Control and Cybernetics

vol. **23** (1994) No. 4

The maximum principle and the convexification of optimal control problems

by

Uldis Raitums

Institute of Mathematics and Computer Science
University of Latvia
Boulevard Rainis 29
LV–1459 Riga
Latvia

We study the relation between the relaxability of optimal control problems and the type of necessary optimality conditions. For a given optimal control problem P we show the existence of a special family $\mathcal{M}(P)$ of analogous problems with the following property: if solutions of problems for $\mathcal{M}(P)$ satisfy optimality conditions of the form of the integral maximum principle, then the problem P is relaxable via convexificaton. After that, we describe some classes of optimal control problems for elliptic systems or parabolic equations which are not relaxable and whose solutions do not satisfy, in general, optimality conditions of the form of the integral maximum principle.

1. Introduction

The convexification (the passage to the convex hull of the set of admissible operators and functionals) of optimal control problems was very successfully employed to derive necessary conditions for optimality, see, for instance, Gamkrelidze (1975) and Warga (1972) for the case of ordinary differential equations and Raitums (1989) for the case of elliptic equation. Crucial for this approach was the relaxability of these problems, i.e., that the price of the original problem is equal to the price of the convexified one.

On the other hand, one can easily see that the classes of optimal control problems for which optimality conditions of the form of the integral maximum principle hold and the classes of relaxable problems often coincide. So, we can suppose that there is a strong relation between the type of optimality conditions and the relaxability of optimal control problems. Following this guideline, in this paper, we consider a specific family $\mathcal{M}(P)$ of optimal control problems, associated with the original problem P, and we show that if solutions of problems from $\mathcal{M}(P)$ satisfy the integral maximum principle then the original problem

P is relaxable. This result enables us to derive the "inverse" statement: if a class $\mathcal{M}$ of optimal control problems contains a typical problem which is not relaxable, then the integral maximum principle is not valid as an optimality condition for problems of this class $\mathcal{M}$.

After that, on the basis of concrete examples, we describe some classes of optimal control problems for elliptic or parabolic systems which are not, in general, relaxable and, therefore, the integral maximum principle can not be assumed as an optimality condition for problems of these classes.

2. An abstract case

Let W be a real Banach space with elements u, v, w and let W^* be the dual space and $\ll \cdot, \cdot \gg$ be the pairing between W and W^*. Let $\mathcal{A}$ be a set of pairs (A, I) of operators $A : W \to W^*$ and functionals $I : W \to R^1$. With $\mathcal{A}$ we associate the optimal problem $P\mathcal{A}$:

$$I(u) \to \min,$$

$$Au = 0, \quad (A, I) \in \mathcal{A}, \quad u \in W.$$

In typical cases of optimal control problems the set $\mathcal{A}$ is given by

$$\mathcal{A} = \{(A(\sigma)(\cdot), I(\sigma)(\cdot)) : \sigma \in \mathcal{S}\} \tag{1}$$

where $\mathcal{S}$ is the set of admissible controls.

Let $\mathcal{Z}$ be a set of functionals ℓ defined on $\mathcal{A}$. We do not specify here the representation of functionals ℓ but for typical cases of the set $\mathcal{A}$ given by (1) the functionals ℓ can be defined directly on the set $\mathcal{S}$. We always suppose that the trivial functional $\ell \equiv 0$ belongs to $\mathcal{Z}$.

Definition 2.1 *By $co\mathcal{A}$ we will denote the convex hull of the set $\mathcal{A}$, i.e. the set of all convex combinations of pairs $(A, I) \in \mathcal{A}$.*

Analogously, as for the set $\mathcal{A}$, we define for the set $co\mathcal{A}$ the optimal control problem $Pco\mathcal{A}$:

$$I(u) \to \min,$$

$$Au = 0, \quad (A, I) \in co\mathcal{A}, \quad u \in W.$$

Definition 2.2 *The price of the problem $P\mathcal{A}$ is the value*

$$\inf\{I(u) : (A, I) \in \mathcal{A},\ u \in W,\ Au = 0\}. \tag{2}$$

Analogously is defined the price of the problem $Pco\mathcal{A}$.

DEFINITION 2.3 *The convexification of the problem $P\mathcal{A}$ is the passage from the problem $P\mathcal{A}$ to the problem $Pco\mathcal{A}$. The problem $P\mathcal{A}$ is relaxable if the price of the problem $Pco\mathcal{A}$ is equal to the price of the problem $P\mathcal{A}$.*

In what follows we suppose that the following assumptions hold.

1°. Operators and functionals $(A, I) \in \mathcal{A}$ are defined on W, continuous in W and have Gateaux derivatives $A'(u)$ and $I'(u)$ respectively.

2°. For every pair $(A, I) \in co\mathcal{A}$ the equation

$$Au = 0 \tag{3}$$

with respect to $u \in W$ is uniquely solvable. This solution will be denoted by $u = u(A)$.

3°. For every pair $(A, I) \in co\mathcal{A}$ the variational equality

$$\ll A'(u(A))v, w \gg - \ll I'(u(A)), v \gg = 0 \quad \forall v \in W \tag{4}$$

with respect to $w \in W$ is uniquely solvable. This solution will be denoted by $w = w(A, I)$.

For an arbitrary fixed pair $(A_0, I_0) \in co\mathcal{A}$, a functional $\ell \in \mathcal{Z}$ and a number $\lambda \in [0, 1]$ we define the auxiliary functional

$$J(A_0, I_0, \ell, \lambda) : \mathcal{A} \to R^1,$$

$$J(A_0, I_0, \ell, \lambda)(A, I) \equiv I_0(u_\lambda) + \lambda[I(u_\lambda) - I_0(u_\lambda)] + \lambda\ell(A, I), \tag{5}$$

$$u_\lambda \equiv u(A_0 + \lambda(A - A_0)),$$

and the corresponding optimal problem $P\mathcal{A}(A_0, I_0, \ell, \lambda)$

$$I_0(u) + \lambda[I(u) - I_0(u)] + \lambda\ell(A, I) \to \min,$$

$$A_0 u + \lambda[Au - A_0 u] = 0, \quad (A, I) \in \mathcal{A}, \quad u \in W.$$

DEFINITION 2.4 *The problem $P\mathcal{A}(A_0, I_0, \ell, \lambda)$ satisfies the maximum principle if for every solution (A_*, I_*, u_*) of this problem the following relationship*

$$I_*(u_*) - \ll A_* u_*, w_* \gg + \ell(A_*, I_*) \leq$$

$$\leq I(u_*) - \ll A u_*, w_* \gg + \ell(A, I) \ \forall (A, I) \in \mathcal{A}, \tag{6}$$

$$w_* \equiv w(A_0 + \lambda(A_* - A_0), I_0 + \lambda(I_* - I_0)),$$

holds.

It is easy to see that for $\ell = 0, \lambda = 1$ the inequality (6) coincides with the standard integral maximum principle for the problem $P\mathcal{A}$.

Now, let us introduce the following additional assumptions.

4°. For every fixed pair $(A_0, I_0) \in co\mathcal{A}$, number $\lambda \in [0, 1]$ and number $\varepsilon > 0$ there exists a functional $\ell_\varepsilon \in \mathcal{Z}$ such that

(i) $|\ell_\varepsilon(A,I)| < \varepsilon, \quad \forall (A,I) \in \mathcal{A}$;

(ii) the problem $P\mathcal{A}(A_0, I_0, \ell_\varepsilon, \lambda)$ has a solution.

5°. For every fixed pair $(A_0, I_0) \in co\mathcal{A}$ and functional $\ell \in \mathcal{Z}$ and a chosen $(A,I) \in \mathcal{A}$ the function

$$\begin{aligned}
&\psi = \psi(\lambda), \lambda \in [0,1],\\
&\psi(\lambda) \equiv J(A_0, I_0, \ell, \lambda)(A,I) =\\
&= I_0(u_\lambda) + \lambda[I(u_\lambda) - I_0(u_\lambda)] + \lambda\ell(A,I),\\
&u_\lambda \equiv u(A_0 + \lambda(A - A_0)),
\end{aligned} \tag{7}$$

has the directional derivative

$$\psi'(\lambda) \equiv \lim_{\delta \to +0} \frac{1}{\delta}[\psi(\lambda+\delta) - \psi(\lambda)],$$

this derivative has the representation

$$\begin{aligned}
&\psi'(\lambda) = I(u_\lambda) - I_0(u_\lambda) - \ll Au_\lambda - A_0 u_\lambda, w_\lambda \gg + \ell(A,I),\\
&w_\lambda \equiv w(A_0 + \lambda(A - A_0), I_0 + \lambda(I - I_0)),
\end{aligned} \tag{8}$$

and the function ψ is continuous with respect to $\lambda \in [0,1]$ uniformly with restpect to $(A,I) \in \mathcal{A}$.

REMARK 2.1 The assumption 3° is fulfilled if the space W is reflexive and operators $A'(u(A)) : W \to W^*$ are invertible what often is supposed when concrete optimal control problems are investigated.

REMARK 2.2 The existence of a set $\mathcal{Z}$ for which the assumption 4° is fulfilled can be treated on the basis of results obtained by Ecland, Temam (1976) or Raitums (1976). For the set $\mathcal{A}$ given by (1) it is necessary only to demand a continuous dependence of solutions $u = u(A(\sigma))$ and functionals $I = I(\sigma, u(A(\sigma)))$ upon controls in suitable topologies. After that the set $\mathcal{Z}$ can be defined as a set of some special functionals $\ell : \mathcal{S} \to R^1$. Some results on continuous dependence of solutions of elliptic and parabolic equations upon coefficients of equations one can find, for instance in Raitums (1989, 1992) and Ladyzhenskaya, Solonnikov and Uraltseva (1967).

REMARK 2.3 The existence of the directional derivative $\psi'(\lambda)$ and the representation (8) demand, of course, the invertibility of operators

$$A_0'(u_\lambda) + \lambda[A'(u_\lambda) - A_0'(u_\lambda)] : W \to W^*.$$

If operators and functionals $(A,I) \in \mathcal{A}$ have continuous Frechet derivatives then the representation (8) can be easily obtained by means of the implicit function theorem and standard reasoning. If operators A have only Gateaux derivative then the results of Altman (1979) can be used.

THEOREM 2.1 *Let the assumptions $1° - 5°$ be satisfied.*
If for every fixed $(A_0, I_0) \in co\mathcal{A}, \lambda \in (0,1]$ and $\varepsilon > 0$ there exists a $\ell_\varepsilon \in \mathcal{Z}$ such that $|\ell_\varepsilon(A,I)| < \varepsilon$ for all $(A,I) \in \mathcal{A}$ and the problem $P\mathcal{A}(A_0, I_0, \ell_\varepsilon, \lambda)$ has a solution which satisfies the maximum principle then the price of the problem $Pco\mathcal{A}$ is equal to the price of the problem $P\mathcal{A}$.

PROOF. Let the pair $(A_0, I_0) \in co\mathcal{A}$ be fixed. We will discuss the properties of the function

$$h = h(\lambda),\ \lambda \in [0,1],$$

$$h(\lambda) \equiv \inf_{(A,I)\in\mathcal{A}} J(A_0, I_0, 0, \lambda)(A, I).$$

By the assumptions of the theorem for fixed $\lambda_0 \in (0,1)$ and for every $\varepsilon > 0$ there exists a functional $\ell_\varepsilon \in \mathcal{Z}$ such that $|\ell_\varepsilon(A,I)| < \varepsilon$ for all $(A,I) \in \mathcal{A}$ and the functional $J(A_0, I_0, \ell_\varepsilon, \lambda_0)$ attains its minimum. Let the corresponding solution be (A_*, I_*, u_*) and

$$h_\varepsilon(\lambda) \equiv \inf_{(A,I)\in\mathcal{A}} J(A_0, I_0, \ell_\varepsilon, \lambda)(A, I),\ \lambda \geq \lambda_0.$$

By virtue of assumption 5° we have that

$$\begin{aligned}
&h_\varepsilon(\lambda) - h_\varepsilon(\lambda_0) \leq J(A_0, I_0, \ell_\varepsilon, \lambda)(A_*, I_*)- \\
&-J(A_0, I_0, \ell_\varepsilon, \lambda_0)(A_*, I_*) = \\
&= [I_*(u_*) - I_0(u_*) - \ll A_* u_* - A_0 u_*, w_* \gg + \qquad (9)\\
&+\ell_\varepsilon(A_*, I_*)](\lambda - \lambda_0) + o(\lambda - \lambda_0), \\
&w_* \equiv w(A_0 + \lambda_0(A_* - A_0), I_0 + \lambda_0(I_* - I_0)).
\end{aligned}$$

Because the triple (A_*, I_*, u_*) as a solution of the problem $P\mathcal{A}(A_0, I_0, \ell_\varepsilon, \lambda_0)$ satisfies the maximum principle then

$$\begin{aligned}
&I_*(u_*) - I(u_*) - \ll A_* u_* - A u_*, w_* \gg \leq \\
&\leq \ell_\varepsilon(A, I) - \ell_\varepsilon(A_*, I_*) \leq 2\varepsilon, \qquad (10)\\
&\forall (A, I) \in \mathcal{A}.
\end{aligned}$$

From (9) and (10) it follows that

$$\begin{aligned}
&h_\varepsilon(\lambda) - h_\varepsilon(\lambda_0) \leq [I_*(u_*) - \\
&-I_0(u_*) - \ll A_* u_* - A_0 u_*, w_* \gg + \\
&+\ell_\varepsilon(A_*, I_*)](\lambda - \lambda_0) + o(\lambda - \lambda_0) \leq \qquad (11)\\
&\leq (2\varepsilon + \varepsilon)(\lambda - \lambda_0) + o(\lambda - \lambda_0) \leq 3\varepsilon(\lambda - \lambda_0) + o(\lambda - \lambda_0).
\end{aligned}$$

By construction, $|h_\varepsilon(\lambda) - h(\lambda)| < \varepsilon$, therefore, from (11) and arbitrarity of $\varepsilon > 0$ we have that

$$h(\lambda) - h(\lambda_0) \leq h_\varepsilon(\lambda) - h_\varepsilon(\lambda_0) + 2\varepsilon \leq$$

$$\leq 2\varepsilon + 3\varepsilon(\lambda - \lambda_0) + o(\lambda - \lambda_0) \leq o(\lambda - \lambda_0), \qquad \lambda \geq \lambda_0.$$

From assumption 5° it follows that the function h is continuous on $[0,1]$ and the last relationship is valid for all $0 < \lambda_0 < \lambda \leq 1$, hence,

$$h(\lambda) \leq h(\lambda_0) \text{ if } 0 < \lambda_0 \leq \lambda \leq 1.$$

But $h(1)$ is equal to the price of the problem $P\mathcal{A}$ and $h(0) = I_0(u(A_0))$. Because the pair $(A_0, I_0) \in co\mathcal{A}$ is arbitrary then we have the statement of the theorem. ■

COROLLARY 2.1 *Let the assumptions* 1°-5° *be satisfied. If the price of the problem* $Pco\mathcal{A}$ *is less then the price of the problem* $P\mathcal{A}$ *then for some* $(A_0, I_0) \in co\mathcal{A}, \lambda \in (0,1]$ *there exists an* $\ell \in \mathcal{Z}$ *such that the problem* $P\mathcal{A}(A_0, I_0, \ell, \lambda)$ *has a solution which does not satisfy the maximum principle.*

The meaning of this statement is as follows.

Let us suppose that some class of optimal control problems contains a characteristic problem $P\mathcal{A}_0$ for which the price of the problem $Pco\mathcal{A}_0$ is less than the price of the problem $P\mathcal{A}_0$. Then we cannot expect that the maximum principle will be valid, in general, for this class as a necessary condition of optimality. Besides, sometimes it is easier to construct an example where the convexification is not successful than to prove that the maximum principle is not true.

On the other hand, it is easy to see that if the convexification does not change the price of the problem then under some regularity conditions a necessary condition of optimality will be valid in the form of the maximum principle. Indeed, if the triple (A_*, I_*, u_*) is a solution of the original problem $P\mathcal{A}$ then it will be a solution of the problem $Pco\mathcal{A}$ too. But for the problem $Pco\mathcal{A}$ it is necessary to discuss only the directional derivatives

$$\frac{d}{d\lambda} J(A_*, I_*, 0, \lambda)(A, I)$$

which will be of the form (8).

3. Optimal control for elliptic and parabolic equations

In this section on the basis of concrete examples we will describe some classes of optimal control problems with distributed parameters for which the maximum principle is not valid, in general, as a necessary condition of optimality. More precisely, we will mainly discuss the case of elliptic systems. Results concerning parabolic equations will be some consequence of the elliptic case.

Let n, m, r be integers, let R^n, R^m, R^r be Euclidean spaces and let Ω be a bounded domain in R^n with Lipschitz boundary $\partial\Omega$ and points $x = (x_1, \cdots, x_n) \in \Omega$.

We introduce the spaces

$$L_q^{(r)}(\Omega) \equiv \underbrace{L_q(\Omega) \times \cdots \times L_q(\Omega)}_{r}, 1 \leq q \leq \infty, \tag{12}$$

$$W_m \equiv \underbrace{\overset{\circ}{W}{}_2^1(\Omega) \times \cdots \times \overset{\circ}{W}{}_2^1(\Omega)}_{m},$$

$$V \equiv L_2((0,1); \overset{\circ}{W}{}_2^1(\Omega)) \bigcap L_\infty((0,1); L_2(\Omega)),$$

where $L_q(\Omega), \overset{\circ}{W}{}_2^1(\Omega)$ are standard Lebesgue and Sobolev spaces $\overset{\circ}{W}{}_2^1(\Omega)$ consisting of functions which are equal to zero on the boundary $\partial\Omega$ in the sense of the embedding theorem. For an element $u = (u^1, \cdots, u^m) \in W_m$ we will use the notation $u_x = (u^1_{x_1}, \cdots, u^1_{x_n}, \cdots, u^m_{x_n})$.

Let

$$\Pi : \Omega \to 2^{R^r}$$

be a multivalued map with a countable dense set of measurable selections (see for instance Castaign, Valadier, 1997) such that all $\Pi(x)$ belong to a bounded set of R^r and let

$$\mathcal{S} \equiv \{\sigma = (\sigma^1, \cdots, \sigma^r) \in L_q^{(r)}(\Omega) : \sigma(x) \in \Pi(x), x \in \Omega\} \tag{13}$$

be the set of admissible controls.

With every $\sigma \in \mathcal{S}$ we associate an operator $\mathbf{A}(\sigma) : W_m \to (W_m)^*$ and a functional $I(\sigma) : W_m \to R^1$,

$$\mathbf{A}(\sigma)u := -divA(x, \sigma, u, u_x) + a(x, \sigma, u, u_x), \tag{14}$$

$$I(\sigma) := \int_\Omega g(x, \sigma, u, u_x)dx, \tag{15}$$

with some fixed matrix-function A and vector-function a whose elements are Caratheodory functions.

Our optimal control problem reads as follows

$$I(\sigma)u \to \min,$$

$$\sigma \in \mathcal{S},\ u \in W_m,\ \mathbf{A}(\sigma)u = 0, \tag{16}$$

where operators $\mathbf{A}(\sigma)$ and functionals $I(\sigma)$ are defined by (14) and (15) respectively.

We propose that for $m \geq n$ the maximum principle is not valid, in general, as a necessary condition of optimality for the problem (16). Moreover, if admissible controls do not depend on some spatial variable, for instance x_n then the result is the same for $m = n - 1$ too.

To prove this it is enough, according to the corollary 2.1., to construct examples of the sets $\mathcal{A}, \mathcal{Z}$ for which the convexification reduces the price of the problem $P\mathcal{A}$ and the auxiliary functionals $\ell \in \mathcal{Z}$ are in the form (15) too.

To began with, we recall the following results.

LEMMA 3.1 (RAITUMS, 1976) *Let Z_0 be a bounded closed subset of a real Hilbert space Z, let I be a lower semicontinuous and bounded below functional defined on Z_0. Let there exist a constant c_0 and a functional p defined on some linear manifold $Z_1 \subset Z$ with $Z_0 \subset Z_1$ such that*

$$p(z) \geq 0, \ \forall z \in Z_1,$$

$$p(z) \leq c_0, \ \forall z \in Z_0,$$

$$p(\lambda_1 z_1 + \lambda_2 z_2) \leq |\lambda_1| p(z_1) + |\lambda_2| p(z_2), \quad \forall \lambda_1, \lambda_2 \in R^1,$$

$$\forall z_1, z_2 \in Z_1.$$

Then for every fixed $\xi_0 \in Z$ and $\varepsilon > 0$ there exists an element $\xi \in Z$ such that
(i) $||\xi - \xi_0|| < \varepsilon$;
(ii) the functional $z \to I(z) + \varepsilon \ll \xi, z \gg$
attains its minimum on Z_0.

If, additionally, the value $p(\xi_0)$ is defined and finite then the element ξ can be chosen such that
(iii) $p(\xi - \xi_0) < \varepsilon$.
Here $\ll \cdot, \cdot \gg$ is the scalar product in Z.

Particularly, if $Z = L_2^{(r)}(\Omega)$ and $Z_0 = \mathcal{S}$ then the functional p can be chosen as

$$p(z) = ||z||_{L_\infty^{(r)}(\Omega)}.$$

LEMMA 3.2 (RAITUMS, 1992) *Suppose that the matrix-function $A = A(x, \sigma, u, \zeta)$ and the vector-function $a = a(x, \sigma, u, \zeta)$ are such that*

(i) A is uniformly monotone with respect to ζ (uniformly with respect to $x \in \Omega, \sigma \in \mathcal{S}, u \in W_m$);
(ii) if a sequence $\{u_k\}$ converges weakly in W_m to u_0 and a sequence $\{\sigma_k\}$ converges strongly in $\mathcal{S}$ to σ_0 then the sequence $\{A(\cdot, \sigma_k, u_k, u_{0x})\}$ converges strongly in $(W_m)^*$ to $A(\cdot, \sigma_0, u_0, u_{0x})$;
(iii) if sequences $\{\sigma_k\}$ and $\{u_k\}$ converge strongly in $\mathcal{S}$ and W_m to σ_0 and u_0 respectively then the sequences $\{a(\cdot, \sigma_k, u_k, u_{kx})\}$ and $\{A(\cdot, \sigma_k, u_k, u_{kx})\}$ converge strongly in W_m^* to $a(\cdot, \sigma_0, u_0, u_{0x})$ and $A(\cdot, \sigma_0, u_0, u_{0x})$ respectively;

(iv) A and a define Nemitskii operators which map bounded sets from $\mathcal{S} \times W_m$ into bounded sets in $L_2^{(n\times m)}(\Omega)$ and $L_q^{(m)}(\Omega)$ respectively where q is some constant greater then $2n/(n+2)$.

Then, if the sequence $\{\sigma_k\}$ converges strongly in $\mathcal{S}$ to σ_0 and the corresponding sequence $\{u_k\}$ of solutions of equations

$$\mathbf{A}(\sigma_k)u = 0, \; k = 0, 1, 2, \cdots,$$

is bounded then the sequence $\{u_k\}$ converges to u_0.

REMARK. In Raitums (1992) this result is proved in slightly different formulation.

These results give that for wide classes of optimal control problems of the type (16) (provided there is unique solvability of state equations) we can consider as $\mathcal{Z}$ the set

$$\mathcal{Z} := \{\ell = (\ell^1, \cdots, \ell^r) \in L_\infty^{(r)}(\Omega) | \ell(\sigma) = \int_\Omega \sum_{i=1}^r \ell^i \sigma^i dx, ||\ell|| \leq 1\}.$$

All these desired conditions (including $1^\circ - 5^\circ$) are satisfied for the optimal control problems considered below in examples 3.1-3.4.

Now on the basis of concrete examples of optimal control problems we will describe some essential properties of equations and functionals which can exclude the validity of the maximum principle.

In what follows we take

$$\Omega \equiv \{x \in R^2 : |x| < 3\},$$

$$D_0 \equiv \{x \in \Omega : |x| < 1\}, D \equiv \{x \in \Omega : |x| < 2\},$$

$$\mathcal{S} \equiv \{\sigma \in L_2(\Omega) : \sigma(x) = +1 \text{ or } -1, \; x \in D_0; \sigma(x) = 0, \; x \in \Omega \backslash D_0\}.$$

Let, additionally, φ be a function

$$\varphi(x) \equiv \begin{cases} 1, & |x| < 2, \\ -(|x| - 2)^2 + 1, & |x| \geq 2. \end{cases}$$

EXAMPLE 3.1. Minimize the functional

$$I(\sigma, u) = \int_\Omega [(\nabla u^1 - \nabla g^1)^2 + (\nabla u^2 - \nabla g^2)^2] dx \tag{17}$$

subject to

$$\sigma \in \mathcal{S}, u = (u^1, u^2) \in W_2$$

$$div(2 + \sigma)\nabla u^1 = f^1, \; x \in \Omega, \tag{18}$$

$$div(2+\sigma)\nabla u^2 = f^2,\ x \in \Omega,$$

where

$$g^1(x) \equiv x_1\varphi(x), g^2(x) \equiv x_2\varphi(x), x \in \Omega, \qquad (19)$$
$$f^1 \equiv 2\Delta g^1, f^2 \equiv 2\Delta g^2.$$

Here and what follows we denote by ∇f the gradient of the function f, by Δ — the Laplace operator and by $< \cdot, \cdot >$ the scalar product in Euclidean spaces.

The convexification of the problem (17) and (18) is equal to the passage from the set $\mathcal{S}$ to the set $co\mathcal{S}$. The set $co\mathcal{S}$ contains the element $\sigma = 0$, therefore, the price of the convexified problem is equal to zero and the corresponding solution of the problem is $(\sigma = 0, u^1 = g^1, u^2 = g^2)$.

We wish to show that the price of the original problem (17) and (18) is greater than zero. We argue by contradiction. Let there be a sequence $\{\sigma_k, u_k = u(\sigma_k)\} \subset \mathcal{S} \times W_2$ such that

$$I(\sigma_k, u_k) \to 0 \text{ as } k \to \infty. \qquad (20)$$

Because $u_k \in W_2$ then from the convergence (20) it follows that $u_k \to (g^1, g^2)$ in W_2 as $k \to \infty$ and, as a consequence,

$$(\nabla u_k^1, \nabla u_k^2) \to ((1,0),(0,1)) \text{ in } D \text{ as } k \to \infty.$$

From equations (18) with test functions

$$\eta = (\eta^1, \eta^2) \in W_2,\ \eta^1(x) = \eta^2(x) = 0 \ \text{ for } x \in \Omega \backslash D,$$

we have that

$$\begin{aligned} 0 = \sum_{i=1}^{2} \int_D [(2+\sigma_k) < \nabla u_k^i, \nabla\eta^i > -2 < \nabla g^i, \nabla \eta^i >] dx = \\ = \ 2\sum_{i=1}^{2}\int_D \sigma_k < \nabla g^i, \nabla\eta^i > dx + 2\sum_{i=1}^{2}\int_D < \nabla u_k^i - \nabla g^i, \nabla \eta^i > dx + \\ + \sum_{i=1}^{2}\int_D \sigma_k < \nabla u_k^i - \nabla g^i, \nabla\eta^i > dx = \\ = \ 2\int_D \sigma_k div\, \eta\, dx + \varepsilon_k ||\eta||_{W_2}, \end{aligned} \qquad (21)$$

where $\varepsilon_k \to 0$ as $k \to \infty$.

From Temam (1979) there follows the existence of elements $\eta_k \in W_2$ and constants c_*, c_k such that

$$div\, \eta_k = \sigma_k - c_k,\ c_k = \int_D \sigma_k dx,$$

$$||\eta_k||_{W_2} \leq c_* ||\sigma_k - c_k||_{L_2(D)},$$

$$\eta_k(x) = 0,\ x \in \Omega \backslash D,\ k = 1, 2, \cdots.$$

Hence,

$$0 = \lim_{k\to\infty} \varepsilon_k ||\eta_k|| = \lim_{k\to\infty} 2 \int_D \sigma_k div\, \eta_k dx =$$

$$= \lim_{k\to\infty} 2 \int_D (\sigma_k - c_k) div\, \eta_k dx = \lim_{k\to\infty} 2 \int_D (\sigma_k - c_k)^2 dx \geq \pi,$$

because $\sigma_k(x) = +1$ or -1 in D_0 and $\sigma_k(x) = 0$ in $\Omega \backslash D_0$.

This contradiction shows that the price of the original problem (17) and (18) is greater than zero. Thus, according to corollary 2.1., there exist a $\sigma_0 \in co\mathcal{S}$, a $\lambda \in (0, 1]$ and a $\ell \in L_\infty(\Omega)$ such that the problem

$$\int_D [(\nabla u^1 - \nabla g^1)^2 + (\nabla u^2 - \nabla g^2)^2]\, dx + \lambda \int_\Omega \ell\sigma\, dx \to \min,$$

$$(\sigma, u) \in \mathcal{S} \times W_2,$$

$$div\, (2 + \sigma_0 + \lambda(\sigma - \sigma_0))\nabla u^1 = f^1 \text{ in } \Omega,$$

$$div\, (2 + \sigma_0 + \lambda(\sigma - \sigma_0))\nabla u^2 = f^2 \text{ in } \Omega,$$

has a solution (σ_*, u_*) and for this solution the maximum principle does not hold.

Characteristic properties of this example are

(1) $m = n$;

(2) the principal part of the differential operator depends on controls;

(3) the cost functional is not weakly continuous;

(4) the set $\mathcal{S}$ of admissible controls is not convex.

EXAMPLE 3.2. Minimize the functional

$$I(\sigma, u) = \int_\Omega [(u^1 - g^1)^2 + (u^2 - g^2)^2] dx \tag{22}$$

over $(\sigma, u) \in \mathcal{S} \times W_2$ subject to

$$\triangle u^1 + \frac{\partial}{\partial x_1}\sigma + \sigma div\, u - 2u^1 = f^1,\ x \in \Omega, \tag{23}$$

$$\triangle u^2 + \frac{\partial}{\partial x_2}\sigma + \sigma div\, u - 2u^2 = f^2,\ x \in \Omega,$$

where

$$g^1(x) \equiv -x_1\varphi(x),\ g^2(x) \equiv -x_2\varphi(x), x \in \Omega,$$

$$f^1 \equiv \Delta g^1 - 2g^1,\ f^2 \equiv \Delta g^2 - 2g^2.$$

The convexification of the problem (22) and (23) is equal to the passage to the set $co\mathcal{S}$ as a set of admissible controls. Hence, the price of the convexified problem is equal to zero and the corresponding solution is $(\sigma = 0, u^1 = g^1, u^2 = g^2)$.

Let us denote by $u = u(\sigma)$ the solution of the equation (23) with $\sigma \in \overline{co}\mathcal{S}$. The set $\{(\sigma, u) \in L_2(\Omega) \times W_2 : u = u(\sigma), \sigma \in \overline{co}\mathcal{S}\}$ is bounded, hence, weakly sequentially compact. Let sequences $\{\sigma_k\} \subset \mathcal{S}$ and $\{u_k = u(\sigma_k)\} \subset W_2$ converge weakly to σ_0 and u_0 respectively. We intend to derive an equation which will be satisfied by the element u_0.

It is obvious that properties of the weak convergence of the sequence $\{u_k\}$ are fully determined by terms $\frac{\partial}{\partial x_1}\sigma$ and $\frac{\partial}{\partial x_2}\sigma$. Hence

$$u_k = u_0 + v_k + w_k,\ k = 1, 2, \cdots, \tag{24}$$

where $v_k \to 0$ strongly in W_2 as $k \to \infty$ and elements $w_k \in W_2$ are solutions of equations

$$\begin{aligned} &\Delta w_k^1 + \frac{\partial}{\partial x_1}(\sigma_k - \sigma_0) = 0 \text{ in } \Omega,\\ &\Delta w_k^2 + \frac{\partial}{\partial x_2}(\sigma_k - \sigma_0) = 0 \text{ in } \Omega,\\ &w_k = (w_k^1, w_k^2),\ k = 1, 2, \cdots, \end{aligned} \tag{25}$$

respectively.

If we substitute expressions (25) in equations (23) with $\sigma = \sigma_k$ and pass to the limit $k \to \infty$ then we will obtain

$$\begin{aligned} &\Delta u_0^1 + \frac{\partial}{\partial x_1}\sigma_0 + \sigma_0 div u_0 - 2u_0^1 + w \lim_{k\to\infty}[\sigma_k div w_k] = f^1 \in \Omega,\\ &\Delta u_0^2 + \frac{\partial}{\partial x_2}\sigma_0 + \sigma_0 div u_0 - 2u_0^2 + w \lim_{k\to\infty}[\sigma_k div w_k] = f^2 \in \Omega. \end{aligned} \tag{26}$$

Here by $w\lim$ we denote the limit in the sense of weak convergence.

By means of the Green's function of the Poisson's equation we easily obtain that

$$div w_k = -(\sigma_k - \sigma_0),\ k = 1, 2, \cdots.$$

(If functions σ_k, σ_0 are smooth then this result can be obtained from (25) by simple calculations).

Therefore,

$$w \lim_{k\to\infty}[\sigma_k div w_k] = -w \lim_{k\to\infty}[\sigma_k(\sigma_k - \sigma_0)] =$$

$$= w \lim_{k\to\infty} \sigma_k\sigma_0 - w \lim_{k\to\infty} (\sigma_k)^2 = (\sigma_0)^2 - \chi,$$

where χ is the characteristic function of the set D_0 (in this set $|\sigma_k(x)| = 1$).

Thus, we have that the limit element u_0 is a solution of the system

$$\begin{aligned} &\triangle u_0^1 + \frac{\partial}{\partial x_1}\sigma_0 + \sigma_0 div u_0 - 2u_0^1 + (\sigma_0)^2 - \chi = f^1 \text{ in } \Omega, \\ &\triangle u_0^2 + \frac{\partial}{\partial x_2}\sigma_0 + \sigma_0 div u_0 - 2u_0^2 + (\sigma_0)^2 - \chi = f^2 \text{ in } \Omega. \end{aligned} \tag{27}$$

We point out that the set $\mathcal{G}\mathcal{A}_0$ of all operators $\mathbf{B}(\sigma_0) : W_2 \to (W_2)^*$ corresponding to the equation (27) with $\sigma_0 \in \overline{co}\mathcal{S}$ is in fact the $\mathcal{G}$-closure of the set $\mathcal{A}_0$ of all operators $\mathbf{A}(\sigma) : W_2 \to (W_2)^*$ corresponding to the equation (23) with $\sigma \in \mathcal{S}$. More details of $\mathcal{G}$-convergence (or $S\mathcal{G}$-convergence) of elliptic operators of the type (14) can be obtained in Raitums (1985, 1989).

It is obvious that the set $\mathcal{G}\mathcal{A}_0$ is not convex.

Now we are able to discuss the question of the price of the original problem (22) and (23).

Let us suppose that this price is equal to zero and that the $\{\sigma_k, u_k = u(\sigma_k)\} \subset \mathcal{S} \times W_2$ is the corresponding minimizing sequence.

We can assume that this sequence converges weakly to an element (σ_0, u_0). Then from the weak continuity of the functional in (22) it follows that $u_0 = (g^1, g^2)$.

On the other hand the element u_0 has to satisfy the equation (27) with σ_0. Easy calculations show that this leads to the relationship

$$\int_\Omega \{-\sigma_0 div\eta + [(\sigma_0)^2 - \chi - 2\sigma_0](\eta^1 + \eta^2)\}dx = 0, \ \forall \eta \in W_2,$$

or

$$\begin{aligned} &\frac{\partial}{\partial x_1}\sigma_0 + (\sigma_0)^2 - \chi - 2\sigma_0 = 0 \text{ in } \Omega, \\ &\frac{\partial}{\partial x_2}\sigma_0 + (\sigma_0)^2 - \chi - 2\sigma_0 = 0 \text{ in } \Omega, \end{aligned} \tag{28}$$

in the sense of distributions.

From (28) it follows that $\sigma_0 \in \overset{\circ}{W}{}^1_{n+1}(\Omega)$ and as a consequence — $\sigma_0 \in C(\bar{\Omega})$.

Because $\sigma_0(x) = 0$ in $\Omega \backslash D_0$, standard calculations for equations (28) show that $\max\{|\sigma_0(x)| : x \in D_0\} > 1$ what contradicts the definition of the set $\mathcal{S}$ and elementary properties of the set $\overline{co}\mathcal{S}$.

This contradiction ensures that the price of the original problem (22) and (23) is greater than zero. The construction of a concrete example where the maximum principle is not valid is similar to the case of example 3.1.

Characteristic properties of the example 3.2. are

(1) $m = n$;

(2) the principial part of the differential operator depends on controls;
(3) the set of admissible controls is not convex;
(4) the $\mathcal{G}$-closure of the set of admissible operators is not convex (that allowing consideration weakly continuous cost functionals).

EXAMPLE 3.3. Let $m = 1, Q \equiv (0,1) \times \Omega$. We have to minimize the functional

$$I(\sigma, u) = \int_Q (\nabla u - \nabla g)^2 dt dx \tag{29}$$

over $(\sigma, u) \in \mathcal{S} \times V$ subject to

$$u_t = div(2+\sigma)\nabla u + f, (t,x) \in Q, \tag{30}$$

$$u|_{t=0} = 0 ,$$

where

$$g(t,x) \equiv t x_1 \varphi(x) + t^2 x_2 \varphi(x), \ (t,x) \in Q,$$

$$f \equiv g_t - 2\triangle g.$$

The corresponding convexified problem is defined by the set $co\mathcal{S}$ and has a solution $(\sigma = 0, u = g)$ for which the cost functional $I(0,g) = 0$.

If the price of the original problem (29) and (30) is equal to zero and $\{\sigma_k, u_k = u(\sigma_k)\} \subset \mathcal{S} \times V$ is the corresponding minimizing sequence then arguing analogously as in the example 3.1. we get that

$$\int_Q \sigma_k < \nabla g, \nabla \psi > dt dx = \varepsilon_k ||\psi||, \ \forall \psi \in \overset{\circ}{W}{}_2^1 (\Omega), \tag{31}$$

$$\varepsilon_k \to 0 \text{ as } k \to \infty.$$

Let us choose

$$\psi = \lambda_1(t)\eta^1(x) + \lambda_2(t)\eta^2(x)$$

with arbitrary $\eta = (\eta_1, \eta_2) \in W_2$ and functions λ_1, λ_2 such that

$$\int_0^1 t\lambda_1(t)dt = 1, \int_0^1 t^2\lambda_2(t)dt = 1,$$

$$\int_0^1 t\lambda_2(t)dt = 0, \int_0^1 t^2\lambda_1(t)dt = 0.$$

Then from (31) it follows immediately that (controls σ_k not depending on time variable t),

$$\int_\Omega \sigma_k div\, \eta\, dx = \delta_k ||\eta||, \forall \eta \in W_2,$$

$\delta_k \to 0$ as $k \to \infty$.

This is the same situation as in the example 3.1. Hence, the price of the original problem (29) and (30) is greater than zero.

The continuous dependence of solution of linear parabolic equations on coefficients is discussed, for instance, in Ladyzhenskaya, Solonnikov and Uraltseva (1967). These results suffice for the existence of an example of optimal control problems

$$I(\sigma, u) = \int_Q (\nabla u - \nabla g)^2 dtdx + \lambda \int_Q \ell\sigma dtdx \to \min,$$

$$(\sigma, u) \in \mathcal{S} \times V,$$

$$u_t = div(2 + \sigma_0 + \lambda(\sigma - \sigma_0))\nabla u + f \text{ in } Q$$

$$u|_{t=0} = 0 \,,$$

with some $\sigma_0 \in co\mathcal{S}$, $\lambda \in (0, 1]$, $\ell \in L_\infty(Q)$ for which there exists a solution but this solution does not satisfy the maximum principle.

The same reasoning is valid for cases of elliptic equations where controls do not depend upon one or more spatial variables.

EXAMPLE 3.4. Let $m = 1$ and $\mathcal{S}_0 \equiv \{\sigma \in \mathcal{S} : \sigma$ does not depend on x_2 in $D_0\}$.

We have to minimize

$$I(\sigma, u) = \int_\Omega (u - g)^2 dx \tag{32}$$

over $(\sigma, u) \in \mathcal{S}_0 \times W_1$ subject to

$$\triangle u + \frac{\partial}{\partial x_1}\sigma + \sigma u_{x_1} - 2u = f, \; x \in \Omega, \tag{33}$$

where

$$g(x) \equiv -x_1\varphi(x), x \in \Omega,$$

$$f \equiv \triangle g - 2g.$$

The convexified problem has a solution $(\sigma = 0, u = g)$ and the price is equal to zero.

If $\{\sigma_k, u_k = u(\sigma_k)\} \subset \mathcal{S}_0 \times W_1$ is the minimizing sequence for the problem (32), (33) and $\sigma_k \to \sigma_0$, $u_k \to u_0 = g$ weakly in $L_2(\Omega)$ and W_1 respectively as $k \to \infty$ then analogously as in example 3.2 we get that u_0 satisfies the equation

$$\mathbf{B}(\sigma_0)u \equiv \triangle u + \frac{\partial}{\partial x_1}\sigma_0 + \sigma_0 u_{x_1} + (\sigma_0)^2 - \chi - 2u = f \text{ in } \Omega.$$

Because the function σ_0 does not depend on x_2 in D_0 then almost the same reasoning as in example 3.2 gives that there is no element $\sigma \in co\mathcal{S}_0$ for which

$$\mathbf{B}(\sigma)g = f \text{ in } \Omega.$$

Hence, the price of the original problem (32) and (33) is greater than zero.

For these two examples characteristic properties are

(1) the set of admissible controls is not convex;
(2) the principal part of the differential operator depends upon controls;
(3) admissible controls do not depend upon one or more variables;
(4) either the cost functional is not weakly continuous or the $\mathcal{G}$-closure of the set of admissible operators is not convex.

We point out that a very special form of equations and functionals in these examples is not essential. It is obvious that for small enough perturbations of equations (or functionals) which maintain the continuity of inverse operators and do not lead out of the class given by (15) and (16) the fact of nonequality of prices remains.

References

ALTMAN M. (1979) An application of the method of contractor directions to non–linear programming, *Numerical Functional Analysis and Optimization*, **1**, 6, 647–663.

CASTAIGN C., VALADIER M. (1977) *Convex Analysis and Measurable Multifunctions,* Lecture notes in Mathematics, vol. 580, Berlin–New York, Springer–Verlag.

ECLAND I., TEMAM R. (1976) *Convex Analysis and Variational Problems*, Amsterdam-New York-Oxford, North-Holland.

GAMKRELIDZE R.V. (1975) *Principles of Optimal Control,* Tbilisi, Tbilisi University, (in Russian).

LADYZHENSKAYA O.A., SOLONNIKOV V.A. AND URALTSEVA N.N. (1967) *Linear and Quasi Linear Equations of the Parabolic Type,* Moscow, Nauka, (in Russian).

RAITUMS U. (1976) *Extremal Problems for the Second Order Linear Elliptic Equations, Latvian Mathematical Annual,* bf 19, 198–213, (in Russian).

RAITUMS U. (1985) On a generalization of the *G*-convergence for elliptic systems, *Latvian Mathematical Annual*, **29**, 73–83, (in Russian).

RAITUMS U. (1989) *Optimal Control Problems for Elliptic Equations,* Riga, Zinatne, (in Russian).

RAITUMS U. (1992) On the strong convergence of solutions of equations with multivalued operators, *Acta Universitatis Latviensis*, **570**, 12–23.

TEMAM R. (1979) *Navier-Stokes Equations,* Amsterdam-New York-Oxford, North-Holland.

WARGA J. (1972) *Optimal Control of Differential and Functional Equations,* New York, Academic Press.