Available online at www.sciencedirect.com
ScienceDirect
Journal of Functional Analysis 261 (2011) 1674–1710
JOURNAL OF Functional Analysis
www.elsevier.com/locate/jfa

A universal differentiability set in Banach spaces with separable dual ☆

Michael Doré, Olga Maleva *

School of Mathematics, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Received 23 March 2011; accepted 19 May 2011

Available online 8 June 2011

Communicated by G. Schechtman

Abstract

We show that any non-zero Banach space with a separable dual contains a totally disconnected, closed and bounded subset S of Hausdorff dimension 1 such that every Lipschitz function on the space is Fréchet differentiable somewhere in S.
© 2011 Elsevier Inc. All rights reserved.

Keywords: Universal differentiability set; Fréchet differentiability; Null set

1. Introduction

It is well known that there are quite strong results ensuring the existence of points of differentiability of Lipschitz functions defined on finite and infinite dimensional Banach spaces. Rademacher's theorem implies that real-valued Lipschitz functions on finite dimensional spaces are differentiable almost everywhere in the sense of Lebesgue measure. For the infinite dimensional case, Preiss shows in [12, Theorem 2.5] that every real-valued Lipschitz function defined on an Asplund[1] space is Fréchet differentiable at a dense set of points.

☆ Michael Doré acknowledges the support of the UK EPSRC postdoctoral PhD Plus programme. Olga Maleva acknowledges the support of the EPSRC grant EP/H43004/1.
* Corresponding author.
E-mail addresses: M.J.Dore@bham.ac.uk (M. Doré), O.Maleva@bham.ac.uk (O. Maleva).
[1] This is best possible as any non-Asplund space has an equivalent norm—which of course is a Lipschitz function—that is nowhere Fréchet differentiable; see [2,3].

0022-1236/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jfa.2011.05.016

A natural question then arises as to whether every "small" set S in a finite dimensional or infinite dimensional Asplund space Y gives rise to a real-valued Lipschitz function on Y not differentiable at any point of S. Let us call a subset E of the space Y a universal differentiability set if for every Lipschitz function $f : Y \to \mathbb{R}$, there exists $y \in E$ such that f is Fréchet differentiable at y.

In this paper we show that for non-zero separable Asplund spaces Y, there are always "small" subsets with the universal differentiability property, in the sense that the Hausdorff dimension of the closure of the set can be taken equal to 1. Hence, as we may also take the set to be bounded, in the case in which Y is a finite dimensional space of dimension at least 2 we recover the fact that a universal differentiability set may be taken to be compact and with Lebesgue measure zero, a fact first proved by the authors in [6].

In the case $\dim Y = 1$ it is easy to show that every Lebesgue null subset of $\mathbb{R}$ is not a universal differentiability set; see [13] and [8]. Note also that a separable Asplund space is simply a Banach space with a separable dual. For non-separable spaces Y, any set S of finite Hausdorff dimension is contained in a separable subspace $Y' \subseteq Y$; therefore the distance function $y \mapsto \operatorname{dist}(y, Y')$ is Lipschitz and nowhere differentiable on S.

We note here that for Lipschitz mappings whose codomain has dimension 2 or above, there are many open questions. For example, while Rademacher's Theorem still guarantees that for every $n \geqslant m \geqslant 2$ the set of points where a Lipschitz mapping $f : \mathbb{R}^n \to \mathbb{R}^m$ is not differentiable has Lebesgue measure zero, the answer to the question of whether there are Lebesgue null sets in $\mathbb{R}^n$ containing a differentiability point of every Lipschitz $f : \mathbb{R}^n \to \mathbb{R}^m$ is known only for $m = 2$. The answer for $n = m = 2$ is negative; see [1]. The case $n > m = 2$ is a topic of a forthcoming paper [5] where the authors, building on methods developed in [11] in their study of differentiability problems in infinite dimensional Banach spaces, construct null universal differentiability sets for planar-valued Lipschitz functions.

No similar positive results are known in the case in which the dimension of the codomain is at least 3. However, a partial result was obtained in [4] where it is proved that the union H of all "rational hyperplanes" in $\mathbb{R}^n$ has the property that for every $\varepsilon > 0$ and every Lipschitz mapping $f : \mathbb{R}^n \to \mathbb{R}^{n-1}$ there is a point in H where the function f is ε-Fréchet differentiable. Unfortunately, this is a weaker notion, and the existence of points of ε-Fréchet differentiability does not imply the existence of points of full differentiability. See also [9,10], in which the notion of ε-Fréchet differentiability is studied with the emphasis on the infinite dimensional case.

It follows from the work of Preiss in [12] that Lebesgue null universal differentiability sets exist in any Euclidean space of dimension at least 2. However there is a drawback in the construction by Preiss: any set S covered by [12, Theorem 6.4] is dense in the whole space, and simple refinements of the same approach are only capable of constructing universal differentiability sets that are still dense in some non-empty open set. This can be explained as follows. The proof in [12] makes essential use of the following sufficient condition for S to be a universal differentiability set: S is G_δ and for every $x \in S$ and $\varepsilon > 0$, there is a δ-neighbourhood N of x, for some $\delta = \delta(\varepsilon, x) > 0$, such that for every line segment $I \subseteq N$, the set contains a large portion of a path that approximates I to within $\varepsilon|I|$. Fixing $\varepsilon = 1/2$ say, a simple application of the Baire category theorem shows that one can choose $\delta(1/2, x)$ uniformly over $x \in S \cap U \neq \emptyset$, for some open U. It quickly follows that S itself is dense in U. See also [6, Introduction] for a discussion of this point.

In [7] we improve the result of [12, Corollary 6.5] by constructing, in every finite dimensional space, a compact universal differentiability set that has Hausdorff dimension 1.

The main result of the present paper is that every non-zero Banach space with separable dual contains a closed and bounded universal differentiability set of Hausdorff dimension 1; see Theorem 3.1, Remark 3.4, Lemma 3.5 and Theorem 3.10. The dimension 1 here is optimal; see Lemma 2.1. The universal differentiability set need not contain any non-constant continuous curves; in Theorem 3.10 we show that this set may in fact be chosen to be totally disconnected. In the case in which Y is a finite dimensional space, this result implies the earlier result of [7]. Note that compact subsets of infinite dimensional spaces cannot have the universal differentiability property; indeed if $S \subseteq Y$ is compact then one may even construct a Lipschitz convex function $f : Y \to \mathbb{R}$ not Fréchet differentiable on S, for example

$$f(y) := \operatorname{dist}\big(y, \text{convex hull}(S)\big).$$

See also remark after Lemma 3.8.

The proof of Lemma 3.5 is based on Theorems 3.1 and 3.3, which rely on Sections 4, 5 and 6. Section 6 gives details of the construction of the set. Section 5 explains the procedure for finding the point with almost locally maximal directional derivative. Finally, Section 4 proves any such point is a point of Fréchet differentiability.

Assume we have a closed set S and that we aim to prove S has the universal differentiability property. We describe the details of the construction of S below; at the moment we just say S is going to be defined using a Souslin-like operation on a family of closed "tubes", that is closed neighbourhoods of particular line segments. Consider an arbitrary Lipschitz function $f : Y \to \mathbb{R}$; we would like to show f is Fréchet differentiable at some point of S. The strategy is to, in some sense, almost locally maximise the directional derivative of f; this is done in Theorem 3.2, from within the constructed family of tubes. We then use the Differentiability Lemma 4.2, which gives a sufficient condition for the Lipschitz function to be Fréchet differentiable at a point where it has such an 'almost locally maximal' directional derivative.

In Section 4 we prove that if a Lipschitz function f has a directional derivative L at some point $y \in S$, and this derivative is almost locally maximal in the sense that for every ε, every directional derivative at *any* nearby point from S does not exceed $L + \varepsilon$, then the Lipschitz function is in fact Fréchet differentiable at the original point and the gradient is in the direction e of the almost locally maximal directional derivative. The word *any* in the latter sentence needs in fact to be replaced by a special condition (4.7); see Lemma 4.1 and Lemma 4.2. The proof is then based on the idea that, assuming non-Fréchet differentiability, we can find a wedge—that is a specially chosen union of two line segments—in which the growth of the function contradicts the mean value theorem and the local maximality assumption.

In Section 5 we show how to find such point with 'almost locally maximal' directional derivative. The idea behind the proof is to take a sequence of pairs (y_n, e_n) with the directional derivative $g'(y_n, e_n)$ being very close to the supremum over all directional derivatives $g'(z, u)$ with z close to y_{n-1} and (z, u) satisfying certain additional constraints—see Definition 5.2 and inequality (5.7)—and to argue that the sequence (y_n, e_n) converges to a point-direction pair (y, e) with the desired almost locally maximal directional derivative.

The optimisation method used in the present paper develops ideas from [12] and [6]. The new idea that we use in this paper is that instead of looking at points $y \in Y$, we define a bundle X over Y, where X is a complete topological space and $\pi : X \to Y$ is a continuous mapping, and locally maximise the directional derivative $f'(\pi x, e)$ over $x \in X$. This ensures that during the optimisation iterative procedure we are not thrown to the boundary of the set; if $\pi(X) \subseteq S$ then we are guaranteed that the point we obtain lies inside S.

Another key aspect of the proof of our result is the new set theoretic construction; see Theorem 3.3 and Section 6. First of all, we need to remark that the limit point to be obtained as a result of optimisation procedure must not be a porosity point of the set—see the next section for the definition and reasons. We achieve this by constructing a set in which, for every point x and every $\varepsilon > 0$, sufficiently small δ-neighbourhoods of x contain an $\varepsilon\delta$-dense set of line segments. The set is defined as an intersection of a countable collection $(J_k(\lambda))_{k\geqslant 1}$ of closed sets. Each $J_k(\lambda)$ is in its turn a countable union of "tubes", which are closed neighbourhoods of particular line segments. The construction of $J_k(\lambda)$ is inductive: around every tube in $J_l(\lambda)$ with $l < k$ we add a fine collection of tubes to $J_k(\lambda)$ and replace the original tube with a more narrow tube around the same line segment.

As we are aiming for a final set of Hausdorff dimension 1, we need to ensure the widths of the tubes in $J_k(\lambda)$ tend to 0 as $k \to \infty$. More precisely, we fix upfront a G_δ set O of Hausdorff dimension 1 containing a dense set of straight line segments, and a nested collection of open sets O_k with intersection O. By constructing $J_k(\lambda) \subseteq O_k$ we thereby ensure that $T_\lambda = \bigcap_{k\geqslant 1} J_k(\lambda)$ has Hausdorff dimension at most 1, as required. As $J_k(\lambda)$ are closed sets, so then is T_λ.

The parameter $\lambda \in (0, 1)$ is used to change the widths of all tubes involved in tube sets $J_k(\lambda)$ proportionally, multiplying by λ. We then establish that if $\lambda_1 < \lambda_2$ are fixed and we pick an arbitrary point $y \in T_{\lambda_1}$, then for each $\varepsilon > 0$ every sufficiently small δ-neighbourhood of y has an $\varepsilon\delta$-dense set of line segments that are fully inside T_{λ_2}. In order to achieve this we first find the level N after which, in the construction of tube sets $J_k(\lambda)$ we were choosing new tubes with density finer than ε multiplied by the width of the tube on the previous level. Choose δ to be smaller than the width of a tube on the level N and set $n \geqslant N$ to be the "critical" level on which the width of the tube containing point x multiplied by $\lambda_2 - \lambda_1$ for the first time becomes less than δ. Then the whole δ-neighbourhood of x is guaranteed to be inside the tube sets $J_m(\lambda_2)$, with $m \leqslant n - 1$. For $m \geqslant n + 1$ we find that the new tubes go εw_m-densely around x, where w_m is the width on the tube on level m. Since $w_m \leqslant \varepsilon w_{m-1} \leqslant \varepsilon\delta$ by construction, we find many tubes $\varepsilon\delta$-close to x on those subsequent levels. The problem that remains is that on the level n itself we might not find an appropriate tube at all! We overcome this obstacle by slightly modifying the definition of $J_k(\lambda)$ and taking it to be the union of tubes on a number of levels so that the "one level shift" does not take us outside the tube set $J_k(\lambda)$.

There is extra problem in the infinite dimensional case however. Given a tube T of width w in one of the tube sets J_k, in order to "kill" its porosity points and ensure sufficiently many line segments in the final set, we add tubes w/N_k-densely to J_{k+1}, where $N_k \to \infty$. The problem that immediately arises in the infinite dimensional case is that there is no "minimal" width among all tubes from J_k: since the Banach spaces we are working over are not locally compact, each collection of tubes J_k will have to be infinite, so that the infimum of the widths in J_k may be zero. Therefore we must add such approximating tubes only locally, in a small neighbourhood of each tube from J_k. This forces the length of tubes close to any fixed point $x \in T_\lambda$ to shrink rapidly, and therefore the point x will not have a "safe" neighbourhood $B_r(x)$ in which the set hits every ball $B_{c\|x-y\|}(y)$, i.e. we again get porosity at x.

To overcome this, a new approach is required; see Definition 6.4 and the proof of Lemma 6.6. In brief, on constructing level $k + 1$ approximation of tube T from J_k, we re-visit tubes constructed on each previous levels, $1 \leqslant l \leqslant k$, that form a sequence of ancestors of T. We approximate each tube thus re-visited to level $k + 1$ and include all new tubes in J_{k+1}—see Fig. 1. Approximations of lower level tubes to level $k + 1$ allows us to include longer tubes in J_{k+1}. This makes it possible to find, for each x and $\varepsilon > 0$, the critical value $\delta_1 > 0$ such that $\varepsilon\delta$-close to x there are line segments of length δ, for every $\delta \in (0, \delta_1)$. As explained in the beginning of this

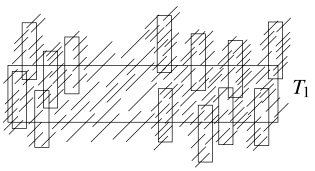

Fig. 1. We show here a horizontal tube T_1 of level 1, vertical tubes of level 2 that approximate points from T_1 and "diagonal" tubes of level 3 that approximate points from tubes of level 2 and points from T_1.

section, this property turns out to be sufficient for the set to have the universal differentiability property.

Theorem 3.3 is stated using more general terms than line segments and tubes; we prove the statement for a general class $(K_r)_{r\in R}$ of compact subsets of an arbitrary metric space (Y, d).

Finally, to get a totally disconnected universal differentiability set, we need to get rid of all these straight line segments that we have included in order to be able to prove the differentiability property inside the set T_λ defined above. For this, we intersect T_λ with a union of parallel hyperplanes obtained as a preimage of a totally disconnected subset of $\mathbb{R}$ under a continuous linear functional. To have this intersection totally disconnected it is enough to ensure that the containing G_δ set has this intersection totally disconnected. To show that the intersection of T_λ with the union of hyperplanes has the universal differentiability property we prove that for every Lipschitz function, its differentiability points inside T_λ form a very dense subset, and then choose the hyperplanes densely enough. See Theorems 3.1 and 3.10 for details.

2. Definitions and notations

In this paper we shall be working with real-valued functions defined on a real Banach space Y with separable dual. If a function $f : E \to \mathbb{R}$ is defined on a subset E of a Banach space Y we say f is locally Lipschitz on its domain E if for every $x \in E$ there exist $r > 0$ and $L \geqslant 0$ such that $|f(y') - f(y)| \leqslant L\|y - y'\|$ for all $y, y' \in E \cap B_r(x)$; the smallest such constant L is called the Lipschitz constant of f in $B_r(x)$ and is denoted $\mathrm{Lip}(f|_{B_r(x)})$. A function $f : Y \to \mathbb{R}$ is simply called Lipschitz if there is a common Lipschitz constant $L < \infty$ for which the Lipschitz condition is satisfied for any pair of points $y, y' \in Y$. The smallest such constant $L \geqslant 0$ is then called the Lipschitz constant of f and is denoted by $\mathrm{Lip}(f)$.

For any $f : Y \to \mathbb{R}$ and $y, e \in Y$, we define the directional derivative of f in the direction e as

$$f'(y, e) = \lim_{t\to 0} \frac{f(y + te) - f(y)}{t} \tag{2.1}$$

if the limit exists. If, for a fixed $y \in Y$, the formula (2.1) defines an element of Y^*, we say f is Gâteaux differentiable at y. Finally, if f is Gâteaux differentiable at y and the convergence in (2.1) is uniform for e in the unit sphere $S(Y)$ of Y, we say that f is Fréchet differentiable at y and call $f'(y)$ the Fréchet derivative of f, where $f'(y)e = f'(y, e)$ for all $e \in Y$.

The main focus of the present paper is on universal differentiability sets (UDS), those subsets of a Banach space Y that contain points of Fréchet differentiability of every Lipschitz function $f : Y \to \mathbb{R}$.

Recall a subset P of Y is called porous if there is a $c > 0$ such that for every $y \in P$ and every $r > 0$ there exist $\rho < r$ and $y' \in B_\rho(y)$ such that $B_{c\rho}(y') \cap P = \emptyset$. It is easy to see that any porous set is not a UDS since the distance function $f(x) = \inf_{y \in P} \|x - y\|$ is 1-Lipschitz and is not Fréchet differentiable at any point of P, provided P is porous; [15]. It turns out that the same is true for any σ-porous set P, that is any set that is a countable union of porous sets; see [3].

The existence, in Euclidean spaces, of a non-σ-porous set without porosity points and with a null closure was first shown in [14]; see also [16,17]. The set we are constructing will, in the finite dimensional case, be a compact null non-σ-porous. In fact, the construction implies that this null set has a universal differentiability subset without porosity points.

We shall be interested in the Hausdorff dimension of the universal differentiability sets we shall construct. Recall, for $s \geqslant 0$ and $A \subseteq Y$

$$\mathscr{H}^s(A) = \lim_{\delta \downarrow 0} \inf\left\{ \sum_i \operatorname{diam}(E_i)^s \text{ where } A \subseteq \bigcup_i E_i,\ \operatorname{diam}(E_i) \leqslant \delta \right\},$$

defines the s-dimensional Hausdorff measure of A, and

$$\dim_{\mathscr{H}}(A) = \inf\bigl\{s \geqslant 0 \text{ such that } \mathscr{H}^s(A) = 0\bigr\}$$

the Hausdorff dimension of A.

Let $(M, \|\cdot\|)$ be a normed space. We call the set $\mathscr{W}_M := M^3$ of triples from M the *wedge space* of M and we define a metric on $\mathscr{W}_M$ by

$$d\bigl(t', t\bigr) = \max_{1 \leqslant i \leqslant 3} \bigl\|t'_i - t_i\bigr\|,$$

where $t = (t_1, t_2, t_3)$ and $t' = (t'_1, t'_2, t'_3)$. Of course the distance d depends on the norm chosen on M.

Given $t \in \mathscr{W}_M$, we call the union of segments $W(t) = [t_1, t_2] \cup [t_2, t_3]$ a *wedge*. Note that triples (t_1, t_2, t_3) and (t_3, t_2, t_1) correspond to the same wedge for any $t_1, t_2, t_3 \in M$.

For $\alpha > 0$ and subsets $S_1, S_2 \subseteq M$ we say S_1 is an *α-wedge approximation* for S_2 in norm $\|\cdot\|$ if for any $t \in \mathscr{W}_M$ with $W(t) \subseteq S_2$, there exists $t' \in \mathscr{W}_M$ with $W(t') \subseteq S_1$ and $d(t', t) \leqslant \alpha$. When it is clear which norm on M is considered we shall just say that S_1 is an *α-wedge approximation* for S_2.

We shall also consider a more general construction when the collection of wedges is replaced by a general family of compact subsets of M, which may now be considered a general metric space. We shall at times make use of the Hausdorff distance between two such compact sets:

$$\mathscr{H}(K_1, K_2) = \inf\bigl\{r > 0\colon\ K_1 \subseteq \overline{B}_r(K_2) \text{ and } K_2 \subseteq \overline{B}_r(K_1)\bigr\}.$$

Here we use $\overline{B}_r(A)$ to denote the closed r-neighbourhood of $A \subseteq M$; we shall also use $B_r(A)$ to denote an open r-neighbourhood of $A \subseteq M$.

As a simple observation we note that if M is a normed space and $t, t' \in \mathscr{W}_M$ then we have

$$\mathscr{H}\bigl(W\bigl(t'\bigr), W(t)\bigr) \leqslant d\bigl(t', t\bigr).$$

In order to construct a UDS we first define a G_δ set O containing a dense set of arbitrarily small wedges and then define a subset S of O as described in Section 1. For an arbitrary Lipschitz function we then apply our optimisation method to S; see Section 5. We remark that any G_δ set is a complete topological space; this lets us conclude that the differentiability point, which we find as a limit point of the iterative construction, belongs to the set S.

As we have already mentioned in Section 1, any UDS has Hausdorff dimension at least 1. We prove this result in the next lemma.

Lemma 2.1. *Let Y be a non-zero Banach space and $S \subseteq Y$ a universal differentiability set. Then the Hausdorff dimension of S is at least* 1.

Proof. Assume $\dim_{\mathscr{H}}(S) < 1$. Fix any nonzero $P \in Y^*$ and $e \in Y$ with $P(e) = 1$. The Hausdorff dimension of $P(S)$ is strictly less than 1, and therefore $P(S)$ has Lebesgue measure 0. Let $g : \mathbb{R} \to \mathbb{R}$ be a Lipschitz function that is not differentiable at any $y \in P(S)$. Then $f := g \circ P : Y \to \mathbb{R}$ is a Lipschitz function that is not differentiable at any $x \in S$, as the directional derivative $f'(y, e)$ does not exist for $y \in S$. $\square$

3. Main results

We begin this section with the statement of a criterion for universal differentiability.

Theorem 3.1. *Let (M, d) be a non-empty complete metric space, $(Y, \|\cdot\|)$ be a Banach space with separable dual and $\pi : M \to Y$ a continuous mapping.*

Suppose that for every $\eta > 0$ and $x \in M$ and every open neighbourhood $N(x)$ of x in M there exists $\delta_0 = \delta_0(x, N(x), \eta) > 0$ such that, for any $\delta \in (0, \delta_0)$ the set $\pi(N(x))$ is a $\delta\eta$-wedge approximation for $B_\delta(\pi(x))$.

Then $\pi(M)$ is a universal differentiability set and, moreover, for every Lipschitz function $g : Y \to \mathbb{R}$ the set $D_g = \{y \in \pi(M)\colon\ g \text{ is Fréchet differentiable at } y\}$ is dense in $\pi(M)$. Furthermore, if $y \in \pi(M)$, $r > 0$ and $P : Y \to \mathbb{R}$ is a non-zero continuous linear map then there exists a finite open interval $I = I_g(y)$ with $Py \in I$ and

$$\mu\big(I \setminus P\big(D_g \cap B_r(y)\big)\big) = 0,$$

where μ denotes the Lebesgue measure.

To prove Theorem 3.1, we need to find points of Fréchet differentiability in $\pi(M)$ for every Lipschitz function defined on Y. To accomplish this, we first apply the next theorem, Theorem 3.2, to obtain a point with almost locally maximal directional derivative, and then use Differentiability Lemma 4.2 to show that the function is in fact Fréchet differentiable at this point.

Theorem 3.2. *Let (M, d) be a non-empty complete metric space, $(Y, \|\cdot\|)$ a Banach space, $\pi : M \to Y$ a continuous map and $\Theta : (0, \infty) \to (0, \infty)$ a real-valued function with $\Theta(t) \to 0$ as $t \to 0^+$. Assume $g : Y \to \mathbb{R}$ is a Lipschitz function and*

$$(x_0, e_0) \in D = \big\{(x, e) \in M \times \big(Y \setminus \{0\}\big) \text{ such that } g'(\pi x, e) \text{ exists}\big\}$$

is such that $\|e_0\| = 1$ and $g'(\pi x_0, e_0) \geqslant 0$.

Then one can define

(1) *a Lipschitz function* $f : Y \to \mathbb{R}$ *by*

$$f = g + 2\,\mathrm{Lip}(g)e_0^*, \tag{3.1}$$

where $e_0^* \in Y^*$ *is a linear functional such that* $\|e_0^*\|_{(Y,\|\cdot\|)^*} = e_0^*(e_0) = 1$,

(2) *a norm* $\|\cdot\|'$ *on* Y, *with* $\|y\| \leqslant \|y\|' \leqslant 2\|y\|$ *for all* $y \in Y$, *and*

(3) *a pair* $(\tilde{x}, \tilde{e}) \in D$ *with* $\|\tilde{e}\|' = 1$

such that $f'(\pi\tilde{x}, \tilde{e}) \geqslant f'(\pi x_0, e_0)$ *and the directional derivative* $f'(\pi\tilde{x}, \tilde{e})$ *is almost locally maximal in the following sense. For any* $\varepsilon > 0$ *there exists an open neighbourhood* N_ε *of* $\tilde{x}$ *in* M *such that whenever* $(x', e') \in D$ *with*

(i) $x' \in N_\varepsilon$, $\|e'\|' = 1$ *and*

(ii) *for any* $t \in \mathbb{R}$

$$\begin{aligned} &\bigl|\bigl(f\bigl(\pi x' + t\tilde{e}\bigr) - f\bigl(\pi x'\bigr)\bigr) - \bigl(f(\pi\tilde{x} + t\tilde{e}) - f(\pi\tilde{x})\bigr)\bigr| \\ &\quad\leqslant \Theta\bigl(f'\bigl(\pi x', e'\bigr) - f'(\pi\tilde{x}, \tilde{e})\bigr)|t|, \end{aligned} \tag{3.2}$$

then we have $f'(\pi x', e') < f'(\pi\tilde{x}, \tilde{e}) + \varepsilon$.

Moreover, if the original norm $\|\cdot\|$ *is Fréchet differentiable on* $Y \setminus \{0\}$ *then the norm* $\|\cdot\|'$ *can be chosen with this property too.*

We prove Theorem 3.2 at the end of Section 5. We will now use its conclusion to prove Theorem 3.1.

Proof of Theorem 3.1. Without loss of generality we may assume that the norm $\|\cdot\|$ is Fréchet differentiable on $Y \setminus \{0\}$, by [2,3], since passing to an equivalent norm keeps the $\delta\eta$-wedge approximation condition and does not change the differentiability property.

Taking arbitrary $x \in M$ and $N_0(x) = M$ we get that the wedge approximation property of $\pi(N_0(x))$ implies that $\pi(M)$ contains a non-degenerate straight line segment $L \subseteq Y$. As any Lipschitz function $g : Y \to \mathbb{R}$ is differentiable at some point $p \in L$ in the direction of L, the set

$$D := \bigl\{(x, e) \in M \times \bigl(Y \setminus \{0\}\bigr) \text{ such that } g'(\pi x, e) \text{ exists}\bigr\}$$

is non-empty.

Without loss of generality we may assume that the Lipschitz constant of g is equal to 1. Picking an arbitrary $(x_0, e_0) \in D$ and $\Theta(s) = 25\sqrt{3s}$, we see that all the conditions of Theorem 3.2 are satisfied if we rescale e_0 in order to have $\|e_0\| = 1$ and replace e_0 with $-e_0$ if necessary so as to have $g'(\pi x_0, e_0) \geqslant 0$. Let the Lipschitz function $f : Y \to \mathbb{R}$, the norm $\|\cdot\|'$ on Y, the pair $(\tilde{x}, \tilde{e}) \in D$ and, for each $\varepsilon > 0$, the open neighbourhood N_ε of $\tilde{x} \in M$ be given by the conclusion of Theorem 3.2. Note that $f'(\pi\tilde{x}, \tilde{e}) \geqslant f'(\pi x_0, e_0)$, $\mathrm{Lip}(f) \leqslant 3$, we may take $\|\cdot\|'$ to be Fréchet differentiable on $Y \setminus \{0\}$ and that

$$\|z\| \leqslant \|z\|' \leqslant 2\|z\| \tag{3.3}$$

for all $z \in Y$, so that $\|\tilde{e}\| \leqslant \|\tilde{e}\|' = 1$.

We claim that $\tilde{y} = \pi\tilde{x}$ is a point of Fréchet differentiability of f.

Since the two norms $\|\cdot\|$, $\|\cdot\|'$ are equivalent, it suffices to verify the conditions of Lemma 4.2 for $(Y, \|\cdot\|')$, applied to the Lipschitz function f, $L = 3$ and the pair $(\tilde{y}, \tilde{e}) = (\pi\tilde{x}, \tilde{e})$. To accomplish this, we let $\varepsilon, \theta > 0$ and show that $F_\varepsilon = \pi(N_\varepsilon)$ and $\delta_* = \delta_0(\tilde{x}, N_\varepsilon, \theta/2)$ are such that (1) and (2) of Lemma 4.2 hold, with the norm $\|\cdot\|$ replaced by $\|\cdot\|'$.

Suppose $\delta \in (0, \delta_*)$, $\|y_i - \tilde{y}\|' < \delta$ for $i = 1, 2, 3$. Then from (3.3) we have $\|y_i - \tilde{y}\| < \delta$ for $i = 1, 2, 3$ as well. Now using that $F_\varepsilon = \pi(N_\varepsilon)$ is a $\delta\theta/2$-wedge approximation for $B_\delta(\tilde{y})$ in $\|\cdot\|$ and the inequality (3.3), we get that F_ε is a $\delta\theta$-wedge approximation for $B_\delta(\tilde{y})$ in norm $\|\cdot\|'$. This verifies condition (1) of Lemma 4.2.

For condition (2) we note that if $y' \in F_\varepsilon$, $\|e'\|' = 1$ and

$$\begin{aligned}&\bigl|\bigl(f\bigl(y' + t\tilde{e}\bigr) - f\bigl(y'\bigr)\bigr) - \bigl(f(\tilde{y} + t\tilde{e}) - f(\tilde{y})\bigr)\bigr| \\ &\quad\leqslant 25\sqrt{\bigl(f'\bigl(y', e'\bigr) - f'(\tilde{y}, \tilde{e})\bigr)L} \cdot |t|\end{aligned}$$

for all $t \in \mathbb{R}$, then as $F_\varepsilon = \pi(N_\varepsilon)$ we may write $y' = \pi x'$ where $x' \in N_\varepsilon$. As $L = 3$ and $\Omega(s) = 25\sqrt{3s}$, the conditions (i) and (ii) of Theorem 3.2 are satisfied, so we deduce that $f'(\pi x', e') < f'(\pi\tilde{x}, \tilde{e}) + \varepsilon$.

As all the conditions of Lemma 4.2 are satisfied we deduce that f is Fréchet differentiable at $\tilde{y} = \pi\tilde{x} \in \pi(M)$. As $f - g$ is linear, we conclude g is also Fréchet differentiable at $\tilde{y} \in \pi(M)$. Hence $\pi(M)$ is indeed a universal differentiability set in Y.

Note moreover we have proved slightly more: namely, if M is any non-empty complete metric space satisfying the wedge approximation property as in the conditions of present theorem, then for any Lipschitz $g : Y \to \mathbb{R}$ and an arbitrary pair $(x_0, e_0) \in M \times (Y \setminus \{0\})$ such that $\|e_0\| = 1$ and $g'(\pi x_0, e_0) \geqslant 0$, there is a Lipschitz function $f : Y \to \mathbb{R}$ defined according to (3.1) and a pair $(\tilde{x}, \tilde{e}) \in M \times (Y \setminus \{0\})$ such that $\|\tilde{e}\| \leqslant 1$, g is Fréchet differentiable at $\pi\tilde{x}$ and $f'(\pi\tilde{x}, \tilde{e}) \geqslant f'(\pi x_0, e_0)$.

To verify the density of the set

$$D_g = \bigl\{y \in \pi(M)\colon\ g \text{ is Fréchet differentiable at } y\bigr\}$$

in $\pi(M)$, it suffices to note that if $y = \pi x \in \pi(M)$ and $\varepsilon > 0$, we may pick a non-empty open set $N \subseteq M$ such that $\pi(N) \subseteq B_\varepsilon(y)$. Then as the restriction bundle $\pi|_N : N \to Y$ satisfies the conditions of the present theorem, any Lipschitz $g : Y \to \mathbb{R}$ contains a point of Fréchet differentiability in $\pi(N) \subseteq \pi(M) \cap B_\varepsilon(y)$.

We now check the last observation of the theorem. We may assume $\|P\| = 1$. Let $y = \pi(x) \in \pi(M)$ and $\delta_0 = \delta_0(x, M, \eta)$, where $\eta \in (0, 1/12)$. Choose also a vector $e_1 \in Y$ such that $Pe_1 = 1$. Fix any $\delta \in (0, \min\{r/2, \delta_0\})$ and find a line segment $L_0 \subseteq \pi(M)$ that is an $\eta\delta$-wedge approximation for $L = [y - \delta e_1, y + \delta e_1]$. It is easy to see that $L_0 \subseteq B_r(y)$ and

$$P(L_0) \supseteq I = \bigl(Py - (1 - \eta)\delta, Py + (1 - \eta)\delta\bigr).$$

Let $L_0 = [z_0, z_0 + l_0 e_0]$ with $\|e_0\| = 1$. As g is Lipschitz, the directional derivative $g'(z, e_0)$ exists for almost all points $z \in L_0$. We note that the set D_g is a $F_{\sigma\delta}$-set:

$$D_g = \bigcap_{n \geqslant 1} \bigcup_{\substack{y^* \in A \\ \delta \in \mathbb{Q}}} \bigcap_{\substack{\|z\| \leqslant 1 \\ |t| < \delta}} \bigl\{y \in Y\colon\ \bigl|g(y + tz) - g(y) - ty^*(z)\bigr| \leqslant |t|/n\bigr\},$$

where A is a countable dense subset of the unit ball of Y^*. Therefore the image $P(D_g \cap B_r(y))$, being a projection of a Borel subset of a Polish space, is an analytic subset of $\mathbb{R}$ and therefore Lebesgue measurable.

Suppose then that the Lebesgue measure $\mu(I \setminus P(D_g \cap B_r(y)))$ is strictly positive. There exists a non-constant everywhere differentiable Lipschitz function $h : I \to \mathbb{R}$ such that $h' = 0$ on $P(D_g \cap B_r(y)) \cap I$. This implies there exists $y_0 \in L_0$ such that $s = Py_0 \in I$, the directional derivative $g'(y_0, e_0)$ exists and $h'(s) \neq 0$. By scaling h if necessary we may assume $h'(s) = 1$. Let $G = g + 3h \circ P$. This is a Lipschitz function defined on Y, and such that the directional derivative $G'(y_0, e_0)$ exists; moreover, $G'(y_0, e_0) = g'(y_0, e_0) + 3P(e_0)$. Note that the vectors $l_0 e_0$ and $2\delta e_1$ which define the line segments L_0 and L_1 have their start and end points $\eta\delta$-close to each other. Therefore

$$\|e_0 - e_1\| = \frac{1}{2\delta}\|2\delta e_0 - 2\delta e_1\| \leqslant \frac{1}{2\delta}\big(|2\delta - l_0| \cdot \|e_0\| + \|l_0 e_0 - 2\delta e_1\|\big) \leqslant 2\eta < \frac{1}{6}.$$

Using $\operatorname{Lip}(g) = 1$, $Pe_1 = 1$ and $\|P\| = 1$ we conclude $G'(y_0, e_0) \geqslant 3/2$.

Let $\widetilde{M} = \pi^{-1}(L_0 \cap P^{-1}(I))$. Note that $\widetilde{M}$ is a G_δ-set and $\widetilde{M} \subseteq M$. As $y_0 \in \pi(\widetilde{M})$ we conclude there is an $x_0 \in \widetilde{M}$ such that $y_0 = \pi x_0$.

Then, using the more general statement we have proved for the first part of the theorem for $\widetilde{M}$ instead of M and G instead of g, we conclude that there is a Lipschitz function $F : Y \to \mathbb{R}$ defined according to (3.1), $F = G + 2\operatorname{Lip}(G)e_0^*$, where $\|e_0^*\| = e_0^*(e_0) = 1$, and a pair $(\tilde{x}, \tilde{e}) \in \widetilde{M} \times (Y \setminus \{0\})$ such that $\|\tilde{e}\| \leqslant 1$, G is Fréchet differentiable at $\pi\tilde{x}$ and $F'(\pi\tilde{x}, \tilde{e}) \geqslant F'(\pi x_0, e_0)$. Then $\tilde{y} = \pi\tilde{x} \in L_0 \subseteq B_r(y)$ is a point of Fréchet differentiability of G and

$$\begin{aligned} G'(\tilde{y}, \tilde{e}) - G'(y_0, e_0) &= F'(\tilde{y}, \tilde{e}) - F'(y_0, e_0) + 2\operatorname{Lip}(G)e_0^*(e_0 - \tilde{e}) \\ &\geqslant F'(\tilde{y}, \tilde{e}) - F'(y_0, e_0) \geqslant 0 \end{aligned}$$

as $e_0^*(e_0) = 1$ and $e_0^*(\tilde{e}) \leqslant \|\tilde{e}\| \leqslant 1$. Together with $G'(y_0, e_0) \geqslant 3/2$ we conclude $G'(\tilde{y}, \tilde{e}) \geqslant 3/2$.

However, since G is Fréchet differentiable at $\tilde{y}$, so is g, and therefore $\tilde{y} \in D_g \cap B_r(y)$. As we also have $\tilde{y} = \pi\tilde{x} \in \pi(\widetilde{M})$, we conclude $P\tilde{y} \in P(D_g \cap B_r(y)) \cap I$; hence $G'(\tilde{y}, \tilde{e}) = g'(\tilde{y}, \tilde{e})$, a contradiction to $G'(\tilde{y}, \tilde{e}) \geqslant 3/2$ as a directional derivative of a 1-Lipschitz function g cannot exceed 1. □

Together with the following statement, Theorem 3.1 implies the existence of a closed universal differentiability set; see Lemma 3.5.

Theorem 3.3. *Let (Y, d) be a metric space and let $(K_r)_{r \in R}$ be a collection of non-empty compact subsets of Y indexed by a non-empty metric space (R, γ) such that the Hausdorff distance $\mathscr{H}(K_r, K_s)$ is bounded from above by $\gamma(r, s)$ for every $r, s \in R$. Assume O is a G_δ subset of Y such that O contains a γ-dense subset of the family $(K_r)_{r \in R}$ and $K_{r_0} \subseteq O$ is one of these compacts. Assume further that there exist $\rho > 0$ and $\varepsilon_0 > 0$ such that for every $\varepsilon \in (0, \varepsilon_0)$ we can find a set of indices $R(\varepsilon) \subseteq R$ such that*

- *for every $s \in R$ there exists $t \in R(\varepsilon)$ with $\gamma(t, s) < \varepsilon$,*
- *for every subset S of Y of diameter at most $\rho\varepsilon$ the set $\{r \in R(\varepsilon)\colon\ S \cap K_r \neq \emptyset\}$ is finite.* (3.4)

Then there exists a nested collection of closed non-empty subsets $(T_\lambda)_{0\leqslant\lambda\leqslant 1}$ *of* O—$T_{\lambda'} \subseteq T_\lambda$ *whenever* $0 \leqslant \lambda' \leqslant \lambda \leqslant 1$—*each containing* K_{r_0} *that satisfies the following. For each* $\eta > 0$, $\lambda \in (0, 1]$ *and* $y \in \bigcup_{0\leqslant\lambda'<\lambda} T_{\lambda'}$ *there exists* $\delta_1 = \delta_1(\eta, \lambda, y) > 0$ *such that if* $\delta \in (0, \delta_1)$ *and* $s \in R$ *with* $K_s \subseteq \overline{B}_\delta(y)$ *there exists* $t \in R$ *such that* $K_t \subseteq T_\lambda$ *and* $\gamma(t, s) < \eta\delta$.

We prove Theorem 3.3 in Section 6.

Remark 3.4. Let now $R = Y^3$ be the wedge space on Y, and for each triple $r = (y_1, y_2, y_3) \in R$ define $K_r = W(r) = [y_1, y_2] \cup [y_2, y_3]$ to be the corresponding wedge. If we further let $\gamma(K_r, K_s)$ be equal to the standard wedge distance, $\gamma(K_r, K_s) = d(W(r), W(s))$, the conclusion of Theorem 3.3 is: there exists $\delta_1 > 0$ such that if $\delta \in (0, \delta_1)$ then T_λ is a $\eta\delta$-wedge approximation for $\overline{B}_\delta(y)$. In Lemma 3.5 we show that this property implies that T_λ are universal differentiability sets. We will later easily get that T_λ has Hausdorff dimension 1 by taking the containing G_δ-set O of Hausdorff dimension 1. See Lemma 3.9 for the list of properties that we require O to satisfy for this.

However, in order to get the conclusion of Theorem 3.3, one needs to verify condition (3.4). In the case in which Y is a finite dimensional space, it is easy to see that since balls in Y are totally bounded sets, $R = Y^3$ satisfies the required condition. In case Y is an infinite dimensional space, we prove this property in Lemma 3.6.

Lemma 3.5. *Let* Y *be a Banach space with separable dual and* $(\mathscr{W}, d) = (\mathscr{W}_Y, d)$ *be the wedge space equipped with the standard wedge distance. Suppose* O *is a* G_δ *subset of* Y *containing* $W(t)$ *for* t *belonging to a* d*-dense subset of* $\mathscr{W}$, *and the nested collection* $(T_\lambda)_{0\leqslant\lambda\leqslant 1}$ *of non-empty closed subsets of* Y, $T_{\lambda'} \subseteq T_\lambda$ *for* $0 \leqslant \lambda' \leqslant \lambda \leqslant 1$, *satisfies the condition that for each* $\eta > 0$, $\lambda \in (0, 1]$ *and* $y \in \bigcup_{0\leqslant\lambda'<\lambda} T_{\lambda'}$ *there is a* $\delta_1 = \delta_1(\eta, \lambda, y) > 0$ *such that for all* $\delta \in (0, \delta_1)$ *the set* T_λ *is a* $\eta\delta$*-wedge approximation for* $B_\delta(y)$.

Then for each $\lambda \in (0, 1]$ *the set* T_λ *is a closed universal differentiability set. Furthermore, for any Lipschitz function* $g : Y \to \mathbb{R}$, *any* $x \in T_{\lambda'}$, $0 \leqslant \lambda' < \lambda \leqslant 1$, $r > 0$ *and any non-zero continuous linear map* $P : Y \to \mathbb{R}$ *there exists a finite open interval* $I = I_g(x)$ *with* $Px \in I$ *and*

$$\mu\big(I \setminus P\big(T_\lambda \cap D_{g,r}(x)\big)\big) = 0,$$

where $D_{g,r}(x)$ *is the set of points of Fréchet differentiability of* g *in the* r*-neighbourhood of* x *and* μ *denotes the Lebesgue measure.*

Proof. For every $\lambda \in (0, 1]$, define a subset of $(0, \lambda) \times Y$

$$X_\lambda = \big\{(\tau, y)\colon\ 0 < \tau < \lambda \text{ and } y \in T_{\tau'} \text{ for every } \tau' \in (\tau, 1)\big\}. \tag{3.5}$$

Note that if $\tau \in (0, \lambda)$ we have $X_\lambda \supseteq \{\tau\} \times T_\tau$, so $X_\lambda \neq \emptyset$; and for every $(\tau, y) \in X_\lambda$ we necessarily have $y \in T_\lambda$. Moreover, if we let Δ denote a complete metric on $(0, \lambda)$, then

$$d\big((\tau', y'), (\tau, y)\big) = \Delta\big(\tau', \tau\big) + \|y' - y\|$$

makes X_λ a complete metric space, since T_λ is closed.

We now check that the conditions of Theorem 3.1 are satisfied for $M = X_\lambda$ and $\pi(\tau, y) = y$. Assume we are given $\eta \in (0, 1)$, a point $x = (\tau, y) \in X_\lambda$ and its open neighbourhood $N(x)$. Without loss of generality we may assume there is $\psi > 0$ such that

$$N(x) = \{(\tau', y') \in X_\lambda\colon \Delta(\tau', \tau) < \psi \text{ and } \|y' - y\| < \psi\}.$$

Then fixing $\tau' \in (\tau, \lambda)$ such that $\Delta(\tau', \tau) < \psi$ we get $\pi(N(x)) \supseteq B_\psi(y) \cap T_{\tau'}$. Define now $\delta_0(x, N(x), \eta) = \min\{\delta_1(\eta, \tau', y), \psi/2\}$ and assume $\delta \in (0, \delta_0)$. Since $T_{\tau'}$ is a $\delta\eta$-wedge approximation for $B_\delta(y)$ and $\delta + \delta\eta < 2\delta < \psi$, we conclude that $T_{\tau'}$, and therefore $\pi(N(x))$ as well is a $\delta\eta$-wedge approximation for $\overline{B}_\delta(x)$.

The conclusion of Theorem 3.1 says that $\pi(X_\lambda)$ is a universal differentiability set. Since $\pi(X_\lambda) \subseteq T_\lambda$ we conclude T_λ is a universal differentiability set, for every $\lambda \in (0, 1]$.

Moreover, if $x \in T_{\lambda'}$ and $0 \leqslant \lambda' < \lambda \leqslant 1$ we conclude $(\lambda', x) \in X_\lambda$ (if $\lambda' = 0$ then find $\lambda'' \in (0, \lambda)$ and get $x \in T_{\lambda''}$ so $(\lambda'', x) \in X_\lambda$). Then the final part of the lemma follows from the conclusion of Theorem 3.1. □

Lemma 3.6 shows that most natural choices of (R, γ) in Y, an infinite dimensional separable Banach space, satisfy the conditions of Theorem 3.3 with $\rho = 1/4$; in particular the conditions are satisfied whenever the collection $(K_r)_{r\in R}$ of compacts is translation invariant, with $\gamma(K_r, x + K_r) \leqslant \|x\|$.

Lemma 3.6. *Suppose $(Y, \|\cdot\|)$ is an infinite dimensional Banach space, (R, γ) is separable and has the property that whenever $r \in R$ and $x \in Y$ then $K_s = x + K_r$ for some $s \in R$ with*

$$\gamma(s, r) \leqslant \frac{1}{4\rho}\|x\|. \tag{3.6}$$

Then for every $\varepsilon > 0$ there exists a set $R(\varepsilon) \subseteq R$ such that

(1) *for all $r \in R$ there exists $s \in R(\varepsilon)$ with $\gamma(s, r) < \varepsilon$,*
(2) *if r, s are distinct elements of $R(\varepsilon)$ then $\operatorname{dist}(K_r, K_s) > \rho\varepsilon$,*

where for compact $K, K' \subseteq Y$, we define

$$\operatorname{dist}(K, K') = \inf\{\|k' - k\| \text{ where } k \in K,\ k' \in K'\}.$$

We establish the lemma in a few short steps.

Lemma 3.7. *If Y is an infinite dimensional Banach space and $K \subseteq Y$ is compact then for every $\varepsilon > 0$ there exists $y \in Y$ with $\|y\| = \varepsilon$ and $\operatorname{dist}(y, K) > \varepsilon/3$.*

Proof. It is well known that one may find an infinite collection $(e_n)_{n\in\mathbb{N}}$ in Y with $\|e_n\| = 1$ and $\|e_n - e_m\| \geqslant 1$ for $m \neq n$. Assuming, for a contradiction, that we cannot find n with $\operatorname{dist}(\varepsilon e_n, K) > \varepsilon/3$ then we can pick $k_n \in K_n$ for each n with $\|k_n - \varepsilon e_n\| \leqslant \varepsilon/3$. It then follows that $\|k_n - k_m\| \geqslant \varepsilon/3$ for all $m \neq n$, contradicting the compactness of K. □

Lemma 3.8. *If Y is an infinite dimensional Banach space and $(K_n)_{n\geqslant 1}$ are compact subsets of Y then for any $\varepsilon > 0$ we can find $y_n \in Y$ with $\|y_n\| = \varepsilon$ for each $n \geqslant 1$ such that $K'_n := y_n + K_n$ satisfy* $\operatorname{dist}(K'_n, K'_m) > \varepsilon/3$ *for $n \neq m$.*

Proof. Suppose $n \geqslant 1$ and we have chosen $(y_m)_{1\leqslant m<n}$ such that $\operatorname{dist}(K'_m, K'_l) > \varepsilon/3$ for $1 \leqslant l < m < n$. It suffices to pick y_n such that $\operatorname{dist}(K'_n, K'_m) > \varepsilon/3$ for $1 \leqslant m < n$.

The difference set

$$K := K_n - \bigcup_{1\leqslant m<n} K'_m = \{k - k' \text{ where } k \in K_n,\ k' \in K_m \text{ for some } m < n\}$$

is compact so that we may find $y \in Y$ with $\|y\| = \varepsilon$ and $\operatorname{dist}(y, K) > \varepsilon/3$, using Lemma 3.7. Then $\operatorname{dist}(0, -y + K) > \varepsilon/3$ so that, choosing $y_n = -y$,

$$\operatorname{dist}\Big(0, K'_n - \bigcup_{1\leqslant m<n} K'_m\Big) > \varepsilon/3. \qquad \square$$

Proof of Lemma 3.6. We may assume $R \neq \emptyset$. Let $(r_n)_{n\geqslant 1}$ be a dense sequence in R. By Lemma 3.8 we can find $y_n \in Y$ with $\|y_n\| = 3\rho\varepsilon$ such that $K'_n := y_n + K_{r_n}$ satisfy $\operatorname{dist}(K'_n, K'_m) > \rho\varepsilon$ for $n \neq m$. Now we may pick r'_n with $K_{r'_n} = y_n + K_{r_n} = K'_n$ and

$$\gamma(r'_n, r_n) \leqslant \frac{1}{4\rho}\|y_n\| = \frac{3}{4}\varepsilon$$

using (3.6). Setting $R(\varepsilon) = \{r'_n \text{ where } n \in \mathbb{N}\}$ we are done. $\square$

Conclusion. We summarise what we have shown and add some further observations. First note, Lemma 3.7 implies that any compact set in an infinite dimensional space is porous. Now, as any porous set is not a UDS, it follows that a UDS cannot be compact in infinite dimensional spaces.

On the other hand, we now show that inside any non-empty open set in Y we can find a closed universal differentiability set of Hausdorff dimension 1 which does not contain any continuous curves: this set can be chosen to be totally disconnected.

Lemma 3.9. *Let Y be a non-zero separable Banach space and $(\mathscr{W}, d) = (\mathscr{W}_Y, d)$ be the wedge space on Y equipped with the standard wedge distance. Then given any $\varphi \in Y^* \setminus \{0\}$ there exists a G_δ subset O of Y of Hausdorff dimension 1 such that O contains the wedges $W(t)$ for t belonging to a d-dense subset of $\mathscr{W}$ and the intersection $O \cap (y + \ker\varphi)$ is totally disconnected for any $y \in Y$.*

Proof. Let $\mathscr{W}_0 \subseteq \mathscr{W}$ be a d-dense countable subset. Note that

$$\mathscr{W}_1 = \{(y_1, y_2, y_3) \in \mathscr{W}_0\colon\ \varphi(y_1) \neq \varphi(y_2) \text{ and } \varphi(y_2) \neq \varphi(y_3)\}$$

is then also d-dense in $\mathscr{W}$.

As $\bigcup_{t\in\mathscr{W}_1} W(t)$ is a countable union of line segments we can cover it, for each n, by $O'_n = \bigcup_{m\geqslant 1} B_{2^{-(m+n)}}(L_m)$, where $\bigcup_{m\geqslant 1} L_m = \bigcup_{t\in\mathscr{W}_1} W(t)$ and each L_m has length less than or equal to 1. Note that this implies that the Hausdorff dimension of the G_δ-set $O' = \bigcap_{n\geqslant 1} O'_n$ is less than or equal to 1. Indeed, for any $s > 1$ we have

$$\mathscr{H}^s(O'_n) \leqslant \sum_{m\geqslant 1} 2^{m+n}\bigl(2\cdot 2^{-(m+n)}\bigr)^s = \frac{2^s}{(2^{s-1})^{n+1}}\frac{1}{1-2^{-(s-1)}},$$

and so $\mathscr{H}^s(O') = 0$. On the other hand, the Hausdorff dimension of O' cannot be less than 1 as it contains non-trivial line segments.

We thus found a G_δ set O' of Hausdorff dimension 1 such that $W(t) \subseteq O'$ for all $t \in \mathscr{W}_1$. Let further $\{t^{(1)}, t^{(2)}, \ldots\}$ be an enumeration of $\mathscr{W}_1$.

Let $L = y_0 + \mathbb{R}e$ be a line through one of the sides of a wedge $W(t)$ where $t \in \mathscr{W}_1$ and $\|e\| = 1$. Let $a > 0$. Then, for any $y \in Y$, the diameter of any connected component of the intersection of $B_a(L)$ with hyperplane $y + \ker\varphi$ does not exceed $2a(1 + \|\varphi\|/|\varphi(e)|)$. Therefore if we let the countable set $(e_{1,i}, e_{2,i})_{i\geqslant 1}$ be the pairs of unit directions of sides of all wedges $W(t^{(i)})$, $i \geqslant 1$, then for any $y \in Y$ each connected component of the intersection of the open set $O_n = \bigcup_{i\geqslant 1} B_{\varepsilon_i}(W(t^{(i)}))$ with $y + \ker\varphi$ has diameter less than $1/n$, whenever

$$0 < \varepsilon_i < 1\bigg/\left(n2^{i+2}\left(1 + \frac{\|\varphi\|}{\min\{|\varphi(e_{1,i})|, |\varphi(e_{2,i})|\}}\right)\right).$$

Thus the conclusion of the lemma is satisfied for $O = O' \cap \bigcap_{n\geqslant 1} O_n$. □

Theorem 3.10. *Let Y be a Banach space with separable dual. Then for every open set $U \subseteq Y$ there is a closed set $S \subseteq U$ of Hausdorff dimension 1 such that every locally Lipschitz function f defined on a domain containing U has a point of Fréchet differentiability inside S. Moreover, the set S may be chosen to be in addition totally disconnected so that it contains no non-constant continuous curves.*

Proof. Fix any non-zero continuous linear map $P : Y \to \mathbb{R}$. Let O be a G_δ subset of Y satisfying Lemma 3.9 for the wedge space $(\mathscr{W}, d) = (\mathscr{W}_Y, d)$ equipped with the standard wedge distance, and $\varphi = P$. By Remark 3.4 and Lemma 3.6 we can apply Theorem 3.3 in order to get a nested sequence of closed sets $T_\lambda \subseteq O$ satisfying the hypothesis of Lemma 3.5.

Fix $y_0 \in T_\lambda$ for some $\lambda \in [0, 1)$. Let $\lambda_0 \in (\lambda, 1]$ and $r_0 > 0$ be such that $B_{r_0}(y_0) \subseteq U$.

Let $C \subseteq [0, 1]$ be a closed totally disconnected set of positive measure, such that every neighbourhood of any of its points intersects C by a set of positive measure. An example of such set could be a Cantor set of positive measure.

Let C_0 be a shift of C such that $Py_0 \in C_0$. Consider a set

$$S = P^{-1}(C_0) \cap T_{\lambda_0} \cap \overline{B}_{r_0/2}(y_0).$$

We clearly have $y_0 \in S \subseteq \overline{B}_{r_0/2}(y_0) \subseteq U$. Note further that as $P^{-1}(C_0) \cap O$ is totally disconnected for every $c \in C_0$, and C_0 is totally disconnected by itself, the set S set is totally disconnected. It is also clear S is closed and $\dim_{\mathscr{H}}(S) \leqslant 1$ as $\dim_{\mathscr{H}}(O) = 1$ and $S \subseteq O$. It remains to verify that every locally Lipschitz function defined on a domain containing U has a point of differentiability in S. By Lemma 2.1 this would also imply $\dim_{\mathscr{H}}(S) = 1$.

Let $f : U' \to \mathbb{R}$ be a locally Lipschitz function with domain U' containing U. Let $r_1 \in (0, r_0)$ be such that the restriction of f to $B_{r_1}(y_0)$ is Lipschitz. Then for the restriction $f|_{B_{r_1}(y_0)}$

there exists a Lipschitz extension $\tilde f$ to the whole space Y; one can take for example $\tilde f(x) = \inf_{y\in B_{r_1}(y_0)}(f(y) + L\|y - x\|)$, where $L \geqslant \operatorname{Lip}(f|_{B_{r_1}(y_0)})$.

Let $D_{\tilde f}$ be the set of points of Fréchet differentiability of $\tilde f$ inside T_{λ_0}. By Lemma 3.5 there exists a finite open interval $I = I_{\tilde f}(y_0) \ni Py_0$ such that almost every point in I belongs to $P(S_1 \cap D_{\tilde f})$, where $S_1 = T_{\lambda_0} \cap \overline{B}_{r_1/2}(y_0)$.

Since $Py_0 \in C_0$, we can find a nearby point that belongs to $C_0 \cap P(S_1 \cap D_{\tilde f})$. This means $P^{-1}(C_0)$ intersects $S_1 \cap D_{\tilde f}$. As f coincides with $\tilde f$ on $B_{r_1/2}(y_0)$ and $D_{\tilde f} \subseteq T_{\lambda_0}$ we conclude there is a point of Fréchet differentiability of f that belongs to $P^{-1}(C_0) \cap S_1 \subseteq P^{-1}(C_0) \cap T_{\lambda_0} \cap \overline{B}_{r_0/2}(y_0) = S$. □

4. Differentiability

We start this section by quoting [6, Lemma 4.2]:

Lemma 4.1. *Let $(Y, \|\cdot\|)$ be a Banach space, $f : Y \to \mathbb{R}$ be a Lipschitz function with Lipschitz constant $\operatorname{Lip}(f) > 0$ and let $\varepsilon \in (0, \operatorname{Lip}(f)/9)$. Suppose $y \in Y$, $e \in S(Y)$ and $s > 0$ are such that the directional derivative $f'(y,e)$ exists, is non-negative and*

$$\bigl|f(y+te) - f(y) - f'(y,e)t\bigr| \leqslant \frac{\varepsilon^2}{160\operatorname{Lip}(f)}|t| \tag{4.1}$$

for $|t| \leqslant s\sqrt{\frac{2\operatorname{Lip}(f)}{\varepsilon}}$. Suppose further $\xi \in (-s/2, s/2)$ and $\lambda \in Y$ satisfy

$$\bigl|f(y+\lambda) - f(y+\xi e)\bigr| \geqslant 240\varepsilon s, \tag{4.2}$$

$$\|\lambda - \xi e\| \leqslant s\sqrt{\frac{\varepsilon}{\operatorname{Lip}(f)}}, \quad \textit{and} \tag{4.3}$$

$$\frac{\|\pi s e + \lambda\|}{|\pi s + \xi|} \leqslant 1 + \frac{\varepsilon}{4\operatorname{Lip}(f)} \tag{4.4}$$

for $\pi = \pm 1$. Then if $s_1, s_2, \lambda' \in Y$ are such that

$$\max\bigl(\|s_1 - se\|, \|s_2 - se\|\bigr) \leqslant \frac{\varepsilon^2}{320\operatorname{Lip}(f)^2}s \tag{4.5}$$

and

$$\bigl\|\lambda' - \lambda\bigr\| \leqslant \frac{\varepsilon s}{16\operatorname{Lip}(f)}, \tag{4.6}$$

we can find $y' \in [y - s_1, y + \lambda'] \cup [y + \lambda', y + s_2]$ and $e' \in S(Y)$ such that the directional derivative $f'(y', e')$ exists, $f'(y',e') \geqslant f'(y,e) + \varepsilon$ and for all $t \in \mathbb{R}$ we have

$$\begin{aligned}&\bigl|\bigl(f\bigl(y'+te\bigr) - f\bigl(y'\bigr)\bigr) - \bigl(f(y+te) - f(y)\bigr)\bigr|\\ &\quad\leqslant 25\sqrt{\bigl(f'\bigl(y',e'\bigr) - f'(y,e)\bigr)\operatorname{Lip}(f)}|t|.\end{aligned} \tag{4.7}$$

Our next lemma is crucial for the proof of Theorem 3.1 and enables us to demonstrate the universal differentiability property of the set by finding a point y with almost maximal directional derivative and a family of sets around y with wedge approximation for arbitrarily small balls around y. See the definition of wedge approximation in Section 2.

Lemma 4.2 *(Differentiability Lemma). Let $(Y, \|\cdot\|)$ be a Banach space such that the norm $\|\cdot\|$ is Fréchet differentiable on $Y \setminus \{0\}$. Let $f : Y \to \mathbb{R}$ be a Lipschitz function and $(y, e) \in Y \times S(Y)$ be such that the directional derivative $f'(y, e)$ exists and is nonnegative. Suppose that there is a family of sets $\{F_\varepsilon \subseteq Y\colon \varepsilon > 0\}$ such that*

(1) *whenever $\varepsilon, \theta > 0$ there exists $\delta_* = \delta_*(\varepsilon, \theta) > 0$ such that for any $\delta \in (0, \delta_*)$ the set F_ε is a $\delta\theta$-wedge approximation for $B_\delta(y)$, and*
(2) *whenever $(y', e') \in F_\varepsilon \times S(Y)$ is such that the directional derivative $f(y', e')$ exists, $f'(y', e') \geqslant f'(y, e)$ and for any $t \in \mathbb{R}$ (4.7) is satisfied, i.e.*

$$\begin{aligned}&\bigl|\bigl(f\bigl(y'+te\bigr)-f\bigl(y'\bigr)\bigr)-\bigl(f(y+te)-f(y)\bigr)\bigr|\\&\quad\leqslant 25\sqrt{\bigl(f'\bigl(y',e'\bigr)-f'(y,e)\bigr)\operatorname{Lip}(f)|t|},\end{aligned}$$

then $f'(y', e') < f'(y, e) + \varepsilon$.

Then f is Fréchet differentiable at y.

Proof. We may assume $\operatorname{Lip}(f) = 1$. Let e^* be the Fréchet derivative of the norm $\|\cdot\|$ at e. We shall prove that f is Fréchet differentiable at y and that $f'(y)$ is given by the formula

$$f'(y)(u) = f'(y, e)e^*(u).$$

Note that $\|e^*\| = 1$ and $e^*(e) = 1$. Fix an arbitrary $\eta \in (0, 1/3)$. Choose $\Delta \in (0, \eta)$ such that

$$\bigl|\|e + th\| - \|e\| - te^*(h)\bigr| \leqslant \eta|t| \tag{4.8}$$

for any $\|h\| \leqslant 1$ and $|t| \leqslant \Delta$.

Let $\varepsilon = \eta\Delta$ and $\theta = \eta^2\Delta^2/320$. We know that the directional derivative $f'(y, e)$ exists so that we may pick $\rho \in (0, \delta_*(\varepsilon, \theta))$ such that whenever $|t| < \rho$,

$$\bigl|f(y + te) - f(y) - f'(y, e)t\bigr| < \frac{\eta^2\Delta^2}{160}|t|. \tag{4.9}$$

Let $\delta = \frac{1}{32}\rho\Delta\sqrt{\Delta\eta}$. We plan to show that

$$\bigl|f(y + ru) - f(y) - f'(y, e)e^*(u)r\bigr| < 5000\eta r \tag{4.10}$$

for any $\|u\| \leqslant 1$ and $r \in (0, \delta)$. This will imply the differentiability of f at y.

Assume for a contradiction, that there exist $r \in (0, \delta)$ and $\|u\| \leqslant 1$ such that the inequality (4.10) does not hold:

$$\bigl|f(y + ru) - f(y) - f'(y, e)e^*(u)r\bigr| \geqslant 5000\eta r. \tag{4.11}$$

Define

$$s = 16r/\Delta, \qquad \lambda = ru \quad \text{and} \quad \xi = re^*(u).$$

We check now that all the conditions of Lemma 4.1 are satisfied with $\varepsilon, s, \xi, \lambda$ defined as above.

First of all, $\varepsilon = \eta\Delta < 1/9$ and condition (4.1) follows from (4.9) as $\varepsilon^2 = \eta^2\Delta^2$ and $s\sqrt{2/\varepsilon} = 16\sqrt{2}r/(\Delta\sqrt{\eta\Delta}) < 32\delta/(\Delta\sqrt{\eta\Delta}) = \rho$.

Next we check $|\xi| < s/2$ and condition (4.2). Indeed, $|\xi| \leqslant r < r/\Delta = s/16 < s/2$. Moreover, $r \leqslant \delta < \rho$, so that we may apply (4.9) with $t = \xi$. Combining this with (4.11) we verify condition (4.2):

$$\begin{aligned}\bigl|f(y+ru) - f(y+\xi e)\bigr| &\geqslant 5000\eta r - \eta r\frac{\eta\Delta^2}{160}\bigl|e^*(u)\bigr| \\ &\geqslant 240 \cdot 16\eta r = 240 s\Delta\eta = 240 s\varepsilon.\end{aligned}$$

Note that $\|\lambda - \xi e\| = r\|u - e^*(u)e\| \leqslant 2r < 16r\sqrt{\eta/\Delta} = s\sqrt{\varepsilon}$; this establishes condition (4.3).

Finally, for $\pi = \pm 1$ we have $|\pi s + \xi| \geqslant s/2$; thus

$$t = \left\|\frac{\pi se + \lambda}{\pi s + \xi} - e\right\| = \left\|\frac{\lambda - \xi e}{\pi s + \xi}\right\| \leqslant \frac{2r}{s/2} = \Delta/4,$$

and so applying (4.8) for $h = (\frac{\lambda - \xi e}{\pi s + \xi})/t$ we get

$$\left\|\frac{\pi se + \lambda}{\pi s + \xi}\right\| \leqslant 1 + e^*\left(\frac{\lambda - \xi e}{\pi s + \xi}\right) + \eta|t|.$$

Note that $e^*(\frac{\lambda - \xi e}{\pi s + \xi}) = 0$ as $e^*(\lambda) = re^*(u) = \xi = e^*(\xi e)$ and hence (4.4):

$$\left\|\frac{\pi se + \lambda}{\pi s + \xi}\right\| \leqslant 1 + \eta\Delta/4 = 1 + \varepsilon/4.$$

Define $u_1 = -e$, $u_2 = e$ and $u_3 = (r/s)u$. Note that $r/s = \Delta/16 < 1$; thus all vectors u_1, u_2, u_3 are in the unit ball. We also have $s < 16\delta/\Delta = \frac{1}{2}\rho\sqrt{\Delta\eta} < \rho < \delta_*(\Delta\eta, \Delta^2\eta^2/320) = \delta_*(\varepsilon, \delta)$, and therefore as F_ε is an $s\theta$-wedge approximation for $B_s(y)$, we can find u_1', u_2', u_3' such that $\|u_i' - u_i\| < \theta = \Delta^2\eta^2/320$ for $i = 1, 2, 3$ and

$$[y - s_1, y + \lambda'] \cup [y + \lambda', y + s_2] \subseteq F_\varepsilon,$$

where $s_1 = -su_1'$, $s_2 = su_2'$ and $\lambda' = su_3'$. We then have $\|s_i - se\| = s\|u_i' - u_i\| \leqslant \Delta^2\eta^2 s/320$ for $i = 1, 2$ and $\|\lambda' - \lambda\| = \|su_3' - ru\| = s\|u_3' - u_3\| \leqslant \Delta^2\eta^2 s/320 < \Delta\eta s/16$, which verifies (4.5) and (4.6).

Therefore all conditions of Lemma 4.1 are satisfied; hence we may find

$$y' \in [y - s_1, y + \lambda'] \cup [y + \lambda', y + s_2] \subseteq F_\varepsilon$$

and a direction $e' \in S(Y)$ such that the directional derivative $f(y', e')$ exists, satisfies $f'(y', e') \geqslant f'(y, e) + \varepsilon$ and for all $t \in \mathbb{R}$ the inequality (4.7) holds. But for every pair (y', e') from $F_\varepsilon \times S(Y)$ that satisfies (4.7) we have $f'(y', e') < f'(y, e) + \varepsilon$, a contradiction. Hence for every $r \in (0, \delta)$ and $\|u\| \leqslant 1$, (4.10) is satisfied. □

5. Optimisation

In this section we prove Theorem 3.2. It describes how, given a Lipschitz function g on a Banach space Y and a bundle $\pi : M \to Y$, where (M, d) is a complete metric space and π is continuous, one finds a point $\tilde{x} \in M$ and direction $\tilde{e}$ in the unit sphere of Y with almost locally maximal directional derivative.

We describe how to choose the desired sequence of pairs

$$(x_n, e_n)_{n\geqslant 0} \subseteq M \times S(Y)$$

as an inductive procedure. While convergence of $(x_n)_{n\geqslant 0}$ simply follows from the fact that x_{n+1} is chosen very close to x_n, we shall need additional work in order to obtain the convergence of e_n. For this, we change the norm on each step; see (5.5) and Lemma 5.4. We then argue in Section 5.5 that the sequence of norms defined in (5.5) converges to the norm $\|\cdot\|'$ specified in Theorem 3.2.

Suppose the assumptions of Theorem 3.2 are satisfied. We thus have a Lipschitz function g acting on a Banach space Y such that the set

$$D = \bigl\{(x, e) \in M \times \bigl(Y \setminus \{0\}\bigr)\text{: the directional derivative } g'(\pi x, e) \text{ exists}\bigr\}$$

is not empty. Assume without loss of generality that $\mathrm{Lip}(g) = 1/3$.

Recall $\|e_0\| = 1$ and $g'(\pi x_0, e_0) \geqslant 0$. Choose $e_0^* \in Y^*$ with $e_0^*(e_0) = 1$ and $\|e_0^*\| = 1$, and define

$$f = g + \frac{2}{3}e_0^* \tag{5.1}$$

so that item (1) of Theorem 3.2 is satisfied. Note that f is a Lipschitz function with $\mathrm{Lip}(f) \leqslant \mathrm{Lip}(g) + \frac{2}{3} = 1$, As $f - g$ is linear, the set D is precisely the set of all $(x, e) \in M \times (Y \setminus \{0\})$ such that $f'(\pi x, e)$ exists. We can make immediately a very simple observation: if $f'(\pi x_0, e_0) \leqslant f'(\pi x, e)$ then

$$g'(\pi x_0, e_0) + \frac{2}{3} \leqslant g'(\pi x, e) + \frac{2}{3}e_0^*(e),$$

so that

$$e_0^*(e) \geqslant 1 - \frac{3}{2}g'(\pi x, e) \geqslant \frac{1}{2}. \tag{5.2}$$

Note that for any Lipschitz function $f : Y \to \mathbb{R}$ with $\mathrm{Lip}(f) \leqslant 1$ and $x, x' \in M$, $e \in Y$ with $\|e\| \leqslant 1$, we have

$$\bigl|\bigl(f\bigl(\pi x' + te\bigr) - f\bigl(\pi x'\bigr)\bigr) - \bigl(f(\pi x + te) - f(\pi x)\bigr)\bigr| \leqslant 2|t|;$$

therefore, we may assume that $\Theta(t) \leqslant 2$ for all $t > 0$.

We now introduce a function $\Omega(t) : (0, \infty) \to (0, \infty)$ that we are going to use instead of $\Theta(t)$ in our subsequent argument.

Lemma 5.1. *If $\Theta : (0, \infty) \to (0, 2]$ satisfies $\Theta(t) \to 0$ as $t \to 0$ then there exists a function $\Omega : (0, \infty) \to (0, \infty)$ such that*

(1) *$\Omega(t) \geqslant 2\Theta(t)$ for all $t \in \mathbb{R}$,*
(2) *$\Omega(t) \to 0$ as $t \to 0^+$,*
(3) *if $A, B > 0$ then $\Omega(A) + 2B \leqslant \Omega(A + B)$.*

Proof. For each $n \in \mathbb{Z}$, define $\beta(2^n) := \sup_{0<t' \leqslant 2^{n+1}} \Theta(t')$. We may uniquely extend β to $(0, \infty)$ by imposing the property that β is affine on each interval of the form $[2^n, 2^{n+1}]$ for $n \in \mathbb{Z}$. Note that β is continuous, increasing and $\beta(t) \geqslant \Theta(t)$ for every $t > 0$. Further for $t \leqslant 2^n$ where $n \in \mathbb{Z}$, we have $\beta(t) \leqslant \beta(2^n) = \sup_{0<t' \leqslant 2^{n+1}} \Theta(t')$ and as $\Theta(t) \to 0$ as $t \to 0^+$ we deduce that $\beta(t) \to 0$ as $t \to 0^+$.

We now let $\Omega(t) = 2\beta(t) + 2t$. Then (1) and (2) are immediate as $\beta(t) \geqslant \Theta(t)$ and $\beta(t) \to 0$ as $t \to 0^+$. Finally for (3) we may use the fact that β is increasing to deduce that for $A, B > 0$, $\Omega(A + B) = 2\beta(A + B) + 2A + 2B \geqslant 2\beta(A) + 2A + 2B = \Omega(A) + 2B$. □

We now define a notion of weight and a class of pairs that weigh more than the given pair.

Definition 5.2. If p is a norm on Y and $(x, e) \in D$ then we call

$$w_p(x, e) = \frac{f'(\pi x, e)}{p(e)}$$

the weight of (x, e) with respect to the norm p.

Further for $\sigma \geqslant 0$ we let $G_p(x, e, \sigma)$ be the set of all $(x', e') \in D$ such that

$$w_p(x, e) \leqslant w_p\big(x', e'\big) \tag{5.3}$$

and

$$\begin{aligned} &\big|\big(f\big(\pi x' + te\big) - f\big(\pi x'\big)\big) - \big(f(\pi x + te) - f(\pi x)\big)\big| \\ &\quad \leqslant \big(\sigma + \Omega\big(w_p\big(x', e'\big) - w_p(x, e)\big)\big)|t| \end{aligned} \tag{5.4}$$

for all $t \in \mathbb{R}$, where the function Ω is given by Lemma 5.1.

In what follows, the notation $\|y - \mathbb{R}e\|$ where $y \in Y$ and $e \in Y \setminus \{0\}$ is used for the distance between the point y and the one-dimensional subspace of Y generated by e. This distance is calculated with the original norm $\|\cdot\|$ on Y.

5.3. Inductive construction

Let $\sigma_0 = 16$, $\delta_0 = 1$, $t_0 \in (0, 1/2)$, the norm $p_0 = \|\cdot\|$ and $w_0 = w_{p_0}$. The pair (x_0, e_0) was chosen earlier. Below we will define various positive parameters $\sigma_n, t_n, \varepsilon_n, \nu_n, \Delta_n, \delta_n$, nested sequence D_n of non-empty subsets of D and pairs $(x_n, e_n) \in D_n$. For every $n \geqslant 1$, we define

$$p_n(y) = \sqrt{\|y\|^2 + \sum_{m=0}^{n-1} t_m^2 \|y - \mathbb{R}e_m\|^2} \tag{5.5}$$

and let $w_n = w_{p_n}$ be the weight function defined on D. It is clear (5.5) defines a norm on Y and $p_n(y) \geqslant \max\{\|y\|, p_{n-1}(y)\}$ for all $y \in Y$. Together with $\mathrm{Lip}(f) \leqslant 1$, this implies $w_n(x, e) \leqslant \min\{1, w_{n-1}(x, e)\}$ for any $(x, e) \in D$.

For every $n \geqslant 1$, choose

$$\begin{aligned} &\sigma_n \in (0, \sigma_{n-1}/16), \quad t_n \in (0, t_{n-1}/2) \quad \text{with } t_n^2 < \sigma_{n-1}/16 \quad \text{and} \\ &\varepsilon_n \in \left(0, t_n^2 \sigma_n^2 / 2^{13}\right). \end{aligned} \tag{5.6}$$

Let D_n to be the set of all pairs $(x, e) \in D$ with $d(x, x_{n-1}) < \delta_{n-1}$, $\|e\| = 1$ and

$$(x, e) \in G_{p_n}(x_{n-1}, e_{n-1}, \sigma_{n-1} - \nu)$$

for some $\nu \in (0, \sigma_{n-1}/2)$. Note that $(x_{n-1}, e_{n-1}) \in D_n$, and so $D_n \neq \emptyset$. Since w_n is bounded by 1 from above we can choose $(x_n, e_n) \in D_n$ such that for every $(x, e) \in D_n$

$$w_n(x, e) \leqslant w_n(x_n, e_n) + \varepsilon_n. \tag{5.7}$$

Note that the definition of D_n then implies $d(x_n, x_{n-1}) < \delta_{n-1}$, and as $(x_n, e_n) \in D_n$ and $p_n(e_{n-1}) = p_{n-1}(e_{n-1})$, we have for every $n \geqslant 1$

$$w_{n-1}(x_{n-1}, e_{n-1}) = w_n(x_{n-1}, e_{n-1}) \leqslant w_n(x_n, e_n). \tag{5.8}$$

This implies $w_n(x, e) \geqslant w_0(x_0, e_0) = f'(\pi x_0, e_0)$ for every $(x, e) \in D_n$; in particular, (5.2) implies

$$e_0^*(e) \geqslant 1/2 \tag{5.9}$$

for any $(x, e) \in D_n$.

Let $\nu_n \in (0, \sigma_{n-1}/2)$ be such that $(x_n, e_n) \in G_{p_n}(x_{n-1}, e_{n-1}, \sigma_{n-1} - \nu_n)$. Finally pick $\Delta_n > 0$ such that

$$\left| f(\pi x_n + t e_n) - f(\pi x_n) - f'(\pi x_n, e_n)t \right| \leqslant \sigma_{n-1}|t|/32, \tag{5.10}$$

$$\left| f(\pi x_{n-1} + t e_{n-1}) - f(\pi x_{n-1}) - f'(\pi x_{n-1}, e_{n-1})t \right| \leqslant \sigma_{n-1}|t|/32 \tag{5.11}$$

for all t with $|t| \leqslant 4\Delta_n/\nu_n$. Choose $\delta_n \in (0, (\delta_{n-1} - d(x_n, x_{n-1}))/2)$ such that $\|\pi x - \pi x_n\| \leqslant \Delta_n$ whenever $d(x, x_n) \leqslant \delta_n$; such a δ_n exists because π is continuous.

Let us make some simple observations. First of all, (5.6) implies that the sequences $\sigma_n, t_n, \varepsilon_n$ all tend to zero. Since $\nu_n < \sigma_{n-1}/2$ and δ_n satisfies the inequality $\delta_n < (\delta_{n-1} - d(x_n, x_{n-1}))/2$ we conclude that ν_n and δ_n tend to zero, too. The latter inequality also implies

$$\overline{B}_{\delta_n}(x_n) \subseteq B_{\delta_{n-1}}(x_{n-1}) \tag{5.12}$$

for every $n \geqslant 1$ and so

$$d(x_k, x_n) < \delta_n \quad \text{for all } k \geqslant n. \tag{5.13}$$

Since M is complete we conclude that the sequence (x_n) converges in M to some point x_∞.

The inequality $t_n < t_{n-1}/2$ also implies $p_n(y)^2 \leqslant \|y\|^2 + 2t_0^2 \cdot \|y\|^2 \leqslant 2\|y\|^2$, so for all $y \in Y$,

$$\|y\| \leqslant p_n(y) \leqslant 2\|y\|. \tag{5.14}$$

Then, using $p_n(e_{n-1}) \leqslant 2$, we get for every $(x, e) \in D$

$$\begin{aligned}
&\left|f'(\pi x, e) - f'(\pi x_{n-1}, e_{n-1})\right| \\
&\quad \leqslant 2\frac{|f'(\pi x, e) - f'(\pi x_{n-1}, e_{n-1})|}{p_n(e_{n-1})} \\
&\quad \leqslant 2\left|\frac{f'(\pi x, e)}{p_n(e)} - \frac{f'(\pi x_{n-1}, e_{n-1})}{p_n(e_{n-1})}\right| + 2\left|f'(\pi x, e)\right|\left|\frac{1}{p_n(e_{n-1})} - \frac{1}{p_n(e)}\right| \\
&\quad \leqslant 2\left|w_n(x, e) - w_n(x_{n-1}, e_{n-1})\right| + 2\frac{\|e\|}{p_n(e_{n-1})p_n(e)}\left|p_n(e) - p_n(e_{n-1})\right| \\
&\quad \leqslant 2\left|w_n(x, e) - w_n(x_{n-1}, e_{n-1})\right| + 4\|e - e_{n-1}\|,
\end{aligned} \tag{5.15}$$

where, in the penultimate line, we are using $\mathrm{Lip}(f) \leqslant 1$ and, in the final line, $p_n(e) \geqslant \|e\|$, $p_n(e_{n-1}) \geqslant \|e_{n-1}\| = 1$ and the fact that

$$\left|p_n(e) - p_n(e_{n-1})\right| \leqslant p_n(e - e_{n-1}) \leqslant 2\|e - e_{n-1}\|.$$

We are now ready to prove a very important property of sets D_n; the "moreover" part of Lemma 5.4 together with (5.6) implies the convergence of the sequence (e_n) to some $e_\infty \in Y$ with $\|e_\infty\| = 1$. We will show later that the pair (x_∞, e_∞) has the properties required by Theorem 3.2.

Lemma 5.4. *For every $n \geqslant 1$, we have $D_{n+1} \subseteq G_{p_n}(x_{n-1}, e_{n-1}, \sigma_{n-1} - \nu_n/2)$ and $D_{n+1} \subseteq D_n$. Moreover, for any $(x, e) \in D_{n+1}$ we have $\|e - e_n\| \leqslant \sigma_n/8$.*

Proof. Notice first that since $\sigma_0 = 16$, the "moreover" statement is satisfied for $n = 0$.

We shall now show that assuming the latter statement is satisfied for $n - 1$, where $n \geqslant 1$, the full conclusion of the present lemma holds for n.

Assume therefore $n \geqslant 1$, the "moreover" part is satisfied for $n - 1$ and $(x, e) \in D_{n+1}$. Since $(x_n, e_n) \in D_n$, we get

$$\|e_n - e_{n-1}\| \leqslant \frac{\sigma_{n-1}}{8}. \tag{5.16}$$

Since $(x,e) \in G_{p_{n+1}}(x_n, e_n, \sigma_n - \nu)$ for some $\nu > 0$, we get that $w_{n+1}(x,e) \geqslant w_{n+1}(x_n, e_n)$; thus using (5.8), we obtain the first defining property of the set $G_{p_n}(x_{n-1}, e_{n-1}, *)$:

$$w_n(x,e) \geqslant w_{n+1}(x,e) \geqslant w_{n+1}(x_n,e_n) = w_n(x_n,e_n) \geqslant w_n(x_{n-1},e_{n-1}).$$

In order to show $(x,e) \in G_{p_n}(x_{n-1}, e_{n-1}, \sigma_{n-1} - \nu_n/2)$, we need to prove the second defining property of the latter set. We prove the inequality separately for $|t| < 4\Delta_n/\nu_n$ and $|t| \geqslant 4\Delta_n/\nu_n$.

If $|t| < 4\Delta_n/\nu_n$, using first (5.10), (5.11) and then $\mathrm{Lip}(f) \leqslant 1$,

$$\begin{aligned}
&\bigl|\bigl(f(\pi x + te_{n-1}) - f(\pi x)\bigr) - \bigl(f(\pi x_{n-1} + te_{n-1}) - f(\pi x_{n-1})\bigr)\bigr| \\
&\quad \leqslant \bigl|\bigl(f(\pi x + te_{n-1}) - f(\pi x)\bigr) - \bigl(f(\pi x_n + te_n) - f(\pi x_n)\bigr)\bigr| \\
&\qquad + \bigl|f'(\pi x_n, e_n) - f'(\pi x_{n-1}, e_{n-1})\bigr| \cdot |t| + \frac{1}{16}\sigma_{n-1}|t| \\
&\quad \leqslant \bigl|\bigl(f(\pi x + te_n) - f(\pi x)\bigr) - \bigl(f(\pi x_n + te_n) - f(\pi x_n)\bigr)\bigr| + \|e_n - e_{n-1}\| \cdot |t| \\
&\qquad + \bigl|f'(\pi x_n, e_n) - f'(\pi x_{n-1}, e_{n-1})\bigr| \cdot |t| + \frac{1}{16}\sigma_{n-1}|t|.
\end{aligned}$$

We may now apply (5.16), $(x,e) \in G_{p_{n+1}}(x_n, e_n, \sigma_n - \nu)$ and (5.15) to deduce that the latter is bounded from above by

$$\begin{aligned}
&|t|\bigl(\sigma_n - \nu + \Omega\bigl(w_{n+1}(x,e) - w_{n+1}(x_n,e_n)\bigr)\bigr) + \frac{3}{16}\sigma_{n-1} \\
&\quad + 2\bigl(w_n(x_n,e_n) - w_n(x_{n-1},e_{n-1}) + 4\|e_n - e_{n-1}\|\bigr).
\end{aligned} \tag{5.17}$$

Recall that Ω is an increasing function and

$$w_{n+1}(x,e) - w_{n+1}(x_n,e_n) = w_{n+1}(x,e) - w_n(x_n,e_n) \leqslant w_n(x,e) - w_n(x_n,e_n);$$

then using again (5.16), $\sigma_n \in (0, \sigma_{n-1}/16)$ and $\nu_n \in (0, \sigma_{n-1}/2)$ so that $\frac{3}{4}\sigma_{n-1} \leqslant \sigma_{n-1} - \nu_n/2$, and Lemma 5.1(3), we have

$$\begin{aligned}
&\bigl|\bigl(f(\pi x + te_{n-1}) - f(\pi x)\bigr) - \bigl(f(\pi x_{n-1} + te_{n-1}) - f(\pi x_{n-1})\bigr)\bigr| \\
&\quad \leqslant \left(\frac{3}{4}\sigma_{n-1} + \Omega\bigl(w_n(x,e) - w_n(x_n,e_n)\bigr) + 2\bigl(w_n(x_n,e_n) - w_n(x_{n-1},e_{n-1})\bigr)\right)|t| \\
&\quad \leqslant \bigl(\sigma_{n-1} - \nu_n/2 + \Omega\bigl(w_n(x,e) - w_n(x_{n-1},e_{n-1})\bigr)\bigr)|t|.
\end{aligned}$$

Now we consider the case $|t| \geqslant 4\Delta_n/\nu_n$. As $(x,e) \in D_{n+1}$, we have $d(x,x_n) < \delta_n$. Therefore, from the definition of δ_n, we have $\|\pi x - \pi x_n\| \leqslant \Delta_n \leqslant \nu_n|t|/4$. Thus, replacing $f(\pi x + te_{n-1})$ with $f(\pi x_n + te_{n-1})$ and $f(\pi x)$ with $f(\pi x_n)$, we get

$$\begin{aligned}
&\bigl|\bigl(f(\pi x + te_{n-1}) - f(\pi x)\bigr) - \bigl(f(\pi x_{n-1} + te_{n-1}) - f(\pi x_{n-1})\bigr)\bigr| \\
&\quad \leqslant \nu_n|t|/2 + \bigl|\bigl(f(\pi x_n + te_{n-1}) - f(\pi x_n)\bigr) - \bigl(f(\pi x_{n-1} + te_{n-1}) - f(\pi x_{n-1})\bigr)\bigr|.
\end{aligned}$$

Now using $(x_n, e_n) \in G_{p_n}(x_{n-1}, e_{n-1}, \sigma_{n-1} - \nu_n)$, we estimate the second term by $(\sigma_{n-1} - \nu_n + \Omega(w_n(x_n, e_n) - w_n(x_{n-1}, e_{n-1})))|t|$. Adding $\nu_n|t|/2$ to this and noting Ω is an increasing function, we estimate this from above by

$$\bigl(\sigma_{n-1} - \nu_n/2 + \Omega\bigl(w_n(x,e) - w_n(x_{n-1}, e_{n-1})\bigr)\bigr)|t|.$$

This finishes the proof of $(x,e) \in G_{p_n}(x_{n-1}, e_{n-1}, \sigma_{n-1} - \nu_n/2)$.

Further, for $(x,e) \in D_{n+1}$ we have $\|e\| = 1$ and $d(x, x_{n-1}) < \delta_{n-1}$, using inequality $d(x, x_n) < \delta_n$ and (5.12). Therefore, $(x,e) \in D_n$; hence $D_{n+1} \subseteq D_n$.

Finally to prove $\|e - e_n\| \leqslant \sigma_n/8$, note that (5.8) together with the definition of (x_n, e_n) implies

$$w_n(x_n, e_n) = w_{n+1}(x_n, e_n) \leqslant \frac{p_n(e)}{p_{n+1}(e)} w_n(x,e) \leqslant \frac{p_n(e)}{p_{n+1}(e)} \bigl(w_n(x_n, e_n) + \varepsilon_n\bigr). \tag{5.18}$$

Writing $p_{n+1}(e) = \sqrt{p_n^2(e) + t_n^2 d^2}$, where $d = \|e - \mathbb{R}e_n\| \leqslant 1$ and using $t_n < t_0 < 1/2$ we deduce

$$\frac{p_n(e)}{p_{n+1}(e)} = 1/\sqrt{1 + t_n^2 d^2/p_n(e)^2} \leqslant 1 - \frac{t_n^2 d^2}{4p_n(e)^2}$$

as $1/\sqrt{1+x} \leqslant 1 - x/4$ for $0 \leqslant x \leqslant 1$. Substituting this inequality into (5.18) and using (5.6) we obtain

$$\frac{t_n^2 d^2}{4p_n(e)^2} w_n(x_n, e_n) \leqslant \varepsilon_n \left(1 - \frac{t_n^2 d^2}{4p_n(e)^2}\right) \leqslant \varepsilon_n < t_n^2 \sigma_n^2/2^{13}.$$

On the other hand, (5.8) and $g'(\pi x_0, e_0) \geqslant 0$ imply

$$w_n(x_n, e_n) \geqslant w_0(x_0, e_0) = f'(\pi x_0, e_0) = g'(\pi x_0, e_0) + \frac{2}{3} > \frac{1}{2},$$

so using $p_n(e) \leqslant 2$ we conclude $d \leqslant \sigma_n/2^4$. This means there is a $t \in \mathbb{R}$ such that

$$\|e - te_n\| \leqslant \frac{\sigma_n}{16}. \tag{5.19}$$

It follows $|e_0^*(e - te_n)| \leqslant \sigma_n/16 \leqslant 1/2$. However, by (5.9), $e_0^*(e), e_0^*(e_n) \geqslant 1/2$, hence $t \geqslant 0$. Then from (5.19) and $\|e_n\| = \|e\| = 1$ we get that $|1 - t| \leqslant \frac{\sigma_n}{16}$ and so

$$\|e - e_n\| \leqslant \frac{\sigma_n}{8}. \qquad \square$$

We note here that Lemma 5.4 implies that $\|e_m - e_n\| \leqslant \sigma_n/8$ whenever $m \geqslant n+1$. Thus (e_n) is a Cauchy sequence, so it converges. Let $e_\infty = \lim e_n$. As $\|e_n\| = 1$ for each $n \geqslant 1$, we have $\|e_\infty\| = 1$.

5.5. Existence of directional derivative $f'(\pi x_\infty, e_\infty)$

From 5.13 we have $d(x_k, x_n) < \delta_n$ for all $k \geqslant n$. We also know that for $k \geqslant n$, $(x_k, e_k) \in D_{k+1} \subseteq D_{n+1}$ using Lemma 5.4, so that $\|e_k - e_n\| \leqslant \sigma_n/8$, again by Lemma 5.4. Hence the sequences x_n and e_n converge to x_∞ and e_∞ respectively, where

$$d(x_\infty, x_n) < \delta_n \quad \text{and} \quad \|e_\infty - e_n\| \leqslant \sigma_n/8 \tag{5.20}$$

are satisfied for every $n \geqslant 1$, the strictness of the first inequality following from (5.12). It is also clear that the sequence of norms p_n converges to

$$p_\infty(y) = \sqrt{\|y\|^2 + \sum_{m=1}^{\infty} t_m^2 \|y - \mathbb{R}e_m\|^2}$$

as this formula defines a norm and

$$p_n^2(y) \leqslant p_\infty^2(y) \leqslant p_n^2(y) + 2t_n^2\|y\|^2 \leqslant \left(1 + 2t_n^2\right)p_n^2(y) \leqslant \left(1 + t_n^2\right)^2 p_n^2(y)$$

implies for all $y \in Y$

$$p_n(y) \leqslant p_\infty(y) \leqslant \left(1 + t_n^2\right)p_n(y). \tag{5.21}$$

This implies for every $(x, e) \in D$

$$\left|w_n(x, e) - w_\infty(x, e)\right| = \left|f'(x, e)\right| \cdot \frac{|p_\infty(e) - p_n(e)|}{p_n(e)p_\infty(e)} \leqslant \|e\| \frac{t_n^2 \cdot p_n(e)}{p_n(e)p_\infty(e)} \leqslant t_n^2 \tag{5.22}$$

using $\mathrm{Lip}(f) \leqslant 1$ and $p_\infty(e) \geqslant \|e\|$.

We will now show that the directional derivative $f'(\pi x_\infty, e_\infty)$ exists and

$$w_m(x_m, e_m) \nearrow w_\infty(x_\infty, e_\infty), \tag{5.23}$$

where $w_\infty = w_{p_\infty}$.

Indeed, for every $n \geqslant 1$, the inequality $p_n(y) \geqslant \|y\|$ and (5.8) imply

$$0 < w_0(x_0, e_0) \leqslant w_n(x_n, e_n) \leqslant \mathrm{Lip}(f) \leqslant 1.$$

Thus there is $L \in (0, 1]$ such that $w_n(x_n, e_n) \nearrow L$. From (5.21) we conclude $w_\infty(x_n, e_n) \to L$ and $w_{n+1}(x_n, e_n) \to L$. Note then

$$w_m(x_n, e_n) - w_m(x_{m-1}, e_{m-1}) \underset{n\to\infty}{\longrightarrow} \frac{p_\infty(e_\infty)}{p_m(e_\infty)}L - w_m(x_{m-1}, e_{m-1}) =: s_m \underset{m\to\infty}{\longrightarrow} 0.$$

Assuming $n \geqslant m$ we get $(x_n, e_n) \in D_n \subseteq D_{m+1}$. The first condition (5.3) of

$$(x_n, e_n) \in G_{p_m}(x_{m-1}, e_{m-1}, \sigma_{m-1} - \nu_m/2) \tag{5.24}$$

says $w_m(x_n, e_n) \geqslant w_m(x_{m-1}, e_{m-1})$, thus $s_m \geqslant 0$ for each m. Taking $n \to \infty$ in the second inequality (5.4) from the definition of (5.24), we obtain

$$\bigl|\bigl(f(\pi x_\infty + te_{m-1}) - f(\pi x_\infty)\bigr) - \bigl(f(\pi x_{m-1} + te_{m-1}) - f(\pi x_{m-1})\bigr)\bigr| \leqslant r_m|t| \quad (5.25)$$

for any $t \in \mathbb{R}$, where

$$r_m := \sigma_{m-1} - \nu_m/2 + \Omega(s_m) \to 0$$

by Lemma 5.1(2). Using $\|e_\infty - e_{m-1}\| \leqslant \sigma_{m-1}$ and $\operatorname{Lip}(f) \leqslant 1$:

$$\begin{aligned}&\bigl|\bigl(f(\pi x_\infty + te_\infty) - f(\pi x_\infty)\bigr) - \bigl(f(\pi x_{m-1} + te_{m-1}) - f(\pi x_{m-1})\bigr)\bigr| \\ &\quad \leqslant (r_m + \sigma_{m-1})|t|. \end{aligned} \quad (5.26)$$

Let $\varepsilon > 0$. Note that as

$$f'(\pi x_{m-1}, e_{m-1}) = p_{m-1}(e_{m-1})w_{m-1}(x_{m-1}, e_{m-1}) \to p_\infty(e_\infty)L$$

we may pick m such that

$$r_m + \sigma_{m-1} \leqslant \varepsilon/3 \quad \text{and} \quad \bigl|f'(\pi x_{m-1}, e_{m-1}) - p_\infty(e_\infty)L\bigr| \leqslant \varepsilon/3 \quad (5.27)$$

and then $\delta > 0$ with

$$\bigl|f(\pi x_{m-1} + te_{m-1}) - f(\pi x_{m-1}) - f'(\pi x_{m-1}, e_{m-1})t\bigr| \leqslant \varepsilon|t|/3 \quad (5.28)$$

for all t with $|t| \leqslant \delta$. Combining (5.26), (5.27) and (5.28) we obtain

$$\bigl|f(\pi x_\infty + te_\infty) - f(\pi x_\infty) - p_\infty(e_\infty)Lt\bigr| \leqslant \varepsilon|t|$$

for $|t| \leqslant \delta$. Hence the directional derivative $f'(\pi x_\infty, e_\infty)$ exists and equals $p_\infty(e_\infty)L$ and $w_\infty(x_\infty, e_\infty) = L$.

The last equality and the definition of s_m implies

$$w_m(x_\infty, e_\infty) - w_m(x_{m-1}, e_{m-1}) = s_m \geqslant 0,$$

so together with (5.25) we get

$$(x_\infty, e_\infty) \in G_{p_m}(x_{m-1}, e_{m-1}, \sigma_{m-1} - \nu_m/2) \quad (5.29)$$

and so $(x_\infty, e_\infty) \in D_m$ for all $m \geqslant 1$.

5.6. Maximality of the weight function at (x_∞, e_∞)

We now verify that the value of the weight function $w_\infty(x_\infty, e_\infty)$ is almost maximal in the following sense: For every $\varepsilon > 0$ there exists $\delta > 0$ such that whenever $(x', e') \in G_{p_\infty}(x_\infty, e_\infty, 0)$ with $d(x', x_\infty) \leqslant \delta$ then we have

$$w_\infty\big(x', e'\big) < w_\infty(x_\infty, e_\infty) + \varepsilon. \tag{5.30}$$

Assume $\varepsilon > 0$ is fixed, choose then $n \geqslant 1$ with $\varepsilon_n + 2t_n^2 < \varepsilon$ and pick $\Delta > 0$ such that for $|t| < 8\Delta/\nu_n$, the following two inequalities are satisfied:

$$\big|f(\pi x_\infty + te_\infty) - f(\pi x_\infty) - f'(\pi x_\infty, e_\infty)t\big| \leqslant \frac{1}{16}\sigma_{n-1}|t|; \tag{5.31}$$

$$\big|f(\pi x_{n-1} + te_{n-1}) - f(\pi x_{n-1}) - f'(\pi x_{n-1}, e_{n-1})t\big| \leqslant \frac{1}{16}\sigma_{n-1}|t|. \tag{5.32}$$

Using (5.20) and the continuity of π we can find

$$\delta \in \big(0, \delta_{n-1} - d(x_\infty, x_{n-1})\big) \tag{5.33}$$

such that whenever $d(x', x_\infty) \leqslant \delta$,

$$\big\|\pi x' - \pi x_\infty\big\| \leqslant \Delta. \tag{5.34}$$

We now suppose, for a contradiction, that $(x', e') \in G_{p_\infty}(x_\infty, e_\infty, 0)$ is such that $d(x', x_\infty) \leqslant \delta$ and contrary to (5.30) we have

$$w_\infty\big(x', e'\big) \geqslant w_\infty(x_\infty, e_\infty) + \varepsilon. \tag{5.35}$$

As $w_\infty(x', e')$ is invariant if we scale e' by a positive factor, as is the membership relation $(x', e') \in G_{p_\infty}(x_\infty, e_\infty, 0)$, we may assume that $\|e'\| = 1$.

First we shall show that $(x', e') \in D_n$. Since (5.33) and $d(x', x_\infty) \leqslant \delta$ imply $d(x', x_{n-1}) < \delta_{n-1}$, by definition of D_n to prove $(x', e') \in D_n$ it is enough to show that

$$\big(x', e'\big) \in G_{p_n}(x_{n-1}, e_{n-1}, \sigma_{n-1} - \nu_n/4). \tag{5.36}$$

Note that from (5.22) we have

$$\begin{aligned} w_n\big(x', e'\big) - w_n(x_\infty, e_\infty) &\geqslant w_\infty\big(x', e'\big) - w_\infty(x_\infty, e_\infty) - 2t_n^2 \\ &\geqslant \varepsilon - 2t_n^2 \geqslant \varepsilon_n > 0; \end{aligned} \tag{5.37}$$

therefore

$$w_n\big(x', e'\big) > w_n(x_\infty, e_\infty) \geqslant w_n(x_{n-1}, e_{n-1})$$

as $s_n = w_n(x_\infty, e_\infty) - w_n(x_{n-1}, e_{n-1}) \geqslant 0$: see the end of Section 5.5.

We now check the second condition of (5.36). Assume $|t| < 8\Delta/\nu_n$; then using first (5.31), (5.32) and then $\mathrm{Lip}(f) \leqslant 1$,

$$\begin{aligned}
&\bigl|\bigl(f\bigl(\pi x' + te_{n-1}\bigr) - f\bigl(\pi x'\bigr)\bigr) - \bigl(f(\pi x_{n-1} + te_{n-1}) - f(\pi x_{n-1})\bigr)\bigr| \\
&\quad \leqslant \bigl|\bigl(f\bigl(\pi x' + te_{n-1}\bigr) - f\bigl(\pi x'\bigr)\bigr) - \bigl(f(\pi x_\infty + te_\infty) - f(\pi x_\infty)\bigr)\bigr| \\
&\qquad + \bigl|f'(\pi x_\infty, e_\infty) - f'(\pi x_{n-1}, e_{n-1})\bigr| \cdot |t| + \frac{1}{8}\sigma_{n-1}|t| \\
&\quad \leqslant \bigl|\bigl(f\bigl(\pi x' + te_\infty\bigr) - f\bigl(\pi x'\bigr)\bigr) - \bigl(f(\pi x_\infty + te_\infty) - f(\pi x_\infty)\bigr)\bigr| + \|e_\infty - e_{n-1}\| \cdot |t| \\
&\qquad + \bigl|f'(\pi x_\infty, e_\infty) - f'(\pi x_{n-1}, e_{n-1})\bigr| \cdot |t| + \frac{1}{8}\sigma_{n-1}|t|.
\end{aligned}$$

In this sum of four terms, we use $(x', e') \in G_{p_\infty}(x_\infty, e_\infty, 0)$ to bound the first term from above by $\Omega(w_\infty(x', e') - w_\infty(x_\infty, e_\infty)) \cdot |t|$, and (5.15) to bound the third term by $(2(w_n(x_\infty, e_\infty) - w_n(x_{n-1}, e_{n-1})) + 4\|e_\infty - e_{n-1}\|)|t|$. Using in addition inequality $\|e_\infty - e_{n-1}\| \leqslant \sigma_{n-1}/8$ from (5.20), we get

$$\begin{aligned}
&\bigl|\bigl(f\bigl(\pi x' + te_{n-1}\bigr) - f\bigl(\pi x'\bigr)\bigr) - \bigl(f(\pi x_{n-1} + te_{n-1}) - f(\pi x_{n-1})\bigr)\bigr| \\
&\quad \leqslant |t|\biggl(\Omega\bigl(w_\infty\bigl(x', e'\bigr) - w_\infty(x_\infty, e_\infty)\bigr) \\
&\qquad + 2\bigl(w_n(x_\infty, e_\infty) - w_n(x_{n-1}, e_{n-1})\bigr) + \frac{3}{4}\sigma_{n-1}\biggr).
\end{aligned} \tag{5.38}$$

We now use the fact that Ω is an increasing function and $w_\infty(x', e') \leqslant w_n(x', e')$ which follows from $p_n \leqslant p_\infty$, and use (5.22) to estimate $2w_n(x_\infty, e_\infty)$ from above by $2w_\infty(x_\infty, e_\infty) + 2t_n^2$. Then the expression in the right hand side of (5.38) is less than or equal to

$$|t|\biggl(\Omega\bigl(w_n\bigl(x', e'\bigr) - w_\infty(x_\infty, e_\infty)\bigr) + 2\bigl(w_\infty(x_\infty, e_\infty) - w_n(x_{n-1}, e_{n-1})\bigr) + \frac{3}{4}\sigma_{n-1} + 2t_n^2\biggr).$$

As $t_n^2 < \sigma_{n-1}/16$ and Ω satisfies property (3) in Lemma 5.1, we finally get

$$\begin{aligned}
&\bigl|\bigl(f\bigl(\pi x' + te_{n-1}\bigr) - f\bigl(\pi x'\bigr)\bigr) - \bigl(f(\pi x_{n-1} + te_{n-1}) - f(\pi x_{n-1})\bigr)\bigr| \\
&\quad \leqslant \biggl(\Omega\bigl(w_n\bigl(x', e'\bigr) - w_n(x_{n-1}, e_{n-1})\bigr) + \frac{7}{8}\sigma_{n-1}\biggr)|t| \\
&\quad \leqslant \bigl(\sigma_{n-1} - \nu_n/4 + \Omega\bigl(w_n\bigl(x', e'\bigr) - w_n(x_{n-1}, e_{n-1})\bigr)\bigr)|t|,
\end{aligned}$$

as $\nu_n < \sigma_{n-1}/2$. Thus we proved the second condition of (5.36) for $|t| < 8\Delta/\nu_n$.

Now we consider the case $|t| \geqslant 8\Delta/\nu_n$. From $d(x', x_\infty) \leqslant \delta$ and (5.34) we have

$$\bigl\|\pi x' - \pi x_\infty\bigr\| \leqslant \Delta \leqslant \nu_n|t|/8$$

so we get, using $(x_\infty, e_\infty) \in G_{p_n}(x_{n-1}, e_{n-1}, \sigma_{n-1} - \nu_n/2)$ from (5.29),

$$\begin{aligned}
&\bigl|\bigl(f\bigl(\pi x'+te_{n-1}\bigr)-f\bigl(\pi x'\bigr)\bigr)-\bigl(f(\pi x_{n-1}+te_{n-1})-f(\pi x_{n-1})\bigr)\bigr|\\
&\quad\leqslant\bigl|\bigl(f(\pi x_\infty+te_{n-1})-f(\pi x_\infty)\bigr)-\bigl(f(\pi x_{n-1}+te_{n-1})-f(\pi x_{n-1})\bigr)\bigr|+2\bigl\|\pi x'-\pi x_\infty\bigr\|\\
&\quad\leqslant\bigl(\sigma_{n-1}-\nu_n/2+\Omega\bigl(w_n(x_\infty,e_\infty)-w_n(x_{n-1},e_{n-1})\bigr)\bigr)|t|+\nu_n|t|/4\\
&\quad\leqslant\bigl(\sigma_{n-1}-\nu_n/4+\Omega\bigl(w_n\bigl(x',e'\bigr)-w_n(x_{n-1},e_{n-1})\bigr)\bigr)|t|,
\end{aligned}$$

where, in the final line, we have used $w_n(x',e')\geqslant w_n(x_\infty,e_\infty)$ from (5.37).

This finishes the proof of $(x',e')\in D_n$ for every $n\geqslant 1$. Recall the property of the pair $(x_n,e_n)\in D_n$ is such that

$$w_n(x,e)\leqslant w_n(x_n,e_n)+\varepsilon_n$$

for all $(x,e)\in D_n$. Notice that by (5.23) the right hand side of this inequality is less than or equal to $w_\infty(x_\infty,e_\infty)+\varepsilon_n$, thus together with (5.35) and (5.22) we finally get

$$w_\infty(x_\infty,e_\infty)+\varepsilon\leqslant w_\infty\bigl(x',e'\bigr)\leqslant w_n\bigl(x',e'\bigr)+t_n^2\leqslant w_\infty(x_\infty,e_\infty)+\varepsilon_n+t_n^2.$$

This is a contradiction as $\varepsilon>\varepsilon_n+t_n^2$. This means that the assumption (5.35) is false, completing the proof of the statement of the present section.

5.7. *Proof of Theorem 3.2*

We first quote [12, Lemma 4.3] for determining the Fréchet differentiability of the norm p_∞:

Lemma. *If the norm of a Banach space Y is Fréchet differentiable on $Y\setminus\{0\}$, $e_m\in Y$ and $t_m\geqslant 0$ with $\sum t_m^2<\infty$, then the function $p:Y\to\mathbb{R}$ defined by the formula*

$$p(y):=\sqrt{\|y\|^2+\sum_{m=1}^{\infty}t_m^2\|y-\mathbb{R}e_m\|^2}$$

is an equivalent norm on Y that is Fréchet differentiable on $Y\setminus\{0\}$.

We verify the conclusions of Theorem 3.2 for Lipschitz function f defined in (5.1) and the norm $\|\cdot\|'=p_\infty$.

The items (1) and (2) of Theorem 3.2 follow from (5.1) and (5.14) as $p_n\to p_\infty$. The "moreover" statement in Theorem 3.2 is a direct consequence of the lemma quoted above and the definition of p_∞ in Section 5.5.

For part (3) of Theorem 3.2 we define $\tilde{x}=x_\infty$ and $\tilde{e}=e_\infty/\|e_\infty\|'$. Then we have $(\tilde{x},\tilde{e})\in D$ and $\|\tilde{e}\|'=1$. Further we have

$$f'(\pi\tilde{x},\tilde{e})=w_\infty(\tilde{x},\tilde{e})=w_\infty(x_\infty,e_\infty)\geqslant w_0(x_0,e_0)=f'(\pi x_0,e_0)$$

by Definition 5.2 and (5.23). Now given any $\varepsilon>0$ we choose $\delta>0$ as in Section 5.6 and then define the open neighbourhood of $\tilde{x}$ in M by $N_\varepsilon=B_\delta(\tilde{x})$.

By replacing $w_\infty(x', e')$ by $f'(\pi x', e')$ and $w_\infty(x_\infty, e_\infty)$ by $f'(\pi\tilde{x}, \tilde{e})$ we get (5.3) from $f'(\pi\tilde{x}, \tilde{e}) \leqslant f'(\pi x', e')$ and (5.4) from (3.2). Here we used $2\Theta \leqslant \Omega$, from Lemma 5.1(1), and $\|e_\infty\|' \leqslant 2\|e_\infty\| = 2$. Hence

$$\big(x', e'\big) \in G_{p_\infty}(x_\infty, e_\infty, 0).$$

Then we have just showed in Section 5.6, as $x' \in B_\delta(x)$,

$$w_\infty\big(x', e'\big) < w_\infty(x_\infty, e_\infty) + \varepsilon$$

by (5.30), and so we conclude

$$f'\big(\pi x', e'\big) < f'(\pi\tilde{x}, \tilde{e}) + \varepsilon. \qquad \square$$

6. Set theory

In this section we shall prove Theorem 3.3. We recall its hypotheses: (Y, d) is a metric space and $(K_r)_{r\in R}$ is a collection of non-empty compact subsets of Y indexed by (R, γ), a non-empty metric space, such that

$$\mathscr{H}(K_r, K_s) \leqslant \gamma(r, s) \tag{6.1}$$

for every $r, s \in R$, where $\mathscr{H}$ denotes the Hausdorff distance.

Further O is a G_δ subset of Y containing every element of the family $(K_r)_{r\in R'}$ where $R' \subseteq R$ is γ-dense and $r_0 \in R'$. We further recall that $\rho, \varepsilon_0 > 0$ are such that for every $\varepsilon \in (0, \varepsilon_0)$ there exists $R(\varepsilon) \subseteq R$ such that (3.4) holds:

- for every $s \in R$ there exists $t \in R(\varepsilon)$ with $\gamma(t, s) < \varepsilon$,
- for every subset S of Y of diameter at most $\rho\varepsilon$ the set $\{r \in R(\varepsilon)\colon S \cap K_r \neq \emptyset\}$ is finite.

We may assume $\rho \in (0, 1)$ is fixed. Write $O = \bigcap_{n=1}^\infty O_n$ where (O_n) is a nested sequence of open subsets of Y, $O_{n+1} \subseteq O_n$ for each $n \geqslant 1$.

We first observe that due to the fact that O contains a γ-dense collection of compacts K_r, we may replace the families $R(\varepsilon)$ of compacts with families $R'(\varepsilon) \subseteq R'$, so that $K_r \subseteq O$ for every $r \in R'(\varepsilon)$, and properties listed in the following lemma are satisfied.

Lemma 6.1. *For every $\varepsilon \in (0, \varepsilon_0)$ we can find $R'(\varepsilon) \subseteq R$ such that $K_r \subseteq O$ for all $r \in R'(\varepsilon)$ and*

- *for every $r \in R$ there exists $t \in R'(\varepsilon)$ with $\gamma(t, r) < \varepsilon$,*
- *for every subset B of Y of diameter at most $\frac{4}{5}\rho\varepsilon$ the set*

$$F^B(\varepsilon) := \big\{t \in R'(\varepsilon) \text{ with } K_t \cap B \neq \emptyset\big\} \tag{6.2}$$

is finite.

Proof. For each $s \in R$ take $t_s \in R'$ with $\gamma(t_s, s) < \rho\varepsilon/10$, using the density of R'. Set

$$R'(\varepsilon) = \big\{t_s\colon s \in R(4\varepsilon/5)\big\}.$$

It is clear that $K_r \subseteq O$ for every $r \in R'(\varepsilon)$ and that for every $r \in R$ we can find $t \in R'(\varepsilon)$ with $\gamma(t,r) < 4\varepsilon/5 + \rho\varepsilon/10 < \varepsilon$.

Now if $t \in F^B(\varepsilon)$ then, writing $t = t_s$ with $s \in R(4\varepsilon/5)$, we see from $\gamma(t_s, s) < \rho\varepsilon/10$ and (6.1) that K_s intersects $\overline{B}_{\rho\varepsilon/10}(B)$; this set has diameter at most $\rho\varepsilon$ so the set $F^B(\varepsilon)$ is finite by (3.4). □

We now define the set

$$\mathbb{T} = \{(r, w, \alpha) \in R \times (0, \varepsilon_0) \times (0, \infty) \text{ such that } K_r \subseteq O \text{ and } w \leqslant \alpha\}. \tag{6.3}$$

Here $w \in (0, \varepsilon_0)$ denotes the width of the neighbourhood $T = \overline{B}_w(K_r)$ around K_r; as we mentioned earlier in Remark 3.4 our main example is the case when K_r is a wedge, then T is an angled tube around K_r. A slightly bigger neighbourhood $\overline{B}_\alpha(K_r)$, defined by the third parameter, is considered as a neighbourhood of the tube T just constructed, in which we plan to choose smaller tubes that approximate T. Therefore each element $(r, w, \alpha) \in \mathbb{T}$ presents a tube $\overline{B}_w(K_r)$ with some "safe" neighbourhood $\overline{B}_\alpha(K_r)$. For convenience, we will use these terms even in the general case when K_r are arbitrary compacts; we will also refer to elements $(r, w, \alpha) \in \mathbb{T}$ as tube triples.

For fixed $r_0 \in R'$ choose $0 < w_0 < \alpha_0 < \varepsilon_0$ so that

$$R_0 = \{(r_0, w_0, \alpha_0)\} \subseteq \mathbb{T}. \tag{6.4}$$

We shall now construct, for each $k \geqslant 1$, a set $R_k \subseteq \mathbb{T}$ inductively by adding, for every $(r, w, \alpha) \in R_l$ where $l < k$, a collection $R_{k,l} = R_{k,l}(r, w, \alpha)$ of tube triples $(t, v, \beta) \in \mathbb{T}$ with $\overline{B}_v(K_t) \subseteq O_k$ such that the collection $(K_t)_{(t,v,\beta)\in R_{k,l}}$ well approximates the collection of all compacts $(K_s)_{s\in R}$ when restricted to the "safe" neighbourhood $\overline{B}_\alpha(K_r)$. First let

$$r_{k,l} \in (0, \rho/10) \tag{6.5}$$

for each $0 \leqslant l < k$, where $\rho \in (0, 1)$ is the number fixed in the beginning of the present section. Later, in (6.14), we will impose additional restrictions on $(r_{k,l})$; however Lemmas 6.2, 6.3 and 6.5 we prove up to that point are valid for any $r_{k,l} \in (0, \rho/10)$.

Lemma 6.2. *If $0 \leqslant l < k$ and $(r, w, \alpha) \in \mathbb{T}$ then there is a set*

$$R_{k,l} = R_{k,l}(r, w, \alpha) \subseteq \mathbb{T}$$

such that

(1) *for every $s \in R$ with $K_s \subseteq \overline{B}_\alpha(K_r)$ there exists $(t, v, \beta) \in R_{k,l}$ such that*

$$\gamma(t, s) \leqslant \frac{10}{\rho} r_{k,l} w,$$

(2) *if $(t, v, \beta) \in R_{k,l}$ then $\beta = r_{k,l} w < \alpha/10$ and $v < \varepsilon_0/k$,*
(3) *if $(t, v, \beta) \in R_{k,l}$ then $\overline{B}_v(K_t) \subseteq O_k$ and $K_t \subseteq \overline{B}_{2\alpha}(K_r)$,*

(4) *if* $B \subseteq Y$ *has diameter at most* $8r_{k,l}w$ *then the set*

$$F = F^B_{k,l}(r, w, \alpha)$$

of all $(t, v, \beta) \in R_{k,l}$ *such that* K_t *intersects* B*, is finite,*

(5) *there exists* $v > 0$ *such that* $(r, v, r_{k,l}w) \in R_{k,l}$.

Proof. For each $t \in R$ with $K_t \subseteq O$ we can pick $v_t \in (0, \varepsilon_0/k)$ such that $v_t \leqslant r_{k,l}w$ and

$$\overline{B}_{v_t}(K_t) \subseteq O_k,$$

as $K_t \subseteq O \subseteq O_k$, K_t is compact and O_k is open. Now let

$$\varepsilon = \frac{10}{\rho} r_{k,l}w.$$

Note that $\varepsilon < w < \varepsilon_0$ from (6.5) and (6.3) and that for any $t \in R'(\varepsilon) \cup \{r\}$ we have $K_t \subseteq O$. So we may set

$$R_{k,l} = \big\{(t, v_t, r_{k,l}w)\colon\ t \in R'(\varepsilon) \cup \{r\} \text{ is such that } K_t \subseteq \overline{B}_{2\alpha}(K_r)\big\}.$$

Observe that $R_{k,l} \subseteq \mathbb{T}$, using the definition of v_t.

To see item (1) of the lemma, for $s \in R$ with $K_s \subseteq \overline{B}_\alpha(K_r)$ we pick $t \in R'(\varepsilon)$ with $\gamma(t, s) < \varepsilon$. Then $\gamma(t, s) \leqslant w \leqslant \alpha$ so that $K_t \subseteq \overline{B}_\alpha(K_s)$ using (6.1). It follows that $K_t \subseteq \overline{B}_{2\alpha}(K_r)$ so that $(t, v_t, r_{k,l}w) \in R_{k,l}$.

Items (2) and (3) are immediate.

For (4) note that if $(t, v_t, r_{k,l}w) \in F$ then as $t \in R'(\varepsilon) \cup \{r\}$ and the set B has diameter at most $\frac{4}{5}\rho\varepsilon$ we have

$$t \in F^B(\varepsilon) \cup \{r\};$$

see (6.2). As this set is finite then so is F.

Finally item (5) is immediate with $v = v_r$. □

Recall from (6.4) that we have defined $R_0 \subseteq \mathbb{T}$. Now for $k \geqslant 1$ define $R_k \subseteq \mathbb{T}$ by the recursion

$$R_k = \bigcup_{l=0}^{k-1} \bigcup_{(r,w,\alpha) \in R_l} R_{k,l}(r, w, \alpha). \tag{6.6}$$

Note that for any $(t, v, \beta) \in R_k$ we have

$$K_t \subseteq O \quad \text{and} \quad \overline{B}_v(K_t) \subseteq O_k \tag{6.7}$$

and

$$0 < v \leqslant \min\left(\beta, \frac{\varepsilon_0}{k}\right) \tag{6.8}$$

using (6.3) and Lemma 6.2, (2) and (3).

Next lemma proves that the collection of tube triples R_k has some local finiteness in its structure; we will use this property later to prove that if we consider unions of all tubes on each level and then intersect these unions up to a certain level then the resulting set is closed, see Definition 6.4 and Lemma 6.5.

Lemma 6.3. *If $y \in Y$ and $k \geqslant 0$ there exists $\delta_k = \delta_k(y) > 0$ such that the set*

$$F_k = F_k(y) := \{(r, w, \alpha) \in R_k \text{ such that } d(y, K_r) \leqslant \delta_k + 3\alpha\}$$

is finite.

Proof. Let $y \in Y$. For any $\delta_0 > 0$ we pick, the set $F_0 \subseteq R_0$ will be finite. Suppose now that $k \geqslant 1$ and we have picked $\delta_l > 0$ for every $0 \leqslant l < k$ such that F_l is finite.

Pick $\delta_k > 0$ such that for every $l < k$ we have $\delta_k < \delta_l$ and, for any $(r, w, \alpha) \in F_l$, $\delta_k < r_{k,l} w$. We shall show that F_k is finite.

Suppose that $(t, v, \beta) \in F_k$. We may write $(t, v, \beta) \in R_{k,l}(r, w, \alpha)$ where $l < k$ and $(r, w, \alpha) \in R_l$, using (6.6). Note that $K_t \subseteq \overline{B}_{2\alpha}(K_r)$ by Lemma 6.2(3). Hence

$$\begin{aligned} d(y, K_r) &\leqslant d(y, K_t) + 2\alpha \\ &\leqslant \delta_k + 3\beta + 2\alpha \\ &\leqslant \delta_l + 3\alpha \end{aligned}$$

using $\delta_k < \delta_l$ and $\beta = r_{k,l} w < \alpha/10$ from Lemma 6.2(2). Hence $(r, w, \alpha) \in F_l$ and so $\delta_k < r_{k,l} w$. We get $d(y, K_t) \leqslant \delta_k + 3\beta < 4 r_{k,l} w$ so that

$$K_t \cap \overline{B}_{4r_{k,l}w}(y) \neq \emptyset$$

and $(t, v, \beta) \in F_{k,l}^{\overline{B}_{4r_{k,l}w}(y)}(r, w, \alpha)$; see Lemma 6.2(4).

We conclude that

$$F_k \subseteq \bigcup_{l=0}^{k-1} \bigcup_{(r,w,\alpha) \in F_l} F_{k,l}^{\overline{B}_{4r_{k,l}w}(y)}(r, w, \alpha),$$

which is finite by Lemma 6.2(4). □

Definition 6.4. If $k \geqslant 1$, $\lambda \in [0, 1]$ and $w > 0$ we define $M_k(\lambda, w)$ to be the set of $y \in Y$ such that there exist integers $n \geqslant 1$, $0 = l_0 < l_1 < \cdots < l_n = k$ and tube triples $(r_m, w_m, \alpha_m) \in R_{l_m}$ for $0 \leqslant m \leqslant n$ with

(1) $(r_m, w_m, \alpha_m) \in R_{l_m, l_{m-1}}(r_{m-1}, w_{m-1}, \alpha_{m-1})$ for $1 \leqslant m \leqslant n$,
(2) $d(y, K_{r_m}) \leqslant \lambda \alpha_m$ for $0 \leqslant m \leqslant n$,
(3) $d(y, K_{r_n}) \leqslant \lambda w_n$,
(4) $w_n = w$.

We then let

$$M_k(\lambda) = \bigcup_{w>0} M_k(\lambda, w).$$

Remark. Note that Definition 6.4(3) implies that $M_k(\lambda)$ is a subset of the union $\bigcup \overline{B}_{\lambda w}(K_r)$, where the union is taken over the collection of all tube triples (r, w, α) in R_k. Since each of those tubes is inside O_k by (6.7), we conclude $M_k(\lambda) \subseteq O_k$. Further from (6.4), (6.6), Lemma 6.2(5), Definition 6.4(2) and (6.7),

$$K_{r_0} \subseteq M_k(\lambda) \tag{6.9}$$

for all $k \geqslant 1$ and $\lambda \in [0, 1]$. Finally if $M_k(\lambda, w) \neq \emptyset$ then by Lemma 6.2(2),

$$w < \varepsilon_0/k. \tag{6.10}$$

Lemma 6.5. *For any $k \geqslant 1$ and $\lambda \in [0, 1]$, the set $M_k(\lambda)$ is a closed subset of (Y, d).*

Proof. Suppose that $y^{(i)} \in M_k(\lambda)$ with $y^{(i)} \to y \in Y$. It suffices to show that $y \in M_k(\lambda)$.

For each $i \geqslant 1$ we have $y^{(i)} \in M_k(\lambda)$, therefore we can find $n^{(i)} \geqslant 1$, $0 = l_0^{(i)} < \cdots < l_{n^{(i)}}^{(i)} = k$ and $(r_m^{(i)}, w_m^{(i)}, \alpha_m^{(i)}) \in R_{l_m^{(i)}}$ for $0 \leqslant m \leqslant n^{(i)}$ such that the conditions in Definition 6.4(1)–(3) are satisfied:

$$\big(r_m^{(i)}, w_m^{(i)}, \alpha_m^{(i)}\big) \in R_{l_m^{(i)}, l_{m-1}^{(i)}}\big(r_{m-1}^{(i)}, w_{m-1}^{(i)}, \alpha_{m-1}^{(i)}\big) \quad \text{for } 1 \leqslant m \leqslant n^{(i)}, \tag{6.11}$$

$$d\big(y^{(i)}, K_{r_m^{(i)}}\big) \leqslant \lambda \alpha_m^{(i)} \quad \text{for } 0 \leqslant m \leqslant n^{(i)}, \tag{6.12}$$

$$d\big(y^{(i)}, K_{r_{n^{(i)}}^{(i)}}\big) \leqslant \lambda w_{n^{(i)}}^{(i)}. \tag{6.13}$$

As $1 \leqslant n^{(i)} \leqslant k$ we may assume, passing to a subsequence if necessary, that $n^{(i)} = n$ is constant. But then as $0 \leqslant l_m^{(i)} \leqslant k$ we may assume, passing to another subsequence, that $l_m^{(i)} = l_m$ is constant for each $0 \leqslant m \leqslant n$ with $0 = l_0 < l_1 < \cdots < l_n = k$.

Fixing m then as $d(y, y^{(i)}) \to 0$, $\lambda \leqslant 1$ and

$$d(y, K_{r_m^{(i)}}) \leqslant d\big(y, y^{(i)}\big) + \lambda \alpha_m^{(i)},$$

from (6.12), we have $(r_m^{(i)}, w_m^{(i)}, \alpha_m^{(i)}) \in F_{l_m}(y)$ for i sufficiently high; see Lemma 6.3. As this set is finite we can assume, passing to another subsequence, that

$$\big(r_m^{(i)}, w_m^{(i)}, \alpha_m^{(i)}\big) = (r_m, w_m, \alpha_m)$$

is constant for each $0 \leqslant m \leqslant n$, with $(r_m, w_m, \alpha_m) \in R_{l_m}$. Further from (6.11)–(6.13) we have

- $(r_m, w_m, \alpha_m) \in R_{l_m, l_{m-1}}(r_{m-1}, w_{m-1}, \alpha_{m-1})$ for $1 \leqslant m \leqslant n$,
- $d(y^{(i)}, K_{r_m}) \leqslant \lambda \alpha_m$ for $0 \leqslant m \leqslant n$,
- $d(y^{(i)}, K_{r_n}) \leqslant \lambda w_n$.

Taking the $i \to \infty$ limit and using $y^{(i)} \to y$ we obtain

- $d(y, K_{r_m}) \leqslant \lambda\alpha_m$ for $0 \leqslant m \leqslant n$,
- $d(y, K_{r_n}) \leqslant \lambda w_n$,

so that $y \in M_k(\lambda)$. □

Up to this point we have let $r_{k,l} \in (0, \rho/10)$ be arbitrary; see (6.5). We now further stipulate that if $0 \leqslant l < l' \leqslant k$ then we have

$$r_{k+1,k} \leqslant \frac{1}{k} \quad \text{and} \quad r_{k+1,l} \leqslant \frac{1}{k} r_{l',l}. \tag{6.14}$$

We now come to the crucial lemma. It proves that if we consider a point y is in $M_k(\lambda, w)$ and $\lambda' > \lambda$, then the whole $(\lambda' - \lambda)w$-neighbourhood of y is inside $M_k(\lambda', w)$. If, however, we want to find compacts K_t close to y of bigger size, $\delta > (\lambda' - \lambda)w/2$, we can accomplish this as long as we agree to consider tube sets constructed on subsequent levels.

Lemma 6.6. *Suppose $k \geqslant 1$, $0 \leqslant \lambda < \lambda + \psi \leqslant 1$, $w > 0$, $\varepsilon \in (0, 1)$ and $y \in M_k(\lambda, w)$. Then*

(1) *$\overline{B}_{\psi w}(y) \subseteq M_k(\lambda + \psi, w)$,*
(2) *if $2\delta \in (\psi w, \psi\alpha_0)$ and $20/(\rho\psi k) < \varepsilon < 1$ then for each $s \in R$ with $K_s \subseteq \overline{B}_\delta(y)$ there exists $t \in R$ with $\gamma(t, s) < \varepsilon\delta$ and $K_t \subseteq M_{k+j}(\lambda + \psi)$ for all $j \geqslant 1$.*

Proof. From Definition 6.4 we can find integers $n \geqslant 1$,

$$0 = l_0 < l_1 < \cdots < l_n = k$$

and tube triples $(r_m, w_m, \alpha_m) \in R_{l_m}$ for $0 \leqslant m \leqslant n$ with

$$(r_m, w_m, \alpha_m) \in R_{l_m, l_{m-1}}(r_{m-1}, w_{m-1}, \alpha_{m-1}) \quad \text{for } 1 \leqslant m \leqslant n, \tag{6.15}$$

$$d(y, K_{r_m}) \leqslant \lambda\alpha_m \quad \text{for } 0 \leqslant m \leqslant n, \tag{6.16}$$

$$d(y, K_{r_n}) \leqslant \lambda w_n, \tag{6.17}$$

$$w_n = w. \tag{6.18}$$

Note that

$$\alpha_m = r_{l_m, l_{m-1}} w_{m-1} < \alpha_{m-1} \tag{6.19}$$

for each $1 \leqslant m \leqslant n$ by Lemma 6.2(2).

To establish (1) of the present lemma, suppose $d(y', y) \leqslant \psi w$; then from (6.16) and (6.17),

$$d\big(y', K_{r_m}\big) \leqslant \lambda\alpha_m + \psi w \quad \text{for } 0 \leqslant m \leqslant n,$$
$$d\big(y', K_{r_n}\big) \leqslant \lambda w_n + \psi w.$$

Using (6.18) and (6.19) we have $w = w_n \leqslant \alpha_n \leqslant \alpha_m$ so that

$$d\bigl(y', K_{r_m}\bigr) \leqslant (\lambda + \psi)\alpha_m \quad \text{for } 0 \leqslant m \leqslant n,$$
$$d\bigl(y', K_{r_n}\bigr) \leqslant (\lambda + \psi)w_n;$$

combining these with (6.15) and (6.18) we get $y' \in M_k(\lambda + \psi, w)$, as required.

We now turn to (2). We claim that we can find m with $0 \leqslant m \leqslant n$ and

$$(t, w, \alpha) \in R_{k+1,l_m}(r_m, w_m, \alpha_m) \tag{6.20}$$

where $2\delta \leqslant \psi\alpha_m$ and $\gamma(t, s) < \varepsilon\delta$.

To see this suffices, note first that as $\mathscr{H}(K_t, K_s) \leqslant \gamma(t, s) < \delta$, using $\varepsilon \leqslant 1$, we have

$$K_t \subseteq \overline{B}_\delta(K_s) \subseteq \overline{B}_{2\delta}(y) \subseteq \overline{B}_{\psi\alpha_m}(y), \tag{6.21}$$

where we have also used $2\delta \leqslant \psi\alpha_m$ from the claim to be proved and $K_s \subseteq \overline{B}_\delta(y)$ from the hypothesis of (2).

Now let $l'_j = l_j$ and $(r'_j, w'_j, \alpha'_j) = (r_j, w_j, \alpha_j)$ for $j \leqslant m$ and $l'_{m+j} = k + j$ for $j \geqslant 1$ and, using (6.20) for $(r'_{m+1}, w'_{m+1}, \alpha'_{m+1})$ and Lemma 6.2(5), pick inductively

$$\bigl(r'_{m+j}, w'_{m+j}, \alpha'_{m+j}\bigr) \in R_{l'_{m+j}, l'_{m+j-1}}\bigl(r'_{m+j-1}, w'_{m+j-1}, \alpha'_{m+j-1}\bigr)$$

for each $j \geqslant 1$, with $r'_{m+j} = t$. Then for any $y' \in K_t$, as

$$d\bigl(y', K_{r_j}\bigr) \leqslant d(y, K_{r_j}) + \psi\alpha_m \leqslant (\lambda + \psi)\alpha'_j$$

for $j \leqslant m$, using (6.16) and (6.21), while $d(y', K_{r'_{m+j}}) = 0$ for $j \geqslant 1$ from $y' \in K_t = K_{r'_{m+j}}$, we have $y' \in M_{k+j}(\lambda + \psi)$ for $j \geqslant 1$ as required.

We now establish the claim. Suppose first that $2\delta \leqslant \psi\alpha_n$. Then as

$$K_s \subseteq \overline{B}_\delta(y) \subseteq \overline{B}_{\lambda\alpha_n+\delta}(K_{r_n}) \subseteq \overline{B}_{\alpha_n}(K_{r_n}),$$

using (6.16), we may pick, by Lemma 6.2(1), $(t, w, \alpha) \in R_{k+1,k}(r_n, w_n, \alpha_n)$ with

$$\gamma(t, s) \leqslant \frac{10}{\rho} r_{k+1,k} w_n \leqslant \frac{10}{\rho} \frac{1}{k} \frac{2\delta}{\psi} < \varepsilon\delta$$

using (6.14) and $2\delta \in (\psi w_n, \psi)$. Thus we can satisfy the claim with $m = n$.

Suppose instead that $\psi\alpha_n < 2\delta$. As $2\delta \leqslant \psi\alpha_0$ we can find m with

$$\psi\alpha_{m+1} < 2\delta \leqslant \psi\alpha_m \tag{6.22}$$

where $0 \leqslant m \leqslant n - 1$. Then as

$$K_s \subseteq \overline{B}_\delta(y) \subseteq \overline{B}_{\lambda\alpha_m+\delta}(K_{r_m}) \subseteq \overline{B}_{\alpha_m}(K_{r_m}),$$

we may pick, by Lemma 6.2(1), $(t, w, \alpha) \in R_{k+1,l_m}(r_m, w_m, \alpha_m)$ with

$$\gamma(t,s) \leqslant \frac{10}{\rho} r_{k+1,l_m} w_m \leqslant \frac{10}{\rho}\frac{1}{k} r_{l_{m+1},l_m} w_m = \frac{10}{\rho}\frac{1}{k}\alpha_{m+1} < \frac{10}{\rho}\frac{1}{k}\frac{2\delta}{\psi} < \varepsilon\delta$$

using (6.14) with $l_m < l_{m+1} \leqslant k$, (6.19) and (6.22). Thus the claim is satisfied. □

6.7. Proof of Theorem 3.3

We are now ready to prove Theorem 3.3.

Assume r_0 used in (6.4) is the one given by hypothesis of Theorem 3.3.

Given $\lambda \in [0,1]$ we set

$$T_\lambda = \bigcap_{k=1}^{\infty} J_k(\lambda), \quad \text{where } J_k(\lambda) = \bigcup_{k \leqslant n \leqslant (1+\lambda)k} M_n(\lambda).$$

Note that as (6.9) implies $K_{r_0} \subseteq M_n(\lambda) \subseteq O_n \subseteq O_k$ for $n \geqslant k$, we have $K_{r_0} \subseteq J_k(\lambda) \subseteq O_k$ for every $k \geqslant 1$ and hence $K_{r_0} \subseteq T_\lambda \subseteq O$ for every $\lambda \in [0,1]$. Similarly as $M_k(\lambda)$ is closed by Lemma 6.5, the set $J_k(\lambda)$ is also closed for every $k \geqslant 1$, and hence T_λ is closed for every $\lambda \in [0,1]$. We further note that if $0 \leqslant \lambda_1 \leqslant \lambda_2 \leqslant 1$ then as $M_k(\lambda_1) \subseteq M_k(\lambda_2)$ from Definition 6.4, we have $J_k(\lambda_1) \subseteq J_k(\lambda_2)$ and hence we have $T_{\lambda_1} \subseteq T_{\lambda_2}$.

Assume $\eta \in (0,1)$, $0 \leqslant \lambda' < \lambda \leqslant 1$ and $y \in T_{\lambda'}$. By the definitions of $T_{\lambda'}$ and $M_k(\lambda')$ and the last part of Definition 6.4, there exists, for each $k \geqslant 1$, an index n_k with $k \leqslant n_k \leqslant (1+\lambda')k$ and $w_k > 0$ such that $y \in M_{n_k}(\lambda', w_k)$. Let $\psi = \lambda - \lambda' > 0$.

Pick $\delta_1 > 0$ with $2\delta_1 < \psi w_k$ for every $k \leqslant 20/(\rho\psi\eta)$, where $\rho \in (0,1)$ is the number fixed in the beginning of the present section. Now suppose that $\delta \in (0,\delta_1)$. We need to show that if $K_s \subseteq \overline{B}_\delta(y)$ for some $s \in R$ then there exists $t \in R$ such that $K_t \subseteq T_\lambda$ and $\gamma(t,s) < \eta\delta$. Let $k_0 \geqslant 1$ be the minimal index k such that $2\delta > \psi w_k$. Such k_0 exists as (6.10) implies $w_k \to 0$. Note that $k_0 > 20/(\rho\psi\eta)$. In particular $\psi k_0 > 1$ and so $k_0 < n_{k_0} + 1 < (1+\lambda')k_0 + \psi k_0 = (1+\lambda)k_0$.

By Lemma 6.6(2) there exists $t \in R$ such that $\gamma(s,t) < \eta\delta$ and $K_t \subseteq M_j(\lambda)$ for every $j \geqslant n_{k_0} + 1$, so that $K_t \subseteq J_k(\lambda)$ for all $k \geqslant k_0$. Note that $\gamma(t,s) < \eta\delta < \delta$ implies $K_t \subseteq \overline{B}_{2\delta}(y) \subseteq \overline{B}_{\psi w_k}(y)$ for every $k < k_0$. By Lemma 6.6(1) we conclude $K_t \subseteq M_{n_k}(\lambda, w_k)$ for every $k < k_0$.

Hence $K_t \subseteq T_\lambda$ as required. □

References

[1] G. Alberti, M. Csörnyei, D. Preiss, Differentiability of Lipschitz functions, structure of null sets, and other problems, in: Proceedings of the International Congress of Mathematicians, Hyderabad, India, 2010.

[2] E. Asplund, Fréchet differentiability of convex functions, Acta Math. 121 (1968) 31–47.

[3] Y. Benyamini, J. Lindenstrauss, Geometric Nonlinear Functional Analysis, vol. 1, Amer. Math. Soc. Colloq. Publ., vol. 48, Amer. Math. Soc., Providence, 2000.

[4] T. de Pauw, P. Huovinen, Points of ε-differentiability of Lipschitz functions from $\mathbb{R}^n$ to $\mathbb{R}^{n-1}$, Bull. Lond. Math. Soc. 34 (5) (2002) 539–550.

[5] M. Doré, O. Maleva, Fréchet differentiability of planar-valued Lipschitz functions on Hilbert spaces, in preparation.

[6] M. Doré, O. Maleva, A compact null set containing a differentiability point of every Lipschitz function, Math. Ann., doi:10.1007/s00208-010-0613-4, in press.

[7] M. Doré, O. Maleva, A compact universal differentiability set of Hausdorff dimension one in Euclidean space, Israel J. Math., in press.

[8] T. Fowler, D. Preiss, A simple proof of Zahorski's description of non-differentiability sets of Lipschitz functions, Real Anal. Exchange 34 (1) (2008/2009) 1–12.

[9] W.B. Johnson, J. Lindenstrauss, D. Preiss, G. Schechtman, Almost Fréchet differentiability of Lipschitz mappings between infinite-dimensional Banach spaces, Proc. Lond. Math. Soc. 84 (3) (2002) 711–746.

[10] J. Lindenstrauss, D. Preiss, On Fréchet differentiability of Lipschitz maps between Banach spaces, Ann. of Math. 157 (2003) 257–288.
[11] J. Lindenstrauss, D. Preiss, J. Tišer, Fréchet differentiability of Lipschitz functions and porous sets in Banach spaces, Monograph in preparation.
[12] D. Preiss, Differentiability of Lipschitz functions on Banach spaces, J. Funct. Anal. 91 (1990) 312–345.
[13] Z. Zahorski, Sur l'ensemble des points de non-derivabilite d'une fonction continue, Bull. Soc. Math. France 74 (1946) 147–178.
[14] L. Zajíček, Sets of σ-porosity and sets of σ-porosity (q), Časopis pro Pěstování Matematiky 101 (1976) 350–359.
[15] L. Zajíček, Differentiability of the distance function and points of multivaluedness of the metric projection in Banach space, Czechoslovak Math. J. 33 (108) (1983) 292–308.
[16] L. Zajíček, Small non-σ-porous sets in topologically complete metric spaces, Colloq. Math. 77 (2) (1998) 293–304.
[17] M. Zelený, J. Pelant, The structure of the σ-ideal of σ-porous sets, Comment. Math. Univ. Carolin. 45 (1) (2004) 37–72.