Serdica

Mathematical Journal

Сердика

Математическо списание

Serdica Math. J. **21** (1995), 185-200

Serdica
Mathematical Journal

Institute of Mathematics
Bulgarian Academy of Sciences

SUFFICIENT CONDITIONS OF OPTIMALITY FOR CONTROL PROBLEMS GOVERNED BY VARIATIONAL INEQUALITIES

James Louis Ndoutoume*

Communicated by R. Lucchetti

Abstract. The author recently introduced a regularity assumption for derivatives of set-valued mappings, in order to obtain first order necessary conditions of optimality, in some generalized sense, for nondifferentiable control problems governed by variational inequalities. It was noticed that this regularity assumption can be viewed as a symmetry condition playing a role parallel to that of the well-known symmetry property of the Hessian of a function at a given point. In this paper, we elaborate this point in a more detailed way and discuss some related questions. The main issue of the paper is to show (using this symmetry condition) that necessary conditions of optimality alluded above can be shown to be also sufficient if a weak pseudo-convexity assumption is made for the subgradient operator governing the control equation. Some examples of application to concrete situations are presented involving obstacle problems.

The symmetry condition we suggest for the proto-derivative set-valued mappings has already been used in our earlier paper Ndoutoume [17] in order to obtain first order necessary conditions of optimality for control problems governed by variational inequalities. It plays a role parallel to that of the well-known symmetry property

1991 *Mathematics Subject Classification:* 49B99, 49A29

Key words: set-valued mapping, proto-derivative, subgradient operator, pseudo-convexity, closed convex process, optimality condition, variational inequality

*This work was completed while the author was visiting the University of Limoges. Support from the laboratoire "Analyse non-linéaire et Optimisation" is gratefully acknowledged.

of the Hessian of a function at a given point and does not seem to be well-known in the literature. This leads to our first concern in this paper (see Section 1). We show that the usual functional operations which preserve the proto-differentiability (cf. Do [14, 15], and Rockafellar [19, 20, 21] for example) also have the property of preserving the symmetry condition. Among these operations, are the conjugacy and the integration operations. The main issue of the paper is to show (using this symmetry condition) that necessary conditions of optimality alluded above can be shown to be also sufficient if a weak pseudo-convexity assumption (notion described in detail in Section 2) is made for the subgradient operator governing the control equation. We assume the reader is familiar with elementary definitions and techniques from set-valued analysis (cf. Aubin and Frankowska [7] and references therein). Throughout the following, X will be a real Hilbert space with scalar product denoted by $\langle\cdot,\cdot\rangle$ and associated norm by $\|\cdot\|$.

1. Symmetry condition for proto-derivative set-valued mappings. Given a set-valued mapping $\Gamma : X \rightrightarrows X$, a point $x \in X$ where $\Gamma(x)$ is nonempty and a point $z \in \Gamma(x)$, consider the difference quotient set-valued mappings

$$(\Delta_t\Gamma)_{x,y} := \frac{\Gamma(x+th)-z}{t} \quad \text{for all} \quad h \in X \quad (t>0). \tag{1.1}$$

If the graphs of $(\Delta_t\Gamma)_{x,z}$ as a family of subsets of $X \times X$ strongly converge as $t \downarrow 0$ to another subset of $X \times X$ in the Painlevé-Kuratowski sense (cf. Aubin and Frankowska [7] for the precise definition), then Γ is said to be proto-differentiable at x relative to z, and the limit set is the graph of another set-valued mapping $\Gamma'_{x,z} : X \rightrightarrows X$ which will be called the proto-derivative of Γ at x relative to z. For more information about this concept and its relationship with other forms of differentiation, the reader is referred to Rockafellar [19, 20, 21]. We simply recall here that the graph of $\Gamma'_{x,z}$ coincides at the same time with the contingent and the intermediate cone to the graph of Γ at (x,z) (cf. Aubin and Frankowska [7] for the precise definitions). Set-valued mappings whose graphs are cones are positively homogeneous. They are called "processes". Hence, proto-derivative set-valued mappings are "closed processes". Set-valued mappings whose graphs are closed convex cones are called "closed convex processes". Closed processes, as continuous linear mappings, can be transposed. Indeed, given a closed process $\Gamma : X \rightrightarrows X$, its transpose $[\Gamma]^*$ is the closed convex process from X to X defined by:

$$p \in [\Gamma]^*(q) \iff \forall x \in X, \forall y \in \Gamma(x), \langle p,x\rangle \le \langle q,y\rangle. \tag{1.2}$$

The graph of $[\Gamma]^*$ is related to the polar cone of the graph of Γ (denoted by $(\operatorname{graph}\Gamma)^-$) in the following way:

$$(1.3) \qquad (p,q) \in \operatorname{graph}[\Gamma]^* \iff (p,-q) \in (\operatorname{graph}\Gamma)^-.$$

We shall say that a proto-derivative set-valued mapping $\Gamma'_{x,z} : X \rightrightarrows X$ satisfies the symmetry condition if and only if

$$(1.4) \qquad \operatorname{graph}\Gamma'_{x,z} \subset \operatorname{graph}[\Gamma'_{x,z}]^*.$$

It is not difficult to see that the latter means

$$(1.5) \qquad (u,v),(w,p) \in \operatorname{graph}\Gamma'_{x,z} \Longrightarrow \langle u,p\rangle = \langle w,v\rangle.$$

In order to clarify the meaning of the symmetry condition (1.4) when Γ happens to be a subgradient operator (i.e. the subdifferential of a lower semicontinuous proper convex function), some results have been laid out in Ndoutoume [17] along with many elementary examples. We refer the reader to that paper and simply recall here the following elementary fact: for a function F that is twice Fréchet differentiable at a point $x \in X$, the proto-derivative of ∇F (the gradient of F) at x relative to $z := \nabla F(x)$, exists and coincides with $\nabla^2 F(x)$ (the Hessian of F at x). In this case, the symmetry condition (1.5) can be rewritten as follows

$$(1.6) \qquad \langle \nabla^2 F(x)u, v\rangle = \langle u, \nabla^2 F(x)v\rangle \quad \text{for all} \quad u,v \in X.$$

This means that the linear mapping $\nabla^2 F(x)$ is self-adjoint. We are all familiar with this property of the second-order Fréchet derivative. However, in a nondifferentiable setting, where $(\partial F)'_{x,z}$ can be considered as a substitute for $\nabla^2 F(x)$, the proto-derivative of a subgradient operator does not always satisfy the symmetry condition described above. The reader will note that the proto-derivative of a subgradient operator satisfies the symmetry condition whenever it is a closed convex process (cf. Ndoutoume [17], Theorem 3.11). Subgradient operators of many common types of convex functions are known to be proto-differentiable such that the corresponding proto-derivative set-valued mapping satisfy the symmetry condition. We refer the reader to Ndoutoume [17] for the precise nature of which.

Now, we are going to show that the symmetry condition is preserved under integration operator. To this end, we consider the following situation: let $(\Omega, A, d\omega)$ be a σ-finite measure space and $L^2(\Omega)$ be the Hilbert space of square-integrable measurable functions from Ω to $\mathbb{R}^n$. Our aim is to show that the symmetry condition on the proto-derivative of the subdifferential of a convex integrand $F : \Omega \times \mathbb{R}^n \to]-\infty, +\infty]$, carries over that on the proto-derivative of the subdifferential of the corresponding integral functional $I_F(x) := \int_\Omega F(\omega, x(\omega))d\omega$ on $x \in L^2(\Omega)$. We recall that when I_F is proper, an element z of $L^2(\Omega)$ belongs to $\partial I_F(x)$ if and only if $z(\omega)$ belongs to $\partial F_\omega(x(\omega))$ for a.e. $\omega \in \Omega$, where $F_\omega(\cdot) := F(\omega, \cdot)$ (cf. Rockafellar [18] for details). The connection between the proto-differentiability of ∂I_F and that of ∂F_ω has been studied by C. N. Do in [15]. More precisely, it has been proved that if for a pair $(x, z) \in \operatorname{graph} \partial I_F$, ∂F_ω is proto-differentiable at $x(\omega)$ relative to $z(\omega)$ for a.e. $\omega \in \Omega$, then ∂I_F is proto-differentiable at x relative to z, furthermore an element v of $L^2(\Omega)$ belongs to $(\partial I_F)'_{x,z}(u)$ if and only if $v(\omega)$ belongs to $(\partial F_\omega)'_{x(\omega),z(\omega)}(u(\omega))$ for a.e. $\omega \in \Omega$ (cf. Do [15], Theorem 5.5). Under these circumstances, it is easy to see using statement (1.5) that if for a.e. $\omega \in \Omega$, $(\partial F_\omega)'_{x(\omega),z(\omega)}$ satisfies the symmetry condition, then $(\partial I_F)'_{x,z}$ automatically also satisfies the symmetry condition.

Duality is addressed next. To this end, we recall that if $F^* : X \to]-\infty, +\infty]$ is the Legendre-Fenchel conjugate of the lower semicontinuous proper convex function $F : X \to]-\infty, +\infty]$, then the subgradient operator ∂F^* is the inverse of ∂F (cf.Rockafellar [18]). Otherwise, for $(x, z) \in \operatorname{graph} \partial F$, the proto-differentiability of ∂F at x relative to z is equivalent to that of ∂F^* at z relative to x. Moreover $(\partial F^*)'_{z,x}$ is the inverse of $(\partial F)'_{x,z}$ (cf. Rockafellar [19], Theorem 2.4). A natural question in the context of the present section is to ask whether the symmetry condition introduced above satisfies an equally strong link between the function F and its conjugate F^*. The answer to this question is simple as shown by the following proposition.

Proposition 1.1. *Let X be a real Hilbert space, $F : X \to]-\infty, +\infty]$ a lower semicontinuous proper convex function, and $(x, z) \in \operatorname{graph} \partial F$. Assume that ∂F is proto-differentiable at x relative to z. Then, the following statements are equivalent*

(i) $\operatorname{graph} (\partial F)'_{x,z} \subset \operatorname{graph} [(\partial F)'_{x,z}]^*$;

(ii) $\operatorname{graph} (\partial F^*)'_{z,x} \subset \operatorname{graph} [(\partial F^*)'_{z,x}]^*$.

P r o o f. It is a simple application of the definition of the symmetry condition. $\square$

Now, consider $\varphi : X \to]-\infty, +\infty]$ a lower semicontinuous proper convex function constructed from F in either the following ways:

(a) $\varphi = F + g$, where g is a convex function of class C^2 on X;

(b) $\varphi = F \circ g$ (infinimal convolution), where g is the conjugate of a convex function h of class C^2 on X.

For a given pair $(x,z) \in \operatorname{graph} \partial F$, the proto-differentiability of ∂F at x relative to z is equivalent to that of $\partial\varphi$ at w relative to p, where $(w,p) := (x, z + \nabla g(x))$ in the case of (a) and $(w,p) := (x + \nabla h(z), z)$ in the case of (b) (cf. Rockafellar [19] Theorem 3.7). With these notations, we observe the following

Theorem 1.1. *Assume that ∂F is proto-differentiable at x relative to z. Then the following statements are equivalent*

(i) $\operatorname{graph} (\partial F)'_{x,z} \subset \operatorname{graph} [(\partial F)'_{x,z}]^*$;

(ii) $\operatorname{graph} (\partial \varphi)'_{w,p} \subset \operatorname{graph} [(\partial \varphi)'_{w,p}]^*$.

Proof. When $\varphi = F + g$, the equivalence between (i) and (ii) comes from the equality: $(\partial\varphi)'_{w,p} = (\partial F)'_{x,z} + \nabla^2 g(x)$, and the fact that the symmetry condition is preserved by addition. On the other hand, when $\varphi = F \circ g$, the equivalence between (i) and (ii) comes from Proposition 1.1 by duality with (a): indeed, one has $F \circ g = (F^* + h^*)^*$, where h is a convex function of class C^2. This ends the proof. □

Corollary 1.1. *Let $F : X \to]-\infty, +\infty]$ be a lower semicontinuous proper convex function and let F_λ (for $\lambda > O$) be the Moreau-Yosida approximate of index λ of F, defined by: $F_\lambda(y) := \min \left\{ F(w) + \frac{\|w-y\|^2}{2\lambda} \ : \ w \in X \right\}$ for all $y \in X$. For a given point $x \in X$, set $J^{F_\lambda} := (I + \lambda\partial F)^{-1}$ and assume that ∂F is proto-differentiable at $u := J^{F_\lambda}(x)$ relative to $z = \frac{x-u}{\lambda}$. Then the following statements are equivalent*

(i) $\operatorname{graph} (\partial F)'_{u,z} \subset \operatorname{graph} [(\partial F)'_{u,z}]^*$;

(ii) $\operatorname{graph} (\nabla F_\lambda)'_{x,z} \subset \operatorname{graph} [(\nabla F_\lambda)'_{x,z}]^*$;

(iii) $\operatorname{graph} (J^{F_\lambda})'_{x,u} \subset \operatorname{graph} [(J^{F_\lambda})'_{x,u}]^*$.

Proof. Using Proposition 1.1, the equivalence between (i) and (ii) comes from the fact that $F_\lambda = F \circ g$, where $g(x) = \frac{\|x\|^2}{2\lambda}$ and $g^*(p) = \frac{\lambda\|p\|^2}{2}$. The equivalence between (ii) and (iii) follows from the fact that p belongs to $(\partial F)'_{u,z}(q)$ if and only if q belongs to $(J^{F_\lambda})'_{x,u}(p+q)$ (cf. Aubin and Frankowska [7], Proposition 5.2.7). □

2. Pseudo-convexity assumption for set-valued mappings. We begin with the following definition which can be found in Aubin and Frankowska [7] (see also Aubin [5] and Aubin and Ekeland [6] for additional descriptions).

Definition 2.1. *Consider a set-valued mapping* $\Gamma : X \rightrightarrows X$*, a point* $x \in X$ *where* $\Gamma(x)$ *is nonempty and a point* $z \in \Gamma(x)$*. We shall say that* Γ *is pseudo-convex at* (x, z) *if and only if*

$$\forall q \in \mathrm{Dom}\,\Gamma,\ \Gamma(q) \subset \mathrm{D}\,\Gamma(x,z)(q-x) + z. \tag{2.1}$$

Here, $\mathrm{Dom}\,\Gamma := \{q \in X : \Gamma(q) \neq \emptyset\}$ and $\mathrm{D}\,\Gamma(x,z)$ stands for the set-valued mapping from X into X whose graph is the contingent cone (cf. Aubin and Frankowska [7] for the precise definition) to the graph of Γ at (x, z). When $\Gamma'_{x,z}$ exists, (2.1) can be rewritten as follows:

$$\forall q \in \mathrm{Dom}\,\Gamma,\ \Gamma(q) \subset \Gamma'_{x,z}(q-x) + z. \tag{2.2}$$

It is not difficult to see that the existence of $\Gamma'_{x,z}$ does not imply that Γ is pseudo-convex at (x, z). It has been established in Aubin [5], Lemma 2.1 that when Γ is starshaped around (x, z) in the following sense:

$$\forall (p,q) \in \mathrm{graph}\,\Gamma,\ \forall t \in [0,1]\ :\ z + t(p-z) \in \Gamma(x + t(q-x)), \tag{2.3}$$

then Γ is pseudo-convex at (x, z), furthermore $\Gamma'_{x,z}$ exists. A number of elementary properties of pseudo-convex set-valued mappings have been furnished in Aubin [5] along with further justification of the concept. Some classes of pseudo-convex set-valued mappings have been identified in Aubin [5], Aubin and Ekeland [6], and Aubin and Frankowska [7]. In the sequel, we deal with subgradient operators. Unfortunately, the assumption that a subgradient operator can be pseudo-convex, even in the simplest cases, is not always feasible. As an illustration of this fact, we present the following simple example

Example 2.1. For $X = \mathbb{R}$ and $F = |\cdot|$, set $x = 0$ and $z = 1$. In this case, it is easy to see that $(\partial F)'_{0,1}$ exists and its graph is given by:

$$\mathrm{graph}\,(\partial F)'_{0,1} = \{0\} \times]-\infty, 0] \cup]0, +\infty[\times \{0\}. \tag{2.4}$$

The latter allows to assert that ∂F is not pseudo-convex at $(0, 1)$. This example shows that the concept of pseudo-convexity such that it has been defined above, is not adapted to subgradient operators. his leads us to introduce the following weak pseudo-convexity concept, which extends the main ideas inherent in the concept of pseudo-convexity described above (cf. Definition 2.1) to situations where statement (2.1) may not hold true.

Definition 2.2. *Consider a set-valued mapping $\Gamma : X \rightrightarrows X$, a point $x \in X$ where $\Gamma(x)$ is nonempty and a point $z \in \Gamma(x)$. We shall say that Γ is weakly pseudo-convex at (x, z) relatively to an element p of X if and only if one has:*

$$\forall q \in \operatorname{Dom}\Gamma,\ \Gamma(q) \subset \mathrm{D}\,\Gamma(x,z)(q-x) + z + C(p) \tag{2.5}$$

where $C(p)$ stands for the closed convex cone defined by:

$$C(p) := \{z \in X \ :\ \langle z, p\rangle \geq 0\}.$$

It is readily seen that the pseudo-convexity assumption on Γ at (x, z), in the sense of Aubin and Frankowska (cf. Definition 2.1), amounts exactly to say that Γ is weakly pseudo-convex at (x, z) relatively to any $p \in X$. It suffices to take in this case $C(p) := \{0\}$. In order to clarify the meaning of the weak pseudo-convexity concept introduced in Definition 2.2, we present the following elementary examples.

Example 2.2. Let $F = \|\cdot\|$ be the Euclidean norm in $X = \mathbb{R}^n$. For $x = 0$, it is well-known that $\partial F(0)$ coincides with B (the closed unit ball of $\mathbb{R}^n$). For $z \in \mathbb{R}^n$ such that $\|z\| < 1$, $(\partial F)'_{0,z}$ exists and its graph is defined by: $\operatorname{graph}(\partial F)'_{0,z} = \{0\} \times \mathbb{R}^n$. On the other hand, for $z \in \mathbb{R}^n$ such that $\|z\| = 1$, $(\partial F)'_{0,z}$ exists and is defined by:

$$(\partial F)'_{0,z}(u) = N_{\mathbb{R}(z)}(u), \quad \text{for all } u \in \operatorname{Dom}(\partial F)'_{0,z} := \mathbb{R}_+(z)$$

where $N_{\mathbb{R}(z)}$ stands for the normal cone to $\mathbb{R}_+(z) := \{tz : t \geq 0\}$. Then, it is easy to see that for any $z \in \mathbb{R}^n$ such that $\|z\| \leq 1$, ∂F is not pseudo-convex at $(0, z)$ (in the sense of Definition 2.1). Nevertheless, ∂F is weakly pseudo-convex at $(0, z)$ relatively to any $p \in -N_B(z)$ (where $N_B(z)$ stands for the normal cone to B at z).

Example 2.3. Let Ω be a bounded and open subset of $\mathbb{R}^n$ with a (sufficiently) smooth boundary T. Let $g : \mathbb{R} \to]-\infty, +\infty]$ be a lower semicontinuous proper convex function, and set $\beta := \partial g$. Define the function $F : L^2(\Omega) \to]-\infty, +\infty]$ by:

$$F(y) = \begin{cases} 2^{-1}\int_\Omega |\nabla y|^2 + \int_T g(y)ds & \text{if } y \in H^1(\Omega),\ g(y) \in L^1(T) \\ +\infty & \text{otherwise.} \end{cases}$$

It is well known (cf. Barbu [8]) that F is a lower semicontinuous proper convex function. Moreover, its subdifferential $\partial F : L^2(\Omega) \to L^2(\Omega)$ is given by:

$$\partial F(y) = -\Delta y \quad \text{for all} \quad y \in \operatorname{Dom} \partial F$$

$$\operatorname{Dom} \partial F := \left\{ y \in H^2(\Omega) | -\frac{\partial y}{\partial s} \in \beta(y) \quad \text{a.e. in } T \right\}.$$

The contingent derivative of ∂F at (x, z) (where $z \in \partial F(x)$) is given by:

$$\mathrm{D}\,\partial F(x,z)(u) = -\Delta u \quad \text{for all} \quad u \in \operatorname{Dom} \mathrm{D}\,\partial F(x,z)$$

$$\operatorname{Dom} \mathrm{D}\,\partial F(x,z) = \left\{ u \in H^2 | -\frac{\partial y}{\partial s} \in \mathrm{D}\,\beta(x, -\frac{\partial x}{\partial s})(u) \text{ a.e. in } T \right\}$$

Then it is easy to see that ∂F is weakly pseudo-convex at (x, z) relatively to any $p \in N_{\operatorname{Im}\Delta}(\Delta x)$ (where $\operatorname{Im}\Delta := \{\Delta u \in L^2(\Omega); u \in L^2(\Omega)\}$ and $N_{\operatorname{Im}\Delta}(\Delta x)$ stands for the normal cone to $\operatorname{Im}\Delta$ at Δx).

Remark 2.1. Coming back to Definition 2.2, if in relation (2.5), $C(p)$ is replaced by any closed convex cone C such that the space X is partially ordered by C, then the set valued-mapping Γ is said to be invex at (x, z). For instance when $X = \mathbb{R}^n$, C may be the nonnegative orthant of $\mathbb{R}^n$. Then the concept of invexity for set-valued mappings is comparable with that of weak pseudo-convexity described in Definition 2.2, but neither contains it nor is contained in it. There have been significant contributions strongly relied on the concept of invexity. The works of B. D. Craven and P. H. Sach [11, 12], D. T. Luc and C. Malivert [16], A. B. Israel and B. Mond [9], and T. W. Reiland [22], among others, must be considered.

3. Optimal control in some variational inequalities. Throughout this section V and H are real Hilbert spaces such that V is dense in H. The latter is identified with its own dual and is then identified with a subspace of the dual V' of V.

3.1. Analytical background. In our earlier paper Ndoutoume [17], we gave conditions guaranteeing the existence of a dual extremal element (in some generalized sense) for nondifferentiable control problems which can be set in the following simplified form

(P) : $$\text{minimize } J(y, u)$$

over all pairs (y, u) subject to the state system (variational inequality)

$$Ay + \partial F(y) \ni Bu + f. \tag{3.1}$$

Here, A is a linear operator from V into V', ∂F is a subgradient operator from H into H, B is a linear operator from the space of controls U (that is also a real Hilbert space) to the state space H, and f is a given element of H. The cost (or objective) function $J : V \times U \to]-\infty, +\infty[$ is assumed to be locally Lipschitz with respect to the variable y. More precisely, the following result can be found in Ndoutoume [17]:

Theorem 3.1 (Necessary conditions). *Let (y, u) be any optimal solution in problem* (P) *and set $r := Bu - Ay + f$. Assume that $(\partial F)'_{y,r}$ exists and satisfies the symmetry condition. Then there exists an unique element p satisfying along with y and u the following system:*

$$-A^*p \in (\partial F)'_{y,r}(p) + \partial_1 J(y, u); \tag{3.2}$$

$$B^*p \in \partial_2 J(y, u) \tag{3.3}$$

Here $\partial_1 J$ and $\partial_2 J$ refer to the partial contingent subdifferential of J in respect to y and to u. We may view the unique element p satisfying along with y and u the optimality system composed of (3.2) and (3.3) as a dual extremal element of problem (P). In this context, (3.2) and (3.3) may be seen as "generalized" first-order necessary conditions of optimality. For more information about this result and its relationship with some other results relative to necessary conditions of optimality for control problems of type (P), the reader is referred to Ndoutoume [17] (see also Barbu [8] and references therein).

Our main purpose here is to look for conditions under which the existence of an element p (satisfying along with y and u the optimality system composed of (3.2) and (3.3)) implies that the pair (y, u) is an optimal solution for (P). To this end, let us first observe through the following simple example that necessary conditions of optimality mentioned above are not always sufficient.

Example 3.1. For $U = X = \mathbb{R}$, $A = 1$, $F = 0$, $B = 1$, $f = 0$, $h(u) = |u|$ and $g(y) = -y^2$, with $J(y, u) = h(u) + g(y)$, problem (P) can be rewritten as follows

$$\text{minimize } -u^2 + |u| \text{ over all } u \in \mathbb{R}.$$

It is easy to see that this problem has no solutions while for $p = y = u = 0$, conditions (3.2) and (3.3) are satisfied. The reader will note that the function g appearing in Example 3.1 is not convex. This could explain why conditions (3.2) and (3.3) in this example are not sufficient. In reality, even when the cost function J appearing in (P) is convex, necessary conditions mentioned above are not always sufficient.

As an illustration of this fact, we present the following elementary example.

Example 3.2. For $U = X = \mathbb{R}$, $A = 0$, $B = 1$, $f = 0$, $h(u) = e^u$, $F(y) = |y|$, and $g(y) = y^2$, with $J(y,u) := g(y) + h(u)$, problem (P) can be rewritten as follows:

$$\text{minimize } y^2 + e^u$$

over all pairs (y,u) such that $u \in \partial F(y)$.

For $p = e$, $y = 0$ and $u = 1$, (3.2) and (3.3) are satisfied. However, the pair $(0,1)$ is not an optimal pair for (P).

It is proved in the next paragraph that conditions (3.2) and (3.3) given above are also sufficient for a pair (y,u) to be optimal whenever the symmetry condition is fulfilled by $(\partial F)'_{y,r}$ along with the weak pseudo-convexity assumption on ∂F.

3.2. Main theorems. Throughout the following, (y,u) is taken to be a fixed pair satisfying state system (3.1), and the cost function J is assumed to be lower semicontinuous proper and convex. In order to simplify our discussion, we first assume that

$$J(y,u) = g(y) + h(u) \tag{3.4}$$

and then, we shall generalized the result obtained from (3.4) to the general case.

Theorem 3.2. *Set $r := Bu - Ay + f$ and assume that*

(i) *$(\partial F)'_{y,r}$ exists and satisfies the symmetry condition*;

(ii) *∂F is weakly pseudo-convex at (y,r) relatively to an element p of X satisfying along with y and u the following system:*

$$-A^*p \in (\partial F)'_{y,r}(p) + \partial g(y); \tag{3.5}$$

$$B^*p \in \partial h(u). \tag{3.6}$$

Then the pair (y,u) is an optimal solution for (P). *In other words, conditions* (3.5) *and* (3.6) *are necessary and sufficient for a pair (y,u) to be optimal.*

P r o o f. Consider (z,v) satisfying state system (3.1) and $\eta \in \partial g(y)$ such that

$$-A^*p - \eta \in (\partial F)'_{y,r}(p). \tag{3.7}$$

Let us first show, using the symmetry condition on $(\partial F)'_{y,r}$ and the weak pseudo-convexity assumption on ∂F, the following equality:

$$\langle c, p\rangle - \langle \eta, z - y\rangle = \langle v - u, B^* p\rangle \tag{3.8}$$

for some $c \in C(p)$. In fact, since $-Az + Bv + f \in \partial F(z)$, we get from the weakly pseudo-convexity assumption on ∂F at (y, r) relatively to p, that there exists an element $c \in C(p)$ such that

$$A(y - z) + B(v - u) - c \in (\partial F)'_{y,r}(z - y). \tag{3.9}$$

Using the symmetry condition on $(\partial F)'_{y,r}$, we get (3.8) from (3.7) and (3.9). Then, since $B^* p \in \partial h(u)$, it follows from (3.8) that

$$h(v) \geq h(u) + \langle c, p\rangle - \langle \eta, z - y\rangle \geq h(u) - \langle \eta, z - y\rangle. \tag{3.10}$$

That implies

$$g(z) + h(v) \geq g(z) + h(u) - \langle \eta, z - y\rangle. \tag{3.11}$$

On the other hand, since $\eta \in \partial g(y)$, we get from (3.11) that

$$g(z) + h(v) \geq g(y) + h(u). \tag{3.12}$$

The latter being true for any (z, v) subject to state system (3.1), we get that (y, u) is an optimal pair for (P). This ends the proof of Theorem 3.2. □

Theorem 3.3. *Set* $r := Bu - Ay + f$ *and assume that*

(i) $(\partial F)'_{y,r}$ *exists and satisfies the symmetry condition;*

(ii) ∂F *is weakly pseudo-convex at* (y, r) *relatively to an element* p *of* X *satisfying along with* y *and* u *the following condition*

$$(-A^* p, B^* p) \in (\partial F)'_{y,r}(p) \times \{0\} + \partial J(y, u). \tag{3.13}$$

Then the pair (y, u) *is an optimal solution for* (P).

P r o o f. Consider (z, v) satisfying state system (3.1) and $\sigma \in (\partial F)'_{y,r}(p)$ such that

$$(-A^* p - \sigma, B^* p) \in \partial J(y, u). \tag{3.14}$$

It follows from (3.14) that

$$J(z,v) - J(y,u) \geq \langle -A^*p - \sigma, z - y\rangle + \langle v - u, B^*p\rangle. \tag{3.15}$$

As for (3.8) in the proof of the preceding theorem, there exists $c \in C(p)$ such that

$$\langle c, p\rangle + \langle \sigma, z - y\rangle = \langle -A^*p, z - y\rangle + \langle v - u, B^*p\rangle. \tag{3.16}$$

From (3.15) and (3.16), using the definition of $C(p)$, we obtain

$$J(z,v) - J(y,u) \geq 0. \tag{3.17}$$

The latter being true for any (z,v) subject to the state system (3.1), we get that (y,u) is an optimal pair for (P). This ends the proof of Theorem 3.3. □

Remark 3.1. It is readily seen that condition (3.13) implies that

$$-A^*p \in (\partial F)'_{y,r}(p) + \partial_1 J(y,u); \tag{3.18}$$

$$B^*p \in \partial_2 J(y,u) \tag{3.19}$$

whenever

$$\partial J(y,u) \subset \partial_1 J(y,u) + \partial_2 J(y,u). \tag{3.20}$$

We refer the reader to Rockafellar [18] for sufficient conditions ensuring that (3.20) holds.

3.3. Applications to control problems governed by semilinear equations. The optimal control problem for a system governed by a semilinear equation is proposed by Berkovitz [9] and discussed in Barbu [8]. The formulation of this problem is as follows:

(D) : $$\text{minimize} \quad g(y) + h(u)$$

on all $y \in H_0^1(\Omega) \cap H_0^2(\Omega)$ and $u \in U$ subject to

$$\begin{aligned} &A_0 y + \beta y \ni f + Bu \quad \text{a.e.in} \quad \Omega \\ &y = 0 \text{ in } \Gamma. \end{aligned} \tag{3.21}$$

Here β is a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ satisfying the condition $0 \in \operatorname{Dom}\beta$, Ω is a bounded and open subset of $\mathbb{R}^n$ having a sufficiently smooth boundary Γ, f belongs to $L^2(\Omega)$ and B is a linear operator from the control space U to $L^2(\Omega)$. A_0 stands for the elliptic differential operator defined by

$$A_0 y = -\sum (a_{ij}(x) y_{x_i})_{x_j} + a_0(x) y \tag{3.22}$$

where $a_{ij} \in C^1(\Omega)$, $a_0 \in L^\infty(\Omega)$, $a_{ij} = a_{ji}$ for all $i, j = 1, \ldots, n$. Problem (D) can be written in the form of (P) where $V = H_0^1(\Omega)$, $H = L^2(\Omega)$, and $A : H_0^1(\Omega) \to H^{-1}(\Omega)$ is given by

$$\langle Ay, v\rangle = -\sum \int_\Omega (a_{ij}(x) y_{x_i}) v_{x_j} dx + \int_\Omega a_0(x) y v dx \text{ for all } y, v \in H_0^1(\Omega). \tag{3.23}$$

The function F in this case is given by

$$F(y) = \int_\Gamma j(y) dr, \quad y \in L^2(\Omega) \tag{3.24}$$

where $\beta = \partial j$. Here the functions $g : L^2(\Omega) \to]-\infty, +\infty[$ and $h : U \to]-\infty, +\infty[$ are assumed to satisfy:

(i) g is locally Lipschitz and convex;

(ii) h is lower semicontinuous and convex.

Under the above notations and assumptions, let us first consider the case where β has a convex graph. Under these circumstances, it is easy to see (from Theorems 3.1 and 3.2) that a pair (y, u) is an optimal solution for problem (D) if and only if there exists $p \in H_0^1(\Omega)$ which satisfies along with y and u the following system

$$A_0 y + \beta y \ni f + Bu \text{ a.e. in } \Omega \text{ and } y = 0 \text{ in } \Gamma \tag{3.25}$$

$$-A_0 p \in (\beta)'_{y,r}(p) + \partial g(y) \quad \text{a.e. in } \Omega \text{ and } p = 0 \text{ in } \Gamma \tag{3.26}$$

$$B^* p \in \partial h(u). \tag{3.27}$$

Consider now the case where the graph of β is given by

$$\operatorname{graph} \beta = \{0\} \times]-\infty, 0] \cup]0, +\infty[\times \{0\}. \tag{3.28}$$

Under these circumstances, system (3.21) reduces to the obstacle problem. Obstacle problems represent an important class of non linear problems and occur in the mathematical description of a large variety of physical problems (cf. Barbu [8] for example).

Thus, the control system (3.21) can be rewritten as follows:

$$(3.29)\qquad \begin{aligned} &(-A_0 y + bu + f)y = 0 \;\text{ a.e. in }\; \Omega \\ &y \geq 0, -A_0 y + Bu + f \leq 0 \;\text{ a.e. in }\; \Omega, \;\text{ and }\; y = 0 \;\text{ in }\; \Gamma. \end{aligned}$$

In the present setting, one has:

$$(3.30)\qquad \operatorname{graph} \beta'_{w,0} = \mathbb{R} \times \{0\} \;\text{ for }\; w > 0,$$

$$(3.31)\qquad \operatorname{graph} \beta'_{00} = \operatorname{graph} \beta,$$

$$(3.32)\qquad \operatorname{graph} \beta'_{0,s} = \{0\} \times \mathbb{R} \;\text{ for }\; s < 0.$$

It is easy to see that for any $(w, s) \in \{0\}\times]-\infty, 0[\cup]0, +\infty[\times\{0\}$, β is weakly pseudo-convex at (w, s) relatively to any $q \leq 0$. Then, when applying Theorem 3.2, we obtain that a pair $(y, u) \in H_0^1(\Omega) \cap H_0^2(\Omega) \times U$ (such that $(y, Bu - A_0 u + f) \neq (0, 0)$ a.e. in Ω) is an optimal solution for problem (D) whenever there exists $p \leq 0$ a.e. in Ω, belonging to $H_0^1(\Omega)$ and satisfying along with y and u the following system:

$$(3.33)\qquad (-A_0 y + Bu + f)y = 0 \;\text{ a.e. in }\; \Omega$$

$$(3.34)\qquad y \geq 0, -A_0 y + Bu + f \leq 0 \;\text{ a.e. in }\; \Omega, \;\text{ and }\; y = 0 \;\text{ in }\; \Gamma$$

$$(3.35)\qquad -A_0 p \in \beta'_{y,r}(p) + \partial g(y) \;\text{ a.e. in }\; \Omega$$

$$(3.36)\qquad B^* p \in \partial h(u).$$

Acknowledgements. I am greatly indebted to the anonymous referees for suggestions that have done much to bring this paper to its present form.

REFERENCES

[1] H. Attouch. Variational convergence for functions and operators. Pitman Research Notes in Mathematics, London, 1984.

[2] H. Attouch, J.-L. Ndoutoume and M. Théra. Epigraphical convergence of functions and convergence of their derivatives in Banach spaces. Séminaire d'Analyse convexe de Montpellier, Exposé No 9, 1990, 9.1-9.45.

[3] J. P. Aubin. Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusion and differential inclusions. Advances in Mathematics, Supplementary Studies (Ed. L. Nachbin), 1981, 180-232.

[4] J. P. Aubin. Lipschitz behavior of solutions to convex minimization problems. *Math. Oper. Res.*, **8** (1983), 87-111.

[5] J. P. Aubin. Differential calculus of set-valued maps. IIASA wp, 1986, 87-93.

[6] J. P. Aubin and I. Ekeland. Applied Nonlinear Analysis. Wiley-Interscience, 1984.

[7] J. P. Aubin and H. Frankowska. Set-valued Analysis. Birkhauser, 1990.

[8] V. Barbu. Optimal control of variational inequalities. Pitman Advanced Publishing Program, 1984.

[9] A. Ben-Israel and B. Mond. What is invexity. *J. Austral. Math. Soc. Ser. B*, **28** (1986), 1-9.

[10] L. D. Berkovitz. Optimal control theory. Spinger, 1974.

[11] B. D. Craven and P. H. Sach. Invexity in multifunction optimization. *Numer. Funct. Anal. Optim.*, **12** (1991), 383-394.

[12] B. D. Craven and P. H. Sach. Invex multifunctions and duality. *Numer. Funct. Anal. Optim.*, **12** (1991), 575-591.

[13] B. D. Craven, P. H. Sach, N. D. Yen and T. D. Phuong. A new class of invex multifunctions. Melbourne University, Dept. of Math., Preprint No 21, 1991.

[14] C. N. Do. Second-order nonsmooth analysis and sensivity in optimization problems involving convex integral functionals. Thesis, Dept of Mathematics, Univ. of Washington, 1989.

[15] C.N. Do. Generalized second derivatives of convex functions in reflexive Banach spaces. *Trans. Amer. Math. Soc.*, to appear.

[16] D.T. Luc and C. Malivert. Conditions d'optimalité en optimisation invexe. *Bull. Austral. Math. Soc.*, **46** (1991), 47-66.

[17] J. L. Ndoutoume. Optimality conditions for control problems governed by variational inequalities. *Math. Oper. Res.*, to appear.

[18] R. T. Rockafellar. Convex Analysis. Princeton Univ. Press., Princeton, NJ, 1970.

[19] R. T. Rockafellar. Generalized second derivatives of convex functions and saddle function. *Trans. Amer. Math. Soc.*, **320** (1990), 810-822.

[20] R. T. Rockafellar. First and second-order epi-differentiability in nonlinear programming. *Trans. Amer. Math. Soc.*, **307**, (1988), 75-108.

[21] R. T. Rockafellar. Proto-differentiability of set-valued mappings and its applications in optimization. In Analyse non linéaire, (eds. H. Attouch et al), Gauthier-Villars, Paris, 1989, 449-482.

[22] T. W. Reiland. Nonsmooth invexity. *Bull. Austral. Math. Soc.*, **42** (1990), 437-446.

Institut Africain d'Informatique
B.P. 2263 Libreville
Gabon
Fax: (241) 720 011

Received July 12, 1994
Revised January 25, 1995