Set-Valued Analysis **4**: 101–117, 1996.

Second-Order Subdifferentials of $C^{1,1}$ Functions and Optimality Conditions*

PANDO GR. GEORGIEV and NADIA P. ZLATEVA
University of Sofia, Department of Mathematics and Informatics, 5 James Bourchier Blvd., 1126 Sofia, Bulgaria

(Received: 14 January 1994; in final form: 23 May 1995)

Abstract. We present second-order subdifferentials of Clarke's type of $C^{1,1}$ functions, defined in Banach spaces with separable duals. One of them is an extension of the generalized Hessian matrix of such functions in $\mathbb{R}^n$, considered by J. B. H.-Urruty, J. J. Strodiot and V. H. Nguyen. Various properties of these subdifferentials are proved. Second-order optimality conditions (necessary, sufficient) for constrained minimization problems with $C^{1,1}$ data are obtained.

Mathematics Subject Classifications (1991). 49J52, 49K30.

Key words: second-order subdifferentials, Clarke's subdifferential, twice strict differentiability, sufficient optimality conditions, necessary optimality conditions.

0. Introduction

One of the motivations of this paper is the article of J. B. H.-Urruty, J. J. Strodiot and V. H. Nguyen [8]. They defined the so-called generalized Hessian matrix of $C^{1,1}$ functions in $\mathbb{R}^n$ and investigated many of its properties, including necessary second-order optimality conditions for minimization problems with $C^{1,1}$ data.

Our goal is to define similar notions in infinite-dimensional Banach spaces and to investigate their properties.

The Rademacher theorem plays a crucial role for defining in $\mathbb{R}^n$ the generalized Hessian matrix to $C^{1,1}$ functions. In infinite-dimensional Banach spaces, we use a weaker generalization of it, due to J. P. R. Christensen [3], Theorem 7.5, but for $C^{1,1}$ functions defined on Banach spaces with separable duals, and we define an extension of the generalized Hessian matrix, called here a second-order subdifferential. Many of the properties of the generalized Hessian matrix are also valid in our case.

Let us note that the above-mentioned result of Christensen was used by L. Thiboult in [16], where he extended the notion of Clarke's subdifferential to mappings acting from separable Banach spaces to reflexive spaces.

* This work was partially supported by the National Foundation for Scientific Investigations in Bulgaria under contract No. MM-406/1994.

We obtain necessary and sufficient conditions for constrained minimization problems with $C^{1,1}$ data. The necessary condition cannot be proved by the method in [8], which does not work in infinite-dimensional Banach spaces.

Our approach is based on the method, described by V. M. Alekseev, V. M. Tikhomirov and S. V. Fomin [1] for obtaining necessary and sufficient conditions for constrained minimization problems, defined by twice Fréchet differentiable functions.

We refer to the following (incomplete) list of publications and references therein concerning the recent development of the second-order derivatives and optimality conditions in nonsmooth analysis: [2, 10–12, 14, 15, 17].

Some of the results in this paper are announced in [6].

1. Basic Definitions and Properties

Let $(E, \|\cdot\|)$ be a real Banach space with separable dual $(E^*, \|\cdot\|)$ (so E is separable too) and G be an open subset of E.

Consider the class $C^{1,1}(G)$ of all functions $f\colon G \to \mathbb{R}$, whose first Gateaux derivatives are locally Lipschitz (then, by the mean value theorem, f is strictly Fréchet differentiable on G). Having in mind that every separable dual space E^* has a Radon–Nikodym property (see [13]), it follows from a theorem of J. P. R. Christensen [3], Theorem 7.5, that for every $f \in C^{1,1}(G)$, f' is Gateaux differentiable on a dense subset $G(f)$ of G. In fact, $G(f)$ is *Haar-null* set (see [3]). We shall say that f is twice Gateaux differentiable on $G(f)$ and denote by $f''(x)$ the Gateaux derivative of f' at $x \in G(f)$.

We shall denote by $x^*[h]$ the value of the linear functional $x^* \in E^*$ on the element $h \in E$ and by $L[h_1, h_2]$ the value of the bilinear functional L defined on $E \times E$ on the pairs of elements $h_1, h_2 \in E$.

Let $\mathcal{L}(E \times E)$ be the Banach space of all bilinear continuous functionals $L\colon E \times E \to \mathbb{R}$ with the norm

$$\|L\| = \sup_{\substack{\|h_1\|=1 \\ \|h_2\|=1}} |L[h_1, h_2]|,$$

and $\mathcal{L}(E, E^*)$ be the Banach space of all linear continuous mappings $L\colon E \to E^*$ with the norm $\|L\| = \sup_{\|h\|=1} \|L(h)\|^*$.

It is well known that $\mathcal{L}(E \times E)$ and $\mathcal{L}(E, E^*)$ are isometrically isomorphic (see [1], Section 2.2.5). So, in the sequel, we shall identify $\mathcal{L}(E \times E)$ and $\mathcal{L}(E, E^*)$.

In the sequel we shall also suppose that the function f belongs to the class $C^{1,1}(G)$.

DEFINITION 1.1. For every $x \in G$, $h_1, h_2 \in E$ we define

$$f^{\infty}(x; h_1, h_2) := \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{f'(y + th_1)[h_2] - f'(y)[h_2]}{t},$$

$$d^2 f(x; h_1, h_2) := \limsup_{G(f)\ni z\to x} f''(z)[h_1, h_2].$$

PROPOSITION 1.2. *For every $x \in G$, $h_1, h_2 \in E$ we have*

$$f^{00}(x; h_1, h_2) = d^2 f(x; h_1, h_2).$$

Proof. From [16], Proposition 2.2, we have

$$\begin{aligned} f^{00}(x; h_1, h_2) &= \langle f'(\cdot), h_2\rangle^0(x; h_1) = \limsup_{G(f)\ni z\to x} \langle f'(\cdot);\ h_2\rangle'(z; h_1) \\ &= \limsup_{G(f)\ni z\to x} f''(z)[h_1, h_2]. \end{aligned}$$

□

PROPOSITION 1.3. *The function $f^{00}(\cdot; h_1, h_2)$ is upper semicontinuous for every $h_1, h_2 \in E$ and*

$$|f^{00}(x;\ h_1, h_2)| \leqslant l_x \|h_1\| \cdot \|h_2\|,$$

where l_x is a Lipschitz constant of f' on a neighbourhood of x.

Proof. Since $f^{00}(x; h_1, h_2) = \langle f'(\cdot), h_2\rangle^0(x; h_1)$, the assertion follows from the properties of the Clarke derivative (see [4, Proposition 2.1.1]). □

Having in mind Proposition 1.2, we claim that $d^2 f(x; h_1, h_2)$ has the same properties.

$\mathcal{L}(E, E^*)$ is a conjugate space (see Holmes [7], Chapter 23B); the w^*-topology of $\mathcal{L}(E, E^*)$ is called weak*-operator topology. The predual of $\mathcal{L}(E, E^*)$ is the linear hull V of all functionals $l_{h_1,h_2} \in (\mathcal{L}(E, E^*))^*$, $h_1, h_2 \in E$ of the form $l_{h_1,h_2}(L) := L[h_1, h_2]$ with the norm in $(\mathcal{L}(E, E^*))^*$. Then V is a separable normed space, therefore we can note the following.

Remark 1.4. Every w^*-compact subset in $\mathcal{L}(E\times E)$ is metrizable (see Holmes [7], Chapter 12F).

We have that a sequence $\{L_n\} \subset \mathcal{L}(E \times E)$ converges in the w^*-topology to some $L_0 \in \mathcal{L}(E \times E)$ iff $L_n[h_1, h_2] \to L_0[h_1, h_2]$ for every $h_1, h_2 \in E$.

There are two natural ways to define second-order subdifferentials of $f \in C^{1,1}$ (by analogy with the first-order Clarke's subdifferential (see [4], p. 27 and [4], Theorem 2.5.1).

DEFINITION 1.5.

$$\partial_c^2 f(x) := \{L \in \mathcal{L}(E \times E)\colon\ L[h_1, h_2] \leqslant f^{00}(x; h_1, h_2),\ \forall (h_1, h_2) \in E \times E\}.$$

DEFINITION 1.6.

$$\partial^2 f(x) := \overline{co}^*\Big\{L \in \mathcal{L}(E \times E)\colon\ L = \underset{G(f)\ni z\to x}{w^*\text{-lim}}\ f''(z)\Big\}.$$

An analogue of Definition 1.6 was introduced by L. Thibault in [16] in order to extend the notion of Clarke's subdifferential to locally Lipschitz mappings, acting from a separable Banach space to a reflexive space.

It is easy to see that $\partial^2 f(x) \subset \partial_c^2 f(x)$, for every $x \in G$. Indeed, let

$$L = \underset{G(f) \ni z_n \to x}{w^*\text{-lim}} f''(z_n).$$

Since $f''(z_n)[h_1, h_2] \leqslant f^{00}(z_n; h_1, h_2)$ for every $h_1, h_2 \in E$, by Proposition 1.3 we have

$$\begin{aligned} L[h_1, h_2] &= \lim_{n\to\infty} f''(z_n)[h_1, h_2] \\ &\leqslant \limsup_{n\to\infty} f^{00}(z_n; h_1, h_2) \leqslant f^{00}(x; h_1, h_2). \end{aligned}$$

Hence, $L \in \partial_c^2 f(x)$, and since $\partial_c^2 f(x)$ is obviously convex and w^*-closed, we obtain $\partial^2 f(x) \subset \partial_c^2 f(x)$.

Also, it is easy to see that

$$d^2 f(x; h_1, h_2) = \sup\{L[h_1, h_2]; L \in \partial^2 f(x)\}. \tag{1.1}$$

In the case when $E = \mathbb{R}^n$, Definition 1.6 was considered and used by J. B. H.-Urruty, J. J. Strodiot and V. H. Nguyen [8] ($\partial^2 f(x)$ was called there a generalized Hessian matrix). They used Rademacher's theorem (instead of Christensen's); then f is twice Fréchet differentiable almost everywhere. Therefore $f''(z)$ (when it exists) is a symmetric matrix (see [1], Section 2.2.5) and, hence, $\partial^2 f(x)$ consists of symmetric matrices. In our infinite-dimensional case, we cannot say that $\partial^2 f(x)$ consists of symmetric bilinear functionals.

Note that in $\mathbb{R}^n$, $\partial_c^2 f(x)$ is in fact the *plenary hull* of $\partial^2 f(x)$ in the terminology of [9]. So $\partial_c^2 f(x)$ can be essentially bigger than $\partial^2 f(x)$, but they coincide when $\partial^2 f(x)$ is a singleton (see Corollary 1.11).

PROPOSITION 1.7. *For every $x \in E$, the sets $\partial^2 f(x)$ and $\partial_c^2 f(x)$ are nonempty convex and w^*-compact. The multivalued mappings $\partial^2 f$ and $\partial_c^2 f$ are locally norm bounded in $\mathcal{L}(E \times E)$.*

Proof. Let $x_n \in G(f)$ and $x_n \to x$, as $n \to \infty$. Since the set

$$D := \{f''(x_n)\colon\ n > \nu\}$$

is norm-bounded in $\mathcal{L}(E \times E)$ for some ν by the Lipschitz constant of f' at x, we can apply the Alaoglou–Bourbaki theorem. Thus, the limit of a w^*-convergent subsequence of D belongs to $\partial^2 f(x)$. Since $\partial^2 f(x) \subset \partial_c^2 f(x)$, then $\partial_c^2 f(x) \neq \emptyset$ too. The convexity and w^*-closedness of $\partial^2 f$ and $\partial_c^2 f$ follows directly from the definitions. The locally boundedness follows from the fact that f' is locally Lipschitz.

Again by the Alaouglou–Bourbaki theorem, we obtain that $\partial^2 f(x)$ and $\partial_c^2 f(x)$ are w^*-compact. □

DEFINITION 1.8. The function $f\colon G \to \mathbb{R}$ is said to be twice strictly Gateaux differentiable at $x \in G$ if there exists $D_s^2 f(x) \in \mathcal{L}(E \times E)$, such that

$$\lim_{\substack{y \to x \\ t \downarrow 0}} \frac{f'(y + th_1)[h_2] - f'(y)[h_2]}{t} = D_s^2 f(x)[h_1, h_2], \quad \forall h_1, h_2 \in E.$$

The following three propositions are analogous of the corresponding ones concerning the first-order case (see Clarke [4]).

PROPOSITION 1.9. *If $f \in C^{1,1}$ is twice strictly Gateaux differentiable at $x \in G$, then $\partial_c^2 f(x) = \{D_s^2 f(x)\}$.*

Proof. By definition, $f^{00}(x; h_1, h_2) = D_s^2 f(x)[h_1, h_2]$ and then $L[h_1, h_2] \leqslant D_s^2 f(x)[h_1, h_2]$ for all $h_1, h_2 \in E$ and $L \in \partial_c^2 f(x)$. Hence,

$$L[h_1, h_2] = D_s^2 f(x)[h_1, h_2]$$

and therefore $L = D_s^2 f(x)$. □

PROPOSITION 1.10. *Let $f \in C^{1,1}(G)$. Then $\partial^2 f(x)$ is a singleton if and only if f is twice strictly Gateaux differentiable at $x \in G$.*

The idea of the proof is from [4].

Proof. Let $\partial^2 f(x) = \{L\}$. Then $L = w^*\text{-}\lim_{G(f) \ni z \to x} f''(z)$ and from Proposition 1.2 and (1.1),

$$f^{00}(x; h_1, h_2) = \limsup_{G(f) \ni z \to x} f''(z)[h_1, h_2] = L[h_1, h_2].$$

We may write

$$\begin{aligned}
&\liminf_{\substack{x' \to x \\ t \downarrow 0}} \frac{(f'(x' + th_1) - f'(x'))[h_2]}{t} \\
&\quad = -\limsup_{\substack{x' \to x \\ t \downarrow 0}} \frac{(f'(x') - f'(x' + th_1))[h_2]}{t} \\
&\quad = -\limsup_{\substack{x' \to x \\ t \downarrow 0}} \frac{(f'(x' + th_1 - th_1) - f'(x' + th_1))[h_2]}{t} \\
&\quad = -f^{00}(x; -h_1, h_2) = -L[-h_1, h_2] = L[h_1, h_2] = f^{00}(x; h_1, h_2),
\end{aligned}$$

whence f is twice strictly differentiable.

The other direction follows from the inclusion $\partial_c^2 f(x) \supset \partial^2 f(x)$ and from Proposition 1.9. □

By Propositions 1.9 and 1.10, we obtain immediately the following.

COROLLARY 1.11. *For a $C^{1,1}$ function $f\colon G \to \mathbb{R}$, $\partial_c^2 f(x)$ is a singleton ($x \in G$) if and only if $\partial^2 f(x)$ is a singleton.*

The arguments for proving the following proposition are now classical (similar to the first order case) and the proof is omitted.

PROPOSITION 1.12. *The multivalued mappings $\partial_c^2 f$ and $\partial^2 f\colon (G, \|\cdot\|) \to (\mathcal{L}(E \times E), w^*)$ are upper semicontinuous.*

Let $\psi\colon \mathbb{R} \to G$ be an affine function, $\phi\colon G \to \mathbb{R}$ be a $C^{1,1}$ function. It is clear that $\phi \circ \psi \in C^{1,1}(\mathbb{R})$.

It is easy to derive in the same way like in [8], Theorem 2.2, the following.

PROPOSITION 1.13. *For all $x_0, u, v \in \mathbb{R}$*

$$(\phi \circ \psi)^{00}(x_0; u, v) = \phi^{00}(\psi(x_0);\ \psi'(x_0)u, \psi'(x_0)v).$$

From Propositions 1.2 and 1.13, we have that

$$d^2(\phi \circ \psi)(x_0; u, v) = d^2\phi(\psi(x_0);\ \psi'(x_0)u, \psi'(x_0)v).$$

Since $d^2 f(x_0; h_1, \cdot)$ is by (1.1) the support function of $\partial^2 f(x_0)[h_1]$, we derive that

$$\partial^2(\phi \circ \psi)(x_0) = \bigcup \left\{ L[\psi'(x_0), \psi'(x_0)]\colon\ L \in \partial^2\phi(\psi(x_0)) \right\}. \tag{1.2}$$

The following proposition is stated in [8] without proof. Here we include the proof for completeness.

PROPOSITION 1.14. *Let I be an open interval containing $[0,1]$ and let $\phi \in C^{1,1}(I)$. Then*

$$\phi(1) - \phi(0) - \phi'(0) \in \tfrac{1}{2}\partial^2\phi(t)$$

for some $t \in (0,1)$.

Proof. Define

$$h(t) = \phi(1) - \phi(t) - \phi'(t)(1-t) - (1-t)^2\lambda, \quad t \in [0,1],$$

where $\lambda = \phi(1) - \phi(0) - \phi'(0)$. So we have $h(1) = h(0) = 0$. Obviously, h is locally Lipschitz on [0,1]. There exists $\xi \in (0,1)$ such that either

(1) ξ is a minimum of h over [0,1], or
(2) ξ is a maximum of h over [0,1].

Let (1) be fulfilled. Then by the necessary condition for a local minimum (see Clarke [4])

$$0 \in \partial h(\xi) = -\partial\phi'(\xi)(1-\xi) + 2(1-\xi)\lambda.$$

Hence, $\lambda \in \frac{1}{2}\partial\phi'(\xi)$. But by Clarke [4], Theorem 2.5.1, and by definition of $\partial^2\phi$, we have $\partial\phi'(\xi) = \partial^2\phi(\xi)$.

Let (2) be fulfilled. Then ξ is a minimum point of the function $-h$ over [0,1] and since $\partial(-h)(\xi) = -\partial h(\xi)$ (Clarke [4]), we have $0 \in \partial(-h)(\xi) = -\partial h(\xi)$, so $0 \in \partial h(\xi)$ and as in case (1) we obtain $\lambda \in \frac{1}{2}\partial^2\phi(\xi)$. □

Using (1.2) and Proposition 1.14 for $\phi(x) = f(x)$ and $\psi(t) = a + t(b-a)$, we immediately obtain the second-order expansion (the same as in [8], Theorem 2.3).

PROPOSITION 1.15. *Let $f \in C^{1,1}(G)$. Then for every $a, b \in G$, with $[a,b] \subset G$ there exists $c \in (a,b)$ and $L_c \in \partial^2 f(c)$ such that*

$$f(b) = f(a) + f'(a)[b-a] + \tfrac{1}{2} L_c[b-a, b-a].$$

We shall use essentially this proposition in the sequel.

2. Necessary and Sufficient Optimality Conditions

We now consider the following constrained minimization problem:

$$P(E)\quad \begin{cases} f_0(x) \to \min, \\ x \in E, \\ f_i(x) \leqslant 0, \quad 1 \leqslant i \leqslant m, \\ F(x) = 0, \end{cases}$$

where $F(x) = (g_1(x), \ldots, g_k(x))^T$ and the functions f_0, f_i, $1 \leqslant i \leqslant m$, g_j, $1 \leqslant j \leqslant k$ are $C^{1,1}(E)$ functions. The Lagrangian function for $P(E)$ is

$$\mathcal{L}(x; \lambda, \mu) = \sum_{i=0}^{m} \lambda_i f_i(x) + \sum_{j=1}^{k} \mu_j g_j(x),$$

where

$$(\lambda, \mu) := (\lambda_0, \ldots, \lambda_m, \mu_1, \ldots, \mu_k) \in \mathbb{R}^{m+1} \times \mathbb{R}^k$$

are the Lagrange multipliers.

Denote by $\operatorname{locmin} P(E)$ the set of all points of local minimum of $P(E)$ and define

$$\Lambda(x) = \Bigg\{ (\lambda, \mu) \in \mathbb{R}^{m+1} \times \mathbb{R}^k \colon \sum_{i=0}^{m} \lambda_i f_i'(x) + \sum_{j=1}^{k} \mu_j g_j'(x) = 0,$$

$$\lambda_i f_i(x) = 0,\ 1 \leqslant i \leqslant m,\ \lambda_i \geqslant 0,\ 0 \leqslant i \leqslant m,\ \sum_{i=0}^{m} \lambda_i = 1 \Bigg\};$$

$$K(x) = \{h \in E\colon\ f_i'(x)[h] \leqslant 0,\ 0 \leqslant i \leqslant m,\ g_j'(x)[h] = 0,\ 1 \leqslant j \leqslant k\}.$$

Further, we need the following facts.

PROPOSITION 2.1 (Necessary condition, [1], Section 3.4.2). *Let*

$$x_0 \in \operatorname{locmin} P(E) \quad \textit{and} \quad \operatorname{Im} F'(x) = \mathbb{R}^k.$$

Then $\Lambda(x_0)$ *is a nonempty convex compact set.*

LEMMA 2.2 (Minimax, [1], Section 3.3.4). *Let* $A\colon E \to \mathbb{R}^k$ *be a linear continuous surjective operator,* $AE = \mathbb{R}^k$, $x_i^* \in E^*$, $0 \leqslant i \leqslant m$, $y \in \mathbb{R}^k$, $a \in \mathbb{R}^{m+1}$,

$$\max_{0\leqslant i\leqslant m} x_i^*[x] \geqslant 0 \quad \forall x \in \operatorname{Ker} A. \tag{2.1}$$

Denote

$$S(a,y) = \inf_{Ax+y=0} \max_{0\leqslant i\leqslant m} (a_i + x_i^*[x]). \tag{2.2}$$

Then

$$\text{(a)} \quad S(a,y) = \sup_{(\lambda,\mu)\in\Lambda} \left(\sum_{i=0}^{m} \lambda_i a_i + \sum_{j=1}^{k} \mu_j y_j\right), \tag{2.3}$$

where

$$\Lambda = \left\{(\lambda,\mu) \in \mathbb{R}^{m+1} \times \mathbb{R}^k\colon\ \lambda_i \geqslant 0,\ \sum_{i=0}^{m} \lambda_i = 1,\ \sum_{i=0}^{m} \lambda_i x_i^* + A^*\mu = 0\right\}.$$

(b) inf *in* (2.2) *and* sup *in* (2.3) *are attained.*

Let $x_0 \in \operatorname{locmin} P(E)$. Without loss of generality, we may assume in the sequel that $f_i(x_0) = 0, 0 \leqslant i \leqslant m$ (see the reduction in [1], Section 3.2.3).

Consider the problem

$$(\mathrm{P}') \quad f(x) := \max\{f_0(x), \ldots, f_m(x)\} \to \min;\ F(x) = 0.$$

It is easy to see the validity of the following.

LEMMA 2.3 ([1], Section 3.4.2). $x_0 \in \operatorname{locmin}(\mathrm{P}')$.

Now we can state the necessary optimality condition for $P(E)$. The proof uses ideas of the proof of [5], Theorem 9.1.2 (see also [1], Theorem 3.4.2), where the functions are assumed to be of class C^2.

THEOREM 2.4. *Let in $P(E)$ $\operatorname{Im} F'(x_0) = \mathbb{R}^k$. If $x_0 \in \operatorname{locmin} P(E)$, then for every $h \in K(x_0)$ there exist $L_i \in \partial^2 f_i(x_0)$, $0 \leqslant i \leqslant m$, $M_j \in \partial^2 g_j(x_0)$, $1 \leqslant j \leqslant k$ such that*

$$\max_{(\lambda,\mu)\in\Lambda(x_0)} \left(\sum_{i=0}^{m} \lambda_i L_i[h,h] + \sum_{j=1}^{k} \mu_j M_j[h,h] \right) \geqslant 0.$$

Proof. By Proposition 2.1, the set $\Lambda(x_0)$ is nonempty.

Let us assume the contrary, i.e. there exists $h \in K(x_0)$ such that for every $L_i \in \partial^2 f_i(x_0)$, $0 \leqslant i \leqslant m$, and every $M_j \in \partial^2 g_j(x_0)$, $1 \leqslant j \leqslant k$, we have

$$\max_{(\lambda,\mu)\in\Lambda(x_0)} \left(\sum_{i=0}^{m} \lambda_i L_i[h,h] + \sum_{j=1}^{k} \mu_j M_j[h,h] \right) < 0. \tag{2.4}$$

It is clear that $\|h\| \neq 0$. Let $0 < t_n \to 0$ as $n \to \infty$. From Proposition 1.15 we have

$$\begin{aligned} g_j(x_0 + t_n h) &= g_j(x_0) + t_n g_j'(x_0)[h] + \frac{t_n^2}{2} M_{j,n}[h,h] \\ &= \frac{t_n^2}{2} M_{j,n}[h,h], \end{aligned}$$

where $M_{j,n} \in \partial^2 g_j(x_0 + \gamma_{j,n} t_n h)$ and $\gamma_{j,n} \in (0,1)$.

Since $x_0 + \gamma_{j,n} t_n h \to x_0$, as $n \to \infty$, and $\partial^2 g_j$ are locally bounded and have w^*-closed graphs, we can choose w^*-convergent subsequences from $\{M_{j,n}\}_{n\geqslant 1}$, whose w^*-limits M_j are in $\partial^2 g_j(x_0)$ for $1 \leqslant j \leqslant k$.

Analoguously, having in mind that $f_i(x_0) = 0$ (see the remark before Lemma 2.3),

$$\begin{aligned} f_i(x_0 + t_n h) &= f_i(x_0) + t_n f_i'(x_0)[h] + \frac{t_n^2}{2} L_{i,n}[h,h] \\ &= t_n f_i'(x_0)[h] + \frac{t_n^2}{2} L_{i,n}[h,h], \end{aligned}$$

where $L_{i,n} \in \partial^2 f_i(x_0 + \eta_{i,n} t_n h)$, $\eta_{i,n} \in (0,1)$ and choose w^*-convergent subsequences from $\{L_{i,n}\}_{n\geqslant 1}$, whose w^*-limits L_i are in $\partial^2 f_i(x_0)$ for $0 \leqslant i \leqslant m$.

For every $\varepsilon > 0$, there exists an integer N_1 such that for every $n \geqslant N_1$ the following inequalities are fulfilled:

$$|M_{j,n}[h,h] - M_j[h,h]| < 2\varepsilon, \qquad |L_{i,n}[h,h] - L_i[h,h]| < 2\varepsilon.$$

Hence, for $n \geqslant N_1$ we have

$$g_j(x_0 + t_n h) = \frac{t_n^2}{2} M_j[h,h] + \phi_{j,n}(\varepsilon), \tag{2.5}$$

where

$$|\phi_{j,n}(\varepsilon)| < t_n^2\varepsilon \quad \text{for } 1 \leqslant j \leqslant k,\ n \geqslant N_1 \tag{2.6}$$

and

$$f_i(x_0 + t_n h) = t_n f_i'(x_0)[h] + \frac{t_n^2}{2} L_i[h,h] + \psi_{i,n}(\varepsilon), \tag{2.7}$$

where

$$|\psi_{i,n}(\varepsilon)| < t_n^2\varepsilon \quad \text{for } 0 \leqslant i \leqslant m,\ n \geqslant N_1. \tag{2.8}$$

Denote

$$x_i^* = f_i'(x_0), \qquad a_i = \tfrac{1}{2} L_i[h,h], \quad 0 \leqslant i \leqslant m,$$

$$y_j = \tfrac{1}{2} M_j[h,h], \quad 1 \leqslant j \leqslant k, \qquad y = (y_1, \dots, y_k), \qquad A = F'(x_0).$$

From [1], Section 3.2.4, Lemma 1, the condition (2.1) of Lemma 2.2 is fulfilled. Applying Lemma 2.2, we can find $\xi = \xi(h) \in E$, such that

$$F'(x_0)\xi + y = 0 \tag{2.9}$$

and

$$\max_{0\leqslant i\leqslant m} (a_i + f_i'(x_0)[\xi]) = \max_{(\lambda,\mu)\in\Lambda(x_0)} \left(\sum_{i=0}^{m} \lambda_i a_i + \sum_{j=1}^{k} \mu_j y_j \right) =: \Psi(h), \tag{2.10}$$

where $\Psi(h) < 0$ from (2.4).

Let l be the maximum of the Lipschitz constants for g_j' and f_i' in a neighborhood U_1 of x_0. There exists a neighbourhood $U_2 \subset U_1$ of x_0 and a constant s such that g_j' and f_i' are norm-bounded by s there.

Using the mean-value theorem, (2.5) and (2.9), for large n, we have

$$\begin{aligned} & g_j(x_0 + t_n h + t_n^2\xi) \\ &\quad = g_j(x_0 + t_n h + t_n^2\xi) - g_j(x_0 + t_n h) + g_j(x_0 + t_n h) \\ &\quad = t_n^2 g_j'(x_0 + t_n h + \kappa_{j,n} t_n^2 \xi)[\xi] + \frac{t_n^2}{2} M_j[h,h] + \phi_{j,n}(\varepsilon) \\ &\quad = t_n^2 g_j'(x_0 + t_n h + \kappa_{j,n} t_n^2 \xi)[\xi] - t_n^2 g_j'(x_0)[\xi] + \phi_{j,n}(\varepsilon) \\ &\quad \leqslant t_n^2 l \|t_n h + \kappa_{j,n} t_n^2 \xi\| \|\xi\| + \phi_{j,n}(\varepsilon) \\ &\quad \leqslant t_n^3 l(\|h\| + t_n\|\xi\|)\|\xi\| + |\phi_{j,n}(\varepsilon)| =: \theta_{j,n}(\varepsilon), \end{aligned}$$

where $\kappa_{j,n} \in (0,1)$.

Using (2.6), we obtain

$$\theta_{j,n}(\varepsilon) \leqslant \mathrm{o}(t_n^2) + t_n^2\varepsilon, \quad \forall j = 1, \ldots, k, \tag{2.11}$$

where $\mathrm{o}(t_n^2)/t_n^2 \to 0$.

Now we apply a generalization of the implicit function theorem (see [1], Theorem 2.3.1) and obtain that there exist a constant q and a map $\Phi\colon U \to \mathbb{R}^k$, where $U \subset U_2$ is a neighbourhood of the point x_0, such that

$$F(x + \Phi(x)) = 0, \qquad \|\Phi(x)\| \leqslant q\|F(x)\|, \quad \forall x \in U.$$

Substitute $r(t_n) = \Phi(x_0 + t_n h + t_n^2\xi)$. Then, for n sufficiently large,

$$x_0 + t_n h + t_n^2\xi \in U, \qquad g_j(x_0 + t_n h + t_n^2\xi + r(t_n)) = 0 \quad \text{for } 1 \leqslant j \leqslant k,$$

and by (2.11), we have

$$\begin{aligned} \|r(t_n)\| &\leqslant q\|F(x_0 + t_n h + t_n^2\xi)\| \\ &= q\left(\sum_{j=1}^{k} g_j^2(x_0 + t_n h + t_n^2\xi)\right)^{1/2} \leqslant q\sqrt{k}\big(\mathrm{o}(t_n^2) + t_n^2\varepsilon\big). \end{aligned} \tag{2.12}$$

Again, using the mean-value theorem, (2.7) and the fact that $h \in K(x_0)$, for large n we obtain

$$\begin{aligned} & f_i(x_0 + t_n h + t_n^2\xi + r(t_n)) \\ &\quad = f_i(x_0 + t_n h + t_n^2\xi + r(t_n)) - f_i(x_0 + t_n h) + f_i(x_0 + t_n h) \\ &\quad = f_i'(x_0 + t_n h + \nu_{i,n}(t_n^2\xi + r(t_n))[t_n^2\xi + r(t_n)] + \\ &\qquad + t_n f_i'(x_0)[h] + \frac{t_n^2}{2}L_i[h,h] + \psi_{i,n}(\varepsilon) \\ &\quad \leqslant t_n^2 l\|t_n h + \nu_{i,n}(t_n^2\xi + r(t_n))\| \cdot \|\xi\| + \\ &\qquad + s\|r(t_n)\| + t_n^2 f_i'(x_0)[\xi] + \frac{t_n^2}{2}L_i[h,h] + \psi_{i,n}(\varepsilon), \end{aligned}$$

where $\nu_{i,n} \in (0,1)$. Hence, from (2.8), (2.10) and (2.12), we obtain

$$\begin{aligned} & f(x_0 + t_n h + t_n^2\xi + r(t_n)) \\ &\quad := \max_{0 \leqslant i \leqslant m} f_i(x_0 + t_n h + t_n^2\xi + r(t_n)) \\ &\quad \leqslant \max_{0 \leqslant i \leqslant m} [t_n^2 l\|t_n h + \nu_{i,n}(t_n^2\xi + r(t_n))\| \cdot \|\xi\| + \\ &\qquad + s\|r(t_n)\| + t_n^2 f_i'(x_0)[\xi] + \frac{t_n^2}{2}L_i[h,h] + \psi_{i,n}(\varepsilon)] \\ &\quad \leqslant \mathrm{o}_1(t_n^2) + (t_n^2 l\|\xi\| + s)\|r(t_n)\| + t_n^2\Psi(h) + t_n^2\varepsilon \\ &\quad \leqslant \mathrm{o}_1(t_n^2) + (t_n^2 l\|\xi\| + s)q\sqrt{k}[\mathrm{o}(t_n^2) + t_n^2\varepsilon] + t_n^2\Psi(h) + t_n^2\varepsilon \\ &\quad = t_n^2\left[\frac{\mathrm{o}_1(t_n^2)}{t_n^2} + (t_n^2 l\|\xi\| + s)q\sqrt{k}\Big(\frac{\mathrm{o}(t_n^2)}{t_n^2} + \varepsilon\Big) + \Psi(h) + \varepsilon\right], \end{aligned}$$

where $o_1(t_n^2)/t_n^2 \to 0$.

Now, since $\Psi(h) < 0$, it is clear that if we chose ε to be sufficiently small, then for sufficiently large n, we have

$$f(x_0 + t_n h + t_n^2 \xi + r(t_n)) < 0,$$

which is in contradiction with Lemma 2.3. □

The second-order sufficient condition for the problem $P(E)$ is the following.

THEOREM 2.5. *Let in $P(E)$ $f_i(x_0) = 0, 0 \leqslant i \leqslant m$, $\operatorname{Im} F'(x_0) = \mathbb{R}^k$, $\Lambda(x_0) \neq \varnothing$ and there exists a constant $\alpha > 0$ such that for every $L_i \in \partial^2 f_i(x_0)$, $0 \leqslant i \leqslant m$, $M_j \in \partial^2 g_j(x_0)$, $1 \leqslant j \leqslant k$ we have*

$$\max_{(\lambda,\mu)\in\Lambda(x_0)} \left(\sum_{i=0}^{m} \lambda_i L_i[h,h] + \sum_{j=1}^{k} \mu_j M_j[h,h] \right) \geqslant \alpha \|h\|^2 \quad \forall h \in K(x_0).$$

Then x_0 is a strict local minimum (i.e. unique minimum in a neighborhood of x_0) of the problem $P(X)$ for every finite-dimensional subspace $X \ni x_0$ of E. *If, in addition, the functions f_i, $1 \leqslant i \leqslant m$, are convex and the functions g_j, $0 \leqslant j \leqslant k$, are affine, then the problem $P(X)$ has unique solution x_0. Also in this case the problem $P(E)$ has unique solution x_0.*

Proof. We use ideas of [5], Theorem 10.1.1, where the functions are assumed to be C^2.

We shall show that for every finite-dimentional subspace $X \ni x_0$ there exists $\delta > 0$ such that the conditions: $h \in \delta B \cap X$ with $h \neq 0$ and

$$f_i(x_0 + h) \leqslant 0, \quad 0 \leqslant i \leqslant m, \qquad F(x_0 + h) = 0, \tag{2.13}$$

where B is the unit ball in E, are inconsistent. From this, we will obtain immediately the desired conclusion.

Let us assume the contrary: there exists a finite-dimentional subspace $X \ni x_0$, such that for every $\delta > 0$ there exists a nonzero $h_\delta \in \delta B \cap X$ such that the conditions (2.13) are fulfilled.

Denote $H = \bigcup_{\delta>0} h_\delta$.

Let $h \in H$. From Proposition 1.15 we have

$$f_i(x_0 + h) = f_i'(x_0)[h] + \tfrac{1}{2} L_i(x_0 + \eta_i h)[h,h], \tag{2.14}$$

where

$$\eta_i \in (0,1), \qquad L_i(x_0 + \eta_i h) \in \partial^2 f_i(x_0 + \eta_i h), \quad 0 \leqslant i \leqslant m;$$

$$g_j(x_0 + h) = g_i'(x_0)[h] + \tfrac{1}{2} M_j(x_0 + \gamma_j h)[h,h], \tag{2.15}$$

where

$$\gamma_j \in (0,1), \qquad M_j(x_0 + \gamma_j h) \in \partial^2 g_j(x_0 + \gamma_j h), \quad 1 \leqslant j \leqslant k.$$

Since $\Lambda(x_0) \neq \emptyset$, it follows from the proof of Theorem in [1], Section 3.4.2, that it is a compact and, therefore, there exists a constant c_1, such that for every $(\lambda, \mu) \in \Lambda(x_0)$ we have $\sum_{i=1}^{k} \mid \mu_i \mid \leqslant c_1$, where $\mu = (\mu_1, \ldots, \mu_k)$.

Substitute

$$x_i^* = f_i'(x_0),$$

$$a_i = \tfrac{1}{2} L_i(x_0 + \eta_i h)[h,h], \quad 0 \leqslant i \leqslant m, \qquad A = F'(x_0)$$

(recall that $F(x) = (g_1(x), \ldots, g_k(x))$),

$$y_j = \tfrac{1}{2} M_j(x_0 + \gamma_j h)[h,h], \quad 1 \leqslant j \leqslant k, \ y = (y_1, \ldots, y_k),$$

$$f(x) = \max_{0 \leqslant i \leqslant m} f_i(x).$$

From (2.14) and (2.15), we obtain that (2.13) is equivalent to

$$x_i^*[h] + a_i = f_i(x_0 + h) \leqslant 0, \quad 0 \leqslant i \leqslant m, \qquad Ah + y = 0. \tag{2.16}$$

Since $\partial^2 f_i$ and $\partial^2 g_j$ are locally bounded, we have

$$\begin{aligned} &\exists c_2 > 0, \ \exists \delta_1 > 0\colon \ x \in B(x_0, \delta_1), \ L_i \in \partial^2 f_i(x), \quad 0 \leqslant i \leqslant m, \\ &M_j \in \partial^2 g_j(x), \quad 1 \leqslant j \leqslant k, \Rightarrow \|L_i\| < c_2, \quad \|M_j\| < c_2 \end{aligned} \tag{2.17}$$

and if we define $x_i^*[h]_+ := \max\{x_i^*[h], 0\}$, then

$$\|h\| < \delta_1 \Rightarrow x_i^*[h]_+ \leqslant |a_i| \leqslant \frac{c_2}{2}\|h\|^2,$$

$$\|Ah\| = \|y\| \leqslant \frac{c_2}{2}\|h\|^2. \tag{2.18}$$

For $(\lambda, \mu) \in \Lambda(x_0)$ we have

$$\sum_{i=0}^{m} \lambda_i x_i^*[x] + \sum_{j=1}^{k} \mu_j g_j'(x_0)[x] = 0 \quad \forall x \in E.$$

Hence $\sum_{i=0}^{m} \lambda_i x_i^*[x] = 0, \forall x \in \operatorname{Ker} A$, and therefore

$$\max_{0 \leqslant i \leqslant m} x_i^*[x] \geqslant 0, \quad \forall x \in \operatorname{Ker} A,$$

which is the condition (2.1) of Lemma 2.2. From (2.16) and Lemma 2.2, we obtain

$$\begin{aligned} &\max_{0 \leqslant i \leqslant m} f_i(x_0 + h) \\ &\quad = \max_{0 \leqslant i \leqslant m} (x_i^*[h] + a_i) \\ &\quad \geqslant \min_{Ax+y=0} \max_{0 \leqslant i \leqslant m} (x_i^*[x] + a_i) \\ &\quad = \max_{(\lambda,\mu) \in \Lambda(x_0)} \frac{1}{2} \left(\sum_{i=0}^{m} \lambda_i L_i(x_0 + \eta_i h)[h,h] + \right. \\ &\qquad\qquad \left. + \sum_{j=1}^{k} \mu_j M_j(x_0 + \gamma_j h)[h,h] \right). \end{aligned} \tag{2.19}$$

We estimate the distance from h to the cone $K(x_0)$ by Hoffman's lemma ([1], Section 3.3.4) and after that by (2.18):

$$\begin{aligned} d(h, K(x_0)) &:= \inf\{\|h-y\|,\ y \in K(x_0)\} \\ &\leqslant c\left(\sum_{i=0}^{m} x_i^*[h]_+ + \|Ah\|\right) < c_3\|h\|^2 \end{aligned}$$

for some constant c, which does not depend on h; $c_3 := c \cdot c_2$.

Hence, h can be represented in the type $h = h' + h''$, where

$$h' \in K(x_0), \qquad \|h''\| < c_3\|h\|^2. \tag{2.20}$$

If $\|h\| < 1/2c_3$, then

$$\tfrac{1}{2}\|h\| < (1 - c_3\|h\|)\|h\| \leqslant \|h'\| \leqslant \|h\|(1 + c_3\|h\|) \leqslant 2\|h\|. \tag{2.21}$$

Since $\partial^2 f_i$ and $\partial^2 g_j$ are w^*-upper semicontinuous and locally bounded, and since the w^*-compact sets in separable dual spaces are sequentially w^*-compact (from Remark 1.4), we can find (applying Alaoglou–Bourbaki theorem) a sequence $\{h_n\} \subset H$, $\|h_n\| \to 0$ and elements $L_i \in \partial^2 f_i(x_0)$, $M_j \in \partial^2 g_j(x_0)$ such that

$$L_i(x_0 + \eta_i h_n) \xrightarrow{w^*} L_i, \qquad M_j(x_0 + \gamma_j h_n) \xrightarrow{w^*} M_j.$$

Hence there exists ν such that for every $h \in \{h_n\}_{n \geqslant \nu}$ we have

$$|L_i(x_0 + \eta_i h)[h,h] - L_i[h,h]| < \frac{\alpha\|h\|^2}{16}, \tag{2.22}$$

$$|M_j(x_0 + \gamma_j h)[h,h] - M_j[h,h]| < \frac{\alpha\|h\|^2}{16c_1} \tag{2.23}$$

(here we use the fact that X is finite-dimensional, i.e. the restrictions of $L_i(x_0 + \eta_i h_n)$ and $M_j(x_0 + \gamma_j h_n)$ to $X \times X$ converge in the norm topology, respectively, to the restrictions of L_i and M_j to $X \times X$, when $n \to \infty$).

Then for every $h \in \{h_n\}_{n \geqslant \nu}$, with $\|h\| < \min\{1/2c_3, \delta_1\}$, having in mind that $h = h' + h''$, where $h' \in K(x_0)$ and h'' satisfy (2.20) and (2.21), and using (2.19), (2.22), and (2.23) we obtain:

$$\begin{aligned} 0 &\geqslant f(x_0 + h) \\ &\geqslant \max_{(\lambda,\mu)\in\Lambda(x_0)} \frac{1}{2}\left(\sum_{i=0}^{m} \lambda_i L_i(x_0 + \eta_i h)[h,h] + \sum_{j=1}^{k} \mu_j M_j(x_0 + \gamma_j h)[h,h]\right) \\ &\geqslant -\frac{\alpha\|h\|^2}{16} + \max_{(\lambda,\mu)\in\Lambda(x_0)} \frac{1}{2}\left(\sum_{i=0}^{m} \lambda_i L_i[h,h] + \sum_{j=1}^{k} \mu_j M_j[h,h]\right) \end{aligned}$$

$$\geqslant -\frac{\alpha\|h\|^2}{16} + \max_{(\lambda,\mu)\in\Lambda(x_0)} \frac{1}{2}\Bigg(\sum_{i=0}^{m}\lambda_i L_i[h',h'] + \sum_{j=1}^{k}\mu_j M_j[h',h'] +$$

$$+\sum_{i=0}^{m}\lambda_i L_i[h',h''] + \sum_{i=0}^{m}\lambda_i L_i[h'',h'] + \sum_{i=0}^{m}\lambda_i L_i[h'',h''] +$$

$$+\sum_{j=1}^{k}\mu_j M_j[h',h''] + \sum_{j=1}^{k}\mu_j M_j[h'',h'] + \sum_{j=1}^{k}\mu_j M_j[h'',h'']\Bigg)$$

$$\geqslant -\frac{\alpha\|h\|^2}{16} + \frac{\alpha}{2}\|h'\|^2 - \Bigg(\sum_{i=0}^{m}\lambda_i\|L_i\|\|h'\|\|h''\| + \frac{1}{2}\sum_{i=0}^{m}\lambda_i\|L_i\|\|h''\|^2 +$$

$$+\sum_{j=1}^{k}|\mu_j|\,\|M_j\|\|h'\|\|h''\| + \frac{1}{2}\sum_{j=1}^{k}|\mu_j|\,\|M_j\|\|h''\|^2\Bigg)$$

$$\geqslant -\frac{\alpha\|h\|^2}{16} + \frac{\alpha}{8}\|h\|^2 - c_2\|h'\|\|h''\|(1+c_1) - \frac{1}{2}c_2\|h''\|^2(1+c_1)$$

$$\geqslant \|h\|^2\left[\frac{\alpha}{16} - 2c_2c_3(1+c_1)\|h\| - \frac{1}{2}c_2c_3^2(1+c_1)\|h\|^2\right].$$

The last expression is positive, when h is sufficiently small. This is a contradiction.

Therefore x_0 is a strict local minimum of $P(X)$. When f_i, $0 \leqslant i \leqslant m$, are convex and g_j, $1 \leqslant j \leqslant k$, are affine functions, then the Lagrange function is convex and the admissible set is convex, therefore the local minimum is global. Obviously x_0 is a strict global minimum of the problem $P(E)$ too. □

The following sufficient condition is a modification of that one in [1], Section 3.4.3.

THEOREM 2.6. *Let in $P(E)$ $f_i(x_0) = 0$, $1 \leqslant i \leqslant m$, $\operatorname{Im} F'(x_0) = \mathbb{R}^k$, there exists a number $\alpha > 0$ and Lagrange's multipliers $(\lambda,\mu) \in \mathbb{R}^{m+1} \times \mathbb{R}^k$, such that $\lambda_0 = 1$, $\lambda_i > 0$, $1 \leqslant i \leqslant m$,*

$$\mathcal{L}'_x(x_0;\lambda,\mu) = f'_0(x_0) + \sum_{i=1}^{m}\lambda_i f'_i(x_0) + \sum_{j=1}^{k}\mu_j g'_j(x_0) = 0 \tag{2.24}$$

and for all $L \in \partial^2\mathcal{L}(x_0;\lambda,\mu)$

$$L[h,h] \geqslant 2\alpha\|h\|^2, \quad \forall h \in C(x_0), \tag{2.25}$$

where

$$C(x_0) = \{h \in E\colon\ f'_i(x_0)[h] = 0,\ 1 \leqslant i \leqslant m,\ F'(x_0)[h] = 0\}.$$

Then x_0 is a local minimum of the problem $P(X)$ for every finite-dimensional subspace $X \ni x_0$.

Proof. We shall follow the proof of Theorem in [1], Section 3.4.3. Instead of the usual Taylor expansion, we use Proposition 1.15.

Assume the contrary, i.e. there exists a finite-dimensional subspace $X \ni x_0$, such that for every $\delta > 0$ there exists $h_\delta \in X \cap \delta B$ (B is the closed unit ball in E) such that

$$f_i(x_0 + h_\delta) \leqslant 0, \quad 1 \leqslant i \leqslant m, \qquad F(x_0 + h_\delta) = 0,$$

and

$$f_0(x_0 + h_\delta) < f_0(x_0).$$

Denote $H := \bigcup_{\delta > 0} h_\delta$ and let $h \in H$, $h \neq 0$.

Since $\sum_{i=1}^{m} \lambda_i f_i(x_0 + h) \leqslant 0$, from Proposition 1.15 applied for the Lagrangian function $\mathcal{L}$, we have

$$f_0(x_0 + h) \geqslant \mathcal{L}(x_0; \lambda, \mu) + \tfrac{1}{2} L(x_0 + \eta_h h)[h, h],$$

where $L(x_0 + \eta_h h) \in \partial^2 \mathcal{L}(x_0 + \eta_h h; \lambda, \mu)$, $\eta_h \in (0, 1)$. The local boundedness of $\partial^2 \mathcal{L}(\cdot; \lambda, \mu)$ allows us to find a sequence $\{h_n\} \subset H$, $\|h_n\| \to 0$, such that $L_n := L(x_0 + \eta_{h_n} h_n)$ is w^*-convergent to some $L \in \partial^2 \mathcal{L}(x_0; \lambda, \mu)$ (by the w^*-upper semicontinuity of $\partial^2 \mathcal{L}(\cdot; \lambda, \mu)$). Since X is finite-dimensional, the restrictions of L_n to $X \times X$ converge to the restriction of L to $X \times X$ in the norm topology.

So we have

$$f_0(x_0 + h_n) \geqslant f_0(x_0) + \frac{1}{2} L[h_n, h_n] - \frac{\alpha}{2} \|h\|^2, \quad \forall n > \nu \tag{2.26}$$

for some ν.

Further, the proof is the same as the proof of Theorem in [1], Section 3.4.3. The only difference is that the inequality (7) from [1, p. 191] is replaced by (2.26). In such a way, for $n > \nu$, we have $f_0(x_0 + h_n) \geqslant f_0(x_0)$, which is a contradiction. □

The authors are very grateful to the referees for their valuable remarks and suggestions, and esspecially for observing the new proof of Proposition 1.2.

References

1. Alekseev, V. M., Tykhomirov, V. M., and Fomin, S. V.: Optimal control, in: R. Gamkrelidze (ed.), *Contemporary Soviet Mathematics*, Consultants Bureau, New York, London, 1987.
2. Aubin, J.-P. and Frankowska, H.: *Set-Valued Analysis*, Birkhauser-Verlag, Boston, 1990.
3. Christensen, J. P. R.: *Topology and Borel Structure*, North-Holland Mathematics Studies 10, Notas de Matematica (51), L. Nachbin (ed.), North-Holland, American Elservier, Amsterdam, London, New York, 1974.
4. Clarke, F. H.: *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.
5. Galeev, E. M. and Tykhomirov, V. M.: *A Brief Course of the Theory of Extremal Problems*, Moscow State Univ., 1989 (in Russian).

6. Georgiev, P. G. and Zlateva, N. P.: Second subdifferentials of $C^{1,1}$ functions: optimality conditions and well posedness, *Compt. Rend. Acad. Bulg. Sci.* **46**(11) (1993), 25–28.
7. Holmes, R. B.: *Geometric Functional Analysis and Its Application*, Graduate Text in Mathematics 24, Springer-Verlag, New York, 1975.
8. Hiriart-Urruty, J. B., Strodiot, J. J. and Nguyen, V. H.: Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data, *Appl. Math. Optim.* **11** (1984), 43–56.
9. Hiriart-Urruty, J. B.: Characterizations of the plenary hull of the generalized Jacobian matrix, *Math. Programming Study* **17** (1982), 1–12.
10. Ioffe, A. D.: On some recent developments in the theory of second order optimality conditions, in: S. Dolecki (ed.), *Optimization, Proc. 5th French-German Conf. Varetz / Fr 1988*, Lecture Notes in Math. 1405, Springer-Verlag, Berlin, 1989, pp. 55–68.
11. Ioffe, A. D.: Composite optimization: Second order conditions, value functions and sensitivity, in: *Analysis and Optimization on Systems, Proc. 9th Int. Conf., Antibes / Fr. 1990*, Lecture Notes Control Inf. Sci. 144, Springer-Verlag, Berlin, 1990, pp. 442–451.
12. Penot, J. P. and Michael, P.: Second order moderate derivatives, *Nonlinear Anal. Theory, Meth. Appl.* **22** (1994), 809–823.
13. Phelps, R. R.: *Convex Functions, Monotone Operators and Differentiability*, 2nd edn, Lecture Notes in Math. 1364, Springer-Verlag, Berlin, 1993.
14. Rockafellar, R. T.: First- and second-order epi-differentiability in nonlinear programming, *Trans. Amer. Math. Soc.* **307** (1988), 75–108.
15. Rockafellar, R. T.: Generalized second derivatives of convex functions and saddle functions, *Trans. Amer. Math. Soc.* **322** (1990), 51–77.
16. Thibault, L.: On generalized differentials and subdifferentials of Lipschitz vector-valued functions, *Nonlinear Anal. Theory, Meth. Appl.* **6** (1982), 1037–1053.
17. Yang, X.: Second-order conditions in $C^{1,1}$ optimization with applications, *Numer. Funct. Anal. Optim.* **14** (1993), 621–633.