Output Least Squares Stability in Elliptic Systems*

F. Colonius[1] and K. Kunisch[2]

[1] Institut für Dynamische Systeme, Universität Bremen, Bibliothekstrasse, Pf. 330440, D-2800 Bremen, West Germany

[2] Institut für Mathematik, Technische Universität Graz, Kopernikusgasse 24, A-8010 Graz, Austria

Abstract. In this paper the stability of the solutions of parameter estimation problems in their output least squares formulation is analyzed. The concepts of output least squares stability (OLS stability) is defined and sufficient conditions for this property are proved for abstract elliptic equations. These results are applied to the estimation of the diffusion, convection, and friction coefficient in second-order elliptic equations in $\mathbb{R}^n$, $n=2,3$. Results on Tikhonov regularization in a nonlinear setting are also given.

1. Introduction

In this paper we study continuous dependence of the solutions of the output least squares (OLS) formulation of parameter estimation problems on the problem data. It is assumed that the model equation is of the form

$$A(q)u=f, \tag{1.1}$$

where $A(q)$ is a linear elliptic operator depending on a functional parameter q and that we have observations z of the system for which (1.1) is the model

* Part of this research was carried out while both authors were visitors at the Division of Applied Mathematics at Brown University. The research of F. Colonius was supported in part by the Deutsche Forschungsgemeinschaft. Both authors also acknowledge support from the Fonds zur Förderung der wissenschaftlichen Forschung, Austria, under Grant No. S3206.

equation. This observation may arise from interpolation of point data. One common approach to determine q from z is given by the OLS formulation

$$\text{minimize } |u(q)-z|^2 \tag{1.2}$$

over some set Q_{ad} of admissible parameters. This problem is not well posed, in general, which means that the solutions of (1.2) do not depend continuously on the data z. For example, consider the special case $z = u(q^*)$, so that z corresponds to some "true" model parameter. Then it is well known that, in general, the mapping $q \to u(q)$ is not continuously invertible. We refer to [2], [16], and [17] for further discussion of these matters.

To obtain stability inspite of this difficulty, we need to change the mathematical formulation of the problem or to make specific assumptions concerning the structure of the original problem (1.2). One possibility is to add a regularization term to the cost functional in (1.2):

$$\min |u(q)-z|^2 + \beta |q|_Q^2 \text{ over } Q_{ad}, \tag{1.3}$$

where $\beta > 0$ is a real "regularization" parameter and $Q_{ad} \subset Q$ with Q a normed linear space. The use of regularization is very common for inversion of linear operators, see [10] and [19], but it is not well studied in the nonlinear environment that is needed for parameter estimation problems. (Note that $q \to u(q)$ is not linear.) Therefore the first part of this paper is devoted to a study of the properties of the global solutions of (1.3), which are not unique, in general, and in particular the behavior of these solution sets as $\beta \to 0^+$. Subsequently, the stability of the regularized problems is studied. The essential technical tool that is used is a stability result from abstract optimization theory, which strongly depends on higher- (here second-) order sufficient optimality conditions. It is instructive to consider formally the second derivative of the Lagrangian of (1.3) evaluated at a solution q^β of (1.3) in direction (h, h):

$$\begin{aligned} F_{qq}(q^\beta)(h, h) = 2|\eta(h)|^2 + 2(u(q^\beta) - z, \xi(h, h)) + 2\beta |h|_Q^2 \\ + (\text{terms involving Lagrange multipliers} \\ \text{and the constraints}); \end{aligned} \tag{1.4}$$

here $\eta(h)$ and $\xi(h, h)$ are the first and second derivative of u at q^β in direction h and (h, h), respectively. From (1.4) it is apparent why regularization is helpful in obtaining positivity of $F_{qq}(q^\beta)$ and consequently stability of (1.3). We also observe that the distance between the solution of (1.1) evaluated at the solution q^β and the observation z enters in an important manner. Of course we will try to take $\beta > 0$ as small as possible and still obtain positivity of $F_{qq}(q^\beta)$. This will depend (among other considerations) on the size of $|u(q^\beta) - z|$ and on possible lower bounds for $\eta(h)$. In general we do not assume the existence of $q^* \in Q_{ad}$ with $u(q^*) = z$.

We mention that in a recent paper [12] Kravaris and Seinfeld have studied the use of Tikhonov regularization for parameter estimation in parabolic partial differential equations. Their approach is based on a variant of Tikhonov's lemma, which states that if a continuous function f between metric spaces X and Y is injective on a precompact subset $K \subset X$, then f is continuously invertible on

$f(K)$. The Tikhonov approach of [12] requires uniqueness of the solution of the unregularized problem, which is not needed in our analysis. In our approach stability is checked at each solution of (1.3). If stability can be guaranteed then such solutions are isolated. Furthermore, we distinguish between the situations where z^0 can be obtained as the solution $u(q^0)$ of the state equation evaluated at the "true" parameter q^0 and where this is not the case.

As indicated above, if stability holds without the use of regularization, then further assumptions must be made. One such assumption is the finite dimensionality of Q, which is feasible for practical applications. Another one is to make assumptions on the values of the solution of (1.1) evaluated at the solution $\bar{q}$ of (OLS). Finally, we point out that we consider not only stability with respect to the observations z but also with respect to the constraints defining Q_{ad}.

Concerning the stability problem in parameter estimation we also mention the work of Chavent [6] who introduced the notion of output least squares identifiability (OLSI). A parameter is OLSI if there exists a neighborhood $\mathcal{N}$ of the attainable set $\mathcal{V}$ ($\mathcal{V} = u(Q_{\mathrm{ad}})$ in our case), such that for every element $z \in \mathcal{N}$ there exists a unique solution $q \in Q_{\mathrm{ad}}$ of OLS depending continuously on z. Chavent derives general sufficient conditions involving $\operatorname{dist}(z, \mathcal{V})$, diam Q_{ad}, and lower and upper bounds on the second derivative of u with respect to the parameter, which imply OLSI. These results are well suited for certain classes of ordinary differential equations and hyperbolic partial differential equations.

The present paper is a continuation of the research started in [8], where we introduced the concepts of OLS stability and ROLS stability (OLS stability by regularization) and applied them in the simple case of estimating c in the two-point boundary-value problem $-(au_x)_x + cu = f$. In the present paper we first consider abstract elliptic equations and then apply our results to the estimation of the diffusion, convection, and friction coefficient in second-order elliptic equations in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Chavent recently extended his results to include the advantages of regularization [7]. The main difference between Chavent's work and our's is given by the fact that Chavent's concept of uniqueness and stability is a global one, whereas our's is local. As a consequence these two approaches employ completely different mathematical tools.

In [8] we also made use of the last term in (1.4) to obtain positivity of $F_{qq}(\bar{q})$ or $F_{qq}(q^\beta)$ and we could repeat this in the general case by making assumptions that guarantee nontriviality of certain Lagrange multipliers. Differently from the one-dimensional case [8], [13], however, it seems to be unfeasible to give explicit conditions in terms of the quantities of the differential equation that imply nontriviality of the Lagrange multipliers.

The paper is organized as follows: Section 2 gives an informal discussion of the method we employ for studying sensitivity with respect to parameter variations. We hope that this motivates the reader to study the abstract theory in the next two sections. Section 3 contains results on Tikhonov regularization in a nonlinear setting and is independent of the elliptic structure of (1.1). In Section 4 we define the concept of ROLS stability and prove sufficient conditions for this property. These results are applied to several examples in Section 5. The first part of Section 6 is devoted to obtaining stability without the benefit of regularization and it closes with an example for local regularization.

2. An Introductory Example

In this section we informally discuss the method that we propose for the study of sensitivity in parameter estimation problems by means of a simple two-point boundary-value problem. Consider

$$\begin{aligned} -(qu_x)_x &= f \quad \text{on } (0, 1), \\ u(0) &= u(1) = 0, \end{aligned} \tag{2.1}$$

where $f \in L^2$. Throughout this section all function spaces are considered over the interval (0, 1). It is well known that for every positive $q \in H^1$ there exists a unique solution $u(q) \in H^2 \cap H_0^1$ of (2.1). The output least squares formulation to determine q from an observation $z^0 \in L^2$ is given by

$$\text{(OLS)} \quad \min |u(q) - z^0|_{L^2}^2 \text{ over } Q_{\text{ad}},$$

where Q_{ad} is the set of admissible parameters. It is chosen in such a way that $q \in Q_{\text{ad}}$ guarantees existence of a solution of (2.1) as well as of (OLS). For the former we require a pointwise bound on q and for the latter a norm bound, i.e., we take

$$Q_{\text{ad}} = \{q \in H^1 \colon q(x) \geq k^0(x), |q|_{H^1} \leq \gamma^0\},$$

where $k^0 \in H^1$, $k^0 > 0$, and $\gamma^0 > 0$. It is then simple to argue the existence of a solution q^0 of (OLS). Our goal is the study of the continuous dependence of q^0 on perturbations in the observation and the set of admissible parameters Q_{ad}, or, equivalently, the continuous dependence of q^0 on $w^0 = (z^0, k^0, \gamma^0) \in L^2 \times H^1 \times \mathbb{R}$. Henceforth we consider w^0 as the unperturbed reference parameter and w as a perturbed parameter. Since the solution q^0 of the unperturbed problem need not be unique we have to specify our notion of continuous dependence: we require the existence of neighborhoods $V(w^0) \subset L^2 \times H^1 \times \mathbb{R}$ and $V(q^0) \subset Q$ of w^0 and q^0 such that for every $w \in V(w^0)$ there exists at least one (local) solution q_w of (OLS) (with w^0 replaced by w) and all local solutions of (OLS) which lie in $V(q^0)$ depend continuously on w. Observe that regardless of the choice of q^0 as a global or local solution of (OLS) we need to allow for the solutions q_w of the perturbed problem to be local solutions only. (For example, consider in $\mathbb{R}^1$ the family of functions $f(x; \varepsilon) = x^4 + \varepsilon x^3 + (\varepsilon - 1)x^2$, $\varepsilon < 1$. Then $f(\cdot; 0)$ has two global minima and $f(\cdot, \varepsilon)$, $\varepsilon \neq 0$, has two local minima, only one of which is also global.)

The sensitivity analysis of optimization problems in finite-dimensional spaces suggests that if the cost functional is sufficiently regular with respect to q and z^0, and if Q_{ad} depends continuously on (k^0, γ^0) in an appropriate sense, then a second-order sufficient optimality condition guarantees continuous dependence of q on (z^0, k^0, γ^0) [9]. This is studied in a rigorous framework in the ensuing sections. Here we only discuss the feasibility of these requirements for the specific estimation problem (OLS) arising from (2.1). First observe that $(q, z) \to |u(q) - z|^2$

from $Q_{ad} \times L^2 \subset H^1 \times L^2$ to $\mathbb{R}$ is continuous. Moreover, $q \to u(q)$ is twice continuously Fréchet differentiable with respect to q. Let η and ξ denote respectively the first and second Fréchet-derivative at q^0 in directions $h \in H^1$ and $(h, h) \in H^1 \times H^1$. Then η and ξ are characterized by

$$-(q^0 \eta_x)_x = (hu_x(q^0))_x, \qquad \eta(0) = \eta(1) = 0, \tag{2.2}$$

and

$$-(q^0 \xi_x)_x = 2(h\eta_x)_x, \qquad \xi(0) = \xi(1) = 0. \tag{2.3}$$

The set Q_{ad} can also be characterized as follows: let $g(\cdot\,; k^0, \gamma^0)$: $H^1 \to H^1 \times \mathbb{R}$ be given by

$$g(q; k^0, \gamma^0) = (k^0 - q, |q|^2_{H^1} - (\gamma^0)^2)$$

and let $K = C^+ \times \mathbb{R}^+$. Clearly, $q \in Q_{ad}$ if and only if $g(q) \in -K$. Moreover, g is seen to be smooth with respect to k^0 and γ^0. Finally observe that there exist constants c_1 and c_1^* such that

$$|u|_{H^2} \le c_1 |A(q)u|_{L^2} \le c_1^* |u|_{H^2} \tag{2.4}$$

for all $q \in Q_{ad}$ and $u \in H^2 \cap H_0^1$. Here we put $A(q)u = (qu_x)_x$.

To investigate a second-order sufficient optimality condition we introduce the Lagrangian associated with (OLS):

$$F(q) = |u(q) - z|^2 + \langle \lambda_1, k^0 - q \rangle_* + \lambda_2(|q|^2_{H^1} - (\gamma^0)^2),$$

where $\langle \cdot, \cdot \rangle_*$ denotes the duality pairing between H^1 and $(H^1)^*$ and $\lambda_2 \ge 0$. A second-order optimality condition is satisfied if $F_{qq}(q^0)$, the second Fréchet derivative of F at q^0, is uniformly positive on an appropriately chosen subset of H^1. Thus we are led to investigate positivity of

$$F_{qq}(q^0)(h, h) = |\eta|^2_{L^2} + \langle u(q^0) - z^0, \xi \rangle_{L^2} + \lambda_2 |h|^2_{H^1}, \tag{2.5}$$

where η and ξ are given in (2.2) and (2.3). If $\lambda_2 > 0$ and $|u(q^0) - z^0|$ is sufficiently small then we can show quite easily the existence of $\kappa > 0$ such that

$$F_{qq}(q^0)(h, h) \ge \kappa |h|^2_{H^1} \qquad \text{for all} \quad h \in H^1, \tag{2.6}$$

which is the desired estimate.

In general it is not a simple matter to give conditions which imply nontriviality of λ_2. Moreover, $\lambda_2 > 0$ implies $|q^0|_{H^1} = \gamma^0$ and thus it is evident that this is only a special case. If $\lambda_2 = 0$, then positivity of $F_{qq}(q^0)$ as in (2.6) is not possible, in general. In fact, $|\eta|^2_{L^2}$ behaves like $|hu_x(q^0)|^2_{H^{-1}}$ and the second term in (2.5) can only be bounded from below in the form

$$\langle u(q^0) - z^0, \xi \rangle_{L^2} \ge -\tilde{\kappa} |u(q^0) - z^0| \, |h|^2_{H^1}, \qquad \tilde{\kappa} > 0.$$

This estimate can be verified by (2.4), for example. The right-hand side in the last inequality is zero if and only if z^0 is in the attainable set, i.e., $z^0 \in \mathcal{V} = \{u(q) \colon q \in Q_{ad}\}$. This, however, will not be satisfied in general.

One remedy to the failure of $F_{qq}(q^0)$ to be positive is the use of a regularization term in the output least squares formulation. Thus we are led to consider

$$(\text{ROLS})^\beta \quad \min |u(q) - z^0|^2 + \beta |q|^2_{H^1} \text{ over } Q_{ad},$$

where β is a (small) positive parameter. The solutions of $(\mathrm{ROLS})^\beta$ are denoted by q^β. Since it appears that regularization theory is not well studied in a nonlinear environment (note that $q \to u(q)$ is not linear) we devote Section 3 to a study of some basic properties of regularized optimization problems. In particular, we shall show that the global solutions q^β of $(\mathrm{ROLS})^\beta$ satisfy

$$\lim_{\beta\to 0^+} \beta^{-1}(|u(q^\beta)-z^0|^2 - |u(q^0)-z^0|^2) = 0. \tag{2.7}$$

In the special case that $z^0 \in \mathcal{V}$ (2.7) becomes

$$\lim_{\beta\to 0^+} \beta^{-1}|u(q^\beta)-z^0|^2 = 0. \tag{2.8}$$

The analogue of (2.5) for $(\mathrm{ROLS})^\beta$ is

$$F^\beta_{qq}(q^\beta)(h,h) = |\eta|^2_{L^2} + \langle u(q^\beta)-z^0, \xi\rangle + 2(\lambda_2+\beta)|h|^2_{H^1}, \tag{2.9}$$

where now η and ξ are calculated at q^β. Positivity of F^β can be obtained if β is chosen appropriately: the term

$$\langle u(q^\beta)-z^0, \xi\rangle = -2\langle [h(A^{-1}(q^\beta)(u(q^\beta)-z^0))_x]_x, \eta\rangle$$

can be bounded from below by $-|\eta|^2 - \tilde{\kappa}|h|^2_{H^1}|u(q^\beta)-z^0|^2_{L^2}$, where $\tilde{\kappa}>0$. If $z^0 \in \mathcal{V}$, then from (2.8) we find that $F^\beta_{qq}(q^\beta)(h,h) \ge [2(\lambda_2+\beta) - \beta\rho(\beta)]|h|^2_{H^1}$ with $\lim_{\beta\to 0}\rho(\beta)=0$. Hence there always exist $\bar{\beta}>0$, $\kappa>0$ such that

$$F^\beta_{qq}(q^\beta)(h,h) \ge \kappa|h|^2_{H^1} \quad \text{for all} \quad h\in H^1 \text{ and } \beta\in(0,\bar{\beta}).$$

If $z^0 \notin \mathcal{V}$, then, as is shown in Section 4, if $\operatorname{dist}(z^0, \mathcal{V})$ is sufficiently small there exist $\underline{\beta}$, $\bar{\beta}$, and κ such that

$$F^\beta_{qq}(q^\beta)(h,h) \ge \kappa|h|^2_{H^1} \quad \text{for all} \quad h\in H^1 \text{ and } \beta\in(\underline{\beta},\bar{\beta}).$$

We point out that, differently from (OLS), in the regularized case we require the unperturbed solution with parameter w^0 to be a *global* solution of $(\mathrm{ROLS})^\beta$. This is necessary since (2.7) and (2.8) are only shown to hold for global solutions of $(\mathrm{ROLS})^\beta$.

In Sections 3 and 4 we show that—besides some technical assumptions like sufficient regularity of $q\to u(q)$ and $q \to g(q)$, and a constraint qualification which guarantees the existence of Lagrange multipliers associated with the inequalities involved in defining Q_{ad}—a second-order sufficient optimality condition suffices to argue continuous dependence at local solutions of (OLS) and global solutions of $(\mathrm{ROLS})^\beta$ on w. The example of this section is reexamined in Section 5 by applying the general theory of Sections 3 and 4.

Finally, let us explain how the results by Chavent in [6] and [7] relate to the present example. The result in [6] (which allows for perturbations of z^0 but not of Q_{ad}) is not applicable since it requires an estimate of the form

$$|\eta|_{L^2} = |A^{-1}(q^0)(hu_x)_x| \ge \kappa|h|^2_1, \qquad \kappa>0,$$

which is not feasible. However, the regularization approach of [7] may be used, provided $\operatorname{dist}(z^0, \mathcal{V})$, as well as $\operatorname{diam} Q_{\mathrm{ad}}$, are sufficiently small.

3. Nonlinear Regularization Theory

Let us consider the problem

(P) minimize $\mathcal{J}(x)$ over $x \in Q_{ad}$,

where $Q_{ad} \subset Q$ with Q a reflexive Banach space and $\mathcal{J}$ is a mapping from Q_{ad} into $\mathbb{R}$. A typical example that we have in mind is $\mathcal{J}(x) = |\Phi x - z|_X^2$ where $\Phi: Q_{ad} \to X$, $z \in X$, and X is some Banach space. In this case (P) is the formulation of the equation $\Phi x = z$ as an optimization problem. If Φ does not have a continuous inverse then we frequently study a regularized form of (P), which is given by

$(\mathrm{RP})^\beta$ minimize $\mathcal{J}(x) + \beta\mathcal{N}(x)$ over $x \in Q_{ad}$,

where $\beta > 0$ and $\mathcal{N}(x)$ is an appropriately defined functional as, for example, $|x|_Q^2$ (see [10] and [19] if $\mathcal{J}$ is quadratic).

The purpose of this section is to summarize some of the basic properties of the solutions of $(\mathrm{RP})^\beta$ as $\beta \to 0$ in the case that $\mathcal{J}$ is not quadratic or, referring back to our example, if Φ is not linear. The nonquadratic case commonly occurs in parameter estimation problems, which are the subject matter of the subsequent sections.

We assume throughout that:

$$\text{(A1)} \quad \begin{cases} Q_{ad} \text{ is a closed and convex subset of the reflexive Banach space } Q, \\ \mathcal{J}: Q_{ad} \to \mathbb{R}^+ \text{ is weakly lower semicontinuous,} \\ \mathcal{N}: Q \to \mathbb{R}^+ \text{ is weakly lower semicontinouus with } \lim_{|x| \to \infty} \mathcal{N}(x) = \infty. \end{cases}$$

(A2) There exists a minimizer x^0 of (P).

Condition (A2) will not be needed before Lemma 3.2. The assumptions on Q_{ad} imply that it is weakly closed (or equivalently weakly sequentially closed) and this is precisely the property that is needed of Q_{ad}. If in addition Q_{ad} is also bounded, then (A1) implies (A2).

Lemma 3.1. *Let* (A1) *hold. Then there exists a solution* x^β *of* $(\mathrm{RP})^\beta$ *for each* $\beta > 0$.

Proof. Let $\delta = \inf\{\mathcal{J}(x) + \beta\mathcal{N}(x): x \in Q_{ad}\} \geq 0$ and let x_n be a minimizing sequence. Since $\mathcal{N}$ is radially unbounded, the sequence $\{x_n\}$ is bounded and, since Q is reflexive, it has a weak cluster point $x^\beta \in Q_{ad}$. By weak lower semicontinuity of $\mathcal{N}$ and $\mathcal{J}$ it follows that $\mathcal{J}(x^\beta) + \beta\mathcal{N}(x^\beta) = \inf\{\mathcal{J}(x) + \beta\mathcal{N}(x): x \in Q_{ad}\}$. □

The solutions x^β of $(\mathrm{RP})^\beta$ are not unique, in general, and we denote by $X^\beta = \{x^\beta: x^\beta \text{ is a solution of } (\mathrm{RP})^\beta\}$ the set of solutions of $(\mathrm{RP})^\beta$ and put $\mathcal{J}(X^\beta) = \{\mathcal{J}(x^\beta): x^\beta \in X^\beta\}$ and $\mathcal{N}(X^\beta) = \{\mathcal{N}(x^\beta): x^\beta \in X^\beta\}$ for any $\beta \geq 0$. The following properties are satisfied by $\mathcal{J}$ and $\mathcal{N}$.

Lemma 3.2. *Let* (A1) *hold. Then for all* $\beta > \beta_0 > 0$ *we have*:

(a) $\sup \mathcal{N}(X^{\beta}) \le \inf \mathcal{N}(X^{\beta_0})$,
(b) $\sup \mathcal{J}(X^{\beta_0}) \le \inf \mathcal{J}(X^{\beta})$.

If in addition (A2) *is satisfied then* (a) *and* (b) *hold for* $\beta > \beta_0 \ge 0$.

Proof. For any $x^{\beta_0} \in X^{\beta_0}$ and $x^{\beta} \in X^{\beta}$ we have

$$\mathcal{J}(x^{\beta_0}) + \beta_0 \mathcal{N}(x^{\beta_0}) \le \mathcal{J}(x^{\beta}) + \beta_0 \mathcal{N}(x^{\beta}). \tag{3.1}$$

Adding $(\beta - \beta_0)\mathcal{N}(x^{\beta})$ on both sides of (3.1) yields, by the definition of x^{β},

$$\begin{aligned} \mathcal{J}(x^{\beta_0}) + \beta \mathcal{N}(x^{\beta}) + \beta_0(\mathcal{N}(x^{\beta_0}) - \mathcal{N}(x^{\beta})) &\le \mathcal{J}(x^{\beta}) + \beta \mathcal{N}(x^{\beta}) \\ &\le \mathcal{J}(x^{\beta_0}) + \beta \mathcal{N}(x^{\beta_0}). \end{aligned} \tag{3.2}$$

Estimating the first by the last term in (3.2) we obtain

$$\beta_0(\mathcal{N}(x^{\beta_0}) - \mathcal{N}(x^{\beta})) \le \beta(\mathcal{N}(x^{\beta_0}) - \mathcal{N}(x^{\beta})).$$

Since $0 \le \beta_0 < \beta$ this inequality implies $\mathcal{N}(x^{\beta}) \le \mathcal{N}(x^{\beta_0})$ and therefore (a) holds. Using (3.2) together with (a) we find

$$\mathcal{J}(x^{\beta_0}) - \mathcal{J}(x^{\beta}) \le \beta_0(\mathcal{N}(x^{\beta}) - \mathcal{N}(x^{\beta_0})) \le 0.$$

Thus (b) is verified. □

Remark 3.1. For estimate (3.1) it is required that x^{β_0} is a (global) solution of $(\mathrm{RP})^{\beta_0}$. For this reason we need to restrict our attention to (global) solutions of $(\mathrm{RP})^{\beta}$ throughout this section.

Lemma 3.3. *Assume that* (A1) *and* (A2) *hold.*

(a) *Let* $\beta_n \to \beta_0 \ge 0$ *and let* $\{x^{\beta_n}\}$ *be any sequence of corresponding solutions of* $(\mathrm{RP})^{\beta_n}$. *Then* $\{x^{\beta_n}\}$ *has a weak cluster point and every weak cluster point of* $\{x^{\beta_n}\}$ *is a solution of* $(\mathrm{RP})^{\beta_0}$.
(b) *If moreover* $\beta_n \to \beta_0^+ \ge 0$, *then* $\lim_{n\to\infty} \mathcal{N}(x^{\beta_n})$ *exists and equals* $\min \mathcal{N}(X^{\beta_0})$.

Proof. By Lemma 3.2(a) and (A2) the set $\{\mathcal{N}(x^{\beta_n})\}_{n=1}^{\infty}$ is bounded by $\mathcal{N}(x^0)$. Hence $\{x^{\beta_n}\}_{n=1}^{\infty}$ is bounded in Q and there exists a subsequence, again denoted by $\{x^{\beta_n}\}_{n=1}^{\infty}$, converging weakly to some $\tilde{x} \in Q_{\mathrm{ad}}$. For all $x \in Q_{\mathrm{ad}}$ we have

$$J(x^{\beta_n}) + \beta_n \mathcal{N}(x^{\beta_n}) \le J(x) + \beta_n \mathcal{N}(x).$$

Weak lower semicontinuity of $\mathcal{J}$ and $\mathcal{N}$ implies

$$\mathcal{J}(\tilde{x}) + \beta_0(\tilde{x}) \le \mathcal{J}(x) + \beta_0 \mathcal{N}(x)$$

for all $x \in Q_{\mathrm{ad}}$. Thus $\tilde{x} \in X^{\beta_0}$ and (a) is proved. Now let $\beta_n \to \beta_0^+ \ge 0$ and let $x^{\beta_{n_k}}$ be any subsequence of x^{β_n} with $x^{\beta_{n_k}}$ converging weakly to some $\tilde{x} \in X^{\beta_0}$. If there were a solution x^{β_0} of $(\mathrm{RP})^{\beta_0}$ with $\mathcal{N}(x^{\beta_0}) < \mathcal{N}(\tilde{x})$ then by Lemma 3.2(a) and weak lower semicontinuity of $\mathcal{N}$ we have

$$\limsup_n \mathcal{N}(x^{\beta_n}) \le \mathcal{N}(x^{\beta_0}) < \mathcal{N}(\tilde{x}) \le \liminf_n \mathcal{N}(x^{\beta_n}),$$

which is impossible. Thus $\mathcal{N}(x^{\beta_0}) = \mathcal{N}(\tilde{x})$ and $\lim_n \mathcal{N}(x^{\beta_n})$ exists with $\lim_n \mathcal{N}(x^{\beta_n}) = \min\{\mathcal{N}(x^{\beta_0}) \colon x^{\beta_0} \in X^{\beta_0}\}$. □

Corollary 3.1. *Assume that* (A1) *and* (A2) *hold, that Q is uniformly convex and that $\mathcal{N}(x)=|x|_Q^p$ for some $p>0$. Let $\beta_n \to \beta_0^+$. Then every weak limit point x^{β_0} of x^{β_n} is a strong limit point and x^{β_0} is a minimum norm solution of* $(\mathrm{RP})^{\beta_0}$.

This follows from Lemma 3.3(b) and the fact that weak convergence together with convergence of the norms implies strong convergence in a uniformly convex Banach space.

In the next corollary we show that in a weak sense all minimum norm solutions of (P) are approximated by solutions of the regularized problems.

Corollary 3.2. *Let the assumptions of Corollary* 3.1 *hold and let C be an open connected component in the weak topology of $X_{\min}^0=\{x^0\colon x^0$ is a minimum norm solution of* (P)$\}$. *Then for every $\beta_n \to 0^+$ there exists a subsequence β_{n_k} of β_n such that $x^{\beta_{n_k}}$ converges to some element of C as $n_k \to \infty$.*

Proof. Observe that $X_{\min}^0$ is weakly compact. Hence C and $X_{\min}^0 \backslash C$ are weakly compact as well. Then $Q\backslash(X_{\min}^0\backslash C)$ is weakly open and contains C. Let $x \in C$. Recall that a basis of weak neighborhoods for x is given by sets of the form

$$V_x=\{y\colon |\langle f_i, y-x\rangle|<\varepsilon,\ i \in I\},$$

where $f_i \in X^*$ and I is a finite index set. Thus there exists a weak neighborhood of x of the form V_x which is contained in $Q\backslash(X_{\min}^0\backslash C)$. Hence also

$$W_x=\{y\colon |\langle f_i, y-x\rangle|<\varepsilon/2,\ i \in I\}$$

is a neighborhood of x and the weak closure of W_x is

$$\bar{W}_x=\{y\colon |\langle f_i, y-x\rangle|\le\varepsilon/2,\ i \in I\}.$$

Since $\bar{W}_x \subset V_x$ we find that for every $x \in C$ there exists a weak neighborhood W_x such that its weak closure $\bar{W}_x$ has void intersection with $X_{\min}^0\backslash C$. Since C is weakly compact the union of finitely many W_{x_i}, $i=1,\ldots,m$, covers C. Moreover, $W=\bigcup_{i=1}^m \bar{W}_{x_i}$ is weakly closed and $W \cap (X_{\min}^0\backslash C)=\varnothing$.

For $\beta>0$ we consider the problems $(\widetilde{\mathrm{RP}})^\beta$ given by

$$(\widetilde{\mathrm{RP}})^\beta \quad \text{minimize } \mathcal{J}(x)+\beta|x|^p \text{ over } Q_{\mathrm{ad}} \cap W.$$

Since $Q_{\mathrm{ad}} \cap W$ is weakly closed we can argue that there exists a solution x^β of $(\widetilde{\mathrm{RP}})^\beta$ for every $\beta>0$. Let $\{\beta_n\}$ be a sequence with $\lim \beta_n=0$, $\beta_n>0$. By Corollary 3.1 applied to $(\widetilde{\mathrm{RP}})^\beta$ there exists a subsequence $\{\beta_{n_k}\}$ of $\{\beta_n\}$ with $x^{\beta_{n_k}} \in Q_{\mathrm{ad}} \cap W$ converging strongly to a minimum norm solution x^0 of $(\widetilde{\mathrm{RP}})^0$. Since the norm of the minimum norm solution of $(\widetilde{\mathrm{RP}})^0$ equals the norm of the minimum norm solution of (P), it follows that $x^0 \in C$. Note that for n_k sufficiently large, $x^{\beta_{n_k}}$ is in the strong interior of W and thus $x^{\beta_{n_k}}$ is a solution of $(\mathrm{RP})^{\beta_{n_k}}$. □

Theorem 3.1. *Let* (A1) *and* (A2) *hold and let x^0 be any global solution of* (P). *Then*:

(a) $\lim_{\beta \to 0^+} \beta^{-1}(\sup \mathcal{J}(X^\beta)-\mathcal{J}(x^0))=0$, *and*
(b) $\sup \mathcal{J}(X^\beta)=o(\beta)$ *if* $\mathcal{J}(x^0)=0$.

(Observe that even if X^0 is not a singleton, $\mathscr{J}(X^0)$ is single valued.)

Proof. We proceed by contradiction and assume that there exist $\delta > 0$ and a sequence of solutions x^{β_n} of $(\mathrm{RP})^{\beta_n}$ with $\beta_n \to 0$ and

$$\beta_n^{-1}(\mathscr{J}(x^{\beta_n}) - \mathscr{J}(x^0)) \geq \delta > 0. \tag{3.3}$$

By Lemma 3.3 there exists a subsequence of $\{x^{\beta_n}\}$, again denoted by $\{x^{\beta_n}\}$ with $x^{\beta_n} \to \tilde{x}^0$ weakly in Q, with $\tilde{x}^0 \in X^0$, and $\lim_n \mathscr{N}(x^{\beta_n}) = \mathscr{N}(\tilde{x}^0)$. Since $J(\tilde{x}^0) = \mathscr{J}(x^0)$ we have by (3.3) that

$$\beta_n^{-1}(\mathscr{J}(x^{\beta_n}) - \mathscr{J}(\tilde{x}^0)) \geq \delta > 0. \tag{3.4}$$

From the second inequality in (3.2) it follows that

$$0 \leq \mathscr{J}(x^{\beta_n}) - \mathscr{J}(\tilde{x}^0) \leq \beta_n(\mathscr{N}(\tilde{x}^0) - \mathscr{N}(x^{\beta_n})). \tag{3.5}$$

But $\mathscr{N}(x^{\beta_n}) \to \mathscr{N}(\tilde{x}^0)$, therefore (3.5) contradicts (3.4) and thus (3.3) cannot hold for any $\delta > 0$. This proves (a). Part (b) is an obvious consequence of (a). □

4. Output Least Squares Stability by Regularization

As described in the introduction we study the problem of estimating an unknown coefficient in an elliptic equation from a measurement z^0 of the solution. The ouput least squares formulation of this problem is given by

$$\text{(OLS)} \quad \text{minimize } |u(q) - z|_H^2 \text{ over } q \in Q_{\text{ad}},$$

where

$$A(q)u(q) = f \quad \text{in } H$$

and

$$Q_{\text{ad}} = \{q \in Q\colon g(q) \in -K\}.$$

Here H is a (real) Hilbert space, $z^0, f \in H$, Q is a (real) Hilbert space, Y is a Banach lattice with ordering induced by the closed convex cone,

$$K = \{y \in Y\colon y \geq 0\},$$

and the specifications for g and $A(q)$ will be given shortly. To investigate the continuous dependence of solutions of (OLS) on z^0 and g, let $(\tilde{W}, \tilde{\delta})$ be a metric space, put

$$W = H \times \tilde{W}$$

endowed with a product metric and consider

$$\text{(OLS)}_w \quad \text{minimize } |u(q) - z|^2 \text{ over } q \in Q_{\text{ad}}^{\alpha},$$

where

$$A(q)u(q) = f,$$

$$Q_{\text{ad}}^{\alpha} = \{q \in Q\colon g(q, \alpha) \in -K\},$$

and

$$w = (z, \alpha).$$

Here w is the perturbation from the reference parameter $w^0 = (z^0, \alpha^0)$. The specification $w^0 = (z^0, \alpha^0)$ will be dropped at times and we write (OLS) = $(\text{OLS})_{w^0}$, $Q_{\text{ad}}^{\alpha^0} = Q_{\text{ad}}$, $g(q) = g(q, \alpha^0)$, etc. The family of regularized problems is given by

$(\text{ROLS})_w^\beta$ minimize $|u(q) - z|^2 + \beta |q|_Q^2$ over $q \in Q_{\text{ad}}^\alpha$
subject to
$A(q)u(q) = f.$

Again we put $(\text{ROLS})^\beta = (\text{ROLS})_{w^0}^\beta$. The following assumptions on g and $A(q)$ are made.

There exists a set $U \subset Q$ and a neighborhood $I(\alpha^0)$ of α^0 in $\tilde{W}$ such that $Q_{\text{ad}}^\alpha \subset \text{int}\, U$ for all $\alpha \in I(\alpha^0)$ and:

(H1) $g(\cdot, \alpha): Q \to Y$ is continuous and K-convex on Q, with first and second continuous Fréchet derivative, for every $\alpha \in I(\alpha^0)$ with the first derivative bounded on bounded subsets of Q. Furthermore, for every open set $U_q \times U_\alpha \subset Q \times \tilde{W}$ with $\alpha^0 \in U_\alpha$ there exists L satisfying

$$|g(q, \alpha) - g(q, \alpha^0)| \le L\delta(\alpha, \alpha^0) \qquad \text{for all} \quad (q, \alpha) \in U_q \times U_\alpha.$$

(H2) $A(q): D(A) \to H$ with $D(A) \subset H$ independent of $q \in U$ and $\overline{D(A)} = H$, $A(q)$ is a linear, closed operator for every $q \in U$.

(H3) The domain $D(A^*)$ of the adjoint operators $A(q)^*$ (in H) of $A(q)^*$ is independent of $q \in U$.

(H4) There exists a Banach space X continuously embedded in H satisfying

$$D(A) \subset X \quad \text{and} \quad D(A^*) \subset X$$

and nondecreasing positive functions c_1 and c_2 such that

$$|x|_X \le c_1(|q|)|A(q)x|_H \qquad \text{for all} \quad x \in D(A), \quad q \in U, \tag{4.1}$$

$$|x|_X \le c_2(|q|)|A(q)^*x|_H \qquad \text{for all} \quad x \in D(A^*), \quad q \in U. \tag{4.2}$$

(H5) For every $q \in Q$ the operator $A(q)$ has the form

$$A(q) = A_0 + A_1(q)$$

with A_0 and $A_1(q)$ linear operators in H satisfying that $A_1(q)$ is linear in $q \in Q$ and

$$|A_0 x|_H \le K_0 |x|_X \qquad \text{for all} \quad x \in D(A),$$

$$|A_1(q)x|_H \le K_1 |q|_Q |x|_X \qquad \text{for all} \quad x \in D(A) \text{ and } q \in Q, \tag{4.3}$$

$$|A_1(q)^*x|_H \le K_2 |q|_Q |x|_X \qquad \text{for all} \quad x \in D(A^*) \text{ and } q \in Q, \tag{4.4}$$

where K_i are independent of x and q.

(H6) $(A_1(q)^*x, y) = (x, A_1(q)y)$ for all $y \in D(A)$, $x \in D(A^*)$ and $q \in Q$.

(H7) $q \to A^{-1}(q)f$ from $U \subset Q$ to H is continuous from the weak to the weak topology.

Concerning (H5) and (H6) we recall that the adjoint of a densely defined closed operator is again densely defined and closed. Moreover, the adjoint of a densely defined operator is well defined, and thus $A_1(q)^*$ is meaningful [11, p. 167]. For elliptic problems we think of the space X as the space of functions in $D(A)$ (resp. $D(A^*)$) where the boundary conditions are omitted.

We point out some of the consequences of these assumptions. As a consequence of (H5) and the closedness of $A(q)$, the set $D(A)$ endowed with the X topology is a Banach space denoted by $D_X(A)$. Similarly, $D(A^*)$ endowed with the X topology is a Banach space which we denote by $D_X(A^*)$. Further, $A(q)$ and $A^*(q)$ are surjective by (H2) and (H4). In particular there exists a unique solution $u(q)\in D(A)$ of

$$A(q)u(q)=f \tag{4.5}$$

for every $q\in U$.

By Y^* we denote the topological dual space of Y and

$$K^{\oplus}=\{y^*\in Y^*: \langle y, y^*\rangle\geq 0 \text{ for all } y\in K\}$$

is the positive conjugate cone of K in Y^*. For every $y^*\in K^{\oplus}$ and $\alpha\in I(\alpha^0)$ the function $x\to\langle g(x,\alpha), y^*\rangle$ is convex, continuously F-differentiable and therefore also weakly lower semicontinuous [20, p. 82]. Therefore the set $\{x\in Q: \langle g(x,\alpha), y^*\rangle\leq 0\}$ is closed and convex in Q. Since

$$Q_{\text{ad}}^{\alpha}=\{x: g(x,\alpha)\leq 0\}=\bigcap_{y^*\in K^{\oplus}}\{x: \langle g(x,\alpha), y^*\rangle\leq 0\},$$

Q_{ad}^{α} is closed and convex (compare [22], pp. 353 and 404), and thus weakly sequentially closed. Moreover, (H1) implies that $g_{qq}(q,\alpha)(h,h)\in K$ for all q and h in Q.

Due to the weak sequential closedness of Q_{ad}^{α}, the weak continuity assumption (H7), and weak lower semicontinuity of the norm, the assumptions of Lemma 3.1 are satisfied and for every $\beta>0$ and $w=(z,\alpha)\in H\times I(\alpha^0)$ existence of a solution q_w^{β} of $(\text{ROLS})_w^{\beta}$ is guaranteed.

The following definition specifies the desired stability property of the solutions of $(\text{ROLS})_w^{\beta}$.

Definition 4.1. We call q in (OLS) output least square stable by regularization (ROLS stable) at $w^0=(z^0,\alpha^0)$ in Q_{ad} for β in the interval $I\subset(0,\infty)$, if for every global solution $q_{w^0}^{\beta}$ of $(\text{ROLS})_{w^0}^{\beta}$ with $\beta\in I$ there exist neighborhoods $V(w^0)$ of w^0 in W and $V(q_{w^0}^{\beta})$ of $q_{w^0}^{\beta}$ in Q and a continuous nondecreasing real-valued function ρ with $\rho(0)=0$ such that for all $w\in V(w^0)$ there exists a local solution $q_w^{\beta}\in V(q_{w^0}^{\beta})$ of $(\text{ROLS})_w^{\beta}$ and for every such local solution q_w^{β} we have $|q_w^{\beta}-q_{w^0}^{\beta}|_Q\leq\rho(|z-z^0|_H+\tilde{\delta}(\alpha,\alpha^0))$.

Remark 4.1. We first used the notion of ROLS stability in [8] to study continuous dependence on problem data for identification of the restoring force coefficient in a second-order boundary-value problem. There we also introduced "output-least-square-stability" (OLS stability) which is the continuous dependence of solutions on the data in $(\mathrm{OLS})_w$ (i.e., $\beta = 0$ here). Due to the nature of the inverse problems that are under consideration OLS stability cannot hold without strong additional assumptions. Some of these cases are discussed in Section 6.

Before we state the two main theorems of this section we require the following definitions:

Definition 4.2. The point $q \in Q_{\mathrm{ad}}$ is called *regular point* (with respect to the constraint Q_{ad}) if

$$0 \in \operatorname{int}\{g(q) + g'(q)Q + K\}.$$

Here and below all derivatives are taken in the sense of Fréchet.

Definition 4.3. The *attainable set* $\mathcal{V}$ is given by $\mathcal{V} = \{u(q)\colon q \in Q_{\mathrm{ad}}\}$. The problems $(\mathrm{OLS}) = (\mathrm{OLS})_{w^0}$ and $(\mathrm{ROLS})^\beta = (\mathrm{ROLS})^\beta_{w^0}$ are special cases of (P) and $(\mathrm{RP})^\beta$ with $\mathcal{J}(q) = |u(q) - z|^2$ and, assuming the existence of a solution q^0 of (OLS), the results of the previous section are applicable. In particular, the solutions q^β of $(\mathrm{ROLS})^\beta$ satisfy

$$|q^\beta| \le |q^0| \tag{4.6}$$

and

$$\lim_{\beta \to 0^+} \sup_{q^\beta \in Q^\beta} |q^\beta| = |q^0|,$$

where q^0 is a minimum norm solution of (OLS) and Q^β is the set of solutions of the unperturbed regularized problem $(\mathrm{ROLS})^\beta$.

Theorem 4.1. *Assume that there exists a solution of (OLS), that* (H1)-(H7) *hold, and that the points of Q_{ad} are regular. Let $\bar\beta > 0$ be chosen such that for a minimum norm solution q^0 of* (OLS)

$$|q^0|^2 - \sup_{q^{\bar\beta} \in Q^{\bar\beta}} |q^{\bar\beta}|^2 < \frac{1}{(K_2 \cdot c_2(|q^0|))^2} \tag{4.7}$$

and define

$$\underline{\beta} = \operatorname{dist}(z^0, \mathcal{V})^2 \left[\frac{1}{(K_2 c_2(|q^0|))^2} - |q^0|^2 + \sup_{q^{\bar\beta} \in Q^{\bar\beta}} |q^{\bar\beta}|^2 \right]^{-1} \ge 0.$$

If $\underline{\beta} < \bar\beta$ then the parameter q in (OLS) *is* ROLS *stable at $w^0 = (z^0, \alpha^0)$ in Q_{ad} for all $\beta \in (\underline{\beta}, \bar\beta)$. Moreover, if $z^0 \in \mathcal{V}$ then q is* ROLS *stable in Q_{ad} for all $\beta \in (0, \bar\beta)$.*

This result, which is proved later, gives conditions that guarantee the continuous dependence of solutions of $(\mathrm{ROLS})^\beta_{w^0}$ on the problem data w. In the first part there is no attainability assumption, but $\underline{\beta}$ may be larger than $\bar\beta$ in which case the theorem is not applicable. In this situation a more accurate measurement or an improved model, decreasing $\operatorname{dist}(z^0, \mathcal{V})$, should lead to success. Thus let $w_n^0 = (z_n^0, \alpha^0) \in W$ with $z_n^0 \to z_0^0$ in H and $z_0^0 \in \mathcal{V}$. We denote the solutions of $(\mathrm{OLS})_{w_n^0}$ and $(\mathrm{ROLS})^\beta_{w_n^0}$ by $q^0_{w_n^0}$ and $q^\beta_{w_n^0}$. The following stability property of solutions of $(\mathrm{ROLS})^\beta_{w_n^0}$ can be obtained for z_n^0 sufficiently close to $z_0^0 \in \mathcal{V}$. (Recall that stability is investigated with respect to the upper index in w_n^0.) By Q_w^0 we denote the set of solutions of $(\mathrm{OLS})_w$.

Theorem 4.2. *Let the assumptions of Theorem* 4.1 *hold and assume that* $q \to u(q)$ *from* $U \subset Q$ *to* H *is continuous from the weak to the strong topology. Choose* $z_n^0 \to z_0^0$ *in* H *with* $z_0^0 \in \mathcal{V}$, *and assume that solutions* $q_{w_n^0}^0$ *of* $(\mathrm{OLS})_{w_n^0}$ *with* $w_n^0 = (z_n^0, \alpha^0)$ *exist and that* $\sup_n \min_{q_{w_n^0}^0 \in Q_{w_n^0}^0} |q_{w_n^0}^0| < \infty$. *Then there exists* $\tilde{\beta} > 0$ *with the following property: for all* $\beta^* \in (0, \tilde{\beta})$ *there is* $N(\beta^*)$ *and a neighborhood* $I(\beta^*)$ *of* β^* *such that for all* $n \geq N(\beta^*)$ *the parameter* q *in* $(\mathrm{OLS})_{w_n^0}$ *is* ROLS *stable in* Q_{ad} *at* $w_n^0 = (z_n^0, \alpha 0)$ *for all* $\beta \in I(\beta^*)$.

The proofs of Theorems 4.1 and 4.2 are based on stability results for abstract nonlinear optimization problems; the essential requirement for such stability results is that higher-order sufficient optimality conditions hold. For the convenience of the reader and for frequent reference we state a result due to Alt [3]. A more detailed discussion is given in Section 3 of [8].

Let Q, Y, $K \subset Y$ be as above and let (W, δ) be a metric space. Further, let $f: D \times W \to \mathbb{R}^+$, $g: Q \times W \to Y$, where $D \subset Q$ is an open set satisfying

$$Q_{\mathrm{ad}} = \{q \in Q: g(q, w^0) \in -K\} \subset D.$$

Again w^0 is a fixed reference parameter which is dropped when no ambiguity can occur. Consider

$$(\mathrm{P})_w \quad \text{minimize } f(q, w) \text{ subject to } g(q, w) \in -K.$$

A functional $\lambda^* \in Y^*$ is called the Lagrange multiplier for $(\mathrm{P})_{w^0}$ at q^0 if

$$f_q(q^0) + \lambda^* g_q(q^0) = 0, \qquad \lambda^* \in K^+ \quad \text{and} \quad \lambda^* g(q^0) = 0, \tag{4.8}$$

and $F: D \to \mathbb{R}$ given by $F(q) = f(q) + \lambda^* g(q)$ is called the associated Lagrange functional. If q^0 is a solution of $(\mathrm{P})_{w^0}$ then under the assumptions of Proposition 4.1 below there exists a Lagrange multiplier associated with $(\mathrm{P})_{w^0}$ [15].

Proposition 4.1. *Let* q^0 *be a solution of* $(\mathrm{P})_{w^0}$ *which is a regular point of* Q_{ad} (*i.e.*, $\mathcal{O} \in \mathrm{int}\{g(q^0, w^0) + g_x(q^0, w^0)Q - K\}$) *and assume that* f *and* g *are twice continuously differentiable with respect to* q *at* (q^0, w^0) *and that there exist constants* $\nu > 0$ *and* $\gamma > 0$ *such that for a Lagrangian functional* F

$$F_{qq}(q^0)(h, h) \geq \gamma |h|^2 \tag{4.9}$$

for all $h \in g_q^{-1}(-K + \mathbb{R}g(q^0)) \cap \{h: \lambda^* g_q(q^0)h \leq \nu |h|\}$. *Moreover, assume that there exists a neighborhood* $U = U_q \times U_w$ *of* (q^0, w^0) *such that for constants* L_f *and* L_g

$$|f(q, w) - f(q', w^0)| \leq L_f(|q - q'| + \delta(w, w^0)), \tag{4.10}$$

$$|g(q, w) - g(q, w^0)| \leq L_g \delta(w, w^0) \tag{4.11}$$

for all $(q, w) \in U$, *and* $q' \in U_q$. *Then there exist* $r > 0$, $d > 0$, *and a neighborhood* V *of* w_0 *such that*:

(i) *The local extremal value function*

$$\mu_r(w) = \{\inf f(q, w): g(q, w) \in -K: |q - q^0| \leq r\}$$

is Lipschitz continuous at w^0.

For every $w \in V$ the following additional statements hold:

(ii) *For any sequence q_n with $g(q_n, w) \in -K$, $|q_n - q^0| \leq r$ and*

$$\lim_n f(q_n, w) = \mu_r(w) \quad \text{it follows that} \quad |q_n - q^0| < r$$

for all sufficiently large n.

(iii) *If there exists q_w with $g(q_w, w) \in -K$, $|q_w - q^0| \leq r$, and $f(q_w, w) = \mu_r(w)$, then $|q_w - q^0| < r$ and $|q_w - q^0| \leq d\delta(w, w^0)^{1/2}$.*

Lemma 4.1. *Let (H1)-(H7) hold. The first and second derivatives $\eta = u_q(q)h$ and $\xi = u_{qq}(q)(h, h)$ of the solution of (4.5) at $q \in Q_{\text{ad}}$ in direction $h \in Q$ are characterized by*

$$A(q)\eta = -A_1(h)u(q),$$

and

$$A(q)\xi = -2A_1(h)\eta.$$

Here $q \to u(q)$ is taken as a mapping from Q to $D_X(A)$.

Proof. We apply the implicit function theorem (see pp. 115 and 134 of [4]) to $F: Q \times D_X(A) \to H$ defined by $F(q, u) = A(q)u - f$. For $q^0 \in Q_{\text{ad}}$ we have $F(q^0, u(q^0)) = 0$. Moreover, F is continuous in a neighborhood of $(q^0, u(q^0))$ by (H5) and $F_u(q, u) = A(q) \in \mathscr{L}(D_x(A), H)$, $F_{uu} = 0$, $F_q(q, u) = A_1(\cdot)u \in \mathscr{L}(Q, H)$, $F_{uq}(q, u) = A_1(\cdot)(\cdot)$, and $F_{qq} = 0$. By (4.3) and linearity of A_1 in q we have that F_u and F_q depend continuously on (q, u) at $(q^0, u(q^0))$. Moreover, by (H2), (H4), and the closed graph theorem, $A(q^0)$ is a homeomorphism between $D_X(A)$ and H. Thus u is twice continuously differentiable at q^0. We obtain, from (4.5) by the implicit function theorem,

$$A(q^0)\eta + A_1(h)u(q^0) = 0$$

and

$$2A_1(h)\eta + A(q^0)\xi = 0.$$

But $A(q^0)^{-1} \in \mathscr{L}(H, D_X(A))$ and the lemma is verified. □

Proof of Theorem 4.1. We apply Proposition 4.1 to $(\text{ROLS})_w^\beta$ and first check the second-order sufficient condition. Let $q^\beta = q_{w^0}^\beta$ be any (global) solution of $(\text{ROLS})_{w^0}^\beta$. The associated Lagrange functional is given by

$$F(q) = F(q, w^0) = |u(q) - z^0|_H^2 + \lambda^* g(q, \alpha^0) + \beta |q|_Q^2,$$

where $\lambda^* \in Y^*$ is a Lagrange multiplier and $w^0 = (z^0, \alpha^0)$. For $h \in Q$ we have

$$F_{qq}(q^\beta)(h, h) = 2|\eta|^2 + 2(u(q^\beta) - z^0, \xi) + \langle g_{qq}(q^\beta, \alpha^0)(h, h), \lambda^* \rangle + 2\beta |h|^2, \tag{4.12}$$

with η and ξ defined in Lemma 4.1 (with $q = q^{\beta}_{w^0}$). Since $g_{qq}(q^{\beta}, \alpha^0)(h, h) \geq 0$ in Y we find

$$F_{qq}(q^{\beta})(h, h) \geq 2|\eta|^2 - 4(u(q^{\beta}) - z^0, A^{-1}(q^{\beta})A_1(h)\eta) + 2\beta|h|^2$$
$$= 2|\eta|^2 - 4(A_1(h)^* A^{-1}(q^{\beta})^*(u(q^{\beta}) - z^0), \eta) + 2\beta|h|^2$$
$$\geq 2\beta|h|^2 - 2|A_1(h)^* A^{-1}(q^{\beta})^*(u(q^{\beta}) - z^0)|^2. \tag{4.13}$$

But by (H5) and (H4)

$$|A_1(h)^* A^{-1}(q^{\beta})^*(u(q^{\beta}) - z^0)| \leq |A_1(h)^*|_{\mathscr{L}(D_X(A^*), H)} |A^{-1}(q^{\beta})^*(u(q^{\beta}) - z^0)|_X$$
$$\leq K_2|h| c_2(|q^0|)|u(q^{\beta}) - z^0|_H.$$

Using this estimate in (4.13) leads to

$$F_{qq}(q^{\beta})(h, h) \geq 2|h|^2(\beta - (K_2 c_2(|q^0|))^2 |u(q^{\beta}) - z^0|_H^2). \tag{4.14}$$

For every minimum norm solution q^0 of $(\mathrm{OLS})_{w^0}$ and for every solution q^{β} of $(\mathrm{ROLS})^{\beta}_{w^0}$ we have

$$-|u(q^{\beta}) - z^0|^2 \geq \beta(|q^{\beta}|^2 - |q^0|^2) - \mathrm{dist}(z^0, \mathscr{V})^2.$$

Therefore with $\kappa = (K_2 c_2(|q^0|))^2$

$$F_{qq}(q^{\beta})(h, h) \geq 2|h|^2 \kappa \left[\beta\left(\frac{1}{\kappa} - |q^0|^2 + |q^{\beta}|^2 \right) - \mathrm{dist}(z^0, \mathscr{V})^2 \right]$$
$$\geq 2|h|^2 \kappa \ \mathrm{dist}(z^0, \mathscr{V})^2 [\beta \cdot \underline{\beta}^{-1} - 1],$$

and $F_{qq}(q^{\beta})(h, h) \geq \mathrm{const}|h|^2$ for all $\beta \in (\underline{\beta}, \bar{\beta})$. This implies (4.9) of Proposition 4.1. Existence of local solutions q^{β}_w of the perturbed problems as required in Proposition 4.1(iii) follows by weak compactness of $Q_{\mathrm{ad}} \cap \{q: |q - q^0| \leq r\}$, $r > 0$, and weak lower semicontinuity of $q \to |u(q) - z|^2$. Condition (4.11) follows at once from (H1) and (4.10) is satisfied since $q \to u(q)$ from $U \subset Q$ to H is Fréchet differentiable and by (4.1), (4.3), and Lemma 4.1 the Fréchet derivative is uniformly bounded on bounded sets of its arguments. □

Proof of Theorem 4.2. As in (4.14) we obtain for $n = 0, 1, \ldots$

$$F_{qq}(q^{\beta}_n, w^0_n)(h, h) \geq 2|h|^2(\beta - \kappa \sup |u(q^{\beta}_n) - z^0_n|^2), \tag{4.15}$$

where we write q^{β}_n for $q^{\beta}_{w^0_n}$ and the supremum is taken over all solutions q^{β}_n of $(\mathrm{ROLS})^{\beta}_{w^0_n}$. Recall that $w^0_n = (z^0_n, \alpha^0)$. For β^* sufficiently small we need to bound $F_{qq}(q^{\beta}_n, w^0_n)$ from below uniformly in $\beta \in I(\beta^*)$ and $n \geq N(\beta^*)$. By Theorem 3.1 we can choose $\tilde{\beta}$ so that for every $\beta^* \in (0, \tilde{\beta})$ there exists $\varepsilon = \varepsilon(\beta^*)$ satisfying

$$\beta^* - \kappa \sup |u(q^{\beta^*}_0) - z^0_0|^2 \geq \varepsilon, \tag{4.16}$$

where the supremum is taken over all solutions $q^{\beta^*}_0$ of $(\mathrm{ROLS})^{\beta^*}_{w^0_0}$. First we show that for all $\beta^* \in (0, \tilde{\beta})$ there exists an $N(\beta^*)$ such that

$$\beta^* - \kappa \sup |u(q^{\beta^*}_n) - z^0_n|^2 \geq \varepsilon/2 \tag{4.17}$$

for all $n \geq N(\beta^*)$. If (4.17) were false then there would exist a sequence $\{n_k\}$ with $\lim n_k = \infty$ and solutions $q_{n_k}^{\beta^*}$ of $(\text{ROLS})_{w_{n_k}}^{\beta_0^*}$ such that

$$\beta^* - \kappa |u(q_{n_k}^{\beta^*}) - z_{n_k}^0|^2 < \varepsilon/2. \tag{4.18}$$

By Lemma 3.2(a) and the assumption on the boundedness of $|q_n^0|$ it follows that $\{q_{n_k}^{\beta^*}\}$ is bounded. Therefore there exists a weakly convergent subsequence of $q_{n_k}^{\beta^*}$, again denoted by $q_{n_k}^{\beta^*}$, with limit $q_0^{\beta^*}$. Since Q_{ad} is weakly sequentially closed, $q_0^\beta \in Q_{\text{ad}}$ and, moreover, q_0^β is a solution of $(\text{ROLS})_{w_0^0}^{\beta^*}$. Taking the limit in (4.18) we obtain

$$\beta^* - \kappa |u(q_0^{\beta^*}) - z_0^0|^2 \leq \varepsilon/2,$$

which contradicts (4.16) and hence (4.17) holds. Next we prove that there exists a closed neighborhood $I(\beta^*) \subset (0, \infty)$ of β^*, such that

$$\beta - \kappa \sup |u(q_n^\beta) - z_n^0|^2 \geq \varepsilon/4 \tag{4.19}$$

for all $n \geq N(\beta^*)$ and all $\beta \in I(\beta^*)$. If (4.19) were not true, then there would exist sequences $\{\beta_m\}$ and $\{n_m\}$ with $\beta_m \to \beta^*$ and $n_m \geq N(\beta^*)$, and solutions $q_{n_m}^{\beta_m}$ of $(\text{ROLS})_{w_{n_m}}^{\beta_0^m}$ such that

$$\beta_m - \kappa |u(q_n^{\beta_m}) - z_{n_m}^0|^2 < \varepsilon/4. \tag{4.20}$$

Concerning the index n_m we have to consider two cases: either infinitely many n_m assume the same value or $\{n_m\}$ is unbounded. First, without loss of generality, assume that $n_m = \tilde{n}$ for all m. Again $\{q_{\tilde{n}}^{\beta_m}\}_{m=1}^\infty$ is a bounded set in Q and by Lemma 3.3(a) there exists a subsequence of $\{q_{\tilde{n}}^{\beta_m}\}_{m=1}^\infty$ converging weakly to a solution $q_{\tilde{n}}^{\beta^*}$ of $(\text{ROLS})_{w_{\tilde{n}}^0}^{\beta^*}$. Taking the limit in (4.20) we obtain

$$\beta^* - \kappa |u(q_{\tilde{n}}^{\beta^*}) - z_{\tilde{n}}^0|^2 \leq \varepsilon/4,$$

where $\tilde{n} \geq N(\beta^*)$. This contradicts (4.17). On the other hand, if $\lim n_m = \infty$, then again $\{q_{n_m}^{\beta_m}\}$ contains a subsequence converging weakly to a solution $q_0^{\beta^*}$ of $(\text{ROLS})_{w_{\tilde{n}}^0}^{\beta^*}$. Taking the limit in (4.20) we arrive at

$$\beta^* - \kappa |u(q_0^{\beta^*}) - z_0^0|^2 \leq \varepsilon/4$$

which contradicts (4.16). Thus (4.19) is verified. Using (4.19) in (4.15) we obtain

$$F_{qq}(q_n^\beta, w_n)(h, h) \geq \frac{\varepsilon}{4} |h|^2 \qquad \text{for} \quad \beta \in I(\beta^*).$$

Proposition 4.1 now implies the result. □

The essential estimate in the proofs of Theorems 4.1 and 4.2 is the lower bound and the second derivative of the Lagrangian. This also implies the following

Corollary 4.1. *Under the assumptions of Theorem 4.1 all solutions q^β of $(\text{ROLS})_{w^0}^{\beta_0}$ with $\beta \in (\underline{\beta}, \bar{\beta})$ are isolated.*

Proof. Under the assumption of Theorem 4.1 we have shown that $F_{qq}(q^\beta)(h, h) \geq \text{const}|h|^2$ for every $\beta \in (\underline{\beta}, \bar{\beta})$. By Theorem 5.6 of [15] there exists $\delta > 0$ and a neighborhood $V(q^\beta)$ of q^β in Q such that

$$|u(q) - z^0|^2 + \beta|q|^2 \geq |u(q^\beta) - z^0|^2 + \beta|q^\beta|^2 + \delta|q - q^\beta|^2$$

for all $q \in V(q^\beta) \cap Q_{\text{ad}}$. This implies the assertion of the corollary. □

Similarly, under the assumptions of Theorem 4.2, all solutions $q^{\beta_0}_{w_n}$ of $(\text{ROLS})^{\beta_0}_{w_n}$ with $\beta \in I(\beta^*)$ and $n \geq N(\beta^*)$ are isolated.

Remark 4.2. Due to the method of our proof involving the second-order sufficient optimality condition we find that Theorems 4.1 and 4.2 guarantee that ρ can be taken as $\rho(x) = \sqrt{x}$ in the definition of ROLS stability.

5. Examples

In this section we apply the abstract framework of the previous section to several concrete examples. Two types of constraints will be used: a pointwise lower bound on the admissible coefficients guarantees the well posedness of the differential equation whereas a norm bound can be used to imply existence of solutions for the unregularized output least squares problem (OLS). We start with a result which implies that the elements of the constraint set satisfy the regular point condition for such a choice of constraints.

As before let Q be a real Hilbert space, and let $K \subset Q$ be a closed convex cone with vertex at zero inducing an ordering on Q such that $K = \{q \in Q: q \geq 0\}$. For $k \in K$ and $\gamma \in \mathbb{R}^+$ we define $g = (g_1, g_2): Q \to Q \times \mathbb{R}$ by

$$g(q) = (g_1(q), g_2(q)) = (k - q, |q|^2 - \gamma^2),$$

so that Y and K of the previous section are $Q \times \mathbb{R}$ and $K \times \mathbb{R}^+$ here.

Lemma 5.1. *Let $\gamma > |k|$. Then every point of the set $Q_{\text{ad}} = \{q \in Q: g(q) \leq 0\}$ is a regular point.*

Proof. Let $q \in Q_{\text{ad}}$. We need to verify that

$$0 \in \text{int}\{k - q - h + K, |q|^2 - \gamma^2 + 2(q, h) + \mathbb{R}^+: h \in Q\}. \tag{5.1}$$

By δ we denote a positive number which will be chosen sufficiently small later. Let $(\tilde{q}, \tilde{r}) \in Q \times \mathbb{R}$ with $|(\tilde{q}, \tilde{r})| < \delta$. Since K is a closed convex cone in a Hilbert space with vertex at zero, there exists a projection mapping q into $q^K \in K$ with $|q^K| \leq |q|$ for all $q \in Q$. Clearly,

$$\tilde{q} = k - q - (k - q - \tilde{q} + \tilde{q}^K) + \tilde{q}^K \tag{5.2}$$

which is of the same form as the first coordinate on the right-hand side of (5.1) with $h = k - q - \tilde{q} + \tilde{q}^K$. Concerning the second component of this set we compute

$$\begin{aligned} |q|^2 - \gamma^2 + 2(q, h) &= |q|^2 - \gamma^2 + 2(q, k) - 2|q|^2 + 2(q, \tilde{q}^K - \tilde{q}) \\ &\leq -\gamma^2 + |k|^2 + 2|q|(|\tilde{q}^K| + |\tilde{q}|) \\ &\leq |k|^2 - \gamma^2 + 4\delta|q|. \end{aligned} \tag{5.3}$$

From (5.3) it follows that for δ sufficiently small there exists $r' \in \mathbb{R}^+$ such that

$$\tilde{r} = |q|^2 - \gamma^2 + 2(q, h) + r'. \tag{5.4}$$

The result follows from (5.2) and (5.4). □

From the proof of this lemma we immediately obtain the following

Corollary 5.1. *Let $\gamma > |k|$. Then every point of the set $Q_{\mathrm{ad}} = \{q \in Q: g_1(q) \leq 0\}$ is a regular point. Similarly, every point of the set $Q_{\mathrm{ad}} = \{q \in Q: q_2(q) \leq 0\}$ is a regular point.*

Example 5.1. We reconsider the introductory example of Section 2 of estimating the diffusion coefficient q in the two-point boundary-value problem

$$\begin{aligned} &-(qu_x)_x + cu = f \quad \text{on } (0, 1), \\ &u(0) = u(1) = 0, \end{aligned} \tag{5.5}$$

when $f \in L^2$, $c \in L^2$ with $c \geq 0$. All function spaces of this example are considered over the interval $(0, 1)$. In the notation of Section 4 we take

$$H = L^2, \quad Q = H^1, \quad Y = H^1 \times \mathbb{R}, \quad K = H^1 \cap C^+ \times \mathbb{R}^+, \quad \tilde{W} = H^1 \times \mathbb{R},$$
$$W = L^2 \times \tilde{W} = L^2 \times H^1 \times \mathbb{R}, \quad \text{and} \quad X = H^2.$$

Here C^+ is the natural cone of nonnegative functions in C. A generic element in W is given by $w = (z, \alpha) = (z, k, \gamma)$ and the unperturbed reference parameter is $w^0 = (z^0, \alpha^0) = (z^0, k^0, \gamma^0)$ with $|k^0|_{H^1} < \gamma^0$ and $k_1 := \min_{x \in [0,1]} k^0(x) > 0$. The function $g: Q \times \tilde{W} \to Y$ is given by

$$g(q, \alpha) = (g_1(q, \alpha), g_2(q, \alpha)) = (k - q, |q|^2_{H^1} - \gamma^2)$$

and

$$Q^{\alpha}_{\mathrm{ad}} = \{q \in H^1: g(q, \alpha) \leq 0\} = \{q \in H^1: k(x) \leq q(x), |q|_{H^1} \leq \gamma\}.$$

The unperturbed output least squares problem is given by

$$\text{(OLS)} \quad \min |u(q) - z^0| \text{ over } Q^{\alpha^0}_{\mathrm{ad}}$$

with $u(q)$ satisfying (5.5). We now show that Theorems 4.1 and 4.2 are applicable. Let

$$U = \{q \in H^1: q(x) \geq k_1/2\}.$$

Then there exists a neighborhood $I(\alpha^0)$ of α^0 such that $Q_{\rm ad}^\alpha \subset \operatorname{int} U$ for all $\alpha \in I(\alpha^0)$. Condition (H1) can easily be checked. Next we define operators $A(q)$ for $q \in U$ by $D(A(q)) = D(A) = H_0^1 \cap H^2$ and

$$A(q)u = -(qu_x)_x + cu.$$

Clearly, $A(q)$ is densely defined, closed, and self-adjoint. An easy calculation shows that $k_1|u_x|_{L^2} \leq |A(q)u|_{L^2}$ for every $u \in D(A)$. Since $|u_x|$ is a norm on H_0^1, the operators $A(q)$ are bijections, cf. [5].

In particular, there exists a unique solution $u(q) \in D(A)$ for every $q \in U$. Moreover, it can be shown by elementary calculations that

$$|u(q)|_{H^2} \leq \text{const}\, |f|_{L^2},$$

where const can be chosen uniformly as q and c vary in bounded sets of H^1 and L^2, respectively. From these observations it follows that (H2)-(H4) and (H6) hold. For $q \in Q$ and $u \in D(A)$ we define

$$A_0 u = cu \quad \text{and} \quad A_1(q)u = -(qu_x)_x.$$

These operators satisfy (H5). Finally, let $\{q_n\}$ be a sequence converging weakly to q in $U \subset Q$. The corresponding sequence $\{u(q_n)\}$ of solutions of (5.5) is bounded in H^2, and therefore there exists a subsequence $\{u(q_{n_k})\}$ converging strongly in H^1 to an element $u \in H_0^1$. It is simple to show that $u = u(q)$ and therefore the sequence $\{u(q_n)\}$ itself converges to $u(q)$. This implies (H7) and the additional assumption of Theorem 4.2. Thus q in (OLS) is ROLS stable at w^0 in $Q_{\rm ad}$ for β in appropriately chosen intervals, that is, q depends Hölder continuously on L^2-perturbations of the observation z, on perturbations of the norm bound γ, and on H^1-perturbations of the lower bound k of the admissible parameter set $Q_{\rm ad}$.

Remark 5.1. The norm constraint $|q|_{H^1} \leq \gamma$ in the previous example implies the existence of a solution of the unperturbed and unregularized problem (OLS) but is not used otherwise. It could be replaced by $z^0 \in \mathscr{V}$, for example. The norm constraint is not needed for the applicability of Theorems 4.1 and 4.2, see Corollary 5.1.

Example 5.2. Here we consider the estimation of $q = \operatorname{col}(q_1, \ldots, q_n)$ with $n = 2$ or 3 in

$$\begin{aligned} -\sum_{i=1}^{n} (q_i u_{x_i})_{x_i} + cu &= f \quad \text{in } \Omega, \\ u|\Gamma &= 0, \end{aligned} \tag{5.6}$$

where $f \in L^2$ and $c \in L^2$ with $c \geq 0$. All function spaces are taken over the bounded domain Ω, which is assumed to have a smooth (C^2-) boundary Γ or to be a parallelepiped. In the notation of Section 4 we take

$$H = L^2, \quad Q = \bigotimes_{i=1}^{n} H^2, \quad Y = Q \times \mathbb{R}, \quad K = Q \cap \bigotimes_{i=1}^{n} C^{\oplus} \times \mathbb{R}^+, \quad \tilde{W} = Q \times \mathbb{R},$$

$$W = L^2 \times \tilde{W}, \quad \text{and} \quad X = H^2.$$

We endow Q with the Hilbert space product topology and denote by q_i the ith coordinate of the vector q. A generic element in W is given by $w=(z,\alpha)=(z,k,\gamma)$ and the unperturbed reference parameter is $w^0=(z^0,\alpha^0)=(z^0,k^0,\gamma^0)$ with $|k^0|_Q<\gamma^0$ and $k_1=\min_i \min_x k_i^0(x)>0$. Here we recall that $H^2\subset C$ is a continuous embedding for $n=2$ or 3. The function $g\colon Q\times\tilde{W}\to Y$ is given by

$$g(q,\alpha)=(g_1(q,\alpha),g_2(q,\alpha))=(\mathrm{col}(k_1-q_1,\ldots,k_n-q_n),|q|_Q^2-\gamma^2)$$

and

$$\begin{aligned}Q_{\mathrm{ad}}^{\alpha}&=\{q\in Q\colon g(q,\alpha)\le 0\}\\&=\{q_i\in H^2\colon k_i(x)\le q_i(x),|q|_Q\le\gamma,\ i=1,\ldots,n\}.\end{aligned}$$

The output least squares problem is given by

$$\text{(OLS)}\quad \min|u(q)-z^0|_{L^2}^2 \text{ over } Q_{\mathrm{ad}}^{\alpha^0},$$

with $u(q)$ satisfying (5.6). Again we show that Theorems 4.1 and 4.2 are applicable. Let

$$U=\{q\in Q\colon q_i(x)\ge k_1/2,\ i=1,\ldots,n\}.$$

Then there exists a neighborhood $I(\alpha^0)$ of α^0 such that $Q_{\mathrm{ad}}^{\alpha}\subset\operatorname{int}U$ for all $\alpha\in I(\alpha^0)$ and (H1) follows. Associated with the left-hand side in (5.6) we define the bilinear forms $l_q\colon H_0^1\times H_0^1\to\mathbb{R}$ by

$$l_q(u,v)=\sum_i (q_iu_{x_i},v_{x_i})+(cu,v).$$

Here and below we drop the index for the inner product and norm in L^2. For $q\in Q$ we have

$$l_q(u,u)\ge k_1\sum_i |u_{x_i}|^2,$$

which by Poincaré's inequality implies

$$l_q(u,u)\ge K|u|_{H_0^1}^2 \tag{5.7}$$

for a constant K depending only on k_1 and Ω [21, p. 120]. Thus by the Lax-Milgram lemma there exists a weak solution $u\in H_0^1$ of (5.6). By [14, pp. 180, 188] we further have that $u\in H_0^1\cap H^2$ and

$$|u|_{H^2}\le c_1(|u|+|f|)$$

and by (5.7)

$$|u|_{H^2}\le c_1|f|(1+K^{-1})=\tilde{c}_1|f|. \tag{5.8}$$

Here c_1 and therefore $\tilde{c}_1$ depend only on k_1, γ, and c. In fact, (5.8) holds with $q_i\in W^{1,p}$, with n arbitrary and $p>n$, but our general theory is restricted to the Hilbert space case. For $q\in U$ we now define operators $A(q)$ in L^2 by $D(A(q))=H_0^1\cap H^2$ and

$$A(q)u=-\sum_i (q_iu_{x_i})_{x_i}+cu.$$

These operators are densely defined, closed, and self-adjoint. From the above considerations it also follows that they are homeomorphisms between $D(A)$ endowed with the H^2-norm and L^2. Assumptions (H2)-(H4) now follow. We turn to (H5) and define for $q \in Q$

$$A_0 u = cu \qquad \text{for} \quad u \in L^2$$

and

$$A_1(q)u = -\sum_i (q_i u_{x_i})_{x_i} \qquad \text{for} \quad u \in D(A).$$

Using Green's formula we can verify (H6) with $A_1^*(q) = A_1(q)$ on $D(A)$. It is now simple to show (H5).

Finally, let $\{q_n\}$ be a sequence converging weakly to q in $U \subset Q$. With an argument analogous to the one in Example 5.1 it follows that $u(q_n) \to u(q)$ in H^1. Thus all the hypotheses of Theorems 4.1 and 4.2 hold.

Example 5.3. This is the problem of estimating the convection coefficient $q = \text{col}(q_1, \ldots, q_n)$ for $n = 2$ or 3 in

$$\begin{aligned} &-\sum_{i,j=1}^{n} (a_{ij} u_{x_i})_{x_j} + \sum_{i=1}^{n} q_i u_{x_i} + cu = f \quad \text{in } \Omega, \\ &u \,|\, \Gamma = 0, \end{aligned} \tag{5.9}$$

where $f \in L^2$, $c \in L^2$, $a_{i,j} \in W^{1,p}$, $p > n$, with $c \geq 0$, $a_{ji} = a_{ij}$, and

$$\nu \sum_i \xi_i^2 \leq \sum_{i,j} a_{ij}(x) \xi_i \xi_j$$

for some $\nu > 0$ and all $x \in \Omega$, $\xi \in \mathbb{R}^n$. The assumptions on Ω and Γ are as in Example 5.2. In this example we only use a norm constraint and take

$$H = L^2, \quad Q = \bigotimes_{i=1}^{n} H^1, \quad Y = \mathbb{R}, \quad K = \mathbb{R}^+, \quad \tilde{W} = \mathbb{R},$$

$$W = L^2 \times \mathbb{R}, \quad \text{and} \quad X = H^2.$$

We associate a bilinear form $l_q \colon H_0^1 \times H_0^1 \to \mathbb{R}$ with (5.9) and put

$$l_q(u, v) = \sum_{i,j} (a_{ij} u_{x_i}, v_{x_j}) + \sum_i (q_i u_{x_i}, v) + (cu, v).$$

Next $\alpha^0 \in \mathbb{R}$ is chosen in such a way that there exists $k_1 > 0$ satisfying

$$l_q(u, u) \geq k_1 |u|_{H^2}^2 \tag{5.10}$$

for all $u \in H_0^1$ and $q = (q_1, \ldots, q_n) \in Q_{\text{ad}}^{\alpha^0} = \{q \in Q \colon |q|_Q \leq \alpha^0\}$. We define $g \colon Q \times \mathbb{R} \to \mathbb{R}$ by $g(q, \alpha) = |q|_Q^2 - \alpha^2$ and

$$Q_{\text{ad}}^{\alpha} = \{q \in Q \colon g(q, \alpha) \leq 0\}.$$

The output least squares problem is given as in the previous example except that $u(q)$ must satisfy (5.9).

To show that (H1)-(H7) and the additional continuity requirement of Theorem 4.2 hold let us observe first that H^1 embeds continuously into L^p for $p \geq 2$ if $n=2$ and for $6 \geq p \geq 2$ if $n=3$. We define

$$U=\left\{q \in Q\colon l_q(u, u) \geq \frac{k_1}{2}|u|_{H^1} \text{ for all } u \in H_0^1\right\}.$$

Since $|(qu_x, u)| \leq \kappa |u|_{H^1}^2 |q|_{H^1}$, where κ is an embedding constant, it is simple to argue that there exists a constant $\varepsilon > 0$ such that $\{q \in Q\colon |q|_Q \leq \alpha^0 + \varepsilon\} \subset U$. Consequently, there exists a neighborhood $I(\alpha^0)$ of α^0 such that $Q_{\text{ad}}^\alpha \subset \operatorname{int} U$ for $\alpha \in I(\alpha^0)$. It is simple to show that for every $q \in U$ there exists a constant K_q such that $l_q(u, v) \leq K_q |u|_{H^1} |v|_{H^1}$ for all $u, v \in H_0^1$. Together with (5.10) this implies the existence of a weak solution $u \in H_0^1$ of (5.9). As in Example 5.2 the results in [14, p. 180, 188] imply that $u \in H_0^1 \cap H^2$ and that

$$|u|_{H^2} \leq c_1 |f|. \tag{5.11}$$

Here c_1 depends only on k_1, a_{ij}, c, and μ where $\sum_i |q_i|_{H^1} \leq \mu$. For $q \in U$ we define the operators $A(q)$ in L^2 by $D(A) = D(A(q)) = H_0^1 \cap H^2$ and

$$A(q)u = -\sum_{i,j} (a_{ij} u_{x_i})_{x_j} + \sum q_i y_{x_i} + cu.$$

For every $q \in U$ the operator $A(q)$ is densely defined, closed, and a homeomorphism between $D(A)$ endowed with the H^2-norm and L^2. The adjoint of $A(q)$ for $q \in U$ is characterized by $D(A^*(q)) = H_0^1 \cap H^2$ and

$$A^*(q)u = -\sum_{i,j} (a_{ij} u_{x_i})_{x_j} + \sum (q_i u)_{x_i} + cu.$$

Again, $A^*(q)$, $q \in Q$, is densely defined, closed, and a homeomorphism between $D(A^*(q))$ endowed with the H^2-norm and L^2. Moreover, there exists a constant c_2, with the same properties as c_1 above, such that

$$|u|_{H^2} \leq c_2 |A^*(q)u| \tag{5.12}$$

for all $u \in H_0^1 \cap H^2$. Thus (H2)-(H4) holds. Next, for every $q \in Q$ we introduce the densely defined operators A_0 and $A_1(q)$ by

$$A_0 u = -\sum_{i,j} (a_{ij} u_{x_i})_{x_j} + cu$$

and

$$A_1(q)u = \sum_i q_i u_{x_i} \qquad \text{for} \quad u \in D(A).$$

Clearly,

$$A_1^*(q)u = \sum_i (q_i u)_{x_i} \qquad \text{for} \quad u \in D(A),$$

so that in particular (H6) holds. Clearly, there exist constants K_0 and K_1 such that $|A_0 u| \leq K_0 |u|_{H^2}$ and $A_1(q)u \leq K_1 |q|_Q |u|_{H^2}$ for all $u \in D(A)$. Moreover, for all $u \in D(A)$ the following estimate holds:

$$|A_1^*(q)u| \leq \sum_i (|q_i|_{L^4} |u_{x_i}|_{L^4} + |(q_i)_x| |u|_{L^\infty}) \leq K_2 |q|_Q |u|_{H^2},$$

where K_2 is an embedding constant. Finally, using (5.11) it is simple to show that $q \to u(q)$ is continuous from U endowed with the weak topology of Q to H^1. This ends our consideration for this example.

Remark 5.2. In all our examples we assumed the availability of observations $z \in L^2(\Omega)$. If only point data are available, our results can be utilized in the following way. Let $Iz \in L^2(\Omega)$ denote an appropriate interpolation of pointwise data $z = \{z_i\}_{i=1}^M$ taken at $x_i \in \Omega$, $i = 1, \ldots, M$. If the interpolation operator is Lipschitz continuous from $\mathbb{R}^M$ to $L^2(\Omega)$, then again we can argue that (4.10) of Proposition 4.1 holds and stability with respect to the point data can be guaranteed. In a similar context this is carried out in more detail in the following section.

6. Output Least Squares Stability

In this section we discuss some conditions which guarantee stability of the solutions of the output least squares problem without the benefit of regularization terms or with only local regularization. We recall that the problems that we investigate are given by

$$(\mathrm{OLS})_w \quad \text{minimize } |u(q) - z|^2 \text{ over } q \in Q_{\mathrm{ad}}^{\alpha}$$

where

$$A(q)u(q) = f, \qquad Q_{\mathrm{ad}}^{\alpha} = \{q \in Q \colon g(q, \alpha) \in -K\}, \quad \text{and} \quad w = (z, \alpha).$$

As before w is the perturbed parameter with reference parameter $w^0 = (z^0, \alpha^0)$, and $(\mathrm{OLS}) = (\mathrm{OLS})_{w^0}$, $Q_{\mathrm{ad}} = Q_{\mathrm{ad}}^{\alpha^0}$. We shall assume throughout the existence of a solution q_{w^0} of (OLS).

Definition 6.1. The parameter q in $(\mathrm{OLS})_{w^0}$ is called output least squares stable (OLS stable) at $w^0 = (z^0, \alpha^0)$ in Q_{ad} at the local solution q_{w^0} of $(\mathrm{OLS})_{w^0}$, if there exist neighborhoods $V(w^0)$ of w^0 in W and $V(q_{w^0})$ of q_{w^0} in Q and a continuous real-valued function ρ with $\rho(0) = 0$ such that for all $w \in V(w^0)$ there exists a local solution $q_w \in V(q_{w^0})$ of $(\mathrm{OLS})_w$ and for every such local solution q_w we have

$$|q_w - q_{w^0}|_Q \le \rho(|z - z^0|_H + \tilde{\delta}(\alpha, \alpha^0)).$$

We point out that OLS stability is a concept which applies at specific local solutions, whereas ROLS stability requires stability of *all* global solutions of the regularized problem.

6.1. *Finite-Dimensional Parameter Space*

In this subsection the case of a finite-dimensional parameter space Q is considered. We write $q^0 = q_{w^0}$.

Theorem 6.1. *Assume that* (H1)–(H7) *holds with Q a finite-dimensional normed linear space and let q^0 be a local solution of* $(\mathrm{OLS})_{w^0}$. *If q^0 is a regular point, if $h \to |A_1(h)u(q^0)|_H$ defines a norm on Q, and if $|z^0 - u(q^0)|$ is sufficiently small, then q is* OLS *stable at w^0 in Q_{ad} at q^0.*

Proof. The Lagrange functional associated with $(\text{OLS})_{w^0}$ is given by $F(q) = F(q, w^0) = |u(q) - z^0|_H^2 + \lambda^* g(q, \alpha^0)$, where $\lambda^* \in Y^*$ is a Lagrange multiplier and $w^0 = (z^0, \alpha^0)$. For all $h \in Q$ we have

$$F_{qq}(q^0)(h, h) = 2|\eta|^2 + 2(u(q^0) - z^0, \xi) + \lambda^* g_{qq}(q^0, \alpha^0)(h, h),$$

where η and ξ are characterized by $A(q^0)\eta = -A_1(h)u(q^0)$ and $A(q^0)\xi = -2A_1(h)\eta$. As in the proof of Theorem 4.1 we have $\lambda^* g_{qq}(q^0, \alpha^0)(h, h) \geq 0$ and therefore

$$\begin{aligned} F_{qq}(q^0)(h, h) &\geq 2|\eta|^2 + 2(u(q^0) - z^0, \xi) \\ &\geq 2|\eta|^2 - 4|A_1(h)^*(A^{-1}(q^0))^*(u(q^0) - z^0)||\eta| \\ &\geq |\eta|^2 - 4|A_1(h)^*(A^{-1}(q^0))^*(u(q^0) - z^0)|^2 \\ &\geq |\eta|^2 - |h|_Q^2[2K_2c_2(|q^0|)|u(q^0) - z^0|_H]^2. \end{aligned}$$

Here we use (H4) and (H5) and proceed similarly as in (4.13). To estimate $|\eta|$ from below let H^- denote the dual space (with respect to H as a pivot space) of $D(A^*)$ endowed with the graph norm of $A(q^0)$, i.e.,

$$|w|_{H^-} = \sup_{v \in D(A^*(q^0))} \frac{(w, v)_H}{|A^*(q^0)v|_H}.$$

We easily check that $|w|_{H^-} \leq |A(q^0)^{-1}w|_H$ for all $w \in H$. Therefore

$$|\eta|_H^2 \geq |A_1(h)u(q^0)|_{H^-}^2.$$

Next we observe that $h \to |A_1(h)u(q^0)|_{H^-}$ is a norm on Q. Homogeneity and the triangle inequality follow from the linearity of $h \to A_1(h)$. Moreover, if $|A_1(h)u(q^0)|_{H^-} = 0$ then $A_1(h)u(q^0) = 0$ in H; but by the assumption that $h \to |A_1(h)u(q^0)|_H$ is a norm on Q we conclude that $h = 0$. Thus $h \to |A_1(h)u(q^0)|_H$ as well as $h \to |A_1(h)u(q^0)|_{H^-}$ are norms on Q, and since Q is finite dimensional these norms are equivalent to the norm on Q. Therefore we have, for some constant κ,

$$F_{qq}(q^0)(h, h) \geq \kappa |h|_Q^2 - |h|_Q^2[2K_2c_2(|q^0|)|u(q^0) - z^0|_H]^2.$$

For $|u(q^0) - z^0|_H$ sufficiently small this implies (4.9) of Proposition 4.1. Conditions (4.10) and (4.11) can easily be checked. The result now follows from Proposition 4.1.

Example 6.1. As in Example 5.1 we consider the estimation of the coefficient q in the two-point boundary-value problem

$$\begin{aligned} &-(qu_x)_x + cu = f \quad \text{on } (0, 1), \\ &u(0) = u(1) = 0, \end{aligned} \tag{6.1}$$

where $f \in L^2$ and $c \in L^2$, with $c \geq 0$. In the notation of Section 4 we take

$$H = L^2, \quad Q \subset H^1 \text{ a finite-dimensional linear space,}$$
$$Y = H^1 \times \mathbb{R}, \quad K = C^+ \times \mathbb{R}^+, \quad \tilde{W} = H^1 \times \mathbb{R},$$
$$W = L^2 \times H^1 \times \mathbb{R}, \quad \text{and} \quad X = H^2.$$

The definitions of w^0, g, and (OLS) are as in Example 5.1, only now we have to assume in addition that there exists at least one element $q \in Q$ such that $q(x) \geq k^0(x)$ on [0, 1], and $|q|_{H^1} \leq \gamma^0$, to guarantee that $Q_{ad}^{\alpha^0}$ is not empty. The existence of at least one solution q^0 of (OLS) is guaranteed since $Q_{ad}^{\alpha^0}$ is bounded, closed, and convex. Moreover, by Lemma 5.1 every point of $Q_{ad}^{\alpha^0}$ satisfies the regular point condition. We now make the following assumption:

$$q^0 \text{ is a local solution of (OLS) and } u(q^0)_x \text{ is different from zero almost everywhere on } [0, 1]. \tag{6.2}$$

Using Theorem 6.1 we now show that q is OLS stable at w^0 in Q_{ad} at the specific local solution q^0 for which (6.2) holds. It suffices to argue that $h \to |(hu_x)_x|_{L^2}$, with $u = u(q^0)$, defines a norm on Q. Homogeneity and the triangle inequality hold. Since $u(0) = u(1) = 0$ there exists at least one $\xi \in (0, 1)$ such that $u_x(\xi) = 0$. Therefore $|(hu_x)_x|_{L^2} \geq |hu_x|_{L^2}$ for all $h \in Q$. Thus, if $|(hu_x)_x| = 0$ then $|hu_x|_{L^2} = 0$ and by assumption $h = 0$ on [0, 1]. We have shown OLS stability of q at q^0. We point out that $u_x \neq 0$ almost everywhere on [0, 1] is a sufficient condition: but in specific cases $u_x = 0$ on sets of positive measure may still allow us to argue that $h \to |(hu_x)_x|$ is a norm on Q. For example, if Q consists of linear spline functions and meas$\{x: u_x(x) = 0\}$ is strictly smaller than the measure of the support of any of the basis elements, then this assumption holds.

Remark 6.1. Further examples can easily be obtained. For instance, q_1 in (5.9) is OLS stable in a local solution q^0 for $Q \subset H^1$ finite dimensional if $u_x(q^0) \neq 0$ almost everywhere in Ω.

6.2. *OLS Stability over Subsets of the Domain*

In this subsection we discuss another case of stability of the solutions of the OLS problem without requiring regularization terms. To do so we need to change somewhat the general framework that we used so far. First, we now only attempt to estimate the unknown parameters away from the "singularities" of the solution u. To give an indication of what we mean by singularity let us consider Example 5.1 and let q^0 be a solution of (OLS). Then $\mathscr{S} = \{x: u_x(q^0) = 0\}$ is the set of singularities for the diffusion coefficient q^0. From (5.5) it is clear that the determination of $q^0(x)$ for $x \in \mathscr{S}$ using an OLS approach is difficult or even impossible (if x is in the interior of $\mathscr{S}$). Secondly, we use a fit-to-data criterion involving the X-norm rather than the H-norm. This requires the availability of measurements z in the X-norm and, in practice, it may result in approximate differentiation of pointwise "real data." We therefore proceed as follows. Let Z be the normed linear space from which we take "real data" z and let $\mathscr{I}: Z \to X$ be an operator satisfying

$$|\mathscr{I}z_1 - \mathscr{I}z_2|_X \leq M_1|z_1 - z_2|_Z \tag{6.3}$$

for a constant M_1 independent of $z_1, z_2 \in Z$. In applications Z may be $\mathbb{R}^M$ with $z = \{z_i\}_{i=1}^M$ point data and $\mathscr{I}$ an interpolation operator. We show stability with respect to the constraints as well as to z.

In addition to the spaces H, Q, and Y, the cone K, and the operators $A(q)$ of Section 4 we require a real Hilbert space $\tilde{Q}$ and an affine operator $E\colon Q \to \tilde{Q}$ which has $\mathscr{E}'$ as a Fréchet derivative independent of where the derivative is taken. The operators $A(q)$ are now of the form

$$A(q) = A_0 + A_1(\mathscr{E}(q)).$$

We require (H1)-(H7) to hold with Q replaced by $\tilde{Q}$. The parameter dependent output least squares problem is

$$(\text{OLS})_w \quad \min |u(q) - \mathscr{I}z|_X^2 \text{ over } q \in Q_{\text{ad}}^{\alpha},$$

where

$$A(q)u(q) = f$$

and

$$Q_{\text{ad}}^{\alpha} = \{q \in Q\colon g(q, \alpha) \in -K\}, \qquad w = (z, \alpha).$$

Just as in Lemma 4.1 we can show that $q \to u(q)$ is twice continuously differentiable from Q to X with $\eta = u_q(q)(h)$ and $\xi = u_{qq}(q)(h, h)$ given by

$$A(q)\eta + A_1(\mathscr{E}'h)u(q) = 0 \tag{6.4}$$

and

$$A(q)\xi + 2A_1(\mathscr{E}'h)\eta = 0. \tag{6.5}$$

Theorem 6.2. *Assume that* (H1)-(H7) *hold with Q replaced by $\tilde{Q}$, that $\mathscr{E}$ and $\mathscr{I}$ are as defined above, and that q^0 is a local solution of* $(\text{OLS})_{w^0}$. *If q^0 is a regular point, if $|\mathscr{I}z^0 - u(q^0)|_X$ is sufficiently small, and if there exists a constant k such that*

$$|A_1(\mathscr{E}'h)u(q^0)|_H \geq k|h|_Q \qquad \text{for all} \quad h \in Q, \tag{6.6}$$

then q is OLS *stable at w^0 in Q_{ad} at q^0.*

Proof. Again we use Proposition 4.1. The Lagrangian is now given by

$$F(q) = F(q, w^0) = |u(q) - \mathscr{I}z^0|_X^2 + \lambda^* g(q, \alpha^0).$$

Using (6.4) and (6.5) the second derivative of F at q_{w^0} is found to be

$$F_{qq}(q^0)(h, h) = 2|\eta|_X^2 + 2(u(q^0) - \mathscr{I}z^0, \xi)_X + \lambda^* g_{qq}(q^0, \alpha^0)(h, h).$$

We therefore have, putting $c_1 = c_1(|q^0|)$, $c_3 = K_0 + K_1\|\mathscr{E}'\|\,|q_{w^0}|$,

$$\begin{aligned}
F_{qq}(q_{w^0})(h, h) &\geq 2|A(q^0)^{-1}A_1(\mathscr{E}'h)u(q^0)|_X^2 \\
&\qquad -4|u(q^0) - \mathscr{I}z^0|_X|A(q^0)^{-1}A_1(\mathscr{E}'h)\eta|_X \\
&\geq 2c_3^{-2}|A_1(\mathscr{E}'h)u(q^0)|_H^2 - 4c_1|u(q^0) - \mathscr{I}z^0|_X|A_1(\mathscr{E}'h)\eta|_H \\
&\geq 2c_3^{-2}k^2|h|_Q^2 \\
&\qquad -4c_1K_1\|\mathscr{E}'\|\,|h|\,|A^{-1}(q^0)A_1(\mathscr{E}'h)u(q^0)|_X|u(q^0) - \mathscr{I}z^0|_X \\
&\geq 2c_3^{-2}k^2|h|_Q^2 - 4c_1^2K_2^2\|\mathscr{E}'\|^2|h|_Q^2|u(q^0)|_X|u(q^0) - \mathscr{I}z^0|_X \\
&= 2|h|_Q^2[c_3^{-2}k^2 - 2c_1^2K_1^2\|\mathscr{E}'\|\,|u(q^0)|_X|u(q^0) - \mathscr{I}z^0|_X].
\end{aligned}$$

For $|u(q^0)-\mathscr{I}z^0|_X$ sufficiently small this implies $F_{qq}(q^0)(h,h)\geq\gamma|h|_Q^2$ for an appropriately defined γ independent of h. This is (4.9) of Proposition 4.1. As before, (4.11) follows from (H1) and (4.10) is a consequence of

$$\begin{aligned}
&||u(q_1)-\mathscr{I}z_1|_X-|u(q_2)-\mathscr{I}z_2|_X|\\
&\quad\leq|u(q_1)-u(q_2)|_X+|\mathscr{I}z_1-\mathscr{I}z_2|_X\\
&\quad\leq\sup_{\tau\in[0,1]}|u_q(q_1+\tau(q_2-q_1))|_{\mathscr{L}(Q,X)}|q_1-q_2|+M_1|z_1-z_2|\\
&\quad\leq c_1(r)^2K_2\|\mathscr{E}\|\,|f|\,|q_1-q_2|+M_1|z_1-z_2|,
\end{aligned}$$

where we used (6.3) and put $r=\max(|q_1|,|q_2|)$. This ends the proof. □

Example 6.2. Here we consider the estimation of q over the measurable set $\Omega\subset(0,1)$ in

$$\begin{aligned}
&-(au_x)_x+qu=f\quad\text{on }(0,1),\\
&u(0)=u(1)=0,
\end{aligned}$$

where q is known on $[0,1]\backslash\Omega$. We take discrete measurements $z=\{z_i\}_{i=0}^M\in\mathbb{R}^{m+1}$ at i/M, $i=0,\dots,M$, and assume that $a\in H^1$, $f\in L^2$ with $a(x)>0$. The general theory will be used by making the following choice for the spaces:

$$\begin{aligned}
&H=L^2(0,1),\quad Q=L^2(\Omega),\quad \tilde{Q}=L^2(0,1),\quad Y=L^2(\Omega)\times\mathbb{R},\\
&K=L_+^2(\Omega)\times\mathbb{R}^+,\quad \tilde{W}=L^2(\Omega)\times\mathbb{R},\quad W=\mathbb{R}^M\times L^2(\Omega)\times\mathbb{R},\\
&X=H^2(0,1),\quad\text{and}\quad Z=\mathbb{R}^{M+1}.
\end{aligned}$$

Here $L_+^2(\Omega)$ denotes the cone of almost everywhere nonnegative functions. It is assumed that q is known on $[0,1]\backslash\Omega$ and has the value $c\in L^2((0,1)\backslash\Omega)$ there with $c\geq 0$. For $q\in Q=L^2(\Omega)$ we define

$$\mathscr{E}(q)=\begin{cases}c&\text{on }[0,1]\backslash\Omega,\\ q&\text{on }\Omega.\end{cases}$$

We easily find

$$\mathscr{E}'(h)=\begin{cases}0&\text{on }[0,1]\backslash\Omega,\\ h&\text{on }\Omega.\end{cases}$$

To specify the constraints we choose $k^0\in L^2(\Omega)$, $k^0\geq 0$, $\gamma^0>0$, and define $g\colon Q\times\tilde{W}\to Y$ by $g(q,\alpha)=(k-q,|q|_{H^1}^2-\gamma^2)$, where $\alpha=(k,\gamma)\in\tilde{W}$. Moreover, we put $Q_{\text{ad}}^\alpha=\{q\in Q\colon g(q,\alpha)\leq 0\}$. We further define $\mathscr{I}\colon\mathbb{R}^{M+1}\to X$ as the cubic spline interpolation operator which maps $z=\{z_i\}$ into the unique element $\mathscr{I}z\in C^2$ satisfying $(\mathscr{I}z)(i/M)=z_i$, $(\mathscr{I}z)'(0)=(\mathscr{I}z)'(1)=0$ and which is a cubic polynomial on all subintervals $(i/M,(i+1)/M)$. It is then straightforward to see that (6.3) holds [18, pp. 46, 47]. We have now fully specified the OLS problem. It is simple to see that there exists at least one solution of $(\text{OLS})_{w^0}$. Let q^0 be any local solution of $(\text{OLS})_{w^0}$ and assume that for some $\mu>0$

$$|u(q^0)(x)|\geq\mu>0\qquad\text{for all}\quad x\in\Omega.\tag{6.7}$$

By Lemma 5.1 we know that q_{w^0} is a regular point. Moreover, (6.7) implies that

$$|A_1(\mathscr{E}'h)u(q^0)|_{L^2(0,1)} = |hu(q^0)|_{L^2(\Omega)} \geq \mu|h|_{L^2(\Omega)}, \tag{6.8}$$

which is assumption (6.6) of Theorem 6.2. Conditions (H1)-(H7) can easily be checked for the present example. Thus all assumptions of Theorem 6.2 are satisfied and OLS stability of q at q^0 is obtained if (6.7) holds and $|\mathscr{I}z^0 - u(q^0)|_{H^2(0,1)}$ is small.

Remark 6.2. The above example can be generalized with appropriate modifications to the multidimensional case. Theorem 6.2 can also be used to study OLS stability of the parameter b in $-(au_x)_x + bu_x + cu = f$, where a natural space for the coefficients is again L^2. However, further generalizations of the theory are required to treat the multidimensional analog or to study OLS stability of the diffusion coefficient a in this equation.

6.3. An Example for Local Regularization

We have seen that the essential requirement in obtaining OLS or ROLS stability is positive definiteness of the second derivative of the Lagrangian. In Sections 4 and 5 this was achieved by means of a regularization term, whereas in the first two subsections of this section positivity was guaranteed by means of specific assumptions on the OLS problem. These two ideas can be combined by only regularizing over certain subsets of the domain of the unknown parameter. We give an example.

Example 6.3. We return to Example 6.2 which is the identification of q in

$$-(au_x)_x + qu = f \quad \text{on } (0,1),$$
$$u(0) = u(1) = 0,$$

with $f \in L^2(0,1)$, $a \in H^1(0,1)$, with $a > 0$. The cost functional involves cubic spline interpolation of point data $z \in \mathbb{R}^{M+1}$ and an H^2-criterion just as in Example 6.2. The aim is to estimate q in $Q_{\text{ad}} = \{q \in L^2(0,1): q(x) \geq 0, |q| \leq \gamma^0\}$, $\gamma > 0$, in such a way that the result depends continuously on the pointwise data and on the constraints defining Q_{ad}.

The OLS problem involving a local regularization term is given by

$$\min |u(q) - \mathscr{I}z|^2_{H^2} + \beta|q|^2_{L^2([0,1]\setminus\Omega)} \quad \text{over } Q_{\text{ad}}. \tag{6.9}$$

In the notation of Section 3 we put

$$H = L^2, \quad Q = L^2, \quad Y = L^2 \times \mathbb{R}, \quad K = L^2_+ \times \mathbb{R}^+, \quad \tilde{W} = L^2 \times \mathbb{R},$$
$$W = \mathbb{R}^{M+1} \times \tilde{W}, \quad \text{and} \quad X = H^2,$$

where the domain is $(0,1)$ for all function spaces. Moreover, $g(q,\alpha) = (k - q, |q|^2 - \gamma^2)$, where $\alpha = (k, \gamma)$ and $\alpha^0 = (0, \gamma^0)$, and $Q^{\alpha}_{\text{ad}} = \{q \in L^2: g(q) \leq 0\}$. Clearly, there exists a solution of (6.9) and every element of Q^{α}_{ad} is a regular point. Let q^β be a local solution of (6.9) for $w = w^0$ and assume that for some constant $\mu > 0$

$$|u(q^\beta)(x)| \geq \mu \qquad \text{for every} \quad x \in \Omega. \tag{6.10}$$

We calculate the second derivative of the Lagrangian at q^β in direction (h, h), for $h \in L^2$, and define $A(q)$: $H_0^1 \cap H^2 \to L^2$ by $A(q)u = -(au_x)_x + qu$ and $A_1(q)u = qu$. In the following estimates the constants k_i are independent of h:

$$\begin{aligned} F_{qq}(q^\beta)(h, h) &\geq 2|\eta|_{H^2}^2 - 4|u(q^\beta) - \mathcal{I}z^0|_{H^2}|A(q^\beta)^{-1}A_1(h)\eta|_{H^2} + 2\beta|h|^2_{L^2([0,1]\setminus\Omega)} \\ &\geq k_1|A_1(h)u(q^\beta)|_{L^2}^2 - k_2|u(q^\beta) - \mathcal{I}z^0|_{H^2}|A_1(h)\eta|_{L^2} \\ &\quad + 2\beta|h|^2_{L^2([0,1]\setminus\Omega)} \\ &\geq k_1\mu^2|h|^2_{L^2(\Omega)} - k_3|h|_{L^2}|u(q^\beta) - \mathcal{I}z^0|_{H^2}|A_2(h)u(q^\beta)|_{L^2} \\ &\quad + 2\beta|h|^2_{L^2([0,1]\setminus\Omega)} \\ &\geq \min(k_1\mu^2, 2\beta)|h|_{L^2}^2 - k_4|h|_{L^2}^2|u(q^\beta) - \mathcal{I}z^0|_{H^2}|u(q^\beta)|_{L^\infty}. \end{aligned}$$

Thus $F_{qq}(q^\beta)(h, h) \geq k_5(\beta)|h|_{L^2}^2$ for some $k_5(\beta) > 0$ independent of h, provided (6.10) holds, β is sufficiently large, and $|u(q^\beta) - \mathcal{I}_{z^0}|_{H^2}$ is sufficiently small. This implies that—in the sense of Proposition 4.1—the local solution q^β is stable with respect to perturbations of the point observations z and the parameter α defining the admissible parameter set.

Clearly, the concept of local regularization requires a systematic analysis, which is not within the scope of this paper.

References

1. R. A. Adams: Sobolev Spaces, Academic Press, New York, 1975
2. G. Alessandrini: An identification problem for an elliptic equation in two variables, Ann Mat Pura Appl, to appear
3. W. Alt: Lipschitzian perturbations of infinite optimization problems, in: Mathematical Programming with Data Pertubations, II, ed. by A. V. Fiacco, Lecture Notes in Pure and Applied Mathematics, 85, Marcel Dekker, New York, 1984, 7-21
4. M. S. Berger: Nonlinearity and Functional Analysis, Academic Press, New York, 1977
5. H. Brezis: Analyse Fonctionelle, Théorie et Applications, Masson, Paris, 1983
6. G. Chavent: Local stability of the output least square parameter estimation technique, Mat Apl Comput 2 (1983), 3-22
7. G. Chavent: On parameter identifiability, Proceedings of the IFAC Symposium on Identification and Parameter Estimation, York, July 1985, 531-536
8. F. Colonius and K. Kunisch: Stability for parameter estimation in two-point boundary-value problems, J Reine Angew Math 370 (1986), 1-29
9. A. V. Fiacco: Introduction to Sensitivity and Stability Analysis in Nonlinear Programming, Academic Press, New York, 1983
10. C. W. Groetsch: The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind, Research Notes in Mathematics, 105, Pitman, Boston, 1984
11. T. Kato: Perturbation Theory for Linear Operators, 2nd edition, Springer-Verlag, Berlin, 1980
12. C. Kravaris and J. H. Seinfeld: Identification of parameters in distributed systems by regularization, SIAM J Control Optim 23 (1985), 217-241
13. K. Kunisch and L. W. White: Regularity properties in parameter estimation of diffusion coefficients in elliptic boundary-value problems, Appl Anal 21 (1986), 71-88
14. O. A. Ladyzhenskaya and N. N. Ural'tseva: Linear and Quasilinear Elliptic Equations, Academic Press, New York, 1968

15. H. Maurer and J. Zowe: First- and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems, Math Programming, 16 (1979), 98–110
16. G. R. Richter: An inverse problem for the steady state diffusion equation, SIAM J Appl Math 41 (1981), 210–221
17. D. L. Russell: Some remarks on numerical aspects of coefficient identification in elliptic systems, in: Optimal Control of Partial Differential Equations, ed. by K. H. Hoffmann and W. Krabs, Birkhäuser, Basel, 1984, 210–228
18. M. H. Schultz: Spline Analysis, Prentice-Hall, Englewood Cliffs, NJ
19. A. N. Tikhonov and V. Y. Arsenin: Solutions of Ill-Posed Problems, Winston, Washington, 1977
20. M. M. Vainberg: Variational Methods for the Study of Nonlinear Operators, Holden Day, San Francisco, 1964
21. J. Wloka: Partielle Differentialgleichungen, Teubner, Stuttgart, 1982
22. A. Wouk: A Course of Applied Functional Analysis, Wiley, New York, 1979