Control and Cybernetics
Vol. 16 (1987) No. 3–4

A Hölder open mapping theorem and necessary conditions of optimality in problems with Hölder data

by
PHAN QUOC KHANH
NGUYEN THIEN LUAN
Institute of Mathematics
Hanoi, Viet-Nam

We introduce Hölder subgradients for mappings of finite dimensional spaces and prove an open mapping theorem for locally Hölder mappings. Using it we obtain necessary conditions for optimality in problems with Hölder data in generalized multiplier rule forms.

1. Introduction

In recent years a great deal of attention has been devoted to nonsmooth and nonconvex problems. To consider such problems many concepts of differentiability in relaxed senses have been proposed: the generalized gradient of Clarke [2], the screen of Halkin [6], [7], the derivate container of Warga [15], the generalized derivative of Pourciau [13], the contingent derivative of Aubin [1], the shield of Dien [3] and the Hölder subgradient of Khanh and Luan [12]. One of the main objectives of this direction of consideration is obtaining necessary conditions for optimality. For this purpose, open mapping theorems are often proved and applied, since each sufficient condition for openness gives us an origin of necessary conditions for optimality. Investigations along this line usually yield multiplier rules, see e.g. [1], [3], [6], [7], [13]. It should be noted that with the exception of screens and Hölder subgradients, mentioned above notions are appropriate mainly to Lipschitz mappings.

In [12] we used variational principle of Ekeland [5] and Hölder subgradients to obtain optimality necessary conditions in problems with Hölder data. But the form of these conditions is far from the multiplier rules. In present paper we prove open mapping theorems for locally Hölder mappings and use them to obtain optimality necessary conditions for more general problems, but in finite dimensional spaces. The conditions we get are nearer to the multiplier rules.

Two following facts should be added. Firstly, most of open mapping theorems have the form of affirming the openness of a mapping in a full sense under assumptions of openness in a weaker sense. In theorems based on concepts of differentiability the weaker openness means an openness of derivatives in relaxed sense. (For a survey and general open mapping theorems of this form see [10], [11].) However, open mapping theorems of this paper have another form. Secondly, although the screen is general (for instance, more general than the generalized gradient and the derivate container) and the open mapping theorems of Halkin in [6], [7] may be applied to continuous mappings (not Lipschitz), they are appropriate only to problems with weakly Lipschitz data, not to problems with Hölder data as shown by Theorem 4.2 in Section 4 below.

2. Hölder subgradients

Throughout the paper X is a finite dimensional space, Y is a m-dimensional one. For $x_0 \in X$ we write $B(x_0, \delta) = \{x \in X / \|x - x_0\| < \delta\}$ and, for $A \subset X$, $B(A, \delta) = U_{x \in A} B(x, \delta)$.

A functional $f\colon X \to R$ is said to be locally Hölder of degree α, $0 < \alpha \leqslant 1$, at x_0 if there is a neighborhood $B(x_0, \delta)$ and $K > 0$ such that, for all x_1, $x_2 \in B(x_0, \delta)$,

$$|f(x_1) - f(x_2)| \leqslant K \|x_1 - x_2\|^{\alpha}.$$

A functional f is called locally Hölder in a subset $A \subset X$ if f is locally Hölder at each $x \in A$. If not otherwise specified, all α mentioned in the paper satisfy $0 < \alpha \leqslant 1$.

We call

$$f_{\alpha}(x_0; v) = \limsup_{\lambda \downarrow 0} \frac{f(x_0 + \lambda v) - f(x_0)}{\lambda^{\alpha}}$$

a directional α-Hölder derivative of f at x_0 and we denote

$$f_{\alpha}^{\downarrow}(x_0; v) = \liminf_{\lambda \downarrow 0} \frac{f(x_0 + \lambda v) - f(x_0)}{\lambda^{\alpha}}.$$

Let Φ^{α} be the set of all continuous functionals φ on X, which are positively homogeneous of degree α (i.e., $\varphi(\lambda x) = \lambda^{\alpha} \varphi(x)$ for $\lambda > 0$), $\varphi(-x) = -\varphi(x)$, and bounded in the sense of $\varphi(x) \leqslant M \|x\|^{\alpha}$. Let us define for φ, $\psi \in \Phi^{\alpha}$ the operations

$$(\varphi + \psi)(x) = \varphi(x) + \psi(x),$$

$$(\gamma\varphi)(x) = \gamma\varphi(x) \quad \text{for } \gamma \in R,$$

$$\|\varphi\|_\alpha = \sup_{x\neq 0}\frac{|\varphi(x)|}{\|x\|^\alpha}.$$

Then it is clear that $(\phi^\alpha, \|\cdot\|_\alpha)$ is a normed space.

DEFINITION 2.1 [12]. α-Hölder subgradient of f at x, or shortly α-subgradient, denoted by $\partial_\alpha f(x)$, is the set of all $\varphi\in\phi^\alpha$ such that $\varphi(v)\leqslant f_\alpha(x;v)$ for all $v\in X$. If $\partial_\alpha f(x)\neq\phi$ we say that f is α-subdifferentiable at x.

Let ϕ be an arbitrary class of functionals on X. We recall that a set $\Delta\subset X$ is said to be ϕ-convex if Δ has the form $\{x\in X/\varphi_i(x)\leqslant\gamma_i,\ \varphi_i\in\phi,\ \gamma_i\in R\}$. A set Δ is called ϕ-closed if $x_n\in\Delta$ and $\varphi(x_n)\to\varphi(x)$ for all $\varphi\in\phi$ imply $x\in\Delta$. In particular, if $\phi=\phi^\alpha$ and for each $x\in X$ we define a functional on ϕ^α: $x(\varphi)=\varphi(x)$, then x is a linear functional. We call the weakest topology on ϕ^α such that all $x\in X$ are continuous X-topology. If $U\subset\phi^\alpha$ is compact in X-topology, we say that U is compact. Since X is also a class of functionals on ϕ^α, we can speak about the X-convexity and the X-closedness.

In [12] we obtained the following properties of α-Hölder subgradients.

1. $\partial_\alpha f(x)$ is a X-convex and closed subset of ϕ^α. If f is locally Hölder of degree α at x with constant K, then $\|\varphi\|_\alpha\leqslant K$ for all $\varphi\in\partial_\alpha f(x)$ and $\partial_\alpha f(x)$ is X-compact.

2. If f is locally Hölder and α-subdifferentiable at x, then for all $v\in X$

$$f_\alpha(x;v)=\max\{\varphi(v)/\varphi\in\partial_\alpha f(x)\}\overset{df}{=}C(v;\partial_\alpha f(x)).$$

3. If f is locally Hölder and α-subdifferentiable at x, and if Ω is a X-convex subset in ϕ^α, then $\partial_\alpha f(x)\subset\Omega$ if and only if

$$f_\alpha(x;v)\leqslant\sup\{\varphi(v)/\varphi\in\Omega\}\quad\text{for all } v\in X.$$

4. If f and g are locally Hölder and α-subdifferentiable at x then

$$\partial_\alpha(f+g)(x)\subset Co_x(\partial_\alpha f(x)+\partial_\alpha g(x)),$$

where Co_x stands for X-convex hulls.

5. If x_0 is a local minimum of f, then f is α-subdifferentiable at x_0 and $0\in\partial_\alpha f(x_0)$.

The following properties of ϕ^α-convexity proved in [9] will also be used later.

6. For $\Omega\subset\phi^\alpha$ we have

$$Co_x\Omega=\{\varphi\in\phi^\alpha/\varphi(v)\leqslant\sup_{\psi\in\Omega}\psi(v),\ \forall v\in X\}.$$

7. We have, for $\Omega\subset\phi^\alpha$ and $v\in X$,

$$\sup\{\varphi(v)/\varphi\in\Omega\}=\sup\{\varphi(v)/\varphi\in Co_x\Omega\}.$$

Now let $F\colon X\to Y$ be a mapping. We define the property of being locally

Hölder in the same way as for functionals and do not recall it. It is obvious that $F=(f^1,\ldots,f^m)$ is locally Hölder at x if and only if so do f^i, $i=1,\ldots,m$.

DEFINITION 2.2. α-Hölder subgradient of F at x, denoted by $\partial_\alpha F(x)$, is the set of all $\theta\in\phi^\alpha_m=\phi^\alpha\times\ldots\times\phi^\alpha$ (m times) such that $\theta=(\theta^1,\ldots,\theta^m)$, $\theta^i\in\partial_\alpha f^i(x) \overset{df}{=} \partial^i_\alpha F(x)$. If $\partial_\alpha F(x)\neq\emptyset$ we say that F is α-subdifferentiable at x.

PROPOSITION 2.1. Let $F\colon X\to Y$ be locally Hölder and α-subdifferentiable at x_0. If $\partial_\alpha(-F)(x_0)\neq\emptyset$ and is contained in $-\partial_\alpha F(x_0)$, and $g\colon Y\to R$ is continuously differentiable at $F(x_0)$, then

$$\partial_\alpha(gF)(x_0)\subset Co_x\,\Sigma^m_{i=1}\left.\frac{\partial g}{\partial y^i}\right|_{F(x_0)}\partial^i_\alpha F(x_0).$$

Proof. For the sake of simplicity we shall write $\frac{\partial g}{\partial y^i}$ instead of $\left.\frac{\partial g}{\partial y^i}\right|_{F(x_0)}$. Let $v\in X$. We have

$$(gF)_\alpha(x_0;v)=\limsup_{\lambda\downarrow 0}\frac{g(f^1(x_0+\lambda v),\ldots,f^m(x_0+\lambda v))-g(f^1(x_0),\ldots,f^m(x_0))}{\lambda^\alpha}$$

$$=\limsup_{\lambda\downarrow 0}\Sigma^m_{i=1}\frac{\partial g}{\partial y^i}\frac{1}{\lambda^\alpha}[f^i(x_0+\lambda v)-f^i(x_0)]$$

$$\leqslant\Sigma^m_{i=1}\left|\frac{\partial g}{\partial y^i}\right|\limsup_{\lambda\downarrow 0}\operatorname{sgn}\left(\frac{\partial g}{\partial y^i}\right)\frac{1}{\lambda^\alpha}[f^i(x_0+\lambda v)-f^i(x_0)]$$

$$=\Sigma^m_{i=1}\left|\frac{\partial g}{\partial y^i}\right|\left(\operatorname{sgn}\left(\frac{\partial g}{\partial y^i}\right)f^i\right)_\alpha(x_0;v)$$

$$=\Sigma^m_{i=1}\left|\frac{\partial g}{\partial y^i}\right|\max\left\{\theta^i(v)/\theta^i\in\partial^i_\alpha\left(\operatorname{sgn}\left(\frac{\partial g}{\partial y^i}\right)F\right)(x_0)\right\}$$

$$=\Sigma^m_{i=1}\max\left\{\varphi^i(v)/\varphi^i\in\left|\frac{\partial g}{\partial y^i}\right|\partial^i\left(\operatorname{sgn}\left(\frac{\partial g}{\partial y^i}\right)F\right)(x_0)\right\}$$

$$\leqslant\Sigma^m_{i=1}\max\left\{\varphi^i(v)/\varphi^i\in\frac{\partial g}{\partial y^i}\partial^i_\alpha F(x_0)\right\}$$

$$=\max\left\{\psi(v)/\psi\in\Sigma^m_{i=1}\frac{\partial g}{\partial y^i}\partial^i_\alpha F(x_0)\right\}$$

$$\leqslant\max\left\{\psi(v)/\psi\in Co_x\,\Sigma^m_{i=1}\frac{\partial g}{\partial y^i}\partial^i_\alpha F(x_0)\right\}.$$

By property 3 above the proof is finished. ■

PROPOSITION 2.2. If F and $-F$ are locally Hölder and α-subdifferentiable at x,

then two following conditions are equivalent
(i) $\partial_\alpha(-F)(x) \subset -\partial_\alpha F(x)$;
(ii) $f_\alpha^{i}(x; -v)+f_\alpha^{i\downarrow}(x; v) \geqslant 0$ for $i = 1, \ldots, m$ and for all $v \in X$.

Proof. To show (i)$\Rightarrow$(ii) we see that $\partial_\alpha(-F)(x) \subset -\partial_\alpha F(x)$ means that $\partial_\alpha(-f^i)(x) \subset -\partial_\alpha f^i(x)$ for $i = 1, \ldots, m$. By the mentioned property 3 we have, for all $v \in X$ and $i = 1, \ldots, m$,

$$(-f^i)_\alpha(x; v) \leqslant \max\{\theta^i(v)/\theta^i(v)/\theta^i \in -\partial_\alpha f^i(x)\}$$
$$= \max\{\theta^i(-v)/\theta^i \in \partial_\alpha f^i(x)\}.$$

Then $f_\alpha^{i}(x; -v) \geqslant (-f^i)_\alpha(x; v) = -f_\alpha^{i\downarrow}(x; v)$.
The proof of (ii)$\Rightarrow$(i) is similar. ∎

PROPOSITION 2.3. If F and $-F$ are locally Hölder and α-subdifferentiable at x, then three following conditions are equivalent
(i) $\partial(-F)(x) = -\partial_\alpha F(x)$;
(ii) $f_\alpha^{i}(x; -v)+f_\alpha^{i\downarrow}(x; v) = 0$ for all $v \in X$, $i = 1, \ldots, m$;
(iii) $f_\alpha^{i\downarrow}(x; v) = \min\{\theta^i(v)/\theta^i \in \partial_\alpha f^i(x)\}$ for all $v \in X$ and $i = 1, \ldots, m$.

Proof. The proposition follows from Proposition 2.2 and property 2 of α-subgradients. ∎

3. Hölder open mapping theorems

In the theory of extremal problems there are two ways to get necessary conditions for extrema. In the first one the objective functional and each constraint functional of the problem are approximated independently and necessary conditions are derived as relations between those aproximations. This direction of investigation is very traditional and so developed that in a certain sense it is almost complete. The second way of study is based upon approximation of all functionals involved in the problem together, i.e., they are considered a unique system. In this case open mapping theorems are often needed, since a point corresponding to a local extremum must not be assigned to an interior point of the image of a neighborhood through a mapping chosen properly characterizing the problem. This line of consideration is rather new but has been very extensively developed recently (see [4], [6], [7], [8], [13], [14]). Most of known open mapping theorems may be applied only to Lipschitz mappings. In this section we shall prove an open mapping theorem for locally Hölder mappings basing on α-subdifferentiability.

DEFINITION 3.1. A multifunction Γ: $X \to 2^Y$ is called zero-separable on C for $C \subset X$ if for each $e \in Y$, $\|e\| = 1$, there exists $x \in C$ such that $(e, y) < 0$ for all $y \in \Gamma(x)$, $(\cdot,\cdot)$ being the scalar product. If $C = X$ we say that Γ is zero separable.

Let $F\colon X \to Y$ be α-subdifferentiable at x_0. Then by $\partial_\alpha F(x_0)(\cdot)$ we denote the multifunction

$$\partial_\alpha F(x_0)(v) = \{\theta(v)/\theta \in \partial_\alpha F(x_0)\}.$$

PROPOSITION 3.1. If $F\colon X \to Y$ is α-subdifferentiable at x_0 and satisfies the following conditions

(i) $\partial_\alpha(-F)(x_0) = -\partial_\alpha F(x_0)$;

(ii) $f^i_\alpha(x_0; v) f^{i\downarrow}_\alpha(x_0; v) > 0$ for all $v \in X$ and $i = 1, \ldots, m$;

(iii) there exists $\varphi_0 \in \partial_\alpha F(x_0)$ such that the mapping $x \to \varphi_0(x)$ is surjective, then $\partial_\alpha F(x_0)(\cdot)$ is zero-separable.

Proof. e, $\|e\| = 1$, one can find by (iii) a $v \in X$ such that $\varphi_0(v) = -e$. It follows from (i) and Proposition 2.3 that for all $\psi = (\psi^1, \ldots, \psi^m) \in \partial_\alpha F(x_0)$ we have

$$f^{i\downarrow}_\alpha(x_0; v) \leqslant \psi^i(v) \leqslant f^i(x_0; v).$$

Therefore by (ii) $\varphi^i_0(v)\psi^i(v) > 0$ and then $e^i\psi^i(v) < 0$. Thus $(e, \psi(v)) < 0$ for all $\psi \in \partial_\alpha F(x_0)$. ■

LEMMA 3.1 [12]. Let $C \subset X$ be convex and compact. Let $f\colon X \to R$ be locally Hölder of degree α and α-subdifferentiable at x_0. If $f(x_0) \leqslant f(x)$ for all $x \in C \cap B(x_0, \delta)$ and for some $\delta > 0$, then

$$\min_{v \in C} \max_{\varphi \in \partial_\alpha f(x_0)} \varphi(v - x_0) \geqslant 0.$$

We recall that a multifunction $\Gamma\colon X \to 2^Y$ is said to be upper Hausdorff semicontinuous (u.H.s.c.) at x_0 if for all $\varepsilon > 0$ there is a neighborhood $V(x_0)$ of x_0 such that $\Gamma(x) \subset B(\Gamma(x_0), \varepsilon)$ for all $x \in V(x_0)$.

THEOREM 3.1. *Suppose that $U \subset X$ is open, that $C \subset X$ is convex, that L is a subspace of X, that $x_0 \in U \cap \bar{C} \cap L$ and that $F\colon U \to Y$ is locally Hölder of degree α, $0 < \alpha < 1$, and satisfies the following conditions*:

(a) *there is a neighbourhood $V(x_0) \subset U$ of x_0 such that $\varnothing \neq \partial_\alpha(-F)(x) \subset -\partial_\alpha F(x)$ for $x \in V(x_0) \cap L$;*

(b) *the multifunctions $x \to \partial^i_\alpha F(x)$, $i = 1, \ldots, m$, are u.H.s.c. on L at x_0;*

(c) *$\partial_\alpha F(x_0)(\cdot)$ is zero-separable on $C \cap L - x_0$.*

Then $F(x_0) \in \operatorname{int} F(U \cap C)$.

Proof. All properties involved in the theorem clearly remain the same when the norm of space Y is replaced by an equivalent norm. As on a finite dimensional space all the norms are equivalent we can assume that the norm of Y is continuously differentiable at every nonzero point.

Take $\bar{B} = B(x_0, r) \subset V(x_0)$. Let C_j, $j = 1, 2, \ldots$, be convex and compact subsets of X such that (cf. Halkin [7])

$$x_0 \in C_j \subset C \cup \{x_0\},$$

$$C_j \subset C_{j+1},$$
$$C \subset \overline{U_{j=1}^{\infty} C_j}.$$

If there is an integer j_0 such that

$$F(x_0) \in \operatorname{int} F(\bar{B} \cap C_{j_0}), \tag{1}$$

then we can easily complete the proof. Indeed, if $F(x_0) \notin \operatorname{int} F(U \cap C)$, there is a sequence $\{y_k\}$ such that $y_k \neq F(x_0)$, $y_k \to F(x_0)$ and $y_k \notin F(U \cap C)$. By (1) $y_k \in F(U \cap C_{j_0})$ for k large enough. So there exist $x_k \in U \cap C_{j_0}$ such that $x_k \neq x_0$, $x_k \notin C$, and $y_k = F(x_k)$. Hence, $x_k \notin C \cup \{x_0\}$, which contradicts the fact that $x_k \in C_{j_0} \subset C \cup \{x_0\}$.

To show (1) we suppose to the contrary that

$$F(x_0) \notin \operatorname{int} F(\bar{B} \cap C_j), \quad j = 1, 2 \ldots .$$

Fix an arbitrary j. We can find a sequence $z_k \to F(x_0)$, $z_k \notin F(\bar{B} \cap C_j)$. Put $\varepsilon_k = \|z_k - F(x_0)\|$ and $\xi_k(x) = \|z_k - F(x)\|$. Since $\bar{B} \cap C_j \cap L$ is a complete metric space, by Ekeland's variational principle there is $v_k \in \bar{B} \cap C_j \cap L$ such that

$$\|v_k - x_0\| < \sqrt{\varepsilon_k},$$

$$\xi_k(v_k) \leqslant \varepsilon_k(x) + \sqrt{\varepsilon_k}\, \|x - v_k\| \quad \text{for } x \in \bar{B} \cap C_j \cap L.$$

Since $v_k \in \operatorname{int} \bar{B}$ for k large enough and the functional $P(x) = \xi_k(x) + \sqrt{\varepsilon_k}\, \|x - v_k\|$ attains the minimum on $\bar{B} \cap C_j \cap L$ at v_k, by Lemma 3.1 we have

$$\max_{\varphi \in \partial_\alpha P(v_k)} \varphi(v - v_k) \geqslant 0 \quad \text{for all } v \in C_j \cap L. \tag{2}$$

Taking assumption (a) and Proposition 2.1. into account we get, for some $l_k = (l_k^1, \ldots, l_k^m)$, $\|l_k\| = 1$,

$$\partial_\alpha P(v_k) \subset Co_x\, \Sigma_{i=1}^m l_k^i \partial_\alpha^i F(v_k). \tag{3}$$

Because of (b) and the continuity of φ, letting $k \to +\infty$, (2), (3) and property 7 of the ϕ^α-convexity together yield

$$\max_{\varphi \in \Sigma_{i=1}^m \beta^i \partial_\alpha^i F(x_0)} \varphi(v - x_0) \geqslant 0 \quad \text{for all } v \in C_j \cap L,$$

where $\beta = (\beta^1, \ldots, \beta^m)$ is some limit of the sequence $\{l_k\}$. Since $C \subset \overline{U_{j=1}^{\infty} C_j}$ we have

$$\inf_{v \in C \cap L} \max_{\varphi \in \Sigma_{i=1}^m \beta^i \partial_\alpha^i F(x_0)} \varphi(v - x_0) \geqslant 0. \tag{4}$$

On the other hand, by the zero separability on $C \cap L - x_0$ of $\partial_\alpha F(x_0)(\cdot)$

there is $\bar{v} \in C \cap L$ such that

$$\Sigma_{i=1}^{m} \beta^i \varphi^i (\bar{v} - x_0) < 0 \quad \text{for all } \varphi \in \partial_\alpha F(x_0).$$

This contradiction to (4) concludes the proof. ■

EXAMPLE 3.1. Let $X = R^2$, $Y = R$, $U = C = X$, $L = \{(x, y) \in X / ax + by = 0\}$, ($a$, b, c and d are real numbers) and

$$F(x, y) = |ax + by|^\alpha \operatorname{sgn}(ax + by) + |cx + dy|^\alpha \operatorname{sgn}(cx + dy),$$

where $0 < \alpha < 1$ and $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$. Then, direct computations give

$$F_\alpha((0, 0); (u, v)) = |au + bv|^\alpha \operatorname{sgn}(au + bv) + |cu + dv|^\alpha \operatorname{sgn}(cu + dv),$$

$$\partial_\alpha F(0, 0)(u, v) = \{|cu + dv|^\alpha \operatorname{sgn}(cu + dv)\} \quad \text{for } (u, v) \in L$$

(this is an one-point set in $Y = R$),

$$\partial_\alpha F(x, y) = \{\varphi_0\} \text{ if } (x, y) \in L \setminus \{(0, 0)\},$$

where $\varphi_0(u, v) = |au + bv|^\alpha \operatorname{sgn}(au + bv)$. Therefore it is easy to verify directly that all the hypotheses (a), (b) and (c) of Theorem 3.1 are satisfied and the conclusion $0 \in \operatorname{int} F(R^2)$ holds.

REMARK 3.1. *Example 3.1 shows a case in which* $\partial_\alpha F(x) \neq \{0\}$ *for all* x *on a subspace* L of X. *The following example assures us that* $\partial_\alpha F(x)$ *may be different from* $\{0\}$ *densely around a point* x_0.

EXAMPLE 3.2. Let φ: $[0, 1] \to R$ be defined by

$$\varphi(x) = \begin{cases} |x|^{1/2} & \text{if } \quad 0 \leqslant x \leqslant \frac{1}{2} \\ |x-1|^{1/2} & \text{if } \quad \frac{1}{2} < \alpha \leqslant 1. \end{cases}$$

We extend φ to the whole R by $\varphi(-x) = -\varphi(x)$ and $\varphi(x+2) = \varphi(x)$. Now let f: $R \to R$ be defined by the formula

$$f(x) = \Sigma_{n=0}^{\infty} 10^{-n} \varphi(10^n x).$$

Then f satisfies the Hölder condition

$$|f(x) - f(y)| \leqslant \sqrt{2}\, H\, |x - y|^{1/2}$$

where $H = \Sigma_{n=0}^{\infty} 10^{-n/2} = (1 - 10^{-1/2})^{-1}$. Direct calculations yield $\partial_{1/2} f(0) = \{\theta_0\}$, where $\theta_0(v) = H\,|v|^{1/2} \operatorname{sgn} v$. So $\partial_{1/2} f(0)(\cdot)$ is zero-separable. Moreover, we can define by direct computations

$$\partial_{1/2} f(2k\, 10^{-s}) = \{\varphi_s\}, \quad \text{for } k, s = 1, 2, \ldots,$$

$$\partial_{1/2} f((2k-1)\, 10^{-s}) = \{\psi_s\}, \quad \text{for } k, s = 1, 2, \ldots,$$

where

$$\varphi_s = \left(H(s) + 10^{-s/2}\right)|v|^{1/2}\operatorname{sgn} v,$$

$$\psi_s = \left(H(s) - 10^{-s/2}\right)|v|^{1/2}\operatorname{sgn} v,$$

$$H(s) = \Sigma_{n=s+1}^{\infty} 10^{-n/2}.$$

4. Necessary conditions of optimality in problems with Hölder data

To our knowledge almost all known necessary conditions for optimality are in force only for problems with convex or Lipschitz data. In [6], [7] Halkin proved open mapping theorems and necessary conditions for problems with continuous data. But as shown by the lines below these conditions are also appropriate only to cases near to Lipschitz ones.

At first we recall a notion and a result of Halkin in [7]. Let $U \subset X$ be open and $F\colon U \to Y$ be a mapping. A set S of linear operators of X into Y is said to be a screen for F at $x_0 \in U$ if for any $\varepsilon > 0$ and any $r > 0$ there is a $\delta \in (0, r]$ with $B(x_0, \delta) \subset U$ and a continuously differentiable mapping $G\colon B(x_0, \delta) \to Y$ such that for all $x \in B(x_0, \delta)$ we have $\|G(x) - F(x)\| \leqslant \varepsilon\delta$ and $G'(x) \in B(S, \varepsilon)$.

Consider the following optimization problem

$$\begin{aligned} &\text{minimize } g_0(x), \\ &g_i(x) \leqslant 0 \quad \text{for } i = -\mu, \ldots, -1, \\ &g_i(x) = 0 \quad \text{for } i = 1, \ldots, n, \\ &x \in \Omega, \end{aligned} \tag{P}$$

where g_i, $i = -\mu, \ldots, n$, are defined on an open subset $U \subset X$, and $\Omega \subset X$ is a convex subset.

The following interesting necessary condition in the multiplier rule form is due to Halkin:

Theorem 4.1 [7]. *Assume that $g_{-\mu}, \ldots, g_{-1}, g_0, g_1, \ldots, g_n$ are continuous in a neighborhood of x_0 and admit compact screens $S_{-\mu}, \ldots, S_n$ at x_0. If x_0 solves problem* (P) *locally, then there exists a nonzero vector $(\lambda_{-\mu}, \ldots, \lambda_n)$ such that*

(1) *for some $p \in \Sigma_{i=-\mu}^{n} \lambda_i S_i$ we have*

$$p(x_0 - x) \geqslant 0 \text{ for all } x \in \Omega,$$

(2) $\lambda_i \leqslant 0$ *for* $i = -\mu, \ldots, 0$, *and*

(3) $\lambda_i g_i(x_0) = 0$ *for* $i = -\mu, \ldots, -1$.

Moreover, if $x_0 \in \operatorname{int}\Omega$, statement (1) can be replaced by

(1*) $0 \in \Sigma_{i=-\mu}^{n} \lambda_i S_i$.

However, the condition that a mapping admits a compact screen is near to that it is locally Lipschitz as we shall see below.

DEFINITION 4.1. A mapping $F: U \to Y$, U being open, is said to be weakly locally Lipschitz at x_0 if there is a sequence $\delta_k \to 0$ such that for all sequences $x_k \to x_0$, $z_k \to x_0$ satisfying $\delta_k \|x_k - z_k\|^{-1} \leqslant N$ for some $N > 0$ we have

$$\limsup_{k \to \infty} \frac{\|F(x_k) - F(z_k)\|}{\|x_k - z_k\|} < +\infty.$$

THEOREM 4.2. *If $F: U \to Y$, U being open, admits a compact screen S at x_0 then F is weakly locally Lipschitz at x_0.*

Proof. Since S is compact, $\|A\| \leqslant M$ for some $M > 0$ and for all $A \in S$. For any $\varepsilon_k \to 0$ and $r_k \to 0$, there is $\delta_k \in (0, r_k)$ with $B(x_0, \delta_k) \subset U$ and a continuously differentiable mapping G_k: $B(x_0, \delta_k) \to Y$ such that for all $x \in B(x_0, \delta_k)$ we have $\|G_k(x) - F(x)\| \leqslant \varepsilon_k \delta_k$ and $G_k'(x) \in B(S, \varepsilon_k)$. For arbitrary x_k, $z_k \in B(x_0, \delta_k)$ we have

$$\limsup_{k \to \infty} \|x_k - z_k\|^{-1} \|G_k(x_k) - G_k(z_k)\| < +\infty. \tag{5}$$

In fact, suppose the contrary: that there exist subsequences, which are denoted by the same notation to avoid double indices, such that

$$\lim_{k \to \infty} \|x_k - z_k\|^{-1} \|G_k(x_k) - G_k(z_k)\| = +\infty.$$

Then for all k large enough

$$\|G_k(x_k) - G_k(z_k)\| > 3M \|x_k - z_k\|. \tag{6}$$

Since $G_k'(z_k) \in B(S, \varepsilon_k)$ for each k, we have

$$\lim_{z \to z_k} \|z - z_k\|^{-1} \|G_k(z) - G_k(z_k)\| = \|G_k'(z_k)\| \leqslant M + \varepsilon_k < 2M.$$

So, for all z close enough to z_k, $\|G_k(z) - G_k(z_k)\| < 3M \|z - z_k\|$, contradicting (6).

Now let x_k, $z_k \in B(x_0, \delta_k)$ satisfy $\delta_k \|x_k - z_k\|^{-1} \leqslant N$. We have

$$\begin{aligned} &\|x_k - z_k\|^{-1} \big| \|G_k(x_k) - G_k(z_k)\| - \|F(x_k) - F(z_k)\| \big| \\ &\quad \leqslant \|x_k - z_k\|^{-1} \big(\|G_k(x_k) - F(x_k)\| + \|G_k(z_k) - F(z_k)\| \big) \\ &\quad\quad \leqslant 2\varepsilon_k \delta_k \|x_k - z_k\|^{-1} \leqslant 2\varepsilon_k N \to 0. \end{aligned}$$

Hence a glance at (5) shows that F is weakly locally Lipschitz at x_0. ■

Moreover, the following proposition says that if F is not locally Lipschitz at x_0 then each compact screen for F at x_0 approximates F not very well.

PROPOSITION 4.1. If F: $U \to Y$ is not locally Lipschitz at x_0 and admits a compact screen S at x_0, then, for any $\varepsilon_k \to 0$, $\delta_k \to 0$ and continuously differentiable mapping G_k with $G_k'(x) \in B(S, \varepsilon_k)$ for all $x \in B(x_0, \delta_k)$, we can find $\bar{x}_k$, $\bar{z}_k \in B(x_0, \delta_k)$ such that

$$\lim_{k \to \infty} \|\bar{x}_k - \bar{z}_k\|^{-1} (\|G_k(\bar{x}_k) - F(\bar{x}_k)\| + \|G_k(\bar{z}_k) - F(\bar{z}_k)\|) = +\infty.$$

Proof. Since F is not locally Lipschitz at x_0 there are $\bar{x}_k$, $\bar{z}_k \in B(x_0, \delta_k)$ for any $\delta_k \to 0$ such that

$$\lim_{k \to \infty} \|\bar{x}_k - \bar{z}_k\|^{-1} \|F(\bar{x}_k) - F(\bar{z}_k)\| = +\infty.$$

Then, we have, by (5),

$$\|\bar{x}_k - \bar{z}_k\|^{-1} (\|G_k(\bar{x}_k) - F(\bar{x}_k)\| + \|G_k(\bar{z}_k) - F(\bar{z}_k)\|)$$
$$\geqslant \|\bar{x}_k - \bar{z}_k\|^{-1} \big|\|G_k(\bar{x}_k) - G_k(\bar{z}_k)\| - \|F(\bar{x}_k) - F(\bar{z}_k)\|\big| \to +\infty.$$

This completes the proof. ■

Having seen that Theorem 4.1 cannot be applied to problems with Hölder data we are sure that two following theorems may be useful.

THEOREM 4.3. *Let, in problem* (P), $\Omega \equiv X$ *and* g_i, $i = -\mu, \ldots, n$, *be locally Hölder of degree* α, $0 < \alpha < 1$, *and satisfy the following conditions in a neighborhood of* x_0:

(i) $\emptyset \neq \partial_\alpha(-g_i)(x) \subset -\partial_\alpha g_i(x)$;

(ii) *the multifunctions* $x \to \partial_\alpha g_i(x)$ *are* u.H.s.c. *at* x_0.

If x_0 *solves problem* (P) *locally, then there exists a vector* $\beta = (\beta_{-\mu}, \ldots, \beta_n)$, $\|\beta\| = 1$, *such that*

$$0 \in Co_x \Sigma_{i=-\mu}^{n} \beta_i \partial_\alpha g_i(x_0). \tag{7}$$

Proof. Denote

$$M(x) = \{i \in \{-\mu, \ldots, -1\} / g_i(x) = 0\},$$
$$A = \{x / g_i(x) < 0 \quad \text{for all } i \in \{-\mu, \ldots, -1\} \setminus M(x_0)\},$$
$$V = A \cap U.$$

Then V is open and $x_0 \in V$. We define a mapping $F\colon V \to R^{\mu+n+1}$ by

$$F(x) = (g_{-\mu}(x), \ldots, g_0(x), \ldots, g_n(x)).$$

F clearly satisfies the assumptions (a) and (b) of Theorem 3.1 with $C = L = X$.

On the other hand, $F(x_0) \notin \operatorname{int} F(V)$. Were this false, there would exist $\varepsilon > 0$ such that

$$(t_{-\mu}, \ldots, t_0, 0, \ldots, 0) \in F(V) \quad \text{for } t_i\ (g_i(x_0) - \varepsilon, g_i(x_0)),$$

for $i \in \{-\mu, \ldots, -1, 0\} \setminus M(x_0)$, i.e., x_0 were not a local solution of (P).

By Theorem 3.1, $\partial_\alpha F(x_0)(\cdot)$ must not be zero-separable. So there exists $\beta = (\beta_{-\mu}, \ldots, \beta_n)$, $\|\beta\| = 1$, such that for each $v \in X$, there is $\theta_i \in \partial_\alpha g_i(x_0)$, $i = -\mu, \ldots, n$, with $\Sigma_{i=-\mu}^{n} \beta_i \theta_i(x) \geqslant 0$. Then by virtue of property 6 of the ϕ^α-convexity we obtain (7). ■

In [12] the case $\Omega = X$ was considered when X is a Banach space (infinite dimensional) and we obtained the following necessary condition

$$0 \in Co_x \bigcup_{\substack{i=-\mu,\ldots,0 \\ j=1,\ldots,m}} (\partial_\alpha g_i(x_0) \cup \partial_\alpha |g_j(x_0)|) \tag{8}$$

without the assumption (i). In general, none of inclusions (7) and (8) is stronger than the other. The advantage of (7) is that it has a form closer to the classical multiplier rules.

The assumptions of Theorem 4.3 are not necessary to obtain (7) as explained by the following example.

EXAMPLE 4.1. Consider the problem

$$g_0(x) = |x-1|^{1/2} + |x-2|^{1/2} \operatorname{sgn}(x-2) \to \min,$$
$$g_1(x) = |x-1|^{1/2} - 1 \leqslant 0.$$

It is easy to see that the solution is $x_0 = 1$, $g_0(x_0) = -1$. Direct calculations give: $g_{01/2}(1; v) = |v|^{1/2}$, $(-g_0)_{1/2}(1; v) = -|v|^{1/2}$, $\partial_{1/2} g_0(1) = \{x \to |ax|^{1/2} \cdot \operatorname{sgn}(ax)/|a| \leqslant 1\}$, $\partial_{1/2}(-g_0)(1) = \emptyset$. So the assumptions of Theorem 4.3 are not satisfied. But

$$0 \in CO_x(1.\partial_{1/2} g_0(1) + 0.\partial_{1/2} g_1(1)).$$

THEOREM 4.4. *Let, in problem* (P), g_i, $i = -\mu, \ldots, n$, *satisfy the conditions* (i) *and* (ii) *of Theorem* 4.3. *If* x_0 *is a local solution, then there is* $\beta = (\beta_{-\mu}, \ldots, \beta_n)$, $\|\beta\| = 1$, *such that*

$$\inf_{v \in \Omega} \max_{\xi \in \Sigma_{i=-\mu}^{n} \beta_i \partial_\alpha g_i(x_0)} \xi(v - x_0) \geqslant 0.$$

Proof. Put $V = U \times R^{\mu+n+1}$ and define $Q\colon V \to R^{\mu+n+1}$ by

$$Q(x, w) = F(x) - w,$$
$$F(x) = (g_{-\mu}(x), \ldots, g_0(x), \ldots, g_n(x)).$$

Set

$$W = \{w = (w_{-\mu}, \ldots, w_n) \in R^{\mu+n+1} / g_i(x_0) + w_i \leqslant 0$$
$$\text{for } i = -\mu, \ldots, -1,\ w_0 \leqslant 0 \text{ and } w_i = 0 \text{ for } i = 1, \ldots, n\}.$$

Denote $C = \Omega \times W$. Then C is convex.

If $Q(x_0, 0) \in \operatorname{int} Q(V \cap C)$, we see that

$$F(x_0) = F(z_0) - \bar{w}$$

for some z_0 in $U \cap \Omega$ and some $\bar{w}$ in W with $\bar{w}_0 < 0$. Therefore, z_0 is a feasible point of problem (P) with

$$g_0(z_0) < g_0(x_0),$$

i.e., x_0 is not a solution. Hence,

$$Q(x_0, 0) \notin \operatorname{int} Q(V \cap C).$$

To define $\partial_\alpha Q(x, w)$ we have, for (x, w) close enough to $(x_0, 0)$,

$$\begin{aligned} Q^i_\alpha((x, w); (v, \tau)) &= \limsup_{\lambda \downarrow 0} \lambda^{-\alpha} [g_i(x+\lambda v) - (w_i + \lambda \tau_i) - (g_i(x) - w_i)] \\ &= g_{i\alpha}(x; v) \quad \text{for } i = -\mu, \ldots, n. \end{aligned}$$

Consequently, $\partial_\alpha Q(x, w) = \partial_\alpha F(x)$, or equivalently, $\partial^i_\alpha Q(x, w) = \partial_\alpha g_i(x)$, $i = -\mu, \ldots, n$. It follows that Q satisfies assumptions (*a*) and (*b*) of Theorem 3.1 with $L \equiv X$ and then $\partial_\alpha Q(x_0, 0)(\cdot)$ must not be zero-separable on $C - (x_0, 0)$. Hence there exists $\beta = (\beta_{-\mu}, \ldots, \beta_n)$ $\|\beta\| = 1$, such that

$$\inf_{v \in \Omega, w \in W} \; \max_{\varphi \in \sum_{i=-\mu}^{n} \beta_i \partial_\alpha Q(x_0, 0)} \varphi(v - x_0, w) \geq 0.$$

Since $0 \in W$, this implies that

$$\inf_{v \in \Omega} \; \max_{\xi \in \sum_{i=-\mu}^{n} \beta_i \partial_\alpha g_i(x_0)} \xi(v - x_0) \geq 0.$$

■

The authors express their sincere thanks to Professor S. Rolewicz for valuable remarks and advice.

References

[1] Aubin J. P. Contingent Derivatives of Set-Valued Maps and Existence of Solutions to Non-Linear Inclusion and Differential Inclusions. Mathematical Analysis and Applications, Part A, Adv. in Math., Suppl. Stud., Vol. 7a, Acad. Press, New York, 1981, 159–229.

[2] Clarke F. H. Generalized Gradients and Applications. *Trans. Amer. Math. Soc.* **205** (1975), 247–252.

[3] Dien P. H. Some Results on Locally Lipschitzian Mappings. *Acta Math. Vietnamica* **6** (1981) 2, 97–105.

[4] Dmitruk A. V., Miljutin A.A., Osmolovskii N. P. Lusternik theorem in the Theory of Extremum. *Uspekhi Mat. Nauk*, **35** (1980) 6, 11–46 (in Russian).

[5] Ekeland I. On the Variational Principle. *J. Math. Anal. Appl.* **47** (1974), 324–353.

[6] Halkin H. Interior Mapping Theorem with Set-Valued Derivatives. *J. d'Anal. Math.* **30** (1976), 200–207.

[7] Halkin H. Mathematical Programming without Differentiability, Calculus of Variations and Control Theory. Edited by D.L. Russell, Acad. Press, New York, 1976, 279–287.

[8] Ioffe A.D., Tihomirov V. M. Theory of Extremal Problems. Amsterdam, North-Holland, 1979.

[9] KHANH P. Q. A Formula for Extremum Involving ϕ-convexity. *Control and Cybernetics*, **13** (1984), 305–311.

[10] KHANH P. Q. An Induction Theorem and General Open Mapping Theorems. *J. Math. Anal. Appl.* **118** (1986), 519–534.

[11] KHANH P. Q. An Open Mapping Theorem for Families of Multifunctions. *J. Math. Anal. Appl.* (to appear).

[12] KHANH P. Q., LUAN N. T. Höder Subgradients and Applications in Optimization. *Control and Cybernetics*, **14** (1985), 4, 323–341.

[13] POURCIAU B. H. Analysis and Optimization of Lipschitz Continuous Mappings. *J. Opti. Theory Appl.* **22** (1977), 311–351.

[14] SACH P. H. Un Théorème Général de Type d'Application Localement Ouverte, Analyse Appliquée et Informatique, Université de Bordeaux 1, U.E.R. de Math. et Informatique 1980, n. 8021.

[15] WARGA J. Derivative Containers, Inverse Functions and Controllability, Calculus of Variations and Control Theory. Edited by D. L. Russell, New York, Acad. Press, 1976, 13–46.

Received, March 1986.

Twierdzenie o otwartym przekształceniu Höldera i warunki konieczne optymalności w zadaniach opisanych funkcjami spełniającymi warunek Höldera

Wprowadzono subgradienty Höldera dla przekształceń w przestrzeniach skończenie wymiarowych i udowodniono twierdzenie o przekształceniu otwartym dla lokalnie hölderowskich przekształceń. Zastosowanie tego twierdzenia pozwoliło na wprowadzenie warunków koniecznych w postaci uogólnionej reguły mnożnika dla zadań opisanych funkcjami spełniającymi warunek Höldera.

Теорема об открытом гельдеровом преобразовании и необходимые условия оптимальности в задачах описываемых функциями удовлетворяющими условие Гельдера

Введены субградиенты Гельдера для преобразований в конечномерных пространствах и доказана теорема об открытом преобразовании для локально гельдеровых преобразований. Применение этой теоремы позволило ввести необходимые условия в виде обобщенного правила множителя для задач описываемых функциями, удовлетворяющими условие Гельдера.