JEMS

J. Lindenstrauss · D. Preiss

A new proof of Fréchet differentiability of Lipschitz functions

Abstract. We give a relatively simple (self-contained) proof that every real-valued Lipschitz function on ℓ_2 (or more generally on an Asplund space) has points of Fréchet differentiability. Somewhat more generally, we show that a real-valued Lipschitz function on a separable Banach space has points of Fréchet differentiability provided that the w^* closure of the set of its points of Gâteaux differentiability is norm separable.

1. Introduction

The purpose of this paper is to give a simpler (self-contained) proof of the main result of [4], namely that a Lipschitz function from ℓ_2 (or more generally from an Asplund space X) into $\mathbb{R}$ has points of Fréchet differentiability.

The proof in [4] used an iterative method for finding a point of differentiability. Each stage of the iteration procedure involved a new renorming of the space. The fact that X is Asplund was used to ensure the existence of suitable good renormings. In this paper we also use an iterative procedure but the fact that X is Asplund is used via slicing properties of bounded sets in X^*.

Before we proceed we recall some background material (for a discussion of it and further references see, e.g., Phelps [3]).

A function f from an open set in a Banach space X to Y is said to be Gâteaux differentiable at x_0 if there is a bounded linear operator $T : X \to Y$ so that for every $e \in X$

$$\lim_{t \to 0}(f(x_0 + te) - f(x_0))/t = Te.$$

If the limit above exists uniformly in e in the unit sphere of X we say that f is Fréchet differentiable at x_0. Alternatively, f is Fréchet differentiable at

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

D. Preiss: Department of Mathematics, University College London, London WC1E 6BT, UK, e-mail: d.preiss@ucl.ac.uk

x_0 if and only if

$$f(x_0 + y) = f(x_0) + Ty + o(\|y\|) \text{ as } \|y\| \to 0.$$

If Y is the real line we denote the Gâteaux derivative T of f at the point x_0 by $f'(x_0) \in X^*$.

A Banach space X is called an Asplund space if every (real valued) continuous convex function on X is Fréchet differentiable on a dense set. A space X is Asplund if and only if every separable subspace of X has a separable dual. If X is not Asplund there is an equivalent norm on X which is nowhere Fréchet differentiable and thus in particular the assumption that X is Asplund in the theorem of [4] quoted above cannot be weakened. If A is a bounded set in X^* a w^*-slice S of A is a set of the form $S = \{x^* \in A : x^*(x) \geq \sigma - \alpha\}$ for some $x \in X, \alpha > 0$ and where $\sigma = \sup\{x^*(x) : x^* \in A\}$. We say in such a situation that S is a w^*-slice defined by the vector x (to be sure S depends also on α).

We shall consider here mainly separable Banach spaces X. It is well known (see, e.g., [2]) that if f is a Lipschitz function from an open set G in a separable X into $\mathbb{R}$ then f is Gâteaux differentiable in a subset D of G so that $G \setminus D$ is a Gauss null set. We shall work with the set $A \subset X^*$ of Gâteaux derivatives of f at the points of D. This set is clearly bounded in norm by the Lipschitz constant of f. In very crude terms our strategy is to find a decreasing sequence $\{S_n\}_{n=1}^{\infty}$ of w^*-slices of subsets of A whose diameters tend to 0 and find points x_n with $f'(x_n) \in S_n$ so that $x_n \to x$ for some vector x (this is easy to achieve) and so that f is Fréchet differentiable at x with $f'(x) = \lim_n f'(x_n)$ (this is delicate). What the proof gives in particular is that starting with any non empty slice S of A we find a point x of Fréchet differentiability with $f'(x) \in S$. This also proves that the mean value theorem holds for Fréchet derivatives.

The precise statement of the result we prove here is

Theorem. *Let X be a separable Banach space and f a Lipschitz function from an open subset G in X into $\mathbb{R}$. Let $D \subset G$ be the set on which f is Gâteaux differentiable and let $A = \{f'(x) : x \in D\}$. Assume that the w^* closure of A is norm separable. Then G contains a point of Fréchet differentiability of f.*

Moreover, for every segment $[u, v] \subset G$ and for every $m < f(v) - f(u)$ there is a point x of Fréchet differentiability of f so that $f'(x)(v - u) > m$.

By the separation theorem (and the definition of Gâteaux derivatives) an obvious equivalent way to state the theorem above is the following

Corollary. *With the notations and assumptions as above we have that*

$$w^* - \text{cl conv}\{f'(x) : x \in D\} = w^* - \text{cl conv}\{f'(x) : x \in \tilde{D}\}$$

where $\tilde{D}$ is the subset of D consisting of all points at which f is Fréchet differentiable.

It should be pointed out that also the nonseparable case of the main result of [4] follows from the theorem stated above. This will be shown in Proposition 2 in the next section.

Besides being simpler than the proof of [4] (and this is the main point) and giving a somewhat more general result (for separable X) there are also some drawbacks to the proof presented here. In particular it does not seem that the method of proof of this paper yields the following result (which was obtained in [4] as a by-product): There is a subset N of the plane $\mathbb{R}^2$ with Lebesgue measure 0 so that every Lipschitz function from $\mathbb{R}^2$ into $\mathbb{R}$ is differentiable at some point of N.

Although the proof given here is definitely simpler than that of [4] it is still quite involved. The inherent reason for the difficulty of the proof is that it provides an algorithm for producing a sequence of points $\{x_n\}_{n=1}^{\infty}$ which converge to a point of Fréchet differentiability. Such an algorithm is apparently not simple to find even in the classical case of maps from the line into itself. The usual (very short and elegant) proofs of Lebesgue's theorem on the line prove automatically that the function is differentiable outside a set of measure 0. Unfortunately, we are not aware of any definition of a non trivial σ-ideal (or even just an ideal) of sets in an infinite dimensional space X (even if $X = \ell_2$) so that every Lipschitz function from X to $\mathbb{R}$ is Fréchet differentiable outside this σ-ideal (resp. ideal). Several natural candidates for such a σ-ideal can be proved not to be suitable. The strongest tool for proving this are the results of [6]. For example, it is not true that the collection of Gauss null sets is such a σ-ideal. For this specific σ-ideal there is now an even more striking counterexample. The Hilbert space ℓ_2 can be renormed so that the new norm (which is of course a convex function) is Fréchet differentiable only on a Gauss null set [1]. Thus if our algorithm is applied to this specific norm in ℓ_2 we invariably end up in a fixed Gauss null set.

The fact that no "almost everywhere" version of the theorem is known leads to another complication. It is not known, for example, if two Lipschitz functions $f, g \colon \ell_2 \to \mathbb{R}$ have a common point of Fréchet differentiability, or equivalently, whether a Lipschitz function from ℓ_2 to the plane has a point of Fréchet differentiability. The second example in [6] shows that the method used here to prove the theorem (i.e. via slices in the set of Gâteaux derivatives) does not generalize in an obvious way to maps into $\mathbb{R}^k$ with $k > 1$. In [6] it is shown that there is a Lipschitz mapping f from ℓ_2 to $\mathbb{R}^3$ so that

$$f_1'(x)(e_1) + f_2'(x)(e_2) + f_3'(x)(e_3) = 0$$

at every point of Fréchet differentiability of f, but so that f is Gâteaux differentiable at the origin and

$$f_1'(0)(e_1) + f_2'(0)(e_2) + f_3'(0)(e_3) = 1.$$

(Here $\{e_i\}_{i=1}^3$ are some vectors in ℓ_2 and $f_i'(x)$ is the i'th component of the derivative of f which is, by definition, a linear map from ℓ_2 to $\mathbb{R}^3$). A similar example is given in [6] for mappings from ℓ_p into $\mathbb{R}^2$ if $1 < p < 2$.

(We mention here that in [6] there is a bad misprint which, unless corrected, makes the proof unreadable. On page 227 in the statement of Lemma 3 and in many places in pages 228 and 229 there is a meaningless symbol $g \simeq$ in the formulas. This symbol should be replaced everywhere by $g\left(\frac{\|\pi_m z\|}{r_m}\right)$.)

The organization of this paper is as follows:

In Sect. 2 we prove some preliminary functional analytic results needed to show the general form of our theorem. For the particular case of separable Asplund space one needs Proposition 1 and Remark 1 only under the additional assumption that X^* is separable; under this assumption a simple proof is provided by Remark 2.

In Sect. 3 we present the proof of a rather technical lemma concerning Lipschitz functions from $\mathbb{R}^2$ to $\mathbb{R}$.

In Sect. 4 we present the algorithm for finding the promised point of Fréchet differentiability of f.

In Sect. 5 we prove (using the result of Sect. 3) that the algorithm presented in Sect. 4 really works.

2. Functional analytic preliminaries

The first result in this section shows how the assumption in the statement of the theorem, that A has a norm separable w^* closure, will be applied. This result is related to several known results concerning the Radon-Nikodým property, but we are not aware of a reference in which the same result is proved. In any case, in order to make the proof of the theorem here self-contained, we prove the proposition in detail.

Proposition 1. *Let A a be a nonempty bounded subset of X^* so that the w^* closure of A is norm separable. Then for every $x \neq 0$ in X and every $0 < \varepsilon < \|x\|$ there is a w^* slice of A which is determined by a vector y with $\|y - x\| \leq \varepsilon$ and has a diameter at most ε.*

Proof. Assume that the assertion is false for some x and $0 < \varepsilon < \|x\|$ and assume also that A is contained in the unit ball of X^*.

Let $\sigma = \sup\{f(x) : f \in A\}$, let $\eta = \varepsilon/2$, $\delta = \eta\varepsilon/3$ and put $S = \{f \in A : f(x) \geq \sigma - \delta - 2\eta\}$. By our assumption there are $g_1, g_{-1} \in \{f \in A :$

$f(x) > \sigma - \delta\}$ so that $\|g_1 - g_{-1}\| > \varepsilon$; i.e. there is a u in the unit sphere of X so that $g_1(u) - g_{-1}(u) > \varepsilon$.

Consider next the set $\{f \in A : f(x + \eta u) > \sigma - \delta + \eta g_1(u)\}$. This set is non empty (it contains g_1) and is contained in S (since $f(x) > \sigma - \delta + \eta(g_1 - f)(u) \geq \sigma - \delta - 2\eta$). Also, for f in this set

$$\sigma + \eta f(u) \geq f(x + \eta u) > \sigma - \delta + \eta g_1(u)$$

and hence

$$f(u) > g_1(u) - \delta/\eta = g_1(u) - \varepsilon/3.$$

Let now $\sigma_1 = \sup\{f(x + \eta u) : f \in A\}$, let δ_1 and η_1 be such that $\eta_1 < \varepsilon/4$, $\eta_1 = 3\delta_1/\varepsilon$, and

$$\sigma_1 - \delta_1 - 2\eta_1 > \sigma - \delta + \eta g_1(u),$$

and put

$$S_1 = \{f \in A : f(x + \eta u) \geq \sigma_1 - \delta_1 - 2\eta_1\}.$$

From the discussion above it follows that the w^* slice S_1 satisfies $S_1 \subset S$ and for $f \in S_1$ we have $f(u) > g_1(u) - \varepsilon/3$.

In a similar way we define $\sigma_{-1} = \sup\{f(x - \eta u) : f \in A\}$ and choose positive δ_{-1} and η_{-1} so that $\eta_{-1} < \varepsilon/4$, $\eta_{-1} = 3\delta_{-1}/\varepsilon$,

$$\sigma_{-1} - \delta_{-1} - 2\eta_{-1} > \sigma - \delta - \eta g_{-1}(u),$$

and put

$$S_{-1} = \{f \in A : f(x - \eta u) \geq \sigma_{-1} - \delta_{-1} - 2\eta_{-1}\}.$$

Then $S_{-1} \subset S$ and for $f \in S_{-1}$, $f(u) < g_{-1}(u) + \varepsilon/3$. In particular if $f \in S_1$ and $\tilde{f} \in S_{-1}$ then $f(u) - \tilde{f}(u) > \varepsilon/3$.

Since $\|x + \eta u - x\| = \eta = \varepsilon/2 < \varepsilon$ it follows from our assumptions that there are $g_{1,1}$ and $g_{1,-1}$ in A so that $g_{1,1}(x + \eta u), g_{1,-1}(x + \eta u) > \sigma_1 - \delta_1$ and $g_{1,1}(u_1) - g_{1,-1}(u_1) > \varepsilon$ for some vector u_1 in the unit sphere of X. Similarly, we choose $g_{-1,1}$, and $g_{-1,-1}$ in A and u_{-1} in the unit sphere of X so that $g_{-1,1}(x - \eta u), g_{-1,-1}(x - \eta u) > \sigma_{-1} - \delta_{-1}$ and $g_{-1,1}(u_{-1}) - g_{-1,-1}(u_{-1}) > \varepsilon$.

We continue in an obvious way and define for every k-tuple $\theta_1, \theta_2, \cdots, \theta_k$ of signs ($k = 1, 2, \cdots$) elements $g_{\theta_1,\cdots,\theta_k} \in A$, vectors $u_{\theta_1,\cdots,\theta_k}$ in the unit sphere of X, numbers $0 < \eta_{\theta_1,\cdots,\theta_k} < 2^{-k}\varepsilon$ and w^* slices $S_{\theta_1,\cdots,\theta_k}$ so that

(i) $g_{\theta_1,\cdots,\theta_k} \in S_{\theta_1,\cdots,\theta_{k-1}}$,
(ii) $S_{\theta_1,\cdots,\theta_{k+1}} \subset S_{\theta_1,\cdots,\theta_k}$,
(iii) the slice $S_{\theta_1,\cdots,\theta_k}$ is determined by the vector $x + \theta_1\eta u + \theta_2\eta_{\theta_1}u_{\theta_1} + \cdots + \theta_k\eta_{\theta_1,\cdots,\theta_{k-1}}u_{\theta_1,\cdots,\theta_{k-1}}$,
(iv) $f \in S_{\theta_1,\cdots,\theta_k,1}, \tilde{f} \in S_{\theta_1,\cdots,\theta_k,-1} \Rightarrow f(u_{\theta_1,\cdots,\theta_k}) - \tilde{f}(u_{\theta_1,\cdots,\theta_k}) > \varepsilon/3.$

Note that the distance of the vector appearing in (iii) from x is smaller than ε. For every infinite sequence of signs $\{\theta_i\}_{i=1}^{\infty} = \theta$ let g_θ be any w^* limit point of the sequence $\{g_{\theta_1,\cdots,\theta_i}\}_{i=1}^{\infty}$. If $\theta \neq \tilde{\theta}$ i.e. if $\theta_i = \tilde{\theta}_i$ for $i \leq k$ and $\theta_{k+1} \neq \tilde{\theta}_{k+1}$ then

$$\|g_\theta - g_{\tilde{\theta}}\| > |(g_\theta - g_{\tilde{\theta}})(u_{\theta_1,\cdots,\theta_k})| \geq \varepsilon/3$$

and this contradicts the separability of the w^* closure of A.

Remark 1. If the w^* closure of A is separable the same clearly holds for every subset of A. It follows that any subset of A has w^* slices of arbitrarily small diameter which are determined by vectors arbitrarily close to any given vector. This assertion is exactly what is used concerning A in the proof of the theorem below.

Remark 2. If we make the stronger assumption that X^* itself is separable then the proof of Proposition 1 becomes much simpler. Let $p(x) = \sup\{f(x) : f \in A\}$. This is a continuous convex function and if X^* is separable, it is Fréchet differentiable at a dense set of vectors x. It is trivial to check that whenever p is Fréchet differentiable at some point x, then this x determines w^* slices of A of arbitrarily small diameter (cf. [3, Lemma 2.18]).

The second proposition we present in this section is a tool for reducing differentiability problems to the separable case. In particular, it implies immediately the main result of [4] in the non separable case, once we know the theorem stated in Sect. 1. This reduction result appears already in [5, Theorem 1]. We reproduce here its proof for the sake of completeness.

Proposition 2. *Let X and Y be Banach spaces, G an open set in X and $f : G \to Y$ a Lipschitz function. Then for every separable subspace V of X there is a separable subspace W of X containing V which has the following property: If $f_{|G\cap W}$ is Fréchet differentiable at some $x \in G \cap W$ then f itself if Fréchet differentiable at the same point.*

Proof. We note first that f is Fréchet differentiable at x if (and obviously only if) for every $\varepsilon > 0$ there is a $\delta > 0$ such that whenever $u, v \in X$ with $\|u\| = \|v\| = 1$ and whenever s and t satisfy $0 < |s|, |t| < \delta$ then

$$\left\| \begin{array}{l} (f(x + t(u + v)) - f(x))/t \\ \quad - (f(x + su) - f(x))/s - (f(x + sv) - f(x))/s \end{array} \right\| \leq \varepsilon. \qquad (\dagger)$$

Indeed, by taking $u = v$ we deduce that $Tu = \lim_{t\to 0}(f(x + tu) - f(x))/t$ exists. It also follows from $(\dagger)$ that T is linear and bounded. Using $(\dagger)$ again for $u = v$, fixed s and $t \to 0$ we get

$$\|Tu - (f(x + su) - f(x))/s\| \leq \varepsilon/2$$

for all u with the norm 1 and $|s| < \delta$. Thus T is the Fréchet derivative of f at x.

For every $x \in G$ we choose a countable subset $A(x)$ with the following property: Whenever s and t are non zero rational numbers with $2\max(|s|, |t|) < \mathrm{dist}(x, X \setminus G)$ there are $u, v \in A(x)$ so that $\|u\| = \|v\| = 1$ and

$$\begin{aligned} &\big\|(f(x + t(u + v)) - f(x))/t \\ &\quad -(f(x + su) - f(x))/s - (f(x + sv) - f(x))/s\big\| \\ &\geq -|s| + \sup\big\{\big\|(f(x + t(z + w)) - f(x))/t - (f(x + sz) - f(x))/s \\ &\quad -(f(x + sw) - f(x))/s\big\| : z, w \in X, \|z\| = \|w\| = 1\big\}. \end{aligned}$$

We now define inductively an increasing sequence of separable subspaces of X. Put $V_0 = V$. Once V_n has been defined let C_n be a dense sequence in it and let V_{n+1} be the closed linear span of $C_n \cup \bigcup\{A(x) : x \in C_n \cap G\}$. We claim that $W = \overline{\bigcup_{r=0}^{\infty} V_n}$ has the desired property. Let $x \in G \cap W$ and assume that $f_{|W}$ is Fréchet differentiable at x with $\tilde{T}$ as its derivative there. Then for every $\varepsilon > 0$ there is a $0 < \delta(\varepsilon) < \min(\varepsilon, \mathrm{dist}(x, X \setminus G))$ so that for every $z \in W$ with $\|z - x\| < \delta(\varepsilon)$

$$\|f(z) - f(x) - \tilde{T}(z - x)\| < \varepsilon\|z - x\|/28.$$

If f is not Fréchet differentiable at x it would follow from the observation in the beginning of the proof that there are $u, v \in X$ of norm 1, an $\varepsilon > 0$ and rational s, t, with $0 < |s|, |t| < \delta(\varepsilon)/4$ so that (†) fails to hold. From the definition of W it follows that there is a $y \in \bigcup_{n=1}^{\infty} C_n$ so that

$$\begin{aligned} \|y - x\| &< \min(|s|, |t|) < \delta(\varepsilon)/4, \\ (1/|t| + 2/|s|)\|y - x\| \, \|\tilde{T}\| &< \varepsilon/4, \\ (1/|t| + 2/|s|)\|f(y) - f(x)\| &< \varepsilon/4 \end{aligned}$$

and

$$\|(f(y + t(u + v)) - f(y))/t - (f(y + su) - f(y))/s - (f(y + sv) - f(y))/s\| > \varepsilon.$$

Since $A(y) \subset W$ and by the definition of $A(y)$, there are $u', v' \in W$ of norm 1 so that

$$\begin{aligned} &\big\|(f(y + t(u' + v')) - f(y))/t - (f(y + su') - f(y))/s \\ &\quad - (f(y + sv') - f(y))/s\big\| > \varepsilon - |s| > 3\varepsilon/4. \end{aligned}$$

On the other hand the left-hand side of the preceding inequality does not exceed

$$\left\| (f(y + t(u' + v')) - f(x))/t - (f(y + su') - f(x))/s \right.$$
$$- (f(y + sv') - f(x))/s \left\| + (1/|t| + 2/|s|) \right\| f(y) - f(x) \|$$
$$< \| \tilde{T}((y + t(u' + v') - x)/t - (y + su' - x)/s - (y + sv' - x)/s) \|$$
$$+ \varepsilon \| y + t(u' + v') - x \| /28|t|$$
$$+ \varepsilon(\| y + su' - x \| + \| y + sv' - x \|)/28|s| + \varepsilon/4$$
$$< (1/|t| + 2/|s|) \| y - x \| \| \tilde{T} \| + \varepsilon(4 + \| y - x \| /|t| + 2\| y - x \| /|s|)/28$$
$$+ \varepsilon/4 < 3\varepsilon/4.$$

Thus we are led to a contradiction and the proposition is proved.

3. Measure theoretic lemmas

The purpose of this section is to prove Lemma 3 below. We start with the well known result that the so called Hardy-Littlewood maximal operator is of weak type (1,1). We will denote by m the Lebesgue measure on the line.

Lemma 1. *Let $g \in L_1(\mathbb{R})$ then for every $\lambda > 0$*

$$m\{t : Mg(t) \geq \lambda\} \leq 2/\lambda \int |g(t)| dt$$

where

$$Mg(t) = \sup \left\{ \int_I |g(s)| ds \cdot m(I)^{-1} : I \text{ is an interval containing } t \right\}.$$

Consequently, if $|g(t)| \leq 1$ for every t then Fubini's Theorem and the bound $0 \leq Mg(t) \leq 1$ give

$$\int Mg(t)^2 dt = \int_0^1 m\{t : Mg(t)^2 > s\} ds = \int_0^1 m\{t : Mg(t) > \sqrt{s}\} ds$$
$$\leq \int_0^1 2/\sqrt{s} \int |g(t)| dt ds = 4 \int |g(t)| dt.$$

The next lemma is a slight variant of a lemma used already in [4, Lemma 3.3].

Lemma 2. *Suppose $a < b$ are real numbers and $h: [a, b] \to \mathbb{R}$ a Lipschitz function whose Lipschitz constant is ≤ 1. Assume that*

$$\int_a^b |h'(t)| dt \geq 2|h(b) - h(a)|.$$

Then there is a measurable subset $A \subset (a, b)$ so that

(i) $m(A) \geq \frac{1}{16} \int_a^b |h'(t)| dt,$

(ii) $h'(s) \geq \int_a^b |h'(t)| dt/8(b-a)$ *for every* $s \in A,$

(iii) $|h(t) - h(s)| \leq 8\sqrt{h'(s)}|t - s|$ *for every* $s \in A$ *and* $t \in [a, b].$

Proof. Let $g(t) = h'(t)$ whenever the derivative exists (and take it as 0 otherwise). Put $g^+ = \max(g, 0)$ and

$$D = \left\{ t \in [a, b] : g(t) > \int_a^b |g(u)| du/8(b-a) \right\}$$

$$B = \{ t \in D : g(t) < (Mg(t))^2/64 \}$$

$$A = D \setminus B.$$

It is evident that every $s \in D$ (and thus $s \in A$) satisfies (ii). For $s \in A$ and $t \in [a, b]$

$$|h(t) - h(s)| = \left| \int_{[t,s]} g(u) du \right| \leq Mg(s)|t - s| \leq 8\sqrt{h'(s)}|t - s|$$

and thus also (iii) holds. It remains to verify (i),

$$\int_a^b |g(t)| dt = 2 \int_a^b g^+(t) dt - \int_a^b g(t) dt = 2 \int_a^b g^+(t) dt - (h(b) - h(a))$$

$$\leq 2 \int_a^b g^+(t) dt + \frac{1}{2} \int_a^b |g(t)| dt.$$

Thus

$$\int_a^b |g(t)| dt \leq 4 \int_a^b g^+(t) dt = 4 \int_D g(t) dt + 4 \int_{[a,b] \setminus D} g^+(t) dt$$

$$\leq 4 \int_D g(t) dt + \frac{1}{2} \int_a^b |g(t)| dt$$

and consequently

$$\int_a^b |g(t)| dt \leq 8 \int_D g(t) dt.$$

In view of Lemma 1,

$$\int_B g(t) dt \leq \int_B Mg(t)^2 dt/64 \leq \int_a^b |g(t)| dt/16$$

and hence

$$\int_a^b |g(t)| dt \leq 8 \int_A g(t) dt + 8 \int_B g(t) dt \leq 8 \int_A g(t) dt + \frac{1}{2} \int_a^b |g(t)| dt$$

or

$$\int_a^b |g(t)| dt \leq 16 \int_A g(t) dt \leq 16m(A).$$

We come now to the lemma which is the purpose of the present section. The lemma has a rather special and quite involved statement. It is tailor-made to its application in Sect. 5. The proof of the lemma is however not very complicated. It uses just the previous two lemmas and Fubini's theorem in the plane.

Lemma 3. *Let $0 < \kappa, \eta, \tau < 1$ with $\eta \geq \tau$, let $r, s > 0$ with $s \leq 10^{-6}\kappa\eta\tau r$. Let $g \colon [-2r, 2r] \times [0, s] \to \mathbb{R}$ be a Lipschitz function which satisfies*

$$|g(0, s)| > \kappa s \quad \text{and} \quad |g(u, 0)| \leq \kappa s/32 \text{ for all } u \in [-2r, 2r]$$

and is such that both partial derivatives $g_1'(u, v)$ and $g_2'(u, v)$ are a.e. of absolute value at most 1. Finally, let $N \subset [-r, r] \times [0, s]$ be a set of (Lebesgue) measure 0.

Then there is a $(u, v) \in [-r, r] \times [0, s] \setminus N$ so that

(i) $g_1'(u, v) > -\tau$
(ii) $g_1'(u, v) > \varepsilon_1 = \kappa s/128r$ *or* $|g_2'(u, v)| > \varepsilon_2 = \kappa/16$
(iii) $|g(u + t, v) - g(u, v)| \leq (\eta + 8\sqrt{g_1'(u, v)^+})|t|$, *if* $|t| \leq r$.

Proof. We may clearly assume that N contains all points of non differentiability of g in $[-r, r] \times [0, s]$. Note that by the assumptions on g

$$|g(u, v)| \leq 2s \text{ for all } (u, v) \in [-2r, 2r] \times [0, s].$$

Let I be the set of all $v \in [0, s]$ for which $N_v = \{u : (u, v) \in N\}$ has measure 0, let $C_1 = 4/\eta$ and $k = 32$. If $\int_{-r}^{r} |g_1'(u, v)|du > kC_1s$ for some $v \in I$ then by Lemma 2 there is a $u \in [-r + C_1s, r - C_1s] \setminus N_v$ so that $g_1'(u, v) \geq C_1s/r$ (which is clearly $> \varepsilon_1$) and

$$|g(u + t, v) - g(u, v)| \leq 8\sqrt{g_1'(u, v)}|t| \text{ whenever } |u + t| \leq r.$$

If $|t| \leq r$ but $|u + t| \geq r$ then by the choice of u, $|t| \geq C_1s$ and

$$|g(u + t, v) - g(u, v)| \leq 4s \leq 4|t|/C_1 \leq \eta|t|$$

and therefore the point (u, v) we found has the desired properties. Thus we may assume from now on that

$$\int_{-r}^{r} |g_1'(u, v)|du \leq kC_1s \text{ for all } v \in I.$$

Let M be the set of points $(u, v) \in [-r, r] \times I$ such that at least one of the following happens

(a) $g_1'(u, v) \leq -\tau$,
(b) There is a t with $|t| \leq r$ so that $|g(t, v) - g(u, v)| > \eta|t - u|$.

By Lemma 1 and Chebyshev's inequality for every $v \in I$,

$$m\{u : (u, v) \in M\} \le kC_1 s/\tau + 2kC_1 s/\eta = C_2 s.$$

By Fubini's theorem there is an $r_0 \in \left[\frac{r}{4}, \frac{r}{2}\right]$ so that the set I_1 of those $v \in I$ such that $(r_0, v) \in M \cup N$ or $(-r_0, v) \in M \cup N$ has measure at most $c_2 s = 8C_2 s^2/r$. Note that by the definition of C_2 and the assumption on s in the statement of the lemma $c_2 < \varepsilon_2/2$.

Every $(u, v) \in D = [-r_0, r_0] \times [0, s] \setminus (M \cup N)$ clearly satisfies (i) in the statement of the Lemma. It also satisfies (iii), even in the stronger form

$$|g(u + t, v) - g(u, v)| \le \eta|t| \text{ if } |t| \le r. \tag{$\star$}$$

Indeed, this is clear by the definition of M if $|u + t| \le r$. If $|t| \le r$ and $|u + t| > r$ then since $|u| \le r_0 \le r/2$ we deduce that $|t| > r/2$. But then

$$|g(u + t, v) - g(u, v)| \le 4s \le (8s/r)(r/2) \le \eta|t|$$

as desired. Consequently, if for some $(u, v) \in D$, $|g_2'(u, v)| > \varepsilon_2$ this point will satisfy all the requirements in the lemma and the proof will be finished.

We can thus assume that $|g_2'(u, v)| \le \varepsilon_2$ for all $(u, v) \in D$. Hence, in particular, for all $v \in [0, s]$,

$$
\begin{aligned}
|g(\pm r_0, v)| &\le |g(\pm r_0, 0)| + \int_0^s |g_2'(\pm r_0, t)| dt \\
&\le \kappa s/32 + \int_{I_1} |g_2'(\pm r_0, t)| dt + \int_{[0,s] \setminus I_1} |g_2'(\pm r_0, t)| dt \\
&\le \kappa s/32 + c_2 s + \varepsilon_2 s \le 2\varepsilon_2 s.
\end{aligned}
$$

Since $m(I_1) \le c_2 s$ there is a $v \in (I \setminus I_1) \cap [(1 - 2c_2)s, s]$ which we fix from now on. We have

$$
\begin{aligned}
\kappa s < |g(0, s)| &\le |g(0, v)| + 2c_2 s \\
&\le |g(r_0, v)| + \int_{-r_0}^{r_0} |g_1'(t, v)| dt + 2c_2 s \\
&\le \int_{-r_0}^{r_0} |g_1'(t, v)| dt + 2(\varepsilon_2 + c_2)s \\
&\le \int_{-r_0}^{r_0} |g_1'(t, v)| dt + \kappa s/2.
\end{aligned}
$$

Hence $\int_{-r_0}^{r_0} |g_1'(t, v)| dt > \kappa s/2$. Since also

$$|g(r_0, v) - g(-r_0, v)| \le 4\varepsilon_2 s \le \kappa s/4$$

we deduce from Lemma 2 that there is a $u \in [-r_0, r_0] \setminus N_v$ so that $g_1'(u, v) > \kappa s/(32 r_0) \geq \varepsilon_1$ and

$$|g(u + t, v) - g(u, v)| \leq 8\sqrt{g_1'(u, v)}|t|, \text{ whenever } u + t \in [-r_0, r_0].$$

In order to verify that for this choice of (u, v) assertion (iii) also holds, we have to consider only points t with $|t| \leq r$ and $|u + t| > r_0$. Let $\tilde{t}$ be of the same sign as t so that $|u + \tilde{t}| = r_0$. Since $v \in I \setminus I_1$ it follows that $(u + \tilde{t}, v) \in D$ and hence, using $(\star)$,

$$|g(u + t, v) - g(u, v)| \leq |g(u + t, v) - g(u + \tilde{t}, v)| + |g(u + \tilde{t}, v) - g(u, v)|$$
$$\leq \eta|t - \tilde{t}| + 8\sqrt{g_1'(u, v)}|\tilde{t}| \leq (\eta + 8\sqrt{g_1'(u, v)})|t|.$$

4. The algorithm

In this section we describe the algorithm we use for proving the theorem stated in the introduction. We use the notation appearing in the statement of the theorem. We just add now for convenience the assumption that the Lipschitz constant of f is at most $\frac{1}{2}$ (and thus also $\|f'(x)\| \leq \frac{1}{2}$ for all $x \in D$).

We denote $e_0 = (v - u)/\|v - u\|$, $\sigma_0 = \sup\{f'(x)(e_0) : x \in D\}$, $\tau_0 = (\sigma_0 - m/\|u - v\|)/3$ and $c_0 = \sigma_0 - 2\tau_0$. Since f is Gâteaux differentiable outside a set of Gaussian measure 0 we can replace the segment $[u, v]$ by a very close parallel segment $[u', v']$ so that f is Gâteaux differentiable a.e. on $[u', v']$ and infer that there is an $\hat{x} \in D$ such that $f'(\hat{x})(e_0) > m/\|u - v\|$. Hence $\tau_0 > 0$ and we may find $x_0 \in D$ so that $f'(x_0)(e_0) > \sigma_0 - 10^{-10}\tau_0$. We let $0 < \delta_0 < 1$ be such that the open ball $B(x_0, \delta_0)$ (with center x_0 and radius δ_0) belongs to G. We let

$$D_0 = \{x : x \in D \cap B(x_0, \delta_0^2) \text{ and } f'(x)(e_0) > c_0\}.$$

All the further points we choose will be in D_0 and note that $\{f'(x) : x \in D_0\}$ is a subset of a slice of A defined by the vector e_0.

We let $\kappa_k = 2^{-k}$ for $k = 0, 1, 2, \ldots$ and will define inductively (in this order) numbers $\alpha_k, \gamma_k, \eta_k$, unit vectors e_k, numbers σ_k, c_k, τ_k, points x_k, numbers ζ_k, δ_k and sets D_k. We will do this choice so that in particular the following hold:

(1) $\qquad\qquad x_j \in D_j$ and $D_j \subset D_{j-1}$ for all $j \geq 0$,

(2) $\qquad\qquad 2\tau_j \leq f'(x_j)(e_j) - c_j$ for $j \geq 0$,

(3) $\quad$ for every $x \in D_j$ there is $\beta = \beta(x, j) < \alpha_j$ such that $\quad$ $|f(x + te_j) - f(x) - tf'(x_j)(e_j)| < \beta|t|$ for all $|t| < \delta_j$.

For $k = 0$ we define the values we have not defined as yet by $\alpha_0 = 2$, $\zeta_0 = 1/2$ and $\gamma_0 = \eta_0 = 1$. Obviously, (1) and (2) hold for $j = 0$ (where D_{-1} means D), and (3) is satisfied with $\beta = 3/2$.

Let $k \geq 1$ and assume that all the above have been defined for values smaller than k. We put

$$(4) \qquad \alpha_k = 10^{-8} \min \left(\kappa_{k-1}^2 \tau_{k-1} \eta_{k-1}, \alpha_{k-1} \right)$$

$$(5) \qquad \gamma_k = 10^{-8} \min \left(\tau_{k-1}, \alpha_k^2, \gamma_{k-1} \right)$$

$$(6) \qquad \eta_k = 10^{-8} \min(\kappa_k, \alpha_k).$$

By Proposition 1 there is a unit vector $e_k \in X$ with $\|e_k - e_{k-1}\| < \gamma_k$ which defines a slice of the set $\{f'(x) : x \in D_{k-1}\}$ of diameter less than η_k. Put

$$\sigma_k = \sup\{f'(x)(e_k) : x \in D_{k-1}\}.$$

Since $\|f'(x)\| \leq 1/2$ for all $x \in D$,

$$(7) \qquad \sigma_k \leq \sigma_{k-1} + \|e_k - e_{k-1}\|/2$$

and, since $x_{k-1} \in D_{k-1}$ ((1) for $j = k - 1$),

$$(8) \qquad \sigma_k \geq f'(x_{k-1})(e_k) \geq f'(x_{k-1})(e_{k-1}) - \|e_k - e_{k-1}\|/2.$$

By (2) (for $j = k - 1$) and (5),

$$\|e_k - e_{k-1}\| < \gamma_k \leq \tau_{k-1} < f'(x_{k-1})(e_{k-1}) - c_{k-1}$$

and thus (8) implies that $\sigma_k > (f'(x_{k-1})(e_{k-1}) + c_{k-1})/2$. Hence we may choose c_k so that

$$(9) \qquad \sigma_k > c_k > (f'(x_{k-1})(e_{k-1}) + c_{k-1})/2,$$

$$(10) \qquad \sigma_k - c_k < 10^{-3} \alpha_k^2,$$

and so that

(11) the diameter of $\{f'(x) : x \in D_{k-1}, f'(x)(e_k) > c_k\}$ is at most η_k.

For future reference we record that by (9) and by (2) (for $j = k - 1$)

$$(12) \qquad c_k - c_{k-1} > (f'(x_{k-1})(e_{k-1}) - c_{k-1})/2 \geq \tau_{k-1}.$$

We now take $\tau_k = \min((\sigma_k - c_k)/4, \eta_k)$ and a point $x_k \in D_{k-1}$ such that

$$(13) \qquad f'(x_k)(e_k) > \max \left(\sigma_k - 10^{-10} \kappa_k^2 \tau_k \eta_k, (\sigma_k + c_k)/2 \right).$$

These two choices ensure in particular that (2) holds for $j = k$. Combining (13) with (7) and (8), we infer that

$$(14) \qquad |f'(x_k)(e_k) - f'(x_{k-1})(e_{k-1})| < 10^{-9}\kappa_{k-1}^2 \tau_{k-1}\eta_{k-1}.$$

By (3) (for $j = 1, \ldots, k - 1$) there is a $\zeta_k > 0$ so that

$$(15) \qquad \begin{aligned} &|f(x_k + te_j) - f(x_k) - tf'(x_j)(e_j)| \le (\alpha_j - \zeta_k)|t| \\ &\text{for all } 1 \le j \le k - 1 \text{ and } |t| \le \delta_j. \end{aligned}$$

Next we choose $\delta_k > 0$ so that $\delta_k < \min(\zeta_k/4, \delta_{k-1}^2)$, $B(x_k, 2\delta_k^2) \subset B(x_{k-1}, \delta_{k-1}^2)$ and

$$(16) \qquad |f(x_k + te_k) - f(x_k) - tf'(x_k)(e_k)| \le 10^{-3}\alpha_k|t| \text{ for } |t| \le \delta_k.$$

Finally we let D_k be those points x in $D_{k-1} \cap B(x_k, \delta_k^2)$ which satisfy $f'(x)(e_k) > c_k$ and for which there is a $\beta = \beta(x, k)$ with $0 < \beta < \alpha_k$ so that

$$|f(x + te_k) - f(x) - tf'(x_k)(e_k)| \le \beta|t| \text{ if } |t| \le \delta_k.$$

Clearly $x_k \in D_k$ so that (1) holds also for $j = k$. By the definition of D_k also (3) holds for $j = k$. Moreover, (11) implies that

$$(17) \qquad \text{the diameter of } \{f'(x) : x \in D_k\} \text{ is at most } \eta_k.$$

Note that we do not claim at this stage that the set D_k is in any sense large. It may very well consist just of one point x_k. Still our inductive choice of all parameters works. Of course once a set D_j reduces to a single point x_j then $x_k = x_j$ for $k \ge j$ and all the sets D_k with $k \ge j$ consist of the single point x_k.

Since $\delta_k \to 0$ as $k \to \infty$ it is clear that $x = \lim_{k\to\infty} x_k$ exists. To prove that this point has the desired property we shall prove (in the next section) the following lemma.

Lemma 4. *There are numbers $w_j > 0$ and $\Delta_j > 0$ for $j = 0, 1, 2, \ldots$ with $w_j \to 0$ so that for every $k > j \ge 0$*

$$(18) \qquad |f(x_k + y) - f(x_k) - f'(x_j)(y)| \le w_j\|y\| \text{ whenever } \|y\| \le \Delta_j$$

where the $\{x_k\}$ are the points constructed above.

Let us show now the simple fact that Lemma 4 implies the theorem. First note that it follows from (18) that for any two positive integers j_1 and j_2, $\|f'(x_{j_1}) - f'(x_{j_2})\| \le w_{j_1} + w_{j_2}$ and hence $x^* = \lim_j f'(x_j)$ exists and $\|x^* - f'(x_j)\| \le w_j$. Consequently, $|f(x_k + y) - f(x_k) - x^*(y)| \le 2w_j\|y\|$ whenever $\|y\| \le \Delta_j$ and $k > j$. Keeping j fixed and letting $k \to \infty$, we infer that

$$|f(x + y) - f(x) - x^*(y)| \le 2w_j\|y\| \text{ if } \|y\| \le \Delta_j.$$

Thus f is Fréchet differentiable at x and $f'(x) = x^*$. Since all the points x_j belong to the set D_0 we get that $f'(x) = \lim_j f'(x_j)$ satisfies $f'(x)(e_0) \geq c_0$, hence $f'(x)(v - u) > m$ as desired.

5. The proof of Lemma 4

In this section we prove Lemma 4 and this will conclude the proof of the theorem. We keep the notation of the previous section and claim that suitable w_j and Δ_j ($j = 0, 1, \dots$) for which (18) holds are given by

$$w_j = 10^3(\kappa_j + \alpha_j)$$
$$\Delta_j = 10^{-9}\kappa_j\tau_j\eta_j\delta_{j+1}\zeta_{j+1}.$$

Indeed, assume that (18) is false for some $k > j \geq 0$. Since $w_0 \geq 1$ it is clear that $j \geq 1$. Let $y \in X$ with $\|y\| \leq \Delta_j$ be such that

$$|f(x_k + y) - f(x_k) - f'(x_j)(y)| > w_j\|y\|.$$

By moving y slightly we may assume that at almost every point of the intersection of G and the affine span of $x_k, x_k + y, x_k + e_j$ the function f is Gâteaux differentiable.

We next pass into a setting where we can apply Lemma 3. Let $s = \|y\|$, $e = y/s$, $r = 10^6 s/\kappa_j\tau_j\eta_j$ and define $g \colon [-2r, 2r] \times [0, s] \to \mathbb{R}$ by

$$g(u, v) = f(x_k + ue_j + ve) - f(x_k) - f'(x_j)(ue_j + ve).$$

We intend to apply Lemma 3 with the given choices of r, s, g, with $\kappa = \kappa_j$, $\tau = \tau_j$, $\eta = \eta_j$ and N the set of (u, v) such that f is not Gâteaux differentiable at $x_k + ue_j + ve$. We check first that all the assumptions of Lemma 3 are satisfied.

To estimate $g(u, 0)$ we cannot use (3) directly since this would give an estimate in terms of r and not in terms of s. We therefore use (14) to replace $f'(x_j)(e_j)$ by $f'(x_{j+1})(e_{j+1})$. We proceed as follows:

Since

$$2r = 2 \cdot 10^6 s/\kappa_j\tau_j\eta_j \leq 2 \cdot 10^6\Delta_j/\kappa_j\tau_j\eta_j \leq \delta_{j+1}$$

we deduce, using (3), (4), (5) and (14), that for $|u| \leq 2r$

$$\begin{aligned}
|g(u, 0)| &\leq |f(x_k + ue_{j+1}) - f(x_k) - f'(x_{j+1})(ue_{j+1})| \\
&\quad + (\|e_{j+1} - e_j\|/2 + |f'(x_{j+1})(e_{j+1}) - f'(x_j)(e_j)|)|u| \\
&\leq (\alpha_{j+1} + \gamma_{j+1}/2 + 10^{-9}\kappa_j^2\tau_j\eta_j)|u| \\
&\leq 2r(10^{-8}\kappa_j^2\tau_j\eta_j + 2 \cdot 10^{-9}\kappa_j^2\tau_j\eta_j) \\
&= 2s(10^{-2} + 2 \cdot 10^{-3})\kappa_j \leq \kappa_j s/32.
\end{aligned}$$

It remains to note that

$$|g(0, s)| = |f(x_k + y) - f(x_k) - f'(x_j)(y)| > w_j\|y\| \geq \kappa_j s.$$

Having verified that all the assumptions of Lemma 3 hold we use it to find a point $(u, v) \in [-r, r] \times [0, s]$ so that

(19) f is Gâteaux differentiable at $x = x_k + ue_j + ve$

(20) $g_1'(u, v) > -\tau_j$

(21) $g_1'(u, v) > \kappa_j s/128r$ or $|g_2'(u, v)| > \kappa_j/16$

(22) $|g(u + t, v) - g(u, v)| \leq (\eta_j + 8\sqrt{g_1'(u, v)^+})|t|$ for $|t| \leq r.$

We now make the following

Claim. $x \in D_{j-1}$ and $f'(x)(e_j) > c_j$.

We show first that this will conclude the proof. Indeed, by (13) and (21) we have either

$$\begin{aligned}
\sigma_j &\geq f'(x)(e_j) = f'(x_j)(e_j) + g_1'(u, v) \\
&\geq f'(x_j)(e_j) + \kappa_j s/128r = f'(x_j)(e_j) + 10^{-6}\kappa_j^2\tau_j\eta_j/128 > \sigma_j
\end{aligned}$$

which is a contradiction, or by (11), (13) and (21)

$$\kappa_j/16 < |g_2'(u, v)| \leq \|f'(x_j) - f'(x)\| \leq \eta_j$$

and this again is a contradiction in view of (6).

It remains to prove the claim. Note first that by (2), $f'(x_j)(e_j) > c_j + \tau_j$ and by (5) and (12), $\gamma_i \leq \tau_{i-1} \leq c_i - c_{i-1}$ for all $i \geq 1$. Hence, for $0 \leq i \leq j$

$$\begin{aligned}
f'(x)(e_i) &\geq f'(x)(e_j) - \|e_i - e_j\| = f'(x_j)(e_j) + g_1'(u, v) - \|e_i - e_j\| \\
&\geq f'(x_j)(e_j) - \tau_j - \sum_{k=i}^{j-1} \gamma_{k+1} \geq f'(x_j)(e_j) - \tau_j - c_j + c_i > c_i.
\end{aligned}$$

This proves the second assertion in the claim and also that $x \in D_0$, since $\|x - x_k\| \leq r + s \leq \delta_{j+1}, \|x_k - x_j\| < \delta_j^2$ and $B(x_j, 2\delta_j^2) \subset B(x_0, \delta_0^2)$.

It remains to prove that $x \in D_i$ for $i = 1, 2, \ldots, j - 1$ and we do this by induction on i. Let $i \leq j - 1$ and assume that $x \in D_{i-1}$. What has to be

shown in order to ensure that $x \in D_i$ is that there is a $\beta_i < \alpha_i$ so that

$$|f(x + te_i) - f(x) - tf'(x_i)(e_i)| \leq \beta_i|t| \text{ for all } |t| \leq \delta_i.$$

We shall verify this by treating separately three different ranges of t.

Case I, $|t| \leq r$. Since $x \in D_{i-1}$ and since $f'(x_j)(e_j) > c_j > c_i$ (by (2) and (12)) we have

$$g'_1(u, v) = f'(x)(e_j) - f'(x_j)(e_j) \leq f'(x)(e_i) - f'(x_j)(e_j) + \frac{1}{2}\|e_j - e_i\|$$

$$\leq \sigma_i - c_i + \frac{1}{2}\sum_{k=i+1}^{j} \gamma_i < \sigma_i - c_i + \gamma_{i+1}.$$

Hence, using that $x_i, x_j \in D_i$ and so (17) implies that $\|f'(x_j) - f'(x_i)\| \leq \eta_i$, we infer from (22) that

$$|f(x + te_i) - f(x) - tf'(x_i)(e_i)|$$
$$\leq (\|e_j - e_i\| + \|f'(x_j) - f'(x_i)\|)|t| + |f(x + te_j) - f(x) - tf'(x_j)(e_j)|$$
$$\leq \left(2\gamma_{i+1} + \eta_i + \eta_j + 8\sqrt{g'_1(u, v)^+}\right)|t|$$
$$\leq \left(2\gamma_{i+1} + 2\eta_i + 8\sqrt{\gamma_{i+1} + \sigma_i - c_i}\right)|t|$$

In view of (4), (5), (6) and (10) the number in front of $|t|$ is strictly smaller than α_i.

Case II, $r < |t| \leq \delta_{j+1}/2$. In this case, since $x_k \in D_{j+1}$, $|v| \leq s$ and $|u| \leq r$ we get

$$|f(x + te_i) - f(x) - tf'(x_i)(e_i)|$$
$$\leq (\|e_j - e_i\| + \|f'(x_i) - f'(x_{j+1})\|$$
$$\quad + |f(x + te_j) - f(x) - tf'(x_{j+1})(e_j)|)|t|$$
$$\leq (2\gamma_{i+1} + \eta_i)|t| + |f(x_k + (t + u)e_j) - f(x_k + ue_j) - tf'(x_{j+1})(e_j)| + s$$
$$\leq (4\gamma_{i+1} + \eta_i)|t|$$
$$\quad + |f(x_k + (t + u)e_{j+1}) - f(x_k + ue_{j+1}) - tf'(x_{j+1})(e_{j+1})| + s$$
$$\leq (4\gamma_{i+1} + \eta_i)|t|$$
$$\quad + |f(x_k + (t + u)e_{j+1}) - f(x_k) - (t + u)f'(x_{j+1})(e_{j+1})|$$
$$\quad + |f(x_k + ue_{j+1}) - f(x_k) - uf'(x_{j+1})(e_{j+1})| + s$$
$$\leq (4\gamma_{i+1} + \eta_i + 3\alpha_{j+1} + s/r)|t|$$
$$= (4\gamma_{i+1} + \eta_i + 3\alpha_{j+1} + 10^{-6}\kappa_j\tau_j\eta_j)|t|$$

and again the number in front of $|t|$ is strictly less than α_i.

Case III, $\delta_{j+1}/2 < |t| \leq \delta_i$. By (15) (in which k is replaced by $j + 1$ and j by i),

$$
\begin{aligned}
|f(x + te_i) &- f(x) - tf'(x_i)(e_i)| \\
&\leq |f(x + te_i) - f(x_{j+1} + te_i)| + |f(x) - f(x_{j+1})| \\
&\quad + |f(x_{j+1} + te_i) - f(x_{j+1}) - tf'(x_i)(e_i)| \\
&\leq \|x - x_{j+1}\| + (\alpha_i - \zeta_{j+1})|t| \\
&\leq \|x - x_k\| + \|x_k - x_{j+1}\| + (\alpha_i - \zeta_{j+1})|t| \\
&\leq \big(2\big(2r + \delta_{j+1}^2\big)/\delta_{j+1} + \alpha_i - \zeta_{j+1}\big)|t|.
\end{aligned}
$$

Since $r \leq 10^6 \Delta_j/\kappa_j \tau_j \eta_j$ it follows from the definitions of Δ_j and δ_{j+1} that also here the number in front of $|t|$ is strictly less than α_i.

This, finally, finishes the proof of Lemma 4 and thus of the theorem.

References

1. J. Matoušek, E. Matoušková: A highly nonsmooth norm on Hilbert space. Israel J. Math. **112**, 1–28 (1999)
2. R.R. Phelps: Gaussian null sets and differentiability of Lipschitz maps on Banach spaces. Pacific J. Math. **77**, 523–531 (1978)
3. R.R. Phelps: Convex functions, monotone operators and differentiability (Springer Lecture Notes in Mathematics, vol. 1364), Berlin, Heidelberg, New York: Springer 1993
4. D. Preiss: Differentiability of Lipschitz functions on Banach spaces. J. Funct. Anal. **91**, 312–345 (1990)
5. D. Preiss: Gâteaux differentiable functions are somewhere Fréchet differentiable. Rend. Circ. Matem. Palermo **33**, 122–133 (1984)
6. D. Preiss, J. Tišer: Two unexpected examples concerning differentiability of Lipschitz functions on Banach spaces, Gafa seminar 92–94, Lindenstrauss J., Milman V.D. (eds), Basel: Birkhäuser 1995, pp 219–238