PACIFIC JOURNAL OF MATHEMATICS
Vol. 103, No. 2, 1982

A NOTE ON ε-SUBGRADIENTS AND MAXIMAL MONOTONICITY

J. M. BORWEIN

It is our desire in this note to provide certain formulae relating subgradients, directional derivatives and ε-subgradients of proper lower semi-continuous convex functions defined on a Banach space.

Our aim is to provide these formulae, which somewhat extend those in [5], [6], [7], as a direct and hopefully straightforward consequence of Ekeland's non convex-version [3], of the Bishop-Phelps-Bronsted-Rockafellar Theorem [1], [2], [3], [4].

As a by-product we obtain somewhat more self contained proofs of the maximality of the subgradient as a monotone relation and of some related results.

1. Preliminaries. Throughout X is a real Hausdorff *locally convex space* (l.c.s) with topological dual X^*. A function $f: X \to [-\infty, \infty]$ is said to be *convex* if its *epigraph*, $\operatorname{Epi} f = \{(x, r) \mid f(x) \leqq r\}$ is a convex subset of $X \times R$. Also f is *lower semi-continuous* (l.s.c.) if Epi f is closed. We will restrict our attention to *proper* convex functions. These are the functions which are somewhere finite and never $-\infty$. The *domain* of f, dom f, is the set of points in X for which $f(x)$ is finite.

With each convex function we associate its (one-sided) *directional derivative* at x in dom f given by

$$(1) \qquad f'(x; h) = \lim_{t \downarrow 0} \frac{f(x + th) - f(x)}{t} .$$

Then $f'(x; \cdot)$ is well defined as a (possibly improper) convex positively homogeneous function. We also define, for each $\varepsilon \geqq 0$, the ε-*subgradient* set for f at x by

$$(2) \qquad \partial_\varepsilon f(x) = \{x^* \in X^* \mid x^*(h) + f(x) \leqq f(x + h) + \varepsilon, \forall h \in X\} .$$

When $\varepsilon = 0$, we supress ε and the object is the ordinary subgradient. We now may also write

$$(3) \qquad \partial f(x) = \{x^* \in X^* \mid x^*(h) \leqq f'(x; h), \forall h \in X\} .$$

For amplification about these concepts the reader is referred to [3], [4], [7].

2. The main result. We begin with a subsidiary proposition which may be found in [5] with a different proof.

PROPOSITION 1. *Let f be a lower semi-continuous proper convex function defined on a locally convex space X. For any x in the domain of f one has the following formula:*

$$f'(x; h) = \inf_{\varepsilon \downarrow 0} \sup \{x^*_\varepsilon(h) \mid x^*_\varepsilon \in \partial_\varepsilon f(x)\} . \tag{4}$$

Proof. Let $\varepsilon > 0$ and let $x^*_\varepsilon \in \partial_\varepsilon f(x)$. Then (2) shows that for $t > 0$

$$x^*_\varepsilon(h) \leqq \frac{f(x + th) - f(x) + \varepsilon}{t} .$$

We let $t = \sqrt{\varepsilon}$ and derive

$$x^*_\varepsilon(h) \leqq \frac{f(x + \sqrt{\varepsilon} h) - f(x)}{\sqrt{\varepsilon}} + \sqrt{\varepsilon} . \tag{5}$$

Then (5) and (1) combine to show that

$$f'(x; h) \geqq \limsup_{\varepsilon \downarrow 0} \{x^*_\varepsilon(h) \mid x^*_\varepsilon \in \partial_\varepsilon f(x)\} . \tag{6}$$

Conversely, let d be any real number less than $f'(x; h)$, and let $\varepsilon > 0$ be given. For $0 \leqq t \leqq 1$ one has

$$f(x + th) \geqq f(x) + td . \tag{7}$$

Thus the line segment

$$L = \{(x, f(x) - \varepsilon) + t(h, d) \mid 0 \leqq t \leqq 1\} \tag{8}$$

can be strictly separated from the closed convex set Epi f, [4]. Simple and standard calculation shows that any separating functional $(x^*, -r^*)$ in $X^* \times R$ satisfies $r^* > 0$ and that

$$\left(\frac{x^*}{r^*}\right)(h) \geqq d - \varepsilon;\ \frac{x^*}{r^*} \in \partial_\varepsilon f(x) . \tag{9}$$

The nature of d and (9) show that

$$\sup \{x^*(h) \mid x^* \in \partial_\varepsilon f(x)\} \geqq f'(x; h) - \varepsilon . \tag{10}$$

It is clear from (6) and (10) that (4) holds. □

If f is actually continuous at x then $\partial_\varepsilon f(x)$ is weak-star compact [4], and (4) reduces to the standard formula

$$f'(x; h) = \sup \{x^*(h) \mid x^* \in \partial f(x)\} . \tag{11}$$

Even in finite dimensions (11) can fail at a point of discontinuity, while in Fréchet space it is possible that ∂f is empty [4], [5]. In Banach space Rockafellar [5], [6], has given formulae replacing (11), in terms of approximations by subgradients at nearby points. Taylor [8] has given an alternative stronger formula. All these results follow from some form of the Bishop-Phelps [1] or Bronsted-Rockafellar [2] theorems. We now proceed to derive a strong version of Taylor's formula which uses Ekeland's variational form of the previously mentioned theorems [3].

THEOREM 1. *Let f be a proper convex lower semi-continuous function defined on a Banach space $(X, \|\cdot\|)$. Suppose that $\varepsilon > 0$ and $t \geqq 0$ are given. Suppose that*

$$x_0^* \in \partial_\varepsilon f(x_0) \ . \tag{12}$$

Then one may find points x_ε and x_ε^* such that

$$x_\varepsilon^* \in \partial f(x_\varepsilon) \ , \tag{13}$$

and such that

$$\|x_\varepsilon - x_0\| \leqq \sqrt{\varepsilon} \ , \tag{14}$$

$$|f(x_\varepsilon) - f(x_0)| \leqq \sqrt{\varepsilon}\left(\sqrt{\varepsilon} + \frac{1}{t}\right) , \tag{15}$$

$$\|x_\varepsilon^* - x_0^*\| \leqq \sqrt{\varepsilon}\ (1 + t\|x_0^*\|) \ , \tag{16}$$

$$|x_\varepsilon^*(h) - x_0^*(h)| \leqq \sqrt{\varepsilon}\ (\|h\| + t|x_0^*(h)|) \ , \tag{17}$$

$$x_\varepsilon^* \in \partial_{2\varepsilon} f(x_0) \ . \tag{18}$$

Proof. We renorm X using the equivalent norm given by

$$\|x\|_t = \|x\| + t|x_0^*(x)| \ . \tag{19}$$

We set $g(x) = f(x) - x_0^*(x)$ and observe that g is l.s-c. and that

$$g(x_0) \leqq \varepsilon + \inf_X g(x) \ . \tag{20}$$

We now apply Ekeland's theorem [3, p. 29] to g and $\|\cdot\|_t$. We are promised the existence of x_ε in X such that, for $x \neq x_\varepsilon$,

$$g(x) + \sqrt{\varepsilon}\ \|x - x_\varepsilon\|_t > g(x_\varepsilon) \tag{21}$$

and

$$g(x_\varepsilon) + \sqrt{\varepsilon}\ \|x_0 - x_\varepsilon\|_t \leqq g(x_0) \ . \tag{22}$$

Now (21) can be read as saying that

$$0 \in \partial(g + \sqrt{\varepsilon h})(x_\varepsilon);\ h(x) = \|x - x_\varepsilon\|_t\ . \tag{23}$$

Since h is continuous, and since

$$\partial h(x_\varepsilon) = \{x^* + \alpha x_0^* \mid \|x^*\| \leqq 1,\ |\alpha| \leqq t\}\ , \tag{24}$$

we may write, using the subgradient sum formula [4],

$$0 \in \partial f(x_\varepsilon) - x_0^* + \sqrt{\varepsilon}\,\partial h(x_\varepsilon)\ . \tag{25}$$

Hence there is some point x_ε^* in $\partial f(x_\varepsilon)$ of the form

$$x_\varepsilon^* = \sqrt{\varepsilon}\,x^* + (1 - \sqrt{\varepsilon}\,\alpha(t))x_0^* \tag{26}$$

with $|\alpha(t)| \leqq t$ and $\|x^*\| \leqq 1$. Thus (16) holds. Since (20) holds (22) shows that

$$\sqrt{\varepsilon}\ \|x_0 - x_\varepsilon\| + \sqrt{\varepsilon}\,t\,|x_0^*(x_0 - x_\varepsilon)| \leqq \varepsilon\ . \tag{27}$$

In particular (14) holds and

$$|x_0^*(x_0 - x_\varepsilon)| \leqq \sqrt{\varepsilon}/t\ . \tag{28}$$

Combined use of (20) and (22) shows that

$$|f(x_\varepsilon) - f(x_0)| \leqq |x_0^*(x_\varepsilon - x_0)| + \varepsilon\ . \tag{29}$$

Now (15) follows from (28) and (29). Also (26) shows that

$$\begin{aligned} \|x_\varepsilon^*(h) - x_0^*(h)\| &\leqq \sqrt{\varepsilon}(\|h\| + |\alpha(t)||x_0^*(h)|) \\ &\leqq \sqrt{\varepsilon}\,(\|h\| + t|x_0^*(h)|)\ . \end{aligned} \tag{30}$$

Finally, since $x_\varepsilon^* \in \partial f(x_\varepsilon)$,

$$\begin{aligned} x_\varepsilon^*(x - x_0) &\leqq x_\varepsilon^*(x - x_\varepsilon) + x_\varepsilon^*(x_\varepsilon - x_0) \\ &\leqq f(x) - f(x_0) + [f(x_0) - f(x_\varepsilon) + x_0^*(x_\varepsilon - x_0)] \\ &\quad + (x_\varepsilon^* - x_0^*)(x_\varepsilon - x_0)\ . \end{aligned} \tag{31}$$

Since (12) holds, $f(x_0) - f(x_\varepsilon) + x_0^*(x_\varepsilon - x_0) \leqq \varepsilon$, and since (26) holds,

$$|(x_\varepsilon^* - x_0^*)(x_\varepsilon - x_0)| \leqq \sqrt{\varepsilon}\,(\|x_\varepsilon - x_0\| + t|x_0^*(x_\varepsilon - x_0)|) \leqq \varepsilon\ ,$$

on using (27). Then (31) establishes (18). Observe that, with the convention that $1/0 = \infty$, the arguments are preserved when $t = 0$. Let us also observe that back substitution of (29) into (27) produces a strengthening of (15) to

$$\|x_\varepsilon - x_0\| + t|f(x_\varepsilon) - f(x_0)| \leqq \sqrt{\varepsilon} + t\varepsilon\ , \tag{15$'$}$$

which is slightly less convenient for application. □

REMARKS. (1) Our purposes in producing this proof with a parameter t are three-fold: (a) it leads to a unified development of the Bronsted-Rockafellar theorem ($t = 0$) and the improvement of the Taylor result ($t = 1$) and allows one to see the differences in the relative approximations in, for example, (15) and (16); (b) since one wishes to approximate in direction x_0^* it is intuitively plausible that $\|\cdot\|_t$ is the appropriate norm to use; (c) for all the details the proof is really very straightforward and essentially reduces to "apply Ekeland's theorem to g and $\|\cdot\|_t$". Notice that (17), which is critical to the next result, is considerably more useful than (16) in relating $x_0^*(h)$ and $x_\varepsilon^*(h)$ as ε varies. This is because while $\|x_\varepsilon^*\|$ typically will grow unboundedly as ε shrinks, $|x_\varepsilon^*(h)|$ can generally be given a uniform bound independent of ε.

THEOREM 2. *Let f be a proper convex lower semi-continuous convex function on a Banach space $(X, \|\cdot\|)$. Then, for any x_0 in the domain of f and any h in X,*

$$f'(x_0; h) = \inf_{\varepsilon \downarrow 0} \sup \{x_\varepsilon^*(h) \mid x_\varepsilon^* \in S_\varepsilon(x_0)\} \tag{32}$$

where

$$x_\varepsilon^* \in S_\varepsilon(x_0) \Longleftrightarrow \begin{cases} \text{(i)} & x_\varepsilon^* \in \partial f(x_\varepsilon)\ , \\ \text{(ii)} & \|x_\varepsilon - x_0\| \leqq \varepsilon\ , \\ \text{(iii)} & |f(x_\varepsilon) - f(x_0)| \leqq \varepsilon\ , \\ \text{(iv)} & x_\varepsilon^* \in \partial_\varepsilon f(x_0)\ . \end{cases} \tag{33}$$

Proof. Since $S_\varepsilon(h) \subset \partial_\varepsilon f(x_0)$, Proposition 1 shows that it suffices to establish that the right hand side of (32) is no smaller than the left hand side. Suppose first that $f'(x_0; h) = d < \infty$. Set $1 > \delta > 0$ and pick x_0^*, using Proposition 1(9), so that $x_0^*(h) \geqq d - \delta$ and $x_0^* \in \partial_\delta f(x_0)$. Let us apply Theorem 1 to this x_0^* with $t = 1$, $\delta = \varepsilon$. Then we obtain points x_δ^* and x_δ with $x_\delta^* \in \partial f(x_\delta)$ which on relabeling satisfy (33) with $\varepsilon = 2\sqrt{\delta}$. Also (17) shows that

$$x_\delta^*(h) \geqq x_0^*(h) - \sqrt{\delta}(\|h\| + |x_0^*(h)|)\ . \tag{34}$$

For sufficiently small δ, $x_0^*(h) \leqq d + 1$, as follows from (5). Thus

$$x_\delta^*(h) \geqq d - \delta - \sqrt{\delta}(\|h\| - |d| - 1)\ . \tag{35}$$

Since the right hand side of this expression tends to d as δ tends to zero, (32) is established in this case. Suppose now that $f'(x_0; h) =$

∞. Proposition 1 shows that we can pick $x_0^* \in \partial_\delta f(x_0)$ and $x_0^*(h) \geqq 1/\delta$. As before (33) holds with $\varepsilon = 2\sqrt{\delta}$. In this case (34) implies that

$$(36) \qquad x_\delta^*(h) \geqq \left(1 - \sqrt{\delta}\right)\frac{1}{\delta} - \|h\| ,$$

and now the right hand side has supremum infinity. Again (32) is established. □

The approximation in (32) is very strong as we may actually pick subgradients at points which are nearer and nearer x_0 and have converging function values. Observe that application of Theorem 1 with $t = 0$ leads to Theorem 2 except for (33) (iii).

One may recover Taylor's formula [8] on replacing (33) (iii) and (iv) by

$$(37) \qquad |x_\varepsilon^*(x_\varepsilon - x_0)| \leqq \varepsilon$$

and observing that (37) follows from (33) (i), (ii), (iv) since

$$|x_\delta^*(x_\delta - x_0)| \leqq |f(x_\delta) - f(x_0)| + \delta$$

if $x_\delta^* \in \partial_\delta f(x_0) \cap \partial f(x_\delta)$. Thus Taylor's approximating subset is a bigger set than ours. Since (37), (33) (i) and (ii) still force $x_\varepsilon^* \in \partial_{2\varepsilon} f(x_0)$ for small ε, (32) still holds. Indeed, except for scale constants our Theorem 2 and Taylor's Corollary 1 are interderivable.

Recall that ∂f is a *monotone relation* [3]: if $x_i^* \in \partial f(x_i)$ $(i = 1, 2)$ then

$$(38) \qquad (x_2^* - x_1^*)(x_2 - x_1) \geqq 0 .$$

Rockafellar [5] produced a proof that ∂f is always maximal as a monotone relation. Rockafellar's proof was irremediably flawed and he subsequently gave a correct proof using conjugate functions in [6]. Taylor [8] then produced an essentially correct proof more in the spirit of [5]. This proof is slightly flawed technically ($d < \infty$ is assumed). We provide here a derivation of the result from Theorem 2.

COROLLARY 1. *If f is a proper lower semi-continuous convex function on a Banach space X then ∂f is maximal as a monotone relation in $X \times X^*$.*

Proof. As in [5], [8] we may assume by translation that $0 \notin \partial f(0)$. A one dimensional argument now produces a point x_0 in $\operatorname{dom} f$ with $f'(x_0; -x_0) > 2\delta > 0$. Note that it may well be that

$f'(x_0; -x_0)$ is infinite, contrary to the implicit assumption in [8]. By any account, we have, from Theorem 2, points x_δ and $x_\delta^* \in \partial f(x_\delta)$ with

$$\text{(i)}\quad x_\delta^*(-x_0) > 2\delta\,,$$
$$\text{(ii)}\quad x_\delta^* \in \partial_\delta f(x_0)\,, \tag{39}$$
$$\text{(iii)}\quad |f(x_\delta) - f(x_0)| \leqq \delta\,.$$

Since (ii) holds

$$x_\delta^*(x_\delta) \leqq x_\delta^*(x_0) + f(x_\delta) - f(x_0) + \delta \tag{40}$$

and thus (i) and (iii) combine to show

$$x_\delta^*(x_\delta) < -2\delta + \delta + \delta < 0\,. \tag{41}$$

Thus one cannot have $(x^* - 0)(x - 0) > 0$ for each $x^* \in \partial f(x)$ and so ∂f is maximal. □

COROLLARY 2. *If f is a proper convex lower semi-continuous function on a Banach space X then*

$$f(x) = \lim_{\varepsilon \downarrow 0} \sup \{x^*(x) - f^*(x^*) \mid x^* \in S_\varepsilon(x)\}\,, \tag{42}$$

where f^* is the conjugate function

$$f^*(x^*) = \sup \{x^*(x) - f(x) \mid x \in \operatorname{dom} f\}\,. \tag{43}$$

Proof. For any x^* in the nonempty set $S_\varepsilon(x)$ one has

$$x^*(y) - f(y) \leqq x^*(x) - f(x) + \varepsilon \tag{44}$$

or

$$x^*(x) - f^*(x^*) \geqq f(x) - \varepsilon\,. \tag{45}$$

Thus the right hand side of (42) dominates $f(x)$. The opposite inequality follows directly from (43) or Young's inequality. □

COROLLARY 3. *If f is a proper, lower semi-continuous convex function on a Banach space X the following mean-value theorem holds. For each x_1 and x_2 in* dom f *one can find z in (x_1, x_2) and sequences of points $\{z_n\}$ in X and $\{z_n^*\}$ in X^* with*

$$\text{(i)}\quad \|z_n - z\| \leqq \frac{1}{n}\,,$$
$$\text{(ii)}\quad |f(z_n) - f(z)| \leqq \frac{1}{n}\,, \tag{46}$$
$$\text{(iii)}\quad z_n^* \in \partial f(z_n)\,,$$

$$\text{(iv)} \quad z_n^* \in \partial \frac{1}{n} f(z) ,$$

and such that

$$\lim_n z_n^*(x_1 - x_2) = f(x_1) - f(x_2) . \tag{47}$$

Proof. It is straightforward to show that for some z in (x_1, x_2) one has

$$f'(z; x_1 - x_2) \geqq f(x_1) - f(x_2) \geqq -f'(z; x_2 - x_1) . \tag{48}$$

The result now follows from Theorem 1. □

In the case that f is continuous at z, as observed before $\partial_\varepsilon f(z)$ is w^* compact, and (47) reduces to the better known

$$f(x_1) - f(x_2) \in \partial f(z)(x_1 - x_2) . \tag{49}$$

REFERENCES

1. F. Bishop and R. R. Phelps, *The support functionals of a convex set,* Convexity, Proceedings Sympos. Pure Math., **7** Amer. Math. Soc., Providence, R.I., (1962), 27–35.
2. A. Bronsted and R. T. Rockafellar, *On the subdifferentiability of convex functions,* Proc. Amer. Math. Soc., **16** (1965), 605–611.
3. I. Ekeland and R. Temam, *Convex analysis and variational problems,* North-Holland, Amsterdam, 1976.
4. R. B. Holmes, *Geometric Functional Analysis,* Springer-Verlag, New York, 1975.
5. R. T. Rockafellar, *Characterization of the subdifferentials of convex functions,* Pacific J. Math., **17** (1966), 497–510.
6. ———, *On the maximal monotonicity of subdifferential mappings,* Pacific J. Math., **33** (1970), 209–216.
7. ———, *Convex analysis,* Princeton University Press, Princeton, N.J., 1970.
8. P. D. Taylor, *Subgradients of a convex function obtained from a directional derivative,* Pacific J. Math., **44** (1973), 739–747.

Received December 19, 1980.

DALHOUSIE UNIVERSITY
HALIFAX, NOVA SCOTIA