GENERALIZED DIFFERENTIALS OF NONSMOOTH FUNCTIONS, AND NECESSARY CONDITIONS FOR AN EXTREMUM

A. Ya. Kruger

UDC 517.988:519.853.4

In this article we introduce an abstract notion of extremeness, which similarly to the corresponding definition of Neustadt [1, 2] embraces different notions of a solution to extremal problems, and establish necessary conditions for extremeness. These conditions are then concretized for the problem of nondifferentiable (locally Lipschitz) programming in a Banach space with equality and inequality constraints. We give conditions for optimality in the form of a generalized condition for the Lagrange function to be stationary, which are generalizations of known results for smooth and convex problems [3-5]. Also we generalize and refine the results of Clarke [6], Warga [7], and Hiriart-Urruty [8].

In contrast to the approach used most widely in the literature [1, 2, 5, 9-12], our derivation of necessary conditions for extremeness is realized without using any local approximations of sets and functions. Here a generalized Euler equation [13, 14] plays the role of the separation theorem for convex sets and its modifications. The result characterizes the extreme position of a system of nonconvex sets in a space, and is a generalization of a theorem of Dubovitskii and Milyutin [9]. It enables us to extend many facts of convex analysis to nonconvex sets and functions.

The necessary conditions are stated in terms of generalized differentials and generalized normals, which are developments of the corresponding finite-dimensional constructions in [15-18]. They are used in [13, 15-17, 19-21] in the formulation of necessary conditions for optimality in nonsmooth problems of mathematical programming and of optimal control. In contrast to the majority of known generalizations of the notion of a derivative, generalized differentials can be nonconvex. They are more refined local characteristics of functions in comparison with Clarke's generalized gradients [6, 22-24] and with the derivative sets of Warga [7, 25].

The definitions and certain properties of generalized differentials are given in Sec. 1. In Sec. 2 we establish the generalized Euler equation. Abstract conditions for extremeness are proved in Sec. 3. Necessary conditions for an extremum for the general problem of nondifferentiable programming are stated in Sec. 4.

We use the following notation and terminology. The topological dual of a Banach space X is denoted by X*. Norms in a space and in its dual are denoted by $\|\cdot\|$ and $\|\cdot\|_*$, respectively. $\langle x^*, x\rangle$ stands for the value of a linear functional $x^* \in X^*$ at $x \in X$. Unless stated otherwise, a dual space is assumed to be equipped with the weak* topology. The symbols int and cl denote, respectively, interior and closure (the closure of a set in a dual space is understood to be the set of all weak* limit points of bounded sequences of elements of this set). The unit ball in a Banach space (with center at the zero element) is denoted by B. The sum of sets and the product of a set by a number are defined, respectively, by $U + V = \{u + v: u \in U, v \in V\}$, $\alpha U = \{\alpha u: u \in U\}$, where $\alpha \neq 0$ if $U = \emptyset$, and $0 \cdot \emptyset = \{0\}$. Operations on functions are defined in the usual way. Also it is assumed that $+\infty + (-\infty) = +\infty$ and $0 \cdot (\pm\infty) = 0$.

1. GENERALIZED DIFFERENTIALS

Let X be a real Banach space, φ an arbitrary function from X into $\overline{R} = [-\infty, \infty]$ that is finite at x^0. For any $\varepsilon \geqslant 0$ we introduce the convex set

$$\partial_\varepsilon^- \varphi(x^0) = \left\{x^* \in X^* : \lim_{x \to x^0} \inf \frac{\varphi(x) - \varphi(x^0) - \langle x^*, x - x^0\rangle}{\|x - x^0\|} \geqslant -\varepsilon\right\}$$

in the dual space. It is called the *lower ε-semidifferential of φ at* x^0. This construction is the simplest generalization of the notion of a Frechet derivative and of the subdifferential of a convex function [2-5]. It is used in the definition of the more complicated

Minsk. Translated from Sibirskii Matematicheskii Zhurnal, Vol. 26, No. 3, pp. 78-90, May-June, 1985. Original article submitted December 1, 1982.

 0037-4466/85/2603-0370$09.50

construction of a *lower generalized semidifferential*:

$$D^-\varphi(x^0) = \bigcap_{\varepsilon>0} \mathrm{cl} \bigcup \left[\partial_\varepsilon^- (\mathrm{cl}^\downarrow \varphi)(x) : \|x - x^0\| \leqslant \varepsilon, |\mathrm{cl}^\downarrow \varphi(x) - \varphi(x^0)| \leqslant \varepsilon\right]. \tag{1}$$

The symbol $\mathrm{cl}^\downarrow\varphi$ in (1) denotes the lower semicontinuous hull of φ, that is, the function defined by $\mathrm{cl}^\downarrow\varphi(x) = \liminf_{y\to x} \varphi(y)$, $x \in X$. When φ is continuous (in a neighborhood of x^0), (1) can obviously be simplified:

$$D^-\varphi(x^0) = \bigcap_{\varepsilon>0} \mathrm{cl} \bigcup \left[\partial_\varepsilon^- \varphi(x) : \|x - x^0\| \leqslant \varepsilon\right].$$

In the characterization of local properties of a function φ we use as well as the lower generalized semidifferential (1) the sets

$$D^+\varphi(x^0) = -D^-(-\varphi)(x^0), \tag{2}$$

$$D^0\varphi(x^0) = D^-\varphi(x^0) \cup D^+\varphi(x^0), \tag{3}$$

$$\widehat{D}\varphi(x^0) = D^-\varphi(x^0) \cup \left[-D^+\varphi(x^0)\right], \tag{4}$$

which are called, respectively, the *upper generalized semidifferential*, the *generalized differential*, and the *symmetric generalized differential of* φ *at* x^0. Let us note that the sets (1) and (2) can be nonempty simultaneously, and are not contained in one another (see the examples in [18, 21]); this justifies the introduction of the constructions (3) and (4).

If φ is continuous and takes values in a Banach space Z, then for each $z^* \in Z^*$ we can associate with it the scalar function $L_{z^*}(x) = \langle z^*, \varphi(x)\rangle$, $x \in X$. By using this we define the sets

$$D\varphi(x^0) = \bigcap_{\varepsilon>0} \mathrm{cl} \bigcup \left[(x^*, z^*) \in X^* \times Z^* : x^* \in \partial_\varepsilon^- L_{z^*}(x), \|x - x^0\| \leqslant \varepsilon\right],$$
$$D_{z^*}\varphi(x^0) = \{x^* \in X^* : (x^*, z^*) \in D\varphi(x^0)\}. \tag{5}$$

The set (5) is called the *generalized differential of* φ *at* x^0 *in the direction* z^*. We note that

$$D^- L_{z^*}(x^0) \subset D_{z^*}\varphi(x^0). \tag{6}$$

In what follows, when referring to the set of constructions (1)-(5) we shall call them *generalized differentials*.

We give below some properties of generalized differentials (for more details of the properties of ε-semidifferentials and generalized differentials see [14, 26]). Unless stated otherwise, all functions are assumed to act from X into $\bar{R}$, and to be finite at x^0.

1) If φ is strictly differentiable [4] at x^0, then

$$D^0\varphi(x^0) = D^-\varphi(x^0) = D^+\varphi(x^0) = \{\varphi'(x^0)\}.$$

2) If φ is convex, then

$$D^0\varphi(x^0) = D^-\varphi(x^0) = \{x^* \in X^* : \varphi(x) - \varphi(x^0) \geqslant \langle x^*, x - x^0\rangle \ \forall x \in X\}.$$

3) If φ satisfies a Lipschitz condition with a constant C in a neighborhood of x^0, then $\|x^*\|_* \leqslant C$ for any $x^* \in D^0\varphi(x^0)$.

4) If x^0 is a point of a local minimum of φ, then $0 \in D^-\varphi(x^0)$.

5) The multivalued map $D^-\varphi(\cdot)$ has the following lower semicontinuity type property:

$$D^-\varphi(x^0) = \bigcap_{\varepsilon>0} \mathrm{cl} \bigcup \left[D^-\mathrm{cl}\,\varphi(x) : \|x - x^0\| \leqslant \varepsilon, |\mathrm{cl}\,\varphi(x) - \varphi(x^0)| \leqslant \varepsilon\right].$$

6) If $\dim X < \infty$, then the set $\partial_\varepsilon^- (\mathrm{cl}^\downarrow\varphi)(x)$ in (1) can be replaced by $\partial_0^- (\mathrm{cl}^\downarrow\varphi)(x)$.

7) If $\lambda \in R$, then the set $D^-(\lambda\varphi)(x^0)$ coincides with $\lambda D^-\varphi(x^0)$ for all $\lambda \geqslant 0$, and with $\lambda D^+\varphi(x^0)$ for $\lambda \leqslant 0$.

8) If φ_2 is strictly differentiable at x^0, then

$$D^-(\varphi_1 + \varphi_2)(x^0) = D^-\varphi_1(x^0) + \varphi_2'(x^0).$$

9) The set $D^-\varphi(x^0)$ is contained in Clarke's generalized gradient $\partial_c\varphi(x^0)$ of φ at x^0 [22, 23], also the set $D^0\varphi(x^0)$ is contained in the symmetric generalized gradient [24].

10) If $\dim X < \infty$ and φ satisfies a Lipschitz condition in a neighborhood of x^0, then $D^0\varphi(x^0)$ is contained in any derivate set of Warga [25] of φ at x^0.

11) If φ is continuous from X to Z, then for equality to hold in (6) it is sufficient that one of the two following conditions holds:

a) $\dim Z < \infty$, and φ satisfies a Lipschitz condition in a neighborhood of x^0;

b) φ is strictly differentiable at x^0.

Properties 12)-16), given later, are established under the assumption that the space X satisfies the following Condition A. This condition is also used in the proof of the generalized Euler equation in the next section.

Condition A. There is on the space X a scalar function ψ such that $\psi(0) = 0$, $\psi(x) > 0$ if $x \neq 0$, ψ is continuous in a neighborhood U of the zero element and Frechet differentiable on $U \setminus \{0\}$, and $\|\psi'(x)\|_* \geqslant 1$ if $x \in U \setminus \{0\}$.

We note that Condition A holds, e.g., in any reflexive space.

12) If φ satisfies a Lipschitz condition in a neighborhood of x^0, then $D^-\varphi(x^0) \neq \varnothing$; also

$$\operatorname{clco} D^0\varphi(x^0) = \partial_C\varphi(x^0).$$

13) If at least one of φ_1, φ_2 satisfies a Lipschitz condition in a neighborhood of x^0, then

$$D^-(\varphi_1 + \varphi_2)(x^0) \subset D^-\varphi_1(x^0) + D^-\varphi_2(x^0).$$

14) If φ_i, $i = 1, 2, \ldots, m$, satisfy a Lipschitz condition in a neighborhood of x^0, and $\varphi(x) = \max\{\varphi_i(x): i = 1, 2, \ldots, m\}$, then

$$D^-\varphi(x^0) \subset \bigcup \left[\sum_{i=1}^m \lambda_i D^-\varphi_i(x^0) : \lambda_i \geqslant 0, \lambda_i(\varphi_i(x^0) - \varphi(x^0)) = 0, \ i = 1, 2, \ldots, m; \sum_{i=1}^m \lambda_i = 1\right].$$

15) If φ satisfies a Lipschitz condition in a domain containing the segment $[x_1, x_2]$, then there are a point $x^0 \in]x_1, x_2[$ and an element $x^* \in D^0\varphi(x^0)$ such that

$$\varphi(x_2) - \varphi(x_1) = \langle x^*, x_2 - x_1\rangle.$$

16) If a space Z satisfies Condition A, and if the functions $\varphi: X \to Z$ and $g: Z \to R$ satisfy a Lipschitz condition in a neighborhood of x^0 and $\varphi(x^0)$, respectively, then

$$D^-(g \circ \varphi)(x^0) \subset \bigcup \left[D_{z^*}\varphi(x^0) : z^* \in D^-g(\varphi(x^0))\right].$$

These properties of generalized differentials enable us to apply the constructions in the analysis of nonsmooth functions and in an investigation of nonsmooth optimization problems. As is clear from properties 9) and 10), generalized differentials have definite advantages compared with Clarke's generalized gradients and the derivate sets of Warga. Ioffe has established the following result, which shows that the lower generalized semidifferential is the best (in a certain sense) construction among all possible generalizations of the concept of a derivative that have the natural properties.

17) If $D^-\varphi(\cdot)$ is a generalized derivative for which properties 2) [not taking $D^0\varphi(x^0)$ into account], 4), 5), and 13) hold, then the set $D^-\varphi(x^0)$ contains the lower generalized semidifferential of φ at x^0.

We note that, in contrast to the majority of known generalizations of the notion of a derivative, the generalized differentials (1)-(4) can be nonconvex. For example, for a function $\varphi: R \to R$ defined by $\varphi(x) = -|x|$, $x \in R$, we have $D^-\varphi(0) = \{-1, 1\}$. Other examples of the calculation of generalized differentials can be found in [14, 18].

Let Ω be a nonempty set in X, and $x^0 \in \operatorname{cl}\Omega$. For any $\varepsilon \geqslant 0$ we define the convex set

$$N_\varepsilon(x^0|\Omega) = \left\{x^* \in X^* : \limsup_{\substack{x \to x^0 \\ x \in \Omega}} \frac{\langle x^*, x - x^0\rangle}{\|x - x^0\|} \leqslant \varepsilon\right\},$$

called the set of ε-*normal elements to* Ω *at* x^0. By using this construction we define the cone

$$K(x^0|\Omega) = \bigcap_{\varepsilon > 0} \operatorname{cl} \bigcup [N_\varepsilon(x|\Omega) : x \in \operatorname{cl}\Omega, \|x - x^0\| \leqslant \varepsilon], \tag{7}$$

which is called the *generalized normal cone to* Ω *at* x^0. The elements of the set (7) are called *generalized normals*.

It is not difficult to establish that the generalized normal cone (7) coincides with the lower generalized semidifferential (and with the generalized differential) at x^0 of the indicator function $\delta(\cdot|\mathrm{cl}\,\Omega)$ of the set $\mathrm{cl}\,\Omega$ ($\delta(x|\mathrm{cl}\,\Omega)=0$, if $x\in\mathrm{cl}\,\Omega$, and $\delta(x|\mathrm{cl}\,\Omega)=\infty$, if $x\notin\mathrm{cl}\,\Omega$). Thus, certain of the above-mentioned properties of generalized differentials carry over to generalized normal cones. In particular, when Ω is convex, the set (7) coincides with the normal cone to Ω in the sense of convex analysis.

On the other hand, the lower generalized semidifferential of a function φ can be defined in terms of the generalized normals to its overgraph $E(\varphi)=\{(x,\ \mu)\in X\times R:\varphi(x)\leqslant\mu\}$. Namely, the following representation is valid:

$$D^-\varphi(x^0)=\{x^*\in X^*:(x^*,\ -1)\in K(x^0,\ \varphi(x^0)|E(\varphi))\}.$$

A similar representation can be written down for the upper generalized semidifferential (in terms of the generalized normals to the subgraph).

Let M be a closed set in Z and φ a function from X to Z. We consider the set

$$\mathscr{E}(\varphi)=\{(x,\ z)\in X\times Z:\varphi(x)-z\in M\}. \tag{8}$$

If M = {0}, then the set (8) coincides with the graph of φ. If M is a convex cone defining a preordering relation in Z (a cone of nonpositive elements), then the set (8) can be treated as the overgraph of φ. The following theorem gives a representation of generalized normals to the set (8); its proof is in [14].

THEOREM 1. Let $(x^0,\ z^0)\in\mathscr{E}(\varphi)$, and let φ satisfy a Lipschitz condition in a neighborhood of x^0. Then

$$K(x^0,z^0|\mathscr{E}(\varphi))\subset\{(x^*,-z^*)\in X^*\times Z^*:x^*\in D_{z^*}\varphi(x^0),z^*\in K(\varphi(x^0)-z^0|M)\}.$$

If, in addition, $\dim Z<\infty$, then in the latter inclusion we have equality.

2. THE GENERALIZED EULER EQUATION

Let $\Omega_1,\ \Omega_2,\ldots,\Omega_n$ be nonempty closed sets in a space X. The system of sets Ω_i, i = 1, 2,...,n, is called an *extreme* system [13, 14, 19] if the following conditions hold: 1) $\bigcap_{i=1}^{n}\Omega_i\neq\emptyset$; 2) there are sequences $\{a_k^i\}\subset X$, i = 1, 2,...,n, converging to the zero element such that $\bigcap_{i=1}^{n}\Omega_{ik}=\varnothing$, where $\Omega_{ik}=\{x-a_k^i:x\in\Omega_i\}$, i = 1, 2,...,n, k = 1, 2,.... Also each point $x^0\in\bigcap_{i=1}^{n}\Omega_i$ is called an extreme point of the system of sets Ω_i, i = 1, 2,...,n.

The notion of an extreme system characterizes the mutual position of sets in a space. There is a close connection between the extremeness of a system of sets and separation. We give below a necessary condition for the extreme property of a system that plays a basic role in the derivation of necessary conditions for optimality in nonsmooth and nonconvex extremum problems. It is assumed that the space X satisfies Condition A.

THEOREM 2. Let x^0 be an extreme point of a system of closed sets Ω_i, i = 1, 2,...,n. Then for any $\varepsilon>0$ there are points $y_i\in\Omega_i\cap[x^0+\varepsilon B]$, $x_i^*\in N_\varepsilon(y_i|\Omega_i)$, i = 1, 2,...,n, such that $\sum_{i=1}^{n}\|x_i^*\|_*=1$ and

$$x_1^*+x_2^*+\ldots+x_n^*=0. \tag{9}$$

If, in addition, $X=X_1\times X_2$, $\dim X_2<\infty$, $x^0=(x_1^0,\ x_2^0)$, and the sets Ω_i, i = 1, 2,...,n − 1, have the form

$$\Omega_i=\{(x_1,\ x_2)\in X_1\times X_2:x_1\in\Omega_{i1},\ x_2=\varphi_i(x_1)\},$$

where Ω_{i1} is a closed set in X_1, epi-Lipschitzian at x_1^0,† and φ is a function from X_1 to X_2 satisfying a Lipschitz condition in a neighborhood of x_1^0, then there are elements $x_i^*\in K(x^0|\Omega_i)$, i = 1, 2,...,n, not simultaneously zero, such that (9) holds.

†A set $\Omega\subset X$ is called *epi-Lipschitzian* at $x^0\in\Omega$, if there are a $\delta>0$ and a nonempty open set $V\subset X$ such that $x+t\bar{x}\in\Omega$ for all $x\in\Omega\cap[x^0+\delta B]$, $t\in(0,\ \delta)$, $\bar{x}\in V$ [27].

For n = 2, Theorem 2 can be regarded as a generalization of the classical theorem on the separation of convex sets (in the statement of Theorem 2 the possibility that $\dim X_2 = 0$ is not excluded). In the general case it generalizes the theorem of Dubovitskii and Milyutin [9]. By analogy with this result the relation (9) is called a generalized Euler equation. The condition that the sets are epi-Lipschitzian is an analogue of the traditional condition of solidness. We note that in the case of a finite-dimensional space X, the assertion in the second part of Theorem 2 follows from the first part without additional assumptions regarding the structure of the sets Ω_i, $i = 1, 2, \ldots, n-1$. Also we obtain a generalization of the corresponding result in [11].

Proof of Theorem 2. We first consider the case n = 2. Let ψ be a scalar function on X whose existence follows from Condition A. It follows from the extreme property of the system of sets Ω_1, Ω_2 that for any $\varepsilon > 0$ there is an $a \in X$ such that $\psi(a) \leqslant \varepsilon^2$ and $\psi(x - y + a) > 0$ for all $x \in \Omega_1$ and $y \in \Omega_2$. By using Ekeland's variational principle [28], we conclude that there are points $y_i \in \Omega_i \cap [x^0 + \varepsilon B]$, $i = 1, 2$, such that $\psi(x - y + a) - \psi(y_1 - y_2 + a) + \varepsilon\|(x, y) - (y_1, y_2)\| \geqslant 0$ for all $x \in \Omega_1$ and $y \in \Omega_2$. Without loss of generality we may assume that ψ is differentiable at $y_1 - y_2 + a$. We set $x^* = -\psi'(y_1 - y_2 + a)$. It follows from the above inequality that $x^* \in N_\varepsilon(y_1|\Omega_1)$ and $-x^* \in N_\varepsilon(y_2|\Omega_2)$. Also $\|x^*\|_* \geqslant 1$. By putting $x_1^* = x^*/2\|x^*\|_*$ and $x_2^* = -x_1^*$ we complete the proof of the first part of Theorem 2.

Now we prove the second part. By what was proved above there are sequences $\{y_{1k}\} \subset \Omega_1$, $\{y_{2k}\} \subset \Omega_2$, $\{x_k^*\} \subset X^*$, and $\{\varepsilon_k\} \subset R_+$, such that $x_k^* \in N_{\varepsilon_k}(y_{1k}|\Omega_1)$, $-x_k^* \in N_{\varepsilon_k}(y_{2k}|\Omega_2)$, $\|x_k^*\|_* = 1$, $k = 1, 2, \ldots$, $y_{1k} \to x^0$, $y_{2k} \to x^0$, and $\varepsilon_k \to 0$ as $k \to \infty$. Let x^* be the limit point of $\{x_k^*\}$. Obviously $x^* \in K(x^0|\Omega_1)$, and $-x^* \in K(x^0|\Omega_2)$. We have to prove that $x^* \neq 0$.

We assume that Ω_1 is epi-Lipschitzian at x^0. Then by definition there are a $\delta > 0$ and an $\bar{x} \in X$ such that $x + t\tilde{x} \in \Omega_1$ for all $t \in (0, \delta)$, $x \in \Omega_1 \cap [x^0 + \delta B]$, and $\tilde{x} \in \bar{x} + \delta B$. For every $k = 1, 2, \ldots$ we can find an $\bar{x}_k \in \bar{x} + \delta B$, such that $\langle x_k^*, \bar{x}_k - \bar{x}\rangle > \frac{\delta}{2}$. Also $\langle x_k^*, \bar{x}_k\rangle = \lim\limits_{t \to +0} t^{-1}\langle x_k^*, (y_{1k} + t\bar{x}_k) - y_{1k}\rangle \leqslant \varepsilon_k\|\bar{x}_k\| \leqslant \frac{\delta}{4}$ for a sufficiently large k; hence $\langle x_k^*, \bar{x}\rangle \leqslant -\frac{\delta}{4}$. It follows from here that $\langle x^*, \bar{x}\rangle \leqslant -\frac{\delta}{4}$. Thus, $x^* \neq 0$.

Now suppose that $X = X_1 \times X_2$, $\dim X_2 < \infty$, $x^0 = (x_1^0, x_2^0)$, and Ω_1 has the form

$$\Omega_1 = \{(x_1, x_2) \in X_1 \times X_2 : x_1 \in \Omega_{11}, \quad x_2 = \varphi(x_1)\},$$

where Ω_{11} is a closed set in X_1 that is epi-Lipschitzian at x_1^0, and φ is a function from X_1 to X_2 that satisfies a Lipschitz condition in a neighborhood of x_1^0. Then $X^* = X_1^* \times X_2^*$, $\dim X_2^* = \dim X_2 < \infty$, $y_{1k} = (x_{1k}, \varphi(x_{1k}))$, $x_k^* = (x_{1k}^*, x_{2k}^*)$, $k = 1, 2, \ldots$, and $x^* = (x_1^*, x_2^*)$. By hypothesis for every $k = 1, 2, \ldots$ we have

$$\limsup_{\substack{x \to x_{1k} \\ x \in \Omega_{11}}} \frac{\langle x_{1k}^*, x - x_{1k}\rangle + \langle x_{2k}^*, \varphi(x) - \varphi(x_{1k})\rangle}{\|(x, \varphi(x)) - (x_{1k}, \varphi(x_{1k}))\|} \leqslant \varepsilon,$$

and so

$$\limsup_{\substack{x \to x_{1k} \\ x \in \Omega_{11}}} \frac{\langle x_{1k}^*, x - x_{1k}\rangle}{\|(x, \varphi(x)) - (x_{1k}, \varphi(x_{1k}))\|} \leqslant \varepsilon_k + \|x_{2k}^*\|_*.$$

Since φ is Lipschitz there is a $C > 1$ such that $\|(x, \varphi(x)) - (x_{1k}, \varphi(x_{1k}))\| \leqslant C\|x - x_{1k}\|$, and so

$$\limsup_{\substack{x \to x_{1k} \\ x \in \Omega_{11}}} \frac{\langle x_{1k}^*, x - x_{1k}\rangle}{\|x - x_{1k}\|} \leqslant \tilde{\varepsilon}_k,$$

where $\tilde{\varepsilon}_k = C(\varepsilon_k + \|x_{2k}^*\|_*)$. Thus, $x_{1k}^* \in N_{\tilde{\varepsilon}_k}(x_{1k}|\Omega_{11})$, $k = 1, 2, \ldots$. We assume that $x_2^* = 0$. Without loss of generality we may assume that $x_{2k}^* \to 0$ as $k \to \infty$. Then $\|x_{1k}^*\|_* \to 1$ and $\tilde{\varepsilon}_k \to 0$ as $k \to \infty$. We put $y_k^* = x_{1k}^*/\|x_{1k}^*\|_*$ and $\hat{\varepsilon}_k = \tilde{\varepsilon}_k/\|x_{1k}^*\|_*$. Obviously, $y_k^* \in N_{\hat{\varepsilon}_k}(x_{1k}|\Omega_{11})$, $\|y_k^*\|_* = 1$, $k = 1, 2, \ldots$, $\hat{\varepsilon}_k \to 0$ as $k \to 0$ and x_1^* is a limit point of $\{y_k^*\}$. Since Ω_{11} is epi-Lipschitzian at x_1^0, by what has been proved above we have $x_1^* \neq 0$. This proves the theorem for n = 2.

The general case reduces to the one we have just proved. For if x^0 is an extreme point of a system $\Omega_1, \Omega_2, \ldots, \Omega_n$ of closed sets, then it is not difficult to establish that $\tilde{x}^0 = (x^0, x^0, \ldots, x^0) \in X^{n-1}$ is extreme for the sets $\widetilde{\Omega}_1 = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_{n-1}$ and $\widetilde{\Omega}_2 = \{(x, x, \ldots, x) \in X^{n-1} : x \in \Omega_n\}$.

It is obvious that X^{n-1} satisfies Condition A. To complete the proof of Theorem 2 it is now enough to use elementary representations of sets of ε-normal elements and generalized normal cones to $\tilde{\Omega}_1$ and $\tilde{\Omega}_2$ [26].

3. ABSTRACT CONDITIONS FOR EXTREMENESS

Let Ω and M be closed sets in Banach spaces X and Z, respectively, and let φ be a function from X to Z. We assume that it is M-*semicontinuous* [this means that the set $\mathscr{E}(\varphi)$ defined by (8) is closed in $X \times Z$]. A point $x^0 \in \Omega$ is called (φ, Ω, M)-*extreme* if $\varphi(x^0) \in M$ and there is a sequence $\{z_k\} \subset Z$, converging to the zero element, such that $[\varphi(\Omega) + z_k] \cap M = \varnothing$, $k = 1, 2, \ldots$.

The concept of a (φ, Ω, M)-extreme point embraces, similarly to extremeness in the sense of Neustadt [1, 2], different concepts of a solution to optimization problems with scalar and vector criteria and also to minimax problems. Theorem 3, proved below, gives necessary conditions for (φ, Ω, M)-extremeness. It is assumed that the spaces X and Z satisfy Condition A.

THEOREM 3. Let x^0 be a (φ, Ω, M)-extreme point. Then for any $\varepsilon > 0$ there are elements $(y_1, z) \in \mathscr{E}(\varphi) \cap [(x^0, 0) + \varepsilon B]$, $y_2 \in \Omega \cap [x^0 + \varepsilon B]$, $x^* \in X^*$, and $z^* \in Z^*$ such that $\|x^*, z^*\|_* = 1$,

$$(x^*, z^*) \in N_\varepsilon(y_1, z | \mathscr{E}(\varphi)), \quad -x^* \in N_\varepsilon(y_2 | \Omega).$$

If, in addition, $Z = Z_1 \times Z_2$, $\dim Z_2 < \infty$, $M = M_1 \times \{0\}$, $\varphi = (\varphi_1, \varphi_2)$ satisfies a Lipschitz condition in a neighborhood of x^0, and M_1 is epi-Lipschitzian at $\varphi_1(x^0)$, then there is a nonzero element $z^* = (z_1^*, z_2^*) \in Z_1^* \times Z_2^*$ such that

$$0 \in D_{z^*}\varphi(x^0) + K(x^0 | \Omega), \tag{10}$$

$$z_1^* \in K(\varphi_1(x^0) | M_1). \tag{11}$$

This theorem generalizes and refines known necessary conditions for an extremum in optimization problems. The component φ_2 of φ corresponds to equality type constraints. In the conditions of the second part of Theorem 3, the assumption that Z_2 is finite-dimensional means that in an explicit form we allow only a finite number of these constraints. The condition that M_1 is epi-Lipschitzian is an analogue of the solidness condition of the cone that is usually used to define inequality type constraints. Condition (10) is an abstract multiplier rule, and (11) contains the traditional conditions on the sign of multipliers and conditions of complementing nonrigidity.

Proof of Theorem 3. We consider in $X \times Z$, as well as the set $\mathscr{E}(\varphi)$, the set $\tilde{\Omega} = \Omega \times \{0\}$. Under our assumptions these sets are closed and form an extreme system; moreover $(x^0, 0)$ is an extreme point for them. Obviously, the space $X \times Z$ satisfies Condition A. The assertion of the first part of Theorem 3 follows from Theorem 2 with due regard for the elementary representation

$$N_\varepsilon(y_2, 0 | \tilde{\Omega}) = N_\varepsilon(y_2 | \Omega) \times Z^*.$$

Suppose that the assumptions of the second part of Theorem 3 hold. We again use Theorem 2. The set $\mathscr{E}(\varphi)$ is representable as

$$\mathscr{E}(\varphi) = \{(x, z_1, z_2) \in X \times Z_1 \times Z_2 : (x, z_1) \in \mathscr{E}(\varphi_1), \quad z_2 = \varphi_2(x)\},$$

where $\mathscr{E}(\varphi_1) = \{(x, z_1) \in X \times Z_1 : \varphi_1(x) - z_1 \in M_1\}$. We show that $\mathscr{E}(\varphi_1)$ is epi-Lipschitzian at $(x^0, 0)$. Since M_1 is epi-Lipschitzian, there are a $\delta > 0$ and a $\bar{z}_1 \in Z_1$ such that $z_1 + t\tilde{z}_1 \in M_1$ for all $t \in (0, \delta)$, $z_1 \in M_1 \cap [\varphi_1(x^0) + \delta B]$, $\tilde{z}_1 \in \bar{z}_1 + \delta B$. We choose $\delta_1 > 0$ so that $\|\varphi_1(x) - \varphi_1(x^0) - z_1\| \leq \delta$, if $(x, z_1) \in (x^0, 0) + \delta_1 B$ and $\delta_1 \leq \frac{\delta}{2}\min(1, C^{-1})$, where C is the Lipschitz constant of φ. Then $\tilde{z}_1 + t^{-1}(\varphi_1(x + t\tilde{x}) - \varphi_1(x)) \in \bar{z}_1 + \delta B$, if $(x, z_1) \in (x^0, 0) + \delta_1 B$, $t \in (0, \delta_1)$, $(\tilde{x}, \tilde{z}_1) \in (0, \bar{z}_1) + \delta_1 B$. If $\varphi_1(x) - z_1 \in M_1$, then since M_1 is epi-Lipschitzian it follows that $\varphi_1(x + t\tilde{x}) - (z_1 - t\tilde{z}_1) \in M_1$. The latter means that the set $\mathscr{E}(\varphi_1)$ is epi-Lipschitzian at $(x^0, 0)$. By Theorem 2 there is a nonzero element $(x^*, -z^*) \in K(x^0, 0 | \mathscr{E}(\varphi))$ such that $(-x^*, z^*) \in K(x^0, 0 | \tilde{\Omega})$. Taking into account the representation of a generalized normal cone to Ω, conditions (10) and (11) follow from Theorem 1.

Now we prove that $z^* \neq 0$. By the definition of a generalized normal cone there are sequences $\{x_k, z_{1k}\} \subset \mathscr{E}(\varphi_1)$, $\{x_k^*, z_{1k}^*, z_{2k}^*\} \subset X^* \times Z_1^* \times Z_2^*$, and $\{\varepsilon_k\} \subset R_+$ such that $(x_k^*, -z_{1k}^*, -z_{2k}^*) \in N_{\varepsilon_k}(x_k, z_{1k}, \varphi_2(x_k) | \mathscr{E}(\varphi))$, $k = 1, 2, \ldots$, $(x_k, z_{1k}) \to (x^0, 0)$, and $\varepsilon_k \to 0$ as $k \to \infty$, and $(x^*, z^*) \in X^* \times Z^*$ is a limit point of $\{x_k^*, z_{1k}^*, z_{2k}^*\}$. Furthermore, let $\delta > 0$ and $\bar{z}_1 \in Z_1$ be such that $z_1 + t\tilde{z}_1 \in M_1$, if $z_1 \in M_1 \cap [\varphi_1(x^0) + \delta B)$, $t \in (0, \delta)$, $\tilde{z}_1 \in \bar{z}_1 + \delta B$, and $\bar{x}$ is an arbitrary point from $(\delta/C)B$. Then

$$\varphi_1(x_k + t\bar{x}) - (z_{1k} - t\bar{z}_1) = \varphi_1(x_k) - z_{1k} + t(\bar{z}_1 + t^{-1}(\varphi_1(x_k + t\bar{x}) - \varphi(x_k))) \in M_1$$

for all $t \in (0, \delta)$ and all sufficiently large k. We have

$$\limsup_{t \to +0} \frac{\langle x_k^*, \bar{x}\rangle + \langle z_{1k}^*, \bar{z}_1\rangle - \langle z_{2k}^*, t^{-1}(\varphi_2(x_k + t\bar{x}) - \varphi_2(x_k))\rangle}{\|\bar{x}, \bar{z}_1, t^{-1}(\varphi_2(x_k + t\bar{x}) - \varphi_2(x_k))\|} \leqslant \varepsilon_k,$$

and so

$$\langle x_k^*, \bar{x}\rangle + \langle z_{1k}^*, \bar{z}_1\rangle \leqslant \varepsilon_k((C+1)\|\bar{x}\| + \|\bar{z}_1\|) + C\|z_{2k}^*\|_* \cdot \|\bar{x}\|.$$

If we now assume that $z^* = 0$, then $x^* \neq 0$, and without loss of generality we may assume that $\|z_{2k}^*\|_* \to 0$, $\langle z_{1k}^*, \bar{z}_1\rangle \to 0$, and $\langle x_k^*, \bar{x}\rangle \to \langle x^*, \bar{x}\rangle$ as $k \to \infty$. From the last inequality we obtain $\langle x^*, \bar{x}\rangle \leqslant 0$, which is impossible since $\bar{x}$ is an arbitrary point from $(\delta/C)B$. This contradiction proves that $z^* \neq 0$, and the theorem is proved.

4. NECESSARY CONDITIONS FOR AN EXTREMUM IN PROBLEMS OF NONDIFFERENTIABLE PROGRAMMING

We illustrate an application of Theorem 3 for the derivation of necessary conditions for an extremum by the following example of a mathematical programming problem:

$$\varphi_0(x) \to \min, \tag{12}$$

$$\varphi_1(x) \in M_1, \tag{13}$$

$$\varphi_2(x) = 0, \tag{14}$$

$$x \in \Omega. \tag{15}$$

Here Ω and M_1 are closed sets in Banach spaces X and Z_1, respectively, $\varphi_0 : X \to R$, $\varphi_1 : X \to Z_1$, and $\varphi_2 : X \to Z_2$. It is assumed that the spaces X and Z_1 satisfy Condition A, and that Z_2 is finite-dimensional.

Let x^0 be a local solution to the problem (12)-(15), where φ_0, φ_1, and φ_2 satisfy a Lipschitz condition in a neighborhood of x^0, and M_1 is epi-Lipschitzian at $\varphi_1(x^0)$.

Obviously, x^0 is $(\tilde{\varphi}, \tilde{\Omega}, \tilde{M})$-extreme for some $\delta > 0$, where $\tilde{\varphi} = (\varphi_0 - \varphi_0(x^0), \varphi_1, \varphi_2)$, $\tilde{\Omega} = \Omega \cap [x^0 + \delta B]$, and $\tilde{M} = R_- \times M_1 \times \{0\}$. Also $R \times Z_1 \times Z_2$ satisfies Condition A, and all the assumptions of the second part of Theorem 3 hold. An application of this theorem leads to the following result.

THEOREM 4. If x^0 is a local solution to the problem (12)-(15), then there is a nonzero $z^* = (\lambda_0, z_1^*, z_2^*) \in R_+ \times Z_1^* \times Z_2^*$ such that

$$0 \in D_{z^*}(\varphi_0, \varphi_1, \varphi_2)(x^0) + K(x^0 \mid \Omega)$$

and condition (11) holds.

Taking account of the properties of generalized differentials in Sec. 1, Theorem 4 generalizes known results for the smooth and the convex cases [3-5]. It also generalizes and refines results of Clarke, Warga (for the case when all the functions are defined in a neighborhood of x^0), and others [6-8, 13, 19, 20].

We note that we can derive from Theorem 3 necessary conditions for an extremum for the problem (12)-(15) without assuming that φ_0, φ_1, and φ_2 are Lipschitz, and that M_1 is epi-Lipschitzian. These conditions are stated in terms of ε-normal elements to the corresponding sets (see [21]).

If M_1 is convex, then the assumption that it is epi-Lipschitzian is equivalent to the solidness condition $\operatorname{int} M_1 \neq \emptyset$. Also (11) is written in the form

$$\langle z_1^*, z_1 - \varphi_1(x^0)\rangle \leqslant 0 \quad \text{for all} \quad z_1 \in M_1.$$

In the special case $Z_1 = R^m$, $M_1 = R_-^m$, and $Z_2 = R^n$, Theorem 4 gives a Lagrange multiplier rule for the problem of nondifferentiable programming with a finite number of equalities and inequalities

$$\varphi_i(x) \leqslant 0, \quad i = 1, 2, \ldots, m, \tag{16}$$

$$\varphi_i(x) = 0, \quad i = m+1, \ldots, m+n, \tag{17}$$

where φ_i, $i = 1, 2, \ldots, m+n$, are scalar functions on X satisfying a Lipschitz condition in a neighborhood of x^0.

COROLLARY 1. If x^0 is a local solution to the problem (12), (15)-(17), then there are numbers λ_i, $i = 0, 1, \ldots, m+n$, not simultaneously zero, such that

$$\lambda_i \geqslant 0, \quad i = 0, 1, \ldots, m, \tag{18}$$

$$\lambda_i \varphi_i(x^0) = 0, \quad i = 1, 2, \ldots, m, \tag{19}$$

$$0 \in D^- L_\lambda(x^0) + K(x^0 | \Omega), \tag{20}$$

where $L_\lambda(x) = \sum_{i=0}^{m+n} \lambda_i \varphi_i(x)$, $x \in X$.

The multiplier rule (20) is a generalized condition for the Lagrange function L_λ to be stationary. By applying the properties 7) and 13) of generalized differentials (Sec. 1), we can deduce from it a multiplier rule in terms of the generalized differentials of the individual functions φ_i, $i = 0, 1, \ldots, m+n$.

COROLLARY 2. If x^0 is a local solution to the problem (12), (15)-(17), then there are non-negative numbers λ_i, $i = 0, 1, \ldots, m+n$, not simultaneously zero, such that (19) and the following condition hold:

$$0 \in \sum_{i=0}^{m} \lambda_i D^- \varphi_i(x^0) + \sum_{i=m+1}^{m+n} \lambda_i \widehat{D} \varphi_i(x^0) + K(x^0 | \Omega). \tag{21}$$

In contrast to the classical rule for Lagrange multipliers and known generalizations of it, in the multiplier rule (21), to equalities and inequalities (in a performance criterion) there correspond different sets, namely, the analogues of a derivative. In addition, the statement of Corollary 2 asserts that all multipliers (corresponding both to inequalities and a performance criterion and to equalities) are non-negative. This is explained by specifying equality type constraints and by a feature of the definition of a symmetric generalized differential (4). The multiplier rule can be stated in a more useful form by using the construction of a generalized differential (3).

COROLLARY 3. If x^0 is a local solution to the problem (12), (15)-(17), then there are numbers λ_i, $i = 0, 1, \ldots, m+n$, not simultaneously zero, such that (18), (19) and the following condition hold:

$$0 \in \sum_{i=0}^{m+n} \lambda_i D^0 \varphi_i(x^0) + K(x^0 | \Omega). \tag{22}$$

In contrast to the differentiable case, the relations (21) and (22), written in terms of generalized differentials of the individual functions, and (20), written by using the Lagrange function, are not equivalent. Corollary 2 gives, in general, weaker necessary conditions for an extremum compared with Corollary 1. In turn, Corollary 3 is obtained as a result of a certain weakening of the assertion of Corollary 2 (nevertheless it gives stronger necessary conditions compared with Clarke's theorem [6]). We can also state a result which is "intermediate" between Corollaries 2 and 3, if we use generalized differentials only for the indices corresponding to equality type constraints, and for the remaining indices we retain, in the statement of the multiplier rule, lower generalized semidifferentials.

For the problems (12)-(15), and (12), (15)-(17) considered above we can state normality conditions that guarantee non-zero multipliers λ_0 in the assertions of Theorem 4 and Corollaries 1-3 (see [21]). These conditions are formulated in terms of the tangent cones of Clarke and Clarke's generalized directional drivatives [22, 23, 27]. The conditions are a generalization of the conditions for the linear dependence of gradients in differentiable programming and Slater's condition in convex programming.

Now we give examples of the application of the necessary conditions for an extremum established in this section. We consider the following two problems of mathematical programming in R^2:

$$(P_1) \begin{cases} |x_1| - |x_2| \to \min \\ 2x_1 + |x_2| = 0, \end{cases} \qquad (P_2) \begin{cases} |x_1| + 2x_2 \to \min, \\ ||x_1| + x_2| + \frac{1}{2} x_2 = 0. \end{cases}$$

The functions $\varphi_1(x) = |x_1| - |x_2|$, $\varphi_2(x) = 2x_1 + |x_2|$, $\varphi_3(x) = |x_1| + 2x_2$, $\varphi_4(x) = ||x_1| + x_2|$, and $\varphi_5(x) = \varphi_4(x) + x_2/2$, $x = (x_1, x_2) \in R^2$, in Problems (P_1) and (P_2) do not have a derivative at the zero element. By using Corollary 2 we clarify whether the point 0 is a solution to either of the Problems

(P_1), (P_2). For this we have to find the lower generalized semidifferentials of φ_1 and φ_3, and the symmetric generalized differentials of φ_2 and φ_5. With due regard for property 8) in Sec. 1, it is easy to establish that $D^-\varphi_3(0)=\{(\alpha,2):-1\leqslant\alpha\leqslant 1\}$ and $\widehat{D}\varphi_2(0)=\{(2,\alpha):-1\leqslant\alpha\leqslant 1\}\cup(-2,-1)\cup(-2,1)$. The generalized semidifferentials of φ_1 and φ_4 are calculated from the definition [18, 21]: $D^-\varphi_1(0)=\{(\alpha,\pm 1):-1\leqslant\alpha\leqslant 1\}$, $D^-\varphi_4(0)=\{(\alpha_1,\ \alpha_2):|\alpha_1|\leqslant\alpha_2\leqslant 1\}\cup\{(\pm\alpha,\ \alpha):-1\leqslant\alpha\leqslant 0\}$, $D^+\varphi_4(0)=\{(\alpha,-1):-1\leqslant\alpha\leqslant 1\}\cup(-1,1)\cup(1,1)$. For φ_5, by property 8) we have $\widehat{D}\varphi_5(0)=\left\{(\alpha_1,\alpha_2):|\alpha_1|+\frac{1}{2}\leqslant\alpha_2\leqslant\frac{3}{2}\right\}\cup\left\{\left(\pm\alpha,\alpha+\frac{1}{2}\right):-1\leqslant\alpha\leqslant 0\right\}\cup\left\{\left(\alpha,\frac{1}{2}\right):-1\leqslant\alpha\leqslant 1\right\}\cup\left(-1,-\frac{3}{2}\right)\cup\left(1,\ -\frac{3}{2}\right)$. It is not difficult to represent all these sets graphically on the plane R^2.

We also observe that 0 does not belong to any of the sets $\lambda_1 D^-\varphi_1(0)+\lambda_2\widehat{D}\varphi_2(0)$, $\lambda_1 D^-\varphi_3(0)+\lambda_2\widehat{D}\varphi_5(0)$ for any $\lambda_1\geqslant 0$, $\lambda_2\geqslant 0$ not simultaneously zero. Hence it follows from Corollary 2 that the point 0 is not a solution to either of the Problems (P_1), (P_2).

We remark that one is unable to make this deduction from either Clarke's theorem [6] (and its strengthened version [8]) or Warga's theorem [7]. For example, in the case of Clarke's theorem, by the definition of a generalized gradient [22] we have $\partial_C\varphi_1(0)=\{(\alpha_1,\alpha_2):-1\leqslant\alpha_1\leqslant 1,\ -1\leqslant\alpha_2\leqslant 1\}$ and $\partial_C\varphi_5(0)=\left\{(\alpha_1,\alpha_2):-1\leqslant\alpha_1\leqslant 1,\ -\frac{1}{2}\leqslant\alpha_2\leqslant\frac{3}{2}\right\}$ [these representations can also be obtained by using property 12) in Sec. 1], and each of these sets contains the zero element. Thus, without calculating the generalized gradients of φ_1 and φ_5, we can make the deduction that Clarke's necessary conditions hold for the Problems (P_1), (P_2).

It is clear from the above arguments that the application of Theorem 4 and its corollaries is connected with the calculation (sometimes unwieldy) of the generalized differentials of various nonsmooth functions. In certain cases these calculations can be avoided by using certain properties in Sec. 1. Thus, the function φ_1 is representable as the sum of the functions $(x_1, x_2)\to|x_1|$ and $(x_1, x_2)\to-|x_2|$, whose generalized differentials are easily calculated directly, and it follows from property 13) that $D^-\varphi_1(0)$ is contained in $\{(\alpha,\pm 1):-1\leqslant\alpha\leqslant 1\}$ (actually, as we remarked above, these sets coincide), and it follows from Corollary 2 that 0 is not a solution to the Problem (P_1). Similarly, φ_4 is representable as the superposition of the functions $(x_1, x_2)\to|x_1|+x_2$ and $x\to|x|$, and the required estimates of its generalized semidifferentials can be obtained from property 16).

Remark. If we use the dual criterion of a covering for nonsmooth operators that is given in [29], then we can obtain a result close to Theorem 3 without assuming that the space X satisfies Condition A. An application of the approximate subdifferentials introduced in [30] enables us, in general, to dispense with Condition A.

LITERATURE CITED

1. L. W. Neustadt, "A general theory of extremals," J. Comput. System Sci., 3, No. 1, 57-92 (1969).
2. L. W. Neustadt, Optimization: A Theory of Necessary Conditions, Princeton Univ. Press (1976).
3. A. D. Ioffe and V. M. Tikhomirov, Theory of Extremal Problems [in Russian], Nauka, Moscow (1974).
4. V. M. Alekseev, V. M. Tikhomirov, and S. V. Fomin, Optimal Control [in Russian], Nauka, Moscow (1979).
5. B. N. Pshenichnyi, Convex Analysis and Extremal Problems [in Russian], Nauka, Moscow (1980).
6. F. H. Clarke, "A new approach to Lagrange multipliers," Math. Oper. Res., 1, No. 2, 165-174 (1976).
7. J. Warga, "Controllability and a multiplier rule for nondifferentiable optimization problems," SIAM J. Control Optim., 16, No. 5, 803-812 (1978).
8. J.-B. Hiriart-Urruty, "Refinements of necessary optimality conditions in nondifferentiable programming. I," Appl. Math. Optim., 5, No. 1, 63-82 (1979).
9. A. Ya. Dubovitskii and A. A. Milyutin, "Extremum problems in the presence of constraints," Zh. Vychisl. Mat. Mat. Fiz., 5, No. 3, 395-453 (1965).
10. H. Halkin, "Necessary conditions in mathematical programming and optimal control theory," Lecture Notes Econ. Math. Syst., 105, Springer-Verlag, Berlin (1974), pp. 113-165.
11. V. G. Boltyanskii, "The method of tents in the theory of extremum problems," Usp. Mat. Nauk, 30, No. 3, 3-65 (1975).

12. A. G. Kusraev, "Convex approximations of nonsmooth operators and necessary conditions for an extremum," in: Optimization [in Russian], No. 21, Novosibirsk (1978), pp. 41-54.
13. A. Ya. Kruger and B. Sh. Mordukhovich, "Extremal points and Euler's equation in nonsmooth optimization problems," Dokl. Akad. Nauk BSSR, 24, No. 8, 684-687 (1980).
14. A. Ya. Kruger, Generalized Differentials of Nonsmooth Functions [in Russian], Minsk (1980), Dep. at VINITI, No. 1332-81.
15. B. Sh. Mordukhovich, "The maximum principle in an optimal speed of response problem with nonsmooth constraints," Prikl. Mat. Mekh., 40, No. 6, 1014-1023 (1976).
16. B. Sh. Mordukhovich and A. Ya. Kruger, "Necessary conditions for optimality in a terminal control problem with nonfunctional constraints," Dokl. Akad. Nauk BSSR, 20, No. 12, 1064-1067 (1976).
17. A. Ya. Kruger and B. Sh. Mordukhovich, "The minimization of nonsmooth functionals in optimal control problems," Izv. Akad. Nauk SSSR, Tekh. Kibern., No. 4, 176-182 (1978).
18. A. Ya. Kruger and B. Sh. Mordukhovich, Generalized Normals and Derivatives, and Necessary Conditions for an Extremum in Problems of Nondifferentiable Programming. I [in Russian], Minsk (1980), Dep. at VINITI, No. 408-80.
19. A. Ya. Kruger and B. Sh. Mordukhovich, Generalized Normals and Derivatives, and Necessary Conditions for an Extremum in Problems of Nondifferentiable Programming. II [in Russian], Minsk (1980), Dep. at VINITI, No. 494-80.
20. B. Sh. Mordukhovich, "Metric approximations and necessary conditions for optimality for general classes of nonsmooth extremal problems," Dokl. Akad. Nauk SSSR, 254, No. 5, 1072-1076 (1980).
21. A. Ya. Kruger, Necessary Conditions for an Extremum in Problems of Nonsmooth Optimization [in Russian], Minsk (1981), Dep. at VINITI, No. 1333-81.
22. F. H. Clarke, "Generalized gradients and applications," Trans. Am. Math. Soc., 204, 247-262 (1975).
23. F. H. Clarke, "Generalized gradients of Lipschitz functions," Adv. Math., 40, No. 1, 52-67 (1981).
24. J.-H. Hiriart-Urruty, "Tangent cones, generalized gradients and mathematical programming in Banach spaces," Math. Oper. Res., 4, No. 1, 79-97 (1979).
25. J. Warga, Optimal Control of Differential and Functional Equations, Academic Press, New York (1972).
26. A. Ya. Kruger, ε-Semidifferentials and ε-Normal Elements [in Russian], Minsk (1981), Dep. at VINITI, No. 1331-81.
27. R. T. Rockafellar, "Generalized directional derivatives and subgradients of nonconvex functions," Can. J. Math., 32, No. 2, 257-280 (1980).
28. I. Ekeland, "On the variational principle," J. Math. Anal. Appl., 47, No. 2, 324-353 (1974).
29. A. Ya. Kruger, "On a characteristic of the covering property for nonsmooth operators," in: School on the Theory of Operators in Function Spaces (Abstracts of Papers) [in Russian], Minsk (1982), pp. 94-95.
30. A. D. Ioffe, "Sous-différentielles approchées de fonctions numériques," C. R. Acad. Sci. Paris, Sér. 1, 292, No. 14, 675-678 (1981).