1

DIFFERENTIABILITY OF CONE-MONOTONE FUNCTIONS ON SEPARABLE BANACH SPACE

JONATHAN M. BORWEIN, JAMES V. BURKE, AND ADRIAN S. LEWIS

ABSTRACT. Motivated by applications to (directionally) Lipschitz functions, we provide a general result on the almost everywhere Gâteaux differentiability of real valued functions on separable Banach spaces, when the function is monotone with respect to an ordering induced by a convex cone with non-empty interior. This seemingly arduous restriction is useful as it covers the case of directionally Lipschitz functions, and necessary: we show by way of example that most results fail more generally.

1. INTRODUCTION

Directionally Lipschitz functions play an important role in optimization theory [6], and more recently in algorithm construction [4]. It is the purpose of this brief note to establish that results about differentiability of these functions are well approached by establishing results of differentiability of real valued functions f that are $K-$*increasing*: $x \geq_K y$ implies $f(x) \geq f(y)$.

We call f $K-$*decreasing* if $-f$ is $K-$increasing, and we call f $K-$*monotone* if it is either $K-$increasing or $K-$decreasing. Clearly, f is $K-$decreasing if and only if it is $-K-$increasing. Furthermore, if f is $K-$increasing, then it remains $S-$increasing for any convex cone $S \subset K$. In particular, if K has nonempty interior, we lose no generality in assuming it is also closed, since it contains a closed cone with nonempty interior.

Thus, we provide a general result on the almost everywhere Gâteaux differentiability of real valued functions on separable Banach spaces, when the function is monotone with respect to an ordering induced by a convex cone with non-empty interior (Theorem 6). This seemingly arduous restriction is useful as it covers afore-mentioned the case of directionally Lipschitz functions, and necessary: we show by way of example that most results fail more generally.

Our note is ordered as follows. Section Two provides preliminary results about null sets and about monotone functions. Section Three establishes our main differentiability result. Section Four establishes the application to directionally Lipschitz functions. This is almost immediate after we observe that each such function locally decomposes as a sum of a monotone function with respect to an appropriate cone and a linear function (Proposition 8). In the fifth section,

[1]J.M. Borwein's research supported by NSERC and by the Canada Research Chair Programme. A.S. Lewis's research supported by NSERC. J.V. Burke's research supported by NSF DMS-9971852 & NIH P41-RR-12609

Date: November 19, 2002.

2000 *Mathematics Subject Classification.* Primary 26B25; Secondary 90C29.

Key words and phrases. Monotonicity, directionally Lipschitz functions, null sets, a.e. differentiability, cones.

an application to the computation of 'random subgradients' is presented. Finally in Section Six various limiting examples are given.

2. Preliminaries on Measure and on Monotonicity

Throughout we consider a separable Banach space Y, partially ordered by a nonempty closed convex cone $K \subset Y$: $x \geq_K y$ if $x - y \in K$. We make the analogous definition for $\leq_K$. Where clear, we write $\leq$ for $\leq_K$. We write $[a, b]_K$ for the *order interval* $\{x : a \leq_K x \leq_K b\}$.

We begin by recapturing the situation in finite dimensions. Recall that a function is *Hadamard* differentiable if it is Gâteaux differentiable uniformly on norm compact sets of directions. In finite dimensions, this coincides with Fréchet differentiability.

Theorem 1 (Monotone functions in finite dimensions [5]). *Suppose that $f : \mathbb{R}^n \mapsto \mathbb{R}$ satisfies $f(x) \leq f(y)$ whenever $x_i \leq y_i$, $i = 1, 2, ...n$. Then:*

(a) *f is measurable.*
(b) *If, for some d with $d_i > 0$ for $i = 1, 2, ...n$, the function $t \mapsto f(x_0 + td)$ is lower semicontinuous at $t = 0$, then f is lower semicontinuous at x_0. Similarly for upper semicontinuity.*
(c) *f is almost everywhere continuous.*
(d) *If f is Gâteaux differentiable at x_0, then it is Hadamard differentiable at x_0.*
(e) *Let $\underline{f}$ be the lower semicontinuous hull of f. Then f is continuous at x_0 if and only if $\underline{f}$ is. Similarly, f is Gâteaux differentiable at x_0 if and only $\underline{f}$ is, and if these functions are Gâteaux differentiable, their derivatives agree.*
(f) *f is almost everywhere Hadamard differentiable.*

Although no Haar measure exists on an infinite dimensional Banach space, various classes of null sets can be defined and exploited: see Benyamini and Lindenstrauss [1].

The Banach space version of (c) and (f) of Theorem 1, proved below in Theorem 6, requires such a notion of a null set in a Banach space. We proceed to make the following definitions—for details on these, and other measure-related notions we use in the paper, we refer the reader to Benyamini and Lindenstrauss [1] (see also Borwein and Moors [3]). Indeed, while (a) of Theorem 1 is not especially well phrased for infinite dimensions, all the parts remain true, appropriately interpreted. We shall only prove in detail the parts central to our main task.

Let X be a separable Banach space. A probability measure μ on X is *Gaussian* if for every $x^* \in X^*$, the measure μ_{x^*} on the real line, defined by $\mu_{x^*}(A) = \mu\{y \mid \langle x^*, y \rangle \in A\}$, has a Gaussian distribution. It is additionally called *nondegenerate* if for every $x^* \neq 0$ the distribution μ_{x^*} is nondegenerate. A Borel set $C \subset X$ is called *Gauss null* if $\mu(C) = 0$ for every nondegenerate Gaussian measure on X. It is known that the set of points where a given Lipschitz function $f : X \mapsto \mathbb{R}$ is not Gâteaux (Hadamard) differentiable is Gauss null. This in fact holds for functions with values in a space with the Radon-Nikodym property (Benyamini and Lindenstrauss [1] Theorem 6.42), while it fails completely for the stronger notion of Fréchet differentiability.

A larger class is that of *Haar null* sets. These are Borel sets A such that there is a Borel probability measure μ satisfying $\mu(A+x)=0$ for all $x \in X$. In finite dimensions these coincide with Gaussian null sets (i.e. Lebesgue null sets).

Each of these classes is closed under countable unions. This and the following weak forms of Fubini's theorem is what gives 'nullness' its utility.

Recall that a set is *universally (Radon) measurable* if it lies in the completion of the Borel $\sigma-$algebra for each Borel probability measure on X. Similarly we say that the set is *Gaussian measurable* if it lies in the completion of the Borel algebra for all Gaussian measures on X.

Theorem 2 ("Fubini"). *(a)* [3] *Suppose* $(H,+,\tau_1)$ *is a completely metrizable Abelian group and* $(T,+,\tau_2)$ *is a locally compact Polish Abelian group. Then for each universally Radon measurable subset* $E \subseteq H \times T$ *(the product group), the following are equivalent:*

(i) $\{t \in T : (h,t) \in E\}$ *is a Haar-null set, for the Haar-measure on* T*, for almost all* $h \in H$*;*
(ii) *The set* E *is a Haar-null set in the product group* $H \times T$*.*

(b) [1]*(Prop. 6.29) Suppose that* E *is a Borel subset of a separable Banach space* X*, and suppose that there is a finite dimensional subspace* F *of* X *such that* $\mu_F((E+x)\cap F)=0$ *for all* $x \in X$*. Then* E *is Gaussian (and so Haar) null. (Here* μ_F *is (equivalent to) Lebesgue measure on* F*.)*

The most immediate application of part (a) is to the case where T and H are Banach spaces and T is finite dimensional.

While our results and definitions can be adjusted to extended real-valued functions, for clarity we consider only finite-valued functions on an open set A in X. Since the results are largely local in nature this presents no real inconvenience. The *upper* and *lower Dini-derivatives* of $f : X \mapsto \mathbb{R}$ at x are given by:

$$f^+(x;v) := \limsup_{t\downarrow 0} \frac{f(x+tv)-f(x)}{t} \quad \text{and } f^-(x;v) := \liminf_{t\downarrow 0} \frac{f(x+tv)-f(x)}{t}.$$

We write $f'(x,v)$ for the common limit when it exists.

We note that both $f^-(x;v)$ and $f^+(x;v)$ are $K-$increasing in v whenever f is K-increasing. More significantly, we have the following two results, generalizing Theorem 1 (b) and (e) respectively. Parts (c), (d) and (f) are established in Theorem 6.

The analogue of (a), given in Corollary 7, is a little more subtle since even on $\mathbb{R}^3$ a cone-monotone function need not be universally measurable. Consider the convex cone K generated by $\{(x,y,1)\colon x^2+y^2<1 \text{ or } (x,y)\in A\}$, where A is any non-measurable (Lebesgue) subset of the unit circle, S, and let $1-\chi_{-K}$ be the $K-$monotone function which is zero for $x \in -K$ and is one otherwise. Let λ_S be normalized Lebesgue measure on S. Then, consideration of the Borel measure defined by $\mu(C)=\lambda_S(S\cap C)$ shows that K is not μ-measurable and so is not universally measurable [3].

Proposition 3. *Let X be a normed space containing a convex cone K with non-empty interior. Suppose f is $K-$monotone on A, and that for some direction $d \in \operatorname{int} K$ the function $t \mapsto f(x_0 + td)$ is lower (resp. upper) semicontinuous at $t = 0$, then f is lower (resp. upper) semicontinuous at x_0. In particular, f is continuous at x_0 whenever $t \mapsto f(x_0 + td)$ is.*

Proof. For $d \in \operatorname{int} K$, the order interval $N := [-d, d]_K$ is a symmetric neighbourhood of zero. We observe that for $u \in N$, and $|t| < \varepsilon$, we have $f(x_0 - \varepsilon d) \leq f(x_0 + tu) \leq f(x_0 + \varepsilon d)$, and the result follows easily. □

Proposition 4. *Let X be a normed space containing a convex cone K with non-empty interior. Suppose f is $K-$monotone. Let*

$$\underline{f}(x) := \sup_{\delta > 0} \inf_{z \in B_\delta(x)} f(z)$$

denote the lower semicontinuous hull of f. (i) Then f is continuous at x_0 if and only if $\underline{f}$ is. (ii) Similarly, f is Gâteaux differentiable at x_0 if and only $\underline{f}$ is, and if these functions are Gâteaux differentiable, their derivatives agree.

Proof. Without loss we may assume K is closed and we note that $\underline{f}$ is then also $K-$monotone. Fix x and $d \in \operatorname{int} K$. Let g denote the function $t \mapsto f(x + td)$. Since $N := [-d, d]_K$ is a neighbourhood of zero, it follows that $\underline{f}(x + td) = \underline{g}(t)$ for all t. By Theorem 1 (e), $\underline{g}$ is continuous at 0 iff g is. In combination with Proposition 3 we have established (i).

To prove (ii), take any two directions in X, say h and k, and consider two directions c and d in $\operatorname{int} K$ such that $\operatorname{span}\{c, d\} = F := \operatorname{span}\{h, k\}$. Let g denote the function $(s, r) \mapsto f(x + sc + rd)$. Much the same argument as above shows $\underline{f}(x + sc + rd) = \underline{g}(s, r)$ for all (s, r). We now see that $f'(x; \cdot)$ exists and is linear on the subspace F if and only if $g'(0; \cdot)$ exists and is linear, and the parallel statement holds for $\underline{f}$ and $\underline{g}$. Now the two-dimensional version of Theorem 1 (e) along with Proposition 3, applied to g, shows that f and $\underline{f}$ restricted to the arbitrary two-dimensional subspace F have the same linear gradients, so we are done. □

3. Main Result

Unlike differentiability results for convex functions, where Baire category is very useful, results about monotone functions seem intrinsically to be of a measure theoretic nature. The following Theorem may be viewed as an extension to separable Banach space of the Lebesgue monotone differentiability theorem, see [8]. This was extended to monotone functions on $\mathbb{R}^2$ by Saks [7] and first explicitly stated for $\mathbb{R}^n$ in [5], although as we shall see this is really immediate from Sak's result. A preparatory lemma is useful.

Lemma 5. *Suppose that K is a cone with nonempty interior in a normed space X and that Y is a dense subspace of X. For each $u \in X$, there are sequences $w_k, z_k \in Y$ such that*

(1) $$w_k \leq_K u \leq_K z_k \quad \text{and} \quad w_k \to u, z_k \to u.$$

Proof. Let $y_k \in Y$ converge to u. Since Y is a dense subspace there is $e \in Y \cap \operatorname{int} K$. Then $-t_k e \leq_K u - y_k \leq_K t_k e$ for a sequences of real numbers $t_k \to 0$, as $[-e,e]_K$ is a zero-neighbourhood. Then $z_k := y_k + t_k e$, $w_k := y_k - t_k e$ are as desired. □

Let us introduce temporary notation, for $d \in X$:

$$f'_{\mathbb{Q}}(x,d) = \lim_{0 \leftarrow q \in \mathbb{Q}} \frac{f(x+qd) - f(x)}{q},$$

when this limit exists. We also observe that when f is lower semicontinuous $-f'_{\mathbb{Q}}(x,-d) = f'_{\mathbb{Q}}(x,d) = f^+(x,d)$, while if also f is continuous, or K-increasing with $d \in K$, then $f'_{\mathbb{Q}}(x,d) = f'(x,d)$, and $\{x\colon f'(x,d) \text{ exists}\}$ is a Borel set since when f is Borel measurable

(2) $$\left\{x\colon f'_{\mathbb{Q}}(x,d) \text{ exists}\right\} =$$

$$\bigcap_{n\in\mathbb{N}} \bigcup_{m\in\mathbb{N}} \left\{x\colon \frac{f(x+rd) - f(x)}{r} < \frac{f(x+sd) - f(x)}{s} + \frac{1}{n},\ \forall r,s \in \mathbb{Q}, |r|,|s| < \frac{1}{m}\right\}.$$

Theorem 6. *Let Y be a separable Banach space and let $K \subset Y$ be a convex cone with non-empty interior. Suppose $f : Y \mapsto \mathbb{R}$ is $K-$monotone. Then f is continuous and (Hadamard) differentiable except at the points of a Gaussian null set.*

Proof. Without loss of generality, suppose f is K-increasing. We begin by assuming f is lower semicontinuous. One argument uses [5], but we proceed directly.

Let S be a countable dense subset of int K that is $\mathbb{Q}-$convex, (i.e., closed under sums and positive rational multiples). For $h,k \in S$ define

(3) $$D(h,k) := \{x\colon f'_Q(x,v) \text{ exists and is linear for } v \in \operatorname{span}_{\mathbb{Q}}\{h,k\}\}.$$

i). We assert that $D(h,k)$ is a Borel set, much as in (2), because f is lower semicontinuous. (We may assume h and k are independent.) A monotonicity argument now shows equation (3) also holds with $\operatorname{span}_{\mathbb{Q}}$ replaced by $\operatorname{span}_{\mathbb{R}} = \operatorname{span}$. In fact, we can deduce

$$D(h,k) = \{x\colon f'(x,v) \text{ exists and is linear for } v \in \operatorname{span}\{h,k\}\}$$

once we show $f^+(x,u) = f^-(x,u)$ for each $x \in D(h,k)$ and $u \in \operatorname{span}\{h,k\}$. From above, we need only consider the case that $u = h - \alpha k$, $\alpha > 0$ (the case $\alpha < 0$ is similar). Consider $\beta \in \mathbb{Q}$, $\alpha < \beta$. Now suppose $r_n \downarrow 0$ satisfies

$$f^-(x,u) = \lim \frac{f(x+r_n u) - f(x)}{r_n}$$

and select rationals $q_n \downarrow 0$ with $r_n(h - \alpha k) \geq_K q_n(h - \beta k)$, and $q_n/r_n \to 1$. It follows that

$$\begin{aligned} f^-(x,u) &= \liminf_n \frac{f(x+r_n u) - f(x)}{r_n} \geq \liminf_n \frac{f(x+q_n(h-\beta k)) - f(x)}{q_n} \\ &= f'_{\mathbb{Q}}(x, h-\beta k) = f'_{\mathbb{Q}}(x,u) + (\alpha-\beta) f'_{\mathbb{Q}}(x,k) = f^+(x,u) + (\alpha - \beta) f'_{\mathbb{Q}}(x,k). \end{aligned}$$

Since $0 < \alpha < \beta$ is arbitrary we are done.

ii). We now show the complement of $D(h,k)$ is Gaussian null. This follows from Theorem 2 (b) and the two dimensional version of Theorem 1. Thus

$$D := \bigcap_{k,h \in S} D(h,k)$$

has a null complement. It remains to show that f is Hadamard differentiable at points of D. By Proposition 4 we can then drop the assumption that f is semicontinuous.

iii). First, we observe that for $x \in D$, $\lambda := v \mapsto f^+(x,v)$ determines a linear functional on any dense subspace, Z containing S, of countable Hamel dimension. To see this, we write $Z := \cup_n Z_n$ for some increasing sequence Z_n of subspaces of dimension n. Now $\lambda \in K^+$ is necessarily linear on any finite dimensional subspace of Z because this holds on all two-dimensional subspaces (3), since S was supposed $\mathbb{Q}$–convex.

Next we observe that λ is continuous on Z. We fix $e \in Z \cap \operatorname{int} K$ and write

$$\lambda(y) = f'_{\mathbb{Q}}(x,y) \leq f^+(x,e) < \infty$$

for all y in the zero neighbourhood $[-e,e]_K \cap Z$. Hence, λ is bounded above on a neighbourhood of the origin and necessarily is continuous.

iv). It follows that λ extends to a continuous linear function on Y, which we still denote by λ . It remains to show that λ is indeed the Gâteaux derivative of f at x. To show this we fix $x \in Y$ and then appeal to Lemma 5. For w_k and z_k as guaranteed by (1) we may write

$$\begin{aligned} f^+(x,u) &\leq f^+(x,z_k) = \lambda(z_k) \to \lambda(u) \\ f^-(x,u) &\geq f^-(x,w_k) = \lambda(w_k) \to \lambda(u), \end{aligned}$$

and in combination these show

$$f^+(x,u) \leq \lambda(u) \leq f^-(x,u).$$

Thus, the Gâteaux derivative indeed exists as claimed.

v). In particular, for d interior to K and for x in D, $t \mapsto f(x+td)$ is continuous at 0. Then, Proposition 3 establishes that f is indeed continuous at each $x \in D$.

vi). Finally, we confirm that (with no semicontinuity or separability assumption), since K has nonempty interior, λ is also the Hadamard derivative of f at x. This relies on the fact that for $\varepsilon > 0$, if $d_n \to d$ and $t_n \to 0$ then

$$\frac{f(x+t_n(d+\varepsilon e)) - f(x)}{t_n} \geq \frac{f(x+t_n\,d_n) - f(x)}{t_n} \geq \frac{f(x+t_n(d-\varepsilon e)) - f(x)}{t_n}$$

for n large and $e \in \operatorname{int} K$. Now we observe that

$$\begin{aligned} \liminf_n \frac{f(x+t_n\,d_n) - f(x)}{t_n} &\geq \lambda(d-\varepsilon e) \\ \limsup_n \frac{f(x+t_n\,d_n) - f(x)}{t_n} &\leq \lambda(d+\varepsilon e), \end{aligned}$$

and let ε go to zero. □

We complete the section by establishing the Banach space version of Theorem 1 (a).

Corollary 7. *Let Y be a separable Banach space and let $K \subset Y$ be a convex cone with non-empty interior. Suppose $f : Y \mapsto \mathbb{R}$ is $K-$monotone. Then f is Gaussian measurable.*

Proof. Fix $r \in \mathbb{R}$ and write $L := \{x\colon f(x) < r\}$ and let N denote a Gaussian null Borel set such that $\{x\colon \underline{f}(x) < f(x)\} \subset N$, as exists by Theorem 6. Let $M := \{x\colon \underline{f}(x) < r, x \notin N\}$. Then M is Borel and

$$M \subset L \subset M \cup N.$$

where N is null for all Gaussian measures. □

We are now ready for our applications.

4. Applications to Directionally Lipschitz Functions

We call a locally lower semicontinuous function $f : A \to \mathbb{R}$ *fully directionally Lipschitz* at x in direction u if there is $\varepsilon > 0$ such that for $\|h-u\| \le \varepsilon$, $\|z-x\| \le \varepsilon$, and $0 < t \le \varepsilon$ one has

$$\frac{f(z+th) - f(z)}{t} < M, \tag{4}$$

for some finite number M. Recall that f is *directionally Lipschitz* if inequality (4) holds under the additional assumption $|f(z) - f(x)| \le \varepsilon$. Thus continuous directionally Lipschitz functions are fully directionally Lipschitz. Notice in particular that every (globally) Lipschitz function is (globally) fully directionally Lipschitz in all directions. It is now an easy matter to establish:

Proposition 8. *A locally lower semicontinuous function f on A is fully directionally Lipschitz if (and only if) it is locally representable as*

$$f = g + l$$

where g is monotone with respect to a convex cone with interior (and bounded base), and l is linear. Moreover, if the function is globally (fully directionally) Lipschitz on A, then such a decomposition holds globally.

Proof. Suppose f is fully directionally Lipschitz at x in direction u, so inequality (4) holds. We can assume $\|u\| = 1$, and $\epsilon < 1/4$. Choose a norm-one continuous linear functional ϕ such that $\phi(u) > 1 - \epsilon$. Denote the closed unit ball in X by B. Then all $y \in u + \epsilon B$ satisfy $\phi(y) > 1 - 2\epsilon$ and $\|y\| \le 1 + \epsilon$, and hence $\phi(y) > \epsilon\|y\|$. The closed cone K generated by $u + \epsilon B$ is convex, and we see that $\phi(k) \ge \epsilon\|k\|$ for all $k \in K$ (so K is certainly pointed).

Now suppose $\|z - x\| \le \epsilon$ and $k \in \mathbb{R}_+(u + \epsilon B) \subset K$ with $\|k\| \le \epsilon/2$. Then $k = th$ for some $t \in \mathbb{R}_+$ and $h \in u + \epsilon B$, so $\|h\| \ge 1 - \epsilon$, and hence

$$t = \frac{\|k\|}{\|h\|} \le \frac{\epsilon}{2(1-\epsilon)} < \epsilon.$$

By our fully directionally Lipschitz assumption, we deduce

$$\begin{aligned} f(z+k)-f(z) &= f(z+th)-f(z) \\ &< Mt \le \frac{M\phi(th)}{\epsilon\|h\|} \le \frac{M\phi(k)}{\epsilon(1-\epsilon)} < \frac{2M}{\epsilon}\phi(k). \end{aligned}$$

Lower semicontinuity now implies

$$f(z+k)-f(z) \le \frac{2M}{\epsilon}\phi(k)$$

whenever $\|z-x\| \le \epsilon$ and $k \in K$ with $\|k\| \le \epsilon/2$. If we now define the continuous linear functional $l := 2\epsilon^{-1}M\phi$, then the function $g = f - l$ is locally K-decreasing.

The converse is immediate, and the global analogue follows by much the same argument. □

Notice that the above result fails for the real function

$$f(x) = \begin{cases} 0 & \text{if } x \le 0 \\ 1-\sqrt{x} & \text{if } x > 0, \end{cases}$$

which is directionally Lipschitz but not fully directionally Lipschitz.

Corollary 9. *Every fully directionally Lipschitz function f on a separable Banach space is almost everywhere Hadamard differentiable.*

Proof. Since separable Banach spaces are Lindelöf spaces, it suffices to show that the function is locally differentiable a.e. Now this is a direct application of Proposition 8 and Theorem 6. □

We thus recapture the Banach space version of Rademacher's result that locally Lipschitz functions on $\mathbb{R}^n$ are differentiable a.e. (see [1]). A first consequence is that the boundary of a convex set C with nonempty interior must be Gaussian null, since the metric distance function, given by $d_C(x) := \inf_{x\in C}\|x-c\|$, is Lipschitz and nondifferentiable at any boundary point of C.

We also, note that part vi). of the proof of Theorem 6 actually establishes that Gâteaux and Hadamard differentiability coincide for functions monotone with respect to a cone with non-empty interior in Banach space, and so for fully directionally Lipschitz functions.

5. Random Subgradients

Our interest in the differentiability of directionally Lipschitz functions arises partly from recent work on stochastic approximations to the Clarke subdifferential. We sketch the ideas below.

A fundamental observation due to Clarke, using the almost everywhere differentiability of every Lipschitz function f on $\mathbb{R}^n$, was that Clarke subgradients are precisely the limits of convex combinations of gradients at nearby points. This property is often the most convenient way to compute the Clarke subdifferential $\partial_C f(\bar{x})$. Since, as we have seen, fully directionally Lipschitz functions share the property of almost everywhere differentiability, we can ask whether the same idea extends to this case. The following result gives an affirmative answer. It can be viewed as legitimating the sampling of *random (sub-)gradients.*

Theorem 10 (Gradient-based approximation). *Suppose that, near the point $\bar{x} \in \mathbb{R}^n$, the continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is directionally Lipschitz and absolutely continuous on lines. If Q is a full-measure subset of a neighbourhood of $\bar{x}$ consisting of points where f is differentiable, then*

$$\partial_C f(\bar{x}) = \bigcap_{\delta>0} \operatorname{cl}\operatorname{conv} \nabla f(Q \cap (\bar{x} + \delta B)).$$

Proof. The proof relies on Theorem 6 via Corollary 9, and is given in full in [4]. □

The example $f(x) := \sqrt{(\|x\| - 1)^+}$ (for any unit vector $\bar{x}$) shows that this result may fail for functions that are not directionally Lipschitz.

Consider, then, a continuous function $f : \mathbb{R}^n \to \mathbb{R}$ that is directionally Lipschitz around a point $\bar{x}$. We can try to approximate $\partial_C f(\bar{x})$ as follows. We choose a small "sampling radius" $\delta > 0$, and a random sequence of independent points $x_1, x_2, \ldots$, uniformly distributed on the ball $\bar{x} + \delta B$. We then construct a corresponding increasing sequence of closed convex random sets

$$C_k = \operatorname{conv}\{\nabla f(x_i) : i = 1, 2, \ldots, k\} \ (k = 1, 2, \ldots).$$

The result above suggests that these sets may approximate $\partial_C f(\bar{x})$ well. From a variety of corroborative results along these lines, we illustrate with one.

Theorem 11 (False positives). *If the continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is directionally Lipschitz around the point $\bar{x}$ and $0 \notin \partial_C f(\bar{x})$, then for any sufficiently small sampling radius we have*

$$\lim_{k\to\infty} \operatorname{dist}(0, C_k) > 0 \textit{ almost surely.}$$

This technique of approximation is useful for functions whose gradients are easy to compute when they exist, but whose Clarke subdifferentials may be less tractable. A good example is the spectral abscissa of a square matrix (the largest of the real parts of the eigenvalues).

6. Limiting Examples and Conclusion

Examples 12. Theorem 6 and Corollary 9 fail completely outside of separable space, when the cone has empty interior, or when the directionally Lipschitz functions are supposed merely Hölder continuous. This is discussed in detail in [2]. Indeed:

1. No useful information is possible if $K - K \neq X$.
2. Even continuous convex functions in non-separable spaces may be nowhere Gâteaux differentiable, and in non-reflexive spaces may be nowhere Fréchet differentiable.
3. In every non-reflexive space there is a non-null (generating) cone K and a K-monotone, quasi-convex, lsc function which is not a.e. Haar continuous on lines, and hence is not a.e. Gâteaux differentiable. The simplest example is the indicator function $1 - \chi_{-c_0^+}$ which is zero for non-positive sequences in c_0 and is 1 elsewhere.

4. Perhaps the most striking open question is: *does there exist a real valued coordinatewise monotone continuous function on $\ell_2(\mathbb{N})$ with no points of Gâteaux differentiability?* By comparison, in ℓ_2, let J be the order interval $[-1/2^n, 0]$. By 6.24 in [1] this not Gaussian null. (It contains $\overline{\mathrm{co}}(-e_n/2^n.)$ Let $f(x) := \sqrt{\|x^+\|_2}$. Then for $x \in J$ and $h := (2^{-n/2})$, $t_n := 2^{1-n/2}$,
$$\frac{f(x+t_nh)-f(x)}{t_n} \geq \frac{\sqrt{x_n+th_n}}{t_n} \geq \frac{1}{2}$$
and so f is nowhere differentiable in direction h on J.

Rademacher's theorem remains valid when the range of the function is a Banach space with the RNP. It is not clear what is true for cone-monotone operators.

We conclude by commenting that as in [5] one can extend the results to an extended real-valued function, say g, if one appropriately defines differentiability at $\pm\infty$. This is quite easily obtained by composing g with a function such as arctan. Relatedly, one may exploit the fact the $\max(f,0)$ and $\min(f,0)$ are $K-$monotone when f is.

References

[1] Y. Benyamini and J. Lindenstrauss, *Geometric Nonlinear Functional Analysis, Volume 1*, Volume 48, AMS Colloquium Publications, Providence RI, 2000.

[2] J.M. Borwein and R. Goebel, "On the nondifferentiability of cone-monotone functions in Banach spaces," CECM preprint 02:179, March 2002.

[3] J.M. Borwein and W.B. Moors, "Null sets and essentially smooth Lipschitz functions," *SIAM J. Optim.* **8** (1998), 309–323.

[4] J.V. Burke, A.S. Lewis and M.L. Overton, "Approximating subdifferentials by random sampling of gradients", *Mathematics of Operations Research*, (2002), in press.

[5] Y. Chabrillac, and J.-P. Crouzeix, "Continuity and differentiability properties of monotone real functions of several real variables." in *Nonlinear analysis and optimization* (Louvain-la-Neuve, 1983). Math. Programming Study No. 30 (1987), 1–16.

[6] R. T. Rockafellar, R. J-B Wets, *Variational Analysis,* Springer-Verlag, New York, 1998.

[7] S. Saks, *Theory of the Integral,* English translation, second edition, Stechert, New York, 1937.

[8] K.R. Stromberg, *An Introduction to Classical Real Analysis*, Wadsworth, 1981.

J.M. Borwein, Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6

E-mail address: `jborwein@cecm.sfu.ca`

J.V. Burke, Department of Mathematics, University of Washington, University of Washington, Department of Mathematics, Box 354350 Seattle, Washington 98195-4350 USA

E-mail address: `burke@math.washington.edu`

A.S. Lewis, Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6

E-mail address: `aslewis@cecm.sfu.ca`