STRICT (ε, δ)-SUBDIFFERENTIALS AND EXTREMALITY CONDITIONS

A. YA. KRUGER[†]

Department of Information Technologies and Mathematics, International Institute of Labour and Social Relations, Kazintsa St. 21/3, 220099 Minsk, Belarus

A new abstract definition of extremality is introduced extending traditional extremality notions in optimization problems. The set of points satisfying the new definition includes points not necessarily optimal in the usual sense but nevertheless having some extremal properties. Necessary and sufficient extremality conditions are derived. Contrary to usual notions of extremality the new one is stable relative to small deformations of the data.

Keywords: nonsmooth analysis, subdifferentials, normal cones, optimality conditions, set-valued mappings, Asplund spaces

Mathematics Subject Classification 1991:
Primary: 49K27, 90C48; Secondary: 58C20, 58E30.

1 Introduction.

If one takes an optimization problem and some necessary optimality conditions, then in general there exist points satisfying these conditions but not being optimal in the usual sense. Nevertheless, if the necessary conditions are strong enough then such points do have some extremality properties. The main idea of the paper is to extend the traditional notion of optimality in order to be able to treat these "almost optimal" points as well. On this way we come to new (extended) definitions of extremality, and the necessary conditions become also sufficient.

[†]E-mail: kruger@iilsr.minsk.by

As a basis for constructing the new abstract notion of extremality we take some modifications of the known necessary extremality conditions from the theory of nonsmooth optimization: *the generalized Euler equation* for *extremal systems of sets* [20] and *Lagrange multipliers rule* for (f, Ω, M)-*extremal* points [11] (see also [13, 15, 16, 17, 18, 21, 25]).

Contrary to usual notions of extremality the one presented below is stable relative to small deformations of the data near the point under consideration. Let us take for example a problem of unconditional minimization of a real valued function $\varphi(u) = u^2$ defined on a real line. $u = 0$ is obviously a point of minimum. If we add to φ an indefinitely small (in a neighborhood of $u = 0$) function $\psi(u) = -|u|^{3/2}$, then $u = 0$ is no longer a point of minimum (Actually it is a point of local maximum of $\varphi+\psi$). When using the extended definition of minimality, $u = 0$ remains a point of minimum for $\varphi + \psi$ as far as it is a point of minimum for φ and $\nabla\psi(0) = 0$ (Proposition 7.3).

The paper is organized as follows. In Section 2 we present definitions of derivative-like objects which are used in statements of extremality conditions. In Section 3 the generalized Euler equation is defined. Sections 4 and 5 are devoted to new definitions of the extremal system of sets and (f, Ω, M)-extremal points. The necessary and sufficient extremality conditions are also presented. Section 6 presents the *covering theorem*. Several examples of problems of unconditional minimization and the definition of *extended minimality* can be found in Section 7.

Some constants are defined in the paper which simplify definitions and statements of the results.

Mostly standard notations are used throughout the paper. X, Y denote Banach spaces and X^*, Y^* denote their topological duals. $\langle\cdot,\cdot\rangle$ is a canonical paring between a space and its dual. $B_\rho(x)$ stands for a closed ball with center x and radius ρ. We shall write B_ρ instead of $B_\rho(0)$. The norms in the primal and the dual spaces will be denoted respectively by $\|\cdot\|$ and $\|\cdot\|_*$.

2 Strict (ε, δ)-subdifferentials.

Let Ω be a set in a Banach space X and let φ be a function from Ω into R, $x \in \Omega$, $\varepsilon \geq 0$, $\delta > 0$.

A convex set

$$\partial_\varepsilon^- \varphi(x) = \left\{ x^* \in X^* : \liminf_{u \xrightarrow{\Omega} x} \frac{\varphi(u) - \varphi(x) - \langle x^*, u - x \rangle}{\|u - x\|} \geq -\varepsilon \right\} \quad (1)$$

is called an *ε-subdifferential* of φ at x. Here $u \xrightarrow{\Omega} x$ means that $u \to x$ with $u \in \Omega$.

When $x \in \operatorname{int} \Omega$, $\varepsilon = 0$ and φ is convex (Fréchet differentiable at x) the set (1) coincides with the subdifferential of convex analysis (reduces to the Fréchet derivative). In general it is a direct extension of the Fréchet subdifferential (see [19] for the survey of different properties of such subdifferentials and their generalizations).

Let us consider another set based on (1):

$$\hat{\partial}_{\varepsilon,\delta}^- \varphi(x) \;=\; \bigcup_{\substack{u \in cl\Omega \cap B_\delta(x) \\ |cl^\downarrow \varphi(u) - \varphi(x)| \leq \delta}} \partial_\varepsilon^- (cl^\downarrow \varphi)(u). \quad (2)$$

$cl^\downarrow \varphi$ denotes here the lower semicontinuous envelope of φ. The set (2) can be nonconvex. We shall call it a *strict (ε,δ)-subdifferential* of φ at x.

An *ε-superdifferential* and a *strict (ε,δ)-superdifferential* of φ at x can be defined in a similar way:

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

$$\hat{\partial}_{\varepsilon,\delta}^+ \varphi(x) \;=\; \bigcup_{\substack{u \in cl\Omega \cap B_\delta(x) \\ |cl^\uparrow \varphi(u) - \varphi(x)| \leq \delta}} \partial_\varepsilon^+ (cl^\uparrow \varphi)(u) \quad (3)$$

($cl^\uparrow \varphi$ is the upper semicontinuous envelope of φ). The equality

$$\hat{\partial}_{\varepsilon,\delta}^+ \varphi(x) = -\hat{\partial}_{\varepsilon,\delta}^- (-\varphi)(x)$$

holds true.

Let us note that strict sub- and superdifferentials can be nonempty simultaneously and can be essentially different as in the nonsmooth case "differential" properties of a function "from below" and "from above" can differ significantly.

The set

$$\hat{\partial}^0_{\varepsilon,\delta}\varphi(x) = \hat{\partial}^-_{\varepsilon,\delta}\varphi(x) \cup \hat{\partial}^+_{\varepsilon,\delta}\varphi(x) \tag{4}$$

is called a *strict (ε,δ)-differential* of φ at x.

The above definitions (2), (3), (4) are some modifications of the definitions of strict ε-semidifferentials introduced in [14]. When $\varepsilon = 0$ we shall omit index ε in corresponding denotations. Strict δ-subdifferentials were used in [17].

All the "strict" sets (2), (3), (4) do have some properties of a strict derivative. This is due to the fact that when defining a strict (ε, δ)-subdifferential we actually calculate subdifferentials in a neighborhood of a point under consideration (see (2), (3)).

Let us recall that a function $\varphi : X \to R$ is called *strictly differentiable* at x (with a *strict derivative* x^*) if

$$\lim_{\substack{u\to x\\ u'\to x}} \frac{\varphi(u) - \varphi(u') - \langle x^*, u - u'\rangle}{\|u - u'\|} = 0.$$

Evidently this definition imposes on φ stronger conditions, than the usual definition of a Fréchet derivative.

Strict (ε, δ)-subdifferentials (strict δ-subdifferentials) and the corresponding geometrical objects are used below in statements of the extremality conditions. Of course, all the results can be formulated in terms of original subdifferentials (without the union operation). Using strict subdifferentials simplifies the statements.

With the help of strict (ε, δ)-subdifferentials it is possible to define limiting constructions – the *generalized differentials* [12, 21] and the *approximative subdifferentials* [9], which have been successfully applied in nonsmooth analysis and optimization for many years.

Some geometrical objects can be defined similarly: a *set of ε-normals*

$$N_\varepsilon(x|\Omega) = \left\{ x^* \in X^* : \limsup_{u \xrightarrow{\Omega} x} \frac{\langle x^*, u - x\rangle}{\|u - x\|} \le \varepsilon \right\}$$

and a *strict set of (ε, δ)-normals*

$$\hat{N}_{\varepsilon,\delta}(x|\Omega) \;=\; \bigcup \left[N_\varepsilon(u|cl\Omega) : \; u \in cl\Omega \cap B_\delta(x) \right],$$

and (when $\varepsilon = 0$) a *normal cone* $N(x|\Omega)$ and a *strict δ-normal cone* $\hat{N}_\delta(x|\Omega)$ to Ω at x. They are closely related to ε-subdifferentials and strict (ε, δ)-subdifferentials.

For a set-valued mapping $F: \ \Omega \rightrightarrows Y$ we can define a *strict (ε, δ)-coderivative* and a *strict δ-coderivative* [17]. They are defined by the equalities

$$\hat{\partial}_{\varepsilon,\delta} F(x, y; y^*) = \{x^* \in X^* : \ (x^*, -y^*) \in \hat{N}_{\varepsilon,\delta}(x, y|\operatorname{gph} F)\},$$

$$\hat{\partial}_{\delta} F(x, y; y^*) = \{x^* \in X^* : \ (x^*, -y^*) \in \hat{N}_{\delta}(x, y|\operatorname{gph} F)\},$$

where $\operatorname{gph} F$ is the graph of F.

3 The generalized Euler equation.

Let Ω_1, Ω_2 be closed sets in X. Following [15], let us introduce a constant

$$\theta(\Omega_1, \Omega_2) = \sup\{r \geq 0 : \ B_r \subset \Omega_1 - \Omega_2\}, \tag{5}$$

describing the rate of their "overlapping". The difference of sets in (5) is understood in the algebraic sense: $\Omega_1 - \Omega_2 = \{\omega_1 - \omega_2 : \omega_1 \in \Omega_1, \ \ \omega_2 \in \Omega_2\}$. When $\Omega_1 \cap \Omega_2 = \emptyset$, we suppose, that $\theta(\Omega_1, \Omega_2) = -\infty$.

The definition of an *extremal set system* introduced in [20] can now be formulated as follows.

Definition 3.1 *A system of sets Ω_1, Ω_2 is extremal if $\theta(\Omega_1, \Omega_2) = 0$. In this case any $x \in \Omega_1 \cap \Omega_2$ is called an extremal point of the system Ω_1, Ω_2.*

Definition 3.2 *A system of sets Ω_1, Ω_2 is locally extremal in a neighborhood of $x \in \Omega_1 \cap \Omega_2$, if for some $\rho > 0$ the system $\Omega_1 \cap B_\rho(x)$, $\Omega_2 \cap B_\rho(x)$ is extremal.*

The concept of an extremal system characterizes mutual arrangement of sets in space. In convex case the definition of an extremal system reduces to the classical form.

Proposition 3.1 *Let Ω_1, Ω_2 be convex sets,* $\operatorname{int} \Omega_2 \neq \emptyset$. *A system of sets Ω_1, Ω_2 is extremal if and only if $\Omega_1 \cap \operatorname{int} \Omega_2 = \emptyset$.*

In general this concept covers different notions of optimality in scalar and/or vector constrained optimization problems (see various examples in [11, 20, 21] and the recent survey paper [22]). A simple example of an extremal system is provided by the pair $\{x\}, \Omega$, where x is a boundary point of Ω.

In case of an Asplund space it is possible to formulate a dual criterion of extremality (see [10, 16, 20, 21, 25]).

Let us recall that a Banach space is called *Asplund* [1] (see [27]) if any continuous convex function on it is Fréchet differentiable on a dense set of points. Asplund spaces form a rather broad subclass of Banach spaces. See [3, 27] for various properties and characterizations of Asplund spaces. This class includes e. g. all spaces which admit Fréchet differentiable bump functions (in particular, Fréchet smooth spaces). Reflexive spaces are examples of Fréchet smooth spaces.

Asplund spaces have proved to be very convenient for investigating different properties of nonsmooth functions. Actually Asplund property of Banach spaces 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 [7, 8, 24, 26]).

DEFINITION 3.3 *Let us say that the generalized Euler equation holds true at* $x \in \Omega_1 \cap \Omega_2$ *if for any* $\varepsilon > 0$, $\delta > 0$ *there exist such elements* $x_1^* \in \hat{N}_\delta(x|\Omega_1)$, $x_2^* \in \hat{N}_\delta(x|\Omega_2)$ *that*

$$\|x_1^* + x_2^*\|_* < \varepsilon, \tag{6}$$

$$\|x_1^*\|_* + \|x_2^*\|_* = 1. \tag{7}$$

THEOREM 3.1 *Let* X *be an Asplund space and let a system of sets* Ω_1, Ω_2 *be locally extremal near* x. *Then the generalized Euler equation holds true at* x.

Theorem 3.1 can be viewed as a certain generalization of the classical separation theorem for convex sets. The generalized Euler equation characterizes extremal properties of set systems. It was used in [11, 20, 21, 25] as a main tool for deducing calculus formulas and necessary optimality conditions.

Strict δ-normal cones are used in Theorem 3.1 (see Definition 3.3) instead of strict sets of (ε, δ)-normals, which were used in the corresponding statement in [16], and the theorem is formulated not for smooth but for an arbitrary Asplund space. This became possible due to replacing in its proof the variational principle of Ekeland [5] by the smooth variational principle of Borwein and Preiss [2] and applying the separable reduction method of Fabian [6].

The statement of Theorem 3.1 cannot be extended to nonasplund spaces. Actually it can be considered as a form of extremal characterization of Asplund spaces.

4 Extremal set systems.

Let us introduce a local constant based on (5):

$$\hat{\theta}_x(\Omega_1, \Omega_2) = \liminf_{\substack{\omega_1 \xrightarrow{\Omega_1} x,\ \omega_2 \xrightarrow{\Omega_2} x \\ \rho \to +0}} \frac{\theta([\Omega_1 - \omega_1] \cap B_\rho, [\Omega_2 - \omega_2] \cap B_\rho)}{\rho}. \tag{8}$$

Definition 4.1 *A system of sets Ω_1, Ω_2 is extended extremal (e-extremal) near $x \in \Omega_1 \cap \Omega_2$ if $\hat{\theta}_x(\Omega_1, \Omega_2) = 0$.*

Proposition 4.1 *If a system of sets Ω_1, Ω_2 is locally extremal in a neighborhood of x, then it is e-extremal near x.*

The definition (8) can be slightly simplified if we use another local constant:

$$\theta_x(\Omega_1, \Omega_2) = \liminf_{\rho \to +0} \frac{\theta(\Omega_1 \cap B_\rho(x),\ \Omega_2 \cap B_\rho(x))}{\rho}. \tag{9}$$

Then we have

$$\hat{\theta}_x(\Omega_1, \Omega_2) = \liminf_{\omega_1 \xrightarrow{\Omega_1} x,\ \omega_2 \xrightarrow{\Omega_2} x} \theta_0(\Omega_1 - \omega_1,\ \Omega_2 - \omega_2). \tag{10}$$

In general, Definition 4.1 differs from Definition 3.2 in two important points. First, it is required not strict "nonoverlapping" of sets but up to an indefinitely small deformation (see (9)). Second, the sets can "nonoverlap" not in the point x: it is sufficient that in any its neighborhood there exist such points $\omega_1 \in \Omega_1$, $\omega_2 \in \Omega_2$, that the sets $\Omega_1 - \omega_1$, $\Omega_2 - \omega_2$ "almost nonoverlap" (see (10)).

THEOREM 4.1 *Let X be an Asplund space. A system of sets Ω_1, Ω_2 is e-extremal near $x \in \Omega_1 \cap \Omega_2$ if and only if the generalized Euler equation holds true at x.*

Proof. The proof of the necessity is based on the same ideas, which were used to prove Theorem 3.1. Let the system Ω_1, Ω_2 be e-extremal near $x \in \Omega_1 \cap \Omega_2$, i. e. $\hat{\theta}_x(\Omega_1, \Omega_2) = 0$, and let numbers $\varepsilon > 0$, $\delta > 0$ be given. Without loss of generality we will suppose that $\varepsilon \leq 1$. Then due to (8) there exist points $\hat{\omega}_1 \in \Omega_1 \cap B_{\delta/2}(x)$, $\hat{\omega}_2 \in \Omega_2 \cap B_{\delta/2}(x)$ and a positive number $\rho < \delta$, such that $\theta([\Omega_1 - \hat{\omega}_1] \cap B_\rho, [\Omega_2 - \hat{\omega}_2] \cap B_\rho) < \rho\varepsilon/8$. Taking into account the definition (5) the last inequality yields the existence of a point $a \in B_{\rho\varepsilon/8}$, such that $a \notin ([\Omega_1 - \hat{\omega}_1] \cap B_\rho) - ([\Omega_2 - \hat{\omega}_2] \cap B_\rho)$. This means that $\|\omega_1 - \hat{\omega}_1 - \omega_2 + \hat{\omega}_2 - a\| > 0$ for any $\omega_1 \in \Omega_1 \cap B_\rho(\hat{\omega}_1)$, $\omega_2 \in \Omega_2 \cap B_\rho(\hat{\omega}_2)$.

Let us consider a real valued function $\psi_1(\omega_1, \omega_2) = \|\omega_1 - \hat{\omega}_1 - \omega_2 + \hat{\omega}_2 - a\|$ defined on $X \times X$ (we suppose that the latter space is equipped with the maximum norm: $\|u_1, u_2\| = \max(\|u_1\|, \|u_2\|)$). Then $\psi_1(\omega_1, \omega_2) > 0$ for all $(\omega_1, \omega_2) \in \Omega = [\Omega_1 \cap B_\rho(\hat{\omega}_1)] \times [\Omega_2 \cap B_\rho(\hat{\omega}_2)]$ and $\psi_1(\hat{\omega}_1, \hat{\omega}_2) = \|a\| \leq \rho\varepsilon/8$. It follows from the Ekeland's variational principle [5] that there exist points $\tilde{\omega}_1 \in \Omega_1 \cap B_{\rho/4}(\hat{\omega}_1)$, $\tilde{\omega}_2 \in \Omega_2 \cap B_{\rho/4}(\hat{\omega}_2)$, such that $(\tilde{\omega}_1, \tilde{\omega}_2)$ is a point of minimum for the extended real valued function $\psi_1 + \psi_2 + \psi_3$, where $\psi_2(u_1, u_2) = (\varepsilon/2)\|(u_1, u_2) - (\tilde{\omega}_1, \tilde{\omega}_2)\|$, $\psi_3(u_1, u_2) = 0$ if $(u_1, u_2) \in \Omega$ and $\psi_3(u_1, u_2) = \infty$ if $(u_1, u_2) \notin \Omega$ (the indicator function of Ω).

The functions ψ_1 and ψ_2 are convex and Lipschitzian, and we can apply lemma 4 from [6] ("zero" fuzzy sum rule). There exist points $(x_{1j}, x_{2j}) \in B_{\rho/4}(\tilde{\omega}_1, \tilde{\omega}_2)$ and $(x_{1j}^*, x_{2j}^*) \in \partial^- \psi_j(x_{1j}, x_{2j})$, $j = 1, 2, 3$, such that $(x_{13}, x_{23}) \in \Omega$ and $\|(x_{11}^*, x_{21}^*) + (x_{12}^*, x_{22}^*) + (x_{13}^*, x_{23}^*)\|_* \leq \varepsilon/2$. Evidently we have $(x_{1j}, x_{2j}) \in B_{\rho/2}(\hat{\omega}_1, \hat{\omega}_2)$, $\|x_{12}^*, x_{22}^*\|_* \leq \varepsilon/2$, $x_{13}^* \in N(x_{13}|\Omega_1)$, $x_{23}^* \in N(x_{23}|\Omega_2)$ and $x_{11}^* = x^*$, $x_{21}^* = -x^*$, where x^* is the subgradient of the norm, calculated in the nonnull point $x_{11} - \hat{\omega}_1 - x_{21} + \hat{\omega}_2 - a$. Thus $\|x^*\|_* = 1$ and $\|x_{13}^* + x_{23}^*\|_* \leq \|x_{13}^* + x^*\|_* + \|x_{23}^* - x^*\|_* = \|(x_{11}^*, x_{21}^*) + (x_{13}^*, x_{23}^*)\|_* \leq \varepsilon$. We have also $2 - \varepsilon \leq \|x_{13}^*\| + \|x_{23}^*\|_* \leq 2 + \varepsilon$.

Let us denote $x_1 = x_{13}$, $x_2 = x_{23}$, $x_1^* = x_{13}^*/(\|x_{13}^*\|_* + \|x_{23}^*\|_*)$, $x_2^* = x_{23}^*/(\|x_{13}^*\|_* + \|x_{23}^*\|_*)$. Then $x_1 \in \Omega_1 \cap B_\delta(x)$, $x_2 \in \Omega_2 \cap B_\delta(x)$,

$x_1^* \in N(x_1|\Omega_1) \subset \hat{N}_\delta(x|\Omega_1)$, $x_2^* \in N(x_2|\Omega_2) \subset \hat{N}_\delta(x|\Omega_2)$, $\|x_1^*\|_* + \|x_2^*\|_* = 1$, $\|x_1^* + x_2^*\|_* \le \varepsilon/(2-\varepsilon) \le \varepsilon$. This completes the proof of the necessity.

Let a system of sets Ω_1, Ω_2 be not e-extremal near $x \in \Omega_1 \cap \Omega_2$, i. e. $\hat{\theta}_x(\Omega_1, \Omega_2) = \alpha > 0$. Then there exist $r > 0$, such that

$$\theta([\Omega_1 - \omega_1] \cap B_\rho, [\Omega_2 - \omega_2] \cap B_\rho) \ge \alpha\rho/2 \tag{11}$$

for any positive $\rho \le r$ and any $\omega_1 \in \Omega_1 \cap B_r$, $\omega_2 \in \Omega_2 \cap B_r$. If the generalized Euler equation holds true at x, then there exist $\omega_1 \in \Omega_1 \cap B_r$, $\omega_2 \in \Omega_2 \cap B_r$ and $x_1^* \in N(\omega_1|\Omega_1)$, $x_2^* \in N(\omega_2|\Omega_2)$, such that conditions (6), (7) hold true for $\varepsilon = \min(\alpha/16, 1/4)$. It follows from the definition of a normal cone that there exist a positive $\rho \le r$, such that $\langle x_i^*, \omega\rangle \le \frac{\alpha}{32}\|\omega\|$ for any $\omega \in [\Omega_i - \omega_i] \cap B_\rho$, $i = 1, 2$. But $x_2^* = -x_1^* + (x_1^* + x_2^*)$. Consequently we have $-\langle x_1^*, \omega\rangle \le (\frac{\alpha}{32} + \varepsilon)\|\omega\| \le \frac{3\alpha}{32}\|\omega\|$ for any $\omega \in [\Omega_2 - \omega_2] \cap B_\rho$ and $\langle x_1^*, \omega\rangle \le \alpha\rho/8$ for any $\omega \in ([\Omega_1 - \omega_1] \cap B_\rho) - ([\Omega_2 - \omega_2] \cap B_\rho)$. Now it follows from (11) and the definition (5) that $\|x_1^*\|_* \le 1/4$ and $\|x_2^*\|_* \le \|x_1^*\|_* + \varepsilon \le 1/2$. This contradicts (7). ∎

Let us note that the sufficient part of Theorem 4.1 is true for any (not necessary Asplund) Banach space.

The above theorem can serve as a tool for deducing necessary and sufficient conditions of “extended extremality” in various extremal problems.

The definition of the extremal system can be expanded for the case of n sets.

Definition 4.2 *A system of sets Ω_i, $i = 1, 2, \dots, n$, is e-extremal near x if the system of sets $\tilde{\Omega}_1 = \prod_{i=1}^n \Omega_i$ and $\tilde{\Omega}_2 = \{(\omega, \omega, \dots, \omega) \in X^n\}$ is e-extremal near $(x, x, \dots, x) \in X^n$.*

Proposition 4.2 *Let $I = \{1, 2, \dots, n\}$, $j \in I$. A system of sets Ω_i, $i \in I$, is e-extremal near x if the system of sets $\tilde{\Omega}_1 = \prod_{i \in I\setminus\{j\}} \Omega_i$ and $\tilde{\Omega}_2 = \{(\omega, \omega, \dots, \omega) \in X^{n-1} : \omega \in \Omega_j\}$ is e-extremal near $(x, x, \dots, x) \in X^{n-1}$.*

The following theorem gives a dual criterion of extremality (the generalized Euler equation).

THEOREM 4.2 *Let X be an Asplund space. A system of sets Ω_i, $i = 1, 2, \ldots, n$, is e-extremal near x if and only if for any $\varepsilon > 0$, $\delta > 0$ there exist such elements $x_i^* \in \hat{N}_\delta(x|\Omega_i)$, $i = 1, 2, \ldots, n$, that*

$$||x_1^* + x_2^* + \cdots + x_n^*||_* < \varepsilon,$$

$$||x_1^*||_* + ||x_2^*||_* + \cdots + ||x_n^*||_* = 1.$$

The proof of Theorem 4.2 reduces to calculating the cones $N(\tilde{x}|\tilde{\Omega}_1)$ and $N(\tilde{x}|\tilde{\Omega}_2)$.

5 Extremality Conditions.

Let us consider closed sets $\Omega \subset X$, $M \subset Y$ and a function $f : \Omega \to Y$. We shall suppose that f is M-closed, i. e. the graph of the set-valued mapping

$$F(u) = f(u) - M, \; u \in \Omega \tag{12}$$

is closed in $X \times Y$. Let us introduce a constant

$$\hat{\theta}_x(f, \Omega, M) = \liminf_{\substack{u \xrightarrow{\Omega} x \\ v - f(u) \overset{v \in M}{\to} 0 \\ \rho \to +0}} \frac{\theta(f(\Omega \cap B_\rho(u)) - f(u), M - v)}{\rho}.$$

DEFINITION 5.1 *A point $x \in \Omega$ is (f, Ω, M)-extremal if $f(x) \in M$ and $\hat{\theta}_x(f, \Omega, M) = 0$.*

The above definition is a modification of the corresponding definitions from [11, 13]. Contrary to the known abstract notions of extremality there are no conditions on the set M in Definition 5.1: it does not need to be convex and/or to have nonempty interior.

Definition 5.1 is in a sense a counterpart of Definition 4.1 of an extended extremal system of sets. It also gives a rather general notion of (extended) extremality, embracing different optimality notions in optimization problems. If e. g. x is a local solution of the nonlinear programming problem

$$
\begin{array}{ll}
\text{minimize} & f_0(u) \\
\text{subject to} & f_i(u) \leq 0,\ i = 1, 2, \ldots, m, \\
& f_i(u) = 0,\ i = m+1, \ldots, n, \\
& u \in \Omega,
\end{array}
$$

where Ω is a closed set, m, n are nonnegative integers, $m \leq n$, functions f_i are lower semicontinuous for $i = 0, 1, \ldots, m$ and continuous for $i = m+1, \ldots, n$, then x is (f, Ω, M)-extremal if one takes $f = (f_0 - f_0(x), f_1, f_2, \ldots, f_n) : X \to R^{n+1}$, $M = R_-^{m+1} \times 0$.

Theorem 5.1 *Let X, Y be Asplund. A point $x \in \Omega$ is (f, Ω, M)-extremal if and only if $f(x) \in M$ and for any $\delta > 0$ there exists an element $y^* \in Y^*$, such that $||y^*||_* = 1$ and*

$$0 \in \hat{\partial}_\delta F(x, 0; y^*), \tag{13}$$

where F is defined by (12).

Theorem 5.1 can be deduced from Theorem 4.1 or it can be proved independently. Actually both results are equivalent.

The above theorem gives general extremality conditions. The element y^* can be viewed as an analog of the Lagrange multipliers vector in the classical problem of nonlinear programming. (13) yields the inclusion

$$y^* \in \hat{N}_\delta(f(x)|M), \tag{14}$$

generalizing conditions on signs of multipliers and complementarity slackness conditions.

The following statement is a corollary of Theorem 5.1 under the additional assumption that f is Lipschitz.

Theorem 5.2 *Let X, Y be Asplund and let f be Lipschitz continuous near x. A point $x \in \Omega$ is (f, Ω, M)-extremal if and only if $f(x) \in M$ and for any $\delta > 0$ there exists such an element $y^* \in Y^*$ that $||y^*||_* = 1$,*

$$0 \in \hat{\partial}_\delta^- \langle y^*, f \rangle (x) \tag{15}$$

and the condition (14) holds true.

The inclusion (15) generalizes the classical Lagrange multipliers rule. Here the function $\langle y^*, f\rangle : X \to R$ is defined by the equality $\langle y^*, f\rangle(u) = \langle y^*, f(u)\rangle$.

Some examples of necessary optimality conditions, derived from Theorem 5.2, can be found in [15].

6 The Covering Theorem.

The last "extremal" result of the paper is the *covering theorem* (see [4, 13, 15]), which can be viewed as a certain generalization of the classical open mapping theorem to the case of a set-valued mapping.

Let us consider a set-valued mapping $F:\ \Omega \rightrightarrows Y$ with a closed graph $\operatorname{gph} F$. Let us denote

$$\theta_F(x, y, \rho) = \sup\{r \geq 0 : B_r(y) \subset F(\Omega \cap B_\rho(x))\}, \tag{16}$$

$$\hat{\theta}_F(x, y) = \liminf_{\substack{u \xrightarrow{\Omega} x,\ v \xrightarrow{F(u)} y \\ \rho \to +0}} \frac{\theta_F(u, v, \rho)}{\rho}. \tag{17}$$

DEFINITION 6.1 *F covers (at a linear rate [23]) near $(x, y) \in \operatorname{gph} F$ if $\hat{\theta}_F(x, y) > 0$.*

The following constant is used for characterizing the covering property:

$$b_F(x, y) = \sup_{\delta > 0} \inf\{\|x^*\|_* :\ x^* \in \hat{\partial}_\delta F(x, y; y^*),\ \|y^*\|_* = 1\}. \tag{18}$$

THEOREM 6.1 *Let X, Y be Asplund spaces. F covers in a neighborhood of $(x, y) \in \operatorname{gph} F$ if and only if $b_F(x, y) > 0$.*

Theorem 6.1 can be proved in a similar way as Theorem 4.1. On the other hand it can be deduced from Theorem 4.1 (or Theorem 5.1). In their turn Theorems 4.1 and 5.1 can be proved as corollaries of Theorem 6.1.

In case of an ordinary (single-valued) continuous mapping $f : \Omega \to Y$ the following constants are used in the definition of the covering property instead of (16), (17):

$$\theta_f(x,\rho) = \sup\{r \geq 0 : B_r(f(x)) \subset f(\Omega \cap B_\rho(x))\},$$

$$\hat{\theta}_f(x) = \liminf_{\substack{u \xrightarrow{\Omega} x \\ \rho \to +0}} \frac{\theta_f(u,\rho)}{\rho},$$

and in the formulation of the covering criterion instead of (18) it is necessary to use the constant

$$b_f(x) = \sup_{\delta > 0} \inf\{\|x^*\|_* : \ x^* \in \hat{\partial}_\delta^- \langle y^*, f\rangle(x), \ \|y^*\|_* = 1\}.$$

THEOREM 6.2 *Let X, Y be Asplund and let f be Lipschitz continuous near x. f covers near $x \in \Omega$ if and only if $b_f(x) > 0$.*

7 Some Examples

In this section we present some examples illustrating the abstract definition of extremality introduced in Section 5. We shall consider a problem of unconditional minimization of a lower semicontinuous function $\varphi : X \to R$.

Let us fix $x \in X$ and set $f(u) = \varphi(u) - \varphi(x)$, $\Omega = X$, $M = R_-$. In this setting Definition 5.1 of a (φ, Ω, M)-extremal point is equivalent to the following one.

DEFINITION 7.1 *x is a point of extended minimum (e-minimum) of φ if*

$$\limsup_{\substack{u \xrightarrow{\varphi} x \\ \rho \to +0}} \inf_{v \in B_\rho(u)} \frac{\varphi(v) - \varphi(u)}{\rho} = 0.$$

Hear $u \xrightarrow{\varphi} x$ means that $u \to x$ and $\varphi(u) \to \varphi(x)$.

The following statement is an easy consequence of the definition.

PROPOSITION 7.1 *If x is a point of local minimum of φ then it is a point of e-minimum of φ.*

Application of Theorem 5.1 leads to the following result.

PROPOSITION 7.2 *Let X be Asplund. x is a point of e-minimum for φ if and only if $0 \in \hat{\partial}_\delta^- \varphi(x)$ for any $\delta > 0$.*

The definition of extended minimality is stable relative to small deformations of the data.

PROPOSITION 7.3 *Let ψ be strictly differentiable at x with $\nabla\psi(x) = 0$. If x is a point of e-minimum for φ then it is a point of e-minimum for $\varphi + \psi$.*

In the smooth case a set of e-minimal points coincides with a set of stationary points.

PROPOSITION 7.4 *Let φ be strictly differentiable at x. x is a point of e-minimum for φ if and only if $\nabla\varphi(x) = 0$.*

In each of the examples below $x = 0$ is a point of e-minimum for $\varphi:\ R \to R$.

EXAMPLE 7.1 $\varphi(x) = x^2,\ x \in R.$

EXAMPLE 7.2 $\varphi(x) = -x^2,\ x \in R.$

EXAMPLE 7.3 $\varphi(x) = x^3,\ x \in R.$

EXAMPLE 7.4
$$\varphi(x) = \begin{cases} -\dfrac{1}{n}, & -\dfrac{1}{n} < x \le -\dfrac{1}{n+1},\ n = 1, 2, \ldots, \\ 0, & x = 0, \\ \dfrac{1}{n}, & \dfrac{1}{n+1} < x \le \dfrac{1}{n},\ n = 1, 2, \ldots. \end{cases}$$

EXAMPLE 7.5
$$\varphi(x) = \begin{cases} -\dfrac{1}{n}, & -\dfrac{1}{n} < x \le -\dfrac{1}{n+1},\ n = 1, 2, \ldots, \\ 0, & x = 0, \\ -\dfrac{1}{n}, & \dfrac{1}{n+1} \le x < \dfrac{1}{n},\ n = 1, 2, \ldots. \end{cases}$$

EXAMPLE 7.6
$$\varphi(x) = \begin{cases} -\dfrac{1}{n} + \dfrac{1}{n^2}(x + \dfrac{1}{n}), & -\dfrac{1}{n} < x \le -\dfrac{1}{n+1},\ n = 1, 2, \ldots, \\ 0, & x = 0, \\ \dfrac{1}{n} + \dfrac{1}{n^2}(x - \dfrac{1}{n+1}), & \dfrac{1}{n+1} < x \le \dfrac{1}{n},\ n = 1, 2, \ldots. \end{cases}$$

Examples 7.1, 7.2, 7.3 are traditional examples of stationary points. The functions in Examples 7.4, 7.5 are piecewise constant. In both cases $x = 0$ is not a point of local minimum. But in every its neighborhood there exists a point of local minimum. The function in Example 7.6 is discontinuous and piecewise linear and it does not have local minimums at all.

Acknowledgements

Research supported by the Belarusian Republican Fund of Fundamental Research grant F97-196 and the INTAS grant 97-1050.

References

[1] E. Asplund (1968) Fréchet differentiability of convex functions, *Acta Math.* **121**, 31-47.

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

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

[4] A. V. Dmitruk, A. A. Miljutin and N. P. Osmolovskii (1980) Ljusternik's theorem and the theory of extrema, *Russ. Math. Surv.* **35**, 11-51.

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

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

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

[8] M. Fabian and B. S. Mordukhovich (1999) Separable reduction and supporting properties of Fréchet-like normals in Banach spaces, *Canad. J. Math.* **51**, 26-48.

[9] A. D. Ioffe (1984) Approximate subdifferentials and applications. 1: the finite dimensional theory, *Trans. Amer. Math. Soc.* **281**, 389-416.

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

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

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

[13] A. Y. Kruger (1988) Covering theorem for set-valued mappings, *Optimization* **19**, 763-780.

[14] A. Y. Kruger (1996) On calculus of strict ε-semidifferentials, *Dokl. Akad. Nauk Belarus* **40**, no. 4, 34-39 (in Russian).

[15] A. Y. Kruger (1997) Strict ε-semidifferentials and extremality conditions, *Dokl. Akad. Nauk Belarus* **41**, no. 3, 21-26 (in Russian).

[16] A. Y. Kruger (1998) On extremality of set systems, *Dokl. Nat. Akad. Nauk Belarus* **42**, no. 1, 24-28 (in Russian).

[17] A. Y. Kruger (1998) On necessary and sufficient extremality conditions, in: *Dynamical Systems: Stability, Control, Optimization, Abstracts, Vol. 2, Minsk*, 165-168.

[18] A. Y. Kruger (2000) Strict (ε, δ)-semidifferentials and extremality of sets and functions, *Dokl. Nat. Akad. Nauk Belarus* **44**, no.2, 21-24 (in Russian).

[19] A. Y. Kruger (2002) On Fréchet subdifferentials, *J. Math. Sci.* (to appear).

[20] A. Y. Kruger and B. S. Mordukhovich (1980) Extremal points and the Euler equation in nonsmooth optimization problems, *Dokl. Akad. Nauk BSSR* **24**, no. 8, 684-687 (in Russian).

[21] B. S. Mordukhovich (1988) *Approximate Methods in Problems of Optimization and Control*, Nauka, Moscow (in Russian).

[22] B. S. Mordukhovich (2001) *The extremal principle and its applications to optimization and economics*, in: *Optimization and Related Topics*, A. Rubinov and B. Glover, eds., Kluwer Academic Publishers, 343-369.

[23] B. S. Mordukhovich and Y. Shao (1995) Differential characterizations of covering, metric regularity, and Lipschitzian properties of multifunctions between Banach spaces, *Nonlin. Anal.* **25**, 1401-1424.

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

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

[26] B. S. Mordukhovich and B. Wang (2000) On variational characterizations of Asplund spaces, in: *Constructive, Experimental and Nonlinear Analysis*, M. Thera, ed., Canad. Math. Soc. Conf. Proceed. Vol. 27, 245-254.

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