Canad. Math. Bull. Vol. **23** (1), 1980

UPPER SEMI-CONTINUITY OF SUBDIFFERENTIAL MAPPINGS

BY
DAVID A. GREGORY

ABSTRACT. Characterizations of the upper semi-continuity of the subdifferential mapping of a continuous convex function are given.

1. **Notation.** The following notation will be used throughout the paper. Let (E, τ) denote a vector space E (over the real numbers R) with a locally convex Hausdorff topology τ. We shall also use τ to denote the product topology on $E \times R$. Let E^* be the space of all continuous linear functionals x^* on E. For x in E and x^* in E^* let $\langle x, x^*\rangle \equiv x^*(x)$.

Let $\mathscr{A}$ be a class of weakly bounded absolutely convex subsets A of E such that $E = \bigcup_{A \in \mathscr{A}} A$ and $\lambda A \in \mathscr{A}$ whenever $A \in \mathscr{A}$ and $\lambda > 0$. Let $\tau_{\mathscr{A}}$ be the locally convex topology on E^* of uniform convergence on the members of $\mathscr{A}$; that is, the topology determined by the seminorms p_A, $A \in \mathscr{A}$ where $p_A(y^*) \equiv \sup\langle A, y^*\rangle$ for each y^* in E^*. Equivalently, τ_A is the vector space topology that has the sets $A^0 \equiv \{y^* \in E^* : \sup\langle A, y^*\rangle \leq 1\}$, $A \in \mathscr{A}$ as a neighborhood subbase of the origin in E^*. In particular, if $\mathscr{A}$ is the class of all balanced line segments in E, then $\tau_{\mathscr{A}}$ is the weak* topology. Let $\mathscr{I}$ be the class of all finite closed subintervals of R and let $\mathscr{A} \times \mathscr{I} \equiv \{A \times I : A \in \mathscr{A}, I \in \mathscr{I}\}$.

If J is an index set, we say that a net y_j, $j \in J$ $\tau_{\mathscr{A}}$*-converges* to y if $p_A(y_j - y)$, $j \in J$ converges to zero for each A in $\mathscr{A}$. We say that the net $\tau_{\mathscr{A}}$*-approaches* Y if $\inf p_A(y_j - Y)$, $j \in J$ (each infimum is taken over all y in Y with j, A fixed) converges to zero for each A in $\mathscr{A}$. The index set J will usually be dropped in statements about convergence.

Let f be a function on E with values in $R \cup \{\infty\}$ and suppose that f is convex; that is, $\operatorname{epi} f \equiv \{(y, r) \in E \times R : f(y) \leq r\}$ is a convex subset of $E \times R$. We shall also assume that the convex function f is (finite and) continuous at a point x in (E, τ) and so continuous in some τ-neighborhood of x.

2. **Definitions.** A *subgradient* of f at x is any x^* in E^* such that $x^*(y - x) \leq f(y) - f(x)$ for all y in E. The *subdifferential* of f at x is the set $\partial f(x)$ of all subgradients x^* of f at x. For $\varepsilon > 0$, the ε*-approximate subdifferential* of f at x is the set $\partial_\varepsilon f(x) \equiv \{z^* \in E^* : z^*(y - x) \leq f(y) - f(x) + \varepsilon$ for all y in $E\}$. The

Received by the editors September 12, 1978.

conjugate function f^* of f is defined by

$$f^*(z^*) \equiv \sup\{z^*(y) - f(y) : y \in E\} \quad \text{for} \quad z^* \quad \text{in} \quad E^*.$$

The *directional derivative* of f at x in the direction y is

$$f'(x; y) \equiv \lim_{\lambda \to 0^+} \frac{f(x+\lambda y) - f(x)}{\lambda}$$

for each y in E. If x^*, $(x+\lambda y)^*$ are subgradients of f at x and $x+\lambda y$ respectively, then by the convexity of f, the inequalities

$$x^*(y) \le f'(x; y) \le \lambda^{-1}[f(x+\lambda y) - f(x)] \le (x+\lambda y)^*(y)$$

hold for all y in E and all sufficiently small $\lambda > 0$. In particular, the subgradients of f at x are precisely those x^* in E^* which are dominated by the continuous sublinear functional $f'(x; \cdot)$. Thus, by the Hahn–Banach theorem, for each y in E there is an x^* in $\partial f(x)$ such that $x^*(y) = f'(x; y)$. If for each A in $\mathscr{A}$ the convergence in the limit is uniform for y in A as $\lambda \to 0^+$, we shall say that f is *$\tau_{\mathscr{A}}$-directionally differentiable* at x. The limits here are one-sided; although $-f'(x; y) \le f'(x; -y)$, equality need not hold.

We shall say that a set valued mapping T from E to the set 2^{E^*} of all subsets of E^* is $\tau - \overline{\tau_{\mathscr{A}}}$ *upper semi-continuous* (respectively, *lower semi-continuous*) at x if for each $\tau_{\mathscr{A}}$-neighborhood V of 0 in E^*, there is a τ-neighborhood U of x in E such that $T(y) \subset T(x) + V$ (respectively, $T(x) \subset T(y) + V$) whenever y is in U. The mapping T is $\tau - \overline{\tau_{\mathscr{A}}}$ *continuous* at x if it is both $\tau - \overline{\tau_{\mathscr{A}}}$ u.s.c. and l.s.c. there. (The notation $\overline{\tau_{\mathscr{A}}}$ may be considered to represent the power set uniformity on 2^{E^*} determined by $\tau_{\mathscr{A}}$). In the definitions of u.s.c. and l.s.c. in the literature, the uniform structure is usually ignored; in that case we shall say that T is $\tau - \tau_{\mathscr{A}}$ u.s.c. (respectively, l.s.c.) at x if for each $\tau_{\mathscr{A}}$-open set G such that $T(x) \subset G$ (respectively, $T(x) \cap G \neq \emptyset$) there is a τ-neighborhood U of x such that $T(y) \subset G$ (respectively, $T(y) \cap G \neq \emptyset$) whenever y is in U. If $T(x)$ is $\tau_{\mathscr{A}}$-compact (in particular, if it is a singleton), the definitions agree.

A subset C of $E \times R$ is *strictly above* the function $f'(x, \cdot)$, where x is fixed, if there is an $\varepsilon > 0$ such that $f'(x; y) \le r - \varepsilon$ for all (y, r) in C.

3. **Upper semi-continuity of the subdifferential mapping.** Many of the equivalences in the following theorem can be regarded as an extension (to the case where $\partial f(x)$ is not a singleton) of results of Asplund and Rockafellar [1] on the $\mathscr{A}$-differentiability of convex functions. The implications $4 \Leftrightarrow 3 \Rightarrow 5 \Rightarrow 6$ in the proof are simple extensions of arguments in [1].

3.1. THEOREM. *The following conditions are equivalent for a convex function f which is continuous in a neighborhood of x in* (E, τ).

1. *The function f is $\tau_{\mathscr{A}}$-directionally differentiable at x.*

https://doi.org/10.4153/CMB-1980-002-9 Published online by Cambridge University Press

2. *Whenever a subset of a member of* $\mathcal{A}\times\mathcal{I}$ *is strictly above* $f'(x;\cdot)$, *then it is contained in* $\lambda[\operatorname{epi} f-(x, f(x)]$ *for some* $\lambda>0$.

3. *The mapping* $\lambda\to\partial_\lambda f(x)$ *is* $\overline{\tau_{\mathcal{A}}}$ *u.s.c. at* 0 *in* $[0,\infty)$; *that is, for each* A *in* $\mathcal{A}$, *there is an* $\eta>0$ *such that* $\partial_\eta f(x)\subset\partial f(x)+A^0$.

4. *The function* $y^*\to\langle x, y^*\rangle-f^*(y^*)$ *attains its supremum strictly at* $\partial f(x)$ *with respect to* $\tau_{\mathcal{A}}$; *that is, whenever* y^* *in* E^* *are such that* $\langle x, y^*\rangle-f^*(y^*)$ *converges to* $f(x)$, *then* y^* $\tau_{\mathcal{A}}$*-approaches* $\partial f(x)$.

5. *The approximate subdifferential mapping* $(y,\lambda)\to\partial_\lambda f(y)$ *is* $\tau-\overline{\tau_{\mathcal{A}}}$ *u.s.c. at* $(x, 0)$; *that is for each* A *in* $\mathcal{A}$, *there are an* $\eta>0$ *and a* τ*-neighborhood* U *of* x *such that* $\partial_\eta f(y)\subset\partial f(x)+A^0$ *whenever* y *is in* U.

6. *The subdifferential mapping* ∂f *is* $\tau-\overline{\tau_{\mathcal{A}}}$ *u.s.c. at* x; *that is, for each* A *in* $\mathcal{A}$, *there is a* τ*-neighborhood* U *of* x *such that* $\partial f(y)\subset\partial f(x)+A^0$ *whenever* y *is in* U.

7. *Whenever* y *converges to* x *in* (E,τ), *then for each* A *in* $\mathcal{A}$, $\inf p_A(\partial f(y)-\partial f(x))$ *converges to zero.*

Proof. $(1\Rightarrow 2)$ Suppose that $C\subset A\times[a, b]$ and, for some $\varepsilon>0$, $f'(x; y)\le r-\varepsilon$ for all (y, r) in C. Since $y\in A$ when $(y, r)\in C$, it follows from 1 that there is a $\lambda>0$ such that for all (y, r) in C, $\lambda^{-1}[f(x+\lambda y)-f(x)]\le r$; equivalently, $f(x+\lambda y)\le f(x)+\lambda r$. Thus, $\lambda C+(x, f(x))\subset\operatorname{epi} f$ and condition 2 follows.

$(2\Rightarrow 1)$ Let $A\in\mathcal{A}$ and $\varepsilon>0$ be given. Since f is continuous at x, $f'(x,\cdot)$ is continuous on E and is therefore bounded on A. Thus, $C\equiv\{(y, r_y): y\in A, r_y=\varepsilon+f'(x; y)\}$ is a subset of a member of $\mathcal{A}\times\mathcal{I}$. Also, C is strictly above $f'(x;\cdot)$ and so, by assumption, contained in $\lambda^{-1}[\operatorname{epi} f-(x, f(x))]$ for some $\lambda>0$. Therefore, for y in A,

$$0\le\lambda^{-1}[f(x+\lambda y)-f(x)]-f'(x; y)\le r_y-f'(x; y)=\varepsilon.$$

Since the difference quotient is monotone decreasing with λ, condition 1 follows.

$(1\Rightarrow 3)$ Given A in $\mathcal{A}$, choose $\lambda>0$ so that $|\lambda^{-1}[f(x+\lambda y)-f(x)]-f'(x; y)|<\frac{1}{2}$ for all y in A. Let $0<\eta\le\lambda/2$. Then for y^* in $\partial_\eta f(x)$, we have $f(x+\lambda y)\ge f(x)+\lambda y^*(y)-\eta$ for all y in E. Thus, for any x^* in $\partial f(x)$ we have

$$(y^*-x^*)(y)\le\frac{f(x+\lambda y)-f(x)}{\lambda}-x^*(y)+\frac{\eta}{\lambda}.$$

Now, for each y there is an x^* in $\partial f(x)$ such that $x^*(y)=f'(x; y)$. Thus $\inf\{(y^*-x^*)(y): x^*\in\partial f(x)\}<1$ for each y in A. From the definition, it is clear that $\partial f(x)$ is weak* closed and convex and contained in the polar of the convex body $\{y-x: f(y)-f(x)\le 1\}$. Thus $\partial f(x)$ is weak* compact and convex. It now follows by a simple contradiction argument using the separation theorem that $(y^*-\partial f(x))\cap A^0\ne\emptyset$; that is, $y^*\in\partial f(x)+A^0$.

$(3\Leftrightarrow 4)$ This follows directly from the fact that $\partial_\eta f(x)=\{y^*\in E^*:\langle x, y^*\rangle-f^*(y^*)\ge f(x)-\eta\}$ and $f(x)\ge\langle x, y^*\rangle-f^*(y^*)$ for all y^* in E^*.

(3 ⇒ 5) It is sufficient to show that if $0<\eta<\gamma/2$ then there is a τ-neighborhood U of x such that $\partial_\eta f(y)\subset\partial_\gamma f(x)$ whenever y is in U. Let $y^*\in\partial_\eta f(y)$. Then $f(z)-f(y)\geq y^*(z-y)-\eta$ for all z in E. Thus, for all z in E,

$$\begin{aligned} f(z)-f(x) &= f(z)-f(y)+f(y)-f(x) \\ &\geq y^*(z-y)-\eta+f(y)-f(x) \\ &\geq y^*(z-x)-\eta+y^*(x-y)+f(y)-f(x) \\ &\geq y^*(z-x)-2\eta+[f(y)-f(2y-x)]+[f(y)-f(x)]. \end{aligned}$$

Since f is continuous at x, there is an absolutely convex τ-neighborhood V of 0 such that $|f(y)-f(z)|\leq\gamma/2-\eta$ whenever $y, z\in x+2V$. Then, for $y\in U=x+V$ we have $f(z)-f(x)\geq y^*(z-x)-\gamma$ for all z; that is, $y^*\in\partial_\gamma f(x)$.

(5 ⇒ 6 ⇒ 7) These implications are immediate.

(7 ⇒ 1) For all x^* in $\partial f(x)$ and $(x+\lambda y)^*$ in $\partial f(x+\lambda y)$ we have for y in A and $\lambda>0$ that

$$\begin{aligned} 0 &\leq \lambda^{-1}[f(x+\lambda y)-f(x)]-f'(x;y) \\ &\leq [(x+\lambda y)^*-x^*](y) \\ &\leq p_A[(x+\lambda y)^*-x^*] \end{aligned}$$

Thus, the first difference is non-negative and less than or equal to inf $p_A(\partial f(x+\lambda y)-\partial f(x))$. Since A is bounded, $x+\lambda y$ converges to x uniformly for y in A as $\lambda\to 0^+$. Thus, 1 follows from 7.

3.2. Remarks. 1. By the tangent cone to epi f at $(x, f(x))$ we mean the smallest closed convex cone $K_{f,x}$ in $(E\times R, \tau)$ which contains epi $f-(x, f(x))$; that is, $K_{f,x}\equiv\overline{\bigcup_{\lambda>0}\lambda[\text{epi}\, f-(x, f(x))]}$. It follows from the separation theorem that

$$K_{f,x}=\{(y, r)\in E\times R: x^*(y)\leq r \quad \text{for all} \quad x^*\in\partial f(x)\}=\text{epi}\, f'(x;\cdot).$$

Also, C is strictly above $f'(x;\cdot)$ if and only if $C-(0,\varepsilon)\subset K_{f,x}$ for some $\varepsilon>0$; equivalently, if and only if $C+U\times[-\varepsilon,\varepsilon]\subset K_{f,x}$ for some $\varepsilon>0$ and τ-neighborhood U of the origin. If the later hold, we say that C is strictly inside the tangent cone. Condition 2 can then be replaced by the following condition on the 'rate of formation' of the tangent cone $K_{f,x}$:

> "2′. Whenever a subset of a member of $\mathscr{A}\times\mathscr{I}$ is strictly inside in the tangent cone $K_{f,x}$, then it is eventually engulfed by the increasing family $\lambda[\text{epi}\, f-(x, f(x))]$, $\lambda>0$ which yields $K_{f,x}$."

Each member of the family $\lambda[\text{epi}\, f-(x, f(x))]+(0,\varepsilon)$, $\varepsilon>0$, $\lambda>0$ is engulfed (same λ), and each point strictly inside $K_{f,x}$ is in the interior of such a set. Therefore, the τ-compact sets strictly inside $K_{f,x}$ are always engulfed. Thus, the conditions of Theorem 3.1 all hold. When the members of $\mathscr{A}$ are all τ-compact (in particular, if $\tau_{\mathscr{A}}$ is the weak* topology, or the topology $\tau_{\mathscr{C}}$ of uniform convergence on the class $\mathscr{C}$ of all τ-compact convex subsets of E). Condition 5 then gives the result of Moreau [1; p. 458] that the approximate subdifferential mapping $\partial_\lambda f$ is $\tau-\overline{\tau_{\mathscr{C}}}$ u.s.c. at (x, λ) for all $\lambda \geq 0$ (it is always $\tau-\overline{\tau_{\mathscr{A}}}$ continuous at (x, λ) for $\lambda > 0$ [1; p. 456]).

2. Let ϕ be a *selection* of ∂f near x; that is, $\phi(y) \in \partial f(y)$ for each y in a τ-neighborhood of x. Condition 7 implies that in order to have $\tau-\overline{\tau_{\mathscr{A}}}$ u.s.c. of ∂f at x, it is sufficient to have that for each A in $\mathscr{A}$ there is an $\eta > 0$ and a τ-neighborhood U of 0 such that $\phi(y) \in \partial f(x) + A^0$ for y in U. In fact, a different selection ϕ may be used for each A. It follows from this that if ∂f is $\tau-\tau_{\mathscr{A}}$ or $\tau-\overline{\tau_{\mathscr{A}}}$ l.s.c. at x, then it is $\tau-\overline{\tau_{\mathscr{A}}}$ u.s.c. at x, and, consequently, $\tau-\overline{\tau_{\mathscr{A}}}$ continuous at x. Moreover, $\partial f(x)$ must then be a singleton [4; p. 67]. Consequently, ∂f is $\tau-\tau_{\mathscr{A}}$ or $\tau-\overline{\tau_{\mathscr{A}}}$ l.s.c. at x if and only if it is $\tau-\tau_{\mathscr{A}}$ or $\tau-\overline{\tau_{\mathscr{A}}}$ u.s.c. at x and $\partial f(x)$ is a singleton. It is for this reason that upper semi-continuity is examined in this paper. (For completeness, we comment without proof that this remark holds for any maximal monotone mapping; that is for any mapping $T: E \to 2^{E^*}$ which is maximal with respect to the monotone property $\langle y - x, y^* - x^* \rangle \geq 0$ for all x, y in E and $x^* \in T(x)$, $y^* \in T(y)$.)

Most of the equivalent statements in the corollary below can be found in [1].

3.3. Corollary. *Let f be a convex function which is continuous in a neighborhood of x in (E, τ), and let $x^* \in E^*$. Then the following conditions are all equivalent and imply that $\partial f(x)$ is the singleton $\{x^*\}$.*

1. *The convex function f is $\tau_{\mathscr{A}}$-differentiable at x with $\tau_{\mathscr{A}}$-differential x^*; that is, for each A in $\mathscr{A}$, $\lambda^{-1}[f(x+\lambda y) - f(x)]$ converges to $x^*(y)$ uniformly for y in A as $\lambda \to 0$.*

2. *Whenever a subset of a member of $\mathscr{A} \times \mathscr{I}$ is strictly above the functional x^*, then it is contained in $\lambda[\operatorname{epi} f - (x, f(x))]$ for some $\lambda > 0$.*

3. *The conjugate function f^* is $\tau_{\mathscr{A}}$-rotund at x^* relative to x; that is* [1; p. 445] *for each A in $\mathscr{A}$ there is an $\eta > 0$ such that*

$$\{y^* : f^*(x^* + y^*) - f^*(x^*) - \langle x, y^* \rangle \leq \eta\} \subset A^0.$$

4. *The function $y^* \to \langle x, y^* \rangle - f^*(y^*)$ attains its supremum strictly at x^* with respect to $\tau_{\mathscr{A}}$; that is, whenever y^* in E^* are such that $\langle x, y^* \rangle - f^*(y^*)$ converges to $f(x)$, then y^* $\tau_{\mathscr{A}}$-converges to x^*.*

5. *The approximate subdifferential mapping $\partial_\lambda f$ is $\tau-\overline{\tau_{\mathscr{A}}}$ continuous at $(x, 0)$ and x^* is in $\partial f(x)$.*

6. *The subdifferential mapping ∂f is $\tau-\overline{\tau_{\mathscr{A}}}$ continuous at x and x^* is in $\partial f(x)$.*

7. *Whenever y converges to x in (E, τ), there is a selection $\phi(y) \in \partial f(y)$ for y near x such that $\phi(y)$ $\tau_{\mathscr{A}}$-converges to x^*.*

Proof. The implications shown in Theorem 3.1 can also be proved here either by mimicking the proofs or by showing that the condition assumed implies that $\partial f(x)$ is a singleton and appealing to Theorem 3.1. Note also that the set to be included in A^0 in condition 3 is just $\partial_\eta f(x)-x^*$.

4. **Upper semi-continuity of support face mappings.** In this section, we examine the consequences of Theorem 3.1 for a continuous convex function which is everywhere finite and non-negative and is positively homogeneous. Such a function is a continuous Minkowski functional; that is, f is specified by any of the convex bodies $U_r=\{y\in E: f(y)\le r\}$, $r>0$; for example, $f(y)=\inf\{\lambda>0: y\in\lambda U_1\}$ or $f(y)=\sup\langle y, U_1^0\rangle$ for y in E. Under these assumptions the equivalent statements in Theorem 3.1 all have natural geometric formulations.

Let $M=M(U_1^0, x)\equiv\sup\langle x, U_1^0\rangle=f(x)$. The subgradients of f at x are the (normalized) *support* functionals to U_M at x; that is, the x^* in E^* such that $x^*(y)\le x^*(x)=M$ for all y in U_M. Dually, the subdifferential mapping ∂f associates to x in E the set $F(U_1^0, x)\equiv\{x^*\in U_1^0: x^*(x)=M\}$; that is, the (weak* compact convex non-empty) *face* of U_1^0 *supported by* x.

The conjugate function f^* is zero on U_1^0 and ∞ elsewhere on E. Consequently, the ε-approximate subdifferential of f at x, $\partial_\varepsilon f(x)$, is the set $S(U_1^0, x, \varepsilon)\equiv\{y^*\in U_1^0: y^*(x)\ge M(U_1^0, x)-\varepsilon\}$; this set is called a (closed) x-*slice* of U_1^0 if $0<\varepsilon<M$.

Suppose that $M=f(x)>0$. Because of the positive homogeneity of f, just as epi f is determined by U_M, so is the tangent cone $K_{f,x}$ to epi f at $(x, f(x))$ determined by the tangent cone K_x to U_M at x, where

$$K_x\equiv\overline{\bigcup_{\lambda>0}\lambda(U_M-x)}=\{y\in E: x^*(y)\le 0 \text{ for all } x^*\in\partial f(x)\}.$$

It turns out that the 'engulfing' condition 2 of Theorem 3.1 can be replaced by a condition on K_x if we make an additional assumption on the class $\mathscr{A}$:

"Whenever $A\in\mathscr{A}$ and $I\in\mathscr{I}$ there is an $A'\in\mathscr{A}$ such that $A+Ix\subset A'$."

We say that a subset C of E is *strictly inside* K_x if $C+\varepsilon x\subset K_x$ (or $C+\varepsilon U_M\subset K_x$) for some $\varepsilon>0$; equivalently, if $C+U\subset K_x$ for some τ-neighborhood U of zero.

4.1. THEOREM. *Let f be a continuous Minkowski functional on (E, τ), let $f(x)=M>0$ and suppose that $\mathscr{A}$ satisfies the additional assumption above. Then the following are equivalent.*

1. *The Minkowski functional f is $\tau_{\mathscr{A}}$-directionally differentiable at x.*

2. *If a subset of a member of $\mathscr{A}$ is strictly inside K_x, then it is contained in $\lambda(U_M-x)$ for some $\lambda>0$.*

3. *The face $F(U_1^0, x)$ is weak* $\tau_{\mathscr{A}}$-exposed in U_1^0 by x; that is, for each A in $\mathscr{A}$, $F(U_1^0, x)+A^0$ contains an x-slice of U_1^0.*

4. *The linear functional x on E^* attains its supremum strictly on U_1^0 at $F(U_1^0, x)$ with respect to $\tau_{\mathscr{A}}$; that is, whenever y^* in U_1^0 are such that $\langle x, y^*\rangle$ converges to M, then y^* $\tau_{\mathscr{A}}$-approaches $F(U_1^0, x)$.*

5. *The slice mapping* $(y, \lambda) \to S(U_1^0, y, \lambda)$ *is* $\tau - \overline{\tau_{\mathscr{A}}}$ *u.s.c. at* $(x, 0)$.

6. *The support face mapping* $F(U_1^0, \cdot)$ *is* $\tau - \overline{\tau_{\mathscr{A}}}$ *u.s.c. at* x.

7. *If* y *converges to* x *in* (E, τ), *then for each* A *in* $\mathscr{A}$, $\inf p_A(F(U_1^0, y) - F(U_1^0, x))$ *converges to zero.*

Proof. All of the conditions except 2 are simple restatements of the corresponding conditions in Theorem 3.1. Let $\pi : E \times R \to E$ be defined by $\pi(y, r) = y - rx/M$. The extra assumption on $\mathscr{A}$ ensures that if $C \subset A \times [a, b]$ for some A in $\mathscr{A}$ and a, b in R, then $\pi(C) \subset A'$ for some A' in $\mathscr{A}$. The equivalence of 2 with the corresponding condition in Theorem 3.1 follows immediately from the following relations, all of which can be verified directly. Let $C \subset A \times [a, b]$, $B \subset E$, and $\lambda, \varepsilon > 0$. Then

$$C - (0, \varepsilon) \subset K_{f,x} \quad \text{if and only if} \quad \pi(C) + \frac{\varepsilon x}{M} \subset K_x;$$

$\pi(C) \subset \lambda(U_M - x)$ implies

$$C \subset \left(\lambda + \frac{b-a}{M}\right)[\operatorname{epi} f - (x, f(x))];$$

and

$$B \times \{0\} \subset \lambda[\operatorname{epi} f - (x, f(x))]$$

implies $B \subset \lambda(U_M - x)$.

4.2. Remarks. 1. It is important to use subsets of members of $\mathscr{A}$ in condition 2 rather than say, translates, for the engulfing depends very much on the shape of the set used. In particular, points strictly inside K_x are always engulfed and so any translated U_r that is strictly inside K_x is always engulfed too ($z + U_r$ will be contained in $\lambda(U_M - x)$ if $z + rx/M$ is).

2. If f is a seminorm (that is, if f also satisfies the condition $f(-y) = -f(y)$ for all y in E), then $\partial f(x) = U_1^0$ whenever $f(x) = 0$. Thus, for seminorms, the support mapping is always $\tau - \overline{\tau_{\mathscr{A}}}$ u.s.c. at those x for which $f(x) = 0$.

As in the previous section we have a 'single-valued' form of Theorem 4.1.

4.3. Corollary. *Under the same assumptions as in Theorem* 4.1, *the following conditions are equivalent and imply that* $F(U_1^0, x) = \{x^*\}$.

1. *The Minkowski functional* f *is* $\tau_{\mathscr{A}}$*-differentiable at* x *with* $\tau_{\mathscr{A}}$*-differential* x^*.

2. *If a subset of a member of* $\mathscr{A}$ *is strictly inside the half space* $\{y : x^*(y) \leq 0\}$, *then it is contained in* $\lambda(U_M - x)$ *for some* $\lambda > 0$. *Also,* $\langle x, x^* \rangle = M$.

3. *The point* x^* *is weak** $\tau_{\mathscr{A}}$*-exposed in* U_1^0 *by* x; *that is, each* $\tau_{\mathscr{A}}$*-neighborhood of* x^* *contains an* x*-slice of* U_1^0.

4. *The linear functional* x *on* E^* *attains its supremum strictly on* U_1^0 *at* x^* *with respect to* $\tau_{\mathscr{A}}$; *that is, whenever* y^* *in* U_1^0 *are such that* $\langle x, y^* \rangle$ *converges to* M *then* y^* $\tau_{\mathscr{A}}$*-converges to* x^*.

5. *The slice mapping* $(y, \lambda) \to S(U_1^0, y, \lambda)$ *is* $\tau - \overline{\tau_{\mathscr{A}}}$ *continuous at* $(x, 0)$ *and* $x^* \in F(U_1^0, x)$.

6. *The support face mapping* $F(U_1^0, \cdot)$ *is* $\tau - \overline{\tau_{\mathscr{A}}}$ *continuous at* x *and* $x^* \in F(U_1^0, x)$.

7. *If* y *converges to* x *in* (E, τ), *then there is a selection* $\phi(y) \in \partial f(y)$ *such that* $\phi(y)$ $\tau_{\mathscr{A}}$*-converges to* x^*.

If $f = \|\cdot\|$ is the norm on a normed linear space $E = X$, it is customary to call ∂f the duality mapping D and denote a typical member of the support face $D(x)$ by f_x. The unit ball U_1 is denoted by $B(X)$, the dual ball U_1^0 by $B(X^*)$. The norm topology on the dual X^* is the topology $\tau_{\mathscr{A}}$ where $\mathscr{A}$ is the class of all weakly (or norm) bounded absolutely convex subsets of X; that is, it is the *strong* topology on X^*. In this special case we get the following corollary of Theorem 4.1. Most of the equivalences have already been shown in [2] under the additional assumption that X is a Banach space.

4.4. Corollary. *Let* X *be a normed linear space and let* $\|x\| = 1$. *Then the following conditions are equivalent.*

1. *The norm is strongly directionally differentiable at* x.

2. *Each bounded set strictly inside* K_x *is contained in* $\lambda(B(X) - x)$ *for some* $\lambda > 0$

3. *The support face* $D(x)$ *is weak* strongly exposed in* $B(X^*)$ *by* x; *that is, for each* $\varepsilon > 0$, $D(x) + \varepsilon B(X^*)$ *contains an* x*-slice of* $B(X^*)$.

4. *The linear functional* x *on* E^* *attains its supremum strictly on* U_1^0 *at* $D(x)$ *with respect to the strong* $(=$ *norm*$)$ *topology.*

5. *The slice mapping* $(y, \lambda) \to S(U_1^0, y, \lambda)$ *is norm*$-\overline{\text{norm}}$ *u.s.c. at* $(x, 0)$.

6. *The duality mapping* D *is norm*$-\overline{\text{norm}}$ *u.s.c. at* x.

7. *Whenever* y *strongly converges to* x, *there is a selection* $f_y \in D(y)$ *such that* f_y *strongly approaches* $D(x)$.

In the special case that $D(x)$ is a singleton, we get the following corollary. Some of the equivalences are classical [5]. The equivalence of conditions 1 and 2 has been shown by J. R. Giles [6].

4.5. Corollary. *Let* X *be a normed linear space, let* f_x *be an element of* X^* *and let* $\|x\| = 1$. *Then the following conditions are equivalent and imply that* $D(x) = \{f_x\}$.

1. *The norm is strongly* $(\equiv$ *Fréchet*$)$ *differentiable at* x *with strong differential* f_x.

2. *Each bounded set strictly inside the half space* $\{y : f_x(y) \le 0\}$ *is contained in* $\lambda(B(X) - x)$ *for some* $\lambda > 0$.

3. *The point* f_x *is weak* strongly exposed in* $B(X^*)$ *by* x; *that is, each norm neighborhood of* f_x *contains an* x*-slice of* $B(X^*)$. *Also* $f_x(x) = 1$.

4. *The linear functional* x *on* X^* *attains its supremum strictly on* $B(X^*)$ *at* f_x *with respect to the strong* $(=$ *norm*$)$ *topology.*

5. *The slice mapping* $(y, \lambda) \to S(B(X^*), y, \lambda)$ *is norm–$\overline{\text{norm}}$ continuous at* $(x, 0)$ *and* $f_x \in D(x)$.

6. *The duality mapping* D *is norm–$\overline{\text{norm}}$ continuous at* x *and* $f_x \in D(x)$.

7. *Whenever* y *strongly converges to* x, *there is a selection* $f_y \in D(y)$ *such that* f_y *strongly converges to* f_x.

4.5 Remarks. 1. We have observed that all of the theorems and corollaries hold if $\tau_{\mathscr{A}}$ is the weak* topology. If the weak* rather than the strong topology is used for $\tau_{\mathscr{A}}$ in Corollary 4.5, we get equivalent conditions for the weak* ($\equiv$Gateau or weak) differentiability of the norm at x. In condition 3, it is then customary to say simply that f_x is weak* exposed (rather than weak* weak* exposed).

2. It is well-known that the set of points of strong differentiability of the norm

$$\bigcap_n \left\{ x : \sup_{y \in B(X)} \frac{\|x+\lambda y\| + \|x-\lambda y\| - 2\|x\|}{\lambda} < \frac{1}{n} \quad \text{for some} \quad \lambda > 0 \right\}$$

is a G_δ subset of X. This is not the case for the set G of points at which the duality mapping is norm–$\overline{\text{norm}}$ upper semi-continuous. For example, let X be the Banach space m of bounded sequences $x = (x_n)$ with the supremum norm. Then from condition 2 of Corollary 4.4 it is clear that G is the set of points x for which $\|x\|$ is not an accumulation point of $\{|x_n| : |x_n| \neq \|x\|\}$. The set G is dense and its complement is a dense G_δ subset. Since the intersection of two dense G_δ subsets of a Baire space must be dense, G cannot be a G_δ subset.

Acknowledgements. These results were originally stated only for the duality mapping on a Banach space [2]. I would like to thank J. Borwein for his helpful suggestions which led to Theorem 3.1 in its present general form. I would also like to thank J. R. Giles for some interesting discussions, and R. T. Rockafellar for a series of lectures (Queen's University, November, 1977) that refined my appreciation of convex functions.

References

1. E. Asplund and R. T. Rockafellar, *Gradients of convex functions*, Trans. A.M.S., **139** (1969), 443–467.

2. J. R. Giles, D. A. Gregory, and Brailey Sims, *Geometrical implications of upper semi-continuity of the duality mapping on a Banach space*, to appear in Pacific J. Math.

3. R. B. Holmes, *Geometric functional analysis and its applications*, Springer-Verlag, New York (1975).

4. P. Kenderov, *Semi-continuity of set-valued monotone mappings*, Fund. Math. **88** (1975), 61–69.

5. V. L. Smulian, *Sur la derivabilité de la norme dans l'espace de Banach*, Dokl. Akad. Nauk. SSSR., **27** (1940), 643–648.

6. J. R. Giles, *Strong differentiability of the norm and rotundity of the dual*, to appear.

Department of Mathematics and Statistics
Queen's University
Kingston, Ontario
K7L 3N6