ON GRADIENT MAPPINGS IN BANACH SPACES
TSUYOSHI ANDÔ

1. Introduction. Let $E$ be a real Banach space, $V$ a convex open subset of $E$. A real valued functional $I(x)$ defined on $V$ is said to have the Fréchet derivative $D(x, h)$ if, for arbitrary fixed $x \in V$, $D(x, h)$ is a bounded linear functional on $E$ in the variable $h$ and

$$I(x + h) - I(x) = D(x, h) + o(h),$$

where

$$o(h) = o(||h||).$$

Thus $x \rightarrow D(x, \cdot)$ is a mapping of $V$ into the space $E^*$ conjugate to $E$, and is called a gradient mapping. It is said to be compact, if the image of each bounded set of $V$ has a compact closure in $E^*$.

Recently E. H. Rothe [1] gave a (necessary and) sufficient condition for the compactness of gradient mappings under the condition that $E$ has property $(P)$, and showed that every reflexive Banach space with a basis has it. Here $E$ is said to have property $(P)$, if there exists a sequence $\{\psi_i\}$ of linearly independent elements of $E^*$ and a positive number $M$ with the following properties: the closed linear span of $\{\psi_i\}$ coincides with $E^*$ and for each positive integer $n$ there exists a linear projection of norm at most $M$ on the intersection $\bigcap_{i=1}^n N_i$, where $N_i = \{x \in E | \psi_i(x) = 0\}$. In this note a theorem with the conclusion that the gradient mapping is compact is proved under conditions differing somewhat from those of Rothe, and without using anything like the property $(P)$.

**Theorem.** Let $I(x)$ have the following two properties: (a) for each $\alpha > 0$ there exists a positive number $\beta$ such that

$$|I(x) - I(y)| \leq \beta \|x - y\|, \quad x, y \in V, \|x\|, \|y\| \leq \alpha;$$

(b) to each $\epsilon > 0$ corresponds a finite number of elements $\phi_1, \phi_2, \ldots, \phi_n$ of $E^*$ such that

$$|I(x + h) - I(x)| \leq \epsilon \|h\|, \quad x, x + h \in V,$$

for all $h$ for which $\|h\| \leq \epsilon$ and $\phi_i(h) = 0 \ (i = 1, 2, \ldots, n)$. Then the gradient mapping $D(x, \cdot)$ is compact.

2. Proof of theorem. Our proof is based on the following decomposition lemma instead of on the use of linear projections.

Received by the editors April 18, 1960.
Lemma. Let \( N \) be a closed linear subspace of \( E \) with a finite co-dimension. Then there exists a compact set \( A \) with the properties that \( \|x\| \leq 1 \) (\( x \in A \)), and each \( h \in S \) (the open unit ball) admits a decomposition (not unique) \( h = a + b \) with \( a \in A \), \( b \in N \).

Proof. We consider as usual the quotient space \( E/N \) with the norm defined by \( \|P(x)\| = \inf_{y \in N} \|x - y\| \) where \( x \rightarrow P(x) \) is the canonical mapping of \( E \) onto \( E/N \). Since \( E/N \) is finite dimensional, the image \( P(S) \) of \( S \) has a compact closure in it. We shall construct by induction a sequence \( \{a_{ij}\}, i = 0, 1, 2, \ldots, j = 1, 2, \ldots, k_i \), of elements of \( S \) such that

\[
(2) \quad \min_{k} \|a_{ik} - a_{i,k+1}\| < 1/2^i, \quad i = 0, 1, 2, \ldots, j = 1, 2, \ldots, k_i,
\]

and for each \( x \in S \) and each \( i \) there exists \( y \in E \) (depending on \( x \) and \( i \)) such that

\[
(3) \quad P(x) = P(y) \quad \text{and} \quad \min_{j} \|a_{ij} - y\| < 1/2^i.
\]

Put \( k_0 = 1 \) and \( a_{01} = 0 \) (the origin). Suppose that

\[
\{a_{ij}\}, \quad i = 0, 1, 2, \ldots, n, \quad j = 1, 2, \ldots, k_i,
\]

have been constructed. Denote, for convenience, by \( S_n(x) \) the open ball with center \( x \) and radius \( 1/2^n \). Since \( P(S_n(a_{nj}) \cap S) \) is totally bounded in \( E/N \) and \( P(S) \subset \bigcup_j P(S_n(a_{nj})) \) because of \( (3) \), there exists a finite number of elements of \( S \), denoted by \( a_{n+1,j}, j = 1, 2, \ldots, k_{n+1} \), such that \( a_{n+1,j} \in \bigcup_k S_n(a_{nk}), j = 1, 2, \ldots, k_{n+1} \), and

\[
\min_{j} \|P(x) - P(a_{n+1,j})\| < 1/2^{n+1} \quad \text{for all} \quad x \in S.
\]

Then on account of the definition of norm in \( E/N \), for each \( x \in S \) there can be chosen \( y \in E \) such that \( P(x) = P(y) \) and \( \min_{j} \|a_{n+1,j} - y\| < 1/2^{n+1} \). Thus the induction is completed. Notice that the set \( \{P(a_{ij})\} \) is dense in \( P(S) \) by construction. We claim that the closure \( A \) of the set \( \{a_{ij}\} \) is compact. In fact, from \( (2) \) it follows that for each \( n, m \) with \( m \geq n \)

\[
\min_{k} \|a_{mk} - a_{nk}\| < \sum_{i=n}^{m} 1/2^i < 1/2^{n-1}.
\]

Since \( \{a_{ij}\}_{i \leq n} \) is a finite set, this means that the set \( \{a_{ij}\} \) is totally bounded; consequently its closure \( A \) is compact. Since \( P(A) \) is compact in \( E/N \) and contains a dense subset of \( P(S) \), it follows that \( P(A) \supset P(S) \), that is, for each \( h \in S \) there exists \( a \in A \) with \( P(a) = P(h) \), i.e., \( h - a \in N \).
Now we turn to the proof of the theorem. If \( D(x, \cdot) \) is not compact, there exist a positive number \( \varepsilon \) and a sequence \( \{x_i\} \) of elements of \( V \) such that \( \|x_i\| \leq 1/\varepsilon \), \( i = 1, 2, \ldots \), and \( \|D(x_i, \cdot) - D(x_j, \cdot)\| > \varepsilon \) for \( i \neq j \). Let \( \phi_1, \phi_2, \ldots, \phi_n \) be chosen as in (b) with \( \varepsilon/8 \) instead of \( \varepsilon \). Since \( N = \{ x \in E | \phi_i(x) = 0, i = 1, 2, \ldots, n \} \) has a finite co-dimension, by the lemma there exists a compact set \( A \) of \( S \) (the closure of \( S \)) such that each \( h \in S \) admits a decomposition \( h = a + b \) with \( a \in A, b \in N \). It follows that \( \|b\| \leq \|a\| + \|h\| \leq 2 \). Since property (a) combined with definition (1) implies \( \sup_i \|D(x_i, \cdot)\| < \infty \), the family \( \{D(x_i, h)\} \), considered as continuous functions on the compact set \( A \) in the variable \( h \), is equi-continuous; consequently Ascoli-Arzela’s theorem tells us that there exists a subsequence \( \{x'_i\} \) for which

\[
\sup_{a \in A} |D(x'_i, a) - D(x'_j, a)| < \varepsilon/2, \quad i, j = 1, 2, \ldots .
\]

On the other hand, on account of the definition of norm of linear functionals,

\[
\|D(x'_i, \cdot) - D(x'_j, \cdot)\| = \sup_{h \in S} \left| D(x'_i, h) - D(x'_j, h) \right|
\]

\[
\leq \sup_{a \in A} \left| D(x'_i, a) - D(x'_j, a) \right| + 2 \cdot \sup_{k, b \in N, \|b\| \leq 2} \left| D(x'_k, b) \right| .
\]

Since \( b \in N \) and \( \|b\| \leq 2 \), property (b) (with \( \varepsilon/8 \)) combined with (1) implies \( \|D(x'_k, b)\| \leq \varepsilon/4 \), consequently

\[
\|D(x'_i, \cdot) - D(x'_j, \cdot)\| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon .
\]

This contradiction establishes the theorem.

**Reference**


**Hokkaido University, Sapporo, Japan**