Null sets and essentially smooth Lipschitz functions

Jonathan M. Borwein*
CECM, Department of Mathematics and Statistics,
Simon Fraser University,
Burnaby, BC V5A 1S6, Canada jborwein@cecm.sfu.ca

Warren B. Moors†
Department of Mathematics,
The University of Auckland,
Private Bag 92019, Auckland, New Zealand

June 4th 1996; revised January 17th 1997

Abstract
In this paper we extend the notion of a Lebesgue-null set to a notion which is valid in any completely metrizable Abelian topological group. We then use this definition to introduce and study the class of essentially smooth functions. These are, roughly speaking, those Lipschitz functions which are smooth (in each direction) almost everywhere.

Keywords: Lipschitz functions, Haar-null sets, Essentially smooth functions


*Research supported by NSERC and the Shrum Endowment at Simon Fraser University.
†Research supported by a New Zealand Science and Technology Post-Doctoral Fellowship.
1 Introduction.

The primary purpose of this paper is to introduce a new class of Lipschitz functions and present some of their most fundamental properties. In order to accomplish this goal, we must first extend the notion of a Lebesgue-null set to notion which is valid in any Banach space. In the case of separable Banach spaces this task has already been achieved. In fact, for separable Banach spaces, there have been several successful solutions to this problem (see, [20], [15] and [7] to name just a few). For our purposes the most useful generalization of a “null” set is that of the Haar-null (or zero) set introduced by J. P. R. Christensen in [7]. Indeed, in section two of this paper we retrace and extend the results of Christensen to spaces which are not necessarily separable. While some of our extensions need additional work, many follow naturally from their separable counter-parts.

In section three we introduce the class of “essentially smooth” Lipschitz functions. Loosely speaking, these are those functions which are smooth (in each direction) almost everywhere, that is, smooth (in each direction) everywhere, except on a “null” set. As we shall come to see, the essentially smooth functions enjoy particularly desirable differentiability properties. For instance, they are all integrable and $D$-representable (on class($\mathcal{S}$) spaces) in the sense defined within.

This class of functions is of central importance in nonsmooth optimization. The importance of the class stems (see [2, 4, 5, 6]) from the fact that it provably provides the largest robust class of functions

- stable under the most significant operations of calculus and of formation of marginal functions ([4]);
- uniquely determined by their Clarke subgradients ([4]);
- smooth enough for algorithmic purposes ([4, 18]);
- possessing strong chain rules ([5]).

A primary utility of this paper is that it provides coherent tools for the use of “measure-theoretic” techniques in non-separable spaces, such as $L^\infty$, which are frequently natural settings for optimization and control problems.
2 Haar-null sets

Throughout this section of the paper, $(G, +, \tau)$ will denote a completely metrizable Abelian topological group, that is, $(G, \tau)$ is homeomorphic to a complete metric space. Due to a result of V. Klee, (see, [14]) we may and do, assume that the topology $\tau$ is generated by a metric $d$, which is both complete and invariant. (Actually, all invariant metrics on a completely metrizable Abelian group are complete. This, however, is not generally the case in non-Abelian groups.)

In our terminology a measure will always be non-negative and not identically zero. By a Borel measure on a topological space $X$ we mean any measure defined on $\mathcal{B}(X)$ – the Borel subsets of $X$. By a Radon measure $\mu$ on $X$ we mean any Borel measure on $X$, extended to its completion on $X$, which satisfies:

(i) $\mu(K) < \infty$ for each compact subset $K \subseteq X$;

(ii) $\mu(A) = \sup\{\mu(K) : K \subseteq A, K$ compact$\}$ for each $A \in \mathcal{B}(X)$.

Note that, (ii) actually hold for each $A$ in the $\mu$-completion of $\mathcal{B}(X)$. In addition, if $\mu$ is finite (that is, $\mu(X) < \infty$) then, $\mu(A) = \inf\{\mu(U) : A \subseteq U, U$ open$\}$ for each $A$ in the $\mu$-completion of $\mathcal{B}(X)$.

We will say that a subset $A$ of a topological space $X$ is universally (Radon) measurable if it belongs to the $\mu$-completion of each finite (Radon) Borel measure $\mu$ on $X$ and we shall denote by $\mathcal{U}(X)$ ($\mathcal{U}_R(X)$) the family of all universally (Radon) measurable subsets on $X$. Observe that $\mathcal{B}(X) \subseteq \mathcal{U}(X) \subseteq \mathcal{U}_R(X)$. (In the case when $X$ is Polish, $\mathcal{U}(X) = \mathcal{U}_R(X)$ see, Theorem 2.2.) It follows immediately from this definition that on any topological group, the universally (Radon) measurable sets form a translation invariant $\sigma$-algebra.

A natural question to ask at this point, is whether the family of all universally measurable subsets properly contains the Borel sets. One way to see the affirmative answer to this, is the following.

Given a completely regular topological space $X$, let $\mathcal{D}(X)$ denote the smallest $\sigma$-algebra on $X$ which contains all the Borel subsets and is stable under the Souslin operation. (Let us refresh our memory on the definition of the Souslin operation. We say that a family of sets $\mathcal{A}$ is closed under the Souslin operation if all sets of the form:

$$\bigcup\bigcap\{A(\sigma|n) : n \geq 1\} : \sigma \in \mathcal{N}^N\}, \ A(\sigma|n) \in \mathcal{A}$$

[Here we use $\sigma|n$ to denote $\sigma_1, \sigma_2, \ldots, \sigma_n$ when $(\sigma_1, \sigma_2, \ldots, \sigma_n, \ldots) \in \mathcal{N}^N$]
are contained in $A$.) By using the fact, that for each complete finite measure $\mu$ on $X$, the family of all $\mu$-measurable sets is closed under the Souslin operation, we may deduce that each member of $\mathcal{D}(X)$ is universally measurable (see, [21]). Now, it is known that $\mathcal{D}(X)$ contains all the Čech-analytic subsets of $X$, which in turn, contain all the analytic subsets of $X$ (see, [12]). (Recall, that a subset $A \subseteq X$ is analytic if it is the continuous image of a complete, separable metric space, that is, the continuous image of a Polish space.) Finally, it is also known that each uncountable Polish space contains an analytic set which is not a Borel set, (see, [13] p.201).

We are now in a position to define our "null" sets. Let $(G,+,\tau)$ be a completely metrizable Abelian topological group. We will call a universally Radon measurable set $A \subseteq G$ a Haar-null set if there exists a (not necessarily unique) Radon probability measure $p$ on $G$ such that $p(g + A) = 0$ for each $g \in G$. (In such a case, we shall call the measure $p$ a test-measure for $A$.) More generally, we will say that a subset $A \subseteq G$ is a Haar-null set if it is contained in a universally Radon measurable Haar-null set. It will sometimes be useful to remember that if $p$ is a test-measure for $A$, then $g + A$ is $p$-measurable for each $g \in G$, even though the set $A$ may not be universally Radon measurable. We should also note that the Dirac (point-mass) measures are rarely useful as test-measures, since the only set they can test is the empty set.

In order to simplify the later part of this section, we will take this opportunity to present some of the fundamental properties possessed by Radon measures. All of the results presented in the next theorem are either straightforward or maybe found (more-or-less as stated) in Chapter two of [1].

**Theorem 2.1** Let $X$ and $Y$ be Hausdorff topological spaces and let $Z$ be a Borel subset of $X$. Further, let $\mu$ be a Radon measure on $X$ and $\nu$ be a Radon measure on $Y$. Then we have the following:

(a) The space $X$ contains a smallest closed subset $A \subseteq X$, (called the support of $\mu$) such that $\mu(X \setminus A) = 0$. It follows immediately from this definition, that for each open subset $U \subseteq X$, $\mu(U \cap A) > 0$ whenever, $U \cap A \neq 0$. Furthermore, if $X$ is metrizable and $\mu(X) < \infty$, then the support of $\mu$ is separable.

(b) If we define $\mu_Z : \mathcal{B}(Z) \to [0, \infty]$ by, $\mu_Z(A) \equiv \mu(A)$, for each $A \in \mathcal{B}(Z)$ then (the completion of) $\mu_Z$ defines a Radon measure on $Z$, which we will call, the restriction of $\mu$

(c) Conversely, if we are given a finite Radon measure $\mu_Z$ on $Z$, and we
Null sets and essentially smooth Lipschitz functions

define \( \eta : \mathcal{B}(X) \to [0, \infty] \) by \( \eta(A) \equiv \mu_Z(A \cap Z) \), for each \( A \in \mathcal{B}(X) \) then (the completion of) \( \eta \) is a Radon measure on \( X \), which in future, we will call the extension of \( \mu_Z \). (Note that, if \( Z \) is a closed subset of \( X \), then we need not assume that \( \mu_Z \) is finite.)

(d) \( \mathcal{U}_R(Z) = \{ A \in \mathcal{U}_R(X) : A \subseteq Z \} \).

(e) If \( \mu \) is \( \sigma \)-finite, that is, \( X = \bigcup \{ X_n : n \in \mathbb{N} \} \), where each \( X_n \in \mathcal{B}(X) \) and \( \mu(X_n) < \infty \), then each member of \( \mathcal{U}_R(X) \) is \( \mu \)-measurable.

(f) If \( A \) is a universally Radon measurable subset of \( X \times Y \) and both the measures \( \mu \) and \( \nu \) are \( \sigma \)-finite then the mapping \( x \to \nu([A]_X(x)) \) is universally Radon measurable on \( X \) and \( y \to \mu([A]_Y(y)) \) is universally Radon measurable on \( Y \). (Here, \([A]_X(x) \equiv \{ y \in Y : (x, y) \in A \} \) and \([A]_Y(y) \equiv \{ x \in X : (x, y) \in A \} \).

Moreover, there exists a unique Radon product measure on \( X \times Y \), denoted \( \mu \otimes \nu \), that restricts to the usual product measure on \( \mathcal{B}(X) \times \mathcal{B}(Y) \) and for which

\[
\int_A d(\mu \otimes \nu) = \int_X \nu([A]_X(x))d\mu = \int_Y \mu([A]_Y(y))d\nu
\]

(g) If \( f \) is a continuous mapping from \( X \) into \( Y \) and \( \mu(X) < \infty \), then the measure \( \mu_f : \mathcal{B}(Y) \to [0, \infty] \) defined by \( \mu_f(A) \equiv \mu(f^{-1}(A)) \) extends to define a Radon measure on \( Y \). Moreover, \( f^{-1}(A) \in \mathcal{U}_R(X) \) for each \( A \in \mathcal{U}_R(Y) \).

(h) If \( \mu \) and \( \nu \) are finite measures and \( X = Y \) then we may define the convolution of \( \mu \) and \( \nu \) as, \( \mu \ast \nu \equiv (\mu \otimes \nu)_+ \), that is,

\[ \mu \ast \nu(A) \equiv \mu \otimes \nu(\{(x, y) : x + y \in A\}) \text{ for each } A \in \mathcal{B}(X) \]

The completion of this measure is a finite Radon measure on \( X \).

The convolution defined above plays a central role in what follows.

Remark 2.1 It follows from (f) and (h) above, that if \( p_A \) is a test-measure for a set \( A \) and \( p_B \) is a test-measure for a set \( B \) then, \( p_A \ast p_B \) is a test-measure for \( A \cup B \). Indeed, since \( p_A \ast p_B(A \cup B) \leq p_A \ast p_B(A) + p_A \ast p_B(B) \) it suffices to observe, that \( p_A \ast p_B(A) = 0 \) and \( p_A \ast p_B(B) = 0 \). As a consequence of this, we may deduce that the Haar-null sets are closed under finite unions. Later, we shall see that this remains true of countable unions.

Before we present the elementary properties possessed by Haar-null sets, we will digress for a moment, so that we may justify calling our “null” sets, Haar-null sets, as in [7].
**Theorem 2.2** (Ulam’s Theorem, [9] p.176) Let \((X, \tau)\) be a separable, and completely metrizable topological space, (that is, \(X\) is a Polish space). Then the completion of each finite Borel measure on \(X\), defines a finite Radon measure on \(X\).

A generalization of the above theorem is the following.

**Remark 2.2** If the support of a finite Borel measure is Polish then its completion is a finite Radon measure. In particular, if \((X, \tau)\) is metrizable, then each finite Borel measure on \(X\), with compact support, extends to be a finite Radon measure on \(X\).

So we see then, that the definition of a Haar-null set given in this paper is consistent with that given in [7]. Hence, in a Polish setting, it also coincides with the definition of a shy set, as given in [11].

A key set of facts is listed next:

**Proposition 2.1** Let \((G, +, \tau)\) be a completely metrizable Abelian group. Then we have the following:
(a) Every subset of a Haar-null set is Haar-null;
(b) If \(A\) is a Haar-null set, then so is \(g + A\), for every \(g \in G\);
(c) If \(A\) is a Haar-null set, then there exists a test-measure for \(A\), with compact support;
(d) Suppose that \(A\) is a subset of a universally Radon measurable set \(B\). Then to show that \(A\) is a Haar-null set, it is sufficient to show that there exists a test-measure for \(B\) which is \(\sigma\)-finite. Note that if \(G\) is Polish, then all Radon measures on \(G\) are \(\sigma\)-finite, (see, [1] p.62) and so in this case it is sufficient to use any Radon measure. If \(B\) is a Borel set, then we may use any Borel measure \(\mu\) (as a test-measure) for which there exists a compact subset \(K \subseteq G\) with \(0 < \mu(K) < \infty\);
(e) If \(A\) is a Haar-null set, then \(G\setminus A\) is dense in \(G\);
(f) If \(\{A_j : j \in N\}\) are Haar-null sets, then so is \(\bigcup\{A_j : j \in N\}\);
(g) If \(G\) is a locally compact Polish group, then a subset \(A \subseteq G\) is a Haar-null set if, and only if, \(\lambda(A) = 0\), for the Haar-measure \(\lambda\) on \(G\).

**Proof.** The proof of the facts (a)-(e) are obvious and hence left to the reader.
(d) Suppose that \(\mu\) is a test-measure for the set \(B\). Let \(K\) be any compact subset of \(G\) such that \(0 < \mu(K) < \infty\). It is easy to see that the finite Radon
measure, μₖ, defined by μₖ(X) = μ(∩X) for each X ∈ ℬ(G) is also a test-measure for the set B.

To prove (c) it is sufficient to show that there are no (non-empty) open Haar-null sets. To this end, let U be a non-empty open subset of G and suppose, to the contrary, that there does exist a test-measure p for U on G. Let A denote the support of p. It is easy to see that for some g₀ ∈ G, (g₀ + U) ∩ A ≠ 0. Hence, p(g₀ + U) ≥ p((g₀ + U) ∩ A) > 0; which contradicts the fact that p is a test-measure for U. Hence, no non-empty open subset of G is a Haar-null set.

(f) Let us first observe, that without loss of generality we may assume that each set Aⱼ is universally Radon measurable. For each j ∈ N, let pⱼ be a test-measure for Aⱼ on G. Let H be the smallest closed subgroup of G that contains the support of each pⱼ. Since the support of each pⱼ is separable (see, Theorem 2.1 part (a)) it is not too difficult to see that H is Polish. Next, let pⱼ* denote the restriction of pⱼ to H. We know from Theorem 1 in [7] that there exists a Radon probability measure p* on H that is a test-measure for each set of the form \( \bigcup \{ B_j : j \in N \} \), provided that \( B_j \in ℬ(H) \) and \( p_j^* \) is a test-measure for \( B_j \). (The measure \( p^* \) is essentially the infinite convolution of the \( p_j^* \).) Let \( p \) be the extension, to G, of \( p^* \). We claim that \( p \) is a test-measure for \( \bigcup \{ A_j : j \in N \} \). To prove this, we must show that for each \( g \in G \), \( p(g + \bigcup \{ A_j : j \in N \}) = 0 \). So let us fix \( g \in G \). Then,

\[
p(g + \bigcup_{j=1}^{∞} A_j) = p^*((g + \bigcup_{j=1}^{∞} A_j) \cap H) = p^*(\bigcup_{j=1}^{∞} (g + A_j) \cap H)
\]

However, each \( p_j^* \) is a test-measure for \( (g + A_j) \cap H \) since each \( p_j \) is a test-measure for \( A_j \) and \( h + ((g + A_j) \cap H) = ((h + g) + A_j) \cap H \subseteq (h + g) + A_j \) for each \( h \in H \). Therefore, \( p^* \) is a test-measure for \( \bigcup \{ (g + A_j) \cap H : j \in N \} \), and so \( p^*(\bigcup \{ (g + A_j) \cap H : j \in N \}) = 0 \), which gives that \( p(g + \bigcup \{ A_j : j \in N \}) = 0 \).

Let us also note, that an alternative proof of this is given in [11]. We should, however, lend a note-of-caution here, that in their proof they implicitly assume that the metric under consideration is both translation invariant and complete. (Of course, from our remark at the start of this section, we see that such an assumptions is indeed valid.)

(g) The proof of this fact maybe found in [7].

One of the most powerful theorems in measure theory is Fubini’s theorem, and so it is natural to want to determine whether there is a version of Fubini’s theorem that holds for Haar-null sets. Disappointingly, in [7], the author
gives an example to show that we must banish all hope of obtaining a full version of Fubini’s theorem. Nonetheless, in this same paper, the author indicates (without proof) that a weaker version of Fubini’s theorem does hold (in a Polish Abelian group). We next take the opportunity to record a proof of this theorem, in the setting of completely metrizable Abelian groups.

**Theorem 2.3** If \((H, +, \tau_1)\) is a completely metrizable Abelian group and \((T, +, \tau_2)\) is a locally compact Polish Abelian group, then for each universally Radon measurable subset \(A \subseteq H \times T\) (the product group), the following are equivalent:

(i) \([A]_H(h)\) is a Haar-null set, for the Haar-measure on \(T\), for almost all \(h \in H\);

(ii) The set \(A\) is a Haar-null set in the product group \(H \times T\).

**Proof.** Let \(p_T\) be a Radon probability measure on \(T\) that is equivalent to the Haar-measure on \(T\), that is, for each \(B \in \mathcal{U}_R(T)\), \(p_T(B) = 0\) if, and only if, \(B\) is a null set for the Haar-measure on \(T\). (Note: such a measure \(p_T\) exists since \(T\), with the Haar-measure, is \(\sigma\)-finite.)

Suppose that (i) holds. Let \(A_H \equiv \{h \in H : \mu([A]_H(h)) > 0\}\). Let \(p\) be a test-measure for the set \(A_H\). We claim that \(p \otimes p_T\) is a test-measure for \(A\). Fix \((h_0, t_0) \in H \times T\). Then,

\[
\{h \in H : p_T([h_0, t_0] + A]_H(h)) > 0\} = h_0 + A_H
\]

The result now follows from Theorem 2.1 part (f).

Suppose that (ii) holds. Let \(\delta_0\) denote the Dirac (point-mass) measure at \(0_H\) and define \(\mu \equiv \delta_0 \otimes p_T\). Let \(p\) be a test-measure for the set \(A\) and set \(A_H \equiv \{h \in H : p_T([A]_H(h)) > 0\}\). We know from Theorem 2.1 part (f) that \(A_H\) is universally Radon measurable, so it suffices to construct a test-measure for \(A_H\). To this end, we define a Radon measure \(p_H\) on \(H\) by \(p_H(B) \equiv p(B \times T)\) for each \(B \in \mathcal{U}_R(H)\) (Note: \(p_H(B) \equiv p(P^{-1}(B))\) where \(P\) is the natural projection of \(H \times T\) onto \(H\)). Fix \(h_0 \in H\), then,

\[
0 = \mu * p((h_0, 0_T) + A) = \int_{H \times T} \mu(A + (h_0 - h, -t))dp(h, t)
\]

Now, the mapping \((h, t) \mapsto \mu(A + (h_0 - h, -t)) = p_T([A + (0_H, -t)]_H(h - h_0))\) is universally Radon measurable and hence \(p\)-measurable (see, Theorem 2.1 parts (f) and (g)) and strictly positive on \((h_0 + A_H) \times T\).
Null sets and essentially smooth Lipschitz functions

Hence, $p_H(h_0 + A_H) = p((h_0 + A_H) \times T) = 0$. 

There is one significant problem which, up until this point, we have neglected. That is, how does one actually determine whether a given set is Haar-null? In general, this is a difficult question to answer. In some situations however, the following facts are helpful.

**Proposition 2.2** Let $A$ be a Borel subset of a completely metrizable Abelian group $(G, +, \tau)$. Then $A$ is a Haar-null set if for some locally compact subgroup $H$ of $G$, $\lambda((g + A) \cap H) = 0$ for each $g$ in some transversal $T$ to $H$ in $G$. Here, $\lambda$ denotes the Haar-measure on $H$.

Let us recall the definition of a transversal. A subset $T$ of a group $G$, with normal subgroup $H$, which contains just one element from each coset of $H$ is called a transversal to $H$ in $G$.

The next condition, is stated in [7], for Polish groups. However, we note here that the proof in no-way relies on the fact that the group is separable.

**Proposition 2.3** Let $A$ and $B$ be arbitrary universally Radon measurable sets in a completely metrizable Abelian group $G$. If we put

$$F(A, B) \equiv \{g \in G : (g + A) \cap B \text{ is not a Haar-null set}\}$$

Then $F(A, B)$ is an open subset of $G$ (possibly empty). In particular, if $A$ is not a Haar-null set, then $F(A, A)$ is a open neighbourhood of $0_G$.

**Corollary 2.1** If $(G, +, \tau)$ is a completely metrizable Abelian group, which is not locally compact, then each compact subset is Haar-null.

**Proof.** Suppose that to the contrary that there exists a compact subset $K \subseteq G$ which is not a Haar-null set. Then, by Proposition 2.3, $F(K, K)$ is an open neighbourhood of $0_G$. However, $0_G \in F(K, K) \subseteq K - K$; which is compact. Therefore, all compact subsets of $G$ must be Haar-null.

By employing considerably more effort, one can show, in the setting of Banach spaces, that certain more complicated sets are Haar-null. For instance, in [3] the authors show that the boundary of every closed convex subset of every separable super-reflexive Banach space is Haar-null. Moreover, it is shown in [16] that the boundary of every closed convex nowhere-dense set in any super-reflexive space is Haar-null (Note: In [16] the author does not use the Haar-null terminology). Whereas, it is known (see, [17]) that every
separable non-reflexive Banach space contains a closed convex subset whose boundary is not Haar-null (Note: such a set must have no interior).

We end this section with an example that shows that the completeness of the group $G$, in the proof of Proposition 2.1 part $(f)$ is indispensible.

**Example 2.1** Let $(X, || \cdot ||)$ be a infinite dimensional normed linear space such that, $X = \bigcup \{X_n : n \in N \}$, where each $X_n$ is a finite dimensional subspace of $X$. It follows from Proposition 2.2 that each subspace $X_n$ is a Haar-null subset of $X$. (To see this, consider the following. For each $n \in N$, let $H_n$ be a one-dimensional subspace of $X$ such that $H_n \not\subseteq X_n$ and let $\lambda_n$ denote the Lebesgue measure on $H_n$, then $\lambda_n((x + X_n) \cap H_n) = 0$ for each $x \in X$.) However, $p(\bigcup \{X_n : n \in N \}) = 1$, for each probability measure $p$ on $X$.

3 **Essentially smooth Lipschitz functions**

In this section of the paper we will introduce the class of essentially smooth Lipschitz functions and develop some of their most fundamental properties. We begin by recalling some preliminary definitions regarding the Clarke generalized directional derivative [8]. A real-valued function $f$ defined on a non-empty open subset $U$ of a Banach space $X$ is said to be *locally Lipschitz* on $U$, if for each $x_0 \in U$ there exists a $K > 0$ and a $\delta > 0$ such that

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

For functions in this class, we will consider the following three directional derivatives:

(i) The **upper Dini derivative** at $x \in U$, in the direction $y$, given by

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

(ii) The **lower Dini derivative** at $x \in U$, in the direction $y$, given by

$$f^-(x; y) = \liminf_{\lambda \to 0^+} \frac{f(x + \lambda y) - f(x)}{\lambda}$$
(iii) The Clarke generalized directional derivative at \( x \in U \), in the direction \( y \), given by
\[
\dot{f}^0(x; y) \equiv \limsup_{\lambda \to 0} \frac{f(z + \lambda y) - f(z)}{\lambda}
\]
It is immediate from these definitions that for each \( x \in U \) and each \( y \in X \)
\[
f^-(x; y) \leq f^+(x; y) \leq f^0(x; y)
\]
An initially surprising, but important, fact regarding the Clarke generalized
directional derivative is that for each \( x_0 \in U \), the mapping, \( y \to \dot{f}^0(x_0; y) \)
is a continuous sublinear functional on \( X \) (see, [8]). Associated with the
Clarke generalized directional derivative is the Clarke subdifferential mapping,
which is defined by,
\[
\partial f(x) \equiv \{ x^* \in X^* : x^*(y) \leq \dot{f}^0(x; y) \text{ for each } y \in X \} \text{ for each } x \in U
\]
Next, we shall examine some notions of differentiability. We say that a real-valued
locally Lipschitz function \( f \) is differentiable at \( x \), in the direction \( y \), if
\[
f'(x; y) \equiv \lim_{\lambda \to 0} \frac{f(x + \lambda y) - f(x)}{\lambda}
\] exists
and we say that \( f \) is Gateaux differentiable at \( x \), if
\[
\nabla f(x)(y) \equiv \lim_{\lambda \to 0} \frac{f(x + \lambda y) - f(x)}{\lambda}
\] exists
for each \( y \in X \) and \( \nabla f(x) \) is a continuous linear functional on \( X \).
We shall also be interested in a slightly stronger notion of differentiability.
A locally Lipschitz function \( f \) is said to be strictly differentiable at \( x \), in the
direction \( y \), if
\[
\lim_{\lambda \to 0} \frac{f(z + \lambda y) - f(z)}{\lambda}
\]
exists
and we say that \( f \) is strictly differentiable at \( x \), if \( f \) is strictly differentiable
at \( x \), in every direction \( y \in X \). We recall that a locally Lipschitz function \( f \)
is strictly differentiable at \( x \), in the direction \( y \) if, and only if,
\[
f^0(x; y) = f'(x; y) = -f^0(x; -y)
\]
Our first theorem in this section shows us that the Clarke generalized
directional derivative is insensitive to Haar-null sets.
Lemma 3.1 \textit{(Mean-value Theorem)} Let $f$ be a real-valued locally Lipschitz function defined on a non-empty open subset of the real line, which contains the non-degenerate interval $[a, b]$. Then there exists a Borel subset $M$ of $[a, b]$, with positive measure, such that for each $t \in M$, $f'(t)$ exists and
\[
f'(t) \geq \frac{f(b) - f(a)}{b - a}.
\]

Theorem 3.1 Let $f$ be a real-valued locally Lipschitz function defined on a non-empty open subset $U$ of a Banach space $X$. Then for each $x \in U$, $y \in X \setminus \{0\}$ and Haar-null set $N \subseteq X$, we have that,
\[
f^0(x; y) = \limsup_{z \to x} f^-(z; y) - f^0(x; -y) = \liminf_{z \to x} f^+(z; y)
\]

\textbf{Proof.} Since the proof of the second case is identical to that of the first, we shall only consider the first case. Fix $x \in U$, $y \in X \setminus \{0\}$ and let $N$ be a Haar-null subset of $X$. (Note that, by possibly making $N$ bigger, we may assume that it is universally Radon measurable.) We begin by observing that
\[
\limsup_{z \to x} f^-(z; y) \leq \limsup_{z \to x} f^+(z; y) = f^0(x; y)
\]
So it is sufficient to show that
\[
f^0(x; y) \leq \limsup_{z \to x} f^-(z; y)
\]
Indeed, to accomplish this, it will be sufficient to show that for each $\epsilon > 0$ and each $\delta > 0$
\[
f^0(x; y) - \epsilon < \sup\{f^-(z; y) : ||z - x|| \leq \delta \text{ and } z \in U \setminus N\}
\]
So suppose $\epsilon > 0$ and $\delta > 0$ are given. Note that, by possibly making $\delta$ smaller, we may assume that $B(x, \delta) \subseteq U$. Let $H$ be any closed hyperplane in $X$, that does not contain the point $y$, and consider the isomorphism $T : H \times R \to X$ defined by, $T(h, r) \equiv h + ry$. If we set $N_H \equiv \{h \in H : \lambda(\{r \in R : T(h, r) \in N\}) > 0\}$ then by Theorem 2.3 we have that $N_H$ is a Haar-null subset of $H$.
Now, by the definition of $f^0(x; y)$, there exists a point $z \equiv T(h_z, r_z) \in U$ and a $\lambda \in (0, 1)$ such that, $||(z + \lambda y) - x|| < \delta$, $||z - x|| < \delta$ and
\[
\frac{f(z + \lambda y) - f(z)}{\lambda} > \frac{f(z + \lambda y) - f(z)}{\lambda} > f^0(x; y) - \epsilon
\]
Null sets and essentially smooth Lipschitz functions

Since $f$ is locally Lipschitz on $U$, we may, by possibly moving $z$ slightly, assume that $h_z \in H \setminus N_H$. Now, by the Mean-value Theorem there exists a set $M$ of positive measure in the interval $(z, z + \lambda y)$ where $f'(z'; y)$ exists and

$$f'(z'; y) \geq \frac{f(z + \lambda y) - f(z)}{\lambda} > f^0(x; y) - \epsilon$$

Choose $z_0 \in M \setminus N \subseteq U$ (this is possible since $h_z \not\in N_H$). Then $||z_0 - x|| < \delta$, and so

$$f^0(x; y) - \epsilon < f'(z_0; y) \leq \sup \{f^-(z; y) : ||z - x|| < \delta \text{ and } z \in U \setminus N\}$$

This completes the proof. ☺

We now isolate the class of essentially smooth functions. A real-valued locally Lipschitz function $f$ defined on a non-empty open subset $U$ of a Banach space $X$ is called essentially smooth on $U$ if for each $y \in X \setminus \{0\}$

$$\{x \in U : f^0(x; y) \neq -f^0(x; -y)\}$$

is a Haar-null subset of $X$

Furthermore, we shall denote by $\mathcal{S}_n(U)$ the family of all real-valued essentially smooth locally Lipschitz functions on $U$.

Our first observation concerning these functions is that on separable Banach spaces they are just those locally Lipschitz functions which are strictly differentiable almost everywhere. (Note: in [2] and [6] such functions functions were called essentially strictly differentiable.)

**Theorem 3.2** Let $U$ be a non-empty open subset of a separable Banach space $X$. Then each member of $\mathcal{S}_n(U)$ is strictly differentiable almost everywhere in $U$.

**Proof.** Let $\{y_n : n \in N\}$ be a dense subset of $X \setminus \{0\}$. For each $n \in N$, let $N_n \equiv \{x \in U : f^0(x, y_n) \neq -f^0(x, -y_n)\}$. We claim that $f$ is strictly differentiable at each point of $U \setminus \bigcup \{N_n : n \in N\}$. To see this, consider a fixed point $x_0 \in U \setminus \bigcup \{N_n : n \in N\}$. Now, as both mappings $y \to f^0(x_0; y)$ and $y \to -f^0(x_0; -y)$ are continuous on $X$ and $f^0(x_0; y_n) = -f^0(x_0; -y_n)$ for each $n \in N$, we must have that $f^0(x_0; y) = -f^0(x_0; -y)$ for each $y \in X$. This shows that $f$ is strictly differentiable at $x_0$, and hence, almost everywhere in $U$. ☺

In order to establish that the essentially smooth functions are $D$-representable (see below) we shall need some further definitions.

A set-valued mapping $\Phi$ from a topological space $U$ into subsets of a linear topological space $(X, \tau)$ is called a $\tau$-cusc on $U$ if:
(i) for each $x \in U$, $\Phi(x)$ is non-empty, convex and compact;

(ii) for each open subset $W$ of $X$, $\{x \in U : \Phi(x) \subseteq W\}$ is open in $U$.

(Set-valued mappings satisfying only (ii) are called upper or outer semi-continuous.)

Our interest in cusco mappings arises from the fact that the Clarke subdifferential mapping of a locally Lipschitz function defined on a non-empty open subset of a Banach space is, a fortiori, a weak$^*$ cusco. Amongst the class of cusco mappings, special attention is given to the class of minimal cuscoks. A cusco mapping $\Phi$ from a topological space $U$ into subsets of a linear topological space $(X, \tau)$ is called a minimal $\tau$-cusco if its graph does not strictly contain the graph of any other $\tau$-cusco defined on $U$.

We may now present the following central theorem.

**Theorem 3.3** Let $U$ be a non-empty open subset of a Banach space $X$. Then the Clarke subdifferential mapping of each member of $\mathcal{S}_c(U)$ is a minimal weak$^*$ cusco on $U$.

**Proof.** Let $f \in \mathcal{S}_c(U)$ and let us suppose, for the purpose of obtaining a contradiction, that $x \to \partial f(x)$, is not a minimal weak$^*$ cusco on $U$. That is, let us suppose that there exists a weak$^*$ cusco $\Phi$ on $U$, whose graph is strictly contained in that of $\partial f$. It follows then, via a separation argument in $(X^*, weak^*)$, that we may find a point $x_0 \in U$ and a direction $\dot{y} \in \dot{X}$ (here $\dot{X}$ denotes the natural embedding of $X$ into $X^{**}$) such that:

$$\max \dot{y}(\Phi(x_0)) < \max \dot{y}(\partial f(x_0))$$

Now consider the two cusco mappings $T_y : U \to 2^R$ and $S_y : U \to 2^R$ defined by,

$$S_y(x) \equiv \dot{y}(\Phi(x)) \quad \text{and} \quad T_y(x) \equiv \dot{y}(\partial f(x))$$

Clearly, $S_y(x) \subseteq T_y(x)$ for each $x \in U$. However, $S_y(x_0) \neq T_y(x_0)$ since

$$\max S_y(x_0) = \max \dot{y}(\Phi(x_0)) < \max \dot{y}(\partial f(x_0)) = \max T_y(x_0)$$

So we will obtain our desired contradiction if we can show that $T_y = S_y$. Let

$N \equiv \{x \in U : f^0(x; y) \neq -f^0(x; -y)\}$. It is clear then, from the definition of $T_y$, that

$$S_y(x) = T_y(x) = [-f^0(x; -y), f^0(x; y)] = \{f^1(x; y)\} \quad \text{at each point of } U \setminus N.$$
Null sets and essentially smooth Lipschitz functions

We may now use the upper semi-continuity of $S_y$, a simple separation argument (in $\mathbb{R}$) and Theorem 3.1 to deduce that $S_y = T_y$. 

The importance of the previous result is that minimality of the Clarke subdifferential mapping is closely related to the, about to be defined, notion of $D$-representability. Moreover, minimality of the Clarke subdifferential mapping also provides a means for deducing differentiability results. Indeed, let us recall that a Banach space $X$ is said to be of class$(S)$ if every minimal weak* cusco that maps from a Baire space, into subsets of $X^*$, is single-valued at the points of a dense and $G_δ$ subset of its domain (see, [10]). In consequence, we see that in these spaces the essentially smooth functions are strictly differentiable at the points of a dense and $G_δ$ subset of their domain. The family of all class$(S)$ spaces is rather large. Indeed, all smooth Banach spaces (that is, spaces which admit an equivalent norm that is Gateaux differentiable everywhere, except at $0$) belong to this class (see, [10]), as do, all those Banach spaces that contain, as a dense subspace, the continuous linear image of an Asplund space (such spaces are called GSG spaces, [10]).

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

$$\partial f(x) = \{ x^* \in X^* : x^* = weak^* - \lim_{x_n \to x} \nabla f(x_n) \}$$

The next theorem now follows directly from Corollary 2.2 and Theorem 2.5 in [4].

**Theorem 3.4** Let $U$ be a non-empty open subset of a class$(S)$ Banach space. Then each member of $\mathcal{S}_c(U)$ is $D$-representable.

The next property that we shall address is that of integrability. Let $f$ be a real-valued locally Lipschitz function defined on a non-empty open subset $U$ of a Banach space $X$. Then we say that $f$ is integrable on $U$ if $\partial(f - g) \equiv \{0\}$ for each real-valued locally Lipschitz function $g$, defined on $U$, with $\partial g(x) \subseteq \partial f(x)$ for each $x \in U$.

We now show that the essentially smooth functions are integrable.

**Proposition 3.1** Suppose that $U$ is a non-empty open subset of a Banach space $X$. Let $f$ and $g$ be real-valued locally Lipschitz function on $U$ such
that \( \partial g(x) \subseteq \partial f(x) \) for each \( x \in U \). Then \( \partial (f - g) \equiv \{0\} \) if, and only if, \( x \to \partial (f - g)(x) \), is a minimal weak* cusc\( \circ \) on \( U \).

**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 \( x \in U \). So if, \( x \to \partial(f - g)(x) \), is a minimal weak* cusc\( \circ \) on \( U \), then \( \partial(f - g) \equiv \{0\} \). The converse is obvious.

**Corollary 3.1** Suppose that \( U \) is a non-empty open connected subset of a Banach space \( X \). Let \( f \) and \( g \) be real-valued locally Lipschitz functions defined on \( U \) such that \( \partial g(x) \subseteq \partial f(x) \) for each \( x \in U \). Then \( f - g \) is constant on \( U \) if, and only if, \( x \to \partial(f - g)(x) \), is a minimal weak* cusc\( \circ \) on \( U \).

**Proof.** Suppose that \( x \to \partial(f - g)(x) \), is a minimal weak* cusc\( \circ \) on \( U \). Let \( x_0 \in U \) and let \( U_0 \equiv \{ x \in U : (f - g)(x) = (f - g)(x_0) \} \). It follows from the Mean-value Theorem and the result above, that both \( U_0 \) and \( U \setminus U_0 \) are open subsets of \( U \). Now, \( U_0 \neq \emptyset \) since \( x_0 \in U_0 \). Therefore, \( U_0 = U \); which shows that \( f - g \) is constant on \( U \). The converse is again obvious.

**Theorem 3.5** Let \( U \) be a non-empty open subset of a Banach space \( X \). Then each member of \( \mathcal{S}_c(U) \) is integrable.

**Proof.** Suppose that \( f \in \mathcal{S}_c(U) \) and \( g \) is any real-valued locally Lipschitz function defined on \( U \) such that \( \partial g(x) \subseteq \partial f(x) \) for all \( x \in U \). Clearly then \( g \in \mathcal{S}_c(U) \). Moreover, since \( \partial(f - g)(x) \subseteq \partial f(x) - \partial g(x) \) we see that \( f - g \in \mathcal{S}_c(U) \), and so the subdifferential mapping, \( x \to \partial(f - g)(x) \), is a minimal weak* cusc\( \circ \) on \( U \). Hence, by the previous Proposition, \( \partial(f - g) \equiv \{0\} \) on \( U \).

We say that a locally Lipschitz function \( f \) is hereditarily integrable on an open subset \( U \) if its restriction to every open subset is integrable (on that open set). Hence, we see that the members of \( \mathcal{S}_c(U) \) are not only integrable, but also hereditarily integrable on \( U \). It is appropriate to recall that there are Lipschitz functions which are integrable but not hereditarily integrable (see, [4] Section 7).

**Theorem 3.6** (Identity Theorem) Suppose that \( f \) and \( g \) are real-valued locally Lipschitz functions defined on a non-empty open connected subset \( U \) of a Banach space \( X \). If \( f \in \mathcal{S}_c(U) \) and \( x \to \partial g(x) \), is a minimal weak* cusc\( \circ \), then \( f - g \equiv \text{constant on } U \) if, and only if, \( \{ x \in U : \partial g(x) \cap \partial f(x) \neq \emptyset \} \) is dense in \( U \).
Null sets and essentially smooth Lipschitz functions

Proof. Consider the set-valued mapping $T : U \to 2^{X^*}$ defined by, $T(x) \equiv \partial g(x) \cap \partial f(x)$. Since both mappings, $x \to \partial f(x)$ and $x \to \partial g(x)$, are upper semi-continuous on $U$ (and possess compact images) $T$ possesses non-empty weak* compact, convex images. Moreover, since the graphs of both $\partial f$ and $\partial g$ are closed in $U \times X^*$, with $X^*$ equipped with the weak* topology, so is the graph of $T$. Hence, we may deduce that $T$ is a cuso on $U$. However, for each $x \in U$, $T(x) \subseteq \partial f(x)$ and $T(x) \subseteq \partial g(x)$. Therefore, by the minimality of $\partial f$ and $\partial g$ we must have that $\partial g = T = \partial f$. The result now follows from Theorem 3.5. Yet again, the converse is obvious. \(\textcircled{8}\)

In the following theorem we provide a condition which is sufficient to ensure membership in $S_c(U)$.

Theorem 3.7 Let $f$ be a real-valued locally Lipschitz function defined on a non-empty open subset $U$ of a Banach space $X$. Let $B$ be any subset of $X$ such that $\overline{spB} = X$. If for each $b \in B$, $\{x \in U : f^0(x; b) \neq -f^0(x; -b)\}$ is a Haar-null set, then $f \in S_c(U)$.

Proof. Let $y \in X \setminus \{0\}$. We will show that $\{x \in U : f^0(x; y) \neq -f^0(x; -y)\}$ is a Haar-null set. We may choose a countable set $B' \equiv \{b_n : n \in N\} \subseteq B$ so that $y \in \overline{spB'}$. For each $n \in N$, let $S_n \equiv \{x \in U : f^0(x; b_n) = -f^0(x; -b_n)\}$. It is clear from the hypothesis that each set $U \setminus S_n$ is Haar-null. Let $S \equiv \bigcap\{S_n : n \in N\}$. We claim that $f$ is strictly differentiable, in the direction $y$, at each point of $S$. So let us consider an arbitrary point $x_0 \in S$. To show that $f$ is strictly differentiable at $x_0$ in the direction $y$, we need only show that if $x_1$ and $x_2$ are Clarke subgradients of $f$ at $x_0$, then $x_1^*(y) = x_2^*(y)$ (this is because, if $f^0(x_0; y) \neq -f^0(x_0; -y)$, then we could construct, using the Hahn-Banach extension theorem, two distinct subgradients $x_1$ and $x_2$ of $f$, at $x_0$, such that $x_1^*(y) = f^0(x_0; y)$ and $x_2^*(y) = -f^0(x_0; -y)$). To this end, let $x_1$ and $x_2$ be Clarke subgradients of $f$ at $x_0$. Then, $x_1^*(b_n) \leq f^0(x_0; b_n)$ and $x_1^*(-b_n) \leq f^0(x_0; -b_n)$ for each $i \in \{1, 2\}$ and $n \in N$. Therefore,

$$x_i^*(b_n) \leq f^0(x_0; b_n) = -f^0(x_0; -b_n) \leq -x_i^*(-b_n) = x_i^*(b_n)$$

for each $i \in \{1, 2\}$ and $n \in N$. Hence, $x_1^*(b_n) = f^0(x_0; b_n) = x_2^*(b_n)$ for each $n \in N$. Now, both $x_1$ and $x_2$ are linear on $X$ and so we have that $x_1^*(b) = x_2^*(b)$ for each $b \in spB'$. However, $x_1^*$ and $x_2^*$ are also both continuous on $X$, therefore $x_1^* = x_2^*$ on $\overline{spB'}$. In particular, $x_1^*(y) = x_2^*(y)$. \(\textcircled{8}\)

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

\[
f^+(x; y) \geq \limsup_{t \to 0^+} f^0(x + ty; y) \quad \left( f^-(x; y) \leq \liminf_{t \to 0^+} f^0(x + ty; -y) \right)
\]

If \( f \) is both upper hemi-smooth and lower hemi-smooth at \( x \in U \), in the direction \( y \), then we say that \( f \) is hemi-smooth at \( x \), in the direction \( y \).

**Remark 3.1** If we define

\[
T_y : U \to R \text{ by, } T_y(x) \equiv \limsup_{t \to 0^+} f^0(x + ty; y)
\]

and

\[
S_y : U \to R \text{ by, } S_y(x) \equiv \liminf_{t \to 0^+} f^0(x + ty; -y)
\]

then it is easy to check that both \( T_y \) and \( S_y \) are Borel measurable on \( U \). Hence, the set of points in \( U \) where \( f \) is upper (lower) hemi-smooth in the direction \( y \), is always a Borel subset of \( U \). Indeed, to see that \( T_y \) is Borel measurable, it suffices to observe that:

\[
T_y(x) = \lim_{n \to \infty} g_n(x) \text{ where, } g_n(x) \equiv \sup \{f^0(x + ty; y) : 0 < t \leq 1/n\}
\]

and

\[
g_n(x) \equiv \lim_{m \to \infty} f^m_n(x) \text{ where, } f^m_n(x) \equiv \max \{f^0(x + ty; y) : 1/m \leq t \leq 1/n\}
\]

(for each \( m > n \)) is upper semi-continuous on \( U \). A similar argument shows that \( S_y \) is also Borel measurable.

Further, we say that \( f \) is essentially upper hemi-smooth (essentially lower hemi-smooth) on \( U \), in the direction \( y \), if the set of all points in \( U \) where \( f \) is not upper hemi-smooth (lower hemi-smooth) in the direction \( y \), is a Haar-null set. We shall also say that \( f \) is pseudo-regular at \( x \) in the direction \( y \) if, \( f^0(x; y) = f^+(x; y) \) and we shall say that \( f \) is pseudo-regular at \( x \), if it is pseudo-regular at \( x \), in every direction \( y \).

**Lemma 3.2** Let \( f \) be a real-valued locally Lipschitz function defined on a non-empty open subset \( U \) of a Banach space \( X \). Then, in the notation of the previous remark, for each \( y \in S(X) \) the (Borel) set

\[
F_y \equiv \{ x \in U : f^0(x; y) > T_y(x) \} \quad \left( E_y \equiv \{ x \in U : -f^0(x; y) < S_y(x) \} \right)
\]

has the property that for each \( x_0 \in U \), \( F_y(x_0) \equiv \{ r \in R : x_0 + ry \in F_y \} \)

\( E_y(x_0) \equiv \{ r \in R : x_0 + ry \in E_y \} \) is at most countable.
Null sets and essentially smooth Lipschitz functions

Proof. Fix \( y \in S(X) \) and \( x_0 \in U \). We will show that \( F_y(x_0) \) is at most countable (the proof that \( E_y(x_0) \) is countable is identical to this). Note that without loss of generality we may assume that \( F_y(x_0) \) is non-empty. So in this case, we define \( s : F_y(x_0) \to \mathbb{Q}^2 \) by \( s(t) \equiv (r_1, r_2) \) where \( r_1 \in (T_y(x_0 + ty), f^0(x_0 + ty; y)) \cap Q \) and \( r_2 \in (t, \infty) \cap Q \) is chosen so that

\[
\sup \{ f^0(x_0 + ry; y) : t < r < r_2 \} < r_1
\]

It is easy to see that \( s \) is 1-to-1 and so, \( F_y \) must be at most countable. Indeed, if \( t_1 < t_2 \) and \( s(t_1) = s(t_2) = (r_1, r_2) \) then \( t_1 < t_2 < r_2 \) and so \( f^0(x_0 + t_2y; y) < r_1 \). However this is impossible since it contradicts the definition of \( s \) at \( t_2 \). (here, \( Q \) denotes the rational numbers) \( \Box \)

Remark 3.2 If \( X \) is a Banach space, then for each \( y \in S(X), \{ x \in U : f \text{ is upper hemi-smooth, but not pseudo-regular}, \text{in the direction } y \} \text{ is contained in } F_y \) and hence is a Haar-null set (in this case, we may take the normalised Lebesgue measure, supported on \( sp\{y\} \), as a test-measure for the Borel set \( F_y \)).

Lemma 3.3 Let \( f \) be a real-valued locally Lipschitz function defined on a non-empty open subset \( U \) of a Banach space \( X \). If for some \( y \in X \setminus \{0\}, f \) is essentially upper hemi-smooth (essentially lower hemi-smooth) in the direction \( y \), then \( f^0(x; y) = f(x; y) \) (\( -f^0(x; -y) = f(x; y) \)) almost everywhere in \( U \).

Proof. Suppose that \( f \) is essentially upper hemi-smooth in the direction \( y \) (the proof for the case when \( f \) is essentially lower hemi-smooth in the direction \( y \), is similar). It follows from Remark 3.2 that \( f^0(x; y) = f^+(x; y) \) almost everywhere in \( U \). It now only remains to observe that \( f^+(x; y) = f(x; y) \) almost everywhere in \( U \) (to see this, take the normalised Lebesgue measure, supported on \( sp\{y\} \), as a test-measure for the set \( \{ x \in U : f(x; y) \text{ does not exist } \} \)). \( \Box \)

Now from Lemma 3.3 and Theorem 3.7 we may deduce the following theorem.

Theorem 3.8 Let \( f \) be a real-valued locally Lipschitz function defined on a non-empty open subset \( U \) of a Banach space \( X \). Let \( B \) be any subset of \( X \) such that \( spB = X \). If for each \( b \in B \) either, (i) \( f \) is essentially upper hemi-smooth in the directions \( b \) and \( -b \); (ii) \( f \) is essentially lower hemi-smooth in the directions \( b \) and \( -b \); (iii) \( f \) is essentially hemi-smooth in the direction \( b \) or (iv) \( f \) is essentially hemi-smooth in the direction \( -b \), then \( f \in \mathcal{S}_e(U) \).
It follows from the above theorem, that the class of essentially smooth functions is quite large, and as we shall see next, also, quite robust. More precisely, we will show that the essentially smooth functions are closed under addition, subtraction, multiplication and division (when this is defined) as well as the lattice operations. Actually, it not unreasonable to expect that the essentially smooth functions obey closure properties considerably stronger than those just mentioned. In fact, a first, but naive guess, might be that if \( f_1, f_2, \ldots, f_n \in S_c(U) \) and \( g \in S_c(R^n) \), then \( g \circ f \in S_c(U) \), where \( f \equiv (f_1, f_2, \ldots, f_n) \). To the contrary, in [5], the authors give an example that shows that this is not true, when \( n \geq 2 \).

For the final result of the paper we will need to consider vector-valued functions. Let \( U \) be a non-empty open subset of a Banach space \( X \) and let \( V \) be a non-empty open subset of \( R^n \). If \( x: U \to V \) is defined by,

\[
x(t) \equiv (x_1(t), x_2(t), \ldots, x_n(t)) \quad \text{with} \quad x_j: U \to R
\]

then we say that the vector-valued function \( x \) is essentially smooth on \( U \) if \( x_j \in S_c(U) \) for each \( 1 \leq j \leq n \) and in this case we write: \( x \in S_c(U; V) \). Further to this, we will say that a real-valued locally Lipschitz function \( f \) defined on a non-empty open subset \( V \) on \( R^n \) is arc-wise essentially smooth on \( V \), if for each locally Lipschitz function \( x \in S_c((0,1); V) \)

\[
\lambda\{t \in (0,1) : f^0(x(t); x'(t)) \neq -f^0(x(t); -x'(t))\} = 0
\]

[Here \( x'(t) \equiv (x'_1(t), x'_2(t), \ldots, x'_n(t)). \]

We shall denote by \( A_c(V) \), the family of all arc-wise essentially smooth Lipschitz functions on \( V \). A relevant fact for the next theorem is that \( A_c(V) \) contains all the pseudo-regular and semi-smooth (see [18] for definition) functions on \( V \). In particular, \( A_c(V) \) contains all the continuous convex and \( C^1 \) functions on \( V \) (see [5]).

**Theorem 3.9** ([5] Theorem 2.3) Let \( U \) be a non-empty open subset of a Banach space \( X \) and let \( f \equiv (f_1, f_2, \ldots, f_n) \in S_c(U; V) \), where \( V \) is any non-empty open subset of \( R^n \) that contains \( f(U) \). Then \( g \circ f \in S_c(U) \) whenever \( g \in A_c(V) \).

**Remark 3.8** In [4] it is shown that if \( U \) is a non-empty open subset of \( R \), then \( S_c(U) = A_c(U) \). However, it is also shown in [4] that if \( U \) is a non-empty open subset of \( R^n \) (\( n \geq 2 \)) then \( A_c(U) \) is a proper subset of \( S_c(U) \).
References


