

Approximate Rolle's theorems for the proximal subgradient and the generalized gradient

Daniel Azagra,¹ Juan Ferrera,² and Fernando López-Mesas^{*,3}

*Departamento de Análisis Matemático, Facultad de Matemáticas, Universidad Complutense,
28040 Madrid, Spain*

Received 1 April 2002

Submitted by K. Jarosz

Abstract

We establish approximate Rolle's theorems for the proximal subgradient and for the generalized gradient. We also show that an exact Rolle's theorem for the generalized gradient is completely false in all infinite-dimensional Banach spaces (even when they do not possess smooth bump functions). © 2003 Elsevier Inc. All rights reserved.

Keywords: Rolle's theorem; Proximal subgradient; Generalized gradient

1. Introduction

Rolle's theorem in finite-dimensional spaces states that, for every bounded open subset U of \mathbb{R}^n and for every continuous function $f: \bar{U} \rightarrow \mathbb{R}$ such that f is differentiable in U and constant on the boundary ∂U , there exists a point $x \in U$ at which the differential of f vanishes. Rolle's theorem does not remain true in infinite-dimensional Banach spaces. It was Shkarin [12] that first showed that this theorem fails for infinite-dimensional super-reflexive spaces and for nonreflexive spaces with equivalent Fréchet differentiable norms.

* Corresponding author.

E-mail addresses: daniel_azagra@mat.ucm.es (D. Azagra), ferrera@mat.ucm.es (J. Ferrera), flopez_mesa@mat.ucm.es (F. López-Mesas).

¹ D. Azagra was supported by a Marie Curie Fellowship of the European Community Training and Mobility of Researchers Programme under Contract HPMF-CT-2001-01175.

² J. Ferrera was partially supported by BFM Grant 2000/0609.

³ F. López-Mesas enjoyed a study licence granted by the Consejería de Educación de la Comunidad de Madrid.

Other explicit counterexamples were constructed for c_0 and ℓ_2 by Bès and Ferrera [5], and independently by Ferrer [10]. The class of spaces for which Rolle's theorem fails was enlarged in [1], where it is shown that Rolle's theorem fails in all infinite-dimensional Banach spaces which have smooth norms. On the other hand, Rolle's theorem is trivially true in all Banach spaces which do not admit smooth bump functions. Therefore, in many cases, Rolle's theorem is either trivially true or (nontrivially) false. In this setting it has been recently proved [4] that in fact this is what happens in all infinite-dimensional Banach spaces, that is, a Banach space X has a C^p smooth (and Lipschitz) bump function if and only if there exists a bounded open (contractible) subset U of X and a C^p smooth (and Lipschitz) function $f: \bar{X} \rightarrow \mathbb{R}$ such that $f = 0$ on $X \setminus U$ and yet $f'(x) \neq 0$ for all $x \in U$ (that is, Rolle's theorem fails in X); here $p \in \mathbb{N} \cup \{\infty\}$. Despite the failure of Rolle's theorem in infinite dimensions, the following approximate version of the result remains true in all Banach spaces, as it was proved in [3].

Theorem 1.1. *Let U be a bounded connected open subset of a Banach space X . Let $f: \bar{U} \rightarrow \mathbb{R}$ be a continuous bounded function which is Gâteaux differentiable on U . Let $R > 0$ and $x_0 \in U$ be such that $\text{dist}(x_0, \partial U) = R$, and suppose that $f(\partial U) \subseteq [-\varepsilon, \varepsilon]$ for some $\varepsilon > 0$. Then there exists some $x_\varepsilon \in U$ so that $\|f'(x_\varepsilon)\| \leq \varepsilon/R$.*

Natural extensions of this result are worth exploring within the various theories of subdifferentiability. In [1,2], a version of this result for Fréchet and Gâteaux subdifferentials was proved (together with a subdifferential mean value inequality theorem which was later improved by Godefroy [11], see also [7]), for the class of Banach spaces which possess (Fréchet or Gâteaux) smooth Lipschitz bump functions. In particular it was shown that for every Banach space X with a Fréchet smooth and Lipschitz bump, every continuous bounded function $f: \bar{B}(0, 1) \rightarrow \mathbb{R}$ which oscillates between $-\varepsilon$ and ε on the unit sphere $S(0, 1)$ must satisfy that $\inf\{\|p\|: p \in D^- f(x) \cup D^+ f(x), \|x\| < 1\} \leq 2\varepsilon$. Here $D^- f(x)$ and $D^+ f(x)$ stand for the sets of Fréchet subdifferentials and superdifferentials, respectively, at a point x , and $\bar{B}(0, 1)$ is the unit ball of the space X . In this paper we will establish similar results for other important kinds of subdifferentials. In Section 2 we obtain an approximate Rolle's theorem for the proximal subgradient in real Hilbert spaces. In Section 3 we first prove that an exact Rolle's theorem for the generalized gradient is false in all infinite-dimensional real Banach spaces, even for spaces which do not possess any smooth bump functions. More specifically, we show that if X is an infinite-dimensional real Banach space, there are Lipschitz functions defined on the closure of a bounded connected open set U which vanish on the boundary ∂U and yet, for all $x \in U$ and all functionals p in the generalized gradient $\partial f(x)$ of the function f at x , we have that $p \neq 0$. That is, Rolle's theorem also fails when the differentiability assumptions on both the space and the function are weakened and replaced by mere Lipschitzness of the function, and all the generalized gradients are considered. Notice that, since the generalized gradient contains all the known subdifferentials and superdifferentials, this is close to be the most radical form of failure that an exact Rolle's theorem for subdifferentials may suffer. It is thus necessary to consider alternative approximate results: in the last part of the paper we deal with an approximate version of Rolle's theorem for the generalized gradient, which we show to be true in all real Banach spaces. To finish this introduction let us quote one of the versions of

Ekeland's variational principle, which is an important ingredient in many of the proofs of this paper. A proof can be found in [9], for instance.

Theorem 1.2 (Ekeland's variational principle). *Let X be a Banach space, and let $f : X \rightarrow [-\infty, +\infty)$ be a proper, upper semicontinuous function which is bounded above. Let $\varepsilon > 0$ and $x_0 \in X$ be such that $f(x_0) > \sup\{f(x) : x \in X\} - \varepsilon$. Then, for each λ with $0 < \lambda < 1$, there exists a point $x_1 \in \text{Dom}(f)$ such that*

- (i) $\lambda\|x_1 - x_0\| \leq f(x_1) - f(x_0)$;
- (ii) $\|x_1 - x_0\| < \varepsilon/\lambda$;
- (iii) $\lambda\|x_1 - x\| + f(x_1) > f(x)$ whenever $x \neq x_1$.

Throughout the paper, $B(x, r)$ and $S(x, r)$ stand for the open ball and the sphere of center x and radius r , with respect to the norm under consideration, while $\overline{B}(x, r)$ is the closed ball of center x and radius r .

2. An approximate Rolle's theorem for the proximal subgradient

Definition 2.1. Let X be a real Hilbert space. A vector $\zeta \in X$ is called a *proximal subgradient* of a lower semicontinuous function f at $x \in \text{Dom}(f)$ provided there exist positive numbers σ and η such that

$$f(y) \geq f(x) + \langle \zeta, y - x \rangle - \sigma\|y - x\|^2 \quad \text{for all } y \in B(x, \eta).$$

The set of all such ζ is denoted $\partial_p f(x)$, and is referred to as the *proximal subgradient*, or P-subdifferential. In a similar way, we may introduce the *proximal supergradient*. For an upper semicontinuous function f , we say that $\zeta \in X$ is a *proximal supergradient* of f at $x \in \text{Dom}(f)$ provided there exist positive numbers σ and η such that

$$f(y) \leq f(x) + \langle \zeta, y - x \rangle + \sigma\|y - x\|^2 \quad \text{for all } y \in B(x, \eta).$$

We will denote the set of all such ζ by $\partial^p f(x)$.

In the proof of one of the main results of this section we will use the second-order smooth variational principle of Deville et al. The following theorem is a weak restatement of this variational principle in the case when X is the Hilbert space. For the general statement and a proof, see [8]. The following notation is used: $\|\varphi\|_\infty = \sup\{|\varphi(x)| : x \in X\}$, $\|\varphi'\|_\infty = \sup\{\|\varphi'(x)\| : x \in X\}$.

Theorem 2.2. *Let X be a real Hilbert space, $F : X \rightarrow (-\infty, \infty]$ be a proper, lower semicontinuous function which is bounded below. Then, for every $\delta > 0$ there exist a C^2 smooth function φ with bounded derivatives, and a point $x_0 \in X$ such that*

- (1) $F - \varphi$ attains its minimum on X at the point x_0 ;
- (2) $\|\varphi\|_\infty < \delta$ and $\|\varphi'\|_\infty < \delta$.

Finally, we will also need the following fact.

Lemma 2.3. *Let $f : X \rightarrow (-\infty, \infty]$ be a proper, lower semicontinuous function. If $f - \varphi$ attains a minimum at a point x_0 and φ is twice differentiable at x_0 , then $\varphi'(x_0)$ is a proximal subgradient of f at x_0 , that is $\varphi'(x_0) \in \partial_p f(x_0)$. Similarly, if $g : X \rightarrow [-\infty, \infty)$ is a proper upper semicontinuous function, ψ is a function which is twice differentiable at x_0 , and $g + \psi$ attains a maximum at x_0 , then $\psi'(x_0) \in \partial^p g(x_0)$.*

Proof. We know that

$$f(y) - f(x_0) \geq \varphi(y) - \varphi(x_0) \tag{1}$$

for all y . Since φ is twice differentiable at x_0 , for a given $\varepsilon > 0$ we can find $\delta > 0$ such that

$$|\varphi(y) - \varphi(x_0) - \varphi'(x_0)(y - x_0) - \frac{1}{2}\varphi''(x_0)(y - x_0)^2| \leq \varepsilon \|y - x_0\|^2$$

whenever $\|y - x_0\| \leq \delta$. In particular,

$$\varphi(y) - \varphi(x_0) \geq \varphi'(x_0)(y - x_0) + \frac{1}{2}\varphi''(x_0)(y - x_0)^2 - \varepsilon \|y - x_0\|^2$$

for $\|y - x_0\| \leq \delta$, and therefore

$$\varphi(y) - \varphi(x_0) \geq \varphi'(x_0)(y - x_0) - (\|\varphi''(x_0)\| + \varepsilon) \|y - x_0\|^2 \tag{2}$$

whenever $\|y - x_0\| \leq \delta$. By combining (1) and (2) we get that

$$f(y) - f(x_0) \geq \langle p, y - x_0 \rangle - \sigma \|y - x_0\|^2$$

for all $y \in B(x_0, \delta)$, where $\sigma = (\|\varphi''(x_0)\| + \varepsilon)$ and $p = \varphi'(x_0)$, and this means that $p \in \partial_p f(x_0)$. \square

Taking into account this lemma and the very definition of $\partial_p f(x)$ and $\partial^p f(x)$, we can immediately deduce the following

Corollary 2.4. *Let $f : X \rightarrow (-\infty, \infty]$ be a proper, lower semicontinuous function. Then*

$$\partial_p f(x) = \{\varphi'(x) : \varphi \in C^2(X, \mathbb{R}), f - \varphi \text{ attains a local minimum at } x\}.$$

Similarly, if $g : X \rightarrow [-\infty, \infty)$ is a proper upper semicontinuous function, then

$$\partial^p g(x) = \{\varphi'(x) : \varphi \in C^2(X, \mathbb{R}), g + \varphi \text{ attains a local maximum at } x\}.$$

This corollary suggests a natural extension of the definition of proximal subgradients for Banach spaces which are not Hilbertian but do have C^2 smooth norms. For such spaces, defining $\partial_p f$ and $\partial^p f$ as in the corollary (or equivalently through the subdifferential proximal inequality), all the results that we present in this section remain true. Let us now prove some approximate versions of Rolle’s theorem for proximal subgradients and supergradients.

Theorem 2.5. *Let X be a real Hilbert space, $B = \overline{B}(0, R)$, $S = S(0, R)$, and $f : B \rightarrow \mathbb{R}$ be a bounded continuous function such that $f(S) \subseteq [-\varepsilon, \varepsilon]$ for some $\varepsilon > 0$.*

- (1) If $\inf f(B) < \inf f(S)$ then, for every $\alpha > 0$ there exist $x_0 \in \text{int}(B)$ and $\zeta \in \partial_p f(x_0)$ such that $\|\zeta\| < \alpha$.
- (2) If $\sup f(B) > \sup f(S)$ then, for every $\alpha > 0$ there exist $x_0 \in \text{int}(B)$ and $\zeta \in \partial^p f(x_0)$ such that $\|\zeta\| < \alpha$.

Otherwise,

- (3) If $f(B) \subseteq [-\varepsilon, \varepsilon]$ then, for every $\alpha > 0$ there exist $x_1, x_2 \in \text{int}(B)$ and $\zeta_1 \in \partial_p f(x_1)$, $\zeta_2 \in \partial^p f(x_2)$ such that $\|\zeta_1\|, \|\zeta_2\| < 2\varepsilon/R + \alpha$.

Proof. (1) Let $\eta = \inf f(S) - \inf f(B) > 0$. Consider the function F defined as $F(x) = f(x)$ if $x \in B$, $F(x) = +\infty$ otherwise; this function is obviously lower semicontinuous and bounded below. Then, the Deville–Godefroy–Zizler variational principle (Theorem 2.2) provides us with a C^2 smooth function g such that $\|g\|_\infty < \eta/3$, $\|g'\|_\infty < \alpha$, and $F - g$ attains its minimum at a point $x_0 \in B$. We claim that $x_0 \in \text{int}(B)$. Indeed, if $x_0 \in S$ then we could take $a \in B$ such that $f(a) < \inf f(B) + \eta/3$, and then we would get

$$\inf f(B) + 2\eta/3 > f(a) - g(a) \geq F(x_0) - g(x_0) \geq \inf f(S) - \eta/3,$$

that is, $\inf f(B) + \eta > \inf f(S)$, a contradiction. Since $f - g$ attains its minimum at x_0 , Lemma 2.3 ensures that $\zeta := g'(x_0) \in \partial_p f(x_0)$. On the other hand, $\|\zeta\| < \alpha$ because $\|g'\|_\infty < \alpha$.

(2) It suffices to apply (1) to the function $-f$.

(3) Take $\beta > 0$ small enough so that $\beta/2 + \beta/R < \alpha$ and $\beta < R$, and then choose $N > 1$ large enough so that

$$\frac{2\varepsilon + \beta}{R} \frac{\beta}{N} < \beta.$$

Let $a: \mathbb{R} \rightarrow \mathbb{R}$ be a C^∞ smooth convex function such that

- (i) $a(t) = t$ if $t \geq \beta/N$;
(ii) $a(t) = a(0) > 0$ if $t \leq \beta/4N$;
(iii) $a'(t) > 0$ if $t > \beta/4N$;
(iv) $a''(t) > 0$ if and only if $t \in (\beta/4N, \beta/N)$.

Such a function a can easily be constructed by integrating twice a C^∞ smooth nonnegative real function b whose support is the interval $[\beta/4N, \beta/N]$ and is such that $\int_{-\infty}^{\infty} b(t) dt = 1$, and then adding a suitable positive constant to obtain the properties that $a(0) > 0$ and $a(t) = t$ for $t \geq \beta/N$. Define then the function $h: X \rightarrow (0, \infty)$ by $h(x) = a(\|x\|)$. It is clear that h is C^∞ smooth, $h(0) = a(0) \in (0, \beta/N)$, and its derivative satisfies $h'(x) = 0$ for $\|x\| \leq \beta/4N$, and $h'(x) = a'(\|x\|)x/\|x\|$ for $\|x\| \geq \beta/4N$. In particular we see that

$$\|h'(x)\| \leq a'(\|x\|) \leq 1 \quad \text{for all } x \in X \quad \text{and} \quad h(x) = \|x\| \quad \text{if } \|x\| \geq \beta/N.$$

Let us consider the function $G: B \rightarrow \mathbb{R}$ defined by

$$G(x) = f(x) + \frac{2\varepsilon + \beta}{R} h(x).$$

The function G is continuous and G satisfies that $\inf G(B) < \inf G(S)$, as is easily checked. Then, by applying case (1) to G we obtain the required point x_1 . The point x_2 can be obtained by replacing f with $-f$. \square

From this result we can immediately deduce the following

Theorem 2.6. *Let U be a bounded connected open subset of a real Hilbert space X , and let $f : \bar{U} \rightarrow \mathbb{R}$ be a bounded continuous function. Let $R > 0$ and $x_0 \in U$ be such that $\text{dist}(x_0, \partial U) = R$. Suppose that $f(\partial U) \subset [-\varepsilon, \varepsilon]$ for some $\varepsilon > 0$.*

- (1) *If $\inf f(U) < \inf f(\partial U)$ then, for every $\alpha > 0$ there exist $x_1 \in U$ and $\zeta \in \partial_p f(x_1)$ such that $\|\zeta\| < \alpha$.*
- (2) *If $\sup f(U) > \sup f(\partial U)$ then, for every $\alpha > 0$ there exist $x_2 \in U$ and $\zeta \in \partial^p f(x_2)$ such that $\|\zeta\| < \alpha$.*
- (3) *If $f(U) \subseteq [-\varepsilon, \varepsilon]$ then, for every $\alpha > 0$ there exist $x_1, x_2 \in U$ and $\zeta_1 \in \partial_p f(x_1)$, $\zeta_2 \in \partial^p f(x_2)$ such that $\|\zeta_1\|, \|\zeta_2\| < 2\varepsilon/R + \alpha$.*

In any case, $\inf\{\|\zeta\| : \zeta \in \partial_p f(x) \cup \partial^p f(x), x \in U\} \leq 2\varepsilon/R$.

Remark 2.7. The infimum considered in Theorem 2.6 can well be strictly positive, as the following example shows: $f(x) = \varepsilon x$, defined on $\bar{U} = [-1, 1] \subset \mathbb{R}$. In this case, $\{f'(x)\} = \{\varepsilon\} = \partial_p f(x) = \partial^p f(x)$ for all $x \in U$.

If, in the conditions of the preceding theorem, we additionally assume that $\partial_p f(x) \neq \emptyset$ at every $x \in U$, then we can guarantee that $\inf\{\|\zeta\| : \zeta \in \partial_p f(x), x \in U\} \leq 2\varepsilon/R$. Indeed, it is immediately seen that, if for some point x we have $\partial_p f(x) \neq \emptyset \neq \partial^p f(x)$, then the function f is differentiable at x , and

$$\partial_p f(x) = \partial^p f(x) = \{f'(x)\}.$$

Remark 2.8. These results cannot be improved to get a point such that the norm of every proximal subgradient at this point is smaller than $2\varepsilon/R + \alpha$, as the following example shows: $f : [-1, 1] \rightarrow \mathbb{R}$, $f(x) = |x|$.

If we wish to guarantee that there exists a point such that all the proximal subgradients at this point have norm smaller than or equal to $2\varepsilon/R$, we have to be under conditions (2) or (3) of Theorem 2.6 (under condition (1) this additional demand is impossible to meet, as the above example shows). Next we give some results in this direction.

Lemma 2.9. *Let X be a real Hilbert space, $x_1 \in X$, and $f : X \rightarrow \mathbb{R}$ a lower semicontinuous function. Suppose that for some $\lambda > 0$ we have that $\lambda\|x_1 - x\| + f(x_1) > f(x)$ whenever $x \neq x_1$. Then $\|\zeta\| < \lambda$ for all $\zeta \in \partial_p f(x_1)$.*

Proof. Indeed, for all h with $\|h\| = 1$, setting $x = x_1 + th$ we have that

$$\frac{f(x_1 + th) - f(x_1)}{|t|} < \lambda.$$

On the other hand, for any $\zeta \in \partial_p f(x_1)$, there exist $\eta > 0$ and $\sigma > 0$ so that if $|t| < \eta$ then

$$f(x_1 + th) \geq f(x_1) + \langle \zeta, th \rangle - \sigma \|th\|^2,$$

and therefore

$$\langle \zeta, th \rangle \leq f(x_1 + th) - f(x_1) + \sigma t^2,$$

that is

$$\frac{t}{|t|} \langle \zeta, h \rangle \leq \frac{f(x_1 + th) - f(x_1)}{|t|} + \sigma |t|,$$

and in this manner we get

$$|\langle \zeta, h \rangle| \leq \lambda + \sigma |t| \quad \text{for all } |t| < \eta,$$

which implies that $\|\zeta\| \leq \lambda$. \square

Proposition 2.10. *Let $f: \overline{B}(0, R) \rightarrow \mathbb{R}$ be a continuous bounded function. Assume that $f(\overline{B}(0, R)) \subset f(S(0, R)) \subset [-\varepsilon, +\varepsilon]$. Then there exists $x \in B(0, R)$ such that $\|\zeta\| \leq 2\varepsilon/R$ for all $\zeta \in \partial_p f(x)$.*

Proof. Assume first that $2\varepsilon < R$.

Case I. Suppose $f(0) > -\varepsilon$, and let $\lambda = 2\varepsilon/R$. Since $\sup\{f(x) \mid x \in \overline{B}(0, R)\} - 2\varepsilon < f(0)$, we can apply Ekeland's variational principle to the function $F: X \rightarrow [-\infty, +\infty)$ defined by $F(x) = f(x)$ if $x \in \overline{B}(0, R)$ and $F(x) = -\infty$ elsewhere (which is clearly upper semicontinuous), to get some $x_1 \in \overline{B}(0, R)$ such that

- (i) $\lambda \|x_1\| \leq f(x_1) - f(0)$;
- (ii) $\|x_1\| < 2\varepsilon/\lambda$;
- (iii) $\lambda \|x - x_1\| + f(x_1) > f(x)$ whenever $x \neq x_1$.

Then (ii) tells us that $x_1 \in B(0, R)$ and, for every $\zeta \in \partial_p f(x_1)$, property (iii) combined with the preceding lemma implies that $\|\zeta\| \leq 2\varepsilon/R$.

Case II. Suppose now that $f(0) = -\varepsilon$, and choose any $\zeta \in \partial_p f(0)$. We can assume that $\|\zeta\| > 2\varepsilon/R$ (otherwise we are done). Then there exists h with $\|h\| = 1$ such that $\langle \zeta, h \rangle > 2\varepsilon/R$. On the other hand there exist $\eta > 0$ and $\sigma > 0$ such that

$$f(th) \geq f(0) + \langle \zeta, th \rangle - \sigma \|th\|^2$$

for all t with $|t| < \eta$, hence $f(th) + \varepsilon - t\langle \zeta, h \rangle \geq -\sigma t^2$, that is

$$\frac{f(th) + \varepsilon - t\langle \zeta, h \rangle}{|t|} \geq -\sigma |t|.$$

Bearing in mind the facts that $2\varepsilon/R - \langle \zeta, h \rangle < 0$ and that there exists $\delta > 0$ such that $2\varepsilon/R - \langle \zeta, h \rangle < -\sigma\delta$, we get that

$$\frac{f(\delta h) + \varepsilon - \delta\langle \zeta, h \rangle}{\delta} > \frac{2\varepsilon}{R} - \langle \zeta, h \rangle,$$

which implies $f(\delta h) + \varepsilon > 2\delta\varepsilon/R$, and therefore $f(\delta h) > \sup f(\overline{B}(0, R)) - 2\varepsilon$. Now, by setting $\lambda = 2\varepsilon/R$ and applying Ekeland’s variational principle we obtain some $x_1 \in \overline{B}(0, R)$ such that

- (i) $\lambda\|x_1 - \delta h\| \leq f(x_1) - f(\delta h)$;
- (ii) $\|x_1 - \delta h\| \leq \varepsilon/\lambda$;
- (iii) $\lambda\|x - x_1\| + f(x_1) > f(x)$ for all $x \neq x_1$.

According to (i) and taking into account that $f(\delta h) + \varepsilon > 2\varepsilon\delta/R$ we obtain that

$$\|x_1 - \delta h\| \leq \frac{f(x_1) - f(\delta h)}{2\varepsilon/R} \leq \frac{\varepsilon - f(\delta h)}{2\varepsilon/R} < \frac{2\varepsilon - 2\varepsilon\delta/R}{2\varepsilon/R} = R - \delta,$$

from which it follows $\|x_1\| \leq \|x_1 - \delta h\| + \delta < (R - \delta) + \delta = R$, and therefore $x_1 \in B(0, R)$.

From (iii) and the preceding lemma we get that $\|\zeta\| \leq 2\varepsilon/R$ for all $\zeta \in \partial_p f(x)$.

Finally, in the case $2\varepsilon \geq R$, bearing in mind that $\zeta \in \partial_p f(x)$ if and only if $r\zeta \in \partial_p(rf)(x)$ for all $r > 0$, and considering the function $g = \varepsilon' f/\varepsilon$, where $2\varepsilon' < R$, we may easily deduce from the above argument that there exists $x \in B(0, R)$ such that $\|\zeta\| \leq 2\varepsilon/R$ for all $\zeta \in \partial_p f(x)$. \square

Note that, as a consequence of the preceding proposition, for any continuous bounded function $f : X \rightarrow \mathbb{R}$ defined on the Hilbert space and satisfying $\partial_p f(x) \neq \emptyset$ for all $x \in X$, we have that

$$\inf\{\sup\{\|\zeta\| : \zeta \in \partial_p f(x)\}, x \in X\} = 0.$$

Proposition 2.11. *Let U be a connected bounded open subset of a real Hilbert space X . Let $f : \overline{U} \rightarrow \mathbb{R}$ be a bounded continuous function such that $\sup f(\overline{U}) > \sup f(\partial U)$. Then, for every $\alpha > 0$ there exists some $x \in U$ such that $\|\zeta\| < \alpha$ for all $\zeta \in \partial_p f(x)$.*

Proof. For a given $\alpha > 0$, consider the function $F : X \rightarrow [-\infty, +\infty)$ defined by $F(x) = f(x)$ if $x \in \overline{U}$ and $F(x) = -\infty$ if $x \notin \overline{U}$ (which is clearly upper semicontinuous and bounded above), a point $x_0 \in U$ such that $f(x_0) > \sup f(\partial U)$, and a number λ with $0 < \lambda < \min\{\alpha, 1\}$. Then, applying Ekeland’s variational principle (Theorem 1.2), we get a point $x_1 \in \overline{U}$ such that, from (i), $f(x_1) > f(x_0)$, and hence $x_1 \in U$, and from (iii) and according to Lemma 2.9, $\|\zeta\| \leq \lambda$ for all $\zeta \in \partial_p f(x_1)$, and therefore $\|\zeta\| \leq \alpha$. \square

As a consequence of the preceding results we can slightly improve the estimate on the norm of the subgradients.

Theorem 2.12. *Let U be a bounded connected open subset of a Hilbert space X . Let $f : \overline{U} \rightarrow \mathbb{R}$ be a bounded continuous function such that $\partial_p f(x) \neq \emptyset$ for all $x \in X$. Let $R > 0$ and $x_0 \in U$ be such that $\text{dist}(x_0, \partial U) = R$. Suppose that $f(\partial U) \subset [-\varepsilon, \varepsilon]$. Then there exist $x_\varepsilon \in U$ and $\zeta \in \partial_p f(x_\varepsilon)$ such that $\|\zeta\| \leq 2\varepsilon/R$.*

When f is constant on ∂U , we get $\inf\{\|\zeta\| : \zeta \in \partial_p f(x), x \in U\} = 0$.

3. An approximate Rolle's theorem for the generalized gradient

Definition 3.1. Let X be real Banach space and $f : X \rightarrow \mathbb{R}$ be a function such that f is Lipschitz on a neighborhood of a given point $x \in X$. The *generalized directional derivative* of f at x in the direction v , denoted $f^0(x; v)$, is defined as follows:

$$f^0(x; v) = \limsup_{(y,t) \rightarrow (x,0)} \frac{f(y+tv) - f(y)}{t},$$

where of course y is a vector in X and t is a positive real number. We define the *generalized gradient* $\partial f(x)$ of f at x as the set of all $\zeta \in X^*$ such that $f^0(x; v) \geq \langle \zeta, v \rangle$ for all v .

In the proofs of the results in this section we will need the rule for the generalized gradient of the sum, which we next state (a proof can be found in [6, p. 75]).

Proposition 3.2. Let f_i ($i = 1, 2, \dots, n$) be Lipschitz near x , and λ_i ($i = 1, 2, \dots, n$) be real numbers. Then $f = \sum_{i=1}^n \lambda_i f_i$ is Lipschitz near x , and we have

$$\partial \left(\sum_{i=1}^n \lambda_i f_i \right) (x) \subset \sum_{i=1}^n \lambda_i \partial f_i(x).$$

Before proceeding to prove an approximate Rolle's theorem for the generalized gradient, we are going to see that an exact Rolle's theorem for the generalized gradient fails completely in all infinite-dimensional Banach spaces, even if they do not have smooth bump functions. The main result from [4] tells us that Rolle's theorem (for smooth Lipschitz functions) fails in all Banach spaces which have smooth Lipschitz bumps, and is trivially true in those spaces which do not possess any such bumps. In particular, since for C^1 smooth and locally Lipschitz functions the generalized gradient is reduced to the usual differential, an exact Rolle's theorem is also false for the generalized gradient, in all spaces with C^1 smooth Lipschitz bumps. In this setting one could think that, if one takes a Banach space X with no C^1 smooth Lipschitz bump, one considers all locally Lipschitz functions f , and one looks at all of the generalized gradients $\partial f(x)$, then Rolle's theorem might be true, in the sense that if $f = 0$ on the boundary of a bounded connected open set U then there should exist one point $x \in U$ such that $0 \in \partial f(x)$. We next show that this is not the case.

Theorem 3.3. For every infinite-dimensional Banach space X there exists a bounded Lipschitz function f , defined on a bounded convex body U , such that f vanishes on ∂U and yet $0 \notin \partial f(x)$ for all $x \in \text{int}(U)$.

Proof. All reflexive spaces have equivalent C^1 smooth norms (see [8], for instance), and in every infinite-dimensional space with a C^1 smooth norm there is a bounded convex body U and a C^1 smooth function $f : U \rightarrow \mathbb{R}$ such that $f'(x) \neq 0$ for all $x \in \text{int}(U)$ (see [1]). Hence the result is true when X is reflexive, and we may assume that X is nonreflexive.

Then we can take a continuous linear functional $x^* \in X^*$ such that x^* does not attain its norm $\|x^*\| = 1$. Consider the function

$$f(x) = \begin{cases} x^*(x) & \text{if } x \in \overline{B}(0, 1), \\ \frac{2-\|x\|}{\|x\|}x^*(x) & \text{if } x \notin \overline{B}(0, 1), \end{cases}$$

defined on $\overline{B}(0, 2)$ and taking values in \mathbb{R} . The function f clearly vanishes on $S(0, 2)$. We have to prove that $0 \notin \partial f(x_0)$ for every $x_0 \in B(0, 2)$, which is equivalent to seeing that for every $x_0 \in B(0, 2)$ there exists $v \in X$ such that $f^0(x_0, v) < 0$. In the case when $x_0 \in B = B(0, 1)$ we have that $\partial f(x_0) = \{x^*\}$ and the result is obvious. In the case when $x_0 \in S = S(0, 1)$, we may consider the following situations.

Case I. If $x^*(x_0) > 0$, we may choose $x_1 \in S$ such that $x^*(x_1) > x^*(x_0)$ and $[x_0, x_1] \not\subset S$, in order to define a vector $v = x_0 - x_1$ which satisfies $x^*(v) < 0$. Let observe first that there exists $\varepsilon > 0$ such that

$$\|y + tv\| \geq \|y\| \quad \text{for every } y \in B(x_0, \varepsilon) \setminus B \text{ and } t > 0.$$

Indeed, the condition $[x_0, x_1] \not\subset S$ tells us that there is $t_0 > 0$ such that $x_0 - t_0v \in B$ and consequently $y - t_0v \in B \subset B(0, \|y\|)$ for y near x_0 , which implies $y + tv \notin B(0, \|y\|)$, equivalently $\|y + tv\| \geq \|y\|$ for every $t > 0$. To prove that $f^0(x_0, v) < 0$ we consider $(f(y + tv) - f(y))/t$ and three different situations.

(i) $y \in \overline{B}$ and $y + tv \in \overline{B}$. In this case $f(y + tv) - f(y)/t = x^*(y + tv) - x^*(y)/t = x^*(v)$.

(ii) $y \in \overline{B}$ and $y + tv \notin \overline{B}$. Then we have

$$\begin{aligned} \frac{f(y + tv) - f(y)}{t} &= \frac{1}{t} \left[\frac{2 - \|y + tv\|}{\|y + tv\|} x^*(y + tv) - x^*(y) \right] \\ &= \frac{1}{t} \left[\frac{2 - 2\|y + tv\|}{\|y + tv\|} x^*(y) \right] + \frac{2 - \|y + tv\|}{\|y + tv\|} x^*(v) \\ &\leq \frac{2 - \|y + tv\|}{\|y + tv\|} x^*(v) \leq \frac{x^*(v)}{2} \end{aligned}$$

if y is close enough to x_0 and $t > 0$ small, since $x^*(v) < 0$.

(iii) $y \notin \overline{B}$ and $y + tv \notin \overline{B}$. In this case we have

$$\begin{aligned} \frac{f(y + tv) - f(y)}{t} &= \frac{1}{t} \left[\frac{2 - \|y + tv\|}{\|y + tv\|} x^*(y + tv) - \frac{2 - \|y\|}{\|y\|} x^*(y) \right] \\ &= x^*(v) \frac{2 - \|y + tv\|}{\|y + tv\|} + \frac{1}{t} \left[\frac{2 - \|y + tv\|}{\|y + tv\|} - \frac{2 - \|y\|}{\|y\|} \right] x^*(y) \\ &= x^*(v) \frac{2 - \|y + tv\|}{\|y + tv\|} + \frac{2\|y\| - 2\|y + tv\|}{t\|y\|\|y + tv\|} x^*(y) \\ &\leq x^*(v) \frac{2 - \|y + tv\|}{\|y + tv\|} \leq \frac{x^*(v)}{2}. \end{aligned}$$

The case $y \notin \overline{B}$ and $y + tv \in \overline{B}$ is not allowed if y is close enough to x_0 . Taking lim sup we get that $f^0(x_0, v) \leq x^*(v)/2 < 0$.

Case II. If $x^*(x_0) < 0$, it is enough to apply Case I to the function $-f$, and remember that $\partial(-f)(x) = -\partial f(x)$.

Case III. If $x^*(x_0) = 0$, we can take a point $x_1 \in S$ such that $x^*(x_1) > 0$. Define $v = x_0 - x_1$, so that $x^*(v) < 0$. By considering the same situations as in Case I, and proceeding in a similar manner, it is easy to see that $f^0(x_0, v) < 0$.

Finally, when $x_0 \notin \bar{B}$, we may consider two cases.

(i) If $x^*(x_0) = 0$ we take x_1 such that $x^*(x_1) > 0$ and define $v = x_0 - x_1$. Then we have

$$\begin{aligned} \frac{f(y+tv) - f(y)}{t} &= \frac{1}{t} \left[\frac{2 - \|y+tv\|}{\|y+tv\|} x^*(y+tv) - \frac{2 - \|y\|}{\|y\|} x^*(y) \right] \\ &= \frac{2\|y\| - \|y+tv\|}{t\|y\|\|y+tv\|} x^*(y) + \frac{2 - \|y+tv\|}{\|y+tv\|} x^*(v) \\ &< \frac{2 - \|x_0\|}{2\|x_0\|} x^*(v) \end{aligned}$$

bearing in mind the facts that

$$\frac{2\|y\| - \|y+tv\|}{t\|y\|\|y+tv\|}$$

is bounded and $\lim_{y \rightarrow x_0} x^*(y) = 0$. It follows that $f^0(x_0, v) < 0$.

(ii) $x^*(x_0) \neq 0$ is similar to Cases I and II above, but considering only the situation $y \notin \bar{B}$ and $y+tv \notin \bar{B}$. \square

Let us now prove an approximate version of Rolle's theorem for the generalized gradient.

Theorem 3.4 (Rolle's theorem for the generalized gradient). *Let U be a bounded connected open subset of a real Banach space X , $f: \bar{U} \rightarrow \mathbb{R}$ be a bounded, locally Lipschitz function such that $f(\partial U) \subset [-\varepsilon, \varepsilon]$, and $R > 0$ and $x_0 \in U$ be such that $\text{dist}(x_0, \partial U) = R$. Then, $\inf\{\|\zeta\|: \zeta \in \partial f(x), x \in X\} \leq 2\varepsilon/R$.*

Note that, since the generalized gradient contains the proximal subgradient, for Hilbert spaces the statement is a straightforward consequence of Theorem 2.6. However, for Banach spaces which are not Hilbertian or do not possess any C^2 smooth bump functions, a different proof is required. We will split the proof into two easy propositions.

Proposition 3.5. *Let U be a bounded open subset of a real Banach space X and $f: \bar{U} \rightarrow \mathbb{R}$ be a bounded locally Lipschitz function satisfying that $\sup f(U) > \sup f(\partial U)$ or $\inf f(U) < \inf f(\partial U)$. Then, for every $\alpha > 0$ there exist $x \in U$ and $\zeta \in \partial f(x)$ such that $\|\zeta\| < \alpha$.*

Proof. Assume first that $\sup f(U) > \sup f(\partial U)$. Consider the function F defined as $F(x) = f(x)$ for $x \in \bar{U}$ and $F(x) = -\infty$ if $x \notin \bar{U}$. Let $\eta = \sup f(U) - \sup f(\partial U)$ and choose $x_0 \in U$ so that $f(x_0) > \sup f(U) - \eta$. By Ekeland's variational principle, for each α with $0 < \alpha < 1$ we can find $x_1 \in \text{Dom}(F)$ such that $\alpha\|x_1 - x_0\| \leq f(x_1) - f(x_0)$,

$\|x_1 - x_0\| < \eta/\alpha$, and $\alpha\|x - x_1\| + f(x_1) > f(x)$ whenever $x_1 \neq x$. These inequalities yield that $f(x_1) > f(x_0)$, hence $x_1 \in U$, and that the function $\Phi(x) = f(x) - f(x_1) - \alpha\|x - x_1\|$ attains a maximum at $x = x_1$, which gives $0 \in \partial\Phi(x_1)$ and, by applying the rule for the generalized gradient of the sum (Proposition 3.2), we obtain that $0 \in \partial f(x_1) + \partial(-\alpha\|x - x_1\|)$; that is, there exist $\zeta \in \partial f(x_1)$ and $\vartheta \in -\alpha\partial\|\cdot\|(x - x_1)$ with $0 = \zeta + \vartheta$, and, since $\|\vartheta\| \leq \alpha$, we conclude that $\|\zeta\| \leq \alpha$. \square

Proposition 3.6. *Let X be a real Banach space, let $B = B(0, R)$ and $f: \overline{B} \rightarrow \mathbb{R}$ be a locally Lipschitz function so that $f(\overline{B}) \subset [-\varepsilon, \varepsilon]$. Then, for every $\alpha > 0$ there exist $x \in \text{int}(B)$ and $\zeta \in \partial f(x)$ such that $\|\zeta\| < 2\varepsilon/R + \alpha$.*

Proof. Consider the function $\Phi(x) = f(x) - ((2\varepsilon + \alpha')/R)\|x\|$, with $\alpha' > 0$. For all $x \in \partial B$ we have that $\Phi(x) = f(x) - (2\varepsilon + \alpha') < f(0)$. Then we may apply the preceding proposition to the function Φ and obtain a point $x \in B$ and some subgradient $\vartheta_1 \in \partial\Phi(x)$ such that $\|\vartheta_1\| < \alpha'$. Then, according to the rule for the generalized gradient of the sum (Proposition 3.2), $\vartheta_1 \in \partial f(x) - ((2\varepsilon + \alpha')/R)\partial\|\cdot\|(x)$, and therefore $\vartheta_1 = \zeta - ((2\varepsilon + \alpha')/R)\vartheta_2$, where $\|\vartheta_2\| \leq 1$, from which we deduce that $\|\zeta\| \leq \|\vartheta_1\| + (2\varepsilon + \alpha')/R \leq \alpha' + (2\varepsilon + \alpha')/R = 2\varepsilon/R + [\alpha' + \alpha'/R]$. By taking α' such that $\alpha' + \alpha'/R < \alpha$ the result follows. \square

References

- [1] D. Azagra, Smooth negligibility and subdifferential calculus in Banach spaces, with applications, Ph.D. dissertation, Universidad Complutense de Madrid, October 1997.
- [2] D. Azagra, R. Deville, Subdifferential Rolle's and mean value inequality theorems, Bull. Austral. Math. Soc. 56 (1997) 319–329.
- [3] D. Azagra, J. Gómez, J.A. Jaramillo, Rolle's theorem and negligibility of points in infinite-dimensional Banach spaces, J. Math. Anal. Appl. 213 (1997) 487–495.
- [4] D. Azagra, M. Jiménez-Sevilla, The failure of Rolle's theorem in infinite-dimensional Banach spaces, J. Funct. Anal. 182 (2001) 207–226.
- [5] J. Bes, J. Ferrera, On a multidimensional version of Rolle's theorem, Publ. Dto. Análisis Mat. UCM 1 (1996/1997) 21–27.
- [6] F.H. Clark, Yu.S. Ledyev, R.J. Stern, P.R. Wolenski, Nonsmooth Analysis and Control Theory, in: Graduate Texts in Mathematics, Vol. 178, Springer, 1998.
- [7] R. Deville, A mean value theorem for nondifferentiable mappings in Banach spaces, Serdica Math. J. 21 (1995) 59–66.
- [8] R. Deville, G. Godefroy, V. Zizler, Smoothness and Renormings in Banach Spaces, in: Pitman Monographs and Surveys in Pure and Applied Mathematics, Vol. 64, 1993.
- [9] I. Ekeland, Nonconvex minimization problems, Bull. Amer. Math. Soc. (N.S.) 1 (1979) 443–474.
- [10] J. Ferrer, Rolle's, theorem fail in l_2 , Amer. Math. Monthly 103 (1996) 161–165.
- [11] G. Godefroy, Some remarks on subdifferential calculus, Rev. Mat. Complut. 11 (1998) 269–279.
- [12] S.A. Shkarin, On Rolle's theorem in infinite-dimensional Banach spaces, transl. from Mat. Zametki 51 (1992) 128–136.