Control and Cybernetics

vol. **30** (2001) No. 1

Counterexample to extension via convexification of optimal control problems for elliptic systems

by

Uldis Raitums

Institute of Mathematics and Computer Science
29 Rainis boulevard, LV-1459 Riga, Latvia

Abstract: We present a counterexample, which shows that convexification, i.e. the passage to the convex hull of the set of admissible operators, does not preserve the convex hull of the set of feasible states of the corresponding family of elliptic systems.

Keywords: optimal control, elliptic system, relaxation, counterexample.

1. Introduction

We are interested in possible extensions of optimal control problems of the kind

$$
\begin{aligned}
&I_0(\overline{u}) = \sum_{i=1}^{m} \int_\Omega \langle f_i, \nabla u_i \rangle \, dx \rightarrow min \\
&\mathcal{L}(\sigma)\overline{u} \equiv div A(x,\sigma)\nabla\overline{u} = div\overline{f} \text{ in } \Omega, \\
&\overline{u} = (u_1, \ldots, u_m) \in H_0^1(\Omega; \mathbf{R}^m), \\
&\sigma \in S_0 = \{\sigma \in L_\infty(\Omega, \mathbf{R}^r) \mid \sigma(x) \in U_0 \text{ a.e. } x \in \Omega\},
\end{aligned}
\tag{1.1}
$$

where $\Omega \subset \mathbf{R}^n$ is a given bounded Lipschitz domain and $U_0 \subset \mathbf{R}^r$ is a given bounded set (not necessary convex). This class includes some characteristic problems of the optimal material layout, see, for instance, Bendsoe et al. (1993) and references therein.

The interest in extensions of optimal control problems is caused by two main factors: most of the optimal control problems do not possess solutions; the sets of admissible controls are often not convex, which leads to serious difficulties in numerical resolution. It is clear that the passage to the closure of the set $\{\mathcal{L}(\sigma) \mid \sigma \in S_0\}$ with respect to G-convergence (see, for instance, Zhikov et al., 1994) theoretically solves the existence of a solution for the problem (1.1). Unfortunately, until now no effective descriptions of G-closure for the case of elliptic systems have been found. In addition, the G-closed sets are not necessarily convex even for the simplest cases of a single equation, see, for instance, Zhikov et al. (1994). Here the question arises whether is it possible to use the convexification approach to optimal control problems for elliptic systems.

For the case of a single equation, i.e. $m = 1$ in the problem (1.1), it is well known, see, for instance, Raitums (1985), that the passage from the set $\{\mathcal{L}(\sigma) \mid \sigma \in S_0\}$ of admissible operators to its convex hull preserves the closure of the set of feasible states in both strong and weak topologies of $H_0^1(\Omega)$. Moreover, for the most interesting case of the optimal material layout problem (for $m = 1$) Tartar (1997) has given a construction of a convex set $\mathcal{B}$ of matrix-valued functions $B : \Omega \to \mathbf{R}^{n\times n}$ such that the set of all solutions of the corresponding equation in (1.1) with $B \in \mathcal{B}$ instead of $A(\cdot,\sigma)$, $\sigma \in S_0$, gives the weak closure of the initial set of feasible states.

On the other hand, in Raitums (1994) it was shown that optimal control problems for elliptic systems with $m \geq n$ and with controls in the leading coefficients of the corresponding differential operators do not admit the relaxation via convexification, i.e. the passage from $\{\mathcal{L}(\sigma) \mid \sigma \in S_0\}$ to its convex hull does not preserve the closure of the set of feasible states even in the weak topology of $H_0^1(\Omega;\mathbf{R}^m)$. Since linear cost functionals, as in the problem (1.1), demand only the preservation of the convex hull of the set of feasible states, then, for a complete picture, it is necessary to find an answer to the question: does the passage from $\{\mathcal{L}(\sigma) \mid \sigma \in S_0\}$ to its convex hull preserve the convex hull of the set of feasible states or not? Until now no satisfactory answer to that question is known, moreover, as far as I know, the problem of the preservation of the convex hull of the set of feasible states in relaxation procedures for elliptic equations has not been investigated yet, at least for the situations analogous to (1.1). Of course, if a relaxation procedure preserves the weak closure of the set of feasible states, then this procedure preserves the corresponding convex hull too, but not *vice versa*. We point out here that such questions have a meaning only if the sets of admissible controls are rich enough, like for instance, the set S_0 in (1.1).

In what follows we will present a simple problem of the kind of (1.1), with $m = n = 2$, for which the convexification does not preserve the value of the problem. Obviously, that implies that for elliptic systems with $m \geq n \geq 2$ the procedure of convexification does not preserve, in general, the convex hull of the set of feasible states.

2. Statement of the problem

Let $\Omega \subset \mathbf{R}^2$ be a bounded Lipschitz domain, let $D \subset \Omega$ be a cube, $\overline{D} \subset \Omega$, and let

$$S = \left\{ \sigma \in L_\infty(\Omega) \,\middle|\, \sigma(x) = 0 \text{ in } \Omega\setminus\overline{D},\ \sigma(x) = +\varepsilon \text{ or } -\varepsilon \text{ in } D, \right.$$
$$\left. \int_D \sigma(x)\,dx = 0 \right\},$$

where the exact value of $\varepsilon \in (0,1)$ will be specified later.

Consider the optimal control problem

$$I(\overline{u}) \equiv \int_\Omega [\langle \nabla u_1, f_1\rangle + \langle \nabla u_2, f_2\rangle]\, dx \to min, \tag{2.1}$$

$$\begin{cases} div(1+\sigma(x))\nabla u_1 = div f_1 \text{ in } \Omega, \\ div(1-\sigma(x))\nabla u_2 = div f_2 \text{ in } \Omega, \end{cases} \tag{2.2}$$

$$\overline{u} = (u_1, u_2) \in H_0^1(\Omega; \mathbf{R}^2), \ \sigma \in S, \tag{2.3}$$

where $f_1, f_2 \in L_2(\Omega; \mathbf{R}^2)$ are given as

$$\begin{array}{l} f_1 = \nabla \varphi_1, \ f_2 = \nabla \varphi_2; \ \varphi_1, \varphi_2 \in H_0^1(\Omega), \\ \nabla \varphi_1(x) = (1,0) \text{ in } D, \ \nabla \varphi_2(x) = (0,1) \text{ in } D. \end{array} \tag{2.4}$$

Denote by $\overline{u}(\sigma) = (u_1(\sigma), u_2(\sigma)) \in H_0^1(\Omega; \mathbf{R}^2)$ the solution of (2.2) corresponding to a chosen $\sigma \in \overline{co}\, S$ and denote by $J(\sigma)$ the value of $I(\overline{u}(\sigma))$. The special type of the functional and equations in (2.1), (2.2) yield

$$\begin{aligned} J(\sigma) = - \inf_{u_1, u_2 \in H_0^1(\Omega)} \int_\Omega & [(1+\sigma(x))\nabla u_1^2 - 2\langle f_1, \nabla u_1\rangle \\ & + (1-\sigma(x))\nabla u_2^2 - 2\langle f_2, \nabla u_2\rangle]\, dx. \end{aligned} \tag{2.5}$$

Thus, the problem (2.1)–(2.3) is equivalent to minimization of the functional J on the set S. Our aim is to show that

$$\inf_{\sigma \in \overline{co}\, S} J(\sigma) < \inf_{\sigma \in S} J(\sigma).$$

3. Solution of the confexified problem

It is easy to see that the functional J is Fréchet differentiable on $\overline{co}\, S$ (we consider $\overline{co}\, S$ in the strong topology of $L_\infty(\Omega)$). Indeed, the operator

$$\mathcal{A} : \overline{co}\, S \times H_0^1(\Omega; \mathbf{R}^2) \to H^{-1}(\Omega; \mathbf{R}^2), \ \mathcal{A}(\sigma, \overline{u}) = \begin{pmatrix} div(1+\sigma)\nabla u_1 \\ div(1-\sigma)\nabla u_2 \end{pmatrix},$$

satisfies all assumptions of the implicit function theorem, which gives the Fréchet differentiability of the mapping $\sigma \to \overline{u}(\sigma)$. From here and the linearity of I with respect to $\overline{u}$ Fréchet differentiability of J immediately follows.

The functional J is convex on $\overline{co}\, S$ as a supremum $(-inf\, F = sup(-F))$ of functionals linear with respect to σ. Therefore, if $\sigma_0 \in \overline{co}\, S$ is such that

$$J'(\sigma_0)(\sigma - \sigma_0) \geq 0 \ \forall \sigma \in \overline{co}\, S$$

then

$$J(\sigma_0) \leq J(\sigma) \ \forall \sigma \in \overline{co}\, S.$$

Here, by $J'(\sigma)$ we denote the Fréchet derivative of J on the element σ.

Straightforward standard calculations give

$$J'(\sigma)\delta\sigma = -\int_\Omega [|\nabla u_1(\sigma)|^2 - |\nabla u_2(\sigma)|^2]\delta\sigma\, dx.$$

The element $\sigma_0 = 0$ belongs to $\overline{co}\, S$ and the special choice of f_1 and f_2 (relationships (5)) yields

$$u_1(\sigma_0) = \varphi_1,\ u_2(\sigma_0) = \varphi_2,$$
$$\nabla u_1(\sigma_0)(x) = (1,0) \text{ in } D,\ \nabla u_2(\sigma_0)(x) = (0,1) \text{ in } D.$$

Hence,

$$J'(\sigma_0) = 0 \text{ and } J(\sigma_0) \leq J(\sigma)\ \forall \sigma \in \overline{co}\, S.$$

4. Evaluation of $J(\sigma)$

We start with an abstract result.

Let H be a real Hilbert space with the inner product $\langle\cdot,\cdot\rangle$, let M be a set (not necessarily convex) of linear continuous symmetric operators $A : H \to H$ and let there exist positive constants $0 < \nu < \mu$ such that for all $A \in M$ and all $w \in H$

$$\langle Aw, w\rangle \geq \nu\|w\|^2,\ \|Aw\| \leq \mu\|w\|. \tag{4.1}$$

Let the elements $f, g \in H$ be fixed and let

$$I(A) = \inf_{w\in H}[\langle A(w+f), w+f\rangle - 2\langle g, w\rangle].$$

THEOREM 4.1 *Let $A, A_0 \in M$. Then*

$$\begin{aligned} I(A) &= I(A_0) + \langle (A - A_0)(w_0 + f), w_0 + f\rangle \\ &+ \inf_{w'\in H}[\langle A_0 w', w'\rangle + 2\langle (A - A_0)(w_0+f), w'\rangle] \\ &+ \inf_{w''\in H}[\langle A w'', w''\rangle + 2\langle (A - A_0)w_0', w''\rangle] \\ &+ \langle (A - A_0)w_0', w_0'\rangle, \end{aligned} \tag{4.2}$$

where $w_0 \in H$ and $w_0' \in H$ satisfy

$$A_0(w_0 + f) = g,\ A_0 w_0' + (A - A_0)(w_0 + f) = 0. \tag{4.3}$$

Proof. By virtue of (4.1) the operators $A_0 + t(A - A_0) : H \to H$ with $0 \leq t \leq 1$ are invertible. Denote by w_t the solution of the equation

$$(A_0 + t(A - A_0))(w + f) = g.$$

Simple calculations show that w_t has the representation

$$w_t = w_0 + tw_0' + w_0'',$$

where w_0 and w_0' are given by (4.3), but w_0'' satisfies

$$(A_0 + t(A - A_0))w_0'' + t^2(A - A_0)w_0' = 0.$$

Since

$$\inf_{w \in H}[\langle (A_0 + t(A - A_0))(w + f), w + f\rangle - 2\langle g, w\rangle]$$
$$= \langle (A_0 + t(A - A_0))(w_t + f), w_t + f\rangle - 2\langle g, w_t\rangle$$

and analogous relationships hold for w_0' and w_0'', straightforward calculations give

$$I(A_0 + t(A - A_0)) = \langle A_0(w_0 + f), w_0 + f\rangle - 2\langle g, w_0\rangle$$
$$+ t\langle (A - A_0)(w_0 + f), w_0 + f\rangle$$
$$+ t^2 \inf_{w' \in H}[\langle A_0 w', w'\rangle + 2\langle (A - A_0)(w_0 + f), w'\rangle]$$
$$+ \inf_{w'' \in H}[\langle (A_0 + t(A - A_0))w'', w''\rangle + t^2 2\langle (A - A_0)w_0', w''\rangle]$$
$$+ t^3\langle (A - A_0)w_0', w_0'\rangle.$$

This relationship is valid for all $0 \le t \le 1$ and from here with $t = 1$ immediately follows (4.2). ■

COROLLARY 4.1 *Let A, $A_0 \in M$. Then*

$$I(A) \le I(A_0) + \langle (A - A_0)(w_0 + f), w_0 + f\rangle$$
$$+ \langle (A - A_0)(w_0 + f), w_0'\rangle + \langle (A - A_0)w_0', w_0'\rangle, \tag{4.4}$$

where w_0 and w_0' satisfy (4.3).

Proof. The proof immediately follows from the fact that both infimums on the right hand side of (4.2) are less than or equal to zero and that

$$\inf_{w'}[\langle A_0 w', w'\rangle + 2\langle (A - A_0)(w_0 + f), w'\rangle] = \langle (A - A_0)(w_0 + f), w_0'\rangle. \quad ■$$

We want to apply these results to the functional J given by (2.5). We have

$$-J(\sigma) = \inf_{w \in H} \langle Aw, w\rangle - 2\langle g, w\rangle$$

where

$$H = H_0^1(\Omega) \times H_0^1(\Omega),\ w = (w_1, w_2) \in H,\ g = P(f_1, f_2),$$
$$Aw = P((1 + \sigma(\cdot))\nabla w_1, (1 - \sigma(\cdot))\nabla w_2),$$
$$P : L_2(\Omega, \mathbf{R}^2) \times L_2(\Omega; \mathbf{R}^2) \to H,$$

P is the orthogonal projector and for every $f = (f_1, f_2) \in L_2(\Omega; \mathbf{R}^2) \times L_2(\Omega; \mathbf{R}^2)$ and every $w = (w_1, w_2) \in H$

$$\langle Pf, w\rangle = \int_\Omega [\langle f_1, \nabla w_1\rangle + \langle f_2, \nabla w_2\rangle]\, dx.$$

Let A_0 correspond to $\sigma_0 = 0$. Then, the elements w_0 and w_0' from (4.3) are equal to

$$w_0 = (\varphi_1, \varphi_2), \; w_0' = (v_1, -v_2),$$

where $\overline{v} = (v_1, v_2) \in H$ and

$$\triangle v_i + div((\sigma - \sigma_0)\nabla\varphi_i) = 0 \text{ in } \Omega, \; i = 1, 2,$$

or, taking into account the special choice of φ_1 and φ_2 and that $\sigma(x) = \sigma_0(x) = 0$ in $\Omega \setminus D$,

$$\triangle v_i + \frac{\partial}{\partial x_i}\sigma = 0 \text{ in } \Omega, \; i = 1, 2. \tag{4.5}$$

Since the term $\langle (A - A_0)(w_0 + f), w_0 + f\rangle$ in (4.4) represents the Gateaux derivativee of I at A_0 and we know that $J'(\sigma_0) = 0$, from (4.4) we get

$$-J(\sigma) \le -J(\sigma_0) + \int_D \sigma div \overline{v}\, dx + \int_D \sigma[|\nabla v_1|^2 - |\nabla v_2|^2]\, dx. \tag{4.6}$$

Elements $\sigma \in S$ have a zero mean value, hence, according to Nečas (1965) there exists a positive constant c_0 such that all $\sigma \in S$ satisfy

$$||\sigma||_{L_2(\Omega)} \le c_0 \sup_{\overline{u} \in H_0^1(\Omega;\mathbf{R}^2);\, ||\overline{u}|| \le 1} \int \sigma div\, \overline{u}\, dx.$$

From here and from (4.5) it follows immediately that

$$||\,|\nabla\overline{v}|\,||_{L_2(\Omega)} \ge \frac{1}{c_0}||\sigma||_{L_2(\Omega)}, \; \int_D \sigma div\, \overline{v}\, dx \le -\Big(\frac{1}{c_0}||\sigma||_{L_2(\Omega)}\Big)^2. \tag{4.7}$$

On the other hand, the standard *a priori* estimates for solutions of elliptic equations and (4.5) give

$$\int_\Omega |\nabla\overline{v}|^2\, dx \le \int_\Omega \sigma^2\, dx,$$

hence,

$$\Big|\int_D \sigma(\nabla v_1^2 - \nabla v_2^2)\, dx\Big| \le \varepsilon \int_\Omega \sigma^2\, dx \tag{4.8}$$

(we recall that $\sigma \in S$ takes values $+\varepsilon$ or $-\varepsilon$ in D).

Relationships (4.6), (4.7) and (4.8) yield for $\sigma \in S$

$$J(\sigma) \ge J(\sigma_0) + c_0^{-2}\int_\Omega \sigma^2\, dx - \varepsilon\int_\Omega \sigma^2\, dx = J(\sigma_0) + c_0^{-2}\varepsilon^2|D| - \varepsilon^3|D|,$$

where by $|D|$ we denote Lebesgue measure of D.

Obviously, we can choose $1 > \varepsilon > 0$ such that the expression

$$c_0^{-2}\varepsilon^2|D| - \varepsilon^3|D|$$

is strictly positive. Therefore, for such $\varepsilon > 0$

$$\inf_{\sigma\in S} J(\sigma) > J(\sigma_0) = \inf_{\sigma\in\overline{co}\, S} J(\sigma).$$

In this manner, the set of all solutions of (2.2) with $\sigma \in S$ is strictly separated by a linear continuous functional from the solution $\overline{u}(\sigma_0)$ of (2.2) with $\sigma_0 \in \overline{co}\, S$, which shows that the passage to the convex hull of the set of admissible operators does not preserve, for the case of elliptic systems, the convex hull of the set of feasible states of the original problem.

References

BENDSOE, M.P., DIAZ, A. and KIKUCHI, N. (1993) Topology and generalized layout optimization of elastic structures, In: Bendsoe, M.P. and Mota Soares, C.A., eds., *Topology Design of Structures.* Kluwer, Dodrecht, 159–205.

NEČAS, J. (1965) *Equations aux Dérivées Partielles.* Presses de l'Université de Montréal, Montréal.

RAITUMS, U. (1985) On passage to the convex hull of a set of admissible operators in optimal control problems. *Soviet. Math. Dokl.*, **32**, 3, 680–683.

RAITUMS, U. (1994) The maximum principle and the convexification of optimal control problems. *Control and Cybernetics*, **24**, 4, 745–760.

TARTAR, L. (1997) Remarks on the homogenization method in optimal design method. In: Cioranescu, D. et al., eds., *Homogenization and Applications to Material Science.* GAKUTO Internat. Ser. Math. Sci. Appl. 9, Gakkotosho, Tokyo, 279–296.

ZHIKOV, V.V., KOZLOV, S.M. and OLEINIK, O.A. (1994) *Homogenization of Differential Operators and Integral Functionals.* Springer-Verlag, New York.