II), Collège de France, Paris, 1966–67.
5. A. L. Peressini, *Ordered topological vector spaces*, Harper and Row, New York, 1967.
6. H. H. Schaefer, *Topological vector spaces*, Macmillan, New York, 1966.
7. J. Stoer, C. W. Witzgall, *Convexity and optimization in finite dimensions* I (Grundlehren Math. Wissensch. 163), Springer-Verlag, Berlin · Heidelberg · New York, 1970.
8. M. Valadier, *Sous-différentiabilité de fonctions convexes à valeurs dans un espace vectoriel ordonné*, Math. Scand. 30 (1972), 65–74.

INSTITUT FÜR ANGEWANDTE MATHEMATIK
DER UNIVERSITÄT WÜRZBURG, WEST-GERMANY