Journal of Convex Analysis
Volume 5 (1998), No. 2, 221–235

# Limiting Convex Examples for Nonconvex Subdifferential Calculus

**Jonathan M. Borwein***
*CECM, Department of Mathematics and Statistics,
Simon Fraser University, Burnaby, B. C. V5A 1S6, Canada.*
*e-mail: jborwein@cecm.sfu.ca*

**Qiji J. Zhu**†
*Department of Mathematics and Statistics, Western Michigan University,
Kalamazoo, MI 49008, USA.*
*e-mail: zhu@math-stat.wmich.edu*

Received June 20, 1997
Revised manuscript received January 6, 1998

We show, largely using convex examples, that most of the core results for limiting subdifferential calculus fail without additional restrictions in infinite dimensional Banach spaces.

*Keywords:* Nonsmooth analysis, subdifferentials, coderivatives, extremal principle, open mapping theorem, metric regularity, multiplier rules, compactly Lipschitzian conditions

*1991 Mathematics Subject Classification:* 26B05

## 1. Introduction

Through the work of [3, 4, 26, 27] it has become clear that smooth subdifferentials characterize many important generalized derivative concepts such as Clarke's generalized gradient, Ioffe's geometric subdifferential and Mordukhovich'es limiting subdifferential. This renewed interest in smooth subdifferentials has led to many useful results in the calculus of subdifferentials and their applications. Using a limiting process, results stated in terms of the smooth subdifferentials can be rephrased in terms of the limiting subdifferential, the singular subderivative and the limiting normal cone *in finite dimensional spaces*. In infinite dimensional spaces similar limiting results have also been attacked either through a limiting process from corresponding results for smooth subdifferentials or directly by using various constructions for generalized derivatives or normal cones. However, these corresponding results in infinite dimensional spaces are always accompanied by some additional conditions. There are essentially two types of such condition: (a) *local Lipschitz or directional Lipschitz* conditions (see e.g. [6, 7, 11, 12]) and (b) *compactly epi-Lipschitzian, partially normal compactness and codirectional compactness* conditions (see e.g. [2, 14, 18, 24, 25]). It is natural to ask whether these limiting results hold without those additional assumptions. We show below through examples that the answer is negative in almost all cases. First we recall some related definitions.

*Research was supported by NSERC and by the Shrum Endowment at Simon Fraser University.
†Research was supported by the National Science Foundation under grant DMS-9704203 and by funds from the Faculty Research and Creative Activities Support Fund, Western Michigan University.

ISSN 0944-6532 / $ 2.50 

Let $X$ be a Banach space with closed unit ball $B_X$. We use $\bar{R}$ to denote the extended real line $R \cup \{+\infty\}$ and $2^X$ to denote the collection of all subsets of $X$. Let $f : X \to \bar{R}$ be a lower semicontinuous function with $f(x) < +\infty$. We say $f$ is (Fréchet)-*subdifferentiable* and $x^*$ is a (Fréchet)-*subderivative* of $f$ at $x$ if there exists a function $g$ such that $g$ has continuous Fréchet derivative in a neighborhood of $x$, $\nabla g(x) = x^*$ and $f - g$ attains a local minimum at $x$. We denote the set of all Fréchet subderivatives of $f$ at $x$ by $D_F f(x)$. For a closed subset $C$ of $X$ the *Fréchet normal cone* to $C$ at $x \in C$ is defined by $N_F(C, x) := D_F \delta_C(x)$. Here $\delta_C$ is the indicator function of $C$ defined by $\delta_C(x) = 0$ if $x \in C$ and $\delta_C(x) = +\infty$ otherwise. We turn to the definition of the corresponding *limiting objects*. In what follows $w^*$-lim signifies the weak-star sequential limit.

**Definition 1.1 ([19, 24]).** Let $f : X \to \bar{R}$ be a lower semicontinuous function. Define

$$\partial f(x) := \{w^* - \lim_{i\to\infty} v_i : v_i \in D_F f(x_i),\ (x_i, f(x_i)) \to (x, f(x))\},$$

and

$$\partial^\infty f(x) := \{w^* - \lim_{i\to\infty} t_i v_i : v_i \in D_F f(x_i),\ t_i \to 0^+,\ (x_i, f(x_i)) \to (x, f(x))\}$$

and call $\partial f(x)$ and $\partial^\infty f(x)$ the subdifferential and singular subdifferential of $f$ at $x$ respectively.

Secondly, let $C$ be a closed subset of $X$. Define

$$N(C, x) := \{w^* - \lim_{i\to\infty} v_i : v_i \in N_F(C, x_i),\ C \ni x_i \to x\}$$

and call $N(C, x)$ the normal cone of $C$ at $x$.

Let $F : X \to 2^Y$ be a multifunction with closed graph and let $y \in F(x)$. We say $x^* \in X^*$ is a coderivative of $F$ at $(x, y)$ corresponding to $y^* \in Y^*$ provided that

$$(x^*, -y^*) \in N(\mathrm{graph}(F), (x, y)).$$

We denote the collection of all coderivatives of $F$ at $(x, y)$ corresponding to $y^*$ by $\partial^* F(x, y)$ $(y^*)$.

**Remark 1.2.** (a) Denote by $\partial_c$, $\partial_g$ and $\partial_g^\infty$ the Clarke generalized gradient [7], the regular and singular geometric subgradient [3, 11], respectively. If $X$ has an equivalent Fréchet smooth norm then $\partial_c f(x) = \mathrm{cl}^*\ \mathrm{co}[\partial f(x) + \partial^\infty f(x)]$, $\partial_g f(x) = \mathrm{cl}^*\, \partial f(x)$ and $\partial_g^\infty f(x) = \mathrm{cl}^*\, \partial^\infty f(x)$, where $\mathrm{cl}^*$ signifies the weak-star closure (see e.g. [3]). The subdifferential in the above definition has the advantage of being the smallest among the sequentially upper semicontinuous subdifferentials. For this reason we state the finite dimensional positive results below in terms of this limiting subdifferential.

(b) When $f$ is convex $\partial$, $\partial_g$ and $\partial_c$ coincide with the usual subderivative in convex analysis and $\partial^\infty$, $\partial_g^\infty$ and $\partial_c^\infty$ coincide with the recession subdifferential in convex analysis. Therefore the convex examples given below also provide examples for the Clarke generalized gradient and Ioffe's geometric subdifferential.

(c) Many valuable results in terms of the limiting subdifferential, singular subdifferential and normal cone require the underlying Banach spaces to have certain smoothness properties, such as to have a Fréchet smooth equivalent norm. The corresponding results in terms of Clarke generalized gradient and the geometric subdifferential can be derived in general Banach spaces. Later when we discuss examples in general Banach spaces we will use Clarke's generalized gradient. We recall the definition below.

Let $X$ be a Banach space, let $C$ be a closed subset of $X$ and let $x \in C$. The Clarke normal cone of $C$ at $x$ is defined by

$$N_c(C, x) := \operatorname{cone}\{x^* \in X^* : \langle x^*, h\rangle \le \limsup_{x' \to x, t \to 0+} d_C(x' + th)/t\}$$

where $d_C(x) := \inf\{\|x - y\| : y \in C\}$, is the (metric) *distance function* to the set $C$. For a lower semicontinuous function $f : X \to \bar{R}$, the Clarke generalized gradient and singular generalized gradient are defined by using the normal cone to the epigraph of $f$ as follows:

$$\partial_c f(x) := \{x^* \in X^* : (x^*, -1) \in N_c(\mathrm{epi}(f), (x, f(x)))\}$$

and

$$\partial_c^\infty f(x) := \{x^* \in X^* : (x^*, 0) \in N_c(\mathrm{epi}(f), (x, f(x)))\}.$$

We emphasis that most of the examples we construct are convex. This has several useful consequences:

(a) It shows that convex objects are already complicated enough to exhibit most pathological behaviour in infinite dimensional nonsmooth analysis.
(b) The underlying geometric features of the counter-examples are laid bare.
(c) The simple geometric nature of those examples facilitates various manipulations to create more delicate (nonconvex) examples.

## 2. Calculus for subdifferentials

The sum rule is the most basic in the subdifferential calculus. In finite dimensional spaces the following sum rule holds (see e.g. [8, Proposition 1.5], [17, Proposition 5A.4] and [23, Corollary 4.6]).

**Theorem 2.1.** *Let $X$ be a finite dimensional Banach space and let $f_i : X \to \bar{R}$, $i = 1, 2$ be lower semicontinuous functions finite at $x$ with $\partial^\infty f_1(x) \cap (-\partial^\infty f_2(x)) = \{0\}$. Then*

(a) $\partial(f_1 + f_2)(x) \subset \partial f_1(x) + \partial f_2(x)$, *and*
(b) $\partial^\infty(f_1 + f_2)(x) \subset \partial^\infty f_1(x) + \partial^\infty f_2(x)$.

Our first example shows that this sum rule fails in infinite dimensional spaces without additional assumptions. Here and below we usually construct the examples in (always infinite dimensional) *separable Banach spaces*. Moreover, we primarily construct convex examples. Thus, they are also examples for the Clarke generalized gradient and the geometric subdifferential. At the end, we will comment on generalizations.

**Example 2.2.** Let $X$ be a separable Banach space. According to [16, Proposition 1.f.3], $X$ admits a *Markushevich basis* (a fundamental and total biorthogonal sequence), i.e., there exists a biorthogonal collection $\{x_n, f_n\}_{n=1}^\infty$ such that $f_n \in X^*$, $\|x_n\| = 1$ for all $n$, $\{\|x_n\|\|f_n\|\}_{n=1}^\infty$ bounded and $\operatorname{span}(\{x_n\}_{n=1}^\infty)$ is norm dense in $X$, and $f_n(x) \equiv 0$ implies $x = 0$.

Define

$$A := \{f : f(x_{2n}) = 0, n = 1, 2, \cdots\}$$

and

$$B := \{f : f(x_{2n+1} - 2^n x_{2n}) = 0, n = 1, 2, \cdots\}.$$

Then clearly $A$ and $B$ are $w^*$-closed and since the $\{x_k\}$ are densely spanning $A \cap B = 0$. Now consider

$$a_k := f_{2k+1}, \quad b_k := f_{2k+1} - 2^{-k} f_{2k}$$

$$c_k := f_{2k+1} - (f_{2k+1} - 2^{-k} f_{2k})$$

It is easy to check that $a_k \in A$ and $b_k \in B$. Moreover since $A \cap B = 0$ the representation of $c_k$ is unique from $A + B$. Thus, since $c_k \to 0$ while $\|a_k\| \geq 1$ (by biorthogonality), we see that $A + B$ is not closed, otherwise the convex Open Mapping Theorem [1] gives $a_k \to 0$.

Finally, since $\{f_k\}$ are total, $A+B$ is $w^*$-dense in $X^*$. So if we let $M_1 := \{x \in X : \langle a, x\rangle = 0,\ \forall\ a \in A\}$ and $M_2 := \{x \in X : \langle b, x\rangle = 0,\ \forall\ b \in B\}$ we have $M_1^\perp + M_2^\perp$ is $w^*$-dense in $X^*$ but not closed and $M_1^\perp \cap M_2^\perp = 0$. Define $f_1 := \delta_{M_1}$ and $f_2 := \delta_{M_2} + \langle v, \cdot\rangle$ where $-v \in X^* \backslash (M_1^\perp + M_2^\perp)$. Since $M_1^\perp + M_2^\perp$ $w^*$-dense implies that $M_1 \cap M_2 = \{0\}$, $f_1 + f_2$ attains a minimum at 0. However, it is easy to check that $\partial f_1(0) = \partial^\infty f_1(0) = M_1^\perp$, $\partial f_2(0) = M_2^\perp + v$ and $\partial^\infty f_2(0) = M_2^\perp$. Thus,

$$0 \notin \partial f_1(0) + \partial f_2(0)$$

and

$$\partial^\infty f_1(0) \cap (-\partial^\infty f_2(0)) = \{0\}$$

or equivalently

$$0 \in \partial^\infty f_1(0) + \partial^\infty f_2(0)$$

holds only in the trivial case.

**Remark 2.3.** The subspaces $M_1$ and $M_2$ in Example 2.2 are $w^*$-closed *quasi-complements*. Whether such quasi-complements exist in all Banach spaces remains open.

A main application of the subdifferential sum rule is to derive necessary conditions for constrained minimization problems. Necessary conditions for constrained minimization problems with Lipschitz data can be found in [7, 10, 17]. General necessary conditions for problems with lower semicontinuous and continuous data can be found in [20, Theorem 1(b)] (see also [21, Section 7]) and [5, Corollary 2.3]. We recall this necessary condition below.

Let $C \subset X$ and $g_i : X \to \bar{R}$, $\quad i = 0, 1, \ldots, N$. Consider the following *optimization* problem:

$$\begin{array}{rll} \text{minimize} & g_0(x) & \\ \text{subject to} & g_i(x) \leq 0, & i = 1, 2, \ldots, M, \\ & g_i(x) = 0, & i = M+1, \ldots, N, \\ & x \in C. & \end{array} \tag{$\mathcal{P}$}$$

To simplify notation we introduce the quantities $\tau_i, i = 0, 1, \ldots, N$. The $\tau_i's$ associated with the inequality constraints and the cost function are always 1, i.e., $\tau_i = 1, i =$

$0, 1, \ldots, M$. This corresponds to nonnegative multipliers. The $\tau_i's$ associated with the equality constraints are either 1 or $-1$, corresponding to multipliers with arbitrary sign, i.e., $\tau_i \in \{-1, 1\}, i = M+1, \ldots, N$.

**Theorem 2.4.** *Let $X$ be a finite dimensional Banach space, let $C$ be a closed subset of $X$ and let $g_i$ be lower semicontinuous for $i = 0, 1, \ldots, M$ and continuous for $i = M+1, \ldots, N$. Suppose that $\bar{x}$ is a local solution of problem $(\mathcal{P})$. Then either:*

(A1) *there exist $v_i^\infty \in \partial^\infty(\tau_i g_i)(\bar{x})$, $i = 0, 1, \ldots, N$ and $v_{N+1}^\infty \in N(C, \bar{x})$ such that $0 = \sum_{i=0}^{N+1} v_i^\infty$ and $\sum_{i=0}^{N+1} \|v_i^\infty\| = 1$.*

*or there exist $\mu_j \geq 0$, $j = 0, ..., N$ satisfying $\sum_{j=0}^{N} \mu_j = 1$ such that*

(A2)
$$0 \in \sum_{j \in \{i : \mu_i > 0\}} \mu_j \partial(\tau_j g_j)(\bar{x}) + \sum_{j \in \{i : \mu_i = 0\}} \partial^\infty(\tau_j g_j)(\bar{x}) + N(C, \bar{x}).$$

**Remark 2.5.** When all the functions in the constraints are Lipschitz the singular subdifferential part in (A2) superfluous. In general, it cannot be eliminated. This is demonstrated by the following elementary example:

**Example 2.6.** Consider problem $(\mathcal{P})$ with $X = C = R$, $N = M = 1$, $g_0(x) = x$ and $g_1(x) = -x^{1/3}$. Then 0 is the only solution. We can calculate directly that $\partial g_0(0) = \{1\}$, $\partial^\infty g_0(0) = \{0\}$, $\partial g_1(0) = \emptyset$, $\partial^\infty g_1(0) = (-\infty, 0]$. It is clear that relation A1 is impossible at 0 and A2 can be satisfied at 0 only if $\mu_0 = 1$ and $\mu_1 = 0$. In that case $0 \in \partial g_0(0) + \partial^\infty g_1(0) = \{1\} + (-\infty, 0]$.

Using a standard trick we can convert Example 2.2 to show Theorem 2.4 does not hold without additional assumptions in infinite dimensional spaces.

**Example 2.7.** Again let $X$ be a separable Banach space and let $M_1$ and $M_2$ be closed subspaces of $X$ such that $M_1^\perp \cap M_2^\perp = 0$ and $M_1^\perp + M_2^\perp$ $w^*$-dense but not closed. Let $v \notin X^* \backslash (M_1^\perp + M_2^\perp)$ and $g_0(x) = \delta_{M_1}(x) - \langle v, x \rangle$. Consider the problem of minimizing $g_0(x)$ subject to $x \in M_2$. Then 0 is a solution. The normal optimality condition $0 \in \partial g_0(0) + N(M_2, 0) = M_1^\perp + M_2^\perp - v$ does not hold and the singular condition $0 \in \partial^\infty g_0(0) + N(M_2, 0) = M_1^\perp + M_2^\perp$ holds only in the trivial case.

The chain rule is another important calculus rule for subdifferentials. Chain rules with a Lipschitz condition or other additional assumptions can be found in [7, 17, 21, 23, 28] with further references and discussions. A chain rule for lower semicontinuous functions and continuous functions can be deduced from Theorem 2.4 using a method in [29]. Consider the composition $f \circ g$ where $f : R^N \to \bar{R}$ and $g = (g_1, ..., g_N) : X \to R^N$. Suppose that $f$ is nondecreasing in its first $M$ components. Consider a local minimum $\bar{x}$ of $f \circ g$. Then $(\bar{x}, g(\bar{x}))$ is a local solution of:

$$\begin{array}{rll} \text{minimize} & f(y) & \\ \text{subject to} & g_i(x) - y_i \leq 0, & i = 1, 2, \ldots, M, \\ & g_i(x) - y_i = 0, & i = M+1, \ldots, N. \end{array}$$

An application of Theorem 2.4 yields the following chain rule.

**Theorem 2.8.** *Let $X$ be a finite dimensional Banach space, let $f : R^N \to \bar{R}$ and $g_i : X \to \bar{R}, i = 1, ..., M$ be lower semicontinuous functions and let $g_i : X \to R, i = M+1, ..., N$ be continuous functions. Suppose that $f(g_1, ..., g_N)$ attains a minimum at $\bar{x}$ and that $f$ is nondecreasing in its first $M$ components. Then either:*

(A1) *there exist $v_i^\infty \in \partial^\infty(\tau_i g_i)(\bar{x})$, $i = 1, \dots, N$ such that $0 = \sum_{i=1}^N v_i^\infty$ and $\sum_{i=1}^N \|v_i^\infty\| = 1$.*

*or there exist $\mu = (\mu_1, ..., \mu_N) \in \partial f(g(\bar{x}))$ such that*

(A2)
$$0 \in \sum_{j \in \{i:\mu_i \neq 0\}} \mu_j \partial(\tau_j g_j)(\bar{x}) + \sum_{j \in \{i:\mu_i = 0\}} \partial^\infty(\tau_j g_j)(\bar{x}).$$

Setting $M = N = 2$ and $f(y_1, y_2) := y_1 + y_2$ in Theorem 2.8 we obtain a special case of the sum rule of Theorem 2.1 for which Example 2.2 still applies. Thus, Example 2.2 shows that the chain rule Theorem 2.8 fails in infinite dimensional spaces. Moreover, since the chain rule follows from Theorem 2.4 for minimization problems with no set constraint, it a fortiori follows that Theorem 2.4 without a set constraint also fails in infinite dimensional spaces.

The normal cone relation for the intersection of sets (see e.g. [11, Theorem 5.4] and [24, Corollary 4.5]) is a useful geometric consequence of the sum rule for subdifferentials. It asserts that in finite dimensional spaces $\bar{x} \in S_1 \cap S_2$ and $N(S_1, \bar{x}) \bigcap (-N(S_2, \bar{x})) = \{0\}$ implies that
$$N(S_1 \cap S_2, \bar{x}) \subset N(S_1, \bar{x}) + N(S_2, \bar{x}).$$
In infinite dimensional Asplund spaces this result holds under an additional normal compactness condition (see [24, Corollary 4.5]). This assertion also fails in infinite dimensional spaces without additional assumptions.

**Example 2.9.** Let $X$ be a separable Banach space and let $M_1$ and $M_2$ be closed subspaces of $X$ such that $M_1^\perp \cap M_2^\perp = 0$ with $M_1^\perp + M_2^\perp$ $w^*$-dense but not closed. Then $M_1 \cap M_2 = \{0\}$ and
$$X^* = N(M_1 \cap M_2, 0) \not\subset N(M_1, 0) + N(M_2, 0) = M_1^\perp + M_2^\perp.$$

We emphasize that the above phenomenon is due to the behavior of the normal cones of the sets. Even imposing norm compactness of the sets involved will not help as is shown by the following example.

**Example 2.10.** Let $X$ be an infinite dimensional Banach space and $\{e_n\}_{n=1}^\infty$ unit independent vectors in $X$. Define $S_1 := \operatorname{cl}\operatorname{co}\{\frac{\pm e_n}{2^n}\}$ and $S_2 := \{tv : t \in [-1, 1]\}$ where $v := (\sum_{n=1}^\infty \frac{e_n}{n^2}) \in X$. Then both $S_1$ and $S_2$ are norm compact and $S_1 \cap S_2 = \{0\}$. However, noting that $v \in \operatorname{cl}\operatorname{span}(S_1)$,
$$X^* = N(S_1 \cap S_2, 0) \not\subset N(S_1, 0) + N(S_2, 0) = \operatorname{span}(S_1)^\perp + \operatorname{span}(S_2)^\perp = \operatorname{span}(S_1)^\perp$$

## 3. Calculus for coderivatives

Since there are standard methods of deducing calculus for subdifferentials from calculus for coderivatives of multifunctions [25], limiting examples for the calculus of subdifferentials also provide limiting examples for coderivative calculus. However, direct construction seems much easier. We provide two such examples below for the sum rule and the chain rule respectively. First let us state a finite dimensional version of the sum rule for coderivatives from [23, Theorem 4.1]. Recall that a multifunction $G : X \to 2^Y$ is *lower semicompact* around $\bar{x}$ if there exists a neighborhood $U$ of $\bar{x}$ such that for any $x \in U$ and any sequence $x_k \to x$, there exists a sequence $y_k \in G(x_k)$ that contains a norm convergent subsequence.

**Theorem 3.1.** *Let $X$ and $Y$ be finite dimensional Banach spaces, let $F_1$ and $F_2$ be multifunctions from $X$ to $Y$ with closed graphs, and let $\bar{y} \in F_1(\bar{x}) + F_2(\bar{x})$. Assume that the multifunction*

$$S(x, y) := \{(y_1, y_2) : y_1 \in F_1(x), y_2 \in F_2(x), y_1 + y_2 = y\}$$

*is lower semicompact around $(\bar{x}, \bar{y})$, and that the following condition is fulfilled:*

$$\partial^* F_1(x; y_1)(0) \cap (-\partial^* F_2(x; y_2)(0)) = \{0\}, \forall (y_1, y_2) \in S(\bar{x}, \bar{y}).$$

*then*

$$\partial^*(F_1 + F_2)(\bar{x}, \bar{y})(y^*) \subset \bigcup_{(y_1,y_2)\in S(\bar{x},\bar{y})} [\partial^* F_1(\bar{x}, y_1)(y^*) + \partial^* F_2(\bar{x}, y_2)(y^*)].$$

**Remark 3.2.** The results in [23, Theorem 4.1] are derived under the weaker condition that the multifunction $S(x, y)$ is locally bounded. We stated the theorem with the lower semicompactness condition because it is more commonly used for results in infinite dimensional spaces (see [22, 24, 25]).

We adapt Example 2.2 to provide an example showing the coderivative sum rule may fail in infinite dimensional spaces.

**Example 3.3.** Again let $X$ be a separable Banach space and let $M_1$ and $M_2$ be closed subspaces of $X$ such that $M_1^\perp \cap M_2^\perp = 0$ and $M_1^\perp + M_2^\perp$ is $w^*$-dense but not closed. Recall that $M_1^\perp + M_2^\perp$ $w^*$-dense implies that $M_1 \cap M_2 = \{0\}$. Define multifunctions $F_1, F_2 : H \to 2^R$ by $\text{graph}(F_i) := M_i \times R^+$, $i = 1, 2$. Then $\text{graph}(F_1 + F_2) = \{0\} \times R^+$. Consider $0 \in F_1(0)+F_2(0)$. The set $S(x, y) := \{(y_1, y_2) \in R^2 : y_1 \in F_1(x), y_2 \in F_2(x), y_1+y_2 = y\}$ is $\{(0, 0)\}$ at $(x, y) = (0, 0)$ and $\emptyset$ elsewhere. It is obviously lower semicompact around $(0, 0)$. Easy calculation shows that $\partial^* F_i(0, 0)(0) = M_i^\perp$ for $i = 1, 2$ and $\partial^*(F_1+F_2)(0, 0)(0) = X^*$. Thus the regularity condition

$$\partial^* F_1(0, 0)(0) \cap (-\partial^* F_2(0, 0)(0)) = \{0\}$$

holds yet the sum rule

$$\partial^*(F_1 + F_2)(0, 0)(0) \subset \partial^* F_1(0, 0)(0) + \partial^* F_2(0, 0)(0)$$

fails.

Now we turn to the chain rule. The following is a finite dimensional space version of the chain rule [23, Theorem 5.1]. Again note that in the finite dimensional result given in [23, Theorem 5.1] the weaker locally boundedness condition is used for $M(x, z)$.

**Theorem 3.4.** *Let $X$,$Y$ and $Z$ be finite dimensional Banach spaces and let $F : X \times Y \to 2^Z$ and $G : X \to 2^Y$ be multifunctions with closed graphs. Assume that the multifunction*

$$M(x, z) := G(x) \cap F^{-1}(z) = \{y \in G(x) : z \in F(x, y)\}$$

*is lower semicompact around $(\bar{x}, \bar{z})$. Assume also that for any $\bar{y} \in M(\bar{x}, \bar{z})$ the regularity condition*

$$[(x^*, y^*) \in \partial^* F((\bar{x}, \bar{y}); \bar{z})(0) \& -x^* \in \partial^* G(\bar{x}; \bar{y})(y^*)] \Rightarrow x^* = 0 \ \& \ y^* = 0$$

*holds. Then, for all $z^* \in Z^*$,*

$$\partial^*(F \circ G)(\bar{x}; \bar{z})(z^*) \quad \subset \bigcup_{\bar{y} \in M(\bar{x}, \bar{z})} [x_1^* + x_2^* : x_1^* \in \partial^* G(\bar{x}; \bar{y})(y^*),$$
$$(x_2^*, y^*) \in \partial^* F((\bar{x}, \bar{y}); \bar{z})(z^*)].$$

**Example 3.5.** Let the separable Banach space $X$ and its subsets $M_1$ and $M_2$ be as in the previous example. Let $Y = Z = R$. Define multifunctions $F$ and $G$ by

$$G(x) := \begin{cases} R^+ & x \in M_1, \\ \emptyset & \text{otherwise} \end{cases}$$

and

$$F(x, r) := \begin{cases} R^+ & (x, r) \in M_2 \times R^+, \\ \emptyset & \text{otherwise.} \end{cases}$$

Then

$$F(x, G(x)) = \begin{cases} R^+ & x = 0, \\ \emptyset & \text{otherwise.} \end{cases}$$

When $\bar{x} = 0$ and $\bar{z} = 0$ we have $M(0, 0) = \{0\}$ and it is the only value for $(x, z)$ that makes $M(x, z) \neq \emptyset$. Thus, $M$ is lower semicompact around $(0, 0)$. Next we check that the regularity condition is satisfied. In fact, $x^*$ and $y^*$ satisfying the regularity condition amounts to

$$(x^*, y^*, 0) \in N(\text{graph}(F), (0, 0, 0)) = M_2^\perp \times (-R^+) \times (-R^+)$$

and

$$(-x^*, -y^*) \in N(\text{graph}(G), (0, 0)) = M_1^\perp \times (-R^+).$$

This obviously implies that $x^* = 0$ and $y^* = 0$.

Nevertheless, the chain rule does not hold because

$$\partial^*(F \circ G)(0; 0)(0) = X^*$$

while

$$x_1^* \in \partial^* G(0; 0)(y^*)$$

and

$$(x_2^*, y^*) \in \partial^* F((0, 0); 0)(0)$$

implies that $x_1^* \in M_1^\perp$ and $x_2^* \in M_2^\perp$.

We refer to [15, 24, 25] and the references therein for detailed discussions on calculus for coderivatives in infinite dimensional spaces.

## 4. The extremal principle

The extremal principle can be traced back to Mordukhovich'es work [19] (see also [20, 21]). In Fréchet smooth Banach spaces this result was established in Kruger and Mordukhovich [9] while the term "extremal principle" was coined in [23]. It extends the Separation Theorem for convex sets to nonconvex sets and is useful in many applications. The following is a finite dimensional form of the extremal principle [23, Theorem 3.2]. We recall the (equivalent) definition of an extremal point first. Let $S_1$ and $S_2$ be closed sets in a Banach space $X$. A point $\bar{x}$ is called a *local extremal point* of $(S_1, S_2)$ provided that $\bar{x} \in S_1 \cap S_2$ and there is a neighborhood $U$ of $\bar{x}$ and a sequence $\{a_k\}$ in $X$ such that $a_k \to 0$ and

$$S_1 \cap (S_2 - a_k) \cap U = \emptyset.$$

**Theorem 4.1.** *Let $X$ be a finite dimensional Banach space, let $S_1$ and $S_2$ be closed subsets of $X$ and let $\bar{x}$ be a local extremal point of $(S_1, S_2)$. Then*

$$N(S_1, \bar{x}) \cap (-N(S_2, \bar{x})) \neq \{0\}.$$

An easy modification of Example 2.2 give us the following example that shows this extremal principle does not hold in infinite dimensional space without additional assumptions.

**Example 4.2.** Let $X$ be a separable Banach space and let $M_1$ and $M_2$ be closed subspaces of $X$ such that $M_1^\perp \cap M_2^\perp = 0$ and $M_1^\perp + M_2^\perp$ dense but not closed. Again note that $M_1^\perp + M_2^\perp$ $w^*$-dense implies that $M_1 \cap M_2 = \{0\}$. Observe that $M_1 + M_2$ is not $X$ otherwise $M_1^\perp + M_2^\perp$ is closed. Let $v \notin M_1 + M_2$. Then, for any sequence of real numbers $r_k \to 0$, $r_k v \notin M_1 + M_2$. That is to say $(r_k v + M_1) \cap M_2 = \emptyset$ so that 0 is an extremal point for $(M_1, M_2)$. However, $N(M_1, 0) \cap (-N(M_2, 0)) = M_1^\perp \cap (-M_2^\perp) = M_1^\perp \cap M_2^\perp = \{0\}$.

Here the extremal principle's failure is also due to the behavior of the normal cones of the sets. The following example adapted from Example 2.10 shows that the extremal principle may even fail with two norm compact sets.

**Example 4.3.** Let $X$ be a separable Banach space and $\{e_n\}_{n=1}^\infty$ unit independent vectors that densely span $X$. Define $S_1 := \mathrm{cl}\,\mathrm{co}\{\frac{\pm e_n}{2^n}\}$ and $S_2 := \{0\}$. Then both $S_1$ and $S_2$ are norm compact. Let $v := (\sum_{n=1}^\infty \frac{e_n}{n^2}) \in X$. Note that for any sequence of nonzero real numbers $r_k \to 0$, $(r_k v + S_2) \cap S_1 = \{r_k v\} \cap S_1 = \emptyset$. Thus, 0 is an extremal point for $(S_1, S_2)$. However, the extremal principle does not holds at 0 because $N(S_1, 0) = \{0\}$.

## 5. The Open Mapping Theorem, metric regularity and the pseudo-lipschitz property

These are important concepts that lead to many applications and are the focus of much continuing research. In finite dimensional spaces Morduhovich [22] is a definitive paper that contains a thorough discussion of the equivalent relations of openness, metric regularity and Lipschitz properties of multifunctions as well as a complete characterization of those properties in terms of the limiting coderivatives. In this section we construct examples showing these characterizations fail in infinite dimensional spaces.

### 5.1. Open Mapping Theorem

First recall a finite dimensional form of the Open Mapping Theorem [22, Theorem 3.3]. A multifunction $F : X \to 2^Y$ is said to have an *open covering property* with linear rate around $(x, y) \in \text{graph}(F)$ provided that there exists a real number $a > 0$ and open neighborhood $U$ of $x$ and $V$ of $y$ such that, for any $x'$ and $r > 0$ satisfying $x' + rB_X \subset U$, we have $F(x') \cap V + arB_Y \subset F(x' + rB_X)$.

**Theorem 5.1.** *Let $X$ and $Y$ be finite dimensional Banach spaces and let $F : X \to 2^Y$ be locally bounded around $x$. Then $F$ has an open covering property with linear rate around $(x, y) \in \text{graph}(F)$ if and only if*

$$\operatorname{Ker} \partial^* F(x, y) := \{y^* \in Y^* : 0 \in \partial^* F(x, y)(y^*)\} = \{0\}.$$

**Example 5.2.** Let $X$ be any separable Banach space and $\{e_n\}_{n=1}^{\infty}$ unit independent vectors that densely span $X$. Define $S_1 := \operatorname{cl}\operatorname{co}\{\frac{\pm e_n}{2^n}\}$ and $S_2 := \{tv : t \in [-1, 1]\}$ where $v := (\sum_{n=1}^{\infty} \frac{e_n}{n^2}) \in X$. Then both $S_1$ and $S_2$ are norm compact and $S_1 \cap S_2 = S_1 \cap (-S_2) = \{0\}$. Define

$$F(x) := \begin{cases} x + S_1 & \text{if } x \in S_2 \\ \emptyset & \text{otherwise} \end{cases}$$

It is easy to see that $(0, 0) \in \text{graph}(F)$. Since $\text{span}(S_1)$ is dense in $X$ we have

$$N(\text{graph}(F), (0, 0)) \subset [\{0\} \times S_1]^{\perp} = X^* \times \{0\}.$$

Therefore, $\operatorname{Ker} \partial^* F(0; 0) = \{0\}$.

It remains to show that $F$ does not have an open covering property with linear rate around $(0, 0)$. In fact, for any $r > 0$,

$$F(rB_X) = \bigcup_{\gamma \in [0, r/\|v\|]} [\gamma v + S_1]$$

does not contain any open ball around 0. To see this let $u := (\sum_{n=1}^{\infty} \frac{e_n}{n^3})$ and $\alpha$ be an arbitrary positive number. Then $\alpha u \in F(rB_X)$ implies that, for some $\gamma \in [0, r/\|v\|]$, $\alpha u - \gamma v \in S_1$ as can only happen when $\alpha u = \gamma v = 0$.

### 5.2. Metric regularity

A multifunction $F : X \to 2^Y$ is called locally *metrically regular* around $(x, y) \in \text{graph}(F)$ if there exist neighborhoods $U$ of $x$ and $V$ of $y$ and constants $a, b$ such that $d_{F^{-1}(y)}(x) \leq a \cdot d_{F(x)}(y)$ for any $x \in U$ and $y \in V$ satisfying $d_{F(x)}(y) < b$.

By equivalence of metric regularity and openess with linear rate in [22] the same condition

$$\operatorname{Ker} \partial^* F(x, y) := \{y^* \in Y^* : 0 \in \partial^* F(x, y)(y^*)\} = \{0\}$$

also characterizes the local metric regularity of $F$ around $(x, y)$. Example 5.2 then also give us an example where the above characterization of local metric regularity of $F$ fails in infinite dimensional spaces.

**Example 5.3.** Let $X$ be separable Banach space and define $F$ as in Example 5.2. We have already shown that $\operatorname{Ker} \partial^* F(0;0) = \{0\}$. Now we verify that $F$ is not locally regular around $(0,0)$. In fact, it is easy to check that $F^{-1}(0) = \{0\}$. Consider $x := 2^{-k}v$. Then $d_{F(x)}(0) \geq \|2^{-k} \sum_{n=k+1}^{\infty} \frac{e_n}{n^2}\|$. Therefore,

$$\frac{d_{F^{-1}(0)}(x)}{d_{F(x)}(0)} \geq \frac{2^{-k}\|v\|}{\|2^{-k} \sum_{n=k+1}^{\infty} \frac{e_n}{n^2}\|} = \frac{\|v\|}{\| \sum_{n=k+1}^{\infty} \frac{e_n}{n^2}\|}.$$

When $k \to \infty$ the right hand side of the above inequality diverges to $+\infty$. Thus, the relation $d_{F^{-1}(0)}(x) \leq a \cdot d_{F(x)}(0)$ does not hold for any $a$.

The equivalence of metric regularity, the open mapping property and the pseudo-Lipschitz property for the inverse multifunction in [22] then implies that the characterization of the pseudo-Lipschitz property for multifunctions given in [22, Theorem 5.7] does not hold in infinite dimensional spaces without additional assumptions.

## 6. Generalizations

Many examples discussed in the previous sections can be constructed in more general Banach spaces and can be refined. In this section we briefly discuss some methods to do so. As promised in Remark 1.2 (c) we will construct examples for the Clarke generalized gradient.

### 6.1. Constructing examples in more general Banach spaces

Suppose $T$ is a bounded linear mapping from a Banach space $X$ onto a separable (quotient) Banach space $Y$. Then $T^*$ is an isomorphism from $Y^*$ into $X^*$. Consider a lower semicontinuous function $f : Y \to \bar{R}$. Define $g : X \to \bar{R}$ by $g(x) := f(Tx)$. Then it is easy to check that, $\partial_c g(x) = T^* \partial_c f(Tx)$ and $\partial_c^\infty g(x) = T^* \partial_c^\infty f(Tx)$. Using these relations all the examples constructed in $Y$ can be lifted to $X$. We illustrate by embedding Example 2.2 in a Banach space with a separable quotient.

**Example 6.1.** Let $X$ be a Banach space with a separable quotient $Y$, i.e., there exists a linear bounded quotient mapping $T$ from $X$ onto a separable Banach space $Y$. As in Example 2.2 we construct closed subspaces $M_1$ and $M_2$ in $Y$ such that $M_1^\perp + M_2^\perp$ is $w^*$-dense in $Y^*$ but not closed and $M_1^\perp \cap M_2^\perp = \{0\}$. Define $f_1(x) := \delta_{M_1}(Tx)$ and $f_2(x) := \delta_{M_2}(Tx) + \langle v, Tx \rangle$ where $-v \in Y^* \backslash (M_1^\perp + M_2^\perp)$ and, therefore, $-T^*v \in X^* \backslash (T^* M_1^\perp + T^* M_2^\perp)$. Since $M_1 \cap M_2 = \{0\}$, $f_1 + f_2$ attains a minimum at $x = 0$. However, by an open mapping argument, $\partial f_1(0) = \partial^\infty f_1(0) = T^* M_1^\perp$, $\partial f_2(0) = T^* M_2^\perp + T^* v$ and $\partial^\infty f_2(0) = T^* M_2^\perp$. Thus, $0 \notin \partial f_1(0) + \partial f_2(0)$ and $\partial^\infty f_1(0) \cap (-\partial^\infty f_2(0)) = \{0\}$ holds only in the trivial case.

In particular, all *weakly compactly generated* Banach spaces, $l_\infty$ and $L_\infty$ have separable quotients [16] and, therefore, all the examples in the previous sections can be constructed in them.

### 6.2. Constructing continuous examples

In Example 2.2 we use the indicator function $\delta_{M_i}, i = 1, 2$ of the closed subspace $M_i$. These indicator functions are lower semicontinuous extended valued convex functions.

We will show that we can in fact trade convexity for continuity. The trick is to replace the indicator function $\delta_M$ by a power of the distance function, $d^\gamma_M(x)$, for $\gamma \in (0,1)$.

**Example 6.2.** As in Example 2.2 let $X$ be a separable Banach space and let $M_1$ and $M_2$ be closed subspaces with $M_1^\perp \cap M_2^\perp = \{0\}$ and $M_1^\perp + M_2^\perp$ $w^*$-dense in $X^*$ but not closed. Define $f_1 := d^\gamma_{M_1}$ and $f_2 := d^\gamma_{M_2} - \langle v, \cdot\rangle$ where $v \in X^* \backslash (M_1^\perp + M_2^\perp)$ and $\gamma \in (0,1)$. Then the restriction of $f_1 + f_2$ on any finite dimensional subspaces $L$ of $X$ attains a minimum of 0 at $x = 0$. In fact, assume the contrary that there exists a sequence $x_n \in L$ converges to 0 with $f_1(x_n) + f_2(x_n) < 0$, i.e., $0 < d^\gamma_{M_1}(x_n) + d^\gamma_{M_2}(x_n) < \langle v, x_n\rangle$. Since $d^\gamma_{M_i}, i = 1, 2$ are $\gamma$-homogeneous we can rewrite this inequality as

$$0 < d^\gamma_{M_1}(\frac{x_n}{\|x_n\|}) + d^\gamma_{M_2}(\frac{x_n}{\|x_n\|}) < \|x_n\|^{1-\gamma}\langle v, \frac{x_n}{\|x_n\|}\rangle.$$

The right hand side of the inequality goes to 0 as $n \to \infty$. However,

$$d^\gamma_{M_1}(\frac{x_n}{\|x_n\|}) + d^\gamma_{M_2}(\frac{x_n}{\|x_n\|}) \geq \inf\{d^\gamma_{M_1}(u) + d^\gamma_{M_2}(u) : u \in L, \|u\| = 1\} > 0.$$

This is absurd. The $g$-subdifferential characterization of the Clarke generalized gradient [3, 11] then yields $0 \in \partial_c(f_1 + f_2)(0)$.

Observing that $d^\gamma_{M_i}(x) \leq \delta_{M_i}(x)$ we have $\mathrm{epi}(\delta_{M_i}) \subset \mathrm{epi}(d^\gamma_{M_i})$. Therefore $N_c(\mathrm{epi}(d^\gamma_{M_i}), (0,0)) \subset M_i^\perp \times (-R+)$. Thus, $\partial_c f_1(0) \subset M_1^\perp$, $\partial^\infty_c f_1(0) \subset M_1^\perp$, $\partial_c f_2(0) \subset M_2^\perp - v$ and $\partial^\infty_c f_2(0) \subset M_2^\perp$. As in Example 2.2 we can see that the sum rule does not hold.

**Remark 6.3.** In fact the relation between $\partial_c d^\gamma_{M_i}(0)$ and $N_c(M_i, 0)$ that we used in the above example is a special case of the following more general result which is interesting in itself.

**Lemma 6.4 (Nonconvex penalization).** *Let $X$ be a Banach space, let $C$ be a closed subset of $X$ and $x \in C$. Then, for any $\gamma \in (0,1)$,*

$$\partial_c d^\gamma_C(x) = \partial^\infty_c d^\gamma_C(x) = N_c(C, x).$$

**Proof.** For any $\lambda > 0$ there exists a neighborhood $U$ of $x$ such that, for all $y \in U$, $\lambda d_C(y) \leq d^\gamma_C(y) \leq \delta_C(y)$, i.e., $\mathrm{epi}(\delta_{C\cap U}) \subset \mathrm{epi}(d^\gamma_{C\cap U}) \subset \mathrm{epi}(\lambda d_{C\cap U})$. Therefore, $N_c(\mathrm{epi}(\lambda d_C), (x,0)) \subset N_c(\mathrm{epi}(d^\gamma_C), (x,0)) \subset N_c(\mathrm{epi}(\delta_C), (x,0))$. Since $\lambda$ is arbitrary we obtain $N_c(\mathrm{epi}(d^\gamma_C), (x,0)) = N_c(\mathrm{epi}(\delta_C), (x,0)) = N_c(C, x) \times (-R^+)$, which completes the proof. $\square$

We use the following concrete example to illustrate the construction discussed in Example 6.2.

**Example 6.5.** Let $H := \ell_2$ and denote the unit vectors by $\{u_n\}$. Suppose $\{\alpha_n\}$ is a sequence of positive real numbers with $1 > \alpha_n \geq \sqrt{1 - \frac{1}{n^2}}$. Set $e_n := u_{2n-1}$, $f_n := \alpha_n u_{2n-1} + \sqrt{1 - \alpha_n^2}\, u_{2n}$. Define

$$M_1 := \mathrm{cl\,span}\{e_1, e_2, \dots\} \text{ and } M_2 := \mathrm{cl\,span}\{f_1, f_2, \dots\}.$$

Then, for any $x = \sum_{n=1}^{\infty} x_n u_n \in H$, the partial sum

$$\sum_{n=1}^{2N} x_n u_n = \sum_{n=1}^{N}(x_{2n-1} - \frac{x_{2n}\alpha_n}{\sqrt{1-\alpha_n^2}})e_n + \sum_{n=1}^{N} \frac{x_{2n}}{\sqrt{1-\alpha_n^2}} f_n \in M_1 + M_2.$$

Therefore, $M_1 + M_2$ is dense in $H$. This implies that $M_1^{\perp} \cap M_2^{\perp} = \{0\}$. Noting that $M_1^{\perp} = \mathrm{cl}\ \ \mathrm{span}\{h_1, h_2, ...\}$ and $M_2^{\perp} = \mathrm{cl}\ \ \mathrm{span}\{g_1, g_2, ...\}$ where $h_n := u_{2n}$ and $g_n := \sqrt{1-\alpha_n^2}u_{2n-1} - \alpha_n u_{2n}$, we can show by a similar argument that $M_1^{\perp} + M_2^{\perp}$ is dense in $H$ which implies that $M_1 \cap M_2 = 0$. It remains to show that $M_1 + M_2 \neq H$. Consider

$$v := \sum_{n=1}^{\infty} \sqrt{1-\alpha_n^2}u_{2n}.$$

If $v = y + z$ with $y \in M_1$ and $z \in M_2$ then $y = \sum_{n=1}^{\infty} y_n e_n$ and $z = \sum_{n=1}^{\infty} z_n f_n$ because $\{e_n\}$ and $\{f_n\}$ are orthonormal basis for $M_1$ and $M_2$ respectively. Then we must have $z_n = 1$ and $y_n = z_n\alpha_n = \alpha_n \to 1$ which is impossible.

Fix $0 < \gamma < 1$. For arbitrary $x = \sum_{n=1}^{\infty} x_n u_n \in H$, the explicit form of the functions in Example 6.2 is

$$f_1(x) := \big[\sum_{n=1}^{\infty} x_{2n}^2\big]^{\gamma/2}$$

and

$$f_2(x) := \big[\sum_{n=1}^{\infty}(\sqrt{1-\alpha_n^2}x_{2n-1} - \alpha_n x_{2n})^2\big]^{\gamma/2} - \sum_{n=1}^{\infty} \sqrt{1-\alpha_n^2}x_{2n}.$$

By picking $\gamma$ arbitrarily close to 1, we obtain examples that are Hölder continuous of any desired modulus less than one, but of course not Lipschitz (modulus one). This method also applies to Example 2.7.

## References


[1] J. M. Borwein: Adjoint process duality, Math. of Oper. Res. 8 (1983) 403–434.

[2] J. M. Borwein: Epi-Lipschitz like sets in Banach space: Theorems and examples, Nonlinear Analysis: TMA 10 (1987) 1207–1217.

[3] J. M. Borwein, A. D. Ioffe: Proximal analysis in smooth spaces, CECM Research Report 93-04 (1993), Set-Valued Analysis 4 (1996) 1–24.

[4] J. M. Borwein, H. M. Strojwas: Proximal analysis and boundaries of closed sets in Banach space, Part I: Theory, Can. J. Math. 38 (1986) 431–452; Part II: Applications, Can. J. Math. 39 (1987) 428–472.

[5] J. M. Borwein, J. S. Treiman, Q. J. Zhu: Necessary conditions for constrained optimization problems with semicontinuous and continuous data, to appear in: Trans. Amer. Math. Soc.

[6] F. H. Clarke: Generalized gradients and applications, Trans. Amer. Math. Soc. 344 (1975) 307–324.

[7] F. H. Clarke: Optimization and Nonsmooth Analysis, John Wiley & Sons, Inc., New York, 1983.

[8] F. H. Clarke: Methods of dynamic and nonsmooth optimization, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1989.

[9] A. Y. Kruger, B. S. Mordukhovich: Extremal points and the Euler equation in nonsmooth optimization, Dokl. Akad. Nauk BSSR 24 (1980) 684–687.

[10] A. D. Ioffe: Necessary conditions in nonsmooth optimization, Math. Oper. Res. 9 (1984) 159–189.

[11] A. D. Ioffe: Approximate subdifferentials and applications, Part 3, Mathematica 36 (1989) 1–38.

[12] A. D. Ioffe: Proximal analysis and approximate subdifferentials, J. London Math. Soc. 41 (1990) 175–192.

[13] A. D. Ioffe: Non-smooth subdifferentials: their calculus and applications, Proceedings of the First World Congress of Nonlinear Analysis, Tempa, Florida, August 19-26, 1992, Walter de Gruyter, Berlin, New York (1996) 2299–2310.

[14] A. D. Ioffe: Codirectional compactness, metric regularity and subdifferential calculus, preprint, 1996.

[15] A. Jourani, L. Thibault: Qualification conditions for calculus rules of coderivatives of multivalued mappings, preprint, 1996.

[16] J. Lindenstrauss, L. Tzafriri: Classical Banach Spaces I: Sequence Spaces, Springer-Verlag, Berlin, 1979.

[17] P. D. Loewen: Optimal Control via Nonsmooth Analysis, CRM Lecture Notes Series, Amer. Math. Soc., Summer School on Control, CRM, Université de Montréal, 1992, Amer. Math. Soc., Providence, 1993.

[18] P. D. Loewen: Limits of Fréchet normals in nonsmooth analysis, In: A. D. Ioffe et al. (eds.): Optimization and Nonlinear Analysis, Pitman Research Notes in Math. Series 244, Longman, Harlow, Essex (1992) 178–188.

[19] B. S. Mordukhovich: Maximum principle in problems of time optimal control with nonsmooth constraints, J. Appl. Math. Mech. 40 (1976) 960–969.

[20] B. S. Mordukhovich: Metric approximations and necessary optimality conditions for general classes of nonsmooth extremal problems, Soviet Math. Dokl. 22 (1980) 526–530.

[21] B. S. Mordukhovich: Approximation Methods in Problems of Optimization and Control, Nauk, Moscow, 1988. (English transl., Wiley/Interscience, to appear)

[22] B. S. Mordukhovich: Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions, Trans. Amer. Math. Soc. 340 (1993) 1–35.

[23] B. S. Mordukhovich: Generalized differential calculus for nonsmooth and set-valued mappings, J. Math. Anal. Appl. 183 (1994) 250–288.

[24] B. S. Mordukhovich, Y. Shao: Nonsmooth sequential analysis in Asplund spaces, Trans. Amer. Math. Soc. 348 (1996) 1235–1280.

[25] B. S. Mordukhovich, Y. Shao: Nonconvex differential calculus for infinite-dimensional multifunctions, Set-Valued Analysis 4 (1996) 205–236.

[26] R. T. Rockafellar: Proximal subgradients, marginal values and augmented Lagrangians in nonconvex optimization, Math. Oper. Res. 6 (1981) 424–436.

[27] J. S. Treiman: Clarke's gradients and epsilon-subgradients in Banach spaces, Trans. Amer. Math. Soc. 294 (1986) 65–78.

[28] J. Warga: Fat homeomorphisms and unbounded derivate containers, J. Math. Anal. Appl. 81 (1981) 545–560; (Errata 90 (1982) 582–583).

[29] Q. J. Zhu: Subderivatives and applications, Proceedings of International Conference on Dynamical Systems and Differential Equations, Springfield, MO, May 29–June 1, 1996.