Directional upper derivatives and the chain rule formula for locally Lipschitz functions on Banach spaces

Maleva, Olga; Preiss, David

DOI:
10.1090/tran/6480

DIRECTIONAL UPPER DERIVATIVES AND THE CHAIN RULE FORMULA FOR LOCALLY LIPSCHITZ FUNCTIONS ON BANACH SPACES

OLGA MALEVA AND DAVID PREISS

Abstract. Motivated by an attempt to find a general chain rule formula for differentiating the composition $f \circ g$ of Lipschitz functions f and g that would be as close as possible to the standard formula $(f \circ g)'(x) = f'(g(x)) \circ g'(x)$, we show that this formula holds without any artificial assumptions provided derivatives are replaced by complete derivative assignments. The idea behind these assignments is that the derivative of f at y is understood as defined only in the direction of a suitable "tangent space" $U(f, y)$ (and so it exists at every point), but these tangent spaces are chosen in such a way that for any g they contain the range of $g'(x)$ for almost every x. Showing the existence of such assignments leads us to detailed study of derived sets and the ways in which they describe pointwise behavior of Lipschitz functions.

Introduction

The main motivation for the research presented here is to provide a general chain rule formula for differentiating the composition $f \circ g$ of Lipschitz functions f and g between finite-dimensional or infinite-dimensional spaces that would be as close as possible to the basic formula

$$(f \circ g)'(x) = f'(g(x)) \circ g'(x), \tag{0.1}$$

where $h'(x)$ denotes the derivative of a map $h : X \to Y$, which is considered as an element of the space $L(X, Y)$ of continuous linear maps of X to Y. Recall that, as observed by a number of authors, even for Lipschitz functions this formula may become invalid at every x. A typical example is given by $g : \mathbb{R} \to \mathbb{R}^2$ and $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$g(x) = (x, 0) \quad \text{and} \quad f(x, y) = |y|;$$

the chain rule fails because f is not differentiable at any $g(x)$.

For mappings between finite-dimensional spaces this question was previously studied by Ambrosio and Dal Maso [2]. In the Lipschitz setting (we will discuss their more general setting allowing g to be of bounded variation later) they prove that, given Lipschitz $g : \mathbb{R}^n \to \mathbb{R}^m$ and letting for a.e. $x \in \mathbb{R}^n$,

$$T_x^g = \Big\{ y \in \mathbb{R}^m : y = g(x) + g'(x; z) \quad \text{for some } z \in \mathbb{R}^n \Big\}, \tag{0.2}$$

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n.2011-ADG-20110209.

then for every Lipschitz $f : \mathbb{R}^m \to \mathbb{R}^k$ the derivative $(f|_{T_x^g})'(g(x))$ exists for a.e. $x \in \mathbb{R}^n$ and

$$(f \circ g)'(x) = (f|_{T_x^g})'(g(x)) \circ g'(x) \quad \text{for a.e. } x \in \mathbb{R}^n. \tag{0.3}$$

Comparing (0.3) with (0.1), we notice a significant difference: while in (0.1) the derivative of f on the right hand side depends only on f and the point at which it is taken, in (0.3) it depends also on the function g. A more natural analogue of (0.1) would be

$$(f \circ g)'(x) = (f|_{T_{g(x)}})'(g(x)) \circ g'(x) \quad \text{for a.e. } x \in \mathbb{R}^n, \tag{0.4}$$

where the "tangent space" T_y depends only on the point at which it is taken and on the function f, and is such that $(f|_{T_y})'(y)$ exists for every y. We call a (Borel) family of spaces $(T_y)_{y\in\mathbb{R}^m}$ for which (0.4) holds a *tangent space assignment* for f and the mapping $f^{\blacktriangledown}(y) := (f|_{T_y})'(y)$ a *derivative assignment* for f. (Thus there is essentially no difference between the concepts of tangent space assignments and derivative assignments for f, but the former is better suited to questions of existence and the latter to applicability to the chain rule formula.) Of course, the validity of (0.4) requires that for a.a. x, the image of $g'(x)$ be contained in $T_{g(x)}$; in other words that for every g, the tangent space T_x^g defined in (0.2) be contained in $T_{g(x)}$ for a.a. $x \in \mathbb{R}^n$. We call derivative assignments for f that satisfy this condition for every Lipschitz g *complete* (Definition 4.1). It does not follow immediately that complete derivative assignments for f exist. Our main results, Theorems 5.2 and 5.7, show that they indeed exist, even in the infinite dimensional situation of spaces having the Radon-Nikodym property, in which case "almost all" is understood in the sense of the σ-ideal $\mathcal{L}$ defined in Section 1.

There is no canonical way to choose a tangent space assignment for a given f. We present several ways of defining them. The simplest to describe is defining T_y as a maximal linear subspace in the direction of which f is differentiable. However, this assignment can be highly nonmeasurable, which seriously limits its applicability. In Example 5.4 we point out that under the continuum hypothesis there are even less constructive choices. We therefore take great care to establish measurability of the tangent space assignments $(T_y)_{y\in\mathbb{R}^m}$ and the corresponding "generalized derivatives" $f^{\blacktriangledown}(y)$. A more constructive example is obtained, for instance, by choosing T_y as the space of all vectors e such that the directional derivative $f'(y;\, e)$ exists and whenever $f'(y;\, e')$ exists, one has

$$f'(y;\, e + e') = f'(y;\, e) + f'(y;\, e').$$

In Remark 5.8 we show that this assignment is measurable with respect to the σ-algebra generated by Suslin sets (and hence universally measurable). Although this may already be useful in applications, the natural measurability requirement in this context is Borel measurability. In Proposition 5.1 we therefore define several tangent space assignments by replacing the additivity of the directional derivative with inclusions between so-called derived sets, which are the limit sets of decreasing sequences of sets of divided differences with the upper bound for the denominator going to zero, and in Theorem 5.7 we show that this approach indeed leads to Borel measurable tangent space assignments.

The above program leads us naturally to a detailed study of derived sets of Lipschitz mappings. In particular, in the infinite dimensional situation lack of compactness means that the derived sets understood as limit sets may not describe

the behavior of the function; for example, they may be empty. We therefore base our considerations on so-called approximating derived sets, explain how one should understand the inclusion between them, and give rather precise information on smallness (in the sense of porosity) of sets in which various notions of derived sets give different results.

We also address the following natural question. Even if one has a working chain rule formula (0.4) for a composition of two functions with tangent space assignment $(T_y)_{y\in\mathbb{R}^n}$ depending on the outer function f, this still does not settle the problem for the composition of three or more functions. Indeed, as presented above, the chain rule formula gives the derivative of $f\circ g:\mathbb{R}^n\to\mathbb{R}^k$ almost everywhere, so the whole image of a mapping $w:\mathbb{R}^l\to\mathbb{R}^n$ may lie in the set of points where the chain rule for $f\circ g$ does not hold. We answer this problem by stating and proving in Theorem 4.2 the chain rule formula for obtaining a complete derivative assignment for $f\circ g$ by composing complete derivative assignments for f and g. This immediately implies a chain rule formula for any finite composition of mappings, also in the case of mappings between separable Banach spaces with the Radon-Nikodym property.

A natural question which was pointed out to us by J. Borwein is whether a chain rule for mappings $f:Y\mapsto\mathbb{R}$ holds with a naturally defined subdifferential, in particular, with the Michel-Penot subdifferential. We show that with the Michel-Penot subdifferential is it false, but that it holds with the upper Dini subdifferential $\partial^D f(y)$ (defined in 6.7). The chain rule then takes the form

$$\partial^D(f\circ g)=(\partial^D f)\circ g' \quad \text{a.e.}$$

and corresponds to defining the tangent space assignment T_y as the set of directions at which all elements of $\partial^D f(y)$ attain the same value. This result generalizes earlier work by Craven, Ralph and Glover in [11] where the chain rule via the upper Dini subdifferential is proved for the composition of a Lipschitz function and a Gâteaux differentiable function.

Finally, we briefly explain that our results show that the g dependent "tangent spaces" T^g_x may be replaced by the (f dependent) derivative assignments also in the chain rule of Ambrosio and Dal Maso [2] in which the inner function g is assumed to be only of bounded variation. (This rule generalized a number of previous results, for whose discussion we refer to [2].) To describe it, we assume without loss of generality that the given function $g:\mathbb{R}^n\to\mathbb{R}^m$ of bounded variation has been already modified so that it is approximate continuous at every point of the set E of points at which it has an approximate limit and denote by $\nabla g/|\nabla g|$ the Radon-Nikodym derivative of the $L(\mathbb{R}^n,\mathbb{R}^n)$-valued measure ∇g, the distributional derivative of g, with respect to its variation $|\nabla g|$. Then, analogously to (0.2),

$$T^g_x=\Big\{y\in\mathbb{R}^m : y=g(x)+\frac{\nabla g}{|\nabla g|}(x)\cdot z \quad \text{for some } z\in\mathbb{R}^n\Big\}, \tag{0.5}$$

is well-defined $|\nabla g|$ almost everywhere, and [2] shows that for every Lipschitz $f:\mathbb{R}^m\to\mathbb{R}^k$ with $f(0)=0$ the derivative $\nabla(f|_{T^g_x})(g(x))$ exists for $|\nabla g|$ a.e. $x\in\mathbb{R}^n$ and for every Borel $A\subset E$,

$$\nabla(f\circ g)(A)=\int_A \nabla(f|_{T^g_x})(g(x))\circ\frac{\nabla g}{|\nabla g|}(x)\,d|\nabla g|(x). \tag{0.6}$$

We show in Theorem 7.4 that for any complete derivative assignment $(T_y)_{y\in\mathbb{R}^m}$ and any function $g:\mathbb{R}^n\to\mathbb{R}^m$ of bounded variation the inclusion $T^g_x\subset T_{g(x)}$

holds for $|\nabla g|$ a.e. $x \in \mathbb{R}^n$. Hence (0.6) holds also with $\nabla(f|_{T_x^g})(g(x))$ replaced by $(f|_{T_{g(x)}})'(g(x))$.

Acknowledgment. We wish to thank Jonathan Borwein and Warren Moors who, probably motivated by their chain rule in the special case of essentially smooth Lipschitz functions [7], brought some of the problems treated here to our attention, and to Luděk Zajíček for a number of useful observations and comments.

1. Preliminaries

To make our results independent of future developments in the study of differentiability of Lipschitz functions $f : X \to Y$ between separable Banach spaces X and Y, we will understand the notion "almost everywhere" in the sense of the σ-ideal generated by the sets of non-differentiability of such functions. The main notions of differentiability we will be interested in are the *one-sided derivative of f at a point $x \in X$ in the direction of $e \in X$,* which is defined by

$$f'_+(x;\, e) = \lim_{r \searrow 0} \frac{f(x+re) - f(x)}{r};$$

the *bilateral derivative of f at x in the direction of e,* defined by

$$f'(x;\, e) = \lim_{r \to 0} \frac{f(x+re) - f(x)}{r};$$

in both cases, of course, provided the limit exists; and the *Gâteaux derivative of f at x,* which is, by definition, the mapping $f'(x) : X \to Y$, $f'(x)(e) = f'(x;\, e)$, provided all these derivatives exist and the map $e \to f'(x)(e)$ belongs to $L(X, Y)$, that is, it is continuous and linear.

Notice that $f'(x;\, e)$ exists if and only if both $f'_+(x;\, e)$ an $f'_+(x;\, -e)$ exist and $f'_+(x;\, -e) = -f'_+(x;\, e)$; and if this is the case, $f'(x;\, e) = f'_+(x;\, e)$. In particular, in the definition of the Gâteaux derivative $f'(x;\, e)$ may be replaced by $f'_+(x;\, e)$.

The σ-ideal $\mathcal{L}$. Let X be a separable Banach space, We denote by $\mathcal{L}$ (or $\mathcal{L}(X)$) the σ-ideal generated by sets of points of Gâteaux non-differentiability of Lipschitz mappings of X to Banach spaces with the Radon-Nikodym property. Hence $E \subset X$ belongs to $\mathcal{L}$ (or, as we say, is *Lipschitz null*, or, shortly, $\mathcal{L}$ *null*) if there are Lipschitz mappings $f_i : X \to Y_i$, where Y_i have the Radon-Nikodym property, such that

$$E \subset \bigcup_{i=1}^{\infty} \{x \in X : f_i \text{ is not Gâteaux differentiable at } x\}.$$

As usual, we will use expressions such as "a property $P(x)$ holds for $\mathcal{L}$ almost every $x \in X$" instead of $\{x \in X : \text{not } P(x)\} \in \mathcal{L}$.

It is easy to see that $E \subset X$ is $\mathcal{L}$ null if and only if there is a (single) space Y with the Radon-Nikodym property and a (single) map $f : X \to Y$, such that

$$E \subset \{x \in X : f \text{ is not Gâteaux differentiable at } x\}.$$

More substantially, [37] shows that $\mathcal{L}$ is generated by the sets of directional non-differentiability of Lipschitz maps into spaces with the Radon-Nikodym property. So $E \subset X$ is $\mathcal{L}$ null if and only if there are $e_i \in X$ and Lipschitz mappings $f_i : X \to Y_i$, where Y_i have the Radon-Nikodym property, such that

$$E \subset \bigcup_{i=1}^{\infty} \{x \in X : (f_i)'_+(x;\, e_i) \text{ does not exist}\}.$$

However, even on this fairly basic level most questions are open. One of many interesting open problems about the σ-ideal $\mathcal{L}$ is whether it is generated by the sets of points of (Gâteaux or directional) nondifferentiability of real-valued functions. Interestingly, unlike in the one dimensional case, it is not true that every set from $\mathcal{L}(\mathbb{R}^2)$ is contained in the set of non-differentiability of a single real-valued Lipschitz function (see [35, Theorem 6.4] or, for more detailed results on this phenomenon, [13, 14, 15]), but by [1] it is contained in the union of the sets of non-differentiability of two such functions.

Very recently, significant progress had been made on the description of the σ-ideal $\mathcal{L}$ in the finite dimensional situation. In one dimensional spaces it coincides with the σ-ideal of sets of Lebesgue measure zero. (The fact that the former is included in the latter follows from Lebesgue's Theorem on differentiability of monotone functions, the converse can be found, e.g., in [8] and a full description of non-differentiability sets on $\mathbb{R}$ is in [39] or [19].) The key differentiability result, Rademacher's Theorem ([38] or [18, Theorem 3.1.6]), shows that in any finite dimensional space $\mathcal{L}$ is contained in the σ-ideal of sets of Lebesgue measure zero. In the converse direction, a geometric description of sets from $\mathcal{L}$ in finite dimensional spaces and the coincidence of $\mathcal{L}$ with Lebesgue null sets in the two dimensional case was obtained in [1], and very recently announced (but not yet published) results of Csörnyei and Jones [12] use this description to show that $\mathcal{L}$ coincides with the σ-ideal of Lebesgue null sets in every finite dimensional space.

In infinite dimensional spaces, several estimates of the size of the sets from $\mathcal{L}$ are known. Aronszajn's Theorem [3] (Theorem 8.3 of this note) gives a rather strong result, and in this sense contains all earlier results estimating the size of $\mathcal{L}$ (including Christensen's [9] very appealing notion of "sets of Haar measure zero"). However, [37] provides some σ-ideals that are strictly contained in Aronszajn null sets and still contain $\mathcal{L}$.

Notice that results on, say, differentiability of Lipschitz functions may be easily extended to locally Lipschitz functions. For example, if G is an open subset of a separable Banach space X, Y has the Radon-Nikodym property and $f : G \mapsto Y$ is locally Lipschitz, then f is differentiable $\mathcal{L}$ almost everywhere on G. This may be seen as follows: We write $G = \bigcup_{i=1}^{\infty} B(x_i, r_i)$, where f is Lipschitz on $\overline{B(x_i, s_i)} \subset G$, define $f_i(x) = f(x)$ if $x \in B(x_i, s_i)$ and $f_i(x) = f(x_i + s_i(x - x_i)/\|x - x_i\|)$ if $x \notin B(x_i, s_i)$, and observe that the set of points of non-differentiability of f is covered by the union of the sets of points of non-differentiability of f_i. The same argument applies in a number of situations we investigate, and we will therefore often state our results for locally Lipschitz functions on open sets, but prove them for everywhere defined Lipschitz functions only.

Porosity. We now introduce the concept of (directional) porosity, which will often become the main tool in our investigations. We say that $P \subset X$ *is porous in direction of* $e \in X$ if there is a $c > 0$ such that for every $x \in P$ and $\varepsilon > 0$ there is $0 < t < \varepsilon$ for which $B(x + te, ct) \cap P = \emptyset$. Here $B(z, r)$ denotes the open ball of radius r centered at z.

When $V \subset X$, we say that P is *σ-V-directionally porous* if there are $e_i \in V$ such that $P = \bigcup_{i=1}^{\infty} P_i$ where P_i is porous in direction of e_i. In case $V = X$ we speak about *σ-directionally porous sets.*

Many of our results on $\mathcal{L}$ null sets will stem from the obvious fact that if P is porous in the direction of a vector $e \in X$ then the function $f(x) = \operatorname{dist}(x, P)$ is

not differentiable in the direction of e at any point of P. For the sake of reference, we record an immediate consequence of this observation for σ-directionally porous sets.

1.1. **Lemma.** *Every σ-directionally porous subset of a separable Banach space is Lipschitz null.*

The role of porosity has often turned out to be crucial in obtaining deep results on differentiability. The reason for this can be seen in the research monograph [29]: these sets tend to form a subclass of nondifferentiability sets that captures the worst type of local behavior. However, while [29] is mainly concerned with Fréchet differentiability and the related (ordinary) porosity (in whose definition one allows the unit vector e to depend on x and ε), here we treat questions related to directional or Gâteaux differentiability, and so we are concerned with directional porosity. Because of that our definition of directional porosity is sligthy finer than in [29] in that the directions are oriented. This is a genuine difference, as by an example in [32] a set porous in the direction of e need not be a countable union of sets porous in the direction of $-e$. We will, however, sometimes need to consider the directions as not oriented. This could be implemented by requiring that the set V in the definition of σ-V-porosity be symmetric about zero, but we find it more convenient to speak about σ-$(\pm V)$-porosity, where $\pm V = V \cup -V$ and $-V = \{-v : v \in V\}$.

Notice also that our notion of σ-V-porosity is related to directional differentiability and so differs from the notion of σ-directional porosity in the direction of V, which has been used to investigate Fréchet differentiability in [29]. However, the following lemma implies that this difference disappears when V is finite dimensional; it shows that no new notion of (unoriented) σ-porosity would be obtained had we defined porosity using finite dimensional directions instead of one direction. Its proof is a modification of the arguments from [41] to our situation.

1.2. **Lemma.** *Suppose that X is a Banach space, U its subspace spanned by a finite set V, and that a set $E \subset X$ has the property that for every $x \in E$ there is $c_x > 0$ such that for every $\delta > 0$ one may find $u \in U$ and $t > 0$ such that $\delta > t > c_x\|u\|$ and $E \cap B(x+u,t) = \emptyset$. Then E is σ-$(\pm V)$-porous.*

Proof. Let $v_1, \dots, v_n$ be a basis for U chosen from elements of V. We write X as a direct sum of U and of a closed subspace W, denote the corresponding projections π_U and π_W and, noting that the statement does not depend on the choice of (equivalent) norm, we assume that the norm of X is the ℓ_1 sum of the norm $\|\sum_i \lambda_i v_i\| = \sum_i |\lambda_i|$ on U and of the original norm on W. We will also assume that there is $0 < \kappa < 1$ such that $c_x > \kappa$ for all $x \in E$; clearly E is a countable union of such sets.

Let $1 - 1/n < q < 1$ and $m \in \mathbb{N}$ be such that $q^m < \kappa$. For $y \in X$ and $r, s > 0$ let

$$K(y,r,s) = y + \{u \in U : \|u\| < r\} + \{w \in W : \|w\| < s\}$$

and for each $i, j = 0, 1, \dots$ let

$$F_{i,j} = E \cap \bigcup \{K(y, q^{-i}r, q^i r) : K(y,r,r) \cap E = \emptyset, r < 2^{-j}\} \text{ and}$$

$$G_{i,j} = \bigcap_{k=0}^{\infty} F_{i,k} \setminus F_{i-1,j} \text{ for } i \geq 1.$$

Observing that $E = \bigcap_{j=0}^{\infty} F_{m,j}$ and $F_{0,j} = \emptyset$ for each j, we infer that

$$E = \bigcup_{i=1}^{m} \Big(\bigcap_{j=0}^{\infty} F_{i,j} \setminus \bigcap_{j=0}^{\infty} F_{i-1,j} \Big) = \bigcup_{i=1}^{m} \bigcup_{j=1}^{\infty} G_{i,j}.$$

Let $c = \min\big(q^{2m-1}(1-q), q-1-1/n\big)$. If $x \in G_{i,j}$ and $k > j$, we find $y \in X$ and $0 < r < 2^{-k}$ such that $K(y,r,r) \cap E = \emptyset$ and $x \in K(y, q^{-i}r, q^i r)$. We write $\pi_U(y-x) = \sum_i \lambda_i v_i$, choose l such that $|\lambda_l|$ is maximal, and show that $B(x + \lambda_l v_l, c|\lambda_l|) \subset K(y, q^{-(i-1)}r, q^{(i-1)}r)$. Indeed, if $z \in B(x + \lambda_l v_l, c|\lambda_l|)$, then

$$\begin{aligned} \|\pi_U(z-y)\| &\le \|z-x\| + \|\pi_U(x-y)\| \le c|\lambda_l| + \sum_{i \ne l} |\lambda_i| \\ &\le c|\lambda_l| + (1-1/n)\sum_i |\lambda_i| \le cq^{-i} + (1-1/n)q^{-i} \le q^{-(i-1)} \end{aligned}$$

and

$$\begin{aligned} \|\pi_W(z-y)\| &\le \|z-x\| + \|\pi_W(x-y)\| \le c|\lambda_l| + q^i r \\ &\le q^{2m-1}(1-q)q^{-i}r + q^i r \le q^{(i-1)}r. \end{aligned}$$

Using also that $r < 2^{-j}$ and $G_{i,j} \cap F_{i-1,j} = \emptyset$, we infer

$$G_{i,j} \cap B(x + \lambda_l v_l, c|\lambda_l|) \subset G_{i,j} \cap K(y, q^{-(i-1)}r, q^{(i-1)}r) = \emptyset.$$

Since the number of possible directions v_l is finite, one of them has to occur for infinitely many k with the same sign of λ_l, and we conclude that $G_{i,j}$ is a union of $2n$ sets each of which is porous in one of the directions $\pm v_l$. □

1.3. **Corollary.** *Suppose that X is a separable Banach space and that the linear span of a set $V \subset X$ is dense in X. Then every σ-directionally porous subset of X is σ-$(\pm V)$-porous.*

Proof. Let $w_k \in X$ and $E = \bigcup_k E_k$ be such that for every $x \in E_k$ there is a sequence $t_i \searrow 0$ such that $B(x + t_i w_k, t_i/k) \cap E_k = \emptyset$. To prove that E_k is σ-$(\pm V)$-porous it suffices to choose $v_1, \dots, v_n \in V$ and $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ such that $\|w_k - \sum_i \lambda_i v_i\| < 1/(2k)$, to observe that E_k is porous in the direction of $\sum_i \lambda_i v_i$ and to use Lemma 1.2. □

Measurabilities, etc. Here we collect several facts about measurable or Lipschitz dependence of derivatives on variables. We omit their simple standard proofs.

1.4. **Lemma.** *Let f be a real-valued locally Lipschitz function on an open subset G of a Banach space X. Then*

(i) the function $(x,v) \mapsto \overline{D}f(x,v)$ is a Borel measurable function on $G \times X$, and
(ii) for every $x \in X$ the function $v \mapsto \overline{D}f(x,v)$ is Lipschitz; it is K-Lipschitz if f is K-Lipschitz a neighborhood of x.

1.5. **Lemma.** *Let Y and Z be Banach spaces, Y separable, and let f be a locally Lipschitz mapping of an open subset G of Y to Z. Then*

(i) the set $\{(y,v) \in G \times Y : f'_+(y;v)$ exists $\}$ is a Borel subset of $G \times Y$,
(ii) the mapping $(y,v) \mapsto f'_+(y;v)$ (whose domain is the set from (i)) is Borel measurable, and
(iii) for every $y \in G$ the set $\{v \in Y : f'_+(y;v)$ exists $\}$ is closed in Y.

1.6. **Corollary.** *The σ-ideal $\mathcal{L}$ is generated by Borel sets, i.e., every $N \in \mathcal{L}$ is contained in a Borel set $B \in \mathcal{L}$.*

2. Basic notions and results

In this section we deduce crucial estimates of the size of derived sets of Lipschitz mappings between Banach spaces. These results have quite satisfactory formulation in the finite dimensional case (e.g., Corollary 2.15 and Corollary 2.24), but their formulation is not obvious in infinite dimensional spaces, partly because the usual notion of derived sets needs compactness, and we are forced to replace the study of derived sets by the study of their δ approximating sets. Three main types of comparison of approximating derived sets are then considered. The simplest one is essentially known and concerns the chain rule in a fixed direction, a slightly more complicated concerns the comparison of derived sets based on different (subdifferential-type) ideas, and the most complicated, and most applicable, concerns dependence of derived sets on directions.

Derived sets. Suppose that Y and Z are Banach spaces, $f : Y \mapsto Z$, and $y, v \in Y$. The *derived set of f at the point y in the direction of v* is defined as the set $\mathcal{D}f(y, v)$ of all existing limits $\lim_{n\to\infty}(f(y + t_n v) - f(y))/t_n$, where $t_n \searrow 0$. The *δ-approximating derived set of f at y in the direction of v* is defined, for $\delta > 0$, by

$$\mathcal{D}_\delta f(y, v) = \left\{ \frac{f(y + tv) - f(y)}{t} : 0 < t < \delta \right\}.$$

It is easy to see that

$$\mathcal{D}f(y, v) = \bigcap_{\delta>0} \overline{\mathcal{D}_\delta f(y, v)}. \tag{2.1}$$

Other, from our point of view less important, notions of derived sets have been usually implicitly considered by a number of authors (see, for example, [31, 40, 10]).

The first, closest to the previous one, is obtained by defining, for each $\delta, c > 0$ the *(δ, c)-approximating Zahorski derived set of f at y in the direction of v* by

$$\mathcal{Z}_{\delta,c} f(y, v) = \left\{ \frac{f(y + sv + tv) - f(y + sv)}{t} : 0 < t < \delta, \operatorname{dist}(0, [s, s + t]) < ct \right\}.$$

Similar ideas in multidimensional situation lead to defining for each $\delta, c > 0$ the *(δ, c)-approximating Michel–Penot derived set of f at y in the direction of v* by

$$\mathcal{P}_{\delta,c} f(y, v) = \left\{ \frac{f(\hat{y} + tv) - f(\hat{y})}{t} : 0 < t < \delta, \operatorname{dist}(y, [\hat{y}, \hat{y} + tv]) < ct\|v\| \right\}.$$

The ideas behind Clarke's subdifferential provide our last notion of derived sets by defining, for each $\delta > 0$ the *δ-approximating strict (or Clarke–Rockafellar) derived set of f at y in the direction of v* by

$$\mathcal{C}_\delta f(y, v) = \left\{ \frac{f(\hat{y} + tv) - f(\hat{y})}{t} : 0 < t < \delta, \|\hat{y} - y\| < \delta \right\}.$$

All these notions lead to the notion of corresponding derived sets by formulas analogous to (2.1).

$$\mathcal{Z}f(y,v) = \bigcup_{c>0}\bigcap_{\delta>0}\overline{\mathcal{Z}_{\delta,c}f(y,v)}, \quad \mathcal{P}f(y,v) = \bigcup_{c>0}\bigcap_{\delta>0}\overline{\mathcal{P}_{\delta,c}f(y,v)}, \text{ and}$$
$$\mathcal{C}f(y,v) = \bigcap_{\delta>0}\overline{\mathcal{C}_{\delta}f(y,v)}.$$

Immediately from the definition we can see how the approximating derived sets behave when the direction is multiplied by a scalar as well as connections between various derived sets. We state most of them precisely in the case of $\mathcal{D}_\delta$ only.

2.1. **Lemma.** *Let $f : Y \mapsto Z$ and $\lambda, c > 0$.*

(i) $\mathcal{D}_\delta f(y,0) = \{0\}$.
(ii) $\mathcal{D}_\delta f(y,\lambda v) = \lambda\mathcal{D}_{\lambda\delta} f(y,v)$.
(iii) $\mathcal{D}_\delta f(y,v) \subset \mathcal{Z}_{\delta,c} f(y,v) = -\mathcal{Z}_{\delta,c} f(y,-v)$.

Proof. Only the last equality in (iii) may need an argument. If $0 < t < \delta$ then

$$\frac{f(y+sv+tv)-f(y+sv)}{t} = -\frac{f(y-(s+t)(-v)+t(-v))-f(y-(s+t)(-v))}{t}$$

and $\operatorname{dist}(0,[s,s+t]) < ct$ is equivalent to $\operatorname{dist}(0,[-(s+t),-(s+t)+t]) < ct$. □

We denote by $\mathcal{H}(Z)$ the family of non-empty bounded subsets of Z and recall that the *Hausdorff distance* of $S,T \in \mathcal{H}(Z)$ is defined by

$$\varrho(S,T) = \inf\{\varepsilon > 0 : S \subset B(T;\varepsilon) \text{ and } T \subset B(S;\varepsilon)\},$$

where $B(S;\varepsilon) = \{z \in Z : \operatorname{dist}(z,S) < \varepsilon\}$. We also let

$$\varrho_+(S,T) = \sup_{x\in S}\operatorname{dist}(x,T)$$

and observe that $\varrho(S,T) = \max\big(\varrho_+(S,T), \varrho_+(T,S)\big) = \varrho(T,S)$.

2.2. *Remark.* Clearly, if Z is finite dimensional and if f is Lipschitz on a neighborhood of y, the intersection in formula (2.1) may be replaced by the limit of $\mathcal{D}_\delta f(y,v)$ as $\delta \searrow 0$; analogous observations apply to the definitions of the remaining derived sets. No comparable statements hold if Z is infinite dimensional.

We will also need the following simple facts.

2.3. **Lemma.** *Suppose that a mapping f of a Banach space Y to a Banach space Z is K-Lipschitz on a neighborhood of a point $y \in Y$. Then for every $R > 0$ there is δ_0 such that for every $0 < \delta < \delta_0$ and every $c > 0$,*

$$v \mapsto \mathcal{D}_\delta f(y,v), \quad v \mapsto \mathcal{Z}_{\delta,c} f(y,v), \quad v \mapsto \mathcal{P}_{\delta,c} f(y,v), \quad \text{and} \quad v \mapsto \mathcal{C}_\delta f(y,v)$$

are K-Lipschitz mappings of $\{v \in Y : \|v\| \le R\}$ to $\mathcal{H}(Z)$ equipped with the Hausdorff metric.

Proof. If f is K-Lipschitz on $B(y,\Delta)$, we let $\delta_0 = \frac{\Delta}{2(1+R)}$ and infer that if $0 < t < \delta_0$, $\|w\| \le R$, and $\hat{y} \in B(y,\delta_0)$, then

$$\|(f(\hat{y}+tw)-f(\hat{y}))/t - (f(\hat{y}+tv)-f(\hat{y}))/t\| \le K\,\|w-v\|\,,$$

which immediately implies the statement. □

2.4. **Lemma.** *Suppose that a mapping f of a Banach space Y to a Banach space Z is K-Lipschitz on a neighborhood of a point $y \in Y$. Then for every $R > 0$ there is $\delta_0 > 0$ such that for $0 < \delta_2 < \delta_1 < \delta_0$ and $\|v\| \le R$ we have*

$$\rho(\mathcal{D}_{\delta_1} f(y,v), \mathcal{D}_{\delta_2} f(y,v)) \le 2K\|v\| \frac{\delta_1 - \delta_2}{\delta_1}$$

Proof. Choosing δ_0 as in the proof of the previous lemma, we see that for any $\delta_0 > \delta_1 > t_1 \ge \delta_2 > t_2 > 0$,

$$\begin{aligned}
&\Big\| \frac{f(y+t_1 v) - f(y)}{t_1} - \frac{f(y+t_2 v) - f(y)}{t_2} \Big\| \\
&\qquad \le \Big\| \frac{f(y+t_1 v) - f(y+t_2 v)}{t_1} \Big\| + \|f(y+t_2 v) - f(y)\| \Big| \frac{1}{t_1} - \frac{1}{t_2} \Big| \\
&\qquad \le 2K\|v\| \frac{t_1 - t_2}{t_1}.
\end{aligned}$$

Taking first the infimum over t_2 and then the supremum over t_1 gives the result. □

Chain rule formula in given direction.

2.5. **Proposition.** *Suppose that X, Y and Z are Banach spaces, that $g : X \mapsto Y$ is differentiable at a point $x \in X$ in the direction of $e \in X$, and that $f : Y \mapsto Z$ is Lipschitz on a neighborhood of $g(x)$. Then*

$$\lim_{\delta \searrow 0} \varrho(\mathcal{D}_\delta (f \circ g)(x,e), \mathcal{D}_\delta f(g(x), g'_+(x;\, e))) = 0.$$

Proof. Let $0 < K < \infty$ and $\Delta > 0$ be such that f is K-Lipschitz on $B(g(x), \Delta)$. For any $\varepsilon > 0$ we find $\eta > 0$ such that

$$\big\| g(x+te) - g(x) - t g'_+(x;\, e) \big\| \le \varepsilon t / K$$

for every $0 < t < \eta$, let $\kappa = \min\{\eta, \Delta\}/(1 + \big\| g'_+(x;\, e) \big\|)$, and infer that

$$\begin{aligned}
&\|(f(g(x) + t g'_+(x;\, e)) - f(g(x)) - ((f \circ g)(x+te) - (f \circ g)(x))\| \\
&\qquad\qquad \le K \big\| g(x) + t g'_+(x;\, e) - g(x+te) \big\| \le \varepsilon t
\end{aligned} \tag{2.2}$$

whenever $0 < t < \kappa$. Consequently, if $0 < t < \delta < \kappa$, we have that

$$\begin{aligned}
&\operatorname{dist}((f(g(x) + t g'_+(x;\, e)) - f(g(x)))/t, \mathcal{D}_\delta (f \circ g)(x,e)) \le \varepsilon, \text{ and} \\
&\operatorname{dist}(((f \circ g)(x+te) - (f \circ g)(x))/t, \mathcal{D}_\delta f(g(x), g'_+(x;\, e))) \le \varepsilon,
\end{aligned}$$

and we conclude that

$$\varrho(\mathcal{D}_\delta (f \circ g)(x,e), \mathcal{D}_\delta f(g(x), g'_+(x;\, e))) \le \varepsilon$$

if $0 < \delta < \kappa$. □

2.6. **Corollary.** *Suppose that $g : X \mapsto Y$ is differentiable at a point x in the direction of e and that $f : Y \mapsto Z$ is Lipschitz on a neighborhood of $g(x)$. Then*

$$\mathcal{D}(f \circ g)(x,e) = \mathcal{D} f(g(x), g'_+(x;\, e)).$$

Proof. By the formula (2.2) from the proof of Proposition 2.5,

$$\lim_{t \searrow 0} \frac{f(g(x) + t g'_+(x;\, e)) - f(g(x))}{t} - \frac{(f \circ g)(x+te) - (f \circ g)(x)}{t} = 0.$$

□

We note that in the simple example of Lipschitz mappings $g : \mathbb{R} \mapsto \mathbb{R}^2$ and $f : \mathbb{R}^2 \mapsto \mathbb{R}$ defined by $g(x) = (x, x\sin(\log|x|))$ and $f(x,y) = y - x\sin(\log|x|)$ (where we let $x\sin(\log|x|) = 0$ for $x = 0$) we have $f \circ g = 0$, $(1,0) \in \mathcal{D}g(0,1)$, and $1 \in \mathcal{D}f((0,0),(1,0))$. So the assumption of existence of $g'_+(x;e)$ is necessary for the ability to find one of the derived sets knowing the other. However, if Y and Z are finite dimensional and g is Lipschitz on a neighborhood of x, we have at least that $\mathcal{D}(f\circ g)(x,e) \subset \overline{\bigcup_{v\in\mathcal{D}g(x,e)} \mathcal{D}f(g(x),v)}$, which follows from the following statement by Remark 2.2.

2.7. **Proposition.** *Suppose that X, Y and Z are Banach spaces, that $g : X \mapsto Y$ is such that the sets $\mathcal{D}_\delta g(x,e)$ converge to $V \in \mathcal{H}(Y)$ as $\delta \searrow 0$, and that $f : Y \mapsto Z$ is Lipschitz on a neighborhood of $g(x)$. Then*

$$\lim_{\delta\searrow 0} \varrho_+\big(\mathcal{D}_\delta(f\circ g)(x,e), \textstyle\bigcup_{v\in V} \mathcal{D}_\delta f(g(x),v)\big) = 0.$$

Proof. Let $1 \le K < \infty$ and $\Delta > 0$ be such that f is K-Lipschitz on $B(x,\Delta)$. For any $\varepsilon > 0$ we find $\eta > 0$ such that for every $0 < t < \eta$ there is $v_t \in V$ for which

$$\|g(x+te) - g(x) - tv_t\| \le \varepsilon t/K,$$

let $\kappa = \min\{\eta,\Delta\}/(1 + \sup\{\|v\| : v \in V\} + \|e\|)$, and infer that

$$\|(f(g(x)+tv_t) - f(g(x))) - ((f\circ g)(x+te) - (f\circ g)(x))\| \le \varepsilon t$$

whenever $0 < t < \kappa$. Consequently, if $0 < t < \delta < \kappa$, we have that

$$\operatorname{dist}\big(((f\circ g)(x+te) - (f\circ g)(x))/t, \textstyle\bigcup_{v\in V} \mathcal{D}_\delta f(g(x),v)\big) \le \varepsilon,$$

and we conclude that

$$\varrho_+\big(\mathcal{D}_\delta(f\circ g)(x,e), \textstyle\bigcup_{v\in V} \mathcal{D}_\delta f(g(x),v)\big) \le 2\varepsilon$$

if $0 < \delta < \kappa$. □

Comparison of derived sets. We will see in Example 2.17 that in infinite dimensional situation the Hausdorff distance (and even the much weaker distance $|\operatorname{diam}(S) - \operatorname{diam}(T)|$) does not have the expected property that $\mathcal{D}_\delta f(x,-e)$ is often close to $-\mathcal{D}_\delta f(x,e)$. The basic reason behind this is that while in the final dimensional situation $\mathcal{H}(Z)$ equipped with the Hausdorff distance is separable, in the infinite dimensional situation it is not. For us, a more revealing version of this observation is that the Hausdorff distance on $\mathcal{H}(Z)$ is

$$\varrho_+(S,T) = \sup_{\phi\in\mathrm{Lip}_1(Z)} \varrho_\phi(S,T), \tag{2.3}$$

where $\mathrm{Lip}_1(Z)$ is the set of real-valued functions ϕ on Z with $\mathrm{Lip}(\phi) \le 1$, and

$$\varrho_\phi(S,T) = \sup_{z\in S\cup T} \phi(z) - \sup_{z\in T} \phi(z); \tag{2.4}$$

and that we can restrict the supremum in (2.3) to countably many $\phi \in \mathrm{Lip}_1(Z)$ if Z is finite dimensional, but cannot do so if it is not. In the following, we avoid this problem by taking countably many Lipschitz functions ϕ on Z and considering $\mathcal{H}(Z)$ equipped with countably many distances ϱ_ϕ.

Notice that ϱ_ϕ defined by (2.4) should be thought of as a weaker analogue of ϱ_+, not of ϱ. In particular, $\varrho_\phi(S,T) = 0$ whenever $S \subset \overline{T}$ and $\varrho_\phi(S_0,T) \le \varrho_\phi(S_1,T)$ whenever $S_0 \subset S_1$. One can also define, in the same way as for the Hausdorff

distance, a symmetrized version of this distance as $\max\{\varrho_\phi(T,S), \varrho_\phi(S,T)\}$. We, however, do not do it, since we do not have any genuine use for it.

It is obvious that $\varrho_\phi(S,T) \le \varrho_+(S,T)$ and, since

$$\sup_{z\in S\cup U} \phi(z) - \sup_{z\in U} \phi(z) \le \sup_{z\in S\cup T} \phi(z) - \sup_{z\in T} \phi(z) + \sup_{z\in T\cup U} \phi(z) - \sup_{z\in U} \phi(z),$$

we see that ϱ_ϕ satisfies (the nonsymmetric version of) the triangle inequality

$$\varrho_\phi(S,U) \le \varrho_\phi(S,T) + \varrho_\phi(T,U).$$

As in the use of the distance ϱ_+, where it is often convenient to use the inclusion $S \subset B(T;\varepsilon) = \{z : \operatorname{dist}(z,T) < \varepsilon\}$ instead of the (essentially equivalent) inequality $\varrho_+(S,T) < \varepsilon$, it will be often convenient to denote

$$\phi(T) = \sup_{z\in T} \phi(z)$$

and

$$B_\phi(T;\varepsilon) = \{z \in Z : \phi(z) < \phi(T) + \varepsilon\}$$

and use $S \subset B_\phi(T;\varepsilon)$ instead of the (essentially equivalent) inequality $\varrho_\phi(S,T) < \varepsilon$.

Note that (2.3) immediately implies $B(T;\varepsilon) \subset B_\phi(T;\varepsilon)$ for any $T \in \mathcal{H}(Z)$ and $\phi \in \operatorname{Lip}_1(Z)$.

2.8. *Examples.* The following choices of ϕ and the description of the corresponding sets $B_\phi(T;\varepsilon)$ in a Banach space Z will be used in the following text.

(i) If $Q \in \mathcal{H}(Z)$ and $\phi(w) = \operatorname{dist}(w,Q)$ then $B_\phi(S;\varepsilon) = B(Q;\varrho_+(S,Q)+\varepsilon)$.

(ii) If $z \in Z$ and $\phi(w) = \|w-z\|$ (which is a special case of (i) with $Q = \{z\}$), then $B_\phi(T;\varepsilon) = B(z, r_z(T)+\varepsilon)$, where $r_z(T) := \inf\{r > 0 : T \subset B(z,r)\}$ is the outer radius of T about z.

The set $\operatorname{Lip}_1(Z)$ will be equipped with the topology of uniform convergence on bounded sets. Its main use stems from the fact that the distances ϱ_ϕ do not change much with uniform changes of ϕ, which is recorded in the following lemma.

2.9. **Lemma.** *If $\phi, \psi \in \operatorname{Lip}_1(Z)$ and $S, T \in \mathcal{H}(Z)$, then*

$$B_\psi(T;\varepsilon) \subset B_\phi\big(S;\, \varrho_\phi(T,S) + \sup_{z\in T} |\phi(z) - \psi(z)| + \varepsilon\big)$$

and

$$\varrho_\phi(S,T) \le \varrho_\psi(S,T) + 2\sup_{z\in S\cup T} |\phi(z) - \psi(z)|.$$

Proof. The first statement follows by observing that for $z \in B_\psi(T;\varepsilon)$,

$$\begin{aligned}\psi(z) < \psi(T) + \varepsilon \le \phi(T) + \sup_{z\in T} |\phi(z) - \psi(z)| + \varepsilon \\ \le \phi(S) + \varrho_\phi(T,S) + \sup_{z\in T} |\phi(z) - \psi(z)| + \varepsilon.\end{aligned}$$

and the second by estimating

$$\varrho_\phi(S,T) = \sup_{z\in S\cup T} \phi(z) - \sup_{z\in T} \phi(z) \le \sup_{z\in S\cup T} \psi(z) - \sup_{z\in T} \psi(z) + 2\sup_{z\in S\cup T} |\phi(z) - \psi(z)|. \quad \square$$

2.10. **Lemma.** *Suppose a mapping f of a Banach space Y to a Banach space Z is Lipschitz on a neighborhood of a point $y \in Y$. Then for any $c > 0$ the sets*

$$\{(v,\phi) \in Y \times \operatorname{Lip}_1(Z) : \lim_{\delta\searrow 0} \varrho_\phi(\mathcal{Z}_{\delta,c} f(y,v), \mathcal{D}_\delta f(y,v)) = 0\},$$
$$\{(v,\phi) \in Y \times \operatorname{Lip}_1(Z) : \lim_{\delta\searrow 0} \varrho_\phi(\mathcal{P}_{\delta,c} f(y,v), \mathcal{D}_\delta f(y,v)) = 0\},$$

and

$$\{(v,\phi)\in Y\times \mathrm{Lip}_1(Z) : \lim_{\delta\searrow 0}\varrho_\phi(\mathcal{C}_\delta f(y,v),\mathcal{D}_\delta f(y,v))=0\}$$

are closed in $Y\times \mathrm{Lip}_1(Z)$.

Proof. For small enough δ the sets whose ϱ_ϕ we are limiting all lie in a fixed ball $B(0,r)$. Hence, metrizing $Y\times \mathrm{Lip}_1(Z)$ by

$$\mathrm{dist}((u,\phi),(v,\psi)) = \|u-v\| + \sup_{z\in B(0,r)} |\phi(z)-\psi(z)|,$$

we see from Lemma 2.3 and Lemma 2.9 that the $\varrho_\phi(\cdot,\cdot)$ involved in these limits are Lipschitz as functions of (v,ϕ), with a bound on their Lipschitz constant independent of δ. The statement follows. □

2.11. **Lemma.** *Suppose that* $f:Y\to Z$ *is a* K*-Lipschitz mapping between Banach spaces* Y,Z*. Suppose further that* $\phi\in \mathrm{Lip}_1(Z)$*,* $v\in Y$*,* $\varepsilon,\delta>0$*,* $c\in\mathbb{R}$*, and that* E *is the set of* $y\in Y$ *such that* $c\le \phi(\mathcal{D}_\delta f(y,v))<c+\varepsilon/2$*. If* $\hat y\in Y$ *and* $0<t<\delta$ *are such that* $\phi\big(\frac{f(\hat y+tv)-f(\hat y)}{t}\big)>c+\varepsilon$ *then* $B\big(\hat y,\frac{\varepsilon t}{4K}\big)\cap E=\emptyset$.

Proof. If $x\in B\big(\hat y,\frac{\varepsilon t}{4K}\big)\cap E$, we would infer from

$$\Big\|\frac{f(\hat y+tv)-f(\hat y)}{t}-\frac{f(x+tv)-f(x)}{t}\Big\|\le \frac{2K\,\|x-\hat y\|}{t}<\frac{\varepsilon}{2}$$

that

$$c+\varepsilon<\phi\Big(\frac{f(\hat y+tv)-f(\hat y)}{t}\Big)<\phi\Big(\frac{f(x+tv)-f(x)}{t}\Big)+\frac{\varepsilon}{2}<c+\varepsilon. \qquad \square$$

2.12. **Proposition.** *Suppose that* Y *and* Z *are Banach spaces,* V *is a separable subset of* Y*, and* Φ *is a separable subset of* $\mathrm{Lip}_1(Z)$*. Suppose further that* f *is a locally Lipschitz mapping of an open subset* $G\subset Y$ *to* Z*.*

(i) *There is a* σ*-*$(\pm V)$*-directionally porous set* $P_{\mathcal{Z}}\subset Y$ *such that*

$$\lim_{\delta\searrow 0}\varrho_\phi(\mathcal{Z}_{\delta,c}f(y,v),\mathcal{D}_\delta f(y,v))=0$$

for every $\phi\in\Phi$*,* $y\in G\setminus P_{\mathcal{Z}}$*,* $v\in V$*, and* $c>0$*.*

(ii) *There is a* σ*-porous set* $P_{\mathcal{P}}\subset Y$ *such that*

$$\lim_{\delta\searrow 0}\varrho_\phi(\mathcal{P}_{\delta,c}f(y,v),\mathcal{D}_\delta f(y,v))=0$$

for every $\phi\in\Phi$*,* $y\in G\setminus P_{\mathcal{P}}$*,* $v\in V$*, and* $c>0$*.*

(iii) *There is a meager set* $P_{\mathcal{C}}\subset Y$ *such that*

$$\lim_{\delta\searrow 0}\varrho_\phi(\mathcal{C}_\delta f(y,v),\mathcal{D}_\delta f(y,v))=0$$

for every $\phi\in\Phi$*,* $y\in G\setminus P_{\mathcal{C}}$ *and* $v\in V$*.*

Proof. We may assume that f is K-Lipschitz on $Y=G$. Finding countable dense subsets $W\subset V$ and $\Psi\subset\Phi$, we infer from Lemma 2.10 that it suffices to prove the statements with V,Φ replaced by W,Ψ. Hence we may reduce the proof to the case when V and Φ are countable, and we may reduce it further to the case when V and Φ are one element sets, since the sets we are interested in are unions of sets corresponding to the pairs $(v,\phi)\in V\times\Phi$. The same reasoning shows that we may assume that $c>0$ in (i) and (ii) is fixed.

Assume therefore that $V=\{v\}$, $\Phi=\{\phi\}$ and $c>0$, and let $0<\varepsilon<1$. For $n=1,2,\dots$ and $k=0,\pm1,\pm2,\dots$ let $E_{n,k}$ be the set of those points $y\in Y$

such that $k\varepsilon/2 \le \phi(\mathcal{D}_\eta f(y,v)) < (k+1)\varepsilon/2$ for every $0<\eta<1/n$. We prove that $Y=\bigcup_{n,k} E_{n,k}$. Given any $y\in Y$ we use that f is Lipschitz to infer that the function $\eta\mapsto \phi(\mathcal{D}_\eta f(y,v))$ is bounded. Hence $s:=\limsup_{\eta\searrow 0}\phi(\mathcal{D}_\eta f(y,v))$ is finite, and we may choose $k=0,\pm1,\dots$ such that $k\varepsilon/2\le s<(k+1)\varepsilon/2$. Finding $n\in\mathbb{N}$ such that $\phi(\mathcal{D}_\eta f(y,v))<(k+1)\varepsilon/2$ for any $0<\eta<1/n$, and using the monotonicity of ϕ, we have that $k\varepsilon/2\le s\le \phi(\mathcal{D}_\eta f(y,v))<(k+1)\varepsilon/2$ for all such η, implying that $y\in E_{n,k}$.

(i) Let

$$P_{n,k}=\{y\in E_{n,k}:\limsup_{\delta\searrow 0}\varrho_\phi(\mathcal{Z}_{\delta,c}f(y,v),\mathcal{D}_\delta f(y,v))>\varepsilon\}.$$

If $y\in P_{n,k}$, there are arbitrarily small $0<\delta<1/n$ for which one may find $0<t<\delta$ and $s\in\mathbb{R}$ such that $\operatorname{dist}(y,[y+sv,y+sv+tv])<ct\|v\|$ and

$$\phi\Big(\frac{f(y+sv+tv)-f(y+sv)}{t}\Big)>\phi(\mathcal{D}_\delta f(y,v))+\varepsilon\ge k\frac{\varepsilon}{2}+\varepsilon.$$

Hence Lemma 2.11 with $\hat y=y+sv$ implies that $B(y+sv,\varepsilon t/(4K))\cap P_{n,k}=\emptyset$. Since $\|sv\|\le \operatorname{dist}(y,[y+sv,y+sv+tv])+\|tv\|\le (c+1)t\,\|v\|$, we infer that $B(y+sv,\varepsilon s/(4(c+1)K))\cap P_{n,k}=\emptyset$, which means that the set $P_{n,k}$ is porous at y in the direction of v or $-v$, where the sign depends on whether s was positive for arbitrarily small δ or not.

(ii) Let

$$P_{n,k}=\{y\in E_{n,k}:\limsup_{\delta\searrow 0}\varrho_\phi(\mathcal{P}_{\delta,c}f(y,v),\mathcal{D}_\delta f(y,v))>\varepsilon\}.$$

If $y\in P_{n,k}$, there are arbitrarily small $0<\delta<1/n$ for which one may find $0<t<\delta$ and $\hat y\in Y$ such that $\operatorname{dist}(y,[\hat y,\hat y+tv])<ct\|v\|$ and

$$\phi\Big(\frac{f(\hat y+tv)-f(\hat y)}{t}\Big)>\phi(\mathcal{D}_\delta f(y,v))+\varepsilon\ge k\frac{\varepsilon}{2}+\varepsilon.$$

Since $\|\hat y-y\|\le \operatorname{dist}(y,[\hat y,\hat y+tv])+\|tv\|\le (c+1)t\,\|v\|$ and Lemma 2.11 implies that $B(\hat y,\varepsilon t/(4K))\cap P_{n,k}=\emptyset$, the set $P_{n,k}$ is porous at y.

(iii) Let

$$P_{n,k}=\{y\in E_{n,k}:\limsup_{\delta\searrow 0}\varrho_\phi(\mathcal{C}_\delta f(y,v)),\mathcal{D}_\delta f(y,v))>\varepsilon\}.$$

If $y\in P_{n,k}$, there are arbitrarily small $0<\delta<1/n$ for which one may find $0<t<\delta$ and $\hat y\in Y$ such that $\|y-\hat y\|<\delta$ and

$$\phi\Big(\frac{f(\hat y+tv)-f(\hat y)}{t}\Big)>\phi(\mathcal{D}_\delta f(y,v))+\varepsilon\ge k\frac{\varepsilon}{2}+\varepsilon.$$

Then Lemma 2.11 implies $B(\hat y,\varepsilon t/(4K))\cap P_{n,k}=\emptyset$, which means that y does not belong to the interior of the closure of $P_{n,k}$. □

2.13. **Corollary.** *Let f be a locally Lipschitz mapping of an open subset G of a Banach space Y to a separable Banach space Z and let $V\subset Y$ be separable. Then there is a σ-$(\pm V)$-porous set $P\subset Y$ such that for every $y\in Y\setminus P$ and $v\in V$,*

$$\lim_{\delta\searrow 0}\operatorname{diam}(\mathcal{D}_\delta f(y,-v))\le 2\lim_{\delta\searrow 0}\operatorname{diam}(\mathcal{D}_\delta f(y,v)).$$

Proof. Let $P\subset Y$ be the σ-$(V\cup -V)$-directionally porous set obtained by the use of Proposition 2.12(i) with Φ consisting of functions $\phi_z(w)=\|w-z\|$, where $z\in Z$. Denote $c=\lim_{\delta\searrow 0}\operatorname{diam}(\mathcal{D}_\delta f(y,v))$. Given $y\in Y\setminus P$ and $\varepsilon>0$, find $\delta_0>0$ such that $\operatorname{diam}(\mathcal{D}_\delta f(y,v))<c+\varepsilon$ for every $0<\delta\le\delta_0$ and pick any

$z \in \mathcal{D}_{\delta_0} f(y,v)$. Using that $\mathcal{D}_\delta f(y,-v) \subset \mathcal{Z}_{\delta,1} f(y,-v) = -\mathcal{Z}_{\delta,1} f(y,v)$ and that $B_{\phi_z}(\mathcal{D}_\delta f(y,v);\,\varepsilon) \subset B(z, c+2\varepsilon)$ for $\delta \le \delta_0$, we estimate

$$\begin{aligned}\lim_{\delta\searrow 0} \operatorname{diam}(\mathcal{D}_\delta f(y,-v)) &\le \lim_{\delta\searrow 0} \operatorname{diam}(\mathcal{Z}_{\delta,1} f(y,v)) \\ &\le \limsup_{\delta\searrow 0} \operatorname{diam}(B_{\phi_z}(\mathcal{D}_\delta f(y,v);\,\varepsilon)) \le 2c + 4\varepsilon. \qquad \square\end{aligned}$$

2.14. **Corollary.** *Let f be a locally Lipschitz mapping of an open subset G of a Banach space Y to a separable Banach space Z and let $V \subset Y$ be separable. Then the set of those points $y \in G$ for which there is a direction of $v \in V$ at which f is unilaterally but not bilaterally differentiable is σ-$(V \cup -V)$-porous.*

Proof. Since existence of $f'(y;\,v)$ is equivalent to $\lim_{\delta\searrow 0} \operatorname{diam}(\mathcal{D}_\delta f(y,v)) = 0$, this follows from Corollary 2.13. $\square$

2.15. **Corollary.** *Suppose that Y is a separable Banach space and that f is a locally Lipschitz mapping of an open subset $G \subset Y$ to a finite dimensional space Z.*

(i) *There is a σ-directionally porous set $P_{\mathcal{Z}} \subset Y$ such that $\mathcal{Z} f(y,v) = \mathcal{D} f(y,v)$ for every $y \in G \setminus P_{\mathcal{Z}}$ and every $v \in Y$.*
(ii) *There is a σ-porous set $P_{\mathcal{P}} \subset Y$ such that $\mathcal{P} f(y,v) = \mathcal{D} f(y,v)$ for every $y \in G \setminus P_{\mathcal{P}}$ and every $v \in Y$.*
(iii) *There is a meager set $P_{\mathcal{C}} \subset Y$ such that $\mathcal{C} f(y,v) = \mathcal{D} f(y,v)$ for every $y \in G \setminus P_{\mathcal{C}}$ and every $v \in Y$.*

Proof. Since Z is finite dimensional, we may use Proposition 2.12 with $\Phi = \operatorname{Lip}_1(Z)$. Thus, to prove (i), we use Proposition 2.12(i) to infer that there is a σ-directionally porous set $P_{\mathcal{Z}} \subset Y$ such that

$$\lim_{\delta\searrow 0} \varrho_\phi(\mathcal{Z}_{\delta,c} f(y,v)), \mathcal{D}_\delta f(y,v)) = 0$$

for every $\phi \in \Phi$, $y \in G \setminus P_{\mathcal{Z}}$, $v \in V$, and every $c > 0$. Using this with $\phi(z) = \operatorname{dist}(z, \mathcal{D} f(y,v))$, we immediately infer (i). The proofs of the remaining statements are similar. $\square$

We have already seen in Corollary 2.13 that the approximating derived sets enjoy good, but not perfect, symmetry properties. If the range is finite dimensional, Corollary 2.15 (i) gives a better result, namely that the derived sets are symmetric.

2.16. **Corollary.** *Let f be a locally Lipschitz mapping of an open subset G of a separable Banach space Y to a finite dimensional space Z. Then there is a σ-directionally porous set $P \subset Y$ such that $\mathcal{D} f(y,v) = -\mathcal{D} f(y,-v)$ for every $y \in G \setminus P$ and every $v \in Y$.*

2.17. *Example.* The need for replacing the Hausdorff distance by weaker distances is illustrated by an example of a Lipschitz mapping $f : \mathbb{R} \mapsto \ell_\infty(\mathbb{R})$ such that for every $y \in \mathbb{R}$,

$$\lim_{\delta\searrow 0} \operatorname{diam}(\mathcal{D}_\delta f(y,-1)) \le 1 < 2 \le \lim_{\delta\searrow 0} \operatorname{diam}(\mathcal{D}_\delta f(y,1)). \tag{2.5}$$

Notice that this example does not use nonseparability of ℓ_∞, since ℓ_∞ may be replaced by the separable space spanned by the range of f. In particular, this example also shows that the constant 2 in Corollary 2.13 cannot be improved.

To construct f, we first define $g : \mathbb{R} \to \mathbb{R}$ by $g(x) = x \sin(\phi(x))$ where $\phi : \mathbb{R} \to \mathbb{R}$ is continuously differentiable on $(0,\infty)$ and such that $\phi(x) = 0$ if $x \le 0$ or $x \ge 1$,

and $\phi(x) = 2\sqrt{\log(1/x)}$ if $0 < x < 1/e$. It is obvious that g is a bounded Lipschitz function and that $\mathcal{D}_\delta g(0,1)$ contains both values 1 and -1, thus

$$\lim_{\delta\searrow 0} \operatorname{diam}(\mathcal{D}_\delta g(0,1)) \geq 2. \tag{2.6}$$

We show that

$$\lim_{\delta\searrow 0} \sup_{x\in\mathbb{R}} \operatorname{diam}(\mathcal{D}_\delta g(x,-1)) \leq 1. \tag{2.7}$$

Clearly, it suffices to consider $x \geq 0$ only. Given $0 < \varepsilon < 1$, we use uniform continuity of g' on $[\frac{1}{2}e^{-1/\varepsilon^2}, \infty)$ to find $0 < \delta < \frac{1}{2}e^{-1/\varepsilon^2}$ so that $|g'(x) - g'(y)| < \varepsilon$ if $x, y \geq \frac{1}{2}e^{-1/\varepsilon^2}$ and $|x - y| < \delta$. Consider first the case when $x \geq e^{-1/\varepsilon^2}$. Then for any $x - \delta < y < x$ the mean value theorem provides $y < \xi < x$ such that

$$|(g(y) - g(x))/(y - x) - g'(x)| = |g'(\xi) - g'(x)| < \varepsilon,$$

so all values of $(g(y)-g(x))/(y-x)$, $y < x$ belong to the interval $[g'(x)-\varepsilon, g'(x)+\varepsilon]$ of length 2ε. In the case when $0 < x < e^{-1/\varepsilon^2}$ we distinguish two possibilities. If $0 < y < x$, we get

$$\begin{aligned} |(g(y) - g(x))/(y - x) - g(x)/x| &= |y(\sin(\phi(y)) - \sin(\phi(x)))/(y - x)| \\ &\leq |y(\phi(y) - \phi(x))/(y - x)| \\ &\leq y|\phi'(y)| = 1/\sqrt{\log(1/y)} < \varepsilon. \end{aligned}$$

Finally, if $y \leq 0$, $(g(y) - g(x))/(y - x) = g(x)/(x - y)$ belongs to the interval with end points 0 and $g(x)/x$. Together, the two cases show that all values of $(g(y) - g(x))/(y - x)$, $y < x$ belong to the smallest interval containing 0 and $g(x)/x \pm \varepsilon$, whose maximal possible length is $1 + 2\varepsilon$.

Having proved (2.7), we denote $g_u(x) = g(u + x)$ and define $f : \mathbb{R} \to \ell_\infty(\mathbb{R})$ by $f(u) = g_u$, $u \in \mathbb{R}$. Since $\|g_u - g_v\|_\infty \leq \operatorname{Lip}(g)|u - v|$, f is Lipschitz. The xth coordinate of an element from $\mathcal{D}_\delta f(y, -1)$ belongs to $\mathcal{D}_\delta g(y + x, -1)$, and so the inequality (2.7) implies the first inequality in (2.5). For the last inequality in (2.5) we notice that $x = (-y)$th coordinates of elements from $\mathcal{D}_\delta f(y, 1)$ run through whole $\mathcal{D}_\delta g(0, 1)$ and use inequality (2.6).

Exactness of porosity descriptions. We show that the main results describing the size of the sets of points at which derived sets may differ are exact. For the case of the strict or Michel–Penot derived set these are variants of existing relatively simple constructions, and although we will present them, we will treat them briefly and not always in the optimal way. For the Zahorski derived sets our results are new, and the arguments are substantially more involved; we therefore handle this case in detail.

A predecessor of these results, which is a special case of Theorem 2.20(ii), is that for every σ-porous set P there is a Lipschitz function that is not Fréchet differentiable at any point of P; see [29, Theorem 3.4.3]. In fact, our argument for 2.20(ii) differs from theirs only slightly. Also, [29, Remark 3.4.4] indicates an argument showing a special case of Theorem 2.20(i) that for every σ-directionally porous set P there is a Lipschitz function that is not Gâteaux differentiable at any point of P. The argument in [29, Remark 3.4.4] is not completely clear, since it effectively reduces the complement of a set to a disjoint union of balls, and this may not preserve directional porosity. However, we may slightly modify it in the following way. Write P as a union of directionally porous sets P_k, deduce from [6, Theorem 7] or [24] that there are Lipschitz $g_k : X \to [0, 1]$ such that

$\min(1, \mathrm{dist}(x, P_k)) \le g_k(x) \le 2\min(1, \mathrm{dist}(x, P_k))$ and g_k is Gâteaux differentiable on the complement of $\overline{P_k}$. Now observe that the method of proof of [29, Theorem 3.4.3] shows that $g := \sum_k c_k g_k$, where $c_k > 0$ are sufficiently small constants, is Lipschitz and Gâteaux non-differentiable at any point of P. Incidentally, if X^* is separable, which is the main case of the proof of [29, Theorem 3.4.3], we may use [20] (or [22, Corollary 19]) to replace in the above argument Gâteaux by Fréchet, thereby providing an alternative proof of this result.

Before continuing, we should remind ourselves that σ-porous sets need not be σ-directionally porous, and that meager sets need not be σ-porous. Indeed, for the latter consider any meager set on the line whose complement is Lebesgue null, and for the former recall that in any infinite dimensional separable Banach space there is a σ-porous set whose complement is null on every line, see [36, Theorem 1] of [4, Theorem 6.39].

As a technical tool, we will use the following notions and results. A function $f : Y \to \mathbb{R}$ is called uniformly Gâteaux differentiable if it is Gâteaux differentiable on G and for every $u \in Y$ the convergence of $(f(y + tu) - f(y))/t$ to $f'(y)(u)$ is uniform in $y \in Y$. By a result of [24], for any closed set $F \subset Y$ and $r, \varepsilon > 0$ there is a uniformly Gâteaux differentiable $f : Y \to [0, 1]$ such that $\mathrm{Lip}(f) \le \varepsilon + 1/r$, $f(x) = 1$ for $x \in F$, and $f(x) = 0$ if $\mathrm{dist}(x, F) \ge r$. The function f is called regularly Gâteaux differentiable at y if it is Gâteaux differentiable at y and $\mathcal{P}f(y, u) = f'(y)(u)$ for every $u \in Y$. It is immediate that a continuous uniformly Gâteaux differentiable function is regularly Gâteaux differentiable at every point and that the function $x \to \mathrm{dist}(x, P)$ is regularly Gâteaux differentiable at every point of P that is not a porosity point of P.

2.18. **Lemma.** *Let $P \subset Q$ be closed subsets of a Banach space Y and $u \in Y$. There are a function $h : Y \to [0, \infty)$ and a closed set $\tilde{Q} \supset \mathrm{spt}(h)$ such that*

(i) $0 \le h(x) \le \min(1, \mathrm{dist}(x, Q)) \le \min(1, \mathrm{dist}(x, P))$*;*
(ii) $\mathrm{Lip}(h) \le 120$*;*
(iii) *h is regularly Gâteaux differentiable at every point of $Y \setminus P$;*
(iv) *$Q \cap \tilde{Q} \subset P$;*
(v) *whenever $y \in P$ and Q is porous at y in the direction of some vector $v \in Y$, then $\tilde{Q} \cup Q$ is porous at y in the direction of v as well;*
(vi) *whenever $y \in P$ and Q is porous at y in the direction of the vector u, then there are $c > 0$ and $t_i \searrow 0$ such that $y + t_i u + ct_i u, y + t_i u \in \tilde{Q} \setminus Q$ and $\liminf_{i\to\infty} \frac{h(y+t_i u+ct_i u)-h(y+t_i u)}{ct_i} \ge 10\|u\|$.*

Proof. We assume that $\|u\| = 1$, find $u^* \in Y^*$ with $\|u^*\| = u^*(u) = 1$, let $\kappa = \frac{5}{6}$, $\eta_j = \frac{1}{3}(\kappa^j - \kappa^{j+1})$ and $r_j = \frac{1}{8}\eta_j$, and for $i \in \mathbb{Z}$ and $j \in \mathbb{N}$ denote

$$P_{i,j} = \{y \in Y : u^*(y) = i\eta_j,\ \kappa^j - \eta_j \le \mathrm{dist}(y, Q) \le \kappa^j + \eta_j,\ \mathrm{dist}(y, P) \le \kappa^{j/2}\}.$$

We observe that $P_{i,j}$ are closed sets and $\mathrm{dist}(P_{i,j}, P_{k,l}) \ge \max(\eta_j, \eta_l)$ for $(i, j) \ne (k, l)$. Indeed, if $l = j$ the u^* images of these two sets are at least η_j apart, and if $l > j$ then $\mathrm{dist}(y, Q) \le \kappa^l + \eta_l \le \kappa^{j+1} + \eta_j$ for every $y \in P_{k,l}$ and $\mathrm{dist}(z, Q) \ge \kappa^j - \eta_j$ for every $z \in P_{i,j}$. Hence the distance of $P_{i,j}$ and $P_{k,l}$ is at least $\kappa^j - \kappa^{j+1} - 2\eta_j = \eta_j$.

Denote by $\mathcal{I}$ the set of pairs (i, j) for which $P_{i,j} \ne \emptyset$. For every $(i, j) \in \mathcal{I}$ denote $B_{i,j} = \{y : \mathrm{dist}(y, P_{i,j}) \le 2r_j\}$ and choose a uniformly Gâteaux differentiable function $h_{i,j} : Y \to \mathbb{R}$ such that $0 \le h_{i,j} \le \kappa^{j+1}$, $h_{i,j}(y) = \kappa^{j+1}$ for $y \in P_{i,j}$, $h_{i,j}(y) = 0$ when $\mathrm{dist}(y, P_{i,j}) \ge 2r_j$ and $\mathrm{Lip}(h_{i,j}) \le \kappa^{j+1}/r_j = 24\kappa/(1 - \kappa) = 120$.

Define $h = \sum_{(i,j)\in\mathcal{I}} h_{i,j}$ and $\tilde{Q} = P \cup \bigcup_{i,j\in\mathcal{I}} B_{i,j}$. Notice that

$$\operatorname{dist}(B_{i,j}, B_{k,l}) \geq \operatorname{dist}(P_{i,j}, P_{k,l}) - 2(r_j + r_l) \geq \tfrac{1}{2}(\eta_j + \eta_l) - 2(r_j + r_l) = 2(r_j + r_l)$$

whenever $(i,j) \neq (k,l)$. Since $B_{i,j} \subset \{y, \operatorname{dist}(y,P) \leq \kappa^{j/2} + 2r_j\}$ are closed sets, this implies that every point of $Y \setminus P$ has a neighborhood meeting at most one of the sets $B_{i,j}$. Recalling that $\operatorname{spt}(h_{i,j}) \subset B_{i,j}$, and that $h_{i,j}(x) \leq \kappa^{j+1} \leq \kappa^j - \eta_j \leq \operatorname{dist}(x,Q) \leq \operatorname{dist}(x,P)$ for $x \in B_{i,j}$, we get (i)–(iv).

To prove (v), let Q be porous at $y \in P$ in the direction of a unit vector v with constant $0 < c \leq 1$. Let $\alpha > 0$ be so small that $\alpha(\kappa^j + \eta_j + 2r_j) < (1-\alpha)r_j$ for every j. We show that $\tilde{Q} \cup Q$ is porous at y in the direction of v with constant αc. Consider any $t > 0$ such that $ct \leq \operatorname{dist}(y + tv, Q) \leq t$. If $\operatorname{dist}(y + tv, \tilde{Q}) \geq \alpha ct$, we are done, so suppose

$$\operatorname{dist}(y + tv, \tilde{Q} \setminus Q) = \operatorname{dist}(y + tv, \tilde{Q}) < \alpha ct.$$

Then there are i, j such that $(i,j) \in \mathcal{I}$ and $\operatorname{dist}(y + tv, P_{i,j}) < 2r_j + \alpha ct$. Since $ct \leq \operatorname{dist}(y+tv, Q) \leq \kappa^j + \eta_j + 2r_j + \alpha ct$, we get $(1-\alpha)ct \leq \kappa^j + \eta_j + 2r_j$. Therefore $\alpha ct < \alpha(\kappa^j + \eta_j + 2r_j)/(1-\alpha) < r_j$ and so $\operatorname{dist}(y + tv, P_{i,j}) < 3r_j$. Since $y \in Q$, $\operatorname{dist}(y, P_{i,j}) \geq \kappa^j - \eta_j > 3r_j$, and so there is $0 < s \leq t$ so that $\operatorname{dist}(y+sv, P_{i,j}) = 3r_j$. Then every z with $\operatorname{dist}(z, y+sv) < r_j$ satisfies $2r_j < \operatorname{dist}(z, P_{i,j}) < 4r_j$. For $(k,l) \neq (i,j)$ we have $\operatorname{dist}(z, P_{k,l}) \geq \operatorname{dist}(P_{i,j}, P_{k,l}) - 4r_j \geq \frac{1}{2}\eta_l > 2r_l$. This means $z \notin \tilde{Q} \setminus Q$. As $\operatorname{dist}(z, Q) \geq \operatorname{dist}(P_{i,j}, Q) - \operatorname{dist}(y + sv, P_{i,j}) - r_j = \operatorname{dist}(P_{i,j}, Q) - 4r_j > 0$, we get $z \notin Q$. It follows that $\operatorname{dist}(y + sv, \tilde{Q}) \geq r_j > \alpha cs$.

To prove (vi), let Q be porous at $y \in P$ in the direction of u and find $c_0 > 0$ and $s_k \searrow 0$ so that $B(y + s_k u, c_0 s_k) \cap Q = \emptyset$. Let $j_0 \in \mathbb{N}$ be such that $\kappa^{j_0/2-1} < c_0$ and fix for a while an index k with $s_k < \kappa^{j_0}$. Let $j = j(k) \in \mathbb{N}$ be such that $\kappa^j \leq \operatorname{dist}(y + s_k u, Q) < \kappa^{j-1}$. Then $j > j_0$, so we have $\kappa^{j/2-1} s_k < c_0 s_k < \kappa^{j-1}$. Find $\tilde{s}_k \in (0, s_k]$ such that $\operatorname{dist}(y + \tilde{s}_k u, Q) = \kappa^j$ and observe that this choice implies $\kappa^j \leq \tilde{s}_k \leq s_k < \kappa^{j-1}/c_0 < \kappa^{j/2}$. Choose the largest $\tilde{t}_k \leq \tilde{s}_k$ so that $u^*(y + \tilde{t}_k u)$ is an integer multiple of η_j. Then $\tilde{s}_k \geq \tilde{t}_k \geq \tilde{s}_k - \eta_j \geq \kappa^j - \eta_j > \frac{2}{3}\kappa^j$. In particular, $\operatorname{dist}(y + \tilde{t}_k u, P) \leq s_k < \kappa^{j/2}$, from which we see that $y + \tilde{t}_k u \in P_{i,j}$ for some $i = i(j,k) = i(k) \in \mathbb{Z}$. We also recall that this implies $h(y + \tilde{t}_k u) = \kappa^{j+1}$

We let $t_k = \tilde{t}_k - 2r_j$ and show that $h(y + t_k u) = 0$. For this, first recall that on $P_{i,j}$, u^* is identically equal to $i\eta_j$. Since $u^*(y + t_k u) = i\eta_j - 2r_j = i\eta_j - \eta_j/4$, we infer from $y + \tilde{t}_k u \in P_{i,j}$ that $2r_j \geq \operatorname{dist}(y + t_k u, P_{i,j}) \geq \operatorname{dist}(u^*(y + t_k u), u^*(P_{i,j})) = 2r_j$. Consequently, $h_{i,j}(y + t_k u) = 0$. For (i', j') different from (i,j) we have

$$\operatorname{dist}(y + t_k u, P_{i',j'}) \geq \operatorname{dist}(P_{i',j'}, P_{i,j}) - 2r_j \geq \max(\eta'_j, \eta_j) - 2r_j \geq 2r'_j$$

and so $h_{i',j'}(y + t_k u) = 0$ as well. Hence $h(y + t_k u) = 0$ and consequently

$$\frac{h(y + t_k u + 2r_j u) - h(y + t_k u)}{2r_j} = \frac{h(y + \tilde{t}_k u) - h(y + t_k u)}{2r_j} = \frac{\kappa^{j+1}}{2r_j} = 60.$$

Letting $c_k := 2r_j/t_k$ and using that h is Lipschitz, we see that (vi) holds provided a subsequence of c_k converges to some $c > 0$. But for this it suffices to observe that $c_k \geq 2r_j/s_k \geq c_0 \frac{1}{4}\eta_j/\kappa^{j-1} = \frac{1}{12}c_0\kappa(1-\kappa)$ and $c_k \leq 2r_j/(\tilde{s}_k - \eta_j - 2r_j) \leq 1$. □

2.19. **Proposition.** *Let P be a closed subset of a Banach space Y. Then for any $u_1, u_2, \dots \in Y$ there is a Lipschitz function $g : Y \to [0, \infty)$ such that*

(i) $0 \leq g(x) \leq \min(1, \operatorname{dist}(x, P))$;
(ii) *g is regularly Gâteaux differentiable at every point of $Y \setminus P$;*

(iii) *whenever P is porous at a point $y \in P$ in the direction of u_i, there are $c > 0$ and $t_k \searrow 0$ such that* $\liminf_{k\to\infty} \frac{g(y+t_k u_i + ct_k u_i) - g(y+t_k u_i)}{ct_k} \geq 4\|u_i\|$.

(iv) *whenever P is porous at a point $y \in P$ and $u \in Y$, there are $c > 0$, $y_k \to y$ and $t_k > c\|y_k - y\|$ such that $t_k \searrow 0$ and* $\liminf_{k\to\infty} \frac{g(y_k+t_k u) - g(y_k)}{t_k} \geq 4\|u\|$.

(v) *whenever $y \in P$ is not an interior point of P and $u \in Y$, there are $y_k \to y$ and $t_k \searrow 0$ such that* $\liminf_{k\to\infty} \frac{g(y_k+t_k u) - g(y_k)}{t_k} \geq 4\|u\|$.

Proof. Apply the $5r$-covering theorem to the family of balls $B(x, r(x))$, $x \in Y \setminus P$, where $r(x) = \frac{1}{6}\min(1, \operatorname{dist}(x, P))$ to find disjoint balls $B(x_i, r(x_i))$ such that for every $x \in Y \setminus P$ there is i with $B(x, r(x)) \subset B(x_i, 5r(x_i))$. Let

$$Q_0 = Y \setminus \bigcup_i B(x_i, \tfrac{1}{2}r(x_i)).$$

Clearly, $P \subset Q_0$. We observe that for every $x \in Y \setminus P$, $B(x, \frac{1}{3}r(x))$ meets at most one of the balls $B(x_i, \frac{1}{2}r(x_i))$. Indeed, if $B(x, \frac{1}{3}r(x)) \cap B(x_j, \frac{1}{2}r(x_j)) \neq \emptyset$, then, as follows from the definition of $r(x)$ we necessarily have $2r(x) \leq \frac{1}{3}r(x) + \frac{1}{2}r(x_j) + 2r(x_j)$, implying $r(x) \leq \frac{3}{2}r(x_j)$ and $B(x, \frac{1}{3}r(x)) \subset B(x_j, r(x_j))$. Hence j is unique by the disjointness of $B(x_i, r(x_i))$.

We show that for every $x \in P$ and $y \in Y \setminus P$ there is $z \in (x, y]$ so that $B(z, \frac{1}{5}r(y)) \cap Q_0 = \emptyset$. For this, assume that $B(y, \frac{1}{3}r(y)) \cap B(x_j, \frac{1}{2}r(x_j)) \neq \emptyset$ for some j, infer from $2r(y) \leq \frac{1}{3}r(y) + \frac{5}{2}r(x_j)$ that $r(y) \leq \frac{3}{2}r(x_j)$ and so that $\|y - x_j\| \leq r(x_j)$ and, using that $\|x - x_j\| \geq 2r(x_j)$, find $z \in (x, y]$ so that $\|z - x_j\| = \frac{3}{2}r(x_j)$. Hence $B(z, \frac{1}{2}r(x_j)) \cap Q_0 = \emptyset$, and the claim follows by recalling that $\frac{1}{3}r(y) \leq \frac{1}{2}r(x_j)$. Notice that this implies that porosity points of P are porosity points of Q_0, directional porosity points of P are directional porosity points of Q_0 in the same directions, and non-interior points of P remain non-interior points of Q_0.

We finish the starting part of the construction by choosing a Lipschitz, uniformly Gâteaux differentiable function $\phi : Y \to [0, 1]$ such that $\phi(0) = 1$ and $\phi(y) = 0$ for $\|y\| \geq 1$, and defining $h_0(x) = \sum_j \frac{1}{4}r(x_j)\phi(16(x - x_j)/r(x_j))$.

Letting $\tilde{Q}_0 = Q_0$, we recursively define for $k = 1, 2, \ldots$ closed sets Q_k and $\tilde{Q}_k$ and functions $h_k : Y \to \mathbb{R}$ by letting first $Q_k = Q_{k-1} \cup \tilde{Q}_{k-1} \cup \{x : \operatorname{dist}(x, P) \geq 2^{-k-1}\}$ and then using Lemma 2.18 with the sets P, Q_k and the vector u_k to define the set $\tilde{Q}_k$ and the function h_k.

Observing that 2.18(i) implies that $0 \leq h_k \leq 2^{-k-1}$, we see that the function $g := \sum_{k=0}^{\infty} h_k$ is well defined and $0 \leq g \leq 1$. Moreover, each h_k is zero on P and the sets $\operatorname{spt}(h_k) \setminus P$ are disjoint. As a consequence of this we see that (i) and (ii) hold and that g is Lipschitz with constant bounded by 120 on every segment contained in $Y \setminus P$. The last fact together with the already proved (i) shows that $\operatorname{Lip}(g) \leq 120$.

Let P be porous at $y \in P$ in the direction of u_j. Use 2.18(vi) to find $c > 0$ and $t_k \searrow 0$ such that $y + t_k u_j + ct_k u_j, y + t_k u_j \in \tilde{Q}_j \setminus Q_j$ and

$$\liminf_{k\to\infty} \frac{h_j(y + t_k u_j + ct_k u_j) - h_j(y + t_k u_j)}{ct_k} \geq 4\|u_j\|.$$

Since $g(y + t_k u_j + ct_k u_j) = h_j(y + t_k u_j + ct_k u_j)$ and $g(y + t_k u_j) = h_j(y + t_k u_j)$, we conclude that

$$\liminf_{k\to\infty} \frac{g(y + t_k u_j + ct_k u_j) - g(y + t_k u_j)}{ct_k} \geq 4\|u_j\|.$$

This proves (iii).

To prove (iv) and (v), consider any $u \in Y \setminus \{0\}$ and any sequence $z_k \in Y \setminus P$ such that $z_k \to y \in P$. Find i_k so that $B(z_k, r(z_k)) \subset B(x_{i_k}, 5r(x_{i_k}))$ and let $y_k = x_{i_k} - t_k u$ where $t_k = \frac{r(x_{i_k})}{16\|u\|}$. Then $6r(x_{i_k}) \le \|x_{i_k} - y\| \le 5r(x_{i_k}) + \|z_k - y\|$, so $0 < r(x_{i_k}) \le \|z_k - y\|$ and $\|x_{i_k} - y\| \le 6\|z_k - y\|$. Since $g(y_k + t_k u) = h_0(y_k + t_k u)$ and $g(y_k) = h_0(y_k)$, we conclude that $y_k \to y$, $t_k \searrow 0$ and

$$\liminf_{k\to\infty} \frac{g(y_k + t_k u) - g(y_k)}{t_k} = \liminf_{k\to\infty} \frac{h_0(y_k + t_k u) - h_0(y_k)}{t_k} = \frac{r(x_{i_k})}{4t_k} = 4\|u\|.$$

This proves (v). For (iv) we choose the sequence z_k so that $r(z_k) > c_0\|z_k - y\|$. Then, using $5r(x_{i_k}) \ge r(z_k) \ge c_0\|z_k - y\|$ and letting $C = \|u\|(1 + 80 + 80/c_0)$,

$$\|y_k - y\| \le t_k\|u\| + \|x_{i_k} - z_k\| + \|z_k - y\| \le t_k\|u\| + 5r(x_{i_k}) + 5r(x_{i_k})/c_0 = Ct_k.$$

Hence (iv) holds with any $0 < c < 1/C$. □

It is now straightforward to combine the functions g_k obtained by using the above result with a sequence of sets P_k to get functions with irregular behavior on a prescribed σ-directionally porous, σ-porous or meager set. In the following theorem we therefore state only the main such result. However, several additional observations about the function we construct in 2.20(ii) may be useful: it is Gâteaux differentiable at every point of P with the exception on those belonging to a σ-directionally porous set $P_d \subset P$; for every point $y \in P_d$ of which there is $v \in V$ with $\mathcal{Z}f(y, v) \ne \mathcal{D}f(y, v)$; and at every point of $P \setminus P_d$ the function f is irregularly differentiable in every direction in the sense used in [29]. This shows that it would not make sense to improve the sufficient condition for Γ-almost everywhere Fréchet differentiability results, namely the Γ-nullness of porous sets, to Γ-nullness of the sets of point of irregular differentiability, although the latter is what is actually used in [28] and [29].

2.20. **Theorem.** *Let P be a subset of a separable Banach space Y.*

(i) *If P is σ-directionally porous, there is a Lipschitz $f : Y \to \mathbb{R}$ such that for every $y \in P$ there is $v \in V$ with $\mathcal{Z}f(y, v) \ne \mathcal{D}f(y, v)$.*
(ii) *If P is σ-porous, there is a Lipschitz $f : Y \to \mathbb{R}$ such that $\mathcal{P}f(y, v) \ne \mathcal{D}f(y, v)$ for every $y \in P$ and $v \in Y \setminus \{0\}$.*
(iii) *If P is meager, then there is a Lipschitz $f : Y \to \mathbb{R}$ such that $\mathcal{C}f(y, v) \ne \mathcal{D}f(y, v)$ for every $y \in P$ and $v \in Y \setminus \{0\}$.*

Proof. Let $u_1, u_2, \dots$ be a sequence of unit vectors dense in the unit sphere of Y. (i) Write $P = \bigcup_{k=1}^\infty P_k$ where P_k is porous in direction of u_k. Let g_k be the functions constructed in Proposition 2.19 with $P = \overline{P_k}$. Choose $0 < \lambda_k \le 2^{-k}$ so that for every j, $\lambda_j \ge \sum_{k=j+1}^\infty \lambda_k \mathrm{Lip}(g_k)$. The functions $f_j = \sum_{k=j}^\infty \lambda_k g_k$ are clearly well-defined and Lipschitz.

Given $y \in P$ find the least j for which there is i such that $y \in P_j$ and P_j is porous at y in the direction of u_i. By 2.19(iii) find $c > 0$ and $t_k \searrow 0$ such that $\liminf_{k\to\infty} \frac{g_j(y+t_k u_i + ct_k u_i) - g_j(y+t_k u_i)}{ct_k} \ge 4$. Then

$$\liminf_{k\to\infty} \frac{f_j(y + t_k u_i + ct_k u_i) - f_j(y + t_k u_i)}{ct_k} \ge 4\lambda_j - \sum_{k=j+1}^\infty \lambda_k \mathrm{Lip}(g_k) \ge 3\lambda_j$$

and by 2.19(i)

$$\limsup_{t\to 0} \frac{f_j(y+tu_i)-f_j(y)}{t} \le \lambda_j + \sum_{k=j+1}^{\infty} \lambda_k \mathrm{Lip}(g_k) \le 2\lambda_j.$$

If $k<j$ and $y\in P_k$, 2.19(i) implies that g_k is differentiable at y in the direction of u_i, and if $y\notin P_k$, the same conclusion follows from 2.19(ii). Hence the statement holds with $f=f_1$.

(ii) Write $P=\bigcup_{k=1}^{\infty} P_k$ where P_k are porous and define g_k, λ_k and f_k as in (i). Given $y\in P$ find the least j such that $y\in P_j$ and P_j is porous at y. For every $v\in Y\setminus\{0\}$ replace $y+t_ku_i$ in the above argument by the y_k from 2.19(iv) and u_i by v to infer that $\mathcal{P}f_j(y,v)\neq \mathcal{D}f_j(y,v)$. If $k<j$ and $y\in P_k$, 2.19(i) implies that g_k is regularly Gâteaux differentiable at y, and if $y\notin P_k$, the same conclusion follows from 2.19(ii). Hence the statement holds with $f=f_1$.

(iii) The argument is similar to the previous ones and can be omitted. □

Subadditivity of derived sets. Though we postpone the case of one dimensional range to the next section, we should remark that it is now easy to deduce from 2.12(iii) or from 2.15(iii) the result of [21] on generic convexity of upper derivative; indeed one can improve it by using 2.12(ii) or 2.15(ii) to a result on convexity of upper derivative with a σ-porous exceptional set. However, even this improvement is not sufficient to obtain useful information about derivative of composite functions, and we need a considerable refinement of our previous results to achieve this. As in the previous part of our study, the Hausdorff distance is too fine for our purpose (see Example 2.25), and we will continue using the idea behind the distances ϱ_ϕ introduced above. However, even they become too coarse for our purpose, and we will therefore state the result using the neighborhoods $B_\phi(S;\varepsilon)$.

2.21. **Lemma.** *Let $f:Y\to Z$ be a K-Lipschitz map between separable Banach spaces Y and Z, $u,v\in Y$, $\phi\in \mathrm{Lip}_1(Z)$ and $\varepsilon>0$. Then there is a σ-u-porous set $P\subset Y$ such that for every $y\in Y\setminus P$ the inclusion*

$$\mathcal{D}_\delta f(y,u+v)\subset \mathcal{D}_\delta f(y,u)+B_\phi(\mathcal{D}_\delta f(y,v);\varepsilon) \tag{2.8}$$

holds for all sufficiently small $\delta>0$.

Proof. Since the function $\delta\to\phi(\mathcal{D}_\delta f(y,v))$ is bounded and monotonic, for every $y\in Y$ there are rational numbers $c,\tau>0$ such that $c\le\phi(\mathcal{D}_\delta f(y,v))<c+\varepsilon/2$ for every $0<\delta<\tau$. Let $P_{c,\tau}$ be the set of $y\in Y$ that have this property and for which (2.8) fails for arbitrarily small δ.

It suffices to show that each $P_{c,\tau}$ is porous in the direction of u. For this, consider any $y\in P_{c,\tau}$ and, given any $0<\delta<\tau$ for which (2.8) fails, find $0<t<\delta$ with

$$(f(y+t(u+v))-f(y))/t\notin \mathcal{D}_\delta f(y,u)+B_\phi(\mathcal{D}_\delta f(y,v);\varepsilon).$$

Since $(f(y+tu)-f(y))/t\in\mathcal{D}_\delta f(y,u)$, we infer that

$$(f(y+tu+tv)-f(y+tu))/t\notin B_\phi(\mathcal{D}_\delta f(y,v);\varepsilon).$$

Observing that $P_{c,\tau}$ is contained in the set E from Lemma 2.11, we use this lemma with $\hat{y}=y+tu$ to infer that $B(y+tu,\varepsilon t/(4K))\cap P_{c,\tau}=\emptyset$. It follows that $P_{c,\tau}$ is porous at y in the direction of u, as needed. □

2.22. **Proposition.** *Suppose that Y and Z are separable Banach spaces and that $\Phi \subset \mathrm{Lip}_1(Z)$ is separable. Suppose further that V and W are subsets of Y and that f is a locally Lipschitz mapping of an open subset G of Y to Z. Then there is a σ-V-directionally porous set $P \subset Y$ such that for every $\phi \in \Phi$, $\varepsilon > 0$, $y \in G \setminus P$, $v \in V$, and $w \in W$ one may find $\delta_0 > 0$ such that*

$$\mathcal{D}_\delta f(y, v + w) \subset \mathcal{D}_\delta f(y, v) + B_\phi(\mathcal{D}_\delta f(y, w);\, \varepsilon)$$

for every $0 < \delta < \delta_0$.

Proof. As explained in the introduction, we may assume that f is K-Lipschitz on Y. Choose countable dense sets $V_1 \subset V$, $W_1 \subset W$, and $\Phi_1 \subset \Phi$. Fix $\varepsilon > 0$ and for $v_1 \in V_1$, $w_1 \in W_1$, and $\phi_1 \in \Phi_1$ let P_{v_1,w_1,ϕ_1} be the σ-v_1-porous sets from Lemma 2.21 used with ε replaced by $\tau := \varepsilon/7$. We show that the statement holds with $P = \bigcup_{v_1 \in V_1, w_1 \in W_1, \phi_1 \in \Phi_1} P_{v_1,w_1,\phi_1}$, which is clearly a σ-V-porous set.

Let $y \notin P$, $v \in V$, $w \in W$, and $\phi \in \Phi$. Find $v_1 \in V_1$, $w_1 \in W_1$ and $\phi_1 \in \Phi_1$ such that $\|v - v_1\| < \tau/K$, $\|w - w_1\| < \tau/K$, and $\sup_{z \in B(0,K)} |\phi(z) - \phi_1(z)| < \tau$. By Lemmas 2.3 and 2.9,

$$\begin{aligned}
\mathcal{D}_\delta f(y, v + w) &\subset B\big(\mathcal{D}_\delta f(y, v_1 + w_1); 2\tau\big) \\
&\subset B\big(\mathcal{D}_\delta f(y, v_1) + B_{\phi_1}(\mathcal{D}_\delta f(y, w_1);\, \tau); 2\tau\big) \\
&\subset B\big(B(\mathcal{D}_\delta f(y, v);\, \tau) + B_{\phi_1}(\mathcal{D}_\delta f(y, w);\, 2\tau);\, 2\tau\big) \\
&\subset \mathcal{D}_\delta f(y, v) + B(B_{\phi_1}(\mathcal{D}_\delta f(y, w); 4\tau);\, 3\tau\big) \\
&\subset \mathcal{D}_\delta f(y, v) + B_\phi(\mathcal{D}_\delta f(y, w);\, 7\tau).
\end{aligned}$$

□

2.23. **Corollary.** *Suppose that X, Y, Z are separable Banach spaces, $\Phi \subset \mathrm{Lip}_1(Z)$ is separable, $V \subset X$ and $W \subset Y$. If $g : X \mapsto Y$ and $f : Y \mapsto Z$ are locally Lipschitz mappings, then there is a σ-V-directionally porous set $P \subset X$ such that for every $\phi \in \Phi$, $\varepsilon > 0$, $x \in X \setminus P$, $w \in W$, and for every direction of $v \in V$ for which $g'_+(x; v)$ exists there is δ_0 such that*

$$\mathcal{D}_\delta f(g(x), g'_+(x;\, v) + w) \subset \mathcal{D}_\delta f(g(x), g'_+(x;\, v)) + B_\phi(\mathcal{D}_\delta f(g(x), w);\, \varepsilon)$$

whenever $0 < \delta < \delta_0$.

Proof. Let $\tilde{g} : X \oplus Y \mapsto Y$ be defined by $\tilde{g}(x \oplus y) = g(x) + y$, and let $h = f \circ \tilde{g}$. We now treat V and W as subsets of $X \oplus \{0\}$ and $\{0\} \oplus Y$ respectively. Using Proposition 2.22, we find a σ-V-directionally porous set $Q \subset X \oplus Y$ such that for every $\phi \in \Phi$, $\varepsilon > 0$, $x \oplus y \in X \oplus Y \setminus Q$, $v \in V$, and $w \in W$ one may find $\delta_1 > 0$ for which

$$\mathcal{D}_\delta h(x \oplus y, v \oplus w) \subset \mathcal{D}_\delta h(x \oplus y, v) + B_\phi(\mathcal{D}_\delta h(x \oplus y, w);\, \varepsilon/2)$$

for every $0 < \delta < \delta_1$.

Note that $P = \{x \in X : x \oplus 0 \in Q\} = Q \cap X$ is a σ-V-directionally porous subset of X. Let $\phi \in \Phi$, $\varepsilon > 0$, $x \in X \setminus P$, $v \in V$, $w \in W$, and let $g'_+(x; v)$ exist. By the above consequence of Proposition 2.22 used for $x = x \oplus 0 \in X \setminus P$, we find $\delta_1 > 0$ such that

$$\mathcal{D}_\delta h(x, v \oplus w) \subset \mathcal{D}_\delta h(x, v) + B_\phi(\mathcal{D}_\delta h(x, w);\, \varepsilon/2) \tag{2.9}$$

holds for every $0 < \delta < \delta_1$. Since $x, v \in X$, we have that $\tilde{g}(x) = g(x)$ and $\tilde{g}'_+(x;\, av + bw) = ag'_+(x;\, v) + bw$ exists for any $a, b \geq 0$. Hence Proposition 2.5

implies that there is $0 < \delta_0 < \delta_1$ such that

$$\begin{aligned}(2.10)\quad &\varrho(\mathcal{D}_\delta h(x, av + bw), \mathcal{D}_\delta f(g(x), ag'_+(x;\, v) + bw)) \\ &\qquad = \varrho(\mathcal{D}_\delta h(x, av + bw), \mathcal{D}_\delta f(\tilde{g}(x), \tilde{g}'_+(x;\, av + bw))) < \varepsilon/6\end{aligned}$$

for every $0 < \delta < \delta_0$ and $a, b = 0, 1$. Hence (2.9) and (2.10) give that for any $0 < \delta < \delta_0$,

$$\mathcal{D}_\delta f(g(x), g'_+(x;\, v) + w) \subset \mathcal{D}_\delta f(g(x), g'_+(x;\, v)) + B_\phi(\mathcal{D}_\delta f(g(x), w);\, \varepsilon). \qquad \square$$

2.24. **Corollary.** *Suppose that f is a locally Lipschitz mapping of a separable Banach space Y to a finite dimensional space Z. Then there is a σ-directionally porous set $P \subset Y$ such that*

$$\mathcal{D}f(y, v + w) \subset \mathcal{D}f(y, v) + \mathcal{D}f(y, w)$$

for every $y \in Y \setminus P$, and $v, w \in Y$.

If, in addition, X is a separable Banach space and $g : X \mapsto Y$ is locally Lipschitz, then there is a σ-directionally porous set $Q \subset X$ such that

$$\mathcal{D}f(g(x), g'(x;\, e) + w) \subset \mathcal{D}f(g(x), g'(x;\, e)) + \mathcal{D}f(g(x), w)$$

whenever $x \in X \setminus Q$, $e \in X$ is such that $g'(x;\, e)$ exists, and $w \in Y$.

Proof. We prove just the additional part, since the former follows from it by considering the identity mapping g. From Corollary 2.23 with $\Phi = \mathrm{Lip}_1(Z)$ we infer that there is a σ-directionally porous set $Q \subset X$ such that for every $\varepsilon > 0$, $x \in X \setminus Q$, $w \in Y$, and for every direction of $e \in X$ at which g is differentiable there is δ_0 such that

$$\mathcal{D}_\delta f(g(x), g'(x;\, e) + w) \subset \mathcal{D}_\delta f(g(x), g'(x;\, e)) + B_\phi(\mathcal{D}_\delta f(g(x), w);\, \varepsilon)$$

whenever $0 < \delta < \delta_0$ and $\phi \in \mathrm{Lip}_1(Z)$. The statement follows by taking the limits as $\delta \searrow 0$ and with $\varepsilon \searrow 0$, and by using that $\mathcal{D}f(g(x), g'(x;\, e)) + \mathcal{D}f(g(x), w)$ is compact. $\square$

The following example illustrates the need for the finite dimensionality assumption on Z, even in the situation when $X = \mathbb{R}$, $Y = \mathbb{R}^2$, $g(x) = (x, 0)$ and $Z = \ell_2$.

2.25. *Example.* There is $f : \mathbb{R}^2 \mapsto \ell_2$ such that $f(x, 0) = 0$ and

$$\limsup_{\delta \searrow 0} \operatorname{diam}(\mathcal{D}_\delta f((x, 0), (1, 1))) \geq 2 \ > \sqrt{2} \geq \limsup_{\delta \searrow 0} \operatorname{diam}(\mathcal{D}_\delta f((x, 0), (0, 1)))$$

for every $x \in \mathbb{R}$.

Proof. Let e_j, $j = 1, 2, \ldots$, be an orthonormal basis of ℓ_2. We define

$$\begin{aligned} f(x, y) &= 0 \quad \text{if } y \leq 0 \text{ or } y \geq 1 \\ f(x, 2^{-2i}) &= 2^{-2i}(e_{2i} \cos(2^{-2i}\pi x) + e_{2i+1} \sin(2^{-2i}\pi x)), \\ f(x, 2^{-2i+1}) &= 2^{-2i+1}(e_{2i} \cos(2^{-2i}\pi x) + e_{2i+1} \sin(2^{-2i}\pi x)), \end{aligned}$$

and extend f to a function which is affine on each segment $[(x, 2^{-j-1}), (x, 2^{-j})]$, $j = 0, 1, \ldots$.

The first inequality follows by observing that for every $x \in \mathbb{R}$ the unit vectors $(f(x + 2^{-2i}, 2^{-2i}) - f(x, 0))/2^{-2i}$ and $(f(x + 2^{-2i+1}, 2^{-2i+1}) - f(x, 0))/2^{-2i+1}$ are opposite to each other, and hence $\operatorname{diam}(\mathcal{D}_\delta f((x, 0), (1, 1))) \geq 2$ for each $\delta > 0$.

To prove the second inequality, let $x \in \mathbb{R}$ and $0 < t < 1/2$ and observe that for $t \in [2^{-j}, 2^{-j+1}]$ the slope $(f(x, t) - f(x, 0))/t$ is a convex combination

of the slopes $(f(x,2^{-j})-f(x,0))/2^{-j}$ and $(f(x,2^{-j+1})-f(x,0))/2^{-j+1}$. Hence $\mathcal{D}_\delta f((x,0),(0,1))$ is contained in the convex hull of the mutually orthogonal unit vectors $e_{2i}\cos(2^{-2i}\pi x)+e_{2i+1}\sin(2^{-2i}\pi x)$ which has diameter $\sqrt{2}$. □

3. Convexity of upper derivative and other results for one dimensional target

In this section we intend to treat the special case of real-valued functions. The results, including the main Theorem 3.6, are easy corollaries of those from the previous section. However, the special nature of the problems, in particular the possibility of considering various upper derivatives, deserves a separate consideration. We recall that to each of the δ-approximating derived sets of a real-valued locally Lipschitz function f on a Banach space Y there corresponds an upper derivative by

$$\overline{D}f(y,v)=\sup\mathcal{D}f(y,v),\qquad \overline{B}f(y,v)=\sup\mathcal{Z}f(y,v),$$
$$\overline{Z}f(y,v)=\sup\mathcal{P}f(y,v),\qquad \overline{C}f(y,v)=\sup\mathcal{C}f(y,v).$$

We will use the names upper (Dini) derivative, Zahorski's upper derivative, Michel–Penot's (or Michel-Penot) upper derivative, and strict (or Clarke) upper derivative, respectively.

It is easy to see directly or via the fact that $\mathcal{D}_\delta f(y,v)$ converge to $\mathcal{D}f(y,v)$ in the Hausdorff metric that $\overline{D}f(y,v)=\lim_{\delta\searrow 0}\sup\mathcal{D}_\delta f(y,v)$, which is clearly the same as the more usual definition from the introduction. Similarly, the definitions of the remaining upper derivatives may be transformed to their more familiar forms.

Taking suprema in 2.15, we immediately get the result of [21] for strict upper derivative and its counterpart for other upper derivatives.

3.1. **Proposition.** *Let f be a real-valued locally Lipschitz function on an open subset H of a separable Banach space Y.*

(i) *There is a σ-directionally porous set $Q_{\mathcal{Z}}\subset Y$ such that $\overline{B}f(y,v)=\overline{D}f(y,v)$ for every $y\in H\setminus Q_{\mathcal{Z}}$ and every $v\in Y$.*
(ii) *There is a σ-porous set $Q_{\mathcal{P}}\subset Y$ such that $\overline{Z}f(y,v)=\overline{D}f(y,v)$ for every $y\in H\setminus Q_{\mathcal{P}}$ and every $v\in Y$.*
(iii) *There is a meager set $Q_{\mathcal{C}}\subset Y$ such that $\overline{C}f(y,v)=\overline{D}f(y,v)$ for every $y\in H\setminus Q_{\mathcal{C}}$ and every $v\in Y$.*

Symmetry of Dini derivatives of real-valued functions on separable spaces follows immediately from 2.16, or from (i) above.

3.2. **Theorem.** *Let f be a real-valued locally Lipschitz function on an open subset H of a separable Banach space Y. Then there is a σ-directionally porous set $Q\subset Y$ such that $\overline{D}f(y,v)=-\underline{D}f(y,-v)$ for every $y\in H\setminus Q$ and every $v\in Y$.*

Having recapitulated results following from the general investigations in Section 2, we come to the main theme of this section. In addition to proving, on a separable Banach space Y, convexity of the upper derivative except for a σ-directionally porous set, we also address a related question: Can one define a "new" upper derivative which would be convex at every point and which would coincide with the upper derivative except a "very small" (e.g., σ-directionally porous) set? (This is just the first step; ultimately, we wish to have a notion that would allow natural

formulations of the chain rule.) The strict upper derivative is clearly not the answer, as is shown by many well known examples (including the one in 2.20(iii)). The Zahorski's upper derivative coincides with the upper derivative except for a σ-directionally porous set, but easy examples show that it need not be convex at every point, and the Michel–Penot's upper derivative may differ from the upper derivative in a too large set, as shown by the example mentioned in 2.20(ii). In the case of Michel–Penot's derivative this example may not seem to be convincing, since we still get an agreement except for a σ-porous set. However, in connection with the chain rule formula, σ-porous sets may turn out to be too large, and we will see in Example 6.6 that this upper derivative indeed does not lead to a satisfactory chain rule formula. Since the problem seems to be that these upper derivatives may attain values much larger than the Dini derivatives, a natural candidate for the notion of upper derivative that would satisfy our requirements is the smallest convexification of upper derivatives. In similar context this is usually presented in dual form (and in the language of subdifferentials), and so we first recall its construction.

3.3. **Lemma.** *Let p be a Lipschitz positively homogeneous function on Y. Then the function*

$$\tilde{p}(u) = \sup_{v\in Y}(p(u+v) - p(v))$$

is well-defined, Lipschitz, positively homogeneous, subadditive, and therefore convex.

Proof. Since $p(u+v) - p(v) \le C\|u\|$, where C is the Lipschitz constant of p, $\tilde{p}$ is well-defined. If $t > 0$, we use the positive homogeneity of p to infer that $\tilde{p}(u) = \sup_{v\in Y}(p(u+v) - p(v)) = t\sup_{v\in Y}(p(u/t + v/t) - p(v/t)) = t\tilde{p}(u/t)$. If $u_1, u_2 \in Y$ and $\varepsilon > 0$, we find $v \in Y$ such that $p(u_1+u_2+v) - p(v) > \tilde{p}(u_1+u_2) - \varepsilon$. Then $\tilde{p}(u_1+u_2) - \varepsilon < p(u_1+u_2+v) - p(v) = (p(u_1+u_2+v) - p(u_2+v)) + (p(u_2+v) - p(v)) \le \tilde{p}(u_1) + \tilde{p}(u_2)$. Finally, the subadditivity of $\tilde{p}$ implies that $\tilde{p}(u_1) - \tilde{p}(u_2) \le \tilde{p}(u_1 - u_2) \le C\|u_1 - u_2\|$, so $\tilde{p}$ is Lipschitz. □

If f is a real-valued function on a Banach space Y which is Lipschitz on a neighborhood of a point $y \in Y$, we define

$$\widetilde{D}f(y,u) = \sup_{v\in Y}\big(\overline{D}f(y,u+v) - \overline{D}f(y,v)\big).$$

3.4. **Lemma.** *Let f be a locally Lipschitz real-valued function on an open subset H of a Banach space Y. Then*

(i) *for every $y \in H$ the function $u \mapsto \widetilde{D}f(y,u)$ is Lipschitz, convex, positively homogeneous, and subadditive,*
(ii) *if f is Gâteaux differentiable at $y \in H$, then $\widetilde{D}f(y,u) = f'(y;\,u)$ for every $u \in Y$,*
(iii) *f is Gâteaux differentiable at $y \in H$ if and only if $\widetilde{D}f(y,u) = -\widetilde{D}(-f)(y,u)$ for every $u \in Y$,*
(iv) *if Y is separable, then the mapping $(y,u) \mapsto \widetilde{D}f(y,u)$ is Borel measurable.*

Proof. (i) Follows from Lemma 3.3.

(ii) If f is Gâteaux differentiable at y, then $\overline{D}f(y,w) = f'(y;\,w)$ for all $w \in Y$. Therefore,

$$\widetilde{D}f(y,u) = \sup_{v\in Y}(\overline{D}f(y,u+v) - \overline{D}f(y,v)) = \sup_{v\in Y}(f'(y;\,u+v) - f'(y;\,v)) = f'(y;\,u).$$

(iii) If f is Gâteaux differentiable at y, this follows from (ii) used for functions $\pm f$. For the opposite direction, observe that $u \to \widetilde{D}(\pm f)(y,u)$ are both convex and Lipschitz, so the assumed equality implies that they are linear and continuous. Since the upper and lower derivatives of f at y are between these two values, $f'(y;\, u)$ exists and is equal to $\widetilde{D}f(y,u)$.

(iv) Let V be a countable dense subset of Y. Consider any real number a. Then

$$\begin{aligned}\{(y,u) : \widetilde{D}f(y,u) \le a\} &= \{(y,u) : \overline{D}f(y,u+v) - \overline{D}f(y,v) \le a \ \forall v \in Y\} \\ &= \bigcap_{v\in V} \{(y,u) : \overline{D}f(y,u+v) - \overline{D}f(y,v) \le a\}.\end{aligned}$$

Hence it suffices to show that for each fixed $v \in Y$ and $\lambda = 0, 1$, the function $h_\lambda(y,u) = \overline{D}f(y, \lambda u + v)$ is Borel measurable. But this follows by observing that for any fixed $b \in \mathbb{R}$, the set $\{(y,u) : \overline{D}f(y,\lambda u + v) > b\}$ is

$$\bigcup_{n=1}^{\infty} \bigcap_{m=1}^{\infty} \bigcup_{t\in(0,1/m)\cap\mathbb{Q}} \Big\{(y,u) : \frac{f(y+t(\lambda u+v)) - f(y)}{t} > b + \frac{1}{n}\Big\},$$

and the sets involved in this expression are open. □

3.5. **Proposition.** *Let f be a real-valued locally Lipschitz function on an open subset H of a separable Banach space Y. Then there is a σ-directionally porous set $Q \subset Y$ such that $\widetilde{D}f(y,v) = \overline{D}f(y,v)$ for every $y \in H \setminus Q$ and every $v \in Y$.*

More generally, if G is an open subset of a separable Banach space X and $g : G \to H$ is locally Lipschitz, then there is a σ-directionally porous set $P \subset X$ such that $\widetilde{D}f(g(x), g'(x;\, e)) = \overline{D}f(g(x), g'(x;\, e))$ whenever $x \in G \setminus P$ and $e \in X$ is such that $g'(x;\, e)$ exists.

Proof. To prove the second statement, we use 2.24 to find a σ-directionally porous set $P \subset X$ such that

$$\mathcal{D}f(g(x), g'(x;\, e) + w) \subset \mathcal{D}f(g(x), g'(x;\, e)) + \mathcal{D}f(g(x), w)$$

whenever $x \in G \setminus P$, $e \in X$ is such that $g'(x;\, e)$ exists, and $w \in Y$. Taking suprema, we infer that for any such x, e and w,

$$\overline{D}f(g(x), g'(x;\, e) + w) \le \overline{D}f(g(x), g'(x;\, e)) + \overline{D}f(g(x), w).$$

Hence $\widetilde{D}f(g(x), g'(x;\, e)) = \overline{D}f(g(x), g'(x;\, e))$, as claimed. The first statement follows from the second with $X = Y$, $G = H$ and g the identity. □

Convexity of upper derivatives follows immediately from Proposition 3.5 and Lemma 3.4(i).

3.6. **Proposition.** *Let f be a real-valued locally Lipschitz function on an open subset H of a separable Banach space Y. Then there is a σ-directionally porous subset Q of Y such that for every $y \in H \setminus Q$ the function $v \mapsto \overline{D}f(y,v)$ is convex on Y.*

We also remark that Theorem 3.6 and Theorem 3.2 easily give that, for a real-valued locally Lipschitz function f on a separable Banach space Y the set of points y at which the directions of differentiability do not form a linear subspace of Y is σ-directionally porous and that, consequently, with exception of points of a σ-directionally porous set Gâteaux differentiability of f may be deduced from its

directional differentiability in a spanning set of directions. Since we prove more in 8.1 and 8.2 we omit the (simple) details.

4. Chain rule formula

As explained in the introduction, the idea behind our notion of "generalized derivatives" or, as we will call them, "derivative assignments" is simple: given a Lipschitz mapping f of a separable Banach space Y to a (separable) Banach space Z, we choose for each $y \in Y$ a closed linear subspace $U(f, y)$ of Y such that $f'(y; u)$ exists for each $u \in U(f, y)$ and depends linearly on u. The generalized derivative is then defined as the map $u \in U(f, y) \to f'(y; u)$. Thus the simplest, though not very useful, example of a derivative assignment is to assign to every y the trivial subspace $\{0\}$. The assignment $U(f, y) = Y$ if f is Gâteaux differentiable at y and $U(f, y) = \{0\}$ otherwise is also not useful for obtaining a chain rule: since our problem is precisely that the image of g may lie in an $\mathcal{L}$ null set of points at which f is not differentiable, we need to assign non-trivial spaces also to (some) such points. Notice however, that for a nowhere differentiable Lipschitz function $f : \mathbb{R} \to Z$ (which exists provided Z fails to have the Radon-Nikodym property) these two assignments coincide and in fact such an f has only this trivial derivative assignment.

We provide more substantial examples of derivative assignments in the next section. Before that we give precise definitions of these assignments and of the notion of their completeness. The notion of completeness is exactly what we need in order to obtain a natural formulation of the chain rule formula, and we therefore immediately state and prove chain rules involving complete derivative assignments, although their existence will be established only in the next section.

4.1. **Definition.** Let Y, Z be separable Banach spaces and $f : Y \mapsto Z$.

If U is a linear subspace of Y and $y \in Y$, we say that f is *Gâteaux differentiable at y in the direction of U* if there is a continuous linear mapping $L(y) : U \to Z$, which is termed the *Gâteaux derivative of f at y in the direction of U* such that $\lim_{t\to 0}(f(y + tu) - f(y))/t = L(y)(u)$ for every $u \in U$.

A *derivative assignment* for f assigns to every $y \in Y$ the Gâteaux derivative $f^{\blacktriangledown}(y)$ of f at y in the direction of some closed linear subspace $U(f, y)$ of Y. It is often more convenient to speak about the assignment $U(f, y)$; this is justified by observing that $f^{\blacktriangledown}(y)$ is uniquely determined by $U(f, y)$.

The derivative assignment $y \in Y \mapsto f^{\blacktriangledown}(y)$ is said to be *complete* if for every separable Banach space X and every Lipschitz mapping $g : X \mapsto Y$ there is an $\mathcal{L}$ null set $N \subset X$ such that $g'(x; e)$ belongs to the domain of $f^{\blacktriangledown}(g(x))$ whenever $x \in X \setminus N$, $e \in X$, and $g'(x; e)$ exists.

A *complete derivative* of f assigns to every $y \in Y$ a continuous linear mapping $f^{\S}(y) : Y \to Z$ in such a way that for every separable Banach space X and every Lipschitz mapping $g : X \mapsto Y$ there is an $\mathcal{L}$ null set $N \subset X$ such that $f'(g(x); g'(x; e))$ exists and is equal to $f^{\S}(g(x))(g'(x; e))$ whenever $x \in X \setminus N$, $e \in X$, and $g'(x; e)$ exists. In particular, a complete derivative may be obtained from a complete derivative assignment $f^{\blacktriangledown}$ by extending, for each $y \in Y$, $f^{\blacktriangledown}(y)$ to a linear operator $f^{\S}(y)$ defined on the whole Y, provided all these extensions exist. Conversely, every complete derivative $f^{\S}$ is obtained in this way provided that f has a complete derivative assignment, since if $U(f, y)$ is a complete derivative

assignment then $\{u \in U(f,y) : f'(y;u) = f^{\S}(y)(u)\}$ is also a complete derivative assignment.

The importance of complete derivative assignments stems from the validity of natural statements of the chain rule formula. We first state and prove it as a rule for finding a complete derivative assignment for a composition of functions with given complete derivative assignments.

4.2. **Theorem.** *Suppose that X, Y, Z are separable Banach spaces and $g : X \to Y$ and $f : Y \to Z$ are locally Lipschitz functions having complete derivative assignments $x \in X \to g^{\blacktriangledown}(x)$ and $y \in Y \to f^{\blacktriangledown}(y)$. Then*

$$x \in X \to f^{\blacktriangledown}(g(x)) \circ g^{\blacktriangledown}(x), \tag{4.1}$$

is a complete derivative assignment for $f \circ g : X \to Z$.

Proof. We first notice that the composition of derivative assignments is a derivative assignment (without assuming completeness): the domain of $f^{\blacktriangledown}(g(x)) \circ g^{\blacktriangledown}(x)$ is the closed linear subspace $U_x := \{e \in U(g,x) : g^{\blacktriangledown}(x)(e) \in U(f, g(x))\}$, the composition $f^{\blacktriangledown}(g(x)) \circ g^{\blacktriangledown}(x)$ is continuous and linear on its domain and by the chain rule for Gâteaux derivatives it is the Gâteaux derivative of $f \circ g$ in the direction of U_x.

Let W be a separable Banach space and $h : W \to X$ a Lipschitz map. To prove completeness of our assignment, we have to find an $\mathcal{L}$ null set $N \subset W$ such that $h'(w; e)$ belongs to $U_{h(w)}$ whenever $w \in W \setminus N$, $e \in W$, and $h'(w; e)$ exists.

By completeness of the derivative assignment $g^{\blacktriangledown}$ there is an $\mathcal{L}$ null set $N_1 \subset W$ such that $h'(w; e)$ belongs to the domain of $g^{\blacktriangledown}(h(w))$ whenever $w \in W \setminus N_1$, $e \in W$, and $h'(w; e)$ exists. Similarly, by completeness of the derivative assignment $f^{\blacktriangledown}$ there is and $\mathcal{L}$ null set $N_2 \subset W$ such that $(g \circ h)'(w; e)$ belongs to the domain of $f^{\blacktriangledown}((g \circ h)(w))$ whenever $w \in W \setminus N_2$, $e \in W$, and $(g \circ h)'(w; e)$ exists.

Let $N = N_1 \cup N_2$ and suppose that $w \in W \setminus N$ and $e \in W$ are such that $h'(w; e)$ exists. Since $w \notin N_1$, $h'(w; e)$ belongs to the domain of $g^{\blacktriangledown}(h(w))$ and the chain rule for the composition $g \circ h$ in direction of e implies that $(g \circ h)'(w; e)$ exists and $(g \circ h)'(w; e) = g'(h(w); h'(w; e)) = g^{\blacktriangledown}(h(w))(h'(w; e))$. Since $w \in X \setminus N_2$, it follows that $(g \circ h)'(w; e)$ and so $g^{\blacktriangledown}(h(w))(h'(w; e))$ belongs to the domain of $f^{\blacktriangledown}(g(h(w)))$. Hence $h'(w; e) \in U_{h(w)}$, as required. □

The notion of derivative assignments has been chosen so that this rule immediately extends to a composition of any number of functions and to complete derivatives.

4.3. **Theorem.** *Suppose that X_i, $i = 1, \dots, n+1$, are separable Banach spaces and $g_i : X_i \to X_{i+1}$, $i = 1, \dots, n$, are Lipschitz functions.*

(i) *If $x \in X_i \mapsto g_i^{\blacktriangledown}(x)$, $i = 1, \dots, n$, are complete derivative assignments, then*

$$x_1 \in X_1 \mapsto g_n^{\blacktriangledown}(x_n) \circ \cdots \circ g_1^{\blacktriangledown}(x_1),$$

where $x_i = g_{i-1}(x_{i-1})$, is a complete derivative assignment for $g_n \circ \cdots \circ g_1$.

(ii) *If $x \in X_i \mapsto g_i^{\S}(x)$, $i = 1, \dots, n$, are complete derivatives, then*

$$x_1 \in X_1 \mapsto g_n^{\S}(x_n) \circ \cdots \circ g_1^{\S}(x_1)$$

is a complete derivative of $g_n \circ \cdots \circ g_1$.

These chain rules easily imply the chain rules for finding the Gâteaux derivative of a composition.

4.4. **Theorem.** *Suppose X_i, $i = 0, \dots, n+1$, are separable Banach spaces, and $g_i : X_i \to X_{i+1}$, $i = 0, \dots, n$ are Lipschitz functions such that g_i, $i = 1, \dots, n$, have complete derivative assignments $x \in X_i \mapsto g_i^{\blacktriangledown}(x)$. Then for $\mathcal{L}$ almost every $x_0 \in X_0$ at which g_0 is Gâteaux differentiable, the composition $g_n \circ \cdots \circ g_1 \circ g_0$ is Gâteaux differentiable at x_0 and*

$$(g_n \circ \cdots \circ g_1 \circ g_0)'(x_0) = g_n^{\blacktriangledown}(x_n) \circ \cdots \circ g_1^{\blacktriangledown}(x_1) \circ g_0'(x_0), \tag{4.2}$$

where $x_i = g_{i-1}(x_{i-1})$. This, in particular, means that for each $u_0 \in X_0$ the direction of $u_1 = g_0'(x_0)(u_0)$ belongs to the domain of $g_1^{\blacktriangledown}(x_1)$, the direction of $u_2 = g_1^{\blacktriangledown}(x_1)(u_1)$ belongs to the domain of $g_2^{\blacktriangledown}(x_2)$, etc.

Similarly, if $x \in X_i \mapsto g_i^{\S}(x)$, $i = 1, \dots, n$, are complete derivatives, then

$$(g_n \circ \cdots \circ g_0)'(x_0) = g_n^{\S}(x_n) \circ \cdots \circ g_1^{\S}(x_1) \circ g_0'(x_0). \tag{4.3}$$

for $\mathcal{L}$ almost every $x_0 \in X_0$ at which g_0 is Gâteaux differentiable.

Proof. Since by Theorem 4.3, $g_n^{\blacktriangledown}(x_n) \circ \cdots \circ g_1^{\blacktriangledown}(x_1)$ is a complete derivative assignment for $g_n \circ \cdots \circ g_1$, for $\mathcal{L}$ almost every $x_0 \in X_0$ at which g_0 is Gâteaux differentiable the image of $g_0'(x)$ lies in the domain of $g_n^{\blacktriangledown}(x_n) \circ \cdots \circ g_1^{\blacktriangledown}(x_1)$ in the direction of which $g_n \circ \cdots \circ g_1$ is Gâteaux differentiable with derivative $g_n^{\blacktriangledown}(x_n) \circ \cdots \circ g_1^{\blacktriangledown}(x_1)$. Hence the first statement follows from the chain rule for Gâteaux derivatives and the second statement is an immediate consequence of the first. □

4.5. *Remark.* If X_1 has the Radon-Nikodym property, then g_0 is Gâteaux differentiable $\mathcal{L}$ almost everywhere, and hence the chain rule formulas (4.2) and (4.3) hold for $\mathcal{L}$ almost every x_0.

5. Existence and measurability of complete derivative assignments

To show that the results of the previous section give non-trivial information, we construct a number of complete derivative assignments and establish their measurability properties. We will see in Proposition 5.2 that all assignments defined in the following proposition are complete provided Z has the Radon-Nikodym property. The assumption that Z has the Radon-Nikodym property is natural, since testing completeness of a derivative assignment with $X = Y$ and g the identity, we see that for a complete assignment to exist f has to be Gâteaux differentiable $\mathcal{L}$ almost everywhere. However, it does not seem reasonable to require the Radon-Nikodym property already in the definition of derivative assignments; the assignment depends on f, and there are maps between spaces without the Radon-Nikodym property for which a complete derivative assignment exists. (In this connection, see Remark 5.3.)

5.1. **Proposition.** *For any Lipschitz mapping f of a separable Banach space Y to a separable Banach space Z each of the following defines a derivative assignment.*

(i) *To each $y \in Y$ we may assign a maximal subspace $U(f, y)$ in the direction of which f is Gâteaux differentiable at y.*
(ii) *For each $y \in Y$ we define $U_a(f, y)$ as the set of those directions $u \in Y$ such that $f'(y; u)$ exists and such that, whenever $f'(y; v)$ exists, then $f'(y; u+v)$ does and $f'(y; u+v) = f'(y; u) + f'(y; v)$.*
(iii) *If $Z = \mathbb{R}$, we define $U_t(f, y)$ as the set of all directions $u \in Y$ such that $\widetilde{D}f(y, u) = -\widetilde{D}(f)(y, -u) = -\widetilde{D}(-f)(y, u) = \widetilde{D}(-f)(y, -u)$.*

(iv) *If Z is finite dimensional, we define $U_c(f,y)$ as the set of all directions $u \in Y$ such that $f'(y;\, u)$ exists and $\mathcal{D}f(y, \pm u + v) \subset f'(y;\, \pm u) + \mathcal{D}f(y,v)$ for every $v \in Y$.*

(v) *With the radius $r_z(S)$ of a set $S \in \mathcal{H}(Z)$ about a point $z \in Z$ defined as the infimum of those $r > 0$ for which $S \subset B(z,r)$, we let $U_r(f,y)$ be the set of those directions $u \in Y$ for which $f'(y;\, u)$ exists and which have the property that for every $\varepsilon > 0$, $v \in Y$, and $z \in Z$ there is $\delta_0 > 0$ such that for every $0 < \delta < \delta_0$,*

$$\mathcal{D}_\delta f(y, \pm u + v) \subset f'(y;\, \pm u) + B\big(z, r_z(\mathcal{D}_\delta f(y,v)) + \varepsilon\big).$$

(vi) *The previous example may be generalized in the following way. Suppose that $\Phi \subset \mathrm{Lip}_1(Z)$ is separable in the topology of uniform convergence on bounded sets, and satisfies the condition that for every $z \in Z$ and $\tau > 0$ there are $\phi \in \Phi$ and $\varepsilon > 0$ such that $B_\phi(z;\, \varepsilon) \subset B(z, \tau)$. We define $U_\Phi(f,y)$ as the set of those directions $u \in Y$ for which $f'(y;\, u)$ exists and which have the property that for every $\varepsilon > 0$, $v \in Y$, and $\phi \in \Phi$, there is $\delta_0 > 0$ such that for every $0 < \delta < \delta_0$*

$$\mathcal{D}_\delta f(y, \pm u + v) \subset f'(y;\, \pm u) + B_\phi(\mathcal{D}_\delta f(y,v);\, \varepsilon).$$

Moreover,

(a) *$U_a(f,y)$ is the intersection of all maximal subspaces U of Y in the direction of which f is Gâteaux differentiable at y; in particular $U(f,y) \supset U_a(f,y)$ for any choice of $U(f,y)$ in (i);*

(b) *$U_a(f,y) \supset U_r(f,y)$;*

(c) *$U_r(f,y)$ coincides with $U_{\Phi_r}(f,y)$ where $\Phi_r := \{\|\cdot - z\| : z \in Y\}$;*

(d) *if Z is finite dimensional, $U_c(f,y)$ coincides with $U_\Phi(f,y)$ for $\Phi = \mathrm{Lip}_1(Z)$;*

(e) *if $Z = \mathbb{R}$, $U_t(f,y)$, $U_r(f,y)$, $U_c(f,y)$ all coincide;*

Proof. (i) Since f is Lipschitz, Lemma 2.3 implies that every subspace in the direction of which f is Gâteaux differentiable at a point y, is contained in a maximal subspace having the same property, and that such maximal subspaces are closed. Hence $U(f,y)$ is a closed linear subspace of Y in the direction of which f is Gâteaux differentiable.

(a) If f is Gâteaux differentiable at y in the direction of a subspace U and $u \in U_a(f,y)$, it is Gâteaux differentiable at y in the direction of the linear span of $U \cup \{u\}$. If U is maximal, it follow that it contains u. Conversely, if $u \notin U_a(f,y)$ then either $f'(y;u)$ does not exist and u does not belong to any $U(f,y)$, or $f'(y;u)$ exists and there is v such that $f'(y;v)$ exists and f is not Gâteaux differentiable in the direction of the linear span of $\{u,v\}$. Taking for $U(f,a)$ a maximal subspace of Y containing v in the direction of which f is Gâteaux differentiable, we have $u \notin U(f,y)$.

(ii) By (i) and (a), $U_a(f,y)$ is a closed subspace of Y in the direction of which f is Gâteaux differentiable.

(b) It suffices to use the definition of $U_r(f,y)$ with $z = f'(y,v)$ and observe that, if $f'(y,v)$ exists, $r_z(\mathcal{D}_\delta f(y,v)) \to 0$ as $\delta \to 0$.

(c) The definition of $U_r(f,y)$ is a slightly more geometric description of the special case of (vi) in which Φ is the collection of functions $\phi_z(w) = \|w - z\|$.

(d) If $u \in U_c(f,y)$, $v \in Y$ and $\varepsilon > 0$, we use the finite dimensionality of Z to infer that there is $\delta_0 > 0$ such that $\mathcal{D}_\delta f(y, u+v) \subset B(\mathcal{D}f(y,u+v), \varepsilon)$ and

$\mathcal{D}f(y,v) \subset B(\mathcal{D}_\delta f(y,v),\varepsilon)$ for every $0<\delta<\delta_0$. Hence

$$\mathcal{D}_\delta f(y,u+v) \subset f'(y;\,u) + B(\mathcal{D}_\delta f(y,v),2\varepsilon),$$

and the requirement of (vi) holds for every $\phi \in \mathrm{Lip}_1(Z)$.

If $u \in U_{\mathrm{Lip}_1(Z)}(f,y)$ and $v \in Y$, we get that for $\phi(z) := \mathrm{dist}(z,\mathcal{D}f(y,v))$ and any $\varepsilon>0$ there is $\delta_0>0$ so that $\mathcal{D}_\delta f(y,u+v) \subset f'(y;\,u) + B_\phi(\mathcal{D}_\delta f(y,v);\,\varepsilon)$ for $0<\delta<\delta_0$. Taking limit as $\delta \searrow 0$, we get by Example 2.8(i)

$$\mathcal{D}f(y,u+v) \subset f'(y;\,u) + B(\mathcal{D}f(y,v);\,2\varepsilon),$$

which, since $\varepsilon>0$ is arbitrary, gives $\mathcal{D}f(y,u+v) \subset f'(y;\,u) + \mathcal{D}f(y,v)$.

The above arguments apply also to $-u$ instead of u, and so $U_c(f,y) = U_\Phi(f,y)$.

(e) We show $U_t(f,y) \subset U_r(f,y) \subset U_c(f,y) \subset U_t(f,y)$.

Suppose $u \in U_t(f,y)$. Using the inequality $\widetilde{D}g(y,u) \geq \overline{D}g(y,u)$ with $g=f$ and $g=-f$, we infer that $f'_+(y;\,u)$ exists and is equal to $\widetilde{D}f(y,u) = -\widetilde{D}(-f)(y,u)$. Together with the same argument used in the direction of $-u$ we get that $f'(y;\,u)$ exists. Let $\varepsilon>0$, $v \in Y$, and $z \in Z$. Since

$$\overline{D}f(y,u+v) \leq \widetilde{D}f(y,u) + \overline{D}f(y,v) = f'(y;\,u) + \overline{D}f(y,v)$$

and

$$\underline{D}f(y,u+v) \geq -\widetilde{D}(-f)(y,u) - \underline{D}f(y,v) = f'(y;\,u) - \underline{D}f(y,v),$$

we infer that for sufficiently small $\delta>0$,

$$\begin{aligned}\mathcal{D}_\delta f(y,u+v) &\subset [f'(y;\,u) + \overline{D}f(y,v) - \varepsilon, f'(y;\,u) - \underline{D}f(y,v) + \varepsilon] \\ &\subset f'(y;\,u) + B\big(z, r_z(\mathcal{D}_\delta f(y,v)) + \varepsilon\big),\end{aligned}$$

where $\underline{D}f(y,w) = \inf \mathcal{D}f(y,w)$ for $y,w \in Y$. This and a similar argument for $-u$ show that the requirements of the definition of $u \in U_r(f,y)$ hold.

Suppose $u \in U_r(f,y)$, $\varepsilon>0$, $v \in Y$ and $\phi \in \mathrm{Lip}_1(Z)$. By definition of $U_r(f,y)$ with $z = \frac{1}{2}(\underline{D}f(y,v) + \overline{D}f(y,v))$ there is $\delta_0>0$ such that for every $0<\delta<\delta_0$,

$$\mathcal{D}_\delta f(y,u+v) \subset [f'(y;\,u) + \underline{D}f(y,v) - \varepsilon, f'(y;\,u) + \overline{D}f(y,v) + \varepsilon].$$

Since $\mathcal{D}f(y,v) \supset [\underline{D}f(y,v), \overline{D}f(y,v)]$, $\mathcal{D}_\delta f(y,u+v) \subset f'(y;\,u) + B(\mathcal{D}f(y,v),\varepsilon)$ for $0<\delta<\delta_0$. Consequently, $\mathcal{D}f(y,u+v) \subset f'(y;\,u) + B(\mathcal{D}f(y,v),\varepsilon)$, which, together with a similar inclusion for $-u$ shows that $u \in U_c(f,y)$.

Finally, suppose $u \in U_c(f,y)$. Given $\varepsilon>0$ find $v \in Y$ so that

$$\widetilde{D}f(y,u) < \overline{D}f(y,u+v) - \overline{D}f(y,v) + \varepsilon.$$

By the definition of $U_c(f,y)$, the right side is at most $f'(y;\,u)+\varepsilon$, and we infer that $f'(y;\,u) \leq \widetilde{D}f(y,u) < f'(y;\,u) + \varepsilon$. Hence $\widetilde{D}f(y,u) = f'(y;\,u)$. Similarly we show that $-\widetilde{D}(f)(y,-u)$, $-\widetilde{D}(-f)(y,u)$ and $\widetilde{D}(-f)(y,-u)$ are all equal to $f'(y;\,u)$.

(iii–vi) It remains to show that $U_\Phi(f,y)$ is a closed linear space in the direction of which f is Gâteaux differentiable. Suppose $u,v \in U_\Phi(f,y)$. We first show that

$$f'(y;\,u+v) \text{ exists and is equal to } f'(y;\,u) + f'(y;\,v). \tag{5.1}$$

Given $\tau>0$, choose $\phi \in \Phi$ and $\varepsilon>0$ such that $B_\phi(f'(y;\,v);\,2\varepsilon) \subset B(f'(y;\,v),\tau)$. By assumption, there is $\delta_0>0$ such that for every $0<\delta<\delta_0$,

$$\mathcal{D}_\delta f(y,u+v) \subset f'(y;\,u) + B_\phi(\mathcal{D}_\delta f(y,v);\,\varepsilon).$$

Since $f'(y;\, v)$ exists, $\mathcal{D}_{\delta_0} f(y, v) \subset B(f'(y;\, v);\, \varepsilon)$ provided δ_0 is small enough, and then for any $0 < \delta < \delta_0$,

$$\mathcal{D}_\delta f(y, u + v) \subset f'(y;\, u) + B_\phi(B(f'(y;\, v);\, \varepsilon);\, \varepsilon) \subset B(f'(y;\, u) + f'(y;\, v);\, \tau).$$

Hence $f'_+(y;\, u+v)$ exists and is equal to $f'(y;\, u)+f'(y;\, v)$, and the same argument in the direction of $-u$ shows (5.1).

Next we show that $u + v \in U_\Phi(f, y)$. Given $\varepsilon > 0$ and $w \in Y$, find $\delta_0 > 0$ such that for every $0 < \delta < \delta_0$,

$$\mathcal{D}_\delta f(y, u + (v + w)) \subset f'(y;\, u) + B_\phi(\mathcal{D}_\delta f(y, v + w);\, \varepsilon/2)$$

and

$$\mathcal{D}_\delta f(y, v + w) \subset f'(y;\, v) + B_\phi(\mathcal{D}_\delta f(y, w);\, \varepsilon/2).$$

Hence

$$\begin{aligned}\mathcal{D}_\delta f(y, u + v + w) &\subset f'(y;\, u) + B_\phi(\mathcal{D}_\delta f(y, v + w);\, \varepsilon/2)\\ &\subset f'(y;\, u) + f'(y;\, v) + B_\phi(\mathcal{D}_\delta f(y, w);\, \varepsilon)\\ &= f'(y;\, u + v) + B_\phi(\mathcal{D}_\delta f(y, w);\, \varepsilon).\end{aligned}$$

This and a similar argument for $-u$ show that, indeed, $u + v \in U_\Phi(f, y)$.

Finally, suppose $\tilde{u} \in \overline{U_\Phi(f, y)}$ and $v \in Y$. Given $\varepsilon > 0$, find $u \in U_\Phi(f, y)$ such that $\|u - \tilde{u}\| < \varepsilon$, and for this u we find $\delta_0 > 0$ such that for every $0 < \delta < \delta_0$,

$$\mathcal{D}_\delta f(y, u + v) \subset f'(y;\, u) + B_\phi(\mathcal{D}_\delta f(y, v);\, \varepsilon).$$

Then by Lemma 2.3,

$$\begin{aligned}\mathcal{D}_\delta f(y, \tilde{u} + v) &\subset B(\mathcal{D}_\delta f(y, u + v), \mathrm{Lip}(f)\varepsilon)\\ &\subset B(f'(y;\, u) + B_\phi(\mathcal{D}_\delta f(y, v);\, \varepsilon), \mathrm{Lip}(f)\varepsilon)\\ &\subset B(f'(y;\, \tilde{u}), 2\mathrm{Lip}(f)\varepsilon) + B(B_\phi(\mathcal{D}_\delta f(y, v);\, \varepsilon), \mathrm{Lip}(f)\varepsilon)\\ &\subset f'(y;\, \tilde{u}) + B_\phi(\mathcal{D}_\delta f(y, v);\, (3\mathrm{Lip}(f) + 1)\varepsilon).\end{aligned}$$

This and a similar argument for $-u$ show $\tilde{u} \in U_\Phi(f, y)$, and so finish the proof. □

5.2. **Proposition.** *If in Proposition 5.1 the space Z has the Radon-Nikodym property, each of the derivative assignments constructed there is complete.*

Proof. By the additional statement in Proposition 5.1 we just need to consider the assignment $U_\Phi(f, y)$ defined in 5.1(vi). To prove its completeness, let X be a separable Banach space and $g : X \to Y$ a Lipschitz map.

We first check that there is an $\mathcal{L}$ null set $Q \subset X$ such that $f'(g(x);\, g'(x;\, e))$ exists and is equal to $(f \circ g)'(x;\, e)$ whenever $x \in X \setminus Q$, $e \in X$, and $g'(x;\, e)$ exists. Indeed, since Z has the Radon-Nikodym property, and the mapping $f \circ g : X \mapsto Z$ is Lipschitz, the set N of points where $f \circ g$ is not Gâteaux differentiable, is $\mathcal{L}$ null. For all $x \in X \setminus N$ and all $e \in X$, Proposition 2.5 shows that

$$\lim_{\delta \searrow 0} \varrho(\mathcal{D}_\delta(f \circ g)(x, e), \mathcal{D}_\delta f(g(x), g'_+(x;\, e))) = 0.$$

Since $\mathcal{D}_\delta(f \circ g)(x, e)$ converges to $(f \circ g)'_+(x;\, e)$ in the Hausdorff metric, and since this argument can be used also in the direction of $-e$, this shows that $f'(g(x);\, g'(x;\, e))$ exists and is equal to $(f \circ g)'(x;\, e)$, as required.

We are now ready to finish the proof of completeness of U_Φ. Let $v \in Y$ and $\varepsilon > 0$. By Corollary 2.23 there exists a σ-directionally porous set $P \subset X$ such that

for every $\varepsilon > 0$, $x \in X \setminus P$, $v \in Y$ and for every direction $e \in X$ at which g is differentiable one can find $\delta_0 > 0$ such that for every $0 < \delta < \delta_0$,

$$\mathcal{D}_\delta f(g(x), g'_+(x;\, e) + v) \subset \mathcal{D}_\delta f(g(x), g'_+(x;\, e)) + B_\phi(\mathcal{D}_\delta f(g(x), v);\, \varepsilon). \tag{5.2}$$

If $x \in X \setminus (N \cup P)$ and $e \in X$ are such that $g'(x;\, e)$ exists, we may assume δ_0 has been chosen small enough to guarantee $\mathcal{D}_\delta f(g(x), g'_+(x;\, e)) \subset B(f'(g(x);\, g'(x;\, e)), \varepsilon)$. Hence (5.2) implies that

$$\mathcal{D}_\delta f(g(x), g'(x;\, e) + v) \subset f'(g(x);\, g'(x;\, e)) + B_\phi(\mathcal{D}_\delta f(g(x), v);\, 2\varepsilon).$$

The same argument may be used in the direction of $-e$, and we conclude that indeed $g'(x;\, e) \in U_\Phi(f, g(x))$. □

5.3. *Remark.* This proof shows that the assumption that the Banach space Z has the Radon-Nikodym property may be replaced by a weaker one, which depends not on the quality of the space Z but of the function f: we only need that for every separable Banach space X and every Lipschitz mapping $g : X \mapsto Y$ there is an $\mathcal{L}$ null set $N \subset X$ such that $(f \circ g)'(x;\, e)$ exists whenever $x \in X \setminus N$, $e \in X$.

We now turn our attention to measurability of derivative assignments. The reason why we care so much about their (Borel) measurability is apparent from the following example, which points out that "generalized derivatives" for which a chain rule formula holds may be defined in a highly non-constructive way.

5.4. *Example.* Assuming the Continuum Hypothesis, we order all Lipschitz mappings of $\mathbb{R}$ to $\mathbb{R}^2$ into a transfinite sequence g_α indexed by countable ordinals. For each $x \in \mathbb{R}^2$ we find the first ordinal $\alpha = \alpha_x$ for which there is $t \in \mathbb{R}$ such that $g_\alpha(t) = x$, g_α is differentiable at t and $g'_\alpha(t) \neq 0$, and we denote $v_x = g'_\alpha(t)$. Whenever $f : \mathbb{R}^2 \mapsto \mathbb{R}$ we choose for each $x \in \mathbb{R}^2$ any linear map $f^\S(x) : \mathbb{R}^2 \to \mathbb{R}$ such that $f^\S(x)(v_x) = f'(x;\, v_x)$ if $f'(x;\, v_x)$ exists and such that $f^\S(x) \neq f'(x)$ if f is differentiable at x. We show that for any Lipschitz mapping $g : \mathbb{R} \mapsto \mathbb{R}^2$, $(f \circ g)'(t) = f^\S(g(t))(g'(t))$ holds for almost all t. Since this formula clearly holds for almost all t for which $g'(t)$ is a multiple of $v_{g(t)}$, we just have to show that the set T of those t for which $g'(t)$ exists and is not a multiple of $v_{g(t)}$ has measure zero. Find α so that $g = g_\alpha$ and denote by T_β the set of those $(s, t) \in \mathbb{R}^2$ for which $g_\alpha(t) = g_\beta(s)$, $g'_\alpha(t) \neq 0$, $g'_\beta(t) \neq 0$ and $g'_\alpha(t)$ is not a multiple of $g'_\beta(s)$. Since T_β consists only of isolated points, it is countable. If $t \in T$, the minimality of the choice of $\alpha_{g(t)}$ guarantees that $\alpha_{g(t)} \leq \alpha$. Also, we cannot have $\alpha_{g(t)} = \alpha$, since then $g'(t) = v_{g(t)}$. Hence $T \subset \bigcup_{\beta<\alpha} \{t : (\exists s)(t, s) \in T_\beta\}$ is countable, and it follows that the chain rule formula holds, even though it *never* happens that $f^\S(x)$ is the derivative of f at x!

We say that the derivative assignment $y \mapsto f^\blacktriangledown(y)$ is *Borel measurable* if the set of triples $(y, u, f^\blacktriangledown(y)(u))$ such that u belongs to the domain of $f^\blacktriangledown(y)$ is Borel measurable in $Y \times Y \times Z$.

5.5. *Remark.* By standard descriptive set theoretic arguments (see, for example, [25, Theorem 14.12]) the requirement of this definition is equivalent to saying that the set U of the pairs (y, u) such that u belongs to the domain of $f^\blacktriangledown(y)$ is Borel measurable in $Y \times Y$ and that the mapping $(y, u) \mapsto f'(y;\, u)$ is Borel measurable on U. Because of Lemma 1.5, the latter requirement is automatically satisfied if f is Lipschitz.

We intend to show next that the assignment U_Φ from Proposition 5.1(vi), and so also those from (iii), (iv), (v), are Borel measurable. For this, we first establish Borel measurability of sets from which U_Φ is built.

5.6. **Lemma.** *Whenever $f : Y \to Z$ is Lipschitz, $u, v, w \in Y$, $\phi \in \mathrm{Lip}_1(Z)$ and $\alpha, \beta, \gamma, \varepsilon, \tau > 0$, there is a Borel set $E \subset Y$ such that*

$$\{y \in Y : \mathcal{D}_\alpha f(y,u) \subset \mathcal{D}_\beta f(y,v) + B_\phi(\mathcal{D}_\gamma f(y,w);\, \varepsilon)\} \\ \subset E \subset \{y \in Y : \mathcal{D}_\alpha f(y,u) \subset \mathcal{D}_\beta f(y,v) + B_\phi(\mathcal{D}_\gamma f(y,w);\, \varepsilon+\tau)\}$$

Proof. Given any $m \in \mathbb{N}$, we denote by F_n the set of those $y \in Y$ for which one can find $1/n \le r \le \alpha - 1/n$ such that

$$\operatorname{dist}\Big(\frac{f(y+ru)-f(y)}{r}, \mathcal{D}_\beta f(y,v) + B_\phi(\mathcal{D}_\gamma f(y,w);\, \varepsilon)\Big) \ge \tau, \tag{5.3}$$

and we show that each F_n is closed. Let $y_k \in F_n$, $y_k \to y$, and let $1/n \le r_k \le \alpha - 1/n$ be such that (5.3) holds with y replaced by y_k and r by r_k. It suffices to assume that $r_k \to r$ and show (5.3) for this r. Suppose, for a contradiction, that (5.3) fails,

$$\operatorname{dist}\Big(\frac{f(y+ru)-f(y)}{r}, \mathcal{D}_\beta f(y,v) + B_\phi(\mathcal{D}_\gamma f(y,w);\, \varepsilon)\Big) < \tau.$$

This means that there are $0 < s < \beta$, $0 < t < \gamma$ and $z \in Z$ such that

$$\phi(z) < \phi\Big(\frac{f(y+tw)-f(y)}{t}\Big) + \varepsilon$$

and

$$\Big\|\frac{f(y+ru)-f(y)}{r} - \frac{f(y+sv)-f(y)}{s} - z\Big\| < \tau.$$

By continuity, both these inequalities remain valid when we replace y by y_k and r by r_k for k large enough, providing a contradiction with (5.3).

Let $E = Y \setminus \bigcup_{n=1}^\infty F_n$. If $\mathcal{D}_\alpha f(y,u) \subset \mathcal{D}_\beta f(y,v) + B_\phi(\mathcal{D}_\gamma f(y,w);\, \varepsilon)$, then clearly $y \in E$. Conversely, if $y \in E$, we choose for every $0 < r < \alpha$ an $n \in \mathbb{N}$ such that $1/n \le r \le \alpha - 1/n$ and use that $y \notin F_n$ to infer

$$\frac{f(y+ru)-f(y)}{r} \in B(\mathcal{D}_\beta f(y,v) + B_\phi(\mathcal{D}_\gamma f(y,w), \varepsilon);\, \tau) \\ \subset \mathcal{D}_\beta f(y,v) + B_\phi(\mathcal{D}_\gamma f(y,w);\, \varepsilon+\tau). \qquad \square$$

5.7. **Theorem.** *Each of the assignments from Proposition 5.1 (iii)–(vi) is Borel measurable.*

Proof. By the additional statement in Proposition 5.1 we just need to consider the assignment $U_\Phi(f,y)$ defined in 5.1(vi), and by Remark 5.5 it is enough to show that the set A of pairs $(y,u) \in Y \times Y$, such that $u \in U_\Phi(f,y)$, is Borel. Let W and Ψ be countable dense subsets of Y and Φ, respectively. For $\sigma = \pm 1$, $u, v \in W$, $\psi \in \Psi$ and $p, q \in \mathbb{N}$ use Lemma 5.6 to find a Borel set $E_{\sigma,u,v,\psi,p,q}$ such that

$$\{y \in Y : \mathcal{D}_\delta f(y, \sigma u + v) \subset \mathcal{D}_\delta f(y, \sigma u) + B_\phi(\mathcal{D}_\delta f(y,v);\, \varepsilon)\} \\ \subset E_{\sigma,u,v,\psi,p,q} \subset \{y \in Y : \mathcal{D}_\delta f(y, \sigma u + v) \subset \mathcal{D}_\delta f(y, \sigma u) + B_\phi(\mathcal{D}_\delta f(y,v);\, 2\varepsilon)\}$$

where $\varepsilon = 1/p$ and $\delta = 1/q$.

We show that

$$A = \Big\{(y,u) : f'(y;\,u) \text{ exists and } y \in \bigcap_{\sigma=\pm1}\bigcap_{v\in W}\bigcap_{\psi\in\Psi}\bigcap_{p=1}^{\infty}\bigcup_{r=1}^{\infty}\bigcap_{q=r}^{\infty} E_{\sigma,u,v,\psi,p,q}\Big\}. \tag{5.4}$$

If $(y,u)\in A$, $f'(y;\,u)$ exists, and for every $\sigma=\pm1$, $v\in W$, $\psi\in\Psi$ and $p\in\mathbb{N}$ we may find $\delta_0>0$ such that for every $0<\delta<\delta_0$,

$$\mathcal{D}_\delta f(y,\sigma u+w)\subset f'(y;\,\sigma u)+B_\psi(\mathcal{D}_\delta f(y,w);\,1/2p).$$

Choosing $r\in\mathbb{N}$ such that $r>1/\delta_0$, we infer from $f'(y;\,\sigma u)\in\overline{\mathcal{D}_\delta f(y,\sigma u)}$ that $y\in E_{\sigma,u,v,\psi,p,q}$ for every $q\ge r$. Hence (y,u) belongs to the right side of (5.4).

Assume now that (y,u) belongs to the right side of (5.4), $\sigma=\pm1$, $\varepsilon>0$, $v\in Y$, and $\phi\in\Phi$. Find $p\in\mathbb{N}$, $\tilde v\in W$ and $\psi\in\Psi$ such that $p>(3+4\mathrm{Lip}(f))/\varepsilon$, $\|\tilde v-v\|<1/p$, and $\sup_{\|z\|\le\mathrm{Lip}(f)}|\psi(z)-\phi(z)|<1/p$. By our assumption, there is $r\in\mathbb{N}$ such that $y\in E_{\sigma,\tilde u,\tilde v,\psi,p,q}$ for every $q\ge r$. Let $\delta_0=1/r_0$, where $r_0\ge r$ is such that $(\|u\|+\|v\|)/r_0<1/p$. For every $0<\delta<\delta_0$ we find $q\ge r_0$ such that $\delta_q:=1/(q+1)\le\delta<1/q$ and using Lemma 2.4 conclude that

$$\begin{aligned}
\mathcal{D}_\delta f(y,\sigma u+v) &\subset B(\mathcal{D}_{\delta_q} f(y,\sigma u+v), 2\mathrm{Lip}(f)(\|u\|+\|v\|)/q)\\
&\subset B(\mathcal{D}_{\delta_q} f(y,\sigma u+\tilde v), 3\mathrm{Lip}(f)/p)\\
&\subset B(\mathcal{D}_{\delta_q} f(y,\sigma u)+B_\psi(\mathcal{D}_{\delta_q} f(y,\tilde v);\,2/p), 3\mathrm{Lip}(f)/p)\\
&\subset B(\mathcal{D}_{\delta_q} f(y,\sigma u)+B_\phi(\mathcal{D}_{\delta_q} f(y,v), 3/p);\,4\mathrm{Lip}(f)/p)\\
&\subset \mathcal{D}_{\delta_q} f(y,\sigma u)+B_\phi(\mathcal{D}_{\delta_q} f(y,v);\,(3+4\mathrm{Lip}(f))/p)\\
&\subset \mathcal{D}_{\delta} f(y,\sigma u)+B_\phi(\mathcal{D}_{\delta} f(y,v);\,\varepsilon).
\end{aligned}$$

Having established (5.4), we recall that the set $\{(y,u) : f'(y;\,u) \text{ exists}\}$ is Borel by Lemma 1.5 (i), and conclude that (5.4) shows that A is Borel. □

5.8. *Remark.* While the assignment from Proposition 5.1 (i) has non-constructive features reminiscent of Example 5.3, the assignment $U_a(f,x)$ from 5.1 (ii) is defined in a way suggesting that is may be measurable. In fact, it is measurable with respect to the σ-algebra generated by Suslin sets (and hence universally measurable). To see this, let E denote the set of $(y,u)\in Y\times Y$ such that $f'(y;\,u)$ exists, and let F be the set of $(y,u,v,w)\in Y^4$ that satisfy the following three conditions.

- $(y,u),(y,v)\in E$;
- $w=u+v$;
- either $(y,w)\notin E$, or $(y,w)\in E$ and $f'(y;\,w)\ne f'(y;\,u)+f'(y;\,v)$.

Then Lemma 1.5 implies that F is Borel, and so

$$\{(u,y) : u\in U_a(f,y)\} = E\setminus\{(y,u) : (\exists v,w)(y,u,v,w)\in F\}$$

is a complement of a Suslin set.

6. Existence and measurability of complete derivatives

Finally, we want to address briefly the problem of defining complete derivatives. As already pointed out, this is an extension problem, since we construct complete derivatives by considering a complete derivative assignment $y\mapsto f^{\blacktriangledown}(y)$ and attempting to extend, for each y, the mapping $f^{\blacktriangledown}(y)$ from its domain to the whole space; of course, we would also prefer to do this in a measurable way. There are three cases we consider, since firstly, such an extension is easy to establish in

Hilbert spaces, secondly, we have the Hahn-Banach Theorem if the target space is one dimensional, and thirdly, an abstract selection may be used if there is enough (weak) compactness in the target space. We recall that the Borel measurability of $y \mapsto f^{\S}(y)$ means that the mapping $(y, v) \mapsto f^{\S}(y)(v)$ of $Y \times Y$ to Z is Borel measurable; since we are in complete separable metric spaces, this is equivalent to requiring that the set of triples $(y, v, f^{\S}(y)(v))$ is a Borel measurable subset of $Y \times Y \times Z$.

6.1. **Lemma.** *Suppose that Y, Z are separable Banach spaces and subspaces $U(f, y)$ of Y define a measurable derivative assignment for a Lipschitz map $f : Y \to Z$. If Y is reflexive, then for every bounded closed convex subset G of Y the sets $\{y : G \cap U(f, y) \neq \emptyset\}$ and $\{y : G \cap U(f, y) = \emptyset\}$ are Borel measurable subsets of Y.*

Proof. Let $U = \{(y, u) : u \in U(f, y)\}$. The set $V = U \cap (Y \times G)$ is a Borel subset of $Y \times G$ with closed convex sections $V_y = \{u \in G : (y, u) \in V\}$. Endowing G with the (compact metrizable) weak topology, we see that V is a Borel subset of $Y \times G$ with compact sections, hence, according to [25, Theorem 28.8], its projection $\{y \in Y : V_y \neq \emptyset\} = \{y : G \cap U(f, y) \neq \emptyset\}$ is a Borel subset of Y. The second set is complementary to the first. □

The special case of Hilbert spaces, including finite dimensional spaces. We emphasize that finite dimensional spaces are included in this simple approach; this is, of course, achieved by choosing an arbitrary scalar product.

6.2. **Proposition.** *Let Y be a separable Hilbert space and let Z be a separable Banach space. Whenever $y \in Y \mapsto f^{\blacktriangledown}(y)$ is a measurable complete derivative assignment for a Lipschitz map $f : Y \to Z$, then $f^{\S}(y) = f^{\blacktriangledown} \circ P_y$, where P_y is the orthogonal projection onto the domain of $f^{\blacktriangledown}(y)$, is a Borel measurable complete derivative of f.*

Proof. Since the mapping $(y, v) \mapsto f'(y; v)$ is measurable on its domain which is a Borel subset of $Y \times Y$ (see Lemma 1.5), it suffices to show that the mapping $(y, v) \mapsto (y, P_y(v))$ is Borel measurable, which is the same as saying that the mapping $(y, v) \mapsto P_y(v)$ is Borel measurable.

Let $U(y)$ denote the domain of $f^{\blacktriangledown}(y)$, $U := \{(y, u) : u \in U(y)\}$, and let (v_i, r_i) be a sequence dense in $Y \times (0, \infty)$. We fix a closed ball B in Y and put

$$V_{i,n} = \{(y, v) : \|v - v_i\| < 1/n, \overline{B(v_i, r_i)} \cap U(y) = \emptyset, \overline{B(v_i, r_i + 1/n)} \cap B \cap U(y) \neq \emptyset\}.$$

Observing that Lemma 6.1 implies that the sets $V_{i,n}$ are Borel measurable, we see that

$$\{(y, v) : P_y(v) \in B\} = \bigcap_{n=1}^{\infty} \bigcup_{i=1}^{\infty} V_{i,n},$$

is a Borel subset of $Y \times Y$, which implies the statement, since every open subset of Y is a countable union of closed balls. □

Note. By using Proposition 6.2 with the derivative assignments from Proposition 5.1 (v) or (vi), or with Proposition 5.1(iv) if Z is finite dimensional, or Proposition 5.1(iii) if Y is one dimensional, we thus obtain a measurable complete derivative which extends the Gâteaux derivative in the sense that if f is Gâteaux differentiable at y, then $f^{\S}(y) = f'(y)$.

The special case of one dimensional target spaces. We show that the completeness of the assignment from Proposition 5.1(iii) enables relatively simple constructions of complete derivatives for real-valued mappings. Obviously, applied to each coordinate separately this also gives a proof of existence of complete derivatives for mappings into finite dimensional spaces.

6.3. **Proposition.** *Let Y be a separable Banach space and let $f : Y \to \mathbb{R}$ be locally Lipschitz. Suppose that, for each $y \in Y$, $f^{\S}(y) \in Y^*$ is chosen such that $f^{\S}(y)(u) \le \widetilde{D}f(y,u)$ for every $u \in Y$. Then $y \mapsto f^{\S}(y)$ is a complete derivative.*

Proof. Because the derivative assignment from Proposition 5.1(iii) is complete, in order to verify that $y \mapsto f^{\S}(y)$ is a complete derivative it suffices to show that $f^{\S}(y)(u) = \widetilde{D}f(y,u)$ for every $u \in Y$ such that $\widetilde{D}f(y,u) = -\widetilde{D}(f)(y,-u)$. But this is obvious, since $f^{\S}(y)(\pm u) \le \widetilde{D}f(y,\pm u)$ implies that $f^{\S}(y)(u) \le \widetilde{D}f(y,u)$ and $f^{\S}(y)(u) = -f^{\S}(y)(-u) \ge -\widetilde{D}f(y,-u) = \widetilde{D}f(y,u)$. □

Since $\widetilde{D}f(v)$ is convex and subadditive, the Hahn-Banach Theorem implies that Proposition 6.3 may be used to find complete derivatives of all Lipschitz functions. However, we are again left with a measurability problem, which is answered in the following statement.

6.4. **Proposition.** *Let Y be a separable Banach space and let $f : Y \mapsto R$ be locally Lipschitz. Then there is a norm to w^* Borel measurable mapping $f^{\S} : Y \to Y^*$ such that $f^{\S}(y)(u) \le \widetilde{D}f(y,u)$ for every $u \in Y$ for all $y, u \in Y$.*

If Y^* is separable, any such $f^{\S}$ is norm to norm Borel measurable. In particular, $y \mapsto f^{\S}(y)$ is a measurable complete derivative.

Proof. For $y \in Y$ put

$$T(y) = \{y^* \in Y^* : \forall v \in Y \langle y^*, v\rangle \le \widetilde{D}f(y,v)\}.$$

Since the functions $v \to \widetilde{D}f(y,v)$ are Lipschitz, convex, and subadditive, the Hahn-Banach Theorem implies that $T(y) \ne \emptyset$ for each $y \in Y$. Moreover, each $T(y)$ is clearly w^*-closed and bounded, hence w^* compact. We prove that T is a norm to w^* Borel measurable multivalued mapping. To this end it is sufficient to prove that for every $v \in Y$ and every $c \in \mathbb{R}$ the set

$$A_{v,c} = \{y \in Y : T(y) \cap \{y^* : \langle y^*, v\rangle > c\} \ne \emptyset\}$$

is Borel measurable. By the Hahn-Banach Theorem and by Lemma 1.4,

$$A_{v,c} = \{y : \widetilde{D}f(y,v) > c\} = \bigcup_i \{y : \widetilde{D}f(y, v + w_i) - \widetilde{D}f(y,v) > c\},$$

where w_i is any sequence of elements of Y dense in Y. Hence Lemma 1.4 implies that $A_{v,c}$ is Borel measurable. The first statement of the proposition now follows by the Kuratowski-Ryll-Nardzewski selection theorem (see [26]).

If Y^* is separable, w^* and norm Borel sets in Y^* coincide, which implies the second statement of the Proposition. □

6.5. *Remark.* On may, of course, try to impose other conditions on $f^{\S}$ in Proposition 6.4. The most natural would be to require the lower bound by the corresponding lower derivative, i.e., that for all $y, v \in Y$,

$$-\widetilde{D}(-f)(y,v) \le \langle f^{\S}(y), v\rangle \le \widetilde{D}f(y,v).$$

Since $-\widetilde{D}(-f)(y,v) \leq \widetilde{D}f(y,v)$, the only difference in the proof would be the difference in non-emptiness criteria and therefore in the use of the Hahn-Banach Theorem.

Chain rule and subdifferential. We first give an example showing that the chain rule fails with the Michel-Penot subdifferential and then show, as promised in the introduction, that it holds with the upper Dini subdifferential.

6.6. *Example.* For $x \in \mathbb{R}$ let $\psi(x)$ be its distance to the nearest integer. Denote e_1, e_2 the standard basis of $\mathbb{R}^2$, define $f(xe_1 + 2^{-k}e_2) = 2^{-k}\psi(2^{k-2}x)$ for $k = 0, 1, 2, \dots$ and $x \in \mathbb{R}$, $f(xe_1 + ye_2) = 0$ if $y \leq 0$, observe that f is Lipschitz on its domain, and extend it to a Lipschitz function from $\mathbb{R}^2$ to $\mathbb{R}$. Then, for a.e. $x \in \mathbb{R}$,

$$\limsup_{k\to\infty} 2^k(f(xe_1 + 2^{-k}e_2 + 2^{-k}e_1) - f(xe_1 + 2^{-k}e_2)) \geq 1/4.$$

Indeed, for almost every $x \in \mathbb{R}$ there are infinitely many k for which one can find an integer n such that $n \leq 2^{k-2}x \leq n + 1/4$, and for any such k we have $2^k(f(xe_1 + 2^{-k}e_2 + 2^{-k}e_1) - f(xe_1 + 2^{-k}e_2)) = \psi(2^{k-2}x + 1/4) - \psi(2^{k-2}x) = 1/4$.

It follows that for a.e. $x \in \mathbb{R}$ there is $\xi(x)$ belonging to the Michel-Penot subdifferential of f at xe_1 such that $\xi(x)(e_1) = 1/4$. So with $g : \mathbb{R} \mapsto \mathbb{R}^2$ defined by $g(x) = xe_1$ we do not get that $(f \circ g)' = \xi(g(x))(g')$ a.e.

6.7. **Definition.** The *upper Dini subdifferential* $\partial^D f(y)$ of a locally Lipschitz function $f : Y \to \mathbb{R}$ at a point $y \in Y$ is

$$\partial^D f(y) = \{y^* \in Y^* : y^*(v) \leq \widetilde{D}f(y,v) \text{ for every } v \in Y.\}$$

6.8. **Theorem.** *Suppose $f : Y \to \mathbb{R}$ and $g : X \to Y$ are locally Lipschitz maps, where X, Y are separable Banach spaces and Y has the Radon-Nikodym property, Then for $\mathcal{L}$ almost all $x \in X$,*

$$\partial^D(f \circ g)(x) = (\partial^D f(g(x))) \circ g'(x).$$

Proof. For each $y \in Y$ choose some $f^{\S}(y) \in \partial^D f(y)$. Let $U(f, y)$ be the assignment from Proposition 5.1(iii). Since $f^{\S}(y)$ is a complete derivative and $U(f, y)$ a complete assignment, for $\mathcal{L}$ almost all $x \in X$ both $g'(x)$ and $(f \circ g)'(x)$ exist, $(f \circ g)'(x) = f^{\S}(g(x)) \circ g'(x)$ and the range of $g'(x)$ is contained in $U(f, g(x))$. Since $\widetilde{D}f(g(x))$ is linear on $U(f, g(x))$, the restrictions of $f^{\S}(g(x))$ and any $y^* \in \partial^D f(g(x)))$ to $U(f, g(x))$ coincide. Hence $y^* \circ g'(x) = f^{\S}(g(x)) \circ g'(x) = (f \circ g)'(x)$ for every $y^* \in \partial^D f(g(x)))$. □

Existence and measurability for reflexive target spaces. We start with a special case of a result due to Lindenstrauss [27] whose simple proof is due to Pelczynski [34]. Both proofs may be found in [4, Theorem 7.2].

6.9. **Lemma.** *If Y and Z are Banach spaces, Z is reflexive, and $f : Y \mapsto Z$ is K-Lipschitz on a neighborhood of y and Gâteaux differentiable in the direction of a linear subspace $U \subset Y$, then there is a linear mapping $L : Y \mapsto Z$ of norm at most K such that $f'(y; v) = L(v)$ for all $v \in U$.*

6.10. **Proposition.** *Let Y be a separable Banach space and let Z be a separable reflexive space. Whenever $y \in Y \mapsto f^{\blacktriangledown}(y)$ is a measurable complete derivative assignment for a Lipschitz $f : Y \mapsto Z$, then there is a Borel measurable complete derivative $f^{\S}(y)$ such that $f^{\S}(y)(v) = f^{\blacktriangledown}(y)(v)$ for every $y \in Y$ and every v in the domain of $f^{\blacktriangledown}(y)$.*

Proof. Suppose f is K-Lipschitz. The set $\mathfrak{L}$ of linear mappings of Y to Z of norm at most K considered in the topology of pointwise weak convergence is a compact metrizable space. Denote, as always, by $U(f,y)$ the domain of $f^{\blacktriangledown}(y)$ and let

$$\mathfrak{S} := \{(y,L) \in Y \times \mathfrak{L} : L(u) = f'(y;\, u) \text{ for all } u \in U(f,y)\}.$$

For every $y \in Y$ the set $\{L \in \mathfrak{L} : (y,L) \in \mathfrak{S}\}$ is compact and, according to Lemma 6.9, non-empty. Provided $\mathfrak{S}$ is Borel, we infer from the Kuratowski-Ryll-Nardzewski selection theorem (see [26]) that there is Borel measurable $f^{\S} : Y \mapsto \mathfrak{L}$ such that $(y, f^{\S}(y)) \in \mathfrak{S}$ for every $y \in Y$. Clearly, this $f^{\S}$ has all properties required from the complete derivative of f.

It remains to show that $\mathfrak{S}$ is Borel. For this, let (w_i, s_i) be a sequence dense in $Y \times (0,\infty)$, and denote

$$\mathfrak{S}_i = \{(y,L) \in Y \times \mathfrak{L} : \overline{B(w_i,s_i)} \cap U(f,y) = \emptyset\}$$

and

$$\mathfrak{T}_i = \{(y,L) \in Y \times \mathfrak{L} : \operatorname{dist}(L(w_i), \mathcal{D}_\delta f(y,w_i)) \le 4Ks_i + \operatorname{diam}(\mathcal{D}_\delta f(y,w_i)) \forall \delta > 0\}.$$

These sets are Borel, $\mathfrak{S}_i$ by Lemma 6.1, and $\mathfrak{T}_i$ by Lemma 1.4. We let

$$\mathfrak{T} = \bigcap_{i=1}^{\infty} (\mathfrak{S}_i \cup \mathfrak{T}_i)$$

and finish the proof by showing $\mathfrak{S} = \mathfrak{T}$.

If $(y,L) \in \mathfrak{S}$ and $\overline{B(w_i,s_i)} \cap U(f,y) \neq \emptyset$, we choose $u \in \overline{B(w_i,s_i)} \cap U(f,y)$ and use Lemma 2.3 to estimate

$$\begin{aligned}
\operatorname{dist}(L(w_i), \mathcal{D}_\delta f(y,w_i)) &\le 2K \|u - w_i\| + \operatorname{dist}(L(u), \mathcal{D}_\delta f(y,u)) \\
&\le 2K \|u - w_i\| + \operatorname{diam}(\mathcal{D}_\delta f(y,u)) \\
&\le 4K \|u - w_i\| + \operatorname{diam}(\mathcal{D}_\delta f(y,w_i)) \\
&\le 4Ks_i + \operatorname{diam}(\mathcal{D}_\delta f(y,w_i)).
\end{aligned}$$

If $(y,L) \in \mathfrak{T}$, $u \in U(f,y)$, and $\varepsilon > 0$, we find i such that $\|w_i - u\| < \varepsilon/2 < s_i < \varepsilon$ and infer from $\overline{B(w_i,s_i)} \cap U(f,y) \neq \emptyset$ that

$$\begin{aligned}
\operatorname{dist}(L(u), \mathcal{D}_\delta f(y,u)) &\le 2Ks_i + \operatorname{dist}(L(w_i), \mathcal{D}_\delta f(y,w_i)) \\
&\le 6Ks_i + \operatorname{diam}(\mathcal{D}_\delta f(y,w_i)) \\
&\le 8Ks_i + \operatorname{diam}(\mathcal{D}_\delta f(y,u)).
\end{aligned}$$

Choosing δ small enough we conclude $\|L(u) - f'(x;\, u)\| < 8K\varepsilon$, implying $L(u) = f'(y;\, u)$, and so $(y,L) \in \mathfrak{S}$. □

7. Chain rule when the inner function is not Lipschitz

This section is devoted to the situation in which a finite dimensionality assumption helps to strengthen the chain rule to cases when the innermost function is not assumed to be Lipschitz. Two such extensions are of interest: pointwise chain rule at almost all points at which the inner function is differentiable and the chain rule treated in [2] for the weak derivative when the inner function is of bounded variation. Because of Theorem 4.3 it suffices to state the results for composition of two functions only.

The pointwise case is handled by reducing it to the Lipschitz situation for countably many mappings.

7.1. **Lemma.** *Let f be a Borel measurable map of a separable Banach space X to a Banach space Y. Then the set*

$$E := \Big\{x \in X : \limsup_{t \searrow 0} \frac{\|f(x+tu) - f(x)\|}{t} < \infty \text{ for every } u \in X\Big\}.$$

can be covered by countably many Borel sets on each of which f is Lipschitz.

Proof. For $x \in X$ and $k, l \in \mathbb{N}$ denote

$$V_{k,l}(x) := \{v \in X : \big\|f(x + 2^{-l}v) - f(x)\big\| \le 2^{-l}k\} \text{ and } V_k(x) = \bigcap_{l=0}^{\infty} V_{k,l}(x).$$

We first show that for every $u, z \in X$, $r > 0$, and $k \ge 0$, f is Lipschitz on the set A of those $x \in B(z, r/4)$ for which $B(u, r) \setminus V_k(x)$ is meager. To see this, consider any $x, y \in A$, $x \ne y$ and find $l \ge 0$ so that $2^{-l-2}r < \|x - y\| \le 2^{-l-1}r$. Since $(B(u, r) \setminus V_k(x))$ and $(B(u, r) \setminus V_k(y)) - 2^{-l}(x - y)$ are both meager and $B(u, r) - 2^{-l}(x - y) \supset B(u, r/2)$, there is

$$v \in B(u, r/2) \cap V_k(x) \cap (V_k(y) - 2^{-l}(x - y)).$$

Since $v \in V_k(x)$, $\big\|f(x + 2^{-l}v) - f(x)\big\| \le 2^{-l}k$, and, since $w := v + 2^l(x - y) \in V_k(y)$, we also have $\big\|f(y + 2^{-l}w) - f(y)\big\| \le 2^{-l}k$. But $x + 2^{-l}v = y + 2^{-l}w$, implying that $\|f(x) - f(y)\| \le 2^{-l+1}k \le (8k/r)\,\|x - y\|$.

Next we show that the set E defined in the hypothesis of the Lemma can be covered by countably many sets of the above form. Let $u_i \in X$ be dense in X and denote by $E_{i,j,k}$ the set of $x \in X$ for which $B(u_i, 2^{-j}) \setminus V_k(x)$ is meager. Given any $x \in E$, the definition of E shows that $\bigcup_{k=0}^{\infty} V_k(x) = X$. By Baire category theorem there is k such that $V_k(x)$ is not meager. Since $V_k(x)$, being Borel, has the Baire property, its complement is meager in some nonempty open set. Hence there are $i, j \in \mathbb{N}$ such that $B(u_i, 2^{-j}) \setminus V_k(x)$ is meager. It follows that $E_{i,j,k}$ cover E, and so do the sets $E_{i,j,k} \cap B(u_l, 2^{-j-2})$ on which, as shown above, f is Lipschitz.

From the above argument it may not be apparent that the sets we have constructed are Borel. However, they can be easily made Borel, since whenever f is Lipschitz on A, we can extend it to a Lipschitz function g on $\overline{A}$ and observe that f is Lipschitz on the set $\{x \in \overline{A} : f(x) = g(x)\}$ which is Borel and contains A. □

7.2. **Lemma.** *Let $f, g : X \to Y$ be Borel measurable maps between separable Banach spaces X and Y. Then the set of $x \in X$ at which $f(x) = g(x)$, both f and g are Gâteaux differentiable, and $f'(x) \ne g'(x)$ is σ-directionally porous.*

Proof. Replacing f by $f - g$, we may assume that $g = 0$. Let $E = \{x : f(x) = 0\}$, and, for $x \in E$, and $k, l, m \in \mathbb{N}$ denote

$$V_{k,z}(x) := \{v \in X : \|f(x + tv) - tz\| < \tfrac{1}{2}|t|\,\|z\| \text{ for } |t| < 2^{-k}\}.$$

Fix for a while $u \in X$, $r > 0$, $z \in Y$ and $k \in \mathbb{N}$ and let A be the set of those $x \in E$ for which $B(u, r) \setminus V_{k,z}(x)$ is meager. We show that A is porous in the direction of u. To see this, consider any $x \in A$ and suppose there are $0 < t < 2^{-k}$ and $y \in A$ such that $\|y - x - tu\| < rt/2$, in other words $\|(y - x)/t - u\| < r/2$. Since $B(u, r) \setminus V_{k,z}(x)$ and $(y - x)/t - (B(u, r) \setminus V_{k,z}(x))$ are both meager and $(y - x)/t - B(u, r) \supset B(u, r/2)$, there is

$$v \in B(u, r/2) \cap V_{k,z}(x) \cap ((y - x)/t - V_{k,z}(y)).$$

Since $|t| < 2^{-k}$ and $v \in V_{k,z}(x)$, $\|f(x+tv) - tz\| < \frac{1}{2}|t|\,\|z\|$. Observing that $w := -v + (y-x)/t \in V_{k,z}(y)$, we also have $\|f(y-tw) + tz\| < \frac{1}{2}|t|\,\|z\|$. But $x + tv = y - tw$, and so $0 = \|2tz + f(x+tv) - tz - (f(y-tw) + tz)\| > |t|\,\|z\|$ provides the required contradiction.

It remains to show that E can be covered by countably many sets of the above form. Let $u_i \in X$ be dense in X, $z_l \in Y$ dense in Y and denote by $E_{i,j,k,l}$ the set of $x \in X$ for which $B(u_i, 2^{-j}) \cap V_{k,z_l}(x)$ is meager. Given any $x \in E$, any $v \in X$ for which $f'(x;\, v) \neq 0$ belongs to $V_{k,z_l}(x)$ once k is sufficiently large and z_l is close enough to $f'(x;\, v)$. Hence $\bigcup_{k=0}^{\infty} \bigcup_{l=0}^{\infty} V_{k,z_l}(x)$ contains all $v \in X$ with $f'(x;\, v) \neq 0$, and by Baire category theorem there are k, l such that $V_{k,z_l}(x)$ is not meager. Since $V_{k,z_l}(x)$, being co-Suslin, has the Baire property, its complement is meager in some nonempty open set. Hence there are $i, j \in \mathbb{N}$ such that $B(u_i, 2^{-j}) \setminus V_{k,z_l}(x)$ is meager. It follows that $E_{i,j,k,l}$ are directionally porous sets covering E. □

7.3. **Proposition.** *Suppose Y, Z are separable Banach spaces and $f : Y \to Z$ is a Lipschitz function having a derivative assignment $y \in Y \to f^{\blacktriangledown}(y)$. Then for every finite dimensional X and Borel measurable $g : X \to Y$, $f \circ g$ is Gâteaux differentiable at $\mathcal{L}$ almost every point x at which g is, and $(f \circ g)'(x) = f^{\blacktriangledown}(g(x)) \circ g'(x)$.*

Proof. By Lemma 7.1 there are Borel sets $F_k \subset X$ such that g is Lipschitz on each F_k and the union of the F_k contains all points of Gâteaux differentiability of g. By [29, Lemma 5.5.3] we may extend the restriction of g to F_k to a Lipschitz function $g_k : X \to Y$. Since the chain rule holds for the compositions $f \circ g_k$, we just have to establish that $g_k'(x) = g'(x)$ for all $x \in F_k$ except for a set belonging to $\mathcal{L}$. But this was shown in Lemma 7.2. □

In the case when $X = \mathbb{R}$, Y has the Radon-Nikodym property, $g : \mathbb{R} \to Y$ has bounded variation and $f : Y \to Z$ is a Lipschitz map with a derivative assignment $f^{\blacktriangledown}$, g is differentiable almost everywhere and so the previous proposition implies that $(f \circ g)'(x) = f^{\blacktriangledown}(g(x)) \circ g'(x)$ at almost every $x \in \mathbb{R}$ (where almost every means with respect to the Lebesgue measure). However, in contrast to the case when g is Lipschitz, $f \circ g$ cannot be fully recovered from its derivative. We therefore show that a natural formula transforming the distributional derivative of g (which is a measure) into the distributional derivative of $f \circ g$ also holds, thus replacing the g-dependent direction of differentiability in the result of [2] by an f-dependent choice of the domain of $f^{\blacktriangledown}$. We treat only the case of one dimensional domain, since the extension to higher dimensions can be obtained by repeating the arguments of [2], but we treat the problem in somewhat more general setting. Assuming that a function $g : \mathbb{R} \to Y$ is such that for some Radon measure μ in $\mathbb{R}$ and μ-Bochner integrable $\phi : \mathbb{R} \to Y$, one has $g(v) - g(u) = \int_{(u,v]} \phi(s)\, d\mu(s)$ for every interval $(u, v] \subset \mathbb{R}$, we show how to find a μ-Bochner integrable function that integrates to $(f \circ g)(v) - (f \circ g)(u)$, where f is a Lipschitz map having a derivative assignment. Notice that such g is necessarily left continuous and is of bounded variation, and that for any left continuous function g of bounded variation with values in a space with the Radon-Nikodym property such ϕ and μ necessarily exist.

To find a "derivative" of $f \circ g$, we will transform the given function g of bounded variation into a Lipschitz function. This will be done with the help of the fact that every Radon measure ν in $\mathbb{R}$ satisfying $\nu \geq \lambda$ (where λ is the Lebesgue measure) is an image of λ under a non-decreasing Lipschitz surjection $q : \mathbb{R} \to \mathbb{R}$. (So, by definition of image measures, $\nu(E) = \lambda(q^{-1}(E))$ for every Borel set $E \subset \mathbb{R}$.) Such

a function q is easy to define: choosing s_0 with $\mu(\{s_0\}) = 0$, we let

$$q(t) = \sup\{s \geq s_0 : \nu[s_0, s] \leq t\} \qquad \text{for } t \geq 0, \text{ and}$$
$$q(t) = \inf\{s \leq s_0 : \nu[s, s_0] \leq -t\} \qquad \text{for } t < 0.$$

We also recall the formula for integration with respect to the image measure,

$$\int_E h(s)\, d\nu(s) = \int_{q^{-1}(E)} h(q(t))\, dt \tag{7.1}$$

whenever one of these integrals exists.

7.4. **Theorem.** *Suppose that $g : \mathbb{R} \to Y$ is such that for some Radon measure μ in $\mathbb{R}$ and μ-Bochner integrable $\phi : \mathbb{R} \to Y$, $g(v) - g(u) = \int_{(u,v]} \phi(s)\, d\mu(s)$ for every interval $(u, v] \subset \mathbb{R}$. Letting $S = \{s \in \mathbb{R} : \mu(\{s\}) > 0\}$, we have that for any Lipschitz $f : Y \to Z$ with a derivative assignment $f^{\blacktriangledown}$,*

$$\begin{aligned} &f(g(v)) - f(g(u)) \\ &\qquad = \int_{(u,v]\cap S} \frac{f(g(s+)) - f(g(s-))}{\mu(s)}\, d\mu(s) + \int_{(u,v]\setminus S} f^{\blacktriangledown}(g(s))\phi(s) d\mu(s) \end{aligned}$$

for every interval $(u, v] \subset \mathbb{R}$.

Proof. Define a new Radon measure ν on $\mathbb{R}$ by $\nu(E) = \lambda(E) + \int_E (1 + \|\phi\|)\, d\mu$. Observing that μ is absolutely continuous with respect to ν and its Radon-Nikodym derivative satisfies $\dfrac{d\mu}{d\nu} \leq \dfrac{1}{1 + \|\phi\|}$, we let $\psi = \phi \dfrac{d\mu}{d\nu}$ and infer that $\|\psi\| \leq 1$ and $g(v) - g(u) = \int_{(u,v]} \psi(s)\, d\nu(s)$ for every interval $(u, v] \subset \mathbb{R}$.

Since $\nu \geq \lambda$, it is an image of λ under a non-decreasing Lipschitz surjection $q : \mathbb{R} \to \mathbb{R}$. The function $\psi \circ q$ is bounded and measurable; let h be its indefinite Bochner integral. Then h is Lipschitz and $h' = \psi \circ q$, λ-a.e. For $\alpha, \beta \in q^{-1}(\mathbb{R} \setminus S)$, $\alpha < \beta$ we have $q^{-1}(q(\alpha), q(\beta)] = (\alpha, \beta]$, which implies that

$$h(\beta) - h(\alpha) = \int_\alpha^\beta \psi(q(t))\, dt = g(q(\beta)) - g(q(\alpha)),$$

using (7.1). Hence we may assume that $h = g \circ q$ on $q^{-1}(\mathbb{R} \setminus S)$. When $s \in S$, the function h has constant derivative $\psi(s)$ on the interval $q^{-1}(s)$, and we infer that h is affine with values at the left and right end-points $g(s-)$ and $g(s+)$, respectively.

Let $f : Y \to Z$ be Lipschitz with a derivative assignment $f^{\blacktriangledown}$. Then $f \circ h$, being Lipschitz and a.e. differentiable, is an indefinite Bochner integral of its derivative. Hence, given any $u, v \in \mathbb{R}$, $u < v$ and denoting $(\alpha, \beta] = q^{-1}(u, v]$,

$$\begin{aligned} &f(g(u)) - f(g(v)) \\ &\quad = f(h(\beta)) - f(h(\alpha)) = \int_\alpha^\beta (f \circ h)'(t)\, dt \\ &\quad = \sum_{s\in(u,v]\cap S} \int_{q^{-1}(s)} (f \circ h)'(t)\, dt + \int_{q^{-1}((u,v]\setminus S)} f^{\blacktriangledown}(g(q(t)))\psi(q(t))\, dt \\ &\quad = \sum_{s\in(u,v]\cap S} (f(g(s+)) - f(g(s-))) + \int_{(u,v]\setminus S} f^{\blacktriangledown}(g(s))\psi(s) d\nu(s) \\ &\quad = \int_{(u,v]\cap S} \frac{f(g(s+)) - f(g(s-))}{\mu(s)}\, d\mu(s) + \int_{(u,v]\setminus S} f^{\blacktriangledown}(g(s))\phi(s) d\mu(s). \qquad \square \end{aligned}$$

8. Final remarks

In this section we briefly indicate how our results may be used to prove some known differentiability results, including Rademacher's and Aronszajn's theorems and existence of intermediate derivatives. The ideas of these proofs are by now standard: the main point, Lebesgue differentiation theorem, has been key already to the first infinite dimensional differentiability results in [3, 9, 30], the finite dimensional version is very close to [33], and porosity approach has been already used also in the infinite dimensional case, both in studying differentiability and intermediate differentiability, for example in [5, 23, 37].

8.1. **Corollary.** *Let f be a locally Lipschitz mapping of a separable Banach space X to a Banach space Y. Then there is a σ-directionally porous subset P of X such that for every $x \in X \setminus P$ the set V_x of those directions in which f is differentiable is a closed linear subspace of X and the mapping $v \to f'(x; v)$ is a continuous linear mapping of V_x to Z.*

Proof. The statement follows immediately from Corollary 2.14, Proposition 2.22 used with Φ being the collection of functions $\|y - y_0\|$ where $y_0 \in Y$, and Example 2.8(ii). □

As a special case of 8.1 we have

8.2. **Corollary.** *Let f be a locally Lipschitz mapping of a separable Banach space X to a Banach space Y. Let A be the set of all points $x \in X$ for which there is a subset W_x of X such that*

(i) *$f'(x; w)$ exists for all $w \in W_x$,*
(ii) *the closed linear span of W_x is X, and*
(iii) *f is not Gâteaux differentiable at x.*

Then A is σ-directionally porous.

From the point of view of this paper, Aronszajn's theorem [3] says that in separable Banach spaces $\mathcal{L}$ null sets are small in the sense of Aronszajn [3] The more familiar version that Lipschitz mappings of separable Banach spaces into spaces having the Radon-Nikodym property are Gâteaux differentiable except for an Aronszajn null set is, of course, equivalent to it (and is what we actually prove).

8.3. **Theorem.** *For every $\mathcal{L}$ null set E in a separable Banach space X and for every sequence $v_n \in X$ whose linear span is dense in X there are Borel sets $E_n \subset X$ such that $E \subset \bigcup_n E_n$ and, for every n, the set E_n has one dimensional measure zero on every line in the direction of v_n.*

Proof. It clearly suffices to show that for any Lipschitz mapping f of X to a Banach space Y with the Radon-Nikodym, the set E of points at which f is not Gâteaux differentiable has the claimed property. Let A be the σ-directionally porous set from 8.2. Using Corollary 1.3, we write $A = \bigcup(A_n^+ \cup A_n^-)$, where $A_n^\pm$ is porous in the direction of $\pm v_n$. By [29, Lemma 10.1.4] we may find Borel sets $B_n^\pm \supset A_n^\pm$ which are also porous in the directions $\pm v_n$. We observe that the intersection of $B_n^\pm$ with every line L in direction of v_n in a porous set in L, therefore of one dimensional measure zero.

Let C_n be the set of points $x \in X$ at which f is not differentiable in the direction of v_n. From Lemma 1.4 we infer that C_n are Borel sets and from the definition

of the Radon-Nikodym property that they have one dimensional measure zero on every line in direction of v_n. Hence the sets $E_n = B_n^+ \cup B_n^- \cup C_n$ have the required property. □

In another direction, a simple application of Theorem 3.6 concerns the somewhat exotic notion of intermediate derivative: $y^* \in Y^*$ is said to be an intermediate derivative of a function $f : Y \mapsto \mathbb{R}$ at a point $y \in Y$ if $\underline{D}f(y,v) \leq \langle y^*, v\rangle \leq \overline{D}f(y,v)$ for every $v \in V$. Locally Lipschitz functions on separable, and even some non-separable, Banach spaces are known to possess the intermediate derivative generically, see, e.g., [17, 16]. In the separable case we improve the existence statement from a generic result to a σ-directionally porous one.

8.4. **Corollary.** *Let f be a real-valued locally Lipschitz function on an open subset H of a separable Banach space Y. Then f has an intermediate derivative at all points of H, except those which belong to a σ-directionally porous set.*

Proof. Let $Q \subset Y$ be a σ-directionally porous set with the properties from Theorem 3.6 and Theorem 3.2. If $y \in H \setminus Q$, the function $v \mapsto \overline{D}f(y,v)$ is convex, continuous and positively homogeneous, so the Hahn-Banach Theorem provides us with $y^* \in Y^*$ such that $\langle y^*, v\rangle \leq \overline{D}f(y,v)$ for every $v \in Y$. Multiplying the inequality $\langle y^*, -v\rangle \leq \overline{D}f(y,-v)$ by -1 and using that $\underline{D}f(y,v) = -\overline{D}f(y,-v)$ by Theorem 3.2, we obtain that $\underline{D}f(y,v) \leq \langle y^*, v\rangle$, so y^* verifies the requirements from the definition of the intermediate derivative. □

References

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

[2] L. Ambrosio and G. Dal Maso, A general chain rule for distributional derivatives, *Proceedings of the AMS*, **108**, no. 3, 691–701 (1990)

[3] N. Aronszajn, Differentiability of Lipschitzian mappings between Banach spaces, *Studia Math.* **57**, no. 2, 147–190 (1976)

[4] Y. Benyamini and J. Lindenstrauss, *Geometric nonlinear functional analysis. Vol. 1.* American Mathematical Society Colloquium Publications, 48. American Mathematical Society, Providence, RI, 2000

[5] D. N. Bessis and F. H. Clarke, Partial subdifferentials, derivates and Rademacher's theorem. *Trans. Amer. Math. Soc.*, **351** (1999), no. 7, 2899–2926

[6] V. I. Bogachev and S. A. Shkarin, Differentiable and Lipschitzian mappings of Banach spaces, *Math. Notes* **44** (1988), 790–798.

[7] J. M. Borwein and W. Moors, A chain rule for essentially smooth Lipschitz functions, *SIAM J. Optim.* 8 (1998), no. 2, 300–308.

[8] G. Choquet, Applications des propriétés descriptives de la fonction contingent à la théorie des fonctions de variable réelle et à la géometrie différentielle des variétés cartésiennes, *J. Math. Pures Appl.* **26** (1947), 115–226

[9] J. P. R. Christensen, *Topology and Borel Structure*, North-Holland, Amsterdam (1974)

[10] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York 1983

[11] B. D. Craven, D. Ralph and B. M. Glover, Small convex-valued subdifferentials in mathematical programming, *Optimization*, **32**, 1–21 (1995)

[12] M. Csörnyei and P. W. Jones, *Product Formulas for Measures and Applications to Analysis and Geometry*, announced at `http://www.math.nyu.edu/research/peterjonesabstract.pdf`

[13] M. Doré and O. Maleva, A compact null set containing a differentiability point of every Lipschitz function, *Mathematische Annalen* **351** (2011), no. 3, 633–663

[14] M. Doré and O. Maleva, A compact universal differentiability set of Hausdorff dimension one in Euclidean space, *Israel Journal of Mathematics* **191** (2012), 889–900

[15] M. Doré and O. Maleva, A universal differentiability set in Banach spaces with separable dual, *Journal of Functional Analysis* **261** (2011), 1674–1710

[16] M. Fabián and D. Preiss, On intermediate differentiability of Lipschitz functions on certain Banach spaces. *Proc. Amer. Math. Soc.* **113** (1991), no. 3, 733–740

[17] M. Fabián and N. V. Zhivkov, A characterization of Asplund spaces with the help of local ε-supports of Ekeland and Lebourg. *C. R. Acad. Bulgare Sci.* **38** (1985), no. 6, 671–674

[18] H. Federer, *Geometric Measure Theory*, Springer, New York (1969)

[19] T. Fowler and D. Preiss, A simple proof of Zahorski's description of non-differentiability sets of Lipschitz functions. *Real Anal. Exchange* **34** (2009), no. 1, 127138

[20] R. Fry, Approximation by functions with bounded derivative on Banach spaces. *Bull. Aust. Math. Soc.* **69** (2004), 125–131

[21] J. R. Giles and Scott Sciffer, Locally Lipschitz functions are generically pseudo-regular on separable Banach spaces. *Bull. Austral. Math. Soc.* **47** (1993), no. 2, 205–212

[22] P. Hájek and M. Johanis, Smooth approximations. *J. Funct. Anal.* **259** (2010), no. 3, 561–582.

[23] A. D. Ioffe, Typical convexity (concavity) of Dini-Hadamard upper (lower) directional derivatives of functions on separable Banach spaces. *J. Convex Anal.* **17** (2010), no. 3-4, 1019–1032

[24] M. Johanis, Approximation of Lipschitz mappings, *Serdica Math. J.* **29** (2003), no. 2, 141–148.

[25] A. S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995.

[26] K. Kuratowski and C. Ryll-Nardzewski, A general theorem on selectors, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **13** (1965), 397–403

[27] J. Lindenstrauss, On nonlinear projections in Banach spaces. *Michigan Math. J.* **11** (1964), 263–287

[28] J. Lindenstrauss and D. Preiss, On Fréchet differentiability of Lipschitz maps between Banach spaces, *Annals of Math.* **157** (2003), no. 1, 257–288

[29] J. Lindenstrauss, D. Preiss and J. Tišer, *Fréchet Differentiability of Lipschitz Functions and Porous Sets in Banach Spaces*, Annals of Mathematics Studies, No. 179, Princeton University Press 2012

[30] P. Mankiewicz, On the differentiability of Lipschitz mappings in Fréchet spaces, *Studia Mathematica* **45** (1973), 15–29

[31] P. Michel and J.-P. Penot, Calcul sous-différentiel pour des fonctions lipschitziennes et non-lipschitziennes, *C.R. Acad. Sc. Paris, Ser.I*, **298** (1984), 269–272

[32] R. J. Nájares and L. Zajíček, A σ-porous set need not be σ-bilaterally porous, *Comment. Math. Univ. Carolin.* **35** (1994), no. 4, 697–703

[33] A. Nekvinda and L. Zajíček, A simple proof of the Rademacher theorem, *Časopis pro Pěstování matematiky* **113** (1988), 337–34l

[34] A. Pełczyński, *Linear extensions, linear averagings, and their applications to linear topological classification of spaces of continuous functions.* Dissertationes Mathematicae (Rozprawy Matematyczne) **58** (1968)

[35] D. Preiss, Differentiability of Lipschitz Functions on Banach Spaces, *Journal of Functional Analysis*, **91** (1990), 312–345

[36] D. Preiss and J. Tišer, *Two unexpected examples concerning differentiability of Lipschitz functions on Banach spaces*, GAFA Israel Seminar 92-94, Birkhäuser (1995), 219–238

[37] D. Preiss and L. Zajíček, Directional derivatives of Lipschitz functions. *Israel J. Math.* **125** (2001), 1–27.

[38] H. Rademacher, Über partielle und totale Differenzierbarkeit I, *Math. Ann.* **79** (1919), 340–359

[39] Z. Zahorski, Punktmengen, in welchen eine stetige Funktion nicht differenzierbar ist. *Rec. Math. [Mat. Sbornik] N. S.* **9 (51)** (1941), 487–510.

[40] Z. Zahorski, Sur la première dérivée, *Trans. Amer. Math. Soc.* **69** (1950), 1-54

[41] L. Zajíček, Sets of σ-porosity and sets of σ-porosity (q), *Časopis Pěst. Mat.* **101** (1976), no. 4, 350–359