A NEW PROOF OF THE INTEGRABILITY
OF THE SUBDIFFERENTIAL
OF A CONVEX FUNCTION ON A BANACH SPACE

MILEN IVANOV AND NADIA ZLATEVA

(Communicated by Jonathan M. Borwein)

Abstract. We provide a simple proof of the Moreau-Rockafellar theorem that
a proper lower semicontinuous convex function on a Banach space is deter-
mined up to a constant by its subdifferential.

1. Introduction

Recall the following famous theorem of Rockafellar; see [8, 9]:

Theorem 1.1 (Rockafellar). Let $X$ be a Banach space. Let $f$ and $g$ be proper
lower semicontinuous convex functions from $X$ to $\mathbb{R} \cup \{+\infty\}$. If

\begin{equation}
\partial f \subset \partial g,
\end{equation}

then

$$g = f + \text{const.}$$

The reader is referred to Section 3 of Phelps book [7] for a possible context of
this theorem and some of its important implications.

The purpose of this note is to present a new and more direct proof of Theorem 1.1.

First of all, however, we present a brief historical overview. The integrability
of the subdifferential of proper lower semicontinuous convex functions on Hilbert
space is stated and proved by Moreau in [5]. His proof uses Moreau-Yosida regulari-
sation and works also in reflexive Banach space as mentioned on p. 87 of [6]. While
the first complete proof of Theorem 1.1 – that of Rockafellar in [9] – uses duality
arguments, most of the others rely on approximating the directional derivative
and subsequent reduction to the trivial one-dimensional case. This second line is
followed by the original Rockafellar proof in [8]. Although there are some gaps in
the latter proof, the work of Taylor [10] contains both a remedy and a different
proof (cf. [2]). The idea of directional derivative approximation/one-dimensional
reduction is most clearly outlined in the proof of Thibault [11]. A different proof
using the mean-value theorem of Zagrodny is due to Thibault and Zagrodny [12]
(see also [13]).

Received by the editors January 8, 2007.
2000 Mathematics Subject Classification. Primary 52A41, 49J53; Secondary 26E15, 47H05.
Key words and phrases. Convex function, subdifferential.

The first author was supported in part by the Research and Development Fund of Sofia Uni-

©2008 American Mathematical Society
Reverts to public domain 28 years from publication

1787
The proof we are about to present follows a different route. It is similar to the proof of the classical calculus theorem that a monotone function is Riemann integrable. It uses neither duality nor explicit one-dimensional arguments.

The main step in our proof is to show directly that a proper lower semicontinuous convex function on a Banach space differs by a constant from the Rockafellar function (see [1]) of its subdifferential. Namely, we prove the following

**Theorem 1.2** (Rockafellar [8, 9]). Let \( f : X \to \mathbb{R} \cup \{+\infty\} \) be a proper lower semicontinuous convex function on the Banach space \( X \). Let \( x_0 \in \text{dom} \partial f \) and \( p_0 \in \partial f(x_0) \) be arbitrary. Then for all \( x \in X \),

\[
(1.2) \quad f(x) = f(x_0) + R_{\partial f,(x_0,p_0)}(x),
\]

where

\[
(1.3) \quad R_{\partial f,(x_0,p_0)}(x) = \sup \left\{ \sum_{i=0}^{n-1} p_i(x_{i+1} - x_i) : x_n = x, p_i \in \partial f(x_i), n \in \mathbb{N} \right\};
\]

that is, the supremum is taken over all finite sequences in \( \text{gph} \partial f \), starting at \((x_0,p_0)\).

With the aid of the novel technical Lemma 3.3 we derive first the case \( x \in \text{dom} \partial f \) of (1.2). The general case is completed by reference to the following well-known result of Brøndsted and Rockafellar [3]:

**Proposition 1.3.** Let \( f \) be a proper lower semicontinuous convex function from a Banach space \( X \) into \( \mathbb{R} \cup \{+\infty\} \). Then for all \( x \in X \),

\[
(1.4) \quad f(x) = \sup \{ f(\bar{x}) + \bar{p}(x - \bar{x}) : (\bar{x}, \bar{p}) \in \text{gph} \partial f \}.
\]

From Theorem 1.2 it readily follows that (1.1) implies

\[
f(x) - f(x_0) \leq g(x) - g(x_0)
\]

for any \( x_0 \in \text{dom} \partial f \) and all \( x \in X \). In particular, \( f - g \equiv \text{const} \) on \( \text{dom} \partial f \).

The proof of Theorem 1.1 is completed by a reference to lower semicontinuity and graphical density of points of subdifferentiability (see [3] and [2]):

**Proposition 1.4.** Let \( f \) be a proper lower semicontinuous convex function from a Banach space \( X \) into \( \mathbb{R} \cup \{+\infty\} \). Then for any \( \bar{x} \in \text{dom} f \) and any \( \varepsilon > 0 \) there exists \( x \in \text{dom} \partial f \) such that \( \|x - \bar{x}\| + |f(x) - f(\bar{x})| < \varepsilon \).

The remaining content is organized as follows. After a short section on notation, we state and prove Lemma 3.3 in Section 3. In Section 4 the proofs of Theorem 1.2 and Theorem 1.1 are assembled.

### 2. Notation

The notation used throughout the paper is fairly standard, so we recall it briefly.

Usually \( X \) denotes a real Banach space, that is, a complete normed space over \( \mathbb{R} \). The norm of \( X \) is denoted by \( \| \cdot \| \). The closed unit ball \( \{ x \in X : \|x\| \leq 1 \} \) of \( X \) is denoted by \( B_X \). The dual space \( X^* \) of \( X \) is the Banach space of all continuous linear functionals \( p \) from \( X \) to \( \mathbb{R} \). The natural norm of \( X^* \) is again denoted by \( \| \cdot \| \). The value of \( p \in X^* \) at \( x \in X \) is denoted by both \( p(x) \) and \( \langle p, x \rangle \), wherever convenient.
The effective domain $\text{dom } f$ of an extended real-valued function $f : X \to \mathbb{R} \cup \{+\infty\}$ is the set of points $x$ where $f(x) \in \mathbb{R}$. The function $f$ is proper if $\text{dom } f \neq \emptyset$. It is lower semicontinuous if
\[
 f(\bar{x}) \leq \liminf_{x \to \bar{x}} f(x)
\]
for all $\bar{x} \in X$. The epigraph of $f$ is the set $\text{epi } f \subset X \times \mathbb{R}$ defined by
\[
 \text{epi } f = \{(x, r) : x \in \text{dom } f, \ r \geq f(x)\}.
\]
It is clear that $\text{epi } f$ is nonempty whenever $f$ is proper, and $\text{epi } f$ is closed iff $f$ is lower semicontinuous.

The indicator function of a set $C \subset X$ is the function $I_C : X \to \mathbb{R} \cup \{+\infty\}$ defined by:
\[
 I_C(x) = 0, \text{ if } x \in C; \quad I_C(x) = \infty, \text{ otherwise.}
\]
Obviously, $I_C$ is proper iff $C$ is nonempty and $I_C$ is lower semicontinuous iff $C$ is closed.

If $f : X \to \mathbb{R} \cup \{+\infty\}$ is a convex function, then $\partial f(\bar{x})$ denotes the subdifferential of $f$ at $\bar{x} \in \text{dom } f$ in the sense of Convex Analysis, that is,
\[
 \partial f(\bar{x}) = \{p \in X^* : \langle p, x - \bar{x} \rangle \leq f(x) - f(\bar{x}), \ \forall x \in X\},
\]
and $\partial f = \emptyset$ on $X \setminus \text{dom } f$. It is a well-known and easy consequence of convexity that if $\langle p, x - \bar{x} \rangle \leq f(x) - f(\bar{x})$ holds on some neighbourhood of $\bar{x}$, rather than on the whole space, then again $p \in \partial f(\bar{x})$.

The domain $\text{dom } \partial f$ of the multi-valued map $\partial f : X \to X^*$ consists of all points $x \in X$ such that $\partial f(x)$ is nonempty. The graph of $\partial f$ is
\[
 \text{gph } \partial f = \{(x, p) : x \in \text{dom } \partial f, \ p \in \partial f(x)\}.
\]

3. Variational arguments

This section is devoted to Lemma 3.3, which can be viewed as a particular enhancement of Proposition 1.4.

Although the technique of proof is standard, we provide full details for the reader’s convenience. First, however, we recall the famous Ekeland Variational Principle (e.g. [7], p.45):

**Theorem 3.1** (Ekeland Variational Principle). Let $f$ be a proper lower semicontinuous function from a Banach space $X$ into $\mathbb{R} \cup \{+\infty\}$. Let $f$ be bounded below, $\varepsilon > 0$, and $\bar{y} \in \text{dom } f$ be such that $f(\bar{y}) \leq \inf f + \varepsilon$. Then for each $\lambda > 0$ there is $\bar{x} \in \text{dom } f$ such that:

(i) $\lambda \|x - \bar{y}\| \leq f(\bar{y}) - f(\bar{x})$,
(ii) $\|x - \bar{y}\| \leq \varepsilon / \lambda$, and
(iii) $\lambda \|x - \bar{x}\| + f(x) > f(\bar{x})$ for all $x \neq \bar{x}$.

We also use the following well-known sum theorem (e.g. [7], p.47):

**Theorem 3.2** (Moreau-Rockafellar Formula). Let $f$ be a proper lower semicontinuous convex function from a Banach space $X$ into $\mathbb{R} \cup \{+\infty\}$, and let $g : X \to \mathbb{R}$ be convex and continuous. Then for any $x \in X$,
\[
 \partial (f + g)(x) = \partial f(x) + \partial g(x).
\]
(The right-hand side is the usual sum of sets.)
Lemma 3.3. Let \( f \) be a proper lower semicontinuous convex function from a Banach space \( X \) into \( \mathbb{R} \cup \{+\infty\} \). Let \((y_i)_{i=1}^k \subset \text{dom} \ f\). Then for each \( \varepsilon > 0 \) there are \((x_i, p_i) \in \text{gph} \partial f\) such that

\[
\|x_i - y_i\| \leq \varepsilon \text{ and } \|p_j\| \|x_i - y_i\| \leq \varepsilon, \quad \forall i, j = 1, \ldots, k.
\]

Proof. Due to the lower semicontinuity of \( f \) there is \( \delta \in (0, \varepsilon) \) such that for all \( i = 1, \ldots, k \),

\[
f(y_i + \delta B_X) \geq f(y_i) - \varepsilon.
\]

Let \( f_i = f + I_{\{y_i + \delta B_X\}} \). So, \( y_i \) is an \( \varepsilon \)-minimum of the proper lower semicontinuous \( f_i \). The Ekeland Variational Principle applied to \( f_i \) (with this \( \varepsilon \) and \( \lambda = 2\varepsilon / \delta \)) shows that there is \( x_i \in \text{dom} f_i \cap \{y_i + (\varepsilon / \lambda)B_X\} \) such that

\[
f_i(x) + \lambda \|x - x_i\| \geq f_i(x_i), \quad \forall x \in X.
\]

Since \( x_i \in y_i + (\varepsilon / \lambda)B_X \) and \( \varepsilon / \lambda = \delta / 2 < \varepsilon \), we have that

\[
\|x_i - y_i\| \leq \varepsilon / \lambda = \delta / 2 < \varepsilon.
\]

On the other hand, (3.2) means that the minimum of the proper convex and lower semicontinuous function \( f_i(\cdot) + \lambda \|\cdot - x_i\| \) is attained at \( x_i \). Therefore,

\[
0 \in \partial f_i(x_i) = \partial f(x_i).
\]

Next, as \(-p_i\) is a subdifferential of \( \lambda \|\cdot - x_i\| \) at \( x_i \), we get

\[
p_i(x_i - x') \leq \lambda \|x' - x_i\|, \quad \forall x' \in X.
\]

So, \( \|p_i\| \leq \lambda \).

Therefore, for each \( i = 1, \ldots, k \) we have that \( \|p_i\| \leq \lambda \) and \( \|x_i - y_i\| \leq \varepsilon / \lambda \); cf. (3.3). Both of these clearly imply

\[
\|p_j\| \|x_i - y_i\| \leq \lambda \frac{\varepsilon}{\lambda} = \varepsilon, \quad \forall i, j = 1, \ldots, k.
\]

This, (3.4) and (3.3) are all we had to prove. \( \square \)

4. Integrability

We are now ready to present the proof mentioned in the title of this note.

Proof of Theorem 1.2. For each finite sequence \((x_i, p_i)_{i=1}^{n-1} \subset \text{gph} \partial f\), and each \( x \in X \) we have that

\[
f(x) \geq f(x_0) + \sum_{i=0}^{n-1} p_i(x_{i+1} - x_i),
\]

where \( x_n = x \).

Indeed, by the definition of the subdifferential, \( f(x_{i+1}) - f(x_i) \geq p_i(x_{i+1} - x_i), \quad i = 0, \ldots, n - 1 \). Summing these, we get \( f(x_n) - f(x_0) \geq \sum_{i=0}^{n-1} p_i(x_{i+1} - x_i) \), which is (4.1). So, the left-hand side of (1.2) is greater than its right-hand side.

Assume now that \( x \in \text{dom} \partial f \). Fix \( p \in \partial f(x) \) and let \((x_n, p_n) = (x, p)\).
Again using the definition of the subdifferential, \( f(x_i) - f(x_{i+1}) \geq p_{i+1}(x_i - x_{i+1}) \), that is, \( f(x_{i+1}) - f(x_i) \leq p_{i+1}(x_{i+1} - x_i) \). So,

\[
f(x_{i+1}) - f(x_i) - p_i(x_{i+1} - x_i) \leq \langle p_{i+1} - p_i, x_{i+1} - x_i \rangle,
\]

for \( i = 0, \ldots, n - 1 \). Therefore,

\[
0 \leq f(x_n) - f(x_0) - \sum_{i=0}^{n-1} p_i(x_{i+1} - x_i) \leq \sum_{i=0}^{n-1} \langle p_{i+1} - p_i, x_{i+1} - x_i \rangle,
\]

the first inequality being (4.1).

The job will be done when we find such a sequence \((x_i, p_i)_{i=1}^{n-1} \subset \text{gph } \partial f\) that the right-hand side of (4.2) is arbitrarily small.

Fix \( \varepsilon > 0 \).

Fix \( n \in \mathbb{N} \) so large that

\[
n > 2 \langle p - p_0, x - x_0 \rangle / \varepsilon
\]

and let \((x_n, p_n) = (x, p)\).

Let \( y_i = x_0 + \frac{i}{n} (x - x_0), i = 0, \ldots, n \). Since \( f \) is convex, \( y_i \in \text{dom } f \).

Using Lemma 3.3, for each \( i = 1, \ldots, n - 1 \) we get \((x_i, p_i) \in \text{gph } \partial f\) such that

\[
\|x_i - y_i\| \leq \frac{\varepsilon}{8n(1 + \|p_0\| + \|p_n\|)} \quad \text{and} \quad \|p_j\|\|x_i - y_i\| \leq \frac{\varepsilon}{8n}
\]

for all \( i, j = 1, \ldots, n - 1 \). Then, using the above and \( x_0 = y_0, x_n = y_n \),

\[
\|p_j\|\|x_i - y_i\| \leq \frac{\varepsilon}{8n} \quad \forall i, j = 0, \ldots, n.
\]

Now, for each \( i = 0, \ldots, n - 1 \) we can write

\[
\langle p_{i+1} - p_i, x_{i+1} - x_i \rangle \leq \langle p_{i+1} - p_i, y_{i+1} - y_i \rangle + \langle \|p_i\| + \|p_{i+1}\|, \|x_i - y_i\| + \|x_{i+1} - y_{i+1}\| \rangle
\]

\[
\leq \langle p_{i+1} - p_i, y_{i+1} - y_i \rangle + \frac{\varepsilon}{2n},
\]

the second inequality because of (4.4). Therefore,

\[
\sum_{i=0}^{n-1} \langle p_{i+1} - p_i, x_{i+1} - x_i \rangle \leq \sum_{i=0}^{n-1} \langle p_{i+1} - p_i, y_{i+1} - y_i \rangle + \frac{\varepsilon}{2}.
\]

But, \( y_{i+1} - y_i = \frac{x - x_0}{n} \), so

\[
\sum_{i=0}^{n-1} \langle p_{i+1} - p_i, y_{i+1} - y_i \rangle = \frac{1}{n} \sum_{i=0}^{n-1} \langle p_{i+1} - p_i, x - x_0 \rangle
\]

\[
= \frac{1}{n} \langle p_n - p_0, x - x_0 \rangle = \frac{1}{n} \langle p - p_0, x - x_0 \rangle < \frac{\varepsilon}{2},
\]

because of (4.3). Therefore, \( \sum_{i=0}^{n-1} \langle p_{i+1} - p_i, x_{i+1} - x_i \rangle < \varepsilon \) and, in the case when \( x \in \text{dom } \partial f \), (1.2) is established.

Now fix an arbitrary \( x \in X \) and a real number \( r \) such that \( r < f(x) \). By Proposition 1.3 there is \( (\bar{x}, \bar{p}) \in \text{gph } \partial f \) such that

\[
r < f(\bar{x}) + \bar{p}(\bar{x} - \bar{x}).
\]
Fix $\varepsilon > 0$. Since $\bar{x} \in \text{dom } \partial f$, referring to the case just established we find a finite sequence $(x_i, p_i)_{i=1}^{n-2} \in \text{gph } \partial f$ such that
\[
f(\bar{x}) < f(x_0) + \sum_{i=0}^{n-2} p_i(x_{i+1} - x_i) + \varepsilon,
\]
where $x_{n-1} = \bar{x}$.

By letting $p_{n-1} = \bar{p}$ and $x_n = x$ and combining the two above we obtain
\[
r < f(x_0) + \sum_{i=0}^{n-1} p_i(x_{i+1} - x_i) + \varepsilon.
\]

Since $r < f(x)$ and $\varepsilon > 0$ were arbitrary, the proof is completed. □

\textbf{Proof of Theorem 1.1.} Fix $x_0 \in \text{dom } \partial f$ and let $c = g(x_0) - f(x_0)$. From Theorem 1.2 it follows that, provided (1.1),
\[
f(x) - f(x_0) \leq g(x) - g(x_0) \iff g(x) \geq f(x) + c, \forall x \in X.
\]
From this it follows that
\[
(4.6) \quad \text{dom } g \subset \text{dom } f,
\]
and, since we can swap $x$ and $x_0$ if $x$ is also in $\text{dom } \partial f$,
\[
(4.7) \quad g(x) = f(x) + c, \forall x \in \text{dom } \partial f.
\]

Now let $x \in \text{dom } f$ be arbitrary. From Proposition 1.4 we can find a sequence $(x_n)_{n=1}^{\infty} \subset \text{dom } \partial f$ such that $\lim x_n = x$ and $\lim f(x_n) = f(x)$. The lower semi-continuity of $g$ together with (4.7) gives $g(x) \leq \lim inf g(x_n) = \lim inf(f(x_n) + c) = f(x) + c$. That is,
\[
g(x) \leq f(x) + c, \forall x \in \text{dom } f.
\]

In particular $\text{dom } g \supset \text{dom } f$, meaning that $\text{dom } g = \text{dom } f$ because of (4.6), and $g = f + c$ on $\text{dom } f$ because of (4.5). Outside of their common domain both functions are equal to $\infty$ by definition, and thus the proof is completed. □

\textbf{Acknowledgements}

We would like to express our gratitude to Prof. L. Thibault for bringing to our attention the pioneering work of Moreau [5].

\textbf{References}


Faculty of Mathematics and Informatics, University of Sofia, 5, James Bourchier Blvd., 1164 Sofia, Bulgaria
E-mail address: milen@fmi.uni-sofia.bg

Faculty of Mathematics and Informatics, University of Sofia, 5, James Bourchier Blvd., 1164 Sofia, Bulgaria
E-mail address: zlateva@fmi.uni-sofia.bg