SIAM J. CONTROL AND OPTIMIZATION
Vol. 28, No. 1, pp. 113-136, January 1990

006

THE AUGMENTED LAGRANGIAN METHOD FOR PARAMETER ESTIMATION IN ELLIPTIC SYSTEMS*

KAZUFUMI ITO† AND KARL KUNISCH‡

Abstract. In this paper a new technique for the estimation of parameters in elliptic partial differential equations is developed. It is a hybrid method combining the output-least-squares and the equation error method. The new method is realized by an augmented Lagrangian formulation, and convergence as well as rate of convergence proofs are provided. Technically the critical step is the verification of a coercivity estimate of an appropriately defined Lagrangian functional. To obtain this coercivity estimate a seminorm regularization technique is used.

Key words. augmented Lagrangian method, parameter estimation, least squares, elliptic system

AMS(MOS) subject classifications. 35R30, 49D29

1. Introduction. In this paper we consider the problem of determining the unknown functional coefficient q in the elliptic partial differential equation

$$-\operatorname{div}(q \operatorname{grad} u)=f \quad \text{in } \Omega \qquad u=0 \quad \text{on } \Gamma, \tag{1.1}$$

from an observation z of the solution u, where Ω is a bounded domain in $\mathbb{R}^n$, $n=1,2$, or 3, with piecewise smooth boundary Γ and $f\in H^{-1}$ is given. In applications, the function z might be constructed by interpolation of pointwise measurements. We propose and analyze a hybrid method that combines the output least squares and the equation error formulation [2], [17] within the mathematical framework given by the augmented Lagrangian technique.

The output least squares (OLS) approach is used most commonly and in our example for $n=2$ or 3, it is stated, for instance, as the minimization problem in H^2:

$$\text{Minimize} \quad \frac{1}{2}|u(q)-z|_H^2+\frac{\beta}{2}N(q) \tag{1.2}$$

$$\text{over} \quad Q_{\text{ad}}=\{q\in H^2(\Omega)\colon q\geqq \alpha \text{ and } |q|_{H^2}\leqq \gamma\}$$

where α and γ are positive constants chosen a priori, $u(q)$ is the solution of (1.1), and H is chosen as H^i, $i=0$ or 1, for example. The second term in the cost functional represents a regularization term and Q_{ad} is chosen so that (1.2) has a solution for every $\beta\geqq 0$. The use of a regularization term guarantees the continuity of the mappings from the observation $z\in H$ ($H=H_0^1$ or L^2, for example) to a solution $q^\beta(z)\in Q_{\text{ad}}\subset \mathrm{H}^2$ for an appropriate choice of N and $\beta>0$ and, in general, β cannot be taken equal to zero [3], [4], [9], [17]. In this paper we will use a regularization term such that $(N(q))^{1/2}$ is a seminorm on H^2. The use of seminorm regularization is very common for the inversion of linear operators [6], but it is not well studied in nonlinear problems such

* Received by the editors December 21, 1987; accepted for publication (in revised form) January 25, 1989.

† Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912. The research of this author was supported in part by U.S. Air Force Office of Scientific Research grants AFOSR-84-0398 and AFOSR-85-0303, National Aeronautics and Space Administration grant NAG-1-1517, and National Science Foundation grant UINT-8521208.

‡ Institut für Mathematik, Technische Universität Graz, Graz, Austria. The research of this author was supported in part by the Fonds zur Förderung der wissenschaftlichen Forschung, under grants S3206 and P6005, and by the U.S. Air Force Office of Scientific Research grant AFOSR-84-0398. Part of this work was performed while the author was visiting the Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912.

as the one presented by estimating q in (1.1). The OLS formulation is quite flexible with regard to the availability of the data. The OLS term in (1.2) can be adjusted in case data are only available over a subset of the domain, or are given as point measurements or as measurements of the flux at the boundary. As its form indicates, the OLS approach is less sensitive with respect to noise in the data when compared to the equation error method to be specified below. However, the minimization in (1.2) is an indirect method to determine the unknown q, and any iterative algorithm for solving (1.2) requires the solution of (1.1) for every update of $q \in Q_{\mathrm{ad}}$.

An alternative to the OLS formulation is the equation error formulation. For our problem it can be stated as follows:

$$\text{(1.3)} \qquad \text{Minimize} \quad \tfrac{1}{2}|\nabla \cdot (q\nabla z)+f|_H^2$$
$$\text{subject to} \quad q \in H^2, \quad q \geqq \alpha$$

where H is either $H^{-1}(=(H_0^1)^*)$ or H^0. Since in computations H^{-1} requires a lesser amount of numerical differentiations of z, it should be preferred. An obvious disadvantage of this formulation is that it needs a fairly accurate observation of z defined over the entire domain Ω and it may be sensitive to noise in the data. On the other hand, it leads to efficient algorithms, since the minimization in (1.3) is quadratic.

The hybrid method that we propose not only combines both these formulations, but it also inherits the flexibility of the OLS approach and the quadratic structure of the equation error approach. This is achieved by viewing (1.2) as the following constrained minimization problem:

$$\text{(1.4)} \qquad \text{minimize} \quad F(q,u)=\frac{1}{2}|u-z|_{H_0^1}^2+\frac{\beta}{2}N(q)$$

subject to

$$\text{(1.5)} \qquad -\nabla \cdot (q\nabla u)=f \quad \text{in } H^{-1},$$

$$\text{(1.6)} \qquad |q|_{H^2} \leqq \gamma,$$

$$\text{(1.7)} \qquad \alpha \leqq q \quad \text{on } \Omega,$$

in the two independent variables q and u. To solve (1.4)-(1.7) we apply the augmented Lagrangian algorithm (see, e.g., [1], [7], [8], [16]). It essentially involves minimizing a sequence of functionals of the form

$$\text{(1.8)} \qquad L_{c_k}(q,u;\lambda^k)=F(q,u)+\langle \lambda^k, e(q,u)\rangle_{H_0^1}+\frac{c_k}{2}|e(q,u)|_{H_0^1}^2 \quad \text{over } q \in Q_{\mathrm{ad}},$$

and the multiplier sequence $\{\lambda^k\}$ in H_0^1 is given by

$$\text{(1.9)} \qquad \lambda^{k+1}=\lambda^k+c_k e(q_k,u_k),$$

where $\Delta: H_0^1 \to H^{-1}$ is the Laplacian, the function $e: H^2 \times H_0^1 \to H_0^1$ is defined by

$$\text{(1.10)} \qquad e(q,u)=(-\Delta)^{-1}(\nabla \cdot (q\nabla u)+f),$$

and the pair (q_k, u_k) minimizes (1.8). To carry out this iterative scheme a (possibly constant) sequence of positive real numbers $\{c_k\}$ and a startup $\lambda^1 \in H_0^1$ for the Lagrange multiplier need to be chosen. We suggest $\lambda^1=0$ but convergence will be guaranteed for any other choice of λ^1 as well. The inequality constraint $|q|_{H^2} \leqq \gamma$ (see (1.6)) can be augmented in a manner similar to the equality constraint $e(q,u)=0$ (see § 2 for details). Convergence of this algorithm will be shown in Theorem 2.2 by employing the results of [8], where a general framework for the analysis of the augmented Lagrangian method in infinite-dimensional spaces with equality constraints, as well as inequality constraints with finite-dimensional image space, is given. More precisely, convergence and rate of convergence of the pair (q_k, u_k) to a solution (q^β, u^β) of (1.4)–(1.7) in $H^2 \times H_0^1$, as well as of λ^k to λ^*, the Lagrange multiplier associated with the equality constraint (1.5), in H_0^1 will be proved. This result will be obtained under the assumption that the H^1-error between the data z and the nonregularized OLS-solution u^0 is sufficiently small and that the penalty parameters c_k are sufficiently large. It is not required that $\lim c_k = \infty$ as $k \to \infty$.

A number of remarks are in order.

(1) The minimization of the function in (1.8) requires the solution of a Poisson equation. For the discretized problem several efficient numerical techniques are readily available and any variant can be chosen. As a comparison, in the OLS approach (1.1) must be solved for $u = u(q)$ whenever a change in q occurs.

(2) Note that

$$|\varphi|_{H^{-1}}^2 = \langle(-\Delta)^{-1}\varphi, \varphi\rangle = |(-\Delta)^{-1}\varphi|_{H_0^1}^2 \quad \text{for all } \varphi \in H^{-1} = (H_0^1)^*.$$

Thus the minimization of the cost functional in (1.8) is a combination of the OLS-problem (1.2) and the equation error problem (1.3) where $H = H^{-1}$, with the aid of the multiplier method. The choice of the H_0^1-topology for the OLS-term and the H^{-1}-topology for the equation error term (equality constraint (1.5)) is natural from the point of view of the second-order sufficient optimality condition for (1.4) that will be used below, and the choice of these topologies leads to a method that requires the same amount of numerical differentiations in both the OLS and the equation error term.

(3) Note that $e(q, u)$ is a bilinear function in q and u. Thus for fixed q (respectively, u) (1.8) becomes quadratic in u (respectively, q) and we can take advantage of this structure numerically (see § 4 for details).

(4) As will be shown the penalty term $(c/2)|e(q, u)|_{H_0^1}^2$ enhances the convexity in the neighborhood of local minima.

(5) There is some arbitrariness in the choice of the topologies for the OLS and the equation error term. We can use $|\nabla(q\nabla u)+f|^2$ instead of $|e(q, u)|_{H_0^1}^2$ while simultaneously $|u-z|_{H_0^1}^2$ is replaced by $|u-z|_{H_0^1\cap H^2}^2$. This would require a different analysis of the coercivity estimate (see § 3) and would lead to a different numerical implementation. This aspect is not exploited further within this paper.

The paper is organized as follows. In § 2 we describe in detail the augmented Lagrangian algorithm for the estimation of q in (1.1) and state the convergence results in Theorems 2.2 and 2.3. The essential technical tool that guarantees convergence is the positivity of the second Fréchet-derivative of the Lagrange functional (see (2.4) below). This coercivity condition is analyzed for various situations (Propositions 3.4 and 3.5) in § 3. Section 4 is devoted to a brief summary of the numerical experience that has been obtained with the augmented Lagrangian algorithm in parameter estimation. A comprehensive study of our numerical experience will appear elsewhere.

The notation that we use is rather standard. Unless otherwise specified, all function spaces are considered over the domain Ω. We use $\langle\cdot,\cdot\rangle$ to denote the inner product in L^2 and $|\cdot|$ to denote the norm in L^2 and $\mathbb{R}^n$, $n \geqq 1$. For other inner products and norms we use an index, as for instance $|v|_{H^1}$ denotes the common norm in H^1. The space $\mathbb{R}^n$ is endowed with the Euclidean norm. The inner product in H_0^1 is given by $\langle v, w\rangle_{H_0^1} = \langle\nabla v, \nabla w\rangle$ and the associated norm is defined through $|v|_{H_0^1}^2 = \langle v, v\rangle_{H_0^1}$.

2. Problem formulation and convergence results. Let us formulate the problem for the multidimensional case ($n = 2$ or 3) first. We consider the constrained minimization problem in the variables $(q, u) \in H^2 \times H_0^1$:

$$(\mathrm{P}^\beta) \qquad \text{Minimize} \quad \frac{1}{2}|u-z|_{H_0^1}^2 + \frac{\beta}{2}\left(\sum_{i_1,i_2} |q_{x_{i_1}x_{i_2}}|^2 + |\nabla q|^2\right),$$

$$\text{subject to} \quad e(q,u) = (-\Delta)^{-1}(\nabla \cdot (q\nabla u) + f) = 0 \text{ in } H_0^1, \text{ and}$$

$$|q|_{H^2} \leqq \gamma, q \geqq \alpha \quad \text{on } \Omega$$

where Δ is the Laplacian (considered as operator from H_0^1 onto H^{-1}), $f \in H^{-1}$, $z \in H_0^1$, $\beta \geqq 0$, and α, γ are given constants satisfying $\alpha^2 \int_\Omega dx < \gamma^2$. Note that (q^β, u^β) is a global solution of (P^β) if and only if q^β is a solution of (1.2) with $u(q^\beta) = u^\beta$. To argue existence of a solution q^β of (1.2) observe that Q_{ad} is a bounded, closed, and convex subset of H^2 and hence it is weakly sequentially compact. The parameter-to-solution mapping $q \to u(q)$, $q \in Q_{\mathrm{ad}}$, is continuous from the space of continuous functions C to the H_0^1-topology. Moreover, H^2 is compactly embedded in C and norms are weakly lower semicontinuous functionals. Hence it can be shown that for any $\beta \geqq 0$ there exists at least one solution to (1.2) or equivalently (P^β).

Subsequently we will use the closed convex cone $\mathscr{C}$ with vertex at the origin of H^2 defined by

$$\mathscr{C} = \{w \in H^2 \colon w \leqq 0\}.$$

Let

$$\mathscr{C}^+ = \{\phi \in H^2 \colon \langle \phi, h\rangle_{H^2} \leqq 0 \text{ for all } h \in \mathscr{C}\}$$

be the positive dual cone and let

$$\mathbb{R}^- = \{x \in \mathbb{R} \colon x \leqq 0\}.$$

Then (P^β) can be written as follows:

$$(2.1) \qquad \text{Minimize} \quad F(q,u) = \frac{1}{2}|u-z|_{H_0^1}^2 + \frac{\beta}{2} N(q)$$

$$\text{subject to} \quad e(q,u) = (-\Delta)^{-1}(\nabla \cdot (q\nabla u) + f) = 0,$$

$$g(q) = \tfrac{1}{2}(|q|_{H^2}^2 - \gamma^2) \in \mathbb{R}^-,$$

$$l(q) = \alpha - q \in \mathscr{C}.$$

Henceforth (q^β, u^β) denotes a solution of (2.1). The next theorem (that will be proved in the latter part of this section) shows the existence and the uniqueness of a Lagrange multiplier $(\lambda^*, \mu^*, \eta^*) \in H_0^1 \times R^+ \times \mathscr{C}^+$ associated with a solution (q^β, u^β) of (P^β). We suppress the dependence of $(\lambda^*, \mu^*, \eta^*)$ on β.

THEOREM 2.1. *There exists a Lagrange multiplier* $(\lambda^*, \mu^*, \eta^*) \in H_0^1 \times \mathbb{R}^+ \times \mathscr{C}^+$ *such that*

$$L(q, u; \lambda^*, \mu^*, \eta^*) = F(q,u) + \langle \lambda^*, e(q,u)\rangle_{H_0^1} + \frac{\mu^*}{2}(|q|_{H^2}^2 - \gamma^2) + \langle \eta^*, \alpha - q\rangle_{H^2}$$

satisfies

$$(2.2\text{a}) \qquad \nabla L(q^\beta, u^\beta; \lambda^*, \mu^*, \eta^*)(h, v) = 0 \quad \text{for all } (h,v) \in H^2 \times H_0^1$$

$$(2.2\text{b}) \qquad \mu^*(|q^\beta|_{H^2}^2 - \gamma^2) = 0, \qquad \langle \eta^*, \alpha - q^\beta\rangle_{H^2} = 0.$$

Moreover, the Lagrange multiplier is unique and $\lambda^* = \Delta(q^\beta)^{-1}\Delta(u^\beta - z)$ *in* H_0^1 *where* $\Delta(q)u = \nabla \cdot (q\nabla u)$. *Here* $\nabla L(q^\beta, u^\beta; \lambda^*, \mu^*, \eta^*)(h,v)$ *denotes the Fréchet derivative of* $L(\cdot, \cdot; \lambda^*, \mu^*, \eta^*)$ *at* (q^β, u^β) *in direction* $(h, v) \in H^2 \times H_0^1$.

To solve (P^β) (or equivalently (2.1)) we will apply the augmented Lagrangian method. This method, due to Hestenes and Powell, has been studied extensively in the finite-dimensional case and, with equality constraints only, also in the infinite-dimensional case (cf. [1], [5], [7], and [16], for example). The infinite-dimensional case with equality as well as inequality constraints has been studied in [8]. To explain this method we require several reformulations of (P^β). For $c \geqq 0$ consider the augmented problem:

$$(\mathrm{P})_c \qquad \text{Minimize} \quad F(q,u)+\frac{c}{2}|e(q,u)|^2_{H_0^1}+\frac{c}{2}|g(q)+w|^2$$

$$\text{subject to} \quad e(q,u)=0 \quad \text{in } H_0^1,$$

$$g(q)+w=0, \qquad w\in\mathbb{R},$$

$$w \geqq 0,$$

$$\alpha-q\in\mathscr{C}.$$

In the notation of $(\mathrm{P})_c$ as well as F we suppress the dependence on β. We observe that (q^β, u^β) is a solution of (P^β) if and only if $(q^\beta, u^\beta, w^\beta=-g(q^\beta))$ is a solution of $(\mathrm{P})_c$. Moreover, it is simple to verify that $(\lambda^*, \mu^*, \mu^*, \eta^*)$ is a Lagrange multiplier for $(\mathrm{P})_c$, i.e.,

$$\nabla L_c(q^\beta, u^\beta, w^\beta; \lambda^*, \mu^*, \eta^*)=0 \quad \text{and} \quad \langle \eta^*, \alpha-q^\beta\rangle_{H^2}=0, \quad \mu^* w^\beta=0$$

where the Lagrangian $L_c(q, u, w; \lambda^*, \mu^*, \eta^*)$ is given by

$$\begin{aligned} L_c(q,u,w;\lambda^*,\mu^*,\eta^*) &= F(q,u)+\langle\lambda^*, e(q,u)\rangle_{H_0^1}+\mu^* g(q) \\ &\quad +\langle \eta^*, \alpha-q\rangle_{H^2}+\frac{c}{2}|e(q,u)|^2_{H_0^1}+\frac{c}{2}|g(q)+w|^2, \end{aligned} \tag{2.3}$$

and f, e, g are defined in (2.1). It can also be shown that any solution $(q^\beta, u^\beta, w^\beta)$ of $(\mathrm{P})_c$ is a regular point in the sense of ([14], cf. also Theorem 2.1 and its proof), but since we will not use this fact we do not give its proof here.

Henceforth the following second-order sufficient optimality condition will be used:

> There exist constants $\sigma>0$ and $c_0 \geqq 0$ such that the second Fréchet derivative $\nabla^2 L_{c_0}$ of L_{c_0} satisfies the coercivity condition
>
> $$\nabla^2 L_{c_0}(q^\beta, u^\beta, w^\beta; \lambda^*, \mu^*, \eta^*)((h,v,y)(h,v,y)) \geqq \sigma(|h|^2_{H^2}+|v|^2_{H_0^1}+|y|^2) \tag{2.4}$$
>
> for all $(h, v, y)\in H^2\times H_0^1\times\mathbb{R}$, where the Lagrangian L_c is defined by (2.3) and $w^\beta=-g(q^\beta)$.

In the next section we will establish this condition for several specific cases. Under (2.4), it can be shown that $(q^\beta, u^\beta, w^\beta)$ is a solution of $(\mathrm{P})_c$ if and only if it is a solution of

$$\min F(q,u)+\langle\lambda^*, e(q,u)\rangle_{H_0^1}+\mu^*(g(q)+w)+\frac{c}{2}|e(q,u)|^2_{H_0^1}+\frac{c}{2}|g(q)+w|^2, \tag{2.5}$$

with $w\geqq 0$, $\alpha \leqq q$. In (2.5) the equality constraint and the inequality constraint with finite-dimensional image space are eliminated from the explicit constraints. However, (2.5) contains the unknown Lagrange multipliers (λ^*, μ^*). The augmented Lagrangian algorithm applied to (P^β) involves solving iteratively for (λ^*, μ^*) and (q^β, u^β) and requires the solution of the following minimization problem:

(2.6) Given $\lambda \in H_0^1$ and $\mu \in \mathbb{R}^+$

$$\text{minimize} \quad F(q,u)+\langle\lambda, e\rangle_{H_0^1}+\mu(g(q)+w)+\frac{c}{2}|e(q,u)|^2_{H_0^1}+\frac{c}{2}|g(q)+w|^2$$

$$\text{subject to} \quad w\geqq 0,\ \alpha\leqq q \text{ and } (q,u,w)\in H^2\times H_0^1\times\mathbb{R}.$$

This problem is equivalent to the problem of minimizing

$$(2.7)\qquad F(q,u)+\langle\lambda, e\rangle_{H_0^1}+\mu\hat{g}(q,\mu,c)+\frac{c}{2}|e|^2_{H_0^1}+\frac{c}{2}\hat{g}(q,\mu,c)^2$$
$$=F(q,u)+\langle\lambda, e\rangle_{H_0^1}+\frac{c}{2}|e(q,u)|^2_{H_0^1}+\frac{1}{2c}(|\max(0, cg+\mu)|^2-|\mu|^2)$$

subject to $\alpha\leqq q$, where the constraint $w\geqq 0$ is eliminated. Here we put

$$\hat{g}(q,\mu,c)=\max\left(-\frac{\mu}{c}, g(q)\right)$$

and we used the equality

$$c\hat{g}(q,\mu,c)+\mu=\max(-\mu, cg(q))+\mu=\max(0, cg(q)+\mu).$$

Observe that $(q^\beta, u^\beta, w^\beta=\max(0, -g(q^\beta)-\mu/c))$ is a solution of (2.6) if and only if (q^β, u^β) is a solution of (2.7).

We are now prepared to specify the augmented Lagrangian algorithm to solve $(P)^\beta$. Choose a monotonically increasing sequence $\{c_k\}$ of positive real numbers with $c_1>c_0$ and $(\lambda^1, \mu^1)\in H_0^1\times\mathbb{R}^+$. In practice we suggest choosing $(\lambda^1,\mu^1)=(0,0)$, where the choice of λ^1 is based on Theorem 2.1: λ^* is close to 0 if u^β is close to z.

For $k\geqq 1$ determine (q_k, u_k) by solving the following:

$$(2.8)\qquad \text{Minimize}\quad \mathscr{L}_k(q,u),$$
$$\text{subject to}\quad (q,u)\in H^2\times H_0^1,\ \alpha\leqq q,$$

and define

$$(2.9)\qquad \lambda^{k+1}=\lambda^k+(c_k-c_0)e(q_k,u_k),\qquad \mu^{k+1}=\mu^k+(c_k-c_0)\hat{g}(q_k,\mu^k,c_k)$$

where

$$\mathscr{L}_k(q,u)=F(q,u)+\langle\lambda^k, e\rangle_{H_0^1}+\frac{c_k}{2}|e|^2_{H^1}+\frac{1}{2c_k}(|\max(0, c_kg(q)+\mu^k)|^2-|\mu^k|^2).$$

In the following result we will ascertain local convexity of the cost functional appearing in (2.7) and (2.8) in a closed ball $\bar{B}$ containing the solution (q^β, u^β) of (P^β). We call (q_k,u_k) a solution of (2.8) in $\bar{B}$ if $\mathscr{L}_k(q_k,u_k)\leqq\mathscr{L}_k(q,u)$ for all $(q,u)\in\bar{B}$ with $\alpha\leqq q$. Existence of a solution of (2.8) in $\bar{B}$ and a Lagrange multiplier η^k associated with the inequality constraint $\alpha\leqq q$ can easily be verified. We will prove convergence of the solutions $(q_k,u_k)\in\bar{B}$ of (2.8) as $k_{\to\infty}$.

It is useful to observe that

$$\mu^{k+1}=\max\left(\frac{\mu^k c_0}{c_k}, \mu^k+(c_k-c_0)g(q_k)\right)$$

and, since $\mu^1\geqq 0$, this implies that $\mu^k\geqq 0$ for all $k\geqq 1$.

THEOREM 2.2. (a) *Suppose that the coercivity condition* (2.4) *holds. Then for every* $r \geqq \mu^*$ *there exist constants* $\tilde{c} = \tilde{c}(r) > c_0$ *and* $\bar{\sigma} > 0$, *and an open bounded ball* B *in* $H^2 \times H_0^1$ *centered at* (q^β, u^β) *such that*

$$F(q,u) + \langle \lambda^*, e\rangle_{H_0^1} + \mu^* \hat{g}(q,\mu,c) + \langle \eta^*, \alpha - q\rangle_{H^2} + \frac{c_0}{2}|e|^2_{H_0^1} + \frac{c_0}{2}|\hat{g}(q,\mu,c)|^2 - F(q^\beta, u^\beta)$$

$$\geqq \bar{\sigma}(|q - q^\beta|^2_{H^2} + |u - u^\beta|^2_{H_0^1})$$

for all $(q,u) \in \bar{B}$, $c \geqq \tilde{c}$ *and* $\mu \in [0, r]$, *where* $\hat{g}(q,\mu,c) = \max(-\mu/c, g(q))$.

(b) *Suppose that in addition* $r \geqq \mu^* + (|\lambda^1 - \lambda^*|^2 + |\mu^1 - \mu^*|^2)^{1/2}$, *and* $c_1 \geqq \tilde{c}(r)$ *is chosen sufficiently large. Then every solution of* (2.8) *in* $\bar{B}$ *satisfies* $(q_k, u_k) \in B$, $\mu^k \in [0, r]$ *and*

$$\begin{aligned} &\bar{\sigma}(|q_k - q^\beta|^2_{H^2} + |u_k - u^\beta|^2_{H_0^1}) + \frac{1}{2(c_k - c_0)}(|\lambda^{k+1} - \lambda^*|^2_{H_0^1} + |\mu^{k+1} - \mu^*|^2) \\ &\leqq \frac{1}{2(c_k - c_0)}(|\lambda^k - \lambda^*|^2_{H_0^1} + |\mu^k - \mu^*|^2), \end{aligned} \tag{2.10}$$

for every $k \geqq 1$. *Moreover, there exists a constant* $K > 0$ *such that*

$$\begin{aligned} &|q_k - q^\beta|^2_{H^2} + |u_k - u^\beta|^2_{H_0^1} \leqq \frac{1}{2\bar{\sigma}(c_k - c_0)}(|\lambda^k - \lambda^*|^2_{H_0^1} + |\mu^k - \mu^*|^2) \quad \textit{for } k \geqq 1, \\ &|\lambda^k - \lambda^*|^2_{H_0^1} + |\mu^k - \mu^*|^2 + |\eta^{k-1} - \eta^*|^2_{H^2} \\ &\leqq \left(\frac{K}{\bar{\sigma}}\right)^{k-1} \prod_{i=1}^{k-1} \frac{1}{c_i - c_0}(|\lambda^1 - \lambda^*|^2_{H_0^1} + |\mu^1 - \mu^*|^2) \quad \text{for } k \geqq 2. \end{aligned} \tag{2.11}$$

The proof of Theorem 2.2 will be given in the latter part of this section.

Next we formulate the problem for the one-dimensional case. Without loss of generality we can assume that $\Omega = (0, 1)$. We take $(q, u) \in H^1 \times H_0^1$ and the regularization term $N(q)$ is chosen as

$$N(q) = \int_0^1 |q_x|^2\, dx.$$

Thus, for $\beta \geqq 0$, the analogue of (P^β) is defined as follows:

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}|u - z|^2_{H_0^1} + \frac{\beta}{2}|q_x|^2 \\ &\text{subject to} \quad e(q,u) = (-\Delta)^{-1}((qu_x)_x + f) = 0 \quad \text{in } H_0^1, \\ &\qquad |q|_{H^1} \leqq \gamma \quad \text{and} \quad \alpha \leqq q \text{ on } [0,1] \end{aligned} \tag{2.12}$$

where $\Delta: H_0^1 \to H^{-1}$ is given by $\Delta u = u_{xx}$. The results corresponding to Theorem 2.1 hold with $q \in H^2$ replaced by $q \in H^1$. In particular, if $(q^\beta, u^\beta) \in H^1 \times H_0^1$ is a solution of (2.12), then there exists a unique Lagrange multiplier $(\lambda^*, \mu^*, \eta^*) \in H_0^1 \times R^+ \times \mathscr{C}^+$ such that the Lagrangian

$$L(q; u; \lambda^*, \mu^*, \eta^*) = \tfrac{1}{2}|u - z|^2_{H_0^1} + \frac{\beta}{2}|q_x|^2 + \langle \lambda^*, e\rangle_{H_0^1} + \frac{\mu}{2}(|q|^2_{H_0^1} - \gamma^2) + \langle \eta^*, \alpha - q\rangle_{H^1}$$

satisfies $\nabla L(q^\beta, u^\beta; \lambda^*, \mu^*, \eta^*) = 0$ and $\mu^*(|q^\beta|^2_{H_0^1} - \gamma^2) = \langle \eta^*, \alpha - q^\beta\rangle_{H^1} = 0$, where $\langle \cdot, \cdot \rangle_{H^1}$ is the duality pairing on H^1, $\mathscr{C} = \{h \in H^1 : h \leqq 0\}$ and $\mathscr{C}^+ = \{\varphi \in H^1 : \langle \varphi, h\rangle_{H^1} \leqq 0$

for all $h \in \mathscr{C}\}$ is the positive dual cone of $\mathscr{C}$. The Lagrange multiplier associated with the equality constraint can be expressed as

$$\lambda^* = \Delta(q^\beta)^{-1}\Delta(u^\beta - z) \quad \text{in } H_0^1$$

where $\Delta(q^\beta): H_0^1 \to H^{-1}$ is given by $\Delta(q^\beta)u = (q^\beta u_x)_x$. The Lagrangian for the augmented problem (compare (2.3)) is given by

$$\begin{aligned} L_c(q, u, w; \lambda^*, \mu^*, \eta^*) &= \tfrac{1}{2}|u - z|^2_{H_0^1} + \frac{\beta}{2}|q_x|^2 + \langle \lambda^*, e\rangle_{H_0^1} \\ &\quad + \langle \eta^*, \alpha - q\rangle_{H^1} + \mu^* g(q) + \frac{c}{2}|e|^2_{H_0^1} + \frac{1}{2}|g(q) + w|^2 \end{aligned} \tag{2.13}$$

where $e = e(q, u)$, and $g(q) = \frac{1}{2}(|q|^2_{H^1} - \gamma^2)$.

The augmented Lagrangian method for the solution of (2.12) now proceeds precisely as in the multidimensional case described in (2.8), (2.9) with $q \in H^2$ replaced by $q \in H^1$, and the regularization term is chosen as $\beta|q_x|^2$. In particular, $\hat{g}(q, \mu, c) = \max(-\mu/c, \frac{1}{2}(|q|^2_{H^1} - \gamma))$ and (2.8) becomes

$$\text{(2.14)} \qquad \begin{aligned} &\text{Minimize} \quad \mathscr{L}_k(q, u) \\ &\text{subject to} \quad (q, u) \in H^1 \times H_0^1, \alpha \leqq q, \end{aligned}$$

where

$$\begin{aligned} \mathscr{L}_k(q, u) &= \frac{1}{2}|u - z|^2_{H_0^1} + \frac{\beta}{2}|q_x|^2 + \langle \lambda^k, e\rangle_{H_0^1} + \frac{c}{2}|e|^2_{H_0^1} \\ &\quad + \frac{1}{2c_k}(|\max(0, c_k g(q) + \mu^k)|^2 - |\mu^k|^2). \end{aligned}$$

The following analogue of Theorem 2.2 holds for the one-dimensional case.

THEOREM 2.3. *Suppose that there exist constants $\sigma > 0$ and $c_0 \geqq 0$ such that the second Fréchet derivative $\nabla^2 L_{c_0}$ of L_{c_0} satisfies*

$$\nabla^2 L_{c_0}(q^\beta, u^\beta, w^\beta; \lambda^*, \mu^*, \eta^*)((h, v, y), (h, v, y)) \geqq \sigma(|h|^2_{H^1} + |v|^2_{H_0^1} + |y|^2) \tag{2.15}$$

for all $(h, v, y) \in H^1 \times H_0^1 \times \mathbb{R}$. Then for every $r \geqq \mu^*$ there exist constants $\bar{c} = \bar{c}(r) > c_0$ and $\bar{\sigma} > 0$, and an open ball B in $H^1 \times H_0^1$ centered at (q^β, u^β) such that

$$\begin{aligned} &\frac{1}{2}|u - z|^2_{H_0^1} + \frac{\beta}{2}|q_x|^2 + \langle \lambda^*, e\rangle_{H_0^1} + \mu^* \hat{g}(q, \mu, c) + \langle \eta^*, \alpha - q\rangle_{H^1} + \frac{c_0}{2}|e|^2_{H_0^1} + \frac{c_0}{2}|\hat{g}(q, \mu, c)|^2 \\ &\quad \geqq \frac{1}{2}|u^\beta - z|^2_{H_0^1} + \frac{\beta}{2}|q_x^\beta|^2 + \bar{\sigma}(|q - q^\beta|^2_{H^1} + |u - u^\beta|^2_{H_0^1}) \end{aligned}$$

for all $(q, u) \in \bar{B}$, $c \geqq \bar{c}$, and $\mu \in [0, r]$. Similarly, the assertions analogous to Theorem 2.2(b) *hold with H^2 replaced by H^1.*

We now come to the proofs of the results of this section.

Proof of Theorem 2.1. Let $\tilde{M}: H^2 \times H_0^1 \to H_0^1 \times \mathbb{R} \times H^2$ be defined by

$$\tilde{M}(h, v) = ((-\Delta)^{-1}(\nabla \cdot (q^\beta \nabla v + h \nabla u^\beta)), \langle q^\beta, h\rangle_{H^2}, -h).$$

The Fréchet derivatives ∇e, ∇g, and ∇l at the minimizer (q^β, u^β) in direction (h, v) are given by

$$\begin{aligned} &\nabla e(q^\beta, u^\beta)(h, v) = (-\Delta)^{-1}(\nabla \cdot (q^\beta \nabla v + h \nabla u^\beta)) \in H_0^1, \\ &\nabla g(q^\beta, u^\beta)(h, v) = \langle q^\beta, h\rangle_{H^2} \in \mathbb{R}, \qquad \nabla l(q^\beta, u^\beta)(h, v) = -h \in H^2. \end{aligned}$$

These are the coordinates of $\tilde{M}$. The existence of a Lagrange multiplier satisfying (2.2) will follow directly from the regular point condition [14, p. 100] that for the problem under consideration is given by

$$\begin{aligned}\{\tilde{M}(h, v)+(0, r, k)+\lambda(0, \tfrac{1}{2}(|q^\beta|^2_{H^2}-\gamma^2), \alpha-q^\beta):\\ (h, v)\in H^2\times H^1_0, r\in\mathbb{R}^+, -k\in\mathscr{C}, \lambda\in\mathbb{R}\}=H^1_0\times\mathbb{R}\times H^2.\end{aligned}\tag{2.16}$$

We thus turn to the verification of (2.16) and choose $(w_1, w_2, w_3)\in H^1_0\times\mathbb{R}\times H^2$ arbitrarily.

Let

$$k=w_3-\min w_3$$

where the minimum is taken over $\bar{\Omega}$ and is well defined since H^2 embeds continuously into C. Observe that $-k\in\mathscr{C}$. Further define

$$h=\lambda(\alpha-q^\beta)-\min w_3,$$

with $\lambda\in\mathbb{R}$ to be fixed below. Clearly, $h\in H^2$ and

$$-h+k+\lambda(\alpha-q^\beta)=w_3,$$

and thus, independently of $\lambda\in\mathbb{R}$, the third coordinate in (2.16) is satisfied for this choice of k and h. Next we consider the second coordinate of (2.16):

$$\langle q^\beta, h\rangle_{H^2}+r+\frac{\lambda}{2}(|q^\beta|^2_{H^2}-\gamma^2)=w_2$$

or by the choice of h

$$r+\lambda(\langle q^\beta, \alpha\rangle-\tfrac{1}{2}|q^\beta|^2_{H^2}-\tfrac{1}{2}\gamma^2)=w_2+\langle q^\beta, \min w_3\rangle.\tag{2.17}$$

Since $\alpha^2\int_\Omega dx<\gamma^2$, the factor $\langle q^\beta, \alpha\rangle-\frac{1}{2}|q^\beta|^2_{H^2}-\frac{1}{2}\gamma^2$ in (2.17) is negative and thus there exist $\lambda\in\mathbb{R}^+$ and $r\in\mathbb{R}^+$ that satisfy (2.17). With $(\lambda, h, -k, r)\in\mathbb{R}\in H^2\times\mathscr{C}\times\mathbb{R}^+$ fixed we turn to the first coordinate in (2.16) that requires solving

$$\nabla(q^\beta\nabla v)=-\Delta w_1-\nabla(h\nabla u^\beta)\quad\text{in } H^{-1},$$

for $v\in H^1_0$. This is clearly possible and hence (2.16) is verified.

Next we will show that

$$\lambda^*=\Delta(q^\beta)^{-1}\Delta(u^\beta-z)\quad\text{in } H^1_0 \text{ where } \Delta(q)u=\nabla(q\nabla u).\tag{2.18}$$

Note that

$$\begin{aligned}&\nabla L(q^\beta, u^\beta; \lambda^*, \mu^*, \eta^*)(h, v)\\ &=\langle\nabla(u^\beta-z), \nabla v\rangle+\beta\Big(\langle\nabla q^\beta, \nabla h\rangle+\sum_{i_1,i_2}\langle(q_k)_{x_{i_1}x_{i_2}}, h_{x_{i_1}x_{i_2}}\rangle\Big)\\ &\quad-\langle\nabla\lambda^*, h\nabla u^\beta+q^\beta\nabla v\rangle+\mu^*\langle q^\beta, h\rangle_{H^2}-\langle\eta^*, h\rangle_{H^2}=0\end{aligned}\tag{2.19}$$

for all $(h, v)\in H^2\times H^1_0$. Let $h=0$ and $v\in H^1_0$ be arbitrary. Then we have

$$\langle\Delta(u^\beta-z), v\rangle-\langle\nabla\cdot(q^\beta\nabla\lambda^*), v\rangle=0$$

for all $v\in H^1_0$. This implies (2.18) and the uniqueness of the Lagrange multiplier λ^*.

To show the uniqueness of Lagrange multipliers μ^* and η^*, assume that $(\lambda^*, \mu^*_i, \eta^*_i)$, $i=1, 2$ are two Lagrange multipliers satisfying (2.2a) and (2.2b). Let $\mu=\mu^*_1-\mu^*_2$ and $\eta=\eta^*_1-\eta^*_2$. By (2.19) and (2.2b) we find

$$\mu\langle q^\beta, h\rangle_{H^2}-\langle\eta, h\rangle_{H^2}=0\quad\text{for all } h\in H^2$$

and moreover

$$\langle \eta, \alpha - q^\beta \rangle_{H^2} = 0.$$

First consider the case of $q^\beta \equiv \alpha$. Since $\alpha^2 \int_\Omega dx < \gamma^2$, the norm constraint is not active and therefore $\mu_1^* = \mu_2^* = 0$ in this case. Moreover, by the first equation we find that $\eta = 0$ so that $\eta_1^* = \eta_2^*$. Next assume that q^β is not identically α. Putting $h = \alpha - q^\beta$ in the first equation and using the second equation, we obtain $\mu \langle q^\beta, \alpha - q^\beta \rangle_{H^2} = 0$ that implies $\mu = 0$. Thus, $\langle \eta, h \rangle_{H^2} = 0$ for all $h \in H^2$. This implies $\eta = 0$ and the proof is completed.

Remark 2.4. If $\mu^* = \beta = 0$ and $\eta^* = 0$, then

$$\langle \nabla(u^\beta - z), \nabla u^\beta \rangle = 0. \tag{2.20}$$

In fact, suppose that $\mu^* = \beta = 0$ and $\eta^* = 0$. Then (2.19) with $v = 0$ implies $\langle \nabla \lambda^*, h \nabla u^\beta \rangle = 0$ for all $h \in H^2$. This equation with $h = q^\beta$ and λ^* expressed as in (2.18) implies

$$\langle \Delta(q^\beta)^{-1} \Delta(u^\beta - z), \Delta(q^\beta) u^\beta \rangle = \langle \nabla(u^\beta - z), \nabla u^\beta \rangle = 0,$$

which is the desired equality.

To prove Theorem 2.2 we need the following lemma.

LEMMA 2.5. *The adjoint operator $\tilde{M}^* : H_0^1 \times \mathbb{R} \times H^2 \to H^2 \times H_0^1$ of $\tilde{M}$ is surjective. The kernel of $\tilde{M}^*$ is one-dimensional and it is characterized by*

$$\ker \tilde{M}^* = \operatorname{span}\{(0, 1, q^\beta)\}.$$

Moreover, $\tilde{M}\tilde{M}^*$ has a bounded inverse as an operator on range $\tilde{M}$.

Proof. It is simple to show that the range of $\tilde{M}$ is closed. Next suppose that $\tilde{M}(h, v) = 0$ for some $(h, v) \in H^2 \times H_0^1$. Then $h = 0$ and thus $(-\Delta)^{-1} \nabla \cdot (q^\beta \nabla v) = 0$. Since $(-\Delta)^{-1} : H^{-1} \to H_0^1$ is an isomorphism and since $q^\beta \geqq \alpha$, this implies that $v = 0$. Hence $\tilde{M}$ is injective and by the closed range theorem $\tilde{M}^*$ is surjective. A short calculation shows that the kernel of $\tilde{M}^*$ is given as the set of elements $(x, y, z) \in H^1 \times \mathbb{R} \times H^2$ that satisfy

$$\langle y q^\beta, h \rangle_{H^2} = \langle z, h \rangle_{H^2} \quad \text{for all } h \in H^2.$$

In particular, dim $(\ker \tilde{M}^*) = 1$. By the closed range theorem, $(\ker \tilde{M}^*)^\perp = \text{range } \tilde{M}$ and

$$H^1 \times R \times H^2 = \ker \tilde{M}^* \oplus \text{range } \tilde{M}.$$

To show that $\tilde{M}\tilde{M}^*$ has a bounded inverse on range $\tilde{M}$, we observe that $\tilde{M}\tilde{M}^*$ is injective and surjective on range $\tilde{M}$. This completes the proof.

Now we turn to a proof of Theorem 2.2.

Proof of Theorem 2.2. The augmentability estimate in (a) follows from the proof of Theorem 2.1 and from Corollary 2.2 of [8] (with (2.4) replacing Theorem 2.1 of [8]). In addition, if the norm constraint is not active, then $\tilde{c}(r)$ and B can be chosen so that

$$g(q) + \frac{\mu}{c} < 0 \tag{2.21}$$

for all $\mu \leqq r$, $c \geqq \tilde{c}(r)$, and all $(q, u) \in \bar{B}$. Estimate (2.10) is a special case of Proposition 4.1 of [8]. We now turn to the proof of (2.11). First observe that by (2.19)

$$\langle \nabla(u^\beta - z), \nabla v \rangle + \beta \left(\langle \nabla q^\beta, \nabla h \rangle + \sum_{i_1, i_2} \langle q^\beta_{x_{i_1} x_{i_2}}, h_{x_{i_1} x_{i_2}} \rangle \right)$$
$$- \langle \nabla \lambda^*, h \nabla u^\beta + q^\beta \nabla v \rangle + \mu^* \langle q^\beta, h \rangle_{H^2} - \langle \eta^*, h \rangle_{H^2} = 0$$

for all $(h, v) \in H^2 \times H_0^1$ and the necessary optimality condition for (q_k, u_k) yields

$$\begin{aligned}(2.22)\qquad &\langle \nabla(u_k - z), \nabla v\rangle + \beta\left(\langle \nabla q_k, \nabla h\rangle + \sum_{i_1, i_2} \langle (q_k)_{x_{i_1}x_{i_2}}, h_{x_{i_1}x_{i_2}}\rangle\right)\\ &-\langle \nabla \tilde{\lambda}^{k+1}, h\nabla u_k + q_k \nabla v\rangle + \tilde{\mu}^{k+1}\langle q^k, h\rangle_{H^2} - \langle \eta^k, h\rangle_{H^2} = 0\end{aligned}$$

for all $(h, v) \in H^2 \times H_0^1$, where

$$\tilde{\lambda}^{k+1} = \lambda^k + c_k e(q_k, u_k) \quad \text{and} \quad \tilde{\mu}^{k+1} = \mu^k + c_k \hat{g}(q_k, \mu^k, c_k).$$

Subtracting these two equalities, rearranging terms, and using the definition of $\tilde{M}$, we obtain

$$\begin{aligned}&\langle (\lambda^* - \tilde{\lambda}^{k+1}, \mu^* - \tilde{\mu}^{k+1}, \eta^* - \eta^k), \tilde{M}(h, v)\rangle_{H_0^1 \times \mathbb{R} \times H^2}\\ (2.23)\qquad &= \langle \nabla(u_k - u^\beta), \nabla v\rangle + \beta\left(\langle \nabla(q_k - q^\beta), \nabla h\rangle + \sum_{i_1, i_2} \langle (q_k)_{x_{i_1}x_{i_2}} - q^\beta_{x_{i_1}x_{i_2}}, h_{x_{i_1}x_{i_2}}\rangle\right)\\ &\quad -\langle \nabla \tilde{\lambda}^{k+1} \cdot \nabla(u^\beta - u_k), h\rangle + \langle (q^\beta - q_k)\nabla \tilde{\lambda}^{k+1}, \nabla v\rangle + \tilde{\mu}^{k+1}\langle q^k - q^\beta, h\rangle_{H^2}.\end{aligned}$$

From (2.10), the sequences $\{\lambda^k, \mu^k\}$ and $\{q_k, u_k\}$ are uniformly bounded in $H_0^1 \times \mathbb{R}$ and $H^2 \times H_0^1$, respectively. This implies uniform boundedness of the sequence $\{\tilde{\lambda}^k, \tilde{\mu}^k\}$ in $H_0^1 \times \mathbb{R}$. By the Riesz Representation Theorem, the right-hand side of (2.23) can be represented by $\langle b_k, (h, v)\rangle_{H^2 \times H_0^1}$ where $b_k \in H^2 \times H_0^1$ and

$$|b_k|_{H^2 \times H_0^1} \leqq K_1 |(q_k, u_k) - (q^\beta, u^\beta)|_{H^2 \times H_0^1}$$

for a constant K_1 independent of k. We find from (2.23) that

$$\tilde{M}^*(\lambda^* - \tilde{\lambda}^{k+1}, \mu^* - \tilde{\mu}^{k+1}, \eta^* - \eta^k) = b_k$$

in $H^2 \times H_0^1$. If P_R denotes the orthogonal projection of $H_0^1 \times R \times H^2$ onto $(\ker \tilde{M}^*)^\perp$, then

$$\tilde{M}\tilde{M}^* P_R(\lambda^* - \tilde{\lambda}^{k+1}, \mu^* - \tilde{\mu}^{k+1}, \eta^* - \eta^k) = \tilde{M} b_k.$$

Since from Lemma 2.5 $\tilde{M}\tilde{M}^*$ is continuously invertible on range $\tilde{M} = (\ker \tilde{M}^*)^\perp$,

$$(2.24)\qquad |P_R(\lambda^* - \tilde{\lambda}^{k+1}, \mu^* - \tilde{\mu}^{k+1}, \eta^* - \eta^k)|_{H_0^1 \times \mathbb{R} \times H^2} \leqq K_2 |(q_k, u_k) - (q^\beta, u^\beta)|_{H^2 \times H_0^1},$$

for a constant K_2 independent of k.

Next the complementarity conditions imply that $\langle \eta^*, \alpha - q^\beta\rangle_{H^2} = 0$ and $\langle \eta^k, \alpha - q_k\rangle_{H^2} = 0$. Since $\{\tilde{\lambda}^k, \tilde{\mu}^k\}$ is uniformly bounded in $H_0^1 \times \mathbb{R}$, it follows from (2.22) that $\{\eta^k\}$ is uniformly bounded. Thus, these equalities yield the estimate

$$(2.25)\qquad |\langle \eta^* - \eta^k, \alpha - q^\beta\rangle_{H^2}| = |\langle \eta^k, q^\beta - q_k\rangle_{H^2}| \leqq K_3 |q^\beta - q_k|_{H^2},$$

for a constant K_3 independent of k. Now let us assume that $q^\beta \not\equiv \alpha$. Then $\langle q^\beta, \alpha - q^\beta\rangle_{H^2} \neq 0$ and it follows from (2.24), (2.25), and the characterization of $\ker \tilde{M}^*$ in Lemma 2.5 that

$$(2.26)\quad |(\lambda^* - \tilde{\lambda}^{k+1}, \mu^* - \tilde{\mu}^{k+1}, \eta^* - \eta^k)|_{H_0^1 \times \mathbb{R} \times H^1} \leqq K_4 |(q_k, u_k) - (q^\beta, u^\beta)|_{H^2 \times H_0^1},$$

for a constant K_4 independent of k. In the case $q^\beta \equiv \alpha$ the norm constraint is inactive and hence $\mu^* = 0$. By (2.10) we have $\mu^k \leqq r$ for all $k \geqq 1$, and (2.21) then implies that $\tilde{\mu}^k = 0$ for all $k \geqq 2$. Consequently, (2.26) also holds for the case $q^\beta \equiv \alpha$. Next we will show that (2.26) holds when $(\tilde{\lambda}^{k+1}, \tilde{\mu}^{k+1})$ is replaced by $(\lambda^{k+1}, \mu^{k+1})$. Observe that there exists a constant K independent of k such that

$$|\lambda^{k+1} - \tilde{\lambda}^{k+1}|_{H_0^1} \leqq c_0 |\Delta^{-1}\nabla \cdot (q_k \nabla u_k - q^\beta \nabla u^\beta)|_{H_0^1} \leqq K |(q_k, u_k) - (q^\beta, u^\beta)|_{H^2 \times H_0^1},$$

and

$$|\mu^{k+1}-\tilde{\mu}^{k+1}|=c_0\left|\max\left(-\frac{\mu^k}{c_k}, g(q_k)\right)\right|\leqq K|q_k-q^\beta|_{H^2}$$

where for the last estimate we assumed that the norm constraint is active and we used the fact that $\mu^k \geqq 0$. Combining these estimates with (2.26), we obtain

$$|(\lambda^*-\lambda^{k+1}, \mu^*-\mu^{k+1}, \eta^*-\eta^k)|_{H_0^1\times\mathbb{R}\times H^2}\leqq K_5|(q_k, u_k)-(q^\beta, u^\beta)|_{H^2\times H_0^1} \tag{2.27}$$

for a constant K_5 independent of k, provided that $|q^\beta|_{H^2}=\gamma$. If the norm constraint is not active, then by (2.10) we have

$$\mu^{k+1}\leqq \mu^*+(|\lambda^1-\lambda^*|^2_{H_0^1}+|\mu^1-\mu^*|^2)^{1/2}\leqq r \quad \text{for all } k\geqq 1$$

and therefore by (2.21)

$$\begin{aligned}\mu^{k+1}&=\max\left(\frac{\mu^k c_0}{c_k}, \mu^k+(c_k-c_0)g(q_k)\right)\\ &\leqq \max\left(\frac{\mu^k c_0}{c_k}, \mu^k-(c_k-c_0)\frac{\mu^k}{c_k}\right)=\frac{\mu^k c_0}{c_k}.\end{aligned} \tag{2.28}$$

The estimates (2.11) now follow from (2.10), (2.27), and (2.28). This completes the proof.

Remark 2.6. The assumption of Theorem 2.2(b) that $c_1\geqq \tilde{c}(r)$ is sufficiently large is used to guarantee that (q_k, u_k) is in the open ball B for all $k\geqq 1$. If we only assume $c_1\geqq\tilde{c}(r)$, then the estimates of Theorem 2.2 need to be modified. First (2.10) holds with (q_k, u_k) replaced by $(\tilde{q}_k, \tilde{u}_k)$, where $(\tilde{q}_k, \tilde{u}_k)$ is a solution of (2.8) in $\bar{B}$ (compare [8]). In particular, this implies that $(\tilde{q}_k, \tilde{u}_k)\to(q^\beta, u^\beta)$ and that $\{(\lambda^k, \mu^k)\}_{k=1}^\infty$ is bounded in $H_0'\times R$. Let k_0 be chosen such that $(\tilde{q}_k, \tilde{u}_k)\in B$ for all $k\geqq k_0$. Then the analysis of Theorem 2.2 can be repeated to show that for $k\geqq 1$

$$\begin{aligned}&|\lambda^{k+k_0}-\lambda^*|^2_{H_0^1}+|\mu^{k+k_0}-\mu^*|^2+|\eta^{k+k_0-1}-\eta^*|_{H^2}\\ &\quad\leqq\left(\frac{K}{\bar{\sigma}}\right)^k\prod_{i=k_0}^{k+k_0-1}\frac{1}{c_i-c_0}(|\lambda^{k_0}-\lambda^*|^2_{H_0^1}+|\mu^{k_0}-\mu^*|^2)\end{aligned}$$

The proof of Theorem 2.3 can be given along the lines of that for Theorem 2.2.

3. The coercivity condition. In this section we establish the coercivity condition (2.4) for specific cases. To achieve this goal it is necessary to study the behavior of the solutions (q^β, u^β) to (P^β) as $\beta\to 0^+$. Let

$$N(q)=\sum_{i_1,i_2}|q_{x_{i_1}x_{i_2}}|^2+|\nabla q|^2, \tag{3.1}$$

representing the seminorm regularization.

LEMMA 3.1. *Let (q^β, u^β) be any solution of (P^β). For $\beta>\beta^0\geqq 0$, we have*

$$|u^\beta-z|^2_{H_0^1}\leqq|u^{\beta_0}-z|^2_{H_0^1}+\beta(N(q^{\beta_0})-N(q^\beta)), \tag{3.2}$$

$$\sup_{Q^\beta} N(q^\beta)\leqq\inf_{Q^{\beta_0}} N(q^{\beta_0}), \tag{3.3}$$

$$\sup_{U^{\beta_0}}|u^{\beta_0}-z|_{H_0^1}\leqq\inf_{U^\beta}|u^\beta-z|_{H_0^1} \tag{3.4}$$

where for $\beta\geqq 0$, $Q^\beta=\{q^\beta: (q^\beta, u^\beta)$ is a solution of $(\mathrm{P}^\beta)\}$ and $U^\beta=\{u^\beta: (q^\beta, u^\beta)$ is a solution of $(\mathrm{P}^\beta)\}$. If $\beta_n\to 0^+$ and $\{q^{\beta_n}\}$ is any sequence of corresponding solution of (P^{β_n}), then $\{q^{\beta_n}\}$ has a weak cluster point, and every weak cluster point is a solution of (P^0), and we have

$$\limsup_{n\to\infty}\sup_{Q^{\beta_n}} N(q^{\beta_n})=\min_{Q^0} N(q^0). \tag{3.5}$$

Moreover, every weak cluster point of a sequence of solutions q^{β_n} is a strong cluster point in H^2 and it is a minimum norm solution of (1.2).

Proof. The proof of (3.2)-(3.5) is a simple consequence of the above remark on the equivalence between (1.2) and (P^β) and the results in § 2 of [4]. In fact, (3.2) follows from (2.2) in [4], (3.3) and (3.4) from Lemma 2.2, and (3.5) from Lemma 2.3 of [4]. Assumption (A2), requiring existence of a minimizer of (P^0) in [4], is guaranteed by the properties of Q_{ad}. The coercivity assumption for N in (A1) of [4] is replaced by the norm constraint. We will show the last statement of the lemma. If q^{β_n} converges weakly to q, then from (3.5) $N(q^{\beta_n}) \to N(q)$ where q is a minimum norm solution of (P^0). Since the embedding from H^2 into L^2 is compact, q^{β_n} converges strongly to q in L^2. Thus we obtain

$$N(q^{\beta_n}) + |q^{\beta_n}|_{L^2} \to N(q) + |q|^2_{L^2}.$$

Since $N(q)+|q|^2_{L^2}$ defines a norm that is equivalent to the common H^2-norm [15, p. 13], this implies $|q^{\beta_n}|^2_{H^2} \to |q|^2_{H^2}$ so that $\{q^n\}$ converges strongly to q in H^2.

From (3.2), (3.3), (3.5), and observing that $|u^0 - z|^2_{H_0^1}$ is independent of $u^0 \in U^0$, we find the following corollary to Lemma 3.1.

COROLLARY 3.2. *There exists a real-valued monotonically increasing function $\rho(\beta)$ with* $\lim \rho(\beta) \to 0$ *as* $\beta \to 0^+$ *such that for* $\beta \geqq 0$

$$\begin{aligned}\sup_{U^\beta} |u^\beta - z|^2_{H_0^1} &\leqq |u^0 - z|^2_{H_0^1} + \beta \left(\min_{Q^0} N(q^0) - N(q^\beta)\right) \\ &\leqq |u^0 - z|^2_{H_0^1} + \beta\rho(\beta).\end{aligned}$$

The second Fréchet derivative of L_c at $(q^\beta, u^\beta, w^\beta)$ in direction $(h, v, y) \in H^2 \times H_0^1 \times \mathbb{R}$ appearing in (2.4) is given by

$$\begin{aligned}&\nabla^2 L_c(q^\beta, u^\beta, w^\beta; \lambda^*, \mu^*, \eta^*)((h, v, y), (h, v, y)) \\ &\quad = |v|^2_{H_0^1} + \beta N(h) - 2\langle \nabla\lambda^*, h\nabla v\rangle \\ &\qquad + \mu^*|h|^2_{H^2} + c|(-\Delta)^{-1}\nabla \cdot (q^\beta \nabla v + h\nabla u^\beta)|^2_{H_0^1} + c|\langle q^\beta, h\rangle_{H^2} + y|^2\end{aligned} \tag{3.6}$$

where we used

$$|(-\Delta)^{-1}\nabla\varphi|^2_{H_0^1} = \langle(-\Delta)^{-1}\nabla\varphi, \nabla\varphi\rangle = \langle\nabla\Delta^{-1}\nabla\varphi, \varphi\rangle = \langle P\varphi, \varphi\rangle_{L^2(\Omega,\mathbb{R}^n)} \tag{3.7}$$

for $\varphi \in H^1(\Omega; \mathbb{R}^n)$, which can be verified by Green's formula [15, p. 28]. Here the operator $P = \operatorname{grad} \Delta^{-1} \operatorname{div}$ defines an orthogonal projection in $L^2(\Omega, \mathbb{R}^n)$.

To prove the coercivity condition (2.4) in Proposition 3.4 we use some well-known estimates [15, pp. 18, 20, 72].

LEMMA 3.3. *There exist positive constants K_i, $i = 1, 2, 3, 4$, depending only on Ω such that*

(a) $|\varphi|_{L^\infty} \leqq K_1 |\varphi|_{H^2}$ *for all* $\varphi \in H^2$,

(b) $|\varphi| \leqq K_2 |\nabla\varphi|$ *for all* $\varphi \in H^1$ *with* $\int_\Omega \varphi\, dx = 0$,

(c) $|\varphi|_{H^2} \leqq K_3(\sum_{i_1,i_2} |\varphi_{x_{i_1}x_{i_2}}|^2 + |\nabla\varphi|^2)^{1/2}$ *for all* $\varphi \in H^2$ *with* $\int_\Omega \varphi\, dx = 0$.

The proposition also involves the constants α and γ that define Q_{ad}, the function $\rho(\beta)$ from Corollary 3.2, and the constant ω that is defined as follows. For $f \in H^{-1}$ with $f \neq 0$ and $q \in Q_{\text{ad}}$ we have

$$|f|_{H^{-1}} = \sup_{v \in H_0^1} \frac{|\langle q\nabla u, \nabla v\rangle|}{|v|_{H_0^1}} \leqq |q|_{L^\infty} |\nabla u| \leqq K_1 \gamma |u|_{H_0^1}$$

and therefore

$$|u|_{H_0^1} \geqq \frac{|f|_{H^{-1}}}{\gamma K_1} =: \omega > 0. \tag{3.8}$$

PROPOSITION 3.4. *Let $f \in H^{-1}$ satisfy $f \neq 0$ and let $k = (1 + K_2^2 + 4K_1^2K_3^2)^{-1}$. Let (q^0, u^0) be a solution of* (P^0) *and suppose that for a constant $\beta_0 \in (0, 1)$*

$$|u^0 - z|^2_{H_0^1} < \beta_0\left[\frac{\alpha^2 k}{K_1^2}\left(1 - \frac{4k}{\omega} K_1^2\gamma^2\beta_0\right) - \rho(\beta_0)\right]. \tag{3.9}$$

Then there exists a nontrivial compact interval $I = [\underline{\beta}, \beta_0] \subset (0, 1)$ and constants $c_0 > 0$, $\sigma_0 > 0$ such that

$$\nabla^2 L_c(q^\beta, u^\beta, w^\beta; \lambda^*, \mu^*, \eta^*)((h, v, y), (h, v, y)) \geq \sigma_0(|h|^2_{H^2} + |v|^2_{H_0^1} + |y|^2)$$

for all $c \geq c_0$, $(h, v, y) \in H^2 \times H_0^1 \times \mathbb{R}$ and any solution (q^β, u^β) of (P$^\beta$) *with $\beta \in I$. Moreover, if $u^0 = z$, then such a constant β_0 always exists and I can be chosen as any compact subset of $(0, \beta_0]$.*

We point out that in Proposition 3.4 the solutions (q^0, u^0) and (q^β, u^β) are assumed to be global. This is necessary since the proof requires the estimates of Lemma 3.1 and Corollary 3.2 that are given for (global) solutions. The assumption regarding existence of β_0 such that (3.9) holds represents a smallness condition of the error between z and the nonregularized OLS solution u^0. All quantities appearing on the right-hand side of (3.9) except for $\rho(\beta)$ in principle can be given explicitly.

Proof. Define a quadratic form on $H^2 \times H_0^1$ by

$$\begin{aligned} M_c(h, v) &= |v|^2_{H_0^1} + \beta N(h) - 2\langle \nabla\lambda^*, h\nabla v\rangle \\ &\quad + c\langle(-\Delta)^{-1}\nabla \cdot (q^\beta \nabla v + h\nabla u^\beta), \nabla(q^\beta\nabla v + h\nabla u^\beta)\rangle \end{aligned} \tag{3.10}$$

where $c \geq 0$ and the dependence of $M_c(h, v)$ on β is suppressed. For $(h, v, y) \in H^2 \times H_0^1 \times \mathbb{R}$ it follows that

$$\nabla^2 L_c((h, v, y), (h, v, y)) = M_c(h, v) + c(\langle q^\beta, h\rangle_{H^2} + y)^2 + \mu^*|h|^2_{H^2}. \tag{3.11}$$

We first concentrate on M_c. By Theorem 2.1 and Lemma 3.3(a)

$$\begin{aligned} \langle\nabla\lambda^*, h\nabla v\rangle &= -\langle\lambda^*, \nabla \cdot (h\nabla v)\rangle = \langle\nabla\Delta(q^\beta)^{-1}\Delta(u^\beta - z), h\nabla v\rangle \\ &\geq -|\nabla\Delta(q^\beta)^{-1}\Delta(u^\beta - z)||h\nabla v| \\ &\geq -\frac{K_1}{\alpha}|\Delta(u^\beta - z)|_{H^{-1}}|h|_{H^2}|v|_{H_0^1} \\ &= -\frac{K_1}{\alpha}|u^\beta - z|_{H_0^1}|h|_{H^2}|v|_{H_0^1}. \end{aligned} \tag{3.12}$$

By (3.7) and Lemma 3.3 (suppressing the index β)

$$\begin{aligned} \langle(-\Delta)^{-1}\nabla \cdot (q\nabla v + h\nabla u), \nabla \cdot (q\nabla v + h\nabla u)\rangle &= \langle P(q\nabla v + h\nabla u), q\nabla v + h\nabla u\rangle \\ &\geq \tfrac{1}{2}\langle P(h\nabla u), h\nabla u\rangle - \langle P(q\nabla v), q\nabla v\rangle \\ &\geq \tfrac{1}{2}\langle P(h\nabla u), h\nabla u\rangle - |q|^2_{L^\infty}|\nabla v|^2 \\ &\geq \tfrac{1}{2}\langle P(h\nabla u), h\nabla u\rangle - K_1^2\gamma^2|v|^2_{H_0^1}. \end{aligned} \tag{3.13}$$

Here and in (3.14) below we suppress the superscript β. Each $h \in L^2$ can be uniquely decomposed as $h = h_1 + h_2$, where $h_1 = \int_\Omega h(x)\,dx$ and $h_2 \in \{h \in L^2 : \int_\Omega h\,dx = 0\}$. Observe that h_1 and h_2 are orthogonal in L^2 and if $h \in H^2$, then $h_2 \in H^2$. By definition of P and by Lemma 3.3(a) we find

$$\begin{aligned} \langle P(h\nabla u), h\nabla u\rangle &= \langle P(h_1\nabla u), h_1\nabla u\rangle + 2\langle P(h_1\nabla u), h_2\nabla u\rangle + \langle P(h_2\nabla u), h_2\nabla u\rangle \\ &\geq |h_1|^2|\nabla u|^2 - 2|h_1||h_2|_{L^\infty}|\nabla u|^2 \\ &\geq |h_1|^2|\nabla u|^2 - 2K_1|h_1||h_2|_{H^2}|\nabla u|^2. \end{aligned} \tag{3.14}$$

Let us put $\ell := |u^\beta|^2_{H_0^1}$ and $c = \delta\beta$ with $\delta > 0$ to be chosen below. Then from (3.10), (3.12)-(3.14) we have

$$M_{\delta\beta}(h, v) \geqq (1 - \delta\beta K_1^2\gamma^2)|v|^2_{H_0^1} + \beta\left(N(h) + \frac{\delta\ell}{2}|h_1|^2 - K_1\delta\ell|h_1||h_2|_{H^2}\right) \tag{3.15}$$
$$- \frac{2K_1}{\alpha}|u^\beta - z|_{H_0^1}|h|_{H^2}|v|_{H_0^1}.$$

Next we will show that there exist constants k and δ such that

$$N(h) + \frac{\delta\ell}{2}|h_1|^2 - K_1\delta\ell|h_1||h_2|_{H^2} \geqq k(N(h) + |h|^2) \tag{3.16}$$

for all $h \in H^2$.

From Lemma 3.3 it follows that for any $0 < A < 1$ and $B > 0$

$$N(h) + \frac{\delta\ell}{2}|h_1|^2 - K_1\delta\ell|h_1||h_2|_{H^2}$$
$$\geqq (1 - A)N(h) + \frac{\delta\ell}{2}|h_1|^2 + \frac{A}{K_2^2}|h_2|^2 - \frac{(\delta\ell)^2}{4}B^2K_1^2|h_1|^2 - \frac{1}{B^2}|h_2|^2_{H^2}$$
$$\geqq \left(1 - A - \frac{K_3^2}{B^2}\right)N(h) + \left(\frac{\delta\ell}{2} - \frac{(\delta\ell)^2}{4}B^2K_1^2\right)|h_1|^2 + \frac{A}{K_2^2}|h_2|^2.$$

Since $|h|^2 = |h_1|^2 + |h_1|^2$, inequality (3.16) holds if there exist positive constants δ, k, B, and $A \in (0, 1)$ such that

$$1 - A - \frac{K_3^2}{B^2} \geqq k, \frac{\delta\ell}{2} - \frac{(\delta\ell)^2}{4}B^2K_1^2 \geqq k \quad \text{and} \quad \frac{A}{K_2^2} \geqq k.$$

A calculation shows that these inequalities are satisfied if we take

$$k = (1 + K_2^2 + 4K_1^2K_3^2)^{-1}, \quad A = K_2^2k, \quad B^2 = (4kK_1^2)^{-1}, \quad \delta\ell = 4k.$$

This is the choice of k contained in the statement of the proposition. From (3.15) and (3.16) we have

$$M_c(h, c) \geqq k\beta(N(h) + |h|^2) + \left(1 - \frac{4k}{\ell}\beta K_1^2\gamma^2\right)|v|^2_{H_0^1}$$
$$- \frac{2K_1}{\alpha}|u^\beta - z|_{H_1^0}|h|_{H^2}|v|_{H_0^1}$$

where $c = 4k\beta/\ell$ and $\ell = \ell(\beta) = |u^\beta|^2_{H_0^1}$. By the choice of ω the last inequality implies

$$M_c(h, v) \geqq k\beta|h|^2_{H^2} + \left(1 - \frac{4k}{\omega}\beta K_1^2\gamma^2\right)|v|^2_{H_0^1} - \varepsilon k\beta|h|^2_{H^2} - \frac{K_1^2}{\varepsilon\alpha^2k\beta}|u^\beta - z|^2_{H_0^1}|v|^2_{H_0^1} \tag{3.17}$$
$$= (1 - \varepsilon)k\beta|h|^2_{H^2} + \left(1 - \frac{4k}{\omega}\beta K_1^2\gamma^2 - \frac{K_1^2}{\varepsilon\beta\alpha^2k}|u^\beta - z|^2_{H_0^1}\right)|v|^2_{H_0^1}$$

where $c = 4k\beta/\omega$ and $\varepsilon \in (0, 1)$ is arbitrary. By (3.9) and Corollary 3.2 we find

$$\sup_{U^{\beta_0}} |u^{\beta_0} - z|^2_{H_0^1} < \frac{\varepsilon_1\beta_0\alpha^2k}{K_1^2}\left(1 - \frac{4k}{\omega}K_1^2\gamma^2\beta_0\right)$$

for some $\varepsilon_1 \in (0, 1)$. Furthermore, by (3.4) of Lemma 3.1 there exist constants $\bar{\beta} \in (0, \beta_0)$ and $\eta > 0$ such that

$$\sup_{U^\beta} |u^\beta - z|^2_{H_0^1} \leqq \frac{\varepsilon_1 \beta \alpha^2 k}{K_1^2}\left(1 - \frac{4k}{\omega} K_1^2 \gamma^3 \beta\right) - \eta$$

for all $\beta \in I = [\bar{\beta}, \beta_0]$ and by (3.17) with $\varepsilon = \varepsilon_1$ this implies

$$M_c(h, v) \geqq (1 - \varepsilon_1) k\beta |h|^2_{H^2} + \eta |v|^2_{H_0^1} \quad \text{where } c = \frac{4k\beta}{\omega}$$

for all $\beta \in I$. From the definition of M_c and the last inequality, it follows that there exists a constant $\sigma > 0$ such that

$$M_c(h, v) \geqq \sigma\beta(|h|^2_{H^2} + |v|^2_{H_0^1}) \quad \text{for all } c \geqq \frac{4k\beta}{\omega} \text{ and } \beta \in I. \tag{3.18}$$

If $u^0 = z$ then there always exists $\beta_0 > 0$ such that (3.9) is satisfied, and using Corollary 3.2 we can verify (3.18) with I any compact subset of $(0, \beta_0]$.

Next note that

$$|\langle q^\beta, h\rangle_{H^2} + y|^2 \geqq \tfrac{1}{2}|y|^2 - |\langle q^\beta, h\rangle_{H^2}|^2 \geqq \tfrac{1}{2}|y|^2 - \gamma^2 |h|^2_{H^2}$$

and consequently

$$\begin{aligned} \sigma\beta|h|^2_{H^2} + c_2|\langle q^\beta, h\rangle_{H^2} + y|^2 &\geqq (\sigma\beta - \gamma^2 c_2)|h|^2_{H^2} + \frac{1}{2} c_2 |y|^2 \\ &= \frac{\sigma\beta}{1 + 2\gamma^2}(|h|^2_{H^2} + |y|^2) \end{aligned} \tag{3.19}$$

where $c_2 = 2\sigma\beta/(1 + 2\gamma^2)$. Hence from (3.11), (3.18), and (3.19) we obtain

$$\nabla^2 L_c(q^\beta, u^\beta, w^\beta; \lambda^*, \mu^*, \eta^*)((h, v, y), (h, v, y)) \geqq \frac{\sigma\beta}{1 + 2\gamma^2}(|h|^2 + |v|^2_{H^1} + |y|^2)$$

for all $(h, v, y) \in H^2 \times H_0^1 \times \mathbb{R}$, $c \geqq \max(4k\beta/\omega, 2\sigma\beta/(1 + 2\gamma^2))$ and $\beta \in I$. This implies the claim.

In special cases the coercivity estimate of Proposition 3.4 can be obtained with $\beta = 0$.

PROPOSITION 3.5. *Let (q^0, u^0) be a local solution of (P_0).*

(a) *If $\mu^* > 0$ and $\text{dist}^2 := |u^0 - z|^2_{H_0^1} < (\alpha/K_1)^2 \mu^*$, then there exist positive constants σ_1 and c_1 such that*

$$\nabla^2 L_c(q^0, u^0, w^0)((h, v, y), (h, v, y)) \geqq \sigma_1(|h|^2_{H^2} + |v|^2_{H_0^1} + |y|^2)$$

for all $c \geqq c_1$ and $(h, v, y) \in H^2 \times H_0^1 \times \mathbb{R}$.

(b) *Let $\tilde{L}_c$ be given by (1.8) where the norm constraint is not augmented (i.e., $L_c = \tilde{L}_c + (c/2)|g(q) + w|^2$). If $\text{dist} = 0$, then there exist positive constants σ_2 and c_2 such that*

$$\nabla^2 \tilde{L}_c(q^0, u^0)((h, v), (h, v)) \geqq \sigma_2(|P(h\nabla u^0)|^2_{L^2} + |v|^2_{H_0^1}) \tag{3.20}$$

for all $c \geqq c_2$ and $(h, v) \in H^2 \times H_0^1$.

(c) *Let $\{\varphi_i\}_{i=1}^M$ be (curved) linear elements [19] or indicator functions such that $0 \leqq \varphi_i \leqq 1$ on Ω and let $V^M = \{q = \sum_{i=1}^M q_i\varphi_i : q_i \in \mathbb{R}\} \subset L^\infty$. Let $(P^0)_c^M$ be the problem of minimizing*

$$\frac{1}{2}|u - z|^2_{H_0^1} + \frac{c}{2}|e(q, u)|^2_{H_0^1}$$

subject to $e(q, u) = 0$ *in* H_0^1 *and* $q \in Q_{\text{ad}} = \{q = \sum_{i=1}^M q_i\varphi_i \in V^M : q_i \geqq \alpha$ *and* $q^T W^M q < \gamma^2\}$, *where* $|(W^M)^{1/2}\alpha|^2 < \gamma^2$ *and* W^M *is a symmetric positive definite matrix on* $\mathbb{R}^M$. *Furthermore, assume that* $q \in Q_{\text{ad}}$ *implies* $q(x) \geqq \alpha$ *on* Ω. *Then* $(P^0)_c^M$ *has a solution* $(q^0, u^0) \in V^M \times H_0^1$ *with an associated Lagrange multiplier* $(\lambda^*, \mu^*, \eta^*) \in H_0^1 \times \mathbb{R}_+ \times \mathbb{R}_+^M$ *such that if we define the augmented Lagrangian* (*compare* (2.13))

$$\hat{L}_c(q, u, w; \lambda^*, \mu^*, \eta^*) = \frac{1}{2}|u - z|_{H_0^1}^2 + \langle \lambda^*, e\rangle_{H_0^1} + \frac{\mu^*}{2}(q^T W^M q - \gamma^2)$$

$$+ \sum_{i=1}^M \eta_i^*(\alpha - q_i) + \frac{c}{2}|e|_{H_0^1}^2 + \frac{c}{2}\left(\frac{1}{2}(q^T W^M q - \gamma^2) + w\right)^2,$$

then $\nabla \hat{L}_c(q^0, u^0, w^0; \lambda^*, \mu^*, \eta^*)(h, v, y) = 0$ *for all* $(h, v, y) \in V^M \times H_0^1 \times \mathbb{R}$ *and* $\mu^*(q^{0^T} W^M q^0 - \gamma^2) = \sum_{i=1}^M \eta_i^*(\alpha - q_i^0) = 0$. *Moreover, if* $h \to |P(h\nabla u^0)|$ *defines a norm on* V^M *with* $|P(h\nabla u^0)| \geqq b|h|_{L^\infty}$ *for some* $b > 0$ *and all* $h \in V^M$, *and if* $\text{dist} < \alpha b^2(2|q^0|_{L^\infty}^2 + b^2)^{-1}$, *then there exist positive constants* σ_3 *and* c_3 *such that*

$$\nabla^2 \hat{L}_c(q^0, u^0, w^0; \lambda^*, \mu^*, \eta^*)((h, v, y), (h, v, y)) \geqq \sigma_3(|P(h\nabla u^0)|^2 + |v|_{H_0^1}^2 + y^2)$$

for all $c \geqq c_3$ *and* $(h, v, y) \in V^M \times H_0^1 \times \mathbb{R}$.

We precede the proof with a brief discussion of this proposition. Part (a) presents the most desirable situation. In this case $\mu^* > 0$ takes over the role of the regularization parameter $\beta > 0$ of Proposition 3.4. In general, however, it is difficult to give conditions that guarantee $\mu^* > 0$ (see, however, [13] for the one-dimensional case). Part (b) is not directly applicable for the results of § 2, but it exhibits clearly the difficulties that are involved in obtaining a lower bound on the second derivative of the augmented Lagrangian: First the norm involved in (3.20) is only the L^2 rather than the H^2-norm; second, we obtain an estimate only in terms of $P(h\nabla u^0)$, where the kernel of P is the set of all divergence free vector fields. However, (3.20) also indicates how further assumptions can be made to obtain the desired coercivity estimate. To give an example, let us assume that q is known to be constant a priori, i.e., we take $q \in \{q \in \mathbb{R}: \alpha \leqq q \leqq \gamma(\int_\Omega dx)^{1/2}\}$. Then $h \in \mathbb{R}$ and $P(h\nabla u^0)$ becomes $h\nabla u^0$ and the desired coercivity estimate holds, with the H^2-norm replaced by the norm in $\mathbb{R}$, if $f \neq 0$. A less trivial case is considered in part (c) of the proposition. In the statement of (c) we did not distinguish between a function q and its coordinate expansion q in terms of φ_i. Moreover, we used α to also stand for $\text{col}(\alpha, \cdots, \alpha) \in \mathbb{R}^M$ and we recall that $2w^0 = \gamma^2 - q^{0^T} W^M q^0$.

Proof of Proposition 3.5. (a) First observe that Theorem 2.1 is applicable for local solutions (q^0, u^0) of (P^0). From (3.10) with $\beta = 0$, (3.12) and (3.13) we conclude that

$$M_c(h, v) + \mu^*|h|_{H^2}^2 = |v|_{H_0^1}^2 - 2\langle \nabla\lambda^*, h\nabla v\rangle + c|P(q^0\nabla v + h\nabla u^0)|^2 + \mu^*|h|_{H^2}^2$$

$$\geqq |v|_{H_0^1}^2 - \frac{2K_1}{\alpha}\,\text{dist}\,|h|_{H^2}|v|_{H_0^1} + \mu^*|h|_{H^2}^2$$

$$\geqq |v|_{H_0^1}^2(1 - \varepsilon) + |h|_{H^2}^2\left(\mu^* - \frac{K_1^2\,\text{dist}^2}{\varepsilon\alpha^2}\right)$$

for any $\varepsilon > 0$. The assumption on dist implies the existence of $\sigma > 0$ such that

$$M(h, v) + \mu^*|h|_{H^2}^2 \geqq \sigma(|h|_{H^2}^2 + |v|_{H_0^1}^2)$$

for all $(h, v) \in H^2 \times H_0^1$. The claim now follows with the same argument as at the end of Proposition 3.4.

(b) Since dist$=0$, Theorem 2.1 implies that $\lambda^*=0$, and further

$$\begin{aligned}\nabla^2\tilde{L}_c(q^0, u^0)((h, v), (h, v)) &= |v|^2_{H_0^1}+c|P(q^0\nabla v+h\nabla u^0)|^2+\mu^*|h|^2_{H^2}\\ &\geqq |v|^2_{H_0^1}(1-c|q^0|^2_{L^\infty})+\frac{c}{2}|P(h\nabla u^0)|^2\\ &\geqq |v|^2_{H_0^1}(1-cK_1^2\gamma^2)+\frac{c}{2}|P(h\nabla u^0)|^2.\end{aligned}$$

This estimate implies the claim.

(c) The assumptions on α, q, and φ_i imply that Q_{ad} is nonempty and that $q(x)\geqq\alpha$ for every $q\in Q_{\mathrm{ad}}$. It is simple to argue existence of a solution (q^0, u^0) of $(P^0)^M_c$. Moreover, the conclusions of Theorem 2.1 and Lemma 2.5 remain valid if h and $q^\beta\in H^2$ are replaced by h and $q^0\in\mathbb{R}^M$, if $\mathbb{R}^M$ is endowed with the inner product $\langle h, W^M h\rangle$, and if $\mathscr{C}=\mathbb{R}^M_-$, $\ell(q)=\alpha-q\in\mathbb{R}^M$. In particular, there exists a Lagrange muitiplier $(\lambda^*, \mu^*, \eta^*)$ with the specified properties.

For $h=\sum_{i=1}^M h_i\varphi_i\in V^M$ and $v\in H_0^1$ we find

$$\begin{aligned}M_c(h, v) &:= |v|^2_{H_0^1}+2\langle\nabla\lambda^*, h\nabla u\rangle+c|P(q^0\nabla v+h\nabla u^0)|^2\\ &\geqq |v|^2_{H_0^1}-\frac{2}{\alpha}\,\mathrm{dist}\,|h|_{L^\infty}|v|_{H_0^1}+\frac{c}{2}|P(h\nabla u^0)|^2-c|P(q^0\nabla v)|^2\\ &\geqq (1-c|q^0|^2_{L^\infty})|v|^2_{H_0^1}-\frac{2}{\alpha b}\,\mathrm{dist}\,|v|_{H_0^1}|P(h\nabla u^0)|+\frac{c}{2}|P(h\nabla u^0)|^2.\end{aligned}$$

If $\tilde{c}_3=2(b^2+2|q^0|_{L^\infty})^{-1}$ then for $c\geqq c_3$ the following inequality holds:

$$M_c(h, v)\geqq\sigma|v|^2_{H_0^1}-\frac{2}{\alpha b}\,\mathrm{dist}\,|v|_{H_0^1}|P(h\nabla u^0)|+\frac{\sigma}{b^2}|P(h\nabla u^0)|^2$$

where $\sigma=b^2(b^2+2|q^0|^2_{L^\infty})^{-1}$. Thus, if dist$<\sigma\alpha$, then there exists a constant $\tilde{\sigma}_3>0$ such that

$$M_c(h, v)\geqq\tilde{\sigma}_3(|v|^2_{H_0^1}+|P(h\nabla u^0)|^2) \tag{3.21}$$

for all $c\geqq\tilde{c}_3$. Since $\nabla\hat{L}_c^2(q^0, u^0, w^0; \lambda^*, \mu^*, \eta^*)((h, v, y)(h, v, y))=M_c(h, v)+\mu^*h^TW^Mh+c(q^{0^T}W^Mh+y)^2$, the final claim follows from (3.21) and an argument analogous to the one at the end of Proposition 3.4.

Next we consider the one-dimensional case where we have an explicit formula for the orthogonal projection P and explicit values for the estimates in Lemma 3.3. By D we denote differentiation and the domain Ω is $(0, 1)$.

LEMMA 3.6. *The operator $P=D\Delta^{-1}D$ is an orthogonal projection on L^2. Moreover,* ker P *is the set of all constant functions on* $(0, 1)$ *and* $|P\varphi|^2=|\varphi|^2-(\int_0^1\varphi\,dx)^2$ *for* $\varphi\in L^2$.

Proof. The first part of the lemma is obvious. Next recall that $\{1, \sqrt{2}\cos\pi x, \sqrt{2}\cos 2\pi x, \cdots\}$ is a complete orthonormal system in L^2. For $\varphi\in H^1$ with $\varphi=a_0+\sum_{k=1}^\infty a_k\sqrt{2}\cos k\pi x$ it is easy to see that $P\varphi=\sum_{k=1}^\infty a_k\sqrt{2}\cos k\pi x$ and $(I-P)\varphi=a_0$. Since H^1 is dense in L^2 this follows for all $\varphi\in L^2$ and, moreover, $|P\varphi|^2=|\varphi|^2-|(I-P)\varphi|^2=|\varphi|^2-(\int_0^1\varphi\,dx)^2$.

Lemma 3.7. (a) $|\varphi|_{L^\infty}\leqq\sqrt{2}|\varphi|_{H^1}$ *for all* $\varphi\in H^1$,

(b) $|\varphi|_{L^\infty}\leqq 1/\sqrt{3}|D\varphi|$ *for all* $\varphi\in H^1$ *with* $\int_0^1\varphi\,dx=0$,

(c) $\pi|\varphi|\leqq|D\varphi|$ *for all* $\varphi\in H^1$ *with* $\int_0^1\varphi\,dx=0$.

Proof. (a) Let $\varphi\in H^1$. By the Mean Value Theorem there exists $\zeta\in[0, 1]$ such that $\varphi(\zeta)=\int_0^1\varphi(x)\,dx$. For every $x\in[0, 1]$ we have

$$\varphi(x)=\int_\zeta^x D\varphi(s)\,ds+\varphi(\zeta).$$

This implies

$$|\varphi(x)| \leqq \int_0^1 (|D\varphi(s)|+|\varphi(s)|)\,ds \leqq \left(\int_0^1 (|D\varphi(s)+|\varphi(s)|)^2\,ds\right)^{1/2} \leqq \sqrt{2}|\varphi|_{H^1}.$$

(b) By assumption, φ can be expressed as $\varphi=\sum_{k=1}^{\infty} a_k\sqrt{2}\cos k\pi x$. Therefore $|D\varphi|^2=\pi^2\sum_{k=1}^{\infty} k^2 a_k^2$, and

$$|\varphi|_{L^\infty} \leqq \sqrt{2}\sum_{k=1}^{\infty}|a_k| = \sqrt{2}\left(\sum_{k=1}^{\infty} k^2|a_k|^2\right)^{1/2} (\sum k^{-2})^{1/2}$$

$$=\sqrt{2}\frac{|D\varphi|}{\pi}\cdot\frac{\pi}{\sqrt{6}}=\frac{|D\varphi|}{\sqrt{3}},$$

which was to be proved.

(c) The assertion immediately follows from

$$|\varphi|^2=\sum_{k=1}^{\infty}|a_k|^2 \quad\text{and}\quad |D\varphi|^2=\pi^2\sum_{k=1}^{\infty}k^2|a_k|^2 \geqq \pi^2\sum_{k=1}^{\infty}|\alpha_k|^2.$$

Analogously to the multidimensional case we find a lower bound on the solutions of $-D(qDu)=f$ for $q\in Q_{\text{ad}}$:

$$|f|_{H^{-1}}=\sup_{V\in H_0^1}\frac{\langle qDu, Dv\rangle}{|v|_{H_0^1}} \leqq |q|_{L^\infty}|u|_{H_0^1} \leqq \sqrt{2}|q|_{H^1}|u|_{H_0^1} \leqq \sqrt{2}\gamma|u|_{H_0^1},$$

and thus

$$|u(q)|_{H_0^1} \geqq \omega_1 := \frac{|f|_{H^{-1}}}{\sqrt{2}\gamma} \quad \text{for all } q\in Q_{\text{ad}}. \tag{3.22}$$

PROPOSITION 3.8. *Let $f\in H^{-1}$ satisfy $f\neq 0$ and let $k=3\pi^2/(7\pi^2+3)$. Let (q^0, u^0) be a solution of* (P^0) *and suppose that for a constant $\beta_0\in(0,1)$*

$$|u^0-z|_{H_0^1}^2 < \beta_0\left[\frac{\alpha^2 k}{2}\left(1-\frac{8k}{\omega_1}\gamma^2 B_0\right)-\rho(\beta_0)\right]. \tag{3.23}$$

Then there exists a nontrivial compact interval $I=[\underline{\beta}, \beta_0]\subset(0,1)$ and constants $c_0>0$, $\sigma_0>0$ such that

$$\nabla^2 L_c(q^\beta, u^\beta, w^\beta; \lambda^*, \mu^*, \eta^*)((h, v, y), (h, v, y)) \geqq \sigma_0(|h|_{H^1}^2+|v|_{H_0^1}^2+|y|^2)$$

for all $c\geqq c_0$, $(h, v, y)\in H^2\times H_0^1\times\mathbb{R}$ and any solution (q^β, u^β) of (P^β) with $\beta\in I$. Moreover, if $u^0=z$, then such a constant β_0 always exists and I can be chosen as any compact subinterval of $(0, \beta_0]$.

Using Lemmas 3.6 and 3.7 the proof of this proposition is quite analogous to that of Proposition 3.4. In the present case (K_1, K_2, K_3) is replaced by $(\sqrt{2}, 1/\pi, \sqrt{(1/\pi^2)+1})$.

Special cases in which the coercivity estimate holds with $\beta=0$ are quite similar to the multidimensional case and hence we will not explicitly state the analogue of Proposition 3.5 in dimension one.

4. Numerical results. In this section we briefly report on our practical experience with the augmented Lagrangian technique (2.8), (2.9) to determine q in (1.1) from data for u. We carried out extensive testing in dimensions one and two. These results will be presented in a forthcoming paper [8*] (see also [10], [11]) and we will therefore give only two typical examples.

Example 1. This is the problem of determining q in

$$-(qu_x)_x = f \quad \text{on } (0,1), \qquad u(0)=u(1)=0 \tag{4.1}$$

where

$$f(x)=\begin{cases} (18x-6)\dfrac{3\pi}{2}\cos\dfrac{3\pi x}{2}+18\left(\dfrac{3}{2}+\sin\dfrac{3\pi x}{2}\right) & \text{for } x\in[0,\frac{1}{3}), \\ 0 \quad \text{for } x\in(\frac{1}{3},\frac{2}{3}), \\ (18x-12)\dfrac{3\pi}{2}\cos\dfrac{3\pi x}{2}+18\left(\dfrac{3}{2}+\sin\dfrac{3\pi x}{2}\right) & \text{for } x\in[\frac{2}{3},1], \end{cases}$$

and the "true coefficient" q^* is given by

$$q^*(x)=\frac{3}{2}+\sin\frac{3\pi x}{2}.$$

The corresponding solution $u(q^*)=z$ of (4.1) is

$$u(q^*)=\begin{cases} -9x^2+6x \quad \text{for } x\in[0,\frac{1}{3}], \\ 1 \quad \text{for } x\in(\frac{1}{3},\frac{2}{3}], \\ -9x^2+12x-3 \quad \text{for } x\in(\frac{2}{3},1]. \end{cases}$$

With f and z specified, it is immediately clear that q is not unique within the class of positive H^1 functions that satisfy $u(q)=z$, since its value over the interval $S=(1/3,2/3)$ does not effect the solution u there. On the other hand, with the techniques of [12] it can easily be argued that $u(q)=u(q^*)$, $q\in H^1$, $q^*\in H^1$ implies $q=q^*$ on $[0,1]\backslash S$. Thus we expect a different behavior of the algorithm over S than over the complement of S.

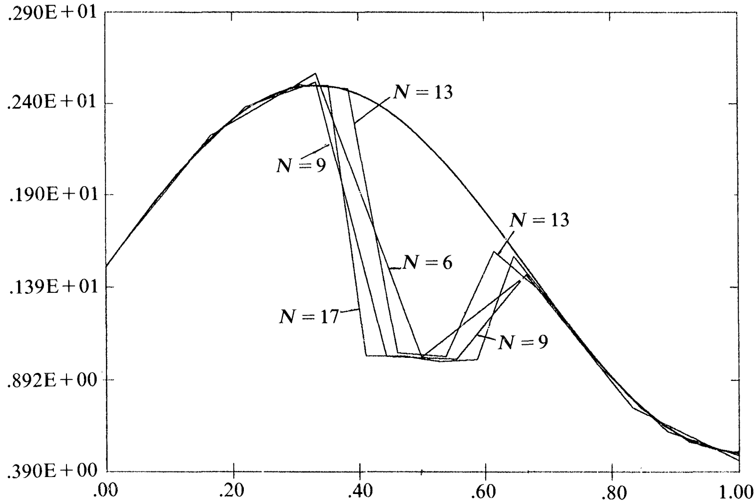
In Fig. 1 we give the numerical results for various values of N. Here N represents the index of discretization of the infinite-dimensional problem (2.8) by finite-dimensional ones involving linear spline subspaces H^N for the statespace H_0^1 and V^N for the parameter space H^1. More precisely we take

$$H^N=\operatorname{span}\{B_i^{2N}\}_{i=1}^{2N-1} \quad\text{and}\quad V^N=\operatorname{span}\{B_j^N\}_{j=0}^N,$$

where B_i^N is the usual first-order B-spline on the interval $[0,1]$ corresponding to the mesh $\{x_k^N=k/N\}$, $k=0,\cdots,N$:

$$B_k(x)=\begin{cases} N(x-x_{k-1}^N) & \text{for } x_{k-1}^N\leqq x\leqq x_k^N, \\ N(x_{k+1}^N-x) & \text{for } x_k^N\leqq x\leqq x_{k+1}^N, \\ 0 \quad \text{elsewhere} \end{cases}$$

where $x_{-1}^N=0$, $x_{N+1}^N=1$. Figure 1 gives the results after one iteration of the augmented Langrangian algorithm where $\lambda^1=\mu^1=0$, $\gamma^2=100{,}000$ and $\beta=0$. The start-up value for the minimization routine to solve (2.8) was taken as $(q^0,u^0)=(1,0)$. The corresponding value for u^{17} is indistinguishable from $u(q^*)(=z)$ on all of $(0,1)$. For this example the use of the regularization term did not change the results significantly. In other examples with the same value for z, but with different values for q^* (and thus of f) the use of the regularization term decreased the L^2-error for q^N-q^*. The penalty



FUNCTION q METHOD ALG1
$n=1$, $\gamma^2=100{,}000$, $N=6, 9, 13, 17$, $\lambda_0=0$,
$\mu_0=0$, L^2-error $=4.134\mathrm{E}-1$, $5.164\mathrm{E}-1$, $4.286\mathrm{E}-1$, $5.514\mathrm{E}-1$

FIG. 1

parameters c_k were taken to be 1 for all k. The augmented Lagrangian algorithm for this example (as well as for the other examples that we tried) was quite insensitive to the choice of u^0 and q^0 as well as the choice of λ^1 and μ^1 and c_k. However, $\lambda^1 \neq 0$, $\mu^1 \neq 0$ increased the number of iterations that were required before convergence was obtained. We compared the augmented Lagrangian method with the output least squares approach and found that it is less sensitive with respect to the choice of the regularization parameter [10].

Example 2. Here we estimate q in

$$-(qu_x)_x-(qu_y)_y=f \quad \text{on } \Omega \tag{4.2}$$
$$u=0 \quad \text{on } \Gamma$$

where $\Omega=[0,1]\times[0,1]$ and

$$f=8\pi^2 \sin 2\pi x \sin 2\pi y(1+6x^2y(1-y))-24\pi xy(1-y)\cos 2\pi x \sin 2\pi y$$
$$-12\pi x^2(1-2y)\sin 2\pi x \cos 2\pi y.$$

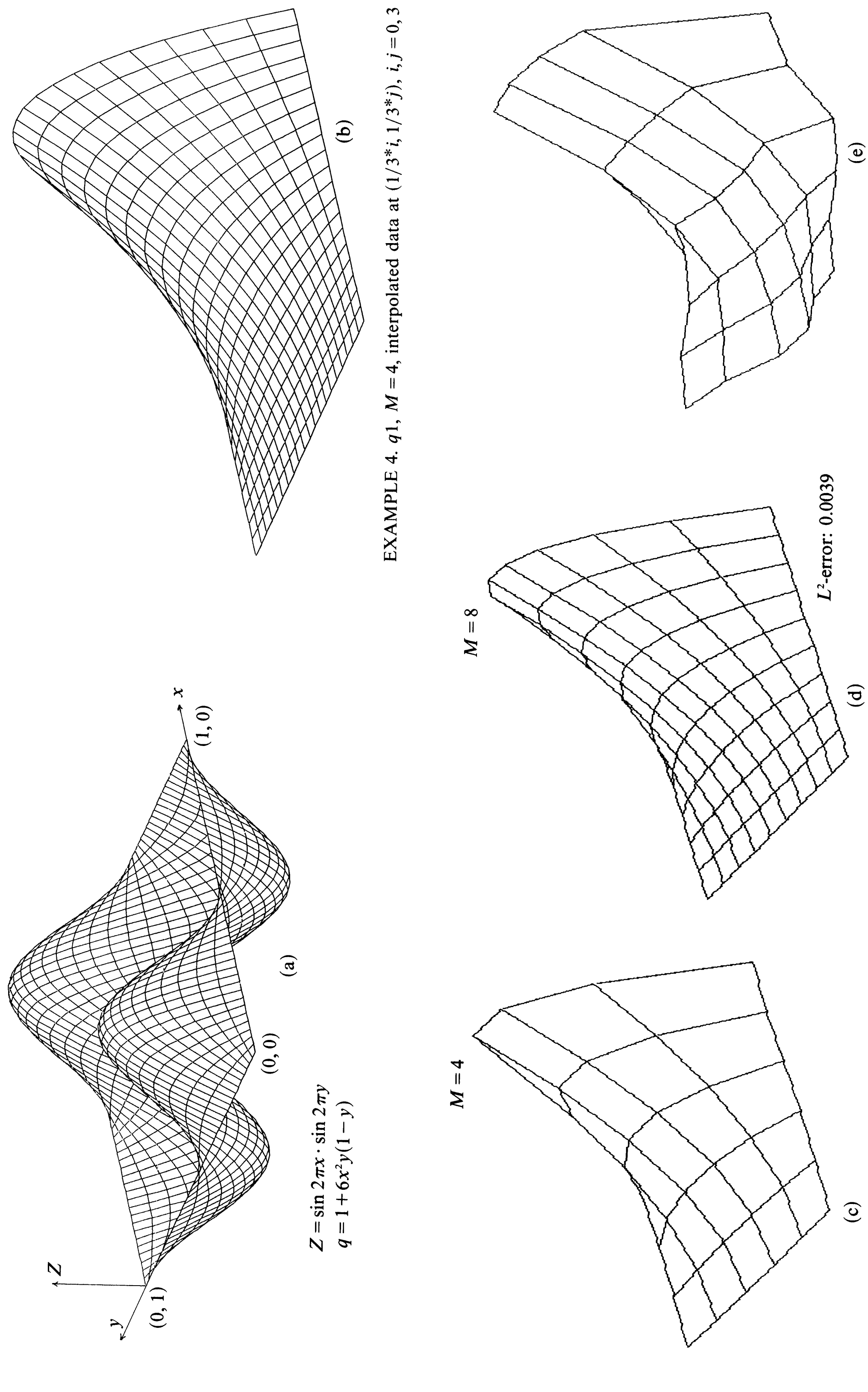
The true solution q^* is

$$q^*=1+6x^2y(1-y)$$

and the corresponding solution $u(q^*)=z$ for (4.2) is given by

$$u(q^*)=\sin 2\pi x \sin 2\pi y.$$

The discretization (4.2) is carried out by taking tensor linear spline subspaces $H^N \otimes H^N$ for the statespace H_0^1 [18], and tensor linear spline subspaces $V^N \otimes V^N$ for the coefficient space. Figures 2(a) and 2(b) give the graphs for z and q^*, respectively. The results after eight iterations of the augmented Lagrangian algorithm with $N=5$ and $N=9$ are given in Figs. 2(c) and 2(d). The results after just one iteration are essentially identical. These results are obtained with $\beta=\lambda^1=\mu^1=0$ and $q_0=1$. We also carried out calculations where we assumed that only partial observations are available.



EXAMPLE 4. $q1$, $M = 4$, interpolated data at $(1/3^*i, 1/3^*j)$, $i, j = 0, 3$

FIG. 2

Specifically, we took the values of $u(q^*)$ at the grid $[(.2i, .2j): i, j = 0, \cdots, 5\}$ and calculated a bicubic interpolation z. Using this z in the augmented Lagrangian algorithm (2.8), (2.9) the resulting plot for q^5 is almost indistinguishable from Fig. 2(c). Then we tried the same procedure with data at $\{(i/3, j/3): i, j = 0, \cdots, 3\}$ and the result for q^5 from these interpolated data is shown in Fig. 2(e).

Overall the augmented Lagrangian approach to estimate q in (1.1) proved to be very effective. This is especially true for the two-dimensional problem, where earlier experiments with the output least squares technique were not very encouraging numerically. Clearly, there is a wide variety of choices for implementing (2.8), (2.9). One variant of (2.8), (2.9) that proved to be effective numerically is the following (we specify it for $n = 2$ or 3).

Step 1. Choose $\lambda^1 = \mu^1 = 0$, $\{c_k\}_{k=1}^\infty$ monotonically increasing $c_k > c_0$.

Step 2. Put $k = 1$, $u_0 = z$.

Step 3. Determine q_k from

$$(P_{\text{equ}}) \text{ minimize } \quad \frac{\beta}{2} N(q) + \langle \lambda^k, e(q, u_{k-1}) \rangle_{H_0^1} + \frac{c_k}{2} |e(q, u_{k-1})|^2_{H_0^1}$$

$$+ \mu^k \hat{g}(q, \mu^k, c_k) + \frac{c_k}{2} \hat{g}(q, \mu^k, c_k)^2$$

over $q \in H^2$ subject to $q \geqq \alpha$.

Step 4. Determine u_k from

$$(P_{\text{out}}) \underset{u \in H_0^1}{\text{minimize}} \quad \frac{1}{2} |u - z|^2_{H_0^1} + \langle \lambda^k, e(q_k, u) \rangle_{H_0^1} + \frac{c_k}{2} |e(q_k, u)|^2_{H_0^1}.$$

Step 5. $\lambda^{k+1} = \lambda^k + c_k e(q_k, u_k)$ and $\mu^{k+1} = \mu^k + c_k \hat{g}(q_k, \mu^k, c_k)$.

Step 6. If convergence is achieved, stop; otherwise put $k = k + 1$ and go to Step 3.

In our implementation for the two-dimensional problem we dropped the second-order derivative terms in the regularization functional of Step 3. This is partially because we used piecewise linear and piecewise constant functions to approximate q and hence second derivatives could only be taken approximately. Moreover, we expect that the second-order derivatives are only required analytically since we choose our coefficients q as H^2 functions, but we expect that this is not essential numerically.

Acknowledgment. We are grateful to Dr. M. Kroller who carried out the numerical calculations that we reported in this section.

REFERENCES

[1] D. P. BERTSEKAS, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.

[2] G. CHAVENT, *Identification of distributed parameter systems: about the output least square method, its implementation and identifiability*, in Proc. 5th IFAC Symposium on Identification and System Parameter Estimation, Pergamon Press, New York, 1979, pp. 85-97.

[3] F. COLONIUS AND K. KUNISCH, *Stability for parameter estimation in two point boundary value problems*, J. Reine Angew. Math., 370 (1986), pp. 1-29.

[4] ———, *Output least squares stability in elliptic systems*, Appl. Math. Optim., 19 (1989), pp. 33-63.

[5] M. FORTIN AND R. GLOWINSKI, *Augmented Lagrangian Methods: Applications to Numerical Solutions of Boundary Value Problems*, North-Holland, Amsterdam, 1983.

[6] C. W. GROETSCH, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Research Notes in Mathematics 105, Pitman, Boston, 1984.
[7] M. R. Hestenes, *Optimization Theory, The Finite Dimensional Case*, John Wiley, New York, 1975.
[8] K. ITO AND K. KUNISCH, *The augmented Lagrangian method for equality and inequality constraints in Hilbert spaces*, Math. Programming, to appear.
[8*] K. ITO, M. KROLLER, AND K. KUNISCH, *A numerical study of the augmented Lagrangian method for the estimation of parameters in elliptic systems*, submitted.
[9] C. KRAVARIS AND J. H. SEINFELD, *Identification of parameters in distributed systems by regularization*, SIAM J. Control Optim., 23 (1985), pp. 217–241.
[10] M. KROLLER AND K. KUNISCH, *A numerical study of augmented Lagrangian method for the estimation of parameters in a two point boundary value problem*, Technical Report No. 87, Institute for Mathematics, Technical University of Graz, Graz, Austria, March 1987.
[11] ———, *A numerical study of an augmented Lagrangian method for the estimation of parameters in elliptic systems*, Technical Report, No. 101, Institute for Mathematics, Technical University of Graz, Graz, Austria, September 1987.
[12] K. KUNISCH, *Inherent identifiability of parameters in elliptic equations*, J. Math. Anal. Appl., 132 (1988), pp. 453–472.
[13] K. KUNISCH AND L. W. WHITE, *Regularity properties in parameter estimation of diffusion coefficients in elliptic boundary value problems*, Appl. Anal., 21 (1986), pp. 71–88.
[14] H. MAURER AND J. ZOWE, *First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, Math. Programming, 16 (1979), pp. 98–110.
[15] J. NEĆAS, *Les Méthodes Directes en Théorie des Équations Elliptiques*, Masson, Paris, 1967.
[16] V. T. POLYAK AND N. Y. TRET'YAKOV, *The method of penalty estimates for conditional extremum problems*, Z. Vychisl. Mat. I Mat. Fiz., 13 (1973), pp. 34–36.
[17] D. L. RUSSELL, *Some remarks on numerical aspects of coefficient identification in elliptic systems*, in Optimal Control of Partial Differential Equations, K. H. Hoffmann and W. Krabs, eds., Birkhäuser, Boston, 1984, pp. 210–228.
[18] M. H. SCHULZ, *Spline Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1973.
[19] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, New York, 1978.