Pacific Journal of Mathematics

# ON THE MAXIMAL MONOTONICITY OF SUBDIFFERENTIAL MAPPINGS

Ralph Tyrrell Rockafellar

Vol. 33, No. 1 March 1970

PACIFIC JOURNAL OF MATHEMATICS
Vol. 33, No. 1, 1970

# ON THE MAXIMAL MONOTONICITY OF SUBDIFFERENTIAL MAPPINGS

R. T. ROCKAFELLAR

**The subdifferential of a lower semicontinuous proper convex function on a Banach space is a maximal monotone operator, as well as a maximal cyclically monotone operator. This result was announced by the author in a previous paper, but the argument given there was incomplete; the result is proved here by a different method, which is simpler in the case of reflexive Banach spaces. At the same time, a new fact is established about the relationship between the subdifferential of a convex function and the subdifferential of its conjugate in the nonreflexive case.**

Let $E$ be a real Banach space with dual $E^*$. A *proper convex function* on $E$ is a function $f$ from $E$ to $(-\infty, +\infty]$, not identically $+\infty$, such that

$$f((1-\lambda)x + \lambda y) \leqq (1-\lambda)f(x) + \lambda f(y)$$

whenever $x \in E, y \in E$ and $0 < \lambda < 1$. The *subdifferential* of such a function $f$ is the (generally multivalued) mapping $\partial f: E \rightarrow E^*$ defined by

$$\partial f(x) = \{x^* \in E^* \mid f(y) \geqq f(x) + \langle y - x, x^* \rangle, \forall y \in E\} ,$$

where $\langle \cdot, \cdot \rangle$ denotes the canonical pairing between $E$ and $E^*$.

A multivalued mapping $T: E \rightarrow E^*$ is said to be a *monotone operator* if

$$\langle x_0 - x_1, x_0^* - x_1^* \rangle \geqq 0 \quad \text{whenever} \quad x_0^* \in T(x_0), x_1^* \in T(x_1) .$$

It is said to be a *cyclically monotone operator* if

$$\langle x_0 - x_1, x_0^* \rangle + \cdots + \langle x_{n-1} - x_n, x_{n-1}^* \rangle + \langle x_n - x_0, x_n^* \rangle \geqq 0$$
$$\text{whenever} \quad x_i^* \in T(x_i), i = 0, \cdots, n .$$

It is called a *maximal* monotone operator (resp. maximal cyclically monotone operator) if, in addition, its graph

$$G(T) = \{(x, x^*) \mid x^* \in T(x)\} \subset E \times E^*$$

is not properly contained in the graph of any other monotone (resp. cyclically monotone) operator $T': E \rightarrow E^*$.

This note is concerned with proving the following theorems.

THEOREM A. *If $f$ is a lower semicontinuous proper convex function on $E$, then $\partial f$ is a maximal monotone operator from $E$ to $E^*$.*

THEOREM B. *Let $T: E \to E^*$ be a multivalued mapping. In order that there exist a lower semicontinuous proper convex function $f$ on $E$ such that $T = \partial f$, it is necessary and sufficient that $T$ be a maximal cyclically monotone operator. Moreover, in this case $T$ determines $f$ uniquely up to an additive constant.*

These theorems have previously been stated by us in [4] as Theorem 4 and Theorem 3, respectively. However, a gap occurs in the proofs in [4], as has kindly been brought to our attention recently by H. Brézis. (It is not clear whether formula (4.7) in the proof of Theorem 3 of [4] will hold for $\varepsilon$ sufficiently small, because $x_i^*$ depends on $\varepsilon$ and could conceivably increase unboundedly in norm as $\varepsilon$ decreases to 0. The same oversight appears in the penultimate sentence of the proof of Theorem 4 of [4]). In view of this oversight, the proofs in [4] are incomplete; further arguments must be given before the maximality in Theorem A, the maximality in the necessary condition in Theorem B, and the uniqueness in Theorem B can be regarded as established. Such arguments will be given here.

**2. Preliminary result.** Let $f$ be a lower semicontinuous proper convex function on $E$. (For proper convex functions, lower semicontinuity in the strong topology of $E$ is the same as lower semicontinuity in the weak topology.) The conjugate of $f$ is the function $f^*$ on $E^*$ defined by

$$f^*(x^*) = \sup\{\langle x, x^*\rangle - f(x) \mid x \in E\} . \tag{2.1}$$

It is known that $f^*$ is a weak* lower semicontinuous (and hence strongly lower semicontinuous) proper convex function on $E^*$, and that

$$\begin{gathered} f(x) + f^*(x^*) - \langle x, x^*\rangle \geqq 0,\ \forall x \in E,\ \forall x^* \in E^* , \\ \text{with equality if and only if } x^* \in \partial f(x) \end{gathered} \tag{2.2}$$

(see Moreau [3, § 6]). The subdifferential $\partial f^*$, which is a multivalued mapping from $E^*$ to the bidual $E^{**}$, can be compared with the subdifferential $\partial f$ from $E$ to $E^*$, when $E$ is regarded in the canonical way as a weak** dense subspace of $E^{**}$ (the weak** topology being the weak topology induced on $E^{**}$ by $E^*$). Facts about the relationship between $\partial f^*$ and $\partial f$ will be used below in proving Theorems A and B.

In terms of the conjugate $f^{**}$ of $f^*$, which is the weak** lower semicontinuous proper convex function on $E^{**}$ defined by

(2.3) $$f^{**}(x^{**}) = \sup\{\langle x^{**}, x^*\rangle - f^*(x^*) \mid x^* \in E^*\}\ ,$$

we have, as in (2.2),

(2.4) $$f^{**}(x^{**}) + f^*(x^*) - \langle x^{**}, x^*\rangle \geqq 0,\ \forall x^{**} \in E^{**},\ \forall x^* \in E^*\ ,$$ $$\text{with equality if and only if } x^{**} \in \partial f^*(x^*)\ .$$

Moreover, the restriction of $f^{**}$ to $E$ is $f$(see [3, § 6]). Thus, if $E$ is reflexive, we can identify $f^{**}$ with $f$, and it follows from (2.2) and (2.4) that $\partial f^*$ is just the "inverse" of $\partial f$, in other words one has $x \in \partial f^*(x^*)$ if and only if $x^* \in \partial f(x)$. If $E$ is not reflexive, the relationship between $\partial f^*$ and $\partial f$ is more complicated, but $\partial f^*$ and $\partial f$ still completely determine each other, according to the following result.

PROPOSITION 1. *Let $f$ be a lower semicontinuous proper convex function on $E$, and let $x^* \in E^*$ and $x^{**} \in E^{**}$. Then $x^{**} \in \partial f^*(x^*)$ if and only if there exists a net $\{x_i^* \mid i \in I\}$ in $E^*$ converging to $x^*$ in the strong topology and a bounded net $\{x_i \mid i \in I\}$ in $E$ (with the same partially ordered index set $I$) converging to $x^{**}$ in the weak$^{**}$ topology, such that $x_i^* \in \partial f(x_i)$ for every $i \in I$.*

*Proof.* The sufficiency of the condition is easy to prove. Given nets as described, we have

$$f(x_i) + f^*(x_i^*) = \langle x_i, x_i^*\rangle,\ \forall i \in I$$

by (2.2), where $f(x_i) = f^{**}(x_i)$. Then by the lower semicontinuity of $f^*$ and $f^{**}$ we have

$$\begin{aligned} f^{**}(x^{**}) + f^*(x^*) &\leqq \liminf\{f^{**}(x_i) + f^*(x_i^*)\} \\ &= \lim\langle x_i, x_i^*\rangle = \langle x^{**}, x^*\rangle\ . \end{aligned}$$

(The last equality makes use of the boundedness of the norms $\|x_i\|$, $i \in I$.) Thus $x^{**} \in \partial f^*(x^*)$ by (2.4).

To prove the necessity of the condition, we demonstrate first that, given any $x^{**} \in E^{**}$, there exists a *bounded* net $\{y_i \mid i \in I\}$ in $E$ such that $y_i$ converges to $x^{**}$ in the weak$^{**}$ topology and

(2.5) $$\lim f(y_i) = f^{**}(x^{**})\ .$$

Consider $f + h_\alpha$, where $\alpha$ is a positive real number and $h_\alpha$ is the lower semicontinuous proper convex function on $E$ defined by

(2.6) $$h_\alpha(x) = 0 \quad \text{if} \quad \|x\| \leqq \alpha,\ h_\alpha(x) = +\infty \quad \text{if} \quad \|x\| > \alpha\ .$$

Assuming that $\alpha$ is sufficiently large, there exist points $x$ at which $f$ and $h_\alpha$ are both finite and $h_\alpha$ is continuous (i.e., points $x$ such that $f(x) < +\infty$ and $\|x\| < \alpha$). Then, by the formulas for conjugates of sums of convex functions (see Moreau [3, pp. 38, 56, 57] or Rockafellar [5, Th. 3]), we have $(f + h_\alpha)^* = f^* \square h_\alpha^*$ (infimal convolution), and consequently

$$(f + h_\alpha)^{**} = (f^* \square h_\alpha^*)^* = f^{**} + h_\alpha^{**} . \tag{2.7}$$

Moreover $h_\alpha^*(x^*) = \alpha \|x^*\|$ for ever $x^* \in E^*$, so that

$$h_\alpha^{**}(x^{**}) = \sup\{\langle x^{**}, x^*\rangle - \alpha \|x^*\| \mid x^* \in E^*\}$$
$$= \begin{cases} 0 & \text{if } \|x^{**}\| \leqq \alpha , \\ +\infty & \text{if } \|x^{**}\| > \alpha . \end{cases}$$

Hence by (2.7), given any $x^{**} \in E^{**}$, we have

$$f^{**}(x^{**}) = (f + h_\alpha)^{**}(x^{**}) \tag{2.8}$$

for sufficiently large $\alpha > 0$. On the other hand, it is known that, for any lower semicontinuous proper convex function $g$ on $E$, $g^{**}$ is the greatest weak** lower semicontinuous function on $E^{**}$ majorized by $g$ on $E$ (see [3, § 6]), so that

$$g^{**}(x^{**}) = \liminf_{y \to x^{**}} g(y) , \tag{2.9}$$

where the "lim inf" is taken over all nets in $E$ converging to $x^{**}$ in the weak** topology. Taking $g = f + h_\alpha$, we see from (2.8) and (2.9) that

$$f^{**}(x^{**}) = \liminf_{y \to x^{**}} [f(y) + h_\alpha(y)] ,$$

implying that (2.5) holds as desired for some net $\{y_i \mid i \in I\}$ in $E$ such that $y_i$ converges to $x^{**}$ in the weak** topology and $\|y_i\| \leqq \alpha$ for every $i \in I$.

Now, given any $x^* \in E^*$ and $x^{**} \in \partial f^*(x^*)$, let $\{y_i \mid i \in I\}$ be a bounded net in $E$ such that $y_i$ converges to $x^{**}$ in the weak** topology and (2.5) holds. Define $\varepsilon_i \geqq 0$ by

$$\varepsilon_i^2 = f(y_i) + f^*(x^*) - \langle y_i, x^*\rangle .$$

Note that $\lim \varepsilon_i = 0$ by (2.5) and (2.4). According to a lemma of Brøndsted and Rockafellar [1, p. 608], there exist for each $i \in I$ an $x_i \in E$ and an $x_i^* \in E^*$ such that

$$x_i^* \in \partial f(x_i),\ \|x_i - y_i\| \leqq \varepsilon_i,\ \|x_i^* - x^*\| \leqq \varepsilon_i .$$

The latter two conditions imply that the net $\{x_i^* \mid i \in I\}$ converges to $x^*$ in the strong topology of $E^*$, while the net $\{x_i \mid i \in I\}$ is bounded and converges to $x^{**}$ in the weak** topology of $E^{**}$. This completes the proof of Proposition 1.

**3. Proofs of Theorems *A* and *B*.** In the sequel, $f$ denotes a lower semicontinuous proper convex function on $E$, and $j$ denotes the continuous convex function $E$ defined by $j(x) = (1/2)\|x\|^2$. We shall make use of the fact that, for each $x \in E$, $\partial f(x)$ is by definition a certain (possibly empty, possibly unbounded) weak* closed convex subset of $E^*$, whereas $\partial j(x)$ is (by the finiteness and continuity of $j$, see [3, p. 60]) a certain nonempty weak* compact convex subset of $E^*$. Furthermore

$$\partial(f + j) = \partial f(x) + \partial j(x), \ \forall x \in E \tag{3.1}$$

(see [3, p. 62] or [5, Th. 3]). The conjugate of $j$ is given by $j^*(x^*) = (1/2)\|x^*\|^2$, and since

$$(f + j)^*(x^*) = (f^* \square j^*)(x^*) = \min_{y^* \in E^*} \{f^*(y^*) + j^*(x^* - y^*)\}$$

([3, § 9] or [5, Th. 3]) the conjugate function $(f + j)^*$ is finite and continuous throughout $E^*$.

*Proof of Theorem* A. Theorem A has already been established by Minty [2] in the case of convex functions which, like $j$, are everywhere finite and continuous. Applying Minty's result to the function $(f + j)^*$, we may conclude that $\partial(f + j)^*$ is a maximal monotone operator from $E^*$ to $E^{**}$. We shall show this implies that $\partial f$ is a maximal monotone operator from $E$ to $E^*$.

Let $T$ be a monotone operator from $E$ to $E^*$ such that the graph of $T$ includes the graph of $\partial f$, i.e.,

$$T(x) \supset \partial f(x), \ \forall x \in E . \tag{3.2}$$

We must show that equality necessarily holds in (3.2).

The mapping $T + \partial j$ defined by

$$\begin{aligned}(T + \partial j)(x) &= T(x) + \partial j(x) \\ &= \{x_1^* + x_2^* \mid x_1^* \in T(x), x_2^* \in \partial j(x)\}\end{aligned}$$

is a monotone operator from $E$ to $E^*$, since $T$ and $\partial j$ are, and by (3.1) and (3.2) we have

$$(T + \partial j)(x) \supset \partial(f + j)(x), \ \forall x \in E . \tag{3.3}$$

Let $S$ be the multivalued mapping from $E^*$ to $E^{**}$ defined as follows: $x^{**} \in S(x^*)$ if and only if there exists a net $\{x_i^* \mid i \in I\}$ in $E^*$ converging to $x^*$ in the strong topology, and a bounded net $\{x_i \mid i \in I\}$ in $E$ (with the same partially ordered index set $I$) converging to $x^{**}$ in the weak** topology, such that

$$x_i^* \in (T + \partial j)(x_i), \ \forall i \in I .$$

It is readily verified that $S$ is a monotone operator. (The boundedness of the nets $\{x_i \mid i \in I\}$ enters in here.) Moreover

$$(3.4) \qquad S(x^*) \supset \partial(f + j)^*(x^*), \ \forall x^* \in E^* ,$$

by (3.3) and Proposition 1. Since $\partial(f + j)^*$ is a maximal monotone operator, equality must actually hold in (3.4). This shows that one has $x \in \partial(f + j)^*(x^*)$ whenever $x \in E$ and $x \in S(x^*)$, hence in particular whenever $x^* \in (T + \partial j)(x)$. On the other hand, one always has $x^* \in \partial(f + j)(x)$ if $x \in \partial(f + j)^*(x^*)$ and $x \in E$. (This follows from applying (2.2) and (2.4) to $f + j$ in place of $f$.) Thus one has $x^* \in \partial(f + j)(x)$ if $x^* \in (T + \partial j)(x)$, implying by (3.3) and (3.1) that

$$(3.5) \qquad T(x) + \partial j(x) = \partial f(x) + \partial j(x), \ \forall x \in E .$$

We shall show now from (3.5) that actually

$$T(x) = \partial f(x), \ \forall x \in E ,$$

so that $\partial f$ must be a maximal monotone operator as claimed. Suppose that $x \in E$ is such that the inclusion in (3.2) is proper. This will lead to a contradiction. Since $\partial f(x)$ is a weak* closed convex subset of $E^*$, there must exist some point of $T(x)$ which can be separated strictly from $\partial f(x)$ be a weak* closed hyperplane. Thus, for a certain $y \in E$, we have

$$\sup \{\langle y, x^*\rangle \mid x^* \in T(x)\} > \sup \{\langle y, x^*\rangle \mid x^* \in \partial f(x)\} .$$

But then

$$\begin{aligned} &\sup \{\langle y, z^*\rangle \mid z^* \in T(x) + \partial j(x)\} \\ &= \sup \{\langle y, x^*\rangle \mid x^* \in T(x)\} + \sup \{\langle y, y^*\rangle \mid y^* \in \partial j(x)\} \\ &> \sup \{\langle y, x^*\rangle \mid x^* \in \partial f(x)\} + \sup \{\langle y, y^*\rangle \mid y^* \in \partial j(x)\} \\ &= \sup \{\langle y, z^*\rangle \mid z^* \in \partial f(x) + \partial j(x)\} , \end{aligned}$$

inasmuch as $\partial j(x)$ is a nonempty bounded set, and this inequality is incompatible with (3.5).

*Proof of Theorem* B. Let $g$ be a lower semicontinuous proper convex function on $E$ such that

$$(3.6) \qquad \partial g(x) \supset \partial f(x), \ \forall x \in E .$$

As noted at the beginning of the proof Theorem 3 of [4], to prove Theorem B it suffices, in view of Theorem 1 of [4] and its Corollary 2, to demonstrate that $g = f + \text{const.}$

We consider first the case where $f$ and $g$ are everywhere finite and continuous. Then, for each $x \in E$, $\partial f(x)$ is a nonempty weak* compact set, and

$$f'(x; u) = \max\{\langle u, x^*\rangle \mid x^* \in \partial f(x)\},\ \forall u \in E\,, \tag{3.7}$$

where

$$f'(x; u) = \lim_{\lambda \downarrow 0} [f(x + \lambda u) - f(x)]/\lambda$$

[3, p. 65]. Similarly, $\partial g(x)$ is a nonempty weak* compact set, and

$$g'(x; u) = \max\{\langle u, x^*\rangle \mid x^* \in \partial g(x)\},\ \forall u \in E\,. \tag{3.8}$$

It follows from (3.6), (3.7) and (3.8) that

$$f'(x; u) \leqq g'(x; u),\ \forall x \in E,\ \forall u \in E\,. \tag{3.9}$$

On the other hand, for any $x \in E$ and $y \in E$, we have

$$f(y) - f(x) = \int_0^1 f'((1-\lambda)x + \lambda y; y - x)d\lambda\,,$$

$$g(y) - g(x) = \int_0^1 g'((1-\lambda)x + \lambda y; y - x)d\lambda$$

(see [6, § 24]), so that by (3.9) we have

$$f(y) - f(x) \leqq g(y) - g(x),\ \forall x \in E,\ \forall y \in E\,.$$

Of course, the latter can hold only if $g = f + \text{const.}$

In the general case, we observe from (3.6) that

$$\partial g(x) + \partial j(x) \supset \partial f(x) + \partial j(x),\ \forall x \in E\,,$$

and consequently

$$\partial(g + j)(x) \supset \partial(f + j)(x),\ \forall x \in E\,,$$

by (3.1)(and its counterpart for $g$). This implies by Proposition 1 that

$$\partial(g + j)^*(x^*) \supset \partial(f + j)^*(x^*),\ \forall x^* \in E^*\,. \tag{3.10}$$

The functions $(f + j)^*$ and $(g + j)^*$ are finite and continuous on $E^*$, so we may conclude from (3.10) and the case already considered that

$$(g + j)^* = (f + j)^* + \alpha$$

for a certain real constant $\alpha$. Taking conjugates, we then have

$$(g + j)^{**} = (f + j)^{**} - \alpha\,. \tag{3.11}$$

Since $(g + j)^{**}$ and $(f + j)^{**}$ agree on $E$ with $g + j$ and $f + j$, respectively, (3.11) implies that

$$g + j = f + j - \alpha\,,$$

and hence that $g = f + \text{const.}$

REMARK. The preceding proofs become much simpler if $E$ is reflexive, since then $\partial f^*$ and $\partial(f + j)^*$ are just the "inverses" of $\partial f$ and $\partial(f + j)$, respectively, and Proposition 1 is superfluous. In this case, $S$ may be replaced by the inverse of $T + \partial j$ in the proof of Theorem A.

## REFERENCES

1. A. Brøndsted and R. T. Rockafellar, *On the subdifferentiability of convex functions*, Proc. Amer. Math. Soc. **16** (1965), 605-611.

2. G. J. Minty, *On the monotonicity of the gradient of a convex function*, Pacific J. Math. **14** (1964), 243-247.

3. J.-J. Moreau, *Fonctionelles convexes*, mimeographed lecture notes, Collège de France, 1967.

4. R. T. Rockafeller, *Characterization of the subdifferentials of convex functions*, Pacific J. Math. **17** (1966), 497-510.

5. ———, *An extension of Fenchel's duality theorem*, Duke Math. J. **33** (1966), 81-90.

6. ———, *Convex analysis*, Princeton University Press, 1969.

Received July 17, 1969. This work was supported in part by the Air Force Office of Scientific Research under grant AF-AFOSR-1202-67.

THE UNIVERSITY OF WASHINGTON, SEATTLE

Printed in Japan by International Academic Printing Co., Ltd., Tokyo, Japan

# Pacific Journal of Mathematics

Vol. 33, No. 1 March, 1970