Pathological Lipschitz Functions In R^N

by

Xianfu Wang

B.Sc., Lanzhou University, 1986

M.Sc., Central South University of Technology, 1989

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in the Department
of
Mathematics & Statistics

SIMON FRASER UNIVERSITY
June 1995

APPROVAL

Name: Xianfu Wang

Degree: Master of Science

Title of thesis: Pathological Lipschitz Functions In R^n

Examining Committee: Dr. G. A. C. Graham
Chair

Dr. J. M. Borwein
Senior Supervisor

Dr. P. B. Borwein

Dr. A. Das

Dr. J. D. Vanderwerff
External Examiner

Date Approved: June 19, 1995

Abstract

In recent years four subdifferential maps have been widely used: the **Clarke subdifferential**, the **Michel–Penot subdifferential**, the **Ioffe–Mordukhovich–Kruger approximate subdifferential**, and the **Dini subdifferential**. We denote these four notions by 'C', 'MP', 'A', and 'D' respectively. Each of them is a generalization from convex to locally Lipschitz functions and each of them generalizes different aspects of the convex situation. In this thesis, we construct Lipschitz functions with pathological properties and study the differences among these four subdifferential maps.

Chapter 1 contains some basic concepts and notation from nonsmooth analysis. These are: subdifferentiability, subderivative, minimal usco, minimal cusco, regularity, integrability, and saineness of a Lipschitz function.

In Chapter 2, we consider our subdifferential maps on the real line. For differentiable functions we prove that most functions have C–subdifferentials which are singleton almost nowhere and most have C–subdifferentials and MP–subdifferentials which differ almost everywhere. We also show that C–integrability and A–integrability coincide on the real line for any locally Lipschitz function. We then show that the C–subdifferential and A–subdifferential can be different at most on a countable set for any locally Lipschitz function. Finally using Borwein and Fitzpatrick's theorem we construct Lipschitz functions which are regular on sets with small measure.

In Chapter 3, we consider the inverse problem. We provide a technique for constructing Lipschitz functions with prescribed subdifferentials. More precisely we show that if f_1, f_2,..., f_k possess minimal C–subdifferential mappings on an open set U, then there exists a real–valued locally Lipschitz function g defined on U such that:

$$\partial_c g(x) = conv\{\partial_c f_1(x), \partial_c f_2(x), \cdots, \partial_c f_k(x)\} \quad \text{for each } x \in U.$$

As a result of this, we deduce that given a finite family of maximal **cyclically monotone** operators $\{T_1, T_2, ..., T_k\}$ on an open set U there exists a real–valued locally Lipschitz function g defined on U such that:

$$\partial_c g(x) = conv\{T_1(x), T_2(x), \cdots, T_k(x)\} \quad \text{for each } x \in U.$$

In particular, we obtain that given any convex polytope P in a finite dimensional space X there is a locally Lipschitz function f such that $\partial_c f(x) = P$ for each $x \in U$. This shows that without restrictions, the C–subdifferential of a locally Lipschitz function can be a somewhat unwieldy beast.

In Chapter 4, we investigate **bump functions**. We begin by showing that any strictly convex body containing 0 is the gradient range of a smooth bump function and that the ranges of C–subdifferentials and A–subdifferentials always contain 0 as an interior point. We use bump functions to construct a Lipschitz function in R^2 such that its C–subdifferential and A–subdifferential differ on a set with large measure. Then we show that there is a Lipschitz function in R^2 such that its A–subdifferential has nonconvex images almost everywhere. Finally we construct two Lipschitz functions with the same C–subdifferential but with their A–subdifferentials differing except on a set with small measure.

Dedication

Dedicated to my son and my wife.

Acknowledgements

I am very grateful to my supervisor J. M. Borwein for giving me the opportunity to study under his direction. During my study and writing this thesis, he has given me a lot of help and encouragement.

I would like to thank Dr. P. B. Borwein, Dr. A. Das, Dr. J. D. Vanderwerff, and Dr. H. H. Bauschke for their valuable advice and suggestions on my thesis. I would like also to thank the many people in the CECM and the department who make such a pleasant study environment.

Contents

Chapter 1

Introduction

1.1 Directional derivatives and subgradients

Let $f : U \subset R^N \to R$ be a locally Lipschitz function on an open set U. That is, for each $x \in U$ there is a neighbourhood $N(x) \subset U$ satisfying:

$$\|f(z) - f(y)\| \leq L\|z - y\| \text{ for all } z, y \in N(x).$$

Assume that a set $\partial f(x) \subset X^*$ (possibly empty) is associated with f and every $x \in U$ in such a way that the following is true:

(1) $0 \in \partial f(x)$ if f attains a local minimum at x;

(2) ∂f is upper semicontinuous at x in the following sense:

$$\partial f(x) = \limsup_{u \to x} \partial f(u) = \bigcap_{\delta > 0} \bigcup_{u \in x + \delta B(0)} \partial f(u)$$

where $B(0)$ is the open unit ball in X;

(3) For a C^1 function $\partial f = \nabla f$. For a convex function ∂f is the **subdifferential** in convex analysis, namely

$$\partial f(x) := \{x^* \in X^* : \langle x^*, h \rangle \leq f(x + h) - f(x), \text{ for all } h \in X\};$$

(4) $\partial(f + g)(x) \subset \partial f(x) + \partial g(x)$, provided that g is convex continuous or is C^1.

If these conditions are satisfied, we call ∂f a subdifferential map. In recent years four subdifferential maps have been widely used: the **Clarke subdifferential, Michel–Penot subdifferential, Ioffe–Mordukhovich–Kruger approximate subdifferential**, and **Dini subdifferential.** We denote these four notions by 'C' , 'MP', 'A', and 'D' respectively. Each of them is a generalization from convex to locally Lipschitz functions and each of them generalizes different aspects of the convex situation.

The Clarke subdifferential map satisfies (1), (2), (3), (4). $\partial f(x)$ is w^* closed, convex, and nonempty for each $x \in U$. It is singleton if and only if f is strictly differentiable at x. In R^N it can be defined by

$$\partial_c f(x) = conv\{\lim \nabla f(x_i) : x_i \to x, x_i \notin S, x_i \notin \Omega_f\}$$

where S is any set of Lebesgue measure 0 in R^N and Ω_f is the set on which f is not Gateaux differentiable [13].

The Michel–Penot subdifferential satisfies (1), (3), (4) but (2) fails. $\partial f(x)$ is convex, w^* closed, and nonempty for each $x \in U$. It is singleton if and only if f is Gateaux differentiable at x. Often $\partial_{mp} f$ is much smaller than $\partial_c f$.

The Dini subdifferential map satisfies (1), (3) but not (2) and (4). $\partial^- f$ is w^* closed, convex and much smaller than the approximate subdifferential. Even for Lipschitz functions it is possibly empty.

The Ioffe–Mordukhovich–Kruger approximate subdifferential map satisfies (1), (2), (3), (4). In R^N

$$\partial_a f(x) := \{\lim_{n\to\infty} x_n^* \in \partial^- f(x_n),\ x_n \to x\}$$

and

$$\overline{conv}\partial_a f(x) = \partial_c f(x).$$

If f is Lipschitz at x then $\partial_a f(x)$ is not empty. It is the minimal subdifferential that satisfies (1), (2), (3), (4) in the sense that $\partial_a f(x) \subset \partial f(x)$ for any $x \in U$ [18, 20].

Associated with these subdifferential maps are four directional derivatives:

- the **Clarke derivative** at point x in the direction h is given by

$$f^o(x;h) := \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{f(y+th)-f(y)}{t}.$$

- **Michel-Penot derivative** at x in the direction h is given by
$$f^{\diamond}(x;h) := \sup_{y} \limsup_{t\downarrow 0} \frac{f(x+ty+th)-f(x+ty)}{t}.$$
- **Upper Dini derivative** at x in the direction h is given by
$$f^{+}(x;h) := \limsup_{t\downarrow 0} \frac{f(x+th)-f(x)}{t}.$$
- **Lower Dini derivative** at x in the direction h is given by
$$f^{-}(x;h) := \liminf_{t\downarrow 0} \frac{f(x+th)-f(x)}{t}.$$

The first two derivatives are Lipschitz and sublinear functions of h while the latter two derivatives are only Lipschitz functions of h. Moreover
$$f^{-}(x;\cdot) \le f^{+}(x;\cdot) \le f^{\diamond}(x;\cdot) \le f^{o}(x;\cdot).$$

If f is convex, then
$$f^{o}(x;\cdot) = f^{\diamond}(x;\cdot) = f^{+}(x;\cdot) = f^{-}(x;\cdot).$$

The link between the generalized subdifferential map and the directional derivative is:
$$\partial^{\sharp} f(x) = \{x^* \in E^* : \langle x^*, h\rangle \le f^{\sharp}(x;h) \text{ for all } h \in E\}$$
where '$\sharp$' is one of '$+$', '$-$', '$\diamond$', 'o'. If we set further $U(x) := \{z|\ \|x-z\| < \delta\}$, then
$$\partial_a f(x) = \bigcap_{\delta>0} \bigcup_{z\in U(x)} \partial^{-} f(z) = \limsup_{z\to x} \partial^{-} f(z).$$

Let us begin with some examples to illustrate the advantages and disadvantages of these subdifferentials.

Example 1 Define
$$R \ni x \mapsto f(x) := |x|.$$

This function achieves its minimum at 0 and $-f$ achieves its maximum at 0. $\partial_c f(0) = [-1,1] = -\partial_c(-f)(0)$ and $\partial_a f(0) = [-1,1]$ but $\partial_a(-f)(0) = \{-1,1\}$. This shows $0 \in \partial_c f(0)$ and $0 \in \partial_c(-f)(0)$ whereas $0 \in \partial_a f(0)$ and $0 \notin \partial_a(-f)(0)$. For the C–subdifferential $0 \in \partial_c f(x)$ is always true whenever f achieves its minimum or maximum at x; for the A–subdifferential it is always true $0 \in \partial_a f(x)$ whenever f attains its minimum but not its maximum at x. $\square$

Example 2 Define

$$f(x) := \begin{cases} x^2 \cdot \sin(1/x), & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{cases}$$

Then f is differentiable everywhere on R, but $\partial_c f(0) = \partial_a f(0) = [-1, 1]$ whereas $\partial_{mp} f(0) = \partial^- f(0) = \{0\}$.

Moreover, for any $0 < \epsilon < 1$ there is an everywhere differentiable and globally Lipschitz function $f : [0,1] \to R$ such that:

$$\begin{aligned} & \mu\{x : \ \partial_c f(x) = [-1,1], \ x \in [0,1]\} \\ = \ & \mu\{x : \ f \text{ is not strictly differentiable at } x, \, x \in [0,1]\} \\ = \ & \epsilon. \end{aligned}$$

Indeed, we take a Cantor set $P \subset [0,1]$ with $\mu(P) = \epsilon$. As shown in Lemma 2.4, there is an everywhere differentiable and globally Lipschitz $g : [0,1] \to R$ such that:

$$\partial_c g(x) = [0,2] \quad \text{for all } x \in P$$

and

$$\partial_c g(x) = g'(x) \quad \text{for all } x \in [0,1] \setminus P.$$

Define f by $f(x) := g(x) - x$. Then $\partial_c f(x) = \partial_c g(x) - 1$. Hence $\partial_c f(x) = [-1,1]$ for all $x \in P$ and $\partial_c f(x) = g'(x) - 1$ for all $x \in [0,1] \setminus P$. By contrast, $\partial_{mp} f(x) = \partial^- f(x) = g'(x) - 1$ for any $x \in [0,1]$. □

Example 3 Define

$$f(x) := \begin{cases} x \cdot \sin(\log |x|), & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{cases}$$

At $x = 0$, $\partial_a f = \partial_c f = [-\epsilon - \sqrt{2}, \epsilon + \sqrt{2}]$ and $\partial^- f(0) = \emptyset$. In particular $0 \in \partial_a f(0) = \partial_c f(0)$. But for any $0 \leq \epsilon < 1$, $f + \epsilon |\cdot|$ does not attain a local minimum at 0.

On the other hand, the Dini subdifferential has one important property. That is, for any locally Lipschitz function f (even for lower semicontinuous function) the function $f + \epsilon \|\cdot - x\|$ attains a strict local minimum at x for any $\epsilon > 0$ if $0 \in \partial^- f(x)$. □

The C–subdifferential and the MP–subdifferential are symmetric but the A–subdifferential and the D–subdifferential are not symmetric.

Example 4 Let $\|\cdot\|$ be the l^2 norm. Define $f := (2 - \|\cdot\|)^+$. Computing its Dini subgradient at 0 we get

$$\partial^- f(0) = \emptyset \text{ and } \partial^-(-f)(0) = B(0,1).$$

Moreover,

$$\partial_a f(0) = \mathrm{bd}B(0,1) \text{ and } \partial_a(-f)(0) = B(0,1),$$

$$\partial_{mp} f(x) = -\partial_{mp}(-f)(x) = \partial_c f(0) = -\partial_c(-f)(0) = B(0,1).$$

□

Example 5 Let $f : R^N \to R$ be a concave continuous function. Then for any $x \in R^N$

$$\partial_c f(x) = conv\{\limsup_{y \to x}\{\nabla f(y)\}\}$$

$$\partial_a f(x) = \limsup_{y \to x}\{\nabla f(y)\}$$

and $\partial^- f(x)$ is non–empty if and only if f is Gateaux differentiable at x. In this case $\partial^- f(x) = \{\nabla f(x)\}$ (from which we can easily see the Dini subdifferential lacks upper semicontinuity).

Let $f : R^N \to R$ be a convex function. Then

$$\partial_c f(x) = \partial_a f(x) = \partial^- f(x) = \partial_{mp} f(x)$$

for all $x \in R^N$. □

We close this section with two pathological examples.

Example 6 We follow the construction in [10, page 190]. The function $f(x) := (x-a)^{\frac{1}{3}}$ has an infinite derivative at $x = a$, and a finite derivative elsewhere. Let q_1, q_2,... be an enumeration of $Q \cap [0,1]$, and for each $n \in N$ let $f_n(x) := (x - q_n)^{\frac{1}{3}}$. Let

$$F(x) := \sum_{n=1}^{\infty} \frac{f_n(x)}{10^n}.$$

Then F is continuous on $[0,1]$. Since each f_n is strictly increasing, so is F. It can be shown that

$$F'(x) = \sum_{n=1}^{\infty} \frac{f_n'(x)}{10^n} = \sum_{n=1}^{\infty} \frac{(x-q_n)^{-\frac{2}{3}}}{3 \cdot 10^n}.$$

In particular $F'(x) = \infty$ for all $x \in Q \cap [0,1]$.

Let $g(x) := F(x) + x$. Then $g'(x) = F'(x) + 1$ on $[0,1]$ and g is a homeomorphism from $[0,1]$ to a nondegenerate interval $[a,b]$. Define h by $h(x) := g^{-1}(x)$, that is the inverse function of g. Then h is strictly increasing and continuous on $[a,b]$, and $h' = 0$ on a dense set of $[a,b]$. Since $g'(x) \geq 1$ for all $x \in [0,1]$, $h'(x) \leq 1$ for all $x \in [a,b]$. Hence we have constructed a strictly increasing function h on $[a,b]$ such that h is differentiable, $0 \leq h'(x) \leq 1$ for all $x \in [a,b]$ and $\{x : h'(x) = 0\}$ is dense in $[a,b]$; thus h is a globally Lipschitz function on $[a,b]$. Since $\partial_a h$ and $\partial_c h$ are uscos, we have

$$0 \in \partial_a h(x) \text{ and } 0 \in \partial_c h(x) \text{ for all } x \in [a,b].$$

However

$$\partial_{mp} h = \partial^- h = h'.$$

In this case only the MP–subdifferential and the D–subdifferential can adequately reflect the properties of h. We note that $\partial_c h = \partial_a h$ and that they are generically single–valued. ◫

Example 7 Let $A \subset R$ be measurable such that

$$0 < \mu(A \bigcap I) < \mu(I)$$

for every nonvoid open interval $I \subset R$ (see Exercise 6 [28, page 307]). Then $A' := R \setminus A$, the complement of A, has the same property as A. Let χ_A and $\chi_{A'}$ be the characteristic functions of these two sets respectively. Now define $G : R \to R$ by

$$G(x) := \int_0^x [\chi_A(s) - \chi_{A'}(s)]ds.$$

Then G is globally Lipschitz but is monotonic on no interval on R and

$$\partial_c G(x) = \partial_a G(x) = [-1,1] \text{ for any } x \in R.$$

It is clear that we get no information about the function G from the C–subdifferential map and A–subdifferential map. In fact there are uncountably many such subsets like A in R. Taking different such subsets we get different globally Lipschitz functions. All of these functions share the same A–subdifferential and C–subdifferential.

In R^2 we define $F : R^2 \to R$ by $F(x,y) := G(x) + G(y)$. Then

$$\partial_c F(x,y) = \partial_a F(x,y) = [-1,1] \times [-1,1] \text{ on } R^2.$$

F takes the l^2_∞ unit ball as its C–subdifferential and A–subdifferential identically and they are too big. □

1.2 Minimal uscos and minimal cuscos

Definition 1.1 *A multifunction F between topological spaces X and Y is called an* **usco** *(***cusco***) if it is compact valued (convex and compact) and upper semicontinuous. That is, $\{t : F(t) \subset W\}$ is open in X if W is open in Y. It is called a* **minimal usco** *(***minimal cusco***) if it is an usco (cusco) whose graph is minimal with respect to set containment among uscos (cuscos). When $\partial^\sharp f$ is a minimal w^*-usco (w^*-cusco), we will say f possesses a* **minimal $\sharp$–subdifferential**.

Example 8 Let $\Omega : R^N \to R^N$ be a densely defined locally bounded multifunction. By Proposition 1.2 and 1.3 [1] we have:

(i) There is a unique minimal cusco containing Ω and it is given by

$$CSC(\Omega) := conv\{a : \exists a_n \to a, \exists x_n \to x, a_n \in \Omega(x_n)\}.$$

(ii) There is a unique minimal usco containing Ω and it is given by

$$USC(\Omega) := \{a : \exists a_n \to a, x_n \to x, a_n \in \Omega(x_n)\}.$$

□

The relationship between minimal cuscos and minimal uscos is the following:

Theorem 1.1 *[2] If $T : X \to Y$ is an usco then $T^*(x) := \overline{conv}T(x)$ defines a cusco T^* on X. If T is a minimal usco then T^* is a minimal cusco.*

In finite dimensions every minimal cusco or usco defined on an open set is generically single–valued but a generically single–valued usco or cusco needs not be minimal (see Example 6). A cusco (usco) T defined on an open set U is minimal if and only if for each non–empty open subset $W \subset U$ the restriction of T to W is a minimal cusco (minimal usco) on W [6]. That is, minimality is inherited by open subsets.

Example 9 Let $f : U \to R$ be locally Lipschitz on U with $\partial_c f$ almost everywhere single–valued. Then $\partial_c f$ is a minimal cusco. In finite dimensions we have

$$\partial_c f(x) = conv\{\lim \nabla f(x_i) : x_i \to x, x_i \notin S, x_i \notin \Omega_f\}$$

where S is any set of Lebesgue measure 0 and Ω_f is the set of points at which f fails to be differentiable. Let S be the null set where $\partial_c f$ fails to be single–valued. Set $N := S \cup \Omega_f$. Then $\partial_c f = CSC(\nabla f|(U \setminus N)) = CSC(\partial_c f)_S$. By Theorem 2.2(b) [1], $\partial_c f$ is minimal. □

Example 10 Let R^N be endowed with a smooth norm. For each non–empty closed subset $C \subset R^N$, we define its distance function $d_C : R^N \to R$ by

$$d_C(x) := \inf\{\|x - c\| : \ c \in C\}.$$

Then d_C has a minimal C–subdifferential on R^N. Indeed by Theorem 4.10 [6], d_C possesses a minimal C–subdifferential on R^N if and only if $\partial_c d_C$ is minimal on $R^N \setminus C$. However $-d_C$ is C–regular on $R^N \setminus C$, so d_C is essentially strictly differentiable on $R^N \setminus C$. Hence d_C is C–minimal on $R^N \setminus C$. □

Definition 1.2 *A set–valued map T from R^N into R^N is said to be a* **monotone operator** *provided*

$$\langle x^* - y^*, x - y\rangle \geq 0$$

whenever $x, y \in R^N$ and $x^* \in T(x)$, $y^* \in T(y)$. We say that a monotone operator T is **maximal monotone** *if its graph is a maximal monotone set with respect to set inclusion.*

Let $T : R^N \to 2^{R^N}$ be monotone and $D \subset R^N$ be an open set, we say that T is maximal monotone in D provided the monotone set

$$G(T) \cap (D \times R^N) \equiv \{(x, x^*) \in D \times R^N : x \in D \text{ and } x^* \in T(x)\}$$

is maximal (under set inclusion) in the family of all monotone sets contained in $D \times R^N$.

Any monotonic function from R into itself is a monotone operator. The subdifferential of a convex function defined on an open set is a maximal monotone operator [26]. As shown in Theorem 7.9 [25], every maximal monotone operator defined on an open set is a minimal cusco in D but it needs not be a minimal usco. We should also note that not every minimal cusco is monotone since any continuous non–monotonic function from R into itself is a counter–example.

Example 11 Any minimal cusco Ω from R into R is a C–subdifferential.

To see this, define $f(t) := \sup \Omega(t)$ and $g(t) := \inf \Omega(t)$. Then f and g are locally–bounded measurable functions and $\{f(t)\} \cup \{g(t)\} \subset \Omega(t)$ for all t. Set

$$F(x) := \int_0^x f(t)dt \text{ and } G(x) := \int_0^x g(t)dt.$$

Since Ω is minimal we get $\partial_c F = \partial_c G = \Omega$.

Let I be an open interval in R and suppose $T : I \to 2^R$ has nonempty values and is maximal monotone in I. As shown in Theorem 7.9 [25] T is a minimal cusco in I. By Theorem 2.28 [25] T is locally bounded in I. Rockafellar has shown that if T is a maximal cyclically monotone operator on a Banach space E with $D(T) \neq \emptyset$ then there exists a proper lower semicontinuous convex function on E such that $T = \partial_c f$. He constructed the function using the cyclical property of T. The proof is much simplified for a maximal cyclically monotone T from R to 2^R. In fact, letting $s(t) := \sup \Omega(t)$ we see that s is a locally bounded and measurable selection of T since $T(x)$ is closed for each $x \in I$ and locally bounded. Set $f(x) := \int_0^x s(t)dt$. Then $\partial_c f = T$ in I. Proposition 2.2.9 [13] says that a locally Lipschitz function defined on an open convex set U is convex if and only if $\partial_c f$ is monotone on U. Since T is monotone, so f is actually convex on I. □

As shown in [1], convex functions, saddle functions, and the difference of any two continuous convex functions defined on an open set always have minimal C–subdifferential maps. In any separable Banach space regular, pseudo–regular, and essentially strictly differentiable locally Lipschitz functions (see Definition 1.5 and Definiton 1.6) have minimal C–subdifferential maps. Moreover the C–subdifferential map is the smallest cusco containing the gradient but the A–subdifferential map is not the smallest usco containing the gradient:

Example 12 Consider $f(x) := |x|$. Then

$$\partial_a f(x) = \begin{cases} -1, & \text{if } x < 0 \\ [-1,1], & x = 0 \\ 1, & \text{if } x > 0 \end{cases}$$

So $\partial_a f$ is the smallest cusco containing ∇f but is not the smallest usco containing ∇f since we can remove $(-1,1)$ from $\partial_a f(0)$; f is C–minimal and $-f$ is A–minimal. □

Definition 1.3 *A locally Lipschitz function f defined on an open convex set U is* $\sharp$**–integrable** *if for any g satisfying $\partial_\sharp g(x) \subset \partial_\sharp f(x)$ for all $x \in U$ we may deduce that f and g differ only by a constant.*

Example 13 Let '$\sharp$' denote 'c', 'mp'. Following Proposition 4.4 [1], we give a similar result for 'mp'–integrability. Suppose $U \subset R^N$ is open and convex and that f and g are locally Lipschitz functions such that

$$\partial_\sharp g(x) \subset \partial_\sharp f(x) \text{ for each } x \in U.$$

Then f is $\sharp$–integrable if and only if $\partial_\sharp(f-g)$ is $\sharp$–minimal.

To see this, letting $h := f - g$, we have $\partial_\sharp f \subset \partial_\sharp h(x) + \partial_\sharp g(x) \subset \partial_\sharp h(x) + \partial_\sharp f(x)$. Since $\partial_\sharp f(x)$ and $\partial_\sharp h(x)$ are closed, bounded convex sets so $0 \in \partial_\sharp h(x)$ for all $x \in U$. If $\partial_\sharp h$ is minimal then $\partial_\sharp h(x) = \{0\}$ and the Mean Value Theorem [13, 4] shows h is constant. □

In a separable Banach space every function that is strictly Gateaux differentiable except possibly at points of a Haar–null set is C–integrable and A–integrable [1, 18]. In particular convex functions are C–integrable and A–integrable. We note that in a Banach space if a locally Lipschitz function is C–integrable then it is A–integrable [18]. In R both integrabilities coincide for all locally Lipschitz functions. It is very important to note that integrability is not inherited by open subsets (see [6]). Minimality and integrability are closely related:

Example 14 Every C–integrable Lipschitz function on R has a minimal C–subdifferential. To see this, suppose $\partial_c f$ is not C–minimal then there is a minimal cusco $s : R \to 2^R$ such that $s(t) \subset \partial_c f(t)$ for all t. Just as in Example 11, we can find a locally Lipschitz function g satisfying $\partial_c g(t) = s(t)$ for any $t \in R$. Thus $\partial_c g \subset \partial_c f$ but $f - g$ is not a constant, a contradiction. □

Example 15 Any function f defined on R with **Riemann** integrable f' is both C–integrable and C–minimal. Since f' is Riemann integrable if and only if f' is continuous almost everywhere, this implies f is strictly differentiable almost everywhere. □

We close this section with the Goffman function showing that a function with minimal subdifferentials needs not be integrable. For examples in several dimensions see Theorem 4.6.

Example 16 We follow the construction given in [1]. Let C be a Cantor set in $[0,1]$ corresponding to the sequence $a_n = 2^{-n-1} + 3^{-n} \cdot 2^{-1}$ and let I_n's be an enumeration of the component intervals of $[0,1] \setminus C$. Then $\mu(C) = \frac{1}{2}$ and $\sum_{n=1}^{\infty} |I_n| = \frac{1}{2}$. Let J_n be the open interval concentric with I_n for which $|J_n| = |I_n|^2$. Let $f_n : [0,1] \to [0,1]$ be a function that is continuous on $[0,1]$, is equal to 1 at the midpoint c_n of I_n, and is identically 0 on $[0,1] \setminus J_n$. Define

$$f(x) := \sum_{n=1}^{\infty} f_n(x).$$

Then f is continuous on each I_n, discontinuous at each point of C, and is not Riemann integrable over $[0,1]$. However, f is Lebesgue integrable. Define the **Goffman function** by

$$F(x) := \int_0^x f(s)ds.$$

As shown in [28, page 279] $F'(x) = 0 = f(x)$ for all $x \in C$ and $F'(x) = f_n(x) = f(x)$ if $x \in I_n$ for some n. Since each point of C is a limit point of the set $\{c_n\}_{n=1}^{\infty}$ we obtain:

$$\partial_c F(x) = [0,1] \ \text{ if } x \in C$$

and

$$\partial_c F(x) = f(x) \text{ if } x \notin C.$$

Moreover, ∂F is both C–minimal and A–minimal since any selection agrees with ∂F on $[0,1] \setminus C$. We show F is not C–integrable. Indeed, let $h : [0,1] \to [0,1]$ be any Lebesgue measurable function with $\text{supp}(h) \subset C$ and let $S(x) := \int_0^x h(s)ds$. Then since $\nabla S = 0$ in $[0,1] \setminus C$ and $0 \le h \le 1$ in C, $\partial_c(F+S) = \partial_c F$ but $S = (F+S) - F$ is not a constant on $[0,1]$.

Let $h := \chi_C$. Since F is differentiable, $\partial^-(F+S)(x) = F'(x) + \partial^- S(x)$ for all $x \in [0,1]$. On $[0,1] \setminus C$, $\partial^-(F+S) = F'$; On C, $\partial^-(F+S) = \partial^- S$. Noting that g_n is continuous for each n we have

$$\partial_a(F+S)(x) = \partial_c(F+S)(x) = [0,1] \text{ if } x \in C$$

and

$$\partial_a(F+S)(x) = \partial_c(F+S)(x) = f(x) \text{ if } x \notin C.$$

This shows F is not A–integrable either even though F has a minimal A–subdifferential. □

1.3 Regularity and saineness of Lipschitz functions and the essentially strictly differentiable functions

In this section we suppose that U is a connected open set in R^N.

Definition 1.4 *We say a function $f : U \to R$ is* **saine** *provided f is locally Lipschitz and $\langle \partial_c f(x(t)), x'(t)\rangle$ is almost everywhere single–valued for all absolutely continuous $x : [0,1] \to U$. We denote the set of saine Lipschitz functions by $L_s(U)$.*

Definition 1.5 *A locally Lipschitz function defined on U is called C–***regular** *at $x \in U$ if $f^-(x;\cdot) = f^+(x;\cdot) = f^\diamond(x;\cdot) = f^o(x;\cdot)$ and C–***pseudo-regular** *if the latter three coincide ($f^+(x;\cdot) = f^o(x;\cdot)$). We say f is C–regular or C–pseudo-regular on U if f is C–regular or C–pseudo-regular at each point $x \in U$. We denote the set of C–regular functions by $L_r(U)$.*

Definition 1.6 *A locally Lipschitz function defined on U is* **essentially strictly differentiable** *on U if it is strictly Gateaux differentiable on U except possibly at points of a Lebesgue–null set. We denote the set of essentially strictly differentiable functions by $S_e(U)$.*

Any convex function $f : U \to R$ is saine, regular, and essentially strictly differentiable [13]. The three classes of functions are well-behaved. By Corollary 3.10 and Corollary 4.6 in [1] all functions in the three classes are C–integrable and C–minimal. By Corollary 5.16 [6] and Theorem 2 [32] we know both $S_e(U)$ and $L_s(U)$ are closed under addition, subtraction, multiplication and division (when this is defined), as well as, the lattice operations. We give some facts from [31] and [32] to show other relations.

Fact 1.1 In R we have $f \in S_e(U) \Leftrightarrow f \in L_S(U)$.

To see this, letting $f \in L_s(U)$, then we take $x(t) := t$. Hence $\langle \partial_c f(x(t)), x'(t)\rangle = \partial_c f(t)$. By definition $\partial_c f$ is single–valued almost everywhere and so $f \in S_e(U)$. Conversely, letting $x : [0,1] \to U$ be any absolutely continuous curve, by Theorem 6.93 [28] we have:

$$(f \circ x)'(t) = f'(x(t)) \cdot x'(t) \text{ a.e. on } [0,1].$$

Since $f \in S_e(U)$ we have $\partial_c f(x(t)) = f'(x(t))$ whenever the derivative of f exists at $x(t)$. Therefore

$$(f \circ x)'(t) = \langle \partial_c f(x(t)), x'(t)\rangle \text{ a.e. on } [0,1]$$

and $f \in L_s(U)$. □

Fact 1.2 Let $f \in L_r(U)$. Then $f \in L_s(U)$.

To see this, by the Lebourg Mean Value Theorem [13] there exists a point u in the line segment $[x(t+h), x(t)]$ such that:

$$f(x(t+h)) - f(x(t)) \in \langle \partial_c f(u), x(t+h) - x(t) \rangle.$$

Since f is locally Lipschitz and x is absolutely continuous, $(f \circ x)'$ exists a.e. on $[0,1]$. Let $t \in [0,1]$ be such a point, then letting $h \to 0$ we get:

$$\begin{aligned} (f \circ x)'(t) &\in \langle \partial_c f(x(t)), x'(t) \rangle \\ &= [-f^o(x(t); -x'(t)), f^o(x(t); x'(t))] \\ &= [-f'(x(t); -x'(t)), f'(x(t); x'(t))] \\ &= [(f \circ x)'(t-), (f \circ x)'(t+)]. \end{aligned}$$

Since $(f \circ x)'(t-) = (f \circ x)'(t+)$ a.e on $[0,1]$, so $f \in L_s(U)$. □

Fact 1.3 Let $f \in L_s(U)$. Then $f \in S_e(U)$.

To see this, it suffices to consider affine maps $x_k(t) := \bar{x} + t \cdot e_k$ for $k = 1, 2, \ldots, N$. Then

$$\langle \partial_c f(\bar{x} + t \cdot e_k), e_k \rangle = (\partial_c f)_k(\bar{x} + t \cdot e_k) \text{ is singleton a.e. on } [0,1].$$

So $(\partial_c f)_k(x_1, \ldots, x_{k-1}, t, x_{k+1}, \ldots, x_\eta)$ is single–valued if $t \notin N_k$ which is a Lebesgue null set in R. Let $M := \cup_{k=1}^N \{(N_k \times R^{N-1}) \cap U\}$. It is clear M is a Lebesgue null set in R^N. It follows that if $x \in U \setminus M$ then $\partial_c f(x)$ is singleton, and so $f \in S_e(U)$. □

There are many locally Lipschitz functions which are not in the three classes.

Example 17 Let $C \subset R^N$ be nonempty. We consider its distance function d_C.

(i) Let $C \subset U$ be convex. Then d_C is convex and $d_C \in L_r(U)$;

(ii) Take $C \subset U$ with $\mu(C) > 0$, $\mathrm{int} C = \emptyset$, and C is closed. Then d_C is a.e. C–irregular on C. To see this, letting $x \in C$, we have $0 \in \partial_c d_C(x)$. Since d_C is Lipschitz it is

Gateaux differentiable almost everywhere on C. Let d_C be Gateaux differentiable at $x \in C$ then d_C is not C–regular. Suppose it is C–regular then $\partial_c d_C$ is singleton, that is $\partial_c d_C(x) = \{0\}$. Then $N_C(x) = \{0\}$. But any $x \in C$ is a boundary point of C. By Corollary 1 of Theorem 2.5.6 in [13] $N_C(x)$ contains nonzero points, which is a contradiction. Thus d_C is C–irregular almost everywhere on C. Note that d_C is not strictly differentiable on a positive measure set so $d_C \notin S_e(U)$;

(iii) Let $\|\cdot\|$ be a smooth norm on R^N. Then for each nonempty closed subset C of R^N, we have that $d_C \in S_e(R^N)$ if and only if $\mu(\partial C) = 0$. Indeed, since no point of ∂C can be a point of strict differentiability we immediately have $\mu(\partial C) = 0$. Conversely by Theorem 8 [4], we see that $-d_C$ is C–regular on $(R^N \setminus C) \cup \text{int}C$. Therefore if ∂C is a Lebesgue-null set then d_C is strictly differentiable almost everywhere in R^N.

□

Example 18 In R we will see later that $f \in S_e(U)$ if and only if f' is Riemann integrable. There are many functions which are everywhere differentiable but their derivatives are not Riemann integrable.

Let $P \subset [0,1]$ be a Cantor set with $\mu(P) > 0$. We can construct a Volterra function $F : [0,1] \to R$ as follows (see [28, page 312]). Let $\phi(t) := t^2 \sin(\frac{1}{t})$ for $t \neq 0$. For $x \in P$, put $F(x) := 0$. If (a,b) is a component interval of $[0,1] \setminus P$, let $c := \sup\{t : 0 < t \leq \frac{(b-a)}{2},\ \phi'(t) = 0\}$ and define $F(a+t) = F(b-t) := \phi(t)$ if $0 < t \leq c$ and $F(x) := \phi(c)$ if $a + c \leq x \leq b - c$. Then

(i) F is differentiable on $[0,1]$;

(ii) $F'(x) = 0$ for every $x \in P$;

(iii) $|F'(x)| < 3$ for every $x \in [0,1]$;

(iv) F' is discontinuous at every point of P;

(v) $F(x) = \int_0^x F'(s)ds$ for all $x \in [0,1]$.

Thus $F \notin S_e([0,1])$.

In R^2 we define $F : [0,1] \times [0,1] \to R$ by $F(x,y) := f(x) + f(y)$ where f is a Volterra function. Then F is not strictly Gateaux differentiable on $P \times P$. Therefore $F \notin S_e([0,1] \times [0,1])$. □

Chapter 2

Subdifferentials on the real line

2.1 Everywhere differentiable functions

It is very instructive to discuss everywhere differentiable functions on R to understand what nonsmooth analysts have done. We have different subdifferentials: A natural question is: "Are they really different?" Because nonsmooth analysts generalize the notion of differentiability from a convex analysis point of view, they lose some essential properties for differentiable functions.

Definition 2.1 *A real-valued function f defined on an interval I is called a* **Darboux function** *if it has the intermediate value property. That is, if whenever x_1 and x_2 are in I, and y is any number between $f(x_1)$ and $f(x_2)$, there is a number x_3 between x_1 and x_2 such that $f(x_3) = y$.*

Theorem 2.1 *(Darboux) Let $f : [a, b] \to R$ be differentiable on $[a, b]$ and suppose that ν is a number strictly between $f'_+(a)$ and $f'_-(b)$. Then there exists $\xi \in]a, b[$ such that $f'(\xi) = \nu$.*

Since the proof of the Darboux Theorem essentially uses the Rolle Theorem (see [28, page 186]), we omit it. By the Darboux Theorem we can show that the A–subdifferential and the C–subdifferential coincide for differentiable functions on R.

To see this, let f be an everywhere differentiable function. Then

$$\partial_c f(x) = [\liminf_{y\to x} f'(y), \limsup_{y\to x} f'(y)] \quad \text{and} \quad \partial_a f(x) = \{\lim_{y\to x} f'(y)\}.$$

We claim $\partial_a f(x) = \partial_c f(x)$. Let $\alpha \in \partial_c f(x)$. Then there exists x_n and y_n such that $f'(x_n) \leq \alpha \leq f'(y_n)$ for sufficiently large n. Since f' is a Darboux function we see that there exists z_n with $f'(z_n) = \alpha$. Note that $z_n \in [x_n, y_n]$ so $z_n \to x$. Therefore $\alpha \in \partial_a f(x)$ and $\partial_c f(x) \subset \partial_a f(x)$. Hence $\partial_a f(x) = \partial_c f(x)$.

We can also give a direct proof via the characterization of the Approximate subdifferential in R given in [5]. By Theorem 2.2 [5] we have

$$\partial_a g(x) = [\liminf_{\substack{y \to x \\ y \notin N}} g'(y), \limsup_{\substack{y \to x^+ \\ y \notin N}} g'(y)] \cup [\liminf_{\substack{y \to x^- \\ y \notin N}} g'(y), \limsup_{\substack{y \to x \\ y \notin N}} g'(y)]$$

for any Lebesgue null set N. By the Darboux property there exists two sequences $\{x_n\}$ and $\{y_n\}$ such that $x_n \downarrow x$ with $g'(x_n) \to g'(x)$ and $y_n \uparrow x$ with $g'(y_n) \to g'(x)$. Hence $\partial_a g(x)$ is connected and so $\partial_a g(x) = \partial_c g(x)$. Let $\partial_s g$ denote the **symmetric subdifferential** defined by $\partial_s g(x) := \partial_a g(x) \cap (-\partial_a(-g)(x))$. Then $\partial_s g(x) = \partial_c g(x) = \partial_a g(x)$. Hence we have shown the following:

Theorem 2.2 *Let $f : I \to R$ be differentiable. Then*

$$\partial_c f(x) = \partial_a f(x) = \partial_s f(x) \text{ for all } x \in I.$$

Corollary 2.1 *Let f be differentiable everywhere. Then f is C–integrable if and only if f is A–integrable.*

Proof. By Proposition 3.4 and Proposition 3.3 [18] we know the C-integrability implies A–integrability in finite dimensions. So it suffices to show the other direction. Suppose f is A–integrable. Let g be any locally Lipschitz function such that $\partial_c g(x) \subset \partial_c f(x)$ for any $x \in I$. By Theorem 2.2 we deduce $\partial_a g \subset \partial_c g \subset \partial_a f$. Thus $f - g$ is constant. ☺

Example 19 Stromberg and Katznelson [28, pages 217–218] have constructed an everywhere differentiable $F : R \to R$ which is monotone on no nonvoid open interval of R and $|F'(x)| < 1$ for all $x \in R$. Since F is nowhere monotonic both $\{x \in R : F'(x) < 0\}$ and $\{x : F'(x) > 0\}$ are dense in R. Hence $0 \in \partial_c f(x) = \partial_a f(x)$ for all $x \in R$. This implies F is not C–minimal and not C–integrable and so it is not A–minimal and not A–integrable. However F' is Lebesgue integrable. Here we note that the Lebesgue integrability of F' is not sufficient for F to be C–integrable or A–integrable (see Theorem 2.9). □

2.1.1 The MP–subdifferential and the C–subdifferential may be different almost everywhere

Since the derivative function of any locally Lipschitz everywhere differentiable function f is generically continuous, we know $\partial_c f$ and $\partial_{mp} f$ must agree generically. However we do have the following theorem.

Theorem 2.3 *There exists an everywhere differentiable function defined on* $[0,1]$ *such that* $\partial_c f$ *and* $\partial_{mp} f$ *are different almost everywhere.*

Proof. Let $\{C_m\}_{m=1}^{\infty}$ be a sequence of Cantor sets in $[0,1]$ satisfying:

(i) $C_m \subset C_{m+1}$ for each $m \in \mathbf{N}$;

(ii) $\mu(C_m) = 1 - 2^{-m}$.

Let f_m be the Goffman function associated with C_m (see Example 16) such that:

(1) $f_m'(x)$ exists with $0 \leq f_m'(x) \leq 1$ for each $x \in [0,1]$;

(2) f_m' is continuous on each open interval in $[0,1] \setminus C_m$;

(3) f_m' is discontinuous at each point of C_m and $\partial_c f_m(x) = [0,1]$ for all $x \in C_m$;

(4) $f_m(x) = \int_0^x f_m'(s)ds$ for all $x \in [0,1]$.

Define

$$f(x) := \sum_{m=1}^{\infty} 4^{-m} f_m(x).$$

Set $O_m := [0,1] \setminus C_m$ and $G := \bigcap_{m=1}^{\infty} O_m$. Since $0 \leq f_m' \leq 1$ and $f_m(0) = 0$, by Theorem 4.56 in [28] we have

$$f'(x) = \sum_{m=1}^{\infty} f_m'(x) \quad \text{for all } x \in [0,1].$$

Since f_m is continuously differentiable at each point of G and so f is continuously differentiable at $x \in G$. If $x \notin G$ then choose the first m with $x \notin O_m$. Since $\partial f_m(x) = [0,1]$ and $\partial f_k(x) = f_k'(x)$ for $k < m$, it follows that $\operatorname{diam}[\partial f(x)] > \frac{4^{-m}}{3}$. Thus f is a differentiable Lipschitz function which is strictly differentiable exactly on the null G_δ–set G. In particular the C–subdifferential and MP–subdifferential differ on $[0,1] \setminus G$ with $\mu([0,1] \setminus G) = 1$. ☺

Corollary 2.2 *There is an everywhere differentiable Lipschitz function on* $[0,1]$ *that is almost everywhere C–irregular on* $[0,1]$.

Proof. Since regularity and Gateaux differentiability together imply $\partial_c f$ is singleton, it follows that f is C–irregular on $[0,1]\setminus G$. ☺

2.2 Typical properties

Definition 2.2 *A function* $f:[0,1]\to R$ *is said to be of* **Baire** *class 1 on* $[0,1]$ *if there exists some sequence of real-valued functions which are continuous on* $[0,1]$ *that converges to* f *at every point of* $[0,1]$.

Definition 2.3 *A function* $f:[0,1]\to R$ *is said to be a* **derivative function** *if there is* $F:[0,1]\to R$ *such that* $F'(x)=f(x)$ *for all* $x\in[0,1]$.

It is clear that every derivative function is of Baire class 1. A derivative function can be very badly discontinuous but it can not be discontinuous everywhere [9, 10].

2.2.1 Functions with subdifferentials containing 0 identically

Lemma 2.1 *Let* Δ' *denote the class of derivative functions. Suppose* $f_n\in\Delta'$, $n=1,2,3,\ldots$ *and* $f_n\to f$ *(uniformly), then* $f\in\Delta'$.

Since the proof of this lemma can be found in any of a number of standard texts (see Theorem 4.56 [28]), we omit it.

Let $M\Delta'$ denote the space of bounded derivative functions on $[0,1]$ with

$$\rho(f,g):=\sup_{x\in[0,1]}|f(x)-g(x)|.$$

By Lemma 2.1 we know $(M\Delta',\rho)$ is complete. Let

$$M\Delta'_o:=\{f\in M\Delta':f=0 \text{ on a dense set}\}.$$

Weil has proven the following lemma [9].

Lemma 2.2 *The set of functions in $M\Delta_o'[0,1]$ which are positive on one dense subset of $[0,1]$ and negative on another dense subset of $[0,1]$, forms a residual subset of $(M\Delta_o', \rho)$.*

Proof. *Step 1.* We show that $M\Delta_o'$ is closed under addition and complete topologically. Let $f, g \in M\Delta_o'$. Since f, g are Baire–1 functions, the set $\{x : f(x) = 0\}$ and $\{x : g(x) = 0\}$ are of type G_δ. But the intersection of two dense sets of type G_δ is also of type G_δ and dense. Thus $f + g \in M\Delta_o'$. Similarly, if $\{f_n\}$ is a sequence from $M\Delta_o'$ and $f_n \to f$ uniformly, then $f \in M\Delta_o'$. To see this, let $A_n = \{x : f_n(x) = 0\}$, and $A = \bigcap_{n=1}^{\infty} A_n$. Each of the sets is dense and of type G_δ, so the same is true of A. But $A \subset \{x : f(x) = 0\}$. It follows that $f \in M\Delta_o'$. Thus $M\Delta_o'$ is closed in $M\Delta'$, and therefore complete topologically.

Step 2. Note that the derivative of Stromberg and Katznelson's function (Example 19) belongs to $M\Delta_o'$. Thus $M\Delta_o' \neq \emptyset$. Let I be an interval in $[0,1]$, and let

$$P := \{p \in M\Delta_o' : p \geq 0 \text{ on } I\}.$$

By Step 1, P is closed in $M\Delta_o'$.

Now we show P is nowhere dense in $M\Delta_o'$. Let $B(f, \epsilon)$ be an open ball in $M\Delta_o'$. If $f \notin P$, we have shown that $\tilde{P} \cap B(f; \epsilon) \neq \emptyset$. If $f \in P$, then f is Baire–1 and continuous on a dense G_δ set. Let x_o be a point of continuity of f in the interval I. It is clear that $f(x_o) = 0$, since $\{x : f(x) = 0\}$ is dense in $[0,1]$. Choose an open interval $J \subset I$ such that $f(x) < \frac{\epsilon}{2}$ on J. By Zahorski's Theorem (see Lemma 2.9) we can choose $g \in M\Delta_o'$ such that $-g \in P$ and

$$\sup_{[0,1]}(-g(x)) = \sup_{J}(-g(x)) = \epsilon.$$

By Step 1, $f + g \in M\Delta_o'$ and

$$\rho(f, f+g) = \rho(g, 0) = \epsilon.$$

On J we have $0 \leq f(x) < \frac{\epsilon}{2}$. In addition, there exists $x_1 \in J$ such that $-g(x_1) > \frac{\epsilon}{2}$, so $g(x_1) < -\frac{\epsilon}{2}$. Thus $f(x_1) + g(x_1) < 0$. It follows that $f + g \in B(f; \epsilon) \setminus P$, and that P is nowhere dense in $M\Delta_o'$. In a similar way, we show that

$$N := \{f \in M\Delta_o' : f \leq 0 \ \text{ on } I\}$$

is closed and nowhere dense in $M\Delta_o'$.

Step 3. By Step 2, given any open interval $I \subset [0,1]$, the set

$$A(I) := \{f \in M\triangle_o^{'} : \exists x_1, x_2 \in I \text{ such that } f(x_1) > 0 \text{ and } f(x_2) < 0\}$$

is a dense open subset of $M\triangle_o^{'}$. Let $\{I_k\}$ be an enumeration of the open interval in $[0,1]$ with rational endpoints. Let $A_k := A(I_k)$. Then $A := \bigcap_{k=1}^{\infty} A_k$ is a dense subset of type G_δ in $M\triangle_o^{'}$. If $f \in A$, then $\{x : f(x) > 0\}$ and $\{x : f(x) < 0\}$ are dense in $[0,1]$. Thus f is a bounded derivative which takes both signs in every open interval contained in $[0,1]$. ☺

Remark 2.1 *By the Fundamental Theorem of Calculus we know for any nondegenerate interval I both $I \cap \{x : f(x) > 0\}$ and $I \cap \{x : f(x) < 0\}$ have positive Lebesgue measures if $f \in A$ given in the previous proof.*

Let $f \in M\triangle_o^{'}$. Define $F(x) := \int_0^x f(s)ds$. Then F is globally Lipschitz and $F^{'} = f$ on $[0,1]$.

Theorem 2.4 *Let $\triangle_o$ denote the set of differentiable functions F on $[0,1]$ such that $F(0) = 0$ and $F^{'} \in M\triangle_o^{'}$. For $F, G \in \triangle_o$, let*

$$\rho(F,G) = \sup_{x\in[0,1]} |F^{'}(x) - G^{'}(x)|.$$

Then:

(i) $(\triangle_o, \rho)$ *is a complete metric space;*

(ii) Every member in the space has a C-subdifferential with 0 *in it identically;*

(iii) For any nondegenerate interval $I \subset [0,1]$, a typical F in the space has a C-subdifferential which is not singleton on a positive measure set, hence F is not C–regular on the corresponding positive measure set.

Corollary 2.3 *Let $f \in C[0,1]$ and $F \in \triangle_o$. Define $\tilde{f}(x) := \int_0^x f(s)ds$ and $g := \tilde{f} + F$. Then $\partial_c g$ contains f and in the topologically complete space $(\tilde{f} + \triangle_o, \rho)$ every member has a C–subdifferential containing f and a typical g in the space is C–irregular on a positive measure set within each nondegenerate interval.*

Corollary 2.4 *A typical $F \in \triangle_o$ has the following properties:*

(i) F is not C–minimal and not C–integrable;

(ii) F is not A–minimal and not A–integrable.

However every member in Δ_o is MP–minimal and MP–integrable.

2.2.2 Regularity

We begin with an example showing that a derivative function can be discontinuous almost everywhere.

Example 20 Let $E \dot{\subset} [0,1]$ be a union of a sequence of closed nowhere dense sets with Lebesgue measure $\lambda(E) = 1$. Suppose for any $x \in E$ we have $d(x, E) = 1$ (where d represents the metric density; see Definition 2.5). Then by Zahorski's Theorem (see Lemma 2.9) there exists a differentiable Lipschitz function $F : [0,1] \to R$ such that

$$0 < F'(x) \leq 1 \text{ for all } x \in E$$

and

$$F'(x) = 0 \text{ for all } x \in [0,1] \setminus E.$$

Therefore F' is discontinuous almost everywhere on $[0,1]$. □

As in section 2.2.1, let $M\Delta'$ denote the space of bounded derivative functions on $[0,1]$ with

$$\rho(f,g) = \sup_{x\in[0,1]} |f(x) - g(x)|.$$

We have seen that $(M\Delta', \rho)$ is a complete metric space. We will show that a typical derivative f in $M\Delta'$ is almost everywhere discontinuous (due to Bruckner and Petruska [11]). Define $F(x) := \int_0^x f(s)ds$. Then F is not regular at x whenever f is not continuous at x. In a measure space X, we say that a property P is true almost everywhere on X provided there exists $A \subset X$ such that $\mu(X \setminus A) = 0$ and P is true on A. In a complete metric space X we say a property P is true typically provided that all members in X satisfy P except for a set of the first category. In this section for any function f we denote by C_f the set of continuity points. Lebesgue measure is denoted by λ and any Borel measure is denoted by μ.

Definition 2.4 *A measurable function f is said to be* **approximately continuous** *at x if for every pair of real numbers k_1 and k_2 such that $k_1 < f(x) < k_2$ the set $\{y : k_1 < f(y) < k_2\}$ has metric density 1 at x, that is*

$$\lim_{r \to 0+} \frac{\lambda((x-r, x+r) \cap \{y : k_1 < f(y) < k_2\})}{2r} = 1.$$

We write this as $\lim_{y \to x} app\, f(y) = f(x)$.

It is obvious that if f is continuous at x, then f is approximately continuous at x. As shown in [9] a real–valued function f on I is almost everywhere approximately continuous if and only if f is Lebesgue measurable. Let f be bounded in a neighborhood I of x and lower (or upper) semicontinuous at x. Then f is approximately continuous at x if and only if f is the derivative of its integral at x.
We give some lemmas from [11] which will be used in the proof of the main Theorem.

Lemma 2.3 *Let μ be an arbitrary Borel measure on $[0,1]$ and $\delta > 0$. Define*

$$A_{\mu,\delta} := \{f \in M\Delta' \mid \mu(C_f) \geq \delta\}.$$

Then $A_{\mu,\delta}$ is closed in $M\Delta'$.

Proof. Let $f_n \in A_{\mu,\delta}$ and $f_n \to f$ uniformly. We consider $C := \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} C_{f_n}$. For any $x \in C$ we have $x \in C_{f_n}$ for infinitely many n, hence $C \subset C_f$. Thus $\mu(C_f) \geq \mu(C) \geq \delta$ and this proves $f \in A_{\mu,\delta}$. ☺

Lemma 2.4 *Let $I_n := (a_n, b_n) = (c_n - h_n, c_n + h_n)$ be a sequence of pairwise disjoint open intervals in $[0,1]$ such that the open set $H := \bigcup_{n=1}^{\infty} I_n$ is everywhere dense in $[0,1]$. Let g_n $(n = 1, 2, \ldots)$ be the piecewise linear continuous function, for which*

$$g_n(x) := \begin{cases} 1, & \text{if } x = c_n \\ 0, & \text{if } x \leq c_n - \frac{1}{n}h_n \text{ or } x \geq c_n + \frac{1}{n}h_n \end{cases}$$

and g_n connects 0 and 1 linearly on $[c_n - \frac{1}{n}h_n, c_n]$ and $[c_n, c_n + \frac{1}{n}h_n]$. Let $g := \sum_{n=1}^{\infty} g_n$. Then

(i) $g|_{[a_n,b_n]}$ is continuous on $[a_n, b_n]$ for $n = 1, 2, \ldots$, $C_g = H$ and $g(x) = 0$ for $x \notin H$;

(ii) g is approximately continuous on $[0,1]$, that is $g \in M\Delta'$.

Proof. For any x there exists at most one n with $g_n(x) \neq 0$, thus the definition of g makes sense. Obviously $g|_{[a_n,b_n]} = g_n|_{[a_n,b_n]}$, hence it is continuous on $[a_n, b_n]$. Therefore $C_g \supset H$. For any $x \notin H$, $g_n(x) = 0$ for any n, this implies $g(x) = 0$; since H is an everywhere dense open set, any point $x \notin H$ is the limit point of some subsequence of $\{c_n\}$, thus $x \notin C_g$ and this proves $C_g = H$. Since g is continuous on H, it suffices to show g is approximately continuous for $x \notin H$. We consider two cases:

Case 1. If $x = a_n$ for some n then g is continuous from the right at a_n.

Case 2. Every right–hand side neighborhood $(x, x+h)$ meets infinitely many intervals I_n. Let $n = N$ be the smallest index with $I_n \cap (x, x+h) \neq \emptyset$. Then

$$\frac{1}{h}\lambda(\{t : x < t < x+h, g(t) > 0\}) \leq$$
$$\begin{cases} \frac{\sum \frac{2}{n} h_n}{\sum 2h_n}, & \text{if } x+h \notin H, \text{ or } x+h \in I_\nu \text{ but } a_\nu < x+h \leq c_\nu - \frac{1}{\nu}h_\nu \\ \frac{\sum \frac{2}{n}h_n + \frac{2}{\nu}h_\nu}{\sum 2h_n + \frac{\nu-1}{\nu}h_\nu}, & \text{if } c_\nu - \frac{1}{\nu}h_\nu \leq x+h \leq b_\nu \end{cases}$$

where the summation is always extended to the indices n such that $I_n \subset (x, x+h)$. Since $n \geq N$, $\nu \geq N$, $2 \geq \frac{N-1}{N}, \frac{\nu-1}{\nu} \geq \frac{N-1}{N}$ in both cases we have

$$\frac{1}{h}\lambda(\{t : x < t < x+h, g(t) > 0\}) < \frac{2}{N-1}$$

and we obtain $\lim_{y \to x+0} app\; g(y) = 0 = g(x)$, because $h \to 0$ implies $N \to \infty$. Similarly we can show g is left–hand approximately continuous and hence the proof is complete. ☺

Lemma 2.5 *Let $F := \{f \in M\Delta' : C_f$ does not contain any open interval$\}$. Then F is a dense G_δ set of $M\Delta'$.*

Proof. Let $F_I := \{f \in M\Delta' : C_f \supset I\}$ where I is a given open interval. It is clear that F_I is closed and $M\Delta' \setminus F_I$ is everywhere dense in $M\Delta'$ with the supremum norm. Let I range over the open intervals with rational endpoints, then $\bigcup_I F_I$ is a first category and F_σ set. Hence the result follows by noticing that $F = \bigcap_I (M\Delta' \setminus F_I)$. ☺

Lemma 2.6 *Let μ be an arbitrary Borel measure on $[0,1]$. Then*

$$A := \{f \in M\Delta' : \mu(C_f) = 0\}$$

is a dense G_δ set in $(M\Delta', \rho)$.

Proof. By Lemma 2.3 the set $A_{\mu,\delta}$ is closed. We show that it is nowhere dense in $M\triangle'$. If not, then there exists an open ball $B(f,\epsilon) \subset A_{\mu,\delta}$. By Lemma 2.5 we can also suppose that $f \in F$ since F is dense in $M\triangle'$. Now we cover the everywhere dense set $[0,1] \setminus C_f$ by an open set H such that $\mu(C_f \cap H) < \delta$. By Lemma 2.4 we can define a bounded derivative g on H such that g is continuous on H and discontinuous on $[0,1] \setminus H$. Define $h := f + \frac{\epsilon}{2}g$. Then $h \in M\triangle'$ and $\|h - f\| \leq \frac{\epsilon}{2}$, that is $h \in B(f,\epsilon)$. Furthermore, g is discontinuous out of H, whereas f is continuous there, and hence $C_h \subset C_f \cap H$. This implies $\mu(C_h) < \delta$, that is $h \notin A_{\mu,\delta}$, a contradiction. Let $A := \bigcap_{n=1}^{\infty}(M\triangle' \setminus A_{\mu,\frac{1}{n}})$. Then A is a dense G_δ set in $M\triangle'$ and each function in A is discontinuous almost everywhere on $[0,1]$. ☺

Let $f \in M\triangle'$. Define $F(x) := \int_0^x f(s)ds$. Then F is globally Lipschitz on $[0,1]$.

Theorem 2.5 *Let $M\triangle$ denote the set of differentiable functions F on $[0,1]$ such that $F(0) = 0$ and $F' \in M\triangle'$. For F, $G \in M\triangle$, let*

$$\rho(F,G) = \sup_{x\in[0,1]} |F'(x) - G'(x)|.$$

Then $(M\triangle, \rho)$ is a complete metric space in which a typical member has a C–subdifferential map which is nonsingleton almost everywhere. Hence a typical member is C–irregular almost everywhere on $[0,1]$.

Proof. Note that $\partial_c F$ is a singleton at x if and only if F' is continuous at x. Thus if F' is not continuous almost everywhere then $\partial_c F$ is not a singleton almost everywhere. Since differentiability and C–regularity together imply that F' is continuous, we know that if F' is discontinuous almost everywhere then F must be C–irregular almost everywhere. Hence Lemma 2.6 applies. ☺

Corollary 2.5 *A typical member F in $M\triangle$ has different $\partial_c F$ and $\partial_{mp} F$ almost everywhere on $[0,1]$.*

Corollary 2.6 *A typical member in $M\triangle$ is not sainely Lipschitz.*

Remark 2.2 *The Baire Theorem shows that every Baire–1 function is continuous except at the points of a set of the first category. Therefore every member in $M\triangle$ is generically regular and has a C–subdifferential which is generically single–valued. This doesn't contradict Theorem 2.5 since a dense G_δ set can have measure 0.*

2.3 Two differentiable Lipschitz functions with the same C–subdifferential

In this section we construct two globally Lipschitz functions which are differentiable everywhere and different by more than a constant but have the same C–subdifferential.

Definition 2.5 *Let $A \subset R$ be measurable and $x_o \in R$. The upper metric density of A at x_o is*

$$\overline{d}(x_o, A) = \limsup_{I \to x_o} \frac{\lambda(A \cap I)}{\lambda(I)}$$

and the lower density is defined by

$$\underline{d}(x_o, A) = \liminf_{I \to x_o} \frac{\lambda(A \cap I)}{\lambda(I)}.$$

If $\overline{d}(x_o, A) = \underline{d}(x_o, A)$, we call this number the density of A at x_o and denote it by $d(x_o, A)$.

Since the following lemmas can be found in [9, 27], we omit their proofs.

Lemma 2.7 *(Lebesgue Density Theorem) Let $A \subset R$ be measurable. Then*

$$\mu(A \setminus \{x \in A : \; d(x, A) = 1\}) = 0.$$

That is for almost all $x \in A$ it has metric density 1.

Lemma 2.8 *Let $A \subset R$ be measurable. Then there is an F_σ set $F \subset A$ such that $\mu(A \setminus F) = 0$.*

Lemma 2.9 *(Zahorski's Theorem) Let E be a set of type F_σ with $d(x, E) = 1$ for all $x \in E$. Then there exists an approximately continuous function f such that*

$$0 < f(x) \leq 1 \text{ for all } x \in E$$

and

$$f(x) = 0 \text{ for all } x \notin E.$$

The function f is also upper semicontinuous.

Lemma 2.10 *Let f and g be approximately continuous at x_o, then the same is true for $f + g$.*

2.3.1 Construction

Let $P \subset [0,1]$ be a Cantor set with $\mu(P) > 0$. We can construct a Cantor–like differentiable function g which is constant on each interval contiguous to P.

Step 1. Let $A := \{x : d(x,P) = 1\}$. Then $A \subset P$ and $\mu(P) = \mu(A)$. Thus $\mu(A) > 0$. Let E be a set of type F_σ such that $E \subset A$ and $\mu(E) = \mu(A)$. Then $d(x,E) = d(x,P) = 1$ for all $x \in E$.

Step 2. By Lemma 2.9, there exists an approximately continuous f such that

$$0 < f(x) \leq 1 \text{ for all } x \in E$$

and

$$f(x) = 0 \text{ for all } x \notin E.$$

Define $F : [0,1] \to R$ by $F(x) := \int_0^x f(t)dt$. Then $F'(x) = f(x)$ for all $x \in [0,1]$. In particular, $F' = 0$ on each interval contiguous to P, so F is constant on each such interval. Let I be an open interval intersecting P with $\mu(P \cap I) > 0$. Then $\mu(E \cap I) > 0$ and we see that $F'(x) > 0$ for all $x \in E \cap I$ and so F is not constant on any such interval. Since $\mu(P) > 0$ there must be such an interval and we know F is not a constant function.

Step 3. Set $\tilde{P} := [0,1] \setminus P$. Then $\tilde{P} = \bigcup_{n=1}^{\infty}(a_n, b_n)$ and $\tilde{P}$ is everywhere dense in $[0,1]$. By Lemma 2.4, we can find g on $[0,1]$ such that

(i) g is approximately continuous on $[0,1]$;

(ii) g is continuous on $[a_n, b_n]$ for each $n = 1, 2, \ldots$ and $C_g = \tilde{P}$ and $g(x) = 0$ for $x \in P$;

(iii) On $(a_n, b_n) = (c_n - h_n, c_n + h_n)$ we have $g(c_n) = 1$ and g_n connects 0 and 1 linearly on $[c_n - \frac{1}{n}h_n, c_n]$ and $[c_n, c_n + \frac{1}{n}h_n]$ and 0 otherwise.

Consider $h := f+g$. Then h is approximately continuous and so it is a bounded derivative on $[0,1]$. Define $G : [0,1] \to R$ by $G(x) := \int_0^x g(t)dt$ and $H : [0,1] \to R$ by $H(x) := \int_0^x h(t)dt$. Thus $H = G + F$.

Step 4. We claim $\partial_c G(x) = \partial_c H(x)$ for all $x \in [0,1]$. Indeed, $h(x) = g(x)$ for all $x \in \tilde{P}$ and $h(x) = f(x)$ for all $x \in P$. Noting that $0 < f(x) \leq 1$ we have

$$\partial_c H(x) = \partial_c G(x) = [0,1] \text{ for all } x \in P.$$

Since F is continuously differentiable on $\tilde{P}$ we have $\partial_c H(x) = \partial_c G(x) + \nabla F(x)$. However $\nabla F(x) = 0$ for all $x \in \tilde{P}$. Therefore

$$\partial_c H(x) = \partial_c G(x) = g(x) \text{ for all } x \in \tilde{P}.$$

We summarize our construction as a theorem.

Theorem 2.6 *There are two Lipschitz and everywhere differentiable functions $H : [0,1] \to R$ and $G : [0,1] \to R$ differing by more than a constant such that*

$$\partial_c H(x) = \partial_c G(x) \quad \textit{for all } x \in [0,1].$$

Remark 2.3 *From the construction, each Cantor set with positive measure gives a pair of such functions. Choosing different Cantor sets with positive measure will give us different pairs of such functions. In fact, given one Cantor set with positive measure we can make uncountably many differentiable functions differing by more than constants such that they share the same C–subdifferential. It is not because the functions themselves are pathological, such a pathological situation arises from the upper semicontinuity which nonsmooth analysts impose on the C–subdifferential map.*

Combining Theorem 2.6 and Theorem 2.2 we have:

Corollary 2.7 *There are two Lipschitz and everywhere differentiable functions $H : [0,1] \to R$ and $G : [0,1] \to R$ differing by more than a constant such that*

$$\partial_a H(x) = \partial_a G(x) \textit{ for all } x \in [0,1].$$

2.4 Locally Lipschitz functions

It is easy to construct a locally Lipschitz function on R whose $\partial_c f$ and $\partial_a f$ differ on a given countable set. However for a locally Lipschitz function on R its C–subdifferential and A–subdifferential must coincide except on a countable set (due to Katriel [22]).

Lemma 2.11 *Let g be a locally Lipschitz function defined on I, $x \in I$ and $r \in \partial_c g(x) \setminus \partial_a g(x)$ and let $g_r(y) := g(y) - r \cdot y$. Then there is a $\epsilon > 0$ such that g_r is strictly increasing on $[x - \epsilon, x]$ and g_r is strictly decreasing on $[x, x + \epsilon]$.*

Proof. By assumption $0 \in \partial_c g_r(x)$ and $0 \notin \partial_a g_r(x)$. Denote g_r by f. Suppose f is not decreasing on $(x, x+\epsilon)$ for any $\epsilon > 0$, then there are y, y' with $x \leq y < y' \leq x+\epsilon$ and $f(y) \leq f(y')$. Since $\partial_c f(x) = \overline{conv}\partial_a f(x)$ and $\partial_a f(x)$ is compact, we have $\partial_c f(x) = conv\partial_a f(x)$ by Theorem 1.4.3 [17]. Hence $0 \in \partial_c f(x)$ implies $\partial_a f(x)$ contains a negative number. Thus each neighborhood of x contains a point z such that $\alpha \in \partial^- f(z)$ for some $\alpha < 0$ because $\partial_a f(x) = \limsup_{z \to x} \partial^- f(z)$. In particular we may choose $z \in (x - \epsilon, y)$. The fact that $\alpha \in \partial^- f(z)$ and $\alpha < 0$ implies that for some $z' < z$ close enough to z, $f(z') \geq f(z)$. We choose any such z' which also satisfies $z' \in (x - \epsilon, y)$. Therefore we have the following situation: $x - \epsilon < z' < z < y < y' < x + \epsilon$ and $f(z') \geq f(z)$, $f(y) \leq f(y')$. Thus f attains its minimum at an interior point $w \in (z', y')$, so $0 \in \partial_a f(w)$ while $|w - x| < \epsilon$. Since such w can be found for every $\epsilon > 0$ and that $\partial_a f$ is upper semicontinuous we get $0 \in \partial_a f(x)$. This contradicts the assumption $0 \notin \partial_a f(x)$. Therefore for some $\epsilon > 0$, f is strictly decreasing on $(x, x + \epsilon)$. The fact that for some $\epsilon' > 0$, f is strictly increasing on $(x - \epsilon', x)$ is proved in an analogous way. ☺

Theorem 2.7 *Let $f : R \to R$ be locally Lipschitz. Then*

$$\{x \in R : \partial_c f(x) \neq \partial_a f(x)\}$$

is at most countable.

Proof. Let $A := \{x \in R| \quad \partial_c f(x) \neq \partial_a f(x)\}$. Letting Q denote the rationals, we define for each $r \in Q$:

$$A_r := \{x \in R| \quad r \in \partial_c f(x), r \notin \partial_a f(x)\}.$$

We claim $A = \bigcup_{r \in Q} A_r$. Suppose that $x \in A$ but $x \notin \bigcup_{r \in Q} A_r$. Then for any $r \in Q$ we have $r \in \partial_a f(x)$ if $r \in \partial_c f(x)$, that is $\partial_c f(x) \bigcap Q = \partial_a f(x) \bigcap Q$. Since $\partial_c f(x)$ and $\partial_a f(x)$ are both closed, so $\partial_a f(x) = \partial_c f(x)$, a contradiction. Applying Lemma 2.11 we see that each A_r is at most countable so is A. ☺

In fact for a locally Lipschitz function on R if we know $\partial_c f$ we actually know $\partial_a f$. That is, if two locally Lipschitz functions have the same C–subdifferentials then they have the same A–subdifferentials. To see this we use Borwein and Fitzpatrick's Theorem [5]. It states that $\partial_a f$ is either a single closed interval or a union of two closed intervals and we can calculate $\partial_a f$ from $\partial_c f$. [Compare Theorem 4.8]

Theorem 2.8 *Let g be locally Lipschitz on I and $x \in I$. Suppose $\partial_c g(y) = [\alpha(y), \beta(y)]$ then*

$$\partial_a g(x) = [\liminf_{\substack{y \to x \\ y \notin N}} \alpha(y), \limsup_{\substack{y \to x^+ \\ y \notin N}} \beta(y)] \cup [\liminf_{\substack{y \to x^- \\ y \notin N}} \alpha(y), \limsup_{\substack{y \to x \\ y \notin N}} \beta(y)]$$

for each Lebesgue null set N.

Proof. We need to show that $r \in \partial_c g(x) \setminus \partial_a g(x)$ if and only if

$$\limsup_{y \to x^+} \beta(y) < r < \liminf_{y \to x^-} \alpha(y).$$

Let $\limsup_{y \to x^+} \beta(y) < r < \liminf_{y \to x^-} \alpha(y)$. Then there are $\epsilon > 0$ and $\delta > 0$ such that $\beta(y) \leq r - \epsilon$ for $x < y < x + \delta$ and $\alpha(y) \geq r + \epsilon$ for $x - \delta < y < x$. Now let $y^* \in \partial^- g((x - \delta, x + \delta))$. Then there is $y \in (x - \delta, x + \delta)$ such that $g_-(y) \leq y^* \leq g^-(y)$. We consider two cases:

(i) If $y \leq x$ then for $t < 0$ with $x - \delta < y + t$ we have

$$\frac{g(y+t) - g(y)}{t} \in \partial_c g(z)$$

for some $z \in (x - \delta, y)$ by the Lebourg Mean Value Theorem [13] and so $y^* \geq r + \epsilon$.

(ii) If $y > x$ then for $t > 0$ with $y + t < x + \delta$ we have

$$\frac{g(y+t) - g(y)}{t} \in \partial_c g(z)$$

for some $z \in (y, x + \delta)$ by the Lebourg Mean Value Theorem again, so $y^* \leq r - \epsilon$.

Thus $|y^* - r| \geq \epsilon$ and so

$$r \notin \overline{\partial^- g((x - \delta, x + \delta))}.$$

Since this is true for any sufficiently small $\delta > 0$, it follows that $r \in \partial_c g(x) \setminus \partial_a g(x)$.

Conversely suppose $r \in \partial_c g(x) \setminus \partial_a g(x)$. Since $\partial_a g(x)$ is closed there are $q < r < s$ such that $(q, s) \subset \partial_c g(x) \setminus \partial_a g(x)$. By Lemma 2.11 there is $\epsilon > 0$ such that both $g_q(y) := g(y) - qy$ and $g_s(y) := g(y) - sy$ are strictly increasing on $[x - \epsilon, x]$ and strictly decreasing on $[x, x + \epsilon]$. Thus $\partial_c g_s(y) \subset [0, +\infty)$ for any $x - \epsilon < y < x$ and $\partial_q g(y) \subset (-\infty, 0]$ for any $x + \epsilon > y > x$. Thus

$$\partial_c g(y) \subset [s, +\infty) \quad \text{for } x - \epsilon < y < x$$

and

$$\partial_c g(y) \subset (-\infty, q] \quad \text{for } x+\epsilon > y > x.$$

Hence

$$\limsup_{y \to x^+} \beta(y) \leq q < r < s \leq \liminf_{y \to x^-} \alpha(y)$$

as required. ☺

Example 21 Let us take the Rockafellar function $f : R \to R$ such that $\partial_c f(x) = [0,1]$ for all $x \in R$ (see Example 7). By Theorem 2.8 we get $\partial_a f(x) = \partial_c f(x) = [0,1]$ identically for all $x \in R$.

Let α and β be any two continuous functions defined on I such that $\alpha \leq \beta$. Then by Theorem 1.2 [5] there is a locally Lipschitz function $g : I \to R$ such that $\partial_c g(x) = [\alpha(x), \beta(x)]$. By Theorem 2.8 we have

$$\partial_a g(x) = \partial_c g(x) = [\alpha(x), \beta(x)] \quad \text{for all } x \in I.$$

□

In fact we can say much more for a locally Lipschitz function on R.

Theorem 2.9 *Let f be locally Lipschitz on I. Then the following are equivalent:*

(i) f is A–integrable on $I \subset R$;

(ii) f is C–integrable on $I \subset R$;

(iii) f is almost everywhere strictly differentiable;

(iv) f' is Riemann integrable.

Proof. *(iii)* $\to$ *(ii)*: Let f be almost everywhere strictly differentiable. Then f is an essentially strictly differentiable function. By Corollary 4.6 [1] f is C–integrable.

(ii) $\to$ *(iii)*: Let f be C–integrable. Then $\partial_c f$ is singleton almost everywhere. Indeed if not, set $g_1(t) := \sup \partial_c f(t)$ and $g_2(t) := \inf \partial_c f(t)$. Since $\partial_c f$ is a cusco we deduce g_1 and g_2 are locally–bounded measurable selections of $\partial_c f$. Then $\{t \in I : g_1(t) \neq g_2(t)\}$ is of positive measure. Define

$$F(x) := \int_0^x g_1(t)dt \text{ and } G(x) := \int_0^x g_2(t)dt.$$

Thus F and G are locally Lipschitz functions and their C–subdifferentials lie in $\partial_c f$ but $F - G \neq constant$. Since $\partial_c f$ is C–integrable it has a minimal subdifferential and we get

$$\partial_c F = \partial_c f = \partial_c G.$$

This implies f is not C–integrable, a contradiction.

$(i) \Leftrightarrow (ii)$: Let f be C-integrable. Suppose that $\partial_a g \subset \partial_a f$. Then $conv\partial_a g(x) \subset conv\partial_a f(x)$, that is $\partial_c g(x) \subset \partial_c f(x)$. Thus $f - g$ is constant. Conversely let f be A–integrable. Suppose $\partial_c g \subset \partial_c f$. By Theorem 0.2 [5] there exist α and β, $\alpha \leq \beta$, which are essentially lower semicontinuous and essentially upper semicontinuous respectively such that $\partial_c f(x) = [\alpha(x), \beta(x)]$ for all $x \in I$. Similarly for g there are $\tilde{\alpha}$ and $\tilde{\beta}$, $\tilde{\alpha} \leq \tilde{\beta}$, which are essentially lower semicontinuous and essentially upper semicontinuous respectively such that $\partial_c g(x) = [\tilde{\alpha}(x), \tilde{\beta}(x)]$ for all $x \in I$. Thus $\alpha \leq \tilde{\alpha}$ and $\tilde{\beta} \leq \beta$. By Theorem 2.8 we know

$$\partial_a f(x) = [\liminf_{\substack{y \to x \\ y \notin N}} \alpha(y), \limsup_{\substack{y \to x^+ \\ y \notin N}} \beta(y)] \cup [\liminf_{\substack{y \to x^- \\ y \notin N}} \alpha(y), \limsup_{\substack{y \to x \\ y \notin N}} \beta(y)]$$

and

$$\partial_a g(x) = [\liminf_{\substack{y \to x \\ y \notin N}} \tilde{\alpha}(y), \limsup_{\substack{y \to x^+ \\ y \notin N}} \tilde{\beta}(y)] \cup [\liminf_{\substack{y \to x^- \\ y \notin N}} \tilde{\alpha}(y), \limsup_{\substack{y \to x \\ y \notin N}} \tilde{\beta}(y)].$$

It is clear that

$$\liminf_{\substack{y \to x^- \\ y \notin N}} \alpha(y) \leq \liminf_{\substack{y \to x^- \\ y \notin N}} \tilde{\alpha}(y) \text{ and } \limsup_{\substack{y \to x^+ \\ y \notin N}} \beta(y) \geq \limsup_{\substack{y \to x^+ \\ y \notin N}} \tilde{\beta}(y).$$

Therefore $\partial_a g \subset \partial_a f$. Hence $f - g$ is a constant on I.

$(iii) \Leftrightarrow (iv)$ is Theorem 6.29 in [28]. ☺

Combining Theorem 2.9, Theorem 2.5 and Fact 1.1, we are now ready to state two important results.

Corollary 2.8 *In the complete metric space $(M\triangle, \rho)$, the set of functions which are C–integrable forms a set of first category.*

Corollary 2.9 *In R the four classes of locally Lipschitz functions are equivalent:*

(i) Sainely Lipschitz functions;

(ii) Essentially strictly differentiable Lipschitz functions;

(iii) C–integrable Lipschitz functions;

(iv) A–integrable Lipschitz functions.

2.5 Locally Lipschitz functions with prescribed subdifferentials

It is natural to ask now whether we can construct locally Lipschitz functions with several continuous curves inside their subdifferentials. Here is a result due to Borwein and Fitzpatrick [5].

Theorem 2.10 *Let I be an open interval in R and α and β on I such that $\alpha \leq \beta$, α is essentially lower semicontinuous and β is essentially upper semicontinuous. Then there is a locally Lipschitz function $g : I \to R$ such that $\partial_c g(x) = [\alpha(x), \beta(x)]$ for all $x \in I$.*

They also give a generalization of the classical result about the uniform convergence of derivatives. By adding the C–integrability condition, we get an improved result.

Theorem 2.11 *Let $\{h_k\}$ be a sequence of locally Lipschitz functions on an open interval I such that $\partial_c h_k$ converges uniformly in the Hausdorff metric to an upper semicontinuous compact interval-valued multifunction Ω on I. Let h_k be C–integrable and $h_k(0) = 0$ for each k. Then there exists a locally Lipschitz function g_0 on I which is C–integrable such that $\partial_c g_0 = \Omega$ and h_k converges uniformly to g_0 on compact subintervals of I.*

Proof. Let $\partial_c h_k = [\alpha_k, \beta_k]$ for each $k = 1, 2, \ldots$. Then α_k and β_k converge to α_0 and β_0 uniformly with $\Omega = [\alpha_0, \beta_0]$. By Theorem 1.3 [5], we can find a sequence of locally Lipschitz functions g_k with $\partial_c g_k = [\alpha_k, \beta_k]$ and $g_k(0) = 0$ such that g_k converges to g_0 uniformly on compact subintervals of I and $\partial_c g_0 = [\alpha_0, \beta_0]$. Since h_k is C–integrable, we see that $h_k - g_k = c_k$. Note that $g_k(0) = h_k(0) = 0$ and so $h_k = g_k$ on I.

To see g_0 is integrable we can use Theorem 2.9. Since each h_k is C–integrable, h_k is almost everywhere strictly differentiable and so $\partial_c h_k$ is single–valued except a null set N_k. Letting $N := \cup_{k=1}^{\infty} N_k$, we know Ω is single–valued on $I \setminus N$ and so Ω is single–valued almost everywhere. By Theorem 2.9 we know g_0 is C–integrable. ☺

Example 22 In Theorem 2.11 the uniform convergence in the Hausdorff metric is essential. Let $\{f_n : n \in N\}$ be defined on R such that $\partial_c f_n(x) = [0, \beta_n(x)]$ with $\beta_n(x) := \min\{n|x|, 1\}$ for each $x \in R$, as ensured by Theorem 2.10. Then the point–wise limit of β_n is β_0 defined by

$$\beta_0(x) := \begin{cases} 0, & \text{if } x = 0 \\ 1, & \text{otherwise.} \end{cases}$$

Now let $\Omega(x) := [0, \beta_0(x)]$. $\partial_c f_n$ does not converge uniformly in the Hausdorff metric to Ω. It is obvious that Ω is not a C–subdifferential since Ω is not an usco or β_0 is not essentially upper semicontinuous at 0. □

We can use Theorem 2.10 to construct a large class of pathological Lipschitz functions.

Example 23 Let $\alpha < \beta$ and α and β are continuous on I. Then by Theorem 2.10 there exists a locally Lipschitz function $g : I \to R$ such that

$$\partial_c g(x) = [\alpha(x), \beta(x)]$$

for all $x \in I$. Then g is not sainely Lipschitz. By Theorem 1.1 in [29] we know the set of points where g is C–regular but not differentiable is at most countable. Hence g is also C–regular at no more than countably many points.

It is clear that g is nowhere strictly differentiable on I. Giles and Sciffer have shown that for a locally Lipschitz function defined on an open set the set of points where it is Gateaux differentiable but not strictly differentiable is of first category [15]. Hence g is differentiable only on a first category subset of I. The set of points where g is C–pseudo–regular is residual but is disjoint from the set of points where g is differentiable. □

Example 24 Every G_δ subset of (a, b) is the set of points of strict differentiability of a Lipschitz function on (a, b).

To see this, let G be a G_δ subset of (a, b) then there are closed subsets F_n of (a, b) such that $G := (a, b) \setminus \bigcup_n F_n$. Let $\Omega(x) := [0, \sum_n \chi_{F_n}/2^n]$. Ω is cusco because $\sum_n \chi_{F_n}/2^n$ is upper semicontinuous as a uniform limit of upper semicontinuous functions. $\sum_n \chi_{F_n}/2^n$ can be decomposed into two bounded essentially upper semicontinuous continuous functions. By Theorem 2.10 there is a locally Lipschitz $g : I \to R$ such that diam$[\partial_c g(x)]$=diam$[\Omega(x)]$ for all $x \in (a, b)$. Now g is strictly differentiable at x if and only if $0 =$ diam$[\partial_c g(x)] =$

diam$[\Omega(x)]$. But diam$[\Omega(x)] = 0$ if and only if $x \notin F_n$ for all n, so G is the set of points at which g is strictly differentiable.

Theorem 2.2 [9] shows that a set $A \subset [a, b]$ is the set of points of continuity of a bounded derivative if and only if A is dense and of type G_δ. Hence if G is dense in (a, b), we can in fact find a locally Lipschitz function whose derivative function is continuous at each point of G and discontinuous at each point of $(a, b) \setminus G$. □

Example 25 We can use Example 24 to construct a locally Lipschitz function on R such that it is almost everywhere C–irregular but C–regular on a set of second category.

Indeed, by Theorem 1.6 in [24] we can decompose R into two sets: one is of first category, the other is a G_δ set with zero measure. Denote them by F and G respectively. Then there is a Lipschitz function f on R such that f is strictly differentiable on G. Obviously f is C–regular on G. On F we see that f is almost everywhere C–irregular since C–regularity and Gateaux differentiability together result in $\partial_c f$ singleton. It follows that f is C–regular at least on a second category set but C–irregular almost everywhere. □

Example 26 Let C be a Cantor set in I. Let $\alpha := 0$ and $\beta(x) := d_C(x)$. Then there exists a locally Lipschitz g on I such that $\partial_c g(x) = [0, d_C(x)]$ for each $x \in I$. It follows that g is nondecreasing and strictly differentiable exactly at the points of C. □

We close with a result by Sciffer [29] which says that a locally Lipschitz function can be C–irregular everywhere.

Theorem 2.12 *There exists a locally Lipschitz function $f : (0,1) \to R$ which is nowhere C–regular.*

We sketch his construction. Let $E \subset R$ be a **ubiquitous** set (see Example 7). That is, for any nondegenerate open interval $I \subset R$ we have $0 < \mu(E \cap I) < \mu(I)$. Define

$$f(x) := \int_0^x \chi_E(s)ds.$$

Then f is locally Lipschitz and C–regular at most at countably many points. The metric interior of E corresponds to those points where $f'(x) = 1$, and the metric interior of $(0,1)\setminus E$ to where $f'(x) = 0$. The metric boundary of E is those points where either f or $-f$ is C–regular but not differentiable. Sciffer constructed E in a special way such that the metric boundary of E in $(0,1)$ is empty. Thus f is nowhere C–regular on $(0,1)$. However, we should note that the set of points where f is C–pseudo–regular but not differentiable is a dense G_δ set with measure 0. These points where f is C–pseudo–regular lie in the "fuzzy metric" boundary of E.

Chapter 3

Clarke subdifferential

We begin with two examples to illustrate that given any countable set we can construct a globally Lipschitz function with its C–subdifferential not singleton exactly on the countable set. For examples in several dimensions see Example 40 and Corollary 4.2.

Example 27 Let $\{r_n\}_{n=1}^{\infty}$ be an enumeration of $\mathbf{Q}$. Define f_n by $f_n(x) := 1$ if $x > r_n$ and $f_n(x) := 0$ for all other $x \in R$ (see [16] page 113). Let $f(x) := \sum_{n=1}^{\infty} 2^{-n} f_n(x)$ for all $x \in R$. Then f is Lebesgue measurable and satisfies the following properties:

(a) f is strictly increasing;

(b) f is left continuous;

(c) f is continuous at each irrational number;

(d) f is discontinuous at each rational number;

(e) $\lim_{x\to\infty} f(x) = 1$;

(f) $\lim_{x\to-\infty} f(x) = 0$.

Define $g(x) := \int_0^x f(s)ds$. By the Criterion of Increasing Slopes [17] we know g is convex and Lipschitz on R. It is clear that $\partial_c g(r_k) = \left[f(r_k), f(r_k) + \frac{1}{2^k}\right]$ for each r_k and at any irrational point $\partial_c g$ is singleton. □

Example 28 Define

$$\phi(u) := \begin{cases} u, & \text{if } u \le 0 \\ \frac{1}{k+1}, & \text{if } \frac{1}{k+1} < u \le \frac{1}{k} \\ 1, & \text{if } u > 1. \end{cases}$$

Consider the function $g(x) := \int_0^x \phi(u)du$. Again g is convex and $\partial_c g(\frac{1}{k}) = [\frac{1}{k+1}, \frac{1}{k}]$ for $k = 1, 2, \ldots$ □

Definition 3.1 *A set–valued map $T : R^N \to 2^{R^N}$ is said to be n–***cyclically monotone** *provided*

$$\sum_{k=1}^{n} \langle x_k^*, x_k - x_{k-1} \rangle \ge 0$$

whenever $n \ge 2$ and $x_0, x_1, \ldots, x_n \in R^N$, $x_n = x_0$, and $x_k^* \in T(x_k)$, $k = 1, 2, 3, \ldots, n$. We say T is **cyclically monotone** *if it is n–cyclically monotone for every n. A monotone operator T is said to be* **maximal cyclically monotone** *provided $T = S$ whenever S is cyclically monotone and $G(T) \subset G(S)$.*

Clearly, a 2–cyclically monotone operator is monotone; and a maximal monotone operator which is cyclically monotone is necessarily maximal cyclically monotone. As shown in [26], on a Banach space the subdifferential of every lower semicontinuous proper convex function is a maximal cyclically monotone operator and every maximal cyclically monotone operator is a maximal monotone operator. Every C–subdifferential is a cusco but our next example shows the converse may fail.

Example 29 Following Example 2.21 in [25], we set $T(x, y) := (y, -x)$. Then T is a cusco and a monotone operator. We claim there is no locally Lipschitz function in R^2 such that $\partial_c f = T$. Indeed, by Proposition 2.2.4 [13] we know $\partial_c f(x)$ reduces to a singleton at x if and only if f is strictly differentiable at x, and so $T = \nabla f = \partial_c f$. Since T is monotone so f is convex. It is known that $\partial_c f$ is cyclically monotone. However, checking the points $(1, 1)$, $(0, 1)$ and $(1, 0)$, we find T is not 3–cyclically monotone, a contradiction. □

In convex analysis, a convex function defined on an open set U is characterized by its subdifferential $\partial_c f$. That is, f is convex on U if and only if $\partial_c f$ is a maximal cyclically monotone operator. Most importantly f can be uniquely determined by its $\partial_c f$ up to a constant [26]. In other words, convex functions are integrable. For locally Lipschitz functions which are not convex, there may be many functions whose differences are not a constant but they share the same C–subdifferential (see Example 16).

3.1 Convex case

Let us suppose that T is a C–subdifferential map, can we recover a locally Lipschitz function such that $\partial_c f = T$? The following theorem is due to Rockafellar [26].

Theorem 3.1 *If $T : U \to 2^{R^N}$ is a maximal cyclically monotone operator, with $D(T) \neq \emptyset$, then there exists a proper convex lower semicontinuous function f on U such that $T = \partial_c f$.*

In his proof, Rockafellar used the cyclical property of T to get the following function:

$$f(x) := \sup\{\langle x_n^*, x - x_n\rangle + \langle x_{n-1}^*, x_n - x_{n-1}\rangle + \cdots + \langle x_0^*, x_1 - x_0\rangle\}$$

where the supremum is taken over all finite sets of elements $x_k \in D(T)$ and $x_k^* \in T(x_k)$ for $k = 1, 2, ..., n$, $n = 1, 2, 3, ...$ He showed that f is proper convex lower semicontinuous and $\partial_c f = T$. If we require $T(x) \neq \emptyset$ for each $x \in U$, then f is locally Lipschitz on U (see Theorem 3.1.2 [17]).

3.2 Lipschitz case

Our main goal in this section is to provide a technique for constructing a class of Lipschitz functions with prescribed C–subdifferentials in several dimensions. Using this construction we construct some 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 $\partial_c f = P$ identically. Here is the main result by Borwein, Moors, and Wang [7]. The theorem is actually true in the separable Banach spaces, we only give a version in R^N.

Theorem 3.2 *Let $f_1, f_2, \cdots, f_n$ be real–valued locally Lipschitz functions defined on a non-empty open subset $U \subset R^N$. If each function f_j possesses a minimal C–subdifferential mapping on U, then there exists a real-valued locally Lipschitz function g defined on U such that $\partial_c g(x) = conv\{\partial_c f_1(x), \partial_c f_2(x), \cdots, \partial_c f_n(x)\}$ for each $x \in U$.*

3.2.1 Lemmata

Lemma 3.1 *Let g, f_1, f_2, $\cdots$, f_n be real-valued locally Lipschitz functions defined on a non-empty open subset $U \subset R^N$. If $\nabla g(x) \in \{\nabla f_1(x), \nabla f_2(x), \cdots, \nabla f_n(x)\}$ almost everywhere in U, then $\partial_c g(x) \subset conv\{\partial_c f_1(x), \partial_c f_2(x), \cdots, \partial_c f_n(x)\}$ for all $x \in U$.*

Proof. Consider the set-valued mapping $T : U \to 2^{R^N}$ defined by $T(x) \equiv \bigcup\{\partial_c f_j(x) : 1 \leq j \leq n\}$. Clearly T is an usco mapping on U, hence by Lemma 7.12 in [25] the mapping $T^* : U \to 2^{X^*}$ defined by

$$T^*(x) = \overline{conv}\{\partial_c f_1(x), \partial_c f_2(x), \cdots, \partial_c f_n(x)\}$$

is a cusco on U. By Theorem 1.4.3 [17] the closure operation is superfluous since $T(x)$ is a compact set for each x. Now from the hypothesis we have that $\nabla g(x) \in T^*(x)$ almost everywhere in U. Since C–subdifferential is the minimal cusco containing ∇g, so $\partial_c g(x) \subset T^*(x)$ for all $x \in U$. ☺

Lemma 3.2 *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 and $int\overline{G_n} \neq \emptyset$, 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 and nowhere dense in R;

(ii) For each $n \in \mathbf{N}$, $\mu(G_n \cap E) > 0$ and $G_n \setminus (\bigcup\{E_j : 1 \leq j \leq n\}) \neq \emptyset$;

(iii) If $\mu(G_n \setminus (\bigcup\{E_j : 1 \leq j \leq n\})) > 0$, then $\mu(G_n \setminus E) > 0$.

Proof. We proceed by induction.
Step 1. Set $E_o \equiv \emptyset$ and $r_o \equiv 1$. Now suppose the compact nowhere dense sets E_o, $\cdots$, E_n and positive numbers r_o, $\cdots$, r_n have been chosen.
Step n+1. We consider two cases:

(a) if $\mu(G_n \setminus (\bigcup_{m=0}^n E_m)) = 0$, let $E_{n+1} = \emptyset$ and $r_{n+1} = r_n$.

(b) if $\mu(G_n \setminus (\bigcup_{m=0}^n E_m)) > 0$, then choose $0 < r_{n+1} < \min\{r_o, ..., r_n, \mu(G_n \setminus (\bigcup_{m=0}^n E_m))\}$. By Exercise 5 on page 307 in [28], there exists a compact nowhere dense set $E_{n+1} \subset G_n \setminus (\bigcup_{m=0}^n E_m)$ with $0 < \mu(E_{n+1}) < \frac{r_{n+1}}{2^{n+1}}$.

We claim that the set $E \equiv \bigcup_{m=0}^{\infty} E_m$ satisfies the requirements of the Lemma. By the construction, each set E_m is compact and nowhere dense. Therefore $G_n \setminus \bigcup\{E_j : 0 \leq j \leq n\} \neq \emptyset$. That $\mu(G_n \cap E) > 0$ for each n follows, in the first case, from the fact that

$$\mu(G_n \cap E) \geq \mu(G_n \cap (\bigcup_{m=0}^{n} E_m)) = \mu(G_n) > 0$$

and in the second case from the fact $\mu(G_n \cap E) \geq \mu(E_{n+1}) > 0$. So to complete the proof we need only show that the set E satisfies condition (iii). This follows from the following calculation:

$$\begin{aligned} \mu(G_n \cap E) &= \mu((G_n \cap (\bigcup_{m=0}^{n} E_m)) \cup G_n \cap (\bigcup_{m=n+1}^{\infty} E_m)) \\ &= \mu(G_n) - \mu(G_n \setminus (\bigcup_{m=0}^{n} E_m)) + \mu(G_n \cap (\bigcup_{m=n+1}^{\infty} E_m)) \\ &< \mu(G_n) - r_{n+1} + \sum_{m=n+1}^{\infty} \mu(E_m) < \mu(G_n). \end{aligned}$$

☺

Remark 3.1 *For us, the most important application of Lemma 3.2 is to sets constructed in the following manner. Let f be a real-valued locally Lipschitz function defined on an open interval I. Let M be a dense and Lebesgue measurable subset of I. Then*

(i) $f(M)$ is Lebesgue measurable;

(ii) if $\mu(f(M)) > 0$, then $int\overline{f(M)} \neq \emptyset$.

(i) Since Lebesgue measure on R is regular, there exists a family $\{C_n : n \in \mathbf{N}\}$ of compact subsets of M such that $\mu(M \setminus \bigcup\{C_n : n \in \mathbf{N}\}) = 0$. Now $f(M) = f(\bigcup\{C_n : n \in \mathbf{N}\}) \cup f(M \setminus \bigcup\{C_n : n \in \mathbf{N}\})$ and

$$f(\cup\{C_n : n \in \mathbf{N}\}) = \cup\{f(C_n) : n \in \mathbf{N}\}.$$

The latter is σ-compact. Therefore to see that $f(M)$ is Lebesgue measurable, we need only observe that since f is locally Lipschitz $f(M \setminus \bigcup\{C_n : n \in \mathbf{N}\})$ is Lebesgue measurable, with Lebesgue measure zero by Lemma 6.87 [28].

(ii) Since f is Lipschitz, $f(I)$ is a connected subset of R. However, since $\mu(f(I)) \geq \mu(f(M)) > 0$, $f(I)$ must be a non-degenerate interval in R. Since M is dense in I, $f(M)$ is dense in $f(I)$ and so $int\overline{f(M)} = int(f(I)) \neq \emptyset$. ☺

Lemma 3.3 *Let $\lambda : R \to R$ be locally Lipschitz on R. Suppose $\Phi : U \to R$ is locally Lipschitz where U is an open set in R^N. Define $f : U \to R$ by $f(x) := \lambda(\Phi(x))$. Then*

$$\nabla f(x) = \lambda'(\Phi(x))\nabla\Phi(x) \quad \text{a.e. on } U.$$

Proof. By the Rademacher Theorem [14] $\nabla f(x)$ and $\nabla\Phi(x)$ exist on U a.e. and λ' exists a.e. on R.
Suppose that $\nabla f(x) \neq \lambda'(\Phi(x))\nabla\Phi(x)$ on $E \subset U$ with $\mu(E) > 0$, then for some i, we have

$$E_i := \{x| \quad \frac{\partial f}{\partial x_i}(x) \neq \lambda'(\Phi(x))\frac{\partial \Phi}{\partial x_i}(x),\ x \in U\}$$

has positive measure. Express R^N as $R(e_i) \bigoplus e_i^{\perp}$. By the Fubini Theorem (Theorem 6.124 [28]), for some $\bar{x} \in e_i^{\perp}$, $A := \{t| \quad \bar{x} + te_i \in E_i\} \subset R$ has positive measure. Note that λ is absolutely continuous and Φ is locally Lipschitz. By Lemma 6.87 [28] and the Rademacher Theorem, we can use Theorem 6.93 [28] to get

$$\frac{\partial f}{\partial x_i}(\bar{x} + te_i) = \lambda'(\Phi(\bar{x} + te_i))\frac{\partial \Phi}{\partial x_i}(\bar{x} + te_i) \quad \text{a.e. on } A.$$

This is a contradiction. ☺

Note that the function λ' is really short–hand notation for any function agreeing a.e. with λ'. We now give an example showing that Lemma 3.3 is not true if the range of Φ is in higher dimensions.

Example 30 Let $\Phi : R \to R^2$, $\Phi(x) := (x, 0)$ for any $x \in R$ and $\lambda : R^2 \to R$, $\lambda(a, b) \equiv 0$ for any $(a, b) \in R^2$. Now take $h : R^2 \to R^2$, $h(a, b) = (1, 0)$ if $b = 0$ and $(0, 0)$ otherwise. Then $\nabla\lambda = h$ a.e. on R^2, $f' \equiv 0$ on R, but $\langle h(\Phi(x)), \nabla\Phi(x)\rangle = 1$ for any $x \in R$. □

The next Lemma is a special case of Lemma 6.92 in [28].

Lemma 3.4 *Suppose that f is a locally Lipschitz 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 on E.*

3.2.2 Proof of the main theorem

Proof. The proof is presented in two parts.

Part I. Let $\{y_n : n \in \mathbf{N}\}$ be a dense subset of $S(R^N)$ 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 $\{h_1, h_2, \cdots, h_j\}$ defined on U, there exists a real-valued locally Lipschitz function g defined on U such that:

($\mathbf{a_j}$) $\partial_c g(x) \subset conv\{\partial_c h_1(x), \partial_c h_2(x), \cdots, \partial_c h_j(x)\}$ for each $x \in U$ and $\nabla g(x) \in \{\nabla h_1(x), \nabla h_2(x), \cdots, \nabla h_j(x)\}$ almost everywhere in U;

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

We proceed by induction.

Step 1. Let h_1 be any real-valued locally Lipschitz function defined on U and let $g \equiv h_1$. Then clearly g satisfies ($\mathbf{a_1}$) and ($\mathbf{b_1}$) 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 $k_1, k_2, \cdots, k_m$ defined on U, there exists a locally Lipschitz function g defined on U such that ($\mathbf{a_m}$) and ($\mathbf{b_m}$) are satisfied with respect to the functions $k_1, k_2, \cdots, k_m$.

Step m+1. Let $h_1, h_2, \cdots, h_{m+1}$ 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 ($\mathbf{a_m}$) and ($\mathbf{b_m}$) with respect to the locally Lipschitz functions $c_1, c_2, \cdots, c_m$.

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

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

Let us also set

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

It follows from Remark 3.1 that each member of G is Lebesgue measurable and that $\mathrm{int}\overline{G_n} \neq \emptyset$ for each $n \in \mathbf{N}$. Therefore the family of sets G satisfies the hypothesis of Lemma 3.2. Let $E \equiv \bigcup\{E_n : n \in \mathbf{N}\}$ be the subset of R given in Lemma 3.2 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

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

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

To see that g_{m+1} satisfies $(\mathbf{a_{m+1}})$. Observing that since E is a Borel subset of R, we may apply Lemma 3.3 to get that $\nabla g_{m+1}(x) = \chi_E(g(x)) \cdot \nabla g(x)$ almost everywhere in U. Now by assumption $\nabla g(x) \in \{\nabla c_1(x), \nabla c_2(x), \cdots, \nabla c_m(x)\}$ almost everywhere in U. Therefore,

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

almost everywhere in U. Furthermore, by Lemma 3.1 we also have that

$$\partial_c g_{m+1}(x) \subset conv\{\partial_c c_1(x), \partial_c c_2(x), \cdots, \partial_c c_{m+1}(x)\} \quad \text{for each } x \in U.$$

Next we show that g_{m+1} satisfies $(\mathbf{b_{m+1}})$. 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 3.2, $\mu(G(n,p,r,k) \cap E) > 0$. 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 locally Lipschitz on $U_r(n,p)$, so $\mu(A) > 0$. Therefore by Lemma 3.3

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

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 3.4, $g'(x_n + ty_p; y_p) = 0$ for almost all $t \in M_m(n,p,k) \cap U_r(n,p)$ and so by Lemma 3.3, 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 $t \in M_m(n,p,k) \bigcap U_r(n,p)$. 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 three more cases:

(iii) Suppose that $\mu(G(n,p,r,1)) = 0$. Then by Lemma 3.4, $g'(x_n + ty_p; y_p) = 0$ for almost all $t \in M_m(n,p,1) \cap U_r(n,p)$ and so by Lemma 3.3

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

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 $\mu(G(n,p,r,1) \setminus E) > 0$. Let

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

Since the mapping $t \to g(x_n + ty_p)$ is locally Lipschitz on $U_r(n,p)$, $\mu(A) > 0$. Now by Lemma 3.3

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

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

(v) Suppose that $\mu(G(n,p,r,1)) > 0$ and $\mu(G(n,p,r,1) \setminus E) = 0$. Recall that since $\mu(G(n,p,r,1)) > 0$, $G(n,p,r,1) = G_s$ for some $s \in \mathbf{N}$ and so by the construction of the set E given in Lemma 3.2, we know that $\mu(G_s \setminus \bigcup\{E_j : 1 \le j \le s\}) = 0$ and $G_s \setminus \bigcup\{E_j : 1 \le j \le s\} \ne \emptyset$. Let

$$A \equiv \{t \in M_m(n,p,1) \cap U_r(n,p) : g(x_n + ty_p) \in G_s \setminus (\bigcup\{E_j : 1 \le j \le s\})\}.$$

We claim that A has positive measure. To see this, observe that since g is locally Lipschitz and $\bigcup\{E_j : 1 \leq j \leq s\}$ is closed, $\{t \in U_r(n,p) : g(x_n + ty_p) \notin \bigcup\{E_j : 1 \leq j \leq s\}\} \equiv V$ is a non-empty open subset of $U_r(n,p)$. The proof of the claim is completed by noticing that $A = M_m(n,p,1) \bigcap V$. Now, by Lemma 3.4 we must have $g'(x_n + ty_p; y_p) = 0$ for almost all $t \in A$ and so

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

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

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

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

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

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

almost everywhere in U. Now by Lemma 3.1 we also have

$$\partial_c e(x) \subset conv\{\partial_c h_1(x), \partial_c h_2(x), \cdots, \partial_c h_{m+1}(x)\}$$

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 $(\mathbf{a_{m+1}})$ and $(\mathbf{b_{m+1}})$ with respect to the locally Lipschitz functions h_1, h_2, $\cdots$, h_{m+1} defined on U; which completes the induction.

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

$$\partial_c g(x) = conv\{\partial_c f_1(x), \partial_c f_2(x), \cdots, \partial_c f_n(x)\} \text{ for each } x \in U.$$

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

$$conv\{\partial_c f_1(x), \partial_c f_2(x), \cdots, \partial_c f_n(x)\} \subset \partial_c g(x) \text{ for each } x \in U.$$

Since each function f_j, $1 \leq j \leq n$, possesses a minimal subdifferential mapping, by Theorem 1.2 [2] there exists a dense G_δ subset G_j of U such that f_j is strictly differentiable at each point of G_j. Let $G \equiv \bigcap\{G_j : 1 \leq j \leq n\}$. Clearly G is a dense G_δ subset of U. We show that $\{\nabla f_1(x), \nabla f_2(x), \cdots, \nabla f_n(x)\} \subset \partial g(x)$ for each $x \in G$.

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

$$\langle \nabla f_j(x_o), y \rangle > \alpha > \max_{\xi \in \partial_c g(x_o)} \langle y, \xi \rangle = g^o(x_o; y).$$

Moreover, since $\partial_c g(x_o)$ is a bounded subset we may choose $y = y_p \in \{y_n : n \in \mathbf{N}\}$. By the definitions of $g^o(x_o; y)$ and strict differentiability, we know that there exists an open neighbourhood V of x_o 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 $x_m \in \{x_k : k \in \mathbf{N}\} \bigcap V$ 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

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

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

Next, we fix $1 \leq k \leq n$ and consider the function f_k. Since $\partial_c f_k$ is minimal, by Theorem 2.2 [1] we have $\partial_c f_k = CSC(\sigma)$ where σ is any densely defined selection $\sigma \in \partial_c f_k$. Define $\sigma(x) := \nabla f_k(x)$ when $x \in G$. Then σ is a densely defined selection of $\partial_c g$ and $\partial_c f_k$. As shown in Corollary 4.2 [2] $CSC(\sigma)(x) \subset \partial_c g(x)$ for all $x \in U$. Thus $\partial_c f_k(x) \subset \partial_c g(x)$ for each $x \in U$. Since k, $1 \leq k \leq n$, is arbitrary and $\partial_c g(x)$ is convex, we must have

$$conv\{\partial_c f_1(x), \partial_c f_2(x), \cdots, \partial_c f_n(x)\} \subset \partial_c g(x) \text{ for each } x \in U.$$

☺

3.2.3 Application

Theorem 3.2 provides a rich source of pathological examples.

Example 31 Let f_1, f_2, $\cdots$, $f_n \in S_e(U)$. Then there exists a real-valued locally Lipschitz function g defined on U such that

$$\partial_c g(x) = conv\{\partial_c f_1(x), \partial_c f_2(x), \cdots, \partial_c f_n(x)\} \text{ for each } x \in U.$$

Moreover g is C–minimal and C–integrable if and only if $\partial_c f_1 = \partial_c f_2 = \cdots = \partial_c f_n$.

To see this, it suffices to know that all functions in $S_e(U)$ are C–minimal and C–integrable. □

Example 32 Let a_1, $a_2, \cdots$, $a_n \in R^N$. That is any finite number of vectors in R^N. For each $1 \leq k \leq n$, define $f_k : R^N \to R$ by $f_k(x) := \langle a_k, x\rangle$. Then we know $f_k \in S_e(R^N)$ for $1 \leq k \leq n$ and so there exists a real–valued locally Lipschitz function $g : R^N \to R$ such that

$$\partial_c g(x) = conv\{a_1, a_2, \cdots, a_n\} \text{ for all } x \in R^N.$$

Therefore every polytope arises as the constant C–subdifferential of a class of globally Lipschitz functions. Moreover, g has prescribed derivatives. That is, ∇g takes every point a_i, $1 \leq i \leq n$, positively in every neighborhood and takes no other values except possibly on sets of measure zero.

Let $f : R^N \to R$ be locally Lipschitz and strictly Gateaux differentiable on U. Then by Proposition 2.3.3 [13]

$$\partial_c(f + g)(x) = \nabla f(x) + P \text{ for all } x \in U.$$

Thus any C^1 polytope is a C–subdifferential map. □

Example 33 Let $\{T_1, T_2, \cdots, T_n\}$ be a finite family of maximal cyclically monotone operators defined on an open set $U \subset R^N$. Then T_k is a minimal cusco on U for each $1 \leq k \leq n$ by Theorem 7.9 [25]. By Theorem 3.1, for each $1 \leq k \leq n$ we can find a locally Lipschitz function $f_k : U \to R$ such that $\partial f_k(x) = T_k(x)$ for any $x \in U$. By Theorem 3.2 there exists a real-valued locally Lipschitz function g defined on U such that

$$\partial_c g(x) = conv\{T_1(x), T_2(x), \cdots, T_n(x)\} \text{ for each } x \in U.$$

Thus the convex hull of a finite family of maximal cyclically monotone operators is still a C–subdifferential. □

Example 34 Let $f_1, f_2, \cdots, f_n$ be continuously Gateaux differentiable on U and such that for all $x \in U$ there are $i \neq j$ such that $\nabla f_i(x) \neq \nabla f_j(x)$. By Theorem 3.2 there exists $g : U \to R$ locally Lipschitz such that

$$\partial_c g(x) = conv\{\nabla f_1(x), \nabla f_2(x), \cdots, \nabla f_n(x)\} \text{ for all } x \in U.$$

Note that $\partial_c g$ is nowhere singleton on U. In [15] Giles and Sciffer proved that every locally Lipschitz function is generically C–pseudo-regular on the separable Banach spaces. Observe that if f is pseudo-regular and Gateaux differentiable at x_o, then $\partial_c f(x_o)$ is singleton. The function we constructed is locally Lipschitz and Gateaux differentiable almost everywhere. It shows that g is Gateaux differentiable only on a first category set even though this set is big in measure. In addition, g is almost everywhere C–irregular. We also note that g is nowhere strictly differentiable. □

Example 35 Let $f : R^N \to R$ be a locally Lipschitz function with $\partial_c f$ being a polytope, as ensured by Theorem 3.2. By Proposition 2.9.6 [13, page 102], we have

$$N_{epif}(x, f(x)) = \bigcup_{\lambda > 0} \lambda[\partial_c f(x), -1].$$

Thus N_{epif} is a constant set map. This gives us a globally Lipschitz function on R^N whose C–normal cone to its epigraph is always a constant set. □

Example 36 Theorem 3.2 provides us a technique to construct the globally Lipschitz functions which are nowhere monotonic.

Let $[a, b] = [-1, 1]$ in R. Then there exists a globally Lipschitz function f on R with $\partial_c f(x) = [-1, 1]$. Our construction shows that both $\{x : f'(x) = 1, x \in R\}$ and $\{x : f'(x) = -1, x \in R\}$ are dense in R. Thus f is nowhere monotonic on R. □

3.3 A convergence theorem for C–subdifferentials

Example 37 Let $\{f_k\}$ be a sequence of convex functions converging pointwise to a convex function f and take $x \in \text{dom}f$. For any sequence $s_k \in \partial_c f_k(x)$, the cluster points of $\{s_k\}$

are all in $\partial_c f(x)$. But the converse inclusion is not true in general. Define $f_k : R \to R$ by $f_k(x) := \sqrt{x^2 + \frac{1}{k}}$. When $k \to +\infty$, f_k converges uniformly to f given by $f(x) := |x|$. However

$$\partial_c f_k(0) = \{0\} \text{ and } \partial_c f(0) = [-1, 1].$$

□

In order to discuss the convergence of convex sets, we need to have a measure of the "distance" between two subsets A and B of R^N. We would like the distance between two sets to be small only if the two sets are nearly the same, both in shape and position.

Definition 3.2 *Let A be a nonempty convex subset of R^N. The* **parallel body** *A_ϵ is defined to be*

$$A_\epsilon := A + B_\epsilon(0)$$

where $B_\epsilon(0)$ is the ball centered at 0 with radius ϵ.

Definition 3.3 *Let A and B be nonempty compact convex subsets of R^N. Then the distance D between A and B is defined as*

$$D(A, B) := \inf\{\epsilon : A \subset B_\epsilon \text{ and } B \subset A_\epsilon\}.$$

As shown in [23] D is a metric on the collection $\mathcal{C}$ of all nonempty compact convex subsets of R^N. D is a special case of Hausdorff metric for compact sets. Knowing that D makes $\mathcal{C}$ into a metric space enables us to consider the convergence of sequences of convex sets.

Definition 3.4 *A sequence $\{A_i\}$ of compact convex subsets of R^N is said to converge to a set A if*

$$\lim_{i \to \infty} D(A_i, A) = 0$$

We say that A is the limit of the sequence $\{A_i\}$ and $\lim_{i\to\infty} A_i = A$.

The limit of a sequence of convex sets is also a convex set. The Blaschke Selection Theorem [23] asserts that a uniformly bounded infinite subcollection of $\mathcal{C}$ contains a sequence that converges to a member of $\mathcal{C}$. Using this metric, we obtain the following convergence theorem for C–subdifferentials.

Theorem 3.3 *Suppose* $\{f_i\}_{i=1}^{\infty}$ *are Lipschitz on* R^N, $\sum_{i=1}^{\infty} f_i(x_o) < \infty$ *for some* $x_o \in R^N$ *and* $\sum_{i=1}^{\infty} \|\nabla f_i\|_\infty < \infty$ *in* R^N. *Then* $f := \sum_{n=1}^{\infty} f_n$ *is Lipschitz in* R^N

$$\nabla f = \sum_{i=1}^{\infty} \nabla f_i \quad a.e.\ in\ R^N,$$

and

$$\partial_c f(x) = \lim_{k\to\infty} \partial_c(\sum_{i=1}^{k} f_i)(x) \quad for\ any\ x \in R^N.$$

Proof. First we prove $s_k : R^N \to R$ defined by $s_k(x) := \sum_{i=1}^{k} f_i(x)$ converges uniformly on any compact set in R^N as $k \to \infty$. Applying the Fundamental Theorem of Calculus to the function $s_m - s_n$ to obtain

$$\begin{aligned} |(s_m - s_n)(x) - (s_m - s_n)(y)| &\leq \int_0^1 |\langle(\nabla s_m - \nabla s_n)(x + t(y-x)), (y-x)\rangle| dt \\ &\leq \sqrt{N} \sum_{i=m+1}^{n} \|\nabla f_i\|_\infty \|x - y\|. \end{aligned}$$

If we take $y = x_o$, we have

$$\begin{aligned} |(s_m - s_n)(x)| &\leq |(s_m - s_n)(x) - (s_m - s_n)(x_o)| + |(s_m - s_n)(x_o)| \\ &\leq \sqrt{N} \sum_{i=m+1}^{n} \|\nabla f_i\|_\infty \|x - x_o\| + |\sum_{i=m+1}^{n} f_i(x_o)|. \end{aligned}$$

It follows that on any compact subset of R^N, $\sum_{i=1}^{k} f_i$ converges uniformly to a function f. Since $\{s_n\}_{n=1}^{\infty}$ are equi-Lipschitz, it is obvious that f is Lipschitz globally.
To prove the second claim, we consider two cases. On R, let $s_n(x) := \sum_{i=1}^{n} f_i(x)$. Using the Rademacher Theorem, we have

$$s_n'(x) = \sum_{i=1}^{n} f'(x) \quad \text{a.e. on } R.$$

Then for any $t \in R$, using Theorem 6.85 [28]

$$\begin{aligned} \int_a^t s_n'(x)dx &= \int_a^t \sum_{i=1}^{n} f_i'(x)dx \\ &= \sum_{i=1}^{n} \int_a^t f_i'(x)dx \\ &= \sum_{i=1}^{n} (f_i(t) - f_i(a)) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} f_i(t) - \sum_{i=1}^{n} f_i(a) \\
&= s_n(t) - s_n(a).
\end{aligned}
$$

Since $s_n'(x) \leq \sum_{i=1}^{\infty} \|\nabla f_i\|_\infty$ a.e. for all n, we use the Lebesgue Dominated Convergence Theorem (Theorem 6.22 [28]) to get

$$f(t) - f(a) = \int_a^t g(x)dx \quad \text{where } g(x) = \textstyle\sum_{i=1}^{\infty} f_i'(x) \text{ a.e. on } R.$$

Hence $f' = g$ a.e. on R.
On R^N, consider the partials for f. By the above,

$$S_i = \{x : \quad \frac{\partial f}{\partial x_i}(x) = \sum_{k=1}^{\infty} \frac{\partial f_k}{\partial x_i}(x), \ x \in R^N\}$$

is full measure in R^N. If S_i is not full measure in R^N, then there exists a $E \subset R^N$ with $\mu(E) > 0$ such that $\frac{\partial f}{\partial x_i}(x) \neq \sum_{k=1}^{\infty} \frac{\partial f_k}{\partial x_i}(x)$ on E. Since R^N can be expressed as $R(e_i) + e_i^\perp$, using the Fubini Theorem, we have for some $\bar{x} \in e^\perp$, the set $A := \{t |\ \bar{x} + te_i \in E\}$ has positive measure. Since $A \subset R$, $\frac{\partial f}{\partial x_i}(\bar{x} + te_i) = \sum_{k=1}^{\infty} \frac{\partial f}{\partial x_i}(\bar{x} + te_i)$ a.e. on A, which is a contradiction. So $S := \bigcap_{i=1}^{N} S_i$ is full measure in R^N. By the Rademacher Theorem, ∇f exists a.e. Denoting these differentiability points in R^N by $G(f)$, we see that $G(f)$ is full measure in R^N. Thus $G(f) \bigcap S$ is full measure in R^N, which means

$$\nabla f(x) = \sum_{i=1}^{\infty} \nabla f_i(x) \quad \text{a.e. on } R^N .$$

Now for any $\epsilon > 0$, there exists N_o such that whenever $k \geq N_o$, $\|\nabla f - \nabla(\sum_{i=1}^{k} f_i)\|_\infty \leq \epsilon$ in R^N. This means for $k \geq N_o$ and fixed x

$$\partial_c f(x) \subset \partial_c(\sum_{i=1}^{k} f_i)(x) + B_\epsilon(0) \quad \text{and} \quad \partial_c(\sum_{i=1}^{k} f_i)(x) \subset \partial_c f(x) + B_\epsilon(0).$$

Hence $\partial_c f(x) = \lim_{k\to\infty} \partial_c(\sum_{i=1}^{k} f_i)(x)$ in the Hausdorff metric. ☺

Conjecture 1 *Since any compact convex body in R^N can be approximated by polytopes in the Hausdorff metric, we wonder whether we can construct a locally Lipschitz function on R^N with any given compact convex body as its C-subdifferential identically by using the Lipschitz functions whose C–subdifferentials are polytopes.*

Jouini [21] has given a construction showing that there is a locally Lipschitz function with any given convex compact set as its C–subdifferential identically in R^N. However his proof has some difficulty using the canonical projection from l^1 onto its subspace and seems irrevocably flawed.

Remark 3.2 *It is impossible to construct a globally Lipschitz function with a polytope as its MP–subdifferential identically since $\partial_{mp}f$ is singleton if and only if f is Gateaux differentiable. Thus it is also impossible to construct two locally Lipschitz functions such that they have the same MP–subdifferential but their difference is a not constant. This shows that MP–subdifferential determines the original function uniquely up to a constant.*

Chapter 4

Approximate subdifferential

In this chapter we construct some Lipschitz functions with prescribed A–subdifferentials.

4.1 Examples

We begin with some Lipschitz functions whose C–subdifferential and A–subdifferential are different on a given countable set.

Example 38 Let $h : R \to R$ be a monotonically decreasing function which is discontinuous at a countable set of points $\{a_n\}$. Define $g(x) := \int_0^x h(s)ds$. Then g is a concave continuous function and at each a_n we have

$$\partial_c g(a_n) = [h(a_n^+), h(a_n^-)]$$

and

$$\partial_a g(a_n) = \{h(a_n^+), h(a_n^-)\}.$$

Thus for a concave function $\{x : \partial_a g(x) \neq \partial_c g(x)\} = \{x : \ g'(x) \text{ does not exist}\}$. □

For a nonconcave Lipschitz function $\partial_c f$ and $\partial_a f$ can be different exactly at a given countable set while $\partial_a f(a_n)$ is the union of disjoint intervals for each n, as we show in the next example.

Example 39 We follow the construction in [5]. Suppose we are given a sequence $\{a_n : n = 1, 2, \ldots\}$ of $[0, 1]$. Assume $a_m \neq a_n$ for $m \neq n$. Let

$$\phi(x) := \begin{cases} -|x| + \epsilon x^2 \cdot \sin(1/x), & \text{if } x \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

where $0 \leq \epsilon \leq 1$ and define

$$g(x) := \sum_{n=1}^{\infty} \phi(x - a_n)/2^n.$$

Then g is strictly differentiable except at a_n. Also $\partial_a\phi(0) = [-1-\epsilon, -1+\epsilon] \cup [1-\epsilon, 1+\epsilon]$ and $\partial\phi_c(0) = [-1-\epsilon, 1+\epsilon]$. Letting

$$g_m(x) := \sum_{n \neq m}^{\infty} \phi(x - a_n)/2^n.$$

Since g_m is strictly differentiable at a_m, we have

$$\partial_a g(a_m) = g_m'(a_m) + \partial_a\phi(0)/2^m,$$

and

$$\partial_c g(a_m) = g_m'(a_m) + \partial_c\phi(0)/2^m.$$

Then

(i) When $\epsilon = 0$, $\partial_a g(a_n) \neq \partial_c g(a_n)$ while $\partial_a g(a_n)$ consists of only the extreme points of $\partial_c g(a_n)$;

(ii) When $0 < \epsilon < 1$, $\partial_a g(a_n) \neq \partial_c g(a_n)$ while $\partial_a g(a_n)$ is a union of two disjoint intervals;

(iii) When $\epsilon = 1$, $\partial_a g(x) = \partial_c g(x)$ for all $x \in [0, 1]$.

□

Definition 4.1 *A set $E \subset X$ is called connected if there do not exist open sets O_1 and O_2 in X, with $E \subset O_1 \cup O_2$ and $E \cap O_1 \cap O_2 = \emptyset$.*

The next example shows that in R^2 we can make $\partial_a f(a_n) \neq \partial_c f(a_n)$ even when $\partial_a f(a_n)$ is connected.

Example 40 Let $f(x) := -\|x\|$ on R^2. If the norm is the l_2 norm then

$$\partial_a f(0) = S(B_{l^2}(0)) \text{ and } \partial_c f(0) = B_{l^2}(0)$$

where $S(B_{l^2}(0))$ and $B_{l^2}(0)$ are the unit ball sphere and the unit ball of the l^2 norm respectively.
If the norm is the l_1 norm then

$$\partial_a f(0) = \{(-1,-1),(-1,1),(1,-1),(1,1)\} \text{ and } \partial_c f(0) = B_{l^\infty}(0).$$

Let the norm be the l^2 norm. Define $f : [0,1] \times [0,1] \to R$ by

$$f(x) := -\sum_{n=1}^{\infty} \frac{1}{3^n}\|x - a_n\|$$

where the a_n's are points in $[0,1] \times [0,1]$ with rational coordinates. Then f is Lipschitz and is strictly differentiable except at each a_n. Letting $g_m(x) := -\sum_{n\neq m}^{\infty} \frac{1}{3^n}\|x - a_n\|$, then g_m is strictly differentiable at a_m and we see that $\partial_c f(a_n) \neq \partial_a f(a_n)$ since

$$\partial_c f(a_n) = \nabla g_n(a_n) + \frac{1}{3^n} B_{l^2}(0),$$

and

$$\partial_a f(a_n) = \nabla g_n(a_n) + \frac{1}{3^n} S(B_{l^2}(0)).$$

□

4.2 The gradient ranges of differentiable bump functions

Definition 4.2 *A* **convex body** *in R^N is a bounded convex subset C such that $int(C) \neq \emptyset$.*

Lemma 4.1 *Let S be a strictly convex nonempty closed set. Then for each $d \neq 0$ the* **face**

$$F_S(d) := \{s \in S : \langle s, d\rangle = \sigma_S(d)\}$$

is at most singleton.

Proof. Let $s_1, s_2 \in F_S(d)$. Then we have $\langle s_1, d\rangle = \sigma_S(d)$ and $\langle s_2, d\rangle = \sigma_S(d)$. Thus $\langle \frac{s_1+s_2}{2}, d\rangle = \sigma_S(d)$. Since s is strictly convex, $\frac{s_1+s_2}{2} \in intS$. Taking $\epsilon > 0$ sufficiently small such that $\frac{s_1+s_2}{2} + \epsilon d \in S$ we obtain $\frac{s_1+s_2}{2}$ is not a maximizer of $\langle \cdot, d\rangle$, a contradiction. Hence $s_1 = s_2$. ☺

Definition 4.3 *Let C be a closed convex set containing the origin. The function ν_C defined by*

$$\nu_C(x) := \inf\{\lambda > 0 : x \in \lambda C\}$$

is called the **gauge** *of C. As usual, we set $\nu_C(x) := +\infty$ if $x \in \lambda C$ for no $\lambda > 0$.*

As shown by Theorem 1.2.5 [17] if $0 \in \text{int}C$ then ν_C is finite everywhere and ν_C is a nonnegative closed sublinear function. Our constructions used later are based on the bump functions built from gauges. The following two lemmas from [17] show the relationship between gauges and support functions.

Lemma 4.2 *Let C be a closed convex set containing the origin. Its gauge ν_C is the support function of a closed convex set containing the origin, namely*

$$C^o := \{s \in R^N : \langle s, d\rangle \leq 1 \quad \text{for all } d \in C\}$$

which defined the **polar** *of C.*

Proof. Since ν_C is closed, sublinear and nonnegative, it is the support function of some closed convex set containing 0, say D. Then

$$D := \{s \in R^N : \langle s, d\rangle \leq r \text{ for all } (d, r) \in epi\ \nu_C\}.$$

However $epi\ \nu_C$ is the closed convex conical hull of $C \times \{1\}$; we can use the homogeneity to write

$$D = \{s \in R^N : \langle s, d\rangle \leq 1 \text{ for all } d \text{ such that } \nu_C(d) \leq 1\}.$$

The result follows from the observation that $C = \{d \in R^N : \ \nu_C(d) \leq 1\}$. ☺

Lemma 4.3 *Let $f : R^N \to R$ be convex. For all x and d in R^N, we have*

$$F_{\partial_c f(x)}(d) = \partial_c[f'(x, \cdot)](d).$$

Proof. If $s \in \partial_c f(x)$ then

$$f'(x, d') \geq \langle s, d'\rangle \quad \text{for all } d' \in R^N$$

simply because $f'(x, \cdot)$ is the support function of $\partial_c f(x)$. If, in addition, $\langle s, d\rangle = f'(x, d)$, we get

$$f'(x, d') \geq f'(x, d) + \langle s, d' - d\rangle \quad \text{for all } d' \in R^N$$

which proves the inclusion $F_{\partial_c f(x)}(d) \subset \partial_c[f'(x,\cdot)](d)$.

Conversely, let $s \in \partial_c[f'(x,\cdot)](d)$. Then

$$f'(x,d') \geq f'(x,d) + \langle s, d' - d\rangle \quad \text{for all } d' \in R^N. \tag{4.1}$$

Set $d'' := d' - d$ and deduce from the subadditivity

$$f'(x,d) + f'(x,d'') \geq f'(x,d') \geq f'(x,d) + \langle s, d''\rangle \text{ for all } d'' \in R^N$$

which implies $f'(x,\cdot) \geq \langle s,\cdot\rangle$, hence $s \in \partial_c f(x)$. Also, putting $d' = 0$ in Equation 4.1 shows that $\langle s, d\rangle \geq f'(x,d)$. Altogether, we have $s \in F_{\partial_c f(x)}(d)$. ☺

Let C be a convex body. Define $F(x) := \frac{3\sqrt{3}}{8}[(1-\nu_C^2(x))^+]^2$. That is,

$$F(x) = \begin{cases} \frac{3\sqrt{3}}{8}(1-\nu_C^2(x))^2, & \text{if } x \in C \\ 0, & \text{if } x \notin C. \end{cases}$$

Theorem 4.1 *Let F be defined as above. If C^o is strictly convex then F is smooth and $R(\nabla F) = C^o$.*

Proof. Let $x \neq 0$. By Lemma 4.1, Lemma 4.2 and Lemma 4.3 we have:

$$\partial_c \nu_C(x) = F_{C^o}(x).$$

By Theorem 2.3.10 [13, page 45]

$$\begin{aligned} \partial_c F(x) &= \frac{3\sqrt{3}}{2}(1-\nu_C^2(x))^+\nu_C(x)\partial_c\nu_C(x) \\ &= \frac{3\sqrt{3}}{2}(1-\nu_C^2(x))^+\nu_C(x)\nabla\nu_C(x). \end{aligned}$$

Hence

$$\begin{aligned} R(\nabla F) &= \bigcup_{x\in R^N} \frac{3\sqrt{3}}{2}(1-\nu_C^2(x))^+\nu_C(x)\nabla\nu_C(x) \\ &= \bigcup_{\sigma\geq 0} \frac{3\sqrt{3}}{2}(1-\sigma^2)^+\sigma\nabla\nu_C(x) \\ &= \bigcup_{1\geq\sigma\geq 0} \frac{3\sqrt{3}}{2}\sigma(1-\sigma^2) \bigcup_{\substack{x\in R^N \\ x\neq 0}} F_{C^o}(x) \\ &= \bigcup_{1\geq\sigma\geq 0} \frac{3\sqrt{3}}{2}\sigma(1-\sigma^2)\mathrm{bd}C^o \\ &= C^o. \end{aligned}$$

where $\frac{3\sqrt{3}}{2}\sigma(1-\sigma^2)$ is maximized at $\frac{1}{\sqrt{3}}$ with value 1. Observing that $\nabla F(0)=\{0\}$, we see that

$$R(\nabla F) = \bigcup_{\substack{x\in C\\ \nu_C(x)\le\frac{1}{\sqrt{3}}}} \nabla F(x) = C^o.$$

☺

Definition 4.4 *A bump function on R^N is a real-valued function ϕ which is bounded and has nonempty bounded support $supp(\phi) := \{x \in R^N : \phi(x) \neq 0\}$.*

The following lemma is from [23].

Lemma 4.4 *Let C be a closed convex body with $0 \in int(C)$. Then C^o is a closed convex body and $0 \in int(C^o)$.*

Proof. Since $0 \in \text{int}(C)$, there exists $r > 0$ such that the closed ball $B(0,r) \subset C$. Thus

$$C^o \subset [B(0,r)]^o = B(0,1/r)$$

which implies C^o is bounded. It is clear that C^o is closed and convex. Similarly, since C is bounded, there exists $R > 0$ such that $C \subset B(0,R)$. But then $B(0,1/R) \subset C^o$, so $0 \in \text{int}(C^o)$. ☺

Theorem 4.2 *Every strictly convex closed body containing 0 in its interior is the gradient range of a continuous Gateaux differentiable bump function.*

Proof. Let C be any strictly convex closed body with $0 \in \text{int}(C)$. Define ν_{C^o} and F respectively by

$$\nu_{C^o}(x) := \inf\{t > 0 : x \in tC^o\}$$

and

$$F(x) := \frac{3\sqrt{3}}{8}[(1-\nu_{C^o}^2(x))^+]^2.$$

Then by Theorem 4.1, F is strictly Gateaux differentiable and $R(\nabla F) = C$. By Lemma 4.4, F is a bump function. ☺

Theorem 4.3 *Let C_i be strictly convex closed bodies with $0 \in int(C_i)$ for $i \in I$ (a finite set). Then there is a continuous Gateaux differentiable bump function F such that*

$$R(\nabla F) = \bigcup_{i \in I} C_i.$$

Proof. Let F_i be a bump function with $R(\nabla F_i) = C_i$ and with $suppF_i \subset C_i^o$. Define G_i by

$$G_i(x) := \epsilon_i F_i(\frac{x - x_i}{\epsilon_i}).$$

Since C_i^o is bounded and contains 0, we can choose x_i and ϵ_i appropriately such that $suppG_i \cap suppG_j = \emptyset$ if $i \neq j$. Set $A_i := suppG_i$ and define

$$G(x) := \begin{cases} G_i(x), & \text{if } x \in A_i \\ 0, & \text{otherwise.} \end{cases}$$

Then G is a continuously Gateaux differentiable bump function and $R(\nabla G) = \bigcup_{i \in I} C_i$. ☺

Remark 4.1 *It is the shape of the range of the gradient of a C^1 bump function that determines the images of the subdifferentials which are constructed later. Theorem 4.3 shows we can intersperse different bumps to get interesting images.*

Example 41 Gauges and support functions of elliptic sets deserve a more detailed study. Given a positive definite operator Q, define

$$R^N \ni x \mapsto f(x) := \sqrt{\langle Qx, x \rangle}.$$

Then f is the gauge function of the sublevel–set $E_Q := \{x : f(x) \leq 1\}$. To see this, we write

$$\begin{aligned} f(x) &= \inf\{\lambda > 0 : \langle Qx, x \rangle \leq \lambda^2\} \\ &= \inf\{\lambda > 0 : \langle Q\frac{x}{\lambda}, \frac{x}{\lambda} \rangle \leq 1\} \\ &= \inf\{\lambda > 0 : \frac{x}{\lambda} \in E_Q\}. \end{aligned}$$

Consider the polar of E_Q:

$$E_Q^o := \{y : \langle y, x \rangle \leq 1 \text{ for all } x \text{ satisfying } \langle Qx, x \rangle \leq 1\}.$$

Letting $Q^{1/2}$ be the square root of Q, the change of variable $p = Q^{1/2}x$ gives

$$\begin{aligned} E_Q^o &= \{y : \langle p, Q^{-1/2}y\rangle \leq 1 : \ \|p\|^2 \leq 1\} \\ &= \{y : \ \|Q^{-1/2}y\| \leq 1\} \end{aligned}$$

which is the dual ball of E_Q. By Lemma 4.2 the support function $\sigma_{E_Q^o}$ is exactly f. Let $F(x) := \frac{3\sqrt{3}}{8}[(1 - f(x)^2)^+]^2$. Then

$$R(\nabla F) = \{y : \ \langle y, Q^{-1}y\rangle \leq 1\}$$

which is an elliptic set. By choosing different positive definite operators Q we can get different elliptic sets as the gradient ranges of bump functions. In particular, when $Q = I_n$ we get $E_Q = E_Q^o$ and this is the only norm on R^N having this property. □

A more general fact about nonsmooth bump functions is as follows:

Theorem 4.4 *Let $b : R^N \to R$ be a Lipschitz bump function with $supp(b) \subset B_K(0)$ for some $K > 0$. Suppose there exists y_o with $b(y_o) \neq 0$. Then the following hold:*

(I) For any $\delta \leq \max_{y \in B_K(0)} \frac{|b(y)|}{K + \|y\|}$ we have

$$B_\delta(0) \subset \partial_c b(B_K(0)),$$

(II) If there is y_o with $b(y_o) < 0$ then for any $\delta \leq \max_{\substack{y \in B_K(0) \\ b(y) < 0}} \frac{-b(y)}{K + \|y\|}$ we have

$$B_\delta(0) \subset \partial_a b(B_K(0)).$$

Proof. Part I. Fix $\phi \in B_\delta(0)$ and $l > 1$. Suppose $b(y_o) > 0$. Consider

$$f(x) := \begin{cases} -b(x) + \langle \phi, x\rangle, & \text{if } x \in B_{lK}(0) \\ +\infty, & \text{otherwise} \end{cases}$$

Since $B_{lK}(0)$ is compact, f is lower semicontinuous and bounded below. By the Deville, Godefroy and Zizler version of the Smooth Variational Principle [25] for all $0 < \epsilon < 1$ there is a Lipschitz function ψ having a bounded Gateaux derivative such that:

(i) $\|\psi\|_\infty + \|\nabla\psi\|_\infty < \epsilon$;

(ii) $g := f + \psi$ attains its minimum at some $x_o \in R^N$ with $\|x_o\| \le lK$.

Since $\|y_o\| \le K$ we have

$$\begin{aligned}(-b+\phi+\psi)(y_o) &< -b(y_o) + \|\phi\|\cdot\|y_o\| + \epsilon \\ &\le -b(y_o) + \delta\|y_o\| + \epsilon.\end{aligned}$$

If $K \le \|x\| \le lK$ we have

$$\begin{aligned}(-b+\phi+\psi)(x) &= \phi(x) + \psi(x) \\ &\ge -\|\phi\|\cdot lK - \epsilon \\ &\ge -\delta\cdot lK - \epsilon.\end{aligned}$$

Observing that ϵ can be arbitrarily small, to show $x_o \in B_K(0)$ it suffices to show

$$-\delta\cdot lK > -b(y_o) + \delta\|y_o\|.$$

Let $\delta < \frac{b(y_o)}{lK+\|y_o\|}$. Then $\|x_o\| \le K$ and so $0 \in \partial_c(f+\psi)(x_o)$. By Proposition 2.3.3 [13]

$$0 \in -\partial_c b(x_o) + \partial_c\psi(x_o) + \phi.$$

Then there exists $\xi \in \partial_c b(x_o)$ such that $\|\phi - \xi\| \le \epsilon$. Since $\epsilon > 0$ is arbitrary and $\partial_c b$ is upper semicontinuous we have $\phi \in \partial_c b(B_K(0))$ and so $B_\delta(0) \subset \partial_c b(B_K(0))$.

If $b(y_o) < 0$ consider

$$f(x) := \begin{cases} b(x) - \langle\phi, x\rangle, & \text{if } x \in B_{lK}(0) \\ +\infty, & \text{otherwise}\end{cases}$$

Similarly we get

$$0 \in \partial_c b(x_o) - \phi + \partial_c\psi(x_o).$$

For the same reason we get $B_\delta(0) \subset \partial_c b(B_K(0))$. In either case δ is majorized by $\frac{|b(y_o)|}{lK+\|y_o\|}$ for any y_o with $b(y_o) \ne 0$. Then for any $\delta < \max_{y\in B_{lK}(0)} \frac{|b(y)|}{lK+\|y\|}$ the inclusion is true. Noting that $\partial_c b$ is a cusco and $B_K(0)$ is compact, so $\partial_c b(B_K(0))$ is compact [8]. Letting $l \to 1$ we see that for any $\delta \le \max_{y\in B_K(0)} \frac{|b(y)|}{K+\|y\|}$ the inclusion is true.

Part II. Let $b(y_o) < 0$. Consider

$$f(x) := \begin{cases} b(x) - \langle\phi, x\rangle, & \text{if } x \in B_{lK}(0) \\ +\infty, & \text{otherwise}\end{cases}$$

Again by the Smooth Variational Principle for all $0 < \epsilon < 1$ there is a locally Lipschitz function having a bounded Gateaux derivative such that:

(i) $\|\psi\|_\infty + \|\nabla\psi\|_\infty < \epsilon$;

(ii) $g := f + \psi$ attains its minimum at some $x_o \in R^N$ with $\|x_o\| \leq lK$.

Since $\|y_o\| \leq K$ we have

$$\begin{aligned}(b - \phi + \psi)(y_o) &< b(y_o) + \|\phi\| \cdot \|y_o\| + \epsilon \\ &\leq b(y_o) + \delta\|y_o\| + \epsilon.\end{aligned}$$

If $K \leq \|x\| \leq lK$ we have

$$\begin{aligned}(b - \phi + \psi)(x) &= -\phi(x) + \psi(x) \\ &\geq -\|\phi\| \cdot lK - \epsilon \\ &\geq -\delta \cdot lK - \epsilon.\end{aligned}$$

Observing that ϵ can be arbitrarily small, to show $x_o \in B_K(0)$ it again suffices to show

$$-\delta \cdot lK > b(y_o) + \delta\|y_o\|.$$

Let $\delta < \frac{-b(y_o)}{lK+\|y_o\|}$. Then $x_o \in B_K(0)$. By Theorem 5.6 [20] we have

$$0 \in \partial_a b(x_o) + \partial_a \psi(x_o) - \phi.$$

Therefore there exists $\xi \in \partial_a b(x_o)$ such that $\|\phi - \xi\| \leq \epsilon$. Since $\epsilon > 0$ is arbitrary and $\partial_a b$ is upper semicontinuous so $\phi \in \partial_a b(B_K(0))$, which implies $B_\delta(0) \subset \partial_a b(B_K(0))$.
Since δ is majorized by $\frac{-b(y_o)}{lK+\|y_o\|}$ for any y_o with $b(y_o) < 0$, thus for any $\delta < \max_{\substack{y \in B_{lK}(0) \\ b(y)<0}} \frac{-b(y)}{lK+\|y\|}$ the inclusion is true. Noting that $\partial_a b$ is an usco and $B_K(0)$ is compact, so $\partial_a b(B_K(0))$ is compact (see Proposition 6.2.11 [8]). Letting $l \to 1$ we get for any $\delta \leq \max_{\substack{y \in B_K(0) \\ b(y)<0}} \frac{-b(y)}{K+\|y\|}$ the inclusion is true. ☺

Remark 4.2 *The Theorem shows that if* $0 \notin int\partial_a b(B_K(0))$ *then* $b \geq 0$ *everywhere and* $0 \in int\partial_a(-b)(B_K(0))$. *We have no example with* $b \geq 0$ *but* $0 \notin int\partial_a b(B_K(0))$. *By contrast 0 is always in the interior of the range of C–subdifferential of a locally Lipschitz bump function.*

Remark 4.3 *We say that a Banach space E has property (H_G) provided there exists on E a bump function b which is Gateaux differentiable and globally Lipschitz. The DGZ version of the Smooth Variational Principle holds on a Banach space with (H_G) property [25]. Let E admit an equivalent Gateaux differentiable norm (at nonzero point), then E is called smooth and necessarily has property (H_G). Hence the results of Theorem 4.4 are still true for the Banach spaces with the (H_G) property and the smooth Banach spaces if we substitute $\partial_c b(B_K(0))$ and $\partial_a b(B_K(0))$ by their norm closures respectively. Note that if a set in E^* is w^*–closed then it is norm closed, thus the results of Theorem 4.4 hold if we replace the norm closure by w^*–closure.*

Lemma 4.5 *Let $\Omega : X \to Y$ be an usco. Suppose for each $x \in X$, $\Omega(x)$ is connected. Then $\Omega(S)$ is connected in Y if S is connected in X.*

Proof. Suppose $\Omega(S)$ is not connected, then there exist two open sets O_1 and O_2 such that

$$\Omega(S) \subset O_1 \cup O_2 \ \text{ and } \Omega(S) \cap O_1 \cap O_2 = \emptyset.$$

We show $K := \Omega^{-1}(O_1) \cap \Omega^{-1}(O_2) = \emptyset$. Since $\Omega(x)$ is connected for any x, we see that if $\Omega(x) \cap O_1 \neq \emptyset$ then $\Omega(x) \subset O_1$. Otherwise $\Omega(x)$ is not connected. As Ω is an usco, both $\Omega^{-1}(O_1)$ and $\Omega^{-1}(O_2)$ are nonempty and open. Let $x \in K$. Then there exists $y_1 \in O_1$ and $y_2 \in O_2$ such that $y_1 \in \Omega(x)$ and $y_2 \in \Omega(x)$. Thus we have $\Omega(x) \subset O_1 \cup O_2$ which implies $\Omega(x)$ is not connected. Hence $K = \emptyset$ and $S \subset \Omega^{-1}(O_1) \cup \Omega^{-1}(O_2)$ and so S is not connected. The Lemma follows by contraposition. ☺

Theorem 4.5 *Let $b : R^N \to R$ be a Lipschitz bump function with supp$(b) \subset B_K(0)$. Then $\partial_c b(B_K(0))$ is compact connected with nonempty interior.*

Proof. Since $\partial_c b$ is a cusco, $\partial_c b(x)$ is connected for each $x \in R^N$. Hence Theorem 4.4 and Lemma 4.5 apply. ☺

Remark 4.4 *In R we saw that for an everywhere differentiable function f we have $\partial_c f = \partial_a f$ by the Darboux property . Thus the range of the A–subdifferential of a differentiable bump function is always connected compact with nonempty interior. We don't know whether this is true in R^2.*

We close this section with an open question: Is the range of the C–subdifferential of a Lipschitz bump semi–closed? That is $cl[int(\partial_c b(B_K(0)))] = \partial_c b(B_K(0))$? We note that in R this is true.

4.3 Constructions of the nonconvex A–subdifferentials

In R^2, following [22] we use bump functions to construct the Lipschitz functions whose A–subdifferentials are not convex on sets with large measure.

Theorem 4.6 *Let $\epsilon > 0$. Then there exists a Lipschitz function defined on R^2 such that*

$$\mu\{x : \partial_a f(x) = \partial_c f(x)\} < \epsilon.$$

Proof. By Theorem 4.3 we choose a bump function $h : \overline{B}(0,1) \to R$ satisfying the following conditions:

(1) $h(x) \geq 0$ on $\overline{B}(0,1)$;

(2) $h(x) = 0$ for $\|x\| = 1$;

(3) $h'(x) = 0$ for $\|x\| = 1$;

(4) The set $h'(\overline{B}(0,1)) = \{h'(x) : \ \|x\| \leq 1\}$ is not convex.

Now let $\{B_n\}$ be a sequence of closed balls, $B_n := \overline{B}(z_n, r_n)$, $r_n > 0$, such that:

(5) The B_i's are pairwise disjoint;

(6) $\bigcup_{n=1}^{\infty} B_n$ is dense in R^2;

(7) $\mu(\bigcup_{n=1}^{\infty} B_n) < \epsilon$.

Let $S := \bigcup_{n=1}^{\infty} B_n^o$. We define:

$$f(x) := \begin{cases} r_n h((x - z_n)/r_n), & \text{if } x \in B_n^o \\ 0, & \text{if } x \notin S. \end{cases}$$

If $x \in S$ then $x \in B_n^o$ for some n, so f is continuously differentiable around x, with

$$\partial^- f(x) = h'((x - z_n)/r_n).$$

If $x \notin S$, then $f(x) = 0$ so x is a local minimum by assumption so $f^-(x; v) \geq 0$ for any v. Fixing a direction v, there are two possibilities: Either there is a decreasing sequence of positive numbers $t_n \to 0$ such that $x + t_n v \notin S$ or there is a $\epsilon > 0$ such that $x + tv \in S$ for all $t \in (0, \epsilon)$. In the first case we have

$$f^-(x; v) \leq \liminf_{n \to \infty} \frac{f(x + t_n v) - f(x)}{t_n} = 0.$$

Since we already know $f^-(x; v) \geq 0$, we have $f^-(x; v) = 0$. In the second case, there must be some n such that $x + tv \in B_n$ for all $t \in (0, \epsilon)$. Since the balls are closed we have $x \in B_n$ and since $x \notin S$, $x \in \partial B_n$. Using the fact that $x + tv \in B_n$ for small t we get

$$f^-(x; v) = h'((x - z_n)/r_n)v = 0.$$

Therefore in each case when $x \notin S$ we get $f^-(x; v) = 0$ in any direction, so $\partial^- f(x) = \{0\}$. For every $x \in R^2$ we have

$$\partial_a f(x) \subset h'(\overline{B}(0, 1)) \cup \{0\} = h'(\overline{B}(0, 1)).$$

If $x \notin S$, and U is any neighborhood of x then U contains a ball B_n. Therefore $h'(\overline{B}(0, 1)) \subset \partial_a f(x)$. So for $x \notin S$ we have $\partial_a f(x) = h'(\overline{B}(0, 1))$, which is nonconvex, so in particular $\partial_a f(x) \neq \partial_c f(x)$ for $x \notin S$. Since $\mu(S) < \epsilon$, we are done. ☺

Corollary 4.1 *f has the following properties:*

(I) not A–integrable;

(II) A–minimal;

(III) not C–integrable;

(IV) C–minimal.

Proof. Here we only prove (I), (II), the other two cases are similar. Let $h'(\overline{B}(0, 1)) = \bigcup_{i=1}^2 C_i$ where C_i's are elliptic sets. Take a Cantor set C in $[0, 1]$ with positive measure. Let $\bigcup_{n=1}^\infty B_n$ be dense in R^2 but miss $C \times R$. We define:

$$f_\epsilon(x, y) := f(x, y) + \epsilon \int_{-\infty}^x \chi_C(s) ds.$$

Let $g(x) := \int_{-\infty}^{x} \chi_C(s)ds$. Choose ϵ appropriately such that $\epsilon(\partial^- g(x), 0) \subset \bigcup_{i=1}^{2} C_i$. Then $\partial_a f_\epsilon(x,y) = \partial_a f(x,y)$ for all $(x,y) \in R^2$. However $f_\epsilon - f \neq constant$.

Let $S := \bigcup_{i=1}^{\infty} B_i^o$. From the construction we get $\partial_a f = USC(\partial_a f|S)$. Suppose Ω is an usco and $\Omega \subset \partial_a f$. As $\partial_a f$ is single–valued on S, $\partial_a f|S = \Omega|S$. Then $\partial_a f = USC(\partial_a f|S) = USC(\Omega|S) \subset \Omega$. Thus $\partial_a f$ is a minimal usco. ☺

Remark 4.5 *On the Cantor set C with $\mu(C) > 0$, we can use the Zahorski Theorem (see Lemma 2.9) to get an everywhere differentiable g instead of using the indefinite integral of an indicator function.*

Corollary 4.2 *Let $\epsilon > 0$ and $C \subset R^2$ be a strictly convex closed body with $0 \in int(C)$. Then there exists a Lipschitz function defined on R^2 such that*

$$\mu\{x : \partial_c f(x) \neq C\} < \epsilon.$$

Proof. From the proof of Theorem 4.6 we see that $\partial_a f(x) = h'(\overline{B}(0,1))$ when $x \notin S$. By Theorem 4.2 we can find a bump function $\tilde{h}$ with $\tilde{h}'(\overline{B}(0,1)) = C$. Substituting h by $\tilde{h}$ we get a Lipschitz function $\tilde{f}$ such that:

$$\partial_c \tilde{f}(x) = C \quad \text{if } x \notin S$$

where $\mu(S) < \epsilon$. ☺

4.3.1 An A–subdifferential that is almost always nonconvex

In this section using bump functions we construct the locally Lipschitz functions defined on R^2 with nonconvex A–subdifferentials almost everywhere. We follow a proof given by Borwein (private communication).

Theorem 4.7 *There is a Lipschitz function $f : R^2 \to R$ such that*

$$\mu\{x : \partial_a f(x) \ \ \textit{is convex}\} = 0.$$

Proof. *Step 1.* Construct sequences of closed balls $\{B_{in}\}$ such that

(1) For fixed i, B_{in}'s are pairwise disjoint;

(2) $\bigcup_{n=1}^{\infty} B_{(i+1)n}$ is dense in $S_i := \bigcup_{n=1}^{\infty} B_{in}^o$;

(3) $\mu(\bigcup_{n=1}^{\infty} B_{in}) \to 0$ as $i \to \infty$;

(4) $\bigcup_{n=1}^{\infty} B_{0n}$ is dense in R^2.

For fixed i on each $B_{in} := \overline{B}(z_n^i, r_n^i)$ we define

$$f_i(x) := \begin{cases} r_n^i h((x - z_n^i)/r_n^i), & \text{if } x \in B_{in}^o \\ 0, & \text{if } x \notin S_i \end{cases}$$

where h is any bump function with $R(\nabla h)$ nonconvex.

Since $R(\nabla h)$ is nonconvex we take $p \in conv(h'(B))$ such that $dist(p, h'(B)) = d > 0$. Let $M := \sup\{\|x^*\| : x^* \in h'(B)\}$ and $0 < k < \min(d/4M, 1/2)$. Set

$$f(x) := \sum_{n=0}^{\infty} k^n f_n(x) \quad \text{and} \quad F_m(x) := \sum_{n=m}^{\infty} k^n f_n(x).$$

We prove that $\{x : \partial_a f(x) = \partial_c f(x)\} \subset \bigcap_{n=0}^{\infty} S_n$.

Step 2. If $x \notin S_m$ then $\partial^- F_m(x) = \{0\}$. In fact, $F_m(x) = 0$ and, therefore, x is a local minimum of F_m and $F_m^-(x; v) \geq 0$ for all v. Fixing a direction v, there are two possibilities: Either there exists a $t_n \to 0$ such that $x + t_n v \notin S_m$ or there exists a $\epsilon > 0$ such that $x + tv \in S_m$ for all $t \in (0, \epsilon)$. In case one $F_m^-(x; v) = 0$. In case two, there exists an n such that $x + tv \in B_{mn}$ for all $t \in (0, \epsilon)$ because $\{B_{mn}\}$ are pairwise disjoint and closed. As B_{mn} is closed we obtain $x \in \partial B_{mn}$. Then $x \notin \bigcup_{n=1}^{\infty} B_{(i)n}$ for any $i \geq m + 1$. Therefore, there must be $t_n \to 0$ such that $x + t_n v \notin S_{m+1}$. Thus

$$\begin{aligned} F_m^-(x; v) &= \liminf_{t \to 0+} \frac{F_m(x + tv) - F_m(x)}{t} \\ &\leq \liminf_{n \to \infty} \frac{F_m(x + t_n v) - F_m(x)}{t_n} \\ &= k^m \liminf_{n \to \infty} \frac{f_m(x + t_n v) - f_m(x)}{t_n} \\ &= k^m h'((x - z_n^m)/r_n^m)v \\ &= 0. \end{aligned}$$

Therefore $F_m^-(x; v) = 0$ for all v, which is to say $\partial^- F_m(x) = \{0\}$. For any $x \in R^2$ we have

$$\partial_a F_m(x) \subset \partial^- F_m(B_{mn}) \cup \{0\} = \partial^- F_m(B_{mn}).$$

If $x \notin S_m$, and U is any neighbourhood of x then U contains a ball B_{mn}. Therefore $\partial^- F_m(B_{mn}) \subset \partial_a F_m(x)$. So for any $x \notin S_m$ we have $\partial_a F_m(x) = \partial^- F_m(B_{mn})$.

Step 3. If $x \notin S_m$ then $\partial_a F_m(x)$ is nonconvex. Since all F_m have the same structure we prove this only for $m = 0$. Let $x \in R^2 \setminus S_0$. By the Step 2, we deduce that $\partial_a f(x) = \partial^- f(B_{0n})$. We only need to show $\partial^- f(B_{0n})$ is not convex. Indeed, for any neighborhood U of x there is $B_{0n} \in S_0$ such that $B_{0n} \subset U$. For any such B_{0n}, by the definition of f_0, there exists r_1^n, $r_2^n \in B_{0n}$ such that $p \in [f_0'(r_1^n), f_0'(r_2^n)]$ and $dist(p, f_0'(B_{0n})) = d$. For any such B_{0n} and $y \in B_{0n}$, $x^* \in \partial^- F_1(y)$, we have $\|x^*\| \leq 2kM < d/2$. Since f_0 is continuously differentiable on B_{0n} we have

$$\partial^- f(y) = f_0'(y) + \partial^- F_1(y) \quad \text{for all } y \in B_{0n}.$$

Thus $dist(p, \partial^- f(B_{0n})) > d/2$. On the other hand,

$$\partial^- f(r_i^n) = f_0'(r_i^n) + \partial^- F_1(r_i^n) \ \text{ for } i = 1, 2.$$

This implies that there exists a q such that

$$\|q - p\| < d/2 \quad \text{and } q \in conv \partial^- f(B_{0n}).$$

Therefore $q \notin \partial^- f(B_{0n})$. This shows $\partial^- f(B_{0n})$ is not convex.

To conclude the proof, it suffices to consider any $x \notin \bigcap_{n=1}^{\infty} S_n$, but there exists an N such that $x \in S_N \setminus S_{N+1}$. We have

$$\partial_a f(x) = \sum_{n=1}^{N} k^n f_n'(x) + \partial_a F_{N+1}(x).$$

Since $\partial_a F_{N+1}(x)$ is not convex so is $\partial_a f(x)$. Hence $\partial_a f(x)$ is not convex for any $x \notin \bigcap_{n=1}^{\infty} S_n$. ☺

Definition 4.5 *We call f_{i+1} a generalization of f_i if they are constructed from the same bump function and satisfy:*

(I) For fixed i on each $B_{in} := \overline{B}(z_n^i, r_n^i)$

$$f_i(x) := \begin{cases} r_n^i h((x - z_n^i)/r_n^i), & \text{if } x \in B_{in}^o \\ 0, & \text{if } x \notin S_i \end{cases}$$

where h is any bump function with $R(\nabla h)$ nonconvex;

(II) For each i, $\bigcup_{n=1}^{\infty} B_{(i+1)n}$ is dense in $S_i := \bigcup_{n=1}^{\infty} B_{in}^{o}$.

Corollary 4.3 *Let f_i be defined as above. Then $g := \sum_{i=1}^{m} k^i f_i$ is A–minimal and C–minimal for each m.*

Proof. By Corollary 3.2 [6] we know that to show $\partial_a g$ is a minimal usco is equivalent to showing that for each nonempty open subset W in R^2 the restriction of $\partial_a g$ to W is a minimal usco on W.

By the definition of g it suffices to show g is minimal on B_{1n} for any n since g has the same structure on different balls. Let Ω be an usco and $\Omega \subset \partial_a g$. On B_{mn}, $\partial_a g$ is singleton because all f_i's are continuously differentiable on the ball. Then $\Omega = \partial_a g$ on B_{mn}. On $B_{(m-1)n} \setminus B_{mn}$ we have

$$\partial_a g(x) = \sum_{i=1}^{m-1} k^i f_i'(x) + k^m \partial_a f_m(x).$$

Since $\partial_a f_m$ is minimal so it is minimal on $B_{(m-1)n} \setminus B_{mn}$. Observing that $\sum_{i=1}^{m-1} k^i f_i'(x)$ is single–valued it is obvious that

$$\Omega(x) = \sum_{i=1}^{m-1} k^i f_i'(x) + k^m \partial_a f_m(x) \quad \text{on } B_{(m-1)n} \setminus B_{mn}.$$

Inductively we get

$$\Omega(x) = \sum_{i=1}^{l} k^i f_i'(x) + \partial_a\Big(\sum_{i=l+1}^{m} k^i f_i(x)\Big) \quad \text{on } B_{ln} \setminus B_{(l+1)n}$$

for $1 \le l \le m-1$. Hence $\partial_a g = \Omega$ on B_{1n}. Let $S := \bigcup_{n=1}^{\infty} B_{1n}$. Note that $\partial_a g = USC(\partial_a g|S) = USC(\Omega|S) \subset \Omega$. Therefore $\partial_a g$ is minimal. It is known that if F is a minimal usco then G defined by $G(x) := \overline{conv}[F(x)]$ defines a minimal cusco (see Proposition 3.4 [6]). Since $\partial_c g = conv[\partial_a g]$ we see that $\partial_c g$ is minimal. ☺

4.3.2 Two functions with the same C–subdifferential but different A–subdifferentials except on a small measure set

Theorem 4.8 *For any given $1 > \epsilon > 0$ there exist two locally Lipschitz functions f, f_ϵ defined on the set $[0,1] \times [0,1]$ such that*

$$\partial_c f(x,y) = \partial_c f_\epsilon(x,y) \quad \textit{for all } (x,y) \in [0,1] \times [0,1],$$

and

$$\mu\{(x,y) \in [0,1] \times [0,1] : \partial_a f(x,y) = \partial_a f_\epsilon(x,y)\} < \epsilon.$$

Proof. *Step 1.* Let $h : \overline{B}(0;1) \to R$ be a bump function with $(1,0) \notin h'(\overline{B}(0,1))$ but $(1,0) \in conv(h'(\overline{B}(0,1))$. Suppose C is a Cantor set with positive measure in $[0,1]$. Let $\{B_n\}$ be a sequence of closed balls, $B_n = \overline{B}(z_n, r_n)$, $r_n > 0$, such that:

(1) The B_i's are pairwise disjoint;

(2) $\bigcup_{n=1}^{\infty} B_n$ is dense in $[0,1] \times [0,1]$ but misses $C \times [0,1]$;

(3) $meas(\bigcup_{n=1}^{\infty} B_n) < \epsilon$.

Step 2. Let $S := \bigcup_{n=1}^{\infty} B_n^o$. We define:

$$f(x,y) := \begin{cases} r_n h(((x,y) - z_n)/r_n), & \text{if } (x,y) \in B_n^o(z_n) \\ 0, & \text{if } (x,y) \notin S \end{cases}$$

and

$$f_\epsilon(x,y) := f(x,y) + g(x) \quad \text{and} \quad g(x) := \int_0^x \chi_C(s)ds.$$

We show that $\partial_c f_\epsilon(x,y) = \partial_c f(x,y)$ for all $(x,y) \in [0,1] \times [0,1]$ but $\partial_a f_\epsilon(x,y) \neq \partial_a f(x,y)$ almost everywhere on $C \times [0,1]$.

Indeed if $x \notin C$ we have $\partial^- f_\epsilon(x,y) = \partial^- f(x,y)$; if $x \in C$, $\partial^- f_\epsilon(x,y) = (\partial^- g(x), 0)$. Noting that for almost all $x \in C$ we have $g'(x) = 1$. Thus $(1,0) \in (\partial^- g(x), 0)$ almost everywhere on $C \times [0,1]$. Since for $(x,y) \in C \times [0,1]$

$$\partial_a f_\epsilon(x,y) = \partial_a f(x,y) \cup (\partial_a g(x), 0).$$

It follows that $\partial_a f(x,y) \neq \partial_a f_\epsilon(x,y)$ almost everywhere on $C \times [0,1]$ since $(1,0) \in \partial_a f_\epsilon(x,y)$ almost everywhere on $C \times [0,1]$ whereas $(1,0) \notin \partial_a f(x,y) = h'(\overline{B}(0,1))$ for all $(x,y) \in C \times [0,1]$. By Theorem 2.82 [28] for any $0 \leq \beta < 1$ we can make $\mu(C) = \beta$. Observing that $\mu(C \times [0,1]) = \beta$ we can make $\beta > 1 - \epsilon$. Then $\partial_a f$ and $\partial_a f_\epsilon$ differ on a set with measure bigger than $1 - \epsilon$. By contrast $(1,0) \in conv(\partial_a f(x,y))$ for any $(x,y) \in C \times [0,1]$ and $\partial_a f(x,y) = \partial_a f_\epsilon(x,y)$ for any $(x,y) \in ([0,1] \setminus C) \times [0,1]$. Therefore

$$\partial_c f_\epsilon(x,y) = \partial_c f(x,y) \quad \text{for any } (x,y) \in [0,1] \times [0,1].$$

☺

Corollary 4.4 *Suppose f is defined as in the previous proof. Then f is not C–integrable.*

Bibliography

[1] J. M. Borwein, Minimal CUSCOS and Subgradients of Lipschitz Functions, in Fixed Point Theory and its Applications, Pitman Research Notes 252 (1991), 57–81.

[2] J. M. Borwein, S. P. Fitzpatrick, P. Kenderov, Minimal Convex Uscos and Monotone Operators on Small Sets, *Canadian Journal of Mathematics* **Vol. 43(3)** (1991), 461–476.

[3] J. M. Borwein, Marian Fabian, A Note on Regularity of Sets and of Distance Functions in Banach Space, *Journal of Mathematical Analysis and Applications* **182** (1994), 560–566.

[4] J. M. Borwein, S. P. Fitzpatrick, J. R. Giles, The Differentiability of Real Functions on Normed Linear Space Using Generalized Subgradients, *Journal of Mathematical Analysis and Applications* **128** (1987), 512–534.

[5] J. M. Borwein and S. P. Fitzpatrick, Characterization of Clarke Subdifferentials Among One-dimensional Multifunctions, CECM Preprint 94–006 (1994).

[6] J. M. Borwein, W. B. Moors, Essentially Strictly Differentiable Lipschitz Functions, CECM Preprint 95–029 (1995).

[7] J. M. Borwein, W. B. Moors, Wang Xianfu, Lipschitz Functions with Prescribed Derivatives and Subderivatives, CECM Preprint 94–026 (1994).

[8] G. Beer, Topologies on Closed and Closed Convex Sets, Mathematics and Its Applications, Kluwer Academic Publishers, 1993.

[9] A. M. Bruckner, Differentiation of Real Functions, Lecture Notes in Mathematics edited by A. Dold and B. Eckmann, Springer-Verlag, 1978.

[10] A. M. Bruckner, J. B. Bruckner, B. S. Thomson, Real Analysis, (to appear).

[11] A. M. Bruckner and G. Petruska, Some Typical Results on Bounded Baire–1 Functions, *Acta Math. Hung.* **43 (3–4)** (1984), 325–333.

[12] J. V. Burke, M. C. Ferris, Maijian Qian, On the Clarke Subdifferential of the Distance Function of a Closed Set, *Journal of Mathematical Analysis and Applications* **166** (1992), 199–213.

[13] F. H. Clarke, Optimization and Nonsmooth Analysis, Wiley Interscience, New York, 1983.

[14] L. C. Evans and R. F. Gariepy, Measure Theory and Fine Properties of Functions, Studies in Advanced Mathematics, 1992.

[15] J. R. Giles and S. Sciffer, Locally Lipschitz Functions are Generically Pseudo-regular on Separable Banach Spaces, *Bull. Austral. Math. Soc.* **47** (1993), 205–212.

[16] E. Hewitt, K. R. Stromberg, Real and Abstract Analysis, Springer-Verlag, New York, 1965.

[17] J. Hiriart–Urruty, C. Lemarechal, Convex Analysis and Minimization Algorithms I, Springer–Verlag Berlin Heidelberg, 1993.

[18] A. D. Ioffe, Approximate Subdifferentials and Applications 3: The Metric Theory, *Mathematika* **Vol. 36**, No. 71 (1989), 1–38.

[19] A. D. Ioffe, Approximate Subdifferentials and Applications II, *Mathematika* **33** (1986), 111–128.

[20] A. D. Ioffe, Approximate Subdifferentials and Applications I: The Finite Dimensional Theory, *Transaction of The American Mathematical Society* **281** (1984), 390–416.

[21] E. Jouini, Functions with Constant Generalized Gradients, *Journal of Mathematical Analysis and Applications* **148** (1990), 121–130.

[22] G. Katriel, Are the Approximate and the Clarke Subgradients Generically Equal?, *Journal of Mathematical Analysis and Applications* (to appear).

[23] S. R. Lay, Convex Sets and Their Applications, Jon Wiley & Sons, 1982.

[24] J. C. Oxtoby, Measure and Category, Springer-Verlag, 1971.

[25] R. R. Phelps, Convex Functions, Monotone Operators and Differentiability, Lecture Notes in Mathematics, Springer-Verlag Berlin Heidelberg, 1993.

[26] R. T. Rockafellar, Characterization of the Subdifferentials of Convex Functions, *Pacific Journal of Mathematics* **17** (1966), 497–510.

[27] H. L. Royden, Real Analysis, Macmillan Publishing Company, New York, 1988.

[28] K. R. Stromberg, An Introduction to Classical Real Analysis, Wadsworth International Mathematics Series, 1981.

[29] S. Sciffer, Regularity of Locally Lipschitz Functions on the Line, CECM Preprint 94–011 (1994).

[30] T. Zolezzi, Convergence of Generalized Gradients, *Set–Valued Analysis* **2** (1994), 381–393.

[31] M. Valadier, Entrainement Unilateral, Lignes de Descente, Fonctions Lipschitziennes non Pathologiques, *C. R. Acad. Sci. Paris*, t. 308, Serie I (1989), 241–244.

[32] M. Valadier, Lignes de Descente de Fonctions Lipschitziennes non Pathologiques, Seminaire d'Analyse Convexe Montpellier (1988), Expose no. 9.