Pergamon

Nonlinear Analysis, Theory, Methods & Applications, Vol. 29, No. 1, pp. 53–63, 1997

0362-546X/97 $17.00 + 0.00

PII: S0362-546X(96)00050-8

LIPSCHITZ FUNCTIONS WITH PRESCRIBED DERIVATIVES AND SUBDERIVATIVES

JONATHAN M. BORWEIN†, WARREN B. MOORS‡ and XIANFU WANG
Simon Fraser University, CECM, Burnaby, BC, Canada V5A 1S6

(Received 31 July 1995; received for publication 28 March 1996)

Key words and phrases: Lipschitz functions, differentiability, Clarke subgradient, minimal cusco, Haar-null set, maximal cyclically monotone operator.

1. INTRODUCTION

It is natural to ask when a given set-valued mapping T, which maps from a nonempty open subset U of a Banach space X into subsets of its dual, is the Clarke subdifferential mapping of some real-valued locally Lipschitz functions defined on U. In the case when $X = R$ and U is an open interval the answer is known (see [1]). However, the general question still remains. Even the simpler question of how to construct nontrivial Lipschitz functions which are not built-up from either convex or distance functions has yet to be satisfactorily resolved. In this paper we present a technique for constructing such real-valued locally Lipschitz functions defined on separable Banach spaces. Using this construction we are able to recreate many known examples of pathological locally Lipschitz functions. For example, we can show that given any polytope $P \subset R^n$ there exists a real-valued globally Lipschitz function g, defined on R^n, such that the Clarke subdifferential, $x \to \partial g(x)$, of g is identically equal to P. This example extends the main result of [2], which in turn generalizes an example given in [3].

As another special case of our construction we will see that given any finite family $\{T_1, T_2, \ldots, T_n\}$ of maximal cyclically monotone operators defined on a separable Banach space X, there exists a real-valued locally Lipschitz function g appropriately defined on X such that

$$\partial g(x) = \mathrm{co}\{T_1(x), T_2(x), \ldots, T_n(x)\} \qquad \text{for each } x \in X.$$

We begin with some preliminary definitions. A real-valued function f defined on a nonempty open subset A of a Banach space X is *locally Lipschitz* on A if for each $x_0 \in A$ there exists an $M > 0$ and a $\delta > 0$ such that

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

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

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

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

† Research supported by NSERC and the Shrum Endowment at Simon Fraser University.
‡ Research supported by a Post-Doctoral Fellowship at CECM Simon Fraser University.

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

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

Associated with the Clarke generalized directional derivative is the *Clarke subdifferential* (*or subgradient*) mapping, which is defined by

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

Distinct from the notion of a directional derivative is that of a derivative. We say a function f is differentiable at x in the direction y if

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

We say that f is *Gâteaux 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 are also interested in a stronger notion of differentiability. A locally Lipschitz function f is said to be *strictly* (*Gâteaux*) *differentiable* at x if for each $y \in X$ and each $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\left| \frac{f(z + \lambda y) - f(z)}{\lambda} - \nabla f(x)(y) \right| < \varepsilon \qquad \text{whenever } 0 < \lambda < \delta,\ \|z - x\| < \delta.$$

Apart from the various notions of differentiability (all of which are discussed in [4]), the other key concept we need is that of a minimal cusco. A set-valued mapping Φ from a topological space A into subsets of a topological space X is *upper semi-continuous* on A if for each open subset $W \subset X$, $\{t \in A : \Phi(t) \subset W\}$ is an open subset of A. If in addition, Φ has the property that for each $t \in A$, $\Phi(t)$ is nonempty and compact (convex), then Φ is an *usco* (*cusco*) on A. 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. We say that a locally Lipschitz function has a *minimal* Clarke subdifferential mapping if its Clarke subdifferential mapping is a minimal weak* cusco. Details may be found in [5].

2. CONSTRUCTION OF LOCALLY 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 (see [6]) 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 $x \in X$. In finite dimensions, the Haar-null sets coincide with the universally measurable Lebesgue null sets. In general, however, if N is a Haar-null set then $X \backslash N$ is dense in X. The Haar-null sets are also closed under translation and countable unions.

Lemma 1. Let $g, f_1, f_2, \ldots, f_n$ be real-valued locally Lipschitz functions defined on a nonempty open subset U of a separable Banach space X. If $\nabla g(x) \in \{\nabla f_1(x), \nabla f_2(x), \ldots, \nabla f_n(x)\}$ almost everywhere in U, then $\partial g(x) \subset \mathrm{co}\{\partial f_1(x), \partial f_2(x), \ldots, \partial f_n(x)\}$ for all $x \in U$.

Proof. Consider the set-valued mapping $T: U \to 2^{X^*}$ defined by $T(x) \equiv \bigcup\{\partial f_j(x) : 1 \le j \le n\}$. Clearly T is a weak* usco mapping on U, hence by Lemma 7.12 in [7] the mapping $T^*: U \to 2^{X^*}$ defined by

$$T^*(x) = \mathrm{co}\{\partial f_1(x), \partial f_2(x), \ldots, \partial f_n(x)\}$$

is a weak* cusco on U. Now from the hypothesis we have that $\nabla g(x) \in T^*(x)$ almost everywhere in U. Therefore by [8] $\partial g(x) \subset T^*(x)$ for all $x \in U$. ∎

In the proof of Theorem 1 we use the following well-known result, whose proof we include for the sake of completeness.

Lemma 2. Let $\{G_n : n \in \mathbf{N}\}$ be a family of Lebesgue measurable subsets of R. If for each $n \in \mathbf{N}$, G_n has positive measure. Then there exists a subset $E \equiv \bigcup\{E_n : n \in \mathbf{N}\}$ of R such that:
(i) Each set E_n is compact;
(ii) For each $n \in \mathbf{N}$, $\mu(G_n \cap E) > 0$ and $\mu(G_n \backslash E) > 0$.

Proof. We first construct disjoint compact sets $\{C_j : j \in \mathbf{N}\}$ with $\mu(C_j) > 0$ and $C_j \subset G_j$ for each j. We proceed by induction.

Step 1. By the regularity of Lebesgue measure we may choose a compact subset $C_1^1 \subset G_1$ with $\mu(C_1^1) > 0$. Define $m_1 \equiv \mu(C_1^1)$.

Suppose the first n steps have been completed. Then we will have constructed positive real numbers $\{m_1, m_2, \ldots, m_n\}$ and disjoint compact subsets $\{C_1^j, C_2^j, \ldots, C_j^j\}$ of R for $1 \le j \le n$ such that $m_j \equiv \mu(C_j^j)$ and $C_j^k \subset C_j^j \subset G_j$ for $1 \le j \le k \le n$.

Step $n+1$. Choose a compact set $C \subset G_{n+1}$ with $\mu(C) > 0$, and cover C with a finite number of intervals $\{I_1, I_2, \ldots, I_m\}$ of diameter at most r, where $0 < r < (1/2^{n+1}) \min\{m_1, m_2, \ldots, m_n\}$. Now for some k, $1 \le k \le m$, we have $\mu(I_k \cap C) > 0$. Select a compact subset $C_{n+1}^{n+1} \subset I_k \cap C$ with $\mu(C_{n+1}^{n+1}) > 0$. Define $m_{n+1} \equiv \mu(C_{n+1}^{n+1})$ and $C_j^{n+1} \equiv C_j^n \backslash I_k$ for $1 \le j \le n$. This ends the induction.

Let $C_j \equiv \bigcap_{n=j}^{\infty} C_j^n$ for each $j \in \mathbf{N}$. Then C_j is compact and $C_j \subset G_j$. By the selection of the I_k's we see that $\mu(C_j) > (1 - 1/2^j) \cdot \mu(C_j^j) \ge m_j/2 > 0$ for $j \ge 1$, and by the overall construction, we see that the sets $\{C_j : j \in \mathbf{N}\}$ are pairwise disjoint.

Next, we construct the set E. For each j, choose a compact set $E_j \subset C_j$ such that $0 < \mu(E_j) < \mu(C_j)$ and define $E \equiv \bigcup_{j=1}^{\infty} E_j$. Then for each j we have

$$\mu(E \cap G_j) \ge \mu(E \cap C_j) = \mu(E_j) > 0$$

and

$$\mu(G_j \backslash E) \ge \mu(C_j \backslash E) = \mu(C_j) - \mu(E_j) > 0. \quad ∎$$

LEMMA 3. Let f be a real-valued locally Lipschitz function defined on a nonempty open subset A of a separable Banach space X. Let λ be a real-valued locally Lipschitz function defined on R. Then for any real-valued Borel measurable function λ^* defined on R such that $\lambda^* = \lambda'$ almost everywhere

$$A\setminus\{x \in A : \nabla(\lambda \circ f)(x) = \lambda^*(f(x)) \cdot \nabla f(x)\}$$

is a Haar-null set.

Proof. Let $E \equiv \{x \in A : \nabla(\lambda \circ f)(x) \text{ exists}\}$ and let $F \equiv \{x \in A : \nabla f(x) \text{ exists}\}$. By [6] we have that both $A\setminus E$ and $A\setminus F$ are Haar-null sets. For each y in the unit sphere $S(X)$ of X, let $D_y \equiv \{x \in A : (\lambda \circ f)^+(x; y) = \lambda^*(f(x)) \cdot f^+(x; y)\}$. Clearly D_y is a Borel subset of A, since each of the mappings $x \to (\lambda \circ f)^+(x; y)$ and $x \to \lambda^*(f(x)) \cdot f^+(x; y)$ are Borel measurable.

Let $\{y_n : n \in \mathbf{N}\}$ be a dense subset of $S(X)$ and $T \equiv E \cap F \cap (\bigcap\{D_{y_n} : n \in \mathbf{N}\})$. We claim that $A\setminus T$ is a Haar-null set. To prove this, it is sufficient to show that each subset $A\setminus D_{y_n}$ is a Haar-null set. To accomplish this, we define a probability measure P_n on the universally measurable subsets of X as follows:

$$P_n(M) \equiv \frac{1}{\sqrt{2\pi}} \int_{M^*} \exp(-x^2/2)\, \mathrm{d}x \qquad \text{where } M^* \equiv \{t \in R : ty_n \in M\}.$$

Take any $x_0 \in X$ and let $U = \{t \in R : ty_n - x_0 \in A\}$. If $U = \varnothing$, then $P(x_0 + A\setminus D_{y_n}) = 0$. If $U \neq \varnothing$, then define $g: U \to R$ by $g(t) \equiv \lambda(f(ty_n - x_0))$. Clearly U is open and g is locally Lipschitz on U. Therefore by Theorem 6.93 in [9]

$$\mu(U\setminus\{t \in U: g'(t) = \lambda^*(f(ty_n - x_0)) \cdot f'(ty_n - x_0; y_n)\}) = 0.$$

Now observe that $s \in \{t \in U : g'(t) = \lambda^*(f(ty_n - x_0)) \cdot f'(ty_n - x_0; y_n)\}$ if and only if $sy_n - x_0 \in D_{y_n}$, that is, if and only if $sy_n \in x_0 + D_{y_n}$. Therefore

$$\{t \in R : ty_n \in x_0 + A\setminus D_{y_n}\} = U\setminus\{t \in U: g'(t) = \lambda^*(f(ty_n - x_0)) \cdot f'(ty_n - x_0; y_n)\}$$

and so $P_n(x_0 + A\setminus D_{y_n}) = 0$. Hence we may conclude that $A\setminus D_{y_n}$ is a Haar-null set. It now only remains to observe that $\nabla(\lambda \circ f)(x) = \lambda^*(f(x)) \cdot \nabla f(x)$ if and only if $x \in T$. In fact, for $x \in T$, we know $\nabla(\lambda \circ f)(x)$ and $\nabla f(x)$ exist, and $\nabla(\lambda \circ f)(x)(y_n) = \lambda^*(f(x)) \cdot \nabla f(x)(y_n)$ for all $n \in \mathbf{N}$. Since $\{y_n : n \in \mathbf{N}\}$ is a dense subset of $S(X)$ and each side is continuous in y, we have that

$$\nabla(\lambda \circ f)(x)(y) = \lambda^*(f(x)) \cdot \nabla f(x)(y) \qquad \text{for all } y \in X.$$

This ends the proof. ∎

The next lemma is a special case of Lemma 6.92 in [9].

LEMMA 4. Suppose that f is an absolutely continuous real-valued function defined on an open interval (a, b) of R. If E is a Lebesgue measurable subset of (a, b) and $\mu(f(E)) = 0$, then $f' = 0$ almost everywhere in E.

We may now establish our main result.

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

Proof. The proof is presented in two parts.

Part I. Let $\{y_n : n \in \mathbf{N}\}$ be a dense subset of $S(X)$ and let $\{x_n : n \in \mathbf{N}\}$ be a dense subset of U. In this part we show that given any finite family of real-valued locally Lipschitz functions $\{h_1, h_2, \ldots, h_j\}$ defined on U, there exists a real-valued locally Lipschitz function g defined on U such that

(a_j) $\partial g(x) \subset \text{co}\{\partial h_1(x), \partial h_2(x), \ldots, \partial h_j(x)\}$ for each $x \in U$ and $\nabla g(x) \in \{\nabla h_1(x), \nabla h_2(x), \ldots, \nabla h_j(x)\}$ almost everywhere in U, i.e. everywhere in U except possibly on a Haar-null set.

(b_j) For each $1 \leq k \leq j$ and $n, p \in \mathbf{N}$, the subsets $M_j(n,p,k) \subset R$ defined by $M_j(n,p,k) \equiv \{t \in R : g'(x_n + ty_p; y_p) = h_k'(x_n + ty_p; y_p)\}$ have positive measures in every open interval in $\{t \in R : x_n + ty_p \in U\}$.

We proceed by induction.

Step 1. Let h_1 be any real-valued locally Lipschitz function defined on U and $g \equiv h_1$. Then clearly g satisfies (a_1) and (b_1) with respect to the locally Lipschitz function h_1.

Suppose the first m steps of the induction have been completed. That is, suppose that given any m locally Lipschitz functions $k_1, k_2, \ldots, k_m$ defined on U, there exists a locally Lipschitz function g defined on U such that (a_m) and (b_m) are satisfied with respect to the functions $k_1, k_2, \ldots, k_m$.

Step m + 1. Let $h_1, h_2, \ldots, h_{m+1}$ be real-valued locally Lipschitz functions defined on U. For each $1 \leq i \leq m$, define $c_i: U \to R$ by $c_i \equiv h_i - h_{m+1}$ and $c_{m+1}: U \to R$ by $c_{m+1} \equiv 0$.

By the induction hypothesis, there exists a real-valued locally Lipschitz function g defined on U such that g satisfies (a_m) and (b_m) with respect to the locally Lipschitz functions $c_1, c_2, \ldots, c_m$.

For each $n, p \in \mathbf{N}$, let $\{U_r(n,p) : r \in \mathbf{N}\}$ be a family of bounded open intervals in $\{t \in R : x_n + ty_p \in U\}$, which form a topological base for the relative topology on $\{t \in R : x_n + ty_p \in U\}$. Note that without loss of generality, we may assume that for each $n, p, r \in \mathbf{N}$, $\overline{U_r(n,p)} \subset \{t \in R : x_n + ty_p \in U\}$.

For each $1 \leq k \leq m$ and each $n, p, r \in \mathbf{N}$, let

$$G(n,p,r,k) \equiv g(\{x_n + ty_p \in U : t \in U_r(n,p) \cap M_m(n,p,k)\}).$$

Let us also set

$$\begin{aligned} G &\equiv \{G_n : n \in \mathbf{N}\} \\ &\equiv \{G(n,p,r,k) : 1 \leq k \leq m, n, p, r \in \mathbf{N} \text{ and } \mu(G(n,p,r,k)) > 0\}. \end{aligned}$$

Here μ is the Lebesgue measure on R. Let $E \equiv \bigcup \{E_n : n \in \mathbf{N}\}$ be the subset of R given in Lemma 2 associated with the family of sets G. Define $g_{m+1}: U \to R$ by $g_{m+1}(x) \equiv \lambda_E(g(x))$ where

$$\lambda_E(t) \equiv \int_0^t \chi_E(s)\,\mathrm{d}s \qquad \text{and} \qquad \chi_E(s) = \begin{cases} 1 & \text{if } s \in E \\ 0 & \text{otherwise.} \end{cases}$$

Clearly g_{m+1} is real-valued and locally Lipschitz on U. We claim that g_{m+1} satisfies (a_{m+1}) and (b_{m+1}) with respect to the locally Lipschitz functions $c_1, c_2, \ldots, c_{m+1}$ defined on U.

To see that g_{m+1} satisfies (a_{m+1}). Observe that since E is a Borel subset of R, we may apply Lemma 3 to get that $\nabla g_{m+1}(x) = \chi_E(g(x)) \cdot \nabla g(x)$ almost everywhere in U. Now by assumption $\nabla g(x) \in \{\nabla c_1(x), \nabla c_2(x), \ldots, \nabla c_m(x)\}$ almost everywhere in U. Therefore,

$$\nabla g_{m+1}(x) \in \{\nabla c_1(x), \nabla c_2(x), \ldots, \nabla c_m(x), \nabla c_{m+1}(x)\}$$

almost everywhere in U. Furthermore, by Lemma 1 we also have that

$$\partial g_{m+1}(x) \subset \mathrm{co}\{\partial c_1(x), \partial c_2(x), \ldots, \partial c_{m+1}(x)\} \qquad \text{for each } x \in U.$$

Next we show that g_{m+1} satisfies (b_{m+1}). To this end, fix $1 \le k \le m$ and $n, p \in \mathbf{N}$. Also fix $r \in \mathbf{N}$, corresponding to the open interval $U_r(n,p)$. We consider two cases:

(i) Suppose $\mu(G(n,p,r,k)) > 0$. Then by the construction of the set E given in Lemma 2, $\mu(G(n,p,r,k) \cap E) > 0$. Let

$$A \equiv \{t \in M_m(n,p,k) \cap U_r(n,p) : g(x_n + ty_p) \in E\}.$$

Since the mapping $t \to g(x_n + ty_p)$ is absolutely continuous on $U_r(n,p)$, $\mu(A) > 0$. (Actually, the mapping $t \to g(x_n + ty_p)$ is Lipschitz on $U_r(n,p)$.) Therefore by Lemma 3

$$g'_{m+1}(x_n + ty_p; y_p) = \chi_E(g(x_n + ty_p)) \cdot g'(x_n + ty_p; y_p) = c'_k(x_n + ty_p; y_p)$$

for almost all $t \in A$. Hence, $\mu(M_{m+1}(n,p,k) \cap U_r(n,p)) > 0$.

(ii) Suppose that $\mu(G(n,p,r,k)) = 0$. Then by Lemma 4, $g'(x_n + ty_p; y_p) = 0$ for almost all $t \in M_m(n,p,k) \cap U_r(n,p)$ and so by Lemma 3, we have that

$$g'_{m+1}(x_n + ty_p; y_p) = \chi_E(g(x_n + ty_p)) \cdot g'(x_n + ty_p; y_p) = 0$$

for almost all $t \in M_m(n,p,k) \cap U_r(n,p)$. From this, it follows that

$$g'_{m+1}(x_n + ty_p; y_p) = g'(x_n + ty_p; y_p) = c'_k(x_n + ty_p; y_p) = 0$$

for almost all $t \in M_m(n,p,k) \cap U_r(n,p)$; so $\mu(M_{m+1}(n,p,k) \cap U_r(n,p)) > 0$.

We now show that $\mu(M_{m+1}(n,p,m+1) \cap U_r(n,p)) > 0$ for each $n, p, r \in \mathbf{N}$. Again, fix $n, p, r \in \mathbf{N}$ and consider the set $G(n,p,r,1)$ (in fact, it suffices to consider any one of the sets $G(n,p,r,k)$ with $1 \le k \le m$). We examine two more cases:

(iii) Suppose that $\mu(G(n,p,r,1)) = 0$. Then by Lemma 4, $g'(x_n + ty_p; y_p) = 0$ for almost all $t \in M_m(n,p,1) \cap U_r(n,p)$ and so by Lemma 3

$$\begin{aligned} g'_{m+1}(x_n + ty_p; y_p) &= \chi_E(g(x_n + ty_p)) \cdot g'(x_n + ty_p; y_p) \\ &= 0 = c'_{m+1}(x_n + ty_p; y_p) \end{aligned}$$

for almost all $t \in M_m(n,p,1) \cap U_r(n,p)$. Hence $\mu(M_{m+1}(n,p,m+1) \cap U_r(n,p)) > 0$.

(iv) Suppose that $\mu(G(n,p,r,1)\backslash E) > 0$. Let

$$A \equiv \{t \in M_m(n,p,1) \cap U_r(n,p) : g(x_n + ty_p) \in G(n,p,r,1)\backslash E\}.$$

Since the mapping $t \to g(x_n + ty_p)$ is absolutely continuous on $U_r(n,p)$, $\mu(A) > 0$. Now by Lemma 3

$$\begin{aligned} g'_{m+1}(x_n + ty_p; y_p) &= \chi_E(g(x_n + ty_p)) \cdot g'(x_n + ty_p; y_p) \\ &= 0 = c'_{m+1}(x_n + ty_p; y_p) \end{aligned}$$

for almost all $t \in A$. Therefore $\mu(M_{m+1}(n,p,m+1) \cap U_r(n,p)) > 0$.

At this stage, we have shown that g_{m+1} satisfies (a_{m+1}) and (b_{m+1}) with respect to the locally Lipschitz functions $c_1, c_2, \ldots, c_m, c_{m+1}$.

Define $e: U \to R$ by $e(x) \equiv g_{m+1}(x) + h_{m+1}(x)$. It is clear that $\nabla e(x) = \nabla g_{m+1}(x) + \nabla h_{m+1}(x)$ almost everywhere in U. Hence by the above argument

$$\nabla e(x) \in \{\nabla c_1(x) + \nabla h_{m+1}(x), \ldots, \nabla c_m(x) + \nabla h_{m+1}(x), \nabla c_{m+1}(x) + \nabla h_{m+1}(x)\}$$

almost everywhere in U. In addition to this, we note that for each $1 \le i \le m$, $\nabla c_i(x) = \nabla h_i(x) - \nabla h_{m+1}(x)$ almost everywhere in U. Thus

$$\nabla e(x) \in \{\nabla h_1(x), \nabla h_2(x), \ldots, \nabla h_{m+1}(x)\}$$

almost everywhere in U. Now by Lemma 1 we also have

$$\partial e(x) \subset \mathrm{co}\{\partial h_1(x), \partial h_2(x), \ldots, \partial h_{m+1}(x)\}$$

for each $x \in U$. Further to this, for each $n, p \in \mathbf{N}$

$$e'(x_n + ty_p; y_p) = g'(x_n + ty_p; y_p) + h'_{m+1}(x_n + ty_p; y_p)$$

for almost all $t \in \{t \in R : x_n + ty_p \in U\}$. It now follows that the function e satisfies (a_{m+1}) and (b_{m+1}) with respect to the locally Lipschitz functions $h_1, h_2, \ldots, h_{m+1}$ defined on U; which completes the induction.

Part II. Let $f_1, f_2, \ldots, f_n$ be real-valued locally Lipschitz functions defined on U whose Clarke subdifferential mappings are minimal. By *Part I* we know that there exists a real-valued locally Lipschitz function defined on U which satisfies (a_n) and (b_n) with respect to the locally Lipschitz functions $f_1, f_2, \ldots, f_n$. In this part we shall show that for this function g

$$\partial g(x) = \mathrm{co}\{\partial f_1(x), \partial f_2(x), \ldots, \partial f_n(x)\} \qquad \text{for each } x \in U.$$

Of course, it suffices from *Part I* to show that

$$\mathrm{co}\{\partial f_1(x), \partial f_2(x), \ldots, \partial f_n(x)\} \subset \partial g(x) \qquad \text{for each } x \in U.$$

Since each function f_j, $1 \le j \le n$, possesses a minimal subdifferential mapping, there exists a dense G_δ subset, G_j of U such that f_j is strictly differentiable at each point of G_j [5]. Let $G \equiv \bigcap \{G_j : 1 \le j \le n\}$. Clearly G is a dense and G_δ subset of U. We show that $\{\nabla f_1(x), \nabla f_2(x), \ldots, \nabla f_n(x)\} \subset \partial g(x)$ for each $x \in G$.

Suppose that this is not the case. Then there exists an element $x_0 \in G$, $y \in S(X)$, $\alpha \in R$ and $1 \leq j \leq n$ such that

$$\langle \nabla f_j(x_0), y \rangle > \alpha > \max_{\xi \in \partial g(x_0)} \langle y, \xi \rangle = g^0(x_0; y).$$

Moreover, since $\partial g(x_0)$ is a bounded subset of X^* we may choose $y = y_p \in \{y_n : n \in \mathbf{N}\}$. By the definitions of $g^0(x_0; y)$ and strict differentiability, we know that there exists an open neighborhood V of x_0 contained in U such that $g^+(z; y_p) < \alpha$ for all $z \in V$ and $f^+(z; y_p) > \alpha$ for all $z \in V$. Choose $x_m \in \{x_k : k \in \mathbf{N}\} \cap V$ and $U_r(m,p)$ such that $\{x_m + ty_p : t \in U_r(m,p)\} \subset V$. Now consider a point $x_m + ty_p$, where $t \in M_n(m,p,j) \cap U_r(m,p)$. Then

$$\alpha < f_j'(x_m + ty_p; y_p) = g'(x_m + ty_p; y_p) < \alpha \qquad \text{which is a contradiction.}$$

Therefore $\{\nabla f_1(x), \nabla f_2(x), \ldots, \nabla f_n(x)\} \subset \partial g(x)$ for each $x \in G$.

Next, we fix $1 \leq k \leq n$ and consider the function f_k. By the weak* upper semi-continuity of the mappings $x \to \partial g(x)$ and $x \to \partial f_k(x)$ on U, we see that $\partial g(x) \cap \partial f_k(x) \neq \varnothing$ for all $x \in U$. Define $T: U \to 2^{X^*}$ by $T(x) \equiv \partial g(x) \cap \partial f_k(x)$. Clearly $T(x)$ is nonempty, weak* compact and convex for each $x \in U$. Furthermore, it is not difficult to show, using a standard net argument, that T is weak* upper semi-continuous on U (see Proposition 1.3 [10]). (Actually, since T is locally bounded and the weak* topology on X^* is metrizable on bounded subsets, it is sufficient to use a sequential argument.)

However, since $x \to \partial f_k(x)$ is a minimal weak* cusco on U and $T(x) \subset \partial f_k(x)$ for each $x \in U$, we must have that $T \equiv \partial f_k$. Hence, for each $x \in U$, $\partial f_k(x) \subset \partial g(x)$. Since k, $1 \leq k \leq n$, was arbitrary and $\partial g(x)$ is convex, we must have that $\text{co}\{\partial f_1(x), \partial f_2(x), \ldots, \partial f_n(x)\} \subset \partial g(x)$ for each $x \in U$. ■

We note that the constructed function has a minimal subgradient if, and only if, each of the underlying functions shares the same minimal subgradient. We also note that each of the underlying functions being minimal is generically (on a dense G_δ) strictly differentiable. Thus the constructed function has polyhedral subgradient images generically though perhaps only on a set small in measure.

Remark 1. With some more work, but not significantly more, it can be shown that the function g constructed in the previous theorem is not unique up to a constant, except possibly when $\partial f_1 = \partial f_2 = \cdots = \partial f_n$.

The important observations required to prove this are:

(i) $\mu(M_{m+1}(n,p,k) \cap U_r(n,p) \cap M_m(n,p,k)) > 0$ for all $n, p, r \in \mathbf{N}$ and $1 \leq k \leq m$.

(ii) $\mu(M_{m+1}(n,p,m+1) \cap U_r(n,p) \cap M_m(n,p,k)) > 0$ for all $n, p, r \in \mathbf{N}$ and $1 \leq k \leq m$.

(iii) If $\text{co}\{\partial f_1(x), \ldots, \partial f_{n-1}(x)\} = \partial g(x)$ for all $x \in U$, then $\partial g_{n-1} = \partial g$ and $g_{n-1} - g$ is not a constant unless $\partial f_1(x) = \cdots = \partial f_n(x)$ for all x.

(iv) If $\text{co}\{\partial f_1(x), \ldots, \partial f_{n-1}(x)\} \neq \partial g(x)$ for some $x \in U$, then consider the function $h: U \to R$ defined by $h \equiv g_{n-1} + f_n - g$, then $\partial h \equiv \partial g$ but $h - g$ is not a constant.

Remark 2. It is explicit in the above construction that the function g has very circumscribed derivatives. Indeed (a_n) shows

$$\partial g(x) = \text{co}\{\partial f_1(x), \partial f_2(x), \ldots, \partial f_n(x)\} \qquad \text{for every } x \in U$$

and $\nabla g(x) \in \{\nabla f_1(x), \nabla f_2(x), \ldots, \nabla f_n(x)\}$ almost everywhere in U.

3. APPLICATIONS AND EXAMPLES

To see that Theorem 1 provides a rich source of examples, we need to establish a large class of Lipschitz functions whose Clarke subdifferential mappings are minimal weak* cuscos. Here the results in [10] aid us. In [10] the authors show that on a separable Banach space every real-valued locally Lipschitz function which is strictly differentiable almost everywhere in its domain possesses a minimal subdifferential mapping and they also show that this family of functions is closed under addition, multiplication, and the two lattice operations. Moreover, they also show that any locally Lipschitz function which is either *semi-smooth* almost everywhere or *pseudo-regular* almost everywhere belongs to this class. However, we should note that there are many examples of locally Lipschitz functions whose Clarke subdifferential mappings are minimal, but which are not strictly differentiable almost everywhere. In another direction, the authors in [10] show that if the norm on X is uniformly Gâteaux differentiable then each distance function on X possesses a minimal subdifferential mapping. Let us also note that in finite dimensions all smooth norms are uniformly Gâteaux differentiable.

Corollary 1. Let A be a nonempty open subset of a separable Banach space X. If $\{f_1, f_2, \ldots, f_n\}$ are real-valued strictly differentiable (or equivalently, continuously Gâteaux differentiable) locally Lipschitz functions defined on A, then there exists a real-valued locally Lipschitz function g defined on A such that $\partial g(x) = \text{co}\{\nabla f_1(x), \nabla f_2(x), \ldots, \nabla f_n(x)\}$ for each $x \in A$. In particular, when $\{f_1, f_2, \ldots, f_n\}$ are continuous linear functions on X, then $\partial g(x) \equiv \text{co}\{f_1, f_2, \ldots, f_n\}$ for each $x \in A$.

Thus every polytope arises as the constant Clarke subgradient of some globally Lipschitz functions. Our next application is to *cyclically monotone operators* [7].

Corollary 2. Let A be a nonempty open convex subset of a separable Banach space X and $\{T_1, T_2, \ldots, T_n\}$ be a finite family of maximal cyclically monotone operators from A into nonempty subsets of X^*. Then there exists a real-valued locally Lipschitz function g defined on A such that

$$\partial g(x) = \text{co}\{T_1(x), T_2(x), \ldots, T_n(x)\} \qquad \text{for each } x \in A.$$

Moreover, g is convex if, and only if, $T_1 = T_2 = \cdots = T_n$.

Proof. By [11] each T_j is the Clarke subgradient of a proper lower semi-continuous convex function on A. However, as Phelps observes in [7] such a convex function is finite valued on the domain on T_j and so we may conclude that it is continuous on A. Therefore by Theorem 1, there exists a real-valued locally Lipschitz function g defined on A such that $\partial g(x) = \text{co}\{T_1(x), T_2(x), \ldots, T_n(x)\}$ for each $x \in A$.

If g is convex, then by [10], $x \to \partial g(x)$, is a minimal weak* cusco on A and so $\partial g = T_1 = T_2 = \cdots = T_n$. On the other hand if $T_1 = T_2 = \cdots = T_n$, then by [11] and our previous argument, there exists a continuous convex function g_1 defined on A such that $\partial g_1 = T_1$. However, by an easy result in [12], it follows that $g_1 - g$ is a constant and so g must also be convex on A. ■

Note that the previous corollary shows that the convex hull of a finite family of maximal cyclically monotone or antitone operators is a subgradient. Next we ask whether the conclusions of Theorem 1 hold without the minimality assumptions? To give some support to this question, we establish the following special case.

THEOREM 2. Let $\{f_1, f_2, \ldots, f_n\}$ be real-valued locally Lipschitz functions defined on an open subset A of R. Then there exists a real-valued locally Lipschitz function g defined on A such that

$$\partial g(x) = \text{co}\{\partial f_1(x), \partial f_2(x), \ldots, \partial f_n(x)\} \qquad \text{for each } x \in A.$$

Proof. For each $1 \le j \le n$, let $\partial f_j(x) = [\alpha_j(x), \beta_j(x)]$ for each $x \in A$. Now let $\alpha(x) = \min\{\alpha_j(x) : 1 \le j \le n\}$ and $\beta(x) = \max\{\beta_j(x) : 1 \le j \le n\}$.

Clearly $\text{co}\{\partial f_1(x), \partial f_2(x), \ldots, \partial f_n(x)\} = [\alpha(x), \beta(x)]$ for each $x \in A$. The result now follows from Theorem 3.2 in [1]. ■

A special case of Theorems 1 and 2 recaptures a significant special case of the main result in [1].

COROLLARY 3. Let α and β be real-valued continuous functions defined on an open interval (a, b). Then there exists a real-valued locally Lipschitz function g defined on (a, b) such that $\partial g(x) = \text{co}\{\alpha(x), \beta(x)\}$ for all x in (a, b).

Proof. Let A and B be anti-derivatives of α and β respectively. Then both A and B are locally Lipschitz on (a, b) and the result now follows from either Theorems 1 or 2. ■

Remark 3. We should note that Theorems 1 and 2 fail for countable families of equi-Lipschitz functions.

For example, consider the following family of Lipschitz-1 functions $\{f_n : n \in \mathbf{N}\}$ defined on R such that $\partial f_n(x) = [0, \min\{n|x|, 1\}]$ for each $x \in R$. Then

$$\overline{\text{co}}\left\{\bigcup_{n=1}^{\infty} \partial f_n(x)\right\} = \begin{cases} \{0\} & \text{if } x = 0 \\ [0, 1] & \text{otherwise.} \end{cases}$$

Clearly, $x \to \overline{\text{co}}\{\bigcup_{n=1}^{\infty} \partial f_n(x)\}$ is not upper semi-continuous at $x = 0$ and so not a Clarke subdifferential mapping of any Lipschitz function on R.

Finally we end with a concrete example, which indicates the type of pathological behavior that the functions constructed using Theorem 1 may possess.

Example 1. Consider the Lipschitz function $k: R^2 \to R$ defined by

$$k(x, y) = \frac{1 - \sqrt{2}\cos(4y) \cdot \sin(4x + \pi/4)}{1 + x^2 + y^2}.$$

Let P be a polytope in R^2 and let g be a locally Lipschitz function defined on R^2 such that $\partial g(x, y) \equiv P$.

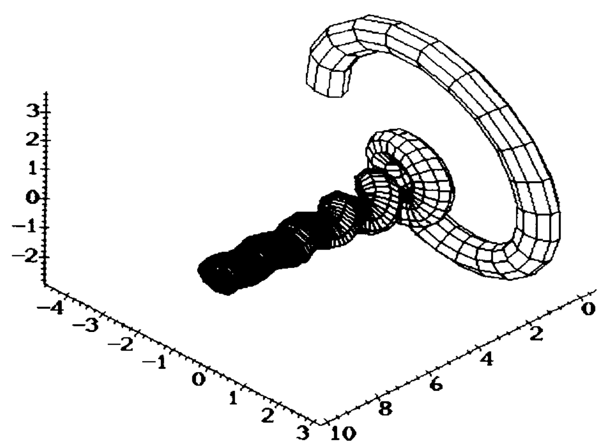

Fig. 1.

In Fig. 1 we show a tube plot of the mapping $(x, y) \to \partial(k + g)(x, y) = \nabla k(x, y) + P$ restricted to the line segment $\{(0, y) : 0 \leq y \leq 8\}$. In the diagram the radius of P is $1/2$.

Observation 1. On any separable Banach space, it is impossible to construct a locally Lipschitz function whose Michel–Penot subgradient is identically equal to a polytope. The reason is that on separable Banach spaces the Michel–Penot subgradients are single-valued almost everywhere (in the Haar-null sense).

Acknowledgement—The authors would like to thank Jens Happe for help in preparing Fig. 1.

REFERENCES

1. Borwein, J. M. and Fitzpatrick, S., Characterization of Clarke subgradients among one-dimensional multifunctions. In *Proc. Optim. Miniconference II*, ed. B. M. Glover and V. Jeyakumar. University of Ballarat, Vic., 1995, pp. 61–73.
2. Jouini, E., A remark on Clarke's normal cone and the marginal cost pricing rule, *Journal of Mathematical Economics*, 1989, **18**, 95–101.
3. Rockafellar, R. T., *The Theory of the Subgradient and its Applications: Convex and Nonconvex Functions*. Helderman, Berlin, 1981.
4. Clarke, F. H., *Optimization and Nonsmooth Analysis*. Wiley Interscience, New York, 1983.
5. Borwein, J. M., Fitzpatrick, S. and Kenderov, P., Minimal convex uscos and monotone operators on small sets, *Canadian Journal of Mathematics*, 1991, **43**(3), 461–476.
6. Christensen, J. P. R., *Topological and Borel Structure*. American Elsevier, New York, 1974.
7. Phelps, R. R. *Convex Functions, Monotone Operators and Differentiability*. Springer-Verlag, New York, 1993.
8. Thibault, L. On generalized differentials and subdifferentials of Lipschitz vector-valued functions, *Journal of Nonlinear Analysis: Theory, Methods, Applications*, 1982, **6**, 1037–1053.
9. Stromberg, K. R., *An Introduction to Classical Real Analysis*. Wadsworth International Mathematics Series, 1981.
10. Borwein, J. M. and Moors, B. S., Essentially strictly differentiable Lipschitz functions, *J. Funct. Anal.* (to appear).
11. Rockafellar, R. T., Characterization of the subdifferentials of convex functions, *Pacific Journal of Mathematics*, 1966, **17**, 497–510.
12. Borwein, J. M., Minimal cuscos and subgradients of Lipschitz functions. In *Fixed Point Theory and Its Applications*, ed. by J.-B. Baillon and M. Thera. Pitman Lecture Notes in Math., Longman, Essex, 1991, pp. 57–82.