Nonsmooth analysis: Fréchet subdifferentials

*Alexander Y. Kruger*
Centre for Informatics and Applied Optimization
School of Information Technology and Mathematical Sciences
University of Ballarat
POB 663, Ballarat, Vic 3350, Australia
*E-mail address: a.kruger@ballarat.edu.au*
*URL: http://uob-community.ballarat.edu.au/~akruger*

# Introduction

Being natural generalizations of the classical Fréchet derivative and the subdifferential in the sense of convex analysis, *Fréchet subdifferentials* have been known for more than thirty years. They were probably first introduced in finite dimensions in [1] (under the name "lower semidifferentials"). Some of their properties in the infinite dimensional setting were investigated in [18, 21]. During the first decade Fréchet subdifferentials were not widely used because of rather poor (direct) calculus. They mostly served as building blocks for more sophisticated *limiting (Fréchet) subdifferentials* (see [19, 23, 30, 31]).

The discovery of the "fuzzy rules" in the 1980-es [11, 12, 15, 29] re-vitalized the interest to Fréchet subdifferentials. It was shown that calculus results and optimality conditions can be formulated in terms of Fréchet and other "simple" subdifferentials computed not at the given point, but at some points arbitrarily close to it, thus incorporating "differential" properties of the function at nearby points. Such results are actually at the core of the corresponding statements for limiting subdifferentials.

Being the smallest among all "simple" subdifferentials with reasonable properties, the Fréchet subdifferentials have proved to be convenient tools for the analysis of nondifferentiable functions on *Asplund* spaces, a very important subclass of general Banach spaces. Furthermore, the main fuzzy results in terms of Fréchet subdifferentials present characterizations of Asplund spaces themselves (see [6, 12, 13, 34, 35, 38, 45]).

The article contains no proofs. A more detailed survey of Fréchet subdifferentials can be found in [25].

Mostly standard notations are used throughout the article. $X, Y$ denote normed linear spaces and $X^*$, $Y^*$ denote their topological duals. $\langle \cdot, \cdot \rangle$ is a bilinear form defining a canonical paring between a space and its dual. $B_\rho(x)$ stands for a closed ball with center $x$ and radius $\rho$. We write $B_\rho$ instead of $B_\rho(0)$ and just $B$ if $\rho = 1$ (unit ball).

# Definitions

## Fréchet subdifferentials

Let $f : X \to \mathbb{R}_\infty = \mathbb{R} \cup \{+\infty\}$, $f(x) < \infty$. The set

$$\partial f(x) = \left\{ x^* \in X^* : \ \liminf_{u \to x} \frac{f(u) - f(x) - \langle x^*, u - x \rangle}{\|u - x\|} \geq 0 \right\} \qquad (1)$$

is called the *Fréchet subdifferential* of $f$ at $x$. This set is (norm) closed and convex. The next proposition shows that it generalizes the notions of the Fréchet derivative and the subdifferential in the sense of convex analysis.

**Proposition 1.** *(i) If $f$ is Fréchet differentiable at $x$ with the derivative $\nabla f(x)$ then $\partial f(x) = \{\nabla f(x)\}$.*

*(ii) If $f$ is convex then*

$$\partial f(x) = \{x^* \in X^* : \ f(u) - f(x) \geq \langle x^*, u - x \rangle, \forall u \in X\}.$$

Note that the Fréchet subdifferential does not change if another equivalent norm on $X$ is used in (1).

*Example* 1. The set (1) can be empty. Take $f : \mathbb{R} \to \mathbb{R} : f(u) = -|u|$, $u \in \mathbb{R}$.

One can also consider the *Fréchet superdifferential*

$$\partial^+ f(x) = \left\{ x^* \in X^* : \ \limsup_{u \to x} \frac{f(u) - f(x) - \langle x^*, u - x \rangle}{\|u - x\|} \leq 0 \right\}. \qquad (2)$$

While the set (1) consists of linear continuous functionals "supporting" $f$ from below, the functionals from (2) "support" $f$ from above. Unlike the classical case, the existence of two different derivative-like objects is quite natural for nonsmooth analysis: "differential" properties of a function "from below" and "from above" could be essentially different.

Sub- and superdifferentials are related by the equality

$$\partial(-f)(x) = -\partial^+ f(x). \qquad (3)$$

Surely, in the nondifferentiable case at least one of the sets (1) and (2) must be empty.

**Proposition 2.** *$\partial f(x) \neq \emptyset$ and $\partial^+ f(x) \neq \emptyset$ if and only if $f$ is Fréchet differentiable at $x$. In this case one has $\partial f(x) = \partial^+ f(x) = \{\nabla f(x)\}$.*

*Example* 2. Both sets (1) and (2) can be empty simultaneously. Take $f : \mathbb{R} \to \mathbb{R} : f(u) = u \sin(1/u)$ if $u \neq 0$, and $f(0) = 0$.

*Example* 3. The fact that the set (1) is a singleton does not imply differentiability. Take $f : \mathbb{R} \to \mathbb{R} : f(u) = \max(u \sin(1/u), 0)$ if $u \neq 0$, and $f(0) = 0$. Then $f$ is nondifferentiable at 0, although one evidently has $\partial f(0) = \{0\}$.

*Example* 4. Fréchet differentiability is essential in Proposition 1 (i) . Gâteaux differentiable functions can be non-subdifferentiable in the Fréchet sense. Take $f : \mathbb{R}^2 \to \mathbb{R} : f(u_1, u_2) = -\sqrt{|u_1|^2 + |u_2|^2}$ if $u_2 = u_1^2$, $f(u_1, u_2) = 0$ otherwise. The Gâteaux derivative of $f$ is 0, while $\partial f(0) = \emptyset$.

*Remark* 1. One can define the *Gâteaux subdifferential* based on the notion of the Gâteaux differentiability. For this subdifferential, the analogs of Propositions 1 and 2 and some other results hold true. Considering Gâteaux (and other types of) "simple" subdifferentials can be useful in some applications. The Gâteaux subdifferential always contains the Fréchet subdifferential.

If $\dim X < \infty$ the Fréchet subdifferential can be expressed equivalently in terms of certain generalized directional derivatives (see [1, 16, 18, 39, 42, 43]).

## Fréchet Normal Cone

Now consider a set $\Omega \subset X$ and let $x \in \Omega$. Similarly to the definition (1) of the Fréchet subdifferential one can define the *Fréchet normal cone*

$$N(x|\Omega) = \left\{ x^* \in X^* : \limsup_{u \overset{\Omega}{\to} x} \frac{\langle x^*, u - x \rangle}{\|u - x\|} \leq 0 \right\} \tag{4}$$

to $\Omega$ at $x$. Here $u \overset{\Omega}{\to} x$ means that $u \to x$ with $u \in \Omega$.

It is a norm closed and convex cone closely related to the subdifferential defined above. It is actually the Fréchet subdifferential $\partial \delta_\Omega(x)$ of the indicator function $\delta_\Omega$ of $\Omega$ ($\delta_\Omega(u) = 0$ if $u \in \Omega$ and $\delta_\Omega(u) = \infty$ otherwise).

This fact allows one to deduce some properties of normal cones from the corresponding statements about subdifferentials. Thus, it follows from Proposition 1 (ii), that the normal cone (4) generalizes the corresponding notion of convex analysis.

In finite dimensions the Fréchet normal cone coincides with the polar of the *tangent* (*contingent*, *Bouligand tangent*) cone to $\Omega$ at $x$. If $X$ is reflexive it coincides with the polar of the *weak tangent* cone (see [1, 2, 9, 14, 43]).

The relationship between Fréchet subdifferentials and normal cones is bilateral. For a function $f : X \to \mathbb{R}_\infty$ one can consider its *epigraph* $\operatorname{epi} f = \{(u, \mu) \in X \times \mathbb{R} : f(u) \leq \mu\}$. If the norm on $X \times \mathbb{R}$ is compatible with that on $X$ (that is $\|(x, 0)\| = \|x\|$), then the following equivalent definition of the Fréchet subdifferential holds true:

$$\partial f(x) = \{x^* \in X^* : (x^*, -1) \in N(x, f(x)|\operatorname{epi} f)\}. \tag{5}$$

"Horizontal" normals to the epigraph can also be of interest. They define the *singular Fréchet subdifferential* of $f$ at $x$:

$$\partial^\infty f(x) = \{x^* \in X^* : (x^*, 0) \in N(x, f(x)|\operatorname{epi} f)\}.$$

Of course, if $f$ is *calm* at $x$ (see [7, 42]), that is $\|f(u) - f(x)\| \le l\|u - x\|$ for some $l > 0$ and for all $u$ in a neighborhood of $x$, then the latter set is empty.

## Strict Fréchet $\delta$-Subdifferentials

As it was mentioned in Introduction, Fréchet subdifferentials have poor calculus and their direct application has been rather limited. There exists a way of enriching the properties of the subdifferentials. It consists in considering differential properties of a function at nearby points.

Consider a new derivative-like object based on the Fréchet subdifferential:

$$\hat\partial_\delta f(x) \;=\; \bigcup_{\substack{u\in B_\delta(x)\\ |\operatorname{cl} f(u)-f(x)|\le\delta}} \partial(\operatorname{cl} f)(u). \tag{6}$$

It depends on some positive $\delta$. $\operatorname{cl} f$ denotes here the lower semicontinuous envelope of $f$ (its epigraph is the closure of the epigraph of $f$ in $X \times \mathbb{R}$). Unlike (1) the set (6) can be nonconvex. It is called the *strict Fréchet $\delta$-subdifferential* of $f$ at $x$ [24].

The *strict Fréchet $\delta$-superdifferential* $\hat\partial^+_\delta f(x)$ of $f$ at $x$ can be defined in a similar way. The equality $\hat\partial^+_\delta f(x) = -\hat\partial_\delta(-f)(x)$ holds true. The strict sub- and superdifferentials can be nonempty simultaneously and can be essentially different. The set $\hat\partial^0_\delta\varphi(x) = \hat\partial_\delta f(x) \cup \hat\partial^+_\delta f(x)$ can be useful in some situations. It is called the *strict Fréchet $\delta$-differential* of $f$ at $x$.

The *strict Fréchet $\delta$-normal cone* to a set $\Omega$ at $x \in \Omega$ is defined similarly:

$$\hat N_\delta(x|\Omega) \;=\; \bigcup_{u\in\operatorname{cl}\Omega\cap B_\delta(x)} N(u|\operatorname{cl}\Omega).$$

The goal of introducing strict Fréchet $\delta$-subdifferentials is mainly notational. They are convenient for formulating "fuzzy" results, but such results can certainly be formulated in terms of ordinary Fréchet subdifferentials.

## Limiting Subdifferentials

The limiting Fréchet subdifferentials are defined as limits of "simple" ones (see [23, 30, 31, 34]). To simplify the definitions we assume in this subsection that $f : X \to \mathbb{R}_\infty$ is lower semicontinuous in a neighborhood of $x$.

The *limiting Fréchet subdifferential* of $f$ at $x$ is defined as

$$\bar{\partial} f(x) = \{x^* \in X^* : \exists \text{ sequences } \{x_k\} \subset X, \{x_k^*\} \subset X^* \text{ such that } x_k \xrightarrow{f} x,\ x_k^* \xrightarrow{w^*} x^* \text{ and } x_k^* \in \partial f(x_k),\ k = 1, 2, \ldots\} \quad (7)$$

The notations $x_k \xrightarrow{f} x$ and $x_k^* \xrightarrow{w^*} x^*$ hear mean respectively that $x_k \to x$ with $f(x_k) \to f(x)$ ($f$-*attentive convergence* [42]), and $x_k^*$ converges to $x^*$ in the *weak*$^*$ *topology* of $X^*$.

$\bar{\partial} f(x)$ is a weakly$^*$ sequentially closed set in $X^*$. In general it is nonconvex. If $f$ is strictly differentiable at $x$ the set (7) reduces to the derivative.

Using strict $\delta$-subdifferentials one can rewrite (7) in the following way:

$$\bar{\partial} f(x) = \bigcap_{\delta > 0} \mathrm{cl}^*\ \hat{\partial}_\delta f(x),$$

where $\mathrm{cl}^*$ denotes the weak$^*$ sequential closure.

Other limiting objects (the *limiting superdifferential*, the *limiting differential*, the *limiting normal cone*, the *singular limiting subdifferential*, and the *limiting coderivative*) can be defined in a similar way.

Thus the limiting normal cone to a closed set $\Omega$ is defined by the equality

$$\bar{N}(x|\Omega) = \bigcap_{\delta > 0} \mathrm{cl}^*\ \hat{N}_\delta(x|\Omega).$$

It coincides with the limiting subdifferential of the indicator function of $\Omega$. The analog of (5) is also valid:

$$\bar{\partial} f(x) = \{x^* \in X^* : (x^*, -1) \in \bar{N}(x, f(x)|\mathrm{epi}\, f)\}.$$

The limiting subdifferentials and normal cones have been well investigated. They possess good calculus (which is the consequence of the fuzzy calculus of Fréchet subdifferentials; see [19, 23, 30, 31, 32, 34, 36, 42] for the properties of these objects and some examples). They have proved to be very efficient for formulating optimality conditions in nonsmooth optimization (see [20, 22, 29, 30, 31, 32, 34, 36]), especially in finite dimensions. When applying limiting subdifferentials in infinite dimensional spaces one must be careful about nontriviality of the limits in the weak$^*$ topology. Additional regularity conditions are needed (*compact epi-Lipschitzness*, *sequential normal compactness*, *partial sequential normal compactness* [4, 34, 36], etc.)

### Fréchet $\varepsilon$-subdifferentials and $\varepsilon$-normals

In some cases it can be convenient to use $\varepsilon$-extensions of the Fréchet subdifferentials and normal cones (see [21, 23, 29, 44]). For instance, the *Fréchet $\varepsilon$-subdifferential* and the *Fréchet $\varepsilon$-superdifferential* of $f$ at $x$ are defined as

$$\partial_\varepsilon f(x) = \left\{ x^* \in X^* : \liminf_{u \to x} \frac{f(u) - f(x) - \langle x^*, u - x \rangle}{\|u - x\|} \geq -\varepsilon \right\},$$

$$\partial_\varepsilon^+ f(x) = \left\{ x^* \in X^* : \limsup_{u \to x} \frac{f(u) - f(x) - \langle x^*, u - x \rangle}{\|u - x\|} \leq \varepsilon \right\}.$$

Unlike (1) and (2), these sets depend on the specific norm on $X$ (when $\varepsilon > 0$).

The next two propositions extend Propositions 1 (ii) and 2 respectively.

**Proposition 3.** *If $f$ is convex then $\partial_\varepsilon f(x) = \partial f(x) + \varepsilon B^* = \{x^* \in X^* : f(u) - f(x) \geq \langle x^*, u - x \rangle - \varepsilon \|u - x\|, \ \forall u \in X\}$.*

*Remark* 2. Note that the above $\varepsilon$-subdifferential differs from the corresponding notinon of convex analysis, which is usually defined [41] as the set of all $x^* \in X^*$, such that $f(u) - f(x) \geq \langle x^*, u - x \rangle - \varepsilon$ for all $u \in X$.

**Proposition 4.** *If $x_1^* \in \partial_{\varepsilon_1} f(x), \ x_2^* \in \partial_{\varepsilon_2}^+ f(x), \ \varepsilon_1 \geq 0, \ \varepsilon_2 \geq 0$ then $\|x_1^* - x_2^*\| \leq \varepsilon_1 + \varepsilon_2$.*

# Formulation

## Direct Calculus

The propositions below present some simple calculus results for Fréchet subdifferentials. Most of them follow directly from the definitions. More advanced statements of fuzzy calculus are presented in the next subsections.

**Proposition 5.** *If $f$ attains a local minimum at $x$ then $0 \in \partial f(x)$.*

**Proposition 6.** *$\partial(\lambda f)(x) = \lambda \partial f(x)$ for any $\lambda > 0$.*

**Proposition 7.** *Let $f_1, f_2 : X \to \mathbb{R}_\infty$ be finite at $x$. Then*

$$\partial(f_1 + f_2)(x) \supset \partial f_1(x) + \partial f_2(x). \tag{8}$$

The above proposition presents an example of a *Sum Rule*. Usually the Sum Rule is the central result of any subdifferential calculus. Unfortunately, the inclusion (8) is almost useless: it does not allow to decompose elements of the subdifferential of the sum of functions in terms of elements of subdifferentials of the original functions. Simple examples show that inclusion (8) can be strict even in the convex case. The next proposition gives two important cases when the equality holds true in (8).

**Proposition 8.** *Let $f_1, f_2 : X \to \mathbb{R}_\infty$ be finite at $x$.*

*(i) If $f_1$ and $f_2$ are convex and one of them is continuous at $x$ then*

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x);$$

*(ii) If $f_1$ is Fréchet differentiable at $x$ then*

$$\partial(f_1 + f_2)(x) = \nabla f_1(x) + \partial f_2(x). \tag{9}$$

Part (i) of Proposition 8 is known as Moreau-Rockafellar theorem (see [41]). Part (ii) is a simple corollary of Proposition 7. It is an interesting example, when an inclusion implies an equality. Indeed, applying Proposition 7 to the sum of the functions $f_1 + f_2$ and $-f_2$ and making use of (3) one gets

$$\partial f_2(x) \supset \partial(f_1 + f_2)(x) - \partial^+ f_1(x). \tag{10}$$

Taking into account Proposition 2, the inclusions (8) and (10) imply (9).

Proposition 8 (ii) yields a simple necessary optimality condition generalizing Proposition 5.

**Proposition 9.** *Let $f_1 : X \to \mathbb{R}$ be Fréchet differentiable at $x$ where $f_2 : X \to \mathbb{R}_\infty$ is finite. If $f_1 + f_2$ attains a local minimum at $x$ then $-\nabla f_1(x) \subset \partial f_2(x)$.*

The next two assertions are corollaries of Propositions 7 and 9.

**Proposition 10.** *Let $x \in \Omega_1 \cap \Omega_2$. Then $N(x|\Omega_1 \cap \Omega_2) \supset N(x|\Omega_1) + N(x|\Omega_2)$.*

**Proposition 11.** *Let $f$ be Fréchet differentiable at $x$. If $f$ attains at $x$ a local minimum on $\Omega$ then $-\nabla f(x) \in N(x|\Omega)$.*

For a set $\Omega \subset X$ and $u \in X$ consider the distance $d_\Omega(u) = \inf_{\omega \in \Omega} \|u - \omega\|$.

**Proposition 12** ([18])**.** *For any $x \in \Omega$ one has $\partial d_\Omega(x) = \{x^* \in N(x|\Omega) : \|x^*\| \le 1\}$.*

## Strict Differentiability

Recall that $f$ is called *strictly differentiable* (see [34, 42]) at $x$ (with the strict derivative $\nabla f(x)$) if

$$\lim_{\substack{u \to x,\ u' \to x \\ u \neq u'}} \frac{f(u') - f(u) - \langle \nabla f(x), u' - u \rangle}{\|u' - u\|} = 0.$$

In the case of a strictly differentiable function the Fréchet sub- and superdifferentials at nearby points cannot differ much from the strict derivative.

**Proposition 13** ([25]). *If $f$ is strictly differentiable at $x$ with the derivative $\nabla f(x)$ then for any $\varepsilon > 0$ there exists a $\delta > 0$ such that*
*(i) $\hat{\partial}^0_\delta f(x) \subset \nabla f(x) + \varepsilon B^*$;*
*(ii) $\nabla f(x) \in \partial_\varepsilon f(u) \cap \partial^+_\varepsilon f(u)$ for all $u \in B_\delta(x)$.*

The rest of the section is devoted to "fuzzy" results in terms of Fréchet subdifferentials and strict Fréchet $\delta$-differentials.

## Variational Principles

The *variational principles* by Ekeland [10], Borwein and Preiss [3] as well as their subsequent followers (see [6, 8, 34, 40, 42]) are very powerful tools of modern variational analysis. They make it possible to substitute an "almost minimal" point (up to $\varepsilon$) by another point, arbitrarily close to the initial one, which is the local minimizer for a slightly perturbed (usually by adding a small term) function. Thus, such principles can be viewed as fuzzy results.

The next assertion is valid for an arbitrary Asplund space. It is known as the *Subdifferential Variational Principle*. Let us recall that a Banach space is called *Asplund* (see [6, 8, 34, 40]) if any continuous convex function on it is Fréchet differentiable on a dense $G_\delta$ set of points. Asplund spaces form a rather broad subclass of Banach spaces. It includes for instance all spaces which admit Fréchet differentiable bump functions (in particular, Fréchet smooth spaces). Reflexive spaces are examples of Fréchet smooth spaces.

Asplund spaces provide a very convenient framework for investigating "differential" properties of nonsmooth functions. Actually the Asplund property of a Banach space is not only sufficient but also a necessary condition for the fulfillment of some basic results in nonsmooth analysis involving Fréchet normals and subdifferentials (see [6, 13, 34, 35, 38] and the statements below).

**Theorem 1** (Mordukhovich and Wang [38]). *Let $X$ be Asplund, $f : X \to \mathbb{R}_\infty$ be lower semicontinuous and bounded below, $\varepsilon > 0$, $\lambda > 0$. Suppose that $f(x) < \inf f + \varepsilon$. Then there exists a $u \in B_\lambda(x)$ and an $x^* \in \partial f(u)$, such that $f(u) \le f(x)$ and $\|x^*\|_* < \varepsilon/\lambda$.*

The following theorem states that the class of Fréchet *subdifferentiability spaces* [15] coincides with Asplund spaces.

**Theorem 2.** *The following assertions are equivalent:*
*(i) $X$ is an Asplund space;*
*(ii) for any lower semicontinuous function $f : X \to \mathbb{R}_\infty$ the set $\{u \in X : \partial f(u) \neq \emptyset\}$ is dense in* $\operatorname{dom} f$*;*
*(iii) for any lower semicontinuous function $f : X \to \mathbb{R}_\infty$ there exists an $x \in \operatorname{dom} f$ such that $\partial f(x) \neq \emptyset$.*

## Sum Rules

After the Sum Rule was first established in the limiting form in [19] (see [23, 31]) the fuzzy versions were derived in [12, 15] (see also [6, 17, 37]). Now two main versions of the fuzzy Sum Rule are known. For simplicity they are formulated below in terms of strict $\delta$-subdifferentials.

**Rule 1** (Weak Fuzzy Sum Rule). *Let $f_1, f_2 : X \to \mathbb{R}_\infty$ be finite at $x$ and lower semicontinuous near $x$. Then $\partial(f_1 + f_2)(x) \subset \hat{\partial}_\delta f_1(x) + \hat{\partial}_\delta f_2(x) + U^*$ for any $\delta > 0$ and any weak$^*$ neighborhood $U^*$ of $0$ in $X^*$.*

**Rule 2** (Strong Fuzzy Sum Rule). *Let $f_1 : X \to \mathbb{R}_\infty$ be finite at $x$ and lower semicontinuous near $x$ and $f_2 : X \to \mathbb{R}$ be Lipschitz continuous near $x$. Then $\partial(f_1 + f_2)(x) \subset \hat{\partial}_\delta f_1(x) + \hat{\partial}_\delta f_2(x) + \delta B^*$ for any $\delta > 0$.*

A Banach space is called a *trustworthy space* [15] (for some kind of a subdifferential) if Rule 1 is valid in it. The following theorem proved by Fabian [12] states that for the Fréchet subdifferential the class of trustworthy spaces coincides with Asplund spaces.

**Theorem 3.** *The following assertions are equivalent:*
*(i) $X$ is an Asplund space;*
*(ii) Weak Fuzzy Sum Rule is valid in $X$;*
*(iii) Strong Fuzzy Sum Rule is valid in $X$.*

Weak Fuzzy Sum Rule yields the following representation of Fréchet normals to the intersection of closed sets.

**Proposition 14.** *Let $\Omega_1$, $\Omega_2$ be closed subsets in an Asplund space $X$ and $x \in \Omega_1 \cap \Omega_2$. Then $N(x, \Omega_1 \cap \Omega_2) \subset \hat{N}_\delta(x, \Omega_1) + \hat{N}_\delta(x, \Omega_2) + U^*$ for any $\delta > 0$ and any weak$^*$ neighborhood $U^*$ of $0$ in $X^*$.*

Other fuzzy calculus results (chain rules, formulas for maximum-type functions, mean value theorems, etc.) for functions and multifunctions can be deduced from (some form of) the Sum Rule (see [5, 16, 24, 37]).

## Extremal Principle

The *Extremal Principle* continues the line of variational principles discussed above and is in a sense equivalent to them as well as to the Sum Rules.

Let $\Omega_1$, $\Omega_2$ be closed subsets in $X$. They are called *locally extremal* [29, 30] near $x \in \Omega_1 \cap \Omega_2$ if there exists a neighborhood $U$ of $x$ and sequences $\{a_{ik}\} \in X$, $i = 1, 2$, $k = 1, 2, \ldots$, such that $a_{ik} \to 0$ when $k \to \infty$ and $(\Omega_1 - a_{1k}) \cap (\Omega_2 - a_{2k}) \cap U = \emptyset$, $k = 1, 2, \ldots$.

This means that by an arbitrarily small shift the sets can be made unintersecting in a neighborhood of $x$. The definition represents a rather general notion of extremality: some locally extremal system corresponds to a local solution of any optimization problem (see various examples in [22, 29, 31, 33]).

The *Extremal Principle*, first established in [29] (see also [22, 30, 31]) for the case of a Fréchet smooth space (and in terms of $\varepsilon$-normals) and in [35] (see [34]) in the Asplund space setting, provides a dual space characterization of locally extremal systems in terms of Fréchet normals. It can be viewed as a fuzzy form of the *separation* property.

**Extremal Principle.** *If a system of sets $\Omega_1$, $\Omega_2$ is locally extremal near $x \in \Omega_1 \cap \Omega_2$ then for any $\delta > 0$ there exist elements $x_1^* \in \hat{N}_\delta(x|\Omega_1)$, $x_2^* \in \hat{N}_\delta(x|\Omega_2)$ such that $\|x_1^* + x_2^*\| < \delta$, $\|x_1^*\| + \|x_2^*\| = 1$.*

The following theorem proved in [35] shows that the Extremal Principle provides an extremal characterization of Asplund spaces.

**Theorem 4.** *The following assertions are equivalent:*
*(i) $X$ is an Asplund space;*
*(ii) The Extremal Principle is valid in $X$.*

Due to Theorems 3, 4 the Extremal principle is equivalent to the Sum Rules. It is also equivalent to some other basic results of nonsmooth analysis (see [6, 34, 45]).

The Extremal Principle can be viewed as a certain extension of the classical separation theorem for convex sets. It was used in [22, 29, 30, 31, 36] and in many other papers as a main tool for deducing calculus formulas and necessary optimality conditions.

The local extremality assumption in the Extremal Principle can be replaced by a weaker *stationarity* condition (see [26, 27, 28]). The resulting *Extended Extremal Principle* is formulated as a necessary and sufficient condition and is also equivalent to the asplundity of the space.

As it was noticed in [35] considering the extremal system provided by the pair $\{x\}, \Omega$, where $x$ is a boundary point of a closed set $\Omega$ makes it possible to deduce from Theorem 4 the following nonconvex generalization of the well-known *Bishop-Phelps theorem* (see [40]).

**Corollary 4.1.** *Let $\Omega$ be a closed subset in an Asplund space $X$ and let $x \in \operatorname{bd}\Omega$. Then for any $\delta > 0$ there exists $x^* \in \hat{N}_\delta(x|\Omega)$ such that $\|x^*\| = 1$.*

See also: **Nonsmooth analysis: Weak stationarity**

# References


[1] Bazaraa, M. S., Goode, J. J., and Nashed, Z. Z. On the cones of tangents with applications to mathematical programming. *J. Optimiz. Theory Appl. 13* (1974), 389–426.

[2] Borwein, J. M. Weak tangent cones and optimization in a Banach space. *SIAM J. Contr. Optim. 16* (1978), 512–522.

[3] Borwein, J. M., and Preiss, D. A smooth variational principle with applications to subdifferentiability and differentiability of convex functions. *Trans. Amer. Math. Soc. 303* (1987), 517–527.

[4] Borwein, J. M., and Strojwas, H. M. Tangential approximations. *Nonlinear Anal. 9* (1985), 1347–1366.

[5] Borwein, J. M., and Zhu, Q. J. A survey of subdifferential calculus with applications. *Nonlinear Anal. 38* (1999), 687–773.

[6] Borwein, J. M., and Zhu, Q. J. *Techniques of Variational Analysis.* CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 20. Springer-Verlag, New York, 2005.

[7] Clarke, F. H. *Optimization and Nonsmooth Analysis.* John Wiley, New York, 1983.

[8] Deville, R., Godefroy, G., and Zizler, V. *Smoothness and Renorming in Banach spaces.* Pitman Monographs and Surveys in Pure and Applied Math., Vol. 64. Longman, Harlow, UK, 1993.

[9] Dubovitskii, A. Y., and Milyutin, A. A. Extremum problems in the presence of restrictions. *U.S.S.R. Comp. Maths. Math. Phys. 5* (1965), 1–80.

[10] Ekeland, I. On the variational principle. *J. Math. Anal. Appl. 47* (1974), 324–353.

[11] Fabian, M. Subdifferentials, local $\varepsilon$-supports and Asplund spaces. *J. London Math. Soc. 34* (1986), 568–576.

[12] Fabian, M. Subdifferentiability and trustworthiness in the light of a new variational principle of Borwein and Preiss. *Acta Univ. Carolinae 30* (1989), 51–56.

[13] Fabian, M., and Mordukhovich, B. S. Nonsmooth characterizations of Asplund spaces and smooth variational principles. *Set-Valued Anal. 6* (1998), 381–406.

[14] Gould, F. J., and Tolle, J. W. Optimality conditions and constraint qualifications in Banach space. *J. Optimiz. Theory Appl. 15* (1975), 667–684.

[15] Ioffe, A. D. On subdifferentiability spaces. *Ann. New York Acad. Sci. 410* (1983), 107–119.

[16] Ioffe, A. D. Calculus of Dini subdifferentials of functions and contingent coderivatives of set-valued maps. *Nonlinear Anal. 8* (1984), 517–539.

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

[18] Kruger, A. Y. *Subdifferentials of nonconvex functions and generalized directional derivatives.* Deposited in VINITI no. 2661-77. Minsk, 1977. in Russian.

[19] Kruger, A. Y. *Generalized differentials of nonsmooth functions.* Deposited in VINITI no. 1332-81. Minsk, 1981. in Russian.

[20] Kruger, A. Y. *Necessary conditions of an extremum in problems of nonsmooth optimization.* Deposited in VINITI no. 1333-81. Minsk, 1981. in Russian.

[21] Kruger, A. Y. *$\varepsilon$-semidifferentials and $\varepsilon$-normal elements.* Deposited in VINITI no. 1331-81. Minsk, 1981. in Russian.

[22] Kruger, A. Y. Generalized differentials of nonsmooth functions and necessary conditions for an extremum. *Siberian Math. J. 26* (1985), 370–379.

[23] Kruger, A. Y. Properties of generalized differentials. *Siberian Math. J. 26* (1985), 822–832.

[24] Kruger, A. Y. On calculus of strict $\varepsilon$-semidifferentials. *Dokl. Akad. Nauk Belarusi 40*, 4 (1996), 34–39. in Russian.

[25] Kruger, A. Y. On Fréchet subdifferentials. *J. Math. Sci. (N. Y.) 116*, 3 (2003), 3325–3358. Optimization and related topics, 3.

[26] Kruger, A. Y. Weak stationarity: eliminating the gap between necessary and sufficient conditions. *Optimization 53*, 2 (2004), 147–164.

[27] Kruger, A. Y. Stationarity and regularity of set systems. *Pac. J. Optim. 1*, 1 (2005), 101–126.

[28] Kruger, A. Y. About regularity of collections of sets. *Set-Valued Anal. 14*, 2 (2006), 187–206.

[29] Kruger, A. Y., and Mordukhovich, B. S. Extremal points and the Euler equation in nonsmooth optimization. *Dokl. Akad. Nauk BSSR 24*, 8 (1980), 684–687. in Russian.

[30] Kruger, A. Y., and Mordukhovich, B. S. *Generalized normals and derivatives and necessary conditions for an extremum in problems of nondifferentiable programming.* Deposited in VINITI, I – no. 408-80, II – no. 494-80. Minsk, 1980. in Russian.

[31] Mordukhovich, B. S. *Approximation Methods in Problems of Optimization and Control.* Nauka, Moscow, 1988. in Russian.

[32] Mordukhovich, B. S. Generalized differential calculus for nonsmooth and set-valued mappings. *J. Math. Anal. Appl. 183*, 1 (1994), 250–288.

[33] Mordukhovich, B. S. The extremal principle and its applications to optimization and economics. In *Optimization and Related Topics*, A. Rubinov and B. Glover, Eds., Applied Optimization, Vol. 47. Kluwer Academic Publishers, Dordrecht, 2001, pp. 343–369.

[34] Mordukhovich, B. S. *Variational Analysis and Generalized Differentiation. I*, vol. 330 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 2006. Basic theory.

[35] Mordukhovich, B. S., and Shao, Y. Extremal characterizations of Asplund spaces. *Proc. Amer. Math. Soc. 124* (1996), 197–205.

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

[37] Mordukhovich, B. S., and Shao, Y. Fuzzy calculus for coderivatives of multifunctions. *Nonlinear Anal. 29* (1997), 605–626.

[38] Mordukhovich, B. S., and Wang, B. Necessary suboptimality and optimality conditions via variational principles. *SIAM J. Control Optim. 41*, 2 (2002), 623–640.

[39] Penot, J.-P. Calcul sous-differentiel et optimisation. *J. Funct. Anal. 27* (1978), 248–276.

[40] Phelps, R. R. *Convex Functions, Monotone Operators and Differentiability, 2nd edition*. Lecture Notes in Mathematics, Vol. 1364. Springer-Verlag, New York, 1993.

[41] Rockafellar, R. T. *Convex Analysis*. Princeton Univ. Press, Princeton, N.J., 1970.

[42] Rockafellar, R. T., and Wets, R. J.-B. *Variational Analysis*. Springer-Verlag, New York, 1998.

[43] Sachs, E. Differentiability in optimization theory. *Math. Operationsforsch. Statist., Ser. Optimiz. 9* (1978), 497–513.

[44] Treiman, J. S. Clarke's gradients and epsilon-subgradients in Banach spaces. *Trans. Amer. Math. Soc. 294*, 1 (1986), 65–78.

[45] Zhu, Q. J. The equivalence of several basic theorems for subdifferentials. *Set-Valued Anal. 81* (1998), 171–185.

MSC 2000: 49K27, 90C48, 58C20, 58E30
*Key words and phrases:* nonsmooth analysis, variational analysis, subdifferential, normal cone, optimality conditions, extremal principle, multifunction, Asplund space.