



J. Lindenstrauss · D. Preiss · J. Tišer

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

$$f'(x; u) = \lim_{t \searrow 0} \frac{f(x + tu) - f(x)}{t} \quad (1)$$

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

$$f(x + u) = f(x) + f'(x)u + o(\|u\|), \quad \|u\| \rightarrow 0. \quad (2)$$

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

$$f(x) = \sin x, \quad x \in L_2[0, 1],$$

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.

J. Lindenstrauss: Institute of Mathematics, The Hebrew University, Jerusalem 91904, Israel;
e-mail: joram@math.huji.ac.il

D. Preiss: Department of Mathematics, University of Warwick, Coventry CV4 7AL,
United Kingdom; e-mail: d.preiss@warwick.ac.uk

J. Tišer: Department of Mathematics, Technical University, 166 27 Praha, Czech Republic;
e-mail: tiser@math.feld.cvut.cz

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 $\|\chi_E\| \rightarrow 0$.

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 $f: G \rightarrow \mathbb{R}$ 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 Banach space X is called an *Asplund space* if the dual of every separable subspace of X is again separable. In particular, every reflexive Banach space is Asplund. The most important special case of Theorem 1.1 is where X is a Hilbert space. The proof of this special case is as hard as the general case. The “moreover” part of Theorem 1.1 represents a mean value theorem for Fréchet derivatives.

The proof in [9] as well as the somewhat simpler proof in [5] involve a delicate iterative construction of a sequence of points which are shown to converge to a point of Fréchet differentiability. This method reminds one of proofs of variational principles. In fact, Fréchet differentiability results are related to the natural domain of variational principles (i.e. the existence of points where some functions attain their maximum). If $f: X \rightarrow Y$ is Lipschitz then it is known that

$$\text{Lip}(f) = \sup \|f'(x)\|$$

where the supremum is taken over all points x where f is Gâteaux differentiable (it is assumed that X is separable and Y has RNP). If there exists a point x_0 such that $\text{Lip}(f) = \|f'(x_0)\|$ then (at least in the case of uniformly smooth space X) it is easy to show that f is Fréchet differentiable at x_0 . However, the existing variational principles (e.g. in [2], [4]) do not seem to apply in our situation. They all assume the completeness of the domain, while the set M of points where f is Gâteaux differentiable (on which we have to maximize $\|f'(x)\|$) is in general not complete and does not even admit an equivalent complete metric. (M is an $F_{\sigma\delta}$ set in X , in general not a G_δ set.) Thus we need a variational principle in which we do not assume the completeness of the domain.

Such a variational principle is proved in Section 2 below. This variational principle has other uses as well. In fact, we first proved it in order to obtain results on the structure

of porous sets (a notion which will not be used in the present paper; see however the book [6].) In Section 3 we show how the variational principle is used to prove the easy special case of Theorem 1.1 where we assume that f is everywhere Gâteaux differentiable.

In Section 4 we prove, using the Hardy–Littlewood maximal lemma, a one-dimensional result which is used in the final step of the proof of Theorem 1.1. Section 5 contains a complete proof of Theorem 1.1 using the variational principle. Although this proof uses the main ideas of the proof in [9] the many difficulties encountered in [9] become much easier to overcome in view of the use of the variational principle and, more importantly, the steps in the proof are now more natural.

The paper ends in Section 6 where several extensions of Theorem 1.1 are announced (which will be proved in the forthcoming book [6]).

2. 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 $\delta_j: M^{j+1} \rightarrow (0, \infty)$, $j \geq 0$, such that every d -Cauchy sequence $(x_j)_{j=0}^\infty$ converges to an element of M provided that*

$$d_0(x_j, x_{j+1}) \leq \delta_j(x_0, \dots, x_j) \quad \text{for each } j \geq 0.$$

A function $f: M \rightarrow \mathbb{R}$ is (d, d_0) -lower semicontinuous if there are functions $\delta_j: M^{j+1} \rightarrow (0, \infty)$, $j \geq 0$, such that

$$f(x) \leq \liminf_{j \rightarrow \infty} f(x_j)$$

whenever $x_j \in M$ converge in metric d to $x \in M$ and

$$d_0(x_j, x_{j+1}) \leq \delta_j(x_0, \dots, x_j) \quad \text{for each } j \geq 0.$$

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 $f: M \rightarrow \mathbb{R}$ has the property that for each $r \in \mathbb{R}$, the set $\{x \in M \mid f(x) \leq r\}$ is a G_δ 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, \dots .

We denote by $B_0[x, r]$ the d_0 -closed ball with centre x and radius r . Let $\delta_0 \equiv 1$. We choose recursively the values $\delta_j(x_0, \dots, x_j) > 0$, $j \geq 1$, so that

- $\delta_j(x_0, \dots, x_j) \leq \frac{1}{2}\delta_{j-1}(x_0, \dots, x_{j-1})$,
- $B_0[x_j, 2\delta_j(x_0, \dots, x_j)] \subset G_i$ whenever $0 \leq i < j$ and $x_j \in G_i$.

Suppose now that the points x_j converge in (M, d) to some $x_\infty \in M$ and satisfy $d_0(x_j, x_{j+1}) \leq \delta_j(x_0, \dots, x_j)$ for each $j \geq 0$. Consider any rational number q with $q < f(x_\infty)$. Since the set $\{x \in M \mid f(x) \leq q\}$ is the intersection of a subcollection of the sets G_i , there is i such that

$$\{x \in M \mid f(x) \leq q\} \subset G_i \quad \text{and} \quad x_\infty \notin G_i.$$

We show that $f(x_j) > q$ for $j > i$. Indeed, $f(x_j) \leq q$ implies that $x_j \in G_i$ and so the ball $B_0[x_j, 2\delta_j(x_0, \dots, x_j)]$ is contained in G_i . But for all $k \geq j$ we have

$$d_0(x_k, x_{k+1}) \leq \delta_k(x_0, \dots, x_k) \leq 2^{j-k}\delta_j(x_0, \dots, x_j),$$

hence

$$\begin{aligned} d_0(x_j, x_k) &\leq d_0(x_j, x_{j+1}) + \dots + d_0(x_{k-1}, x_k) \\ &\leq (1 + 2^{-1} + \dots + 2^{j-k+1})\delta_j(x_0, \dots, x_j) \leq 2\delta_j(x_0, \dots, x_j). \end{aligned}$$

Since the metric d_0 is d -continuous, we can take the limit $k \rightarrow \infty$ to deduce that $d_0(x_j, x_\infty) \leq 2\delta_j(x_0, \dots, x_j)$. This implies

$$x_\infty \in B_0[x_j, 2\delta_j(x_0, \dots, x_j)] \subset G_i,$$

which is a contradiction. Hence $f(x_j) > q$ for all $j > i$, and consequently

$$f(x_\infty) \leq \liminf_{j \rightarrow \infty} f(x_j)$$

as required. \square

Lemma 2.3. *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 $\delta_j: M^{j+1} \rightarrow (0, \infty)$ such that

$$f(x) \leq \liminf_{j \rightarrow \infty} f(x_j)$$

whenever $x_j \in M$ converge to x in (M, d) and $d_0(x_j, x_{j+1}) \leq \delta_j(x_0, \dots, x_j)$ for each $j \geq 0$.

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

$$f(x) \leq \liminf_{j \rightarrow \infty} f(x_j) = 0,$$

implying that $x \in S$, and we are done. \square

The variational principle is contained in the following theorem.

Theorem 2.4. Let $f: M \rightarrow \mathbb{R}$ 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 $F_j: M \times M \rightarrow [0, \infty]$, $j \geq 0$, be functions which are d -lower semicontinuous in the second variable with $F_j(x, x) = 0$ for all $x \in M$ and

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

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 $\varphi: M \rightarrow \mathbb{R}$ 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 $0 \leq j < k + 1 \leq \infty$,

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

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 φ . So one can always replace φ by $\varphi - \varphi(x_\infty)$ to achieve that $\varphi(x_\infty) = 0$. Once we have $\varphi(x_\infty) = 0$, we may replace φ 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 φ 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 $F_i(x, y)$ 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 ε_j for $j \geq 1$, if necessary, we may assume that $\sum_{j=1}^{\infty} \varepsilon_j < \infty$ and

$$\inf_{d(x,y) > r_j} F_j(x, y) > \varepsilon_j \quad (5)$$

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

We will now define by induction points $x_j \in M$, d_0 -continuous functions $\varphi_j: M \rightarrow [0, \infty)$ and lower bounded functions $h_j: M \rightarrow (-\infty, \infty]$ 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) $\varphi_i(x_j) + F_i(x_i, x_j) \leq \varepsilon_i$ 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 , φ_i and h_i have already been defined for $i = 0, 1, \dots, j-1$. We let

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

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

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

Finally, we put

$$\varphi_j(x) = 2\varepsilon_j \min\{1, d_0(x, x_j)/\delta_j(x_0, \dots, x_j)\}.$$

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

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

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), $F_i(x, y) > \varepsilon_i$ 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 $d_0(x_i, x_j) \leq \delta_i(x_0, \dots, x_i)$ for $j > i$: the opposite inequality $d_0(x_i, x_j) > \delta_i(x_0, \dots, x_i)$ 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

$$f(x_\infty) \leq \liminf_{j \rightarrow \infty} f(x_j).$$

Let $\varphi(x) = \sum_{j=0}^{\infty} \varphi_j(x)$. Since the series is uniformly convergent, the function φ 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 ,

$$\begin{aligned} h_0(x_0) &\geq h_1(x_1) \geq \dots \geq \lim_{j \rightarrow \infty} h_j(x_j) \geq \liminf_{j \rightarrow \infty} \left(f(x_j) + \varphi(x_j) + \sum_{i=0}^k F_i(x_i, x_j) \right) \\ &\geq f(x_\infty) + \varphi(x_\infty) + \sum_{i=0}^k F_i(x_i, x_\infty). \end{aligned}$$

Hence

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

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

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

Finally, we may take the limit as $j \rightarrow \infty$ and conclude that

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

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

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 $f : X \rightarrow \mathbb{R}$ 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 $\Theta : X \rightarrow \mathbb{R}$ is called *upper Fréchet differentiable* at $x_0 \in X$ if there is $L \in X^*$ such that

$$\limsup_{\|u\| \rightarrow 0} \frac{\Theta(x_0 + u) + \Theta(x_0) - Lu}{\|u\|} \leq 0.$$

Observation 3.2. *Let $\Theta : X \rightarrow \mathbb{R}$ be everywhere upper Fréchet differentiable, $\psi : X \rightarrow \mathbb{R}$ continuous and $f : X \rightarrow \mathbb{R}$ Lipschitz and everywhere Gâteaux differentiable. Suppose further that the function $h : X \times X \rightarrow \mathbb{R}$ 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 Θ 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 ψ , 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 $|t| \leq \delta/\Delta$.

Assume that $v \in X$ with $\|v\| < \delta$. We denote for the moment $t = \|v\|/\Delta$ and $u = (1/t)v + u_0$. Since $h(x, u)$ attains its minimum at (x_0, u_0) one can write, for every $x \in [x_0 - tu_0, x_0 + v]$, that

$$f'(x; u) + \Theta(u) + \psi(x) \geq f'(x_0; u_0) + \Theta(u_0) + \psi(x_0).$$

Hence, using that $\|x - x_0\| \leq \|v\| + t\|u_0\| \leq \|v\|(1 + \|u_0\|/\Delta) < \delta_0$ and $\|u - u_0\| = \Delta$, we obtain

$$\begin{aligned} f'(x; u) &\geq f'(x_0; u_0) - (\Theta(u) - \Theta(u_0)) - (\psi(x) - \psi(x_0)) \\ &\geq f'(x_0; u_0) - L(u - u_0) - \frac{\varepsilon}{3}\|u - u_0\| - \frac{\varepsilon\Delta}{3} \\ &= f'(x_0; u_0) - \frac{1}{t}L(v) - \frac{\varepsilon}{3t}\|v\| - \frac{\varepsilon}{3t}\|v\| = f'(x_0; u_0) - \frac{1}{t}L(v) - \frac{2\varepsilon}{3t}\|v\|. \end{aligned}$$

So

$$\begin{aligned} f(x_0 + v) - f(x_0 - tu_0) &\geq t \inf\{f'(x; u) \mid x \in [x_0 - tu_0, x_0 + v]\} \\ &\geq f'(x_0; tu_0) - L(v) - \frac{2\varepsilon}{3}\|v\|. \end{aligned} \quad (7)$$

Since $|t| \leq \delta/\Delta$, we also have

$$f(x_0 - tu_0) - f(x_0) \geq -f'(x_0; tu_0) - \frac{\varepsilon\Delta}{3}t = -f'(x_0; tu_0) - \frac{\varepsilon}{3}\|v\|.$$

Adding this to (7), we get

$$f(x_0 + v) - f(x_0) \geq -L(v) - \varepsilon\|v\|.$$

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 $x \in [x_0 + v, x_0 + tu_0]$, using again that

$$f'(x; u) + \Theta(u) + \varphi(x) \geq f'(x_0; u_0) + \Theta(u_0) + \varphi(x_0),$$

we get

$$\begin{aligned} f'(x; u) &\geq f'(x_0; u_0) - (\Theta(u) - \Theta(u_0)) - (\psi(x) - \psi(x_0)) \\ &\geq f'(x_0; u_0) - L(u - u_0) - \frac{\varepsilon}{3}\|u - u_0\| - \frac{\varepsilon\Delta}{3} \\ &= f'(x_0; u_0) + \frac{1}{t}L(v) - \frac{2\varepsilon}{3t}\|v\|. \end{aligned}$$

So

$$\begin{aligned} f(x_0 + tu_0) - f(x_0 + v) &\geq t \inf\{f'(x; u) \mid x \in [x_0 + v, x_0 + tu_0]\} \\ &\geq f'(x_0; tu_0) + L(v) - \frac{2\varepsilon}{3}\|v\|. \end{aligned}$$

Subtracting this from

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

we get

$$f(x_0 + v) - f(x_0) \leq -L(v) + \frac{2\varepsilon}{3} \|v\| + \frac{\varepsilon \Delta}{3} t = -L(v) + \varepsilon \|v\|. \quad \square$$

Clearly, the above argument works under considerably less stringent assumptions. For example, the upper Fréchet differentiability of Θ 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 $f: X \rightarrow \mathbb{R}$ 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 $\delta_j: M^{j+1} \rightarrow (0, \infty)$ witnessing the (d, d_0) -continuity. Let $\delta_0(x_0, u_0) = 1$. Given $j \geq 1$ and the pairs $(x_0, u_0), \dots, (x_j, u_j)$, we find

$$0 < \delta \leq \frac{1}{2} \delta_{j-1}((x_0, u_0), \dots, (x_{j-1}, u_{j-1}))$$

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 $\delta_j((x_0, u_0), \dots, (x_j, u_j)) = \delta$.

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

$$\|x_{j+1} - x_j\| \leq \delta_j((x_0, u_0), \dots, (x_j, u_j)).$$

To simplify the notation we let $\delta_j = \delta_j((x_0, u_0), \dots, (x_j, u_j))$.

Since $\delta_{j+1} \leq \frac{1}{2} \delta_j$, we have $\|x - x_j\| \leq 2\delta_j$ and also $j\delta_j \rightarrow 0$. 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:

$$\begin{aligned} |f'(x; tu) - f'(x_j; tu_j)| &\leq |f(x + tu) - f(x) - f(x_j + tu_j) + f(x_j)| + \varepsilon|t| + |t|/j \\ &\leq |f(x + tu) - f(x_j + tu_j)| + |f(x_j) - f(x)| + \varepsilon|t| + |t|/j \\ &\leq 2\text{Lip}(f)\|x - x_j\| + \text{Lip}(f)\|u - u_j\||t| + \varepsilon|t| + |t|/j \\ &\leq \left(\frac{4\text{Lip}(f)\delta_j}{|t|} + \text{Lip}(f)\|u - u_j\| + \varepsilon + \frac{1}{j} \right) |t| \\ &= \left(\frac{1 + 4\text{Lip}(f)}{j} + \text{Lip}(f)\|u - u_j\| + \varepsilon \right) |t| \leq 3\varepsilon|t|. \quad \square \end{aligned}$$

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 ,

$$d((x, u), (y, v)) = \sqrt{\|x - y\|^2 + \|u - v\|^2},$$

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 ε_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_∞, u_∞) and a d_0 -continuous function $\varphi: X \times X \rightarrow \mathbb{R}$ 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_∞, u_∞) . Notice that d_0 -continuity of φ means that the function φ 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_∞ . □

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 $f : \mathbb{R} \rightarrow \mathbb{R}$ 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 $h : [a, b] \rightarrow \mathbb{R}$ 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

$$\int_{\{t \mid h'(t) > 0\}} h'(t) dt = \int_{\{t \mid h'(t) < 0\}} -h'(t) dt,$$

i.e.

$$\int_a^b |h'(t)| dt = 2 \int_a^b \max\{0, h'(t)\} dt. \quad \square$$

Lemma 4.2. *Let $h : [a, b] \rightarrow \mathbb{R}$ 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 ξ ;
- (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 $\mathcal{L}^1 S < \frac{1}{3\kappa} \|h\|_\infty$. Then for a.e. $t \in [a, b] \setminus S$ either

$$h'(t) < \frac{\|h\|_\infty}{3(b-a)} \quad \text{or} \quad h'(t) \geq \frac{\|h\|_\infty}{3(b-a)} \quad \text{and} \quad H(t) > 48\sqrt{\kappa h'(t)}.$$

Recalling also that $h' \leq \kappa$, we see that for a.e. $t \in [a, b]$,

$$\max\{0, h'(t)\} < \max\{0, h'(t)\} \chi_S(t) + \frac{\|h\|_\infty}{3(b-a)} + \min\left\{\kappa, \frac{1}{48^2\kappa} H^2(t)\right\}.$$

Hence we get with the help of Lemma 4.1 a contradiction by estimating

$$\begin{aligned} & \int_a^b \max\{0, h'(t)\} dt \\ & < \int_S \max\{0, h'(t)\} dt + \int_a^b \frac{\|h\|_\infty}{3(b-a)} dt + \int_a^b \min\left\{\kappa, \frac{1}{48^2\kappa} H^2(t)\right\} dt \\ & \leq \kappa \mathcal{L}^1 S + \frac{1}{3} \|h\|_\infty + \int_0^\kappa \mathcal{L}^1 \left\{t \in [a, b] \mid \frac{1}{48^2\kappa} H^2(t) > \lambda\right\} d\lambda \\ & \leq \frac{2}{3} \|h\|_\infty + \int_0^\kappa \frac{1}{6\sqrt{\kappa\lambda}} d\lambda \int_a^b \max\{0, h'(t)\} dt \\ & \leq \frac{2}{3} \int_a^b \max\{0, h'(t)\} dt + \frac{2}{6} \int_a^b \max\{0, h'(t)\} dt \\ & = \int_a^b \max\{0, h'(t)\} dt. \quad \square \end{aligned}$$

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 κ -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 ξ ;
- (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 $h_0: [-a, a] \rightarrow \mathbb{R}$ 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 ξ ;
- (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

$$\sup_{t \in [-a, a]} |h_0(t) - h(t)| = \max\{|h(a)|, |h(-a)|\} = \max\{|g(a)|, |g(-a)|\} \leq \eta a, \tag{8}$$

and for $t, s \in [-a, a]$,

$$\begin{aligned} |(h(t) - h(s)) - (h_0(t) - h_0(s))| &\leq \frac{|t-s|}{2a}|h(a) - h(-a)| \\ &= \frac{|t-s|}{2a}|g(a) - g(-a)| \leq \eta|t-s|. \end{aligned} \tag{9}$$

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

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

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 (9) and (b) 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 (ii). 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

$$\begin{aligned} |h(t) - h(\xi)| &\leq |h_0(t) - h_0(\xi)| + \eta|t - \xi| \leq (48\sqrt{\kappa_0 h'_0(\xi)} + \eta)|t - \xi| \\ &\leq (48\sqrt{2\kappa 2h'(\xi)} + \eta\sqrt{h'(\xi)/\tau})|t - \xi| \\ &= (96 + \eta/\sqrt{\kappa\tau})\sqrt{\kappa h'(\xi)}|t - \xi| \leq 97\sqrt{\kappa h'(\xi)}|t - \xi|. \end{aligned} \tag{10}$$

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

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

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

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

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 $f: G \rightarrow \mathbb{R}$ 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 $\Theta: X \rightarrow [0, \infty)$ such that*

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

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 $B(x_0, r_0) \subset G$ and Θ is Lipschitz on $B(0, r_0)$. We let $r = \frac{1}{3}r_0$ and denote $B = \{x \in X \mid \|x - x_0\| \leq r\}$, $U = \{u \in X \mid \|u - u_0\| \leq r\}$, and

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

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

$$\mathcal{P} = \{g: (-r, r) \rightarrow \mathbb{R} \mid g \text{ is Lipschitz, } g(0) = 0\}$$

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

$$f_{x,u}(t) = f(x + tu) - f(x). \quad (11)$$

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 ϱ is finite. The wished-for metric d on M is

$$d((x, u), (y, v)) = \max\{\|x - y\|, \|u - v\|, \varrho(f_{x,u}, f_{y,v})\}. \quad (12)$$

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

$$\sup_{t \in (-r, r) \setminus \{0\}} \frac{|f_{x_k, u_k}(t) - g(t)|}{|t|} \rightarrow 0. \quad (13)$$

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}, ϱ) , 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 $f'(x_k; u_k) \rightarrow f'(x; u)$. The fact that

$$\lim_{t \rightarrow 0} \frac{f_{x_k, u_k}(t)}{t} = f'(x_k; u_k)$$

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

$$\lim_{t \rightarrow 0} \frac{f_{x, u}(t)}{t} = f'(x; u)$$

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

The three statements of the next lemma relate some ordinary topological or metric notions in $B \times U$ to those coming from the metric ϱ . 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 ϱ 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 $y \mapsto \varrho(f_{x, u}, f_{y, u})$ is norm lower semicontinuous on B .*
- (ii) *For any $(x, u), (x, v) \in B \times U$, $\varrho(f_{x, u}, f_{x, v}) \leq C \|u - v\|$.*
- (iii) *For every $x, y, z \in B, u \in U$ and $\delta > 0$,*

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

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

$$y \mapsto |f_{x, u}(t) - f_{y, u}(t)|/|t|$$

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

(ii) Obvious by taking $C \geq \text{Lip}(f)$.

(iii) If $|t| \geq \delta$ then

$$\begin{aligned} \frac{|f_{x,u}(t) - f_{y,u}(t)|}{|t|} &\leq \frac{|f_{x,u}(t) - f_{z,u}(t)|}{|t|} + \frac{2\text{Lip}(f)}{|t|} \|y - z\| \\ &\leq \varrho(f_{x,u}, f_{z,u}) + \frac{2\text{Lip}(f)}{\delta} \|y - z\|. \end{aligned}$$

If $|t| < \delta$, then for every $w \in U$,

$$\frac{|f_{x,u}(t) - f_{y,u}(t)|}{|t|} \leq \text{Lip}(f) \|u - w\| + \frac{|f_{x,u}(t) - f_{y,w}(t)|}{|t|}.$$

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

Use of the variational principle. Choose $\varepsilon_0 > 0$ so that $f'(x_0; u_0) \leq \varepsilon_0 + \inf\{f'(x; u) \mid (x, u) \in M\}$ and find $\lambda_0 > \varepsilon_0$ so large that

$$\frac{\varepsilon_0}{\lambda_0} < \frac{r}{2} \quad \text{and} \quad \Theta(z) > \frac{\varepsilon_0}{\lambda_0} \quad \text{for} \quad \|z\| \geq \frac{r}{2}. \quad (14)$$

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:

$$\sum_{i=1}^{\infty} \sigma_i \leq \frac{2}{3} \sigma_0. \quad (15)$$

To define our last sequence of parameters, s_i , we start by letting

$$t_i = \inf\{t > 0 \mid \Theta(z) > \sigma_i^2 \text{ for } \|z\| > t\}$$

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, $\lim_{i \rightarrow \infty} t_i > 0$, 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 $i \geq 1$,

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

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

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

where

$$\begin{aligned} \Phi_i(x, y) &= \lambda_i \|x - y\|, & \Psi_i(u, v) &= \lambda_i \Theta(u - v), \\ Q_i((x, u), (y, v)) &= \sigma_i (f'(x; u) - f'(y; v))^2 \end{aligned}$$

and

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

Notice that the peculiarity in the definition of Δ_i is not a misprint: Δ_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 Δ_i in the second variable follows directly from Lemma 5.3(i). 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

$$\varrho(f_{x,u}, f_{y,u}) \geq \varrho(f_{x,u}, f_{y,v}) - \varrho(f_{y,u}, f_{y,v}) \geq (C + 2)s_i - C\|u - v\| \geq 2s_i.$$

Hence $\Delta_i((x, u), (y, v)) \geq \min\{\lambda_i, s_i\}$. □

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

$$h_i(x, u) = f'(x; u) + \sum_{j=0}^{i-1} F_j((x_j, u_j), (x, u)), \tag{17}$$

we have

$$h_\infty(x_\infty, u_\infty) \leq \min\{h_i(x_i, u_i), \varepsilon_i + \inf_{(x,u) \in M} h_i(x, u)\} \tag{18}$$

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

Recalling the definition of the functions F_i , we have

$$h_\infty(x, u) = f'(x; u) + \Phi(x) + \Psi(u) + Q(x, u) + \Delta(x), \tag{19}$$

where

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

to justify the last definition we recall the independence of Δ_i from the last variable. Observe that all these functions are positive and finite, Φ , Ψ and Q are d -continuous on M (for Q recall Lemma 5.2), and Ψ 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_∞ . We first collect several estimates, all easily following from (18), that will be needed in what follows.

Lemma 5.5. $\|x_\infty - x_0\| < \frac{1}{2}r$, $\|u_\infty - u_0\| < \frac{1}{2}r$ and for each $0 \leq i < \infty$, $\|u_i - u_\infty\| \leq \frac{1}{10C}s_i$ and $|f'(x_\infty; u_\infty) - f'(x_i; u_i)| \leq \frac{1}{5}s_i$.

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

$$\begin{aligned} h_i(x_\infty, u_\infty) + \Phi_i(x_i, x_\infty) + \Psi_i(u_i, u_\infty) + Q_i((x_i, u_i), (x_\infty, u_\infty)) \\ \leq h_\infty(x_\infty, u_\infty) \leq \varepsilon_i + h_i(x_\infty, u_\infty). \end{aligned} \tag{20}$$

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

$$\Theta(u_0 - u_\infty) = \frac{1}{\lambda_0} \Psi_0(u_0, u_\infty) \leq \frac{\varepsilon_0}{\lambda_0},$$

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 $\|u_\infty - u_i\| < \frac{1}{10C}s_i$. 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

$$(f'(x_\infty; u_\infty) - f'(x_i; u_i))^2 = \frac{1}{\sigma_i} Q_i((x_i, u_i), (x_\infty, u_\infty)) \leq \frac{\varepsilon_i}{\sigma_i} = \sigma_i^2 \leq \left(\frac{s_i}{5}\right)^2. \quad \square$$

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_Ψ be an upper Fréchet derivative of Ψ at u_∞ . 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_∞ 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

$$|\varkappa| \leq 2 \sum_{i=0}^{\infty} \sigma_i^2 \leq 2 \left(\sigma_0 \kappa + \sum_{i=1}^{\infty} \sigma_i^2 \right) \leq 2 \left(\sigma_0 \kappa + \frac{2}{3} \sigma_0 \right) \leq 2 \left(\frac{1}{20} + \frac{2}{3} \frac{1}{20} \right) = \frac{1}{6}.$$

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 :

$$\begin{aligned} \|L\| &\leq \frac{6}{5} \|L_\Psi\| \leq \frac{6}{5} \text{Lip}(\Psi|_U) \leq \frac{6}{5} \sum_{i=0}^{\infty} \lambda_i \text{Lip}(\Theta|_{B(0, r_0)}) \\ &= \frac{6}{5} \left(\lambda_0 + \sum_{i=1}^{\infty} \sigma_i \right) \text{Lip}(\Theta|_{B(0, r_0)}) \leq 2\lambda_0 \text{Lip}(\Theta|_{B(0, r_0)}). \end{aligned} \quad (21)$$

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_∞ , giving that the derivative of f at the point x_∞ in the direction u_∞ agrees with the value of the linear form L at u_∞ .

Lemma 5.6. $f'(x_\infty; u_\infty) = L(u_\infty)$.

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

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

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

$$f'(x_\infty; u_\infty) + L_\Psi u_\infty + \varkappa f'(x_\infty; u_\infty) = 0,$$

yielding

$$f'(x_\infty; u_\infty) = -\frac{L_\Psi u_\infty}{1 + \varkappa} = L u_\infty$$

as required. \square

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_∞ 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

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

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 $0 < \delta_0 < \frac{1}{2}r$ so that

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

$$\Psi(u) \leq \Psi(u_\infty) + L_\Psi(u - u_\infty) + \frac{1}{234}\eta \|u - u_\infty\|.$$

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

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

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

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

$$\Phi(x) < \Phi(x_\infty) + \frac{1}{3}\varepsilon_k;$$

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

$$|f(x_i + tu_i) - f(x_i) - tf'(x_i; u_i)| \leq \min\left\{\frac{s_k}{5}, \frac{\tau^2}{\kappa}\right\}|t|.$$

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 $\gamma: [-\delta, \delta] \rightarrow X$ 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 γ is contained in the interior of the ball B . Further, let $g, h: [-\delta, \delta] \rightarrow \mathbb{R}$ 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

$$|h(0) - g(0)| > \eta\|v\| = 13\tau a \quad \text{and} \quad |g(t) - g(0)| \leq \tau^2/\kappa \quad \text{for } |t| \leq \delta$$

follow directly from (22), (β) and from (δ) with $i = \infty$. Also,

$$\tau = 6\varepsilon_k \leq 6\sigma_0 \leq \kappa.$$

The remaining assumption is that h is κ -Lipschitz. Using (21),

$$\begin{aligned} \text{Lip}(h) &\leq \text{Lip}(f) \text{Lip}(\gamma) + \|L\| \text{Lip}(\gamma) \leq (\text{Lip}(f) + \|L\|)(\|u_\infty\| + \beta) \\ &\leq (\text{Lip}(f) + 2\lambda_0 \text{Lip}(\Theta|_{B(0, r_0)}))(1 + r) \leq \kappa. \end{aligned}$$

So we may apply Corollary 4.3 to find $\xi \in (-a, a) \setminus \{0\}$ such that

- (a) h is differentiable at ξ ;
- (b) $h'(\xi) \geq 6\varepsilon_k$;
- (c) $|h(t) - h(\xi)| \leq 99\sqrt{\kappa h'(\xi)}|t - \xi|$ 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 $\|u - u_\infty\| = \beta \leq \frac{1}{2}r$, 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 . □

We wish to show that $h_\infty(x, u) < h_\infty(x_\infty, u_\infty)$ to contradict that h_∞ attains its minimum at (x_∞, u_∞) . For this, we first compare the terms not containing the function Δ . 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

$$\begin{aligned} Q(x, u) &= \sum_{i=0}^{\infty} \sigma_i [f'(x_i; u_i) - f'(x_\infty; u_\infty) + f'(x_\infty; u_\infty) - f'(x; u)]^2 \\ &= Q(x_\infty, u_\infty) + \varkappa(f'(x; u) - f'(x_\infty; u_\infty)) + \sum_{i=0}^{\infty} \sigma_i (f'(x; u) - f'(x_\infty; u_\infty))^2 \\ &= Q(x_\infty, u_\infty) + \varkappa(-h'(\xi) + L(u - u_\infty)) + \sum_{i=0}^{\infty} \sigma_i (-h'(\xi) + L(u - u_\infty))^2 \\ &\leq Q(x_\infty, u_\infty) - \varkappa h'(\xi) + \varkappa L(u - u_\infty) + \frac{10}{3}\sigma_0(h'(\xi))^2 + \frac{10}{3}\sigma_0(L(u - u_\infty))^2. \end{aligned}$$

Since h is κ -Lipschitz the choice of σ_0 gives

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

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

$$Q(x, u) \leq Q(x_\infty, u_\infty) + \varkappa L(u - u_\infty) + \frac{1}{3}h'(\xi) + \frac{1}{3}\varepsilon_k.$$

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

$$\begin{aligned} \Psi(u) &\leq \Psi(u_\infty) + L_\Psi(u - u_\infty) + \frac{1}{234}\eta\|u - u_\infty\| \\ &= \Psi(u_\infty) - (1 + \varkappa)L(u - u_\infty) + \frac{1}{234}\eta\beta \\ &= \Psi(u_\infty) - (1 + \varkappa)L(u - u_\infty) + \frac{1}{3}\varepsilon_k. \end{aligned}$$

Finally, the last needed inequality $\Phi(x) < \Phi(x_\infty) + \frac{1}{3}\varepsilon_k$ follows from (γ) , 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

$$\begin{aligned} G(x, u) &= f'(x; u) + \Phi(x) + \Psi(u) + Q(x, u) \\ &< f'(x; u) + \Phi(x_\infty) + \Psi(u_\infty) + Q(x_\infty, u_\infty) - L(u - u_\infty) + \frac{1}{3}h'(\xi) + \varepsilon_k \\ &= -h'(\xi) + Lu + \Phi(x_\infty) + \Psi(u_\infty) + Q(x_\infty, u_\infty) - Lu + f'(x_\infty; u_\infty) \\ &\quad + \frac{1}{3}h'(\xi) + \varepsilon_k \\ &= -\frac{2}{3}h'(\xi) + G(x_\infty, u_\infty) + \varepsilon_k \leq -\frac{1}{2}h'(\xi) + G(x_\infty, u_\infty). \quad \square \end{aligned}$$

The next lemma provides us with a certain control of the function Δ .

Lemma 5.9. For each $0 \leq i \leq k$,

$$(A_i) \quad 495\sqrt{\kappa h'(\xi)} \leq s_i;$$

$$(B_i) \quad \Delta_i((x_i, u_i), (x, u)) \leq \Delta_i((x_i, u_i), (x_\infty, u_\infty)) + \frac{1}{k+1}\varepsilon_k.$$

Proof. We first prove that for each i , $0 \leq i \leq k$, (A _{i}) implies (B _{i}). By the definition of Δ_i it suffices to show that

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

By Lemma 5.3(iii) with $\delta/2$ instead of δ we have

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

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

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

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

$$C \|u_i - w\| + \left| \frac{f_{x, w}(t) - f_{x_i, u_i}(t)}{t} \right| \quad (24)$$

for $w = \frac{1}{t}(\gamma(\xi + t) - \gamma(\xi))$. (Notice that $\xi + t$ still belongs to the interval $(-\delta, \delta)$ where γ 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 $\Delta_i((x_i, u_i), (x_\infty, u_\infty))$.

Observe first that by the definition of γ , $\|u_\infty - w\| \leq \|v\|/a = \beta$. Together with Lemma 5.5 this gives

$$C \|u_i - w\| \leq C \|u_i - u_\infty\| + C \|u_\infty - w\| \leq \frac{1}{10} s_i + C\beta \leq \frac{1}{5} s_i.$$

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

$$\begin{aligned} \left| \frac{f_{x, w}(t) - f_{x_i, u_i}(t)}{t} \right| &\leq \left| \frac{f_{x, w}(t)}{t} + L(w) \right| + |L(w) + f'(x_\infty; u_\infty)| \\ &\quad + |f'(x_\infty; u_\infty) - f'(x_i; u_i)| + \left| f'(x_i; u_i) - \frac{f_{x_i, u_i}(t)}{t} \right|. \end{aligned}$$

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

$$\left| \frac{f_{x, w}(t)}{t} + L(w) \right| = \frac{|h(\xi + t) - h(\xi)|}{|t|} \leq 99\sqrt{\kappa h'(\xi)} \leq \frac{1}{5} s_i;$$

the choice of β guarantees

$$|L(w) + f'(x_\infty; u_\infty)| = |L(w) - L(u_\infty)| \leq \|L\|\beta \leq \frac{1}{5} s_i;$$

the estimate

$$|f'(x_\infty; u_\infty) - f'(x_i; u_i)| \leq \frac{1}{5}s_i$$

follows from Lemma 5.5; and finally (δ) yields

$$\left| \frac{f_{x_i, u_i}(t)}{t} - f'(x_i; u_i) \right| \leq \frac{1}{5}s_i.$$

Adding these inequalities we get

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

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

To prove (A_i) , we infer from Lemma 5.8 and (18) that

$$\begin{aligned} \frac{1}{2}h'(\xi) &\leq G(x_\infty, u_\infty) - G(x, u) = h_\infty(x_\infty, u_\infty) - \Delta(x_\infty, u_\infty) - G(x, u) \\ &\leq \varepsilon_i + h_i(x, u) - \Delta(x_\infty, u_\infty) - G(x, u) \\ &\leq \varepsilon_i + \sum_{j=0}^{i-1} (\Delta_j((x_j, u_j), (x, u)) - \Delta_j((x_j, u_j), (x_\infty, u_\infty))). \end{aligned}$$

With $i = 0$ this gives $h'(\xi) \leq 2\varepsilon_0$, hence

$$495\sqrt{\kappa h'(\xi)} \leq 495\sqrt{2\kappa\varepsilon_0} \leq s_0$$

by the choice of s_0 . Proceeding by induction, we infer from the first part of this proof that

$$\Delta_j((x_j, u_j), (x, u)) - \Delta_j((x_j, u_j), (x_\infty, u_\infty)) \leq \frac{\varepsilon_k}{k+1} \quad \text{for } j < i.$$

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

$$495\sqrt{\kappa h'(\xi)} \leq 990\sqrt{\kappa\varepsilon_i} \leq s_i. \quad \square$$

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

Proof. First recall that the functions Δ_i are bounded by λ_i by their very definition. Further, the choice of ε_i and λ_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 $i > k$,

$$\Delta_i((x_i, u_i), (x, u)) \leq \lambda_i = \varepsilon_{i-1} \leq 2^{k-i+1}\varepsilon_k.$$

Hence the previous lemma gives

$$\begin{aligned} \Delta(x, u) &= \sum_{i=0}^k \Delta_i((x_i, u_i), (x, u)) + \sum_{i=k+1}^{\infty} \Delta_i((x_i, u_i), (x, u)) \\ &\leq \sum_{i=0}^k \left(\Delta_i((x_i, u_i), (x_{\infty}, u_{\infty})) + \frac{\varepsilon_k}{k+1} \right) + 2\varepsilon_k \\ &\leq \sum_{i=0}^{\infty} \Delta_i((x_i, u_i), (x_{\infty}, u_{\infty})) + 3\varepsilon_k = \Delta(x_{\infty}, u_{\infty}) + 3\varepsilon_k. \quad \square \end{aligned}$$

Adding the inequalities from Lemma 5.8 and Corollary 5.10 and using that $h'(\xi) \geq 6\varepsilon_k$ leads to

$$\begin{aligned} h_{\infty}(x, u) &= G(x, u) + \Delta(x, u) \\ &< \left(-\frac{1}{2}h'(\xi) + G(x_{\infty}, u_{\infty})\right) + (\Delta(x_{\infty}, u_{\infty}) + 3\varepsilon_k) \\ &= -\frac{1}{2}h'(\xi) + 3\varepsilon_k + h(x_{\infty}, u_{\infty}) \leq h_{\infty}(x_{\infty}, u_{\infty}). \end{aligned}$$

Recalling that h_{∞} attains its minimum at (x_{∞}, u_{∞}) 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 $f: X \rightarrow \mathbb{R}$ is called *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 Θ 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 $f : X \rightarrow \mathbb{R}$ 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 $f : X \rightarrow Y$ satisfies

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

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

- [1] Benyamini, Y., Lindenstrauss, J.: Geometric Nonlinear Functional Analysis. Vol. 1, Colloq. Publ. 48, Amer. Math. Soc. (2000) Zbl 0946.46002 MR 1727673
- [2] Borwein, J. M., Preiss, D.: A smooth variational principle with applications to sub-differentiability and to differentiability of convex functions. Trans. Amer. Math. Soc. **303**, 517–528 (1987) Zbl 0632.49008 MR 0902782
- [3] Borwein, J. M., Wang, X.: Cone monotone functions, differentiability and continuity. Canad. J. Math. **57**, 961–982 (2005) Zbl 1095.49015 MR 2164591
- [4] Deville, R., Godefroy, G., Zizler, V.: Smoothness and Renormings in Banach Spaces. Pitman Monogr. Surveys 64, Longman (1993) Zbl 0782.46019 MR 1211634
- [5] Lindenstrauss, J., Preiss, D.: A new proof of Fréchet differentiability of Lipschitz functions, J. Eur. Math. Soc. **2**, 199–216 (2000) Zbl 0968.58006 MR 1776937
- [6] Lindenstrauss, J., Preiss, D., Tišer, J.: Fréchet Differentiability of Lipschitz Maps and Porous Sets in Banach Spaces. A book in preparation
- [7] Preiss, D.: Gâteaux differentiable Lipschitz functions need not be Fréchet differentiable on a residual set. Suppl. Rend. Circ. Mat. Palermo (2) **2**, 217–222 (1982) Zbl 0518.46032 MR 0683783
- [8] Preiss, D.: Gâteaux differentiable functions are somewhere Fréchet differentiable. Rend. Circ. Mat. Palermo (2) **33**, 122–133 (1984) Zbl 0573.46024 MR 0743214
- [9] Preiss, D.: Differentiability of Lipschitz functions on Banach spaces. J. Funct. Anal. **91**, 312–345 (1990) Zbl 0711.46036 MR 1058975
- [10] Preiss, D., Tišer, J.: Two unexpected examples concerning differentiability of Lipschitz functions on Banach spaces. GAFA Israel Seminar 92–94, Birkhäuser, 219–238 (1995) Zbl 0872.46026 MR 1353461