

## SUBGRADIENT REPRESENTATION OF MULTIFUNCTIONS

J. BORWEIN<sup>1</sup>, W. B. MOORS<sup>2</sup> and Y. SHAO<sup>1</sup>

(Received 4 March 1997; revised 10 June 1997)

### Abstract

We provide necessary and sufficient conditions for a minimal upper semicontinuous multifunction defined on a separable Banach space to be the subdifferential mapping of a Lipschitz function.

### 1. Introduction and preliminaries

A set-valued mapping  $\Phi$  from a topological space  $A$  into subsets of a linear topological space  $(Y, \tau)$  is referred to as a  $\tau$ -*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 *upper or outer semi-continuous*.)

A cusco mapping  $\Phi$  from a topological space  $A$  into subsets of a linear topological space  $(Y, \tau)$  is said to be a *minimal  $\tau$ -cusco* if its graph does not strictly contain the graph of any other  $\tau$ -cusco defined on  $A$ .

Let  $X$  be a Banach space with norm dual  $X^*$ . We use  $\langle \cdot, \cdot \rangle$  to specify the pairing between  $X$  and  $X^*$ . A real-valued function  $f$  defined on a non-empty open subset  $A$  of  $X$  is *locally Lipschitz* on  $A$  if for each  $x_0 \in A$  there exists a  $K > 0$  and a neighbourhood  $U$  of  $x_0$  such that

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

For functions in this class, the *Clarke generalized directional derivative* at  $x \in A$ , in the direction  $y$ , is defined by

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

<sup>1</sup>CECM, Department of Mathematics and Statistics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

<sup>2</sup>Department of Mathematics, The University of Auckland, Private Bag 92019, Auckland, New Zealand  
© Australian Mathematical Society 1999, Serial-fee code 0334-2700/99

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

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

The Clarke subdifferential mapping,  $x \rightarrow \partial f(x)$ , has played a crucial role in non-smooth analysis. It is well-known that the Clarke subdifferential mapping of a real-valued locally Lipschitz function defined on a non-empty open subset of  $X$  is a weak\* cusco, with images in  $X^*$ . While not every locally Lipschitz function gives rise to a *minimal* cusco, a very large proportion of non-pathological functions do. A notable example is given by the class of essentially smooth functions (those whose subdifferentials are singleton almost everywhere; see [4, 6]), which form a lattice algebra including the convex and  $C^1$  functions. In this paper we concern ourselves with the converse problem: when can a given (minimal) weak\* cusco be represented as the the Clarke subdifferential mapping of a real-valued locally Lipschitz function?

There has been some considerable work on the subgradient representation of multifunctions. An early result is due to Rockafellar, who proved that a multifunction is the subdifferential of a lower semicontinuous proper convex function if and only if the multifunction is *maximal cyclically monotone*; see [12]. Janin further showed in [10] that a multifunction is *cyclically submonotone* if and only if this multifunction is the Clarke subdifferential mapping of a lower- $C^1$  (locally Lipschitz) function in the sense of Rockafellar [14]. Recently Poliquin [13] proved that a multifunction is the proximal subdifferential mapping of a lower semicontinuous function bounded below by a quadratic if and only if it satisfies a monotone selection property. Finally, in one dimension, a full characterization of when a multifunction is the Clarke subgradient of a Lipschitz function may be found in [3].

In this paper we develop, in terms of line integrals (as in the classical Green's theorem), necessary (and sufficient) conditions for a given minimal weak\* cusco to be the Clarke subdifferential mapping of a locally Lipschitz function. However, it is not our intention to provide a deep or thorough investigation of this problem, but rather, to present some ideas and results that will facilitate further research in this area. Indeed, our main goal has been to present results which we feel may have wide-spread utility. We first provide in Section 2 a sufficient condition ensuring that a minimal weak\* cusco is the Clarke subdifferential mapping of a Lipschitz function. This condition is subsequently applied to recapture the Hilbert space case of Rockafellar's cyclic monotonicity theorem. Moreover, we also derive from this condition a necessary and sufficient condition for a Lipschitz vector-field to be the gradient mapping of a real-valued Lipschitz function. Ultimately, in Section 3 we provide a characterization for a minimal weak\* cusco to be the Clarke subdifferential mapping of a Lipschitz function.

Let us now make precise some basic tools for use in the sequel. A line segment  $[a, b]$  in a Banach space  $X$  is defined as

$$[a, b] \equiv \{tb + (1 - t)a \mid 0 \leq t \leq 1\}.$$

We define the *line integral* on  $[a, b]$  of a single-valued mapping  $\sigma : X \rightarrow X^*$ ,  $\int_{[a,b]} \sigma(z) dz$ , as the Lebesgue integral

$$\int_0^1 \langle \sigma(tb + (1 - t)a), (b - a) \rangle dt,$$

where we implicitly assume that the function  $T : [0, 1] \rightarrow \mathbb{R}$ , defined by  $T(t) \equiv \langle \sigma(tb + (1 - t)a), (b - a) \rangle$ , is Lebesgue-measurable. It can easily be seen that

$$\int_{[a,b]} \sigma(z) dz = - \int_{[b,a]} \sigma(z) dz.$$

A *polygonal path*  $C$  in  $X$  is an ordered collection of line segments  $\{[a_i, a_{i+1}] \mid 1 \leq i \leq n - 1\}$  for some positive integer  $n$ . Such a path is said to be closed when  $a_1 = a_n$ . We write  $-C$  to denote the ordered collection of line segments  $\{[a_{n-i+1}, a_{n-i}] \mid 1 \leq i \leq n - 1\}$ . The line integral  $\int_C \sigma(z) dz$  on a polygonal path  $C$  is defined as

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

We say that a multifunction  $\Phi : X \rightarrow 2^{X^*}$  is *locally bounded*, provided that for each  $x_0 \in X$  there exists a positive number  $L$  and a neighbourhood  $U$  of  $x_0$  such that

$$\|y^*\| \leq L \tag{2}$$

for all  $y^* \in \Phi(u)$  with  $u \in U$ . Moreover,  $\Phi$  is said to be *bounded by  $L$*  if inequality (2) is satisfied by all  $y^*$  in the image of  $\Phi$ .

Our other notation is standard. For example,  $B(X)$  ( $B(X^*)$ ) denotes the closed unit ball of  $X$  (resp.  $X^*$ ), while  $S(X)$  denotes the unit sphere of  $X$ .

## 2. A sufficient condition and its applications

In this section we first provide in an arbitrary Banach space a sufficient condition for a minimal weak\* cusco to be the Clarke subdifferential mapping of a Lipschitz function. We illustrate the use of this result by recapturing, in Hilbert space, Rockafellar's cyclic monotonicity theorem. Additionally, we obtain a necessary and sufficient condition for a Lipschitz vector-field to be the gradient mapping of a Lipschitz function in  $\mathbb{R}^2$ .

**THEOREM 2.1.** *Assume that  $X$  is a Banach space and  $A$  is a non-empty connected open subset of  $X$ . Let  $\Phi : A \rightarrow 2^{X^*}$  be a locally bounded weak\* cusco. Suppose  $\Phi$  possesses a selection  $\sigma : A \rightarrow X^*$  such that*

$$\oint_C \sigma(z) dz \leq 0 \tag{3}$$

*for every closed polygonal path  $C$  in  $A$ . Then there is some locally Lipschitz function  $f$  on  $A$  with  $\partial f \subseteq \Phi$ . Moreover, if  $\Phi$  is minimal, then  $\Phi = \partial f$ .*

**PROOF.** Let us note first that open connected sets are actually polygonally connected. We define  $f : A \rightarrow \mathbb{R}$  by

$$f(x) \equiv \int_{\Gamma} \sigma(z) dz,$$

where  $\Gamma$  is some polygonal path in  $A$  from a given point  $a \in A$  to  $x \in A$ . Observe that (3) holding for each closed polygonal path  $C$  in  $A$  implies

$$\oint_C \sigma(z) dz = 0$$

for every closed polygonal path  $C$  in  $A$ . As in the classical case, we see that  $f$  is well defined. Let us now verify that  $f$  is locally Lipschitz on  $A$ . Given any  $x \in A$ , since  $\Phi$  is locally bounded, there exists a positive number  $L$  such that  $\Phi(U) \subseteq LB(X^*)$  for some convex neighbourhood  $U$  of  $x$  contained in  $A$ . Thus, for any  $y_1, y_2 \in U$ , one has

$$\begin{aligned} |f(y_2) - f(y_1)| &= \left| \int_{[y_1, y_2]} \sigma(z) dz \right| = \left| \int_0^1 \langle \sigma(ty_2 + (1-t)y_1), y_2 - y_1 \rangle dt \right| \\ &\leq \int_0^1 |\langle \sigma(ty_2 + (1-t)y_1), y_2 - y_1 \rangle| dt \leq L \|y_2 - y_1\|. \end{aligned}$$

Finally, we show that  $\partial f(z) \subseteq \Phi(z)$  for any  $z \in X$ . We assume, to obtain a contradiction, that for some  $z$ ,  $\partial f(z) \not\subseteq \Phi(z)$ . Then we may find an  $x^* \in \partial f(z)$  and an  $x \in S(X)$  such that  $\langle x^*, x \rangle > \alpha > \max(\Phi(z), x)$ . This implies  $f^0(z; x) > \alpha > \max(\Phi(z), x)$ . By the upper semicontinuity of  $\Phi$ , we can select a convex neighbourhood  $V$  of  $z$  such that  $\max(\Phi(V), x) < \alpha$ . On the other hand, we may select  $v \in V$  such that  $f'(v; x) > \alpha$ . Hence there exists a positive number  $\epsilon$  such that

$$\frac{f(v + \mu x) - f(v)}{\mu} > \alpha \text{ if } |\mu| < \epsilon.$$

The definition of  $f$  allows us to conclude

$$f(v + \epsilon x) - f(v) = \int_{[v, v+\epsilon x]} \sigma(z) dz.$$

Then, employing the mean value theorem and Lebesgue’s differentiation theorem, we obtain  $\sigma(w)(\epsilon x) > \alpha \epsilon$  for some  $w \in [v, v + \epsilon x]$ , which gives  $\sigma(w)(x) > \alpha$ . This is impossible since we have  $w \in V$  and  $\sigma(w) \in \Phi(w) \subseteq \Phi(V)$ . In consequence we have  $\partial f(z) \subseteq \Phi(z)$  for all  $z \in A$ .

It is now immediate that  $\partial f = \Phi$  whenever  $\Phi$  is minimal, since  $\partial f$  is a weak\* cusco.

We illustrate the first application of Theorem 2.1 by recapturing the Hilbert space case of Rockafellar’s cyclic monotonicity theorem [12] as follows. Recall that a multifunction  $T : X \rightarrow 2^{X^*}$  is *cyclically monotone* if given  $(x_i, x_i^*) \in \text{gph } T$ , ( $i = 0, 1, 2, \dots, n$ , where  $n$  is arbitrary) and  $\text{gph } T$  denotes the graph of  $T$ , we have

$$\langle x_0^*, x_1 - x_0 \rangle + \langle x_1^*, x_2 - x_1 \rangle + \dots + \langle x_n^*, x_0 - x_n \rangle \leq 0.$$

**COROLLARY 2.1.** *Let  $H$  be a Hilbert space and  $T : H \rightarrow 2^H$  be maximal and cyclically monotone. Then there exists a proper convex (closed) function  $f$  defined on  $H$  such that  $T = \partial f$ .*

**PROOF.** (**Special case:**  $D(T) = H$ .) [In this case the reasoning applies in *arbitrary Banach spaces*.] Recall that  $T$  is a locally bounded minimal weak\* cusco on  $H$ . Let  $\sigma$  be any selection of  $T$ . We claim that  $\int_C \sigma(x) dz \leq 0$  for any closed polygonal path  $C$ . Notice that for  $[a, b] \subset H$  the function  $t \rightarrow \sigma(tb + (1 - t)a)(b - a)$  is monotone, and hence Riemann integrable. Also,

$$\begin{aligned} & \int_{[a,b]} \sigma(z) dz \\ &= \sup \left\{ \sum_{i=1}^n \sigma(t_{i-1}b + (1 - t_{i-1})a)((b - a)(t_i - t_{i-1})) \mid 0 = t_1 < t_2 < \dots < t_n = 1 \right\} \\ &\leq \sup \left\{ \sum_{i=1}^n \sigma(x_{i-1})(x_i - x_{i-1}) \mid P = \{[x_{i-1}, x_i] : 1 \leq i \leq n\} \text{ polygonal from } a \text{ to } b \right\}. \end{aligned}$$

Hence  $\int_C \sigma(z) dz \leq 0$  for any closed polygonal path  $C$  in  $H$  because

$$\sigma(x_n)(x_0 - x_n) + \sigma(x_{n-1})(x_n - x_{n-1}) + \dots + \sigma(x_0)(x_1 - x_0) \leq 0$$

for any finite set of points  $\{x_0, x_1, \dots, x_n\} \subset H$ . We see from Theorem 2.1 that the conclusion holds whenever  $D(T) = H$ .

(General Case.) Since  $T$  is maximal monotone, so is  $T^{-1}$ . Using Minty's theorem [11] we then know that the range  $R(T^{-1} + I)$  is the whole space  $H$ . This means that the domain  $D(T_1)$  of the resolvent  $T_1 = (T^{-1} + I)^{-1}$  is  $H$ . Observe that the cyclical monotonicity of  $T^{-1}$  follows from that of  $T$ . This implies the cyclical monotonicity of  $T^{-1} + I$  and then that of  $T_1$ . Thus applying the assertion just proved above for the special case that  $D(T) = H$ , we obtain a locally Lipschitz function  $g$  such that  $T_1 = \partial g$ . By virtue of Theorem 5 of [9] we deduce that  $g$  is convex from the monotonicity of  $T_1$ . We thus infer that  $T_1^{-1} = \partial g^*$ , where  $g^*$  is the conjugate of  $g$ . This allows us to conclude that

$$T^{-1} = \partial g^* - I = \partial(g^* - \|\cdot\|^2/2)$$

by virtue of the subgradient sum rule in non-smooth analysis. We again appeal to Theorem 5 of [9] to see that  $h^* = g^* - \|\cdot\|^2/2$  is convex since  $T^{-1}$  is monotone. We have determined that  $T = \partial h$  with  $h$  given as the conjugate of  $h^*$ . This ends the proof of the Corollary.

We conclude this section by applying Theorem 2.1 to get a necessary and sufficient condition for a Lipschitz vector-field in  $\mathbb{R}^2$  to be the gradient mapping of a real-valued Lipschitz function. To this end, we need to modify the classical Green's theorem as follows.

LEMMA 2.1. *Let  $P$  and  $Q$  be two real-valued functions in  $\mathbb{R}^2$  defined and continuous on a Jordan region  $R$  bounded by a rectifiable Jordan curve  $\Gamma$ . Assume that the partial derivatives  $\frac{\partial P}{\partial y}$  and  $\frac{\partial Q}{\partial x}$  exist almost everywhere on  $R$  and are essentially bounded on the interior of  $R$ . Suppose further that  $\alpha = (\alpha_1, \alpha_2) \in \Gamma(R)$ , where  $\Gamma(R)$  is the positively oriented boundary of  $R$ . Then the line integral  $\int_{\Gamma(R)} P d\alpha_1 + Q d\alpha_2$  exists and satisfies*

$$\int_R \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) d(x, y) = \int_{\Gamma(R)} P d\alpha_1 + Q d\alpha_2.$$

The proof of the above lemma is just a modification of that of Theorem 10-43 of [1]. The definitions of a *Jordan region* and a *rectifiable Jordan curve* can also be found in [1]. To derive the following consequence, we also have to employ Rademacher's theorem which states that real-valued locally Lipschitz functions defined on  $R^n$  are Fréchet differentiable almost everywhere (in their domain), in particular, the partial derivatives exist almost everywhere.

PROPOSITION 2.1. *Suppose that  $F = (P, Q)$  is a locally Lipschitz vector-field defined on open simply connected region  $R$  in  $\mathbb{R}^2$ . Then  $F = \nabla f$  for some locally*

Lipschitz function  $f$  on  $R$  if and only if

$$\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y} \quad (4)$$

almost everywhere on  $R$ .

PROOF. Let us assume first that  $F = \nabla f$  for some locally Lipschitz function  $f$ . Then

$$(P, Q) = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right).$$

Since  $F$  is locally Lipschitz,  $F$  is differentiable on  $R$  almost everywhere, according to Rademacher's Theorem. This means that  $\nabla^2 f$  exists on  $R$  almost everywhere. Notice that

$$\frac{\partial P}{\partial y} = \frac{\partial^2 f}{\partial x \partial y} \quad \text{and} \quad \frac{\partial Q}{\partial x} = \frac{\partial^2 f}{\partial y \partial x}. \quad (5)$$

By virtue of Theorem 12.12 in [2], we see that (5) holds wherever  $\nabla^2 f$  exist. Therefore (4) must hold almost everywhere in  $R$ .

Conversely, we assume that (4) holds on  $R$  almost everywhere. Invoking Lemma 2.1 and using (4), we can show, as in the classical case, that

$$\int_C F d\alpha = \int_C P d\alpha_1 + Q d\alpha_2 = 0$$

for any closed polygonal path  $C$  in  $R$ . Observe that  $(x, y) \rightarrow \{F(x, y)\}$  is single-valued and so minimal. Thus by Theorem 2.1 there exists a locally Lipschitz function  $f$  defined on  $R$  such that  $\partial f = \{F\}$ . Consequently Proposition 2.2.4 of [8] allows us to conclude that  $F = \nabla f$  holds on the whole  $R$ . This ends the proof of the proposition.

REMARK 2.1. We should note here that Proposition 2.1 may be extended to any finite-dimensional Banach space in the following way: "Let  $R$  be an open rectangle in  $\mathbb{R}^n$ . Then a locally Lipschitz vector-field  $F$  on  $R$  is the derivative of a real-valued locally Lipschitz function defined on  $R$  if and only if the Clarke generalized Jacobian of  $F$  is symmetric."

Below we give a simple application of Proposition 2.1, by considering the two regions  $x > |y|y$  and  $x < |y|y$ .

EXAMPLE 2.1. For each  $\lambda \in \mathbb{R}$  define  $F_\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  by  $F_\lambda(x, y) \equiv (P_\lambda(x, y), Q_\lambda(x, y))$  where

$$P_\lambda(x, y) \equiv \lambda/2|x + |y|y| \quad \text{and} \quad Q_\lambda(x, y) \equiv |x + |y|y||\lambda y|.$$

Then by Proposition 2.1 we see that  $F_\lambda$  is the derivative of some real-valued locally Lipschitz function on  $\mathbb{R}^2$  if and only if  $\lambda \geq 0$ .

### 3. A subgradient characterization

In this section we characterize, in line integral terms, when a minimal cusco is the Clarke subdifferential mapping of a Lipschitz function. Let us start by recalling some additional definitions and assertions.

Let  $X$  be a Banach space. A Borel subset  $N \subseteq X$  is 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 are some of the basic properties enjoyed by Haar-null sets.

PROPOSITION 3.1. ([6], Proposition 2.1) *Let  $X$  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.*

The following ‘‘Fubini’’ result, regarding Haar-null sets, will also be utilized later.

PROPOSITION 3.2. ([6]) *Let  $X$  be a Banach space and  $N$  a Haar-null subset in  $X \times \mathbb{R}^n$ . Then for almost every  $b \in X$ , that is, except for  $b$  in a Haar-null subset in  $X$ , the set*

$$N_b = \{z \in \mathbb{R}^n \mid (b, z) \in N\}$$

*is a Lebesgue-null subset of  $\mathbb{R}^n$ .*

Using the language of Haar-null sets, Christensen provided the following generalization of Rademacher’s theorem. A slightly more general version of this result may be found in [5].

PROPOSITION 3.3. ([7], Theorem 7.5) *Assume that  $X$  is a separable Banach space. 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  $f$  is Gateaux differentiable at each  $x \in D$ .*

The assertions below regarding the Clarke generalized directional derivative and subdifferential are due to Thibault and generalize results originally proven in  $\mathbb{R}^n$  by Clarke [8]. For the most general form of these results, see [6].



PROPOSITION 3.4. ([15]) *Let  $X$  be a separable Banach space and  $g$  be a locally Lipschitz function on  $X$ . Suppose that  $M$  is a subset of  $X$  on which  $g$  is Gateaux differentiable and such that  $X \setminus M$  is Haar-null in  $X$ . Then we have*

- (i)  $g^0(\bar{x}, v) = \max\{\langle \xi, v \rangle \mid \xi \in L_M(g, \bar{x})\}$  for every  $v \in X$ ;
- (ii)  $\partial g(\bar{x}) = \text{cl}^* \text{co}[L_M(g, \bar{x})]$ , where

$$L_M(g, \bar{x}) \equiv \{\text{weak}^* - \lim_{n \rightarrow \infty} \nabla g(x_n) \mid \lim_{n \rightarrow \infty} x_n = \bar{x}, x_n \in M\}$$

and “ $\text{cl}^* \text{co}$ ” denotes the weak\* closure of the convex hull.

We now further discuss polygonal paths. Let  $X$  be a Banach space, and  $A$  be a non-empty open connected subset of  $X$ . Suppose also that  $B \subseteq A$  is a Borel set. For a fixed  $\epsilon > 0$  we will call an ordered collection of line segments  $P(\epsilon) \equiv \{[a_i, b_i] \mid 1 \leq i \leq n - 1\}$  an  $\epsilon$ -path from  $a$  to  $b$  provided that

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

Such a path is said to be closed if  $a = b$ . Moreover, we write  $-P(\epsilon) = \{[b_i, a_i] \mid 1 \leq i \leq n - 1\}$ . Furthermore, we say that  $P$  is a  $B$ -admissible  $\epsilon$ -path from  $a$  to  $b$  if  $P$  is an  $\epsilon$ -path from  $a$  to  $b$  and

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

for each  $1 \leq i \leq n - 1$ . Line integrals on a  $\epsilon$ -path are defined similarly to (1). Here ‘ $\lambda$ ’ denotes Lebesgue measure on the line.

We are now ready for our main result.

THEOREM 3.1. *Let  $X$  be a separable Banach space, and let  $A$  be a non-empty open connected subset of  $X$ . Then for a minimal weak\* cusco  $\Phi : A \rightarrow 2^{X^*}$  to be the Clarke subdifferential mapping of a locally Lipschitz function (with uniform rank  $L$ ) it is necessary and sufficient that (i)  $\Phi$  be bounded by  $L$  and (ii) that there exist a Borel set  $B \subset A$  with  $A \setminus B$  Haar-null and a measurable selection  $\sigma : B \rightarrow X^*$  such that for each  $\epsilon > 0$  and each closed  $B$ -admissible  $\epsilon$ -path  $P(\epsilon)$  in  $A$ , one has*

$$\left| \oint_{P(\epsilon)} \sigma(x) dx \right| < L\epsilon. \tag{6}$$

PROOF. (a) “only if”. Let us assume that  $\Phi = \partial f$  for some Lipschitz function  $f$  with rank  $L$ . Then it is evident that  $\Phi$  is bounded by  $L$ . Now we verify (6) for every  $\epsilon > 0$  and each closed  $B$ -admissible  $\epsilon$ -path  $P(\epsilon)$  in  $A$ . Letting

$$B \equiv \{x \in A \mid f \text{ is Gateaux differentiable at } x\},$$

we know from Proposition 3.3 that the set  $N \equiv A \setminus B$  is Haar-null. We define  $\sigma : B \rightarrow X^*$  by  $\sigma(x) = \nabla f(x)$ . It follows that the mapping  $\sigma$  is Borel measurable. Given any  $\epsilon > 0$ , let  $P = \{[a_i, b_i] \mid 1 \leq i \leq n\}$  be a closed  $B$ -admissible  $\epsilon$ -path in  $A$ . Then by virtue of Lebesgue's differentiation theorem, we obtain

$$f(b_i) - f(a_i) = \int_0^1 \langle \nabla f(t b_i + (1-t)a_i), b_i - a_i \rangle dt = \int_{[a_i, b_i]} \sigma(z) dz.$$

We may now derive (6) from the following chain of the inequalities:

$$\begin{aligned} \left| \int_{P(\epsilon)} \sigma(x) dx \right| &= \left| \sum_{i=1}^n f(b_i) - f(a_i) \right| \leq |f(a_1) - f(b_n)| + \sum_{i=1}^{n-1} |f(a_{i+1}) - f(b_i)| \\ &\leq L\|a_1 - b_n\| + \sum_{i=1}^{n-1} L\|a_{i+1} - b_i\| < L\epsilon. \end{aligned}$$

(b) "if". Let us now prove the converse. We first define the function

$$f(x) \equiv \lim_{\epsilon \rightarrow 0^+} \int_{P(\epsilon)} \sigma(z) dz \quad \text{for all } x \in A, \tag{7}$$

where  $P(\epsilon)$  is any  $B$ -admissible  $\epsilon$ -path in  $A$  from some fixed  $a \in A$  to  $x$ . Then  $f$  is well defined. Indeed, for any  $\epsilon_1, \epsilon_2 > 0$ , let  $P_i(\epsilon_i)$  be  $B$ -admissible  $\epsilon_i$ -paths from  $a$  to  $x$ ,  $i = 1, 2$ . Then  $P_1(\epsilon_1) + (-P_2(\epsilon_2))$  is a closed  $B$ -admissible  $(\epsilon_1 + \epsilon_2)$ -path. By assumption we then deduce

$$\left| \int_{P_1(\epsilon_1)} \sigma(z) dz - \int_{P_2(\epsilon_2)} \sigma(z) dz \right| = \left| \int_{P_1(\epsilon_1) + (-P_2(\epsilon_2))} \sigma(z) dz \right| < L(\epsilon_1 + \epsilon_2).$$

Therefore we see that the limit on the right of (7) does exist, and then  $f$  is well defined, as claimed.

Next let us verify that the function  $f$  is locally Lipschitz with rank  $L$ . Let  $x_0$  be any point in  $A$  and let  $U$  be any convex neighborhood of  $x_0$  contained in  $A$ . Consider any two points  $x, y \in U$  and fix  $\delta > 0$ . By the definition of  $f$  we may choose  $\delta > \epsilon_0 > 0$  sufficiently small so that if  $P(\epsilon)$  is any  $B$ -admissible  $\epsilon$ -path ( $0 < \epsilon < \epsilon_0$ ) from  $a$  to  $x$ , then

$$\left| f(x) - \int_{P(\epsilon)} \sigma(z) dz \right| < \delta,$$

and if  $P'(\epsilon)$  is any  $B$ -admissible  $\epsilon$ -path from  $a$  to  $y$ , then one has

$$\left| f(y) - \int_{P'(\epsilon)} \sigma(z) dz \right| < \delta.$$

Now suppose  $P(\epsilon) = \{[a_i, b_i] \mid 1 \leq i \leq n\}$  is a  $B$ -admissible  $\epsilon$ -path from  $a$  to  $x$  (with  $0 < \epsilon < \epsilon_0$ ). By Proposition 3.2 we may extend  $P(\epsilon)$  to a  $B$ -admissible  $\epsilon$ -path ( $P'(\epsilon)$ , say) from  $a$  to  $y$ , where  $P'(\epsilon) = \{[a_i, b_i] \mid 1 \leq i \leq n + 1\}$ . It follows that

$$\begin{aligned} |f(y) - f(x)| &\leq \left| f(y) - \int_{P'(\epsilon)} \sigma(z) dz \right| + \left| \int_{P'(\epsilon)} \sigma(z) dz - \int_{P(\epsilon)} \sigma(z) dz \right| \\ &\quad + \left| \int_{P(\epsilon)} \sigma(z) dz - f(x) \right| \\ &\leq \delta + \left| \int_{[a_{n+1}, b_{n+1}]} \sigma(z) dz \right| + \delta \\ &\leq L \|a_{n+1} - b_{n+1}\| + 2\delta \\ &\leq L(\|a_{n+1} - x\| + \|x - y\| + \|y - b_{n+1}\|) + 2\delta \\ &\leq L\|x - y\| + 2(L + 1)\delta. \end{aligned}$$

This gives  $|f(y) - f(x)| \leq L\|x - y\|$  as  $\delta$  is arbitrary.

Finally, let us show that  $\partial f(x) \subseteq \Phi(x)$  for all  $x \in A$ . Define

$$D \equiv \{x \in X \mid f \text{ is Gateaux differentiable at } x\}.$$

We know from Proposition 3.3 that  $X \setminus D$  is Haar-null. So in view of Proposition 3.4 it is sufficient to show that  $\nabla f(x) \in \Phi(x)$  for all  $x \in D$ . Assume, for the purpose of obtaining contradiction, that  $\nabla f(x) \notin \Phi(x)$  for some  $x \in D$ . Thus, by the Mazur separation theorem, we can choose an  $\alpha \in \mathbb{R}$  and a  $y \in S(X)$  such that

$$\langle \nabla f(x), y \rangle > \alpha > \max\langle \Phi(x), y \rangle.$$

Using the upper semicontinuity of  $\Phi$ , we can find a convex neighbourhood  $U$  of  $x$  such that  $\sup\langle \Phi(U), y \rangle < \alpha$ . Moreover, since  $f$  is Gateaux differentiable at  $x$ , we may select an  $\epsilon > 0$  such that

$$\frac{f(x + \epsilon y) - f(x)}{\epsilon} > \alpha.$$

Let  $G \equiv D \cap B$ . Then we know from Proposition 3.1(iv) that  $A \setminus G$  is also Haar-null. By Proposition 3.2 we can find  $v \in X$  close to  $x$  such that  $[v, v + \epsilon y]$  is  $G$ -admissible and

$$\frac{f(v + \epsilon y) - f(v)}{\epsilon} > \alpha.$$

Moreover, we can assume without loss of generality that  $[v, v + \epsilon y] \subset U$ . From the definition of  $f$  we may deduce the equality

$$f(v + \epsilon y) - f(v) = \int_{[v, v + \epsilon y]} \sigma(z) dz.$$

Thus, it follows from Lebesgue’s differentiation theorem that there exists a  $z \in [v, v + \epsilon y] \cap G$  such that  $\langle \sigma(z), \epsilon y \rangle > \alpha \epsilon$ , which immediately gives us that  $\langle \sigma(z), y \rangle > \alpha$ . This is a contradiction, since  $z \in U$  and  $\sigma(z) \in \Phi(z) \subseteq \Phi(U)$ , while  $\max \langle \Phi(U), y \rangle < \alpha$ . We consequently have  $\partial f \subseteq \Phi$ , and then  $\partial f = \Phi$  due to the minimality of  $\Phi$ .

We note here that from Theorem 2.5 in [4] it follows that the function  $f$  constructed in the second half of Theorem 3.1 is in fact  $D$ -representable in the sense that for each dense subset  $D$  of  $\{x \in A : \nabla f(x) \text{ exists} \}$  we have,

$$\partial f(x) = \overline{\text{co}}^{w^*} \{x^* : x^* = \text{weak}^* - \lim_{x_n \rightarrow x} \nabla f(x_n)\}.$$

**COROLLARY 3.1.** *Let  $X$  be a separable Banach space, and let  $A$  be a non-empty open connected subset of  $X$ . Then for a minimal weak\* cusco  $\Phi : A \rightarrow 2^{X^*}$  to be the Clarke subdifferential mapping of a  $D$ -representable locally Lipschitz function (with uniform rank  $L$ ) it is necessary and sufficient that (i)  $\Phi$  be bounded by  $L$  and (ii) that there exist a Borel set  $B \subset A$  with  $A \setminus B$  Haar-null and a measurable selection  $\sigma : B \rightarrow X^*$  such that for each  $\epsilon > 0$  and each closed  $B$ -admissible  $\epsilon$ -path  $P(\epsilon)$  in  $A$ , one has*

$$\left| \oint_{P(\epsilon)} \sigma(x) dx \right| < L\epsilon.$$

One of the disadvantages of Theorem 3.1 is that a priori it is not clear how to test whether a given set-valued mapping has a well-behaved selection. There is one situation however where this problem is resolved, namely, the situation in which the set-valued mapping is single-valued almost everywhere. In this case the selection is essentially uniquely defined. Moreover, we have the following.

**THEOREM 3.2.** *Let  $A$  be a non-empty open connected subset of a separable Banach space  $X$ . Then any minimal weak\* cusco  $\Phi : A \rightarrow 2^{B(X^*)}$  for which  $N \equiv \{x \in A : \Phi(x) \text{ is not single-valued} \}$  is Haar-null is the Clarke subdifferential mapping of a unique (up to an additive constant) locally Lipschitz function  $f$  provided that for each  $\epsilon > 0$  and each closed  $(A \setminus N)$ -admissible  $\epsilon$ -path  $P(\epsilon)$  in  $A$  we have*

$$\left| \oint_{P(\epsilon)} \Phi(x) dx \right| < \epsilon.$$

**PROOF.** From Theorem 3.1 we have that there exists a locally Lipschitz function  $f$  such that  $\partial f = \Phi$ . The uniqueness of  $f$  follows from the fact that since  $\partial f$  is single-valued almost everywhere it must be essentially smooth (see [4]) and so integrable on  $A$ .

## References

- [1] T. M. Apostol, *Mathematical Analysis*, first ed. (Addison-Wesley, Reading, MA, 1957).
- [2] T. M. Apostol, *Mathematical Analysis*, second ed. (Addison-Wesley, Reading, MA, 1974).
- [3] J. M. Borwein and S. Fitzpatrick, "Characterization of Clarke subgradients among one-dimensional multifunctions", in *Proceedings of the Optimization Miniconference II*, (Univ. Ballarat Press, 1995), 61–73.
- [4] J. M. Borwein and W. B. Moors, "Essentially smooth Lipschitz functions", *J. Functional Analysis* **149** (1997) 305–351.
- [5] J. M. Borwein and W. B. Moors, "Lipschitz functions with minimal Clarke subdifferential mappings", in *Proceedings of the Optimization Miniconference III*, (Univ. Ballarat Press, 1997), 5–12.
- [6] J. M. Borwein and W. B. Moors, "Null sets and essentially smooth Lipschitz functions", *SIAM J. Optimization* **8** (1998), 309–323.
- [7] J. P. R. Christensen, *Topological and Borel Structure* (American Elsevier, New York, 1974).
- [8] F. H. Clarke, *Optimization and Nonsmooth Analysis* (Wiley Inter-science, New York, 1983).
- [9] R. Correa, A. Jofré and L. Thibault, "Subdifferential monotonicity as characterization of convex functions", *Numer. Funct. Anal. Optimiz.* **15** (1994) 531–535.
- [10] R. Janin, "Sur des multiapplications qui sont des gradients généralisés", *C. R. Acad. Sc. Paris* **294** (1982) 115–117.
- [11] G. Minty, "Monotone (nonlinear) operators in Hilbert spaces", *Duke Math. J.* **29** (1962) 341–346.
- [12] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability* (Springer-Verlag, New York, 1993).
- [13] R. A. Poliquin, "A characterization of proximal subgradient set-valued mappings", *Can. Math. Bull.* **36** (1993) 116–122.
- [14] R. T. Rockafellar, "Favorable classes of Lipschitz continuous functions in subgradient optimization", in *Progress in Nondifferentiable Optimization*, (Institute of Applied Systems Analysis, Laxenburg, Austria, 1982) 125–144.
- [15] L. Thibault, "On generalized differentials and subdifferentials of Lipschitz vector-valued functions", *Nonlinear Anal. TMA* **6** (1982) 1037–1053.