UTMS 2005–34 August 17, 2005

Applications of convexity in some identification problems

by

Dan TIBA, Gengsheng WANG
and Masahiro YAMAMOTO

UNIVERSITY OF TOKYO
GRADUATE SCHOOL OF MATHEMATICAL SCIENCES
KOMABA, TOKYO, JAPAN

Applications of convexity in some identification problems

Dan Tiba[1], Gengsheng Wang[2] and Masahiro Yamamoto[3]

[1] Institute of mathematics, Romanian Academy

P.O.Box 1-764, RO-014700 Bucharest, Romania

[2] Department of Mathematics, Normal University of Central China

Wuhan, China

[3] Graduate School of Mathematical Sciences, The University of Tokyo

3-8-1 Komaba Meguro Tokyo 153, Japan

1 Introduction

We study a new reconstruction technique in some elliptic identification problems. Our method applies to distributed observation on some subset of the problem domain or to boundary observation via Dirichlet conditions. Moreover our method is applicable to parabolic cases although we mainly discuss the elliptic case.

The basic idea is to divide the classical least squares approach into two steps: the first one searches for the identification parameters (coefficients, sources) that ensure one inequality with respect to the observed state, while the second one deals with the opposite inequality. In many situations, which are partially discussed in Section 2, the first step has convexity properties, even in the case of nonlinear partial differential equations (including variational inequalities). Such properties have been studied by Lemaire [4], Kawohl and Lang [3], Liu and Tiba [5] in a different context. This is an important advantage for the numerical solution, although uniqueness is not valid, in general.

Moreover the convexity gives also the equivalence of the first minimization problem with the corresponding set of optimality conditions. Then, the second step of the method (which is more difficult) admits again a formulation as an optimal control problem which may be solved by appropriate techniques. The argument is indicated in Section 3.

2 The convexity property

Throughout this paper, we set $f_+(x) = \max\{f(x), 0\}$ and $f_-(x) = -\min\{f(x), 0\}$ for a real-valued function f. We note $|f(x)| = f_+(x) + f_-(x)$.

We fix our attention on the Dirichlet problem for the nonlinear elliptic equation

$$-\triangle y(x) = \varphi(x, y(x), u(x)) \quad \text{in} \quad \Omega, \tag{2.1}$$

$$y = 0 \quad \text{on} \quad \partial\Omega. \tag{2.2}$$

Other boundary conditions such as Neumann, mixed or of Robin types may be also considered. Parabolic equations can be treated by our method.

We assume that $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, is a bounded smooth domain and φ: $\Omega \times \mathbb{R} \times \mathbb{R}^k \to \mathbb{R}$, $k \in \mathbb{N}$, satisfies (2.3) - (2.5):

(2.3) $\varphi(x, \cdot, \cdot)$ is convex for a.e. $x \in \Omega$ and there exist $C > 0$ and a positive function $M \in L^s(\Omega)$, $s > \max\left\{2, \frac{d}{2}\right\}$ such that $|\varphi(x,0,u)| \leq M(x) + C|u|$, $\forall (x,u) \in \Omega \times \mathbb{R}^k$;

(2.4) there exist $C > 0$ and a positive function $M \in L^s(\Omega)$, $s > \max\left\{2, \frac{d}{2}\right\}$, a nondecreasing function $\eta : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ such that $0 \geq \varphi'_y(x,y,u) \geq - (M(x) + C|u|)\, \eta(y)$, $\forall (x,y,u) \in \Omega \times \mathbb{R} \times \mathbb{R}^k$;

(2.5) $\varphi(x,y,u)$ and $\varphi'_y(x,y,u)$ are Caratheodory mappings, i.e., measurable in x and continuous in y, u.

The vector-valued function $u = u(x)$ describes coefficients or source terms in equation (2.1). This will be made clear in the examples. If $u \in L^\infty(\Omega)^k$, then (2.1) and (2.2) possess a unique solution $y = y(u)$ in $W^{2,s}(\Omega) \cap H_0^1(\Omega)$. This is a consequence of (2.3)–(2.5) and the well known theory of nonlinear elliptic equations (e.g., Barbu [1], Neittaanmäki, Sprekels and Tiba [6]). We denote the solution to (2.1) and (2.2) by $y(u) = T(u)$ for $u \in L^\infty(\Omega)^k$.

Proposition 2.1. *Under the above assumptions, the nonlinear operator* T: $L^\infty(\Omega)^k \to L^2(\Omega)$, *is convex.*

Proof.

Let u_1, $u_2 \in L^\infty(\Omega)^k$ and $\lambda \in]0,1[$ be given. Let y_1, y_2, y_λ denote the solutions of (2.1), (2.2) corresponding to u_1, u_2 and to $\lambda u_1 + (1-\lambda)u_2$ respectively. By multiplying the equations corresponding to y_1, y_2 respectively by λ, $(1-\lambda)$ and by adding them, hypothesis (2.3) yields the inequality:

$$-\triangle\left(\lambda y_1 + (1-\lambda)y_2\right) \geq \varphi\left(x, \lambda y_1 + (1-\lambda)y_2, \lambda u_1 + (1-\lambda)u_2\right) \text{ in } \Omega. \tag{2.6}$$

We subtract (2.6) from the equation corresponding to y_λ and we obtain:

$$\begin{aligned} -\triangle\left(y_\lambda - \lambda y_1 - (1-\lambda)y_2\right) &\leq \varphi\left(x, y_\lambda, \lambda u_1 + (1-\lambda)u_2\right) \\ &\quad -\varphi\left(x, \lambda y_1 + (1-\lambda)y_2, \lambda u_1 + (1-\lambda)u_2\right) \quad \text{in} \quad \Omega. \end{aligned} \tag{2.7}$$

Multiply (2.7) by $[y_\lambda - \lambda y_1 - (1-\lambda)y_2]_+ \in H_0^1(\Omega)$ and integrate in Ω, we have

$$\int_\Omega \left(\nabla\left[y_\lambda - \lambda y_1 - (1-\lambda)y_2\right]_+\right)^2 \mathrm{d}x \le 0. \tag{2.8}$$

Here we used

$$\begin{aligned}
&-\int_\Omega \Delta(y_\lambda - \lambda y_1 - (1-\lambda)y_2)[y_\lambda - \lambda y_1 - (1-\lambda)y_2]_+ dx \\
&= \int_\Omega \nabla(y_\lambda - \lambda y_1 - (1-\lambda)y_2)\cdot\nabla[y_\lambda - \lambda y_1 - (1-\lambda)y_2]_+ dx \\
&= \int_\Omega (\nabla[y_\lambda - \lambda y_1 - (1-\lambda)y_2]_+)^2 dx
\end{aligned}$$

and

$$\begin{aligned}
&(\varphi(x, y_\lambda, \lambda u_1 + (1-\lambda)u_2) - \varphi(x, \lambda y_1 + (1-\lambda)y_2, \lambda u_1 + (1-\lambda)u_2)) \times [y_\lambda - \lambda y_1 - (1-\lambda)y_2]_+ \\
&\le 0 \quad \text{if } y_\lambda(x) > (\lambda y_1 + (1-\lambda)y_2)(x), \text{ i.e., } [y_\lambda - \lambda y_1 - (1-\lambda)y_2]_+(x) > 0, \\
&= 0 \quad \text{if } y_\lambda(x) \le (\lambda y_1 + (1-\lambda)y_2)(x),
\end{aligned}$$

by (2.4). Clearly (2.8) signifies that $[y_\lambda - \lambda y_1 - (1-\lambda)y_2]_+ = 0$ a.e. in Ω, that is

$$y_\lambda(x) \le \lambda y_1(x) + (1-\lambda)y_2(x) \quad \text{a. e. in} \quad \Omega. \tag{2.9}$$

Relation (2.9) is the desired infinite dimensional convexity property of T and ends the proof.

Let us now comment briefly on some classes of examples that satisfy assumptions (2.3)-(2.5). A very simple situation is obtained when φ arises as a sum

$$\varphi(x, y, u) = a(x)y + b(x)u \tag{2.10}$$

with a, $b \in L^\infty(\Omega)$. If $a \le 0$ a. e. in Ω, then (2.3)-(2.5) are clearly fulfilled. More generally, φ may be assumed to be the sum of convex mappings:

$$\varphi(x, y, u) = \psi_1(x, y) + \psi_2(x, u). \tag{2.11}$$

If ψ_1, ψ_2 are convex functions and $\psi_1(x,\cdot)$ is nonincreasing for a.e. $x \in \Omega$, then (2.3)-(2.5) are valid. As a limit case, for instance, if $\psi_1(x,y) = -\beta(y)$ with β being the maximal monotone graph

$$\beta(y) = \begin{cases} \emptyset, & y < 0, \\ ]-\infty, 0], & y = 0, \\ 0, & y > 0, \end{cases} \tag{2.12}$$

then equations (2.1) and (2.2) become a variational inequality and Proposition 2.1 remains valid. Such situations have been studied by Lemaine [4] and the extension of Proposition 2.1 to (2.12) is obtained by first considering the case of the Yosida approximation $\beta_\varepsilon(y)$ and then by passing to the limit $\varepsilon \to 0$ as it is standard in the theory of variational inequalities, Barbu [1].

Example 2.2.

We indicate now a situation when u appears as a coefficient in the equation:

$$\triangle \left(u^3(x)\triangle y(x)\right) = f(x) \quad \text{in} \quad \Omega, \tag{2.13}$$

$$y = 0, \quad \triangle y = 0 \quad \text{on} \quad \partial\Omega. \tag{2.14}$$

If $\Omega \subset \mathbb{R}$ or $\Omega \subset \mathbb{R}^2$, then problem (2.13) and (2.14) is a simplified model for the deflection of a beam, respectively a plate, under the load f. The coefficient $u \in L^\infty(\Omega)$ with $u > 0$ represents the thickness and boundary conditions (2.13) correspond to the simply supported case.

We assume that $f \in L^2(\Omega)$ satisfies $f \geq 0$ a. e. in Ω and by $h \in H^2(\Omega) \cap H_0^1(\Omega)$ we denote the unique solution of

$$\triangle h = f \quad \text{in} \quad \Omega, \tag{2.15}$$

$$h = 0 \quad \text{on} \quad \partial\Omega. \tag{2.16}$$

By the maximum principle, we clearly have $h(x) \le 0$ in Ω. It is also easy to see that y given by (2.13), (2.14) can be equivalently defined by

$$\triangle y(x) = \frac{h(x)}{u^3(x)} \quad \text{in} \quad \Omega, \tag{2.17}$$

$$y = 0 \quad \text{on} \quad \partial\Omega. \tag{2.18}$$

Since the mapping $u \mapsto u^{-3}$ is convex for $u > 0$ and h is negative, for any u_1, $u_2 \in L^\infty(\Omega)$ with $u_1, u_2 > 0$, we have the inequality

$$\frac{h(x)}{[\lambda u_1(x) + (1-\lambda)u_2(x)]^3} \ge \lambda \frac{h(x)}{u_1(x)^3} + (1-\lambda)\frac{h(x)}{u_2(x)^3} \quad \text{in} \quad \Omega. \tag{2.19}$$

Let y_1, $y_2 \in H^2(\Omega) \cap H_0^1(\Omega)$ denote the solutions of (2.17), (2.18) corresponding to u_1, respectively u_2 and let y_λ denote the solution of (2.17), (2.18) corresponding to $\lambda u_1 + (1-\lambda)u_2$. By (2.17)–(2.19), we obtain

$$\triangle y_\lambda \ge \lambda \triangle y_1 + (1-\lambda)\triangle y_2 \quad \text{in} \quad \Omega. \tag{2.20}$$

The maximum principle shows

$$y_\lambda(x) \le \lambda y_1(x) + (1-\lambda)y_2(x) \quad \text{a. e. in} \quad \Omega.$$

Consequently, we see that (2.9), i.e., the conclusion of Proposition 2.1 remains valid in the case of equations (2.13) and (2.14) (with f positive) as well.

Let us now formulate the identification problem for equations (2.1) and (2.2), which includes all the examples mentioned above. We assume for simplicity that some distributed observation $\widehat{y}$ can be measured on a given measurable subset $\omega \subset \Omega$. However the discussion below extends to the case of Dirichlet observation, when boundary condition (2.2) is replaced by some other type of boundary condition, as it has been already mentioned.

As a first step in the reconstruction method that we are proposing, we modify the classical least square approach and we introduce the following functional:

$$\min_u \frac{1}{2}\int_\omega (y(u)-\widehat{y})_+^2\,\mathrm{d}x \tag{2.21}$$

subject to (2.1), (2.2) and to the constraint

$$u \in \mathcal{K} \subset L^\infty(\Omega)^k, \tag{2.22}$$

$\mathcal{K}$ being a bounded closed convex subset. We set

$$J(u) = \frac{1}{2}\int_\omega (y(u)-\widehat{y})_+^2\,\mathrm{d}x.$$

Here $y(u)$ is the solution to (2.1) and (2.2) for $u \in L^\infty(\Omega)^k$. One example is given by

$$\mathcal{K} = \{u \in L^\infty(\Omega);\ a \le u(x) \le b \text{ a. e. in } \Omega\} \tag{2.23}$$

with some positive constants a, b and $k = 1$.

If $\widehat{y}$ is exact data without errors, that is, $\widehat{y}$ comes from $y(\widehat{u})$ with some $\widehat{u} \in \mathcal{K}$, then $J(\widehat{u}) = 0$. This means that some optimal control $\widehat{u}$ exists. However the uniqueness of minimizers is in general not valid.

Corollary 2.3. *The problem* (2.1), (2.2), (2.21), (2.22) *is a convex control problem.*

Proof.

This is an immediate consequence of Proposition 2.1 and of the fact that the integrand in (2.21) is convex and nondecreasing in y.

Remark. The convexity property makes the application of descent algorithms very efficient. In case (2.21) is replaced by the usual least square approach, the convexity is lost.

Let $J(u^*) = 0$. Then

$$y(u^*)(x) \le \widehat{y}(x) \quad \text{a. e. in} \quad \omega. \tag{2.24}$$

Thus in our optimization problem we are looking for inputs which cause lower outputs than a given target function $\widehat{y}$. For example, if $\widehat{y}$ corresponds to some critical value, then it is practically significant to steer the state keeping a lower level of the critical value.

However, as $\widehat{y}$ may involve measurement errors, it is possible that (2.24) is violated on some subset of ω. This also shows the necessity to provide an existence argument independently of the definition of $\widehat{y}$. In the next section, we study such questions together with the opposite inequality to (2.24).

3 The complete problem

As problem (2.21) - (2.22) is convex, the first order optimality conditions give a characterization of the minimizers.

Theorem 3.1. *The pair* $[y^*, u^*] \in \left[W^{2,s}(\Omega) \cap H_0^1(\Omega)\right] \times \mathcal{K}$ *is optimal for problem* (2.21) - (2.22) *if and only if there exist* $p^* \in W^{2,s}(\Omega) \cap H_0^1(\Omega)$ *satisfying with* y^*, u^* *the first order necessary conditions*

$$-\triangle y^* = \varphi\left(x, y^*, u^*\right) \quad \text{in} \quad \Omega, \tag{3.1}$$

$$-\triangle p^* = \varphi_y'\left(x, y^*, u^*\right) p^* + \left(y^* - \widehat{y}\right)_+ \chi_\omega \quad \text{in} \quad \Omega, \tag{3.2}$$

$$y^* = 0, \quad p^* = 0 \quad \text{on} \quad \partial\Omega, \tag{3.3}$$

$$0 \in p^*(x)\partial\varphi_u\left(x, y^*(x), u^*(x)\right) + \partial I_{\mathcal{K}}\left(u^*(x)\right) \quad \text{a.e.} \quad \Omega. \tag{3.4}$$

Here χ_ω denotes the characteristic function of ω, we set $I_{\mathcal{K}}(u) = 1$ if $u \in \mathcal{K}$ and $= 0$ if $u \notin \mathcal{K}$, $\left[\varphi_y', \partial\varphi_u\right]$ denotes the components of the subdifferential $\partial\varphi$

with respect to $[y,u]$. We recall that φ is assumed to be differentiable in y and by (2.4), $\varphi'_y \in L^s(\Omega)$, $\varphi'_y \leq 0$, $s > \max\left\{2, \frac{d}{2}\right\}$. Then the Dirichlet problem for equation (3.2) has a unique weak solution $p^* \in H^1_0(\Omega)$. As $(y^* - \widehat{y})_+ \chi_\omega$ is positive and bounded, a comparison technique shows that $p^* \in W^{2,s}(\Omega) \cap H^1_0(\Omega)$.

Proof of the sufficiency.

The necessity of conditions (3.1)–(3.4) is well known (see for instance Neittaanmäki, Sprekels and Tiba [6]). Hence we confine ourselves to the proof of the sufficiency of (3.1)–(3.4), based on the convexity property from Proposition 2.1.

We multiply (3.2) by $(y^* - y)$ and (3.4) by $(u^* - w)$ where $[y,w]$ is any admissible pair for problem (2.21) - (2.22). After addition, we obtain:

$$
\begin{aligned}
&-\int_\Omega \varphi'_y(x,y^*,u^*)(y^*-y)\,p^* dx - \int_\Omega p^* \partial\varphi_u(x,y^*,u^*)(u^*-w)\mathrm{d}x \\
&= \int_\Omega \triangle p^*(y^*-y)\,\mathrm{d}x + \int_\Omega (y^*-\widehat{y})_+(y^*-y)\,\chi_\omega \mathrm{d}x \\
&+ \int_\Omega \partial I_K(u^*)(u^*-w)\,\mathrm{d}x \geq \int_\Omega p^*(\triangle y^* - \triangle y)\mathrm{d}x \\
&+\frac{1}{2}\int_\omega (y^*-\widehat{y})_+^2\,\mathrm{d}x - \frac{1}{2}\int_\omega (y-\widehat{y})_+^2\,\mathrm{d}x \\
&= -\int_\Omega p^*\varphi(x,y^*,u^*)\,\mathrm{d}x + \int_\Omega p^*\varphi(x,y,w)\mathrm{d}x \\
&+\frac{1}{2}\int_\omega (y^*-\widehat{y})_+^2\,\mathrm{d}x - \frac{1}{2}\int_\omega (y-\widehat{y})_+^2\,\mathrm{d}x,
\end{aligned}
\tag{3.5}
$$

where we have used (3.1) and the definition of the subdifferential repeatedly. By combining the first and the last parts in (3.5), we conclude that

$$
\begin{aligned}
&\frac{1}{2}\int_\omega (y^*-\widehat{y})_+^2\,\mathrm{d}x - \frac{1}{2}\int_\omega (y-\widehat{y})_+^2\,\mathrm{d}x \\
&\leq \int_\Omega p^*\left[\varphi(x,y^*,u^*) - \varphi(x,y,w)\right]\mathrm{d}x - \int_\Omega \varphi'_y(x,y^*,u^*)(y^*-y)\,p^*\mathrm{d}x \\
&- \int_\Omega \partial\varphi_u(x,y^*,u^*)\,p^*(u^*-w)\,\mathrm{d}x \leq 0.
\end{aligned}
\tag{3.6}
$$

Relation (3.6) is a consequence of the definition of the subdifferential applied to $\varphi(x,\cdot,\cdot)$ and of the positivity of p^* which may be inferred from (3.2). This ends the proof since $[y, w]$ is an arbitrary admissible pair.

The solution triple $[y^*, u^*, p^*]$ of (3.1)–(3.4) is not unique, in general. The set of $[y^*, u^*]$ generated in this way coincides with the set of optimal pairs for (2.21) and (2.22). In many applications it may be enough to solve (3.1)–(3.4) and to by some ad-hoc means select a solution which is appropriate for solving the identification problem associated to (2.1) - (2.2) and the observation $\widehat{y}$. Otherwise, one should study the second step in the proposed reconstruction procedure, formulated as follows

$$\min_{u \in \mathcal{K}} \frac{1}{2} \int_\omega (y - \widehat{y})_-^2 \, \mathrm{d}x \tag{3.7}$$

subject to (3.1)–(3.4). Then we can derive the corresponding relations to (3.1)–(3.4).

Assuming that the observation $\widehat{y}$ is without errors, that is, $\widehat{y} = y(\widehat{u})|_\omega$ with $\widehat{u} \in \mathcal{K}$, then both problems (2.21) and (3.7) have $\widehat{u}$ as solution with minimal value zero (and not necessarily unique). This shows that both (2.21) and (3.7) solve the original identification problem.

In problem (3.7), the optimality conditions (3.1)–(3.3) play the role of the new state system with state variables y and p and control parameter $u \in \mathcal{K}$. Relation (3.4) is a new constraint of mixed type, which is a state-control constraint involving k equalities. As its form is rather unusual, we examplify by the simpler situation when $\varphi(x, y, \cdot)$ is differentiable on $\mathbb{R}^k$ and the indicator function $I_\mathcal{K}$ is regularized in the Yosida–Moreau sense by I_ε, $\varepsilon > 0$. We obtain

$$p(x)\varphi'_u(x, y(x), u(x)) + I'_\varepsilon(u(x)) = 0 \quad \text{a. e. in} \quad \Omega. \tag{3.8}$$

Relation (3.8) has now a usual form of a nonlinear mixed constraint involving the state y, p and the control u with k components. The approximation properties for $\varepsilon \to 0$ are well known in the literature on optimal control, Barbu [1], Barbu and Precupanu [2].

In the sequel we assume that φ'_u is a Caratheodory mapping and satisfies the growth condition

$$|\varphi'_u(x,y,u)| \le (M(x) + C|u|)\,\eta(y). \tag{3.9}$$

Here $C > 0$ is a constant and M, η satisfy the same conditions as in (2.3). We consider constraint (3.4) in the form:

$$0 \in p(x)\varphi'_u(x, y(x), u(x)) + \partial I_{\mathcal{K}}(u(x)). \tag{3.10}$$

Proposition 3.2. *Assume that $\mathcal{K}$ is compact in $L^\infty(\Omega)^k$. Then problem* (3.1)–(3.3), (3.10) *has at least one solution $[y, u, p]$. Therefore minimization problem (3.7) has at least one solution.*

Proof.

We notice that the admissibility property is automatically fulfilled for system (3.1)–(3.3) and (3.10) as it comes from problem (2.1), (2.2), (2.21), (2.22). Let $\{u_n\}$ be a minimizing sequence. Since $\mathcal{K}$ is compact in $L^\infty(\Omega)^k$, we can extract a subsequence denoted again by u_n such that $u_n \to \widetilde{u} \in \mathcal{K}$ strongly in $L^\infty(\Omega)^k$. Multiplying (3.1) by y_n and integrating by parts, we obtain

$$\int_\Omega |\nabla y_n|^2\,\mathrm{d}x = \int_\Omega \varphi(x, y_n, u_n)\, y_n \mathrm{d}x. \tag{3.11}$$

Relation (3.11), (2.3) and (2.4) show that $\{y_n\}$ is bounded in $H^1_0(\Omega)$. Then, standard techniques in the theory of semilinear elliptic equations and Tiba [7], yield that $\{y_n\}$ is bounded in $W^{2,s}(\Omega)$ and relatively compact in $C\left(\overline{\Omega}\right)$ by the Sobolev embedding theorem.

A similar argument applied to (3.2) proves that $\{p_n\}$ is bounded in $W^{2,s}(\Omega)$ and relatively compact in $C\left(\overline{\Omega}\right)$. We may assume that $y_n \to \widetilde{y}$, $p_n \to \widetilde{p}$ weakly in

$W^{2,s}(\Omega)$ and uniformly on $\overline{\Omega}$. The Caratheodory assumptions on φ and φ'_y (see (2.3)–(2.5)) show that $\varphi(\cdot, y_n, u_n) \to \varphi(\cdot, \widetilde{y}, \widetilde{u})$ and $\varphi'_y(\cdot, y_n, u_n) \to \varphi'_y(\cdot, \widetilde{y}, \widetilde{u})$ in $L^s(\Omega)$, by the Lebesgue theorem. Then one can pass to the limit in (3.1)–(3.3).

The Caratheodory assumption and (3.9) give that $\varphi'_u(x, y_n, u_n) \to \varphi'_u(x, \widetilde{y}, \widetilde{u})$ strongly in $L^s(\Omega)^k$. As $p_n \to \widetilde{p}$ uniformly, we obtain that $p_n\varphi'_u(x, y_n, u_n) \to \widetilde{p}\varphi'_u(x, \widetilde{y}, \widetilde{p})$ strongly in $L^s(\Omega)^k$. By (3.10), we see that $\partial I_{\mathcal{K}}(u_n) \to \alpha$ strongly in $L^s(\Omega)^k$. As $u_n \to \widetilde{u}$ uniformly, the closedness of the maximal monotone operator $\partial I_{\mathcal{K}}$ gives that $\alpha \in \partial I_{\mathcal{K}}(\widetilde{u})$ a. e. in Ω.

One can conclude that $[\widetilde{y}, \widetilde{u}, \widetilde{p}]$ is admissible for problem (3.1)–(3.3) and (3.10). As the passage to the limit in the cost functional (3.7) is obvious, the proof is finished.

We continue with the definition of a penalization and regularization of problem (3.1)–(3.3), (3.7), (3.10). The penalized and regularized cost functional is

$$\min\left\{\frac{1}{2}\int_\omega (y-\widehat{y})_-^2 + \frac{1}{2\varepsilon}\int_\Omega \left[p\varphi'_u(x,y,u) + J'_\varepsilon(u)\right]^2 \mathrm{d}x\right\} \tag{3.12}$$

subject to $u \in \mathcal{K}$ and to the state system (3.1)–(3.3). Here J_ε, $\varepsilon > 0$, is a regularization of the Moreau-Yosida approximation of $I_{\mathcal{K}}$ (via a Friedrichs mollifier, for instance). Consequently, J'_ε exists for any $u \in \mathbb{R}^k$.

The formulation (3.12), (3.1)–(3.3) includes the mixed constraint (3.4) in the cost functional and involves just control constraints $u \in \mathcal{K}$. As in Proposition 3.2, we can prove the existence of at least one optimal triple, denoted by $[y_\varepsilon, u_\epsilon, p_\epsilon]$, $\varepsilon > 0$, if $\mathcal{K}$ is compact in $L^\infty(\Omega)^k$.

Proposition 3.3. *When* $\varepsilon \to 0$, *a sequence* $[y_\varepsilon, u_\epsilon, p_\epsilon]$ *converges to* $[y^*, u^*, p^*]$ strongly *in* $W^{2,s}(\Omega) \times L^\infty(\Omega) \times W^{2,s}(\Omega)$ *by taking a subsequence, and* $[y^*, u^*, p^*]$ *is some optimal triple for problem* (3.1)–(3.3), (3.7), (3.10).

Proof.

Any admissible triple $[\widetilde{y}, \widetilde{u}, \widetilde{p}]$ for the problem (3.1)–(3.3), (3.7), (3.10) is admissible for the problem (3.1)–(3.3), (3.12) and we have inequality:

$$
\begin{aligned}
&\frac{1}{2}\int_\omega (y_\varepsilon - \widehat{y})_-^2 \,\mathrm{d}x + \frac{1}{2\varepsilon}\int_\Omega \left[p_\varepsilon \varphi'_u(x, y_\varepsilon, u_\varepsilon) + J'_\varepsilon(u_\varepsilon)\right]^2 \mathrm{d}x \\
&\le \int_\omega \frac{1}{2}(\widetilde{y} - \widehat{y})_-^2 \,\mathrm{d}x + \frac{1}{2\varepsilon}\int_\Omega \left[\widetilde{p}\varphi'_u(x, \widetilde{y}, \widetilde{u}) + J'_\varepsilon(\widetilde{u})\right]^2 \mathrm{d}x \\
&= \frac{1}{2}\int_\omega (\widetilde{y} - \widehat{y})_-^2 \,\mathrm{d}x + \frac{1}{2\varepsilon}\int_\Omega \left[J'_\varepsilon(\widetilde{u}) - \partial I_{\mathcal{K}}(\widetilde{u})\right]^2 \mathrm{d}x.
\end{aligned} \tag{3.13}
$$

Here we choose the element $\widetilde{z} \in \partial I_{\mathcal{K}}(\widetilde{u})$ which occurs in (3.10), that is, $\widetilde{z} = -\widetilde{p}\varphi'_u(x, \widetilde{y}, \widetilde{u})$ a. e. in Ω and J'_ε may be assumed to be a regularization of this precise mapping defined on $\mathcal{K}$. In the case of (2.23), such a construction is simple. By taking a Friedrichs mollification with a very small parameter ε^m, $m \in \mathbb{N}$, (although we preserve the notation J_ε), we may assume that

$$
\begin{aligned}
&\frac{1}{2\varepsilon}\int_\Omega \left[J'_\varepsilon(\widetilde{u}) - \partial I_{\mathcal{K}}(\widetilde{u})\right]^2 \mathrm{d}x \\
&= \frac{1}{2\varepsilon}\int_\Omega \left[J'_\varepsilon(\widetilde{u}) + \widetilde{p}\varphi'_u(x, \widetilde{y}, \widetilde{u})\right]^2 \mathrm{d}x \to 0
\end{aligned} \tag{3.14}
$$

as $\varepsilon \to 0$. By combining (3.13) and (3.14), we obtain

$$
\begin{aligned}
&\frac{1}{2}\int_\omega (y_\varepsilon - \widehat{y})_-^2 \,\mathrm{d}x + \frac{1}{2\varepsilon}\int_\Omega \left[p_\varepsilon \varphi'_u(x, y_\varepsilon, u_\varepsilon) + J'_\varepsilon(u_\varepsilon)\right]^2 \mathrm{d}x \\
&\le \frac{1}{2}\int_\omega (\widetilde{y} - \widehat{y})_-^2 \,\mathrm{d}x + C(\varepsilon, \widetilde{u}),
\end{aligned} \tag{3.15}
$$

where $\lim_{\varepsilon \to 0} C(\varepsilon, \widetilde{u}) = 0$ for any $\widetilde{u} \in \mathcal{K}$, admissible for the problem (3.1)–(3.3), (3.7), (3.10).

As $\mathcal{K} \subset L^\infty(\Omega)^k$ is compact, we may assume that $u_\varepsilon \to \overline{u} \in \mathcal{K}$ strongly in $L^\infty(\Omega)^k$, on a subsequence. The same argument as in the proof of Proposition 3.2 gives that $y_\varepsilon \to \overline{y}$, $p_\varepsilon \to \overline{p}$ strongly in $W^{2,s}(\Omega)$ and uniformly on $\overline{\Omega}$. Then, by (3.9), we obtain that

$$
p_\varepsilon \varphi'_u(x, y_\varepsilon, u_\varepsilon) \to \overline{p}\varphi'_u(x, \overline{y}, \overline{u}) \tag{3.16}
$$

strongly in $L^s(\Omega)$. The construction of J'_ε ensures that $J'_\varepsilon(u_\varepsilon) \to \widehat{z} \in \partial I_{\mathcal{K}}(\overline{u})$ and that (3.10) is satisfied by $[\overline{y}, \overline{u}, \overline{p}]$. This is a consequence of (3.15), (3.16). The triple $[\overline{y}, \overline{u}, \overline{p}]$ is admissible for problem (3.1)–(3.3), (3.7), (3.10). By passing to the limit in (3.15), we obtain

$$\frac{1}{2}\int_\omega (\overline{y} - \widehat{y})_-^2 \,\mathrm{d}x \le \frac{1}{2}\int_\omega (\widetilde{y} - \widehat{y})_-^2 \,\mathrm{d}x$$

for any admissible $[\widetilde{y}, \widetilde{u}, \widetilde{p}]$. This shows the optimality of $[\overline{y}, \overline{u}, \overline{p}]$ which we denote again by $[y^*, u^*, p^*]$ and the proof is finished.

Remark. If we assume that φ'_u, φ'_y are differentiable, then standard gradient methods may be applied for numerical solutions approximating problem (3.1)–(3.3), (3.12). As the nonconvexity appears in the formulation of the problem (3.1)–(3.3), (3.7), (3.10), just local optimal solutions are to be expected in this second step.

Remark. An alternative formulation of the second step in the solution of the identification problem associated to (2.1), (2.2) and the observation $\widehat{y}$ is:

$$\min\left\{\frac{1}{2}\int_\omega (y - \widehat{y})_-^2 \,\mathrm{d}x\right\}$$

subject to (2.1), (2.2), $u \in \mathcal{K}$ and to the state constraint

$$y \le \widehat{y} \quad \text{a. e. in} \quad \omega. \tag{3.17}$$

Based on the results from Section 2, it is easy to see that the set of admissible controls $u \in \mathcal{K}$ (i. e. such that (3.17) is fulfilled) is convex. This is an important advantage. In view of this remark an interesting open question is to be studied under what assumptions quasiconvexity properties are valid in identification problems. However, by penalizing (3.17) one obtains the classical least square approach in identification problems. That is, this is just a reformulation of the usual approach and does not provide a new solution method.

References

[1] V. Barbu (1984) — "Optimal Control of Variational Inequalities", RNM 100, Pitman, London.

[2] V. Barbu and Th. Precupanu (1986) — "Convexity and Optimization in Banach Spaces", D. Reidel, Dordrecht.

[3] B. Kawohl and J. Lang (1997) — "Are some optimal shape problems convex?", J. Convex Analysis

[4] B. Lemaire (1985) — "Application of a subdifferential of a convex composite functional to optimal control in variational inequalities", LNMEMS 255, Springer-Verlag, Berlin, P. 103–118.

[5] W. B. Liu and D. Tiba (2001) — "Error estimates in the approximation of optimization problem governed by nonlinear operators", Numer. Funct. Anal. Optimiz. 22, p. 953–972.

[6] D. Neittaanmäki, J. Sprekels and D. Tiba (2005) — "Optimization of Elliptic Systems, Theory and Applications", Springer-Verlag, New York.

[7] D. Tiba (1990) — "Optimal Control of Nonsmooth Distributed Parameter Systems", Lecture Notes in Mathematics 1459, Springer-Verlag, Berlin.

UTMS

2005–24 Hirotaka Fushiya: *Limit theorem of a one dimensional Marokov process to sticky reflected Brownian motion.*

2005–25 Jin Cheng, Li Peng, and Masahiro Yamamoto: *The conditional stability in line unique continuation for a wave equation and an inverse wave source problem.*

2005–26 M. Choulli and M. Yamamoto: *Some stability estimates in determining sources and coefficients.*

2005–27 Cecilia Cavaterra, Alfredo Lorenzi and Masahiro Yamamoto: *A stability result Via Carleman estimates for an inverse problem related to a hyperbolic integro-differential equation.*

2005–28 Fumio Kikuchi and Hironobu Saito: *Remarks on a posteriori error estimation for finite element solutions.*

2005–29 Yuuki Tadokoro: *A nontrivial algebraic cycle in the Jacobian variety of the Klein quartic.*

2005–30 Hao Fang, Zhiqin Lu and Ken-ichi Yoshikawa: *Analytic Torsion for Calabi-Yau threefolds.*

2005–31 Ken-ichi Yoshikawa: *On the singularity of Quillen metrics.*

2005–32 Ken-ichi Yoshikawa: *Real $K3$ surfaces without real points, equivariant determinant of the Laplacian, and the Borcherds Φ-function.*

2005–33 Taro Asuke: *Infinitesimal derivative of the Bott class and the Schwarzian derivatives.*

2005–34 Dan Tiba, Gengsheng Wang and Masahiro Yamamoto: *Applications of convexity in some identification problems.*

The Graduate School of Mathematical Sciences was established in the University of Tokyo in April, 1992. Formerly there were two departments of mathematics in the University of Tokyo: one in the Faculty of Science and the other in the College of Arts and Sciences. All faculty members of these two departments have moved to the new graduate school, as well as several members of the Department of Pure and Applied Sciences in the College of Arts and Sciences. In January, 1993, the preprint series of the former two departments of mathematics were unified as the Preprint Series of the Graduate School of Mathematical Sciences, The University of Tokyo. For the information about the preprint series, please write to the preprint series office.

ADDRESS:
Graduate School of Mathematical Sciences, The University of Tokyo
3–8–1 Komaba Meguro-ku, Tokyo 153-8914, JAPAN
TEL +81-3-5465-7001 FAX +81-3-5465-7012