

GENERALIZED SUBDIFFERENTIALS: A BAIRE CATEGORICAL APPROACH

JONATHAN M. BORWEIN, WARREN B. MOORS, AND XIANFU WANG

ABSTRACT. We use Baire categorical arguments to construct pathological locally Lipschitz functions. The origins of this approach can be traced back to Banach and Mazurkiewicz (1931) who independently used similar categorical arguments to show that “almost every continuous real-valued function defined on $[0,1]$ is nowhere differentiable”. As with the results of Banach and Mazurkiewicz, it appears that it is easier to show that almost every function possesses a certain property than to construct a single concrete example. Among the most striking results contained in this paper are: Almost every 1-Lipschitz function defined on a Banach space has a Clarke subdifferential mapping that is identically equal to the dual ball; if $\{T_1, T_2, \dots, T_n\}$ is a family of maximal cyclically monotone operators defined on a Banach space X then there exists a real-valued locally Lipschitz function g such that $\partial_0 g(x) = \text{co}\{T_1(x), T_2(x), \dots, T_n(x)\}$ for each $x \in X$; in a separable Banach space each non-empty weak* compact convex subset in the dual space is identically equal to the approximate subdifferential mapping of some Lipschitz function and for locally Lipschitz functions defined on separable spaces the notions of strong and weak integrability coincide.

1. INTRODUCTION

An important aspect of developing a mathematical theory is in producing both examples and counterexamples that illuminate the content and boundaries of the subject. In this paper we give a general method for constructing examples and counterexamples for the differentiability theory of Lipschitz functions.

The first and perhaps best known counterexample in differentiability theory is the construction of a continuous nowhere differentiable function. The explicit constructions given in the 19th century were later (in 1931) augmented by the use of Baire categorical arguments. Since this time the use of Baire category for the construction of functions (either well behaved or pathological) has been applied to several areas of analysis, (see, [12], [13] and [23] to name but a few). In this paper we continue this tradition by using Baire category arguments to construct Lipschitz functions that have ‘large’ generalized derivatives. The first and most crucial step towards achieving this result is to produce a candidate complete metric space on

Received by the editors March 24, 1999 and, in revised form, February 25, 2000.

1991 *Mathematics Subject Classification.* Primary 49J52, 54E52.

Key words and phrases. Subdifferentials, differentiability, Baire category, upper semi-continuous set-valued map, T -Lipschitz function.

Research of the first author was supported by NSERC and the Shrum endowment of Simon Fraser University.

Research of the second author was supported by a Marsden fund grant, VUW 703, administered by the Royal Society of New Zealand.

which we may apply the Baire category theorem. Here we consider the class of T -Lipschitz functions, denoted \mathcal{X}_T . Loosely speaking (though precise in a smooth space), for a weak* cusco $T : X \rightarrow 2^{X^*}$ from a Banach space X into its dual X^* , we say that a locally Lipschitz function f on X is T -Lipschitz if $\nabla f(x) \in T(x)$ whenever $\nabla f(x)$ exists. In Lemma 5 we show that for a fixed weak* cusco T the set of all T -Lipschitz functions form a complete metric space (under an appropriately defined metric). This simple result provides the basis which enables us to derive a plentiful supply of both examples and counterexamples.

Notation. For a normed linear space $(X, \|\cdot\|)$, we denote by X^* its dual space;

$$\begin{aligned} B_X &:= \{x \in X : \|x\| \leq 1\}; & S_X &:= \{x \in X : \|x\| = 1\}; \\ B_{X^*} &:= \{x^* \in X^* : \|x^*\| \leq 1\}; & B_\delta(x) &:= \{y \in X : \|x - y\| < \delta\}; \\ B_\delta[x] &:= \{y \in X : \|x - y\| \leq \delta\}; & I_{B_\delta[x]}(y) &:= \begin{cases} 0 & \text{if } y \in B_\delta[x], \\ +\infty & \text{otherwise.} \end{cases} \end{aligned}$$

For a non-empty subset E of X^* we denote by \overline{E}^{w^*} the weak* closure of E ; $\overline{\text{co}}^{w^*} E$ the weak* closed convex hull of E . In a topological space (A, τ) we shall denote by \mathcal{B}_A the Borel sets on A , that is, the σ -algebra generated by the open subsets of A and as usual λ will denote the Lebesgue measure.

The structure of the paper is as follows. In the remainder of Section 1 we will review some basic facts concerning the Clarke and approximate subdifferentials, then in Section 2 we will derive the basic properties of the T -Lipschitz functions. In Section 3 we will show that in any separable Banach space $G_f := \{g \in \mathcal{X}_T : \partial_a f(x) \subseteq \partial_a g(x) \text{ for all } x \in A\}$ is residual in (\mathcal{X}_T, ρ) for each $f \in \mathcal{X}_T$ and then derive some of the consequences of this result. Section 4 deals with extending the results from Section 3 to non-separable Banach spaces, while in Section 5 we briefly look at the question of how to determine when $\mathcal{X}_T \neq \emptyset$. This section also examines the question of the existence of Lipschitz functions with ‘minimal’ subdifferential mappings. This provides a contrast to the results contained in Sections 3 and 4.

We begin by recalling some preliminary definitions and properties of locally Lipschitz functions defined on Banach spaces.

1.1. Uscos and cuscusos. Let T be a set-valued mapping from a topological space A into the dual of a normed linear space X . We say that T is *weak* upper semi-continuous* on A if for each weak* open subset W of X^* , $\{x \in A : T(x) \subseteq W\}$ is open in A . When the images of T are non-empty and compact we call T a *weak* usco* and if, in addition, the images of T are also convex, then we call T a *weak* cusco*. We call T a *minimal weak* usco (cusco)* if its graph does not properly contain the graph of any other weak* usco (cusco) on A . By the *graph* of T we mean the set $\text{Gr}(T) := \{(x, x^*) : x^* \in T(x)\}$, which is closed whenever T is an usco. The following result from [1] enables us to generate uscous and cuscous from densely defined set-valued mappings.

Lemma 1 ([1]). *Let T be a densely defined set-valued mapping that maps from a topological space A into the dual of a Banach space X . If T is locally bounded on A then there exists a unique smallest weak* usco (weak* cusco) containing T , denoted $\text{USC}(T)$ ($\text{CSC}(T)$) and given by*

$$\text{USC}(T)(x) := \bigcap \{\overline{T(V)}^{w^*} : V \text{ is an open neighbourhood of } x\},$$

$$CSC(T)(x) := \bigcap \{ \overline{\text{co}}^{w^*} T(V) : V \text{ is an open neighborhood of } x \}.$$

The next result reveals one of the key properties enjoyed by minimal uscós (cuscos).

Lemma 2 ([6]). *Let A be a non-empty open subset of a Banach space X . If $T : A \rightarrow 2^{X^*}$ is a weak* usco (weak* cusco) and $S : A \rightarrow 2^{X^*}$ is a minimal weak* usco (weak* cusco) and $T(x) \cap S(x) \neq \emptyset$ for each $x \in A$, then $S(x) \subseteq T(x)$ for all $x \in A$.*

1.2. Subderivatives and subdifferentials. Let $f : A \subseteq X \rightarrow \mathbb{R}$ be a locally Lipschitz function defined on a non-empty open set A of a Banach space X . The Clarke derivative of f at $x \in A$, [11] is given by

$$f^0(x, v) := \limsup_{t \downarrow 0, y \rightarrow x} \frac{f(y + tv) - f(y)}{t}.$$

The upper and lower Dini-derivatives of f at x are given by

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

The corresponding generalized subdifferentials are defined by

$$\partial_{\sharp} f(x) := \{x^* \in X^* : x^*(v) \leq f^{\sharp}(x, v) \text{ for all } v \in X\},$$

where \sharp is one of $0, +, -$. When X is a smooth Banach space (i.e., has an equivalent Gâteaux differentiable renorm), the approximate subdifferential [3] is given by, $\partial_a f(x) := USC(\partial_- f)(x)$. If in addition (X^*, weak^*) is angelic (e.g. when X is weakly Lindelöf determined [14], which includes weakly compactly generated spaces) then,

$$\partial_a f(x) = \{x^* : x^* = w^*\text{-}\lim_{x_n \rightarrow x} x_n^*, \text{ and } x_n^* \in \partial_- f(x_n)\}.$$

In all cases $\partial_0 f$ is a weak* cusco on A and $\partial_a f$ is a weak* usco on A .

2. BASIC PROPERTIES OF THE SPACE OF T -LIPSCHITZ FUNCTIONS

Let T be weak* cusco that maps from a non-empty open subset A of a Banach space X into its dual space X^* . For such a cusco mapping one may consider the following (possibly empty) set of locally Lipschitz functions defined on A , called the T -Lipschitz functions on A .

$$\mathcal{X}_T := \{f \in \mathbb{R}^A : f \text{ is locally Lipschitz and } \partial_0 f(x) \subseteq T(x) \text{ for all } x \in A\}.$$

When X is smooth, we have by [21] the following simplified definition:

$$\mathcal{X}_T = \{f \in \mathbb{R}^A : f \text{ is locally Lipschitz and } \nabla f(x) \in T(x) \text{ whenever } \nabla f(x) \text{ exists}\}.$$

On \mathcal{X}_T , we may define a metric ρ by $\rho(f, g) := \min\{1, d(f, g)\}$, where $d(f, g) := \sup_{x \in A} |f(x) - g(x)|$. If T is identically equal to some non-empty weak* compact, convex subset C of X^* then we may simply write \mathcal{X}_C in place of \mathcal{X}_T . Some basic properties of T and \mathcal{X}_T are

Lemma 3. *Let X be a Banach space and let A be a metric space, then each weak* usco from A into X^* is locally bounded on A .*

Proof. Let us assume, in order to obtain a contradiction, that T is not locally bounded on A . Then there exists a point $x_0 \in A$ such that for each $n \in \mathbb{N}$ the set $T(B_{1/n}(x_0)) \setminus nB_{X^*} \neq \emptyset$. Using this we may construct two sequences $(x_n : n \in \mathbb{N})$ in A and $(x_n^* : n \in \mathbb{N})$ in X^* so that $(x_n : n \in \mathbb{N})$ converges to x_0 and $x_n^* \in T(x_n) \setminus nB_{X^*}$ for all $n \in \mathbb{N}$. Now if we set $K := \{x_0, x_1, x_2, \dots, x_n, \dots\}$, then K is compact and so $T(K)$ is weak* compact. Therefore, by the uniform boundedness theorem $T(K)$ is bounded, which is impossible since $\{x_n^* : n \in \mathbb{N}\} \subseteq T(K)$ is unbounded. Hence, T must be locally bounded on A . \square

With a little extra effort one can show that each weak* usco mapping from a q -space, [19], into X^* is locally bounded. However, we have no need for this extra generality here.

Lemma 4. *Let A be a non-empty open subset of a normed linear space X and let $T : A \rightarrow 2^{X^*}$ be a weak* cusco on A , then \mathcal{X}_T is a convex sub-lattice of the locally Lipschitz functions defined on A .*

Proof. By Propositions 2.3.12 and 2.3.3 in [11], we have

$$\partial_0(f \vee g)(x) \subseteq \text{co}\{\partial_0 f(x), \partial_0 g(x)\} \subseteq T(x)$$

for all $x \in A$, since $T(x)$ is convex. Similarly, one can show that $\partial_0(f \wedge g)(x) \subseteq T(x)$ for all $x \in A$. Also for any $0 \leq \lambda \leq 1$ we have, $\partial_0(\lambda f + (1 - \lambda)g)(x) \subseteq \lambda \partial_0 f(x) + (1 - \lambda) \partial_0 g(x) \subseteq \lambda T(x) + (1 - \lambda)T(x) = T(x)$ for all $x \in A$. \square

Lemma 5. *Let A be a non-empty open subset of a Banach space X and let $T : A \rightarrow 2^{X^*}$ be a weak* cusco on A , then (\mathcal{X}_T, ρ) is a complete metric space.*

Proof. Let $(f_n : n \in \mathbb{N})$ be a Cauchy sequence in (\mathcal{X}_T, ρ) and let f_∞ be the pointwise limit of the sequence $(f_n : n \in \mathbb{N})$. Note: f_∞ is well-defined since for each $x \in A$, $(f_n(x) : n \in \mathbb{N})$ is a Cauchy sequence in \mathbb{R} . Now since T is locally bounded f_∞ is locally Lipschitz on A . Indeed, if U is a convex open neighbourhood of some point $x_0 \in A$ and $T(U) \subseteq nB_{X^*}$, then for each $f \in \mathcal{X}_T$, $|f(x) - f(y)| \leq n\|x - y\|$ for all $x, y \in U$. So in particular we have

$$|f_\infty(x) - f_\infty(y)| = \lim_{k \rightarrow \infty} |f_k(x) - f_k(y)| \leq n\|x - y\|,$$

for all $x, y \in U$. We now need to show that $\partial_0 f_\infty(x) \subseteq T(x)$ for all $x \in A$. To do this it will suffice to show that for each $x_0 \in A$, $y \in S_X$ and $\varepsilon > 0$, $f_\infty^0(x_0; y) = \max\{x^*(y) : x^* \in \partial f_\infty(x_0)\} \leq \max\{x^*(y) : x^* \in T(x_0)\} + \varepsilon$. So let us fix $x_0 \in A$, $y \in S_X$ and $\varepsilon > 0$. Now since the mapping $x \rightarrow \max\{x^*(y) : x^* \in T(x)\}$ is upper semi-continuous on A there exists a $\delta > 0$ such that $\max\{x^*(y) : x^* \in T(z)\} < \max\{x^*(y) : x^* \in T(x_0)\} + \varepsilon$ for all $z \in B_{2\delta}(x_0)$. From this and the Lebourg mean-value theorem it follows that

$$\frac{f(z + \lambda y) - f(z)}{\lambda} \leq \max\{x^*(y) : x^* \in T(x_0)\} + \varepsilon,$$

for all $f \in \mathcal{X}_T$, $0 < \lambda < \delta$ and $z \in B_\delta(x_0)$. Therefore,

$$\frac{f_\infty(z + \lambda y) - f_\infty(z)}{\lambda} = \lim_{n \rightarrow \infty} \frac{f_n(z + \lambda y) - f_n(z)}{\lambda} \leq \max\{x^*(y) : x^* \in T(x_0)\} + \varepsilon,$$

for all $0 < \lambda < \delta$ and $z \in B_\delta(x_0)$. Hence, $f_\infty^0(x_0, y) \leq \max\{x^*(y) : x^* \in T(x_0)\} + \varepsilon$. This completes the proof. \square

The following result follows from Lemmas 1 and 3.

Proposition 1. *Let \mathcal{F} be a family of real-valued locally Lipschitz functions defined on a non-empty open subset A of a Banach space X . Then the functions in \mathcal{F} are T -Lipschitz for some weak* cusco $T : A \rightarrow 2^{X^*}$ if and only if the family of functions \mathcal{F} is locally equi-Lipschitz on A .*

3. RESULTS ON SEPARABLE BANACH SPACES

Lemma 6. *Let X be a separable Banach space and let f be a locally Lipschitz function defined on a non-empty open subset A of X , then there exists a countable set $C \subseteq \text{Gr}(\partial_- f)$ such that $\text{Gr}(\partial_a f) = \overline{C}$, where the closure is taken with respect to the product topology on $X \times X^*$ and X^* is endowed with weak* topology.*

Proof. For each $m \in \mathbb{N}$ define $A_m := \{x \in A : \partial_- f(x) \subseteq mB_{X^*}\}$. Since B_{X^*} is weak* compact and metrizable, $A_m \times mB_{X^*}$ is hereditarily separable and thus $\text{Gr}(\partial_- f) \cap (A_m \times mB_{X^*})$ is separable. Hence there exists a countable set,

$$C_m \subseteq \text{Gr}(\partial_- f) \cap (A_m \times mB_{X^*}) \quad \text{with} \quad \text{Gr}(\partial_- f) \cap (A_m \times mB_{X^*}) \subseteq \overline{C_m}.$$

Let $C := \bigcup_{m \in \mathbb{N}} C_m$. Then C is countable and

$$\begin{aligned} \text{Gr}(\partial_- f) &= \text{Gr}(\partial_- f) \cap \bigcup_{m \in \mathbb{N}} (A_m \times mB_{X^*}) \\ &= \bigcup_{m \in \mathbb{N}} \text{Gr}(\partial_- f) \cap (A_m \times mB_{X^*}) \subseteq \bigcup_{m \in \mathbb{N}} \overline{C_m} \subseteq \overline{C}. \end{aligned}$$

Since $\text{Gr}(\partial_a f) = \overline{\text{Gr}(\partial_- f)}$ we have $\text{Gr}(\partial_a f) \subseteq \overline{C} \subseteq \overline{\text{Gr}(\partial_a f)} = \text{Gr}(\partial_a f)$. \square

Lemma 7 ([21]). *If X is a smooth Banach space and Y is a finite dimensional subspace of X , then the distance function $x \mapsto d_Y(x) := \min_{y \in Y} \|x - y\|$ is smooth on $X \setminus Y$ and d_Y^2 is smooth on X .*

Lemma 8. *Let Y be a finite dimensional subspace of a smooth Banach space X and let $h : X \rightarrow (-\infty, +\infty]$ be a proper lower semi-continuous function. If $\varepsilon, \delta > 0$ and $z_0 \in X$ are given and h satisfies:*

- (i) *h is bounded below on $B_\delta[z_0]$;*
- (ii) *$h(z) - h(z_0) > -\delta \cdot \varepsilon$ for all $z \in z_0 + \delta B_Y$.*

Then there exists a point $z \in B_\delta(z_0)$ and $z^ \in \partial_- h(z)$ with $\|z^*|_Y\| < 2\varepsilon$.*

Proof. First let us choose K sufficiently large so that

$$\inf\{h(z) + Kd_{z_0+Y}^2(z) : z \in B_\delta[z_0]\} > h(z_0) - \varepsilon \cdot \delta.$$

By Lemma 7, $d_{z_0+Y}^2$ is smooth on X and constant along lines parallel to Y . Therefore we have that $\nabla(d_{z_0+Y}^2)(z)|_Y = 0$ for every $z \in X$. Now by the Borwein-Preiss smooth variational principle [9] we obtain $z \in B_\delta(z_0)$ and $x^* \in X^*$ so that $\|x^*\| < 2\varepsilon$ and

$$0 \in \partial_-(h + Kd_{z_0+Y}^2 + I_{B_\delta[z_0]})(z) + x^* = \partial_- h(z) + K\nabla(d_{z_0+Y}^2)(z) + x^*.$$

Therefore if we set $z^* := -(x^* + K\nabla(d_{z_0+Y}^2)(z))$, then $z^* \in \partial_- h(z)$ and $\|z^*|_Y\| = \|x^*|_Y\| < 2\varepsilon$. \square

Theorem 1 (The approximate subdifferential). *Let A be a non-empty open subset of a separable Banach space X (finite or infinite dimensional) and let $T : A \rightarrow 2^{X^*}$ be a weak* cusco on A , then for each $f \in \mathcal{X}_T$, $\{g \in \mathcal{X}_T : \partial_a f(x) \subseteq \partial_a g(x) \text{ for all } x \in A\}$ is residual in (\mathcal{X}_T, ρ) .*

Proof. Since X is separable, we may choose an increasing sequence of finite dimensional subspaces of X such that $\overline{\bigcup_{n \in \mathbb{N}} X_n} = X$. For each $(x, x^*) \in \text{Gr}(\partial_- f)$ and $n \in \mathbb{N}$ we consider the set,

$$G_{(x, x^*, n)} := \{g \in \mathcal{X}_T : \text{there exists a } z \in B_{1/n}(x) \\ \text{and } z^* \in \partial_- g(z) \text{ so that } \|(x^* - z^*)|_{X_n}\| \leq 4/n\}.$$

The proof is divided into three parts:

(a) For each $(x, x^*) \in \text{Gr}(\partial_- f)$ and $n \in \mathbb{N}$, $\text{int} G_{(x, x^*, n)}$ is dense in (\mathcal{X}_T, ρ) .

Suppose $(g_0, \varepsilon) \in \mathcal{X}_T \times (0, 1)$ is given. We need to verify that $B_\varepsilon(g_0) \cap \text{int} G_{(x, x^*, n)} \neq \emptyset$. Define $h_1 \in \mathcal{X}_T$ by $h_1(z) := f(z) + (g_0(x) - \varepsilon/3 - f(x))$ and the function $h_2 \in \mathcal{X}_T$ by $h_2(z) := \min\{g_0(z), h_1(z)\}$. Clearly, $h_2(z) \leq g_0(z)$ for all $z \in A$. Next we define $h_3 \in \mathcal{X}_T$ by $h_3(z) := \max\{h_2(z), g_0(z) - 2\varepsilon/3\}$ and obtain that $g_0(z) - \frac{2\varepsilon}{3} \leq h_3(z) \leq g_0(z)$ for all $z \in A$, and so $\rho(h_3, g_0) = \min\{1, d(h_3, g_0)\} < \varepsilon$. We claim that $h_3 \in \text{int} G_{(x, x^*, n)}$. To see this, first note that

$$g_0(x) - \frac{2\varepsilon}{3} < h_3(x) = h_2(x) = h_1(x) = g_0(x) - \frac{\varepsilon}{3} < g_0(x).$$

Hence there exists an open neighbourhood U of x so that $h_1 = h_2 = h_3$ on U . Since $x^* \in \partial_- f(x)$, f is Lipschitz around x and S_{X_n} is compact, there exists $0 < \delta < 1/n$ so that (i) $B_\delta[x] \subseteq U$; (ii) f is Lipschitz on $B_\delta[x]$; (iii)

$$(f - x^*)(x + \lambda v) - (f - x^*)(x) > -\frac{\lambda}{n} \geq \frac{-\delta}{n} \quad \text{for all } 0 < \lambda \leq \delta \text{ and } v \in S_{X_n}.$$

We now show that $B_r(h_3) \subseteq G_{(x, x^*, n)}$ for any $0 < r < \delta/(2n)$. To accomplish this, let g be any member of $B_r(h_3)$. Then,

$$(g - x^*)(z) - (g - x^*)(x) > -\frac{2\delta}{n} \quad \text{for all } z \in x + \delta B_{X_n}.$$

By Lemma 8, there exists $z \in B_\delta(x)$ and $y^* \in \partial_-(g - x^*)(z)$ with $\|y^*|_{X_n}\| < 4/n$. Thus if we set $z^* := y^* + x^*$, then $z^* \in \partial_- g(z)$ and $\|(z^* - x^*)|_{X_n}\| < 4/n$. This shows that $g \in G_{(x, x^*, n)}$.

(b) Fix $(x, x^*) \in \text{Gr}(\partial_- f)$, then for each g in $G_{(x, x^*)} := \bigcap_{n \in \mathbb{N}} G_{(x, x^*, n)}$, we have $x^* \in \partial_a g(x)$. If $g \in G_{(x, x^*)}$, then for each $n \in \mathbb{N}$ there exists an $x_n \in B_{1/n}(x)$ and $x_n^* \in \partial_- g(x_n)$ so that $\|x_n^*|_{X_n} - x^*|_{X_n}\| \leq 4/n$. Since the subspaces $(X_n : n \in \mathbb{N})$ are monotonely increasing, we see that $(x_n^* : n \in \mathbb{N})$ converges to x^* pointwise on $\bigcup_{n \in \mathbb{N}} X_n$. Moreover, since g is locally Lipschitz around x , the sequence $(x_n^* : n \in \mathbb{N})$ is norm bounded, and so we have that $(x_n^* : n \in \mathbb{N})$ converges to x^* pointwise on $\overline{\bigcup_{n \in \mathbb{N}} X_n} = X$. Hence $((x_n, x_n^*) : n \in \mathbb{N})$ converges to (x, x^*) in $A \times X^*$, with A endowed with the norm topology and X^* with the weak* topology. However, as $\text{Gr}(\partial_a g)$ is closed in $A \times X^*$ we obtain that $x^* \in \partial_a g(x)$.

(c) By Lemma 6, we may choose a countable set $C \subseteq \text{Gr}(\partial_- f)$ so that $\text{Gr}(\partial_a f) = \overline{C}$. Let

$$G := \bigcap \{G_{(x, x^*)} : (x, x^*) \in C\}.$$

By (a), G is a residual set in (\mathcal{X}_T, ρ) and if $g \in G$, then for every $(x, x^*) \in C$ we have $x^* \in \partial_a g(x)$. That is, $C \subseteq \text{Gr}(\partial_a g)$. Now since $\text{Gr}(\partial_a g)$ is closed in the product topology on $A \times X^*$ we have

$$\text{Gr}(\partial_a f) = \overline{C} \subseteq \text{Gr}(\partial_a g).$$

This shows that if $g \in G$ then $\partial_a f(x) \subseteq \partial_a g(x)$ for all $x \in A$. □

Corollary 1. *Let $\{f_n : n \in \mathbb{N}\}$ be a sequence of locally equi-Lipschitz functions defined on a non-empty open subset A of a separable Banach space X . If we define $T : A \rightarrow 2^{X^*}$ by $T(x) := \bigcup_{n \in \mathbb{N}} \partial_a f_n(x)$, then $\{g \in \mathcal{X}_{CSC(T)} : USC(T)(x) \subseteq \partial_a g(x) \subseteq \partial_0 g(x) \subseteq CSC(T)(x) \text{ for every } x \in A\}$ is residual in $(\mathcal{X}_{CSC(T)}, \rho)$.*

Proof. By Proposition 1, $CSC(T)$ exists. For each $n \in \mathbb{N}$, we may apply Theorem 1 to deduce that the set $G_n := \{g \in \mathcal{X}_{CSC(T)} : \partial_a f_n(x) \subseteq \partial_a g(x) \text{ for all } x \in A\}$ is residual in $(\mathcal{X}_{CSC(T)}, \rho)$. Thus the set $G := \bigcap_{n \in \mathbb{N}} G_n$ is residual in $(\mathcal{X}_{CSC(T)}, \rho)$ and if $g \in G$, then we have $\bigcup_{n \in \mathbb{N}} \partial_a f_n(x) \subseteq \partial_a g(x)$ for every $x \in A$. Hence $USC(T)(x) \subseteq \partial_a g(x) \subseteq \partial_0 g(x) \subseteq CSC(T)(x)$ for all $x \in A$. \square

Theorem 2 (The Clarke subdifferential). *Let A be a non-empty open subset of a separable Banach space X and let $T : A \rightarrow 2^{X^*}$ be a weak*usco on A , then for each $f \in \mathcal{X}_T$, $\{g \in \mathcal{X}_T : \partial_0 f(x) \subseteq \partial_0 g(x) \text{ for all } x \in A\}$ is residual in (\mathcal{X}_T, ρ) .*

Proof. This follows from Theorem 1 and the fact that $\partial_0 f(x) = \overline{\text{co}}^{w^*} \partial_a f(x)$ for all $x \in A$, [17]. \square

Corollary 2. *Let $\{f_n : n \in \mathbb{N}\}$ be a locally equi-Lipschitz family of real-valued functions defined on a non-empty open subset A of a separable Banach space X . If we define $T : A \rightarrow 2^{X^*}$ by $T(x) := \bigcup_{n \in \mathbb{N}} \partial_0 f_n(x)$, then $\{g \in \mathcal{X}_{CSC(T)} : \partial_0 g(x) = CSC(T)(x) \text{ for all } x \in A\}$ is residual in $(\mathcal{X}_{CSC(T)}, \rho)$.*

Corollary 3. *Let f_1, f_2, \dots, f_n be real-valued locally Lipschitz functions defined on a non-empty open subset A of a separable Banach space X . If $T : A \rightarrow 2^{X^*}$ is defined by*

$$T(x) := \text{co}\{\partial_0 f_1(x), \partial_0 f_2(x), \dots, \partial_0 f_n(x)\},$$

then $\{g \in \mathcal{X}_T : \partial_0 g(x) = T(x) \text{ for all } x \in A\}$ is residual in (\mathcal{X}_T, ρ) . In particular, the Clarke subdifferential is closed under the operation of taking finite convex hulls.

Proof. By Corollary 2 it suffices to show that T is a weak*usco on A . To see that this is indeed the case, consider the set-valued mapping $\Omega : A \rightarrow 2^{X^*}$ defined by $\Omega(x) \equiv \bigcup\{\partial_0 f_j(x) : 1 \leq j \leq n\}$. Clearly Ω is a weak*usco on A . Hence, by Lemma 7.12 in [20] the mapping $T : A \rightarrow 2^{X^*}$ defined by $T(x) := \overline{\text{co}}^{w^*} \Omega(x) = \text{co } \Omega(x)$ is a weak*usco on A . This completes the proof. \square

Corollary 3 improves the main result of [8], where the minimality of each $\partial_0 f_j$ was required.

Corollary 4. *Let f be a real-valued locally Lipschitz function defined on a non-empty open connected subset A of a separable Banach space X . Then the following conditions are equivalent:*

- (i) f is “strongly integrable” that is, for each real-valued locally Lipschitz function g defined on A with $\partial_0 g(x) \subseteq \partial_0 f(x)$ for all $x \in A$, $f - g \equiv \text{constant}$.
- (ii) f is “weakly integrable” that is, for each real-valued locally Lipschitz function g defined on A with $\partial_0 g(x) = \partial_0 f(x)$ for all $x \in A$, $f - g \equiv \text{constant}$.

Proof. The fact that (i) implies (ii) is obvious. So it suffices for us to justify that (ii) implies (i). Fix $x_0 \in A$ and let g be any member of $\mathcal{X}_{\partial_0 f}$. By Corollary 3 we may select, for each $0 < \varepsilon < 1$, a function $g_\varepsilon \in \mathcal{X}_{\partial_0 f}$ so that $\rho(g, g_\varepsilon) < \varepsilon$ and

$\partial_0 g_\varepsilon(x) = \partial_0 f(x)$ for all $x \in A$. Then for any $x \in A$,

$$\begin{aligned} |(f - g)(x) - (f - g)(x_0)| &\leq |(f - g_\varepsilon)(x) - (f - g_\varepsilon)(x_0)| + |(g_\varepsilon - g)(x) - (g_\varepsilon - g)(x_0)| \\ &= |(g_\varepsilon - g)(x) - (g_\varepsilon - g)(x_0)| \quad (\text{since } f - g_\varepsilon \equiv \text{constant on } A) \\ &\leq |(g_\varepsilon - g)(x)| + |(g_\varepsilon - g)(x_0)| \leq 2\varepsilon. \end{aligned}$$

However, as our choice of ε was arbitrary, we must have that $(f - g)(x) = (f - g)(x_0)$. This shows that $f - g \equiv \text{constant on } A$. \square

Let I be an open interval in \mathbb{R} and let $f : I \rightarrow \mathbb{R}$. We say that f is *robustly lower (upper) semi-continuous* if,

$$f(x) = \liminf_{\substack{y \rightarrow x \\ y \notin N}} f(y) \quad \left(f(x) = \limsup_{\substack{y \rightarrow x \\ y \notin N}} f(y) \right)$$

for each Lebesgue null set N of I .

Corollary 5. *Let I be an open interval in \mathbb{R} and let α and β be functions on I such that $\alpha \leq \beta$. Then the following are equivalent:*

- (i) α is robustly lower semi-continuous and β is robustly upper semi-continuous;
- (ii) there exists a locally Lipschitz function f on I such that $\partial_0 f(x) = [\alpha(x), \beta(x)]$ for all $x \in I$;
- (iii) if $T : I \rightarrow 2^{\mathbb{R}}$ is defined by $T(x) := [\alpha(x), \beta(x)]$, then \mathcal{X}_T is non-empty and $\{g \in \mathcal{X}_T : \partial_0 g(x) = [\alpha(x), \beta(x)] \text{ for all } x \in I\}$ is residual in \mathcal{X}_T .

Proof. (i) \Rightarrow (iii) Since α is lower semi-continuous and β is upper semi-continuous on I , both functions are Lebesgue integrable. Choose any $p \in I$ and define $f_1, f_2 : I \rightarrow \mathbb{R}$ by

$$f_1(x) := \int_p^x \alpha(t) dt \quad \text{and} \quad f_2(x) := \int_p^x \beta(t) dt.$$

Let $N \subseteq I$ be any Lebesgue null set so that $f'_1(x) = \alpha(x)$ and $f'_2(x) = \beta(x)$ on $I \setminus N$. Then we have

$$\begin{aligned} \partial_0 f_1(x) &= [\liminf_{\substack{t \notin N \\ t \rightarrow x}} f'_1(t), \limsup_{\substack{t \notin N \\ t \rightarrow x}} f'_1(t)] = [\alpha(x), \limsup_{\substack{t \notin N \\ t \rightarrow x}} f'_1(t)] \quad \text{and} \\ \partial_0 f_2(x) &= [\liminf_{\substack{t \notin N \\ t \rightarrow x}} f'_2(t), \limsup_{\substack{t \notin N \\ t \rightarrow x}} f'_2(t)] = [\liminf_{\substack{t \notin N \\ t \rightarrow x}} f'_2(t), \beta(x)]. \end{aligned}$$

Moreover, since $\alpha(x) \leq \beta(x)$ for all $x \in I$ we have

$$\limsup_{\substack{t \notin N \\ t \rightarrow x}} f'_1(t) \leq \limsup_{\substack{t \notin N \\ t \rightarrow x}} \beta(t) = \beta(x) \quad \text{and} \quad \liminf_{\substack{t \notin N \\ t \rightarrow x}} f'_2(t) \geq \liminf_{\substack{t \notin N \\ t \rightarrow x}} \alpha(t) = \alpha(x).$$

Therefore, $T(x) = \text{co}\{\partial_0 f_1(x), \partial_0 f_2(x)\}$ for each $x \in I$ and so the result follows by Corollary 3.

(iii) \Rightarrow (ii) is clear and (ii) \Rightarrow (i) follows from the basic properties of the Clarke subdifferential mapping. \square

This has recovered and improved the main result in [2].

Example 1. Theorem 1 fails if T is only assumed to be a weak* usco. To see this, we consider $T : \mathbb{R} \rightarrow 2^{\mathbb{R}}$ defined by $T(x) := \{0, 1\}$. If Theorem 1 holds for this T , then there would exist a residual set G in \mathcal{X}_T where $\partial_a g = T$ for each $g \in G$. However, T is not an approximate subdifferential map of any Lipschitz function.

Because if there exists an f with $\partial_a f = T$, then $\partial_0 f(x) = \text{co } \partial_a f(x) = [0, 1]$ for all $x \in \mathbb{R}$ and so by Theorem 2.2 in [2] we would have that $\partial_a f(x) = \partial_0 f(x) = [0, 1]$ for all $x \in \mathbb{R}$; a contradiction.

With some extra work one can show that the sum and lattice rules of sub-differential calculus hold with equality for almost all Lipschitz functions. Indeed, if on $\mathcal{X}_{B_{X^*}} \times \mathcal{X}_{B_{X^*}}$ we define the complete metric ρ_1 by $\rho_1((f_1, f_2), (g_1, g_2)) := \rho(f_1, g_1) + \rho(f_2, g_2)$, then we may obtain the following from a more elaborate version of Theorem 1.

Theorem 3 ([26]). *Let A be a non-empty open subset of a separable Banach space X , then there exists a residual set G in $(\mathcal{X}_{B_{X^*}} \times \mathcal{X}_{B_{X^*}}, \rho_1)$ so that for each $(f_1, f_2) \in G$ and $y_1, y_2 \geq 0$,*

$$\partial_0(y_1 f_1 + y_2 f_2)(x) = y_1 \partial_0 f_1(x) + y_2 \partial_0 f_2(x), \quad \partial_0 f_1(x) = \partial_0 f_2(x) = B_{X^*} \text{ and}$$

$$\partial_0[\min\{f_1, f_2\}](x) = \partial_0[\max\{f_1, f_2\}](x) = \text{co}\{\partial_0 f_1(x), \partial_0 f_2(x)\} \text{ for all } x \in A.$$

Example 2. In this example we examine what can be said about the size of the generalized Jacobian of a Lipschitz mapping acting between Banach spaces. For a vector-valued locally Lipschitz function $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ given by $F(x) := [f_1(x), \dots, f_m(x)]$. The *generalized Jacobian* of F at x , denoted by $\partial_0 F(x)$, is defined by

$$\partial_0 F(x) := \text{co}\{\lim \nabla F(x_i) : x_i \rightarrow x, x_i \notin \Omega_F\},$$

where Ω_F denotes the set of points at which F fails to be differentiable. In $\mathcal{X}_{B_{\mathbb{R}^n}} \times \mathcal{X}_{B_{\mathbb{R}^n}}$ there exists a residual set G such that for every $F \in G$, $\partial_0 F$ is not a minimal cusco. Indeed, by Theorem 3 we obtain a residual set $G \subset \mathcal{X}_{B_{\mathbb{R}^n}} \times \mathcal{X}_{B_{\mathbb{R}^n}}$ such that for $F := (f_1, f_2) \in G$ we have $\partial_0(f_1 + f_2) = 2B_{\mathbb{R}^n}$. By Theorem 2.6.6 [11] we have $\partial_0(f_1 + f_2) = (1, 1)\partial_0 F$. Since $2B_{\mathbb{R}^n}$ is not a minimal cusco, $\partial_0 F$ is not minimal for $F \in G$. However, if each f_i is strictly differentiable almost everywhere on A for $i = 1, \dots, m$, then $\partial_0 F$ is a minimal cusco on A .

The scalarization formula for the *coderivative* of F [25, page 366] also shows that the Clarke coderivatives have large images for every $F \in G$ since the Clarke coderivative $D_c^* F(x)(y_1, y_2) \supseteq (y_1 + y_2)B_{\mathbb{R}^n}$ holds for all $x \in A$ and $(y_1, y_2) \in \mathbb{R}_2^+ \setminus \{0\}$. Similar results hold for F having arbitrary m components.

4. RESULTS ON GENERAL BANACH SPACES

In general Banach spaces we obtain weaker, but still highly useful results.

Lemma 9. *Let A be a non-empty open subset of a Banach space X and let $T : A \rightarrow 2^{X^*}$ be a weak* cusco on A . If $f, g \in \mathcal{X}_T$ and $E \subseteq \mathbb{R}$ is Lebesgue measurable, then the function $h : A \rightarrow \mathbb{R}$ defined by $h(x) := \lambda_E((f - g)(x)) + g(x)$ belongs to \mathcal{X}_T and $\rho(h, g) \leq \lambda(E)$, where $\lambda_E : \mathbb{R} \rightarrow \mathbb{R}$ is defined by, $\lambda_E(x) := \int_0^x \chi_E(t) dt$.*

Proof. Suppose $f, g \in \mathcal{X}_T$, we need to show $h \in \mathcal{X}_T$. To this end, let us fix $x \in A$ and choose $\delta > 0$ so that $B_{2\delta}(x) \subseteq A$. For each $v \in S_X$ we define the function $K_v : B_\delta(x) \times (0, \delta) \rightarrow [0, 1]$ by

$$K_v(z, \lambda) := \begin{cases} 0 & \text{if } (f - g)(z + \lambda v) = (f - g)(z), \\ \frac{\int_{(f-g)(z)}^{(f-g)(z+\lambda v)} \chi_E(t) dt}{(f-g)(z+\lambda v) - (f-g)(z)} & \text{otherwise.} \end{cases}$$

Then we have

$$\begin{aligned} \frac{h(z + \lambda v) - h(z)}{\lambda} &= (1 - K_v(z, \lambda)) \frac{g(z + \lambda v) - g(z)}{\lambda} + K_v(z, \lambda) \frac{f(z + \lambda v) - f(z)}{\lambda} \\ &\leq \max \left\{ \frac{g(z + \lambda v) - g(z)}{\lambda}, \frac{f(z + \lambda v) - f(z)}{\lambda} \right\}, \end{aligned}$$

for all $(z, \lambda) \in B_\delta(x) \times (0, \delta)$. Therefore $h^0(x, v) \leq \max\{g^0(x, v), f^0(x, v)\}$ for all $v \in S_X$ and so $\partial_0 h(x) \subseteq \text{co}\{\partial_0 g(x), \partial_0 f(x)\} \subseteq T(x)$. Since the point x was arbitrary we have that $h \in \mathcal{X}_T$. The proof that $\rho(g, h) \leq \lambda(E)$ is obvious. \square

We will say that a Banach space is *smoothable* if it has an equivalent smooth renorm (as is the case in all separable, reflexive or WCG spaces).

Theorem 4 (The approximate subdifferential in smooth Banach space). *Let A be a non-empty open subset of a smoothable Banach space X and let $T : A \rightarrow 2^{X^*}$ be a weak*usco on A . If $f \in \mathcal{X}_T$ and $x \rightarrow \partial_a f(x)$ is a minimal weak*usco, then $\{g \in \mathcal{X}_T : \partial_a f(x) \subseteq \partial_a g(x) \text{ for all } x \in A\}$ is residual in (\mathcal{X}_T, ρ) .*

Proof. For each $m \in \mathbb{N}$, let $A_m := \text{int}\{t \in A : T(t) \subseteq mB_{X^*}\}$. Then by Lemma 3, $A = \bigcup_{m \in \mathbb{N}} A_m$ and each $g \in \mathcal{X}_T$ is m -Lipschitz on each convex subset of A_m . Let $J := \{J_n : n \in \mathbb{N}\}$ be an enumeration of all the open intervals in \mathbb{R} with rational end-points. For each $(m, n, p, \varepsilon) \in \mathbb{N}^3 \times (0, \infty)$ we consider the set

$$\begin{aligned} O_{(m, n, p, \varepsilon)} &:= \{g \in \mathcal{X}_T : \text{for each connected open set } U \text{ with } U + \frac{1}{p}B_X \subseteq A_m \text{ and } J_n \subseteq (f - g)(U) \text{ there} \\ &\text{exists a } z_0 \in U \text{ and } 0 < r_0 < 1/p \text{ so that } (g - f)(z) - (g - f)(z_0) > -\varepsilon r_0 \text{ for all } z \in B_{r_0}(z_0)\}. \end{aligned}$$

The proof is now divided into two parts:

(a) For each $(m, n, p, \varepsilon) \in \mathbb{N}^3 \times (0, \infty)$, $\text{int}O_{(m, n, p, \varepsilon)}$ is dense in (\mathcal{X}_T, ρ) .

Suppose $(g_0, \delta) \in \mathcal{X}_T \times (0, 1)$, we need to verify that $B_\delta(g_0) \cap \text{int}O_{(m, n, p, \varepsilon)} \neq \emptyset$. To this end, suppose $J_n := (r_n, s_n)$ and $0 < \delta' := \min\{(s_n - r_n)/5, \delta\}$. Now let us choose a dense open subset E of \mathbb{R} such that $\mu(E) < \delta'$ and define $h : A \rightarrow \mathbb{R}$ by

$$h(x) := \lambda_E((f - g_0)(x)) + g_0(x) \quad \text{where} \quad \lambda_E(x) := \int_0^x \chi_E(t) dt.$$

By Lemma 9 we have $h \in \mathcal{X}_T$ and $\rho(g_0, h) < \delta' \leq \delta$. We claim that $h \in \text{int}O_{(m, n, p, \varepsilon)} \cap B_\delta(g_0)$. To this end, choose $0 < r < 2m/p$ and $t \in \mathbb{R}$ so that $[t - r, t + r] \subseteq (r_n + 2\delta', s_n - 2\delta') \cap E$ and set $d := \min\{(\varepsilon r)/(4m), \delta'\}$. We will show that $B_d(h) \subseteq O_{(m, n, p, \varepsilon)}$. Let $g \in B_d(h)$ and let U be any connected open subset of A_m with $U + 1/pB_X \subseteq A_m$ and $J_n \subseteq (f - g)(U)$. Then,

$$[t - r, t + r] \subseteq (r_n + 2\delta', s_n - 2\delta') \subseteq (f - g_0)(U),$$

since $(f - g_0)(U)$ is connected (hence convex) and

$$\|(f - g_0) - (f - g)\|_\infty = \|g_0 - g\|_\infty \leq \|g_0 - h\|_\infty + \|g - h\|_\infty < \delta' + d \leq 2\delta'.$$

Choose $z_0 \in U$ so that $(f - g_0)(z_0) = t$ then for any $z \in B_{r_0}(z_0)$ with $r_0 = r/2m < 1/p$, $(f - g_0)(z) \in [t - r, t + r] \subseteq E$ and so $h(z) - h(z_0) = f(z) - f(z_0)$. Therefore by our choice of d ,

$$(g - f)(z) - (g - f)(z_0) = (g - h)(z) - (g - h)(z_0) > -\frac{\varepsilon r_0}{2} - \frac{\varepsilon r_0}{2} = -\varepsilon r_0,$$

for all $z \in B_{r_0}(z_0)$. This shows that $g \in O_{(m, n, p, \varepsilon)}$.

(b) The set $G := \bigcap \{O_{(n_1, n_2, n_3, 1/n_4)} : (n_1, n_2, n_3, n_4) \in \mathbb{N}^4\}$ is residual in (\mathcal{X}_T, ρ) and for each $g \in G$ we have $\partial_a g(x) \cap \partial_a f(x) \neq \emptyset$ for every $x \in A$.

Indeed, if this is not the case then there exists a $g \in G$ and $x_0 \in A$ such that $\partial_a g(x_0) \cap \partial_a f(x_0) = \emptyset$. Moreover, since $\{x \in A : \partial_a g(x) \cap \partial_a f(x) = \emptyset\}$ is open in A we may assume that f is strictly differentiable at $x_0 \in A$, [22]. We may now select a finite set $F \subseteq S_X$ and $\alpha \in \mathbb{R}$ so that if $W := \{x^* : |x^*(y)| < \alpha, y \in F\}$ then $(\partial_a g(x_0) + W) \cap (\partial_a f(x_0) + W) = \emptyset$ and by the weak* upper semi-continuity of $x \rightarrow \partial_a g(x)$, there exists $r > 0$ so that $\partial_a g(B_r(x_0)) \subseteq \partial_a g(x_0) + W$. Next, we choose $n_4 \in \mathbb{N}$ so that $4/n_4 < \alpha$ and set $Y := \text{sp}F$. Then by the strict differentiability of f at x_0 , the Lipschitzness of f and the compactness of S_Y , there exists a $\delta > 0$ so that

$$\frac{f(z + \lambda y) - f(z)}{\lambda} - \nabla f(x_0)(y) > -\frac{1}{n_4}$$

for all $\|z - x_0\| \leq \delta$, $0 < \lambda \leq \delta$ and $y \in S_Y$.

We may now select $n_1, n_3 \in \mathbb{N}$ so that $1/n_3 < \delta$ and $B_{2/n_3}(x_0) \subseteq A_{n_1} \cap B_r(x_0)$. Then we set $U := B_{1/n_3}(x_0)$. Now $(f - g)(U)$ is convex, so either $(f - g)(U) = \{a\}$ for some $a \in \mathbb{R}$ or $J_n \subseteq (f - g)(U)$ for some n . In the first case we get that $f(x) = g(x) + a$ on U which is impossible since $\partial_a f(x_0) \neq \partial_a g(x_0)$. Therefore, there is some $n_2 \in \mathbb{N}$ so that $J_{n_2} \subseteq (f - g)(U)$. However, as $g \in O_{(n_1, n_2, n_3, 1/n_4)}$ there exists a point $z_0 \in U$ and $0 < r_0 \leq 1/n_3$ so that

$$(g - \nabla f(x_0))(z) - (g - \nabla f(x_0))(z_0) > -\frac{2r_0}{n_4} \quad \text{for all } z \in z_0 + r_0 B_Y,$$

since $(f - \nabla f(x_0))(z_0 + \lambda y) - (f - \nabla f(x_0))(z_0) > -\lambda \cdot \frac{1}{n_4}$ for $0 < \lambda \leq \frac{1}{n_3}$ and $y \in S_Y$. By Lemma 8 there exists $z_1 \in B_{r_0}(z_0)$ and $z^* \in X^*$ so that $z^* \in \partial_-(g - \nabla f(x_0))(z_1)$ and $\|z^*|_Y\| < 4/n_4$, i.e., $\nabla f(x_0) + z^* \in \partial_- g(z_1)$. However, $\nabla f(x_0) + z^* \in \nabla f(x_0) + W$, which is impossible since $\partial_- g(z_1) \cap (\nabla f(x_0) + W) = \emptyset$. Therefore, it must be the case that for each $g \in G$, $\partial_a f(x) \cap \partial_a g(x) \neq \emptyset$ for all $x \in A$. The result now follows by Lemma 2. \square

While minimality of $x \rightarrow \partial_a f(x)$ is quite restrictive, it does hold when f is smooth or concave on A . Even this allows for some nice applications:

Corollary 6. *Let A be a non-empty open subset of a smoothable Banach space X and suppose $C \subseteq X^*$ is non-empty, weak* compact, convex and weak* separable, then $\{g \in \mathcal{X}_C : \partial_a g(x) = C \text{ for all } x \in A\}$ is residual in (\mathcal{X}_C, ρ) .*

Proof. Let $\{x_n^* : n \in \mathbb{N}\}$ be a dense subset of (C, weak^*) and for each $n \in \mathbb{N}$, let $G_n := \{g \in \mathcal{X}_C : x_n^* \in \partial_a g(x) \text{ for all } x \in A\}$. Then by Theorem 4, $G := \bigcap_{n \in \mathbb{N}} G_n$ is residual in (\mathcal{X}_C, ρ) and $\partial_a g \equiv C$ for each $g \in G$. \square

Theorem 5 (The Clarke subdifferential in arbitrary Banach space). *Let A be a non-empty open subset of a Banach space X and let $T : A \rightarrow 2^{X^*}$ be a weak*usco on A , then for each $f \in \mathcal{X}_T$, $\{g \in \mathcal{X}_T : \partial_0 g(x) \cap \partial_0 f(x) \neq \emptyset \text{ for all } x \in A\}$ is residual in (\mathcal{X}_T, ρ) . In particular, if $\partial_0 f$ is a minimal weak*usco, then $\{g \in \mathcal{X}_T : \partial_0 f(x) \subseteq \partial_0 g(x) \text{ for all } x \in A\}$ is residual in (\mathcal{X}_T, ρ) .*

Proof. For each $m \in \mathbb{N}$, let $A_m := \text{int}\{t \in A : T(t) \subseteq mB_{X^*}\}$. Then by Lemma 3, $A = \bigcup_{m \in \mathbb{N}} A_m$ and each $g \in \mathcal{X}_T$ is m -Lipschitz on each convex subset of A_m . Let

$J := \{J_n : n \in \mathbb{N}\}$ be an enumeration of all the open intervals in \mathbb{R} with rational end-points. For each $(m, n, \varepsilon) \in \mathbb{N}^2 \times (0, +\infty)$ consider the set

$$O_{(m,n,\varepsilon)} := \{g \in \mathcal{X}_T : \text{for each connected open set } U \\ \text{with } U + \varepsilon B_X \subseteq A_m \text{ and } J_n \subseteq (f - g)(U) \text{ there exist} \\ z_0 \in U \text{ and } 0 < \lambda_0 < \varepsilon \text{ so that } (g - f)(z_0 + \lambda_0 v) - \\ (g - f)(z_0) \geq -\lambda_0 \varepsilon \text{ for all } v \in S_X\}.$$

(a) $\text{int}O_{(m,n,\varepsilon)}$ is dense in (\mathcal{X}_T, ρ) for each $(m, n, \varepsilon) \in \mathbb{N}^2 \times (0, +\infty)$.

Suppose $(g_0, \delta) \in \mathcal{X}_T \times (0, 1)$. We need to verify that $B_\delta(g_0) \cap \text{int}O_{(m,n,\varepsilon)} \neq \emptyset$. To this end, suppose $J_n := (r_n, s_n)$ and $\delta' := \min\{(s_n - r_n)/5, \delta\}$. Now let us choose a dense open subset E of \mathbb{R} such that $\mu(E) < \delta'$ and define $h : A \rightarrow \mathbb{R}$ by

$$h(x) := \lambda_E((f - g_0)(x)) + g_0(x) \quad \text{where} \quad \lambda_E(x) := \int_0^x \chi_E(t) dt.$$

Then by Lemma 9, $h \in \mathcal{X}_T$ and $\rho(g_0, h) < \delta' \leq \delta$. We claim that $h \in \text{int}O_{(m,n,\varepsilon)} \cap B_\delta(g_0)$. To this end, choose $0 < r < 2m\varepsilon$ and $t \in \mathbb{R}$ so that $[t - r, t + r] \subseteq (r_n + 2\delta', s_n - 2\delta') \cap E$ and set $d := \min\{(\varepsilon r)/(4m), \delta'\}$. We show that $B_d(h) \subseteq O_{(m,n,\varepsilon)}$. Let $g \in B_d(h)$ and let U be any connected open subset of A_m with $U + \varepsilon B_X \subseteq A_m$ and $J_n \subseteq (f - g)(U)$. Since $(f - g_0)(U)$ is connected (hence convex) and

$$\|(f - g) - (f - g_0)\|_\infty = \|g_0 - g\|_\infty \leq \|g_0 - h\|_\infty + \|g - h\|_\infty \leq d + \delta' \leq 2\delta',$$

we have $[t - r, t + r] \subseteq (r_n + 2\delta', s_n - 2\delta') \subseteq (f - g_0)(U)$. Choose $z_0 \in U$ so that $(f - g_0)(z_0) = t$, then for every $0 < \lambda \leq r/(2m) < \varepsilon$ and $v \in S_X$ we have $(f - g_0)(z_0 + \lambda v) \in [t - r, t + r] \subseteq E$, thus if we set $\lambda_0 := r/(2m)$

$$\begin{aligned} \frac{g(z_0 + \lambda_0 v) - g(z_0)}{\lambda_0} &\geq \frac{h(z_0 + \lambda_0 v) - h(z_0)}{\lambda_0} - \frac{2d}{\lambda_0} = \frac{f(z_0 + \lambda_0 v) - f(z_0)}{\lambda_0} - \frac{2d}{\frac{r}{2m}} \\ &\geq \frac{f(z_0 + \lambda_0 v) - f(z_0)}{\lambda_0} - \varepsilon. \end{aligned}$$

This shows that $g \in O_{(m,n,\varepsilon)}$.

(b) The set $G := \bigcap \{O_{(n_1, n_2, 1/n_3)} : (n_1, n_2, n_3) \in \mathbb{N}^3\}$ is residual in (\mathcal{X}_T, ρ) and for each $g \in G$ we have $\partial_0 g(x) \cap \partial_0 f(x) \neq \emptyset$ for all $x \in A$.

Indeed, if this is not the case, then there exists a $g \in G$ and $x_0 \in A$ such that $\partial_0 g(x_0) \cap \partial_0 f(x_0) = \emptyset$. By the strong separation theorem, applied in the locally convex space (X^*, weak^*) there exists a $y \in S_X$, $\alpha \in \mathbb{R}$ and $\varepsilon > 0$ such that

$$\begin{aligned} -f^0(x_0, -y) &= \min\{x^*(y) : x^* \in \partial_0 f(x_0)\} > \alpha + \varepsilon \\ &> \alpha - \varepsilon > \max\{x^*(y) : x^* \in \partial_0 g(x_0)\} = g^0(x_0, y). \end{aligned}$$

Now $x_0 \in A_{n_1}$ for some $n_1 \in \mathbb{N}$ and

$$-f^0(x_0, -y) = \liminf_{\substack{z \rightarrow x_0 \\ t \downarrow 0}} \frac{f(z + ty) - f(z)}{t}, \quad g^0(x_0, y) = \limsup_{\substack{z \rightarrow x_0 \\ t \downarrow 0}} \frac{g(z + ty) - g(z)}{t}.$$

Therefore there exists an $n_3 \in \mathbb{N}$ such that $1/n_3 < \varepsilon$, $B_{2/n_3}(x_0) \subseteq A_{n_1}$ and

$$\begin{aligned} (*) \quad \inf \left\{ \frac{f(z + \lambda y) - f(z)}{\lambda} : 0 < \lambda \leq \frac{1}{n_3}, z \in B_{1/n_3}(x_0) \right\} &\geq \alpha + \frac{1}{n_3} \\ &> \alpha - \frac{1}{n_3} \geq \sup \left\{ \frac{g(z + \lambda y) - g(z)}{\lambda} : 0 < \lambda \leq \frac{1}{n_3}, z \in B_{1/n_3}(x_0) \right\} \end{aligned}$$

Let $U := B_{1/n_3}(x_0)$, then $U + 1/n_3 B_X \subseteq A_{n_1}$. Now $(f - g)(U)$ is convex, so either $(f - g)(U) = \{a\}$ for some $a \in \mathbb{R}$ or $J_n \subseteq (f - g)(U)$ for some $n \in \mathbb{N}$. In the first case, we get that $f(x) = g(x) + a$ on U which is impossible since $\partial_0 f(x_0) \neq \partial_0 g(x_0)$. Therefore, there exists some $n_2 \in \mathbb{N}$ so that $J_{n_2} \subseteq (f - g)(U)$. Since $g \in G$, we have that $g \in O_{(n_1, n_2, 1/n_3)}$ and so there exists $z_0 \in U$ and $0 < \lambda_0 < 1/n_3$ so that

$$\frac{g(z_0 + \lambda_0 v) - g(z_0)}{\lambda_0} \geq \frac{f(z_0 + \lambda_0 v) - f(z_0)}{\lambda_0} - \frac{1}{n_3} \quad \text{for every } v \in S_X,$$

which contradicts (*). Therefore, it must be the case that for each $g \in G$ we have $\partial_0 f(x) \cap \partial_0 g(x) \neq \emptyset$ for each $x \in A$. The case when $\partial_0 f$ is a minimal weak* cusco follows from Lemma 2. \square

Corollary 7. *Let A be a non-empty open subset of a Banach space X and let $\{f_n : n \in \mathbb{N}\}$ be a sequence of locally equi-Lipschitz real-valued functions. If $T : A \rightarrow 2^{X^*}$ is defined by $T(x) := \bigcup_{n \in \mathbb{N}} \partial_0 f_n(x)$ and each $\partial_0 f_n$ is a minimal weak* cusco then $\{g \in \mathcal{X}_{CSC(T)} : \partial_0 g(x) = CSC(T)(x) \text{ for each } x \in A\}$ is residual in $(\mathcal{X}_{CSC(T)}, \rho)$.*

Since each maximal cyclically monotone operator defined on a non-empty open convex subset of a Banach space is the Clarke subgradient of some convex locally Lipschitz function we have

Corollary 8. *Let A be a non-empty open convex subset of a Banach space X and let $\{T_1, T_2, \dots, T_n\}$ be a finite family of maximal cyclically monotone operators from A into non-empty subsets of X^* . Then there exists a real-valued locally Lipschitz function f defined on A such that*

$$\partial_0 f(x) = \text{co}\{T_1(x), T_2(x), \dots, T_n(x)\} \quad \text{for every } x \in A.$$

This generalizes Corollary 2 in [8] to non-separable spaces.

Example 3. There are Clarke subdifferentials that can not be expressed as the cusco generated by a countable family of minimal cuscos. Indeed, let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a differentiable and nowhere monotone Lipschitz function [10], then $\{x : f \text{ is strictly differentiable at } x \text{ and } f'(x) = 0\}$ is residual in \mathbb{R} and so the only possible minimal cusco lying inside $\partial_0 f$ is $T \equiv \{0\}$.

Lemma 10 (Lemma 2.5, [15]). *A weak* cusco T from a topological space A into subsets of the dual of a Banach space X is a minimal weak* cusco if, and only if, given any open subset U of A and weak* closed convex subset K of X^* with $T(U) \not\subseteq K$, there exists a non-empty open subset $V \subseteq U$ such that $T(V) \cap K = \emptyset$.*

Theorem 6. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a Banach space X . If $\partial_0 f$ is a minimal weak* cusco and $T : A \rightarrow 2^{X^*}$ is defined by $T(x) := \partial_0 f(x) + B_{X^*}$, then $\{g \in \mathcal{X}_T : \partial_0 g(x) = T(x) \text{ for all } x \in A\}$ is residual in (\mathcal{X}_T, ρ) .*

Proof. For each $n \in \mathbb{N}$ choose a maximal disjoint family of an open ball $\{B_{1/n}(x_n^\alpha) : \alpha \in \Gamma_n\}$ in A and define $f_n : A \rightarrow \mathbb{R}$ by $f_n(x) := f(x) + d_{C_n}(x)$, where $C_n := \{x_n^\alpha : \alpha \in \Gamma_n\}$. Let G_n be any residual set in (\mathcal{X}_T, ρ) such that for each $g \in G_n$, $\partial_0 g(x) \cap \partial_0 f_n(x) \neq \emptyset$ for all $x \in A$. By [5] we know that the sum of a function whose Clarke subdifferential mapping is a minimal weak* cusco and a regular function has a Clarke subgradient that is a minimal weak* cusco. Therefore, $\partial_0 f_n$ is a

minimal weak* cusco on $B_{1/n}(x_n^\alpha)$ and so $\partial_0 f_n(x_n^\alpha) \subseteq \partial_0 g(x_n^\alpha)$ for each $\alpha \in \Gamma_n$. Set $G := \bigcap_{n=1}^\infty G_n$. We will show that for each $g \in G$, $\partial_0 g(x) = T(x)$ for all $x \in A$. To this end, let g be any member of G and let us suppose that $\partial_0 g(x_0) \neq T(x_0)$ for some $x_0 \in A$. Then there is some $y \in S_X$ and $\alpha \in \mathbb{R}$ so that

$$\begin{aligned} g^0(x_0, y) &= \max\{x^*(y) : x^* \in \partial_0 g(x_0)\} < \alpha \\ &< \max\{x^*(y) : x^* \in T(x_0)\} = f^0(x_0, y) + 1. \end{aligned}$$

Now, by the upper semi-continuity of $x \rightarrow g^0(x, y)$ there exists an open neighbourhood U of x_0 in A such that $g^0(z, y) < \alpha$ for all $z \in U$. On the other hand, from the minimality of $x \rightarrow \partial_0 f(x)$ (see Lemma 10) there exists a non-empty open subset V of U so that $\partial_0 f(V) \cap \{x^* \in X^* : x^*(y) \leq \alpha - 1\} = \emptyset$. Therefore,

$$\alpha - 1 < \min\{x^*(y) : x^* \in \partial_0 f(x)\} \equiv -f^0(x, -y) \quad \text{for all } x \in V.$$

Note that in particular, $f^-(x, y) > \alpha - 1$ for all $x \in V$. Next, we may choose $n \in \mathbb{N}$ so that $C_n \cap V \neq \emptyset$ and calculate $f_n^-(z, y) = f^-(z, y) + d'_{C_n}(z, y)$ for $z \in C_n \cap V$, to get $f_n^-(z, y) > \alpha - 1 + 1 = \alpha$. However, this is impossible since $f_n^-(z, y) \leq f_n^0(z, y) \leq g^0(z, y) < \alpha$ for all $z \in U \cap C_n$. Hence for each $g \in G$, $\partial_0 g(x) = T(x)$ for all $x \in A$. \square

An important special case of the above theorem is the following:

Corollary 9 (The dual ball). *Let A be a non-empty open subset of a Banach space X , then $\{g \in \mathcal{X}_{B_{X^*}} : \partial_0 g(x) = B_{X^*} \text{ for all } x \in A\}$ is residual in $(\mathcal{X}_{B_{X^*}}, \rho)$.*

We may also extend a central case of Theorem 1:

Corollary 10. *Let A be a non-empty open subset of an infinite dimensional smoothable Banach space X . Suppose that $\|\cdot\|$ is an equivalent norm on X such that $\text{ext } B_{X^*}$, the extreme points of its associated dual ball B_{X^*} , is weak* dense in B_{X^*} , then $\{g \in \mathcal{X}_{B_{X^*}} : \partial_a g(x) = B_{X^*} \text{ for all } x \in A\}$ is residual in $(\mathcal{X}_{B_{X^*}}, \rho)$.*

Proof. Let $G := \{g \in \mathcal{X}_{B_{X^*}} : \partial_0 g(x) = B_{X^*} \text{ for all } x \in A\}$. We claim that for each $g \in G$, $\partial_a g(x) = B_{X^*}$ for all $x \in A$. To see this, simply note that by the converse of the Krein-Milman theorem $\text{ext } B_{X^*} \subseteq \partial_a g(x)$ for each $g \in G$ and $x \in A$ since, $B_{X^*} = \partial_0 g(x) = \overline{\text{co}}^{w*} \partial_a g(x)$ and $\partial_a g(x)$ is weak* closed for each $x \in A$. \square

Remark 1. The hypotheses of the last corollary are satisfied if the norm $\|\cdot\|$ on X is smooth.

Corollary 11. *Let A be a non-empty open subset of a Banach space X , then*

$$\{g \in \mathcal{X}_{B_{X^*}} : \text{for each } v \in S_X, \{x \in A : g'(x; v) \text{ exists}\} \text{ is first category}\}$$

is residual in $(\mathcal{X}_{B_{X^}}, \rho)$.*

Proof. Let $G := \{g \in \mathcal{X}_{B_{X^*}} : \partial_0 g(x) = B_{X^*} \text{ for all } x \in A\}$. We claim that for each $g \in G$, $D_y := \{x \in A : g'(x; y) \text{ exists}\}$ is first category in A for each $y \in S_X$. To see this, let us fix $y \in S_X$. Then, by [16] there exists a dense G_δ subset P_y of A where $g^0(x; y) = g^+(x; y)$ and $-g^0(x; -y) = -g^+(x; -y)$ for each $x \in P_y$. We will

now show that $D_y \subseteq A \setminus P_y$. Indeed, if $x_0 \in P_y \cap D_y$, then

$$1 = g^0(x_0; y) = g'(x_0; y) = -g'(x_0; -y) = -g^0(x_0; -y) = -1,$$

which is absurd. Therefore, $D_y \subseteq A \setminus P_y$ and so first category in A . \square

5. WHEN IS \mathcal{X}_T NON-EMPTY?

Thus far we have not dwelt too much upon the question of when \mathcal{X}_T is non-empty. However, below we show that this issue is in fact finitely determined, that is, determined by the behaviour of T on finite dimensional subspaces.

Let A be a non-empty open subset of a Banach space X and let $T : A \rightarrow 2^{X^*}$ be a weak* cusco defined on A . Then for each subspace Y of X with $Y \cap A \neq \emptyset$ we define $T_Y : Y \cap A \rightarrow 2^{Y^*}$ by

$$T_Y(x) := \{y^* \in Y^* : y^* = x^*|_Y \text{ and } x^* \in T(x)\}.$$

Theorem 7. *Let A be a non-empty open connected subset of a Banach space X and let $T : A \rightarrow 2^{X^*}$ be a weak* cusco on A . Then $\mathcal{X}_T \neq \emptyset$ if and only if there exists an upwardly directed set (D, \subseteq) of finite dimensional subspaces of X such that (i) $A \subseteq \bigcup_{Y \in D} Y$ and (ii) $\mathcal{X}_{T_Y} \neq \emptyset$ for each $Y \in D$ with $Y \cap A \neq \emptyset$.*

Proof. It is clear that if $\mathcal{X}_T \neq \emptyset$ then (i) and (ii) are satisfied. So we will consider the converse question. Fix $x_0 \in A$ and define $\mathcal{X}_{T_Y}(x_0) := \{g \in \mathcal{X}_{T_Y} : g(x_0) = 0\}$. Note that by possibly making D smaller we may assume that $x_0 \in Y$ for all $Y \in D$. Then for each $Y \in D$ choose $g_Y \in \mathcal{X}_{T_Y}(x_0)$. For each such function we consider the following extension $\tilde{g}_Y : A \rightarrow \mathbb{R}$ defined by

$$\tilde{g}_Y(x) := \begin{cases} g_Y(x) & \text{if } x \in A \cap Y, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $(\tilde{g}_Y : D)$ is a net in $(\mathbb{R}_e)^A$ —which is compact by Tychonoff's theorem. Therefore $(\tilde{g}_Y : D)$ has a convergent subnet which converges to some element $g \in (\mathbb{R}_e)^A$. It is now routine to check that g is real-valued and locally Lipschitz on A . In fact one can show that $\partial_0 g(x) \subseteq T(x)$ for all $x \in A$ and $g(x_0) = 0$. This shows that $g \in \mathcal{X}_T(x_0) \subseteq \mathcal{X}_T$. \square

Remark 2. By applying a similar argument to the above proof we can show that for each $x_0 \in A$, $\mathcal{X}_T(x_0)$ is a pointwise compact, convex sub-lattice of \mathcal{X}_T .

The problem of determining when $\mathcal{X}_T \neq \emptyset$ now reduces to the semi-classical problem of determining when $\mathcal{X}_{T_Y} \neq \emptyset$. Below we give an obvious first step in this direction.

Let (\mathcal{M}, σ) be a measurable space and let X be a normed linear space. We say that a function $f : (\mathcal{M}, \sigma) \rightarrow X^*$ is *weak* measurable* if $f^{-1}(U) \in \sigma$ for each weak* open subset U of X^* . Note: if X is a separable normed linear space, then this is equivalent to demanding that for each $x \in X$, the mapping $\hat{x} \circ f : A \rightarrow \mathbb{R}$ defined by $(\hat{x} \circ f)(t) := f(t)(x)$ is measurable.

For a non-empty open subset A of a normed linear space X the *line integral* along a line segment $[a, b] \subseteq A$ of a weak* measurable function $f : (A, \mathcal{B}_A) \rightarrow X^*$ is the Lebesgue integral

$$\int_{[a,b]} f(z) dz := \int_0^1 f(tb + (1-t)a)(b-a) dt.$$

A *polygonal path* C in A is an ordered collection of line segments $\{[a_i, a_{i+1}] : 1 \leq i \leq n-1\}$ for some integer n . Such a path is said to be closed if $a_1 = a_n$. The *line integral* of f on C is defined as

$$\int_C f(z) dz := \sum_{i=1}^{n-1} \int_{[a_i, a_{i+1}]} f(z) dz.$$

For any fixed $\varepsilon > 0$ we will call an ordered collection of line segments $P(\varepsilon) := \{[a_i, b_i] : 1 \leq i \leq n-1\}$ an ε -path from a to b provided

$$\|a - a_1\| + \sum_{i=1}^{n-1} \|a_{i+1} - b_i\| + \|b_n - b\| < \varepsilon.$$

Such a path is closed if $a = b$. For a Borel subset E of A we say that P is an E -admissible ε -path from a to b if P is an ε -path from a to b and

$$\lambda(\{t \in [0, 1] : tb_i + (1-t)a_i \notin E\}) = 0$$

for $1 \leq i \leq n-1$. Line integrals on an ε -path are defined similarly as above.

Theorem 8 ([7]). *Let A be a non-empty open connected subset of a finite dimensional normed linear space X and let $T : A \rightarrow 2^{X^*}$ be a bounded weak* cusco on A . Then $\mathcal{X}_T \neq \emptyset$ if and only if there exists a Borel set $E \subseteq A$ with $\lambda(A \setminus E) = 0$ and a weak* measurable selection $f : (E, \mathcal{B}_E) \rightarrow X^*$ of T so that*

$$\lim_{\varepsilon \rightarrow 0^+} \int_{P(\varepsilon)} f(z) dz = 0,$$

where $P(\varepsilon)$ is any closed E -admissible ε -path in A .

We now turn to the existence of Lipschitz functions with ‘minimal’ subdifferential mappings. These results form a sharp contrast to the existence of Lipschitz functions with maximal subdifferential mappings given in Sections 3 and 4.

Theorem 9 (Minimal approximate subdifferential). *Let A be a non-empty open connected subset of a smoothable Banach space X and let $f : A \rightarrow \mathbb{R}$ be a locally Lipschitz function on A . Then there exists a locally Lipschitz $g : A \rightarrow \mathbb{R}$ such that $\partial_a g(x) \subseteq \partial_a f(x)$ for all $x \in A$ and $\partial_a g$ is minimal in the sense that for each \hat{g} with $\partial_a \hat{g}(x) \subseteq \partial_a g(x)$ for all $x \in A$ we have that $\partial_a \hat{g} = \partial_a g$.*

Proof. Fix $x_0 \in A$ and define,

$$\mathcal{P} := \{\text{Gr}(G) \subseteq \text{Gr}(\partial_a f) :$$

there exists a locally Lipschitz function g on A with $\partial_a g \equiv G\}.$

On \mathcal{P} , which is clearly non-empty, we may define a partial order ‘ \leq ’ by $G_1 \leq G_2$ if and only if $G_1 \subseteq G_2$. We will use Zorn’s lemma to show that (\mathcal{P}, \leq) has a minimal element. To this end, let C be a chain in \mathcal{P} . We will show that C has a lower bound in \mathcal{P} . For each $G \in C$ let g_G be a locally Lipschitz function on A such that $\partial_a g_G \equiv G$ and $g_G(x_0) = f(x_0)$. Thus, $(g_G : C)$ is a net in $(\mathbb{R}_e)^A$ —which is compact. Therefore $(g_G : C)$ has a subnet which converges to some element $g \in (\mathbb{R}_e)^A$. As with Theorem 7 it is routine to show that g is real-valued and locally Lipschitz on A . We claim that $\partial_a g$ is a lower bound for C . To prove this claim let us consider any $G_0 \in C$. We need to show that $\partial_- g(x) \subseteq \partial_a g_{G_0}(x)$ for all $x \in A$. Indeed, if this is not the case, then there exist an $x_1 \in A$ and $x_1^* \in \partial_- g(x_1)$ such that $x_1^* \notin \partial_a g_{G_0}(x_1)$. We may now select a finite set $F \subseteq S_X$ and $\alpha \in \mathbb{R}$ so that if

$W := \{x^* \in X^* : |x^*(y)| < 4\alpha, y \in F\}$, then $(x_1^* + W) \cap (\partial_a g_{G_0}(x_1) + W) = \emptyset$ and by the weak* upper semi-continuity of $x \rightarrow \partial_a g_{G_0}(x)$ there exists an $r > 0$ so that $\partial_a g_{G_0}(B_r(x_1)) \subseteq \partial_a g_{G_0}(x_1) + W$. Next we set $Y := \text{sp}F$. Then since g is Lipschitz around x_1 and S_Y is compact there exists a $0 < \delta < r$ so that $(g - x_1^*)(x_1 + \lambda v) - (g - x_1^*)(x_1) > -\lambda \cdot \alpha \geq -\delta \cdot \alpha$ for all $0 < \lambda \leq \delta$ and $v \in S_Y$. Now by the Arzela-Ascoli theorem the set $\{g_G|_{x_1 + \delta B_Y} : G \in C\}$ is relatively norm compact in $(C(x_1 + \delta B_Y), \|\cdot\|_\infty)$. Hence the net $(g_G|_{x_1 + \delta B_Y} : G \in C)$ has a subnet which converges to $g|_{x_1 + \delta B_Y}$ with respect to the sup-norm on $x_1 + \delta B_Y$. In particular, this means that we can choose $G \leq G_0$ so that,

$$(g_G - x_1^*)(z) - (g_G - x_1^*)(x_1) > -2\delta \cdot \alpha \quad \text{for all } z \in x_1 + \delta B_Y.$$

Therefore, by Lemma 8 there exists a $z \in B_\delta(x_1)$ and $z^* \in X^*$ with $z^* \in \partial_-(g_G - x_1^*)(z)$ and $\|z^*\|_Y < 4\alpha$. That is, $x_1^* + z^* \in \partial_- g_G(z)$. However, $x_1^* + z^* \in x_1^* + W$; which is impossible since $\partial_- g_G(z) \cap (x_1^* + W) \subseteq \partial_a g_{G_0}(z) \cap (x_1^* + W) = \emptyset$. This shows that $\partial_a g(x) \subseteq \partial_a g_{G_0}(x)$ for all $x \in A$. Hence by Zorn's lemma (\mathcal{P}, \leq) has a minimal element which is the desired subgradient. \square

One may similarly prove the following:

Theorem 10 (Minimal Clarke subdifferential). *Let A be a non-empty open connected subset of Banach space X and let $f : A \rightarrow \mathbb{R}$ be a locally Lipschitz function on A . Then there exists a locally Lipschitz $g : A \rightarrow \mathbb{R}$ such that $\partial_0 g(x) \subseteq \partial_0 f(x)$ for all $x \in A$ and $\partial_0 g$ is minimal in the sense that for each \hat{g} with $\partial_0 \hat{g}(x) \subseteq \partial_0 g(x)$ for all $x \in A$ we have that $\partial_0 \hat{g} = \partial_0 g$.*

As with an usco (cusco) mapping which may contain several minimal uscós (cuscos), an approximate (Clarke) subdifferential mapping may contain several minimal approximate (Clarke) subdifferential mappings. On the real line, Theorem 10 is clear by using the fact that every cusco contains a minimal cusco and that every minimal cusco on the line is a Clarke subdifferential map of some locally Lipschitz function. In fact on the line $\partial_0 g$ is a minimal cusco if and only if $\partial_0 g$ is minimal in the sense of Theorem 10 while the latter is equivalent to $\partial_a g$ being minimal in the sense of Theorem 9 as in one dimension the Clarke subdifferential determines the approximate subdifferential by Theorem 2.2 [2]. However, each strongly integrable function is minimal in the sense defined in Theorems 9 and 10 and there are integrable functions whose generalized subdifferentials are neither minimal weak* uscós nor minimal weak* uscós (see Example 7.1 in [6]).

In Corollary 3 we saw that the Clarke subdifferential mapping is closed under the operation of taking finite convex hulls. Hence it is natural to ask the question of whether the Clarke subdifferential is closed under the operation of taking finite intersections. The next example shows that this fails in a rather strong way.

Example 4. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be an everywhere differentiable and strictly increasing function such that $\{x \in \mathbb{R} : f'(x) = 0\}$ is dense in \mathbb{R} and $|f(x) - f(y)| < |x - y|$ for all $x, y \in \mathbb{R}$. Let $C := \text{epi}(f) := \{(x, y) \in \mathbb{R}^2 : f(x) \leq y\}$. Next, consider the distance function d_C defined on \mathbb{R}^2 by the l_1 norm and the set C . Let g be any Lipschitz function on \mathbb{R}^2 such that $\partial_0 g(x)$ is equivalently equal to the l_2 norm ball. If we define $T : \mathbb{R}^2 \rightarrow 2^{\mathbb{R}^2}$ by $T(x) := \partial_0 g(x) \cap \partial_0 d_C(x)$ for all $x \in \mathbb{R}^2$, then T is a cusco but $\mathcal{X}_T = \emptyset$. Indeed, if $h \in \mathcal{X}_T$, then $\partial_0 h(x) \subseteq \partial_0 d_C(x)$ for all $x \in \mathbb{R}^2$. However, as d_C is integrable [6] we must have that $\partial_0 h(x) = \partial_0 d_C(x)$ for all $x \in \mathbb{R}^2$.

But this is impossible since $\partial_0 d_C(x)$ is not contained in the l_2 ball for all values of $x \in \mathbb{R}^2$.

REFERENCES

1. J. M. Borwein, Minimal CUSCOS and subgradients of Lipschitz functions, in *Fixed Point Theory and its Applications*, Pitman Research Notes 252 (1991), 57–81. MR **92j**:46077
2. J. M. Borwein and S. Fitzpatrick, Characterization of Clarke subgradients among one-dimensional multifunctions, in *Proc. of the Optimization Miniconference II*, edited by B. M. Glover and V. Jeyakumar, (1995), 61–73.
3. J. M. Borwein and A. Ioffe, Proximal analysis in smooth spaces, *Set-Valued Anal.* **4** (1996), 1–24. MR **96m**:49028
4. J. M. Borwein and W. B. Moors, Null sets and essentially smooth Lipschitz functions, *SIAM J. Optim.* **8** (1998), 309–323. MR **99g**:49013
5. J. M. Borwein and W. B. Moors, Separable determination of integrability and minimality of the Clarke subdifferential mapping, *Proc. Amer. Math. Soc.* **128** (2000), 215–221. MR **2000e**:49025
6. J. M. Borwein and W. B. Moors, Essentially smooth Lipschitz functions, *J. Funct. Anal.* **149** (1997), 305–351. MR **98i**:58028
7. J. M. Borwein and W. B. Moors, Y. Shao, Subgradient representation of multi-functions, *J. Austral. Math. Soc. Ser. B.* **40** (1998), 1–13. MR **2001b**:49020
8. J. M. Borwein, W. B. Moors and X. Wang, Lipschitz functions with prescribed derivatives and subderivatives, *Nonlinear Anal.* **29** (1997), 53–64. MR **98j**:49019
9. J. M. Borwein and D. Preiss, A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions, *Trans. Amer. Math. Soc.* **303** (1987), 517–527. MR **88k**:49013
10. A. M. Bruckner, J. B. Bruckner, B. S. Thomson, *Real Analysis*, Prentice-Hall, Inc. 1997.
11. F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York, 1983. MR **85m**:49002
12. B. Dacorogna and P. Marcellini, General existence theorems for Hamilton-Jacobi equations in the scalar and vectorial cases, *Acta Math.* **178** (1997), 1–37. MR **98d**:35029
13. B. Dacorogna and P. Marcellini, Dirichlet problem for nonlinear first order partial differential equations, *Optimization methods in partial differential equations (South Hadley, MA, 1996)*, 43–57, *Contemp. Math.* 209, *Amer. Math. Soc., Providence, RI*, 1997. MR **98h**:35037
14. Marián J. Fabian, *Gâteaux Differentiability of Convex Functions and Topology: Weak Asplund Spaces*, Wiley Interscience, New York, 1997. MR **98h**:46009
15. J. R. Giles and W. B. Moors, A continuity property related to Kuratowski's index of non-compactness, its relevance to the drop property and its implications for differentiability, *J. Math. Anal. and Appl.* **178** (1993), 247–268. MR **94m**:46022
16. J. R. Giles and S. Sciffer, Locally Lipschitz functions are generically pseudo-regular on separable Banach spaces, *Bull. Austral. Math. Soc.* **47** (1993), 205–212. MR **94a**:58022
17. A. D. Ioffe, Approximate subdifferentials and applications II, *Mathematika* **33** (1986), 111–128. MR **87k**:49028
18. A. D. Ioffe, Approximate subdifferentials and applications III. The metric theory, *Mathematika* **36** (1989), 1–38. MR **90g**:49012
19. E. Michael, A note on closed maps and compact sets, *Israel J. Math.* **2** (1964), 173–176. MR **31**:1659
20. R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, Lecture Notes in Mathematics, 1364 Springer-Verlag Berlin Heidelberg, 1993. MR **94f**:46055
21. D. Preiss, Differentiability of Lipschitz functions on Banach spaces, *J. Funct. Anal.* **91** (1990), 312–345. MR **91g**:46051
22. D. Preiss, R. R. Phelps and I. Namioka, Smooth Banach spaces, weak Asplund spaces and monotone or usco mappings, *Israel J. Math.* **72** (1990), 257–279. MR **92h**:46021
23. D. Preiss and J. Tiser, Points of non-differentiability of typical Lipschitz functions, *Real Anal. Exchange* **20** (1995), 219–226. MR **95m**:26006
24. R. T. Rockafellar, *The Theory of Subgradients and Its Applications to Problems of Optimization: Convex and Nonconvex Functions*, Heldermann-Verlag, Berlin, 1981. MR **83b**:90126

25. R. T. Rockafellar, R. J-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1998. MR **98m**:49001
26. Xianfu Wang, Fine and pathological properties of subdifferentials, Ph. D. Thesis, Simon Fraser University, 1999.

CENTRE FOR EXPERIMENTAL AND CONSTRUCTIVE MATHEMATICS, DEPARTMENT OF MATHEMATICS AND STATISTICS, SIMON FRASER UNIVERSITY, BURNABY, B.C. V5A 1S6, CANADA
E-mail address: `jborwein@cecm.sfu.ca`

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF WAIKATO, PRIVATE BAG 3105 HAMILTON, NEW ZEALAND
E-mail address: `moors@math.waikato.ac.nz`

CENTRE FOR EXPERIMENTAL AND CONSTRUCTIVE MATHEMATICS, DEPARTMENT OF MATHEMATICS AND STATISTICS, SIMON FRASER UNIVERSITY, BURNABY, B.C. V5A 1S6, CANADA
E-mail address: `xwang@cecm.sfu.ca`