ESSENTIALLY STRICTLY DIFFERENTIABLE LIPSCHITZ FUNCTIONS

JONATHAN M. BORWEIN AND WARREN B. MOORS

Abstract. In this paper we address some of the most fundamental questions regarding the differentiability structure of locally Lipschitz functions defined on Banach spaces. For example, we examine the relationship between integrability, D-representability and strict differentiability. In addition to this, we show that on a large class of Banach spaces there is a significant family of locally Lipschitz functions which are integrable, D-representable and possess desirable differentiability properties. We also present some striking applications of our results to distance functions.

1. Introduction

The first goal of this paper is to show that there is a significant class of locally Lipschitz functions which possesses the property that each of its members, f, satisfies the following three conditions;

(i)f is D-representable, that is, f is Gateaux differentiable on some dense subset D of its domain and the Clarke subdifferential mapping, $x \to \partial f(x)$, is generated by the derivatives chosen from *any* dense subset of D;

(ii) f is integrable, that is, we may determine the function f, up to an additive constant, from its Clarke subdifferential mapping, $x \to \partial f(x)$, (provided of course, that the domain of f is connected);

(iii) f possesses differentiability properties similar to those enjoyed by continuous convex functions.

In addition to the fore-mentioned properties the class of functions that we exhibit, also possesses very desirable closure properties. For example, it is closed under addition, subtraction, multiplication and division (when this is defined), as well as, the lattice operations. But a further advantage with this class of functions is that it enables us to give a unified presentation of many previously known results, which up until now, appeared unrelated. The second goal of the paper is to examine the relationship between D-representability, integrability and almost everywhere strict differentiability. Of course, this second goal is closely related to our first goal. We mention here, that since this paper incorporates aspects from Set-valued Analysis, Measure Theory and Differentiability Theory we have made a conscious effort to make this paper (as far as possible) self-contained. In doing this,

1991 *Mathematics Subject Classification.* Primary 49J52; Secondary 46N10; 58C20; 58C07.

Key words and phrases. Lipschitz function, Distance function, D-representable, Integrable, Proximal Normal Formula, Minimal cusco, Haar-null set.

we hope to achieve our goal of making the material contained in this paper accessible to as large an audience as possible. We begin by recalling some preliminary definitions regarding the Clarke subdifferential mapping. A real-valued function f defined on a non-empty open subset A of a Banach space X, is *locally Lipschitz* on A, if for each $x_0 \in A$ there exists a $K > 0$ and a $\delta > 0$ such that

$$| f(x) - f(y) | \leq K \| x - y \| \text{ for all } x, y \in B(x_0, \delta).$$

For functions in this class, it is often instructive to consider the following three directional derivatives.

(1) The *upper Dini derivative* at $x \in A$, in the direction y, is given by,

$$f^+(x; y) \equiv \limsup_{\lambda \to 0^+} \frac{f(x + \lambda y) - f(x)}{\lambda}$$

(2) The *lower Dini derivative* at $x \in A$, in the direction y, is given by,

$$f^-(x; y) \equiv \liminf_{\lambda \to 0^+} \frac{f(x + \lambda y) - f(x)}{\lambda}$$

(3) The *Clarke generalized directional derivative* at $x \in A$, in the direction y, is given by,

$$f^0(x; y) \equiv \limsup_{\substack{z \to x \\ \lambda \to 0^+}} \frac{f(z + \lambda y) - f(z)}{\lambda}$$

It is immediate from these three definitions that for each $x \in A$ and each $y \in X$,

$$f^-(x; y) \leq f^+(x; y) \leq f^0(x; y).$$

Associated with the Clarke generalized directional derivative is the *Clarke subdifferential mapping*, which is defined by,

$$\partial f(x) \equiv \{g \in X^* : g(y) \leq f^0(x; y) \text{ for each } y \in X\}$$

The Clarke subdifferential mapping $x \to \partial f(x)$ has played a crucial role in the recent development of non-smooth analysis. Two reasons for this success are;

($*_1$) for each $x \in A$, $\partial f(x)$ is non-empty, convex and weak* compact, and for each weak* open subset W of X^*, $\{x \in A : \partial f(x) \subseteq W\}$ is an open subset of A;

($*_2$) at each point $x \in A$, where the Gateaux derivative of f exists, $\nabla f(x) \in \partial f(x)$.

Let us now examine some notions of differentiability that are associated with locally Lipschitz functions. We say that a function f is *differentiable* at x, *in the direction* y, if,

$$f'(x; y) \equiv \lim_{\lambda \to 0} \frac{f(x + \lambda y) - f(x)}{\lambda} \text{ exists.}$$

We say that f is *Gateaux differentiable* at x, if,

$$\nabla f(x)(y) \equiv \lim_{\lambda \to 0} \frac{f(x + \lambda y) - f(x)}{\lambda} \text{ exists for each } y \in X$$

and $\nabla f(x)$ is a continuous linear functional on X. Note that a locally Lipschitz function f maybe differentiable at a point x, in every direction $y \in X$, and still not be Gateaux differentiable at that point, since the mapping $y \to f'(x;y)$ may not be linear, although, it will be continuous when f is locally Lipschitz. If f is Gateaux differentiable at x and

$$\lim_{\lambda \to 0} \frac{f(x+\lambda y) - f(x)}{\lambda} \text{ is uniform over } y \in S(X); \text{ the unit sphere in } X$$

then f is said to be *Fréchet differentiable* at x (and $\nabla f(x)$ is called the Fréchet derivative of f at x).

In finite dimensions, the notions of Gateaux and Fréchet differentiability coincide for locally Lipschitz functions, ([12], p.30). However, outside of finite dimensions these notions are distinct, [6]. In this paper we are also interested in two slightly stronger notions of differentiability. A locally Lipschitz function f is said to be *strictly differentiable (strictly Fréchet differentiable)* at x, if for each $\varepsilon > 0$ there exists a $\delta > 0$ such that,

$$\left| \frac{f(z+ty) - f(z)}{t} - \nabla f(x)(y) \right| < \varepsilon$$

whenever $0 < t < \delta$ and $\|z - y\| < \delta$ (uniformly over $y \in S(X)$).

For continuous convex functions, Gateaux differentiability coincides with strict differentiability, as does, Fréchet differentiability with strict Fréchet differentiability. In general however, these concepts are different.

Example 1.1. *Let* $f(x) = \begin{cases} x^2 \sin(1/x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$

Then f is differentiable everywhere on $\mathbb{R}$, but f is not strictly differentiable at $x = 0$. In fact $f'(0) = 0$ while $\partial f(0) = [-1, 1]$.

Next, we recall that in general, there is an intimate connection between strict differentiability and single-valuedness of the Clarke subdifferential mapping.

Proposition 1.1. *(*[2]*, Proposition 3.1) Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a Banach space X. Then;*

(a) $\partial f(x)$ is a singleton if, and only if, f is strictly differentiable at x;

(b) $\partial f(x)$ is a singleton and has the property that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that $\partial f(B(x, \delta)) \subseteq \partial f(x) + \varepsilon B(X)$ if, and only if, f is strictly Fréchet differentiable at x.

For further information regarding the Clarke subdifferential mapping see, [12]. Apart from the fore-mentioned notions of differentiability, the other key concept contained in this paper is that of a minimal cusco. A set-valued mapping Φ from a topological space A into subsets of a topological (linear topological) space X is an *usco (cusco)* on A if;

(i) for each $t \in A$, $\Phi(t)$ is non-empty (convex) and compact;
(ii) for each open subset W of X, $\{t \in A : \Phi(t) \subseteq W\}$ is an open subset of A.

It follows from $(*_1)$, that the Clarke subdifferential mapping of any locally Lipschitz function defined on an open subset is always a weak* cusco. Amongst the class of usco (cusco) mappings special attention is given to the so-called minimal uscos (minimal cuscos). An usco (cusco) mapping Φ from a topological space A into subsets of a topological (linear topological) space X is called a *minimal usco (minimal cusco)* if its graph does not strictly contain the graph of any other usco (cusco) defined on A. It is immediate from this definition that all single-valued uscos (cuscos) are minimal, however, there are many important examples of minimal uscos (cuscos) which are not everywhere single-valued. We begin our study of minimal cuscos (minimal uscos) by establishing some of their basic properties. Although most of these facts are well-known, we include their proofs for the sake of completeness.

Proposition 1.2. [11] *Let Φ be an usco (cusco) mapping from a topological space A into subsets of a topological (linear topological) space X. Then there exists a minimal usco (minimal cusco) Ψ defined on A such that $\Psi(t) \subseteq \Phi(t)$ for each $t \in A$.*

Proof. Let $\mathcal{P}$ denote the family of all usco (cusco) mappings defined on A whose graphs are contained in that of Φ. Clearly, $\mathcal{P} \neq \emptyset$, as $\Phi \in \mathcal{P}$. We may partially order $\mathcal{P}$ as follows. Suppose Ψ_1 and Ψ_2 are members of $\mathcal{P}$, then we write $\Psi_1 \leq \Psi_2$ if $\Psi_1(t) \subseteq \Psi_2(t)$ for each $t \in A$.

We shall use Zorn's lemma to show that $(\mathcal{P}, \leq)$ possess a minimal element. To this end, let $\{\Psi_\gamma : \gamma \in \Gamma\}$ be a totally ordered subset of $\mathcal{P}$. Define $\Phi_M : A \to 2^X$ by $\Phi_M(t) \equiv \bigcap\{\Psi_\gamma(t) : \gamma \in \Gamma\}$. Since each $\Psi_\gamma(t)$ is non-empty (convex) and compact $\Phi_M(t)$ is non-empty (convex) and compact. Now, let W be an open subset of X and consider $U \equiv \{t \in A : \Phi_M(t) \subseteq W\}$.

We will show that U is an open subset of A. Note that, without loss of generality, we may assume that $U \neq \emptyset$. Let $t_0 \in U$. By the finite intersection property, for some $\gamma_0 \in \Gamma$, $\Psi_{\gamma_0}(t_0) \subseteq W$. Therefore, there exists an open neighbourhood U_0 of t_0 such that $\Psi_{\gamma_0}(U_0) \subseteq W$, which implies that $\Phi_M(U_0) \subseteq W$.

Thus, $t_0 \in U_0 \subseteq U$, and so U is open in A. From this, we see that $\Phi_M \in \mathcal{P}$ and $\Phi_M \leq \Psi_\gamma$ for each $\gamma \in \Gamma$. Hence, by Zorn's lemma, $(\mathcal{P}, \leq)$ possesses a minimal element. It now only remains to observe that this element is in fact a minimal usco (minimal cusco). ■

Let Ω be a set-valued mapping from a non-empty set A into a non-empty set X. Then by the *graph* of Ω we mean, $\mathrm{Gr}(\Omega) \equiv \{(t, x) \in A \times X : x \in \Omega(t)\}$ and by the (effective) *domain* of Ω we mean $\mathrm{Dom}(\Omega) \equiv \{t \in A : \Omega(t) \neq \emptyset\}$. When the domain of Ω is dense in A we say that Ω is *densely defined.*

It is worthwhile observing that for an usco mapping Φ from a topological space A into subsets of Hausdorff topological space X, the graph of Φ is a closed subset of $A \times X$, when $A \times X$ is endowed with the product topology. It is also interesting to see that to some extent the converse of this observation is true.

Proposition 1.3. [11] *Let Φ be an usco mapping from a topological space A into subsets of a topological space X and let Ω be a set-valued mapping from A into non-empty subsets of X. If $Gr(\Omega)$ is a closed subset of $A \times X$ and $Gr(\Omega) \subseteq Gr(\Phi)$, then Ω is an usco mapping on A.*

Proof. Since the graph of Ω is closed in $A \times X$, $\Omega(t)$ is a closed subset of X for each $t \in A$, and since $\Omega(t) \subseteq \Phi(t)$ for each $t \in A$, $\Omega(t)$ is a compact subset of X, for each $t \in A$. Let W be an open subset of X. Consider $U \equiv \{t \in A : \Omega(t) \subseteq W\}$. We will show that U is an open subset of A. Firstly, observe that without loss of generality, we may assume that $U \neq \emptyset$. Let $t_0 \in U$. We consider two cases.

(i) If $\Phi(t_0) \subseteq W$ then there exists an open neighbourhood U_0 of t_0 such that $\Omega(U_0) \subseteq \Phi(U_0) \subseteq W$, and so $t_0 \in U_0 \subseteq U$.

(ii) On the other hand, suppose that $K \equiv \Phi(t_0)\backslash W \neq \emptyset$. Then for each $x \in K$, choose open sets $U_x \subseteq A$ and $V_x \subseteq X$ such that $(t_0, x) \in U_x \times V_x$ and $(U_x \times V_x) \cap \mathrm{Gr}(\Omega) = \emptyset$. Since K is compact and $K \subseteq \bigcup\{V_x : x \in K\}$ there exists a finite subcover $\{V_{x_j} : 1 \leq j \leq n\}$ of $\{V_x : x \in K\}$ such that $K \subseteq \bigcup\{V_{x_j} : 1 \leq j \leq n\}$. Let $Z_1 \equiv \bigcap\{U_{x_j} : 1 \leq j \leq n\}$, and observe that for each $t \in Z_1$, $\Omega(t) \cap \bigcup\{V_{x_j} : 1 \leq j \leq n\} = \emptyset$. Now, $\Phi(t_0) \subseteq W \cup \bigcup\{V_{x_j} : 1 \leq j \leq n\}$ so there exists an open neighbourhood Z_2 of t_0 such that $\Omega(Z_2) \subseteq \Phi(Z_2) \subseteq \bigcup\{V_{x_j} : 1 \leq j \leq n\} \cup W$. Therefore, $\Omega(t) \subseteq W$ for each $t \in U_0 \equiv Z_1 \cap Z_2$. Hence in both cases, there exists an open neighbourhood U_0 of t_0 such that $t_0 \in U_0 \subseteq W$; which shows that U is open. ■

The next proposition gives further information on the construction of usco (cusco) mappings.

Proposition 1.4. *Let Ω be a densely defined set-valued mapping from a topological space A into subsets of a Hausdorff topological (separated locally convex topological) space X. If the graph of Ω is contained in the graph of a usco (cusco) mapping Φ, then there exists a unique smallest usco (cusco) containing Ω, denoted $USC(\Omega)$ ($CSC(\Omega)$), given by,*

$$USC(\Omega)(x) = \bigcap\{\overline{\Omega(V)} : V \text{ is an open neighbourhood of } x\}$$

$$(CSC(\Omega)(x) = \bigcap\{\overline{co}\Omega(V) : V \text{ is an open neighbourhood of } x\})$$

Proof. Let us begin with the following three observations;

(i) For each $t \in \mathrm{Dom}(\Omega)$, $\quad \Omega(t) \subseteq \mathrm{USC}(\Omega)(t) \quad (\Omega(t) \subseteq \mathrm{CSC}(\Omega)(t))$;
(ii) For any set-valued mapping Ω, $\mathrm{USC}(\Omega)$, $(\mathrm{CSC}(\Omega))$ possesses a closed graph;
(iii) If Ω is an usco (cusco) then $\Omega = \mathrm{USC}(\Omega) \quad (\Omega = \mathrm{CSC}(\Omega))$.

Next, we show that $\mathrm{USC}(\Omega)$ $(\mathrm{CSC}(\Omega))$ is an usco (cusco) mapping on A. From (iii) and the definition of $\mathrm{USC}(\Omega)$ $(\mathrm{CSC}(\Omega))$ it follows that,

$$\mathrm{Gr}(\mathrm{USC}(\Omega)) \subseteq \mathrm{Gr}(\mathrm{USC}(\Phi)) = \mathrm{Gr}(\Phi) \quad (\mathrm{Gr}(\mathrm{CSC}(\Omega)) \subseteq \mathrm{Gr}(\mathrm{CSC}(\Phi)) = \mathrm{Gr}(\Phi))$$

Furthermore, by (ii), we have that the graph of $\mathrm{USC}(\Omega)$ $(\mathrm{CSC}(\Omega))$ is closed, so by Proposition 1.3 it is sufficient to show $\mathrm{Dom}(USC(\Omega)) = A \quad (\mathrm{Dom}(CSC(\Omega)) = A)$.

Suppose, for the purpose of obtaining a contradiction, that there exists an element $t_0 \notin \mathrm{Dom}(\mathrm{USC}(\Omega))$ $(t_0 \notin \mathrm{Dom}(\mathrm{CSC}(\Omega)))$. For each $x \in \Phi(t_0)$ choose open sets $U_x \subseteq A$ and $V_x \subseteq X$ such that $(t_0, x) \in U_x \times V_x$ and $(U_x \times V_x) \cap \mathrm{Gr}(\mathrm{USC}(\Omega)) = \emptyset$ $((U_x \times V_x) \cap \mathrm{Gr}(\mathrm{CSC}(\Omega)) = \emptyset)$. Since $\Phi(t_0)$ is compact and $\Phi(t_0) \subseteq \bigcup\{V_x : x \in$

$\Phi(t_0)\}$ there exists a finite subcover $\{V_{x_j} : 1 \leq j \leq n\}$ of $\{V_x : x \in \Phi(t_0)\}$ such that $\Phi(t_0) \subseteq \bigcup\{V_{x_j} : 1 \leq j \leq n\}$. Let $U_1 \equiv \bigcap\{U_{x_j} : 1 \leq j \leq n\}$, and observe that for each $t \in U_1$, $\mathrm{USC}(\Omega)(t) \cap \bigcup\{V_{x_j} : 1 \leq j \leq n\} = \emptyset$ ($\mathrm{CSC}(\Omega)(t) \cap \bigcup\{V_{x_j} : 1 \leq j \leq n\} = \emptyset$). On the other hand, Φ is an usco (cusco) so there exists an open neighbourhood U_2 of t_0 such that $\Omega(U_2) \subseteq \Phi(U_2) \subseteq \bigcup\{V_{x_j} : 1 \leq j \leq n\}$. Let $U \equiv U_1 \cap U_2 \neq \emptyset$. Now by (i) we have that for each $t \in \mathrm{Dom}(\Omega) \cap U \neq \emptyset$, $\mathrm{USC}(\Omega)(t) \cap \bigcup\{V_{x_j} : 1 \leq j \leq n\} \neq \emptyset$ ($\mathrm{CSC}(\Omega)(t) \cap \bigcup\{V_{x_j} : 1 \leq j \leq n\} \neq \emptyset$). But this contradicts the fact that $\emptyset \neq U \cap \mathrm{Dom}(\Omega) \subseteq U_1$. Hence $\mathrm{Dom}(\mathrm{USC}(\Omega)) = A$ ($\mathrm{Dom}(\mathrm{CSC}(\Omega)) = A$); which shows that $\mathrm{USC}(\Omega)$ ($\mathrm{CSC}(\Omega)$) is an usco (cusco) on A. To see that $\mathrm{USC}(\Omega)$ ($\mathrm{CSC}(\Omega)$) is the smallest usco (cusco) containing Ω it suffices to observe that for any usco (cusco) Ψ containing Ω,

$$\mathrm{Gr}(\mathrm{USC}(\Omega)) \subseteq \mathrm{Gr}(\mathrm{USC}(\Psi)) = \mathrm{Gr}(\Psi) \ (\mathrm{Gr}(\mathrm{CSC}(\Omega)) \subseteq \mathrm{Gr}(\mathrm{CSC}(\Psi)) = \mathrm{Gr}(\Psi)). \quad \blacksquare$$

Note: In the above Proposition, the set-valued mapping, $\mathrm{CSC}(\Omega)$, is called the *cusco generated* by Ω and $\mathrm{USC}(\Omega)$ is called the *usco generated* by Ω.

Now that we have established some of the elementary properties and definitions concerning locally Lipschitz functions and minimal cuscos, we may discuss more precisely the connection between locally Lipschitz functions which possess the properties (i) - (iii), listed at the start of this paper, and minimal cuscos. At the heart of this relationship, is the fact that a locally Lipschitz function f defined on a non-empty open subset A, of a smooth Banach space (or more generally, a class($\mathcal{S}$) space) is D-representable if, and only if, its Clarke subdifferential mapping, $x \to \partial f(x)$, is a minimal weak* cusco on A (see, Corollary 3.8). However, to fully understand this statement we must first make precise what we mean by D-representable. Let f be a real-valued locally Lipschitz defined on a non-empty open subset A of a Banach space X. Then we say that f is *D-representable* on A if;

(a) $D \equiv \{x \in A : \nabla f(x) \text{ exists}\}$ is dense in A and

(b) for each dense subset D^* of D we have that $\partial f = \mathrm{CSC}(\Omega_{D^*})$, where $\Omega_{D^*} : D^* \to 2^{X^*}$ is defined by $\Omega_{D^*}(x) \equiv \{\nabla f(x)\}$.

Note that in particular, when f is D-representable, $\partial f = \mathrm{CSC}(\Omega_D)$.

So we see then, that by a desire to consider locally Lipschitz functions which are D-representable we are inextricably lead to consider locally Lipschitz functions whose Clarke subdifferential mappings are minimal, with respect to the family of weak* cusco mapping. There is however, one significant difference between these definitions, namely, the notion of minimality extends beyond the class of functions which are densely Gateaux differentiable. In addition to the notion of D-representability we need to also make precise what we mean by 'integrable'. Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a Banach space X. Then we say that f is *integrable* on A if, $\partial(f - g) \equiv \{0\}$ for each real-valued locally Lipschitz function g defined on A with $\partial g(x) \subseteq \partial f(x)$ for each $x \in A$. It follows from this, that if A is connected and f is integrable on A, then $f - g \equiv$ constant on A, whenever $\partial g(x) \subseteq \partial f(x)$ for each $x \in A$. In particular, this means that f is determined, up to an additive constant, from its Clarke subdifferential mapping.

Life would be simple if all D-representable functions automatically satisfied the conditions (ii) and (iii) given earlier, however, there are numerous examples (even on $\mathbb{R}$) of locally Lipschitz functions which are D-representable, but which fail to satisfy either (ii) or (iii). Furthermore, the class of D-representable functions is neither closed under addition, multiplication nor either of the lattice operations (see, Example 2.1). Therefore, in order to achieve our goal, we are forced to consider a proper subclass of the D-representable functions.

In this paper, we propose that the appropriate functions to consider (on a separable Banach space) are those functions which are strictly differentiable almost everywhere, that is, strictly differentiable everywhere except on a Haar-null set. That this is a reasonable class of functions to consider, derives from the following facts (a) on the real line, the locally Lipschitz functions which satisfy (i), (ii) and (iii) (actually, on the real line, any locally Lipschitz function which satisfies (ii) automatically satisfies (ii) and (iii)) are exactly those functions which are strictly differentiable almost everywhere on their domain, and (b) on a separable Banach space, continuous convex functions are strictly differentiable almost everywhere.

In Section two we show that; (a) on any Banach space, each member of the vector space generated by the pseudo-regular functions possesses a minimal subdifferential mapping; (b) minimality of the Clarke subdifferential mapping is not preserved under addition, multiplication nor either of the lattice operations; (c) on an Asplund space, those Lipschitz functions which possess a minimal subdifferential mapping are strictly Fréchet differentiable on a dense and G_δ subset of their domain, while those on a class($\mathcal{S}$) Banach space are strictly differentiable on a dense and G_δ subset of their domain; (d) minimality of the Clarke subdifferential mapping is separably determined. Section three begins by recalling some necessary topological prerequisites, which are then used to show that a densely Gateaux differentiable Lipschitz function is D-representable if, and only if, its Clarke subdifferential mapping is a minimal weak* cusco. Next, in Section four, we characterize when the Clarke subdifferential mapping is minimal, in terms of a 'quasi continuity' property possessed by the upper Dini derivative mapping $x \to f^+(x; y)$, (for each $y \in S(X)$). We then use this characterization to show that the distance function d_C generated by a set C possesses a minimal subdifferential mapping on X if, and only if, d_C possesses a minimal subdifferential mapping on $X \backslash C$. We begin Section 5 by showing that on any separable Banach space, the functions which are strictly differentiable almost everywhere satisfy the properties (i), (ii) and (iii) given at the start of this paper. Moreover, we show that the pseudo-regular and semi-smooth functions belong to this class (plus many others). Then, in Section 6, we extend the results from Section 5 to non-separable Banach spaces. In Section 7 we show how the results of Section 6 maybe applied to perturbation functions. Section 8 concerns distance functions; more specifically, in this section we examine when the Clarke subdifferential mapping of a distance function is a minimal weak* cusco. In doing this, we are able to derive a 'Proximal Normal Formula', which holds for all non-empty closed subsets of any reflexive Banach space which possesses a smooth and sequentially Kadec norm. Moreover, we show that if such a proximal normal formula holds for all such subsets, then the space is necessarily reflexive and the norm is necessarily smooth and sequentially Kadec. In Section 9 we re-examine integrability and

D-representability. In particular, we show that D-representability does not imply integrability and that integrability does not imply D-representability. In fact, we show that integrability does not even imply dense strict differentiability. We also show that integrability in *not* inherited by open subsets. Finally, in this section, we give a global property possessed by a Lipschitz function which is sufficient to guarantee integrability. Lastly, in Section 10, we give some examples which highlight some of the behaviour possessed by functions whose Clarke subdifferential mappings are minimal.

2. Basic facts

In this section of the paper we shall address some of the most fundamental questions regarding the minimality of the Clarke subdifferential mappings. For instance, we will establish that on any Banach space there is a significant class of locally Lipschitz functions whose Clarke subdifferential mappings are *minimal* (that is, the mapping $x \to \partial f(x)$ a minimal weak* cusco). Unfortunately, we will also see that the class of functions whose Clarke subdifferential mappings are minimal is not closed under addition, multiplication nor either of the lattice operations. After this, we will answer a natural question, by giving an example of an everywhere differentiable Lipschitz function defined on $\mathbb{R}$ fails to possess a minimal subdifferential mapping. This contrasts strongly with the fact that any locally Lipschitz function which is strictly differentiable everywhere, trivially possesses a minimal subdifferential mapping. Finally, we end this section by examining some of the favourable differentiability properties enjoyed by locally Lipschitz functions whose Clarke subdifferential mappings are minimal. Of course, it is unreasonable to expect that this class of functions to have desirable differentiability properties on all Banach spaces (since in particular, it includes all the continuous convex functions). However, we will show that functions from this class do enjoy differentiability properties very similar to those enjoyed by continuous convex functions, at least in a Baire categorical sense. Though, as we shall see later, this is not necessarily true in a measure theoretic sense, (see, Example 10.1).

Let us begin by examining the most important question concerning the minimality of the Clarke subdifferential mapping. Namely, let us establish that on any Banach there do indeed exist non-trivial examples of locally Lipschitz functions with minimal subdifferential mappings. This question was first answered affirmatively in ([6], Theorem 4.7), in terms of the following definition. A locally Lipschitz function f defined on a non-empty open subset of A of a Banach space X is *pseudo-regular* at a point x, *in the direction* y, if $f^+(x;y) = f^0(x;y)$, and we say that f is *pseudo-regular* at a point x, if f is pseudo-regular at x in every direction y. It is standard, (see, [12], Proposition 2.27) that all continuous convex functions are pseudo-regular at each point of the interior of their domain, and so on any normed linear space there are non-trivial examples of pseudo-regular functions. Of course, it goes without saying that all C^1 functions are also pseudo-regular on their domain.

Proposition 2.1. *([30], Corollary 2.7) Let A be a non-empty open subset of a Banach space X, and let $P(A)$ denote the vector space generated by all the pseudo-regular functions defined on A. Then each member of $P(A)$ possesses a minimal subdifferential mapping. In particular, the difference of any two continuous convex functions defined on A, (that is, any dc function on A) possesses a minimal subdifferential mapping.*

Now that we have established the existence of a reasonably large class of functions whose subdifferential mappings are minimal, it makes sense to ask questions about the various closure properties of this class of functions. For example, is this class of functions closed under addition, multiplication or either of the lattice operationsΓ As previously mentioned, the answer to all of these questions is 'no'.

Example 2.1. *Let C be a Cantor subset of $[0,1]$ (symmetric about $\frac{1}{2}$) with $1 > \mu(C) > 0$, and let $\{(a_n, b_n) : n \in \mathbb{N}\}$ be an enumeration of the disjoint open intervals of $[0,1]\backslash C$. Further, for each $n \in \mathbb{N}$, let $c_n \equiv (a_n + b_n)/2$ and $d_n \equiv (b_n - a_n)^2/2$. Now, consider the Lipschitz functions $f : (0,1) \to [0,1]$ and $g : (0,1) \to [-1,1]$ defined by*

$$g(x) \equiv \begin{cases} 0 & \text{if} \quad \mid x - c_n \mid \geq d_n \text{ for all } n. \\ 2(x - (c_n - d_n)) & \text{if} \quad x \in (c_n - d_n, c_n - \frac{2}{3}d_n] \\ -(x - c_n) & \text{if} \quad x \in (c_n - \frac{2}{3}d_n, c_n] \\ -2(x - c_n) & \text{if} \quad x \in (c_n, c_n + \frac{1}{3}d_n] \\ x - (c_n + d_n) & \text{if} \quad x \in (c_n - \frac{1}{3}d_n, c_n - d_n) \end{cases}$$

$$\text{and } f(x) \equiv \int_{[0,x]} \theta(t)dt \quad \text{where} \quad \theta(t) \equiv \begin{cases} 0 & \text{if} \quad t \in [0,1]\backslash C \\ 2 & \text{if} \quad t \in C \cap [0, 1/2] \\ -2 & \text{if} \quad t \in C \cap (1/2, 1] \end{cases}$$

We shall show later, (see, Example 10.1) that:

(1) g, $f+g$ and $f-g$ possess minimal subdifferential mappings on $(0,1)$.

(2) $\partial g = \partial(f+g)$, but $(f+g) - g$ is not a constant function on $(0,1)$, that is, we cannot determine g, up to an additive constant, from its Clarke subdifferential mapping.

(3) If $h \equiv f+g$ and $k \equiv f-g$, then $h+k$ does not possess a minimal subdifferential mapping.

(4) If M and m are defined by $M(x) \equiv \max\{k(x), h(x)\}$ and $m(x) \equiv \min\{k(x), h(x)\}$, then neither M nor m possesses a minimal subdifferential mapping on $(0,1)$.

(5) If $j : [0,1] \to [0,2]$ is defined by $j(x) \equiv f(x) + 1$, then the functions $j+g$ and $j-g$ possess minimal subdifferential mappings on $(0,1)$, but $(j+g)\cdot(j-g)$ does not.

As promised in the introduction to this section, we give an example on $\mathbb{R}$ of an everywhere differentiable Lipschitz function which fails to possess a minimal subdifferential mapping.

Example 2.2. *On page 216, Example 6 part(e) of [39] an example is given of an everywhere differentiable Lipschitz function $f : \mathbb{R} \to \mathbb{R}$ which is strictly increasing* *on $\mathbb{R}$ and for which the set $\{t \in \mathbb{R} : f'(t) = 0\}$ is dense in $\mathbb{R}$. We claim that the Clarke subdifferential mapping of f is not a minimal cusco on $\mathbb{R}$.*

To see this, consider the mapping $\sigma : \mathbb{R} \to 2^{\mathbb{R}}$ defined by $\sigma(t) \equiv \{0\}$. Since the mapping $t \to \partial f(t)$ is upper semi-continuous on $\mathbb{R}$ we must have that $\sigma(t) \in \partial f(t)$ for each $t \in \mathbb{R}$. However, σ is a cusco on $\mathbb{R}$, and so if $t \to \partial f(t)$ were minimal on $\mathbb{R}$ then $\partial f = \sigma$. But this is not possible, because if $\partial f \equiv \{0\}$, then by the mean-value theorem (for differentiable functions) f would be constant on $\mathbb{R}$; which it is not.

Next, we examine the differentiability properties enjoyed by those functions which possess a minimal subdifferential mapping. To do this, we need to recall some definitions. We call a Banach space X an *Asplund* space if every continuous convex function defined on a non-empty open subset of X is Fréchet differentiable on a dense subset of its domain. It is shown in Chapter 7 of ([32], Lemma 7.14) that if X is an Asplund space then every minimal weak* cusco from a Baire space (for example, an open subset of a complete metric space) into subsets of X^* is single-valued and norm upper semi-continuous at the points of a dense and G_δ subset of its domain. This result, in conjunction with Proposition 1.1 part(b) yields the following fact.

Theorem 2.2. *([2], Theorem 1.6) Let f be a real-valued locally Lipschitz function defined on a non-empty open A subset of a Asplund space X. If the Clarke subdifferential mapping of f is minimal, then f is strictly Fréchet differentiable on a dense and G_δ subset of A.*

We see then that Asplund spaces have desirable Fréchet differentiability implications for a class of locally Lipschitz functions which is considerably larger than just the continuous convex functions. The previous theorem also has significant consequences for D-representability.

Corollary 2.3. *([2], Corollary 2.5) Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a Asplund space X. If the Clarke subdifferential mapping of f is minimal, then f is D-representable on A. In particular, if $D^* \equiv \{x \in A : f$ is strictly Fréchet differentiable at $x\}$, then $\partial f = CSC(\Omega_{D^*})$, where $\Omega^* : D^* \to 2^{X^*}$ is defined by $\Omega_{D^*}(x) \equiv \{\nabla f(x)\}$.*

A Banach X space is said to be of *class*$(\mathcal{S})$ if every minimal weak* cusco from a Baire space into subsets of X^* is single-valued at the points of a dense and G_δ of its domain. It is well-known that if a Banach space X is of class$(\mathcal{S})$ then every continuous convex function defined on a non-empty open convex subset of X is Gateaux differentiable on a dense and G_δ subset of its domain. In fact, this was the original motivation for this class of spaces. In the other direction, it is still an open question as to whether a Banach space X, which has the property that, every continuous convex function defined on a non-empty open convex subset of X is Gateaux differentiable at the points of a dense and G_δ subset of its domain (that is, a *weak Asplund space*), is necessarily of class$(\mathcal{S})$. Our interest in spaces of class$(\mathcal{S})$ is revealed in the next theorem.

Theorem 2.4. *([2], Theorem 1.6) Let f be a real-valued locally Lipschitz function defined on a non-empty open A subset of a class($\mathcal{S}$) space X. If the Clarke subdifferential mapping of f is minimal, then f is strictly differentiable on a dense and G_δ subset of A.*

Proof. It follows from the definition of a class($\mathcal{S}$) space that the mapping $x \to \partial f(x)$ is single-valued on a dense and G_δ subset of A. The result now follows from Proposition 1.1 part(a). ■

As before, this has immediate and important consequences.

Corollary 2.5. *([2], Corollary 2.5) Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a class($\mathcal{S}$) space X. If the Clarke subdifferential mapping of f is minimal, then f is D-representable on A. In particular, if $D^* \equiv \{x \in A : f \text{ is strictly differentiable at } x\}$, then $\partial f = CSC(\Omega_{D^*})$, where $\Omega^* : D^* \to 2^{X^*}$ is defined by $\Omega_{D^*}(x) \equiv \{\nabla f(x)\}$.*

As all Asplund spaces are class($\mathcal{S}$) spaces, this Corollary partially supercedes Corollary 2.3. The family of all class($\mathcal{S}$) spaces is quite large. Indeed, all *smooth* Banach spaces (that is, spaces which admit an equivalent norm which is Gateaux differentiable everywhere, except at 0) belong to this class (see, [35]), as do, all those Banach spaces which contain, as a dense subspace, the continuous linear image of an Asplund space (such spaces are called *GSG* spaces, ([38], p.73)). All reflexive spaces are Asplund. Thus, all separable and WCG (*weakly compactly generated*) spaces are GSG spaces. It is also known, that not every Asplund space admits an equivalent smooth norm.

Remark 2.1. *The significance of Theorems 2.2 and 2.4 will become clearer later on (in Sections 6 and 8), when we establish that on any Banach space there is a large class of locally Lipschitz functions whose Clarke subdifferential mappings are minimal. To give a small taste of what is to come, we mention here, that all distance functions (with respect to any uniformly Gateaux differentiable norm) possess minimal subdifferential mappings.*

We will see in the next section that a locally Lipschitz function f defined on a class($\mathcal{S}$) Banach space is D-representable if, and only if, its Clarke subdifferential mapping, $x \to \partial f(x)$, is a minimal weak* cusco. Let us end this section with the following general fact, which shows that minimality, and so D-representability, of the Clarke subdifferential mapping is separably determined.

Theorem 2.6. *([30], Theorem 2.9) Let f be a locally Lipschitz function defined on a non-empty open subset A of a Banach space X. If for each closed separable subspace Y of X with $A \cap Y \neq \emptyset$ we have that $x \to \partial(f|_{Y\cap A})(x)$ is a minimal weak* cusco on $Y \cap A$, then $x \to \partial f(x)$ is a minimal weak* cusco on A.*

3. Some topological prerequisites

Throughout the remainder of this paper we shall be interested in the topological behaviour of minimal cuscos and to a lesser extent minimal uscos. So in this section we gather-up some further facts concerning minimal uscos and minimal cuscos. Perhaps the most important amongst these is the following characterization.

Theorem 3.1. *([20], Lemma 2.5) A cusco (usco) Φ from a topological space A into subsets of a separated locally convex topological space (Hausdorff topological space) X is a minimal cusco (minimal usco) on X if, and only if, given any open subset U of A and closed and convex subset (closed subset) K in X, with $\Phi(U) \not\subseteq K$, there exists a non-empty open subset V of U such that $\Phi(V) \cap K = \emptyset$.*

Proof. Suppose that Φ is a minimal cusco (minimal usco). Let U be a non-empty open subset of A and let K be a closed and convex subset (closed subset) of X such that $\Phi(U) \not\subseteq K$. If for some $t_0 \in U$, $\Phi(t_0) \cap K = \emptyset$ then there exists an open neighbourhood V of t_0 such that $\Phi(V) \cap K = \emptyset$. So we consider the case when $\Phi(t) \cap K \neq \emptyset$ for each $t \in U$. In this case, we define a cusco (an usco) mapping Φ' on A by,

$$\Phi'(t) \equiv \begin{cases} \Phi(t) \cap K & \text{if} \quad t \in U \\ \Phi(t) & \text{if} \quad t \notin U \end{cases}$$

However, since Φ is a minimal cusco (minimal usco) on A we must have that $\Phi' = \Phi$; in which case, $\Phi(U) \subseteq K$. But this is impossible, since we assumed that $\Phi(U) \not\subseteq K$. Therefore, there must exist a point $t_0 \in U$ such that $\Phi(t_0) \cap K = \emptyset$, and so the result follows, as above. Conversely, let us suppose that Φ satisfies the condition given in the statement of the Theorem but that Φ is not a minimal cusco (minimal usco) on A. Then there exists a cusco (usco) Φ' on A whose graph is contained in that of Φ and $\Phi'(t_0) \neq \Phi(t_0)$ for some $t_0 \in A$. Consider, $x \in \Phi(t_0) \backslash \Phi'(t_0)$. Since $\Phi'(t_0)$ is compact and convex (compact) and X is Hausdorff, there exist disjoint open half spaces (open sets) W_1 and W_2 such that $\Phi'(t_0) \subseteq W_1$ and $x \in W_2$. Now since Φ' is a cusco (an usco) there exists an open neighbourhood V' of t_0 such that $\Phi'(V') \subseteq W_1$. On the other hand, $\Phi(V') \not\subseteq (X \backslash W_2)$ therefore, by the assumed condition there exists a non-empty open subset V of V' such that $\Phi(V) \cap (X \backslash W_2) = \emptyset$. But this is not possible, since $\Phi'(V) \subseteq (X \backslash W_2)$. Hence, we may conclude that Φ is a minimal cusco (minimal usco) on A. ■

From this characterization we may deduce the following useful corollary which says that minimality is locally determined.

Corollary 3.2. *Let Φ be a set-valued mapping from a topological space A into subsets of a separated locally convex topological space (Hausdorff topological space) X. Then Φ is a minimal cusco (usco) on A if, and only if, for each $t_0 \in A$ there exists an open neighbourhood U of t_0 such that Φ restricted to U is a minimal cusco (minimal usco) on U.*

Remark 3.1. *From Corollary 3.2, it follows that* Φ *is a minimal cusco (minimal usco) on* A *if, and only if, for each non-empty open subset* W *of* A*, the restriction of* Φ *to* W *is a minimal cusco (minimal usco) on* W.

We see next that the minimality of a cusco (usco) mapping is preserved under composition with a continuous linear (continuous) function.

Theorem 3.3. *Let* Φ *be a minimal cusco (minimal usco) from a topological space* A *into subsets of a separated locally convex topological space (Hausdorff topological space)* X *and let* f *be a continuous linear mapping (continuous mapping) from* X *into a separated locally convex topological space (Hausdorff topological space)* Y*. Then the mapping* $x \to f(\Phi(x))$ *is a minimal cusco (minimal usco) on* A.

Proof. Clearly, $f \circ \Phi$ is a cusco (an usco) on A, so it remains to show that it is a minimal cusco (minimal usco) on A. Consider a closed and convex subset (closed subset) K of Y and an open set U in A such that $(f \circ \Phi)(U) \not\subseteq K$. Since f is continuous and linear (continuous) on X, $f^{-1}(K)$ is a closed and convex subset (closed subset) of X. Since Φ is a minimal cusco (minimal usco) and $\Phi(U) \not\subseteq f^{-1}(K)$ there exists a non-empty open set $V \subseteq U$ such that $\Phi(V) \cap f^{-1}(K) = \emptyset$. Hence, $(f \circ \Phi)(V) \cap K = \emptyset$. ■

The following proposition shows that in general there is a close connection between minimal uscos and minimal cuscos.

Proposition 3.4. [23] *Suppose* Ψ *is a minimal usco and* Φ *is a cusco, which both map from a topological space* A *into subsets of a separated locally convex topological space* X*. If* $\Psi(t) \subseteq \Phi(t)$ *for each* $t \in A$*, then the set-valued mapping* $\Psi' : A \to 2^X$ *defined by* $\Psi'(t) \equiv \overline{co}\Psi(t)$ *is a minimal cusco on* A*, and* $\Psi'(t) \subseteq \Phi(t)$ *for all* $t \in A$.

Proof. Let us show first that Ψ' is a cusco on A. It is easy to see that for each $t \in A$, $\Psi'(t)$ is non-empty, convex and compact. Next, let W be a non-empty open subset of X and consider the set $U \equiv \{t \in A : \Psi'(t) \subseteq W\}$. We may, without loss of generality, assume that $U \neq \emptyset$. So let $t_0 \in U$. Since X is a separated locally convex topological space and $\Psi'(t_0)$ is compact, there exists a convex open neighbourhood N of 0 in X such that $\Psi(t_0) \subseteq \Psi'(t_0) + N \subseteq \Psi'(t_0) + \overline{N} \subseteq W$. Now, Ψ is an usco on A so there exists an open neighbourhood V of t_0 such that $\Psi(V) \subseteq \Psi'(t_0) + N$. On the other hand, $\Psi'(t_0) + \overline{N}$ is closed and convex and so $\Psi'(t) = \overline{co}\Psi(t) \subseteq \Psi'(t_0) + \overline{N} \subseteq W$ for each $t \in V$. Therefore $t_0 \in V \subseteq U$; which shows that Ψ' is a cusco on A. Next we show that Ψ' is in fact a minimal cusco. To this end, let Ω be a minimal cusco on A whose graph is contained in that of Ψ'. By Proposition 1.2 we know that such a minimal cusco exists. We complete the proof by showing that $\Omega = \Psi'$. Indeed, suppose that this is not the case, then for some $t_0 \in A$, $f \in X^*$ and $x \in \Psi'(t_0) \backslash \Omega(t_0)$, $f(x) > \alpha > \max\{f(y) : y \in \Omega(t_0)\}$. Since Ω is a cusco on A there exists an open neighbourhood U of t_0 such that $\sup\{f(y) : y \in \Omega(U)\} < \alpha$. Now by the definition of Ψ' we must have that $\Psi(U) \not\subseteq \{y \in X : f(y) \leq \alpha\}$. Hence by Theorem 3.1 there exists a non-empty open subset $V \subseteq U$ such that $\inf\{f(y) : y \in \Psi'(V)\} \geq \alpha$, which is impossible, since $\Omega(V) \subseteq \Psi'(V)$. Therefore we may conclude that Ψ' is a minimal cusco on A. ■

Remark 3.2. *In the above proof, the only place where we used the fact that* $\Psi(t) \subseteq \Phi(t)$ *for each* t*, was where we deduced the compactness of* $\overline{co}\Psi(t)$*, and so this condition is not needed when* X *is quasi-complete.*

Theorem 3.5. *Consider a minimal cusco (minimal usco)* Φ *from a topological space* A *into subsets of a separated locally convex topological (Hausdorff topological) space* X*.*

(i) Given a continuous real-valued function g *defined on* A*, the set-valued mapping* $g \cdot \Phi$ *is a minimal cusco (minimal usco) on* A*.*

(ii) Given a continuous mapping T *from* A *into* X*, the set-valued mapping* $T + \Phi$ *is a minimal cusco (minimal usco) on* A*.*

Proof. (i) In the case when Φ is a minimal usco, $g \cdot \Phi$ is the composition of the continuous mapping P, from $\mathbb{R} \times X$ into X defined by $P(t, x) = t \cdot x$ with the minimal usco mapping $t \to (g(t), \Phi(t))$ from A into $\mathbb{R} \times X$. Therefore by Theorem 3.3, $g \cdot \Phi$ is a minimal usco. In the case when Φ is a minimal cusco, consider the following. Let Ψ be a minimal usco whose graph is contained in Graph(Φ). By above, $g \cdot \Psi$ is a minimal usco on A and $g(t)\Psi(t) \subseteq g(t)\Phi(t)$ for all $t \in A$. Now, by Proposition 3.4, $\overline{co}\Psi(t) = \Phi(t)$ for all $t \in A$. Therefore, $g(t)\Phi(t) = g(t)\overline{co}\Psi(t) = \overline{co}(g(t)\Psi(t))$ for all $t \in A$. So by again appealing to Proposition 3.4 we have that $t \to \overline{co}(g(t)\Psi(t))$ is a minimal cusco and so $g \cdot \Phi$ is a minimal cusco.

(ii) The mapping $T + \Phi$ is the composition of the continuous linear mapping $S : X \times X \to X$ defined by $S(x, y) = x + y$ with the minimal cusco (minimal usco) mapping $t \to (T(t), \Phi(t))$ from A into $X \times X$, and so $T + \Phi$ is a minimal cusco (minimal usco) by Theorem 3.3. ■

Recently, the notion of minimality, for a set-valued mapping, has been successfully extended outside the class of cusco (usco) mappings, (see, for example, [19], [21], [30] and [25]). The key to these extensions is Theorem 3.1. A set-valued mapping Φ from a topological space A into non-empty subsets of a linear topological space X is *hyperplane minimal* if for any open half-space W in X and open set U in A with $\Phi(U) \cap W \neq \emptyset$ there exists a non-empty open subset $V \subseteq U$ such that $\Phi(V) \subseteq W$. Similarly, we say that a set-valued mapping Φ from a topological space A into non-empty subsets of a topological space X is *minimal* if for any open set W in X and open set U in A with $\Phi(U) \cap W \neq \emptyset$ there exists a non-empty open subset $V \subseteq U$ such that $\Phi(V) \subseteq W$. It follows then, from Theorem 3.1, that a cusco (usco) mapping from a topological space A into subsets of a separated locally convex topological (Hausdorff topological) space X is a minimal cusco (usco) on A if, and only if, it is hyperplane minimal (minimal) on A.

Another important notion in the analysis of set-valued mappings, and minimal mappings in particular, is that of a selection. Let Φ be a set-valued mapping from a non-empty set A into a non-empty set X. Then a function $f : A \to X$ is called a *selection of* Φ if $f(t) \in \Phi(t)$ for each $t \in A$.

Corollary 3.6. *Let* Ω *be a densely defined set-valued mapping from a topological space* A *into subsets of a separated locally convex topological (Hausdorff topological) space* X*. If the graph of* Ω *is contained in the graph of a cusco (usco)* Φ*,* *then* $CSC(\Omega)$ $(USC(\Omega))$ *is a minimal cusco (minimal usco) if, and only if,* Ω *is hyperplane minimal (minimal).*

Proof. The proof is a straight forward application of Theorem 3.1. ■

Next, we give several *'useful'* characterizations of minimality.

Theorem 3.7. *For a cusco mapping* Φ, *from a topological space* A *into subsets of a separated locally convex topological space* X, *the following conditions are equivalent;*

(i) Φ *is a minimal cusco on* A*;*

(ii) there exists a densely defined, hyperplane minimal selection σ *of* Φ *such that* $CSC(\sigma) = \Phi$;

(iii) for any densely defined selection σ *of* Φ, $CSC(\sigma) = \Phi$*;*

(iv) there exists a densely defined selection σ *of* Φ *such that* $\Phi = CSC(\sigma|_D)$ *for each dense subset* D *of Dom(*σ*).*

Proof. Corollary 3.6 gives us that (i)$\Leftrightarrow$(ii) and clearly (i) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (iv). So it remains to show that (iv) $\Rightarrow$ (i). We proceed via the characterization given in Theorem 3.1 . To this end, let U be a non-empty open subset of A and suppose that $\Phi(U) \not\subseteq K$, where $K \equiv \{x \in X : f(x) \leq \alpha\}$, $\alpha \in \mathbb{R}$ and $f \in X^*$. Choose $x_0 \in \Phi(U)\backslash K$ such that $f(x_0) > \alpha + \varepsilon$, for some $\varepsilon > 0$ and set $D' \equiv \{t \in \mathrm{Dom}(\sigma) \cap U : f(\sigma(t)) \leq \alpha + \varepsilon\}$. Clearly D' is not dense in U, because if D' were dense in U then by hypothesis $\Phi = CSC(\sigma|_{D^*})$ where $D^* \equiv D' \cup \mathrm{Dom}(\sigma)\backslash U$, and this would imply that $\sup\{f(x) : x \in \Phi(U)\} \leq \alpha + \varepsilon$; which is clearly not true. Therefore, there exists a non-empty open subset V of U such that $V \cap D' = \emptyset$. Now consider $CSC(\sigma|_{\mathrm{Dom}(\sigma)})$. Again by hypothesis, $CSC(\sigma|_{\mathrm{Dom}(\sigma)}) = \Phi$, but for each $t \in V \cap \mathrm{Dom}(\sigma)$, $f(\sigma(t)) > \alpha + \varepsilon$, therefore, $\Phi(V) \cap K = CSC(\sigma|_{\mathrm{Dom}(\sigma)})(V) \cap K = \emptyset$. ■

This theorem has some important consequences for differentiability theory.

Corollary 3.8. *Let* f *be a densely Gateaux differentiable real-valued locally Lipschitz function defined on a non-empty open subset* A *of a Banach space* X. *Then* f *possesses a minimal Clarke subdifferential mapping if, and only if,* f *is* D*-representable.*

Proof. This result follows from parts (i) and (iv) in the above theorem. ■

Sometimes it is convenient to express D-representability in terms of sequences. So our next task is to show that on any class($\mathcal{S}$) Banach space, D-representability maybe characterized in terms of sequential limits of Gateaux derivatives.

Lemma 3.9. *(*[2]*, Lemma 1.4 part(b)) Let* X *be a Banach space whose dual ball is weak** *sequentially compact (that is, every sequence in* $B(X^*)$ *possesses a weak** *convergent subsequence) and let* $\{A_n : n \in \mathbb{N}\}$ *be a decreasing sequence of bounded non-empty subsets of* X^*. *Then* $\bigcap\{\overline{co}^{w^*} A_n : n \in \mathbb{N}\} = \overline{co}^{w^*}\{a \in X^* : a = weak^*\text{-}\lim_{n\to\infty} a_n$ *and* $a_n \in A_n\}$.

Proof. The right-hand set, D, is clearly a weak* compact, convex and non-empty subset of the left-hand set, C. If it is strictly smaller then there exists some $c \in C$ and $x \in X$ with $\hat{x}(c) > s > \max\{\hat{x}(d) : d \in D\}$. As $C = \bigcap\{\overline{co}^{w^*} A_n : n \in \mathbb{N}\}$ we can find $a_n \in A_n$ with $\hat{x}(a_n) > s$. Since $B(X^*)$ is weak* sequentially compact there exists some subsequence $\{a_{n_k} : k \in \mathbb{N}\}$ which is weak* convergent to some point a with $\hat{x}(a) \geq s > \max\{\hat{x}(d) : d \in D\}$. As the sets are nested, a lies in D, which is impossible. Therefore $C = D$. ■

It is well known that class($\mathcal{S}$) Banach spaces possess weak* sequentially compact dual balls, (see [27] or [18], p.203). We may now give a sequential characterization of D-representability.

Theorem 3.10. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a class($\mathcal{S}$) Banach space X. Let $D \equiv \{x \in A : \nabla f(x)$ exists at $x\}$. Then f is D-representable if, and only if, for each dense subset D^* of D.*

$$\partial f(x) = \overline{co}^{w^*}\{g \in X^* : g = weak^* - \lim_{x_n \to x} \nabla f(x_n),\ \text{and } x_n \in D^*\}$$

We complete this section by using the results so far obtained, to determine some properties of locally Lipschitz functions whose Clarke subdifferential mappings are minimal. To do this, we need to recall that a real-valued function f defined on a non-empty open subset A of a normed linear space X is strictly differentiable on A if, and only if, $\nabla f(x)$ exists for each $x \in A$ and the mapping $x \to \nabla f(x)$ is continuous on A, with respect to the weak* topology on X^*, (see, [12], p.32).

Theorem 3.11. *Let f and g be real-valued locally Lipschitz functions defined on a non-empty open subset A of a Banach space X. If $x \to \partial f(x)$ is a minimal weak* cusco and g is strictly differentiable on A then;*

(i) $x \to \partial(f+g)(x)$ is minimal on A and $\partial(f+g) = \partial f + \partial g$;

(ii) $x \to \partial(f \cdot g)(x)$ is minimal on A and $\partial(f \cdot g) = f \cdot \partial g + \partial f \cdot g$;

(iii) If $h : \mathbb{R} \to \mathbb{R}$ is a strictly differentiable locally Lipschitz function defined on $\mathbb{R}$ then $x \to \partial(h \circ f)(x)$ is minimal on A and $\partial(h \circ f) = (\nabla h \circ f) \cdot \partial f$.

Proof. (**i**) By Proposition 2.3.3 in [12], we have that $\partial(f+g)(x) \subseteq \partial f(x) + \partial g(x)$ for each $x \in A$. Moreover, since g is strictly differentiable on A, $\partial g(x) = \{\nabla g(x)\}$ for each $x \in A$ and so the mapping $x \to \nabla g(x)$ from A into $(X^*, weak^*)$ is continuous. Hence, from Theorem 3.5 part(ii), the mapping $x \to \partial f(x) + \{\nabla g(x)\}$ is a minimal weak* cusco. On the other hand, the mapping $x \to \partial(f+g)(x)$ is a weak* cusco on A and $\partial(f+g)(x) \subseteq \partial f(x) + \{\nabla g(x)\}$ for each $x \in A$. Therefore, $\partial(f+g)(x) = \partial f(x) + \{\nabla g(x)\}$ for each $x \in A$ and $x \to \partial(f+g)(x)$ is a minimal weak* cusco on A. (**ii**) By Proposition 2.3.13 in [12] we have that $\partial(f \cdot g)(x) \subseteq f(x)\partial g(x) + g(x)\partial f(x)$ for each $x \in A$. As in part(i), the mapping $x \to \{\nabla g(x)\}$ is continuous on A and so the mapping $x \to f(x)\partial g(x)$ is continuous on A. Further to this, we have from Theorem 3.5 part(i), that the mapping $x \to g(x)\partial f(x)$ is a minimal weak* cusco on A. Therefore, we may deduce from Theorem 3.5 part(ii) that the mapping $x \to f(x)\partial g(x) + g(x)\partial f(x)$ is a minimal weak*cusco. However, as $\partial(f \cdot g)(x) \subseteq f(x)\partial g(x) + g(x)\partial f(x)$ for each $x \in A$, it follows that $\partial(f \cdot g)(x) = f(x)\partial g(x) + g(x)\partial f(x)$ for all $x \in A$ and it also follows that $x \to \partial(f \cdot g)(x)$ is a minimal weak* cusco on A. (**iii**) Theorem 2.3.9 part(ii) of [12] says that

$\partial(h \circ f)(x) = \nabla h(f(x))\partial f(x)$ for each $x \in A$. Now the mapping $x \to \nabla h(f(x))$ is continuous on A, therefore with the aid of Theorem 3.5 part(i), we see that $x \to \nabla h(f(x))\partial f(x)$ is a minimal weak* cusco on A. From this we may deduce that $\partial(h \circ f)(x) = \nabla h(f(x))\partial f(x)$ for each $x \in A$ and so also deduce that the subdifferential mapping $x \to \partial(h \circ f)(x)$ is a minimal weak* cusco. ■

Note that equality in (i) and (ii) is usually deduced from regularity. Therefore, the new information contained in (i) and (ii) is that the composite function is minimal. In order to establish some further information about minimal subdifferential mappings, we will need to examine more closely the differential structure of the underlying functions.

4. A characterization of minimal subdifferential mappings

We begin this section by characterizing the minimality of the Clarke subdifferential mapping in terms of a continuity property possessed by the upper Dini directional derivative. We will then use this characterization in conjunction with the results from Section Three to establish some further properties enjoyed by locally Lipschitz functions whose subdifferential mappings are minimal. To accomplish this task, we use a special case of a famous theorem due to Lebesgue.

Theorem 4.1. *(Lebesgue's Differentiation Theorem) Let f be a locally Lipschitz function defined on a non-empty open subset of A of the real line, containing the interval $[a, b]$. Then f is differentiable almost everywhere in A and*

$$\int_{[a,b]} f' d\mu = f(b) - f(a)$$

From this theorem we may easily deduce the following corollary.

Corollary 4.2. *(ε Mean-value Theorem) Let f be a locally Lipschitz function defined on a non-empty open subset of the real line, containing the interval $[a, b]$. Then for each $\varepsilon > 0$ there exists a subset M_ε of positive measure in $[a, b]$ such that for each $t \in M_\varepsilon$, $f'(t)$ exists and*

$$f'(t) > \frac{f(b) - f(a)}{b - a} - \varepsilon$$

Using this corollary we may obtain a well-known characterization of the Clarke generalized directional derivative.

Proposition 4.3. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a Banach space X. Then for each $x \in A$ and each $y \in X$, $f^0(x; y) = \limsup_{z \to x} f^+(z; y)$.*

In order to expedite the rest of this section we will introduce the following definition. Let A be a non-empty subset of a Banach space X. Then a subset S of A is 1-D *almost everywhere* in A, *in the direction* y, if for each $x \in A$

$$\mu(\{t \in \mathbb{R} : x + ty \in A \text{ and } x + ty \notin S\}) = 0$$

(here μ represents the Lebesgue measure on $\mathbb{R}$). Note that, in the above definition, it is implicit that for each $x \in A$, $\{t \in \mathbb{R} : x + ty \in A \text{ and } x + ty \notin S\}$ is Lebesgue measurable. For us, the most important example of a 1-D almost everywhere set is the following.

Proposition 4.4. *Let f be a locally Lipschitz function defined on a non-empty open subset A of a Banach space X. Then for each $y \in S(X)$, $D_y \equiv \{x \in A : f'(x;y)$ exists$\}$ is 1-D almost everywhere in A, in the direction y.*

Proof. Let $x_0 \in A$ and let $U \equiv \{t \in \mathbb{R} : x_0 + ty \in A\}$. Define $g : U \to \mathbb{R}$ by $g(t) = f(x_0 + ty)$. Clearly U is non-empty and open, and g is locally Lipschitz on U. Therefore, by Lebesgue's differentiation theorem $\mu(\{t \in U : g'(t)$ does not exists$\}) = 0$. Now, observe that $t' \in \{t \in U : g'(t)$ exists$\}$ if, and only if, $x_0 + t'y \in D_y$. Therefore, $\mu(\{t \in \mathbb{R} : x_0 + ty \in A \text{ and } x_0 + ty \notin D_y\}) = 0$ and so D_y is 1-D almost everywhere in A, in the direction y. ■

1-D almost everywhere subsets are closed under countable intersections.

Proposition 4.5. *Let $\{S_n : n \in \mathbb{N}\}$ be a family of subsets of a non-empty set A of a Banach space X. Let $y \in S(X)$. If each set S_n, is 1-D almost everywhere in A, in the direction y, then so is the set $\bigcap\{S_n : n \in \mathbb{N}\}$.*

Remark 4.1. *Of course, a subset of an open set A, which is 1-D almost everywhere in A, in some direction y, is necessarily dense in A.*

Theorem 4.6. *(*[30]*, Theorems 2.14 and 2.16) Let f be a locally Lipschitz function defined on a non-empty open subset A of a Banach space X. Then $x \to \partial f(x)$ is a minimal weak* cusco on A if, and only if, for each $y \in S(X)$, one of the following conditions holds.*

(i) The mapping $T_y : A \to 2^{\mathbb{R}}$ defined by $T_y(x) = \widehat{y}(\partial f(x))$ is a minimal cusco.

(ii) The function $D_y : A \to \mathbb{R}$ defined by $D_y(x) = f^+(x;y)$ is hyperplane minimal on A.

(iii) The restriction of D_y to a subset P_y, which is 1-D almost everywhere in A in the direction y, is hyperplane minimal on P_y.

By breaking-down the notion of hyperplane minimality, into its two constituent parts, we are able to refine Theorem 4.6. Let f be a real-valued function defined on a topological space A. Then f is *quasi lower semi-continuous (quasi upper semi-continuous)* on A if for each $t_0 \in A$, $\varepsilon > 0$ and open neighbourhood U of t_0 there exists a non-empty open subset V of U such that $\inf\{f(t) : t \in V\} > f(t_0) - \varepsilon$ ($\sup\{f(t) : t \in V\} < f(t_0) + \varepsilon$), [24]. From these definitions, it follows that f is hyperplane minimal on A if, and only if, it is both quasi upper and quasi lower semi-continuous on A. Let us also make the following observations; (i) f is quasi lower semi-continuous on A if, and only if, $-f$ is quasi upper semi-continuous on A; (ii) if D is a dense subset of A and f is quasi lower semi-continuous on A (quasi upper semi-continuous on A) then the restriction of f to D is quasi lower semi-continuous on D (quasi upper semi-continuous on D).

Theorem 4.7. *Let f be a locally Lipschitz function defined on a non-empty open subset A of a Banach space X. Then, $x \to \partial f(x)$, is a minimal weak* cusco on A if, and only if, for each $y \in S(X)$, there exists a subset P_y of A which is 1-D almost everywhere in A, in the direction y, such that the function $D_y : P_y \to \mathbb{R}$ defined by $D_y(x) \equiv f^+(x;y)$ is quasi lower semi-continuous (quasi upper semi-continuous) on P_y.*

Proof. Suppose that the mapping, $x \to \partial f(x)$ is a minimal weak* cusco on A. Fix $y \in S(X)$ and set $P_y \equiv A$. By Theorem 4.6 part(ii) we have that the mapping $x \to f^+(x;y)$ is hyperplane minimal on A and so quasi lower semi-continuous (quasi upper semi-continuous) on P_y. Conversely, suppose that for each $y \in S(X)$ there exists a subset P_y of A which is 1-D almost everywhere in A, in the direction y, such that the mapping $D_y : P_y \to \mathbb{R}$ defined by $D_y(x) \equiv f^+(x;y)$ is quasi lower semi-continuous (quasi upper semi-continuous) on P_y. Fix $y \in S(X)$, we will show that there exists a subset R_y of A, which is 1-D almost everywhere in A, in the direction y, such that the mapping $x \to f^+(x;y)$ is hyperplane minimal on R_y. Let $S_y \equiv \{t \in A : f'(t;y) \text{ exists}\}$, and define $R_y \equiv P_y \cap S_y \cap P_{-y}$. Since P_y, S_y and P_{-y} are 1-D almost everywhere in A, in the direction y, so is R_y. Now, $R_y \subseteq S_y$, therefore $f^+(x;y) = -f'(x;-y) = -f^+(x;-y)$ and so the mapping D_y, restricted to R_y, is both quasi upper and quasi lower semi-continuous on R_y (that is, D_y is hyperplane minimal on R_y), which completes the proof. ■

Theorem 4.8. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a Banach space X. Let $M \equiv \{x \in A : f(x) = \inf\{f(A)\}\}$. Then $x \to \partial f(x)$ is a minimal weak* cusco on A if, and only if, $x \to \partial f(x)$ is a minimal weak* cusco on $A \backslash M$.*

Proof. It follows directly from Remark 3.1 that if $x \to \partial f(x)$ is a minimal weak* cusco on A then $x \to \partial f(x)$ is a minimal weak* cusco on $A\backslash M$. So now we consider the converse. We proceed via the characterization given in Theorem 4.7. To this end, fix $y \in S(X)$ and let $P_y \equiv \{x \in A : f'(x;y) \text{ exists}\}$. By Proposition 4.4, P_y is 1-D almost everywhere in A, in the direction y. We will show that the mapping $D_y : P_y \to \mathbb{R}$ defined by $D_y(x) \equiv f'(x;y) = f^+(x;y)$ is quasi lower semi-continuous on P_y. We may of course, assume that without loss of generality, $M \neq \emptyset$. Consider a point $x_0 \in P_y$. Clearly, if $x_0 \in (\text{int}M \cup A\backslash M) \cap P_y$ then D_y is quasi lower semi-continuous at x_0 (see, Theorem 4.6 part(ii)). So we consider then case when x_0 is in the boundary of M. Let U be a convex open neighbourhood of x_0 contained in A, and let $\varepsilon > 0$. We may assume, by possibly making U smaller, that f is Lipschitz on U with Lipschitz constant K. Choose $0 < t_0 < 1$ such that $x_0 + t_0 y \in U$, and choose $0 < r < \varepsilon t_0/K$ such that $B(x_0 + t_0 y, r) \subseteq U$. Now since $x_0 \in M \cap P_y$, $D_y(x_0) = 0$. Next, we show that there exists a non-empty open subset $V \subseteq B(x_0 + t_0 y, r) \subseteq U$ such that $D_y(z) > -\varepsilon$ for each $z \in V \cap P_y$. Clearly, if $B(x_0 + t_0 y, r) \cap \text{int}M \neq \emptyset$ then we are done (choose $V \equiv B(x_0 + t_0 y, r) \cap \text{int}M$). In the other case, choose

$x_0 + y' \in B(x_0 + t_0 y, r)\backslash M$. Let $s \equiv \max\{t \in [0,1) : x_0 + ty' \in M\}$. Then,

$$\frac{f(x_0 + y') - f(x_0 + sy')}{1 - s} > 0$$

Hence, by the ε Mean-value Theorem there exists a number $s_0 \in (s, 1)$ such that $f'(x_0 + s_0 y'; y') > 0$. Moreover, since $s_0 > s$, $x_0 + s_0 y' \notin M$. Therefore, by the minimality of $x \to \partial f(x)$ on $A\backslash M$, there exists a non-empty open subset $V \subseteq U\backslash M$ such that $f^+(z; y') > 0$ for each $z \in V$, and by positive homogeneity, $f^+(z; t_0^{-1} y') > 0$ for each $z \in V$. However, by our choice of y', $||t_0 y - y'|| < r < \varepsilon t_0/K$ and so, $D_y(z) = f^+(z; y) = f^+(z; t_0^{-1} y') + (f^+(z; y) - f^+(z; t_0^{-1} y')) \geq f^+(z; t_0^{-1} y') - \varepsilon > -\varepsilon$ for each $z \in V \cap P_y$. This ends the proof. ■

Corollary 4.9. *Let f and g be real-valued locally Lipschitz functions defined on a non-empty open subset A of a Banach space X. If $x \to \partial f(x)$ is a minimal weak* cusco and g is strictly differentiable on A then:*

(i) f^+ and f^- possess minimal subdifferential mappings, (here $f^+(x) \equiv max\{f(x), 0\}$ and $f^-(x) \equiv min\{f(x), 0\}$);.

(ii) $|f|$ possesses a minimal subdifferential mapping;

(iii) $x \to max\{f(x), g(x)\}$ and $x \to min\{f(x), g(x)\}$ possess minimal subdifferential mappings.

Proof. (i) The proof that f^+ and f^- possess minimal subdifferential mappings follows directly from Theorem 4.8 and Remark 3.1. (ii) Similarly, the proof that $|f|$ possesses a minimal subdifferential mapping also follows directly from Theorem 4.8 and Remark 3.1. (iii) Observe that $\max\{f(x), g(x)\} = (f - g)^+(x) + g(x)$ and that $\min\{f(x), g(x)\} = (f - g)^-(x) + g(x)$. Now by Theorem 3.11 part(i) $(f - g)$ possesses a minimal subdifferential mapping and so, by part(i) above, $(f - g)^+$ and $(f - g)^-$ both possess minimal subdifferential mappings. The proof is completed by again appealing to Theorem 3.11 part(i). ■

Remark 4.2. *It is interesting to compare Corollary 4.9 and Theorem 3.11 to Example 2.1.*

By far and away the most important application of Theorem 4.8 is to distance functions. Let C be a non-empty closed subset of a Banach space $(X, ||\cdot||)$. The *distance function* associated with the set C and the norm $||\cdot||$, (denoted by d_C), is defined by, $d_C(x) \equiv \inf\{||x - c|| : c \in C\}$. We may now obtain a notable fact concerning the minimality of the Clarke subdifferential mapping of a distance function.

Theorem 4.10. *Let C be a non-empty closed subset of a Banach space X. Then d_C possesses a minimal subdifferential mapping on X if, and only if, $x \to \partial d_C(x)$ is a minimal weak* cusco on $X\backslash C$.*

Remark 4.3. *In Section 8 of this paper, we examine some important consequences of the above Theorem.*

5. Essentially strictly differentiable locally Lipschitz functions

In this section of the paper we will define a class of locally Lipschitz functions whose subdifferential mappings are both minimal and integrable. This class of functions contains the sub-regular and semi-smooth functions considered in [13] and [29]. In this way, we are able to generalize, in a unified manner, the various results contained in [13], [30], [6], [14], [22], [36], [33] and [2] as well as, [37] (at least in the case of Lipschitz functions).

We say that a subset N of a separable Banach space X is *universally measurable* if it belongs to the m-completion of the Borel subsets, $\mathcal{B}(X)$, for each finite measure m on $\mathcal{B}(X)$. A subset N of X is called a *Haar-null* set if it is universally measurable and there exists a probability measure P on $\mathcal{B}(X)$, (which extends canonically to the universally measurable sets on X) such that $P(x+N)=0$ for all $x \in X$.

The Haar-null sets are closed under translation and countable unions, [9]. It follows therefore, that if N is a Haar-null set then $X\backslash N$ is dense in X. In finite dimensions, the Haar-nulls sets coincide with the Lebesgue null sets. Also, we will say that a property P holds *almost everywhere* in A if $\{t \in A : P(x) \text{ is not true}\}$ is a Haar-null set.

Not surprisingly, Haar-null sets are related to 1-D almost everywhere subsets.

Proposition 5.1. *Let S be a universally measurable subset of a non-empty open subset A of a separable Banach space X. If for some $y \in S(X)$, S is 1-D almost everywhere in A, in the direction y, then $A\backslash S$ is a Haar-null set.*

Proof. Suppose that S is 1-D almost everywhere in A, in the direction y. We define a probability measure on P on $\mathcal{B}(X)$ by,

$$P(M) \equiv \frac{1}{\sqrt{2\pi}} \int_{M^*} \exp(-x^2/2)d\mu \quad \text{where } M^* \equiv \{t \in \mathbb{R} : ty \in M\}$$

We claim that $P(x + A\backslash S) = 0$ for each $x \in X$. To prove this, let us consider $x_0 \in X$. Let $U \equiv \{t \in \mathbb{R} : ty \in x_0 + A\}$. If $U = \emptyset$ then clearly $0 \le P(x_0 + A\backslash S) \le P(x_0 + A) \le 0$. So let us consider when $U \neq \emptyset$. In this case, we choose $t_0 \in U$. Then,

$$\mu(\{t \in \mathbb{R} : ty \in x_0 + A\backslash S\}) = \mu(\{t \in \mathbb{R} : (-x_0 + t_0 y) + ty \in A\backslash S\}) = 0$$

since $-x_0 + t_0 y \in A$ and S is 1-D almost everywhere in A, in the direction y. Therefore, $P(x_0 + A\backslash S) = 0$. ■

Using the results in [40], we may, in the case of a separable Banach space, refine Theorem 4.6 even further.

Proposition 5.2. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a separable Banach space X. Then $x \to \partial f(x)$ is a minimal weak* cusco on A if, and only if, for each $y \in S(X)$ there exists a Haar-null set N_y such that the mapping $D_y : A\backslash N_y \to \mathbb{R}$ defined by $D_y(x) \equiv f^+(x; y)$ is hyperplane minimal on $A\backslash N_y$.*

Proof. Suppose that the mapping $x \to \partial f(x)$ is a minimal weak* cusco on A. Fix $y \in S(X)$ and set $N_y = \emptyset$. By Theorem 4.6 part(ii) we have that the mapping $x \to f^+(x;y)$ is hyperplane minimal on $A \backslash N_y$ Conversely, suppose that for each $y \in S(X)$ there exists a Haar-null set N_y such that the mapping $D_y : A \backslash N_y \to \mathbb{R}$ defined by $D_y(x) \equiv f^+(x;y)$ is hyperplane minimal on $A \backslash N_y$. By [40], we have that $CSC(D_y)(x) = \hat{y}(\partial f(x))$ for each $x \in A$. The result now follows from Theorem 4.6 part(i) and Theorem 3.7 part(ii). ■

Proposition 5.3. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a separable Banach space X. Let $D \equiv \{x \in A : \nabla f(x)$ exists$\}$ and let N be any Haar-null subset of X. If $x \to \nabla f(x)$ is weak* hyperplane minimal on $D \backslash N$, then $x \to \partial f(x)$ is a minimal weak* cusco on A.*

Proof. By [40] we have that $\partial f = CSC(\Omega_{D\backslash N})$ where, $\Omega_{D\backslash N} : D \backslash N \to X^*$ is defined by, $\Omega_{D\backslash N}(x) \equiv \nabla f(x)$. The result now follows from Theorem 3.7 part(ii). ■

The significance of the previous two results is that they entitle us to neglect certain 'small' subsets when determining the global minimality of the Clarke subdifferential mapping. Next, we shall consider an important sub-class of the D-representable locally Lipschitz functions. Let A be a non empty open subset of a separable Banach space X. Following ([2], p.68) we will say that a real-valued locally Lipschitz function f defined on a A is *essentially strictly differentiable* on A, if f is strictly differentiable everywhere on A except possibly on a Haar-null set. We will denote by $S_e(A)$ the family of all real-valued essentially strictly differentiable locally Lipschitz functions defined on A. Let us also note, that this class of functions has also been considered in [36], at least in the case, when X is finite dimensional. Our first two tasks are to show that, each member of $S_e(A)$ possesses a minimal subdifferential mapping and to show that $S_e(A)$ contains a significant class of functions.

Proposition 5.4. *([2], Corollary 3.10) Let A be a non-empty open subset of a separable Banach space X. Then each member of $S_e(A)$ possesses a minimal subdifferential mapping.*

Proof. Suppose that $f \in S_e(A)$. Let $S \equiv \{x \in A : f$ is strictly differentiable at $x\}$. It is easy to see that, $x \to \nabla f(x)$ is weak* continuous on S, and so weak* hyperplane minimal on S. The result now follows directly from Proposition 5.3. ■

Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a Banach space X. Then f is *upper semi-smooth* (*lower semi-smooth*) at a point $x \in A$, *in the direction* y if,

$$f^+(x;y) \geq \limsup_{\substack{t \to 0^+ \\ y' \to y}} f^+(x+ty';y) \left(f^-(x;y) \leq \liminf_{\substack{t \to 0^+ \\ y' \to y}} f^-(x+ty';y) \right)$$

Moreover, we say that f is *semi-smooth* at a point $x \in A$, in the direction y if,

$$\limsup_{\substack{t\to 0^+\\ y'\to y}} f^+(x+ty';y) = f^+(x;y) = f^-(x;y) = \liminf_{\substack{t\to 0^+\\ y'\to y}} f^-(x+ty';y)$$

If X is a separable Banach space then we say that f is *essentially upper semi-smooth* (*essentially lower semi-smooth*) on A, if for each $y \in S(X)$ there exists a subset S_y of A such that $A\backslash S_y$ is a Haar-null set and for each $x \in S_y$, f is upper semi-smooth (lower semi-smooth) at x in the direction y. Note that all pseudo-regular and all semi-smooth functions are upper semi-smooth. In order to accomplish our second task we need the following technical result concerning real-valued measurable functions defined on $\mathbb{R}$. A real-valued function $g : (a,b) \to \mathbb{R}$ is *approximately continuous* at a point $x \in (a,b)$ if for each $\varepsilon > 0$,

$$\lim_{\delta\to 0^+} \frac{\mu(\{t \in [x-\delta, x+\delta] : |g(t)-g(x)| > \varepsilon\})}{2\delta} = 0$$

The following is stated in ([31], Theorem 35.3).

Theorem 5.5. *Let g be a real-valued Lebesgue measurable function defined on (a,b), then g is approximately continuous almost everywhere in (a,b).*

Lemma 5.6. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a separable Banach space X. If for some $y \in S(X)$, f is upper semi-smooth (lower semi-smooth) in the direction y, almost everywhere in A, then f is pseudo-regular in the direction y, almost everywhere in A.*

Proof. Suppose that f is upper semi-smooth in the direction y, almost everywhere in A (the proof for case when f is lower semi-smooth in the direction y is obtained by considering $-f$). For each $x \in A$ define the open subset A_x of $\mathbb{R}$ by $A_x \equiv \{t \in \mathbb{R} : x+ty \in A\}$. Clearly, A_x is non-empty and open. Define $g_x : A_x \to \mathbb{R}$ by $g_x(s) \equiv f^0(x+sy;y)$ and let $E_y \equiv \{x \in A : g_x$ is approximately continuous at $s=0\}$. We claim that E_y is $1-D$ almost everywhere in A, in the direction y. To see this, consider $x_0 \in A$. Observe that g_{x_0} is upper semi-continuous on A_{x_0} and so Lebesgue measurable on A_{x_0}. Therefore by Theorem 5.5 g_{x_0} is approximately continuous almost everywhere in A_{x_0}. Next, let us notice that g_{x_0} is approximately continuous at $t \in A_{x_0}$ if, and only if, g_{x_0+ty} is approximately continuous at $s=0$. Hence $\mu(\{t \in A_{x_0} : x_0+ty \in A$ and $x_0+ty \notin E_y\}) = 0$ and so E_y is indeed $1-D$ almost everywhere in A, in the direction y. Let S_y denote the subset of A on which f is upper semi-smooth in the direction y. Let us show that $f^0(x;y) = f^+(x;y)$ for each $x \in S_y \cap E_y$. To this end, let $x \in S_y \cap E_y$ and suppose $\varepsilon > 0$ is given. Since f is upper semi-smooth at x in the direction y there exists a $\delta > 0$ such that $f^+(x+tw;y) \le f^+(x;y)+\varepsilon/2$ whenever $0<t<\delta$ and $\|w-y\| < \delta$. On the other hand since $x \in E_y$ there exists a $0 < \delta_1 < \delta$ such that $f^0(x;y) \le f^0(x+\delta_1 y;y)+\varepsilon/2$. Let $V \equiv \{z \in A : z = x+tw, 0<t<\delta$ and $\|w-y\|<\delta\}$. Clearly V is open in A and $x+\delta_1 y \in V$. From this it follows that $f^0(x+\delta_1 y;y) \le f^+(x;y)+\varepsilon/2$ and so $f^0(x;y) \le f^+(x;y)+\varepsilon$.

However, since ε was arbitrary we have that $f^+(x;y) = f^0(x;y)$. Let $P_y = \{x \in A : f^+(x;y) = f^0(x;y)\}$. We will show that $A\backslash P_y$ is a Haar-null set (Note, if we knew that E_y was universally measurable then the result would follow immediately, since $S_y \cap E_y \subseteq P_y$). Clearly, P_y is a Borel set. Let H be a closed hyperplane in X such that $y \notin H$. Now consider the isomorphism $T : H \times \mathbb{R} \to X$ defined by, $T(h, r) \equiv h + ry$. By the remark after Theorem 6 in [9] (see, Remark 5.1 below) we see that for almost all $h \in H$, $\mu(\{r \in \mathbb{R} : T(h, r) \in A \backslash S_y\}) = 0$, and since $A \backslash E_y$ is $1 - D$ almost everywhere in A, in the direction y, we have that for almost all $h \in H$,

$$0 \leq \mu(\{r \in \mathbb{R} : T(h, r) \in A \backslash P_y\}) \leq \mu(\{r \in \mathbb{R} : T(h, r) \in A \backslash E_y \cup A \backslash S_y\}) \leq 0$$

The result now follows by again appealing to the remark after Theorem 6 in [9] (see, Remark 5.1 below). ■

Remark 5.1. *The remark in* [9] *just after Theorem 6 says: If H is an arbitrary Abelian Polish group (recall that a topological space, is said to be Polish, if it is homeomorphic to a separable, complete metric space) and T is a locally compact Abelian Polish group (for example, $(\mathbb{R}, +)$) then it can be shown that for any universally measurable set $A \subseteq H \times T$, the following are equivalent. (i) The section $A(h) \equiv \{t \in T : (h, t) \in A\}$ is a Haar-null set, for the Haar measure on T, for almost every $h \in H$. (ii) The set A is a Haar-null set in the product group $H \times T$.*

Proposition 5.7. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a separable Banach space X. Let $\{y_n : n \in \mathbb{N}\}$ be a dense subset of $S(X)$. If for each $n \in \mathbb{N}$, f is pseudo-regular in the direction y_n, almost everywhere in A, then $f \in S_e(A)$.*

Proof. For each $n \in \mathbb{N}$, let P_n be a subset of A such that $A \backslash P_n$ is a Haar-null set and f is pseudo-regular in the direction y_n, at each $x \in P_n$. Let $D \equiv \{x \in A : \nabla f(x)$ exists$\}$. By Theorem 7.5 in [10], $A \backslash D$ is a Haar-null set. Now, let $S \equiv \bigcap\{P_n : n \in \mathbb{N}\} \cap D$. We claim that f is strictly differentiable at each point of S. To see this, consider $x_0 \in S$. Then $f^0(x_0; y_n) = \nabla f(x_0)(y_n)$ for each $n \in \mathbb{N}$. However, since both mappings $y \to f^0(x_0; y)$ and $y \to \nabla f(x_0)(y)$ are continuous on X we must have that $f^0(x_0; y) = \nabla f(x_0)(y)$ for each $y \in S(X)$. This shows that f is strictly differentiable at x_0. ■

We may now establish a fundamental (and surprising) fact.

Corollary 5.8. *If f is a real-valued essentially upper semi-smooth (essentially lower semi-smooth) locally Lipschitz function defined on a non-empty open subset A of a separable Banach space X, then $f \in S_e(A)$.*

Proof. The proof follows from Lemma 5.6 and Proposition 5.7. ■

Next, we show that each member of $S_e(A)$ is integrable.

Proposition 5.9. *(*[2]*, Proposition 4.4) Suppose that A is a non-empty open subset of a Banach space X . Let f and g be real-valued locally Lipschitz functions defined on A such that $\partial g(x) \subseteq \partial f(x)$ for each $x \in A$. Then $\partial(f - g) \equiv 0$ if, and only if, $x \to \partial(f - g)(x)$ is a minimal weak*cusco on A.*

Proof. Since $f = (f - g) + g$ we have that $\partial f(x) \subseteq \partial(f - g)(x) + \partial g(x) \subseteq \partial(f - g)(x) + \partial f(x)$ and "cancellation" of closed, bounded convex sets, implies that $0 \in \partial(f - g)(x)$ for all $x \in A$. So if $x \to \partial(f - g)(x)$ is a minimal weak* cusco on A then $\partial(f - g) \equiv 0$. The converse is obvious. ■

Corollary 5.10. *Suppose that A is a non-empty open connected subset of a Banach space X. Let f and g be real-valued locally Lipschitz functions defined on A such that $\partial g(x) \subseteq \partial f(x)$ for each $x \in A$. Then $f - g$ is a constant function on A if, and only if, $x \to \partial(f - g)(x)$ is a minimal weak*cusco on A.*

Proof. Suppose that $x \to \partial(f - g)(x)$ is a minimal weak*cusco on A. Let $x_0 \in A$ and let $A_0 \equiv \{x \in A : (f - g)(x) = (f - g)(x_0)\}$. It follows from the mean-value theorem and the result above, that both A_0 and $A \backslash A_0$ are open subsets of A. Now, $A_0 \neq \emptyset$, since $x_0 \in A_0$. Therefore $A_0 = A$; which shows that $f - g$ is a constant function on A. The converse, is again obvious. ■

Remark 5.2. *We saw earlier, that in order to investigate D-representability we were forced to consider locally Lipschitz functions whose Clarke subdifferential mappings are minimal. Here again, we see that in order to consider locally Lipschitz functions which are integrable, we are compelled to examine locally Lipschitz functions whose Clarke subdifferential mappings are minimal. Although, as we shall see in Section 9, integrability by itself is not sufficient to ensure minimality of the Clarke subdifferential mapping, except on $\mathbb{R}$.*

Proposition 5.11. *([2], Proposition 4.4) Let A be a non-empty open subset of a separable Banach space X. Then each member of $S_e(A)$ is integrable.*

Proof. Suppose that $f \in S_e(A)$ and g is a real-valued locally Lipschitz function defined on A such that $\partial g(x) \subseteq \partial f(x)$ for all $x \in A$. Clearly then $g \in S_e(A)$. Moreover, since $\partial(f - g)(x) \subseteq \partial f(x) - \partial g(x)$ we see that $f - g \in S_e(A)$, and so the subdifferential $x \to \partial(f - g)(x)$ is a minimal weak*cusco on A.

Hence, by the previous Proposition, $\partial(f - g) \equiv 0$ on A . ■

Theorem 5.12. *(Identity Theorem part(i)) Suppose that f and g are real-valued locally Lipschitz functions defined on a non-empty open connected subset A of a separable Banach space X. If $f \in S_e(A)$, g possesses a minimal subdifferential mapping and $\partial g(x) \cap \partial f(x) \neq \emptyset$ for each x in a dense subset of A, then $f - g$ is a constant function on A.*

Proof. Consider the set-valued mapping $T : A \to 2^{X^*}$ defined by, $T(x) \equiv \partial g(x) \cap \partial f(x)$. Since both $x \to \partial f(x)$ and $x \to \partial g(x)$ are upper semi-continuous on A, T possesses non-empty weak* compact, convex images. Moreover, since the graphs of both ∂f and ∂g are closed in $A \times X^*$, with X^* equipped with the weak* topology, so is the graph of T. Therefore, by Proposition 1.3, T is a cusco on A. But, for each $x \in A$, $T(x) \subseteq \partial f(x)$ and $T(x) \subseteq \partial g(x)$. Hence, by the minimality of ∂f and ∂g we must have that $\partial g = T = \partial f$. The result now follows from Corollary 5.10. ■

Let us now establish some stability properties for $S_e(A)$.

Lemma 5.13. *Let $g_1, g_2, \ldots g_n$ be real-valued locally Lipschitz functions defined on a non-empty open subset A of a Banach space X and let k be a real-valued locally Lipschitz function defined on $\mathbb{R}^n$. Consider the locally Lipschitz function $g : A \to \mathbb{R}^n$ defined by $g(x) \equiv (g_1(x), g_2(x), \ldots g_n(x))$. If for some $y \in X$, $g'(x;y) \equiv (g_1'(x;y), g_2'(x;y), \ldots g_n'(x;y))$ and $(k \circ g)'(x;y)$ all exist at some point $x \in A$, then $(k \circ g)'(x;y) = k'(g(x); g'(x;y))$.*

Proof. We compute,

$$\begin{aligned} k(g(x) + sg(x;y)) - k(g(x)) &= k(g(x+sy)) - k(g(x)) \\ &\quad + k(g(x) + sg'(x;y)) - k(g(x+sy)) \\ &= k(g(x+sy)) - k(g(x)) + o(s) \end{aligned}$$

Therefore, $k'(g(x); g'(x;y)) = (k \circ g)'(x;y)$. ■

One of the most significant consequences of Corollary 5.8 is that it enables us derive some very strong closure properties for the family of functions $S_e(A)$.

Proposition 5.14. *Let $g_1, g_2, \ldots g_n$ be real-valued locally Lipschitz functions defined on a non-empty open subset A of a Banach space X and let k be a locally Lipschitz function defined on $\mathbb{R}^n$. Suppose that for some $x \in A$ and some $y \in X$*

$$g_1^0(x;y) = -g_1^0(x;-y), \; g_2^0(x;y) = -g_2^0(x;-y), \; \ldots . g_n^0(x;y) = -g_n^0(x;-y)$$

Suppose also, that k is upper semi-smooth (lower semi-smooth) at $(g_1(x), g_2(x), \ldots g_n(x))$ and $(k \circ g)'(x;y)$ exists. Then $k \circ g$ is upper semi-smooth (lower semi-smooth) at x in the direction y.

Proof. Let D_y be the subset of A where $(k \circ g)'(x;y) = k'(g(x); g'(x;y))$. By the previous Lemma, D_y is $1 - D$ almost everywhere in A, in the direction y. We now use the fact that k is upper semi-smooth to deduce the following:

$$\begin{aligned} (k \circ g)'(x;y) &= k'(g(x); g'(x;y)) \\ &\geq \limsup_{\substack{t \to 0^+ \\ z \to g'(x;y)}} k^+(g(x) + tz; g'(x;y)) \end{aligned}$$

Since each function g_i, $1 \leq i \leq n$, is strictly differentiable at x, in the direction y, we have that,

$$\limsup_{\substack{t \to 0^+ \\ z \to g'(x;y)}} k^+(g(x) + tz; g'(x;y)) \geq \limsup_{\substack{t \to 0^+ \\ y' \to y}} k^+(g(x+ty'); g'(x;y))$$

Hence,

$$\begin{aligned} (k \circ g)'(x;y) &\geq \limsup\{k^+(g(x+ty'); g'(x;y)) : t \to 0^+, y' \to y\} \\ &= \limsup\{k^+(g(x+ty'); z') : t \to 0^+, z \to g'(x;y), z' \to g'(x;y)\} \\ &\geq \limsup\{k^+(g(x+ty'); g^+(x+ty';y)) : t \to 0^+, y' \to y\} \\ &\geq \limsup\{k'(g(x+ty'); g'(x+ty';y)) : t \to 0^+, y' \to y, x+ty' \in D_y\} \\ &= \limsup\{(k \circ g)'(x+ty';y) : t \to 0^+, y' \to y, x+ty' \in D_y\} \end{aligned}$$

Now, since D_y is $1-D$ almost everywhere in A, in the direction y, we have by the ε Mean-value Theorem that,

$$(k \circ g)'(x; y) \geq \limsup_{\substack{t \to 0^+ \\ y' \to y}} (k \circ g)^+(x + ty'; y)$$

This completes the proof. ■

Corollary 5.15. *Let A be a non-empty open subset of a separable Banach space X and suppose that $g_1, g_2, \ldots g_n \in S_e(A)$. If $k : \mathbb{R}^n \to \mathbb{R}$ is upper semi-smooth (lower semi-smooth) on $g(A)$, where $g \equiv (g_1, g_2, \ldots g_n)$, then the locally Lipschitz function $f : A \to \mathbb{R}$ defined by $f \equiv k \circ g$ is a member of $S_e(A)$.*

Proof. By Corollary 5.8 it is sufficient to show that f is essentially upper semi-smooth (essentially lower semi-smooth) on A. Fix $y \in S(X)$. Let S be a subset of A such that $A \backslash S$ is a Haar-null set and each function g_j, $1 \leq j \leq n$ is strictly differentiable at each point of S. Let $D_y \equiv \{x \in A : (k \circ g)'(x; y) \text{ exists}\}$. It is not difficult to show that D_y is a Borel subset of A. Moreover, by Proposition 4.4 we see that D_y is 1-D almost everywhere in A, in the direction y. Therefore $A \backslash (S \cap D_y)$ is a Haar-null set. It now only remains to observe, that from Lemma 5.14, $k \circ g$ is upper semi-smooth (lower semi-smooth) in the direction y, at each point of $S \cap D_y$. ■

Corollary 5.16. *Let A be a non-empty open subset of a separable Banach space X, then $S_e(A)$ is closed under addition, subtraction, multiplication and division (when this is defined), as well as, the lattice operations. Moreover, $S_e(A)$ contains all the C^1 functions and all the continuous convex functions.*

Proof. In each case k is semi-smooth on $\mathbb{R}^2$. ■

Theorem 5.17. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a separable Banach space X. Let g be a real-valued locally Lipschitz function defined on an open subset U of $\mathbb{R}$ which contains $f(A)$. If $f \in S_e(A)$ and $g \in S_e(U)$ then $g \circ f \in S_e(A)$.*

Proof. Let $y \in S(X)$. We will show that $f \circ g$ is pseudo-regular in the direction y, almost everywhere in A. Let $C \equiv \{t \in U : s \to g^+(s; 1) \text{ is continuous at } t\}$. Then C is a G_δ subset of U (as the points of continuity of any real-valued function always form a G_δ set) and $\mu(U \backslash C) = 0$. Next, consider the following two sets. $D_y \equiv \{x \in A : (g \circ f)^+(x; y) = g^+(f(x); 1) \cdot f^+(x; y)\}$ and $E_y \equiv \{x \in A : f(x) \in C$ or $f^+(x; y) = 0\}$. It is readily verified that D_y and E_y are Borel subsets of A. We will prove that both D_y and E_y are 1-D almost everywhere in A, in the direction y. We consider first the set D_y. Take $x_0 \in A$. Let $V \equiv \{t \in \mathbb{R} : x_0 + ty \in A\}$ and define $h : V \to \mathbb{R}$ by $h(t) \equiv g(f(x_0 + ty))$. Clearly, V is non-empty and open, and g is locally Lipschitz on V. Therefore, by Theorem 6.93 in [39] (see, Remark 5.3 below) $\mu(\{t \in \mathbb{R} : h^+(t; 1) \neq g^+(f(x_0 + ty); 1) \cdot f^+(x_0 + ty; y)\}) = 0$. Now observe that $t' \in \{t \in \mathbb{R} : h^+(t; 1) \neq g^+(f(x_0 + ty); 1) \cdot f^+(x_0 + ty; y)\}$ if, and only if, $x_0 + t'y \in A \backslash D_y$. Therefore, $\mu(\{t \in \mathbb{R} : x_0 + ty \in A \backslash D_y\}) = 0$ and so D_y is 1-D almost everywhere in A, in the direction y. We now consider the set E_y. Take $x_0 \in A$. Let $V \equiv \{t \in \mathbb{R} : x_0 + ty \in A\}$ and define $h : V \to \mathbb{R}$ by $h(t) \equiv f(x_0 + ty)$. By Lemma 6.92 in [39] (see, Remark 5.3 below) $\mu(\{t \in V : h(t) \notin C \text{ and } h^+(t; 1) = 0\}) = 0$.

Therefore, $\mu(\{t \in \mathbb{R} : x_0 + ty \in A \backslash E_y\}) = 0$ and so, E_y is 1-D almost everywhere in A, in the direction y. Let $F_y \equiv D_y \cap E_y$ and let S be the set of all points in A at which f is strictly differentiable. We claim that $g \circ f$ is pseudo-regular in the direction y at each point of $S \cap F_y$. To see this, consider the mapping $T : F_y \to \mathbb{R}$ defined by $T(x) \equiv (g \circ f)^+(x; y)$. Clearly T is continuous at each point of $S \cap F_y$. Now by the ε Mean-value Theorem, $x \to (g \circ f)^+(x; y)$ is upper semi-continuous at each point of $S \cap F_y$. Hence by Proposition 4.3, $g \circ f$ is pseudo-regular in the direction y at each point of $S \cap F_y$; which shows that $f \circ g$ is pseudo-regular in the direction y, almost everywhere in A. The proof is now completed by appealing to Proposition 5.7. ■

Remark 5.3. *A special case of Theorem 6.93 in* [39] *says, that if f is a locally Lipschitz function defined on an open interval (a, b) and g is a locally Lipschitz function defined on an open subset of $\mathbb{R}$ which contains the range of f then $(g \circ f)^+(x) = g^+(f(x)) \cdot f^+(x)$ almost everywhere in (a, b). Lemma 6.92 in* [39] *says that if $E \subseteq (a, b)$ and f is differentiable at each point of E and the Lebesgue outer-measure of $f(E)$ is 0, then $f'(x) = 0$ almost everywhere in E.*

So we see then, that $S_e(A)$ possesses quite remarkable stability properties. We end our investigation of the properties of $S_e(A)$ with the following proposition, which provides a condition, sufficient to ensure membership in $S_e(A)$.

Proposition 5.18. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a separable Banach space X. Let B be a subset of X such that $\overline{sp}B = X$. If for each $b \in B$, $f^0(x; b) = -f^0(x; -b)$ almost everywhere in A, then $f \in S_e(A)$.*

Proof. We may suppose, by possibly making B smaller, that B is a countable set, that is, $B \equiv \{b_n : n \in \mathbb{N}\}$. For each $n \in \mathbb{N}$, let P_n be a subset of A such that $A \backslash P_n$ is a Haar-null set and f is pseudo-regular in the direction b_n, at each point of P_n. Let $S \equiv \bigcap\{P_n : n \in \mathbb{N}\}$. We claim that f is strictly differentiable at each point of S. To prove this, let us consider a point $x_0 \in S$. To show that f is strictly differentiable at x_0 we merely need to show that $\partial f(x_0)$ is a singleton. To this end, let g_1 and g_2 be Clarke subgradients of f at x_0. Then $g_i(b_n) \leq f^0(x_0; b_n)$ and $g_i(-b_n) \leq f^0(x_0; -b_n)$ for each $i \in \{1, 2\}$ and $n \in \mathbb{N}$. Therefore, $g_i(b_n) \leq f^0(x_0; b_n) = -f^0(x_0; -b_n) \leq -g_i(-b_n) = g_i(b_n)$ for each $i \in \{1, 2\}$ and $n \in \mathbb{N}$. Hence, $g_1(b_n) = f^0(x_0; b_n) = g_2(b_n)$ for each $n \in \mathbb{N}$. Now, both g_1 and g_2 are linear on X and so we have that $g_1(b) = g_2(b)$ for each $b \in spB$. But both g_1 and g_2 are also continuous on X, therefore $g_1 = g_2$ on $\overline{sp}B = X$. ■

6. Essentially strictly differentiable Lipschitz functions on non-separable Banach spaces.

This section of the paper, is devoted to extending the results in Section five to arbitrary Banach spaces. The key to this extension is the following definition. Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a Banach space X. Then $f \in M_e(A)$ if; (i) there exist subsets $\{B_n^f : n \in \mathbb{N}\}$ of X such that $\overline{sp}B_n^f = X$ for each $n \in \mathbb{N}$; (ii) for each closed separable subspace Y

of X, $\{x \in Y \cap A :$ for some $y \in S(Y), f^0(x;y) \neq -f^0(x;-y)\}$ is a Haar-null set, whenever, $\overline{sp}(B_n^f \cap Y) = Y$ for each $n \in \mathbb{N}$.

Proposition 6.1. *Let A be a non-empty open subset of a separable Banach space X. Then $M_e(A) = S_e(A)$.*

Proof. It is obvious that $M_e(A) \subseteq S_e(A)$, so it suffices to show that $S_e(A) \subseteq M_e(A)$. Let $f \in S_e(A)$ and let $B_0 \equiv \{b_n : n \in \mathbb{N}\}$ be any dense subset of X. For each $n \in \mathbb{N}$, let $B_n \equiv b_n + 1/2B(X)$. We claim that f satisfies conditions (i) and (ii) for membership in $M_e(A)$, with respect to the sets $\{B_n : n \in \mathbb{N}\}$. Clearly, $\overline{sp}B_n = X$ for each $n \in \mathbb{N}$, therefore it is sufficient to show that f satisfies condition (ii). The way we proceed is to show that the only subspace Y of X, which satisfies the property that $\overline{sp}(B_n \cap Y) = Y$ for each $n \in \mathbb{N}$, is $Y = X$. Indeed, if $\overline{sp}(B_n \cap Y) = Y$ for each n, then $Y \cap B_n \neq \emptyset$ for each n. Therefore, $\mathrm{dist}(b_n, Y) \leq 1/2$ for each $n \in \mathbb{N}$ and so $X \subseteq Y + B(X)$, but this is impossible unless $Y = X$. ■

Next, we show that each member of $M_e(A)$ possesses a minimal subdifferential mapping.

Theorem 6.2. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a Banach space X. If $f \in M_e(A)$ then f possesses a minimal subdifferential mapping.*

Proof. Let $\{B_n : n \in \mathbb{N}\}$ be the subsets of X given in the definition of $M_e(A)$, that are associated with the function f. We will suppose, for the purpose of obtaining a contradiction, that ∂f is not a minimal weak* cusco on A. That is, we will suppose that there exists a cusco $T : A \to 2^{X^*}$ such that $T(y) \subseteq \partial f(y)$ for each $y \in A$ and $T(x_0) \neq \partial f(x_0)$ at some point $x_0 \in A$. We proceed from here, as in Gregory's Method [18], by inductively constructing, an increasing sequence of closed separable subspaces $\{Y^n : n \in \mathbb{N}\}$ of X, and subsets $\{B_j^n : n > 1\}$ of B_j, (for $j \in \mathbb{N}$). For the first step of the induction, choose a $y_0 \in S(X)$ and an $s \in \mathbb{R}$ such that

$$f^0(x_0; y_0) = \max \widehat{y_0}(\partial f(x_0)) > s > \max \widehat{y_0}(T(x_0))$$

Next, we choose a sequence $\{z_n : n \in \mathbb{N}\}$ in A such that

$$\lim_{n\to\infty} z_n = x_0 \text{ and } \lim_{n\to\infty} \frac{f(z_n + \lambda_n y_0) - f(z_n)}{\lambda_n} = f^0(x_0; y_0)$$

for some sequence of positive real numbers $\{\lambda_n : n \in \mathbb{N}\}$ converging to zero and we set $Y^1 \equiv \overline{sp}\{\{x_0, y_0\}, \{z_n : n \in \mathbb{N}\}\}$. Now suppose the first n subspaces have been constructed. For each $j \in \mathbb{N}$, let B_j^{n+1} be a countable subset of B_j, such that $\overline{sp}B_j^{n+1} \supseteq Y^n$ and let $Y^{n+1} \equiv \overline{sp}\{\bigcup\{B_j^{n+1} : j \in \mathbb{N}\}\}$. Now consider, $Y \equiv \overline{\bigcup(Y^n : n \in \mathbb{N}\}}$. It is easy to see that for each $j \in \mathbb{N}$,

$$\overline{sp}(Y \cap B_j) \supseteq \overline{sp}\{\bigcup\{B_j^n : n > 1\} \supseteq \overline{\bigcup\{\overline{sp}B_j^n : n > 1\}} \supseteq \overline{\bigcup\{Y^n : n \in \mathbb{N}\}} \supseteq Y$$

Let $S \equiv \{x \in A \cap Y : f^0(x;y) = -f^0(x;-y)$ for each $y \in S(Y)\}$. By our hypothesis $(A \cap Y) \backslash S$ is a Haar-null set. Our next task, will be to show that for each $x \in S$

and each $y \in S(Y)$, $(f|_Y)^0(x;y) = f^0(x;y)$. We accomplish this, by considering the following. For each $x \in S$ and $y \in S(Y)$ we have that,

$$(f|_Y)^0(x;y) \geq -(f|_Y)^0(x;-y) \geq -f^0(x;-y) = f^0(x;y) \geq (f|_Y)^0(x;y)$$

and so, $(f|_Y)^0(x;y) = -(f|_Y)^0(x;-y) = f^0(x;y)$ for each $x \in S$ and $y \in S(Y)$. Now, by the argument above, it is clear that for each $x \in S$, $y \to (f|_Y)^0(x;y)$ is linear on Y and so $f|_Y \in S_e(A \cap Y)$. By Proposition 5.3, $x \to \partial(f|_Y)(x)$ is a minimal weak* cusco on $A \cap Y$. Let us also observe, that $(f|_Y)^0(x_0;y_0) = f^0(x_0;y_0)$, since $z_n + \lambda_n y_0$ and z_n are members of Y for each $n \in \mathbb{N}$. The next and final step in our proof will be to construct a weak* cusco $T' : A \cap Y \to 2^{Y^*}$ whose graph lies strictly inside that of $\partial(f|_Y) : A \cap Y \to 2^{Y^*}$. This will of course give us our desired contradiction. Consider the restriction mapping $R : X^* \to Y^*$ defined by $R(g) \equiv g|_Y$. It follows directly from the above discussion, that for each $x \in S$, $R(T(x)) \subseteq R(\partial f(x)) \subseteq R(\partial(f|_Y)(x))$. Furthermore, since both set-valued mappings, $x \to R(T(x))$ and $x \to \partial(f|_Y)(x)$ are weak* cuscos on $A \cap Y$ we see that $R(T(x)) \cap R(\partial(f|_Y)(x)) \neq \emptyset$ for each $x \in A \cap Y$. We define the set-valued mapping $T' : A \cap Y \to 2^{Y^*}$ by $T'(x) \equiv R(T(x)) \cap R(\partial(f|_Y)(x))$. It is standard that T' is a weak* cusco on $A \cap Y$, (see, for example, Theorem 5.12). It now only remains to show that the graph of T' is strictly contained in the graph of $\partial(f|_Y)$. To see this, consider the point $x_0 \in A \cap Y$ then,

$$\begin{aligned} \max \widehat{y_0}(T'(x_0)) \leq \max \widehat{y_0}(T(x_0)) < s \quad &< \quad \max \widehat{y_0}(\partial f(x_0)) = f^0(x_0;y_0) \\ &= \quad (f|_Y)^0(x_0;y_0) = \max \widehat{y_0}(\partial(f|_Y)(x_0)) \end{aligned}$$

We have now obtained a contradiction, and so we may indeed conclude that ∂f is a minimal weak* cusco on A. ■

As with $S_e(A)$, the members of $M_e(A)$ are integrable.

Proposition 6.3. *Let A be a non-empty open subset of a Banach space X. Then each member of $M_e(A)$ is integrable.*

Proof. Suppose that $f \in M_e(A)$ and g is a real-valued locally Lipschitz function defined on A such that $\partial g(x) \subseteq \partial f(x)$ for all $x \in A$. Let $\{B_n : n \in \mathbb{N}\}$ be the subsets of X given in the definition of $M_e(A)$. It follows from Proposition 5.9 and Theorem 6.2 that to show f is integrable on A we need only show that $f - g \in M_e(A)$. Indeed, we will show, using the sets $\{B_n : n \in \mathbb{N}\}$, that the function $f - g$ satisfies the requirements for membership in $M_e(A)$. Let Y be a closed, separable subset of X such that $\overline{sp}(B_n \cap Y) = Y$ for each $n \in \mathbb{N}$. Let $S \equiv \{x \in Y \cap A : f^0(x;y) = -f^0(x;-y)$ for each $y \in Y\}$. It follows from the definition of the sets, $\{B_n : n \in \mathbb{N}\}$, that $(Y \cap A)\backslash S$ is a Haar-null set. Now, since $\partial g(x) \subseteq \partial f(x)$ for each $x \in S$ it follows that

$$-f^0(x;-y) = -g^0(x;-y) = g^0(x;y) = f^0(x;y) \quad \text{for each } y \in Y \text{ and each } x \in S.$$

And now, with a little more work, it can be shown that,

$$(f-g)^0(x;y) = -(f-g)^0(x;-y) \quad \text{for each } y \in Y \text{ and each } x \in S.$$

Hence, it follows that $(f - g) \in M_e(A)$. ■

Theorem 6.4. *(Identity Theorem part(ii)) Suppose that f and g are real-valued locally Lipschitz functions defined on a non-empty open connected subset A of a Banach space X. If $f \in M_e(A)$, g possesses a minimal subdifferential mapping and $\partial g(x) \cap \partial f(x) \neq \emptyset$ for each x in a dense subset of A, then $f - g$ is a constant function on A.*

Proof. The proof is identical to that given in Theorem 5.12. ∎

The following proposition provides a condition, similar to that given in Proposition 5.18, which is sufficient to ensure membership in $M_e(A)$.

Proposition 6.5. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a Banach space X. Let B be a subset of X such that $\overline{sp}B = X$. If for each $b \in B$, $f^0(x;b) = -f^0(x;-b)$ 1-D almost everywhere in A, in the direction b, then $f \in M_e(A)$.*

Proof. For each $n \in \mathbb{N}$, let $B_n = B$. We will show that f and $\{B_n : n \in \mathbb{N}\}$ satisfy both condition (i) and condition (ii) given in the definition of $M_e(A)$. The sets $\{B_n : n \in \mathbb{N}\}$ obviously satisfy condition (i) so it is only necessary to check that f and $\{B_n : n \in \mathbb{N}\}$ satisfy condition (ii). Suppose that Y is a separable subspace of X such that $\overline{sp}(B \cap Y) = Y$. Let $B' \equiv B \cap Y$. It is easy to see that for each $b' \in B'$, $f^0(x;b') = -f^0(x,-b')$ almost everywhere in $A \cap Y$ (see, Proposition 5.1). The result now follows exactly as in Proposition 5.18. More specifically, to show that at a given point x, $f^0(x;y) = -f^0(x;-y)$ for each $y \in Y$, we need only show that if g_1 and g_2 are subgradients of f at x then $g_1|_Y = g_2|_Y$. Because, if for some $y \in Y$, $f^0(x;y) \neq -f^0(x;-y)$ then we could construct, using the Hahn-Banach extension theorem, two distinct subgradients g_1 and g_2 of f at x, such that $g_1(y) = f^0(x;y)$ and $g_2(y) = -f^0(x;-y)$. ∎

Corollary 6.6. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a Banach space X. Let B be a subset of X such that $\overline{sp}B = X$. If for each $b \in B$, there exists a subset S_b of X which is 1-D almost everywhere in A, in the direction b, such that f is either, upper semi-smooth in the directions b and $-b$, at each point of S_b, or lower semi-smooth in the directions b and $-b$, at each point of S_b, then $f \in M_e(A)$.*

Proof. It follows from the previous proposition, that to show $f \in M_e(A)$, we need only show, that for each $b \in B$ there exists a subset P_b of A which is $1-D$ almost everywhere in A, in the direction b, such that $f^0(x;b) = -f^0(x;-b)$ at each point of P_b. To this end fix $b \in B$ and let S_b be a subset of A which is 1-D almost everywhere in A, in the direction b, such that f is upper semi-smooth (lower semi-smooth) in the directions b and $-b$ at each point of S_b. For each $x \in A$ define the open subset A_x of $\mathbb{R}$ by,$A_x \equiv \{t \in \mathbb{R} : x + tb \in A\}$. Clearly, A_x is non-empty and open. Define $g_x : A_x \to \mathbb{R}$ by, $g_x(s) \equiv f^0(x + sb; b)$ and let $E_b \equiv \{x \in A : g_x$ is approximately continuous at $s = 0\}$. As in Lemma 5.6 we see that E_b is 1-D almost everywhere in A, in the direction b. Again, as in Lemma 5.6 we see that $f^0(x;b) = f^+(x;b)$ at each point of $S_b \cap E_b$. Now, by repeating this argument on $-b$ we get a subset E_{-b} of A which is 1-D almost every where in A, in the direction b such that $f^0(x;-b) = f^+(x;-b)$ at each point of $S_b \cap E_{-b}$. Now, let $D_b \equiv \{x \in A : f'(x;b)$ exists$\}$. By Proposition 4.4, D_y is 1-D almost everywhere in A, in the direction b.

Let $P_b \equiv E_b \cap E_{-b} \cap S_b \cap D_b$. Clearly, P_y is 1-D almost everywhere in A, in the direction b. So it remains to show that $f^0(x;b) = -f^0(x;-b)$ at each point of P_b. We do this, by performing the following calculation. For each $x \in P_y$ we have,

$$f^0(x;b) = f^+(x;y) = f'(x;b) = -f'(x;-b) = -f^+(x;-y) = -f^0(x;-b). \quad \blacksquare$$

Remark 6.1. *The above result shows that saddle functions are both integrable and possess minimal subdifferential mappings. Moreover, the partially regular functions (as defined in* [14]*) are also both integrable and possesses minimal subdifferential mappings. To convince ourselves of this, we simply observe that the set* $B \equiv \{(x,0) : x \in X\} \cup \{(0,y) : y \in Y\}$ *has the property that* $spB = X \times Y$. *It then follows from the definition of partial regularity, that any partially regular function defined on* $X \times Y$, *satisfies the hypothesis of Corollary 6.6, with respect to the set* B.

Let us also observe, that if the Clarke subdifferential mapping is submonotone, (as defined in [16]*) then the underlying function is lower semi-smooth, and hence a member of* $M_e(A)$. *So in this sense we are able to recover and extend most of the results in* [16].

We end this section of the paper by showing that $M_e(A)$ enjoys similar closure properties to those enjoyed by $S_e(A)$.

Theorem 6.7. *Let* A *be a non-empty open subset of a Banach space* X *and suppose that* $g_1, g_2, \ldots .g_n \in M_e(A)$. *If* $k : \mathbb{R}^n \to \mathbb{R}$ *is upper semi-smooth (lower semi-smooth) on* $g(A)$, *where* $g \equiv (g_1, g_2, \ldots g_n)$, *then the locally Lipschitz function* $f : A \to \mathbb{R}$ *defined by* $f \equiv k \circ g$ *is a member of* $M_e(A)$.

Proof. For each $j \in \{1, 2, ...n\}$ let $\{B_n^{g_j} : n \in \mathbb{N}\}$ be the subsets of X, given in the definition of $M_e(A)$, which are associated with the functions g_j. Let $\{B_n^f : n \in \mathbb{N}\} \equiv \bigcup\{B_m^{g_j} : m \in \mathbb{N} \text{ and } 1 \leq j \leq n\}$. We claim that f and $\{B_n^f : n \in \mathbb{N}\}$ satisfy that conditions (i) and (ii) given in the definition of $M_e(A)$. Clearly, the sets $\{B_n^f : n \in \mathbb{N}\}$ satisfy condition (i), therefore we will content ourselves with showing that f and$\{B_n^f : n \in \mathbb{N}\}$ satisfy condition (ii). Let Y be a closed, separable subspace of X such that $\overline{sp}(B_n^f \cap Y) = Y$ for each $n \in \mathbb{N}$. We will show first, that for each $y \in Y$, f is upper semi-smooth (lower semi-smooth) in the direction y, almost everywhere in $A \cap Y$. Fix $y \in Y$. Let S be a subset of $A \cap Y$ such that $(A \cap Y)\backslash S$ is a Haar-null set and $g_j(x;z) = -g_j(x;-z)$ for each $x \in S$ and each $z \in Y$. Let $D_y \equiv \{x \in A \cap Y : (k \circ g)'(x;y) \text{ exists}\}$. It is not difficult to show that D_y is a Borel subset of $A \cap Y$. Moreover, by Proposition 4.4 we see that D_y is 1-D almost everywhere in $A \cap Y$, in the direction y. Therefore, $(A \cap Y)\backslash(S \cap D_y)$ is a Haar-null set, and by Proposition 5.14, f is upper semi-smooth (lower semi-smooth) in the direction y at each point of $S \cap D_y$. Next, we show that $P_y \equiv \{x \in A \cap Y : f^0(x;y) \neq -f^0(x;-y)\}$ is a Haar-null set. For each $x \in A \cap Y$ and we define the open subset A_x of $\mathbb{R}$ by, $A_x \equiv \{t \in \mathbb{R} : x + ty \in A \cap Y\}$. Clearly, A_x is non-empty and open. Define $g_x : A_x \to \mathbb{R}$ by, $g_x(s) \equiv f^0(x + sy; y)$ and let $E_y \equiv \{x \in A \cap Y : g_x \text{ is approximately continuous at } s = 0\}$. Now, as in Corollary 6.6 and Lemma 5.6 we have that, E_y is 1-D almost everywhere in $A \cap Y$, in the direction y, and $f^0(x;y) = f^+(x;y)$ for each $x \in E_y \cap S \cap D_y$. Let $\{y_n : n \in \mathbb{N}\}$ be a dense subset of $S(Y)$ and let $P_{y_n} \equiv \{x \in A \cap Y : f^+(x;y_n) = f^0(x;y_n)\}$. Clearly, P_{y_n} is a Borel subset of $A \cap Y$. Moreover, by repeating the argument in Lemma 5.6 we have that $(A \cap Y)\backslash P_{y_n}$ is a Haar-null set. Let $P \equiv \bigcap\{P_{y_n} : n \in \mathbb{N}\}$. It follows from the continuity of the mappings $z \to f^+(x;z)$ and $z \to f^0(x;z)$ that in fact $f^+(x;z) = f^0(x;z)$ for each $z \in Y$ and each $x \in P$. Let $T \equiv \{x \in A \cap Y : \nabla(f|_Y)(x)$ exists$\}$. By Theorem 7.5 in [10], $(A \cap Y)\backslash T$ is a Haar-null set. It now only remains to observe that $f^+(x;z) = -f^0(x;-z)$ for each $x \in T \cap P$ and each $z \in Y$. ■

Corollary 6.8. *Let A be a non-empty open subset of a Banach space X, then $M_e(A)$ is closed under addition, subtraction, multiplication and division (when this is defined), as well as, the lattice operations. Moreover, $M_e(A)$ contains all the C^1 functions and all the continuous saddle functions.*

Theorem 6.9. *Let f be a real-valued locally Lipschitz function defined on a non-empty open subset A of a Banach space X. Let g be a real-valued locally Lipschitz function defined on an open subset U of $\mathbb{R}$ which contains $f(A)$. If $f \in M_e(A)$ and $g \in S_e(U)$ then $g \circ f \in M_e(A)$.*

Proof. Let $\{B_n : n \in \mathbb{N}\}$ be the subsets of X, given in the definition of $M_e(A)$. We claim that $g \circ f$ and $\{B_n : n \in \mathbb{N}\}$ satisfy condition (ii) in the definition of $M_e(A)$. Let Y be a closed, separable subspace of X such that $\overline{sp}(B_n \cap Y) = Y$ for each $n \in \mathbb{N}$. Fix $y \in Y$ and let $C \equiv \{t \in U : s \to g^+(s;1)$ is continuous at $t\}$. Now, consider the two sets $D_y \equiv \{x \in A \cap Y : (g \circ f)^+(x;y) = g^+(f(x);1) \cdot f^+(x;y)\}$ and $E_y \equiv \{x \in A \cap Y : f(x) \in C$ or $f^+(x;y) = 0\}$. Let $F_y \equiv D_y \cap E_y$. As in Theorem 5.17, we have that F_y is 1-D almost everywhere in $A \cap Y$, in the direction y. Let $S \equiv \{x \in A \cap Y : f^0(x;y) = -f^0(x;-y)$ for all $y \in Y\}$. We claim that $g \circ f$ is pseudo-regular in the direction y at each point of $S \cap F_y$. To see this, consider the mapping $T : F_y \to \mathbb{R}$ defined by, $T(x) \equiv (g \circ f)^+(x;y)$. Clearly, T is continuous at each point of $S \cap F_y$. Now, by the ε Mean-value Theorem, $x \to (g \circ f)^+(x;y)$ is upper semi-continuous at each point of $S \cap F_y$. Let $\{y_n : n \in \mathbb{N}\}$ be a dense subset of $S(Y)$ and let $F \equiv \bigcap\{F_{y_n} : n \in \mathbb{N}\} \cap S$. It follows from the continuity of the mappings $y \to f^+(x;y)$ and $f^0(x;y)$ that in fact $f^+(x;y) = f^0(x;y)$ for each $y \in S(Y)$ and each $x \in F$. Let $K \equiv \{x \in A \cap Y : \nabla(f|_Y)(x)$ exists$\}$. By Theorem 7.5 in [10], $(A \cap Y)\backslash K$ is a Haar-null set. The proof is completed by observing that $f^0(x;y) = -f^0(x;-y)$ for each $x \in K \cap F$ and $y \in Y$. ■

Remark 6.2. *Let A be a non-empty open subset of a class($\mathcal{S}$) Banach space X and let $\mathcal{M}(A)$ denote the family of all real-valued locally Lipschitz functions defined on A whose Clarke subdifferential mappings are minimal. Furthermore, let $\mathcal{D}(A)$ denote the family of all real-valued locally Lipschitz functions defined on A which are generically strictly differentiable on A (that is, $\{x \in A : f$ is strictly differentiable at $x\}$ contains a dense and G_δ subset of A). Then, $M_e(A) \subseteq \mathcal{M}(A) \subseteq \mathcal{D}(A)$. We saw in Example 2.1 that $\mathcal{M}(A)$ is not closed under addition, multiplication or either of the lattice operations. However, it is easy to show that $\mathcal{D}(A)$ is closed under addition, subtraction, multiplication and division (when this is defined) as well as, the lattice operations. So if one is only interested generic differentiability properties, then it makes sense to consider $\mathcal{D}(A)$ (which is always strictly larger than $M_e(A)$). However, it is necessarily the case, that not every member of $\mathcal{D}(A)$ is D-representable or integrable.*

7. Perturbation functions

In this section of the paper, we apply the results of Section 6 to *perturbation* functions. Let A be a non-empty open subset of a Banach space X and let T be a topological space. We say that a real-valued function $g : A \times T \to \mathbb{R}$ is *locally Lipschitz on A, uniformly in T* if for each $x_0 \in A$ there exists an $K > 0$ and $\delta > 0$ such that,

$$|g(x,t) - g(y,t)| \leq K\|x - y\| \text{ for all } x, y \in B(x_0, \delta) \text{ and } t \in T.$$

Further, we say that an extended real-valued function f defined on A is a *sup-marginal function* if, $f(x) \equiv \sup\{g(x,t) : t \in T\}$ for some function $g : A \times T \to \mathbb{R}$. If more stringently, we have that $f(x) = \max\{g(x,t) : t \in T\}$ and g is locally Lipschitz on A, uniformly in T, then f is real-valued and locally Lipschitz on A. A set-valued mapping M from a topological space A into non-empty subsets of a topological space T will be said to be *semi-continuous* on A if, for each $x \in A$ and each net $(x_\alpha)_{\alpha \in I}$ in A, converging to x, there exists a point $y \in M(x)$ and elements $y_\alpha \in M(x_\alpha)$ such that y is an accumulation point of the set $\{y_\alpha : \alpha \in I\}$, that is, $y \in \overline{\{y_\alpha : \alpha \in I\} \backslash \{y\}}$. (Note, this definition is less arduous than that given in [13] and [14].) The following theorem unifies Theorems 6.1 and 6.2 in [13] and Proposition 2.6 in [14].

Theorem 7.1. *Let A be a non-empty open subset of a Banach space X and let T be a Hausdorff topological space. Let $g : A \times T \to \mathbb{R}$ be locally Lipschitz on A, uniformly in T and let $f : A \to \mathbb{R}$ be defined by $f(x) \equiv \max\{g(x,t) : t \in T\}$. Furthermore, suppose that (i) the set-valued mapping $M : A \to 2^T$, defined by, $M(x) \equiv \{t \in T : f(x) = g(x,t)\}$ is semi-continuous on A and that (ii) for each $x \in A$ and each $y \in B \cup -B$, $(r, y', t) \in \mathbb{R}^+ \times X \times T \to g^+(x + ry', t; y)$ is upper semi-continuous (as a real-valued function) at each point of $\{0\} \times \{y\} \times M(x)$, (here B is any subset of X such that $\overline{sp}B = X$) then $f \in M_e(A)$.*

Proof. To show that $f \in M_e(A)$, it suffices, by Corollary 6.6 to show that f is upper semi-smooth in the direction y, on A, for each $y \in B \cup -B$. Let x be a fixed element of A and y be a fixed element of $B \cup -B$. We will show that for any sequence of positive real numbers $\{s_n : n \in \mathbb{N}\}$ converging to 0 and any sequence $\{y_n : n \in \mathbb{N}\}$ of elements of X converging to y, we have that $\liminf\{f^+(x + s_n y_n; y) : n \to \infty\} \leq f^+(x; y)$. Indeed, by a standard subsequence argument this will show that f is upper semi-smooth at x, in the direction y. So let $\{s_n : n \in \mathbb{N}\}$ be a sequence of positive real numbers converging to 0 and let $\{y_n : n \in \mathbb{N}\}$ be a sequence of elements of X converging to y. For each $n \in \mathbb{N}$, we may choose $0 < \lambda_n < s_n$ such that,

$$f^+(x + s_n y_n; y) < \frac{f(x + s_n y_n + \lambda_n y) - f(x + s_n y_n)}{\lambda_n} + 1/n$$

Since M is semi-continuous on A and $\lim_{n\to\infty}(x + s_n y_n + \lambda_n y) = x$ there exists a point $t \in M(x)$ and a sequence $\{t_n : n \in \mathbb{N}\}$ in T such that $t_n \in M(x + s_n y_n + \lambda_n y)$

for each $n \in \mathbb{N}$ and $t \in \overline{\{t_n : n \in \mathbb{N}\}\backslash\{t\}}$. Now, for each $n \in \mathbb{N}$, we have that,

$$\frac{f(x + s_n y_n + \lambda_n y) - f(x + s_n y_n)}{\lambda_n} \leq \frac{g(x + s_n y_n + \lambda_n y, t_n) - g(x + s_n y_n, t_n)}{\lambda_n}$$

Furthermore, by the ε Mean-value Theorem we have that for each $n \in \mathbb{N}$ there exists a real number s'_n such that $0 < s'_n < \lambda_n$ and

$$\frac{g(x + s_n y_n + \lambda_n y, t_n) - g(x + s_n y_n, t_n)}{\lambda_n} \leq g^+(x + s_n y_n + s'_n y, t_n; y) + 1/n$$

Therefore, for each $n \in \mathbb{N}$,

$$f^+(x + s_n y_n; y) \leq g^+(x + s_n y_n + s'_n y, t_n; y) + 2/n$$

Now, let $s''_n \equiv (s_n + s'_n)$ and $y'_n \equiv (s_n y_n + s'_n y)/s''_n$. Then clearly, $\lim_{n\to\infty} y'_n = y$ and $\lim_{n\to\infty} s''_n = 0$. Hence,

$$\begin{aligned}
\liminf_{n\to\infty} f^+(x + s_n y_n; y) &\leq \liminf_{n\to\infty} g^+(x + s_n y_n + s'_n y, t_n; y) \\
&= \liminf_{n\to\infty} g^+(x + s''_n y'_n, t_n; y) \\
&\leq g^+(x, t; y) \\
&\leq f^+(x; y) \quad (\text{since } t \in M(x)). \ \blacksquare
\end{aligned}$$

In particular, condition (i) holds if T is compact and the function $t \to g(x, t)$ is upper-semi-continuous on T, (or more generally, if M is an usco mapping on A); (ii) is fulfilled if the mapping, $(x, t) \to g^+(x, t; y)$, is upper semi-continuous on $A \times T$, for each $y \in X$.

8. Distance functions

Let us first examine distance functions defined on finite-dimensional Banach spaces. For the most part, we will only consider distance functions that are defined by smooth norms. The reason for this is revealed in the next theorem.

Theorem 8.1. *Let $(X, ||\cdot||)$ be a Banach space. If each distance function on X possesses a minimal subdifferential mapping, then the norm $||\cdot||$ on X is smooth.*

Proof. Suppose that the norm $||\cdot||$ is not smooth at a point $x_0 \in S(X)$, (Note, there is no loss of generality in assuming that $x_0 \in S(X)$). Then there exist two distinct linear functionals g_1 and $g_2 \in S(X^*)$ such that $g_1(x_0) = g_2(x_0) = 1$. Let $g_3 \equiv 1/2(g_1 + g_2)$. Let $K_1 \equiv \ker(g_1)$, $K_2 \equiv \ker(g_2)$ and $K_3 \equiv \ker(g_3)$. Clearly, $K_1 \cap K_2 \subseteq K_3$. Choose $z \in K_3\backslash(K_1 \cap K_2)$ such that $g_1(z) = 1$ and $g_2(z) = -1$. Let us recall that on page 216, Example 6 part(e) of [39], (see also, Example 2.2) an example is given of an everywhere differentiable Lipschitz function $f : \mathbb{R} \to \mathbb{R}$ which is strictly increasing on $\mathbb{R}$ and for which the set $\{x \in \mathbb{R} : f'(x) = 0\}$ is dense in $\mathbb{R}$. Moreover, this function f is a strict contraction on $\mathbb{R}$, that is, $|f(x) - f(y)| < |x - y|$ whenever $x \neq y$. Let us note that each element $x \in X$ can be uniquely expressed as $x = k_x + \lambda_x z + \mu_x x_0$, where $k_x \in K_1 \cap K_2$ and $\lambda_x, \mu_x \in \mathbb{R}$. Furthermore, $\mu_x = g_3(x)$ and $\lambda_x = 1/2(g_1(x) - g_2(x))$, and so, both mappings $x \to \mu_x$ and $x \to \lambda_x$ are continuous and open on X. Let $C \equiv \{x \in X : \mu_x \geq f(\lambda_x)\}$, (it is instructive to think of C as the epigraph of the real-valued function $f_* : K_3 \to \mathbb{R}$, defined by, $f_*(k + \lambda z) \equiv f(\lambda))$. Clearly, C is a proper, non-empty closed subset of X. We will show that, $x \to \partial d_C(x)$, is not a minimal weak* cusco on $X \backslash C$. We claim that $\sigma : X \backslash C \to C$, defined by, $\sigma(x) \equiv x + (f(\lambda_x) - \mu_x)x_0$ is a selection of the metric projection on $X \backslash C$, (Note that, if this is the case, then $d_C(x) = f(\lambda_x) - \mu_x)$. To prove this, consider a point $x \in X \backslash C$. We will show first that $\sigma(x) \in C$. To see this, consider the following,

$$\mu_{\sigma(x)} = \mu_x + (f(\lambda_x) - \mu_x) = f(\lambda_x) = f(\lambda_{\sigma(x)}) \quad \text{since } \lambda_x = \lambda_{\sigma(x)} \qquad (*)$$

Therefore, $f(\lambda_{\sigma(x)}) \leq \mu_{\sigma(x)}$ and so $\sigma(x) \in C$. Next, we show that $d_C(x) = f(\lambda_x) - \mu_x$; which will complete the proof of the claim. Let

$$\begin{aligned} T_x &\equiv \{y \in X : g_1(\sigma(x)) \leq g_1(y) \text{ or } g_2(\sigma(x)) \leq g_2(y)\} \\ &= \{y \in X : \mu_{\sigma(x)} + \lambda_{\sigma(x)} \leq \mu_y + \lambda_y \text{ or } \mu_{\sigma(x)} - \lambda_{\sigma(x)} \leq \mu_y - \lambda_y\} \end{aligned}$$

We will show that $C \subseteq T_x$. To this end, consider $y \in C$, then either,

$$\begin{aligned} &\text{(i)} \qquad f(\lambda_{\sigma(x)}) - f(\lambda_y) \leq \lambda_{\sigma(x)} - \lambda_y \quad \text{or} \\ &\text{(ii)} \qquad f(\lambda_{\sigma(x)}) - f(\lambda_y) \leq \lambda_y - \lambda_{\sigma(x)} \end{aligned}$$

Case(i) $f(\lambda_{\sigma(x)}) - \lambda_{\sigma(x)} \leq f(\lambda_y) - \lambda_y$. By $(*)$ $f(\lambda_{\sigma(x)}) = \mu_{\sigma(x)}$ and since $y \in C$, $f(\lambda_y) \leq \mu_y$. Therefore, $\mu_{\sigma(x)} - \lambda_{\sigma(x)} \leq \mu_y - \lambda_y$. Case(ii) $f(\lambda_{\sigma(x)}) + \lambda_{\sigma(x)} \leq f(\lambda_y) + \lambda_y$. As before, we have that $f(\lambda_{\sigma(x)}) = \mu_{\sigma(x)}$ and $f(\lambda_y) \leq \mu_y$. Therefore, $\mu_{\sigma(x)} + \lambda_{\sigma(x)} \leq \mu_y + \lambda_y$. Hence $y \in T_x$ and so $C \subseteq T_x$.

Now, it is easy to see that $\{y \in X : ||x - y|| < f(\lambda_x) - \mu_x\} \subseteq X \backslash T_x \subseteq X \backslash C$. Indeed, we need only do some arithmetic. Suppose that $||x - y|| < f(\lambda_x) - \mu_x$, then,

$$\begin{aligned} g_1(y) = g_1(x) + g_1(y - x) &< g_1(x) + (f(\lambda_x) - \mu_x) \quad \text{since } ||g_1|| = 1 \\ &= g_1(\sigma(x) - (f(\lambda_x) - \mu_x)x_0) + (f(\lambda_x) - \mu_x) \\ &= g_1(\sigma(x)) \end{aligned}$$

$$\begin{aligned} g_2(y) = g_2(x) + g_2(y - x) &< g_2(x) + (f(\lambda_x) - \mu_x) \quad \text{since } ||g_2|| = 1 \\ &= g_2(\sigma(x) - (f(\lambda_x) - \mu_x)x_0) + (f(\lambda_x) - \mu_x) \\ &= g_2(\sigma(x)) \end{aligned}$$

Therefore, $d_C(x) \geq f(\lambda_x) - \mu_x$, but $\sigma(x) \in C$, and so $d_C(x) = f(\lambda_x) - \mu_x$. Hence, $\nabla d_C(x) = f'(g_0(x)) \cdot g_0 - g_3$ on $X \backslash C$, where $g_0 \equiv 1/2(g_1 - g_2)$. Now, if $x \to \partial d_C(x)$ were a minimal weak* cusco on $X \backslash C$, then $x \to \nabla d_C(x)$ would be hyperplane minimal on $X \backslash C$, but then $x \to f'(g_0(x)) \cdot g_0$ would be hyperplane minimal on $X \backslash C$. However, since g_0 is both continuous and open on X, this would imply that $t \to f'(t)$ is hyperplane minimal on some non-empty open subset of $\mathbb{R}$, (this follows from the general fact if $\Phi \circ T$ is hyperplane minimal and T is both continuous and open, then Φ is hyperplane minimal) but we know this is not true, (by Example 2.2). Therefore we may conclude that, $x \to \partial d_C(x)$, is not a minimal weak* cusco on X. ■

Remark 8.1. *It is interesting to observe the following facts about the set C constructed in Theorem 8.1:*

(a) d_C is Gateaux differentiable on $X\backslash C$;

(b) $\sigma(x) \equiv x + (f(\lambda_x) - \mu_x)x_0$ is Lipschitz-continuous on $X\backslash C$, and this means that C is almost convex, (see [41] or [18], p.240);

(c) $\partial d_C(x) = \partial f(g_0(x)) \cdot g_0 - g_3$ on $X\backslash C$. In particular, d_C is not integrable on $X\backslash C$. For example, let $d_*(x) \equiv h(g_0(x)) - g_3(x)$, where $h : \mathbb{R} \to \mathbb{R}$ is chosen so that $h - f$ is not a constant function on $g_0(X\backslash C)$ and $\partial h(t) \subseteq \partial f(t)$ for each $t \in \mathbb{R}$, then $\partial d_*(x) \subseteq \partial d_C(x)$ for each $x \in X\backslash C$, but $d_* - d_C$ is not a constant function on $X\backslash C$ (note, $h \equiv 0$ will do the job).

So we see then, that even in $\mathbb{R}^2$ there are distance functions whose Clarke subdifferential mappings are not minimal, (of course there are no such examples on $\mathbb{R}$). However, the situation is dramatically better for smooth norms. A normed linear space X is said to have a *uniformly Gateaux differentiable* norm if for each $y \in X$, and each $\varepsilon > 0$, there exists a $\delta(\varepsilon, y) > 0$ such that for every $x \in X$, $||x|| = 1$, there is a continuous linear functional f_x on X and

$$\left| \frac{||x+ty|| - ||x||}{t} - f_x(y) \right| < \varepsilon \quad \text{for all } 0 < t < \delta(\varepsilon, y)$$

Every Hilbert space and L_p space $(1 < p < \infty)$ has a uniformly Gateaux differentiable norm. Furthermore, any separable Banach space can be equivalently renormed to have a uniformly differentiable norm, [42], as can any super-reflexive Banach space.

Proposition 8.2. *([5], Theorem 8) If the norm $||\cdot||$ on a Banach space X is uniformly Gateaux differentiable, then for each non-empty closed subset C of X, $-d_C$ is regular (and hence pseudo-regular) on $X\backslash C$.*

Corollary 8.3. *([2], Theorem 5.2) Let $||\cdot||$ be a uniformly Gateaux differentiable norm on a Banach space X. Then for each non-empty closed subset C of X, d_C is D-representable on X.*

Proof. From Proposition 2.1 we know that $(d_C)|_{X\backslash C}$ possesses a minimal subdifferential mapping, since $(-d_C)|_{X\backslash C}$ is regular on $X\backslash C$. But, $\partial((d_C)|_{X\backslash C})(x) = \partial d_C(x)$ for each $x \in X\backslash C$. Therefore, $x \to \partial d_C(x)$ is a minimal weak* cusco on $X\backslash C$. The result now follows from Theorem 4.10. ■

In finite dimensions all smooth norms are uniformly Gateaux differentiable. Therefore we may deduce the next result.

Proposition 8.4. *The norm $||\cdot||$ on a finite dimensional Banach space X is smooth if, and only if, each distance function defined on X possesses a minimal Clarke subdifferential mapping.*

For a smooth finite dimensional Banach space X we can characterize those subsets C of X such that $d_C \in S_e(X)$. Indeed, since no point of ∂C, (the boundary of C) can be a point of strict differentiability (recall that in a finite dimensional

Banach space the notions of strict Fréchet differentiability and strict Gateaux differentiability coincide) we immediately have a necessary condition for $d_C \in S_e(X)$, namely, ∂C must be a Lebesgue-null set. However, we have from ([5], Theorem 8) that $-d_C$ is regular on $X \backslash C \cup \text{int}C$. Therefore, if ∂C is a Lebesgue-null set then d_C is strictly differentiable almost everywhere in X, since any locally Lipschitz function which is both Gateaux differentiable and pseudo-regular at a given point is necessarily strictly differentiable at that point. Hence we may deduce the following.

Theorem 8.5. *Let $||\cdot||$ be a smooth norm on a finite dimensional Banach space X. Then for each non-empty closed subset C of X, we have that $d_C \in S_e(X)$ if, and only if, ∂C is a Lebesgue-null set.*

It is natural to ask whether the characterization given in Theorem 8.5 still holds for an arbitrary separable Banach space. Unfortunately the answer to this is 'no'. However, we do have the following Corollary.

Corollary 8.6. *Let $||\cdot||$ be a uniformly Gateaux differentiable norm on a separable Banach space X. Then for each non-empty closed subset C of X, $d_C \in S_e(X)$ whenever ∂C is a Haar-null set.*

Next, we show that the converse of this result does not hold.

Example 8.1. *In Example 6.2 part(b) of* [2]*, the author gives an example of a closed and convex subset of $c_0(\mathbb{N})$, such that ∂C is not a Haar-null set. However, as d_C is convex on X (and hence pseudo-regular on X), we must have that $d_C \in S_e(X)$. Furthermore, from* [28] *we know that such sets exist in any separable non-reflexive space. Note also, that such sets necessarily have empty interior.*

We say that a norm $||\cdot||$ on a Banach space X is a *sequentially Kadec* norm if the relative norm and relative weak topologies agree sequentially on the unit sphere, $S(X)$ (that is, if a sequence $\{x_n : n \in \mathbb{N}\} \subseteq S(X)$ converges to an element $x \in S(X)$ in the weak topology, then it converges to x in the norm topology). Using this definition we can prove another important result regarding the minimality of the subdifferential mappings of distance functions.

Theorem 8.7. *Let $||\cdot||$ be a smooth and sequentially Kadec norm on an Asplund space X. Let C be a non-empty closed subset of X such that $C \cap B(0,r)$ is relatively weakly compact for each $r > 0$, (that is, $\overline{C \cap B(0,r)}^{weak}$ is weak compact for each $r > 0$). Then d_C is D-representable on X. In particular, ∂d_C is generated by the strict Fréchet derivatives, and the set of points in $X \backslash C$ which admit a closest point in C contain a dense and G_δ subset of $X \backslash C$.*

Proof. By ([34], Theorem 2.5) we know that $\partial d_C = CSC(\Omega_D)$ where $D \equiv \{x \in X : d_C$ is Fréchet differentiable at $x\}$ and $\Omega_D : D \to 2^{X^*}$ is defined by, $\Omega_D(x) \equiv \{\nabla d_C(x)\}$. Hence, to show that ∂d_C is a minimal weak* cusco on X, we need only show by, Corollary 3.6 and Theorem 4.10 that Ω_D is hyperplane minimal on $D \backslash C$. To this end, we consider the following set-valued mapping $p_C : D\backslash C \to 2^C$ defined by, $p_C(x) \equiv \{z \in C : ||x - z|| = d_C(x)\}$. We proceed from here in two steps.

(i) Our first step is to show that p_C is a norm usco mapping on $D\backslash C$. We recall from Proposition 1.4 in [3] that for each $x \in D\backslash C$ we have that $d_C(x) = \lim_{n\to\infty} \nabla d_C(x)(x - z_n)$ for any sequence $\{z_n : n \in \mathbb{N}\} \subseteq C$ such that $\lim_{n\to\infty} ||z_n - x|| = d_C(x)$. Let us show now that for each $x \in D\backslash C$, $p_C(x)$ is non-empty. Let $x_0 \in D\backslash C$ and let $\{z_n : n \in \mathbb{N}\}$ be any sequence in C such that $\lim_{n\to\infty} ||z_n - x_0|| = d_C(x_0)$. Since the sequence $\{z_n : n \in \mathbb{N}\}$ is bounded there exists a point $z \in X$ and a subsequence $\{z_{n_k} : k \in \mathbb{N}\}$ of $\{z_n : n \in \mathbb{N}\}$ such that weak-$\lim_{n\to\infty} z_{n_k} = z$. Since any norm on X is lower semi-continuous, with respect to the weak topology on X, we have that, $||z - x_0|| \le \lim_{k\to\infty} ||z_{n_k} - x_0|| = d_C(x_0)$. However, by above we have that,

$$||z - x_0|| \ge \nabla d_C(x_0)(z - x_0) = \lim_{k\to\infty} \nabla d_C(x_0)(z_{n_k} - x_0) = d_C(x_0)$$

(note that since d_C is Lipschitz-1, $||\nabla d_C(x_0)|| \le 1$). Hence, $||z - x_0|| = d_C(x_0)$. Now, since the norm on X is sequentially Kadec and $\lim_{k\to\infty} ||z_{n_k} - x_0|| = ||z - x_0||$ we have that $\{z_{n_k} : k \in \mathbb{N}\}$ converges to z in the norm topology on X. In particular, this implies that $z \in C$, (since C is closed). Therefore, $z \in p_C(x_0)$ and so $p_C(x_0)$ is non-empty. Next we show that p_C is an usco mapping on $D\backslash C$. To do this, it suffices to show that for any $x \in D\backslash C$ and any sequences $\{x_n : n \in \mathbb{N}\} \subseteq D\backslash C$ and $\{z_n : n \in \mathbb{N}\} \subseteq C$ such that $\{x_n : n \in \mathbb{N}\}$ converges to x and $z_n \in p_C(x_n)$ for each $n \in \mathbb{N}$, $\{z_n : n \in \mathbb{N}\}$ possesses a subsequence which converges to some element z of $p_C(x)$ (in the norm topology). So let $x \in D\backslash C$ and let $\{x_n : n \in \mathbb{N}\}$ be a sequence in $D\backslash C$ which converges to x. Further, let $\{z_n : n \in \mathbb{N}\}$ be a sequence in C such that $z_n \in p_C(x_n)$ for each $n \in \mathbb{N}$. Now,

$$\begin{aligned} d_C(x) \le \liminf_{n\to\infty} ||z_n - x|| \le \limsup_{n\to\infty} ||z_n - x|| &\le \lim_{n\to\infty} ||z_n - x_n|| + \lim_{n\to\infty} ||x_n - x|| \\ &= \lim_{n\to\infty} d_C(x_n) = d_C(x) \end{aligned}$$

Therefore, $\lim_{n\to\infty} ||z_n - x|| = d_C(x)$. Now, by repeating the argument above, we have that there exists a subsequence $\{z_{n_k} : k \in \mathbb{N}\}$ of $\{z_n : n \in \mathbb{N}\}$ which converges to some point $z \in p_C(x)$ (in the norm topology). This completes part(i) of the proof.

(ii) In this step we show that Ω_D is hyperplane minimal on $D\backslash C$. Let $x \in D\backslash C$ and let z be any element of $p_C(x)$, then for each y

$$\begin{aligned} \nabla d_C(x)(y) = \lim_{\lambda\to 0} \frac{d_C(x + \lambda y) - d_C(x)}{\lambda} &\le \lim_{\lambda\to 0} \frac{||(x + \lambda y) - z|| - ||x - z||}{\lambda} \\ &= ||x - z||'(y) \end{aligned}$$

However, $y \to ||x-z||'(y)$ is linear on X, therefore we must have that $\{\nabla d_C(x)\} = \partial ||x - z||$. Furthermore, since z was an arbitrary element of $p(x)$ we must in fact, have that $\partial ||x - p(x)|| = \{\nabla d_C(x)\}$. Therefore, $x \to \partial ||x - p_C(x)||$ is a single-valued weak* usco on $D\backslash C$ (since it is the composition of two usco mappings) and hence hyperplane minimal on $D\backslash C$. This completes the proof. ■

Recall, that a set C is *densely proximinal* if the set $D(C)$ of X for which best approximations exist is dense in X, that is, if $x \in D(C)$ then there exists a point $p(x) \in C$ such that $d_C(x) = ||x - p(x)||$. When X is reflexive and the norm is sequentially Kadec, Lau's Theorem shows that every closed set is densely proximinal.

Corollary 8.8. *A Banach space $(X, ||\cdot||)$, is reflexive with a sequentially Kadec smooth norm if, and only if, each non-empty closed subset of X is densely proximinal and the corresponding distance function possesses a minimal subdifferential mapping.*

Proof. If X is reflexive and the norm $||\cdot||$ is both sequentially Kadec and smooth, then it follows from above, that each non-empty closed subset C of X is densely proximinal and the corresponding distance function d_C possesses a minimal subdifferential mapping. Conversely, if each non-empty closed subset of X is densely proximinal then by [26], X is reflexive and the norm $||\cdot||$ is sequentially Kadec. However, by Theorem 8.1, if each distance function possesses a minimal subdifferential mapping, then the norm $||\cdot||$ is smooth on X. ■

We define a *proximal normal selection* ρ_D on a subset D of $D(C)$, by setting $\rho_D(x) \equiv f_{x-p(x)}$ (where $f_{x-p(x)}$ is any element of $\partial||x - p(x)||$ for some nearest point $p(x)$), when $x \in D\backslash C$ and setting $\rho_D(x) \equiv 0$ when $x \in D \cap C$. This language is justified by the next Lemma.

Lemma 8.9. *([2], Lemma 5.3) Let C be a non-empty closed subset of a smooth Banach space X. Let D be a dense subset of $D(C)$. Then every proximal selection ρ_D satisfies $\rho_D(x) \in \partial d_C(x)$ for each $x \in D$, and so ρ_D is a densely defined selection of ∂d_C.*

Proof. If $x \in D \cap C$ then $\rho_D(x) = 0 \in \partial d_C(x)$. Suppose $x \in D\backslash C$. Fix $\varepsilon > 0$, and y in X. Then as the norm is Gateaux differentiable at $x - p(x)$, for $0 < t < \delta$

$$\begin{aligned}\frac{d_C(x+ty) - d_C(x)}{t} &\leq \frac{||x - p(x) + ty|| - ||x - p(x)||}{t}\\ &\leq f_{x-p(x)}(y) + \varepsilon \quad \text{where } \{f_{x-p(x)}\} = \partial||x - p(x)||\end{aligned}$$

and so, $f_{x-p(x)}(y) \leq (-d_C)^-(;y) \leq (-d_C)^0(x;y)$. Thus, $-f_{x-p(x)} \in \partial(-d_C)(x) = -\partial d_C(x)$; which shows that $f_{x-p(x)} \in \partial d_C(x)$. ■

Theorem 8.10. *Let C be a non-empty closed subset of a smooth Banach space X. Suppose that C is densely proximinal. Then d_C is D-representable on X if, and only if, for each proximal normal selection ρ_D, $CSC(\rho_D) = \partial d_C$.*

Proof. This follows directly from Lemma 8.9 and Theorem 3.7 part(iii). ■

Corollary 8.11. *(Proximal Normal Formula) Let C be a non-empty closed subset of a reflexive Banach space X. If the norm on X is both smooth and sequentially Kadec, then for each dense subset D of $D(C)$,*

$$\partial d_C(x) = \overline{co}\{g \in X^* : g = \text{ weak} - \lim_{x_n \to x} \rho_D(x_n), \text{ and } x_n \in D\}$$

It follows from Corollary 8.8 and Theorem 8.10 that we cannot weaken the hypothesis in Corollary 8.11 and still have a 'Proximal Normal Formula' holding for all non-empty closed subsets of X.

9. The relationship between integrability, D-representability and strict differentiability

We saw in Example 2.1 part(2) that minimality of the Clarke subdifferential mapping, and so D-representability of the underlying function, is not enough to guarantee integrability. So we begin this section by examining the converse question, namely, does integrability imply D-representability Γ The answer to this question is a little more delicate than one might first expect. Indeed, on $\mathbb{R}$, integrability does imply D-representability (see Corollary 1.3 in [4] or Proposition 10.1 part (b)), in fact on $\mathbb{R}$, integrability implies strict differentiability, almost everywhere. However, we will show next, that in general, integrability does not imply D-representability.

Example 9.1. *Let f be a real-valued Lipschitz function defined on $\mathbb{R}$ such that $\partial f \equiv [0,1]$. Let $C \equiv epi(f) = \{(x,y) \in \mathbb{R}^2 : f(x) \leq y\}$. Next, consider the distance function d_C defined on $\mathbb{R}^2$ by the l_1 norm and the set C. Then d_C is integrable on $\mathbb{R}^2$, but not D-representable on $\mathbb{R}^2$, in fact d_C is not even densely strictly differentiable on $\mathbb{R}^2$.*

Proof. Suppose that g is a real-valued locally Lipschitz function defined on $\mathbb{R}^2$ such that $\partial g(x,y) \subseteq \partial d_C(x,y)$ for each $(x,y) \in \mathbb{R}^2$. Now, $\partial d_C(x,y) = \{0\}$ on $\text{int}C$, and so $\partial g(x,y) = \{0\}$ on $\text{int}C$. But $\text{int}C$ is connected, therefore g is constant on $\text{int}C$, and so constant on C, that is, $g|_C \equiv c_1$ for some real number c_1. Next, we observe that $d_C(x,y) = f(x) - y$ for each $(x,y) \notin C$, (see Theorem 8.1 for a more detailed explanation). Therefore, $\partial d_C(x,y) = \partial f(x) \times \{-1\} = [0,1] \times \{-1\}$ on $\mathbb{R}^2 \backslash C$. Let x_0 be a fixed (but arbitrary) element on $\mathbb{R}$. We know, from above, that $g(x_0, f(x_0)) = c_1$. Therefore, by the mean-value theorem (for differentiable functions) applied to $y \to g(x_0, y)$, we have that $g(x_0, y) = c_1 + (f(x_0) - y)$ for each $y \leq f(x_0)$. Hence, $g(x,y) = (f(x) - y) + c_1 = d_C(x,y) + c_1$ on $\mathbb{R}^2 \backslash C$. But from above, we have that $g(x,y) = c_1 = d_C(x,y) + c_1$ on C. Therefore, $g = d_C + c_1$ on $\mathbb{R}^2$. ■

Remark 9.1. *It is very important to observe, that d_C is not integrable on $\mathbb{R}^2 \backslash C$. Indeed, let $f_1(x) \equiv x - f(x)$, then $\partial f_1 = \partial f$ on $\mathbb{R}$, and so $\partial g_1(x,y) = \partial f(x) \times \{-1\}$ on $\mathbb{R}^2 \backslash C$, where $g_1(x,y) \equiv f_1(x) - y$ on $\mathbb{R}^2 \backslash C$. So we see then, that in general, integrability is not inherited by open subsets. This is a striking contrast with the situation for D-representability.*

The previous example leads us to consider a stronger notion of integrability. We will say that a real-valued locally Lipschitz function f, defined on a non-empty open subset A of a Banach space X is *hereditarily integrable* on A if, for each non-empty open subset U of A the function $f|_U$ is integrable on U. It is immediate, that if f is hereditarily integrable on A then it is integrable on A, however, the previous example shows, that the converse of this is false, even when A is connected. We should also note then, that if $f \in M_e(A)$ then f is not only integrable on A, but also hereditarily integrable on A. Let us also observe, that since integrability is not inherited by open subsets, one cannot expect to characterize this property in terms of a local differentiability property, (as was done for D-representability), but rather, one must expect, such a characterization, to be in terms of some global differentiability property. We give next, a sufficient condition for a distance function to be integrable. It is note worthy, that this condition is, as we mentioned above, expressed in terms of a global property, possessed by the distance function. We say that a subset A of a topological space X is *locally connected* if for each $x \in X$ and each open neighbourhood U of x, there exists an open subset V of U, which contains the point x, such that $V \cap A$ is connected.

Proposition 9.1. *Let $||\cdot||$ be a uniformly Gateaux differentiable norm on a Banach space X. Let C be a non-empty closed subset of X. Then, (a) the distance function d_C, associated with the set C, is integrable if, int$C$ and $X\backslash C$ are both connected subsets of X and (b) the distance function d_C, associated with the set C, is hereditarily integrable if, int$C$ and $X\backslash C$ are both locally connected subsets of X*

Proof. (a)Suppose that g is a real-valued locally Lipschitz function defined on X such that $\partial g(x) \subseteq \partial d_C(x)$ for each $x \in X$. We know from Proposition 8.2 that the restriction of d_C to $X\backslash C$ is a member of $M_e(X\backslash C)$ and so by Proposition 6.3 we can see that, for some $c_1 \in \mathbb{R}$, $g(x) = d_C(x) + c_1$ on $X\backslash C$. On the other hand, $\partial g(x) \subseteq \partial d_C(x) = \{0\}$ on int$C$, and so, $g(x) = d_C(x) + c_2$ on int$C$, for some $c_2 \in \mathbb{R}$. Now, by the continuity of g and d_C, $d_C(x) + c_2 = g(x) = d_C(x) + c_1$ at each point of $\partial C \neq \emptyset$. Therefore, $c_1 = c_2$; which gives us that $g(x) = d_C(x) + c_1$ on X. (b) Let A be a non-empty open subset of X and suppose that g is a real-valued locally Lipschitz function defined on A such that $\partial g(x) \subseteq \partial(d_C)|_A(x)$ for each $x \in A$. We need to show that $\partial(d_C - g)(x) = 0$ for each $x \in A$. To this end, let $x \in A$. It is easy to see that if $x \in (A \cap \text{int}C) \cup (A\backslash C)$ then $\partial(d_C - g)(x) = 0$. Therefore we will consider the case when $x \in \partial C$. By local connectedness of $X\backslash C$ there exists an open neighbourhood V_1 of x (contained in A) such that $V_1 \cap (X\backslash C)$ is connected. Now, as in part(a), we have that $g(x) = d_C(x) + c_1$ on $V_1 \cap (X\backslash C)$. This time, by the local connectedness of int$C$ we have that there exists an neighbourhood V_2 of x (contained in V_1) such that $V_2 \cap \text{int}C$ is connected. And as in part(a) we have that $g(x) = d_C(x) + c_2$ on $V_2 \cap \text{int}C$. Now, by the continuity of g and d_C, $d_C(y) + c_2 = g(y) = d_C(y) + c_1$ for each $y \in V_2 \cap \partial C$. Therefore, $c_1 = c_2$; which gives that $g(x) = d_C(x) + c_1$ on V_2 and so $\partial(f - g)(x) = 0$. ■

We may conclude then, that even for distance functions, with respect to uniformly smooth norms, it is possible to be both integrable and D-representable, while still not being a member of $M_e(X)$. Indeed, with a little more work, we can show the even stronger result:

Example 9.2. *There exists a compact nowhere subset C of $\mathbb{R}^2$ such that (i) d_C is D-representable; (ii) d_C is hereditarily integrable; (iii) d_C is not strictly differentiable almost everywhere in $\mathbb{R}^2$, that is $d_C \notin S_e(\mathbb{R}^2)$. (Actually, there are many such examples.)*

Proof. Let C_1 be a Cantor subset of $[0,1]$ with $\mu(C_1) > 0$. Let $C \equiv C_1 \times C_1 \subseteq \mathbb{R}^2$. Let d_C be the distance function generated by the set C, with the Euclidean norm. Then by Proposition 8.4, d_C is D-representable on $\mathbb{R}^2$. To justify that d_C is hereditarily integrable it suffices by Proposition 9.1 part(b) to show that $X \backslash C$ is locally connected. So let $x \in X$ and U be an open neighbourhood of x. It is easy to see that the only non-trivial case is when $x \in \partial C$. So let us assume that $x \in \partial C$. We may now choose an $r > 0$ such that $B_\infty(x,r) \subseteq U$, where $B_\infty(x,r)$ is the l_∞ ball around x, of radius r. It now only remains to observe that $B_\infty(x,r) \cap X \backslash C$ is a connected subset (in fact, it is polygonally connected). Now, to see that $d_C \notin S_e(\mathbb{R}^2)$ we need only use that standard fact that d_C cannot be strictly differentiable at any point of $\partial C = C$. ■

Let A be a non-empty open subset of a Banach space X. Let $\mathcal{I}(A)$ denote the family of all real-valued, integrable, locally Lipschitz function defined on A, and as before, let $\mathcal{M}(A)$ denote the family of all real-valued locally Lipschitz functions defined on A whose Clarke subdifferential mappings are minimal. It follows then, that $\mathcal{N}(A) \equiv \mathcal{I}(A) \cap \mathcal{M}(A)$ is the largest class of functions which satisfy both the conditions (i) and (ii) given at the start of this paper. So why then, have we not considered the class of functions $\mathcal{N}(A)$ Γ A partial answer to this is revealed in the next example.

Example 9.3. *$\mathcal{N}(\mathbb{R}^2)$ is not closed under addition, multiplication nor either of the lattice operations. (note: we also show that $\mathcal{I}(\mathbb{R}^2)$ is not closed under addition, multiplication nor either of the lattice operations).*

Proof. (a) We show first that $\mathcal{N}(\mathbb{R}^2)$ is not closed under addition. Let f be a non-integrable real-valued Lipschitz-1 function defined on $\mathbb{R}$ such that, $x \to \partial f(x)$, is a minimal cusco. (Note: such functions exists, see Example 2.1.) Let $K_1 \equiv \{(x,y) : f(x) \le y\}$ and $K_2 \equiv \{(x,y) : y \le f(x)\}$. Next, consider the distance functions d_{K_1} and d_{K_2} defined on $\mathbb{R}^2$ by the l_1 norm and the sets K_1 and K_2, respectively. Then,

$$d_{K_1}(x,y) = \begin{cases} f(x) - y & \text{if } (x,y) \notin K_1 \\ 0 & \text{if } (x,y) \in K_1 \end{cases} \quad \text{and}$$

$$d_{K_2}(x,y) = \begin{cases} y - f(x) & \text{if } (x,y) \notin K_2 \\ 0 & \text{if } (x,y) \in K_2 \end{cases}$$

It follows from our earlier work that both d_{K_1} and $-d_{K_2}$ are integrable on $\mathbb{R}^2$ and D-representable on $\mathbb{R}^2$. However, $d \equiv d_{K_1} + (-d_{K_2})$ is not integrable on $\mathbb{R}^2$. In fact, $d(x,y) = f(x) - y$ on $\mathbb{R}^2$ and so $\partial d(x,y) = \partial f(x) \times \{-1\}$ on $\mathbb{R}^2$. Hence for any real-valued function g defined on $\mathbb{R}$ such that $\partial g = \partial f$ and $g - f$ is not a constant function on $\mathbb{R}$, the function $G(x,y) \equiv g(x) - y$, shares that same Clarke subdifferential mapping as d, (while not differing from d by a constant). Therefore,

d is not integrable on $\mathbb{R}^2$. (b) Next, we show that $\mathcal{N}(\mathbb{R}^2)$ is not closed under multiplication. Let $d^*_{K_1} \equiv d_{K_1} + 1$ and $d^*_{K_2} \equiv (-d_{K_2}) + 1$. Then $d^* \equiv d^*_{K_1} \cdot d^*_{K_2}$ is not integrable on $\mathbb{R}^2$. To see this, we compute d^*; $d^*(x,y) = (f(x) - y) + 1$ on $\mathbb{R}^2$. Then as in (a) we see that d^* is not integrable on $\mathbb{R}^2$. (c) Finally, we show that $\mathcal{N}(\mathbb{R}^2)$ is not closed under the lattice operations. Let C be a Cantor subset of $[0,1]$ with $\mu(C) > 0$. We define two sets; $C_1 \equiv \{(x,y) \in \mathbb{R}^2 : x \in C \text{ and } y \leq 0\}$ and $C_2 \equiv \{(x,y) \in \mathbb{R}^2 : x \in C \text{ and } y \geq 0\}$. Now, consider the distance functions d_{C_1} and d_{C_2} defined on $\mathbb{R}^2$ by the Euclidean norm and the sets C_1 and C_2 respectively. Then $d_{C^*}(x,y) \equiv \min\{d_{C_1}(x,y), d_{C_2}(x,y)\}$ is the distance function to the set $C^* \equiv \{(x,y) \in \mathbb{R}^2 : x \in C \text{ and } y \in \mathbb{R}\}$. Moreover, it is easy to see that $d_C(x,y) = d(x)$, where $d : \mathbb{R} \to \mathbb{R}$ is defined by, $d(x) \equiv \min\{|x - c| : c \in C\}$. However, d is not integrable on $\mathbb{R}$ since d is not strictly differentiable almost everywhere on $\mathbb{R}$, in particular, d is not strictly differentiable at any point of C, (see, Proposition 10.1 part(b)). Therefore there exists a Lipschitz function g such that $\partial g = \partial d$ and $g - d$ is not a constant function on $\mathbb{R}$. Let $G(x,y) \equiv g(x)$. Clearly then, $\partial G = \partial d_{C^*}$, but $G - d_{C^*}$is not a constant function on $\mathbb{R}^2$. To show that $\mathcal{N}(\mathbb{R}^2)$ is not closed under 'max' we need only consider $-d_{C^*}$. ■

Another reason why we have not considered the class $\mathcal{N}(A)$ is that thus far, we have not been able to deduce a reasonable characterization for membership in this class of functions.

Let us consider the Venn diagram corresponding to the four sets $\mathcal{D}(A)$, $\mathcal{I}(A)$, $S_e(A)$, $\mathcal{M}(A)$ defined on a non-empty open subset A of a separable Banach space. Then there are seven potentially non-void intersections. $S_e(A) \subseteq \mathcal{M}(A) \subseteq \mathcal{D}(A)$. and our results allow us to supply members of each.

We end this section by giving a comment concerning integrability with respect to the *approximate subdifferential* mapping. It is possible to construct two Lipschitz functions f and g mapping from $\mathbb{R}^2$ into $\mathbb{R}$ such that $\partial f = \partial g$ is minimal, while $\partial_a f$ and $\partial_a g$ differ on a set of positive measure. This cannot happen on the real-line, where ∂f determines $\partial_a f$, [4], (here $\partial_a f$ denotes the approximate subgradient of f). On the other hand we should observe that our conditions for integrability imply integrability with respect to any subdifferential mapping, $x \to \partial f_{\#}(x)$, which has the property that $\overline{co}^{w*} \partial f_{\#}(x) = \partial f(x)$ for each x. Let us also comment that in general $x \to \partial_a f(x)$ is a weak* usco, however, it is very rarely a minimal usco. Indeed, even the approximate subgradient of the absolute value function fails to be a minimal usco.

10. Examples and miscellaneous results

Let us begin this Section by justifying Example 2.1 given in Section 2.

Example 10.1. *(Example 2.1 — revisited)*

Proof. Let us first observe, that by direct calculation one can show that $g'(x) = 0$ for each $x \in C$. (**1**) It is easy to see that $x \to \partial g(x)$, is a minimal cusco on $(0,1)\backslash C$. Indeed, g is strictly differentiable on $(0,1)\backslash C$ except for countably many points. It is also easy to see that $\partial g(x) = [-2,2]$ at each point of C. In fact, $\partial g = CSC((\partial g)|_Y)$. Therefore, we may conclude, from Corollary 3.6 that ∂g is a minimal cusco on $(0,1)$.

Next, we observe that $(f+g)'(x) = f'(x)+g'(x) \in \partial g(x)$ almost everywhere in $(0,1)$. Therefore, by Theorem 2.5 in [12], $\partial(f+g)(x) \subseteq \partial g(x)$ for each $x \in (0,1)$. However, since $x \to \partial g(x)$, is a minimal cusco on $(0,1)$ we must have that $\partial(f+g) = \partial g$. A similar argument shows that $\partial(g-f) = \partial g$ on $(0,1)$. From this we can deduce that $\partial(f-g)$ is a minimal cusco by observing that $\partial(f-g)(x) = (-1)\cdot\partial(g-f)(x)$ at each $x \in (0,1)$. (**2**) This follows immediately from part(1). (**3**) From part(1) we see that both h and k possess minimal subdifferential mappings, but $h+k = 2f$, which clearly does not possess a minimal subdifferential mapping, since in particular, $0 \in \partial f(x)$ for each $x \in (0,1)$, while ∂f is not identically equal to $\{0\}$. (**4**) $M(x) \equiv \max\{k(x), h(x)\} = f(x) + |g(x)|$ and $m(x) \equiv \min\{k(x), h(x)\} = f(x) - |g(x)|$. Moreover,

$$|g|(x) \equiv \begin{cases} 0 & \text{if} \quad |\, x - c_n \,| \geq d_n \text{ for all } n. \\ 2(x-(c_n-d_n)) & \text{if} \quad x \in (c_n - d_n, c_n - \frac{2}{3}d_n] \\ -(x-c_n) & \text{if} \quad x \in (c_n - \frac{2}{3}d_n, c_n] \\ 2(x-c_n) & \text{if} \quad x \in (c_n, c_n + \frac{1}{3}d_n] \\ (c_n+d_n)-x & \text{if} \quad x \in (c_n - \frac{1}{3}d_n, c_n - d_n) \end{cases}$$

Therefore, $M^+(x) \geq -1$ at each point of $(0,1)\backslash C$. However, there exists a set of positive measure $A \subseteq C \cap (1/2, 1)$ such that $M'(x) = f'(x)+|g|'(x) = f'(x) = -2$ at each point of A. Hence, ∂g is not generated by the derivatives chosen from $(0,1)\backslash C$ (which are dense in $[0,1]$), that is, g is not D-representable on $(0,1)$, and so g does not possess a minimal subdifferential mapping on $(0,1)$. A similar argument shows that m does not possess a minimal subdifferential mapping. (**5**) Clearly, both $j+g$ and $j-g$ possess minimal subdifferential mappings. In fact, $\partial(j+g) = \partial(f+g)$ and $\partial(j-g) = \partial(f-g)$. However, $(j+g)\cdot(j-g) = j^2 - g^2$, which does not possess a minimal subdifferential mapping, because $(j^2-g^2)'(x) = 2(j(x)j'(x) - g(x)g'(x))$ almost everywhere in $(0,1)$ and so $(j^2-g^2)'(x) \leq 2$ almost everywhere in $(0,1)\backslash C$ (note: $-1 \leq g(x) \leq 1$ on $(0,1)$). However, there exists a set of positive measure $A \subseteq C \cap (0, 1/2)$ such that $(j^2-g^2)'(x) = 2(j(x)j'(x)+g(x)g'(x)) = 2j(x)j'(x) \geq 4$ at each point of A. Hence, as in (4), it follows that $(j+g)\cdot(j-g)$ is not D-representable, and so $(j+g)\cdot(j-g)$ does not possess a minimal subdifferential mapping. ■

Next, we gather-up a few special fact concerning locally Lipschitz functions defined on $\mathbb{R}$.

Proposition 10.1. *Let I be a non-empty open interval of $\mathbb{R}$, then:*

(a) Each minimal cusco $\Phi : I \to 2^{\mathbb{R}}$ is the Clarke subdifferential mapping of some D-representable locally Lipschitz function defined on I. (Note: not every cusco from I into $2^{\mathbb{R}}$ is the Clarke subdifferential mapping of some locally Lipschitz function defined on I.)

(b) A locally Lipschitz function $f : I \to \mathbb{R}$ is integrable on I if, and only if, $f \in S_e(I)$.

(c) $S_e(I)$ is closed under addition, subtraction, multiplication and division (when this is defined) as well as, the lattice operations. Moreover, $S_e(I)$ is closed under composition (when this is defined).

(d) $f \in S_e(I)$ if, and only if, the mapping $x \to f^+(x;1)$ is Riemann integrable on I.

Proof. (a) Define $h : I \to \mathbb{R}$ by, $h(x) \equiv \max\{\Phi(x)\}$. Then h is upper semi-continuous on I and hence Borel measurable on I. Define $f : I \to \mathbb{R}$ by, $f(x) \equiv \int_{[a,x]} h(t)dt$ (here a is any element of I), then $f'(x) = h(x)$ almost everywhere in I, (let us call this set D). Then, $\partial f(x) = CSC(f^+|_D)(x) \subseteq \Phi(x)$ for each $x \in I$. But since Φ is a minimal cusco, we may deduce that $\partial f = \Phi$. (b) This is Corollary 1.3 in [4]. (c) This is just Corollary 5.16 and Theorem 5.17. (d) A famous Theorem due to Lebesgue says that a real-valued function defined on an interval of $\mathbb{R}$ is Riemann integrable if, and only if, the function is continuous almost everywhere. When we combine this, with the fact that f^+ is continuous at a point $x \in I$ if, and only if, ∂f is single-valued at x, we obtain the desired result. ■

Ironically, D-representable Lipschitz functions are also useful in constructing of highly pathological Lipschitz functions.

Theorem 10.2. *([8], Theorem 1) Let $f_1, f_2, \ldots f_n$ be real-valued locally Lipschitz functions defined on a non-empty open subset A of a separable Banach space X. If each function f_j possesses a minimal Clarke subdifferential mapping on A, then there exists a real-valued locally Lipschitz function g defined on A such that $\partial g(x) = co\{\partial f_1(x), \partial f_2(x), \ldots \partial f_n(x)\}$ for each $x \in A$.*

Remark 10.1. *It was noted in [8] that the function g, given above, is not integrable, except perhaps when, $\partial f_1 = \partial f_2 = \ldots . \partial f_n$.*

We have seen so far in this paper, that those Lipschitz functions which are D-representable, possess very desirable differentiability properties. Hence, the following result due to D. Preiss, [34] is very surprising. In [34] the author show that there is a Lebesgue null set, G_0 which is also a G_δ subset of $\mathbb{R}^n$, (n≥ 2) such that

$$CSC(\partial f|_{G_0 \cap D_f}) = \partial f = CSC(\partial f|_{D_f \backslash G_0})$$

for all locally Lipschitz mappings $f : \mathbb{R}^n \to \mathbb{R}$. Thus, paradoxically, both sets G_0 and $D_f \backslash G_0$ reconstruct any Lipschitz function, where $D_f \equiv \{x \in \mathbb{R}^n : \nabla f(x)$ exists$\}$.

Next, we give a slight improvement of Theorem 3.11 and Corollary 4.9.

Proposition 10.3. *Let f and g be real-valued locally Lipschitz functions defined on a non-empty open subset A of a separable Banach space X. If $x \to \partial f(x)$ is a minimal weak* cusco on A and $g \in S_e(A)$ then;*

(i) $x \to \partial(f+g)(x)$ is a minimal weak* cusco on A;

(ii) $x \to \partial(f \cdot g)(x)$ is a minimal weak* cusco on A;

(iii) $x \to \max\{f(x), g(x)\}$ and $x \to \min\{f(x), g(x)\}$ possess minimal subdifferential mappings on A.

Proof. (**i**) It follows from Theorem 7.5 in [10] that $\nabla(f+g)(x) = \nabla f(x) + \nabla g(x)$ almost everywhere in A, (call this set D). Let S denote the set of all points of strict differentiability of g in A. It now follows, in a similar manner to Theorem 3.5 part(ii), that $x \to \nabla(f+g)(x)$ is hyperplane minimal on $S \cap D$. The result can now be deduced from Proposition 5.3. (**ii**) Again by Theorem 7.5 in [10] we have that $\nabla(f \cdot g)(x) = \nabla f(x)g(x) + f(x)\nabla g(x)$ almost everywhere in A, (call this set D). As before, let S denote the set of all points of strict differentiability of g in A. Then by Theorem 5.3 parts(i) and (ii) we have that $x \to \nabla(f \cdot g)$ is hyperplane minimal on $S \cap D$. The result may now be obtained from Theorem 5.3. (**iii**) $\max\{f(x), g(x)\} = (f-g)^+(x) + g(x)$ and $\min\{f(x), g(x)\} = (f-g)^-(x) + g(x)$. Now, by part(i) $f - g$ possesses a minimal subdifferential mapping. Therefore, by Corollary 4.9 part(ii) $(f-g)^+$ and $(f-g)^-$ possess minimal subdifferential mappings. The proof is completed by again appealing to part(i) above. ■

Remark 10.2. *It follows from this, that in Example 2.2 none of g, h, k, M nor m could belong to $S_e((0,1))$. Of course this is easily checked directly.*

The next result can be considered to be an 'abstract invariance' result.

Theorem 10.4. *(Abstract invariance) Let A be a topological space and let Ω be a minimal usco (minimal cusco) from A into subsets of a Hausdorff topological space (separated linear topological space) X. Let T be a set-valued mapping from A into non-empty (non-empty and convex) subsets of X. If the graph of T is closed in $A \times X$, then the following conditions are equivalent:*
(i) $\{t \in A : \Omega(t) \cap T(t) \neq \emptyset\}$ is dense in A;
(ii) $\Omega(t) \cap T(t) \neq \emptyset$ for each $t \in A$;
(iii) $\Omega(t) \subseteq T(t)$ for each $t \in A$.

Proof. It follows from a standard compactness argument, that (i) $\Rightarrow$(ii). Let $\Omega^*(t) \equiv \Omega(t) \cap T(t)$ for each $t \in T$. Then Ω^* possesses a closed graph in $A \times X$. Moreover, $\Omega^*(t) \subseteq \Omega(t)$ for each $t \in A$. Therefore, by Proposition 1.3 we have that Ω^* is an usco (a cusco) mapping on A. But Φ is a minimal usco (minimal cusco) on A, therefore, $\Omega = \Omega^*$. Hence, $\Omega(t) \subseteq T(t)$ for each $t \in A$; which shows that (ii)$\Rightarrow$(iii). Finally, it is obvious that (iii) implies (i). ■

This last result has interesting more concrete applications to fixed point theory and to differential inclusions. Now standard assumptions force Ω to be a cusco while T needs to be a closed tangent cone set-valued mapping. This is often not the case (for example, C might be the orthant which has problems at the origin) unless C is a reasonable manifold. (However, if $T \equiv N_C$ is the Clarke normal cone and C is epi-Lipschitz, as in [12] or convex the T is closed.) Then (ii) is a *weak inwardness* condition ensuring the existence of a solution in C to $0 \in \Omega(x)$. Also (ii) is a standard hypothesis for *weak invariance* and ensures under appropriate conditions that the differential inclusion $x'(t) \in \Omega(x(t))$, $x(0) \in C$ has a *viable* solution: remaining (locally) in C, (see, [1]). Correspondingly, (iii) is a *strong invariance* condition used to show that every solution remains in C, [1]. Thus for a minimal cusco and closed tangent cone we need only check (i) to determine (iii). Moreover, weak and strong invariance are then effectively co-determinate.

We end the paper by examining the minimality of vector-valued functions. Let $f : \mathbb{R}^n \to \mathbb{R}^k$ be a locally Lipschitz function defined on a non-empty open subset $A \subseteq \mathbb{R}^n$, defined by, $f(x) \equiv (f_1(x), f_2(x), \ldots f_k(x))$. Then the (Clarke) *generalized Jacobian* of f, (denoted ∂f) is defined by, $\partial f(x) \equiv CSC(D(f))$, where $D(f)$ is the classical Jacobian of f defined on $\{x \in A : D(f) \text{ exists}\}$.

Proposition 10.5. *Let $f : \mathbb{R}^n \to \mathbb{R}^k$ be a locally Lipschitz function defined on a non-empty open set $A \subseteq \mathbb{R}^n$, defined by, $f(x) \equiv (f_1(x), f_2(x), \ldots f_k(x))$. If $f_j \in S_e(A)$ for each $j \in \{1, 2, \ldots k\}$, then the (Clarke) generalized Jacobian $x \to \partial f(x)$, is a minimal cusco on A.*

Proof. We may, without loss of generality, assume that $A \equiv B(x, r)$, for some $x \in X$, and $r > 0$. (this is because minimality is locally determined, see Corollary 3.2). Let $S \equiv \{x \in A : \text{each } f_j,\ 1 \leq j \leq k, \text{ is strictly differentiable at } x\}$. By the Theorem in [15] we have that $\partial f = CSC(D(f))$, where $D(f)$ is the classical Jacobian of f restricted to S. It is easy to see that, $x \to D(f)(x)$, is continuous on S, and so the result follows from Corollary 3.6. ■

In the same fashion we may define $\partial f = CSC(D(f))$ whenever $f : E \to F$, E separable and F RNP. In this case it is not known whether ∂f is insensitive to Haar-null sets. Nonetheless, if $f : E \to F$ is $C^{1,1}$while E^* is separable one may define a *generalized Hessian* $\partial^2 f = CSC(\nabla^2 f)$ and study it accordingly. As with first order derivatives $\partial^2 f$ is a cusco, and hence maybe manipulated in a similar manner.

ACKNOWLEDGEMENT

The authors would like to thank John R. Giles for suggesting the refinement of Theorem 4.6, given in Theorem 4.7.

References

1. J. P. Aubin and A. Cellina, *Differential Inclusions*, Springer–Verlag, New York, 1984.
2. J. M. Borwein, Minimal cuscos and subgradients of Lipschitz functions, in : *Fixed Point Theory and its Applications*, (J.– B. Baillion and M Thera eds.), Pitman Lecture Notes in Math., Longman, Essex, (1991) 57–82.
3. J. M. Borwein and Simon Fitzpatrick, Existence onf nearest points in Banach spaces, *Can. J. Math.* **61** (1989) 702-720.
4. Jonathan Borwein and Simon Fitzpatrick, Characterization of Clarke subgradients among one-dimensional multifunctions, *CECM Preprint* 94:006, (1994).
5. J. M. Borwein, S. P. Fitzpatrick and J. R. Giles, The differentiability of real functions on normed linear spaces using generalized subgradients, *J. Math. Anal. and Appl.* **128** (1987) 512–534.
6. Jonathan Borwein, Simon Fitzpatrick and Petàr Kenderov, Minimal convex usco and monotone operators on small sets, *Can. J. Math.* **43**(3) (1991) 461–476.
7. J. Borwein, S. Fitzpatrick and J. Vanderwerff, Examples of convex functions and classification of normed spaces, *J. Convex Anal.* to appear.
8. Jonathan M. Borwein, Warren B. Moors and Wang Xianfu, Lipschitz functions with prescribed derivatives and subderivatives, *CECM Preprint* (1994).
9. Jens Peter Reus Christensen, On sets of Haar measure zero in Abelian Polish groups, *Israel J. Math.* **13** (1972) 255–260.

10. J. P. R. Christensen, *Topological and Borel Structure*, American Elsevier, New York, 1974.
11. Jens Peter Reus, Theorems of Namioka and R. E. Johnson type for upper semicontinuous and compact valued set-valued mappings, *Proc. Amer. Math. Soc.* **86** (1982) 649–655.
12. Frank H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley Inter-science, New York, 1983.
13. R. Correa and A. Jofre, Tangentally continuous directional derivatives in Nonsmooth Analysis, *J. Optim. Theory Appl.* **61**(1) (1989) 1–21.
14. R. Correa and L. Thibault, Subdifferential Analysis of bivariate separately regular functions, *J. Math. Anal. and Appl.* **148** (1990) 157–174.
15. M. Fabian and D. Preiss, On the Clarke's Generalised Jacobian, *Supplemento ai Rendiconti del Circolo Matematico di Palermo*, (series II) **14** (1987).
16. Pando Gr. Georiev, Submonotone mappings in Banach spaces and applications, *Preprint*.
17. J. R. Giles, On a characterisation of Asplund spaces, *J. Austral. Math. Soc. (series A)* **32** (1982) 134–144.
18. J. R. Giles, *Convex Analysis with Applications in Differentiation of Convex Functions,* Res. Notes in Math., Nr **58**, Pitman, Boston–London–Melbourne, 1982.
19. J. R. Giles, P. S. Kenderov, W. B. Moors and S. D. Sciffer, Generic differentiability of convex functions on the dual of a Banach space, *Pacific J. Math.* to appear.
20. John R. Giles and Warren B. Moors, A continuity property related to Kuratowski's index of non-compactness, it relevance to the Drop property, and its implications for differentiability theory, *J. Math. Anal. and Appl.* **178** (1993) 247–268.
21. J. R. Giles and W. B. Moors, Generic continuity of minimal set-valued mappings, *Preprint*.
22. A. Jofre and L. Thibault, D-representation of subdifferentials of directionally Lipschitzian functions, *Proc. Amer. Math. Soc.* **110** (1990) 117–123.
23. L. Jokl, Minimal convex-valued weak* usco correspondences and the Radon-Nikodym property, *Comm. Math. Uni. Carolinae* **28** (1987) 353–375.
24. S. Kempisty, Sur les fonctions quasicontinues, *Fund. Math.* **19** (1932) 184–197.
25. Petàr S. Kenderov and Orihuela, On a generic factorization theorem, *Mathematika* to appear.
26. S. V. Konjagin, On approximation property of closed sets in Banach spaces and the characterization of strongly convex spaces, *Soviet Math. Dokl.* **21** (1980) 418–422.
27. D. G. Larman andR. R. Phelps, Gateaux differentiability of convex functions on Banach spaces, *J. London Math.Soc.* **20** (1979) 115–127.
28. E. Matouskova and C. Stegall, A characterization of reflexive spaces, *Proc. Amer. Math. Soc.* to appear.
29. R. Mifflin, Semismooth and semiconvex functions in Optimization, *SIAM J. Control. Opt.* **15** (1977) 959–972.
30. Warren B. Moors, A characterisation of minimal subdifferential mappings of locally Lipschitz functions, *Set–Valued Anal.* to appear.
31. M. E. Munroe, *Measure and Integration*, Second Edition, Addison-Wesley, Reading, 1971.
32. Robert R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, Springer–Verlag, New York, 1993.
33. R. A. Poliquin, Integration of subdifferentials of non-convex functions, *Nonlinear Anal. Th. Meth. Appl.* **17** (1991) 385–398.
34. D. Preiss, Fréchet derivatives of Lipschitz functions, *J. Funct. Analysis* **91** (1990) 312–345.
35. 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.
36. L. Qi, The maximal normal operator space and integration of subdifferentials of non-convex functions, *Nonlinear Anal. Th. Meth. Appl.* **13** (1989) 1003–1012.
37. R. T. Rockafellar, Favourable classes of Lipschitz-continuous functions in subgradient optimization, *Progress in Nondifferentiable Optimization*, (E. Nurminski ed.), pp. 125–144.IIASA Collaborative Proceedings Series, International Institute of Applied Systems Analysis, Laxenberg, Austria (1982).
38. Scott D. Sciffer, *Differentiability Properties of Locally Lipschitz Functions on Banach Spaces*, Ph.D. Thesis, Newcastle University, Australia 1993.
39. Karl R. Stromberg, *An Introduction to Classical Real Analysis,* Wadsworth International Mathematics Series, 1981.
40. L. Thibault, On Generalized differentials and subdifferentials of Lipschitz vector-valued functions, *Nonlinear Anal. Th. Meth. Appl.* **6** (1982) 1037–1053.

41. L. P. Vlasov, Almost convex and Chebychev sets, *Math. Notes Acad. Sc. USSR* **8** (1970) 776–779.
42. V. Zizler, Some notes on various rotundity and smoothness properties of separable Banach spaces, *Comm. Math. Uni. Carolinae* **10** (1967) 195–206.

Department of Mathematics and Statistics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

Department of Mathematics, The University of Auckland, Private Bag 92019, Auckland, New Zealand