REVISTA MATEMÁTICA COMPLUTENSE
Volumen 11, número 2: 1998
http://dx.doi.org/10.5209/rev_REMA.1998.v11.n2.17240

Some remarks on sub-differential calculus.

Gilles GODEFROY

Abstract

Mean value inequalities are shown for functions which are sub- or super- differentiable at every point.

Mean value inequalities are a classical tool for controlling the range of a real valued function through the behaviour of its derivative. The fundamental theorem of calculus leads to the consideration of the definite integral of the derivative. When the function is *everywhere* differentiable on $\mathbb{R}$ and the derivative is integrable, absolute continuity of the function follows (see [5], Th. 8.21.).

These classical theorems have recently been extended to functions defined on Banach spaces, and which are merely assumed to be sub-differentiable at every point. We refer e.g. to [2] and to D. Azagra's dissertation [1]. In this note, we will use the old-fashioned but powerful tools from [6] to establish slight refinements on this recent progress. Since our arguments rely heavily on [6], we included for the reader's convenience an Appendix with a self-contained proof of a special case of our basic lemma 1. This self-contained proof contains most of the relevant ideas.

I am grateful to D. Preiss who provided me with the fundamental reference [6], and to J. Saint Raymond for useful conversations.

Notation. A function f from a Banach space X to $\mathbb{R}$ is Gateaux-subdifferentiable at $x \in X$ if there exists $p \in X^*$ such that for all $h \in X \backslash \{0\}$,

Mathematics Subject Classification: 46B20.
Servicio Publicaciones Univ. Complutense. Madrid, 1998.

$$\liminf_{t\to 0}\frac{1}{|t|}[f(x+th)-f(x)-p(th)]\geq 0 \qquad (*)$$

When p satisfies $(*)$, we write $p\in D_G^- f(x)$. Similarly, f is Gateaux-superdifferentiable at $x\in X$ and $q\in D_G^+ f(x)$ if

$$\limsup_{t\to 0}\frac{1}{|t|}[f(x+th)-f(x)-q(th)]\leq 0$$

When the above "lim inf" and "lim sup" are uniform on $h\in X$ with $\| h \|=1$, we have Fréchet sub- or superdifferentiability, and we denote $p\in D_F^- f(x)$ or $q\in D_F^+ f(x)$. When $X=\mathbb{R}$, these two notions coincide and we simple write $d^- f(x)$ and $d^+ f(x)$.

If E is a measurable subset of $\mathbb{R}$, we denote by $m[E]$ its Lebesgue measure.

We now proceed to state and prove mean value inequalities, in which we will use the integral of sub- or superdifferentials rather than their supremum.

The following crucial lemma relies heavily on classical results from [6].

Lemma 1. *Let $g:\mathbb{R}\to\mathbb{R}$ be a Borel function such that $d^- g(t)\cup d^+ g(t)\neq\emptyset$ for every $t\in\mathbb{R}$. Define*

$$\varphi(t)=\inf\{|\ell|;\ell\in d^- g(t)\cup d^+ g(t)\}$$

Then φ is a Borel function, and for any $(a,b)\in\mathbb{R}^2$ with $a<b$, one has

$$m[g([a,b])]\leq\int_a^b\varphi(t)dt$$

Proof. One defines one-sided extreme derivatives, with values in $[-\infty,+\infty]$, as follows:

$$\underline{D}^+ g(t)=\lim\inf_{h\to 0+}\frac{1}{h}[g(t+h)-g(t)]$$

and

$$\overline{D^+} g(t)=\lim\sup_{h\to 0+}\frac{1}{h}[g(t+h)-g(t)]$$

These quantities are called the right lower and upper Dini derivatives of g at t. The left Dini derivatives $\underline{D}^- g(t)$ and $\overline{D^-} g(t)$ are defined by substituting 0^- to 0^+. It is easily checked that

$$d^- g(t) = [\overline{D^-} g(t), \underline{D}^+ g(t)] \tag{1}$$

and

$$d^+ g(t) = [\overline{D^+} g(t), \underline{D}^- g(t)] \tag{2}$$

We claim that the Dini derivatives are Borel functions. Indeed, let (x_n) be a sequence in $\mathbb{R}$ such that the sequence $((x_n, g(x_n)))$ is dense in the graph of g. Let u be a given real number. For $(n, k, l) \in \mathbb{N}^3$, we set

$$B_{n,k,l} = \{t \in \mathbb{R}; t \notin (x_n - l^{-1}, x_n) \text{ or } g(x_n) - g(t) \geq (u + k^{-1})(x_n - t)\}$$

It is easy to check that

$$\underline{D}^+ g(t) \geq u \Leftrightarrow t \in \bigcap_{k \geq 1} \bigcup_{l \geq 1} \bigcap_{n \geq 1} B_{n,k,l}$$

it follows from the above equivalence that $\underline{D}^+ g$ is a Borel function. Identical arguments apply to the other Dini derivatives, and its easily follows that the function φ is Borel.

For any $n \geq 1$, we consider the Borel set

$$E_n = \{t \in \mathbb{R}; d^+ g(t) \cap (-n, n) \neq \emptyset\}$$

For all $t \in E_n$, one has by (2)

$$\overline{D^+} g(t) < n; \underline{D}^- g(t) > -n$$

By ([6], lemma VII.6.3), it follows that

$$m[g(E_n)] \leq n \cdot m(E_n) \tag{3}$$

Since we can apply the same argument to $d^- g$, it follows from our assumptions and (3) that for E measurable set,

$$m(E) = 0 \Rightarrow m[g(E)] = 0 \tag{4}$$

Consider now the Borel set

$$G = \{t \in \mathbb{R}; d^- g(t) \neq \emptyset\}$$

From the relation

$$\underline{D}^- g(t) \le \overline{D^-} g(t) \le inf(d^- g(t))$$

it follows that $G = G_1 \cup G_2$, with

$$G_1 = \{t \in G; \overline{D^-} g(t) \in \mathbb{R}\}$$

and

$$G_2 = \{t \in G; \underline{D}^- g(t) = \overline{D^-} g(t) = -\infty\}$$

By ([6], theorem IX.4.1), there exists a subset H_1 of G_1 with $m(H_1) = 0$, such that for all $t \in G_1 \backslash H_1$,

$$\overline{D^-} g(t) = \underline{D}^+ g(t) = \{d^- g(t)\}$$

By the remark in ([6], p.272) for every measurable subset E of $(G_1 \backslash H_1)$, one has

$$m[g(E)] \le \int_E | d^- g(t) | \, dt \tag{5}$$

and by (4) above, $m[g(H_1)] = 0$. Moreover, by ([6], theorem IX.4.4.), one has $m(G_2) = 0$, hence $m[g(G_2)] = 0$. It follows now from (5) that

$$m[g(E)] \le \int_E | d^- g(t) | \, dt \tag{6}$$

for any measurable subset E of G where $d^- g(t)$ can be given arbitrary values when $t \in H_1$ where it is not single-valued.

It is clear that we can apply the same argument to $d^+ g$. It suffices now to write $\mathbb{R}$ as the disjoint union

$$\mathbb{R} = J \cup K$$

where the Borel sets J and K are defined by

$$J = \{t \in \mathbb{R}; \varphi(t) = \inf\{| \ell |; \ell \in d^- g(t)\}\}$$

and

$$K = \{t \in \mathbb{R}; \varphi(t) < \inf\{| \ell |; \ell \in d^- g(t)\}\}$$

to obtain the conclusion by applying (6) on J and the analogue $d^+ g$ inequality on K. ∎

If $h = [a,b] \to I\!R$ is an arbitrarily function, we denote by

$$_{*}\int_a^b h(t)dt$$

its "lower integral" on $[a,b]$, defined by

$$_{*}\int_a^b h(t)dt = \sup\left\{\int_a^b g(t)dt; g \text{ measurable } g \leq h\right\}$$

Our main result is an improvement of ([1], theorem 3.22).

Theorem 2. *Let X be a Banach space, and let $f = X \to I\!R$ be a Borel function such that*

$$D_G^- f(x) \cup D_G^+ f(x) \neq \emptyset$$

for every $x \in X$. Define $\Phi = X \to I\!R$ by

$$\Phi(x) = \inf\{\| p \|; p \in D_G^- f(x) \cup D_G^+ f(x)\}$$

Then for every $(x,y) \in X^2$, one has

$$m[f([x,y])] \leq \| y - x \| \cdot {}_{*}\int_0^1 \Phi(ty + (1-t)x)dt$$

Proof. We pick $(x,y) \in X^2$ with $x \neq y$, and define $g = [0,1] \to I\!R$ by

$$g(t) = f(ty + (1-t)x)$$

Clearly, g is Borel since f is. Moreover, it is easy to check that if

$$p \in D_G^- f(ty + (1-t)x)$$

then

$$p(y-x) \in d^- g(t)$$

Hence, in the notation of lemma 1 and theorem 2, we have for all $t \in [0,1]$

$$\varphi(t) \leq \Phi(ty + (1-t)x \cdot \| y - x \|$$

Since φ is Borel, we have by definition of the lower integral that

$$\int_0^1 \varphi(t)dt \leq \| y - x \| \cdot {}_{*}\int_0^1 \Phi(ty + (1-t)x)dt$$

Moreover, lemma 1 states that

$$m[f([x,y])] = m[g([0,1])] \leq \int_0^1 \varphi(t)dt$$

and this concludes the proof.

■

Let us check what we obtain when we dispense with the continuity assumption in ([1], Th. 3.22) and apply Lemma 1.

Proposition 3. *Let X be a Banach space, and U an open convex subset of X. Let $f : U \to \mathbb{R}$ be a function. Assume that there exists $M \geq 0$ such that for every $x \in U$, there exists $p \in D_G^-(f)(x)$ such that $\| p \| \leq M$. Then for all $(x,y) \in U^2$, one has*

$$m[f([x,y])] \leq M \| x - y \|$$

Proof. Let us call $g = [0,1] \to \mathbb{R}$ the function defined by

$$g(\lambda) = f(\lambda y + (1-\lambda)x)$$

For every $\lambda \in [0,1]$ and every $p \in D_G^- f(\lambda y + (1-\lambda)x)$, one has

$$p(y-x) \in d^- g(x) \tag{7}$$

hence $d^- g(\lambda) \neq \emptyset$ for every $\lambda \in [0,1]$. It follows that g is lower semi-continuous, hence in particular it is a Borel function. Moreover, in the notation of lemma 1, we have by (7) that

$$\varphi(\lambda) \leq M \| x - y \|$$

for every $\lambda \in [0,1]$ and then by lemma 1,

$$m[g([0,1])] = m[f([x,y])] \leq M \| x - y \|$$

■

Remark 4. We recall that a function $f : \mathbb{R} \to \mathbb{R}$ has the Darboux property if $f([a,b])$ contains $[f(a), f(b)] \cup [f(b), f(a)]$ for all $a < b$ (in other words, if f satisfies the intermediate value theorem). By lemma 1, a Borel function g with the Darboux property such that

$$(d^- g(t) \cup d^+ g(t)) \cap [-M, M] \neq \emptyset$$

for every $t \in \mathbb{R}$ is M–Lipchitzian. Note that any derivative is a Borel function with the Darboux property.

Of course, continuous functions have the Darboux property, hence by theorem 2 we can weaken the assumption in ([1], theorem 3.22) and replace "$p \in D_G^- f(x)$" by

$$\text{“}p \in D_G^- f(x) \cup D_G^+ f(x)\text{”}.$$

Remark 5. It is clear that $0 \in d^- g(t)$ for all $t \in \mathbb{R}$ does not imply that g is continuous (take the characteristic function of any open subset of $\mathbb{R}$). However, by [2], if X is a Banach space on which there exists a Lipschitz Fréchet-differentiable bump function and $g : X \to \mathbb{R}$ is such that $D_F^- g(x) = \{0\}$ for all $x \in X$, then g is constant. This assumption on X is actually needed: if X has an equivalent rough norm $\| \cdot \|$ (see [3], chapter III) then the characteristic function h of $\{x; \| x \| < 1\}$ satisfies $D^- F h(x) = \{0\}$ for all $x \in X$. A concrete example is provided by the natural norm of $\ell_1(\mathbb{N})$.

Remark 6. By Proposition 4, if $f : [0,1] \to \mathbb{R}$ is such that $0 \in d^- f(t)$ for all $t \in [0,1]$, then $m[f([0,1])] = 0$. However, there is a function $g = [0,1]^2 \to \mathbb{R}$ such that $(0,0) \in d^- g(v)$ for every $v \in [0,1]^2$ but $g([0,1]^2) = [0,1]$. Indeed, there exists ([7]; see [4]) a continuously differentiable function $w = [0,1]^2 \to \mathbb{R}$ with $w([0,1]^2) = [0,1]$ and a continuous function $\gamma = [0,1] \to [0,1]^2$ such that $w'[\gamma(t)] = 0$ for every $t \in [0,1]$ but $(w \circ \gamma)[0,1] = [0,1]$. The function $g = [0,1]^2 \to [0,1]$ defined by:

(i) $g(v) = w$ if there is $t \in [0,1]$ such that $v = \gamma(t)$.

(ii) $g(v) = 1$ otherwise.

satisfies the required conditions.

Appendix. For the reader's convenience we include a self-contained proof of the following statement: let $f : [0,1] \to \mathbb{R}$ be such that

$d^-f(x) \neq \emptyset$ for every x, and let $\alpha = [0,1] \to [0,+\infty)$ be a Lebesgue measurable function such that

$$[-\alpha(x), \alpha(x)] \cap d^-f(x) \neq \emptyset$$

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

$$m[f([0,1])] \leq \int_0^1 \alpha(t)dt.$$

Proof. For any $A \subset [0,1]$, let

$$d(A) = \sup\{\mid t-u \mid; (t,u) \in A^2\}$$

Fact A: For every $x \in [0,1]$ and $\delta > 0$, there is $\varepsilon > 0$ such that if $x \in A$ and $d(A) < \varepsilon$, then

$$\inf_{y \in A} f(y) \geq f(x) - (\alpha(x) + \delta)d(A).$$

Proof. Pick $t \in d^-f(x)$ with $\mid t \mid \leq \alpha(x)$. There is $\varepsilon > 0$ such that $\mid y - x \mid < \varepsilon$ implies

$$f(y) \geq f(x) + t(y-x) - \delta \mid y - x \mid$$

hence

$$\begin{aligned} f(y) &\geq f(x) - (\mid t \mid + \delta) \mid y - x \mid \\ &\geq f(x) - (\alpha(x) + \delta)d(A) \end{aligned}$$

Fact B: For any $\gamma > 0$, there is a sequence (A_n) of measurable subsets of $[0,1]$ such that

(i) $\cup A_n = [0,1]$

(ii) $\sup_n d(A_n) \leq \gamma$

(iii) $\Sigma d(A_n) < 1 + \gamma$

(iv) $\Sigma d(A_n) \sup_{A_n}(\alpha) \leq \int_0^1 \alpha(t)dt + \gamma$

Proof. For any $k \in \mathbb{N}$ let

$$B_k = \alpha^{-1}([k\gamma, (k+1)\gamma))$$

Since the $B_k's$ are measurable, there are open sets (O_k) with $B_k \subset O_k$ and

$$\Sigma m(O_k) < 1 + \gamma$$

We split the $O_k's$ into disjoint intervals $(I_{k,j})$ of length lees than γ, and we define $A_{k,j} = B_k \cap I_{k,j}$. After reindexing, we write $(A_{k,j}) = (A_n)$. It is easily seen that this sequence (A_n) works.

We now conclude the proof. Pick $\delta > 0$. For all ≥ 1, let $\left(A_n^k\right)_{n\geq 1}$ be a sequence satisfying the conditions of Fact B with $\gamma = k^{-1}$. We set

$$E_k = \bigcup_{n\geq 1} \left[\inf_{A_n^k}(f), \inf_{A_n^k}(f) + (\sup_{A_n^k}(\alpha) + \delta) d\left(A_n^k\right)\right]$$

We have

$$\sup_n d\left(A_n^k\right) \leq k^{-1}$$

Hence by Fact A, for all $x \in [0, 1]$, there exists $K(x) \geq 1$ such that $f(x) \in E_k$ for every $k \geq K(x)$. Hence we have

$$f([0,1]) \subseteq \bigcup_{K\geq 1} \left[\bigcap_{k\geq K} E_k\right]$$

We set

$$F_K = \bigcap_{k\geq K} E_k$$

By conditions (iii) and (iv) of Fact B, we have

$$\begin{aligned} m(E_k) &\leq \Sigma_n d\left(A_n^k\right) \sup_{A_n^k}(\alpha) + \delta \Sigma_n d\left(A_n^k\right) \\ &\leq \int_0^1 \alpha(t)dt + k^{-1} + \delta(1 + k^{-1}) \end{aligned}$$

Therefore

$$m(F_K) \leq \int_0^1 \alpha(t)dt + \delta$$

and since (F_K) is an increasing sequence it follows that

$$m[f([0,1])] \leq m\left[\bigcup_{K\geq 1} F_K\right] \leq \int_0^1 \alpha(t)dt + \delta$$

and the conclusion follows since $\delta > 0$ is arbitrary.

■

It is not difficult to adjust the above proof in order to show that if $f : \mathbb{R} \to \mathbb{R}$ is a function such that

$$O \in d^- f(x) \cup d^+ f(x)$$

for every $x \in \mathbb{R}$, then

$$m[f(\mathbb{R})] = 0$$

This provides a "mean value inequality" for functions which are not necessarily measurable, since e.g. the characteristic function of any set satisfies the assumptions.

References

[1] Azagra D., *Smooth negligibility and subdifferential calculus in Banach spaces, with applications*, Doctoral Dissertation, Univ. Complutense, Madrid (1997).

[2] Deville R., A mean value theorem for the non differentiable mappings, *Serdica Math. J.* 21 (1995), 59-66.

[3] Deville R., Godefroy G., Zizler V., *Smoothness and renormings in Banach spaces*, Pitman Monographs and Surveys 64, Longman Ed. (1993).

[4] Körner T.W., A dense arcwise connected set of critical points: molehills out of mountains, *J. London Math. Soc.* 2, 38, (1988), 3, 442-452.

[5] Rudin W., *Real and complex analysis, Second edition*, Mac Graw Hill series in higher mathematics (1974).

[6] Saks S., *Theory of the integral, Monographie Matematyczne, Tom VII, Second revised edition,* Hafner Publ. Comp., New York (1937).

[7] Whitney H., A function not constant on a connected set of critical points, *Duke Math. J.* 1 (1935), 514-517.

Equipe d'Analyse
Université Paris VI
Boîte 186
4, place Jussieu
75252-Paris Cedex 05
France
gig@ccr.jussieu.fr

Recibido: 1 de Abril de 1998
Revisado: 15 de Junio de 1998