Lipschitz functions with minimal Clarke subdifferential mappings

Jonathan M. Borwein

Centre for Experimental and Constructive Mathematics, Simon Fraser University, Burnaby, B.C., V5A 1S6 Canada

Warren B. Moors

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

Abstract

In this paper we characterise, in terms of the upper Dini derivative, when the Clarke subdifferential mapping of a real-valued locally Lipschitz function is a minimal weak* cusco. We then use this characterisation to deduce some new results concerning Lipschitz functions with minimal subdifferential mappings.

AMS Maths Subject Class (1991): 49J52

In the papers [7] and [1], the authors investigate a class of locally Lipschitz functions that possess ‘generic’ differentiability properties which are similar to those enjoyed by convex functions. This paper, continues this investigation.

Let $(X,||\cdot||)$ be a Banach space. We will call a Borel subset $N \subseteq X$ a Haar-null set if there exists a (not necessarily unique) Radon probability measure $p$ on $X$ such that $p(x + N) = 0$ for each $x \in X$. (In this case, we call $p$ a test-measure for $N$.)

More generally, we say that a subset $N \subseteq X$ is a Haar-null set if it is contained in a Borel Haar-null set. Below, we list some of the properties enjoyed by Haar-null sets.
Proposition 1 ([2], Proposition 2.1)
Let \((X, \| \cdot \|)\) be a Banach space.

(i) If \(M \subseteq N \subseteq X\) and \(N\) is a Haar-null set, then so is \(M\);
(ii) If \(N\) is a Haar-null set, then \(x + N\) is also a Haar-null set for each \(x \in X\);
(iii) If \(N \subseteq X\) is a Haar-null set, then \(X \setminus N\) is dense in \(X\);
(iv) If \(\{N_j \subseteq X : j \in N\}\) are Haar-null sets, then so is \(\bigcup \{N_j : j \in N\}\);
(v) In finite dimensions, the Haar-null sets and Lebesgue-null sets coincide.

A real-valued function \(f\) defined on a non-empty open subset \(A\) of a Banach space \(X\) is said to be locally Lipschitz on \(A\), if for each \(x_0 \in A\) there exists a \(K > 0\) and a \(\delta > 0\) such that

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

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

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

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

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

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

(iii) The Clarke generalized directional derivative at \(x \in A\), in the direction \(y\), given by

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

It is immediate from these 3 definitions that for each \(x \in A\) 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 \(x_0 \in A\), the mapping, \(y \rightarrow f^0(x_0; y)\) is a continuous sublinear functional on \(X\) (see, [4]).

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

\[
\partial f(x) \equiv \{x^* \in X^* : x^*(y) \leq f^0(x; y) \text{ for each } y \in X\}\quad \text{for each } x \in A
\]
With these preliminary definition behind us we may examine some notions of differentiability. We say that a real-valued locally Lipschitz function $f$ is differentiable at $x$, in the direction $y$, if

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

and we say that $f$ is Gateaux differentiable at $x$, if

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

for each $y \in X$ and $\nabla f(x)$ is a continuous linear functional on $X$.

In 1919 Rademacher ("Rademacher's Theorem") proved that real-valued locally Lipschitz functions defined on $R^n$ are Gateaux (and hence Fréchet) differentiable almost everywhere (in their domain). Using the terminology of Haar-null sets, we may reformulate Rademacher's Theorem in a more general form (see also, [3]).

**Theorem 1**

Let $Y$ be a closed separable subspace of a Banach space $X$. Let $f$ be a real-valued locally Lipschitz function defined on a non-empty open subset $A$ of $X$. Then there exists a Borel subset $D \subseteq A$ such that $A \setminus D$ is a Haar-null set and for each $x \in D$ the mapping $r_x : Y \to R$ defined by $r_x(y) \equiv f'(x; y)$ exists and is a continuous linear functional on $Y$.

**Proof.** Let $Y_1 \subseteq Y_2 \subseteq \ldots Y_n \subseteq \ldots$ be an increasing sequence of finite dimensional subspaces of $Y$ whose union is dense in $Y$. For each $n \in N$, let $D_n \subseteq A$ denote the set of all $x \in A$ such that $r_x(y)$ exists for all $y \in Y_n$ and depends linearly on $y \in Y_n$. It is reasonably straightforward to show that $D_n$ is a Borel set, and an application of Rademacher's Theorem on each translate of $Y_n$ shows that $A \setminus D_n$ is a Haar-null set (the normalised Lebesgue measure on $Y_n$ is a test-measure for $D_n$). If we put $D \equiv \bigcap \{D_n : n \in N\}$ then we may check that $D$ fulfills the conditions in the theorem. Finally, it follows from Proposition 1. part(iv) that $A \setminus D$ is a Haar-null set.

In order to obtain some 'generic' differentiability results we will need to consider a slightly stronger notion of differentiability. A locally Lipschitz function $f$ is said to be strictly differentiable at $x$, in the direction $y$, if

$$\lim_{\lambda \to 0^+} \frac{f(z + \lambda y) - f(z)}{\lambda} \text{ exists}$$

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,

$$f^0(x; y) = f'(x; y) = -f^0(x; -y)$$

This definition leads us to consider a particularly interesting class of Lipschitz functions, which do possess 'generic' differentiability properties.
A real-valued locally Lipschitz function \( f \) defined on a non-empty open subset \( A \) of a Banach space \( X \) is called \textit{essentially smooth} on \( A \) if for each \( y \in X \setminus \{0\} \)

\[
\{ x \in A : f^0(x; y) \neq -f^0(x; -y) \}
\]

is a Haar-null subset of \( X \).

We shall denote by \( S_e(A) \) the family of all real-valued essentially smooth locally Lipschitz functions on \( A \).

That the essentially smooth Lipschitz functions possess good generic differentiability properties follows from the fact that their Clarke subdifferential mappings are minimal weak* cuscos, a subject we now discuss.

A set-valued mapping \( \Phi \) from a topological space \( A \) into subsets of a linear topological space \((Y, \tau)\) is called a \( \tau \)-\textit{cusco} on \( A \) if:

(i) for each \( x \in A \), \( \Phi(x) \) is non-empty, convex and compact;

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

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

It is well-known that the Clarke subdifferential mapping of a real-valued locally Lipschitz function defined on a non-empty open subset of a Banach space \( X \) is a weak* cusco, with images in \( X^* \).

Amongst the class of cusco mappings, special attention is given to the class of minimal cuscos. A cusco mapping \( \Phi \) from a topological space \( A \) into subsets of a linear topological space \((Y, \tau)\) is called a \textit{minimal \( \tau \)-cusco} if its graph does not strictly contain the graph of any other \( \tau \)-cusco defined on \( A \).

There is a weak form of continuity for real-valued functions that is closely related to the theory of minimal cuscos. Let \( g \) be a real-valued function defined on a topological space \( A \). Then we say that \( g \) is \textit{hyperplane minimal} on \( A \) if for each open ray \( I \subseteq R \) and open set \( U \subseteq A \) with \( g(U) \cap I \neq \emptyset \), there exists a non-empty open subset \( V \) of \( U \) with \( g(V) \subseteq I \).

This definition leads us further, to consider the following.

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 \textit{essentially minimal} on \( A \) if for each \( y \in X \setminus \{0\} \) there exists a Haar-null set \( N \subseteq X \) such that the mapping, \( x \to f^+(x; y) \), defined on \( A \setminus N \) is hyperplane minimal on \( A \setminus N \).

The main goal of this paper is to show that if \( f \) is essentially minimal, then the Clarke subdifferential mapping of \( f \) is a minimal weak* cusco. To achieve this goal we will need one more preliminary result.

\textbf{Theorem 2} ([2], Theorem 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 for each \( x \in A \), \( y \in X \setminus \{0\} \) and Haar-null set \( N \subseteq X \), we have that,
\[ f^0(x; y) = \limsup_{z \to x} f^+(z; y) - f^0(x; -y) = \liminf_{z \to x} f^+(z; y) \]

**Theorem 3 (Main Theorem)**

Let \( f \) be a real-valued locally Lipschitz function defined on a non-empty open subset \( A \) of a Banach space \( X \). Then the Clarke subdifferential mapping of \( f \) is a minimal weak* cusco on \( A \) if, and only if, \( f \) is essentially minimal on \( A \).

**Proof.** It follows directly from Theorem 2.16 in [7] that if the Clarke subdifferential mapping, \( x \to \partial f(x) \), is a minimal weak* cusco on \( A \), then \( f \) is essentially minimal on \( A \).

Conversely, let us suppose, for the purpose of obtaining a contradiction, that \( f \) is essentially minimal on \( A \) and \( x \to \partial f(x) \), is not a minimal weak* cusco on \( A \). That is, let us suppose that there exists a weak* cusco \( \Phi \) on \( A \) whose graph is strictly contained in that of \( \partial f \). It follows then, via a separation argument in \((X^*, \text{weak}^*)\), that we may find a point \( x_0 \in A \) and a direction \( y \in X \setminus \{0\} \) such that

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

Now consider the two cusco mappings \( T_y : A \to 2^R \) and \( S_y : A \to 2^R \) 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 \( x \in A \). 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 \( N \) be any Haar-null set of \( X \) such that the mapping, \( x \to f^+(x; y) \), is hyperplane minimal on \( A \setminus N \). It is clear, from the definition of \( T_y \), that

\[ T_y(x) = [-f^0(x; -y), f^0(x; y)] \]

at each point of \( A \).

We may now use the upper semi-continuity of \( S_y \), a simple separation argument (in \( R \)), along with Theorem 2 and the hyperplane minimality of, \( x \to \partial f(x) \), on \( A \setminus N \) to deduce that \( S_y = T_y \). \( \Box \)

We will finish this paper with some applications of Theorem 3. For our first application we need some additional (but easily verified) facts concerning hyperplane minimal mappings.

**Proposition 2** ([6], Theorem 1.2)

Let \( f \) and \( g \) be real-valued functions defined on a topological space \( A \). If \( f \) is hyperplane minimal and locally bounded on \( A \) and \( g \) is continuous on \( A \), then both \( f + g \) and \( f \cdot g \) are hyperplane minimal on \( A \).
Corollary 1

Let $f$ and $g$ be real-valued locally Lipschitz functions defined on a non-empty open subset $A$ of a Banach space $X$. If $g \in S_c(A)$ and $x \rightarrow \partial f(x)$, is a minimal weak$^*$ cuso on $A$ then:

(i) $x \rightarrow \partial(f + g)(x)$ and $x \rightarrow \partial(g + f)(x)$ are minimal weak$^*$ cuscors on $A$;

(ii) $x \rightarrow \partial(f \cdot g)(x)$ is a minimal weak$^*$ cuso on $A$;

(iii) $x \rightarrow \partial(g + f)(x)$ is a minimal weak$^*$ cuso on $A \setminus f^{-1}(0)$;

(iv) $x \rightarrow \partial(f + g)(x)$ is a minimal weak$^*$ cuso on $A \setminus g^{-1}(0)$;

(v) $x \rightarrow \partial(\max\{f, g\})(x)$ and $x \rightarrow \partial(\min\{f, g\})(x)$ are minimal weak$^*$ cuscors on $A$.

Proof. The proof of (i) follows from the fact that for each $y \in X$, $(f + g)'(x; y) = f'(x; y) + g'(x; y)$ almost everywhere in $A$ and, $x \rightarrow g^+(x; y)$, is continuous at each point of $A(y) \equiv \{x \in A : g^0(x; y) = -g^0(x; y)\}$. The proof of (ii), (iii) and (iv) are similar to that of (i).

(v) $\max\{f(x), g(x)\} = (f - g)^+(x) + g(x)$ 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) in [1], $(f - g)^+$ and $(f - g)^-$ possess minimal subdifferential mappings. The proof is completed by again appealing to part(i) above.

Remark 1

The conclusions of Corollary 1 fail (in general) if we only assume that both $x \rightarrow \partial f(x)$ and $x \rightarrow \partial g(x)$ are minimal weak$^*$ cuscors on $A$ (see, [1], Example 2.1).

Corollary 2

Let $f$ be a real-valued locally Lipschitz function defined on a non-empty open subset $A$ of the product $X \times Y$, of Banach spaces $X$ and $Y$. If for each $x_0 \in X$, the mapping $(x, y) \rightarrow f^+((x, y); (x_0, 0))$ is continuous almost everywhere in $A$ and for each $y_0 \in Y$, the mapping $(x, y) \rightarrow f^+((x, y); (0, y_0))$ is hyperplane minimal on $A$, then $(x, y) \rightarrow \partial f(x, y)$ is a minimal weak$^*$ cuso on $A$.

Proof. By Theorem 2 it is sufficient to show that $f$ is essentially minimal on $A$. To this end, let $(x_0, y_0) \in X \times Y$. Now, by Theorem 1,

$$f'(((x, y); (x_0, y_0)) = f'((x, y); (x_0, 0)) + f'((x, y); (0, y_0))$$

almost everywhere in $A$. Hence, by the above argument and Proposition 2 we have that there exists a Haar-null set $N \subseteq X$ such that the restriction of $(x, y) \rightarrow f^+((x, y); (x_0, y_0))$ to $A \setminus N$, is hyperplane minimal on $A \setminus N$.

Remark 2

Using the terminology from [5] we can easily verify that the hypothesis of Corollary 2 are satisfied for u-u regular functions; u-l regular functions and l-l regular functions on $A$. Also,
Lipschitz functions with minimal Clarke subdifferential mappings

one can check that if \( f \in S_c(X) \), \( g : Y \rightarrow R \) possesses a minimal subdifferential mapping and \( h : \mathbb{R}^2 \rightarrow \mathbb{R} \) is \( C^1 \) then the mapping \( (x, y) \rightarrow h(f(x), g(y)) \) satisfies the hypothesis of Corollary 2.

A comprehensive discussion of the consequences for \( f \) of \( \partial f \) being minimal may be found in [1].

References


