ACTA MATHEMATICA VIETNAMICA
Volume 18, Number 1, 1993

ON THE MAXIMAL MONOTONICITY OF SUBDIFFERENTIALS

DINH THE LUC

Abstract. We prove that a lower semicontinuous function from a Banach space to the extended real line is convex if and only if its subdifferential is monotone. A simple proof of the maximality of monotone subdifferentials is also given.

1. Introduction

One of the most interesting theorems of convex analysis states that a lower semicontinuous function is convex if and only if its subdifferential (in the sense of convex analysis) is nonvoid and monotone [8], [10]. This theorem has been extended to the case of generalized subdifferentials (in the sense of Clarke) in finite dimensional spaces by Poliquin [9] and in reflexive Banach spaces by Correa et al. [3]. The quadratic conjugate function approach is the main tool used in these works. It relies heavily on the fact that a continuous linear function attains its minimum at any closed convex bounded set, which is of course not true in nonreflexive spaces. Hence their method fails in these spaces. The aim of the present paper is to prove the above discussed result for the case of Banach spaces without the reflexivity assumption. The proof is direct and based on a modified version of Zagrodny's approximate mean value theorem [14]. Moreover, it allows us to validate the result for any kind of subdifferentials which are consistent with the convex analysis one. By this, we establish the equivalence between the monotonicity of subdifferentials and that of directional upper derivatives which were used to characterize convex functions in [5]. A simple proof of Rockafellar's result on the maximal monotonicity of the subdifferential of a proper convex lower semicontinuous function is also given by using our technique.

Received July 14, 1992

2. Preliminaries

Throughout this paper X will denote a Banach space and f a lower semicontinuous function from X to $R \cup \{+\infty\}$. We recall that the generalized subderivative of f at x, where $f(x)$ is finite, with respect to v is defined by

$$f^{\uparrow}(x;v) = \sup_{\epsilon>0} \limsup_{(y,\alpha)\downarrow_f x; t\downarrow 0} \inf_{u\in B(v,\epsilon)} \frac{f(y+tu)-\alpha}{t}$$

and the generalized subdifferential of f at x is

$$\partial f(x) = \{x^* \in X^* : \langle x^*, v\rangle \le f^{\uparrow}(x;v) \text{ for all } v \in X\},$$

where $(y,\alpha)\downarrow_f x$ means that $y \to x, \alpha \to f(x), \alpha \ge f(y), X^*$ is the topological dual of X and $\langle x^*, v\rangle$ is the value of the linear function x^* at v. We say that ∂f is monotone provided for every $x, y \in X, x^* \in \partial f(x), y^* \in \partial f(y)$ one has $\langle x^*, y-x\rangle + \langle y^*, x-y\rangle \le 0$.

In our work we shall make use of the following lemma which can be derived from Zagrodny's approximate mean value theorem [14]. However, for the convenience of the readers a direct proof with less calculation is provided.

LEMMA 2.1. *Assume that $f(b) > f(a)$. There exists a sequence $\{x_k\} \subseteq X$ converging to some $x_0 \in [a,b)$ and $x_k^* \in \partial f(x_k)$ such that for every $c = a+t(b-a)$ with $t \ge 1$ and for every k one has $\langle x_k^*, c-x_k\rangle > 0$.*

PROOF. Assume that $f(b)$ is finite. Following the method of [14], let us consider the functions $g_k, k = 1,2,...$ defined by

$$g_k(x) = f(x) + \frac{f(b)-f(a)}{2||b-a||}||x-b|| + kd_{[a,b]}(x),$$

where $d_{[a,b]}(x)$ denotes the distance from x to the interval $[a,b]$. Sometimes $d_b(x)$ is also used to denote $||x-b||$. Since f is lower semicontinuous, so are the functions g_k. In particular, they are bounded from below on some bounded closed neighborhood (denoted by B) of $[a,b]$. By Ekeland's variational principle [4], for every k there exists a point x_k minimizing the function $g_k + \frac{1}{k}d_{x_k}$ on B. It is clear that $\lim_{k\to\infty} d_{[a,b]}(x_k) = 0$. Hence, one may assume that x_k is in the interior of B and consequently, the subdifferential of $g_k + \frac{1}{k}d_{x_k}$ at x_k must contain the zero vector. Moreover, since the function $d_{[a,b]}, d_b, d_{x_k}$ are convex Lipschitz, using the calculus rules for generalized subdifferentials [2], [11], one has

$$0 \in \partial f(x_k) + \frac{f(b) - f(a)}{2||b-a||}\partial d_b(x_k) + k\partial d_{[a,b]}(x_k) + \frac{1}{k}\partial d_{x_k}(x_k),$$

or in other words, there exist $x_k^* \in \partial f(x_k), u_k^* \in \partial d_b(x_k), v_k^* \in \partial d_{[a,b]}(x_k), w_k^* \in \partial d_{x_k}(x_k)$ such that

$$x_k^* = -\frac{f(b) - f(a)}{2||b-a||}u_k^* - kv_k^* - \frac{1}{k}w_k^*. \tag{1}$$

As noticed above, $\lim_{k\to\infty} d_{[a,b]}(x_k) = 0$. Furthermore, since $[a,b]$ is compact and $g_k(b) > g_k(a)$, we may assume that $\lim_{k\to\infty} x_k = x_0 \in [a,b], x_0 \neq b$. There exists then a number k_1 such that

$$||x_k - x_0|| < \frac{1}{2}||x_0 - b||, \text{ for } k > k_1. \tag{2}$$

The aim now is to estimate $\langle x_k^*, c - x_k\rangle$. Denote by $b_k \in [c, x_k]$ a point minimizing d_b over $[c, x_k]$. It is clear that $\lim_{k\to\infty} b_k = b$. By this and (2), one may assume that $x_k \neq b_k$ and one can obtain the relation $||b - b_k|| \leq d_{[a,b]}(x_k)$. Consequently,

$$d_{[a,b]}(b_k) \leq d_{[a,b]}(x_k).$$

We express $c - x_k = \beta(b_k - x_k)$ for some $\beta \geq 1$ depending on k and calculate

$$\langle u_k^*, c - x_k\rangle = \beta(\langle u_k^*, b_k - b\rangle + \langle u_k^*, b - x_k\rangle). \tag{4}$$

Remembering that $u_k^* \in \partial d_b(x_k)$, we see that $\langle u_k^*, b - x_k\rangle = -||x_k - b||$. Moreover, since $\lim_{k\to\infty} b_k = b$ and $\lim_{k\to\infty} ||x_k - b|| = ||x_0 - b|| \neq 0$, there exists an integer $k_2 > k_1$ such that

$$|\langle u_k^*, b_k - b\rangle| < \frac{1}{4}||x_k - b||, \text{ for } k > k_2.$$

Hence (4) can be evaluated as

$$\langle u_k^*, c - x_k\rangle \leq -\frac{3\beta}{4}||x_k - b||, \text{ for } k > k_2. \tag{5}$$

Furthermore, using (3) one has that

$$\langle v_k^*, c - x_k\rangle = \beta\langle v_k^*, b_k - x_k\rangle \le \beta(d_{[a,b]}(b_k) - d_{[a,b]}(x_k)) \le 0. \tag{6}$$

Notice that $||w_k^*|| \le 1$, there can be found $k_3 > k_2$ such that

$$\frac{1}{k}\langle w_k^*, c - x_k\rangle < \frac{\beta}{8}\frac{f(b) - f(a)}{||b-a||}||x_k - b||, \text{ for } k > k_3.$$

Combine the latter inequality with (1), (5), and (6) to obtain the estimation

$$\langle x_k^*, c - x_k\rangle \ge \frac{\beta}{4}\frac{f(b) - f(a)}{||b-a||}||x_k - b|| > 0,$$

for every $k > k_3$. The sequence $\{x_k\}$ with $k > k_3$ will be such as required in the lemma.

The case $f(b) = \infty$ can be manipulated as follows. Set

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \ne b, \\ f(a) + 1 & \text{otherwise.} \end{cases}$$

Then by the above proof, the lemma is true for $\tilde{f}$. Since $\lim_{k\to\infty} x_k = x_0 \ne b$, we may assume that $x_k \ne b$ for all k. At these points ∂f and $\partial \tilde{f}$ coincide. Hence the lemma is also true for f.

LEMMA 2.2. *Let $a, b, c \in X$ be three distinct points with $b = \lambda a + (1-\lambda)c$ for some $\lambda \in (0,1)$, and $f(b) > \lambda f(a) + (1-\lambda)f(c)$. Then one can find a linear function $x^* \in X^*$ such that*

$$(f + x^*)(b) > (f + x^*)(a) > (f + x^*)(c).$$

PROOF. Let us denote

$$\alpha = \begin{cases} f(b) - \lambda f(a) - (1-\lambda)f(c) & \text{if } f(b) \text{ is finite,} \\ 1 & \text{otherwise.} \end{cases}$$

Construct a linear function x^* on the one-dimensional space $L = \{t(a-c) : t \in R\}$ as follows: $x^*(t(a-c)) = (\alpha + f(c) - f(a))t$ for every $t \in R$. In view of the Hahn-Banach theorem one can extend x^* to a continuous linear function on the whole space X which we denote by the same x^*. Let us calculate the values of $f + x^*$ at the considered points:

$$
\begin{aligned}
(f + x^*)(a) &= f(a) + x^*(c) + x^*(a - c) = f(c) + x^*(c) + \alpha, \\
(f + x^*)(c) &= f(c) + x^*(c), \\
(f + x^*)(b) &= f(b) + x^*(c) + \lambda x^*(a - c) \\
&= f(c) + x^*(c) + f(b) - \lambda f(a) - (1 - \lambda) f(c) + \lambda\alpha \\
&= \begin{cases} f(c) + x^*(c) + (1 + \lambda)\alpha & \text{if } f(b) \text{ is finite,} \\ \infty & \text{otherwise.} \end{cases}
\end{aligned}
$$

Since α is positive, the above equalities imply the required relations.

3. Maximal monotone subdifferential

We are now able to formulate and prove the main result of the paper.

THEOREM 3.1. *The function f is convex if and only if its subdifferential is monotone.*

PROOF. The "only if" part of the theorem is well known. For the "if" part, suppose to the contrary that the function is not convex, i.e. there exist three distinct points $a, b, c \in X$ with $b = \lambda a + (1 - \lambda)c$ for some $\lambda \in (0, 1)$ such that $f(b) > \lambda f(a) + (1 - \lambda) f(c)$. This in particular implies that $f(a)$ and $f(c)$ are finite. Our aim is to find two points $\bar{x}, \bar{y} \in X$ and $x^* \in \partial f(\bar{x}), y^* \in \partial f(\bar{y})$ such that $\langle x^*, \bar{y} - \bar{x}\rangle + \langle y^*, \bar{x} - \bar{y}\rangle \geq 0$ which show that ∂f is not monotone. We may assume without loss of generality that f possesses the following properties:
(i) $f(b) > f(a) > f(c)$,
(ii) $f(y) > f(c)$ for every $y \in [a, b]$,
(iii) $f(b)$ is finite.

In fact, the first property follows from Lemma 2.2 and the calculus rule for generalized subdifferentials: $\partial(f + l)(x) = \partial f(x) + l$, whenever l is a linear continuous function on X. As to the second property, observe that the set $\{y \in [a, b] : f(y) \leq f(a)\}$ is compact and does not contain b. Let a' be the point of this set which is the closest to b. Then $a' \neq b$ and $f(y) \geq f(a) > f(c)$

for every $y \in (a', b]$. It remains to take a' in the role of a to have the wanted situation (in the case $f(a') \le f(c)$, another application of Lemma 2.2 is needed). For the last property, if $f(b) = \infty$, one defines a new function $\tilde{f}$ by the rule

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \neq b, \\ f(a) + 1 & \text{otherwise.} \end{cases}$$

Then at any point $x \neq b$, (the points $\bar{x}$ and $\bar{y}$ to be found are such ones) $\partial f(x)$ and $\partial \tilde{f}(x)$ coincide. Moreover, $\tilde{f}$ satisfies (i)-(iii) if f satisfies (i) and (ii). With these properties of f let us apply Lemma 2.1 first to $[a, b]$. There can be found $\{x_k\} \subseteq X$ with $\lim_{k\to\infty} x_k = x_0 \in [a, b)$ and $x_k^* \in \partial f(x_k)$ such that

$$\langle x_k^*, c - x_k \rangle > 0. \tag{7}$$

By (ii), $f(x_k) > f(c)$ whenever k is large enough. Fix such a point x_k and apply Lemma 2.1 to $[c, x_k]$. There exists then $\{y_n\} \subseteq X$ with $\lim_{n\to\infty} y_n = y_0 \in [c, x_k)$ and $y_n^* \in \partial f(y_n)$ such that $\langle y_n^*, x_k - y_n \rangle > 0$. It follows from (7) that for n sufficiently large, $\langle x_k^*, y_n - x_k \rangle > 0$. The two latter inequalities show that ∂f is not monotone and the proof is complete.

Let us now apply Lemma 2.1 to establish the maximal monotonicity of ∂f when f is convex. To our knowledge, there exist at least four different proofs of this fact [1], [10], [12], [13]. The proof given below seems to be the simplest one.

THEOREM 3.2. *Let f be proper convex. Then ∂f is maximal monotone.*

PROOF. The monotonicity is clear. For the maximality, we have to show that if $x^* \notin \partial f(x)$ for some $x \in X$, there exists $y \in X$ and $y^* \in \partial f(y)$ such that $\langle y^* - x^*, x - y \rangle > 0$. As in [12], one observes that x does not minimize $f - x^*$, hence there is a point $z \in X$ with $(f - x^*)(x) > (f - x^*)(z)$. In view of Lemma 2.1, there can be found a point y in a neighborhood of the interval $[z, x]$ and a vector $y'^* \in \partial(f - x^*)(y)$ such that $\langle y'^*, x - y \rangle > 0$. A vector $y^* \in \partial f(y)$ with $y'^* = y^* - x^*$ will fulfil our requirement.

4. A final remark

We recall that the directional upper derivative of f at x, where $f(x)$ is finite, in the direction v is the map $f'(x;v) : X \to R \cup \{\pm\infty\}$ defined as follows:

$$f'(x;v) = \limsup_{t \downarrow 0} \frac{f(x+tv) - f(x)}{t}.$$

We say that $f'(x;v)$ is monotone if for every $x, y \in X, f'(x; y-x) + f'(y; x-y) \leq 0$, whenever the sum has meaning. In [5] it was shown that a lower semicontinuous function from a real topological vector space X to $R \cup \{+\infty\}$ is convex if and only if its directional upper derivative is monotone. This fact and Theorem 3.1 show that in Banach spaces the monotonicity of $\partial f(x)$ and that of $f'(x;v)$ are equivalent.

It is also interesting to note that Lemma 2.1 is valid for any subdifferential $\bar{\partial}$ which possesses the following properties: (I) it coincides with the convex analysis subdifferential if the function is convex; (II) the subdifferential at a local minimum of a function must contain zero; (III) if g is convex Lipschitz, then $\bar{\partial}(f+g) \subseteq \bar{\partial}f + \bar{\partial}g$.

Consequently, we may state that Theorem 3.1 is true for any subdifferential with the three above listed properties. Mordukhovich's subdifferential, Michel and Penot's subdifferential [7] for instance are examples of such subdifferentials which are different from the Clarke one.

ACKNOWLEDGEMENTS. The author would like to thank Professor J.-P. Penot of the University of Pau for useful comments on the first version [6] of the paper.

REFERENCES

[1] J.B. Borwein, *A note on ε-subgradients and maximal monotonicity*, Pacific Journal of Mathematics **103** (1982), 307–314.

[2] F.H. Clarke, *Optimization and nonsmooth analysis*, Wiley, New York, 1983.

[3] R. Correa, A. Jofre and L. Thibault, *Characterization of lower semicontinuous convex functions*, Proceedings of the American Mathematical Society **116** (1992), 67-72.

[4] I. Ekeland, *On the variational principle*, Journal of Mathematical Analysis and Application **47** (1974), 324–353.

[5] D.T. Luc and S. Swaminathan, *A characterization of convex functions*, Nonlinear Analysis, Theory, Methods and Applications **20** (1993), 697-701 .

[6] D.T. Luc, *On the monotonicity of subgradients*, Manuscript, University of Erlangen-Nurnberg, 1991.

[7] P. Michel and J.-P. Penot, *Calcul sous différentiel pour des fonctions lipschitziennes et non lischiziennes*, C. R. Acad. Sc. Paris **298** (1984), 269–272.

[8] J.J. Moreau, *Fonctionelles convexes*, Lecture Notes, Seminaire Equations aux derivees partielles, College de France, 1966.

[9] R.A. Poliquin, *Subgradient monotonicity and convex functions*, Nonlinear Analysis, Theory, Methods and Applications **14** (1990), 305–317.

[10] R.T. Rockafellar, *On the maximal monotonicity of subdifferential mappings*, Pacific Journal of Mathematics **33** (1970), 209–216.

[11] R.T. Rockafellar, *Generalized derivatives and subgradients of nonconvex functions*, Canadian Journal of Mathematics **XXXII** (1980), 257–280.

[12] S. Simons, *The least slope of a convex function and the maximal monotonicity of its subdifferential*, Journal of Optimization Theorey and Application **71** (1991), 127–136.

[13] P.D. Taylor, *Subgradients of a convex function obtained from a directional derivative*, Pacific Journal of Mathematics **44** (1973), 739–747.

[14] D. Zagrodny, *Approximate mean value theorem for upper subderivatives*, Nonlinear Analysis, Theory, Methods and Applications, **12** (1988), 1413–1428.

INSTITUTE OF MATHEMATICS
P.O. BOX 631, BO HO, 10000 HANOI, VIETNAM