Fréchet differentiability of Lipschitz functions via a variational principle

Received September 25, 2008

Abstract. We prove a new variational principle which in particular does not assume the completeness of the domain. As an application we give a new, more natural, proof of the fact that a real valued Lipschitz function on an Asplund space has points of Fréchet differentiability.

1. Introduction

We recall first some basic definitions concerning differentiability of functions defined on an open set \( G \) of a Banach space \( X \). The function \( f \) defined on \( G \) with values in a Banach space \( Y \) is said to have a directional derivative at \( x \in G \) in the direction \( u \) if

\[
f'(x; u) = \lim_{t \downarrow 0} \frac{f(x + tu) - f(x)}{t}
\]

exists. If \( f'(x; u) \) exists for every \( u \in X \) and defines a bounded linear operator \( f'(x) \) as a function of \( u \) we say that \( f \) is Gâteaux differentiable at \( x \). If, in addition, the limit (1) exists uniformly in \( u \) in the unit sphere of \( X \) we say that \( f \) is Fréchet differentiable at \( x \). Alternatively, \( f \) is Fréchet differentiable at \( x \) iff

\[
f(x + u) = f(x) + f'(x) u + o(\|u\|), \quad \|u\| \to 0.
\]

If \( f \) is a Lipschitz function and \( \dim X < \infty \) then, as easily seen, the notions of Gâteaux differentiability and Fréchet differentiability coincide. However, if \( \dim X = \infty \) the two notions are different. Consider e.g. the function

\[
f(x) = \sin x, \quad x \in \mathbb{L}_2[0, 1],
\]

from \( L_2[0, 1] \) into itself. This function is everywhere Gâteaux differentiable with \( f'(x) = \cos x \) (i.e. \( f'(x)u = u \cos x \)). The function \( f \) is however nowhere Fréchet differentiable.
Take e.g. in the equation (2) the point $x = 0$ and the direction $u = \chi_E$, the characteristic function of a set $E \subset [0, 1]$. Then notice that

$$\sin \chi_E = \chi_E + \chi_E (\sin 1 - 1)$$

and $\chi_E (\sin 1 - 1)$ is not $o(\|\chi_E\|)$ as $\|\chi_E\| \to 0$.

There are well known satisfactory results concerning the existence a.e. of Gâteaux derivatives of Lipschitz functions from $X$ to $Y$ if $X$ is separable and $Y$ has the Radon–Nikodym property (RNP). For a detailed discussion of such results as well as the meaning of a.e. and RNP we refer to the book [1]. On the other hand, there are many natural open problems concerning the existence of Fréchet derivatives, and the proofs of known theorems in this subject are often very hard.

A basic positive result in this direction was proved in [9].

**Theorem 1.1.** Let $f : G \to \mathbb{R}$ be a Lipschitz function defined on a non-empty open subset $G$ of an Asplund space $X$. Then $f$ has points of Fréchet differentiability.

Moreover, for any $a, b \in G$ for which the segment $[a, b]$ lies entirely in $G$ and every $\varepsilon > 0$ there is an $x \in G$ at which $f$ is Fréchet differentiable and

$$f'(x; b - a) < f(b) - f(a) + \varepsilon.$$
A variational principle

We define notions of completeness and lower semicontinuity in an abstract form and in
Theorem 2.4 we prove a general form of the variational principle.

Definition 2.1. Suppose that \((M, d)\) is a metric space and \(d_0\) a continuous pseudometric
on \(M\). We say that \(M\) is \((d, d_0)\)-complete if there are functions \(\delta_j : M^{j+1} \to (0, \infty)\),
j \(\geq 0\), such that every \(d\)-Cauchy sequence \((x_j)_{j=0}^{\infty}\) converges to an element of \(M\) provided
that
\[ d_0(x_j, x_{j+1}) \leq \delta_j(x_0, \ldots, x_j) \quad \text{for each } j \geq 0. \]
A function \(f : M \to \mathbb{R}\) is \((d, d_0)\)-lower semicontinuous if there are functions \(\delta_j : M^{j+1} \to (0, \infty)\),
j \(\geq 0\), such that
\[ f(x) \leq \liminf_{j \to \infty} f(x_j) \]
whenever \(x_j \in M\) converge in metric \(d\) to \(x \in M\) and
\[ d_0(x_j, x_{j+1}) \leq \delta_j(x_0, \ldots, x_j) \quad \text{for each } j \geq 0. \]

Notice that the choice \(d_0 = 0\) in Definition 2.1 yields the standard notions of completeness
and lower semicontinuity.

Lemma 2.2. Let \((M, d)\) be a metric space and \(d_0\) a continuous pseudometric on \(M\). If
\(f : M \to \mathbb{R}\) has the property that for each \(r \in \mathbb{R}\), the set \(\{x \in M \mid f(x) \leq r\}\) is a \(G_\delta\)
subset of \((M, d_0)\), then \(f\) is \((d, d_0)\)-lower semicontinuous.

Proof. For each \(q \in \mathbb{Q}\) choose a sequence \((H_i^q)_{i \geq 1}\) of \(d_0\)-open sets such that
\[ \{x \in M \mid f(x) \leq q\} = \bigcap_{i=1}^{\infty} H_i^q. \]
Order all sets \(H_i^q, i \geq 1, q \in \mathbb{Q}\), into one sequence \(G_0, G_1, \ldots\).

We denote by \(B_0[x, r]\) the \(d_0\)-closed ball with centre \(x\) and radius \(r\). Let \(\delta_0 = 1\). We
choose recursively the values \(\delta_j(x_0, \ldots, x_j) > 0, j \geq 1\), so that
Let $(M, d)$ be a complete metric space, $d_0$ a continuous pseudometric on $M$, and $S$ a $G_δ$ subset of $(M, d_0)$. Then $S$ is $(d, d_0)$-complete.

Proof. Let $f$ be the characteristic function of the complement of $S$, $f = \chi_{M \setminus S}$. Then $f$ satisfies the assumptions of Lemma 2.2. Hence $f$ is $(d, d_0)$-lower semicontinuous, i.e. there are functions $δ_j : M^{j+1} → (0, \infty)$ such that

$$f(x) ≤ \lim_{j → \infty} \inf f(x_j)$$

whenever $x_j ∈ M$ converge to $x$ in $(M, d)$ and $d_0(x_j, x_{j+1}) ≤ δ_j(x_0, x_1, ..., x_j)$ for each $j ≥ 0$.

Suppose now that a Cauchy sequence $(x_j)_{j=0}^{∞}$ in $(S, d)$ satisfies the condition $d_0(x_j, x_{j+1}) ≤ δ_j(x_0, x_1, ..., x_j)$ for each $j ≥ 0$. Since $M$ is complete, the points $x_j$ converge in $(M, d)$ to some $x ∈ M$. Moreover, by the $(d, d_0)$-lower semicontinuity of $f$,

$$f(x) ≤ \lim_{j → \infty} \inf f(x_j) = 0,$$

implying that $x ∈ S$, and we are done. □

The variational principle is contained in the following theorem.
Theorem 2.4. Let $f : M \to \mathbb{R}$ be a function bounded from below on a metric space $(M, d)$. Let $d_0$ be a continuous pseudometric on $M$ such that

- $M$ is $(d, d_0)$-complete;
- $f$ is $(d, d_0)$-lower semicontinuous.

Let $F_j : M \times M \to [0, \infty]$, $j \geq 0$, be functions which are $d$-lower semicontinuous in the second variable with $F_j(x, x) = 0$ for all $x \in M$ and

$$\inf_{d(x, y) > r_j} F_j(x, y) > 0 \text{ for some sequence } r_j \downarrow 0. \quad (3)$$

If $x_0 \in M$ and $(\varepsilon_j)_{j=0}^{\infty}$ is any sequence of positive numbers such that

$$f(x_0) \leq \varepsilon_0 + \inf_{x \in S} f(x),$$

then one may find a sequence $(x_j)_{j=1}^{\infty}$ in $M$ converging in the metric $d$ to some $x_\infty \in M$ and a $d_0$-continuous function $\varphi : M \to \mathbb{R}$ such that the function

$$h(x) := f(x) + \varphi(x) + \sum_{j=0}^{\infty} F_j(x_j, x)$$

attains its minimum on $M$ at $x = x_\infty$. Moreover, if $0 \leq j < k + 1 \leq \infty$,

$$F_j(x_j, x_k) \leq \varepsilon_j \text{ and } h(x_\infty) \leq \varepsilon_j + \inf_{x \in M} \left( f(x) + \varphi(x) + \sum_{i=0}^{j-1} F_i(x_i, x) \right). \quad (4)$$

Before the proof it will be useful to make some comments on Theorem 2.4.

Comments. By considering the case $j = 0$ in the second inequality in (4) it is immediate to see that $h(x_\infty) < \infty$.

Notice also that the statements of Theorem 2.4 do not change when a constant is added to $\varphi$. So one can always replace $\varphi$ by $\varphi - \varphi(x_\infty)$ to achieve that $\varphi(x_\infty) = 0$. Once we have $\varphi(x_\infty) = 0$, we may replace $\varphi$ once more by $\max\{0, \varphi\}$ leaving $h(x_\infty)$ unchanged and not decreasing the other values of $h$. Therefore we may add to the statements of Theorem 2.4 the requirement that $\varphi \geq \varphi(x_\infty) = 0$.

The usual form of the smooth variational principle as considered in [2] is the special case of Theorem 2.4 when $M$ is a Banach space, $d_0 = 0$, $f$ is $\|\cdot\|$-lower semicontinuous and all $F_j$ are constant multiples of a single (smooth) function $F(x - y)$. Then the function $\varphi$ is necessarily constant and so the previous remark says that we can have $\varphi = 0$. In a similar way one can obtain Ekeland’s variational principle putting $F_j(x, y)$ to be a multiple of the metric $d(x, y)$.

In addition to treating non-lower semicontinuous functions $f$ and incomplete spaces $M$, the above variational principle refines the statement of the usual smooth variational principles also in other aspects. The dependence of the perturbation functions $F_j$ on $j$ is allowed to be almost arbitrary and a stronger conclusion than just existence of minimum is stated in (4). Other similar ramifications may be obtained by carefully following
the proof. For example, infinitely many $j$’s are sufficient in (3) or given any $\varepsilon > \varepsilon_0$, we can have $\varepsilon \geq \varphi \geq \varphi(x_{\infty}) = 0$. However, in Theorem 2.4 we have just given those statements that will be used in the following.

**Proof of Theorem 2.4** By diminishing the values of $\varepsilon_j$ for $j \geq 1$, if necessary, we may assume that $\sum_{j=1}^{\infty} \varepsilon_j < \infty$ and

$$\inf_{d(x,y) > r_j} F_j(x, y) > \varepsilon_j$$

for every $j \geq 1$. Find the functions $\delta_j: M_{j+1} \to (0, \infty)$ witnessing the fact $f$ is $(d, d_0)$-lower semicontinuous. We may also assume that the same functions $\delta_j$ witness the $(d, d_0)$-completeness of $M$.

We will now define by induction points $x_j \in M$, $d_0$-continuous functions $\varphi_j: M \to [0, \infty)$ and lower bounded functions $h_j: M \to (-\infty, \infty]$ such that, in particular, the following requirements hold:

(i) $\varphi_j(x_j) = 0$ for $j \geq 0$;
(ii) $h_j(x_j) \leq \varepsilon_j + \inf_{x \in M} h_j(x)$ for $j \geq 0$;
(iii) $h_j(x_j) \leq h_{j-1}(x_j-1)$ for $j \geq 1$;
(iv) $\varphi_j(x_j) + F_{j-1}(x_j, x_j) \leq \varepsilon_j$ whenever $j \geq i \geq 0$.

To start the definitions for $j = 0$, recall that the point $x_0 \in M$ is already given. We let $h_0(x) = f(x)$ and define

$$\varphi_0(x) = 2\varepsilon_0 \min\{1, d_0(x, x_0)/\delta_0(x_0)\}.$$  

Then (i) is satisfied. By assumptions,

$$h_0(x_0) \leq \varepsilon_0 + \inf_{x \in M} h_0(x),$$  

which is the requirement (ii). The condition (iii) requires nothing for $j = 0$ and the remaining condition (iv) is satisfied trivially.

Assume now that for some $j > 0$, $x_i, \varphi_i$ and $h_i$ have already been defined for $i = 0, 1, \ldots, j - 1$. We let

$$h_j(x) = h_{j-1}(x) + \varphi_{j-1}(x) + F_{j-1}(x_j, x_j) = f(x) + \sum_{i=0}^{j-1} (\varphi_i(x) + F_i(x_i, x)).$$

The observation that $h_j(x_{j-1}) = h_{j-1}(x_{j-1})$ allows us to choose $x_j$ so that

$$h_j(x_j) \leq \min\{h_{j-1}(x_{j-1}), \varepsilon_j + \inf_{x \in M} h_j(x)\}.$$  

Finally, we put

$$\varphi_j(x) = 2\varepsilon_j \min\{1, d_0(x, x_j)/\delta_j(x_0, \ldots, x_j)\}.$$
Since (i), (ii) and (iii) hold true we turn our attention to the condition (iv). For \( i = j \) the left hand side is zero. Let \( 0 \leq i < j \). Then

\[
\begin{align*}
&h_i(x_i) \geq h_j(x_j) & \text{since } h_i(x_i) \text{ decrease in } i \\
&\geq h_{i+1}(x_j) & \text{since } h_i(x) \text{ increase in } i \\
&= h_i(x_j) + \varphi_i(x_j) + F_i(x_i, x_j) & \text{by definition of } h_{i+1} \\
&\geq h_i(x_i) + \varphi_i(x_j) + F_i(x_i, x_j) - \varepsilon_i & \text{since } h_i(x_i) \leq \varepsilon_i + h_i(x).
\end{align*}
\]

Subtracting \( h_i(x_i) \) from both sides, we obtain (iv) and the construction is finished.

We now deduce from (iv) that the sequence \((x_j)\) is \(d\)-Cauchy. For any \( \varepsilon > 0 \) find an index \( i \) so that \( r_i \leq \varepsilon \). Hence, using (5), \( F_j(x, y) > \varepsilon_i \) whenever \( d(x, y) > \varepsilon \). Since (iv) gives that for all \( j > i \), \( F_i(x_i, x_j) \leq \varepsilon_i \), this implies that \( d(x_i, x_j) \leq \varepsilon \), as required.

We also notice that the sequence \((x_j)\) satisfies \( d_{0}(x_i, x_j) \leq \delta_i(x_0, \ldots, x_i) \) for \( j > i \); the opposite inequality \( d_{0}(x_i, x_j) > \delta_i(x_0, \ldots, x_i) \) would imply that \( \varphi_i(x_j) = 2\varepsilon_i \), which contradicts (iv).

Hence, by \((d, d_{0})\)-completeness of \( M \) and \((d, d_{0})\)-lower semicontinuity of \( f \), the sequence \((x_j)\) converges in the metric \( d \) to some \( x_{\infty} \in M \) and

\[
f(x_{\infty}) \leq \liminf_{j \to \infty} f(x_j).
\]

Let \( \varphi(x) = \sum_{j=0}^{\infty} \varphi_j(x) \). Since the series is uniformly convergent, the function \( \varphi \) is \( d_{0}\)-continuous. In particular, it is \( d \)-continuous as well. The desired perturbed function \( h(x) \) from the statement of the theorem will be

\[
h(x) = f(x) + \varphi(x) + \sum_{j=0}^{\infty} F_j(x_j, x).
\]

Recalling that the functions \( x \mapsto F(x_j, x) \) are \( d \)-lower semicontinuous, we get, for any fixed \( k \),

\[
h_0(x_0) \geq h_1(x_1) \geq \cdots \geq \lim_{j \to \infty} h_j(x_1) \geq \liminf_{j \to \infty} \left( f(x_j) + \varphi(x_j) + \sum_{i=0}^{k} F_i(x_i, x_j) \right) \\
\geq f(x_{\infty}) + \varphi(x_{\infty}) + \sum_{i=0}^{k} F_i(x_i, x_{\infty}).
\]

Hence

\[
h_j(x_j) \geq h(x_{\infty}) \quad \text{for all } j \geq 0. \tag{6}
\]

The first inequality in (4) follows immediately from (iv), and the second from (6) and (vi).

\[
h(x_{\infty}) \leq h_j(x_j) \leq \varepsilon_j + \inf_{x \in M} h_j(x) \leq \varepsilon_j + \inf_{x \in M} \left( f(x) + \varphi(x) + \sum_{i=0}^{j-1} F_i(x_i, x) \right).
\]
Finally, we may take the limit as \( j \to \infty \) and conclude that

\[
h(x_\infty) \leq \inf_{x \in M} \left( f(x) + \phi(x) + \sum_{i=0}^{\infty} F_i(x_i, x) \right) = \inf_{x \in M} h(x).
\]

Hence \( h(x) \) attains its minimum on \( M \) at \( x = x_\infty \).

\[\Box\]

3. An illustrative special case

To demonstrate the variational approach we prove the following theorem which was first proved in [8].

**Theorem 3.1.** Let \( X \) be a Banach space with separable dual \( X^* \) and let \( f: X \to \mathbb{R} \) be Lipschitz and everywhere Gâteaux differentiable. Then \( f \) has points of Fréchet differentiability.

Besides the variational principle, which is its main ingredient, the proof requires two additional observations. Recall that a function \( \Theta: X \to \mathbb{R} \) is called upper Fréchet differentiable at \( x_0 \in X \) if there is \( L \in X^* \) such that

\[
\limsup_{\|u\| \to 0} \frac{\Theta(x_0 + u) + \Theta(x_0) - Lu}{\|u\|} \leq 0.
\]

**Observation 3.2.** Let \( \Theta: X \to \mathbb{R} \) be everywhere upper Fréchet differentiable, \( \psi: X \to \mathbb{R} \) continuous and \( f: X \to \mathbb{R} \) Lipschitz and everywhere Gâteaux differentiable. Suppose further that the function \( h: X \times X \to \mathbb{R} \) given as

\[h(x, u) = f'(x; u) + \Theta(u) + \psi(x)\]

attains its minimum at \( (x_0, u_0) \). Then \( f \) is Fréchet differentiable at \( x_0 \).

**Proof.** Although it is not necessary, we first notice that after fixing the variable \( x = x_0 \) in \( h(x, u) \), the resulting function of \( u \) is upper differentiable and attains its minimum at \( u = u_0 \). Hence its upper derivative is zero, giving that \( f'(x_0) + L = 0 \), where \( L \) is an upper derivative of \( \Theta \) at \( u_0 \). So it should be no surprise that the formulas below actually show that the Fréchet derivative of \( f \) at \( x_0 \) is equal to \( -L \).

Let \( \varepsilon > 0 \) and find \( \Delta > 0 \) so that

\[
\Theta(u) - \Theta(u_0) \leq L(u - u_0) + \frac{\varepsilon}{3}\|u - u_0\|
\]

for \( \|u - u_0\| \leq \Delta \). By continuity of \( \psi \), there is \( \delta_0 > 0 \) such that

\[
|\psi(x) - \psi(x_0)| < \frac{\varepsilon\Delta}{3}
\]

for \( \|x - x_0\| \leq \delta_0 \). Finally, let \( \delta > 0 \) be such that both \( \delta(1 + \|u_0\|/\Delta) < \delta_0 \) and

\[
|f(x_0 + tu_0) - f(x_0) - f'(x_0; tu_0)| \leq \frac{\varepsilon\Delta}{3}|t|
\]

for \( |t| \leq \delta/\Delta \).
Hence, using that
\[ \| \psi(x) - \psi(x_0) \| = \| f(x) - f(x_0) \| \]
we obtain
\[ f'(x_0; u_0) - L(u - u_0) - \frac{\varepsilon}{3} \| u - u_0 \| - \frac{\varepsilon \Delta}{3} \]
\[ \geq f'(x_0; u_0) - L(u - u_0) - \frac{\varepsilon}{3} \| u - u_0 \| \]
\[ = f'(x_0; u_0) - L(u - u_0) - \frac{\varepsilon}{3} \| u - u_0 \|. \]
Since \(|t| \leq \delta/\Delta\), we also have
\[ f'(x_0; t u_0) \geq f'(x_0; u_0) - \frac{\varepsilon \Delta}{3} - \varepsilon \| u \| - \frac{\varepsilon}{3} \| v \|. \]
Adding this to (7), we get
\[ f'(x_0 + v; u_0) \geq f'(x_0; u_0) - L(v) - \frac{2\varepsilon}{3} \| v \|. \]
To obtain the upper estimate of this increment, we proceed in a completely symmetric way. Let \( t \) be as above but this time we let \( u = (1/t)v + u_0 \). For every \( x \in [x_0 + v, x_0 + t u_0] \), using again that
\[ f'(x_0; u_0) - L(u - u_0) - \frac{\varepsilon}{3} \| u - u_0 \| - \frac{\varepsilon \Delta}{3} \]
we get
\[ f'(x_0; u_0) - L(u - u_0) - \frac{\varepsilon}{3} \| u - u_0 \| \]
\[ = f'(x_0; u_0) + \frac{1}{t} L(v) - \frac{2\varepsilon}{3} \| v \|. \]
So
\[ f'(x_0 + t u_0) - f(x_0 + v) \geq t \inf \{ f'(x_0; u) \mid x \in [x_0 + v, x_0 + t u_0] \} \]
\[ \geq f'(x_0; t u_0) + L(v) - \frac{2\varepsilon}{3} \| v \|. \]
Subtracting this from
\[ f(x_0 + tu_0) - f(x_0) \leq f'(x_0; tu_0) + \frac{\varepsilon \Delta t}{3} \]
we get
\[ f(x_0 + v) - f(x_0) \leq -L(v) + \frac{2\varepsilon}{3} \|v\| + \frac{\varepsilon \Delta t}{3} = -L(v) + \varepsilon \|v\|. \]

Clearly, the above argument works under considerably less stringent assumptions. For example, the upper Fréchet differentiability of \( f \) was used at \( u_0 \) only, and much of the assumed Gâteaux differentiability of \( f \) remained unused. Since in applications \( u_0 \) is not known in advance, the former observation has little influence on what follows, but the latter, after a technical improvement, is in the very heart of our approach to the general case.

The observation below represents the specific feature of our approach. The directional derivative \( f'(x; u) \) is not continuous as a function in two variables \( x, u \) but it is \((d, d_0)\)-continuous for suitable choice of metrics \( d \) and \( d_0 \).

**Observation 3.3.** Let \( X \) be a Banach space and \( f : X \to \mathbb{R} \) a Lipschitz and everywhere Gâteaux differentiable function. Let \( M := X \times X \) be equipped with the metric
\[ d((x, u), (y, v)) = \sqrt{\|x - y\|^2 + \|u - v\|^2} \]
and with the continuous pseudometric
\[ d_0((x, u), (y, v)) = \|x - y\|. \]

Then the map \((x, u) \mapsto f'(x; u)\) from \( M \) to \( \mathbb{R} \) is \((d, d_0)\)-continuous.

**Proof.** We will define the following strategies \( \delta_j : M_{j+1} \to (0, \infty) \) witnessing the \((d, d_0)\)-continuity. Let \( \delta_0(x_0, u_0) = 1 \). Given \( j \geq 1 \) and the pairs \((x_0, u_0), \ldots, (x_j, u_j)\), we find
\[ 0 < \delta \leq \frac{1}{2} \delta_{j-1}((x_0, u_0), \ldots, (x_{j-1}, u_{j-1})) \]
so that
\[ |f(x_j + tu_j) - f(x_j) - f'(x_j; tu_j)| \leq |t|/j \]
whenever \(|t| \leq j\delta\). Then we put \( \delta_j((x_0, u_0), \ldots, (x_j, u_j)) = \delta \).

We have to show that \( f'(x_j; u_j) \) converge to \( f'(x; u) \) whenever the pairs \((x_j, u_j) \in M \) \(d\)-converge to \((x, u) \in M\) and
\[ \|x_{j+1} - x_j\| \leq \delta_j((x_0, u_0), \ldots, (x_j, u_j)). \]
To simplify the notation we let \( \delta_j = \delta_j((x_0, u_0), \ldots, (x_j, u_j)). \)

Since \( \delta_{j+1} \leq \frac{1}{2} \delta_j \), we have \( \|x - x_j\| \leq 2\delta_j \) and also \( j\delta_j \to 0 \). Let \( \varepsilon > 0 \). One can find \( j \in \mathbb{N} \) so that
\[ \text{Lip}(f) \|u - u_j\| < \varepsilon, \quad \frac{1 + 4 \text{Lip}(f)}{j} < \varepsilon, \]
and
\[ |f(x + tu) - f(x) - f'(x; tu)| \leq \varepsilon |t| \quad \text{for } |t| \leq j\delta_j. \]

Let \( t \in \mathbb{R} \) with \( |t| = j\delta_j \). Then we obtain the following estimate:
\[
|f'(x; tu) - f'(x_j; tu_j)| \leq |f(x + tu) - f(x) - f'(x; tu)| + \varepsilon |t|/j \\
\leq 2 \text{Lip}(f) \|x - x_j\| + 2 \text{Lip}(f) \|u - u_j\| |t|/j + \varepsilon |t|/j \\
\leq \left( \frac{4 \text{Lip}(f) \delta_j}{|t|} + \frac{\text{Lip}(f) \|u - u_j\| + \varepsilon}{j} \right) |t| \\
= \left( \frac{1}{j} + \frac{4 \text{Lip}(f)}{j} \right) \|u - u_j\| + \varepsilon |t|/j \leq 3 \varepsilon |t|. \]

**Proof of Theorem 3.1.** Since the dual space \( X^* \) is separable, the space \( X \) admits a Fréchet smooth norm \( \| \cdot \| \) (see e.g. [4, Theorem 3.1]). We plan to apply our variational principle, Theorem 2.4, on the metric space \( M = X \times X \) equipped with the metric
\[
d((x,u), (y,v)) = \sqrt{\|x - y\|^2 + \|u - v\|^2},
\]
and with the pseudometric \( d_0 \),
\[
d_0((x,u), (y,v)) = \|x - y\|.
\]
The space \( (M, d) \) is complete and so \( (d, d_0) \)-complete as well. We choose the functions
\[
F_j((x,u), (y,v)) := 2^{-j} d^2((x,u), (y,v)) = 2^{-j} (\|x - y\|^2 + \|u - v\|^2)
\]
and constants \( r_j = 2^{-j}, j \geq 0 \). Then clearly
\[
\inf \{ F_j((x,u), (y,v)) \mid d((x,u), (y,v)) > r_j \} > 0.
\]
Theorem 2.4 will be used to find a suitable minimum attaining perturbation of the function
\[
g(x, u) := f'(x; u) + \|u\|^2.
\]
For this, we still need to check the remaining assumptions. First, Observation 3.3 guarantees that the function \( g \) is \( (d, d_0) \)-continuous. It is also bounded from below, since \( g(x, u) \geq -\text{Lip}(f) \|u\| + \|u\|^2 \); this was the reason for adding \( \|u\|^2 \). The choice of the starting point and of the parameters \( \varepsilon_j \) controlling the speed of convergence is irrelevant in our situation, with the exception of the case \( j = 0 \) when we have an assumption to verify. Thus we set, e.g., \( \varepsilon_j = 2^{-j} \) and we find the starting point \( (x_0, u_0) \) so that
\[
g(x_0, u_0) < \varepsilon_0 + \inf_{(x,u) \in M} g(x,u).
\]

Theorem 2.4 provides us with a sequence of pairs \((x_j, u_j)\) converging to some \((x_\infty, u_\infty)\) and a \( d_0 \)-continuous function \( \varphi: X \times X \to \mathbb{R} \) such that the function
\[
h(x, u) = f'(x, u) + \|u\|^2 + \varphi(x, u) + \sum_{j=0}^{\infty} 2^{-j} (\|x - x_j\|^2 + \|u - u_j\|^2)
\]
attains its minimum at \((x_{\infty}, u_{\infty})\). Notice that \(d_0\)-continuity of \(\varphi\) means that the function \(\varphi\) depends only on the variable \(x\). An appeal to Observation 3.2 with \(\psi(x) = \varphi(x) + \sum_{j=0}^{\infty} 2^{-j} \|x - x_j\|^2\) and
\[
\Theta(u) = \|u\|^2 + \sum_{j=0}^{\infty} 2^{-j} \|u - u_j\|^2
\]
gives that \(f\) is Fréchet differentiable at \(x_{\infty}\). 

\[\square\]

4. A one-dimensional ‘mean value’ estimate

We will need a certain variant of standard estimates of a maximal operator which is easily deduced from the weak type \((1, 1)\) inequality for the Hardy–Littlewood maximal operator. Recall that the Hardy–Littlewood maximal operator for an integrable function \(f : \mathbb{R} \to \mathbb{R}\) is defined by
\[
Mf(t) = \sup_{r>0} \frac{1}{2r} \int_{t-r}^{t+r} |f(s)| \, ds,
\]
and it satisfies the weak type \((1, 1)\) estimate
\[
\mathcal{L}^1 \{ t \in \mathbb{R} \mid Mf(t) > \lambda \} \leq \frac{2}{\lambda} \| f \|_{L^1} \quad \text{for all } \lambda > 0.
\]

**Lemma 4.1.** Let \(h : [a, b] \to \mathbb{R}\) be a Lipschitz function with \(h(a) = h(b)\). Denote, for \(t \in [a, b]\),
\[
H(t) = \sup \left\{ \frac{|h(\beta) - h(\alpha)|}{\beta - \alpha} \mid a \leq \alpha \leq t \leq \beta \leq b, \alpha < \beta \right\}.
\]
Then for every \(\lambda > 0\),
\[
\mathcal{L}^1 \{ t \in [a, b] \mid H(t) > \lambda \} \leq \frac{8}{\lambda} \int_a^b \max\{0, h'(t)\} \, dt.
\]

**Proof.** Let \(\lambda > 0\). Since clearly \(H(t) \leq 2Mh'(t)\), we obtain
\[
\mathcal{L}^1 \{ t \in [a, b] \mid H(t) > \lambda \} \leq \frac{4}{\lambda} \int_a^b |h'(t)| \, dt.
\]
The condition \(h(a) = h(b)\) implies, however, that
\[
\int_{|h'(t)| > 0} h'(t) \, dt = \int_{|h'(t)| < 0} -h'(t) \, dt,
\]
i.e.
\[
\int_a^b |h'(t)| \, dt = 2 \int_a^b \max\{0, h'(t)\} \, dt. \quad \square
\]

**Lemma 4.2.** Let \(h : [a, b] \to \mathbb{R}\) be a Lipschitz function, \(h(b) = h(a) = 0\), and let \(\kappa = \text{Lip}(h)\). Then there is a set \(S \subset [a, b]\) such that \(3\kappa \mathcal{L}^1 S \geq \|h\|_{\infty}\) and for every \(\xi \in S\).
(i) $h$ is differentiable at $\xi$;
(ii) $h'(\xi) \geq \|h\|_\infty / 3(b - a)$;
(iii) $|h(t) - h(\xi)| \leq 48\sqrt{\kappa h'(\xi)}|t - \xi|$ for every $t \in [a, b]$.

**Proof.** Denote by $S \subset [a, b]$ the set of all points $\xi \in (a, b)$ for which (i)–(iii) hold and suppose to the contrary that $L^1 S < \frac{1}{3\kappa \|h\|_\infty}$. Then for a.e. $t \in [a, b] \setminus S$ either $h'(t) < \|h\|_\infty/3(b - a)$ or $h'(t) \geq \|h\|_\infty/3(b - a)$ and $H(t) > 48\sqrt{\kappa h'(t)}$.

Recalling also that $h' \leq \kappa$, we see that for a.e. $t \in [a, b]$, $\max\{0, h'(t)\} < \max\{0, h'(t)\} \chi_S(t) + \|h\|_\infty/3(b - a) + \min\{\kappa, 1/48^2 \kappa H^2(t)\}$. Hence we get with the help of Lemma 4.1 a contradiction by estimating

$$\int_a^b \max\{0, h'(t)\} \, dt < \int_S \max\{0, h'(t)\} \, dt + \int_a^b \|h\|_\infty/3(b - a) \, dt + \int_a^b \min\{\kappa, 1/48^2 \kappa H^2(t)\} \, dt$$

$$\leq \kappa L^1 S + \frac{1}{3} \|h\|_\infty + \int_0^\kappa L^1\left\{ t \in [a, b] \mid 1/48^2 \kappa H^2(t) > \lambda \right\} d\lambda$$

$$\leq \frac{2}{3} \|h\|_\infty + \int_0^\kappa \frac{1}{6 \kappa \lambda} d\lambda \int_a^b \max\{0, h'(t)\} \, dt$$

$$\leq \frac{2}{3} \int_a^b \max\{0, h'(t)\} \, dt + \frac{2}{6} \int_a^b \max\{0, h'(t)\} \, dt$$

$$= \int_a^b \max\{0, h'(t)\} \, dt. \quad \Box$$

**Corollary 4.3.** Let $g, h: [-\delta, \delta] \rightarrow \mathbb{R}$ be functions, $0 < a \leq \delta$, and let $h(t) = g(t)$ for $a \leq |t| \leq \delta$. Suppose further that $h$ is $\kappa$-Lipschitz and that the parameter $0 < \tau \leq \kappa$ is such that

$$|h(0) - g(0)| \geq 13\tau a \quad \text{and} \quad |g(t) - g(0)| \leq \frac{1}{\kappa} \tau^2 |t| \quad \text{for } |t| \leq \delta.$$

Then there is $\xi \in (-a, a) \setminus \{0\}$ with the following properties:

(i) $h$ is differentiable at $\xi$;
(ii) $h'(\xi) \geq \tau$;
(iii) $|h(t) - h(\xi)| \leq 99\sqrt{\kappa h'(\xi)}|t - \xi|$ for every $t \in [-\delta, \delta]$. 
Proof. Assume that \( \text{Lip}(h) > 0 \). Since neither the assumptions nor the conclusion change when we add the same constant to the functions \( g, h \), we may subtract \( g(0) \) from both of them, and so assume that \( g(0) = 0 \). Define \( h_0: [-a, a] \to \mathbb{R} \) by

\[
h_0(t) = h(t) - \frac{a - t}{2a} h(-a) - \frac{a + t}{2a} h(a).
\]

Notice that \( \text{Lip}(h_0) < 2 \text{Lip}(h) \). We apply Lemma 4.2 to the function \( h_0 \) and \( \kappa_0 = \text{Lip}(h_0) \) to get a subset \( S \subset [-a, a] \) such that \( 3\kappa_0\mathcal{L}^1 S \geq \|h_0\| \) and for every \( \xi \in S \),

(a) \( h_0 \) is differentiable at \( \xi \);
(b) \( h_0'(\xi) \geq \|h_0\|_\infty / 6a \);
(c) \( |h_0(t) - h_0(\xi)| \leq 48\sqrt{\kappa_0 h_0'(\xi)}|t - \xi| \) for every \( t \in [a, b] \).

In order to see what (a), (b), and (c) imply for the original function \( h \), denote \( \eta = \tau^2 / \kappa \) and observe that, clearly, \( \eta \leq \tau \) and \( \eta \leq \kappa \). Now

\[
\sup_{t \in [-a, a]} |h_0(t) - h(t)| = \max(|h(a)|, |h(-a)|) = \max([g(a)], [g(-a)]) \leq \eta a,
\]

and for \( t, s \in [-a, a] \),

\[
|h(t) - h(s)| - (h_0(t) - h_0(s))| \leq \frac{|t - s|}{2a} |h(a) - h(-a)|
= |t - s| \left( \left| g(a) - g(-a) \right| \leq \eta |t - s| \right).
\]

In particular, \( \|h_0\|_\infty \geq \|h|_{[a, a]}\|_{\infty} - \eta a \geq |h(0)| - \eta a \geq 13\tau a - \eta a = 12\tau a \). Since \( \kappa_0 < 2\kappa \) it follows that

\[
\mathcal{L}^1([-a, a] \setminus [-a + \tau a / \kappa, a - \tau a / \kappa]) = \frac{2\tau a}{\kappa} \leq \frac{\|h_0\|_\infty}{6a} \leq \frac{\|h_0\|_\infty}{3\kappa_0} \leq \mathcal{L}^1 S.
\]

Thus we can find a point \( \xi \in S \cap [-a + \tau a / \kappa, a - \tau a / \kappa] \setminus \{0\} \). Since \( h_0'(\xi) \) exists, \( h'(\xi) \) exists as well. By (b) and (c) we also have

\[
h'(\xi) \geq h_0'(\xi) - \eta \geq \frac{\|h_0\|_\infty}{6a} - \eta \geq 2\tau - \eta \geq \tau,
\]

which gives (d). Moreover, it also gives

\[
h_0'(\xi) \leq h'(\xi) + \eta \leq h'(\xi) + \tau \leq 2h'(\xi).
\]

Let \( t \in [-a, a] \). Then by (9) and (c) we obtain

\[
|h(t) - h(\xi)| \leq |h_0(t) - h_0(\xi)| + \eta |t - \xi| \leq (48\sqrt{\kappa_0 h_0'(\xi)} + \eta) |t - \xi| \leq (96 + \eta / \sqrt{\kappa \tau}) \sqrt{\kappa h'(\xi)} |t - \xi| \leq 97 \sqrt{\kappa h'(\xi)} |t - \xi|.
\]
Hence (iii) holds for $t \in [-a, a]$. It remains to show that (iii) holds also for $t \in [-\delta, \delta] \setminus [-a, a]$. Let $s$ be the (unique) point of $(\xi, t) \cap [-a, a]$. Since $\xi \in [-a + \tau a/\kappa, a - \tau a/\kappa]$, we have $|s - \xi| \geq \tau|s|/\kappa$ and so

$$|t - \xi| = |t - s| + |s - \xi| \geq \frac{\tau}{\kappa}|t - s| + \frac{\tau}{\kappa}|s| = \frac{\tau}{\kappa}|t|.$$ 

Hence

$$|h(t) - h(s)| = |g(t) - g(s)| \leq \eta(|t| + |s|) \leq 2\eta|t| \leq \frac{2\eta\kappa}{\tau}|t - \xi| = 2\tau|t - \xi| \leq 2\sqrt{\kappa h'(\xi)}|t - \xi|.$$ 

Since (10) implies

$$|h(s) - h(\xi)| \leq 97\sqrt{\kappa h'(\xi)}|s - \xi| \leq 97\sqrt{\kappa h'(\xi)}|t - \xi|,$$ 

we get the required

$$|h(t) - h(\xi)| \leq |h(t) - h(s)| + |h(s) - h(\xi)| \leq 99\sqrt{\kappa h'(\xi)}|t - \xi|.$$ 

5. Proof of Theorem 1.1

We will prove the following variant of Theorem 1.1. The main difference is that we have replaced the assumption that $X$ be Asplund by existence of a suitably smooth bump function. In fact, our argument uses only the weaker assumption, but in view of the separable reduction arguments (cf. [7] or [5]), such conditions do not bring anything new. However, as pointed out at the final comments, they may be useful in studying other derivatives.

Proposition 5.1. Let $f : G \to \mathbb{R}$ be a Lipschitz function defined on a non-empty open subset $G$ of a Banach space $X$. Assume that $X$ admits a locally Lipschitz, everywhere upper Fréchet differentiable function $2 : X \to [0, \infty)$ such that

$$2(0) = 0, \quad \text{and} \quad \inf_{\|z\| > \tau} \Theta(z) > 0 \quad \text{for every} \quad \tau > 0.$$ 

Then $f$ has a point of Fréchet differentiability.

More precisely, if $x_0 \in G$ and $u_0 \in X$ are such that $f'(x_0; u_0)$ exists, then for every $r_0 > 0$ there are $x \in G$ and $u \in B(u_0, r_0)$ such that $f$ is Fréchet differentiable at $x$ and $f'(x; u) \leq f'(x_0; u_0)$.

The first statement of Theorem 1.1 follows immediately from this. For the second statement, we let $u_0 = b - a$ and notice that the function $t \mapsto f(a + tu_0)$ is Lipschitz. Hence there is a point $x_0 \in [a, b]$ so that $f'(x_0; u_0)$ exists and

$$f'(x_0; u_0) \leq f(b) - f(a).$$
To finish the proof of Theorem 1.1, we use Proposition 5.1 with some $r_0 > 0$ such that $r_0 \text{Lip}(f) < \varepsilon$. Then

\[ f'(x; u_0) \leq f'(x; u) + \text{Lip}(f)\|u - u_0\| < f'(x_0; u_0) + \varepsilon \leq f(b) - f(a) + \varepsilon, \]

as required.

The rest of this section is devoted to the proof of Proposition 5.1.

The space $M$. We shall assume that $\|u_0\| = 1$ and that $0 < r_0 < 1$ is so small that $B(x_0, r_0) \subseteq G$ and $\Theta$ is Lipschitz on $B(0, r_0)$. We let $r = \frac{1}{2}r_0$ and denote $B = \{x \in X | \|x - x_0\| \leq r\}, U = \{u \in X | \|u - u_0\| \leq r\},$ and

\[ M = \{(x, u) \in B \times U | f'(x; u) \text{ exists}\}. \]

We intend to define a new metric on $M$ not topologically equivalent to the metric resulting from its embedding into $X \times X$ which will enable us to use the variational principle. To define it, we map $M$ into the function space $\mathcal{P} = \{g: (-r, r) \rightarrow \mathbb{R} | g \text{ is Lipschitz, } g(0) = 0\}$

by assigning to $(x, u)$ the function $f_{x,u}: (-r, r) \rightarrow \mathbb{R}$ given as

\[ f_{x,u}(t) = f(x + tu) - f(x). \tag{11} \]

We equip $\mathcal{P}$ with the distance $\varrho(g, h)$ defined by

\[ \varrho(g, h) = \sup_{t \in (-r, r) \setminus \{0\}} \frac{|g(t) - h(t)|}{|t|}. \]

Notice that $\varrho(g, h) \leq \text{Lip}(g) + \text{Lip}(h)$, hence $\varrho$ is finite. The wished-for metric $d$ on $M$ is

\[ d((x, u), (y, v)) = \max\{\|x - y\|, \|u - v\|, \varrho(f_{x,u}, f_{y,v})\}. \tag{12} \]

The idea behind this approach is that the convergence in the metric $d$ allows control of the increments $f(x + tu) - f(x)$, thus making it possible to deduce that $f'(x; u)$ exists whenever there is a sequence $(x_k, u_k)$ $d$-converging to $(x, u)$ such that all $f'(x_k; u_k)$ exist.

In the first lemma, we notice that our construction has the additional advantages that $M$ is complete and that the function mapping $(x, u) \in M$ to $f'(x; u)$ is continuous. Hence the variational principle will be used in the simpler form, without the pseudometric $d_0$.

**Lemma 5.2.** The space $(M, d)$ is complete and the real-valued function $(x, u) \mapsto f'(x; u)$ is $d$-continuous and bounded from below on $M$.  


Proof. Suppose that a sequence \((x_k, u_k) \in M\) is \(d\)-Cauchy. Then the points \(x_k\) norm converge to some \(x \in B\), the directions \(u_k\) norm converge to some \(u \in U\) and the functions \(f_{x_k, u_k}\) converge in the space \(\mathcal{P}\) to some function \(g\). The last assertion means that

\[
\sup_{t \in (-r, r) \setminus \{0\}} \frac{|f_{x_k, u_k}(t) - g(t)|}{|t|} \to 0. \tag{13}
\]

Since on the set \((-r, r)\) the functions \(f_{x_k, u_k}\) converge pointwise to \(f_{x, u}\), we see that \(g(t) = f_{x, u}(t)\). Hence \(f_{x_k, u_k}\) converge to \(f_{x, u}\) in \((\mathcal{P}, \rho)\), which implies that \((x_k, u_k)\) converge to \((x, u)\) in \((B \times U, d)\).

It remains to verify that \(f'(x; u)\) exists and that \(f'(x_k; u_k) \to f'(x; u)\). The fact that

\[
\lim_{t \to 0} \frac{f_{x_k, u_k}(t)}{t} = f'(x_k; u_k)
\]

exists for each \(k\) and the condition (13), which says that \(f_{x_k, u_k}(t)/t\) converge to \(f_{x, u}(t)/t\) uniformly, imply that the limit

\[
\lim_{t \to 0} \frac{f_{x, u}(t)}{t} = f'(x; u)
\]

exists, and is equal to \(\lim_{k \to \infty} f'(x_k; u_k)\). So \((x, u) \in M\) and the function \((x, u) \mapsto f'(x; u)\) is \(d\)-continuous on \(M\). \(\Box\)

The three statements of the next lemma relate some ordinary topological or metric notions in \(B \times U\) to those coming from the metric \(\rho\). The first two are quite natural, the third is rather technical and geared toward its use in the final stages of the proof.

**Lemma 5.3.** The metric \(\rho\) has the following three properties, the latter two with a suitable constant \(C \in (0, \infty)\):

(i) For any \((x, u) \in B \times U\), the function \(y \mapsto \rho(f_{x, u}, f_{y, u})\) is norm lower semicontinuous on \(B\).

(ii) For any \((x, u), (x, v) \in B \times U\), \(\rho(f_{x, u}, f_{x, v}) \leq C \|u - v\|\).

(iii) For every \(x, y, z \in B\), \(u \in U\) and \(\delta > 0\),

\[
\rho(f_{x, u}, f_{y, u}) \leq \max \left\{ \rho(f_{x, u}, f_{z, u}) + \frac{C}{\delta} \|y - z\|, \right. \\
\left. \sup_{0 < |t| < \delta} \inf_{w \in U} \left( C \|u - w\| + \left| \frac{f_{x, u}(t) - f_{y, u}(t)}{t} \right| \right) \right\}.
\]

**Proof.** Notice that for any fixed \(t\), the function

\[
y \mapsto \frac{f_{x, u}(t) - f_{y, u}(t)}{|t|}
\]

is continuous. Supremum of continuous functions is lower semicontinuous and we are done.

(ii) Obvious by taking \(C \geq \text{Lip}(f)\).
If $|t| \geq \delta$ then
\[ |f_{x,u}(t) - f_{y,u}(t)| \leq \frac{|f_{x,u}(t) - f_{z,u}(t)|}{|t|} + \frac{2\text{Lip}(f)}{|t|} \|y-z\| \]
\[ \leq \varrho(f_{x,u}, f_{z,u}) + \frac{2\text{Lip}(f)}{\delta} \|y-z\|. \]

If $|t| < \delta$, then for every $w \in U$,
\[ |f_{x,u}(t) - f_{y,u}(t)| \leq \text{Lip}(f) \|u-w\| + |f_{x,u}(t) - f_{y,w}(t)| \]
\[ \leq \varepsilon_0 \lambda_0 \|u-w\| + |f_{x,u}(t) - f_{y,w}(t)|. \]

Hence the statement holds for any constant $C \geq 2\text{Lip}(f)$. \hfill \Box

**Use of the variational principle.** Choose $\varepsilon_0 > 0$ so that $f'(x_0; u_0) \leq \varepsilon_0 + \inf\{f'(x; u) | (x, u) \in M\}$ and find $\lambda_0 > \varepsilon_0$ so large that
\[ \frac{\varepsilon_0}{\lambda_0} < \frac{r}{2} \quad \text{and} \quad \Theta(z) > \frac{\varepsilon_0}{\lambda_0} \quad \text{for } \|z\| \geq \frac{r}{2}. \]

Denote
\[ \kappa = (1 + \text{Lip}(f) + 2\lambda_0 \text{Lip}(\Theta|B(0, r_0)))(2 + r). \]

We will also choose a constant $C > 0$ for which the estimates of Lemma 5.3 [ii] and [iii] hold. Further, let
\[ \sigma_0 = \min\left\{\lambda_0, \frac{1}{20\kappa}\right\} \]
and for $i \geq 1$ define
\[ \sigma_i = \sigma_{i-1}^3, \quad \lambda_i = \sigma_i \quad \text{and} \quad \varepsilon_i = \sigma_i^3. \]

We will often use the following (very rough) estimate:
\[ \sum_{i=1}^{\infty} \sigma_i \leq \frac{2}{3} \sigma_0. \]

To define our last sequence of parameters, $s_i$, we start by letting
\[ t_i = \inf\{t > 0 | \Theta(z) > \sigma_i^2 \text{ for } \|z\| > t\} \]
and observe that $t_i$ is a decreasing sequence, $0 < t_i \leq \frac{1}{2}r$ by the choice of $r$, and $\Theta(z) > \sigma_i^2 = \varepsilon_i/\lambda_i$ for $\|z\| > t_i$. Moreover, $\lim_{t \to \infty} t_i > 0$, since for every $s > 0$ there is $j$ such that $\Theta(z) > \sigma_j^2$ for $\|z\| > s$, hence $t_i \leq s$ for $i \geq j$. It follows that
\[ s_i := \max\{(1 + 10C)t_i, 5(\sigma_0 + \kappa)\sigma_i/\sigma_0, 990\sqrt{\kappa \varepsilon_i}\}. \]

form a sequence decreasing to zero such that
\[ s_0 \geq \max\{5\kappa, 495\sqrt{2\kappa \varepsilon_0}\}. \]
and for $i \geq 1$,

$$s_i \geq \max\{5\sigma_i, 990, \sqrt{K\epsilon_i}\} \quad \text{and} \quad \Theta(z) > \frac{\delta_i}{\lambda_i} \quad \text{for} \quad \|z\| \geq \frac{s_i}{10C}.$$  

We intend to use the variational principle of Theorem 2.4 with the perturbation functions $F_i : M \times M \to [0, \infty)$,

$$F_i((x, u), (y, v)) = \Phi_i(x, y) + \Psi_i(u, v) + Q_i((x, u), (y, v)) + \Delta_i((x, u), (y, v)),$$

where

$$\Phi_i(x, y) = \lambda_i \|x - y\|, \quad \Psi_i(u, v) = \lambda_i \Theta(u - v),$$

$$Q_i((x, u), (y, v)) = \sigma_i (f'(x; u) - f'(y; v))^2$$

and

$$\Delta_i((x, u), (y, v)) = \min\{\lambda_i, \max\{0, \varrho(f_x, u, f_y, v) - s_i\}\}.$$  

Notice that the peculiarity in the definition of $\Delta_i$ is not a misprint: $\Delta_i$ really does not depend on $v$.

We have to verify the assumptions of Theorem 2.4.

**Lemma 5.4.** The functions $F_i$ are non-negative, $d$-lower semicontinuous in the second variable, satisfy $F_i((x, u), (x, u)) = 0$ and there are $r_i \searrow 0$ such that

$$\inf\{F_i((x, u), (y, v)) \mid d((x, u), (y, v)) \geq r_i\} > 0.$$

**Proof.** Clearly, $F_i \geq 0$ and $F_i((x, u), (x, u)) = 0$. Lower semicontinuity of the functions $\Delta_i$ in the second variable follows directly from Lemma 5.3(ii). Since the remaining functions from which $F_i$ consists are continuous (the functions $Q_i$ by Lemma 5.2), this shows the lower semicontinuity of $F_i$ in the second variable.

For the last statement, let $r_i = (C + 2)s_i$. We show that

$$\inf\{F_i((x, u), (y, v)) \mid d((x, u), (y, v)) \geq r_i\} \geq \min\{\lambda_i, s_i, \lambda_i s_i, \inf_{\|z\| \geq s_i} \lambda_i \Theta(z)\}. \quad (16)$$

Recall that

$$d((x, u), (y, v)) = \max\{\|x - y\|, \|u - v\|, \varrho(f_x, u, f_y, v)\}.$$  

The estimate (16) is obvious if $\|x - y\| \geq s_i$ or $\|u - v\| \geq s_i$. If $\|x - y\| < s_i$ and $\|u - v\| < s_i$, then

$$\varrho(f_x, u, f_y, v) \geq (C + 2)s_i.$$  

Lemma 5.3(ii) implies that

$$\varrho(f_x, u, f_y, v) \geq \varrho(f_x, u, f_y, v) - \varrho(f_y, u, f_y, v) \geq (C + 2)s_i - C\|u - v\| \geq 2s_i.$$

Hence $\Delta_i((x, u), (y, v)) \geq \min\{\lambda_i, s_i\}$. \qed
By the variational principle, \((x_0, u_0)\) is the starting term of a sequence \((x_j, u_j) \in M\) which \(d\)-converges to some \((x_\infty, u_\infty) \in M\) and has the properties that, denoting \(\varepsilon_\infty = 0\) and

\[
h_i(x, u) = f'(x; u) + \sum_{j=0}^{i-1} F_j((x_j, u_j), (x, u)),
\]

we have

\[
h_\infty(x_\infty, u_\infty) \leq \min\{h_i(x_i, u_i), \varepsilon_i + \inf_{(x, u) \in M} h_i(x, u)\}
\]

for \(0 \leq i \leq \infty\). Notice that for \(i = \infty\) this inequality is just a complicated way of saying that \(h_\infty\) attains its minimum on \(M\) at \((x_\infty, u_\infty)\).

Recalling the definition of the functions \(F_i\), we have

\[
h_\infty(x, u) = f'(x; u) + \Phi(x) + \Psi(u) + Q(x, u) + \Delta(x),
\]

where

\[
\begin{align*}
\Phi(x) &= \sum_{i=0}^{\infty} \Phi_i(x_i, x) = \sum_{i=0}^{\infty} \lambda_i \|x_i - x\|, \\
\Psi(u) &= \sum_{i=0}^{\infty} \Psi_i(u_i, u) = \sum_{i=0}^{\infty} \lambda_i \|u_i - u\|, \\
Q(x, u) &= \sum_{i=0}^{\infty} Q_i((x_i, u_i), (x, u)) = \sum_{i=0}^{\infty} \sigma_i (f'(x_i; u_i) - f'(x; u))^2, \\
\Delta(x) &= \sum_{i=0}^{\infty} \Delta_i((x_i, u_i), (x, u));
\end{align*}
\]

to justify the last definition we recall the independence of \(\Delta_i\) from the last variable. Observe that all these functions are positive and finite, \(\Phi\), \(\Psi\) and \(Q\) are \(d\)-continuous on \(M\) (for \(Q\) recall Lemma 5.2), and \(\Psi\) is everywhere upper Fréchet differentiable.

Since we already know that \(f'(x_\infty; u_\infty) \leq f'(x_0; u_0)\) by (18) with \(i = 0\), the proof of the proposition will be completed once we show that \(f\) is Fréchet differentiable at \(x_\infty\).

We first collect several estimates, all easily following from (18), that will be needed in what follows.

**Lemma 5.5.** \(\|x_\infty - x_0\| < \frac{1}{2}\varepsilon_0, \|u_\infty - u_0\| < \frac{1}{2}\varepsilon_0\) and for each \(0 \leq i < \infty\), \(\|u_i - u_\infty\| \leq \frac{1}{2}\varepsilon_i\) and \(f'(x_\infty; u_\infty) - f'(x_i; u_i)\) \(\leq \frac{1}{2}\varepsilon_i\).

**Proof.** The definition of \(h_i\) in (17) and the condition (18) imply

\[
h_i(x_\infty, u_\infty) + \Phi_i(x_i, x_\infty) + \Psi_i(u_i, u_\infty) + Q_i((x_i, u_i), (x_\infty, u_\infty)) \\
\leq h_\infty(x_\infty, u_\infty) \leq \varepsilon_i + h_i(x_\infty, u_\infty).
\]

Hence, for \(i = 0\) we obtain

\[
\|x_\infty - x_0\| = \frac{1}{\lambda_0} \Phi_0(x_0, x_\infty) \leq \frac{\varepsilon_0}{\lambda_0} < \frac{r}{2}.
\]
and
\[ \Theta(u_0 - u_\infty) = \frac{1}{\lambda_0} \Psi_0(u_0, u_\infty) \leq \varepsilon_0/\lambda_0, \]
which implies that \( \|u_\infty - u_0\| < \frac{1}{2} r \). For the case of general index \( i \) we deduce that
\[ \Theta(u_i - u_\infty) \leq \varepsilon_i/\lambda_i, \]
and the choice of \( s_i \) implies that \( \|u_\infty - u_i\| < \frac{1}{10} \varepsilon_i s_i \). By the choice of \( s_0 \) we have
\[ |f'(x_\infty; u_\infty) - f'(x_0; u_0)| \leq \text{Lip}(f)(\|u_\infty\| + \|u_0\|) \leq \text{Lip}(f)(2 + r) \leq \kappa \leq s_0/5. \]
Finally, for \( i \geq 1 \) we infer from (20) that
\[ (f'(x_\infty; u_\infty) - f'(x_i; u_i))^2 = \frac{1}{\sigma_i} Q_i((x_i, u_i), (x_\infty, u_\infty)) \leq \frac{\varepsilon_i}{\sigma_i} = \sigma_i^2 \leq \left( \frac{\varepsilon_i}{\lambda_i} \right)^2. \]
Assuming that \( f'(x_\infty) \) exists, we easily guess its value from the fact that the function \( H(u) = h_\infty(x_\infty, u) \) attains its minimum at \( u = u_\infty \). Let \( L_{\psi} \) be an upper Fréchet derivative of \( \Psi \) at \( u_\infty \). By (19), \( H \) is upper differentiable at \( u = u_\infty \), with upper derivative
\[ f'(x_\infty) + L_{\psi} + \varkappa f'(x_\infty), \]
where
\[ \varkappa = 2 \sum_{i=0}^{\infty} \sigma_i (f'(x_\infty; u_\infty) - f'(x_i; u_i)). \]
Since the upper derivative of \( H \) at \( u_\infty \) is equal to zero, this gives
\[ f'(x_\infty) = -\frac{L_{\psi}}{1 + \varkappa}. \]
Recalling that \( |f'(x_\infty; u_\infty) - f'(x_i; u_i)| \leq \sigma_i, i \geq 1 \), we find with the help of (15) that
\[ |\varkappa| \leq 2 \sum_{i=0}^{\infty} \sigma_i^2 \leq 2 \left( \sigma_0 \kappa + \sum_{i=1}^{\infty} \sigma_i^2 \right) \leq 2 \left( \sigma_0 \kappa + \frac{2}{3} \sigma_0 \right) \leq 2 \left( \frac{1}{20} + \frac{2}{3} \right) = \frac{1}{6}. \]
Hence, if we put \( L := -L_{\psi}/(1 + \varkappa) \), the functional \( L \) is well-defined. Since all \( u_i \) belong to \( U \), i.e. \( \|u_i - u_0\| \leq r \), we get the estimate of the norm of \( L \):
\[ \|L\| \leq \frac{6}{5} \|L_{\psi}\| \leq \frac{6}{5} \text{Lip}(\Psi|_U) \leq \frac{6}{5} \sum_{i=0}^{\infty} \lambda_i \text{Lip}(\Theta|_{B(0, r_0)}) \]
\[ = \frac{6}{5} \left( \lambda_0 + \sum_{i=1}^{\infty} \sigma_i \right) \text{Lip}(\Theta|_{B(0, r_0)}) \leq 2 \lambda_0 \text{Lip}(\Theta|_{B(0, r_0)}). \]
We are now ready to prove that, indeed, \( f'(x_\infty) = L \). Before embarking on the main part of the proof, we make a simple but important observation: The above heuristic argument is correct when restricted to the direction \( u_\infty \), giving that the derivative of \( f \) at the point \( x_\infty \) in the direction \( u_\infty \) agrees with the value of the linear form \( L \) at \( u_\infty \).
Lemma 5.6. $f'(x_\infty; u_\infty) = L(u_\infty)$.

Proof. By Lemma 5.5, the vector $su_\infty$ belongs to $U$ for $s \in (1 - \tau, 1 + \tau)$ and suitable $0 < \tau < 1$, so the pair $(x_\infty, su_\infty)$ is in $M$. Hence the function $\psi: (1 - \tau, 1 + \tau) \to \mathbb{R}$,

$$
\psi(s) = f'(x_\infty; su_\infty) + \Phi(x_\infty) + \Psi(su_\infty) + \Delta(x_\infty)
+ \sum_{i=0}^{\infty} \sigma_i(f'(x_\infty; su_\infty) - f'(x_i; u_i))^2
$$

attains its minimum at $s = 1$. Since $\psi$ is upper differentiable at $s = 1$ with upper derivative $f'(x_\infty; u_\infty) + \kappa f'(x_\infty; u_\infty)$, we get

$$
f'(x_\infty; u_\infty) + \kappa f'(x_\infty; u_\infty) = 0,
$$

yielding

$$
f'(x_\infty; u_\infty) = -\kappa f'(x_\infty; u_\infty) = Lu_\infty
$$
as required. \qed

Final step. The rest of the proof is devoted to fulfilling the above stated goal, that is, to showing that $f$ is Fréchet differentiable at $x_\infty$ and $f'(x_\infty) = L$. We will argue by contradiction, and so from now on we make the following

Assumption. There is $\eta > 0$ such that for every $\delta > 0$ there is $v \in X$ with $0 < \|v\| < \delta$

and

$$
|f(x_\infty + v) - f(x_\infty) - L(v)| > \eta \|v\|. \tag{22}
$$

We intend to use this assumption to find a vector $v$ satisfying (22) whose norm is so small that it will beat various error estimates. To quantify these estimates, we start by finding $0 < \delta_0 < \frac{1}{2} \varepsilon_0$ so that

(a) for every $\|u - u_\infty\| \leq \delta_0$,

$$
\Psi(u) \leq \Psi(u_\infty) + L\psi(u - u_\infty) + \frac{1}{2}\varepsilon_0 \|u - u_\infty\|.
$$

Then we choose an index $k > 0$ so large that

(b) $\beta := \frac{78 \varepsilon_k}{\eta} \leq \min\left\{ \frac{\delta_0}{2}, \frac{r}{10C + 5\|L\|} \right\}$ and $10\sigma_0 \|L\|^2 \beta^2 \leq \varepsilon_k$.

Finally, we denote $\tau = 6\varepsilon_k$ and find $0 < \delta < \frac{r}{2(1 + \tau)}$ so that

(c) for every $\|x - x_\infty\| \leq (\|u_\infty\| + \beta)\delta$,

$$
\Phi(x) < \Phi(x_\infty) + \frac{1}{2}\varepsilon_k;
$$

(d) for each $i = 0, 1, \ldots, k$, for $i = \infty$ and for every $|t| \leq \delta$,

$$
|f(x_i + tu_i) - f(x_i) - tf'(x_i; u_i)| \leq \min\left\{ \frac{s_k}{5}, \frac{\tau^2}{\varepsilon_k} \right\} |t|.
$$
Let \( v \in X \) be a vector with
\[
0 < \|v\| < \delta \beta \min \left\{ \frac{1}{2}, \frac{\varepsilon_k}{2(k+1)C(\beta + \|u_\infty\|)} \right\}
\]
for which (22) holds. Let \( a = \frac{1}{\beta} \|v\| \) and define \( \gamma : [-\delta, \delta] \to X \) by
\[
\gamma(t) = x_\infty + tu_\infty + \max\{0, 1 - |t|/a\}v.
\]
We notice that \( \|\gamma(t) - x_\infty\| \leq \delta \|u_\infty\| + \delta \beta \leq \delta (1 + r/2 + r/2) < r/2 \), hence by Lemma 5.5,
\[
\|\gamma(t) - x_0\| \leq \|\gamma(t) - x_\infty\| + \|x_\infty - x_0\| < r.
\]
It follows that the curve \( \gamma \) is contained in the interior of the ball \( B \). Further, let \( g, h : [-\delta, \delta] \to \mathbb{R} \) be defined by
\[
g(t) = -f(x_\infty + tu_\infty) + L(tu_\infty) \quad \text{and} \quad h(t) = -f(\gamma(t)) + L(\gamma(t) - x_\infty).
\]
Then the assumptions of Corollary 4.3 hold: clearly, \( a \leq \delta/2 \) and the inequalities
\[
|h(0) - g(0)| > \eta \|v\| = 13 \tau \varepsilon_k \quad \text{and} \quad |g(t) - g(0)| \leq \tau^2 / \kappa \quad \text{for } |t| \leq \delta
\]
follow directly from (22), (\( \beta \)) and from (\( \delta \)) with \( i = \infty \). Also,
\[
\tau = 6 \varepsilon_k \leq 6 \varepsilon_0 \leq \kappa.
\]
The remaining assumption is that \( h \) is \( \kappa \)-Lipschitz. Using (21),
\[
\operatorname{Lip}(h) \leq \operatorname{Lip}(f) \operatorname{Lip}(\gamma) + \|L\| \operatorname{Lip}(\gamma) \leq (\operatorname{Lip}(f) + \|L\|) (\|u_\infty\| + \beta)
\]
\[
\leq (\operatorname{Lip}(f) + 2\lambda_0 \operatorname{Lip}(\Theta|B(0,r_0))))(1 + r) \leq \kappa.
\]
So we may apply Corollary 4.3 to find \( \xi \in (-a, a) \setminus \{0\} \) such that
(a) \( h \) is differentiable at \( \xi \);
(b) \( h'(\xi) \geq 6 \varepsilon_k \);
(c) \( |h(t) - h(\xi)| \leq 99 \sqrt{K h'(\xi)} |t - \xi| \) for every \( t \in [-\delta, \delta] \).
Let \( x = \gamma(\xi) \) and \( u = \gamma'(\xi) \); so \( u = u_\infty \pm \beta v / \|v\| \).

Lemma 5.7. \( (x, u) \in M \).

Proof. We have already proved that \( \|x - x_0\| = \|\gamma(\xi) - x_0\| < r \) and, using the fact that \( \|u - u_\infty\| = \beta \leq 1/4r \), we see that also
\[
\|u - u_0\| \leq \|u_0 - u_\infty\| + \|u - u_\infty\| < r.
\]
Moreover, since \( h'(\xi) \) exists we find that \( f'(x; u) \) exists as well and
\[
f'(x; u) = -h'(\xi) + Lu.
\]
This says that the pair \( (x, u) \) belongs to \( M \). \( \square \)

We wish to show that \( h_\infty(x, u) < h_\infty(x_\infty, u_\infty) \) to contradict that \( h_\infty \) attains its minimum at \( (x_\infty, u_\infty) \). For this, we first compare the terms not containing the function \( \Delta \). Let us denote them by
\[
G(y, v) = f'(y; v) + \Phi(y) + \Psi(v) + Q(y, v).
\]
Lemma 5.8. \(G(x, u) < -\frac{1}{2}h'(\xi) + G(x_\infty, u_\infty).\)

Proof. We estimate all three terms \(Q(x, u), \Psi(u)\) and \(\Phi(x)\) in \(G(x, u).\) With the help of Lemma 5.6 we have

\[
Q(x, u) = \sum_{i=0}^{\infty} \sigma_i [f'(x_i; u_i) - f'(x_\infty; u_\infty) + f'(x_\infty; u_\infty) - f'(x; u)]^2
\]

\[
= Q(x_\infty, u_\infty) + \kappa(f'(x; u) - f'(x_\infty; u_\infty)) + \sum_{i=0}^{\infty} \sigma_i (f'(x_i; u) - f'(x_\infty; u_\infty))^2
\]

\[
= Q(x_\infty, u_\infty) + \kappa(-h'(\xi) + L(u - u_\infty)) + \sum_{i=0}^{\infty} \sigma_i (-h'(\xi) + L(u - u_\infty))^2
\]

\[
\leq Q(x_\infty, u_\infty) - \kappa h'(\xi) + \kappa L(u - u_\infty) + \frac{10}{3} \sigma_0 (h'(\xi))^2 + \frac{10}{3} \sigma_0 (L(u - u_\infty))^2.
\]

Since \(h\) is \(\kappa\)-Lipschitz the choice of \(\sigma_0\) gives

\[
\frac{10}{3} \sigma_0 h'(\xi) \leq \frac{10}{3} \sigma_0 \kappa \leq \frac{1}{6}.
\]

Also \(|\kappa| \leq \frac{1}{6}\) and \(\frac{10}{3} \sigma_0 L(u - u_\infty)^2 \leq \frac{10}{3} \sigma_0 \|L\|^2 \beta^2 \leq \frac{1}{6} \varepsilon_k\). This allows us to continue the above estimate to obtain

\[
Q(x, u) \leq Q(x_\infty, u_\infty) + \kappa L(u - u_\infty) + \frac{1}{6} h'(\xi) + \frac{1}{6} \varepsilon_k.
\]

Since \(\|u - u_\infty\| = \beta \leq \delta_0\), we see from (5) that

\[
\Psi(u) \leq \Psi(u_\infty) + L \psi(u - u_\infty) + \frac{1}{128 \eta} \|u - u_\infty\|
\]

\[
= \Psi(u_\infty) - (1 + \kappa) L(u - u_\infty) + \frac{1}{128 \eta} \beta
\]

\[
= \Psi(u_\infty) - (1 + \kappa) L(u - u_\infty) + \frac{1}{6} \varepsilon_k.
\]

Finally, the last needed inequality \(\Phi(x) < \Phi(x_\infty) + \frac{1}{6} \varepsilon_k\) follows from (4), since we have \(\|x - x_\infty\| \leq (\|u_\infty\| + \beta) \delta\). Using all three just established estimates of \(Q(x, u), \Psi(u)\) and \(\Phi(x)\) and recalling that \(\varepsilon_k \leq \frac{1}{6} h'(\xi)\), we get

\[
G(x, u) = f'(x; u) + \Phi(x) + \Psi(u) + Q(x, u)
\]

\[
< f'(x; u) + \Phi(x_\infty) + \Psi(u_\infty) + Q(x_\infty, u_\infty) - L(u - u_\infty) + \frac{1}{6} h'(\xi) + \varepsilon_k
\]

\[
= -\frac{1}{2} h'(\xi) + Lu + \Phi(x_\infty) + \Psi(u_\infty) + Q(x_\infty, u_\infty) - L(u + f'(x_\infty; u_\infty))
\]

\[
+ \frac{1}{6} h'(\xi) + \varepsilon_k
\]

\[
= -\frac{1}{2} h'(\xi) + G(x_\infty, u_\infty) + \varepsilon_k \leq -\frac{1}{2} h'(\xi) + G(x_\infty, u_\infty).
\]

The next lemma provides us with a certain control of the function \(\Delta\).

Lemma 5.9. For each \(0 \leq i \leq k\),

\((A_i)\) \[495 \sqrt{k} h'(\xi) \leq \xi_i;\]

\((B_i)\) \[\Delta_i((x_i, u_i), (x, u)) \leq \Delta_i((x_i, u_i), (x_\infty, u_\infty)) + \frac{1}{k+1} \varepsilon_k.\]
Proof. We first prove that for each $i$, $0 \leq i \leq k$, (A) implies (B). By the definition of $\Delta_i$ it suffices to show that

$$q(f_{x_i,u_i}, f_{x,u}) \leq \max \left\{ s_i, q(f_{x_i,u_i}, f_{x_u,u}) + \frac{1}{k+1} \varepsilon_k \right\}. \quad (23)$$

By Lemma [5.3] with $\delta/2$ instead of $\delta$ we have

$$q(f_{x_i,u_i}, f_{x,u}) \leq \max \left\{ q(f_{x_i,u_i}, f_{x_u,u}) + \frac{2C}{\delta} \|x - x_u\|, \sup_{0 < |t| < \delta/2} \inf_{w \in U} \left\{ C\|u_i - w\| + \left| \frac{f_{x,w}(t) - f_{x_i,u_i}(t)}{t} \right| \right\} \right\}. \quad (24)$$

The choice of $\|v\|$ shows that

$$\|x - x_u\| \leq a\|u\| + \|v\| = \left( \frac{1}{\beta} \|u\| + 1 \right) \|v\| \leq \frac{\delta \varepsilon_k}{2(k+1)C}.$$

So the first term in the maximum is at most $q(f_{x_i,u_i}, f_{x_u,u}) + \frac{1}{k+1} \varepsilon_k$. To estimate the second term, we consider any $0 < |t| < \delta/2$ and estimate the expression

$$C\|u_i - w\| + \left| \frac{f_{x,w}(t) - f_{x_i,u_i}(t)}{t} \right|$$

for $w = \frac{1}{t}(\gamma(\xi + t) - \gamma(\xi))$. (Notice that $\xi + t$ still belongs to the interval $(-\delta, \delta)$ where $\gamma$ is defined.) Our goal is to show that it is smaller than $s_i$, since then (23) holds and in fact this term does not contribute to the value of $\Delta_i((x_i, u_i), (x_u, u_u))$.

Observe first that by the definition of $\gamma$, $\|u_u - w\| \leq \|v\|/a = \beta$. Together with Lemma [5.3] this gives

$$C\|u_i - w\| \leq C\|u_i - u_u\| + C\|u_u - w\| \leq \frac{1}{\beta} s_i + C \beta \leq \frac{1}{\beta} s_i.$$

The second summand of (24) is controlled as follows:

$$\left| \frac{f_{x,w}(t) - f_{x_i,u_i}(t)}{t} \right| \leq \left| \frac{f_{x,w}(t)}{t} + L(w) \right| + |L(w) + f'(x_u; u_u)| + |f'(x_u; u_u) - f'(x; u)| + \left| \frac{f'(x; u) - f_{x_i,u_i}(t)}{t} \right|.$$

We will estimate each of the four terms in the sum above. By (A),

$$\left| \frac{f_{x,w}(t)}{t} + L(w) \right| = \frac{|h(\xi + t) - h(\xi)|}{|t|} \leq 99 \sqrt{k h'(\xi)} \leq \frac{1}{\beta} s_i;$$

the choice of $\beta$ guarantees

$$|L(w) + f'(x_u; u_u)| = |L(w) - L(u_u)| \leq \|L\| \beta \leq \frac{1}{\beta} s_i;$$

$$|f'(x; u) - f_{x_i,u_i}(t)| \leq |f'(x; u) - f'(x_i; u)| + |f'(x_i; u) - f_{x_i,u_i}(t)| \leq \frac{1}{\beta} s_i + C \beta \leq \frac{1}{\beta} s_i;$$

and

$$\left| \frac{f_{x_i,u_i}(t)}{t} \right| \leq \frac{1}{\beta} s_i.$$
the estimate
\[ |f'(x_\infty; u_\infty) - f'(x_i; u_i)| \leq \frac{1}{2}s_i \]
follows from Lemma 5.5 and finally (6) yields
\[ \left| \frac{f_{x_i,u_i}(t)}{t} - f'(x_i; u_i) \right| \leq \frac{1}{2}s_i. \]
Adding these inequalities we get
\[ C\|u_i - w\| + \left| \frac{f_{x,w}(t)}{t} - f_{x_i,u_i}(t) \right| \leq s_i. \]

Hence the second term in the maximum is at most \( s_i \) and (B) follows.

To prove (A), we infer from Lemma 5.8 and (18) that
\[ \frac{1}{2}h'(\xi) \leq G(x_\infty, u_\infty) - G(x, u) = h_\infty(x_\infty, u_\infty) - \Delta(x_\infty, u_\infty) - G(x, u) \]
\[ \leq \varepsilon_i + h_i(x, u) - \Delta(x_\infty, u_\infty) - G(x, u) \]
\[ \leq \varepsilon_i + \sum_{j=0}^{i-1} \left( \Delta_j((x_j, u_j), (x, u)) - \Delta_j((x_j, u_j), (x_\infty, u_\infty)) \right). \]

With \( i = 0 \) this gives \( h'(\xi) \leq 2\varepsilon_0 \), hence
\[ 495\sqrt{k h'(\xi)} \leq 495\sqrt{2k \varepsilon_0} \leq s_0 \]
by the choice of \( s_0 \). Proceeding by induction, we infer from the first part of this proof that
\[ \Delta_j((x_j, u_j), (x, u)) - \Delta_j((x_j, u_j), (x_\infty, u_\infty)) \leq \frac{\varepsilon_k}{k+1} \]
for \( j < i \).

Hence we conclude that
\[ h'(\xi) \leq 2\varepsilon_i + 2\sum_{j=0}^{i-1} \frac{\varepsilon_k}{k+1} \leq 4\varepsilon_i, \]
which, by the choice of \( s_i \), gives
\[ 495\sqrt{k h'(\xi)} \leq 990\sqrt{k \varepsilon_i} \leq s_i. \]

**Corollary 5.10.** \( \Delta(x, u) \leq 3\varepsilon_k + \Delta(x_\infty, u_\infty) \).

**Proof.** First recall that the functions \( \Delta_i \) are bounded by \( \lambda_i \) by their very definition. Further, the choice of \( \varepsilon_i \) and \( \lambda_i \) implies that for \( i \geq 1 \), \( \lambda_i = \varepsilon_{i-1} \) and \( \varepsilon_i \leq \frac{1}{2}\varepsilon_{i-1} \). It follows that for \( i > k \),
\[ \Delta_i((x_j, u_j), (x, u)) \leq \lambda_i = \varepsilon_{i-1} \leq 2^{k-i+1}\varepsilon_k. \]
Hence the previous lemma gives
\[
\Delta(x, u) = \sum_{i=0}^{k} \Delta_i((x_i, u_i), (x, u)) + \sum_{i=k+1}^{\infty} \Delta_i((x_i, u_i), (x, u))
\]
\[
\leq \sum_{i=0}^{k} \left( \Delta_i((x_i, u_i), (x_{\infty}, u_{\infty})) + \frac{\varepsilon_k}{k+1} \right) + 2\varepsilon_k
\]
\[
\leq \sum_{i=0}^{\infty} \Delta_i((x_i, u_i), (x_{\infty}, u_{\infty})) + 3\varepsilon_k = \Delta(x_{\infty}, u_{\infty}) + 3\varepsilon_k.
\]

Adding the inequalities from Lemma 5.8 and Corollary 5.10 and using that \( h'(\xi) \geq 6\varepsilon_k \) leads to
\[
h_{\infty}(x, u) = G(x, u) + \Delta(x, u)
\]
\[
< \left( -\frac{1}{2} h'(\xi) + G(x_{\infty}, u_{\infty}) \right) + (\Delta(x_{\infty}, u_{\infty}) + 3\varepsilon_k)
\]
\[
= -\frac{1}{2} h'(\xi) + 3\varepsilon_k + h(x_{\infty}, u_{\infty}) \leq h_{\infty}(x_{\infty}, u_{\infty}).
\]

Recalling that \( h_{\infty} \) attains its minimum at \( (x_{\infty}, u_{\infty}) \) and noticing the strict inequality on the second line, we have the desired contradiction.

6. Generalizations and extensions

In this section we announce several results which generalize Theorem 1.1 and which will be proved in the book [6].

There is a notion of a monotone real-valued function on a Banach space. Let \( X \) be a Banach space with a closed cone \( K \) with vertex 0 and non-empty interior. A function \( f : X \to \mathbb{R} \) is called \textit{cone monotone} (with respect to the cone \( K \)) if \( y \in x + K \) implies \( f(y) \geq f(x) \). This notion goes back to the early 20th century. In [3] this notion was considered from the point of view of differentiability. The authors proved that such a function on a separable space \( X \) is a.e. Gâteaux differentiable. They also raise the question of existence of Fréchet derivatives of such functions if \( X \) is Asplund. The proof given above in Section 5 allows us to prove that indeed every cone monotone function on an Asplund space has points of Fréchet differentiability and that also a natural variant of the mean value theorem for Fréchet derivative holds in this context. The only change in the proof given here concerns the metric used in the space \( M \). In the monotone case we use a metric which is related to the one used by Skorokhod in a different context. This makes the proof a little more involved.

Theorem 1.1 is not an a.e. result. Therefore it does not imply that any two Lipschitz functions on \( X \) have a common point of Fréchet differentiability (or equivalently, that a Lipschitz map from \( X \) to \( \mathbb{R}^2 \) has a point of differentiability). We were however able to prove that any Lipschitz map from a Hilbert space \( X \) into \( \mathbb{R}^2 \) has points of Fréchet differentiability and that even an appropriate mean value theorem for Fréchet derivatives holds. We do not know whether every Lipschitz map from a Hilbert space into \( \mathbb{R}^3 \) has
points of Fréchet differentiability. An example appearing in [10] (or in a refined and stronger form in [6]) shows however that in this case there is no longer a mean value theorem for Fréchet derivatives.

It is easy to see that if the function $\Theta$ in Proposition 5.1 is assumed to be locally Lipschitz and everywhere Gâteaux differentiable our proof provides a point of Gâteaux differentiability for every Lipschitz function. Of course, there are better Gâteaux differentiability results for separable spaces, but this statement has its interest in the non-separable situation, when it was proved before only on spaces with Gâteaux smooth norms.

An everywhere Gâteaux differentiable function $f : X \to \mathbb{R}$ is necessarily Lipschitz on a non-empty open set. Thus in the illustrative example of Section 3 the assumption that $f$ be Lipschitz was redundant. More generally, if $f : X \to Y$ satisfies

$$\limsup_{t \to 0} \frac{\|f(x + tu) - f(x)\|}{|t|} < \infty$$

for every $x$ in a non-empty open subset of a Banach space and for every unit vector $u$, then it is Lipschitz on a non-empty open set. This can be proved by a standard application of the Baire Category Theorem.

Acknowledgments. The research of J. Tišer was supported by grants GAČR 201/07/0394 and MSM 6840770010.

References