SIAM J. CONTROL AND OPTIMIZATION
Vol. 26, No. 4, July 1988

005

SHAPE SENSITIVITY ANALYSIS VIA MIN MAX DIFFERENTIABILITY*

M. C. DELFOUR† AND J.-P. ZOLÉSIO‡

Abstract. The object of this paper is twofold. We introduce a new theorem on the differentiability of a Min Max with respect to a parameter and we show how such a theorem can be applied to compute the material derivative in shape sensitivity analysis problems. We consider the Min Max of a functional which is parametrized by t. We show that, under appropriate conditions, the derivative of the Min Max with respect to t is the Min Max with respect to the points solution of the Min Max problem of the derivative of the original functional with respect to t. To illustrate the use of this theorem, we apply it to the control of an elliptic equation with a nondifferentiable observation and to shape design problems.

Key words. shape sensitivity analysis, minimax, optimal design

AMS(MOS) subject classifications. 35R35, 49A29, 49A52

1. Introduction. Many problems in shape sensitivity analysis can be expressed as a Min Max of some Lagrangian functional which depends on the domain Ω. By introducing a velocity field of deformations V over Ω (cf. Céa [1]-[3], Zolésio [1], [2]), a family of perturbations Ω_t, $t \geqq 0$, of the domain Ω is obtained and the sensitivity analysis reduces to the study of the differentiability of a Min Max functional with respect to the parameter t for fixed velocity fields V and domains Ω.

The object of this paper is twofold. First we present a new theorem on the differentiability of the Min Max:

$$g(t) = \{\text{Min}\,[\text{Max}\; G(t, x, y) \colon y \in \mathcal{B}] \colon x \in \mathcal{A}\}$$

of a functional $G(t, x, y)$ with respect to a real parameter $t \geqq 0$ for some fixed subsets $\mathcal{A} \subset X$ and $\mathcal{B} \subset Y$ of two topological spaces X and Y. We show that, under appropriate hypotheses, the derivative of $g(t)$ with respect to t is the Min Max of the partial derivative $\partial_t G(t, x, y)$ with respect to all points $(x, y) \in \mathcal{A} \times \mathcal{B}$ which are solutions of the Min Max problem. The second objective is to show how such theorems can be applied to shape sensitivity analysis. Thus we give a precise mathematical justification to some results which are usually obtained formally in the literature. For instance, some interesting examples can be found in the book by Haug, Choi, and Komkov [1], Dems and Mroz [1], in the proceedings by Haug and Céa [1], and the recent paper on quick computations by Céa [1].

In order to motivate our approach we shall first consider a special and useful case in § 2 when $G(t, \cdot, \cdot)$ is a convex-concave functional with a unique saddle point (x_t^*, y_t^*) in $\mathcal{A} \times \mathcal{B}$ for each t. We show that under some reasonable hypotheses

$$\frac{dg(t)}{dt} = \partial_t G(t, x_t^*, y_t^*).$$

* Received by the editors November 4, 1986; accepted for publication (in revised form) July 15, 1987. This research was supported in part by Natural Sciences and Engineering Research Council of Canada operating grant A-8730 and a "Fonds pour la formation des chercheurs et l'aide à la recherche" grant from the "Ministère de l'Education du Québec."

† Centre de Recherches Mathématiques and Département de Mathématiques et Statistique, Université de Montréal, Montréal, Québec, Canada H3C 3J7.

‡ Laboratoire de Physique Mathématique, Université des Sciences et Techniques du Languedoc, 34050-Montpellier Cédex, France.

In control theory this is a natural tool in the computation of the directional derivative of the cost function with respect to the control variable when the state is the unique solution of some partial differential equation. In fact this problem can also be expressed as the Min Max of an appropriate Lagrangian. In that context, the existence and uniqueness of the saddle point is equivalent to the well-posedness, existence and uniqueness of the solution to the associated adjoint problem. A very simple nonlinear example will be given in § 3 to illustrate this point.

In § 4 we give the main result where $G(t,\cdot,\cdot)$ is no longer assumed to be a convex-concave functional with saddle points. It is a slight generalization of the earlier result by Delfour and Zolésio [1], [2]. It also extends the work of Zolésio [3, Thm 1.1, p. 1458] on the differentiability of a Min or a Max with respect to a parameter in the shape sensitivity analysis context.

This theorem and its eventual generalizations have many interesting applications. To illustrate that point we describe the associated techniques for two examples. The first, in § 5, is a control or identification problem with a nondifferentiable observation which depends on the state which is the solution of an elliptic equation which itself depends on the control function u. The second example, in § 6, is a shape analysis problem where this technique makes it possible to completely bypass the problem of the existence and interpretation of the Eulerian or material derivative of the state. These two simple examples are given for the purpose of illustration. However, the techniques used here apply to the general linear case and some nonlinear situations (cf. § 7). A first version of the main theorem and its application to the examples of §§ 5 and 6 have been announced in Delfour and Zolésio [1], [2]. They generalize the earlier finite-dimensional result of Dem'yanov [1] for compact subsets $\mathscr{A}$ and $\mathscr{B}$, G continuous in its arguments, and $\partial_t G(t, x, y)$ bounded on $[0, \tau]\times\mathscr{A}\times\mathscr{B}$, $\tau>0$, and continuous with respect to t. Another interesting result has been obtained by Correa and Seeger [1] when the functional $G(t, x, y)$ has saddle points for t in $[0, \tau]$ for some $\tau>0$. It contains as a special case the theorem given in § 2 and it is discussed in § 4 after the presentation of our main result.

Notation. R will denote the field of real numbers, R^+ the subset of positive or zero reals, and R^n ($n\geqq 1$, an integer) the n-fold Cartesian product of R. The inner product and norm in R^n will be defined as

$$x\cdot y=\sum_{l=1,n} x_i y_i, \qquad |x|=(x\cdot x)^{1/2}.$$

The dual operator of a continuous linear operator $A: X\to Y$ will be denoted by A^*. The identity matrix in R^n will be written I_d. The composition of two applications f and g will be denoted by $f\circ g$.

2. Derivative of a Min Max with a unique saddle point with respect to a parameter. Let $\mathscr{A}\subset X$ and $\mathscr{B}\subset Y$ be two nonempty subsets of two topological spaces X and Y, let $\tau>0$ be a real number, and let

$$t, x, y\to G(t, x, y): [0, \tau[\times\mathscr{A}\times\mathscr{B}\to R \tag{1}$$

be a functional which is differentiable with respect to t. Denote by $\partial_t G(t, x, y)$ its partial derivative with respect to t.

THEOREM 1. *Let $\mathscr{A}$, $\mathscr{B}$, and G be given as in* (1) *with G differentiable with respect to t for each $x\in\mathscr{A}$ and $y\in\mathscr{B}$. Assume that for each t in $[0, \tau[$, the functional $G(t,\cdot,\cdot)$ has a unique saddle point (x_t, y_t) and that the following set of hypotheses is verified.*

(HA) *There exist topologies τ_X on X and τ_Y on Y such that*

(i) *the map*

$$t \to (x_t, y_t) : [0, \tau[\to (X, \tau_X) \times (Y, \tau_Y) \tag{2}$$

is continuous (see also Remark 2.2);

(ii) $\forall y \in \mathcal{B}$, $(t, x) \to \partial_t G(t, x, y) : [0, \tau[\times (X, \tau_X) \to R$ *is lower semicontinuous;*

(iii) $\forall x \in \mathcal{A}$, $(t, y) \to \partial_t G(t, x, y) : [0, \tau[\times (Y, \tau_Y) \to R$ *is upper semicontinuous.*

Then the function $g(t) = G(t, x_t, y_t)$ *is differentiable from the right and*

$$dg(t) = \lim_{s \to 0^+} [g(t+s) - g(t)]/s = \partial_t G(t, x_t, y_t). \tag{3}$$

Proof. Write the saddle point conditions at $t+s$ and t for some $s > 0$:

$$G(t+s, x_{t+s}, y) \leqq G(t+s, x_{t+s}, y_{t+s}) \leqq G(t+s, x, y_{t+s}) \tag{4}$$

for all x in $\mathcal{A}$ and y in $\mathcal{B}$, and

$$-G(t, v, y_t) \leqq -G(t, x_t, y_t) \leqq -G(t, x_t, w) \tag{5}$$

for all v in $\mathcal{A}$ and w in $\mathcal{B}$. Denote by ΔG the difference

$$G(t+s, x_{t+s}, y_{t+s}) - G(t, x_t, y_t).$$

Choose

$$v = x_{t+s}, \quad y = y_t, \quad x = x_t, \quad w = y_{t+s}$$

and add (4) and (5). We get

$$G(t+s, x_{t+s}, y_t) - G(t, x_{t+s}, y_t) \leqq \Delta G \leqq G(t+s, x_t, y_{t+s}) - G(t, x_t, y_{t+s}). \tag{6}$$

By the Mean Value Theorem, (6) can be rewritten as

$$\partial_t G(t+\theta_1 s, x_{t+s}, y_t) \leqq \Delta G/s \leqq \partial_t G(t+\theta_2 s, x_t, y_{t+s}) \tag{7}$$

for some θ_1 and θ_2 in $]0, 1[$. By going to the limits as s goes to zero we get

$$\partial_t G(t, x_t, y_t) \leqq \liminf \Delta G/s \leqq \limsup \Delta G/s \leqq \partial_t G(t, x_t, y_t) \tag{8}$$

by using hypothesis (HA). □

Remark 2.1. It is important to emphasize that no differentiability of the map $t \to (x_t, y_t)$ is required to differentiate $g(t)$. So no implicit function theorem is necessary to obtain an equation for the derivatives of x_t and y_t with respect to t. The continuity of the saddle point with respect to t is sufficient.

Remark 2.2. When hypotheses (HA)(ii) and (HA)(iii) are both strengthened from lower and upper semicontinuity to continuity, then g is differentiable for all $t > 0$ in a neighborhood of $t = 0$:

$$\frac{dg(t)}{dt} = \lim_{s \to 0} [g(t+s) - g(t)]/s = \partial_t G(t, x_t, y_t), \qquad t > 0. \tag{3'}$$

This is readily seen if we take s positive or negative in the proof of Theorem 1.

Remark 2.3. In Theorem 1, the continuity hypothesis (HA)(i) can be modified in the following way: there exists $\tau > 0$ such that for all t_0 in $[0, \tau]$:

(i) There exists compact $K_0 \subset X \times Y$ (with respect to the $\tau_X \times \tau_Y$-topology) such that for all $t_n \to t_0$, $t_n > t_0$, there exists a subsequence of $\{t_n\}$, still denoted $\{t_n\}$, such that for all n, $(x_{t_n}, y_{t_n}) \in K_0$;

(ii) For all $y \in \mathcal{B}$, $(t, x) \to G(t, x, y)$ is lower semicontinuous at (t_0, x_{t_0});

(iii) For all $x \in \mathcal{A}$, $(t, y) \to G(t, x, y)$ is upper semicontinuous at (t_0, y_{t_0}).

Under this hypothesis we prove Theorem 1 by using convergent subsequences in K_0 and the following lemma.

LEMMA 1. *Assume that the above three hypotheses are verified. Let $t_n \to t_0$, $t_n > t_0$, and denote by (x^*, y^*) in $X \times Y$ a limit point of (x_{t_n}, y_{t_n}). Then $(x^*, y^*) = (x_{t_0}, y_{t_0})$.* □

In order to make Theorem 1 a computational tool, it is necessary to obtain a characterization of the saddle point. This can obviously be done in several ways and independently. To illustrate that point we give a standard set of hypotheses under which the saddle point is easily characterized.

Assume that

(HB) (i) For all t, $x \to G(t, x, y)$ is convex and Gâteaux differentiable;
(ii) For all t, $y \to G(t, x, y)$ is concave and Gâteaux differentiable.

We quote the following proposition from Ekeland and Temam [1, Props. 1.6, 1.7, pp. 157-158].

PROPOSITION 1. *Assume that $\mathscr{A}$ and $\mathscr{B}$ are convex subsets of two Banach spaces X and Y and that hypothesis* (HB) *is verified. Then for each t, (x_t, y_t) is the unique saddle point of $G(t, \cdot, \cdot)$ if and only if*:

(HC) *There exists a unique solution to the following system of inequalities*:

$$\exists x_t \in \mathscr{A} \quad \forall x \in \mathscr{A}, \quad dG(t, x_t, y_t; x - x_t, 0) \geqq 0, \tag{9}$$

$$\exists y_t \in \mathscr{B} \quad \forall y \in \mathscr{B}, \quad dG(t, x_t, y_t; 0, y - y_t) \leqq 0 \tag{10}$$

where when it exists

$$dG(t, x, y; v, w) = \lim_{s \to 0^+} [G(t, x + sv, y + sw) - G(t, x, y)]/s. \qquad \square$$

So under the hypotheses of Proposition 1 and hypotheses (HB) and (HC), there exists a unique saddle point for each t and under hypothesis (HA) the derivative of $g(t)$ is given by (3).

Remark 2.4. When $\mathscr{A}$ and $\mathscr{B}$ are two linear subspaces of X and Y, respectively, then (9) and (10) are equivalent to

$$\exists x_t \in \mathscr{A} \quad \forall x \in \mathscr{A}, \quad dG(t, x_t, y_t; x, 0) = 0, \tag{9'}$$

$$\exists y_t \in \mathscr{B} \quad \forall y \in \mathscr{B}, \quad dG(t, x_t, y_t; 0, y) = 0. \tag{10'}$$

3. An immediate application to nonlinear control. We have seen at the end of § 2 that under hypotheses (HA), (HB)-(HC), the derivative of the function $g(t)$ is completely characterized by (2.3) provided that system (2.9)-(2.10) has a unique solution.

The above results are readily applicable to classes of nonlinear control problems. Equation (2.9) is usually the state equation which is independent of the y-variable. Equation (2.10) is the adjoint equation and hypothesis (HC) reduces to the well-posedness of that last equation.

So let U, X, and Y be three Banach spaces, U_{ad} a convex subset of U and for each u in U_{ad}

$$A_u : X \to Y' \tag{1}$$

a continuous linear mapping from X into the topological dual Y' of Y. Given f in Im $(A_u) \subset Y'$ and the convex, differentiable functional $F : X \to R$, consider the problem of differentiating the functional

$$J(u) = F(x(u)) \tag{2}$$

where we assume that, for each u in U, $x=x(u)$ is the unique solution in X of the equation

$$A_u x = f. \tag{3}$$

To recast this problem in our framework, introduce the Lagrangian functional

$$\underset{\sim}{G}(u, x, y) = F(x) + \langle A_u x - f, y\rangle_Y \tag{4}$$

where $\langle\ ,\ \rangle_Y$ denotes the duality pairing between Y' and Y. It is well known that the cost function can be rewritten:

$$J(u) = \text{Min}\{\text{Sup}[\underset{\sim}{G}(u, x, y)\colon y\in Y]\colon x\in X\}. \tag{5}$$

We can always formally introduce the adjoint problem:

$$\exists y(u)\in Y \quad \forall x'\in X, \quad dF(x(u); x') + \langle A_u x', y(u)\rangle_Y = 0. \tag{6}$$

Following § 2, assume that for all (x, y) in $X\times Y$, the mapping

$$x\to\langle A_u x, y\rangle_Y \tag{7}$$

has a Gâteaux derivative

$$\langle A'_{u;v}x, y\rangle_Y \tag{8}$$

in all admissible directions v in $U(u\in U_{\text{ad}},\ t>0$ such that $u+tv\in U_{\text{ad}})$. Define

$$G(t, x, y) = \underset{\sim}{G}(u+tv, x, y) \tag{9}$$

and assume that

$$\partial_t G(t, x, y) = \langle A'_{u+tv;v}x, y\rangle \tag{10}$$

verifies hypothesis (HA) for the weak topologies of X and Y.

THEOREM 2. *Assume that problems* (3) *and* (6) *are well posed* (*existence of a unique solution* $(x(u), y(u))\in X\times Y$ *for each* u *in* U_{ad}) *and that the map*

$$u\to(x(u), y(u))\colon U_{\text{ad}}\to(X\text{-weak})\times(Y\text{-weak}) \tag{11}$$

is continuous. Then the cost function $J(u)$ *defined by* (2) *is Gâteaux differentiable and for any admissible direction* v:

$$dJ(u; v) = \langle A'_{u;v}x(u), y(u)\rangle_Y. \tag{12}$$

□

Consider the following illustrative example. Let Ω be a smooth bounded domain in R^N with boundary $\partial\Omega=\Gamma_0\cup\Gamma_1$:

$$X=\{\varphi\in H^1(\Omega)\colon \varphi=0 \text{ on } \Gamma_0\},\quad U=L^\infty(\Gamma_1),\quad U_{\text{ad}}=\{u\in U\colon \exists\alpha,\ u(x)\geqq\alpha>0 \text{ a.e.}\}.$$

To each u in U_{ad}, we associate the boundary value problem

$$\begin{aligned} &-\Delta\varphi(u)=0 \quad \text{in}\,\mathscr{D}'(\Omega), \qquad \varphi(u)\in X,\\ &\partial\varphi(u)/\partial n + u\varphi(u) = g \quad \text{on } \Gamma_1 \end{aligned} \tag{13}$$

for a fixed g in $L^2(\Gamma_1)$.

We can associate with problem (13) the minimization over X of the energy functional

$$E(u, \varphi) = \frac{1}{2}\int_\Omega |\nabla\varphi|^2\,dx + \int_{\Gamma_1}\left[\left(\frac{u}{2}\right)\varphi^2 - g\varphi\right]d\Gamma. \tag{14}$$

For each $x(u)$ define the cost function

$$J(u)=\frac{1}{2}\int_{\Gamma_1}\left[\frac{\partial\varphi(u)}{\partial n}\right]^2 d\Gamma=\frac{1}{2}\int_{\Gamma_1}[g-u\varphi(u)]^2\,d\Gamma \tag{15}$$

and the Lagrangian functional $G:[0,\tau]\times X\times X\to R$,

$$\begin{aligned} G(t,\varphi,\psi)&=F(\varphi)+dE(u+tv,\varphi;\psi)\\ &=\frac{1}{2}\int_{\Gamma_1}[g-(u+tv)\varphi]^2\,d\Gamma+\int_\Omega \nabla\varphi\cdot\nabla\psi\,dx\\ &\quad+\int_{\Gamma_1}[(u+tv)\varphi-g]\psi\,d\Gamma. \end{aligned} \tag{16}$$

Apply Theorem 2 to get

$$\begin{aligned} dJ(u;v)&=\partial_t G(0,\varphi(u),\psi(u))\\ &=\int_{\Gamma_1}[-g+u\varphi(u)+\psi(u)]\varphi(u)v\,d\Gamma \end{aligned} \tag{17}$$

where $\varphi(u)$ is the solution of (13) and $\psi=\psi(u)\in X$ is characterized by

$$dG(0,\varphi(u),\psi;\varphi,0)=0\quad\forall\varphi\in X \tag{18}$$

or

$$\begin{aligned} &=\int_{\Gamma_1}-u[g-u\varphi(u)]\varphi\,d\Gamma+\int_\Omega\nabla\varphi\cdot\nabla\psi\,dx+\int_{\Gamma_1}u\varphi\psi\,d\Gamma\\ &=0\quad\forall\varphi\in X. \end{aligned} \tag{19}$$

But this is equivalent to

$$\begin{aligned} &\Delta\psi=0\quad\text{in }\mathscr{D}'(\Omega),\\ &\partial\psi/\partial n+u\psi=u\partial\varphi(u)/\partial n(=u(g-u\varphi)\in L^2(\Gamma_1)) \end{aligned} \tag{20}$$

which is a well-posed problem. To completely justify (17) it suffices to check the weak continuity of $t\to\varphi(u+tv)$ and $\psi(u+tv)$ which is easily done.

4. Derivative of a Min Max with respect to a parameter. Let $\mathscr{A}\subset X$ and $\mathscr{B}\subset Y$ be subsets of two topological spaces X and Y and let $\tau>0$ be a real number. Given a map

$$G:[0,\tau]\times X\times Y\to R,$$

we consider the following functions:

$$H(t,x)=\operatorname{Sup}\{G(t,x,y):y\in\mathscr{B}\},\qquad t\in[0,\tau],\quad x\in\mathscr{A}, \tag{1}$$

$$g(t)=\operatorname{Inf}\{H(t,x):x\in\mathscr{A}\},\qquad t\in[0,\tau]. \tag{2}$$

As a result

$$g(t)=\operatorname{Inf}\{\operatorname{Sup}[G(t,x,y):y\in\mathscr{B}]:x\in\mathscr{A}\}. \tag{3}$$

We wish to show that under appropriate hypotheses the function g is differentiable at $t=0$ from the right:

$$\lim_{t\to 0^+}(g(t)-g(0))/t\quad\text{exists.} \tag{4}$$

For $t \geqq 0$ we shall need the set

(5) $$A(t) = \{x \in \mathcal{A} : g(t) = H(t, x)\}$$

and for x in $\mathcal{A}$ the set

(6) $$B(t, x) = \{y \in \mathcal{B} : H(t, x) = G(t, x, y)\}.$$

In order to better see the role of each hypothesis in the final result, we proceed in a step-by-step fashion. We first introduce hypotheses to ensure that the Sup and Inf problems have solutions.

(H1) There exists $\tau > 0$, for all t, $0 \leqq t \leqq \tau$:
(i) $A(0) \neq \emptyset$, for all $x_0 \in A(0)$, $B(t, x_0) \neq \emptyset$,
(ii) $A(t) \neq \emptyset$, for all $x_t \in A(t)$, $B(0, x_t) \neq \emptyset$.

LEMMA 2. *Under hypothesis* (H1) *we have the following estimates: for all t*, $0 \leqq t \leqq \tau$

(7) $$g(t) - g(0) \leqq G(t, x_0, y^*) - G(0, x_0, y^*) \quad \forall x_0 \in A(0) \quad \forall y^* \in B(t, x_0)$$

and

(8) $$g(t) - g(0) \geqq G(t, x_t, z^*) - G(0, x_t, z^*) \quad \forall x_t \in A(t) \quad \forall z^* \in B(0, x_t).$$

Proof. The proof uses standard arguments and will be omitted. □

(9) In a second step we obtain upper and lower bounds on the differential quotient

$$\frac{(g(t) - g(0))}{t}.$$

We need the following additional hypothesis:

(H2) (i) For all $x_0 \in A(0)$, the function

(10) $$s \to G(s, x_0, y)$$

is differentiable in a neighborhood of $t = 0$ for all y in

$$\bigcup \{B(t, x_0) : 0 \leqq t \leqq \tau\};$$

(ii) For all t, $0 \leqq t \leqq \tau$, for all $x_t \in A(t)$, the function

$$s \to G(s, x_t, y)$$

is differentiable in a neighborhood of $t = 0$ for all y in

$$\bigcup \{B(0, x_t) : 0 \leqq t \leqq \tau\}.$$

LEMMA 3. *Under hypotheses* (H1) *and* (H2), *for each t*, $0 < t \leqq \tau$:
(i) *There exists* θ_1, $0 < \theta_1 < 1$, *such that*

(11) $$\frac{(g(t) - g(0))}{t} \leqq \partial_t G(\theta_1 t, x_0, y^*) \quad \forall x_0 \in A(0) \quad \forall y^* \in B(t, x_0);$$

(ii) *There exists* θ_2, $0 < \theta_2 < 1$, *such that*

(12) $$\frac{(g(t) - g(0))}{t} \geqq \partial_t G(\theta_2 t, x_t, z^*) \quad \forall x_t \in A(t) \quad \forall z^* \in B(0, x_t).$$

Proof. (i) For $x_0 \in A(0)$ and $y^* \in B(t, x_0)$ define

$$\lambda(s) = G(st, x_0, y^*).$$

By hypothesis (H2), λ is differentiable in a neighborhood of 0. So by Taylor's Theorem there exists $\theta_1 \in]0, 1[$ such that

$$\lambda(1) = \lambda(0) + \frac{d\lambda}{ds}(\theta_1)$$

and

$$G(t, x_0, y^*) - G(0, x_0, y^*) = t\partial_t G(\theta_1 t, x_0, y^*)$$

where $\partial_t G$ denotes the partial derivative of G with respect to the first argument. The proof of part (ii) is similar and will be omitted. □

In the next step we go to the limit in (11) and (12) as t goes to 0. So we introduce

$$\text{(13)} \qquad \bar{d}g(0) = \limsup_{t \to 0^+} \frac{(g(t) - g(0))}{t}, \qquad \underline{d}g(0) = \liminf_{t \to 0^+} \frac{(g(t) - g(0))}{t}.$$

They are the smallest upper and greatest lower bounds of the differential quotient in R. So there exist sequences $\{t_n\}$ and $\{t'_n\}$ of positive numbers in $]0, \tau]$ going to zero as n goes to $+\infty$ such that

$$\text{(14)} \qquad \bar{d}g(0) = \lim_{n \to \infty} \frac{(g(t_n) - g(0))}{t_n},$$

$$\text{(15)} \qquad \underline{d}g(0) = \lim_{n \to \infty} \frac{(g(t'_n) - g(0))}{t'_n}.$$

We first consider the upper bound in (11). We use the following hypotheses of continuity.

(H3) There exists a topology τ_Y on Y such that for all $x_0 \in A(0)$:
- (i) For all sequences $t_n \to 0$, $t_n > 0$, there exists $y_0 \in B(0, x_0)$ and a subsequence of $\{t_n\}$, still denoted $\{t_n\}$, such that for all n, there exists $y_n \in B(t_n, x_0)$, and $y_n \to y_0$ in the τ_Y-topology;
- (ii) The map

$$(t, y) \to \partial_t G(t, x_0, y)$$

is upper semicontinuous in $\{0\} \times \cup \{B(t, x_0): 0 \leqq t \leqq \tau\}$ in the τ_Y-topology.

PROPOSITION 2. *Under hypotheses* (H1)(i), (H2)(i), and (H3)

$$\text{(16)} \qquad \bar{d}g(0) \leqq \inf_{x \in A(0)} \operatorname*{Sup}_{y \in B(0,x)} \partial_t G(0, x, y).$$

Proof. Fix x_0 in $A(0)$ and let $\{t_n\}$ be the sequence in (14). Then by hypothesis (H3)(i), there exists $y_0 \in B(0, x_0)$, there exists a subsequence of $\{t_n\}$, still denoted $\{t_n\}$, such that there exists $y_n \in B(t_n, x_0)$, for all n, and

$$y_n \to y_0 \quad \text{in } \tau_Y\text{-topology}.$$

From (11), using hypothesis (H3)(ii), we obtain

$$\bar{d}g(0) \leqq \limsup_{n \to \infty} \partial_t G(\theta_1 t_n, x_0, y_n) \leqq \partial_t G((0, x_0, y_0)).$$

As a result

$$\bar{d}g(0) \leqq \sup \{\partial_t G(0, x_0, y): y \in B(0, x_0)\}.$$

The last estimate is true for all x_0 in $A(0)$. This is sufficient to establish (16). □

Remark 4.1. (H3)(i) is the Kuratovsky hypothesis at $t=0$ for the set-valued function $t \Rightarrow B(t, x_0)$.

We now turn to the lower bound (12). We formulate the following hypotheses.

(H4) There exist topologies τ_X on X and τ_Y on Y such that:

(i) For all sequences $t_n \to 0$, $t_n > 0$, there exists $x_0 \in A(0)$, for all $y_0 \in B(0, x_0)$, there exists a subsequence of $\{t_n\}$, still denoted $\{t_n\}$, such that for all n, there exists $x_n \in A(t_n)$ and there exists $z_n \in B(0, x_n)$, such that $x_n \to x_0$ in the τ_X-topology and $z_n \to y_0$ in the τ_Y-topology;

(ii) The map $(t, x, y) \to \partial_t G(t, x, y)$ is lower semicontinuous in $\{0\} \times \{(x, y): x \in A(0), y \in B(0, x)\}$ in the $\tau_X \times \tau_Y$-topology.

Remark 4.2. Hypothesis (H4)(i) is verified when the following two hypotheses are verified:

(H4)(i_1) There exists a topology τ_X on X such that for all sequences $t_n \to 0$, $t_n > 0$, there exists $x_0 \in A(0)$, there exists a subsequence of $\{t_n\}$, still denoted $\{t_n\}$, and for all n, there exists $x_n \in A(t_n)$ such that $x_n \to x_0$ in the τ_X-topology.

(H4)(i_2) There exists a topology τ_Y of Y for which the set-valued function $x \Rightarrow B(0, x)$ is lower semicontinuous on $A(0)$ in the sense of Aubin [1, Déf. 9.4, p. 121]: for all convergent sequences $x_n \to x_0$ in X and all z^* in $B(0, x_0)$, there exists a sequence $z_n^* \in B(0, x_n)$ such that $z_n^* \to z^*$ in the τ_Y-topology.

Hypothesis (H4(i_1)) is the Kuratovsky condition at $t=0$ for the set-valued map $t \Rightarrow A(t)$.

We state the analogue of Proposition 2 in the other direction.

PROPOSITION 3. *Under hypotheses* (H1)(ii), (H2)(ii), *and* (H4), *we have*

$$\underline{d}g(0) \geqq \operatorname*{Inf}_{x \in A(0)} \operatorname*{Sup}_{y \in B(0, x)} \partial_t G(0, x, y). \tag{17}$$

Proof. Consider expression (15) and the converging sequence $t'_n \to o^+$, $t'_n > 0$. By (H4)(i), there exist topologies τ_X on X and τ_Y on Y such that there exists $x_0 \in A(0)$, for all $y_0 \in B(0, x_0)$, there exists a subsequence of $\{t'_n\}$, still denoted $\{t'_n\}$, and for all n, there exists $x_n \in A(t'_n)$ and there exists $z_n \in B(0, x_n)$, such that

$$x_n \to x_0 \text{ in the } \tau_X\text{-topology} \quad \text{and} \quad z_n \to y_0 \text{ in the } \tau_Y\text{-topology}.$$

Now from (12)

$$[g(t'_n) - g(0)]/t'_n \geqq \partial_t G(\theta_2 t'_n, x_n, z_n)$$

and, in view of (H4)(ii)

$$\underline{d}g(0) \geqq \liminf_{n \to \infty} \partial_t G(\theta_2 t_n, x_n, z_n) \geqq \partial_t G(0, x_0, y_0)$$

for some $x_0 \in A(0)$ and all $y_0 \in B(0, x_0)$. Finally

$$\underline{d}g(0) \geqq \operatorname*{Sup}_{y_0 \in B(0, x_0)} \partial_t G(0, x_0, y_0) \geqq \operatorname*{Inf}_{x_0 \in A(0)} \operatorname*{Sup}_{y_0 \in B(0, x_0)} \partial_t G(0, x_0, y_0). \qquad \square$$

We now state our main result.

THEOREM 3. *Under hypotheses* (H1)–(H4), *we have*

$$\underline{d}g(0) = \bar{d}g(0) = \operatorname*{Inf}_{x \in A(0)} \operatorname*{Sup}_{y \in B(0(0, x)} \partial_t G(0, x, y) \tag{18}$$

and the function g is differentiable at 0 *from the right*:

$$\lim_{t \to o^+} \frac{(g(t) - g(0))}{t} \quad \textit{exists}. \tag{19}$$

Proof. The proof follows from Propositions 2 and 3. $\square$

The following propositions give sets of sufficient conditions to verify (H3)(i) and (H4)(i_1). They are the conditions given in Delfour and Zolésio [1], [2] in their initial version of Theorem 3.

PROPOSITION 4. *If for each* x_0 *in* $A(0)$:

(a) *There exists a topology* τ_Y *on* Y *and a sequentially compact subset* K_Y *of* Y *such that for all sequences*

$$t_n \to 0, \quad t_n > 0, \quad B(t_n, x_0) \cap K_Y \neq \emptyset;$$

(b) *For all* y *in* $\mathcal{B}$, *the map* $t \to G(t, x_0, y)$ *is lower semicontinuous*;

(c) *The map* $t, y \to G(t, x_0, y)$ *is upper semicontinuous in* $\{0\} \times \bigcup \{B(t, x_0): 0 \leqq t \leqq \tau\}$ *in the* τ_Y*-topology.*

Then hypothesis (H3)(i) *is verified.*

Proof. Fix an arbitrary x_0 in $A(0)$ and let $t_n \to 0$, $t_n > 0$, be an arbitrary converging sequence. By (a) we can choose for each n

$$y_n \in B(t_n, x_0) \cap K_Y.$$

But since K_Y is sequentially compact, there exists a subsequence of $\{y_n\}$, still denoted $\{y_n\}$, such that

$$y_n \to y^* \quad \text{in the } \tau_Y\text{-topology}.$$

By definition of $B(t_n, x_0)$,

$$G(t_n, x_0, y_n) \geqq G(t_n, x_0, y) \quad \forall y \in \mathcal{B}$$

and from (c)

$$G(0, x_0, y^*) \geqq \limsup_{n\to\infty} G(t_n, x_0, y_n) \geqq \limsup_{n\to\infty} G(t_n, x_0, y) \quad \forall y \in \mathcal{B}.$$

But from (b)

$$\limsup_{n\to\infty} G(t_n, x_0, y) \geqq \liminf_{n\to\infty} G(t_n, x_0, y) \geqq G(0, x_0, y) \quad \forall y \in \mathcal{B}.$$

By combining the two inequalities above we conclude that $y^* \in B(0, x_0)$. Thus hypothesis (H3)(i) is verified. □

PROPOSITION 5. *Assume that the following hypotheses are verified*:

(a) *There exists a topology* τ_X *on* X *and a sequentially compact subset* K *of* X *such that* $A(t) \cap K \neq \emptyset$, *for all* t, $0 \leqq t \leqq \tau$;

(b) *For all* x *in* $\mathcal{A}$ *the map* $t \to H(t, x)$ *is upper semicontinuous at* $t = 0$;

(c) *The map* $(t, x) \to H(t, x)$ *is lower semicontinuous in* $\{0\} \times \bigcup \{A(t): 0 \leqq t \leqq \tau\}$ *in the* τ_X*-topology.*

Then hypothesis (H4)(i_1) *is verified.*

Proof. The proof is similar to that of the previous proposition. □

Remark 4.3. In order to obtain the lower bound on $\underline{d}g(0)$, we have used (H4)(i). This hypothesis is to be compared with its counterpart (H3(i)) for the upper bound. In the first we have for all $x_0 \in A(0)$, there exists $y_0 \in B(0, x_0)$ while in the second there exists $x_0 \in A(0)$, for all $y_0 \in B(0, x_0)$. If in (H4)(i) the statement is weakened to: there exists $x_0 \in A(0)$ and there exists $y_0 \in B(0, x_0)$, then the upper and lower bounds on the differential quotient will no longer coincide. This suggests the following weaker form of hypothesis (H4)(i).

(H4)(i)-weak There exist topologies τ_X on X and τ_Y on Y such that for all sequences $t_n \to 0$, $t_n > 0$, there exists $x_0 \in A(0)$, there exists $y_0 \in B(0, x_0)$ and there exists a subsequence of $\{t_n\}$, still denoted $\{t_n\}$, such that for all n, there exists $x_n \in A(t_n)$ and there exists $z_n \in B(0, x_n)$, such that $x_n \to x_0$ in the τ_X-topology and $z_n \to y_0$ in the τ_Y-topology.

THEOREM 4. *Under hypotheses* (H1)–(H3), (H4)(i)-weak *and* (H4)(ii)

$$\inf_{x \in A(0)} \inf_{y \in B(0,x)} \partial_t G(0, x, y) \leqq \underline{d}g(0), \tag{20}$$

$$\bar{d}g(0) \leqq \inf_{x \in A(0)} \sup_{y \in B(0,x)} \partial_t G(0, x, y). \tag{21}$$

COROLLARY. *If, in addition to the hypotheses of Theorem* 4, *the set* $B(0, x)$ *is a singleton for each* x *in* $A(0)$,

$$\forall x \in A(0) \quad B(0, x) = \{y_0(x)\}, \tag{22}$$

then g *is differentiable at* 0 *from the right and*

$$dg(0) = \inf \{\partial_t G(0, x, y_0(x)) : x \in A(0)\}. \tag{23}$$

Remark 4.4. The corollary can also be proved directly by two consecutive applications of the theorem on the differentiability of a Min.

Remark 4.5. A sufficient set of conditions to verify (H4(i)-weak) is given by $(H4)(i_1)$ and a Kuratovsky condition on the set-valued map $x \Rightarrow B(0, x)$:

$(H4)(i_1)$ There exists a topology τ_X on X such that for all sequences $t_n \to 0$, $t_n > 0$, there exists $x_0 \in A(0)$, there exists a subsequence of $\{t_n\}$, still denoted $\{t_n\}$, and for all n, there exists $x_n \in A(t_n)$ such that $x_n \to x_0$ in the τ_X-topology.

$(H4)(i_2)$-weak There exists a topology τ_Y on Y such that for all $x_n \to x_0$ in the τ_X-topology, there exists a subsequence of $\{x_n\}$, still denoted $\{x_n\}$, $z^* \in B(0, x_0)$, for all n, there exists $z_n \in B(0, x_n)$ such that $z_n \to z^*$ for the τ_Y-topology.

Remark 4.6. A set of sufficient conditions to verify $(H4)(i_2)$-weak is given in Delfour and Zolésio [1], [2]:

(a) Given a converging sequence $x_n \to x_0$ in X, there exists a subsequence of $\{x_n\}$, still denoted $\{x_n\}$, there exists $z^* \in Y$ and there exists $z_n \in B(0, x_n)$, for all n, such that $z_n \to z^*$ in the τ_Y-topology;

(b) For all z in $\mathscr{B}$, the map $x \to G(0, x, z)$ is lower semicontinuous on $\mathscr{A}$ and the map $x, z \to G(0, x, z)$ is upper semicontinuous on $\mathscr{A} \times \mathscr{B}$.

In order to complete this section we quote a recent result by Correa and Seeger [1] which assumes the existence of a saddle point of the functional G. First introduce the function

$$h(t) = \sup \{\inf [G(t, x, y) : x \in \mathscr{A}] : y \in \mathscr{B}\}, \tag{24}$$

the associated sets

$$B(t) = \{\underline{y} \in \mathscr{B} : h(t) = \inf [G(t, x, \underline{y}) : x \in \mathscr{A}]\}, \tag{25}$$

$$A(t, y) = \{\underline{x} \in \mathscr{A} : \inf [G(t, x, y) : x \in \mathscr{A}] = G(t, \underline{x}, y)\}, \qquad y \in Y \tag{26}$$

and the set of saddle points

$$S(t) = \{(x_t, y_t) \in \mathscr{A} \times \mathscr{B} : g(t) = G(t, x_t, y_t) = h(t)\}. \tag{27}$$

Then the following lemma is immediate.

LEMMA 4. *If* $S(t) \neq \emptyset$ *for some* $t \geqq 0$, *then*

$$S(t) = A(t) \times B(t), \quad A(t) \neq \emptyset, \quad B(t) \neq \emptyset, \tag{28}$$

$$\forall x_t \in A(t) \quad B(t, x_t) = B(t) \quad \text{and} \quad \forall y_t \in B(t) \quad A(t, y_t) = A(t). \tag{29}$$ □

THEOREM 5 (Correa and Seeger [1]). *Assume that there exists $\tau > 0$ such that the following hypotheses are verified*:

(HH1) $S(t) \neq \emptyset$, $0 \leqq t \leqq \tau$.

(HH2) *For all* (x, y) *in* $\cup\{A(t): 0 \leqq t \leqq \tau\} \times \cup\{B(t): 0 \leqq t \leqq \tau\}$, *the map* $t \to G(t, x, y)$ *is differentiable everywhere in* $[0, \tau]$.

(HH3) *There exists a topology* τ_X *on* X *such that*
(i) *For all* $t_n \to 0$, $0 \leqq t_n \leqq \tau$, *there exists* $x_0 \in A(0)$, *there exists a subsequence of* $\{t_n\}$, *still denoted* $\{t_n\}$, *and for all* n, *there exists* $x_n \in A(t_n)$, *such that* $x_n \to x_0$ *in the* τ_X-*topology*;
(ii) *For all* $y \in \cup\{B(t): 0 \leqq t \leqq \tau\}$,

$$(t, x) \to \partial_t G(t, x, y)$$

is lower semicontinuous at $\{0\} \times A(0)$ *for the* τ_X-*topology.*

(HH4) *There exists a topology* τ_Y *on* Y *such that*
(i) *For all* $t_n \to 0$, $0 \leqq t_n \leqq \tau$, *there exists* $y_0 \in B(0)$, *there exists a subsequence of* $\{t_n\}$, *still denoted* $\{t_n\}$, *and for all* n, *there exists* $y_n \in B(t_n)$, *such that* $y_n \to y_0$ *in the* τ_Y-*topology*;
(ii) *For all* $x \in \cup\{A(t): 0 \leqq t \leqq \tau\}$,

$$(t, y) \to \partial_t G(t, x, y)$$

is upper semicontinuous at $\{0\} \times B(0)$ *for the* τ_Y-*topology.*

Then

$$dg(0) = \lim_{t \to 0^+} (g(t) - g(0))/t = \operatorname*{Inf}_{x \in A(0)} \operatorname*{Sup}_{y \in B(0)} \partial_t G(0, x, t) \tag{30}$$

$$= \operatorname*{Sup}_{x \in B(0)} \operatorname*{Inf}_{y \in A(0)} \partial_t G(0, x, y). \tag{31}$$

This theorem contains as a special case Theorem 1 in § 2 and can be proved by the same technique. The hypotheses are essentially those given in Correa and Seeger [1], except for the fact that we have used Kuratovsky's conditions (HH3)(i) and (HH4)(i) for the set-valued maps instead of the notion of *sequential semicontinuity* for a set-valued map $t \Rightarrow M(t)$ from $[0, \tau]$ to X:

> There exists a topology τ_X on X such that for all sequences $t_n \to 0$, $t_n > 0$, there exists $x_0 \in M(0)$ and for all n, there exists $x_n \in M(t_n)$ such that $x_n \to x_0$ in the τ_X-topology.

5. Derivative of a nondifferentiable observation functional with respect to the control variable. Let Ω be a bounded domain in R^n with smooth boundary Γ, $f \in L^2(\Omega)$ and u be a function on the subset U_{ad} of $U = L^\infty(\Omega)$ defined as

$$U_{\text{ad}} = \{u \in L^\infty(\Omega): \exists \alpha > 0 \text{ such that } u(x) \geqq \alpha \text{ a.e. in } \Omega\}. \tag{1}$$

Consider the solution $y = y(u)$ in $H_0^1(\Omega)$ of the variational problem

$$-\operatorname{div}(u \nabla y) = f \quad \text{in } \Omega, \qquad y = 0 \quad \text{on } \Gamma. \tag{2}$$

Associate with u and y the cost function

$$J(u) = \int_\Omega |y - y_d| \, dx, \qquad y_d \in L^1(\Omega). \tag{3}$$

We want to compute the derivative of $J(u)$ with respect to u subject to (2).

We consider the state equation (2) as a constraint and remove it by introducing a Min Sup. It is easy to check that

$$J(u)=\underset{\varphi\in H_0^1(\Omega)}{\mathrm{Min}}\ \underset{(\mu,\psi)\in M\times H_0^1(\Omega)}{\mathrm{Sup}}\left[\int_\Omega \mu(\varphi-y_d)\,dx+dE(u,\varphi;0,\psi)\right] \tag{4}$$

where $dE(u,\varphi;0,\psi)$ is the Gâteaux derivative of

$$E(u,\varphi)=\frac{1}{2}\int_\Omega [u|\nabla\varphi|^2-2f\varphi]\,dx \tag{5}$$

at (u,φ) in the direction $(0,\psi)$ and

$$M=\{\mu\in L^\infty(\Omega): |\mu(x)|\leqq 1, \text{a.e. in } \Omega\}. \tag{6}$$

In this form, it is not directly possible to apply Theorem 3 in § 4. It is necessary to introduce a perturbed functional indexed by a parameter $r>0$ (which is not necessarily infinitesimally small):

$$-G_r(u,(\mu,\psi),\varphi)=\int_\Omega \mu(\varphi-y_d)\,dx+dE(u,\varphi;0,\psi)+r\{E(u,\varphi)-e(u)\} \tag{7}$$

where

$$e(u)=\mathrm{Inf}\,\{E(u,\varphi): \varphi\in H_0^1(\Omega)\}. \tag{8}$$

Define

$$J_r(u)=\underset{\varphi\in H_0^1(\Omega)}{\mathrm{Min}}\ \underset{(\mu,\psi)\in M\times H_0^1(\Omega)}{\mathrm{Sup}}\ -G_r(u,(\psi,\mu),\varphi) \tag{9}$$

and the dual quantity

$$J_r^*(u)=-\underset{(\mu,\psi)\in M\times H_0^1(\Omega)}{\mathrm{Inf}}\ \underset{\varphi\in H_0^1(\Omega)}{\mathrm{Sup}}\ G_r(u,(\psi,\mu),\varphi). \tag{10}$$

Theorem 3 will be applied to J_r^* with $r>0$. Prior to doing this we show that G_r has saddle points and characterize the associated sets.

PROPOSITION 6. *For each u in U and r, $0<r<2$, the functional $G_r(u,\cdot,\cdot)$ has saddle points and*

$$J_r^*(u)=J_r(u)=-\underset{(\mu,\psi)\in M\times H_0^1(\Omega)}{\mathrm{Min}}\ \underset{\varphi\in H_0^1(\Omega)}{\mathrm{Max}}\ G_r(u,(\mu,\psi),\varphi). \tag{11}$$

Proof. (i) The first part of identity (11) follows from Ekeland and Temam [1, Prop. 2.4, p. 177] applied to the functional

$$F_r(u,\psi,\varphi)=\mathrm{Sup}\,\{-G_r(u,(\mu,\psi),\varphi): \mu\in M\} \tag{12}$$

which is equal to

$$\int_\Omega \{|\varphi-y_d|\,dx+dE(u,\varphi;0,\psi)+r[E(u,\varphi)-e(u)]. \tag{13}$$

It suffices to check the following two conditions:

$$\exists p\in H_0^1(\Omega),\quad \text{such that } \lim_{\|\varphi\|\to\infty} F_r(u,p,\varphi)=+\infty, \tag{14}$$

$$\lim_{\|p\|\to\infty}\ \underset{\varphi\in H_0^1(\Omega)}{\mathrm{Inf}}\ F_r(u,p,\varphi)=-\infty. \tag{15}$$

The first condition is verified for $p=0$. For the second condition, we fix p and choose $\varphi=-p$

$$\operatorname{Inf}\{F_r(u,p,\varphi):\varphi\in H_0^1(\Omega)\}\leqq F_r(u,p,-p)$$

and show that the upper bound goes to $-\infty$ as $\|p\|$ goes to $+\infty$:

$$F_r(u,p,-p)=\int_\Omega\{|-p-y_d|-u|\nabla p|^2-fp+r/2(u|\nabla p|^2+2fp)\}\,dx-re(u). \tag{16}$$

The L^2-norm of ∇p goes to $+\infty$ since it is equivalent to the $H_0^1(\Omega)$-norm. So for r, $0<r<2$, the right-hand side of (16) goes to $-\infty$ and (15) is verified. This shows the existence of a saddle point for $F_r(u,\cdot,\cdot)$:

$$\operatorname*{Min}_{\varphi}\operatorname*{Sup}_{\psi}F_r(u,\psi,\varphi)=\operatorname*{Max}_{\psi}\operatorname*{Inf}_{\varphi}F_r(u,\psi,\varphi). \tag{17}$$

(ii) The next step is to show that for a fixed p,

$$\operatorname*{Inf}_{\varphi}\operatorname*{Sup}_{\mu\in M}-G_r(u,\mu,p,\varphi)=\operatorname*{Max}_{\mu\in M}\operatorname*{Inf}_{\varphi}-G_r(u,\mu,p,\varphi). \tag{18}$$

In view of the properties of $-G_r$ and the fact that M is bounded, this is a consequence of Remark 2.3 and Proposition 2.3 in Ekeland and Temam [1, p. 162]. If we combine (17) and (18)

$$\operatorname*{Min}_{\varphi}\operatorname*{Sup}_{(\mu,\psi)}-G_r=\operatorname*{Max}_{(\mu,\psi)}\operatorname*{Inf}_{\varphi}-G_r,$$

and by Proposition 1.2 in Ekeland and Temam [1, p. 155], $-G_r(u,(\cdot,\cdot),\cdot)$ has saddle points. In view of (10), this is sufficient to establish (11). □

It is now important to note that for all $r\geqq 0$

$$J_r(u)=J_0(u)=J(u). \tag{19}$$

We have shown that for $0<r<2$, $G_r(u,\cdot,\cdot)$ has saddle points and that

$$J_r(u)=J_r^*(u). \tag{20}$$

For $u\in U_{\mathrm{ad}}$ and $v\in U=L^\infty(\Omega)$, there exists $\tau>0$ small enough such that $u+\tau v\in U_{\mathrm{ad}}$. Define for t in $[0,\tau]$

$$G(t,q,\varphi)=G_r(u+tv,q,\varphi) \tag{21}$$

for $q=(\mu,\psi)\in X=M\times H_0^1(\Omega)$ and $\varphi\in Y=H_0^1(\Omega)$. In view of the above proposition, the saddle points of $G(t,\cdot,\cdot)$ are completely characterized by the following set of equations (cf. Ekeland and Temam [1, Prop. 1.6, p. 157]):

$$-\operatorname{div}[(u+tv)\nabla y_t]=f\quad\text{in }\Omega,\qquad y_t=0\quad\text{on }\Gamma, \tag{22}$$

$$-\operatorname{div}[(u+tv)\nabla p_t]+\mu_t=0\quad\text{in }\Omega,\qquad p_t=0\quad\text{on }\Gamma, \tag{23}$$

$$\mu_t\in M_{dt}=\{\operatorname{sgn}(y_t-y_d)-\alpha\chi_{\Omega_{dt}}:\alpha\in M\} \tag{24}$$

where

$$\Omega_{dt}=\{x\in\Omega: y_t(x)=y_d(x)\} \tag{25}$$

is a measurable set.

Remark 5.1. For this special case where the functional F can be expressed as a Sup, we could have completely bypassed the technique with the term in r by noting that the system of equations (22)–(25) has solutions and applying Proposition 1.6 in Ekeland and Temam [1, p. 157] to show that they are saddle points of G_r for all $r\geqq 0$.

We introduce the constants

$$\beta = \tfrac{1}{2}\|u\|_{L^\infty(\Omega)}, \qquad \tau = \beta/\|v\|_{L^\infty(\Omega)}. \tag{26}$$

The sets $\mathcal{A}$, $\mathcal{B}$ are

$$\mathcal{A} = M \times H_0^1(\Omega), \qquad \mathcal{B} = H_0^1(\Omega) \tag{27}$$

and the sets $A(t)$, $0 \leqq t \leqq \tau$, and $B(s, q)$, $0 \leqq s \leqq \tau$, are characterized by the following lemma.

LEMMA 5. (i) *Given $r \geqq 0$, then for all t, $0 \leqq t \leqq \tau$,*

$$A(t) = \{(\mu_t(\alpha), p_t(\alpha)) \colon \mu_t(\alpha) = \operatorname{sgn}(y_t - y_d) - \alpha\chi_{\Omega_{dt}}, \alpha \in M\} \tag{28}$$

where y_t is the solution of (22), *$\mu_t(\alpha)$ is given by* (24) *and $p_t(\alpha)$ by* (23) *with $\mu_t = \mu_t(\alpha)$. For $r > 0$ and all $q = (\psi, \mu) \in H_0^1(\Omega) \times M$ and all s, $0 \leqq s \leqq \tau$,*

$$B(s, q) = \{\varphi_s(q)\} \tag{29}$$

where $\underline{\varphi} = \varphi_s(q)$ is the unique solution in $H_0^1(\Omega)$ of the variational equation

$$\int_\Omega [\mu\varphi + (u + sv)\nabla\varphi \cdot \nabla\psi]\, dx + r\int_\Omega [(u + sv)\nabla\underline{\varphi} \cdot \nabla\varphi - f\varphi]\, dx = 0 \quad \forall \varphi \in H_0^1(\Omega). \tag{30}$$

(ii) *Moreover for all $r \geqq 0$, t, $0 \leqq t \geqq \tau$, and $q_t = (p_t, \mu_t) \in A(t)$,*

$$B(t, q_t) = B(t) = \{y_t\}$$

where y_t is the solution of (22) *which is independent of $q_t \in A(t)$. In particular for all $r > 0$, $S(t) = A(t) \times B(t)$ is the set of saddle points of $G_r(t, q, \varphi)$.*

Proof. The proof follows from previous considerations and Lemma 4 in § 4. □

We now apply Theorem 3 of § 4 to $J_r^*(u)$ with $r > 0$. For $0 \leqq t \leqq \tau$, $A(t) \neq \emptyset$ and (H1)(i) is verified. For each s in $[0, \tau]$, the set $B(s, q)$ reduces to a singleton. So by Lemma 5 it is nonempty and (H1)(ii) is verified. Hypothesis (H2) is obvious. For (H3)(i) we use Proposition 4. For (a) choose for τ_Y the weak topology on $Y = H_0^1(\Omega)$ and for K_Y the weakly compact ball of radius R since for all t in $[0, \tau]$ and $\{y_t\} = B(t)$

$$\beta \int_\Omega |\nabla y_t|^2\, dx \leqq \int_\Omega (u + tv)|\nabla y_t|^2\, dx = \int_\Omega f y_t\, dx \leqq c\|f\|_{L^2}\|\nabla y_t\|_{L^2}$$

$$\Rightarrow \|\nabla y_t\|_{L^2} \leqq R = c/\beta\|f\|_{L^2} \quad \forall t \in [0, \tau].$$

Conditions (b) and (c) and (H3)(ii) are also obvious. For (H4)(i) we use Proposition 5. For (a) we choose for τ_X the weak topology on $X = M \times H_0^1(\Omega)$ and for K the ball of radius $R' = c'/\beta$ (c' as defined below). Indeed for all t in $[0, \tau]$

$$\begin{aligned} \beta\|\nabla p\|^2 &\leqq \int_\Omega (u + tv)\nabla p \cdot \nabla p\, dx = -\int_\Omega [\operatorname{sgn}(y_t - y_d) - \alpha\chi_{\Omega_{dt}}]p\, dx \\ &\leqq c\|p\| \leqq c'\|\nabla p\| \\ &\Rightarrow \|\nabla p\| \leqq R' = c'/\beta \quad \forall t \in [0, \tau]. \end{aligned}$$

For (b)

$$\partial_t G(s, q, \varphi) = -\int_\Omega v\nabla p \cdot \nabla\varphi\, dx + r[dE(u + sv, \varphi; v, 0) - de(u + sv; v)].$$

From Zolésio [2, Thm. 1.1, p. 1458], it is known that

$$de(u + sv; v) = dE(u + sv, y_s; v, 0)$$

where y_s is the solution of (22) with $t=s$. With τ_Y being the strong topology on $Y=H_0^1(\Omega)$, $\partial_t G$ is jointly continuous with respect to its arguments.

Conditions (b) and (c) are on the functional

$$H(t,q)=\operatorname*{Sup}_{,\varphi\in H_0^1(\Omega)} G(t,q,\varphi) \tag{31}$$

where for $q=(\mu,\psi)$

$$G(t,q,\varphi)=-\left\{\int_\Omega \mu(\varphi-y_d)\,dx+dE(u+tv,\varphi;0,\psi)\right.$$
$$\left.+r[E(u+tv,\varphi)-e(u+tv)]\right\}. \tag{32}$$

Since G is lower semicontinuous in the variables (t,q,φ), the functional

$$(t,q)\to H(t,q)$$

is lower semicontinuous and (c) is verified. Condition (b) essentially requires that $t\to H(t,q)$ be continuous at 0. This follows from the continuity with respect to t of the minimizing element φ_t of $-G(t,q,\varphi)$ with respect to t. For (H4)(ii), the map

$$q\to B(0,q)=\{\varphi(q)\}$$

is single valued, affine and continuous with respect to q.

We summarize our results in the next proposition.

PROPOSITION 7. *For all u in U, v in $L^\infty(\Omega)$ and $r\geqq 0$, there exists $\tau>0$ such that hypotheses* (H1)-(H4) *on the functional $G(t,q,\varphi)$ in* (32) *be verified* (*recall that $q=(\mu,p)\in X=M\times H_0^1(\Omega)$ and that $\varphi\in Y=H_0^1(\Omega)$*). *For t in $[0,\tau]$, the sets $A(t)$ and $B(s,q)$ are given by* (28) *and* (29).

THEOREM 6. *For all u in U and v in $L^\infty(\Omega)$, the functional $J(u)$ is Gâteaux semidifferentiable at u in the direction v and*

$$dJ(u;v)=\lim_{t\to 0^+}[J(u+tv)-J(u)]/t=\operatorname{Sup}\left\{\int_\Omega v\nabla p(\alpha)\cdot\nabla y\,dx:\alpha\in M\right\} \tag{33}$$

where y and $p(\alpha)$ are the respective solutions of

$$-\operatorname{div}(u\nabla y)=f \quad \text{in } \Omega, \qquad y=0 \quad \text{on } \Gamma, \tag{34}$$

$$-\operatorname{div}(u\nabla p(\alpha))+\operatorname{sgn}(y-y_d)-\alpha\chi_{\Omega_d}=0 \quad \text{in } \Omega, \qquad p(\alpha)=0 \quad \text{on } \Gamma, \tag{35}$$

$$\Omega_d=\{x\in\Omega: y(x)=y_d(x)\}. \tag{36}$$

Proof. Recall that for $r\geqq 0$ and $0\leqq t\leqq\tau$

$$J_r(u+tv)=J(u+tv).$$

Computing the derivative of J is equivalent to computing the derivative of J_r for some fixed $r>0$. The results of § 4 are now available. It is sufficient to note that the integral in (33) is $\partial_t G(0,(\mu(\alpha),p(\alpha)),y)$, where

$$\mu(\alpha)=\operatorname{sgn}(y-y_d)-\alpha\chi_{\Omega_d}, \qquad \alpha\in M.$$

Expression (33) then follows from Proposition 7 and Theorem 3 in § 4. □

Remark 5.2. Note that the map $\alpha\to p(\alpha)$ is affine and continuous. In (33) the Sup occurs at extremal points of M. By defining

$$M_d=\{\alpha\in L^\infty(\Omega): \alpha(x)=\pm 1 \text{ in } \Omega_d \text{ and } \alpha(x)=0 \text{ elsewhere}\}, \tag{37}$$

we obtain

$$dJ(u;v)=\operatorname{Sup}\left\{\int_\Omega v\nabla p(\alpha)\cdot\nabla y\,dx:\alpha\in M_d\right\}. \tag{38}$$

Remark 5.3. Throughout our analysis, the parameter r is fixed but arbitrary and the saddle points of G_r are independent of r. Thus Theorem 6 could also have been obtained by applying Theorem 5 to $G(t, q, \varphi)$.

Remark 5.4. The interest behind this method for cost functionals of the form

$$J(u)=F(u, y(u)) \tag{39}$$

is to justify the differentiation of $J(u)$ at u in the direction v without using the intermediate step of differentiating the state $y(u)$ at u in the direction v.

The above results formally extend to the class of linear variational problems. For instance, let

$$\underset{\sim}{E}: U_{\mathrm{ad}} \times X \to R, \qquad \underset{\sim}{E}(u, x)=\tfrac{1}{2}a(u, x, x)-L(u, x) \tag{40}$$

for some open set U_{ad} of a Banach space U, a Hilbert space X, a continuous symmetrical bilinear form $a(u, \cdot, \cdot)$, and a continuous linear form $L(u, \cdot)$. For each u in U_{ad}, define

$$\underset{\sim}{e}(u)=\operatorname{Inf}\{\underset{\sim}{E}(u, x): x \in X\} \tag{41}$$

and assume that the minimizing element $y=y(u)$ in X is unique and completely characterized by the variational equation

$$d\underset{\sim}{E}(u, y; 0, \psi)=0 \quad \forall \psi \in X. \tag{42}$$

We can readily extend this method to cost functionals of the form

$$J(u)=\underset{\sim}{F}(u, y(u)) \tag{43}$$

for some map

$$\underset{\sim}{F}: U_{\mathrm{ad}} \times X \to R \tag{44}$$

where $y(u)$ is the solution of (42).

Then the associated functional $\underset{\sim}{G}$ is given by

$$-\underset{\sim}{G}(u, \psi, x)=\underset{\sim}{F}(u, x)+d\underset{\sim}{E}(u, x; 0, \psi) \tag{45}$$

for $u \in U_{\mathrm{ad}}$, $\psi \in X$ and $x \in X$. Assume that F is semidifferentiable at $y(u)$, that is,

$$\forall z \in Z, \quad d\underset{\sim}{F}(u, y(u); z)=\lim_{t \to 0^+}[\underset{\sim}{F}(u, y(u)+tz)-\underset{\sim}{F}(u, y(u))]/t \quad \text{exists.} \tag{46}$$

Let $P(u)$ be the set of solutions $p=p(u)$ of the adjoint inequality

$$d\underset{\sim}{F}(u, y(u); 0, \psi)+d^2\underset{\sim}{E}(u, y(u); 0, p; 0, \psi) \geqq 0 \quad \forall \psi \in X \tag{47}$$

where $y(u)$ is the solution of (42).

THEOREM 7. *Let E and F be as described above. Fix v in V and $\tau>0$ such that for all t, $0 \leqq t \leqq \tau$, $u+tv \in U_{\mathrm{ad}}$. Assume the following*:

(a) *For all t, $0 \leqq t \leqq \tau$, $P(u+tv) \neq \varnothing$;*

(b) *For all φ, $\psi \in X$, $t \to a(u+tv, \varphi, \psi)$ and $t \to L(u+tv, \psi)$ are continuously differentiable in $[0, \tau]$;*

(c) *There exists $\alpha>0$, for all t, $0 \leqq t \leqq \tau$, and for all φ, $a(u+tv, \varphi, \varphi) \geqq \alpha\|\varphi\|^2$;*

(d) *For all ψ, $(t, \varphi) \to d\underset{\sim}{F}(u+tv, \varphi; 0, \psi)$ is upper semicontinuous at $(0, y(u))$;*

(e) *There exists a neighborhood $W(u)$ of u, and there exists $c(u)>0$, such that*

$$\forall \psi \in X, \quad \forall w \in W(u), \quad d\underset{\sim}{F}(w, y(w); 0, \psi) \leqq c(u)\|\psi\|; \tag{48}$$

(f) *For all φ, for all v and for all $t \in [0, \tau]$, the limit*

$$d\underset{\sim}{F}(u+tv, \varphi; v, 0)=\lim_{s \to 0^+}[\underset{\sim}{F}(u+(t+s)v, \varphi)-\underset{\sim}{F}(u+tv, \varphi)]/s \tag{49}$$

exists;

(g) *For all* v, $(t, \varphi) \to d\underline{F}(u+tv, \varphi; v, 0)$ *is lower semicontinuous at* $(0, y_0)$;

(h) *For all* φ *and for all* v, $t \to d\underline{F}(u+tv, \varphi; v, 0)$ *is upper semicontinuous at* $t=0$.

Then $J(u)$ *defined by* (43) *is semidifferentiable at* u *in the direction* v *and*

$$dJ(u; v) = \operatorname{Sup}\{d\underline{F}(u, y(u); v, 0) + d^2\underline{E}(u, y(u); 0, p; v, 0): p \in P(u)\}. \tag{50}$$

The proof of this theorem will be given after a short discussion of the fundamental hypothesis (a).

Remark 5.5. If the map

$$\varphi \to dF(u, y; 0, \varphi) \tag{51}$$

is linear and continuous, the adjoint problem (47) is variational and $P(u)$ reduces to the usual solution of the associated variational problem.

When (54) is nonlinear we can use the augmented Lagrangian technique previously developed.

PROPOSITION 8. *Assume the existence of a number* r, $0<r<2$, *such that*

$$\left.\begin{aligned} &F(u, x) + \frac{r}{2} a(u, x, x) \to +\infty, \\ &-F(u, -x) + (1-r/2) a(u, x, x) \to +\infty \end{aligned}\right\} \quad \text{as } \|x\| \to \infty. \tag{52}$$

Then $P(u)$ *is not empty.* □

Proof of Theorem 7. We show that the hypotheses of Theorem 5 are verified. From (a) the sets of saddle points

$$S(t) = P(u+tv) \times \{y(u+tv)\} \neq \varnothing \quad \forall t \in [0, \tau]$$

and (HH1) is verified. Define

$$-G(t, \psi, \varphi) = \underline{F}(u+tv, \varphi) + d\underline{E}(u+tv, \varphi; 0, \psi).$$

From (b) and (f) for all t in $[0, \tau[$ and φ, ψ in X

$$-\partial_t G(t, \psi, \varphi) = d\underline{F}(u+tv, \varphi; v, 0) + d^2\underline{E}(u+tv, \varphi; 0, \psi; v, 0)$$

and (HH2) is verified. For each t in $[0, \tau]$, $y_t = y(u+tv) \in X$ is the unique solution of

$$a(u+tv, y_t, \psi) - L(u+tv, \psi) = 0 \quad \forall \psi \in X.$$

From the above equation and (b)

$$a(u+tv, y_t - y_0, \psi) + a(u+tv, y_0, \psi) - L(u+tv, \psi) = 0 \quad \forall \psi \in X. \tag{53}$$

But there exists $\theta \in]0, 1[$

$$\begin{aligned} &a(u+tv, y_0, \psi) - L(u+tv, \psi) \\ &\quad = a(u, y_0, \psi) - L(u, \psi) + t[\partial_t a(u+\theta tv, y_0, \psi) - \partial_t L(u+\theta tv, \psi)] \end{aligned}$$

and for t small the term in the square bracket is bounded by $c\|\psi\|$ where $c>0$ is a constant which is independent of t. So

$$a(u+tv, y_t - y_0, \psi) \leqq tc\|\psi\|$$

and by (c) with $\psi = y_t - y_0$

$$\alpha\|y_t - y_0\|^2 \leqq tc\|y_t - y_0\|.$$

So the map $t \to y_t$ is continuous at $t=0$ in X-strong and (HH4)(i) is verified.

(HH4)(ii) reduces to the lower semicontinuity of the map

$$(t, \varphi) \to d\underline{F}(u+tv, \varphi; v, 0) + \partial_t a(u+tv, \varphi, \psi) - \partial_t L(u+tv, \psi)$$

for each ψ in X. For the first term, it is true by (g). As for the other two terms it follows from (b) and the linearity of $\partial_t a$ with respect to φ. Similarly (HH3)(ii) follows from (h), (b) and the linearity of $\partial_t a(u+tv, y, \psi)$ with respect to ψ. As for (HH3)(i) we know that for all t in $[0, \tau]$ there exists at least one solution $p_t \in X$ to the variational inequality

$$dF(u+tv, y_t; 0, \psi)+d^2E(u+tv, y_t; 0, p_t; 0, \psi) \geqq 0 \quad \forall \psi \in X \tag{54}$$

where $y_t = y(u+tv)$ is the unique solution of (53). We first show that p_t is bounded in X. First (54) reduces to

$$dF(u+tv, y_t; 0, \psi)+a(u+tv, p_t, \psi) \geqq 0 \quad \forall \psi \in X \tag{55}$$

and for $\psi = -p_t$

$$\alpha \|p_t\|^2 \leqq dF(u+tv, y_t; 0, -p_t)$$

and by (e) for t small enough

$$\alpha \|p_t\|^2 \leqq c(u)\|p_t\|.$$

So $\{p_t\colon 0 \leqq t \leqq \tau\}$ is bounded in X and there exists p^* and a sequence $t_n \to 0^+$ such that

$$p_n = p_{t_n} \to p^* \quad \text{in } X\text{-weak.}$$

Going back to (55)

$$dF(u+t_n v, y_n; 0, \psi)+a(u+t_n v, p_n, \psi) \geqq 0 \quad \forall \psi \in X.$$

From (b)

$$a(u+t_n v, p_n, \psi) \to a(u, p^*, \psi)$$

and by (d)

$$\limsup_{n\to\infty} dF(u+t_n v, y_n; 0, \psi) \leqq dF(u, y_0; 0, \psi)$$

since $y_n \to y_0$ in X-strong. As a result

$$dF(u, y_0; 0, \psi)+a(u, p^*, \psi) \geqq 0 \quad \forall \psi \in X$$

and $p^* \in P(u)$. So (HH3)(i) is verified and Theorem 5 applies. □

Remark 5.6. Theorem 7 can also be obtained by using for some $r > 0$ the Lagrangian

$$-G_r(u, \psi, \varphi) = F(u, \varphi)+dE(u, \varphi; 0, \psi)+r[E(u, \varphi)-e(u)], \tag{56}$$

hypothesis (52) for $u+tv$ instead of $S(t) \neq \emptyset$ and applying Theorem 3 or 4 to

$$J^*(u) = -\operatorname*{Inf}_{\psi \in X} \operatorname*{Sup}_{\varphi \in X} G_r(u, \psi, \varphi). \tag{57}$$

In general it is not possible to verify hypothesis (H1) in Theorems 3 and 4 for

$$J(u) = \operatorname*{Inf}_{\varphi \in X} \operatorname*{Sup}_{\psi \in X} -G_r(u, \psi, \varphi). \tag{58}$$

Proof of Proposition 8. As in our illustrative example, we construct the augmented Lagrangian

$$-G_r(u, p, x) = -G(u, p, x)+r[E(u, x)-e(u)].$$

It is easy to show that the functional $-G_r$ is convex and lower semicontinuous in x and concave and upper semicontinuous in p. From hypotheses (52),

$$-G_r(u, x, 0) \to +\infty \quad \text{as } \|x\| \to \infty,$$
$$\operatorname*{Inf}_{x \in X} -G_r(u, p, x) \leqq G_r(u, p, -p) \to -\infty \quad \text{as } \|p\| \to \infty.$$

Again by Proposition 2.4 in Ekeland and Temam [1, p. 164], $\underset{\sim}{G}_r(u, \cdot, \cdot)$ has saddle points in $X \times X$. They are completely characterized by the system

$$-d\underset{\sim}{G}_r(u, p, y; 0, 0, \varphi - y) \geqq 0 \quad \forall \varphi \in X, \tag{59}$$

$$-d\underset{\sim}{G}_r(u, p, y; 0, \psi - p, 0) \leqq 0 \quad \forall \psi \in X, \tag{60}$$

which is equivalent to (42) and (47). This is sufficient to establish that $P(u)$ is not empty. □

6. Shape derivative of a functional: a simple example.

6.1. Shape optimization problem. Shape and structural optimal design is quite a broad field of activity. A good account of recent work can be found in the two volumes of the proceedings of the NATO Advanced Study Institute held in Iowa City (cf. Haug and Céa [1]) which contain an enormous amount of material covering the engineering and mathematical aspects of that class of problems. Relatively few books have been published on optimal design problems based on PDE models or in engineering terminology distributed parameter models. According to Haug and Céa [1], we find the book of Prager [1], the proceedings of the symposium held in Warsaw in 1973 edited by Sawczuk and Mroz [1], the books of Rozvany [1] in 1976, Haug and Arora [1], Banichuk [1] (original version in Russian, available in English), and the more recent books by Pironneau [1] and Haug, Choi, and Komkov [1]. The method presented here and in § 7 can be used to obtain a mathematical justification for some parts of the results formally obtained in Céa [1], Dems and Mroz [1], Haug and Céa [1], and Haug, Choi, and Komkov [1]. In this section we consider the shape sensitivity analysis problem. We do not cover variational inequalities. For this we refer to the recent papers by Sokolowski and Zolésio [1], [2] and Sokolowski [1] and their bibliographies.

Consider the following simple example. Let Ω be a bounded open domain in R^n with a smooth boundary Γ. Let $y = y(\Omega)$ be the solution of the variational problem

$$\operatorname{Inf}\{E(\Omega, \varphi) \colon \varphi \in H^1(\Omega)\} \tag{1}$$

where

$$E(\Omega, \varphi) = \frac{1}{2} \int_\Omega [|\nabla \varphi|^2 + |\varphi|^2 - 2f\varphi] \, dx \tag{2}$$

for some fixed function f in $H^1(R^n)$. We associate with y a cost function

$$\mathrm{J}(\Omega) = \mathrm{F}(\Omega, \mathrm{y}(\Omega)). \tag{3}$$

For instance we can choose the standard cost function

$$F(\Omega, y) = \frac{1}{2} \int_\Omega (y - Y_d)^2 \, dx, \qquad Y_d \in H^1(R^n). \tag{4}$$

6.2. The velocity field method. We briefly recall the notion of a shape derivative. Let $V(t, x)$, $t \geqq 0$, $x \in R^n$, be a *velocity field of deformation.* Under the action of V, the points of Ω are transported onto a new domain $\Omega_t = T_t(\Omega)$, where the transformation $T_t \colon R^n \to R^n$ is generated by the solutions of the equation

$$(\partial/\partial t) T_t(x) = V(t, T_t(x)), \qquad t \geqq 0, \quad T_0(x) = x \tag{5}$$

(cf. Zolésio [3]). Let y_t be the solution of problem (1) on the transformed domain Ω_t

$$\operatorname{Inf}\{E(\Omega_t, \varphi) \colon \varphi \in H^1(\Omega_t)\} \tag{1_t}$$

and associate with y_t the cost function

$$J(\Omega_t) = F(\Omega_t, y_t). \tag{3_t}$$

Traditional methods involve the computation of the shape derivative (or partial derivative) Y' or the *Material derivative* $\dot{y}$. The shape derivative is defined as

$$Y'(x) = \lim_{t \to 0^+} [Y(t, x) - Y(0, x)]/t$$

where for some $\tau > 0$ $Y(t, x)$ is an appropriate extension of $y_t(x)$ to $[0, \tau] \times D$ for some fixed domain D containing all perturbations Ω_t of Ω, $0 \leqq t \leqq \tau$. The material derivative is defined as

$$\dot{y} = \lim_{t \to 0^+} [y^t - y]/t$$

in an appropriate function space on Ω, where y^t is the transported solution (from Ω_t to Ω)

$$y^t = y_t \circ T_t.$$

In classical examples Y' is the solution of a boundary value problem which depends on y and the normal component of the velocity field on the boundary Γ. However, in general, the material derivative is also the solution of a boundary value problem on Ω, but it depends on the velocity field in the whole domain. In fact the two derivatives are related through the formula

$$Y' = \dot{y} - \nabla y \cdot V$$

where V is the velocity field at 0, $x \to V(0, x)$. In general Y' is "rougher" than the material derivative.

The next step consists in differentiating $J(\Omega_t)$ using the material derivative or Y'. Then an appropriate adjoint variable p is introduced to eliminate those derivatives and obtain a final expression which depends on Ω, y, p and V. The adjoint variable p is the solution of a boundary value problem which is dual to the corresponding boundary value problem for the material derivative or Y'.

The final expression can then be used for shape sensitivity analysis or as a necessary condition characterizing an eventual minimizing domain Ω^*.

6.3. The Inf Sup formulation of the perturbed problem. In general our objective is the minimization of the cost function J with respect to Ω. In particular we want to compute the shape derivative of J at Ω in the direction of the velocity field of deformations V. To do this we transform the problem (1_t)-(3_t) into an Inf Sup problem. This approach is widespread in the engineering and mathematical literature.

The solution of (1_t) is completely characterized by the variational equation

$$dE(\Omega_t, y_t; \varphi) = 0 \quad \forall \varphi \in H^1(\Omega_t) \tag{6}$$

where

$$dE(\Omega_t, \psi; \varphi) = \int_{\Omega_t} [\nabla \psi \cdot \nabla \varphi + \psi \varphi - f \varphi] \, dx, \qquad \varphi, \psi \in H^1(\Omega_t). \tag{7}$$

Define for $r \geqq 0$

$$-G_r(t, \varphi, p) = F(\Omega_t, \varphi) + dE(\Omega_t, \varphi; p) + r[E(\Omega_t, \varphi) - e(t)] \tag{8}$$

where

$$e(t) = \text{Min} \{E(\Omega_t, \varphi) : \varphi \in H^1(\Omega_t)\}. \tag{9}$$

But

$$\operatorname{Sup}\{-G_r(t,\varphi,p)\colon p\in H^1(\Omega_t)\}=\begin{cases}F(\Omega_t,\varphi) & \text{if } \varphi \text{ is a solution of (6),}\\ +\infty & \text{otherwise.}\end{cases}$$

As a result for all $r \geqq 0$

$$J(\Omega_t)=\operatorname{Inf}\{\operatorname{Sup}[-G_r(t,\varphi,p)\colon p\in H^1(\Omega_t)]\colon \varphi\in H^1(\Omega_t)\}.$$

In this form the spaces depend on the parameter t and it is not readily possible to apply the theorems of § 6. However, for the Neuman problem it is possible to embed everything in $H^1(R^n)$. It is readily seen that

$$\operatorname{Sup}\{-G_r(t,\varphi,p)\colon p\in H^1(R^n)\}=\begin{cases}F(\Omega_t,\varphi) & \text{if } \varphi \text{ is a solution of (6),}\\ +\infty & \text{otherwise.}\end{cases}\tag{10}$$

As a result for all $r \geqq 0$

$$J(\Omega_t)=\operatorname{Inf}\{\operatorname{Sup}[-G_r(t,\varphi,p)\colon p\in H^1(R^n)]\colon \varphi\in H^1(R^n)\}.\tag{11}$$

The spaces involved are now fixed and independent of the parameter $t \geqq 0$.

6.4. Perturbed dual functional J_r^* and existence of saddle points. Our objective is to show the existence of saddle points for $r>0$ and use the results of § 4 together with identity (11).

For $r \geqq 0$, define the functionals

$$J_r(\Omega_t)=-\operatorname{Sup}\{\operatorname{Inf}[G_r(t,\varphi,p)\colon p\in H^1(R^n)]\colon \varphi\in H^1(R^n)\}\tag{12}$$

and

$$J_r^*(\Omega_t)=-\operatorname{Inf}\{\operatorname{Sup}[G_r(t,\varphi,p)\colon \varphi\in H^1(R^n)]\colon p\in H^1(R^n)\}.\tag{13}$$

Recall that in view of (11)

$$J_r(\Omega_t)=J_0(\Omega_t)=J(\Omega_t)\quad \forall r\geqq 0.\tag{14}$$

In general

$$J_r^*(\Omega_t)\leqq J_r(\Omega_t)\tag{15}$$

since J_r^* is the dual functional associated with the perturbed functional G_r.

We have made the above construction in order to apply Theorem 3 to the dual problem for $r>0$ (for $r=0$ hypothesis (H1) would not be verified) or Theorem 5 for $r \geqq 0$.

The functional G_r has a saddle point for all $r \geqq 0$.

PROPOSITION 9. (i) *Given $\tau>0$ small enough, then for all r, $0\leqq r$, and t, $0\leqq t\leqq \tau$, $G_r(t,\cdot,\cdot)$ has saddle points (Y_r, P_r) in $H^1(R^n)\times H^1(R^n)$ and*

$$J_r^*(\Omega_t)=J_r(\Omega_t).\tag{16}$$

(ii) *The restriction of each saddle point to Ω_t*

$$(y_t^r, p_t^r)=(Y_r|_{\Omega_\tau}, P_r|_{\Omega_\tau})\tag{17}$$

coincides with the unique pair (y_t, p_t) solution of the system

$$dE(\Omega_t, y_t;\varphi)=0\quad \forall\varphi\in H^1(\Omega_t),\tag{18}$$

$$dF(\Omega_t,y_t;\psi)+d^2E(\Omega_t,y_t;p_t;\psi)=0\quad \forall\psi\in H^1(\Omega_t)\tag{19}$$

where

$$dF(\Omega_t, y_t; \psi) = \int_{\Omega_t} (y_t - Y_d)\psi \, dx, \tag{20}$$

$$d^2E(\Omega_t, y_t; p_t; \psi) = \int_{\Omega_t} [\nabla p_t \cdot \nabla \psi + p_t \psi] \, dx. \tag{21}$$

Proof. The conditions characterizing a saddle point (Y_r, P_r) of G_r are precisely (18)–(19). Both equations are elliptic with a unique solution in $H^1(\Omega_t)$ which is independent of $r \geqq 0$. □

6.5. Application of Theorem 3. The next step is the application of Theorem 3 from § 4 to the function $g(t) = J_r^*(\Omega_t)$. We first study the differentiability of the functional $G_r(t, \varphi, \psi)$ as defined by (8) with respect to t for all φ and ψ in the space $X = H^1(R^n)$:

$$\begin{aligned} -G_r(\Omega_t, \varphi, \psi) = \frac{1}{2}\int_{\Omega_t} (\varphi - Y_d)^2 \, dx + \int_{\Omega_t} [\nabla \psi \cdot \nabla \varphi + \psi\varphi - f\varphi] \, dx \\ + \frac{r}{2}\int_{\Omega_t} [|\nabla \varphi|^2 + |\varphi|^2 - 2f\varphi] \, dx. \end{aligned} \tag{22}$$

If ψ and φ were smoother (e.g., in $H^2(R^n)$) to make sure that the traces of $|\nabla\varphi|^2$ and $\nabla\psi \cdot \nabla\varphi$ exist, the standard expression for the derivative would be given by a boundary integral of the integrand in (22) (cf. Zolésio [3]). Unfortunately this is not the case. Expression (22) transported from Ω_t onto Ω is given by

$$\begin{aligned} &\int_\Omega A(t)\left[\nabla(\psi \circ T_t) + \frac{r}{2}(\nabla\varphi \circ T_t)\right] \cdot \nabla(\varphi \circ T_t) \, dx \\ &\quad + \int_\Omega \left\{\frac{1}{2}(\varphi \circ T_t - Y_d \circ T_t)^2 \right. \\ &\qquad \left. + \left[(\psi \circ T_t - f \circ T_t) + \frac{r}{2}(\varphi \circ T_t - 2f \circ T_t)\right](\varphi \circ T_t)\right\} J(t) \, dx \end{aligned}$$

where DT_t is the Jacobian matrix of the transformation T_t and

$$J(t) = \det(DT_t), \qquad A(t) = J(t)((DT_t)^{-1})^*(DT_t)^{-1} \tag{23}$$

(* indicates the transposed matrix). Again to differentiate the above expression with respect to t would require that φ and ψ be in $H^2(R^n)$. To get around this difficulty we need the special technique which is described below and which does not seem to have any counterpart in control.

As before, given the smooth velocity field V, define the transformation

$$T_t = T_t(V) : R^n \to R^n \tag{24}$$

which transports R^n onto R^n, $\Omega_0 = \Omega$ onto Ω_t and $\Gamma_0 = \Gamma$ onto $\Gamma_t = \partial\Omega_t$. The space $H^1(\Omega_t)$ is transported in a similar way. As the functions φ and ψ fill in the whole space X, so do the functions $\psi \circ T_t^{-1}$ and $\varphi \circ T_t^{-1}$. As a result

$$g(t) = J_r(\Omega_t) = J_r^*(\Omega_t) = -\operatorname*{Inf}_{\varphi \in X} \operatorname*{Sup}_{\psi \in X} -\underset{\sim}{G}_r(t, \varphi, \psi) \tag{25}$$

where

$$\underset{\sim}{G}_r(t, \varphi, \psi) = G_r(t, \varphi \circ T_t^{-1}, \psi \circ T_t^{-1}). \tag{26}$$

By introducing the quantities

$$J(t)=\det(DT_t), \qquad A(t)=J(t)((DT_t)^{-1})^*(DT_t)^{-1} \tag{27}$$

(DT_t is the Jacobian matrix of the transformation T_t) and the change of variable $x'=T_t(x)$, we obtain

$$-\underset{\sim}{G}_r(t,\varphi,\psi)=-\underset{\sim}{G}_0(t,\varphi,\psi)+r[E(\Omega,\varphi\circ T_t)-e(t)] \tag{28}$$

where

$$e(t)=\operatorname{Inf}\{E(\Omega_t,\varphi):\varphi\in H^1(R^n)\}=\operatorname{Inf}\{E(\Omega,\varphi\circ T_t):\varphi\in H^1(R^n)\} \tag{29}$$

and

$$\begin{aligned}-\underset{\sim}{G}_0(t,\varphi,\psi)=&\int_\Omega[(A(t)\nabla\psi)\cdot\nabla\varphi+\psi\varphi J(t)]\,dx\\&+\int_\Omega\left[\frac{1}{2}(\varphi-Y_d\circ T_2)^2-(f\circ T_t)\psi\right]J(t)\,dx,\end{aligned} \tag{30}$$

$$E(\Omega,\varphi\circ T_t)=\int_\Omega\frac{1}{2}\{(A(t)\nabla\varphi)\cdot\nabla\varphi+J(t)[\varphi\varphi-(f\circ T_t)\varphi]\}\,dx. \tag{31}$$

The derivative with respect to t can easily be obtained for every term in $\underset{\sim}{G}_r(t,\cdot,\cdot)$ except possibly $e(t)$. Fortunately we know from Zolésio [3, Thm. 1.1, p. 1458] that

$$\begin{aligned}d_te(0)&=\operatorname{Inf}\{\partial_tE(\Omega_0,\varphi):\varphi\in X, E(\Omega_0,\varphi)=e(0)\}\\&=\partial_tE(\Omega_0,Y)\end{aligned} \tag{32}$$

where Y is the solution of

$$E(\Omega_0,Y)=\operatorname{Inf}\{E(\Omega_0,\varphi):\varphi\in H^1(R^n)\}$$

or equivalently

$$Y\in H^1(R^n),\quad dE(\Omega,Y;\varphi)=0\quad\forall\varphi\in H^1(R^n). \tag{33}$$

We finally obtain the following intermediate result prior to the application of Theorem 3 in § 4.

PROPOSITION 10. *For all r, $0<r<1$, t, $0\leqq t\leqq\tau$, $\underset{\sim}{G}_r(t,\varphi,\psi)$ is differentiable with respect to t at $t=0$ and*

$$\begin{aligned}-\partial_t\underset{\sim}{G}_r(0,\varphi,\psi)=&\int_\Omega[(A'(0)\nabla\psi)\cdot\nabla\varphi+\psi\varphi\operatorname{div}V(0)]\,dx\\&-\int_\Omega\left[\frac{1}{2}(\varphi-Y_d)\nabla Y_d+\psi\nabla f\right]\cdot V(0)\,dx\\&+\int_\Omega\left[\frac{1}{2}(\varphi-Y_d)^2-f\psi\right]\operatorname{div}V(0)\,dx+r[\partial_tE(\Omega_0,\varphi)-\partial_tE(\Omega_0,Y)]\end{aligned} \tag{34}$$

where

$$A'(0)=[\operatorname{div}V(0)]I_d-[DV(0)+(DV(0))^*]. \tag{35}$$

($(DV(0))^*$ *is the transposed matrix of* $(DV(0))$). □

All the hypotheses of Theorem 3 in § 4 are now verified.

THEOREM 8. *For all r,* $0<r<1$,

$$dJ_r(\Omega;\ V(0)) = -\partial_t \underset{\sim}{G}_r(0, p, y) \tag{36}$$

where y and p are the solutions of

$$dE(\Omega, y;\ \varphi)=0 \quad \forall \varphi \in H^1(\Omega) \tag{37}$$

and

$$dF(\Omega, y;\ \psi)+d^2E(\Omega, y;\ p;\ \psi)=0 \quad \forall \psi \in H^1(\Omega). \tag{38}$$ □

Remark 6.1. The derivative of J_r was obtained without our ever considering the problem of the differentiability of y.

Finally recall identity (14) to obtain the desired result.

THEOREM 9. (i) *The function g is differentiable at* $t=0$ *and*

$$\begin{aligned} dJ(\Omega;\ V(0)) = &\int_\Omega \left[(A'(0)\nabla p)\cdot\nabla y - \left[\frac{1}{2}(y-Y_d)\nabla Y_d + p\nabla f\right]\cdot V(0)\right] dx \\ &+\int_\Omega \left[\frac{1}{2}(y-Y_d)^2 + py - fp\right] \operatorname{div} V(0)\, dx \end{aligned} \tag{39}$$

where y and p are the solutions of (37) *and* (38).

(ii) *If, in addition, p and y belong to* $H^{3/2+\rho}(\Omega)$, $\rho>0$, *then*

$$dJ(\Omega;\ V(0)) = \int_\Gamma \left[\nabla y\cdot\nabla p + yp - fp + \frac{1}{2}(y-Y_d)^2\right] V(0)\cdot n\, d\Gamma. \tag{40}$$

Proof. It suffices to note that for $\varphi = y$ the term which contains the r in identity (36) is identically zero. When y and p are sufficiently smooth, (39) is equivalent to the standard boundary integral formulation in shape optimization. To show that set $\varphi = \nabla p\circ V$ in (37) and $\psi=\nabla y\circ V$ in (38), add the resulting two equations and reorganize the terms. This yields an identity which shows that the right-hand sides of (39) and (40) are equal. □

Remark 6.2. The simple example above contains several techniques which will turn out to be fundamental to the general theory. For instance the introduction of the functional

$$\underset{\sim}{G}_r(t, \varphi, \psi) = G_r(t, \varphi\circ T_t^{-1}, \psi\circ T_t^{-1})$$

followed by the transport of the resulting expression from the domain Ω_t onto Ω makes it possible to keep the test functions in $H^1(\Omega)$ instead of going to the larger space $H^1(R^n)$. For instance this feature is extremely important in the homogeneous Dirichlet problem in $H_0^1(\Omega)$ where it would not be possible to substitute $H^1(R^n)$.

7. Shape derivative of a functional: The general method and other examples. In this last section we formalize the ideas and techniques introduced in § 6 and describe two other examples to further illustrate the applicability of Theorems 3 or 4 and our associated techniques.

The first goes over the discussion at the end of § 6 in Remark 6.2. Key details are provided to show how problems with Dirichlet boundary conditions can be handled. In fact the suggested construction could also have been used right from the beginning in § 6, but we preferred to do it in a different way to better emphasize its importance.

The second example shows that we can handle problems where the smoothness of the solution of the saddle point equations is minimal. Other techniques based on the Implicit Function Theorem would require more smoothness.

7.1. Dirichlet boundary condition. We go back to problem (1)-(4) in § 6 but with $H_0^1(\Omega)$ instead of $H^1(\Omega)$. Let $y = y(\Omega)$ in $H_0^1(\Omega)$ be the solution of the variational problem

$$\operatorname{Inf}\{E(\Omega, \varphi): \varphi \in H_0^1(\Omega)\} \tag{1}$$

where

$$E(\Omega, \varphi) = \frac{1}{2}\int_\Omega [|\nabla \varphi|^2 + |\varphi|^2 - 2f\varphi]\, dx \tag{2}$$

for some fixed function f in $H^1(R^n)$. This is the homogeneous Dirichlet problem. We associate with y a cost function

$$J(\Omega) = F(\Omega, y(\Omega)). \tag{3}$$

Again for simplicity we assume that it is of the form

$$F(\Omega, \varphi) = \frac{1}{2}\int_\Omega (\varphi - Y_d)^2\, dx, \quad \varphi \in H_0^1(\Omega), \quad Y_d \in H^1(R^n). \tag{4}$$

Assume that V is a smooth vector field which transports Ω onto Ω_t, its boundary Γ onto Γ_t and the Sobolev space $H^1(\Omega)$ onto $H^1(\Omega_t)$ at time $t \geqq 0$. As a result it also transports functions in $H_0^1(\Omega)$ onto functions in $H_0^1(\Omega_t)$ and

$$H_0^1(\Omega_t) = \{\varphi \circ T_t^{-1}: \varphi \in H_0^1(\Omega)\}. \tag{5}$$

Here we use techniques described at the end of § 6 in Remark 6.2. We introduce the new functional

$$\varphi \to \underline{E}(t, \varphi) = E(\Omega_t, \varphi \circ T_t^{-1}): H_0^1(\Omega) \to R \tag{6}$$

and notice that

$$\operatorname*{Inf}_{\varphi \in H_0^1(\Omega)} \underline{E}(t, \varphi) = \operatorname*{Inf}_{\psi \in H_0^1(\Omega_t)} E(\Omega_t, \psi). \tag{7}$$

Denote by y^t and y_t the minimizing unique solutions of $\underline{E}(t, \varphi)$ in $H_0^1(\Omega)$ and $E(\Omega_t, \psi)$ in $H_0^1(\Omega_t)$, respectively. Then in view of (5)

$$y_t = y^t \circ T_t^{-1}. \tag{8}$$

The two formulations are equivalent, but the differentiation of $\underline{E}(t, \varphi)$ with respect to t does not require that the function φ be smoother than $H^1(\Omega)$ since

$$\underline{E}(t, \varphi) = \frac{1}{2}\int_{\Omega_t} [|\nabla(\varphi \circ T_t^{-1})|^2 + |\varphi \circ T_t^{-1}|^2 - 2f(\varphi \circ T_t^{-1})]\, dx$$

and after a change of variable

$$\underline{E}(t, \varphi) = \frac{1}{2}\int_\Omega \{(A(t)\nabla\varphi)\cdot\nabla\varphi + [|\varphi|^2 - 2(f \circ T_t)\varphi]J(t)\}\, dx, \tag{9}$$

where DT_t is the Jacobian matrix associated with the transformation T_t,

$$J(t) = \det(DT_t), \qquad A(t) = J(t)((DT_t)^{-1})^*(DT_t)^{-1} \tag{10}$$

and * denotes the transposed matrix.

If we want to work with $\underline{E}(t, \Omega)$ and y^t, we must also transform the functional F into a new functional

$$\varphi \to \underline{F}(t, \varphi) = F(\Omega_t, \varphi \circ T_t^{-1}) : H_0^1(\Omega) \to R. \tag{11}$$

As a result the cost function

$$J(\Omega_t) = F(\Omega_t, y_t) = F(\Omega_t, y^t \circ T_t^{-1}) = \underline{F}(t, y^t). \tag{12}$$

Again the differentiability of $\underline{F}(t, \varphi)$ with respect to t does not require that the function φ be smoother than $H^1(\Omega)$:

$$\underline{F}(t, \varphi) = \frac{1}{2} \int_{\Omega_t} (\varphi \circ T_t^{-1} - Y_d)^2 \, dx$$

and after a change of variable

$$\underline{F}(t, \varphi) = \frac{1}{2} \int_{\Omega} (\varphi - Y_d \circ T_t)^2 J(t) \, dx. \tag{13}$$

Thus we are led to the construction of the functional

$$(\varphi, \psi) \to \underline{G}_r(t, \varphi, \psi) = G_r(\Omega_t, \varphi \circ T_t^{-1}, \psi \circ T_t^{-1}) : H_0^1(\Omega) \times H_0^1(\Omega) \to R \tag{14}$$

and the technique used in § 6.

We do not repeat the details here since the results are the same as those in Theorem 9 except that the functions y and p are the solutions of the variational equations

$$y \in H_0^1(\Omega), \qquad dE(\Omega, y; \varphi) = 0 \quad \forall \varphi \in H_0^1(\Omega), \tag{15}$$

$$p \in H_0^1(\Omega), \; dF(\Omega, y; \psi) + d^2E(\Omega, y; p, \psi) = 0 \quad \forall \psi \in H_0^1(\Omega). \tag{16}$$

So formally it suffices to substitute $H_0^1(\Omega)$ for $H^1(\Omega)$ in Theorem 9.

7.2. An example with less smoothness. In the two previous examples, the solutions (y, p) of the optimality system (42)-(47) or (15)-(16) are smoother than anticipated and belong to $H^2(\Omega)$. So it would be possible to argue that all the results can also be obtained by application of some form of the Implicit Function Theorem.

It is not difficult to slightly modify the example of § 6 to prevent this situation from happening. First change the functionals E and F to

$$E(\Omega, \varphi) = \frac{1}{2} \int_{\Omega} [|\nabla \varphi|^2 + |\varphi|^2 - 2f \cdot \nabla \varphi] \, dx, \qquad \varphi \in H^1(\Omega) \tag{17}$$

where $f \in (H^1(\Omega))^n$

$$F(\Omega, \varphi) = \int_{\Omega} |\nabla \varphi| \, dx, \qquad \varphi \in H^1(\Omega). \tag{18}$$

The minimization problem

$$e(\Omega) = \operatorname{Inf} \{E(\Omega, \varphi) : \varphi \in H^1(\Omega)\} \tag{19}$$

still has a unique solution y in $H^1(\Omega)$ which coincides with the solution of the boundary value problem

$$-\operatorname{div} \nabla y + y - \operatorname{div} f = 0 \quad \text{in } \Omega, \qquad \left(\frac{\partial y}{\partial n}\right) = 0 \quad \text{on } \Gamma. \tag{20}$$

As in § 7.1 we introduce the new functionals

$$\underline{E}(t, \varphi) = E(\Omega_t, \varphi \circ T_t^{-1}), \quad \underline{F}(t, \varphi) = F(\Omega_t, \varphi \circ T_t^{-1}), \quad J(t) = J(\Omega_t) \tag{21}$$

and transport all the integrals from Ω_t to Ω. We are now back to the setup at the end of § 5, and Theorem 7 and Proposition 8 apply with $u = t$.

As a result

$$dJ(\Omega;\, V) = (d/dt)J(t)|_{t=0}$$
$$= \operatorname{Sup}\{d\underline{F}(0, y; 1, 0) + d^2\underline{E}(0, y; 0, p; 1, 0) : p \in P(0)\} \tag{22}$$

where y is the solution of

$$d\underline{E}(0, y; 0, \psi) = 0 \quad \forall \psi \in H^1(\Omega) \tag{23}$$

and $P(0)$ is the set of solutions of the adjoint variational inequality

$$d\underline{F}(0, y; 0, \psi) + d^2\underline{E}(0, y; 0, p; 0, \psi) \geqq 0 \quad \forall \psi \in H^1(\Omega). \tag{24}$$

Note that the set $P(0)$ is not empty since the hypotheses of Proposition 8 are verified. However, the elements of $P(0)$ belong to $H^1(\Omega)$ but again not much more. In fact (24) reduces to

$$\int_{\Omega_+} (\nabla y/|\nabla y|) \cdot \nabla \psi \, dx + \int_{\Omega_0} |\nabla \psi| \, dx + \int_{\Omega} [\nabla p \cdot \nabla \psi + p\psi] \, dx \geqq 0, \tag{25}$$
$$p \in H^1(\Omega) \quad \forall \psi \in H^1(\Omega)$$

where

$$\Omega_+ = \{x \in \Omega: \nabla y(x) \neq 0\}, \qquad \Omega_0 = \{x \in \Omega: \nabla y(x) = 0\}. \tag{26}$$

It is readily seen that (25) has at least one solution since the following variational equation has a unique solution:

$$\int_{\Omega_+} (\nabla y/|\nabla y|) \cdot \nabla \psi \, dx + \int_{\Omega} [\nabla p \cdot \nabla \psi + p\psi] \, dx = 0, \qquad p \in H^1(\Omega) \quad \forall \psi \in H^1(\Omega). \tag{27}$$

Of course, as in § 5, this problem can be solved by replacing the nondifferentiable cost function (18) by a Sup

$$F(\Omega, \varphi) = \operatorname{Sup}\{F(\Omega, \mu, \varphi) : \varphi \in M\}$$

of the functional

$$F(\Omega, \mu, \varphi) = \int_{\Omega} \mu \cdot \nabla \varphi \, dx$$

over the weakly compact subset

$$M = \{\mu \in (L^2(\Omega))^n : |\mu(x)| \leqq 1 \text{ a.e. in } \Omega\}$$

of $(L^2(\Omega))^n$ where $|\ |$ denotes the Euclidean norm in R^n. Instead of the variational inequality (25) we would obtain the set of variational equations

$$p = p(\alpha) \in H^1(\Omega) \quad \forall \psi \in H^1(\Omega),$$
$$\int_{\Omega_+} (\nabla y/|\nabla y|) \cdot \nabla \psi \, dx + \int_{\Omega_0} \alpha \cdot \nabla \psi \, dx + \int_{\Omega} [\nabla p \cdot \nabla \psi + p\psi] \, dx = 0$$

indexed by $\alpha \in M$ and the Sup in (22) would be taken over M:

$$dJ(\Omega;\, V) = \operatorname{Sup}\{d\underline{F}(0, \alpha, y; 1, 0, 0) + d^2\underline{E}(0, y; 0, p(\alpha); 1, 0) : \alpha \in M\}$$

where

$$\underline{F}(t, \alpha, \varphi) = F(t, \alpha \circ T_t^{-1}, \varphi \circ T_t^{-1}).$$

The proof is the same as the one given in § 5 for the nondifferentiable observation.

Note added in proof. The shape sensitivity analysis problem has also recently been studied by a penalization technique. The results apply to some classes of nonlinear problems and some problems governed by variational inequalities (see U. C. Delfour and J. P. Zolésio, Shape Sensitivity Analysis via a Penalization Method, Ann. Mat. Pura Appl., to appear).

REFERENCES

J. P. AUBIN [1], *L'analyse non linéaire et ses motivations économiques*, Masson, Paris, New York, 1984.

N. V. BANICHUK [1], *Optimization of the Shapes of Elastic Bodies*, Nauka, Moscow, 1980 (English translation 1984).

J. CÉA [1], *Conception optimale ou identification de formes. Calcul rapide de la dérivée directionnelle de la fonction coût*, R.A.I.R.O., 20 (1986), pp. 371-402.

——— [2], *Problems of shape optimal design*, in Optimization of Distributed Parameter Structures, Vol. II, E. J. Haug and J. Céa, eds., Sijthoff and Noordhoff, Alphen aan den Rijn, the Netherlands, 1981, pp. 1005-1048.

——— [3], *Numerical methods of shape optimal design*, in Optimization of Distributed Parameter Structures, Vol. II, E. J. Haug and J. Céa, eds., Sijthoff and Noordhoff, Alphen aan den Rijn, the Netherlands, 1981, pp. 1049-1087.

R. CORREA AND A. SEEGER [1], *Directional derivatives of a minimax function*, Nonlinear Anal. Theory Methods and Appl., 9 (1985), pp. 13-22.

M. C. DELFOUR AND J. P. ZOLÉSIO [1], *Dérivation d'un Min Max et application à la dérivation par rapport au contrôle d'une observation non différentiable de l'état*, C. R. Acad. Sci. Paris Sér. I Math., 302 (1986), pp. 571-574.

——— [2], *Differentiability of a Min Max and Application to Optimal Control and Design problems. Parts* I *and* II., in Control Problems for Systems described by Partial Differential Equations and Applications, I. Lasiecka and R. Triggiani, eds., Springer-Verlag, New York, Berlin, 1987, pp. 204-219, pp. 220-229.

K. DEMS AND Z. MROZ [1], *Variational approach by means of adjoint systems to structural optimization and sensitivity analysis, Part* 2. *Structure shape variations*, Internat. J. Solids and Structures, 20 (1984), pp. 527-552.

V. F. DEM'YANOV [1], *Differentiability of a Maximin function.* I, USSR Comput. Math. and Math. Phys., 8 (1968), pp. 1-15. (In English.) Z. Vychisl. Mat. i Mat. Fiz., 8 (1968), pp. 1186-1195. (In Russian.)

I. EKELAND AND R. TEMAM [1], *Analyse convexe et problèmes variationnels*, Dunod, Gauthiers-Villars, Paris, Brussels, Montreal, 1974.

R. H. GALLAGHER AND O. C. ZIENKIEWICZ [1], *Optimum Structural Design Theory and Applications*, John Wiley, New York, 1972; Springer-Verlag, Berlin, New York, 1973.

E. J. HAUG [1], *A review of distributed parameter structural optimization literature*, in Optimization of Distributed Parameter Structures, Vol. I, E. J. Haug and J. Céa, eds., Sijthoff and Noordhoff, Alphen aan den Rijn, the Netherlands, 1981, pp. 3-74.

E. J. HAUG AND J. S. ARORA [1], *Applied Optimal Design*, Wiley-Interscience, New York, 1979.

E. J. HAUG AND J. CÉA, EDS. [1], *Optimization of Distributed Parameter Structures, Vols.* I *and* II, Sijthoff and Noordhoff, Alphen aan den Rijn, the Netherlands, 1981.

E. J. HAUG, K. K. CHOI AND V. KOMKOV [1], *Design Sensitivity Analysis of Structural Systems*, Academic Press, New York, 1986.

A. MYSLINSKI AND J. SOKOLOWSKI [1], *Nondifferentiable optimization problems for elliptic problems*, SIAM J. Control Optim., 23 (1984), pp. 632-648.

O. PIRONNEAU [1], *Optimal Design for Elliptic Systems*, Springer-Verlag, Berlin, New York, 1984.

W. PRAGER [1], *Introduction to Structural Optimization, Courses and Lectures: International Centre for Mechanical Sciences*, Springer-Verlag, Vienna, 1974.

G. I. N. ROZVANY [1], *Optimal Design of Flexural Systems*, Pergamon Press, New York, 1976.

A. SAWCZUK AND Z. MROZ [1], *Optimization in Structural Design*, Springer-Verlag, Berlin, New York, 1975.

J. SOKOLOWSKI [1], *Optimal control in coefficients of boundary value problems with unilateral constraints*, Bull. Polish Acad. Sci. Tech. Sci., 31 (1983), pp. 71-81.

J. SOKOLOWSKI AND J. P. ZOLÉSIO [1], *Dérivée par rapport au domaine de la solution d'un problème unilatéral*, C. R. Acad. Sci. Paris Sér. I, 301 (1985), pp. 103-106.

——— [2], *Shape sensitivity analysis of an elasto-plastic torsion problem*, Bull. Polish Acad. Sci. Tech. Sci., 33 (1985), pp. 579-586.

J. P. ZOLÉSIO [1], *Identification de Domaine*, Thèse de doctorat d'état, Nice, France, 1979.

——— [2], *Semi-derivatives of repeated eigenvalues*, in Optimization of Distributed Parameter Structures, Vol. II, E. J. Haug and J. Céa, eds., Sijthoff and Noordhoff, Alphen aan den Rijn, the Netherlands, 1981, pp. 1457-1473.

——— [3], *An optimal design procedure for optimal control support*, in Convex Analysis and its Applications, A. Auslender, ed., Springer-Verlag, Berlin, New York, 1977, pp. 207-219.