

Necessary Optimality Conditions for Nonsmooth Problems with Operator Constraints

H.-P. SCHEFFLER and W. SCHIROTZEK

Es werden verschiedene Approximationen für nichtglatte Operatoren eingeführt und deren Beziehungen zueinander untersucht. Für nichtglatte Optimierungsprobleme mit Operatorungleichungen als Nebenbedingungen werden F. John- und Karush-Kuhn-Tucker-Bedingungen bewiesen.

Вводятся различные аппроксимации для негладких операторов и исследуются связи между ними. Для негладких задач оптимизации с ограничениями типа операторных неравенств доказываются условия Ф. Йона и Каруша-Куна-Таккера.

For nonsmooth operators, different kinds of approximations are introduced and their relationships are studied. With the aid of these approximations, F. John and Karush-Kuhn-Tucker conditions are established for optimization problems with operator inequality constraints.

1. Introduction

Local optimality conditions for nonsmooth optimization problems are based, in one way or another, on some concept of generalized derivative. NEUSTADT [10] introduced the concept of upper convex approximations which, sometimes in modified form, was also studied by PŠENIČNYJ [12], PENOT [11], GÄHLER [4] and others. In the case of local Lipschitz functionals, the generalized directional derivative of CLARKE [1] is an important instance of an upper convex approximation. In Section 3 of this paper, we consider upper convex approximations for mappings between normed real vector spaces. Here "upper convex" refers to the preorder generated in the range space by a closed convex cone. We also introduce the concept of weak upper convex approximations which is a scalarized variant of the former concept. It is shown that in certain important cases the two concepts coincide (Theorems 1 and 2). Further we consider generalized subdifferentials that extend corresponding concepts studied by CRAVEN and MOND [2], GLÓVER [6], and others. In Section 4 we consider the problem of minimizing an extended real-valued functional subject to operator-inequality and/or operator-equation constraints. We establish necessary optimality conditions not only in terms of upper convex approximations (Theorem 3, cf. GÄHLER [4, 5]) but also in terms of generalized subdifferentials (Theorems 4 and 5). The latter results supplement those obtained by SCHIROTZEK [17, 18] for scalar-valued constraints and extend those obtained by GLOVER [6] to a broader class of operator constraints. An application of these optimality conditions to problems of best approximation is contained in a forthcoming paper by SCHEFFLER [16].

2. Notation

If E is a normed real vector space and C is a cone in E , we denote by E' the topological dual of E and by C^* the polar cone to C , i.e. $C^* = \{u \in E' \mid \langle u, x \rangle \geq 0 \text{ for each } x \in C\}$, where $\langle u, x \rangle$ denotes the value of u at x . Further, for $x_0 \in E$ and $\varepsilon > 0$, $B(x_0; \varepsilon)$

denotes the (closed) ball in E with center x_0 and radius ε . If E and F are normed real vector spaces, $L(E, F)$ denotes the vector space of all continuous linear mappings of E into F . Let M be a non-empty subset of E . Then $\text{int } M$ and $\text{cl } M$ will denote the interior of M and the closure of M , respectively. If M is a subset of E' , then $\text{cl}^* M$ denotes the weak*-closure of M . Moreover, $T(M, x_0)$ denotes the usual tangent cone to M at $x_0 \in M$. The set of all real numbers and all nonnegative real numbers is denoted by \mathbf{R} and \mathbf{R}_+ , respectively. If y_0 is an element of a vector space, then we write $\mathbf{R}_+ y_0 = \{z \mid z = \lambda y_0 \text{ for some } \lambda \in \mathbf{R}_+\}$. If f is an extended real-valued functional on E , i.e., $f: E \rightarrow \mathbf{R} \cup \{+\infty\}$, then $\text{dom } f = \{x \in E \mid f(x) < +\infty\}$. For $\lambda \in \mathbf{R}_+$, $x \in \text{dom } f$, $y \in E$ we write

$$\Delta f(\lambda, x, y) = \frac{1}{\lambda} [f(x + \lambda y) - f(x)];$$

the analogous notation being used for an operator $G: E \rightarrow F$, where E and F are normed real vector spaces.

3. Upper convex approximations

Throughout this section let E and F be normed real vector spaces, let L be a closed convex cone in F , and let G be an operator with domain E and range in F . Recall that an operator $H: E \rightarrow F$ is said to be L -convex if $x, y \in E$ and $\lambda \in (0, 1)$ imply $\lambda H(x) + (1 - \lambda) H(y) - H(\lambda x + (1 - \lambda) y) \in L$. H is said to be L -sublinear if H is L -convex and positively homogeneous. We now define the basic concepts of this paper.

Definition 1: An operator $H: E \rightarrow F$ is said to be an *upper convex approximation* of G at $x_0 \in E$ (with respect to the cone L) if H is L -sublinear and if for each $y \in E \setminus \{o\}$ and each $\varepsilon > 0$ there exist $\delta > 0$ and $\gamma > 0$ such that for each $\lambda \in (0, \delta)$ and each $y' \in B(y; \gamma)$ one has

$$\Delta G(\lambda, x_0, y') - H(y) \in B(o; \varepsilon) - L. \quad (1)$$

The set of all upper convex approximations of G at x_0 will be denoted by $A_L(G; x_0)$.

Remark: Let $f: E \rightarrow \mathbf{R} \cup \{\infty\}$. A functional $h: E \rightarrow \mathbf{R} \cup \{+\infty\}$ is said to be an *upper convex approximation* of f at $x_0 \in \text{dom } f$ in the sense of PŠENIČNYJ [12] if h is sublinear and satisfies $h(o) = 0$ and

$$h(y) \geq \limsup_{(y', \lambda) \rightarrow (y, +0)} \Delta f(\lambda, x_0, y') \quad \text{for each } y \in E \setminus \{o\},$$

where in the limit superior y' varies over neighbourhoods of y in E and λ varies over open intervals $(0, \delta)$ in \mathbf{R}_+ . We shall denote the set of all such h by $A^+(f; x_0)$. It is clear that if f, h are real-valued functionals on E , then $h \in A^+(f; x_0)$ if and only if $h \in A_{\mathbf{R}^+}(f; x_0)$.

In the general case of an operator $G: E \rightarrow F$, it is easy to verify the following

Lemma 1: Let $H \in A_L(G; x_0)$. Then for each $y \in E \setminus \{o\}$ and each $u \in L^*$ one has

$$\langle u, H(y) \rangle \geq \limsup_{(y', \lambda) \rightarrow (y, +0)} \langle u, \Delta G(\lambda, x_0, y') \rangle. \quad (2)$$

This lemma motivates the concept introduced in

Definition 2: An operator $H: E \rightarrow F$ is said to be a *weak upper convex approximation* of G at x_0 (with respect to the cone L) if H is L -sublinear and for each $y \in E \setminus \{o\}$

and $u \in L^*$ inequality (2) holds true. The set of all weak upper convex approximations of G at x_0 will be denoted by $AW_L(G; x_0)$.

According to Lemma 1, $A_L(G; x_0)$ is always a subset of $AW_L(G; x_0)$. Now we shall consider important special cases in which these sets coincide. The first result applies to certain polyhedral cones.

Theorem 1: *Assume that $L = \{z \in F \mid \langle u_i, z \rangle \geq 0 \text{ for } i = 1, \dots, m\}$, where u_1, \dots, u_m are positive-linearly independent elements of F' . Then $A_L(G; x_0) = AW_L(G; x_0)$.*

Proof: By assumption, the convex hull, say M , of $\{u_1, \dots, u_m\}$ does not contain the zero element o of F' . Since M is weak*-compact, M and o can be strongly separated, i.e., there exists a $\rho > 0$ and an element $z_0 \in F$ (considered as a weak*-continuous linear functional on F') such that $\langle v, z_0 \rangle \geq \rho$ for each $v \in M$. Now let $H \in AW_L(G; x_0)$ be given. Further let $y \in E \setminus \{o\}$ and $\varepsilon > 0$. For each $i = 1, \dots, m$ there exist real numbers $\delta_i > 0$ and $\gamma_i > 0$ such that for each $\lambda \in (0, \delta_i)$ and each $y' \in B(y; \gamma_i)$ one has $\langle u_i, \Delta G(\lambda, x_0, y') - H(y) \rangle \leq \varepsilon_0 / \|z_0\|$ and thus, with $z_1 = z_0 / \|z_0\|$, also

$$\langle u_i, \Delta G(\lambda, x_0, y') - H(y) - \varepsilon z_1 \rangle \leq 0. \tag{3}$$

Let $\bar{\delta} = \min \{\delta_1, \dots, \delta_m\}$ and $\bar{\gamma} = \min \{\gamma_1, \dots, \gamma_m\}$. Then (3) holds for each $\lambda \in (0, \bar{\delta})$, each $\lambda \in (0, \bar{\delta})$, each $y' \in B(y; \bar{\gamma})$, and each u_i ($i = 1, \dots, m$) and hence also with u_i replaced by an arbitrary element of the convex cone generated by $\{u_1, \dots, u_m\}$. However, according to the Farkas lemma, the latter cone coincides with L^* . Thus (3) implies that for each $\lambda \in (0, \bar{\delta})$ and each $y' \in B(y; \bar{\gamma})$ one has $\Delta G(\lambda, x_0, y') - H(y) - \varepsilon z_1 \in -L^{**}$. Since $L^{**} = L$ and $\varepsilon z_1 \in B(o; \varepsilon)$, it follows that $H \in A_L(G; x_0)$, and the proof is complete ■

Theorem 1 applies in particular to the case $F = \mathbb{R}^m, L = \mathbb{R}_+^m$. More precisely, for $i = 1, \dots, m$ let $g_i: E \rightarrow \mathbb{R}$ and $h_i \in A_{\mathbb{R}}(g_i; x_0)$. Further let $G = (g_1, \dots, g_m)^T$ and $H = (h_1, \dots, h_m)^T$. Then $H \in AW_{\mathbb{R}, m}(G; x_0)$ and so, by Theorem 1, $H \in A_{\mathbb{R}, m}(G; x_0)$.

Under the assumption of Theorem 1, the convex hull of $\{u_1, \dots, u_m\}$ is obviously a compact base for the cone L^* . The next result shows that $A_L(G; x_0)$ and $AW_L(G; x_0)$ coincide whenever L^* possesses any compact base, provided that F and G satisfy suitable hypotheses. Recall that L^* possesses a compact base if L is generating and L^* is locally compact (cf. JAMESON [8: p. 144]). We shall say that G is *locally Lipschitz* at x_0 if there exist $\varepsilon > 0$ and $\beta > 0$ such that

$$\|G(x_0 + y) - G(x_0)\| \leq \beta \|y\| \quad \text{for each } y \in B(o; \varepsilon).$$

In contrast to this, we shall say that G is *locally Lipschitz around* x_0 if there exist $\varepsilon > 0$ and $\beta > 0$ such that

$$\|G(x_0 + y) - G(x_0 + y')\| \leq \beta \|y - y'\| \quad \text{for all } y, y' \in B(o; \varepsilon).$$

Furthermore, G is said to be *uniformly differentiable* at $x_0 \in E$ (cf. IOFFE and TICHOMIROV [7: p. 209]) if for each $y \in E$ the directional derivative $G'(x_0, y) = \lim_{\lambda \rightarrow +0} \Delta G(\lambda, x_0, y)$ exists and for each $y \in E$ and all $\varepsilon > 0$, there are $\delta > 0$ and $\varepsilon > 0$ such that $\lambda \in (0, \delta), y' \in B(y; \gamma)$ imply

$$\Delta G(\lambda, x_0, y') \in G'(x; y) + B(o; \varepsilon). \tag{4}$$

Theorem 2: *Assume that F is a reflexive Banach space and that L^* possesses a compact base. If G is locally Lipschitz at x_0 or uniformly differentiable at x_0 , then $A_L(G; x_0) = AW_L(G; x_0)$.*

Proof: Let B denote a compact base of L^* and let $H \in AW_L(G; x_0)$. Further let $y \in E \setminus \{o\}$ and $\varepsilon > 0$ be given. Then for each $u \in B$ there exist $\delta(u) > 0$ and $\gamma(u) > 0$ such that $\lambda \in (0, \delta(u))$ and $y' \in B(y; \gamma(u))$ imply

$$\langle u, \Delta G(\lambda, x_0, y') \rangle \leq \langle u, H(y) \rangle + \frac{\varepsilon}{2}. \tag{5}$$

We first consider the case that G is locally Lipschitz at x_0 . Then choosing $\delta(u)$ sufficiently small, one can find $\beta > 0$ such that for each $\lambda \in (0, \delta(u))$ and for each $y' \in B(y; \gamma(u))$

$$\|G(x_0 + \lambda y') - G(x_0)\| \leq \beta \lambda \|y'\|. \tag{6}$$

Now let $c(u) = \beta \|y\| + \beta \gamma(u) + \|H(y)\|$. Since B is compact, there exist $u_1, \dots, u_m \in B$ such that

$$B \subseteq \bigcup_{i=1}^m \left\{ v \in F' \mid \|v - u_i\| < \frac{\varepsilon}{2c(u_i)} \right\}.$$

Let $\bar{\delta} = \min \{\delta(u_1), \dots, \delta(u_m)\}$, $\bar{\gamma} = \min \{\gamma(u_1), \dots, \gamma(u_m)\}$ and take arbitrary elements $\lambda \in (0, \bar{\delta})$, $y' \in B(y; \bar{\gamma})$. For each $u \in B$ there exists $i \in \{1, \dots, m\}$ such that $\|u - u_i\| < \varepsilon/2c(u_i)$. Hence (5), (6) and the definition of $c(u)$ imply

$$\begin{aligned} & \langle u, \Delta G(\lambda, x_0, y') - H(y) \rangle \\ &= \langle u - u_i, \Delta G(\lambda, x_0, y') - H(y) \rangle + \langle u_i, \Delta G(\lambda, x_0, y') - H(y) \rangle \\ &\leq \|u - u_i\| (\beta \|y'\| + \|H(y)\|) + \frac{\varepsilon}{2} < \varepsilon. \end{aligned} \tag{7}$$

Since B is a compact base of L^* and F is reflexive, there exists a $z \in F$ such that $B = \{u \in L^* \mid \langle u, z \rangle = 1\}$. Obviously one may assume that $\|z\| = 1$. For each $u \in B$, (7) implies $\langle u, \Delta G(\lambda, x_0, y') - H(y) - \varepsilon z \rangle \leq 0$, whence

$$\Delta G(\lambda, x_0, y') - H(y) - \varepsilon z \in -L^{**}. \tag{8}$$

The latter set equals $-L$ and (8) holds for all $\lambda \in (0, \bar{\delta})$, $y' \in B(y; \bar{\gamma})$. It follows that $H \in A_L(G; x_0)$, and the proof is complete if G is locally Lipschitz at x_0 .

Suppose now that G is uniformly differentiable at x_0 . Then for each $u \in B$, one can choose $\delta(u)$ and $\gamma(u)$ such that (4) and (5) are satisfied for each $\lambda \in (0, \delta(u))$ and each $y' \in B(y; \gamma(u))$. Now define $c(u)$ by $c(u) = \|G'(x_0, y)\| + \varepsilon + \|H(y)\|$, and define $u_i, \bar{\delta}, \bar{\gamma}$ as above. Then we obtain, instead of (7), the estimation

$$\begin{aligned} & \langle u, \Delta G(\lambda, x_0, y') - H(y) \rangle \\ &= \langle u - u_i, \Delta G(\lambda, x_0, y') - H(y) \rangle + \langle u_i, \Delta G(\lambda, x_0, y') - H(y) \rangle \\ &\leq \|u - u_i\| c(u) + \frac{\varepsilon}{2} \leq \varepsilon \end{aligned}$$

for each $u \in B$, $\lambda \in (0, \bar{\delta})$ and $y' \in B(y; \bar{\gamma})$. Now the proof is completed just as in the first case ■

Using weak upper convex approximations, we now define generalized subdifferentials of the operator G . Notice that for $H \in AW_L(G; x_0)$ and $u \in L^*$, the functional $u \circ H$ is sublinear on E and so the usual subdifferential $\partial(u \circ H)(o)$ is well defined though possibly empty. If $u \circ H$ is lower semicontinuous, then $\partial(u \circ H)(o)$ is known

to be non-empty and to satisfy

$$(u \circ H)(y) = \sup \{ \langle v, y \rangle \mid v \in \partial(u \circ H)(o) \} \tag{9}$$

(cf. IOFFE and TICHOMIROV [7: Chap. 4]). The operator H is said to be L^* -lower semi-continuous if for each $u \in L^*$ the functional $u \circ H$ is lower semicontinuous.

Definition 3: For $H \in AW_L(G, x_0)$ the set

$$\partial_H G(x_0) := \bigcup_{u \in L^*} \partial(u \circ H)(o)$$

is said to be the H -subdifferential of G at x_0 .

Notice that since $o \in L^*$, one always has $o \in \partial_H G(x_0)$ and so the H -subdifferential is never empty. Sets such as $\partial_H G(x_0)$ have been already considered by GLOVER [6]. Among others, GLOVER [6] showed that if H is L -sublinear and L^* -lower semicontinuous, then

$$\text{cl}^* \left(\bigcup_{u \in L^*} \partial(u \circ H)(o) \right) = -(H^{-1}[-L])^* \tag{9a}$$

GLOVER [6] further obtained the remarkable result that if, in addition, E and F are complete and $H[E] + L = F$, then $\bigcup_{u \in L^*} \partial(u \circ H)(o)$ is weak*-closed.

In the following proposition we consider operators $G_1, G_2: E \rightarrow F$.

Proposition 1: If H_i ($i = 1, 2$) is L^* -lower semicontinuous and belongs to $A_L(G_i; x_0)$ or $AW_L(G_i; x_0)$ and α_i are nonnegative real numbers, then $H = \alpha_1 H_1 + \alpha_2 H_2$ belongs to $A_L(G; x_0)$ or $AW_L(G; x_0)$, respectively, where $G = \alpha_1 G_1 + \alpha_2 G_2$, and one has

$$\partial_H G(x_0) = \bigcup_{u \in L^*} \text{cl}^* (\alpha_1 \partial(u \circ H_1)(o) + \alpha_2 \partial(u \circ H_2)(o)) \tag{10}$$

Proof: We only verify (10), the first statement being evident. Let $u \in L^*$. We shall show that

$$\partial(u \circ H)(o) = \text{cl}^* (\alpha_1 \partial(u \circ H_1)(o) + \alpha_2 \partial(u \circ H_2)(o)) \tag{11}$$

It is easy to see that the right-hand side of (11), A for abbreviation, is contained in the left-hand side. Suppose now that $v \in E'$ is not in A . Then by the strong separation theorem, there exist $y \in E$ and $\varepsilon > 0$ such that for all $v_i \in \partial(u \circ H_i)(o)$ ($i = 1, 2$) one has $\langle \alpha_1 v_1 + \alpha_2 v_2, y \rangle + \varepsilon < \langle v, y \rangle$. In view of (9) it follows that $u \circ H(y) < \langle v, y \rangle$ and so, again by (9), v cannot belong to $\partial(u \circ H)(o)$. The proof is thus complete ■

In the notation of Proposition 1, $\partial_H G(x_0)$ can in general not be represented by $\partial_{H_i} G_i(x_0)$. If, however, $u \circ H_i$ is continuous for each $u \in L^*$, then one has

$$\partial_H G(x_0) \subseteq \alpha_1 \partial_{H_1} G_1(x_0) + \alpha_2 \partial_{H_2} G_2(x_0)$$

This follows from (10) since now $\alpha_1 \partial(u \circ H_1)(o)$ is weak*-compact and so $\alpha_1 \partial(u \circ H_1)(o) + \alpha_2 \partial(u \circ H_2)(o)$ is weak* closed. Proposition 1 further implies that if H_1, H_2 are (weak) upper convex approximations of the same operator G , then for each $\alpha \in (0, 1)$, $\alpha H_1 + (1 - \alpha) H_2$ is also a (weak) upper convex approximation of G .

Now we shall consider important special cases in which (weak) upper convex approximations exist. First, it is clear that if G is Fréchet differentiable at x_0 , then the derivative $G'(x_0)$ belongs to $A_L(G; x_0)$, and one has

$$\partial_{G'(x_0)} G(x_0) = \{ u \circ G'(x_0) \mid u \in L^* \}$$

Next we shall show that a rather broad class of mappings studied by GLOVER [6] admits weak upper convex approximations. For this, we recall some definitions (cf. [6]). The operator $G: E \rightarrow F$ is said to be L^* -quasidifferentiable at $x_0 \in E$ if for each $y \in E$ the limit

$$G'(x_0, y) = \lim_{\lambda \rightarrow +0} \Delta G(\lambda, x_0, y)$$

exists in the weak topology of F and for each $u \in L^*$ there exists a non-empty convex weak*-closed subset of E' , denoted by $\hat{\partial}(u \circ G)(x_0)$, such that for each $y \in E$

$$\langle u, G'(x_0, y) \rangle = \sup \{ \langle v, y \rangle \mid v \in \hat{\partial}(u \circ G)(x_0) \}.$$

If E is a Banach space, then using the principle of uniform boundedness it can be shown that $\hat{\partial}(u \circ G)(x_0)$ is weak*-compact, which implies the continuity of $\langle u, G'(x_0, \cdot) \rangle$ for each $u \in L^*$.

Proposition 2: *Let G be L^* -quasidifferentiable at $x_0 \in E$. Assume further that for each $u \in L^*$, $u \circ G$ is uniformly differentiable at x_0 . Then $G'(x_0, \cdot)$ is L^* -lower semicontinuous and belongs to $AW_L(G; x_0)$. Moreover, one has $\partial_{G'(x_0, \cdot)} G(x_0) = \bigcup_{u \in L^*} \hat{\partial}(u \circ G)(x_0)$.*

Proof: Let $u \in L^*$. Then $\langle u, G'(x_0, \cdot) \rangle$, as the support functional of the convex weak*-closed set $\hat{\partial}(u \circ G)(x_0)$, is sublinear and lower semicontinuous, and one has $\partial(\langle u, G'(x_0, \cdot) \rangle)(0) = \hat{\partial}(u \circ G)(x_0)$. It remains to be shown that (2) holds for $H = G'(x_0, \cdot)$. By assumption, the equation

$$\limsup_{(y', \lambda) \rightarrow y, +0} \langle u, \Delta G(\lambda, x_0, y') \rangle = \langle u, G'(x_0, y) \rangle$$

is valid for each $u \in L^*$ and each $y \in E$. This completes the proof ■

In connection with Proposition 2, we mention that if G is L^* -quasidifferentiable at x_0 , then G is uniformly differentiable at x_0 provided that G is locally Lipschitz around x_0 or G is Hadamard differentiable at x_0 with respect to the weak topology on F (what GLOVER [6] calls arcwise directionally differentiable at x_0). DEMJANOV and RUBINOV [3] have introduced another concept of quasidifferentiability for operators. Let E and F be Banach spaces and let $L \subset F$ be a closed convex cone generating a preorder on F such that F is a conditionally complete vector lattice with a monotonic norm. An operator $G: E \rightarrow F$ is called *quasidifferentiable at $x_0 \in E$* if for each $y \in E$ the directional derivative $G'(x_0, y)$ exists and there are continuous sublinear operators $-Q, P: E \rightarrow F$ such that $G'(x_0, \cdot)$ can be represented in the form $G'(x_0, \cdot) = P + Q$. The set $DH(x_0) = [\partial G(x_0), \hat{\partial} G(x_0)]$, where

$$\partial G(x_0) = \{ S \in L(E, F) \mid Py - Sy \in L \text{ for any } y \in E \}$$

and

$$\hat{\partial} G(x_0) = \{ T \in L(E, F) \mid Ty - Qy \in L \text{ for any } y \in E \}$$

is called a *quasidifferential of G at x_0* .

Proposition 3: *Let G be quasidifferentiable and uniformly differentiable at x_0 . Then for each $T \in \hat{\partial} G(x_0)$, the mapping $H_T = P + T$ belongs to $A_L(G; x)$. Furthermore, the relation*

$$\partial_{H_T} G(x_0) = \bigcup_{u \in L^*} \text{cl}^* \{ u \circ (T + S) \mid S \in \partial G(x_0) \} \tag{12}$$

is satisfied.

Proof: The continuity of P and Q implies that for each $y \in E$ (cf. VALADIER [19]) one has

$$P(y) = \max_{S \in \partial G(x_0)} Sy \quad \text{and} \quad Q(y) = \min_{T \in \partial G(x_0)} Ty. \tag{13}$$

Under the assumptions on G the limit $\lim_{(\lambda, y') \rightarrow (+0, y)} \Delta G(\lambda, x_0, y')$ exists for each $y \in E$, and equals $G'(x_0, y)$. Therefore for each $T \in \partial G(x_0)$ it follows that

$$\lim_{(\lambda, y') \rightarrow (+0, y)} \Delta G(\lambda, x_0, y') \leq P(y) + Ty \quad \text{and} \quad H_T \in A_L(G; x_0).$$

Now formula (12) will be verified. It is easily seen that for each $u \in L^*$ one has $\partial(u \circ H_T)(o) \supseteq \text{cl}^* \{u \circ (T + S) \mid S \in \partial G(x_0)\}$. Suppose that, for some $u \in L^*$, there is $v \in \partial(u \circ H_T)(o)$ which does not belong to the right-hand side, denoted by A , of the upper inclusion. Since A is weak*-closed, A and $\partial(u \circ H_T)(o)$ are strongly separable, i.e., there exist $\bar{y} \in E$ and $\varepsilon > 0$ such that for all $T \in \partial G(x_0)$ one has $\langle v, \bar{y} \rangle \geq \langle u, T\bar{y} \rangle + \langle u, S\bar{y} \rangle + \varepsilon$. From this and (13) we obtain the contradiction $\langle u, H_T(\bar{y}) \rangle \geq \langle v, \bar{y} \rangle \geq \langle u, T\bar{y} \rangle + \langle u, S\bar{y} \rangle + \varepsilon \geq \langle u, H_T(\bar{y}) \rangle + \varepsilon$ ■

We remark that if G is quasidifferentiable at x_0 and Lipschitz around x_0 , then G is uniformly differentiable at x_0 .

4. Optimality conditions

Let E and F be normed real vector spaces, let M be a non-empty subset of E and let L be a closed convex cone in F with $\text{int } L \neq \emptyset$. Further let $f: E \rightarrow \mathbf{R} \cup \{+\infty\}$ and $G: E \rightarrow F$. We consider the following optimization problem:

(P) Minimize $f(x)$ subject to $x \in M, G(x) \in -L$.

In all that follows, let x_0 denote a local solution of (P) and let $h \in A'(f; x_0), H \in A_L(G; x_0)$. The following lemma will be the basis for the optimality conditions to be derived in the sequel.

Lemma 2: *There does not exist any $y \in \text{dom } h \cap T(M; x_0)$ such that*

$$h(y) < 0 \quad \text{and} \quad H(y) \in -\text{int } L - R_+ G(x_0). \tag{14}$$

Proof: Suppose there does exist $y \in \text{dom } h \cap T(M; x_0)$ satisfying (14). Then for some $\mu \in \mathbf{R}_+$, we have $H(y) \in -\text{int } L - \mu G(x_0)$. Since the latter set is open, there exists $\varepsilon > 0$ such that

$$H(y) + B(o; \varepsilon) \subset -\text{int } L - \mu G(x_0). \tag{15}$$

Further, since $y \in T(M; x_0)$, there exists a sequence $((\lambda_n, y_n))$ in $(0, +\infty) \times E$ converging to $(0, y)$ such that $x_0 + \lambda_n y_n \in M$ for each n . Since H and h belong to $A_L(G; x_0)$ and $A'(f; x_0)$, respectively, and $y \neq o$, it follows that for all sufficiently large n , say $n \geq n_0$, we have

$$\Delta G(\lambda_n, x_0, y_n) \in H(y) + B(o; \varepsilon) - L, \tag{16}$$

$$\Delta f(\lambda_n, x_0, y_n) \leq h(y) + \frac{1}{2} |h(y)| < 0. \tag{17}$$

From (16) and (15) we conclude that, again for all $n \geq n_0$,

$$G(x_0 + \lambda_n y_n) \in (1 - \lambda_n \mu) G(x_0) - \text{int } L.$$

But if n is large enough, then $1 - \lambda_n \mu \geq 0$ and so $G(x_0 + \lambda_n y_n) \in -L$. It follows that the sequence $(x_0 + \lambda_n y_n)$ eventually satisfies the restrictions of (P). On the other hand, (17) implies that x_0 is not a local solution of (P) which contradicts the hypothesis. This proves the lemma ■

Now we can establish a multiplier rule for (P) in terms of upper convex approximations.

Theorem 3: *Let K be a convex subset of $T(M; x_0)$ with $0 \in K$. Then there exists $(\beta, u) \in \mathbf{R}_+ \times L^*$ such that $(\beta, u) \neq 0$, $\langle u, G(x_0) \rangle = 0$ and*

$$\beta h(y) + \langle u, H(y) \rangle \geq 0 \text{ for each } y \in \text{dom } h \cap K. \quad (18)$$

Proof: Consider the space $F_0 = \mathbf{R} \times F$ equipped with the product topology and let

$$L_0 = \mathbf{R}_+ \times (L + \mathbf{R}_+ G(x_0)),$$

$$K_0 = \{(\alpha, z) \in F_0 \mid y \in \text{dom } h \cap K : (h(y) - \alpha, H(y) - z) \in -\text{int } L_0\}.$$

It is obvious that L_0 is a convex cone with non-empty interior and K_0 is a non-empty convex set. Moreover, Lemma 2 implies $K_0 \cap (-\text{int } L_0) = \emptyset$. Hence K_0 and $-L_0$ can be separated by a closed hyperplane, i.e., there exists $(\beta, u) \in \mathbf{R} \times F'$ such that $(\beta, u) \neq 0$ and

$$\alpha \beta + \langle u, z \rangle \geq \bar{\alpha} \beta + \langle u, \bar{z} \rangle \quad (19)$$

for each $(\alpha, z) \in K_0$ and each $(\bar{\alpha}, \bar{z}) \in -L_0$. Since $-L_0$ is a cone, it follows that $\bar{\alpha} \beta + \langle u, \bar{z} \rangle \leq 0$ for each $(\bar{\alpha}, \bar{z}) \in -L_0$ and so $\beta \geq 0$, $u \in (L + \mathbf{R}_+ G(x_0))^*$. The latter inclusion implies $u \in L^*$ and $\langle u, G(x_0) \rangle \geq 0$. Since, on the other hand, $G(x_0) \in -L$, the condition $\langle u, G(x_0) \rangle = 0$ is verified. Now let $y \in \text{dom } h \cap K$ and choose some $z_0 \in \text{int } L$. Since $z_0 \in \text{int } L$, for each $\delta > 0$ inequality (19) applies with $\alpha = h(y) + \delta$, $z = H(y) + \delta z_0$, $\bar{\alpha} = 0$, $\bar{z} = 0$. By letting $\delta \rightarrow +0$, we finally obtain (18), and the theorem is proved ■

In Theorem 3, a possible choice for K is Clarke's tangent cone to M at x_0 . If $T(M; x_0)$ itself is convex (which is the case if, for instance, M is locally convex at x_0), then $T(M; x_0)$ is of course the "best" choice for K .

Optimality conditions closely related to Theorem 3 have also been established, among others, by GÄHLER [4, 5]. This author allows f to be also vector-valued, but he does not derive the complementary slackness condition $\langle u, G(x_0) \rangle = 0$.

With the aid of Theorem 3 we shall now establish a multiplier rule for (P) in sub-differential form. For this, we need the following sandwich result of LANDSBERG and SCHROTZEK [9: Cor. 3].

Lemma 3: *Let $p, q: E \rightarrow \mathbf{R} \cup \{+\infty\}$ be proper convex functionals such that $-q(y) \leq p(y)$ for each $y \in E$. Assume that the cone generated by $\text{dom } p - \text{dom } q$ is a vector space. Then there exist a linear functional v on E and a real number α such that $-q(y) \leq \langle v, y \rangle + \alpha \leq p(y)$ for each $y \in E$.*

For each $(\beta, u) \in \mathbf{R}_+ \times L^*$ we define a sublinear functional $\varphi_{\beta, u}: E \rightarrow \mathbf{R} \cup \{+\infty\}$ by

$$\varphi_{\beta, u}(y) = \beta h(y) + \langle u, H(y) \rangle \quad \text{for each } y \in E.$$

Theorem 4: Let K be a closed convex subset of $T(M; x_0)$ with $o \in K$. Assume that one of the following conditions (i)–(iii) is satisfied:

- (i) For each $(\beta, u) \in \mathbb{R}_+ \times L^*$, $\varphi_{\beta, u}$ is continuous at some point of $K \cap \text{int dom } h$.
- (ii) $\text{int } K \cap \text{int dom } h \neq \emptyset$.
- (iii) E is complete and $K \cap \text{dom } h$ is a generating cone in E .

Then there exists $(\beta, u) \in \mathbb{R}_+ \times L^*$ such that $(\beta, u) \neq 0$, $\langle u, G(x_0) \rangle = 0$ and $(K \cap \text{dom } h)^* \cap \partial\varphi_{\beta, u}(o) \neq \emptyset$.

Proof: According to Theorem 3 there exists a $(\beta, u) \in \mathbb{R}_+ \times L^*$ such that $(\beta, u) \neq 0$, $\langle u, G(x_0) \rangle = 0$ and $\varphi_{\beta, u}(y) \geq 0$ for each $y \in K \cap \text{dom } h$. It is easy to see that Lemma 3 applies to $p = \varphi_{\beta, u}$ and $q = \delta$, where δ denotes the indicator functional of the convex set $K \cap \text{dom } h$. Hence there exist a linear functional v on E and a real number α such that

$$\langle v, y \rangle + \alpha \begin{cases} \leq \varphi_{\beta, u}(y) & \text{for each } y \in E, \\ \geq 0 & \text{for each } y \in K \cap \text{dom } h. \end{cases}$$

Choosing $y = o$ yields $\alpha = 0$. Moreover, if (i) or (ii) holds, then v is bounded above or below, respectively, on a non-empty open set and so is continuous. If (iii) holds, then continuity of v follows from its nonnegativity on the generating cone $K \cap \text{dom } h$ (see SCHAEFER [14: p. 228]). In any case, we have $v \in (K \cap \text{dom } h)^* \cap \partial\varphi_{\beta, u}(o)$, and the theorem is proved ■

Remarks: 1. The following condition (iv) is obviously sufficient for (i):

- (iv) $K \cap \text{int dom } h$ is non-empty, h is continuous on $\text{int dom } h$, and H is L^* -continuous, i.e., for each $u \in L^*$ the functional $y \rightarrow \langle u, H(y) \rangle$ is continuous on E .

2. If $K \cap \text{int dom } h \neq \emptyset$ or $\text{int } K \cap \text{dom } h \neq \emptyset$, then one has

$$(K \cap \text{dom } h)^* = K^* + (\text{dom } h)^*.$$

3. If h is continuous on the non-empty set $\text{int dom } h$ or H is L^* -continuous, then by a well-known result of convex analysis, one has

$$\partial\varphi_{\beta, u}(o) = \partial(\beta h)(o) + \partial(u \circ H)(o).$$

4. The multiplier β in Theorems 3 and 4 is positive and so can be chosen equal to 1, if the following constraint qualification (C) is satisfied:

- (C) There exists a $y_0 \in K \cap \text{dom } h$ such that $H(y_0) \in -\text{int } L + \mathbb{R}G(x_0)$.

In fact, let (C) hold and suppose that $\beta = 0$. Then $u \neq 0$ and so $\langle u, H(y_0) \rangle < 0$. On the other hand, (18) implies $\langle u, H(y_0) \rangle \geq 0$. Notice that (C) is a generalization of Cottle's constraint qualification in the differentiable case. Furthermore, this regularity condition implies that (18) holds with $\beta = 1$ and for each $y \in K$.

The preceding remarks indicate how to obtain from Theorems 3 and 4 further optimality conditions in terms of upper convex approximations and subdifferentials, respectively, by imposing one or the other additional hypothesis. For instance, we have the following

Corollary: Let K be a closed convex subset of $T(M; x_0)$ with $o \in K$. Assume that (iv) and (C) are satisfied. Then there exist $u \in L^*$ and $v \in K^*$ such that $\langle u, G(x) \rangle = 0$ and $v \in \partial h(o) + \partial(u \circ H)(o)$.

Finally we consider problem (P) with

$$M = G_1^{-1}[-L_1], \tag{20}$$

in other words, we consider the problem

(P1) Minimize $f(x)$ subject to $x \in E, G_1(x) \in -L_1, G(x) \in -L.$

Here G_1 is an operator of E into another normed real vector space F_1 and L_1 is a closed convex cone in F_1 . In contrast to L , the cone L_1 is not assumed to have interior points, thus L_1 may consist of the zero element of F_1 only. We assume that there exists some $H_1 \in AW_{L_1}(G_1; x_0)$ and we now put

$$K = \text{cl } H_1^{-1}[-L - \mathbf{R}_+G_1(x_0)]. \tag{21}$$

It is immediately clear that $H_1 \in AW_{L_1}(G_1; x_0)$ implies $H_1 \in AW_{L_2}(G_1; x_0)$, where L_2 is defined by $\text{cl } \{L_1 + \mathbf{R}_+G_1(x_0)\}$.

Applying the above results with M and K as defined in (20) and (21), respectively, we can derive various optimality conditions for problem (P1). For instance, applying the corollary and noticing (9a), we obtain

Theorem 5: *Assume that, with M and K as defined by (20) and (21), respectively, the conditions $K \subset T(M; x_0)$, (iv) and (C) are satisfied. Then there exists $u \in L^*$ such that $\langle u, G(x_0) \rangle = 0$ and*

$$0 \in \partial h(o) + \partial(u \circ H)(o) + \text{cl}^* \{ \partial(w \circ H_1)(o) \mid w \in L^*, \langle w, G_1(x_0) \rangle = 0 \}. \tag{22}$$

This is an asymptotic optimality condition of Karush-Kuhn-Tucker type. If, in particular, E and F_1 are complete and $H_1[E] + L_1 + \mathbf{R}_+G_1(x_0) = F_1$, then, according to the above-mentioned result of GLOVER [6: Lemma 3], the closure operation in (22) can be dropped and so (22) passes into a nonasymptotic condition.

GLOVER [6] considers vector optimization problems, where the objective and restriction operators are assumed to be L_1^* -quasidifferentiable (and arc-wise directionally differentiable). He establishes F. John and Karush-Kuhn-Tucker conditions with a complementary slackness condition in asymptotic form [6: Cor. 2]. Instead of $K \subset T(M; x_0)$, Glover assumes that the restriction $G_1(x) \in -L_1$ is locally solvable at x_0 .

Beside (C), the crucial hypothesis of Theorem 5 is the regularity condition $K \subset T(M; x_0)$ or

$$H_1^{-1}[-L_1 - \mathbf{R}_+G_1(x_0)] \subset T(G_1^{-1}[-L_1]; x_0). \tag{23}$$

We mention two special cases in which (23) is satisfied. First, if E and F_1 are complete, G_1 is continuously Fréchet differentiable at x_0 and one has $G_1'(x_0)[E] + L_1 + \mathbf{R}_+G_1(x_0) = F_1$, then with $H_1 = G_1'(x_0)$, (23) holds true according to the stability theorem of ROBINSON [13: Theorem 1] (cf. also ZOWE and KURCZYŚ [20]). Second, for the nonsmooth case, we have the following

Proposition 4: *Assume that $\text{int } L_1$ is non-empty, $H_1 \in A_{L_1}(G_1; x_0)$, and there exists $\bar{y} \in E$ such that $H_1(\bar{y}) \in -\text{int } L_1 + \mathbf{R}_+G_1(x_0)$. Then (23) holds.*

The proof runs along familiar lines. First, it is shown that each $z \in E$ satisfying $H_1(z) \in -\text{int } L_1 - \mathbf{R}_+G_1(x_0)$ belongs to the right-hand side of (23). Then, if $y \in H_1^{-1}$

$[-L_1 - \mathbf{R}_+ G_1(x_0)]$, one applies the first step to $z = \alpha \bar{y} + (1 - \alpha)y$, where $0 < \alpha < 1$. Letting $\alpha \rightarrow +0$ yields the desired result ■

As in the differentiable case, the regularity condition in Proposition 4 can be modified if $F_1 = \mathbf{R}^m$ and $L_1 = \mathbf{R}_+^m$. Thus let $G_1(x) = (g_1(x), \dots, g_m(x))^T$ for $x \in E$, where $g_i: E \rightarrow \mathbf{R}$, and let $h_i \in A_{\mathbf{R}_+}(g_i; x_0)$. Then H_1 defined by $H_1(x) = (h_1(x), \dots, h_m(x))^T$ belongs to $A_{\mathbf{R}_+}(G_1; x_0)$ (cf. the remark following the proof of Theorem 1).

Let I denote the set of all $i \in \{1, \dots, m\}$ such that $g_i(x_0) = 0$. It is immediately verified that the existence of $\bar{y} \in E$ satisfying $h_i(\bar{y}) < 0$ for each $i \in I$ implies $H_1(\bar{y}) \in -\text{int } L_1 - \mathbf{R}_+ G_1(x_0)$ and so (23). Here we still need h_i also for $i \notin I$. However, as in the differentiable case, the regularity conditions can be weakened so that they involve upper convex approximations h_i of g_i for $i \in I$ only (cf. SCHIROTZEK [17: Prop. 3.3]).

REFERENCES

- [1] CLARKE, F. H.: A new approach to Lagrange multipliers. *Math. Oper. Res.* 1 (1976) 165–174.
- [2] CRAVEN, B. D., and B. MOND: Lagrangean conditions for quasidifferentiable optimization. In: *Surveys of mathematical programming: Vol. 1, Proc. IX. Intern. Symp. on Math. Programming, Budapest 1976* (Ed.: A. PREKOPA). Budapest: Akadémiai Kiadó, and Amsterdam: North-Holland Publ. Comp. 1979, 177–192.
- [3] DEMJANOV, V. F., and A. M. RUBINOV: On quasidifferentiable mappings. *Math. Operationsforsch. Statist., ser. Optimization* 14 (1983), 3–21.
- [4] GÄHLER, S.: Generalized notions of differentiability and their application in optimization theory. In: *Convergence structures and applications to analysis* (Abhandl. AdW DDR: Bd. 4N (1979)). Berlin: Akademie-Verlag 1980, p. 67–88.
- [5] GÄHLER, S.: A generalization of an optimality theorem. In: *Nonlinear analysis, theory and applications* (Abhandl. AdW DDR: Bd. 2N (1981)). Proc. VII. Intern. Summer School Berlin (GDR). August 27–September 1, 1979. Berlin: Akademie-Verlag 1981, p. 347–349.
- [6] GLOVER, B. M.: A generalized Farkas lemma with applications to quasidifferentiable programming. *Z. Oper. Res.* 26 (1982) 5, 125–141.
- [7] IOFFE, A. D., and V. M. TICHOMIROV: *Theory of extremal problems*. Amsterdam-New York-Oxford: North-Holland Publ. Comp. 1979.
- [8] JAMESON, G.: *Ordered linear spaces* (Lecture Notes in Math.: Bd. 141). Berlin-Heidelberg-New York: Springer-Verlag 1970.
- [9] LANDSBERG, M., and W. SCHIROTZEK: Mazur-Orlicz type theorems with some applications. *Math. Nachr.* 79 (1977), 331–341.
- [10] NEUSTADT, L. W.: *Optimization: A theory of necessary conditions*. Princeton: Princeton University Press 1976.
- [11] PENOT, J. P.: Calcul sous-différentiel et optimisation. *J. Funct. Anal.* 27 (1978), 248–276.
- [12] ПШЕНИЧНЫЙ, Б. Н.: О необходимых условиях экстремума для негладких функций. *Кибернетика* 6 (1977), 92–96.
- [13] ROBINSON, S. M.: Stability theory for systems of inequalities. Part II: Differentiable nonlinear systems. *SIAM J. Numer. Anal.* 13 (1976), 497–513.
- [14] SCHAEFER, H. H.: *Topological vector spaces*. New York: Mac Millan Comp. 1966.
- [15] SCHEFFLER, H.-P.: *Lokale Approximationen von Funktionen und Mengen und deren Anwendung auf nichtglatte Optimierungsprobleme*. Diplomarbeit. Dresden: Techn. Univ. 1984.
- [16] SCHEFFLER, H.-P.: *Nichtglatte Optimierung und Bestapproximation*. *Wiss. Z. Techn. Univ. Dresden* (erscheint).
- [17] SCHIROTZEK, W.: On a theorem of Ky Fan and its application to nondifferentiable optimization. *Optimization* 16 (1985), 352–366.

- [18] SCHIROTZEK, W.: Nonasymptotic necessary conditions for nonsmooth infinite optimization problems. *J. Math. Anal. Appl.* **118** (1986), 535–546.
- [19] VALADIER, M.: Sous-différentiabilité de fonctions convexes à valeurs dans un espace vectoriel ordonné. *Math. Scand.* **30** (1972), 65–74.
- [20] ZOWE, J., and S. KURCZYK: Regularity and stability for the mathematical programming problem in Banach spaces. *Appl. Math. Optim.* **5** (1979), 49–62.

Manuskripteingang: 08. 04. 1987

VERFASSER:

Dr. W. SCHIROTZEK und H.-P. SCHEFFLER
Sektion Mathematik der Technischen Universität
Mommstr. 13
DDR-8027 Dresden