

## DISTINCT DIFFERENTIABLE FUNCTIONS MAY SHARE THE SAME CLARKE SUBDIFFERENTIAL AT ALL POINTS

J. M. BORWEIN AND XIANFU WANG

(Communicated by J. Marshall Ash)

ABSTRACT. We construct, using Zahorski's Theorem, two everywhere differentiable real-valued Lipschitz functions differing by more than a constant but sharing the same Clarke subdifferential and the same approximate subdifferential.

### 1. INTRODUCTION

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 them by ' $\partial_c$ ', ' $\partial_{mp}$ ', ' $\partial_a$ ', and ' $\partial^-$ ' respectively. Each generalizes the notion of differentiability, primarily from a convex analytic point of view. For a convex function defined on an open interval  $I$ , if it is differentiable on  $I$  then it is continuously differentiable on  $I$ . However, a nonconvex but differentiable Lipschitz function defined on  $I$  may be discontinuously differentiable almost everywhere [1, 2]. The Clarke subdifferential and approximate subdifferential (as defined below) are **usc**os: upper semicontinuous multifunctions with compact nonempty images. Upper semicontinuity – a very desirable feature – means that both of these two subdifferentials really generalize continuous differentiability or *smoothness* rather than differentiability. Our purpose in this note is to exhibit how completely this happens.

### 2. PRELIMINARIES

We begin with some preliminary definitions. A real-valued function  $f$  defined on an open interval  $I$  is (*locally*) Lipschitz on  $I$  if for each  $x_o \in I$  there exists  $M > 0$  and  $\delta > 0$  such that  $|f(x) - f(y)| \leq M \cdot |x - y|$  for all  $x, y \in N(x_o, \delta)$ . For the functions in this class, we consider three directional derivatives. (We give these general definitions for locally Lipschitz functions defined on a Banach space  $E$ . Of course on the real line we need only consider  $h := 1$ , or  $-1$ .)

---

Received by the editors July 13, 1995 and, in revised form, September 8, 1995.

1991 *Mathematics Subject Classification*. Primary 49J52; Secondary 26A27, 26A16.

*Key words and phrases*. Lipschitz function, differentiability, integrability, generalized derivative, Clarke subdifferential, approximate continuity, metric density.

The first author's research supported by NSERC and the Shrum Endowment at Simon Fraser University.

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

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

- The **Michel-Penot derivative** at  $x$  in the direction  $h$  is given by

$$f^\diamond(x; h) := \sup_y \limsup_{t \downarrow 0} \frac{f(x + ty + th) - f(x + ty)}{t}.$$

- The **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 third derivative is only a Lipschitz function of  $h$ . Moreover,

$$f^-(x; \cdot) \leq f^\diamond(x; \cdot) \leq f^o(x; \cdot).$$

If  $f$  is convex, then

$$f^o(x; \cdot) = f^\diamond(x; \cdot) = f^-(x; \cdot).$$

The link between the generalized subdifferential map and the directional derivative is:

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

where ‘ $\#$ ’ is one of ‘ $-$ ’, ‘ $\diamond$ ’, ‘ $o$ ’. If we now set  $U(x) := \{z : |x - z| < \delta\}$ , and define

$$\partial_a f(x) := \bigcap_{\delta > 0} \bigcup_{z \in U(x)} \partial^- f(z) = \limsup_{z \rightarrow x} \partial^- f(z),$$

we obtain the **approximate subdifferential**. The Clarke subdifferential at  $x \in I$  is also given by

$$\partial_c f(x) = \text{conv}\{\lim f'(x_i) : x_i \rightarrow x, x_i \notin S, x_i \notin \Omega_f\}$$

where  $S$  is any set of Lebesgue measure 0 in  $I$  and  $\Omega_f$  is the set on which  $f$  is not differentiable. The differential concept linked to the Clarke subdifferential is that of *strict differentiability*. For a real-valued locally Lipschitz function  $f : E \rightarrow R$  we say that  $f$  admits a strict derivative  $\xi \in E^*$  if one has

$$\lim_{\substack{x' \rightarrow x \\ t \downarrow 0}} \frac{f(x' + tv) - f(x')}{t} = \langle \xi, v \rangle \quad \text{for every } v \in E.$$

Now the Clarke subdifferential has closed convex nonempty images for each  $x \in I$ , singleton at  $x$  if and only if  $f$  is strictly differentiable at  $x$ . In finite dimensions, a real-valued locally Lipschitz function  $f$  has a singleton generalized subdifferential for every  $x' \in x + \epsilon B$  if and only if  $f$  is continuously differentiable on  $x + \epsilon B$  [3]. The Michel-Penot subdifferential also has closed convex nonempty images for each  $x \in I$ , singleton at  $x$  if and only if  $f$  is differentiable at  $x$  [6]. The Dini subdifferential has closed and convex images which are often empty. The approximate subdifferential has closed, nonempty, and (usually) nonconvex images for each  $x \in I$ . In finite dimensions we have  $\partial_c f(x) = \text{conv}\{\partial_a f(x)\}$  [4, 5].

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

Throughout the paper, the Lebesgue measure is denoted by  $\lambda$ .

**Definition 2.** Let  $A \subset R$  be measurable and  $x_o \in R$ . The **upper metric density** of  $A$  at  $x_o$  is

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

where  $I \rightarrow x_o$  is read as “ $I$  contracts to  $x_o$ ”. Similarly the **lower metric density** is

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

If  $\bar{d}(x_o, A) = \underline{d}(x_o, A)$ , this number is the **metric density** of  $A$  at  $x_o$  denoted by  $d(x_o, A)$ .

**Definition 3.** A measurable function  $f$  is **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

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

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

It is obvious that if  $f$  is continuous at  $x$ , then  $f$  is approximately continuous at  $x$ . As shown in [1] 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 neighbourhood  $I$  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 will also say that a function is **integrable** if it is determined uniquely up to a constant by its C-subdifferential.

### 3. SOME LEMMAS

**Lemma 1** (Darboux’s Theorem). *Let  $f : [a, b] \rightarrow R$  be differentiable on  $[a, b]$  and suppose that  $\nu$  is a number strictly between  $f'_+(a)$  and  $f'_-(b)$ . Then there exists  $\xi \in ]a, b[$  such that  $f'(\xi) = \nu$ .*

The proof is standard (see [8, page 186]). The Darboux property allows us to show the A-subdifferential and the C-subdifferential coincide for differentiable locally Lipschitz functions on  $R$ . Indeed, let  $f$  be an everywhere differentiable locally Lipschitz function. Then

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

We claim  $\partial_a f(x) = \partial_c f(x)$ . If  $\partial_c f(x)$  is singleton, then it is obvious  $\partial_c f(x) = \partial_a f(x)$ . Let  $\alpha \in \text{int} \partial_c f(x)$ . 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  $z_n \rightarrow x$ . Therefore  $\alpha \in \partial_a f(x)$  and  $\partial_c f(x) \subset \partial_a f(x)$ . Hence  $\partial_a f(x) = \partial_c f(x)$ . Consider the **symmetric subdifferential**:  $\partial_s f(x) := \partial_a f(x) \cap (-\partial_a(-f)(x))$ . Then  $\partial_s f(x) = \partial_c f(x) = \partial_a f(x)$ .  $\square$

Hence we have shown the following:

**Theorem 1.** *Let  $f : I \rightarrow R$  be differentiable and locally Lipschitz. Then*

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

In the sequel, for any function  $f$  we denote by  $C_f$  the set of its **continuity points**. The following lemmas can be found in [1, 7].

**Lemma 2.** [*Lebesgue Density Theorem*] Let  $A \subset \mathbb{R}$  be measurable. Then for almost all  $x \in A$ ,  $A$  has metric density 1 at  $x$ . That is,

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

**Lemma 3.** Let  $A \subset \mathbb{R}$  be measurable. There is an  $F_\sigma$  set  $F \subset A$  such that

$$\lambda(A \setminus F) = 0.$$

**Lemma 4.** Let  $f$  be bounded and approximately continuous on  $I$ . Define  $F$  by  $F(x) := \int_a^x f(t)dt$  where  $a$  is any point in  $I$ . Then  $F'(x) = f(x)$  for all  $x \in I$ .

**Lemma 5.** [*Zahorski's Theorem* [1]] Let  $E$  be a set of type  $F_\sigma$  with  $d(x, E) = 1$  for all  $x \in E$ . Then there is an approximately continuous function  $f$  such that

$$0 < f(x) \leq 1 \quad \text{for all } x \in E \text{ and } f(x) = 0 \text{ for all } x \notin E.$$

The function  $f$  is moreover upper semicontinuous.

**Lemma 6.** Let  $f$  and  $g$  be approximately continuous at  $x_0$ ; then so is  $f + g$ .

We need one more lemma from [2] or [8, page 279].

**Lemma 7.** Let  $I_n := (a_n, b_n) = (c_n - h_n, c_n + h_n)$  be pairwise disjoint open intervals in  $[0, 1]$  such that the open set  $U := \bigcup_{n=1}^{\infty} I_n$  is everywhere dense in  $[0, 1]$  and  $[0, 1] \setminus U$  is a Cantor set. Let  $g_n$  ( $n = 1, 2, \dots$ ) be piecewise linear continuous such that

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

extended 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  $n = 1, 2, \dots$ ,  $C_g = U$  and  $g(x) = 0$  for  $x \notin U$ ;
- (ii)  $g$  is approximately continuous throughout  $[0, 1]$ . Thus,  $g$  is a bounded derivative function.

*Proof.* For any  $x$  there exists at most one  $n$  with  $g_n(x) \neq 0$  and  $g$  is well defined. Obviously  $g|_{[a_n, b_n]} = g_n|_{[a_n, b_n]}$ , hence it is continuous on  $[a_n, b_n]$ . Therefore  $C_g \supset U$ . For any  $x \notin U$ ,  $g_n(x) = 0$  for each  $n$ , this implies  $g(x) = 0$ ; since  $U$  is an everywhere dense open set, any point  $x \notin U$  is the limit point of some subsequence of  $\{c_n\}$ . Thus  $x \notin C_g$  and this proves  $C_g = U$ . Since  $g$  is continuous on  $U$ , it suffices to show that  $g$  is approximately continuous for  $x \notin U$ . 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 neighbourhood  $(x, x + h)$  meets infinitely many intervals  $I_n$ . Let  $n = N$  be the smallest index with  $I_n \cap (x, x + h) \neq \emptyset$ . Then

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

where the summation is over 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  $\lim_{y \rightarrow x+0} \text{app } g(y) = 0 = g(x)$ , because  $h \rightarrow 0$  implies  $N \rightarrow \infty$ . Similarly,  $g$  is approximately continuous from the left and the proof is complete.  $\square$

#### 4. THE MAIN CONSTRUCTION

Let  $P \subset [0, 1]$  be a Cantor set with  $\lambda(P) > 0$ . We will construct a Cantor-like differentiable function  $g$  which is constant on each interval in the complement of  $P$ .

*Step 1.* Let  $A := \{x : d(x, P) = 1\}$ . Then by Lemma 2,  $A \subset P$  and  $\lambda(P) = \lambda(A)$ . Thus  $\lambda(A) > 0$ . Using Lemma 3, select a set  $E$  of type  $F_\sigma$  such that  $E \subset A$  and  $\lambda(E) = \lambda(A)$ . Then  $d(x, E) = d(x, P) = 1$  for all  $x \in E$ .

*Step 2.* By Lemma 5, there exists an approximately continuous  $f$  such that

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

Define  $F : [0, 1] \rightarrow R$  by  $F(x) := \int_0^x f(t) dt$ . Then  $F'(x) = f(x)$  for all  $x \in [0, 1]$  by Lemma 4. In particular,  $F' = 0$  on each open component interval in the complement of  $P$ , so  $F$  is constant on each such interval. Let  $I$  be an open interval intersecting  $P$  with  $\lambda(P \cap I) > 0$ . Then  $\lambda(E \cap I) > 0$  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  $\lambda(P) > 0$  there must be such an interval and  $F$  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  $[0, 1]$ . From Lemma 7, we obtain  $g$  on  $[0, 1]$  such that:

- (i)  $g$  is approximately continuous on  $[0, 1]$ ;
- (ii)  $g$  is continuous on  $[a_n, b_n]$  for each  $n = 1, 2, \dots$  so that  $C_g = \tilde{P}$  and  $g(x) = 0$  for  $x \in P$ ;
- (iii) On  $(a_n, b_n) = (c_n - h_n, c_n + h_n)$  we have  $g(c_n) = 1$  while  $g$  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 is 0 otherwise.

Consider  $h := f + g$ . Then  $h$  is bounded. By Lemma 6 we see that  $h$  is approximately continuous. Hence  $h$  is a bounded derivative function on  $[0, 1]$  by Lemma 4. Define  $G : [0, 1] \rightarrow R$  by  $G(x) := \int_0^x g(t) dt$  and  $H : [0, 1] \rightarrow R$  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  $0 \leq f(x) \leq 1$  on  $P$ , we have from (iii) of Step 3 that

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

Since  $F$  is continuously differentiable on  $\tilde{P}$  we have  $\partial_c H(x) = \partial_c G(x) + F'(x)$ . However  $F'(x) = 0$  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.** *There are two Lipschitz and everywhere differentiable functions  $H : [0, 1] \rightarrow R$  and  $G : [0, 1] \rightarrow R$  differing by more than a constant such that*

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

*Thus  $G$  is differentiable but not integrable.*

Combining Theorem 2 and Theorem 1 we have:

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

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

*Remark 1.* (a) Clearly, by composing with a smooth periodic function, we may replace  $[0, 1]$  by  $R$  in these results. Similarly, after composition with an appropriate projection, on any Banach space there are two everywhere (Fréchet) differentiable functions which share the same A-subdifferentials throughout the space. (b) From the construction, each Cantor set with positive measure gives rise to 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 obtain uncountably many differentiable functions differing by more than constants that share the same C-subdifferentials. (c) This is not because the functions themselves are terribly pathological; such a pathological situation arises precisely from the upper semicontinuity which nonsmooth analysts impose through the C-subdifferential or A-subdifferential map.

*Remark 2.* It was shown in [1, page 80] that there are two real-valued continuous functions  $H$  and  $G$  defined on  $[0, 1]$  such that  $H'(x) = G'(x)$  for every  $x \in [0, 1]$ , but  $H - G$  is not constant. Of course,  $H' = G' = \infty$  on a nondenumerable set.

Let  $0 < \epsilon < 1$ . By Theorem 2.28 [8] we can make the Lebesgue measure of  $P$  equal to  $1 - \epsilon$ . From the construction we know

$$\lambda\{x : H'(x) \neq G'(x)\} = \lambda(E) = \lambda(P) = 1 - \epsilon.$$

By Proposition 1.3 [6] we obtain

$$\partial_{mp} H(x) = \{H'(x)\} \quad \text{and} \quad \partial_{mp} G(x) = \{G'(x)\}$$

for all  $x \in [0, 1]$ . Therefore:

**Corollary 2.** *For  $0 < \epsilon < 1$ , there are everywhere differentiable Lipschitz functions  $H, G : [0, 1] \rightarrow R$  differing by more than a constant with*

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

*while*

$$\lambda\{x \in [0, 1] : \partial_{mp} H(x) = \partial_{mp} G(x)\} = \epsilon.$$

Finally, note that since differentiable functions are generically continuously differentiable  $\{x \in [0, 1] : \partial_{mp} H(x) = \partial_{mp} G(x)\}$  is a residual set. In the language of classical real analysis, Theorem 2 becomes:

**Corollary 3.** *There are two bounded derivative functions  $g$  and  $h$  defined on  $[0, 1]$  differing by more than a constant such that  $g$  and  $h$  have the same cluster sets at every  $x \in [0, 1]$ .*

## ACKNOWLEDGMENTS

The authors wish to thank anonymous referees for helpful suggestions.

## REFERENCES

- [1] A. M. BRUCKNER, Differentiation of Real Functions, Lecture Notes in Mathematics edited by A. Dold and B. Eckmann, Springer-Verlag, 1978. MR **80h**:26002
- [2] A. M. BRUCKNER AND G. PETRUSKA, Some Typical Results on Bounded Baire-1 Functions, *Acta Math. Hung.* **43** (3-4) (1984), 325-333. MR **85h**:26004
- [3] F. H. CLARKE, Optimization and Nonsmooth Analysis, Wiley Interscience, New York, 1983. MR **85m**:49002
- [4] A. D. IOFFE, Approximate Subdifferentials and Applications 3: The Metric Theory, *Mathematika* **Vol. 36**, No. 71 (1989), 1-38. MR **90g**:49012
- [5] ———, Approximate Subdifferentials and Applications I: The Finite Dimensional Theory, *Transactions of The American Mathematical Society* **281** (1984), 390-416. MR **84m**:49029
- [6] P. MICHEL AND J. P. PENOT, Calcul Sous-différentiel Pour les Fonctions Lipschitzienne et non Lipschitzienne, *C. R. Acad. Sci. Paris* **298** (1984), 269-272. MR **85i**:49027
- [7] H. L. ROYDEN, Real Analysis, Macmillan Publishing Company, New York, 1988. MR **90g**:00004
- [8] K. R. STROMBERG, An Introduction to Classical Real Analysis, Wadsworth International Mathematics Series, 1981. MR **82c**:26002

DEPARTMENT OF MATHEMATICS AND STATISTICS, SIMON FRASER UNIVERSITY, BURNABY,  
BRITISH COLUMBIA, CANADA V5A 1S6

*E-mail address:* jborwein@cecm.sfu.ca

*E-mail address:* xwang@cecm.sfu.ca