PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 121, Number 4, August 1994

A NOTE ON THE DIFFERENTIABILITY OF CONVEX FUNCTIONS

WU CONGXIN AND CHENG LIXIN

(Communicated by Palle E. T. Jorgensen)

ABSTRACT. Every real-valued convex and locally Lipschitzian function f defined on a nonempty closed convex set D of a Banach space E is the local restriction of a convex Lipschitzian function defined on E. Moreover, if E is separable and $\operatorname{int} D \neq \varnothing$, then, for each Gateaux differentiability point x ($\in \operatorname{int} D$) of f, there is a closed convex set $C \subset \operatorname{int} D$ with the nonsupport points set $N(C) \neq \varnothing$ and with $x \in N(C)$ such that f_C (the restriction of f on C) is Fréchet differentiable at x.

Rainwater [1] and Verona [2] recently generalized Asplund's theorems [5] in the following manner

Theorem [Rainwater]. *Suppose that E is an Asplund space (a Banach space of class (S)). Then, for every closed convex subset C of E such that the nonsupport points set $N(C)$ of C is nonempty and for every convex function f on C which is locally Lipschitzian on $N(C)$, the set of points Q, where f is Fréchet (Gateaux) differentiable, is a dense G_δ subset of C.*

Notation. We denote by C, C_x, and $N(C)$ a closed convex set of a Banach space E, the cone generated by C from x, and the nonsupport points set of C, respectively. The open and the closed balls centered at x and with radius r are denoted by $B(x, r)$ and $\overline{B}(x, r)$, respectively. All convex functions f, if nothing is added, are assumed real valued.

Definition. We say that a function denoted by f_u is a local extension of f at $u \in \operatorname{dom}(f)$ and that f is a local restriction of f_u at u provided there is a neighborhood U of u such that $f_u = f$ in $U \cap \operatorname{dom}(f)$.

Definition. We say that f is Gateaux (Fréchet) differentiable at $x \in N(C)$ provided there exists a unique $x^* \in E^*$ such that

$$\langle x^*, y - x\rangle \leq f(y) - f(x) \quad \text{for all } y \in C$$

(for all $\varepsilon > 0$ there exists $\delta > 0$ such that

$$0 \leq f(y) - f(x) - \langle x^*, y - x\rangle \leq \varepsilon\|y - x\|$$

whenever $y \in C \cap B(x, \delta)$).

Received by the editors October 28, 1992.
1991 *Mathematics Subject Classification*. Primary 26E15, 46G05.

Lemma 1. *Suppose that f is convex and locally Lipschitzian on C. Then*

(a) *for every $u \in C$, there exist a neighborhood U of u and a convex Lipschitzian function f_u defined on E such that $f = f_u$ in $U \cap C$;*

(b) $\operatorname{dom}(\partial f) = \operatorname{dom}(f) = C$;

(c) *there exists a selection ψ for ∂f on C such that ψ is locally bounded on C; in particular* [1], *if $N(C) \neq \varnothing$, then ∂f is itself locally bounded on $N(C)$.*

Recall that the extended real-valued lower semicontinuous convex function f on E is said to be proper provided $f(x) > -\infty$ for all $x \in E$ and its (convex) essential domain

$$\operatorname{dom} f = \{x : f(x) < +\infty\}$$

is nonempty. For such f, we define the inf-*convolutions as*

$$f_n(x) = \inf\{f(y) + n\|x - y\| : y \in E\}.$$

Then we have (see, for instance, [3, 4, 6]).

Lemma 2. *With f and $\{f_n\}$ as above, the sequence $\{f_n\}$ has the following properties:*

(1) *Each f_n is convex and Lipschitzian on E with Lipschitz constant n;*

(2) *$f_n(x) \leq f_{n+1}(x) \leq f(x)$ for each $x \in E$ and each $n \geq 1$;*

(3) *$f_n(x) = f(x)$ if and only if $\partial f(x) \cap nB^*$ (where B^* denotes the unit ball of E^*) is nonempty or, equivalently, if and only if $\partial f_n(x) = \partial f(x) \cap nB^*$.*

Proof of Lemma 1. Since f is locally Lipschitzian on C, there exists $B(u, r)$ for some $r > 0$ and $L > 0$ such that $|f(y) - f(x)| \leq L\|y - x\|$ whenever $x, y \in \overline{B}(u, r) \cap C \equiv \overline{B}_c(u, r)$. We define the extended real-valued convex function $\hat{f}$ on E by $\hat{f}(x) = f(x)$ if $x \in \overline{B}_c(u, r)$, and $\hat{f}(x) = \infty$, otherwise. Hence f is proper lower semicontinuous convex and bounded below on E and Lipschitzian on $\overline{B}_c(u, r)$ with Lipschitz constant L. Let $\{f_n\}$ be the sequence of the inf-convolutions by $\hat{f}$. Now we will show (a). Suppose, to the contrary, that the neighborhood U of u does not exist. Let $B_n = B_c(u, 1/n)$. Then, for each integer $n \geq 1$, there is $x_n \in B_n$ such that $f_n(x_n) < \hat{f}(x_n)$, by the definition of f_n. For each such x_n, we can choose $y_n \in E$ such that

$$\hat{f}(x_n) > \hat{f}(y_n) + n\|y_n - x_n\|, \qquad n = 1, 2, \ldots,$$

that is, $\hat{f}(x_n) - \hat{f}(y_n) > n\|y_n - x_n\|$ for $n = 1, 2, \ldots$. Clearly, we have $y_n \in \overline{B}_c(u, r)$. Note that $x_n \in B_c(u, r)$ whenever $1/n \leq r$ and $\hat{f}$ is Lipschitzian on $\overline{B}_c(u, r)$ with Lipschitz constant L. We have $|\hat{f}(x_n) - \hat{f}(y_n)| \leq L\|x_n - y_n\|$ whenever $1/n < r$; this is a contradiction which proved assertion (a).

From the proof of assertion (a), it is easy to see that $f_n = \hat{f}$ in $\overline{B}_c(u, r)$ by taking $n = [L] + 1$ (where $[L]$ denotes the maximal integer m satisfying $m \leq L$); that is, by Lemma 2, $\partial \hat{f}(x) \cap nB^*$ is nonempty for each $x \in \overline{B}_c(u, r)$. Hence $\partial f(x) \cap nB^*$ is nonempty for each $x \in B(u, r)$. Now we showed that there is a selection for ∂f on $B(u, r)$ which is bounded by n. The arbitrariness of u says that there is a selection ψ for ∂f on C which is locally bounded; hence we proved assertions (b) and (c).

With f and f_u as in Lemma 1, by Lemma 2, we have $\partial f_u(x) \subset \partial f(x)$ for each $x \in U \cap C$. If C is closed and $N(C) \neq \varnothing$, note that $x \in N(C)$ if

and only if C_x is dense in E [1]. By a simply convexity and differentiability argument, we see $\partial f = \partial f_u$ in $U \cap N(C)$. Hence we have

Proposition 3. *Suppose that C is closed with $N(C) \neq \varnothing$ and f is locally Lipschitzian on $N(C)$. Then the following versions are equivalent:*

(a) *f is Gateaux differentiable at $u \in N(C)$.*
(b) *Every f_u is Gateaux differentiable at $u \in N(C)$.*
(c) *Every selection for ∂f on $N(C)$ is norm-to-weak* continuous at u.*
(d) *There is a selection for ∂f on $N(C)$ which is norm-to-weak* continuous at u.*

Corollary 4. *Suppose that C is a closed convex set of the weak Asplund space E, and suppose that the convex function f is locally Lipschitzian on $N(C)$. Then the set of points G, where f is Gateaux differentiable, is a G_δ set of C.*

It is easy to see that if $u \in N(C)$ is the Fréchet differentiability point of f_u, then so is the one of f. The following example shows that if the interior of C is empty, then the Fréchet differentiability of f_u is really stronger than the one of f.

Example 5. The function f defined by $f(x) = \|x\|_{l^1}$ on $C = \{x \in l^1,\ x_n \geq 0$ for $n = 1, 2, \dots\}$ shows that f is Fréchet differentiable at each point of $N(C) = \{x \in l^1,\ x_n > 0$ for $n = 1, 2, \dots\}$, but the extension f_u of f (where $f_u = \|\cdot\|_{l^1}$ on l^1) is nowhere Fréchet differentiable.

More generally, we have

Theorem 6. *Suppose that E is separable, D is a nonempty open convex set of E, and f is a continuous convex function on D. Then, for each Gateaux differentiability point $x \in D$ of f, there exists a closed convex subset C of D with $N(C) \neq \varnothing$ and $x \in N(C)$ such that f_C (the restriction of f to C) is Fréchet differentiable at x.*

Proof. Let $0 < \varepsilon < \frac{1}{2}$. By [8] there are two sequences $\{x_j\}$ and $\{x_j^*\}$ in E and in E^*, respectively, satisfying

(1) $\|x_j\| = 1$ for $j = 1, 2, \dots$;
(2) $\|x_j^*\| < 1 + \varepsilon$ for $j = 1, 2, \dots$;
(3) $x_i^*(x_j) = \delta_{ij}$ $(= 1$ if $i = j$, $= 0$ otherwise);
(4) $\mathrm{cl}(\mathrm{span}(x_j$ for $j = 1, 2, \dots)) = E$.

Suppose that $x \in D$ is a Gateaux differentiability point of f. Then, for each $n \geq 1$, there exists $1 \geq \delta_n > 0$ $(\delta_n \to 0)$ such that

$$\frac{f(x \pm 2\delta_n x_n) - f(x) - \langle x^*, \pm 2\delta_n x_n\rangle}{2\delta_n} < 2^{-n} \qquad (x^* = \partial f(x)). \tag{1}$$

Let $E_n = \mathrm{span}(x_j$ for $j = 1, 2, \dots, n)$, $F = \overline{co}(\pm y_n$ for $n = 1, 2, \dots)$, and $C = x + F$, where $y_n = \delta_n x_n$ for $n = 1, 2, \dots$. Clearly, C $(\subset D)$ is closed and convex and C_x is dense in E, and hence $x \in N(C)$. Now we prove that point x is a Fréchet differentiability point of f_c. Suppose, to the contrary, that f_c is not Fréchet differentiable at x. Then there is $\varepsilon_0 > 0$ such that for each

δ_n there is $z_n \in C$ with $0 < \|z_n - x\| < \delta_n$ such that

$$\frac{f_c(z_n) - f_c(x) - \langle x^*, z_n - x\rangle}{\|z_n - x\|} \geq \varepsilon_0 \qquad (n = 1, 2, \dots),$$

where $x^* = \partial f(x)$. By the density of $co(\pm y_n$ for $n = 1, 2, \dots)$ in F, we can assume that $\{z_n - x\} \subset co(\pm y_n$ for $n = 1, 2, \dots)$, that is, for each $n \geq 1$, there is a sequence $\{\lambda_i^{(n)}\}$ with $\lambda_i^{(n)} \geq 0$ for $i = 1, 2, \dots$ and $\sum_i \lambda_i^{(n)} = 1$ such that $z_n - x = \sum_i \lambda_i^{(n)}(\pm y_i)$. Therefore, we have, by the properties of $\{x_n\}$ and $\{x_n^*\}$,

$$\begin{aligned} \|z_n - x\| &> \frac{1}{1+\varepsilon} \max_j |x_j^*(z_n - x)| \\ &= \frac{1}{1+\varepsilon} \max_j \lambda_j^{(n)} \|y_j\| = \frac{1}{1+\varepsilon} \max_j \lambda_j^{(n)} \delta_j . \end{aligned} \tag{2}$$

Fix n_0 such that $(1+\varepsilon)\sum_{j=n_0+1}^{\infty} 2^{-j} = (1+\varepsilon)2^{-n_0} < \varepsilon_0/2$. Let

$$u_n = x + 2\sum_{i=1}^{n_0} \lambda_i^{(n)}(\pm y_i) \quad \text{and} \quad v_n = x + 2\sum_{i=n_0+1}^{\infty} \lambda_i^{(n)}(\pm y_i).$$

Since $z_n \to x$, we can claim $u_n, v_n \in D$ for $n = 1, 2, \dots$. Now we have $z_n = (u_n + v_n)/2$, therefore,

$$f_c(z_n) = f(z_n) = f((u_n + v_n)/2) \leq \tfrac{1}{2}[f(u_n) + f(v_n)].$$

Let $w_n = (u_n - x)/\theta_n$ (where $\theta_n = \|z_n - x\|$). Then

$$\begin{aligned} \|w_n\| = \|u_n - x\|/\theta_n &\leq \sum_{i=1}^{n_0} 2\lambda_i^{(n)} \|y_i\|/\theta_n = 2\sum_{i=1}^{n_0} \lambda_i^{(n)} \delta_i/\theta_n \\ &\leq (1+\varepsilon)\cdot 2\sum_{i=1}^{n_0} \lambda_i^{(n)} \delta_i / \max_j \lambda_j^{(n)} \delta_j \leq 2(1+\varepsilon) n_0 ; \end{aligned} \tag{3}$$

that is, $\{w_n\}$ is contained by $2(1+\varepsilon)n_0 B_{n_0}$, where B_{n_0} is the unit ball of the n_0-dimensional subspace E_{n_0} of E.

$$\begin{aligned} f(v_n) &= f\left(x + 2\sum_{i=n_0+1}^{\infty} \lambda_i^{(n)}(\pm y_i)\right) = f\left(\sum_{i=n_0+1}^{\infty} \lambda_i^{(n)}(x \pm 2y_i) + \sum_{i=1}^{n_0} \lambda_i^{(n)} x\right) \\ &\leq \sum_{i=n_0+1}^{\infty} \lambda_i^{(n)} f(x \pm 2y_i) + \sum_{i=1}^{n_0} \lambda_i^{(n)} \cdot f(x). \end{aligned}$$

Hence

$$
\begin{aligned}
\varepsilon_0 &\le \frac{f(z_n) - f(x) - \langle x^*, z_n - x\rangle}{\theta_n} \\
&\le \frac{\{[f(u_n) + f(v_n)]/2 - f(x) - \langle x^*, z_n - x\rangle\}}{\theta_n} \\
&= \frac{f(u_n) - f(x) - \langle x^*, u_n - x\rangle}{2\theta_n} + \frac{f(v_n) - f(x) - \langle x^*, v_n - x\rangle}{2\theta_n} \\
&\le \frac{f(u_n) - f(x) - \langle x^*, u_n - x\rangle}{2\theta_n} \\
&\qquad + \frac{\sum_{i=n_0+1}^{\infty} \lambda_i^{(n)}(f(x \pm 2y_i) - f(x) - \langle x^*, \pm 2y_i\rangle)}{2\theta_n} \\
&\overset{(2)}{\le} \frac{f(u_n) - f(x) - \langle x^*, u_n - x\rangle}{2\theta_n} \\
&\qquad + \frac{(1+\varepsilon)\sum_{i=n_0+1}^{\infty} \lambda_i^{(n)}(f(x \pm 2y_i) - f(x) - \langle x^*, \pm 2y_i\rangle)}{2\max_j \lambda_j^{(n)}\delta_j} \\
&\overset{(1)}{\le} \frac{f(u_n) - f(x) - \langle x^*, u_n - x\rangle}{2\theta_n} + (1+\varepsilon)\sum_{j=n_0+1}^{\infty} 2^{-j} \\
&< \frac{f(u_n) - f(x) - \langle x^*, u_n - x\rangle}{2\theta_n} + \frac{\varepsilon_0}{2} \\
&= \frac{f(x + \theta_n w_n) - f(x) - \langle x^*, \theta_n w_n\rangle}{2\theta_n} + \frac{\varepsilon_0}{2}.
\end{aligned}
$$

Since $2(1+\varepsilon)n_0 B_{n_0}$ is a bounded set of the n_0-dimensional subspace E_{n_0} of E, $w_n \in 2(1+\varepsilon)n_0 B_{n_0}$, and $\theta_n \to 0$, and since f is Gateaux differentiable at x, we must have

$$
\frac{f(x + \theta_n w_n) - f(x) - \langle x^*, \theta_n w_n\rangle}{\theta_n} < \varepsilon_0
$$

for sufficiently large n. Hence

$$
\varepsilon_0 \le \frac{f(x + \theta_n w_n) - f(x) - \langle x^*, \theta_n w_n\rangle}{2\theta_n} + \frac{\varepsilon_0}{2} < \frac{\varepsilon_0}{2} + \frac{\varepsilon_0}{2} = \varepsilon_0
$$

for sufficiently large n, and this is a contradiction which completes our proof.

Remark. The recent work of Preiss, Phelps, and Namioka [7] showed that if E admits an equivalent smooth norm, then it is of class (S). Hence, for the Gateaux differentiability, the Rainwater theorem still holds for smoothable Banach spaces.

Acknowledgment

The second author specially thanks Professors R. R. Phelps and I. Namioka for their hospitality and helpful conversations on this note.

References

1. J. Rainwater, *Yet more on the differentiability of convex functions*, Proc. Amer. Math. Soc. **103** (1988), 773–777.
2. M. E. Verona, *More on the differentiability of convex functions*, Proc. Amer. Math. Soc. **103** (1988), 137–140.
3. J. B. Hiriart Urruty, *Lipschitz r-continuity of the approximate subdifferential of a convex function*, Math. Scand. **47** (1980), 123–134.
4. P. J. Laurent, *Approximation et optimisation*, Hermann, Paris, 1972.
5. E. Asplund, *Fréchet differentiability of convex functions*, Acta. Math. **121** (1968), 31–47.
6. S. Fitzpatrick and R. R. Phelps, *Bounded approximants to monotone operators on Banach spaces* (to appear).
7. D. Preiss, R. R. Phelps, and I. Namioka, *Smooth Banach spaces, weak Asplund spaces and monotone or USCO mappings* (to appear).
8. J. Lindenstrauss and L. Tzafiri, *Classical Banach spaces.* I, Springer-Verlag, New York, 1977.

Department of Mathematics, Harbin Institute of Technology, Harbin, 150001, People's Republic of China