Stability for parameter estimation in two point boundary value problems

Fritz Colonius, K. Kunisch

Angaben zur Veröffentlichung / Publication details:

Colonius, Fritz, and K. Kunisch. 1986. "Stability for parameter estimation in two point boundary value problems." *Journal für die reine und angewandte Mathematik* 370: 1–29. https://doi.org/10.1515/crll.1986.370.1.

Stability for parameter estimation in two point boundary value problems

By *F. Colonius**)**) at Bremen and *K. Kunisch***) at Graz

1. Introduction

In recent years, the use of mathematical models not only in physical or technical sciences, but also for processes in the life sciences like physiology has become a general practice. Often the process of interest can be described by a differential equation the structure of which is determined by general principles, however, the numerical values of certain parameters are unknown [4], [9]. The parameter estimation problem consists of determining these unknown parameters from known observations (data) of the process that is being modelled. In recent years there were many contributions devoted to the numerical aspects of parameter estimation problems (see [3]–[5], [7], [9], [11]–[13] and the references given there, et al.) and to the problem of parameter identifiability; i.e. the injectivity of the map from the parameters to the observations. Furthermore it is wellknown that the parameter-to-observation map is often not continuously invertible and, more generally, the solutions of parameter estimation problems in their output-least-squares formulation do not depend continuously on the observations. However, beyond this general observation that such inverse problems are often illposed, this question received comparatively little attention. The main goal of our paper is to study these stability problems.

Let us describe the situation in a more formal way. By $\mathcal{U}$ we denote the (topological) space of parameters and $\mathcal{U}_{\text{ad}} \subset \mathcal{U}$ is the set of admissible parameters.

The observations are taken in a (normed) space $\mathcal{Z}$. For example, $\mathcal{Z}$ could be a Euclidean space in case of point observations or a function space in case of distributed observations. By $\Phi : \mathcal{U} \longrightarrow \mathcal{Z}$ we denote the mapping from the parameter to the output space. Hence Φ is determined by the specific model equation and the specific observation operator. Finally let $\mathcal{V} := \{\Phi(c) : c \in \mathcal{U}_{\text{ad}}\}$ be the set of attainable observations and let $z \in \mathcal{Z}$ be the actual observation. Note that typically $\mathcal{V}$ is not convex for identification problems.

*) Supported by Deutsche Forschungsgemeinschaft.

**) Both authors acknowledge partial support from the Fonds zur Förderung der wissenschaftlichen Forschung, Austria, # S3206.

If no modelling or observation error were present, then one could assume $z \in \mathcal{V}$. In general, this is not realistic and one must admit the case $z \notin \mathcal{V}$ or "*z close to* $\mathcal{V}$". Even if $z \in \mathcal{V}$, the inverse problem may be illposed (in the sense of Hadamard), since the preimage of z under Φ may not be unique or may not depend continuously on the observation z.

One of the most relevant approaches to the parameter identification problem is the output least squares formulation:

$$\text{(OLS)} \qquad \text{Minimize} \frac{1}{2} |\Phi(c) - z|^2 \text{ over } c \in \mathcal{U}_{\text{ad}}.$$

Again the questions of existence, uniqueness, and continuous dependence of a minimizer c^0 on z must be studied. The answers to these questions will depend on the constraint set $\mathcal{U}_{\text{ad}}$. Since in practical problems, the definition of $\mathcal{U}_{\text{ad}}$ is—to a certain extent—arbitrary, we add the study of continuous dependence of c^0 on the constraints characterizing $\mathcal{U}_{\text{ad}}$ to our list of important questions.

Let us describe two contributions from the literature which are of particular importance in the present context. In [8], G. Chavent defines the concept of output least squares identifiability (OLSI): A parameter identification problem (OLS) is called OLSI if there exists a neighborhood $\tilde{\mathcal{V}}$ of $\mathcal{V}$ such that for every $z \in \tilde{\mathcal{V}}$ there exists a unique solution of (OLS) and this solution depends continuously on z. Sufficient conditions for OLSI involve the diameter of $\mathcal{U}_{\text{ad}}$ and conditions on the first and second Fréchet derivatives of Φ; they appear to be suited primarily for the case where $\mathcal{U}$ is finite dimensional and for specific hyperbolic equations [3], [18]. Observe that uniqueness of a solution of (OLS) requires uniqueness of the projection z_{pr} of z on $\mathcal{V}$ as well as uniqueness of the inverse of Φ at z_{pr}. Since injectivity of Φ itself is difficult to verify, it is not unexpected that OLSI is difficult to obtain.

Recently, C. Kravaris and J. H. Seinfeld [11], [12] used for treatment of identification problems involving partial differential equations the theory of Tikhonov regularization, first suggested for inverse problems involving integral operators. In this case (OLS) is changed to the following regularized output least squares problem:

$$\text{(ROLS)} \qquad \text{Minimize } \frac{1}{2} |\Phi(c) - z|^2 + \beta |c|^2_{\mathcal{U}_{\text{comp}}} \text{ over } \mathcal{U}_{\text{ad}} \cap \mathcal{U}_{\text{comp}}$$

where $\beta > 0$, $\mathcal{U}_{\text{ad}}$ is closed and convex and $\mathcal{U}_{\text{comp}}$ is compactly embedded in $\mathcal{U}$. The assumption is made that z is uniquely attainable, i.e. there exists a unique $c^0 \in \mathcal{U}_{\text{ad}} \cap \mathcal{U}_{\text{comp}}$ such that $\phi(c^0) = z$. Then under appropriate additional conditions, all solutions $c^\beta(\hat{z})$ of

$$\text{minimize } \frac{1}{2} |\Phi(c) - \hat{z}| + \beta |c|^2_{\mathcal{U}_{\text{comp}}} \text{ over } \mathcal{U}_{\text{ad}} \cap \mathcal{U}_{\text{comp}}$$

converge to c^0 as $\hat{z} \to z$ and $\beta \to 0$ in an appropriate way. The key tool of the proofs in [11], [12] is the use of (a variant of) Tikhonov's classical lemma, which states that continuity and injectivity of Φ when restricted to a compact subset $\mathcal{U}_0$ of $\mathcal{U}$ imply continuity of the inverse Φ^{-1} on $\Phi(\mathcal{U}_0)$.

In this paper, we drop the requirement that the solution c^0 of (OLS) is unique and concentrate on the question whether *local solutions* of (OLS) depend continuously on the observation $z \in \mathscr{Z}$ and on the constraints that characterize $\mathscr{U}_{ad}$. We call this property Output-Least-Squares Stability (OLS-stability). In order to guarantee OLS-stability, we do not need any assumption on uniqueness of the projection z_{pr} of z onto $\mathscr{V}$ or on the injectivity of Φ. We also investigate the advantages of regularization for obtaining continuous dependence of the optimal solutions on the problem data. This property will be called ROLS-stability. However, we add a term of the form $\beta|c|^2_{\mathscr{U}}$ to the fit-to-data criteria, instead of $\beta|c|^2_{\mathscr{U}_{comp}}$ as in [11], [12].

The proofs of OLS-stability are based on perturbation theory of infinite dimensional optimization problems as initiated by S. M. Robinson [15]. Specifically, we use recent results by W. Alt [1] on the stability of local solutions under perturbations of the problem data. The applicability of this approach will be demonstrated by analysing a class of two point boundary value problems. First order necessary optimality conditions allow to study the structure of the minimizers. A result in [13] on a smoothing effect for the minimizers is refined and a sufficient condition is given for the alternative that either $z \in \mathscr{V}$ or "c^0 is on the boundary of $\mathscr{U}_{ad}$".

Lower bounds on the second Fréchet derivative of the Lagrangian are the essential tool in establishing OLS-stability. These bounds can be obtained by assuming finite dimensionality of the parameter space, by exploiting the fact that a norm bound on the set of admissible parameters guarantees non-triviality of certain Lagrange multipliers or by adding a regularization term.

Notation. The notation used is rather standard and we only make a few comments. We use the common notation for Sobolev spaces and we drop the domain if it is the interval (0, 1), for example $H^1 = H^1(0, 1)$. A subscript is used to denote the norm of a certain space, as $|\cdot|_{H^1}$, for instance, except with H^0, where we use $|\cdot|$ for the norm and $(\cdot,\cdot)$ for the inner product. The positive, respectively negative cone in a Banach lattice is denoted by a subscript "+", resp. "−". The domain and range of an operator A are denoted by $\mathscr{D}(A)$ and $\mathscr{R}(A)$. A superscript "c" denotes the complement of a set, and a superscript "$*$" denotes the topological dual of a Banach space. The Fréchet derivative of a map with respect to a variable x is denoted by a subscript x.

2. Preliminaries from two-point boundary value problems

Let us consider the equation

(2. 1) $$-(au_x)_x + cu = f \text{ on } (0, 1),\ R_i u = 0,\ i = 1, 2,$$

where $f \in H^0$, $a \in C^1$, $a(x) \geqq \underline{a} > 0$, $R_i u = \alpha_{i1} u(0) + \alpha_{i2} u'(0) + \alpha_{i3} u(1) + \alpha_{i4} u'(1)$, $\alpha_{ij} \in \mathbb{R}$, and the *unknown parameter* c is considered as an element of H^0. Let A be the differential operator in H^0 associated with (2. 1), i.e.

(2. 2) $$\mathscr{D}(A) = \{\phi \in H^0 : \phi \in H^2,\ R_i \phi = 0,\ i = 1, 2\},\quad A\phi = -(a\phi_x)_x + c\phi.$$

If the dependence of A on c is relevant we write $A(c)$. Throughout we make the following assumption:

(H 1) There exist constants $\alpha \geqq 0$ and $k > 0$ such that $(A(c)\,\phi, \phi) \geqq k|\phi|^2_{H^1}$, for all $\phi \in \mathscr{D}(A)$ and $c \in Q := \{c \in H^0 : c(x) \geqq \alpha \text{ a.e.}\}$.

For example, in the case of homogeneous Dirichlet conditions in (2. 1) one can take $k = \underline{a}$ and $\alpha = 0$. Or, in the case of homogeneous Neumann conditions in (2. 1) if $\alpha > 0$ is given, then $k > 0$ can be found so that (H 1) holds, ([2], pg. 31).

We also assume:

(H 2) The boundary conditions R_i, $i = 1, 2$ in (2. 1) are such that $A(c)$ is selfadjoint.

Recall that the operator $A(c)$ is symmetric if and only if

$$a(1)\,(\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}) = a(0)\,(\alpha_{13}\alpha_{24} - \alpha_{14}\alpha_{23}),$$

and

$$\operatorname{rank}\begin{pmatrix}\alpha_{11}, \dots, \alpha_{14}\\ \alpha_{21}, \dots, \alpha_{24}\end{pmatrix} = 2,$$

see [10], pg. 62. If in addition (H 1) holds, then A is selfadjoint and a homeomorphism from $\mathscr{D}(A)$ endowed with the graph norm to H^0.

We denote the unique solution of (2. 1) in $\mathscr{D}(A)$ by u, $u(c)$ or $u(c, f)$ as dictated by the context. Simple perturbation theory implies that an $\tilde{\varepsilon} > 0$ can be chosen such that

$$(A(c)\,\phi, \phi) \geqq \frac{k}{2}\,|\phi|^2_{H^1} \text{ for all } \phi \in \mathscr{D}(A) \text{ and } c \in \{c \in H^0 : \operatorname{dist}(c, Q) \leqq \tilde{\varepsilon}\}.$$

In the optimization problems of Sections 4—5 the unknown parameter c will be allowed to vary in the set

$$\mathscr{U}_{\text{ad}} = \{c \in H^0 : c(x) \geqq \alpha \text{ a.e., } |c| \leqq \gamma\} = \{c \in Q : |c| \leqq \gamma\},$$

where γ is a positive constant with $\alpha < \gamma$. Let $\mathscr{U}$ be the closed set containing $\mathscr{U}_{\text{ad}}$ defined by

$$\mathscr{U} = \{c \in H^0 : |\hat{c} - c| \leqq \tilde{\varepsilon} \text{ for some } \hat{c} \in \mathscr{U}_{\text{ad}}\}.$$

Lemma 2. 1. *Let* (H 1) *hold. Then there are constants* $k_1 > 0$ *and* $k_2 = |a|_{C^1} + \tilde{\varepsilon} + \gamma$ *such that*

$$k_1\,|\phi|_{H^2} \leqq |A(c)\,\phi| \leqq k_2\,|\phi|_{H^2}$$

holds for all $c \in \mathscr{U}$ *and* $\phi \in \mathscr{D}(A)$.

Proof. The estimate from above is obvious. So let us demonstrate the estimate from below. First note that

$$(2.\,3) \qquad |A(c)\,\phi| \geqq \underline{a}\,|\phi_{xx}| - (|a_x|_{L^\infty} + \gamma + \tilde{\varepsilon})\,|\phi|_{H^1} \quad \text{for} \quad c \in \mathscr{U}.$$

Moreover

$$(2.\,4) \qquad |A(c)\,\phi| \geqq \frac{k}{2}\,|\phi|_{H^1} \quad \text{for all} \quad c \in \mathscr{U}.$$

Let $K_1 = |a_x|_{L^\infty} + \gamma + \tilde{\varepsilon}$ and choose $K_2 > \frac{2K_1}{k}$. Multiplying (2. 4) by K_2 and adding (2. 3) we have

$$(1+K_2)\,|A(c)\,\phi| \geqq \underline{a}\,|\phi_{xx}| + \left(K_2\frac{k}{2} - K_1\right)|\phi|_{H^1},$$

which implies the first inequality in Lemma 2. 1.

In the following lemma the dependence of u on c and f is studied. By $\rightharpoonup$ we denote weak convergence.

Lemma 2. 2. *Let* (H 1), (H 2) *hold. If* $f^n \rightharpoonup f$ *in* H^0 *and* $c^n \rightharpoonup c$ *in* H^0 *with* c^n *and* c *in* $\mathscr{U}$ *for* $n = 1, 2, \ldots,$ *then*

$$u(c^n, f^n) \rightharpoonup u(c, f) \text{ in } H^2.$$

Proof. By Lemma 2. 1 the sequence $u(c^n, f^n)$ is uniformly bounded in H^2. Therefore there exist a subsequence n_k and an element $w \in H^2$ such that $u(c^{n_k}, f^{n_k}) \rightharpoonup w$ in H^2 and $u(c^{n_k}, f^{n_k}) \to w$ in H^1. Let $v \in C_0^\infty$. Then

$$(a u(c^{n_k}, f^{n_k})_x, v_x) + (c^{n_k} u(c^{n_k}, f^{n_k}), v) = (f^{n_k}, v).$$

Taking the limit in this equation we obtain

$$(a w_x, v_x) + (c w, v) = (f, v),$$

and since $w \in H^2$

$$-((a w_x)_x, v) + (c w, v) = (f, v) \quad \text{for all} \quad v \in C_0^\infty.$$

But C_0^∞ is dense in H^0 and $c \in \mathscr{U}$; therefore $w = u(c, f)$.

Lemma 2. 3. *Let* (H 1), (H 2) *hold. Then*

$$|u(c_1) - u(c_2)|_{H^2} \leqq k_1^{-2}\,|c_1 - c_2|\,|f|,$$

for all $c_1, c_2 \in \mathscr{U}$.

Proof. Let $v = u(c_1) - u(c_2)$. Then v satisfies

$$A(c_2)\,v = (c_2 - c_1)\,u(c_1).$$

By Lemma 2. 1 we obtain

$$|v|_{H^2} \leqq k_1^{-1}\,|c_1 - c_2|\,|u(c_1)|_{H^2} \leqq k_1^{-2}\,|c_1 - c_2|\,|f|,$$

and the claim is verified.

Lemma 2. 4. *Let* (H 1), (H 2) *hold. The mapping* $c \to u(c)$ *from* $\mathscr{U} \subset H^0$ *to* H^2 *is continuously Fréchet differentiable with Fréchet derivative* $u_c(c)\,h = \eta(h)$ *given as the unique solution of*

$$A(c)\,\eta(h) = -h u(c). \tag{2. 5}$$

Proof. This result is a very special case of [7], Theorem 2. 2. 4. For the sake of completeness, we include the proof. Let $c \in \mathscr{U}$. First we show Gateaux differentiability in direction $h \in H^0$. Let $\varepsilon \in \mathbb{R}$ be such that $c + \varepsilon h \in \mathscr{U}$ and put

$$z^\varepsilon = \varepsilon^{-1}(u(c + \varepsilon h) - u(c)).$$

Then z^ε satisfies

$$-(a z_x^\varepsilon)_x + c z^\varepsilon = -hu(c+\varepsilon h), \quad R_i z^\varepsilon = 0, \quad i = 1, 2, . \tag{2.6}$$

From Lemma 2.3 it follows that $u(c+\varepsilon h) \to u(c)$ in H^2 and therefore $hu(c+\varepsilon h) \to hu(c)$ in H^0 as $\varepsilon \to 0$. By Lemma 2.1 and (2.6) we have that z^ε converges in H^2. We denote the limit by $\eta(h)$ or simply η. Taking the limit as $\varepsilon \to 0$ in (2.6) implies

$$-(a\eta_x)_x + c\eta = -hu(c), \quad R_i \eta = 0, \quad i = 1, 2. \tag{2.7}$$

Next observe that by Lemma 2.1

$$|\eta(h)|_{H^2} \leqq k_1^{-1} |hu(c)| \leqq k_1^{-2} |h| \, |f|, \tag{2.8}$$

and therefore $u_c(c) \in \mathscr{L}(H^0, H^2)$. For c and $\bar{c} \in \mathscr{U}$ we put $v = u_c(c)\,h - u_c(\bar{c})\,h$. Then

$$A(\bar{c})\,v + (c - \bar{c})\,u_c(c)\,h = -h(u(c) - u(\bar{c})).$$

By Lemma 2.1 this implies

$$|v|_{H^2} \leqq k_1^{-1}(|c-\bar{c}| \, |u_c(c)\,h|_{H^2} + |h| \, |u(c) - u(\bar{c})|_{H^2}).$$

Using Lemma 2.1 once again, together with (2.8) and Lemma 2.3 we obtain

$$|u_c(c)\,h - u_c(\bar{c})\,h|_{H^2} = |v|_{H^2} \leqq 2k_1^{-3} |c - \bar{c}| \, |h| \, |f|.$$

Therefore $c \to u_c(c)$ from $\mathscr{U}$ to $\mathscr{L}(H^0, H^2)$ is continuous. This together with (2.8) implies continuous Fréchet differentiability of $c \to u(c)$; (see e.g. [19], pg. 270).

Lemma 2.5. *Let* (H1), (H2) *hold. The mapping $c \to u(c)$ from $\mathscr{U} \subset H^0$ to H^2 is twice continuously Fréchet differentiable and the second Fréchet derivative*

$$u_{cc}(c)\,(h, k) = \xi(h, k) = \xi$$

is the unique solution of

$$A(c)\,\xi = -k u_c(c)\,(h) - h u_c(c)\,(k). \tag{2.9}$$

The proof is quite similar to the one of the previous lemma and is therefore not included here.

Let $V = \{v \in C^\infty : R_i v = 0, \ i = 1, 2\}$. For $u \in C^\infty$ let

$$|u|_{\tilde{H}^{-2}} = \sup_{v \in V} \frac{|(u, v)|}{|v|_{H^2}},$$

and define $\tilde{H}^{-2}$ as the completion of C^∞ with respect to the $|u|_{\tilde{H}^{-2}}$ norm. We have $|(u, v)| \leqq |u|_{\tilde{H}^{-2}} \, |v|_{H^2}$ for all $u \in \tilde{H}^{-2}$ and $v \in V$. For further discussion of this Hilbert space we refer to [16], for example.

Lemma 2.6. *Let* (H1), (H2) *hold. Then*

$$k_1 |\phi| \leqq |A(c)\,\phi|_{\tilde{H}^{-2}} \leqq k_2 |\phi| \tag{2.11}$$

for all $c \in \mathscr{U}$ and $\phi \in \mathscr{D}(A)$.

Proof. The second estimate in (2. 11) is a simple consequence of Lemma 2. 1 and the definition of the norm in $\tilde{H}^{-2}$. We also observe that

(2. 12) $$|(u, v)| \leqq |u|_{\tilde{H}^{-2}} |v|_{H^2} \quad \text{for all} \quad u \in \tilde{H}^{-2} \quad \text{and} \quad v \in \mathcal{D}(A).$$

Next, let $\phi \in \mathcal{D}(A)$ and $v \in \mathcal{D}(A)$ with $A(c)v = \phi$. Then by Lemma 2. 1 we have $k_1 |v|_{H^2} \leqq |\phi|$. Moreover

$$|\phi|^2 = (\phi, A(c)\, v) = (A(c)\, \phi, v) \leqq |A(c)\, \phi|_{\tilde{H}^{-2}} |v|_{H^2}$$
$$\leqq k_1^{-1} |A(c)\, \phi|_{\tilde{H}^{-2}} |\phi|.$$

Therefore $k_1 |\phi| \leqq |A(c)\, \phi|_{\tilde{H}^{-2}}$, which is the first inequality in (2. 11).

3. Tools from optimization theory

In this section we collect some facts from optimization theory in Banach spaces. We cite results on necessary and sufficient optimality conditions and on stability under perturbation of the problem data.

Consider real Banach spaces X and Y, a metric space (W, δ), a closed convex cone $K \subset Y$ with vertex at the origin, and mappings $f: D \times W \to \mathbb{R}$ and $g: X \times W \to Y$, where $D \subset X$ is open. For each $w \in W$ we consider the following optimization problem:

(P)^w $$\text{minimize } f(x, w) \text{ subject to } g(x, w) \in K.$$

Keep $w^0 \in W$ fixed. Then for notational simplicity we often omit the argument w^0 in f and g and refer to (P)$^{w^0}$ as the original *unperturbed problem* (P). We always assume that (for fixed w^0), g is continuous, and f and g are continuously Fréchet differentiable (with respect to x) in a neighborhood of $x^0 \in g^{-1}(K)$.

Throughout this section it is assumed that

(3. 1) $$g^{-1}(K) = \{x \in X : g(x) \in K\} \subset D.$$

The following regularity condition for an element $x^0 \in g^{-1}(K)$ plays a central role:

(3. 2) $$0 \in \operatorname{int}\{g(x^0) + \mathcal{R}(g_x(x^0)) - K\}.$$

Here $\mathcal{R}$ stands for the range of a linear operator. A point x^0 satisfying (3. 2) is called regular. The following first order necessary condition holds.

Theorem 3. 1 ([20], Theorem 3. 1). *Let* $x^0 \in g^{-1}(K)$ *be a local solution of* (P) *satisfying the regularity condition* (3. 2). *Then there exists a Lagrange multiplier* $\lambda \in Y^*$ *such that*

(3. 3) $$f_x(x^0) - \lambda g_x(x^0) = 0,$$

(3. 4) $$\lambda z \geqq 0, \quad \textit{for all} \quad z \in K,$$

(3. 5) $$\lambda g(x^0) = 0.$$

For $\lambda \in Y^*$ *satisfying* (3. 3)—(3. 5) *the function*

(3. 6) $$F(x) = f(x) - \lambda g(x)$$

is called a Lagrangian for (P) *at* x^0.

This theorem as well as the subsequent ones actually differ from the cited results in the literature in that here f is only defined on a subset D of X. But due to (3. 1) these minor generalizations can easily be verified.

The following second order optimality conditions can be found in [14], for example.

Theorem 3. 2 ([14], Theorem 5. 6). *Let $x^0 \in g^{-1}(K)$ be regular. Suppose that f and g are twice Fréchet differentiable (with respect to x) at x^0 and that there are constants $\gamma > 0$ and $\beta > 0$ with*

$$F_{xx}(x^0)(h, h) \geqq \gamma \|h\|^2, \tag{3. 7}$$

for all $h \in g_x^{-1}(K + \mathbb{R} g(x^0)) \cap \{h: \lambda g(x^0) h \leqq \beta \|h\|\}$. Then there exist $\alpha > 0$ and $\varrho > 0$ such that

$$f(x) \geqq f(x^0) + \alpha \|x - x^0\|^2$$

for all $x \in g^{-1}(K)$ with $\|x - x^0\| \leqq \varrho$.

Remark 3. 1. The estimate for the second derivative $F_{xx}(x^0)$ of the Lagrangian is needed only for those h which violate the first order sufficient optimality condition

$$f_x(x^0) h = \lambda g_x(x^0) h \geqq \beta \|h\|,$$

see ([14], Theorem 5. 3).

The regularity condition (3. 2) together with the second order sufficiency condition (3. 6) imply continuous dependence of local solutions on the problem data as is shown by the following result which is a special case of ([1], Theorem 6); see ([1], Theorem 3 and Remark 3; and Theorem 3. 2 above).

Theorem 3. 3. *Let the assumptions of Theorem 3. 2 be satisfied and assume in addition:*

There exists a neighborhood $U = U_x \times U_w$ of (x^0, w^0) such that for some constant $L_f > 0$

$$|f(x, w) - f(x', w^0)| \leqq L_f(\|x - x'\| + \delta(w, w^0)) \tag{3. 8}$$

for all $(x, w) \in U$ and all $x' \in U_x$.

For U as above the mapping $g(x, \cdot)$ is Lipschitz continuous at w^0 for each $x \in U_x$, i.e. there exists some constant $L_g > 0$ such that

$$|g(x, w) - g(x, w_0)| \leqq L_g \delta(w, w_0) \tag{3. 9}$$

for all $(x, w) \in U$.

Then there exist $r > 0$, $d > 0$ and a neighborhood V of w_0 such that:

(i) *The local extremal value function*

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

is Lipschitz continuous at w^0.

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

(ii) *For any sequence x_n with $g(x_n, w) \in K$, $\|x_n - x^0\| \leqq r$ and*

$$\lim_{n \to \infty} f(x_n, w) = \mu_r(w)$$

it follows that $\|x_n - x_0\| < r$ for all sufficiently large n.

(iii) *If there exists x_w with $g(x_w, w) \in K$, $\|x_w - x^0\| \leqq r$ and $\mu_r(w) = f(x_w, w)$, then $\|x_w - x^0\| < r$ and*

$$\|x_w - x^0\| \leqq d\,\delta(w, w^0)^{\frac{1}{2}}.$$

Remark 3. 2. Suppose that W is a subset of a Banach space and that f and g are continuously Fréchet differentiable with respect to x and w in a neighborhood of (x^0, w^0). Then by the mean value theorem conditions (3. 8) and (3. 9) hold.

4. Application of first order optimality conditions

In this section we begin our study of the parameter identification problem formulated as an output least squares problem. We assume to have available an observation $z \in H^0$ and we minimize the quadratic fit-to-data criterion $|u(c) - z|^2$ over $c \in \mathscr{U}_{\mathrm{ad}}$. In applications, very often, information of the modelled system will only be available from a certain subset $\Omega \subset (0, 1)$; on several occasions we will therefore discuss the possibility of generalizing our results to this situation. The observation z can also be thought of as obtained from interpolating point measurements.

In all that follows we assume that the hypotheses (H 1) and (H 2) of Section 2 are satisfied.

The precise formulation of the minimization problem is as follows:

(OLS) $$\text{minimize } \frac{1}{2}|u(c) - z|^2 \text{ over } c \in \mathscr{U}_{\mathrm{ad}},$$

where $\mathscr{U}_{\mathrm{ad}} = \{c \in H^0 : c(x) \geqq \alpha,\ |c| \leqq \gamma\}$ and $u(c)$ is the unique solution of (2. 1).

Define the attainable set by

$$\mathscr{V} := \{u(c) : c \in \mathscr{U}_{\mathrm{ad}}\}.$$

If $z \in \mathscr{V}$ the minimal value of the fit-to-data criterion is zero. We do not impose such an attainability assumption and instead consider the general case. Let us point out that if the observation z is constructed by piecewise linear interpolation of point observations, then such a z cannot be attainable, since $u(c) \in H^2$ for all $c \in \mathscr{U}_{\mathrm{ad}}$.

Remark 4. 1. The pointwise constraint in $\mathscr{U}_{\mathrm{ad}}$ is imposed in order to guarantee unique solvability of the system equation (2. 1). The norm constraint in $\mathscr{U}_{\mathrm{ad}}$ will be needed to prove existence of an optimal solution c^0. In fact, without this latter condition (OLS) may not have a solution, see Remark 4. 7 below. Recall also the assumption $\alpha < \gamma$ which implies $\mathscr{U}_{\mathrm{ad}} \neq \emptyset$.

Proposition 4. 1. *For every $z \in H^0$, there exists a solution c^0 of* (OLS) *over* $\mathscr{U}_{ad}$.

Proof. Observe that $\mathscr{U}_{ad}$ is weakly closed in H^0, since it is convex and closed with respect to the norm topology. The latter property follows from the fact that norm convergence in H^0 implies pointwise convergence almost everywhere of a subsequence. Since $\mathscr{U}_{ad}$ is bounded it is also weakly sequentially compact.

By Lemma 2. 2 the mapping $c \longrightarrow u(c)$ is continuous from the weak to the strong topology of H^0. Therefore $c \longrightarrow |u(c)-z|^2$ is weakly lower semi-continuous. Hence we have to minimize a lower semi-continuous functional on a weakly sequentially compact set, which yields existence of a solution c^0 of (OLS) in $\mathscr{U}_{ad}$.

Remark 4. 2. The proof of the previous proposition can easily be adapted to more general fit-to-data criteria and different sets of admissible parameter values. If only the fit-to-data criterion is weakly lower semi-continuous in the variable u from the weak H^2-topology to $\mathbb{R}$ and the set of admissible parameter values is weakly sequentially compact, existence of a solution c^0 is guaranteed.

To apply the general theory of Section 3 we put

$$W = H^0 \times C \times \mathbb{R}, \quad X = H^0, \quad Y = H^0 \times \mathbb{R}, \quad K = H^0_- \times \mathbb{R}_- \subset Y,$$

where H^0_- and $\mathbb{R}_-$ are the natural negative cones in H^0 and $\mathbb{R}$ respectively. Then X, Y and W are Banach spaces and K is a closed convex cone in Y with vertex at the origin. Let $g: X \times W \longrightarrow Y$ be defined by

$$g(c, w) = (g_1(c, w), g_2(c, w)) = \left(\alpha - c, \frac{1}{2}(|c|^2 - \gamma^2)\right),$$

where $c \in X$ and $w = (z, \alpha, \gamma) \in W$, with $\alpha < \gamma$. Here α stands for the constant function with value α. We henceforth use the more suggestive notation

$$g_1(c, \alpha) = \alpha - c,$$

and

$$g_2(c, \gamma) = \frac{1}{2}(|c|^2 - \gamma^2).$$

Similarly let $f: \mathscr{U} \times W \longrightarrow \mathbb{R}$ be defined by

$$f(c, z) = \frac{1}{2}|u(c) - z|^2.$$

Note, that the minimization criterion is defined on a neighborhood $\mathscr{U}$ of $\mathscr{U}_{ad}$ but not on all of H^0. With the above notation, the OLS-problem on $\mathscr{U}_{ad}$ is equivalent to the problem

$$(\text{OLS})^w \qquad \text{minimize } f(c, z) \text{ subject to } (g_1(c, \alpha), g_2(c, \gamma)) \in K.$$

Lemma 2. 4—2. 5 show that this problem can be considered as a special case of the abstract optimization problem $(\text{P})^w$ in Section 3. We check the regularity condition (3. 2) for (OLS).

Lemma 4. 1. *Every $c \in \mathscr{U}_{ad}$ is regular in the sense of* (3. 2).

Proof. We have to show that

$$0 \in \operatorname{int}\{g(c)+g_c(c)\,H^0 - H^0_- \times \mathbb{R}_-\}. \tag{4. 1}$$

Recall also that

$$g_c(c)\,h = (-h, (c, h)).$$

Hence (4. 1) is equivalent to

$$0 \in \operatorname{int}\left\{\left(\alpha - c - h - H^0_-, \frac{1}{2}(|c|^2-\gamma^2)+(c, h) - \mathbb{R}_-\right) : h \in H^0\right\}. \tag{4. 2}$$

Now let $(\phi, r) \in H^0 \times \mathbb{R}$ with $|(\phi, r)| < \delta$, $\delta > 0$ to be chosen sufficiently small. Denoting the projection of ϕ onto the closed convex cone H^0_- by ϕ^- we get that $\phi = \alpha - c - (\alpha - c - \phi - \phi^-) - \phi^-$, as in the first component of the set in (4. 2). Concerning the second component of this set we observe that

$$\begin{aligned}\frac{1}{2}(|c|^2-\gamma^2)+(c, \alpha - c - \phi - \phi^-) &= -\frac{1}{2}\gamma^2 - \frac{1}{2}|c|^2 + (c, \alpha) - (c, \phi + \phi^-)\\ &\leqq \frac{1}{2}(\alpha^2-\gamma^2) + 2|c|\,\delta.\end{aligned}$$

Since $\alpha < \gamma$ one can always (for δ sufficiently small) choose $\tilde{r} \in \mathbb{R}_+$ such that

$$r = \frac{1}{2}(|c|^2-\gamma^2) + (c, \alpha - c - \phi - \phi^-) + \tilde{r}.$$

This proves regularity of c.

Next we derive first order necessary optimality conditions for (OLS) over $\mathscr{U}_{ad}$.

Proposition 4. 2. *Let c^0 be a local solution of* (OLS) *in $\mathscr{U}_{ad}$. Then there exist $\lambda_1 \in H^0_-$ and $\lambda_2 \in \mathbb{R}_-$ such that for every $h \in H^0$, the solution $\eta = \eta(h)$ of $A(c^0)\,\eta = -h\,u(c^0)$ satisfies*

$$(u(c^0)-z, \eta) + (\lambda_1, h) - \lambda_2(c^0, h) = 0 \tag{4. 3}$$

and

$$(\lambda_1, \alpha - c^0) = 0, \quad \lambda_2(|c^0|^2-\gamma^2) = 0. \tag{4. 4}$$

Proof. The result follows from Theorem 3. 1, Lemmas 4. 1 and 2. 4, equation (4. 1), and an application of the chain rule.

Corollary 4. 1. *Let c^0 be a local solution of* (OLS) *over $\mathscr{U}_{ad}$. Then there exist $\lambda_1 \in H^0_-$ and $\lambda_2 \in \mathbb{R}_-$ such that the unique solution p of*

$$A(c^0)\,p = u(c^0) - z \tag{4. 5}$$

satisfies

$$p\,u(c^0) = \lambda_1 - \lambda_2 c^0, \tag{4. 6}$$

and (4. 4) *holds*

Equations (4. 6) *and* (4. 4) *can, equivalently, be replaced by*

(4. 7) $$(pu(c^0)+\lambda_2 c^0, h) \leqq 0, \quad \textit{for all} \quad h \in H^0$$

with $h(x) \geqq \alpha - c^0(x)$ *a.e. and*

(4. 8) $$\lambda_2(|c^0|^2-\gamma^2)=0.$$

Proof. Proposition 4. 2 and selfadjointness of A imply for all $h \in H^0$ that

$$\begin{aligned} 0 &= (u(c^0)-z, \eta)+(\lambda_1, h)-\lambda_2(c^0, h) \\ &= (A(c^0)\, p, \eta)+(\lambda_1-\lambda_2 c^0, h) \\ &= (-u(c^0)\, p+\lambda_1-\lambda_2 c^0, h). \end{aligned}$$

Thus the first part of the corollary follows. Obviously, (4. 4) and (4. 6) imply (4. 7) and (4. 8). Conversely, define

$$\lambda_1 := pu(c^0)+\lambda_2 c^0.$$

Then $\lambda_1 \in H^0_-$ since any $h \in H^0_+$ is admissible in (4. 7). Appropriate choices of h in (4. 7) yield that $\lambda_1(x)=0$ if $c^0(x)>\alpha$. Thus also (4. 4) holds.

Remark 4. 3. Let $M=\{x: \lambda_1(x)=0\}$. Then from (4. 4) we obtain

$$M^c \subset \{x \in [0, 1]: c^0(x)=\alpha\}, \text{ except for a set of measure zero.}$$

Also note that if $\lambda_2=0$ (e.g. if $|c^0|<\gamma$), then $(pu(c^0), h) \leqq 0$ for all $h \in H^0$ with $h(x) \geqq \alpha - c^0(x)$ a.e. Therefore, the solution of the adjoint equation and $u(c^0)$ have opposite signs or at least one of them is zero.

Remark 4. 4. Let us also consider three different fit-to-data criteria:

(4. 9) $$\frac{1}{2}|u(c)-z|^2_{H^0(\Omega)}$$

with Ω a measurable set in (0, 1) and $z \in H^0(\Omega)$,

(4. 10) $$\frac{1}{2}|u(c)(\tilde{x})-z|^2_{\mathbb{R}}$$

with $\tilde{x} \in [0, 1]$ and $z \in \mathbb{R}$ and

(4. 11) $$\frac{1}{2}|u(c)-z|^2_{H^1}$$

with $z \in H^1$. It can easily be seen that the problem of minimizing one of these criteria over $\mathscr{U}_{\mathrm{ad}}$ subject to (2. 1) holding has a solution.

We extend z to an H^0-function on (0, 1) in case of (4. 9) and to a constant function with value z in case of (4. 10). In either of the cases (4. 9)—(4. 10) the formula analogous to (4. 3) becomes

(4. 3 A) $$(u(c^0)-z, \eta\chi)+(\lambda_1, h)-\lambda_2(c^0, h)=0,$$

where χ is the characteristic function of Ω in case of (4. 9), and the delta function with weight at $\tilde{x}$ in case of (4. 10). In either case $\chi \in H^{-1}$. Let us define p as the unique solution of

(4. 5 A) $$A(c^0)\, p=(u(c^0)-z)\,\chi.$$

Note, that $p \in H^2$ in case of (4. 9) and $p \in H^1$ if (4. 10) is chosen as the fit-to-data criterion. Then the assertions of Corollary 4. 1 in particular (4. 4), (4. 6) and (4. 7), remain correct.

In a similar way, formula (4. 3) becomes

$$(u(c^0)-z, \eta)_{H^1} + (\lambda_1, h)_{H^0} - \lambda_2(c^0, h) = 0 \tag{4. 3B}$$

in case of (4. 11).

Remark 4. 5. At several occasions it will be convenient to consider a finite dimensional space of parameters. Let H_N be a finite dimensional subspace of H^0. Assume further that set theoretically $H_N \subset L^\infty$. Define

$$\mathscr{U}_{\text{ad}}^N := \mathscr{U}_{\text{ad}} \cap H_N,$$

and as before let $W = H^0 \times C \times \mathbb{R}$ be the space of perturbation parameters.

Then, with $X = H_N$, $Y := H_N \times \mathbb{R}$ and $K := (H_-^0 \cap H_N) \times \mathbb{R}_-$, the Output Least Squares Problem over $\mathscr{U}_{\text{ad}}^N$ is a special case of the abstract optimization problem $(\text{P})^w$ in Section 3.

One can easily see that the assertions of Proposition 4. 1 and Lemma 4. 1 remain valid, i.e. an optimal solution c^0 of (OLS) over $\mathscr{U}_{\text{ad}}^N$ exists and each solution is regular. Hence for $h \in H_N$ also the assertions (4. 3) and (4. 4) of Proposition 4. 2 hold. Furthermore, with p defined by (4. 5), it follows that

$$(pu(c^0), h) = (\lambda_1 - \lambda_2 c^0, h) \tag{4. 6B}$$

for all $h \in H_N$.

We now investigate smoothness properties of a local solution c^0. These results are generalizations of ([13], Section 4). Let I denote the set of all x, where $c^0(x)$ does not lie on the boundary α of $\mathscr{U}_{\text{ad}}$, i.e.:

$$I = \{x \in [0, 1] : c^0(x) > \alpha\}.$$

Unless otherwise specified we assume that I has positive Lebesgue measure meas(I). We shall make use of the following alternative assumption on an open subset $V \subset (0, 1)$ with meas$(V \cap I) > 0$.

(A) *Either* f is not a.e. zero on $V \cap I$ and for every subinterval $\tilde{V}$ of V with meas$(\tilde{V} \cap I) > 0$, the function $u(c^0) - z$ is not a.e. zero on $\tilde{V} \cap I$, *or* for any such subinterval $\tilde{V}$ of V with meas$(\tilde{V} \cap I) > 0$, the function f is not a.e. zero on $\tilde{V} \cap I$ and $u(c^0) - z$ is not a.e. zero on $V \cap I$.

Theorem 4. 1. *For a local solution c^0 of* (OLS) *over $\mathscr{U}_{\text{ad}}$, let $V \subset (0, 1)$ be open with* meas$(V \cap I) > 0$ *and* (A) *holding. Then $\lambda_2 < 0$ and $c^0|I = -\lambda_2^{-1} pu(c^0)|I$. Further, for any open set $U \subset (0, 1)$ with* meas$((U \cap I) - I) = 0$ *it follows that $c^0|U \in H^2(U)$.*

Thus Theorem 4. 1 establishes the *local smoothing* property that—under the above assumptions—the local solution c^0 will have jumps only when it "enters or leaves the boundary α" of the admissible set.

Proof of Theorem 4. 1. From (4. 7) in Corollary 4. 1 we obtain

$$pu(c^0) = -\lambda_2 c^0 \text{ a.e. on } I. \tag{4. 12}$$

Suppose next that $\lambda_2 = 0$. Then $pu(c^0) = 0$ a.e. on I. Suppose $u(c^0)(\bar{x}) \neq 0$ for some point of density $\bar{x}$ of $V \cap I$. Then $u(c^0)(x) \neq 0$ for all x in a neighborhood $\tilde{V}$ of $\bar{x}$ with $\text{meas}(\tilde{V} \cap I) > 0$. Hence $p(x) = 0$ for $x \in \tilde{V} \cap I$ and therefore

$$u(c^0) - z = 0 \text{ a.e. on } \tilde{V} \cap I.$$

This contradicts the first part of the alternative (A). Since almost all elements of $\tilde{V} \cap I$ are points of density it follows that $u(c^0)$ is zero a.e. on $V \cap I$. This implies $f = 0$ a.e. on $V \cap I$ and again we arrive at a contradiction to the first part of the alternative.

Hence $\lambda_2 \neq 0$. In the same way also the second part of the alternative implies $\lambda_2 \neq 0$. Thus $c^0 = -\lambda_2^{-1} pu(c^0)$ on I. Since p and $u(c^0)$ are elements of H^2 the final assertion follows easily.

Remark 4. 6. The condition "for any subinterval $\tilde{V}$ of V with $\text{meas}(\tilde{V} \cap I) > 0$ the function f (resp. $u(c^0) - z$) is not a.e. zero on $\tilde{V} \cap I$" in alternative (A) is implied by $\text{meas}\{x \in V \cap I : f(x) = 0\} = 0$ (resp. $\text{meas}\{x \in V \cap I : u(c^0)(x) - z(x) = 0\} = 0$), but not conversely. As an example one may take $V = I = (0, 1)$ and $f = 1 - \chi_C$, where χ_C is the indicator function of a generalized Cantor set C with $0 < \text{meas}\, C < 1$. Hence C is a closed set with no interior points (see e.g. [6], pg. 189). Thus f is not a.e. zero on any subinterval of V but f is not a.e. different from zero.

We immediately obtain from the proof of Theorem 4. 1 the following consequence on "attainability" of the observation z. Observe that $|c^0| < \gamma$ implies $\lambda_2 = 0$.

Corollary 4. 2. *For a local solution c^0 of* (OLS) *over $\mathscr{U}_{\text{ad}}$ let $V \subset (0, 1)$ be open with* $\text{meas}(V \cap I) > 0$ *and suppose that for any subinterval $\tilde{V}$ of V with* $\text{meas}(\tilde{V} \cap I) > 0$ *the function f is not a.e. zero on $\tilde{V} \cap I$. Further assume that $\lambda_2 = 0$. Then*

$$u(c^0) = z \text{ a.e. on } V \cap I.$$

In particular, if f is not a.e. zero on any subinterval of (0, 1) *and $|c^0| < \gamma$ then*

$$u(c^0) = z \text{ a.e. on } I.$$

Remark 4. 7. Suppose that f is not a.e. zero on any subinterval of (0, 1). Then the corollary shows the following: If $u(c^0) \neq z$, then the norm constraint must be active.

Remark 4. 8. Results similar to Theorem 4. 1 hold for other fit-to-data criteria as well. In the case of the point fit-to-data criterion (4. 11), for example, we replace alternative (A) by the following condition:

($\tilde{\text{A}}$) For any subinterval $\tilde{V}$ of (0, 1) with $\text{meas}(\tilde{V} \cap I) > 0$ the function f is not a.e. zero on $\tilde{V} \cap I$. Furthermore $u(c^0)(\tilde{x}) \neq z$ and almost all elements of a neighborhood of $\tilde{x}$ are in I.

Indeed, from Remark 4.4 we know that (4.7) and hence (4.12) hold. If λ_2 were 0, then $pu(c^0)=0$ a.e. on I. Assume that $p(\bar{x})\neq 0$ for some point of density $\bar{x}$ of I. Then $p(x)\neq 0$ for all $x\in\tilde{V}$, where $\tilde{V}$ is an interval with $\operatorname{meas}(\tilde{V}\cap I)>0$. Hence $u(c^0)=0$ a.e. on $\tilde{V}\cap I$ and therefore $f=0$ a.e. on $\tilde{V}\cap I$, which contradicts ($\tilde{\mathrm{A}}$). Thus $p=0$ a.e. on I and ($\tilde{\mathrm{A}}$) implies that $p=0$ in a neighborhood of $\bar{x}$. By (4.5A) this contradicts the assumption $u(c^0)(\bar{x})\neq z$. We conclude that $\lambda_2\neq 0$ and

$$\frac{pu(c^0)}{\lambda_2}=-c^0 \text{ a.e. on } I.$$

Therefore, if (4.11) is used as a fit-to-data criterion and ($\tilde{\mathrm{A}}$) holds, then for any open set $U\subset(0,1)$ with $\operatorname{meas}((U\cap I)-I)=0$ it follows that $c^0|U\in H^1(U)$.

In addition to the identification problem (OLS) over $\mathscr{U}_{\mathrm{ad}}$, we consider the following variant (OLS) over $\mathscr{U}^1_{\mathrm{ad}}$:

Minimize $\frac{1}{2}|u(c)-z|^2$ over $\mathscr{U}^1_{\mathrm{ad}}$, where

$$\mathscr{U}^1_{\mathrm{ad}}:=\{c\in H^1: c(x)\geqq\alpha,\ |c|_{H^1}\leqq\gamma\},$$

and $u(c)$ is the unique solution of (2.1).

In a completely analogous way to the procedure for (OLS) over $\mathscr{U}_{\mathrm{ad}}$ one can establish the existence of an optimal solution c^0 of (OLS) over $\mathscr{U}^1_{\mathrm{ad}}$ and show that this problem falls into the general framework of Section 3 be defining

$$W=H^0\times C\times\mathbb{R},\quad X=H^1,\quad Y=C\times\mathbb{R},\quad K=C_-\times\mathbb{R}_-\subset Y,$$

and

$$g: X\times W\to Y$$

by

$$g(c,w)=(g_1(c,w),g_2(c,w))=\left(\alpha-c,\frac{1}{2}(|c|^2_{H^1}-\gamma^2)\right).$$

Again we use the more suggestive notation

$$g_1(c,\alpha)=\alpha-c,$$

and

$$g_2(c,\gamma)=\frac{1}{2}(|c|^2_{H^1}-\gamma^2).$$

First we verify regularity of an optimal solution.

Lemma 4.2. *Every $c\in\mathscr{U}^1_{\mathrm{ad}}$ is regular in the sense of* (3.2).

Proof. We need to show that

$$(4.13)\qquad 0\in\operatorname{int}\{g(c)+g_x(c)H^1-C_-\times\mathbb{R}_-\}$$
$$=\operatorname{int}\left\{\left(\alpha-c-h+C_+,\frac{1}{2}(|c|^2_{H^1}-\gamma^2)+(c,h)_{H^1}+\mathbb{R}_+\right): h\in H^1\right\}.$$

As in the proof of Lemma 4. 1 consider $(\phi, r) \in C \times \mathbb{R}$ with

$$|(\phi, r)|_{C \times \mathbb{R}} < \delta,$$

and $\delta > 0$ to be chosen sufficiently small. We will show that

$$(\phi, r) \in g(c) + g_x(c)\, H^1 - C_- \times \mathbb{R}_- .$$

Note that the function $\Phi = \phi - \min \phi$ satisfies $\Phi \in C_+$. Hence, considering the first component in (4. 13) we decompose ϕ as

$$\phi = \alpha - c - (\alpha - c - \min \phi) + \Phi,$$

and therefore $\phi \in \alpha - c - H^1 + C_+$. The second component can be treated in a similar manner as in the proof of Lemma 4. 1.

We obtain the following first order optimality conditions for (OLS) over $\mathscr{U}^1_{\mathrm{ad}}$.

Proposition 4. 3. *For every local solution c^0 of* (OLS) *over $\mathscr{U}^1_{\mathrm{ad}}$ there exist $\lambda_1^* \in C_-^*$ and $\lambda_2 \in \mathbb{R}_-$ such that the unique solution p of $A(c^0)\,p = u(c^0) - z$ satisfies*

$$(4.\,14) \qquad (-u(c^0)\,p, h) + \langle \lambda_1^*, h \rangle_C - \lambda_2 (c^0, h)_{H^1} = 0$$

for all $h \in H^1$ and

$$(4.\,15) \qquad \langle \lambda_1^*, \alpha - c^0 \rangle_C = 0, \quad \lambda_2 (|c^0|^2_{H^1} - \gamma^2) = 0.$$

For (OLS) over $\mathscr{U}^1_{\mathrm{ad}}$ we can prove stronger local and even global smoothness results than over $\mathscr{U}_{\mathrm{ad}}$. First observe that a priori c^0 is in H^1 and hence continuous. Thus $I = \{x : c^0(x) > \alpha\}$ is open in [0, 1]. In the following theorem we do not assume that $I \neq \emptyset$.

Theorem 4. 2. *Let c^0 be a local solution of* (OLS) *over $\mathscr{U}^1_{\mathrm{ad}}$ and suppose that the alternative* (A) *holds for $V = I$. Then $c^0_x \in BV$, and in particular $c^0 \in H^{2-s}$ for $s \in \left(\frac{1}{2}, 1\right]$. Moreover, either $c^0(x) \equiv \alpha$ or $|c^0|_{H^1} = \gamma$. In the latter case $I \neq \emptyset$ and $c^0 \in H^3(I)$; if $0 \in I$, then $c^0_x(0) = 0$, if $1 \in I$, then $c^0_x(1) = 0$.*

Proof. Suppose first that $\lambda_2 = 0$. Then

$$(4.\,16) \qquad (-u(c^0)\,p, h) + \langle \lambda_1^*, h \rangle_C = 0, \quad \text{for all} \quad h \in H^1,$$

and

$$(4.\,17) \qquad \langle \lambda_1^*, \alpha - c^0 \rangle_C = 0.$$

Recall that since $\lambda_1^* \in C_-^*$ it can be represented as

$$\langle \lambda_1^*, h \rangle_C = \int_0^1 h(x)\, dv(x), \quad \text{for} \quad h \in C,$$

where v is of bounded variation and monotonically nonincreasing. If $I \neq \emptyset$, then (4. 17) implies that v is constant on each connected component of I. Therefore $pu(c^0) = 0$ on I by (4. 16). This leads to a contradiction to (A) as in the proof of Theorem 4. 1. Since $|c^0|_{H^1} < \gamma$ implies $\lambda_2 = 0$, the alternative $I = \emptyset$ or $|c^0|_{H^1} = \gamma$ holds.

Suppose now that $I \neq \emptyset$. Then necessarily $\lambda_2 \neq 0$ and for every $h \in H^1(I)$ we have

$$(pu(c^0), h)_{H^0(I)} + \lambda_2 (c^0, h)_{H^1(I)} = 0.$$

Thus (taking derivatives in the distributional sense on each of the countably many subintervals constituting I) we obtain

$$c^0 - c^0_{xx} = \lambda_2^{-1} pu(c^0) \text{ on } I.$$

This proves $c^0_{xx} \in H^1(I)$ or $c^0 \in H^3(I)$.

If $0 \in I$, let $\tau \in (0, 1)$ be such that $[0, \tau) \subset I$. Choose $h \in H^1$ with $h \equiv 0$ on $[\tau, 1]$ but otherwise arbitrary. Then

$$\begin{aligned} 0 &= (pu(c^0), h)_{H^0} + \lambda_2 (c^0, h)_{H^1} \\ &= (pu(c^0), h) + \lambda_2 (c^0, h) - \lambda_2 (c^0_{xx}, h) - \lambda_2 c^0_x(0)\, h(0). \end{aligned}$$

This implies $c^0_x(0) = 0$. Similarly, one can show $c^0_x(1) = 0$ if $1 \in I$.

It remains to prove that $c^0_x \in BV$. If $I = \emptyset$, then $c^0(x) = \alpha$ and nothing is to prove. Thus, by the first part of the proof, we can assume $\lambda_2 \neq 0$ and we have for all $h \in H^1$

$$(c^0, h)_{H^1} = -\frac{1}{\lambda_2}(pu(c^0), h) + \frac{1}{\lambda_2}\langle \lambda_1^*, h \rangle_C.$$

This implies

$$(4.18) \qquad (c^0_x, h_x) = -\frac{1}{\lambda_2}(pu(c^0), h) + \frac{1}{\lambda_2}\langle \lambda_1^*, h \rangle_C - (c^0, h).$$

The right hand side in (4.18) can be extended to a continuous linear functional on C, having the representation $\int h\, d\psi$, $h \in C$, and with $\psi \in BV$ and normalized in the usual manner. Therefore, for all $h \in H^1_0$ we obtain

$$(c^0_x, h_x) = \int h\, d\psi = -\int \psi h_x\, dx.$$

Thus $c^0_x = -\psi$ proving the assertion.

5. Output Least Squares stability

In this section, we study continuous dependence of the (not necessarily unique) solutions of the estimation problem on the observation z, the (upper) norm bound γ, and the (lower) pointwise bound α. This continuity property is called Output Least Squares (OLS-)stability and it will be analyzed using perturbation theory of optimization problems.

The sets $\mathscr{U}_{ad}$, $\mathscr{U}_{ad}^1$ and $\mathscr{U}_{ad}^N$ of admissible parameters have been defined in Section 4 (see Remark 4.5 and Lemma 4.2). Consider the output least squares identification problems:

$$(\mathrm{OLS})^w \qquad \text{Minimize } f(c, w) = \frac{1}{2}|u(c) - z|^2 \text{ over } c \in \mathscr{U}_{ad} \text{ or } \mathscr{U}_{ad}^1 \text{ or } \mathscr{U}_{ad}^N,$$

where $w = (z, \alpha, \gamma)$. The family of problems $(\mathrm{OLS})^w$, $w \in W$ are considered as perturbations of the reference problem $(\mathrm{OLS})^{w^0} = (\mathrm{OLS})$, where $w^0 = (z^0, \alpha^0, \gamma^0) \in W$. When calling upon the notation and results of Section 2 and 3 we replace (α, γ) of those sections by (α^0, γ^0).

The following definition should be compared with Chavent's notion of output least squares identifiability (OLSI) [7]. We do not assume uniqueness of the solutions c^0 of $(\mathrm{OLS})^{w^0}$ and consider stability with respect to the constraint set as well as the observations. Moreover, the stability concept of this paper requires only continuity at the fixed parameter value w^0 (instead of continuity for all z in a neighborhood of the attainable set).

Definition 5.1. The unknown parameter c is called Output Least Squares (OLS-)-stable in $\mathscr{U}_{ad}$ (resp. $\mathscr{U}_{ad}^1$ or $\mathscr{U}_{ad}^N$) at the local solution c^0 of $(\mathrm{OLS})^{w^0}$ if there exist a neighborhood V of w^0 in W, a constant $r > 0$, and a nondecreasing continuous real valued function ϱ with $\varrho(0) = 0$, such that for all $w = (z, \alpha, \gamma) \in V$ there exists a local solution c_w^0 of $(\mathrm{OLS})^w$ with $|c_w^0 - c^0| < r$ and for every such local solution c_w^0

$$|c_w^0 - c^0| \leqq \varrho(d(w, w^0)).$$

Remark 5.1. Before we analyse OLS-stability of the parameter c, we note that the *boundedness assumption* itself implies some weak continuous dependence of the solutions c_w^0 of $(\mathrm{OLS})^w$ on w. For let $(z^n, \alpha^n, \gamma^n) = w^n \to w^0 = (z^0, \alpha^0, \gamma^0)$ in $H^0 \times C \times \mathbb{R}$, with $\alpha^0 < \gamma^0$. Assume that there exists a unique solution c^0 $(= c_{w^0}^0)$ of $(\mathrm{OLS})^{w^0}$ and let $c_{w^n}^0$ denote solutions of $(\mathrm{OLS})^{w^n}$. Since $|c_{w^n}^0| \leqq \sup \gamma^n$, the set $|c_{w^n}^0|$ is bounded in H^0. Take any weakly convergent subsequence of $c_{w^n}^0$ with limit $\tilde{c}$. Then $\tilde{c}(x) \geqq \alpha^0$ a.e., and $|\tilde{c}| \leqq \gamma^0$. Moreover

$$|u(c_{w^n}^0) - z^n| \leqq |u(c) - z^n|, \quad \text{for all} \quad c \in \{c \in H^0 : c(x) \geqq \alpha^n, |c| \leqq \gamma^n\}$$

which, by Lemma 2.2 with $(\alpha, \gamma) = (\alpha^0, \gamma^0)$, implies

$$|u(\tilde{c}) - z^0| \leqq |u(c) - z^0|, \quad \text{for all} \quad c \in \{c \in H^0 : c(x) \geqq \alpha^0, |c| \leqq \gamma^0\}.$$

But c^0 is assumed to be the unique solution of $(\mathrm{OLS})^{w^0}$ and therefore $\tilde{c} = c^0$. The usual subsequence argument implies that $c_{w^n}^0 \rightharpoonup c^0$ as $n \to \infty$. If the solution of $(\mathrm{OLS})^{w^0}$ is not unique, then every subsequence of $c_{w^n}^0$ converges weakly to some solution of $(\mathrm{OLS})^{w^0}$.

On the other hand, OLS-stability of a parameter requires more: weak continuous dependence is replaced by strong continuous dependence and if c is OLS-stable at the local solution c^0 then the perturbed problems must have local solutions in a neighborhood of c^0 and these local solutions must depend continuously on the perturbation.

OLS-stability can only be obtained by imposing further assumptions on the problem data or by adding a regularization term to the fit-to-data criterion. The regularization approach will be taken in the next section.

In this section we obtain OLS-stability by either admitting only a finite dimensional parameter space (Theorem 5. 1 and 5. 2) or by exploiting the consequences of the norm bound defining the set $\mathscr{U}_{\mathrm{ad}}$ (Theorem 5. 3). First we consider the OLS-problem over the set

$$\mathscr{U}_{\mathrm{ad}}^N := \mathscr{U}_{\mathrm{ad}} \cap H_N.$$

Define the corresponding set of attainable observations $\mathscr{V}^N$ as

$$\mathscr{V}^N := \{u(c) : c \in \mathscr{U}_{\mathrm{ad}}^N\}.$$

Note that $\mathscr{V}^N \subset H^2$.

Theorem 5. 1. *Let* (H 1), (H 2) *hold and suppose that* $z^0 \in \mathscr{V}^N$, *i.e.* $z^0 = u(c^0)$ *for some element* $c^0 \in \mathscr{U}_{\mathrm{ad}}^N$. *Furthermore, assume that* $u(c^0)$ *does not vanish identically on any interval. Then the parameter* c *is* OLS-*stable in* $\mathscr{U}_{\mathrm{ad}}^N$ *at every such solution* c^0 *of* $(\mathrm{OLS})^{w^0}$.

Proof. As noticed in Remark 4. 5, the problem $(\mathrm{OLS})^w$ over $\mathscr{U}_{\mathrm{ad}}^N$ can be recast as a special case of the parameter dependent optimization problem $(\mathrm{P})^w$ of Section 3. Let c^0 satisfy the requirements of the theorem. Then, by Remark 3. 2, conditions (3. 8) and (3. 9) in Theorem 3. 3 hold, since f and g are continuously Fréchet differentiable. Regularity of c^0 holds by Remark 4. 5. We turn to estimate the second derivative of the Lagrangian from below.

That is, we have to consider the expression

$$F_{cc}(c^0, w^0)(h, h) = |\eta|^2 + (u(c^0) - z^0, \xi) - \lambda_2 |h|^2$$

where $\xi = u_{cc}(c^0)(h, h)$ is given by $A(c^0)\xi = -2h\eta$ and $\eta = u_c(c^0)h$ is given by $A(c^0)\eta = -hu(c^0)$. By assumption $z^0 \in \mathscr{V}^N$ and by Lemma 2. 7 we obtain

$$\begin{aligned} F_{cc}(c^0, w^0)(h, h) &= |\eta|^2 - \lambda_2 |h|^2 \\ &\geqq |\eta|^2 \geqq k_1^{-2} |hu(c^0)|_{\tilde{H}^{-2}}^2. \end{aligned}$$

Since $\operatorname{meas}\{x : u(c^0)(x) = 0\} = 0$, $|hu(c^0)|_{\tilde{H}^{-2}}$ defines a norm for $h \in H_N$. This norm is equivalent to the H^0-norm, because H_N is finite dimensional. This implies

$$F_{cc}(c^0, w^0)(h, h) \geqq \delta |h|_{H^0}^2$$

for an appropriately defined constant $\delta > 0$ independent of $h \in H_N$. Therefore all the hypotheses of Theorem 3. 3 are satisfied and the proof is complete.

Remark 5. 2. In Theorem 5. 1, the function ϱ in the Definition 5. 1 of OLS-stability is given by $\varrho(x) = d\sqrt{x}$, $d > 0$, since ϱ is obtained by an application of Theorem 3. 3. The same remark holds true for all other results on OLS-stability in this paper.

We can obtain OLS-stability also in certain cases, where z is not in the set $\mathscr{V}^N$ of reachable observations. For this purpose, we prepare the following lemma.

Lemma 5. 1. *Let* $c \in \mathscr{U}_{\mathrm{ad}}^N$. *If* $u(c) > 0$ *on* $[0, 1]$, *then for all* $h \in H_N$ *and*

$$\eta = A^{-1}(c)(hu(c)), \quad \xi = -2A^{-1}(c)(h\eta)$$

the following inequality holds:

$$|\eta|^2 + (u(c) - z^0, \xi) \geqq |\eta| \left(k_2^{-1} |hu(c)|_{\tilde{H}^{-2}} - 2k_1^{-1} |u(c) - z^0| \, |h|_{L^\infty}\right).$$

Proof. We compute, employing Lemma 2. 6

$$\begin{aligned}|\eta|^2+(u(c)-z^0,\xi) &\geqq |\eta|^2-2(A^{-1}(c)(u(c)-z^0), h\eta)\\ &\geqq |\eta|^2-2k_1^{-1}|u(c)-z^0|_{\tilde H^{-2}}|h\eta|\\ &\geqq |\eta|^2-2k_1^{-1}|u(c)-z^0|\,|h|_{L^\infty}|\eta|\\ &\geqq |\eta|\,(k_2^{-1}|hu(c)|_{\tilde H^{-2}}-2k_1^{-1}|u(c)-z^0|\,|h|_{L^\infty})\end{aligned}$$

for all $h\in H_N$. This is the desired inequality.

Theorem 5. 2. *Let $c^0\in\mathcal{U}_{ad}^N$ be a local solution of* (OLS)$^{w^0}$ *and suppose that $u(c^0)>0$ or $u(c^0)<0$ on* [0, 1]. *Choose $\kappa>0$ with $|hu(c^0)|_{L^\infty}\leqq\kappa|hu(c^0)|_{\tilde H^{-2}}$, and suppose*

$$|u(c^0)-z|\leqq\frac{1}{2}k_1k_2^{-1}\kappa^{-1}\min_x|u(c^0)(x)|.$$

Then c is OLS-*stable in $\mathcal{U}_{ad}^N$ at the solution c^0 of* (OLS)$^{w^0}$.

Proof. By our assumptions $|hu(c^0)|_{L^\infty}$ and $|hu(c^0)|_{\tilde H^{-2}}$ define norms on H_N. Since all norms on H_N are equivalent, κ exists. By Remark 4. 5 $c^0\in\mathcal{U}_{ad}^N$ is regular. For the second derivative of the Lagrangian we obtain by Lemma 5. 1

$$\begin{aligned}F_{cc}(c^0,w^0)(h,h) &= |\eta|^2+(u(c^0)-z^0,\xi)-\lambda_2|h|^2\\ &\geqq |\eta|\,(k_2^{-1}|hu(c^0)|_{\tilde H^{-2}}-2k_1^{-1}|u(c^0)-z^0|\,|h|_{L^\infty})\\ &\geqq |\eta|\,(k_2^{-1}\kappa^{-1}|hu(c^0)|_{L^\infty}-2k_1^{-1}|u(c^0)-z^0|\,|h|_{L^\infty})\\ &\geqq |\eta|\,|h|_{L^\infty}(k_2^{-1}\kappa^{-1}\min_x u(c^0)(x)-2k_1^{-1}|u(c^0)-z^0|)\\ &\geqq \tilde\delta|h|^2\end{aligned}$$

for some $\tilde\delta>0$ independent of *h*. Hence, as in the proof of Theorem 5. 1, the assertion follows from Theorem 3. 3.

Theorem 5. 3. *Suppose that c^0 is a local solution of* (OLS)$^{w^0}$ *in $\mathcal{U}_{ad}^N$ with $u(c^0)=z$ and that the assumptions of Theorem* 4.1 *are satisfied. Then the parameter c is* OLS-*stable in $\mathcal{U}_{ad}^N$ at c^0.*

Proof. By Theorem 4. 1 one has $\lambda_2<0$. Again we appeal to Theorem 3. 3 and it only remains to establish the lower bound on F_{cc}. We compute

$$\begin{aligned}F_{cc}(c^0,w^0) &= |\eta|^2+(u(c^0)-z,\xi)-\lambda_2|h|^2\\ &\geqq -\lambda_2|h|^2.\end{aligned}$$

Since $\lambda_2<0$ the assertion follows.

6. Output Least Squares stability by regularization

In the last section we imposed further assumptions on the problem data in order to establish OLS-stability. An alternative is to add a regularization term to the fit-to-data criterion (see e.g. [11], [12]). Then this regularized problem is analyzed with respect to continuous dependence on the observation *z* and the constraints defining the set of admissible parameters.

For $\beta>0$ we define the regularized output least squares minimization problem as:

$$\text{(ROLS)} \qquad \text{minimize } \frac{1}{2}|u(c)-z|^2+\beta|c|^2 \text{ over } c\in\tilde{\mathscr{U}}_{ad};$$

here $\tilde{\mathscr{U}}_{ad}=\{c\in H^0: c(x)\geqq\alpha \text{ a.e.}\}$. Of course, α is chosen so that (H 1) holds. It is simple to see that for each $\beta>0$ there exists a solution c^β of (ROLS) in $\tilde{\mathscr{U}}_{ad}$. At times we shall write $(\text{ROLS})_\beta$ to specify the value of the regularization parameter in (ROLS).

In the previous section the existence of a solution c^0 of (OLS) was guaranteed by the norm constraint involved in defining the set $\mathscr{U}_{ad}$. In contrast, in this section we make the following assumption:

(H 3) There exists a minimizer c^0 of $\frac{1}{2}|u(c)-z|^2$ over $c\in\tilde{\mathscr{U}}_{ad}$.

Thus (H 1), (H 2) and (H 3) are assumed to hold throughout this section. Without loss of generality we also assume throughout that a minimum-norm solution c^0 of (OLS) in $\tilde{\mathscr{U}}_{ad}$ satisfies $|c^0|\leqq\gamma$ so that the estimates of Lemma 2.1 are applicable. We again refer to the optimization problem in (H 3) as (OLS) and we define

$$\tilde{\mathscr{U}}_{ad}^N=\tilde{\mathscr{U}}_{ad}\cap H_N,\quad \tilde{\mathscr{V}}=\{u(c):c\in\tilde{\mathscr{U}}_{ad}\}\quad\text{and}\quad \tilde{\mathscr{V}}^N=\{u(c):c\in\tilde{\mathscr{U}}_{ad}^N\}.$$

The set of perturbation parameters is now taken in the space

$$\tilde{W}=H^0\times C.$$

Of course, (H 3) holds, if $\tilde{\mathscr{U}}_{ad}$ is replaced by $\mathscr{U}_{ad}$ and in this case all the results of this section remain correct, if the norm bound γ is added to the set of perturbation parameters.

We first discuss the asymptotic behaviour of the global solutions c^β as $\beta\to 0$. Let c^0 be any solution of (OLS) over $\tilde{\mathscr{U}}_{ad}$. Then clearly, $|u(c^0)-z|=\operatorname{dist}(z,\tilde{\mathscr{V}})$. This is generally not true for solutions c^β of (ROLS).

Lemma 6.1. *For all $\beta>\beta_0\geqq 0$ the following assertions hold:*

(a) $\sup\{|c^\beta|\}\leqq\inf\{|c^{\beta_0}|\}$,

(b) $\sup\{|u(c^{\beta_0})-z|\}\leqq\inf\{|u(c^\beta)-z|\}$,

(c) $\sup\left\{\frac{1}{2}|u(c^\beta)-z|^2+\beta|c^\beta|^2\right\}\leqq\frac{1}{2}\operatorname{dist}(z,\tilde{\mathscr{V}})^2+\beta\inf|c^0|^2$;

here the supremum and the infimum are taken over all global solutions in $\tilde{\mathscr{U}}_{ad}$ of (ROLS) *or* (OLS), *for $\beta\neq 0$ or $\beta=0$, respectively.*

Proof. For all solutions c^{β_0} and c^{β} we have

$$\frac{1}{2}|u(c^{\beta_0})-z|^2+\beta_0|c^{\beta_0}|^2 \leqq \frac{1}{2}|u(c^{\beta})-z|^2+\beta_0|c^{\beta}|^2.$$

Addition of $(\beta-\beta_0)|c^{\beta}|^2$ yields by definition of c^{β}

$$\frac{1}{2}|u(c^{\beta_0})-z|^2+\beta|c^{\beta}|^2+\beta_0(|c^{\beta_0}|^2-|c^{\beta}|^2) \tag{6.1}$$
$$\leqq \frac{1}{2}|u(c^{\beta})-z|^2+\beta|c^{\beta}|^2 \leqq \frac{1}{2}|u(c^{\beta_0})-z|^2+\beta|c^{\beta_0}|^2.$$

Estimating the first by the last term in (6.1) we obtain

$$\beta(|c^{\beta}|^2-|c^{\beta_0}|^2) \leqq \beta_0(|c^{\beta}|^2-|c^{\beta_0}|^2).$$

Since $\beta_0 \leqq \beta$, this implies $|c^{\beta}| \leqq |c^{\beta_0}|$ and therefore (a) is verified. Again by (6.1) we find

$$\frac{1}{2}|u(c^{\beta_0})-z|^2+\beta_0|c^{\beta_0}|^2 \leqq \frac{1}{2}|u(c^{\beta})-z|^2+\beta_0|c^{\beta}|^2,$$

and consequently, using (a) we have

$$\frac{1}{2}|u(c^{\beta_0})-z|^2-\frac{1}{2}|u(c^{\beta})-z|^2 \leqq \beta_0(|c^{\beta}|^2-|c^{\beta_0}|^2) \leqq 0.$$

Therefore (b) holds. Finally (c) follows from the second inequality in (6.1) with $\beta_0=0$.

In the following lemma (H3) is needed essentially.

Lemma 6.2. (a) *Let $\beta_n \to \beta_0 \geqq 0$ and let c^{β_n} be any sequence of corresponding solutions in $\tilde{\mathscr{U}}_{ad}$ of* $(\mathrm{ROLS})_{\beta_n}$. *Then c^{β_n} has a weak limit point and every weak limit point of c^{β_n} is a solution of* $(\mathrm{ROLS})_{\beta_0}$.

(b) *Let $\beta_n \to \beta_0^+ \geqq 0$ and let c^{β_n} be any sequence of solutions in $\tilde{\mathscr{U}}_{ad}$ of* $(\mathrm{ROLS})_{\beta_n}$. *Then there exists a convergent subsequence of c^{β_n}, and every convergent subsequence converges to a minimum-norm solution of* $(\mathrm{ROLS})_{\beta_0}$.

Proof. (a) By Lemma 6.1(a) and (H3) the set $\{|c^{\beta_n}|\}$ is bounded and thus there exists a weakly convergent subsequence $c^{\beta_{n_k}} \rightharpoonup \hat{c} \in \tilde{\mathscr{U}}_{ad}$. For all $c \in \tilde{\mathscr{U}}_{ad}$ we have

$$\frac{1}{2}|u(c^{\beta_n})-z|^2+\beta_n|c^{\beta_n}|^2 \leqq \frac{1}{2}|u(c)-z|^2+\beta_n|c|^2.$$

By weak lower semicontinuity of the norm and Lemma 2.2 we have

$$\frac{1}{2}|u(\hat{c})-z|^2+\beta_0|\hat{c}|^2 \leqq \frac{1}{2}|u(c)-z|^2+\beta_0|c|^2,$$

for all $c \in \tilde{\mathscr{U}}_{ad}$. Thus (a) is proved. Next, let $\beta_n \to \beta_0^+$ and let $c^{\beta_{n_k}}$ be any weakly convergent subsequence of c^{β_n} with $c^{\beta_{n_k}} \rightharpoonup \hat{c} \in \tilde{\mathscr{U}}_{ad}$. Then by Lemma 6.1(a) and lower semicontinuity of the norm

$$\limsup_{n_k} |c^{\beta_{n_k}}| \leqq |\hat{c}| \leqq \liminf_{n_k} |c^{\beta_{n_k}}|.$$

This implies $c^{\beta_{n_k}} \to \hat{c}$. If there were a solution c^{β_0} of $(\text{ROLS})_{\beta_0}$ with $|c^{\beta_0}| < |\hat{c}|$, then

$$\limsup_{n_k} |c^{\beta_{n_k}}| \leqq |c^{\beta_0}| < |\hat{c}| \leqq \liminf_{n_k} |c^{\beta_{n_k}}|,$$

which is impossible and (b) is verified.

Lemma 6. 3. (a) $\lim_{\beta \to 0^+} \beta^{-1}(\sup\{|u(c^\beta) - z|^2\} - \text{dist}(z, \tilde{\mathscr{V}})^2) = 0$,

(b) *if* $z \in \tilde{\mathscr{V}}$, *then* $\sup\{|u(c^\beta) - z|\} = o(\sqrt{\beta})$,

where again the sup *is taken over all solutions* c^β *of* $(\text{ROLS})_\beta$.

Proof. From Lemma 6. 1(b) and (6. 1) we obtain for each solution c^β and c^0

$$0 \leqq |u(c^\beta) - z|^2 - \text{dist}(z, \tilde{\mathscr{V}})^2 \leqq 2\beta(|c^0|^2 - |c^\beta|^2). \tag{6. 2}$$

Assume that there exists $\delta > 0$ and a sequence of solutions c^{β_n} of $(\text{ROLS})_{\beta_n}$, $\beta_n \to 0$, with

$$\beta_n^{-1}(|u(c^{\beta_n}) - z|^2 - \text{dist}(z, \tilde{\mathscr{V}})^2) \geqq \delta > 0. \tag{6. 3}$$

By Lemma 6. 2(b) there exists a subsequence of $\{c^{\beta_n}\}$ converging to a solution c^0 of (OLS). Thus (6. 3) contradicts (6. 2) and (a) is proved. Part (b) is an obvious consequence of (a).

We now investigate further the relationship between solutions of (OLS) and (ROLS). Let C_0 be the set of all minimum-norm solutions of (OLS) and C_β the set of all solutions in $\tilde{\mathscr{U}}_{\text{ad}}$ of $(\text{ROLS})_\beta$. Then, assuming (H 3), C_0 and C_β are nonempty and weakly closed. In general, they are not connected. Recall that a set is connected if it cannot be expressed as the union of two nonempty, relatively open disjoint sets. A connected component is a maximal connected subset. The following result establishes that every connected component of C_0 in the weak topology is approximated by solutions of $(\text{ROLS})_\beta$. In this sense no minimum-norm solutions of (OLS) get lost when regularization is introduced.

Proposition 6. 1. *Let M be a connected component of C_0, when C_0 is endowed with the weak topology of H^0. Then for every sequence $\beta_n \to 0$, $\beta_n > 0$, there exists a corresponding subsequence of local solutions $c^{\beta_{n_k}}$ of $(\text{ROLS})_{\beta_{n_k}}$ in $\tilde{\mathscr{U}}_{\text{ad}}$ converging in H^0 to an element of M as $n_k \to \infty$.*

Proof. Observe that C_0 is weakly compact and hence $C_0 \setminus M$ is weakly compact as well. One can show that then exists a weakly closed set V, which contains C in its strongly interior and which satisfies $V \cap (C_0 \setminus M) = \emptyset$. Now consider for $\beta > 0$ the problems $(\text{P})_\beta$ minimize $\frac{1}{2}|u(c) - z|^2 + \beta|c|^2$ over $\tilde{\mathscr{U}}_{\text{ad}} \cap V$. There exist solutions c^β of $(\text{P})_\beta$. Take a sequence $\{\beta_n\}$ with $\beta_n \to 0$. As in Lemma 6. 2(b) there exists a norm convergent subsequence of solutions $c^{\beta_{n_k}}$ of $(\text{P})_{\beta_{n_k}}$ converging to a minimum norm solution of $(\text{P})_0$. By choice of V we have $c^0 \in M$ and moreover $c^{\beta_{n_k}}$ is a solution of $(\text{ROLS})_{\beta_{n_k}}$ for all β_{n_k} sufficiently large. This ends the proof.

Next we turn to the central question of stability under parameter perturbations. As in Section 5 we employ a superscript w to denote the value of the perturbation parameter $w = (z, \alpha) \in \tilde{W} = H^0 \times C$. Thus (ROLS) becomes $(\text{ROLS})^w$. The unperturbed problem is $(\text{ROLS})^{w^0}$, with $w^0 = (z^0, \alpha^0)$ and α^0 so that (H 1) holds. The next definition formally specifies the property that will be the focus of our investigations.

Definition 6. 1. The unknown parameter c in (OLS) is called *Output Least Squares Stable by Regularization* (ROLS-stable) in $\tilde{\mathscr{U}}_{\text{ad}}$ (resp. $\tilde{\mathscr{U}}^1_{\text{ad}}$ or $\tilde{\mathscr{U}}^N_{\text{ad}}$) at $w^0 \in \tilde{W}$ for $\beta \in J \subset (0, \infty)$, if for every (global) solution $c^\beta_{w^0}$ of $(\text{ROLS})^{w^0}_\beta$ with $\beta \in J$ there exists a neighborhood V of w^0 in $\tilde{W}$, a constant $r > 0$ and a continuous nondecreasing real valued function ϱ with $\varrho(0) = 0$ such that for all $w \in V$ there exists a local solution c^β_w of $(\text{ROLS})^w_\beta$ with $|c^\beta_w - c^\beta_{w^0}| < r$ and for every such local solution c^β_w

$$|c^\beta_w - c^\beta_{w^0}| \leqq \varrho(|w - w^0|_{\tilde{W}}).$$

Remark 6. 1. Similar to $(\text{OLS})^w_1$ we define $(\text{ROLS})^w_1$ by replacing the H^0-norm in the fit-to-data criterion of (ROLS) by the H^1-norm. Correspondingly $(\text{ROLS})_1$-stability is defined in an analogous manner to ROLS-stability with $w \in W^1 = H^1 \times C$.

The following theorem is an analogon to Theorem 5. 2. In reference to the earlier development of this section we have to replace (OLS) by $(\text{OLS})^{w^0}$.

Theorem 6. 1. *Assume that there exists a solution of* $(\text{OLS})^{w^0}$ *in* $\tilde{\mathscr{U}}^N_{\text{ad}}$ *and that for a solution* c^0 *of minimum norm* $|u(c^0)(x)| \geqq \mu > 0$ *on* $[0, 1]$. *Choose* $\kappa > 0$ *such that* $|hu(c^0)|_{L^\infty} \leqq \kappa |hu(c^0)|_{\tilde{H}^{-2}}$ *for all* $h \in H_N$. *If* $\operatorname{dist}(z^0, \tilde{\mathscr{V}}^N) = |u(c^0) - z| < \frac{1}{4} k_1 k_2^{-1} \kappa^{-1} \mu$, *then there exists* $\tilde{\beta} > 0$ *such that* c *is* ROLS-*stable in* $\tilde{\mathscr{U}}^N_{\text{ad}}$ *for* $\beta \in (0, \tilde{\beta})$.

Proof. Note first, that Lemma 6. 1 and Lemma 6. 2 remain correct with $\tilde{\mathscr{U}}_{\text{ad}}$ replaced by $\tilde{\mathscr{U}}^N_{\text{ad}}$. Lemma 6. 2 implies that there exists a $\beta_1 > 0$ such that either $u(c^\beta) \geqq \frac{\mu}{2}$ or $u(c^\beta) \leqq -\frac{\mu}{2}$ on $[0, 1]$ for $\beta \in (0, \beta_1)$. Since every solution c^β is regular, it remains to estimate F_{cc} from below. Recall that $|c^\beta| \leqq |c^0|$ for $\beta > 0$ and that by assumption $|c^0| \leqq \gamma$ for a minimum norm solution of $\tilde{\mathscr{U}}^N_{\text{ad}}$. By Lemma 5. 1 we obtain for $\eta = u_c(c^\beta)(h)$, $\xi = u_{cc}(c^\beta)(h, h)$ and $h \in \tilde{\mathscr{U}}^N_{\text{ad}}$:

$$\begin{aligned} F_{cc}(c^\beta, w^0)(h, h) &= |\eta|^2 + (u(c^\beta) - z^0, \xi) + \beta |h|^2 - \lambda_2 |h|^2 \\ &\geqq |\eta|\, |h|_{L^\infty} \left(k_2^{-1} \kappa^{-1} \frac{\mu}{2} - 2 k_1^{-1} |u(c^\beta) - z^0| \right) + \beta |h|^2 \geqq \beta |h|^2, \end{aligned}$$

provided that

$$k_2^{-1} \mu \kappa^{-1} \geqq 4 k_1^{-1} |u(c^\beta) - z^0|, \tag{6. 4}$$

for all solutions c^β of $(\text{ROLS})^{w^0}_\beta$. By Lemma 6. 1 (c)

$$|u(c^\beta) - z^0|^2 \leqq \operatorname{dist}(z^0, \tilde{\mathscr{V}})^2 + 2\beta |c^0|^2.$$

Hence (6. 4) holds, if $\operatorname{dist}(z^0, \tilde{\mathscr{V}})^2 + 2\beta |c^0|^2 \leqq \left(\frac{1}{4} k_1 k_2^{-1} \mu \kappa^{-1} \right)^2$. Finally (6. 4) holds for $\beta \leqq \beta_2$, where

$$\beta_2 = \frac{\left(\frac{1}{4} k_1 k_2^{-1} \mu \kappa^{-1} \right)^2 - \operatorname{dist}(z^0, \tilde{\mathscr{V}})^2}{2 |c^0|^2}.$$

The assertion of the theorem thus follows for $\tilde{\beta} = \min(\beta_1, \beta_2)$.

Recall from Lemma 6.2 that $\lim_{\beta \to 0^+} |c^\beta| = |c^0|$, where c^0 is a minimum-norm solution of $(\mathrm{OLS})^{w^0}$ and c^β is any solution of $(\mathrm{ROLS})_\beta^{w^0}$.

Theorem 6.2. *Suppose that $\bar\beta > 0$ is chosen such that for a minimum norm solution c^0 of* $(\mathrm{OLS})^{w^0}$

$$|c^0|^2 - \sup |c^{\bar\beta}|^2 < k_1^2,$$

and define

$$\underline{\beta} = \tfrac{1}{2} \operatorname{dist}(z^0, \tilde{\mathscr{V}})^2 (k_1^2 - |c^0|^2 + \sup |c^{\bar\beta}|^2)^{-1} \geqq 0,$$

where the supremum is taken over all solutions $c^{\bar\beta}$ of $(\mathrm{ROLS})_{\bar\beta}^{w^0}$ *in $\tilde{\mathscr{U}}_{\mathrm{ad}}$.*

If $\underline{\beta} < \bar\beta$, then the parameter c is ROLS-*stable at w^0 in $\tilde{\mathscr{U}}_{\mathrm{ad}}$ for all $\beta \in (\underline{\beta}, \bar\beta)$. If $z^0 \in \tilde{\mathscr{V}}$, then c is* ROLS-*stable in $\tilde{\mathscr{U}}_{\mathrm{ad}}$ for all $\beta \in (0, \bar\beta)$.*

Proof. Arguing as in the previous results we only show how to obtain the lower bound on F_{cc}. Let $\eta = u_c(c^\beta) h$ and $\xi = u_{cc}(c^\beta)(h, h)$ with $h \in H^0$. Then

$$\begin{aligned} F_{cc}(c^\beta, w^0)(h, h) &= |\eta|^2 + (u(c^\beta) - z^0, A^{-1}(c^\beta)(-2h\eta)) + 2\beta |h|^2 - \lambda_2 |h|^2 \\ &\geqq |\eta|^2 - 2(hA^{-1}(c^\beta)(u(c^\beta) - z^0), \eta) + 2\beta |h|^2 \\ &\geqq 2\beta |h|^2 - |hA^{-1}(c^\beta)(u(c^\beta) - z^0)|^2 \\ &\geqq 2\beta |h|^2 - |A^{-1}(c^\beta)(u(c^\beta) - z^0)|_{H^2}^2 |h|^2 \\ &\geqq |h|^2 (2\beta - k_1^{-2} |u(c^\beta) - z^0|^2), \end{aligned}$$

by selfadjointness of A^{-1} and Lemma 2.1; here we use the assumption that $|c^0| \leqq \gamma$, which by Lemma 6.1(a) implies $|c^\beta| \leqq \gamma$. Thus we have for all solutions c^β of $(\mathrm{ROLS})_\beta^{w^0}$

$$F_{cc}(c^\beta, w^0)(h, h) \geqq |h|^2 (2\beta - k_1^{-2} |u(c^\beta) - z^0|^2). \tag{6.5}$$

Now we use Lemma 6.1(c):

$$\begin{aligned} F_{cc}(c^\beta, w^0)(h, h) &\geqq |h|^2 k_1^{-2} [2k_1^2 \beta - \operatorname{dist}(z^0, \tilde{\mathscr{V}})^2 - 2\beta(|c^0|^2 - |c^\beta|^2)] \\ &= |h|^2 k_1^{-2} [2\beta(k_1^2 - |c^0|^2 + |c^\beta|^2) - \operatorname{dist}(z^0, \tilde{\mathscr{V}})^2]. \end{aligned}$$

Note that by assumption $k_1^2 - |c^0|^2 + |c^\beta|^2 > 0$ for $\beta \in (0, \bar\beta)$. Thus for $z^0 \in \tilde{\mathscr{V}}$ the theorem is proved. Otherwise we observe that

$$2\beta(k_1^2 - |c^0|^2 + |c^\beta|^2) - \operatorname{dist}(z^0, \tilde{\mathscr{V}})^2 = \operatorname{dist}(z^0, \tilde{\mathscr{V}})^2 (\beta \underline{\beta}^{-1} - 1) > 0,$$

and again the desired estimate on F_{cc} is obtained.

Remark 6.2. If $|c^0|^2 < k_1^2$, then one can choose $\bar\beta$ arbitrarily. In this case c is ROLS-stable at w^0 in $\tilde{\mathscr{U}}_{\mathrm{ad}}$ for all $\beta \in (\underline{\beta}, \infty)$ with $\underline{\beta} = \tfrac{1}{2} \operatorname{dist}(z^0, \tilde{\mathscr{V}})^2 (k_1^2 - |c^0|^2)^{-1}$.

Remark 6.3. Similarly suppose that $|c^0|^2 - \alpha^2 < k_1^2$. Then $|c^0|^2 - |c^\beta|^2 < k_1^2$ for all solutions c^β of $(\mathrm{ROLS})_\beta^{w^0}$, $\beta > 0$ and $\bar{\beta}$ in Theorem 6.2 can be chosen arbitrarily. The parameter c is ROLS-stable at w^0 in $\tilde{\mathscr{U}}_{\mathrm{ad}}$ for $\beta \in (\underline{\beta}, \infty)$ with

$$\underline{\beta} = \tfrac{1}{2}\,\mathrm{dist}(z^0, \tilde{\mathscr{V}})^2\,(k_1^2 - |c^0|^2 + \alpha^2)^{-1}.$$

Remark 6.4. Let

$$C^0 = \{c^0\colon \text{there exist solutions } c^\beta \text{ of } (\mathrm{ROLS})_\beta^{w^0} \text{ such that } \lim_{\beta \to 0} c^\beta = c^0\}$$

and let $z^0 \in \tilde{\mathscr{V}}$. Then Lemma 6.2(b) and Theorem 6.2 clearly imply that for every minimum-norm solution $c^0 \in C^0$ and every $\varepsilon > 0$ there exist a ROLS-stable solution c^β with $|c^0 - c^\beta| < \varepsilon$. If moreover C^0 consists of only one element c^0, then $\lim_{\beta \to 0} c^\beta = c^0$ (where c^β is any solution of $(\mathrm{ROLS})_\beta^{w^0}$) and c^β is ROLS-stable for all β sufficiently small.

The condition $\underline{\beta} < \bar{\beta}$ of Theorem 6.2 in case that $z \notin \tilde{\mathscr{V}}$ constitutes a certain relationship between the convergence rate of c^β to c^0, the bound k_1 and $\mathrm{dist}(z^0, \tilde{\mathscr{V}})$: fast convergence rates of $|c^\beta|$, large bounds k_1 and small distances $\mathrm{dist}(z^0, \tilde{\mathscr{V}})$ are favorable.

If the condition $\underline{\beta} < \bar{\beta}$ is violated, a natural idea is to try to get a better observation z^0, i.e. to lower $\mathrm{dist}(z^0, \tilde{\mathscr{V}})$. Next we show that this is a reasonable strategy which—at least theoretically—leads to success.

Let us first introduce some additional notation. For $w_n^0 = (z_n^0, \alpha^0)$, with $z_n^0 \in H^0$, $n = 0, 1, \ldots$, and α^0 such that (H 1) holds, we denote the solutions of $(\mathrm{OLS})^{w_n^0}$ and $(\mathrm{ROLS})_\beta^{w_n^0}$ by c_n^0 and c_n^β respectively.

Theorem 6.3. *Let $z_n^0 \to z_0^0$ in H^0, with $z_0^0 \in \tilde{\mathscr{V}}$, and assume that minimum-norm solutions c_n^0 of $(\mathrm{OLS})^{w_n}$, with $w_n^0 = (z_n^0, \alpha^0)$ exist with $\sup\{|c_n^0|\colon n = 0, 1, 2, \ldots\} \leqq \gamma$. Then there exists $\tilde{\beta} > 0$ with the following property:*

For all $\beta^* \in (0, \tilde{\beta})$ there exists an $N(\beta^*)$ and a neighborhood $J(\beta^*)$ of β^* such that for all $n \geqq N(\beta^*)$ the parameter c is ROLS-stable in $\tilde{\mathscr{U}}_{\mathrm{ad}}$ at $w_n^0 = (z_n^0, \alpha^0)$ for $\beta \in J(\beta^*)$.

Proof. As in (6.5) we obtain for $n = 0, 1, \ldots$

$$F_{cc}(c_n^\beta, w_n^0) \geqq |h|^2\,(2\beta - \sup k_1^{-2}\,|u(c_n^\beta) - z_n^0|^2), \tag{6.6}$$

where the supremum is taken over all solutions c_n^β of $(\mathrm{ROLS})_\beta^{w_n^0}$. For β^* sufficiently small we need to bound (6.6) from below uniformly in $\beta \in J(\beta^*)$ and $n \geqq N(\beta^*)$. By Lemma 6.3(b) we can choose $\tilde{\beta}$ so that for every $\beta^* \in (0, \tilde{\beta})$ there exists $\varepsilon = \varepsilon(\beta^*) > 0$ with

$$2\beta^* - \sup k_1^{-2}\,|u(c_0^{\beta^*}) - z_0^0|^2 = \varepsilon, \tag{6.7}$$

where the supremum is taken over all solutions $c_0^{\beta^*}$ of $(\mathrm{ROLS})_{\beta^*}^{w_0^0}$.

First we will show that for all $\beta^* \in (0, \tilde{\beta})$ there exists an $N(\beta^*)$ such that

$$2\beta^* - \sup k_1^{-2}\,|u(c_n^{\beta^*}) - z_n^0|^2 \geqq \frac{\varepsilon}{3} \quad \text{for all} \quad n \geqq N(\beta^*). \tag{6.8}$$

If (6. 8) were false, then there would exist a sequence n_k with $\lim n_k = \infty$ and solutions $c_{n_k}^{\beta^*}$ such that

$$(6.9) \qquad 2\beta^* - k_1^{-2}|u(c_{n_k}^{\beta^*}) - z_{n_k}^0|^2 < \frac{\varepsilon}{2},$$

By Lemma 6. 1(a) and the assumption on the boundedness of $|c_n^0|$ it follows that $\{|c_{n_k}^{\beta^*}|\}$ is bounded. Therefore there exists a weakly convergent subsequence of $c_{n_k}^{\beta^*}$, again denoted by $c_{n_k}^{\beta^*}$, with weak limit $c_0^{\beta^*}$. It can easily be shown that $c_0^{\beta^*}$ is a solution of $(\mathrm{ROLS})_{\beta^*}^{w_0^0}$. Then, taking the limit in (6. 9), we obtain $2\beta^* - k_1^{-2}|u(c_0^{\beta^*}) - z_0^0|^2 \leqq \frac{\varepsilon}{2}$, which contradicts (6. 7) and (6. 8) is verified.

Next we show that there exists a neighborhood $J(\beta^*)$ of $\beta^* \in (0, \tilde{\beta})$, such that

$$(6.10) \qquad 2\beta - \sup k_1^{-2}|u(c_n^\beta) - z_n^0|^2 \geqq \frac{\varepsilon}{5} \quad \text{for all} \quad n \geqq n_0(\beta^*) \quad \text{and} \quad \beta \in J(\beta^*).$$

If (6. 10) were not true, then there would exist sequences β_m and n_m, with $\beta_m \to \beta^*$ and $n_m \geqq n_0(\beta^*)$, and solutions $c_{n_m}^{\beta_m}$ of $(\mathrm{ROLS})_{\beta_m}^{w_{n_m}^0}$ such that

$$(6.11) \qquad 2\beta_m - k_1^{-2}|u(c_{n_m}^{\beta_m}) - z_{n_m}^0|^2 < \frac{\varepsilon}{4}.$$

Concerning the index n_m we have to consider two cases: either $\lim n_m = \infty$ or infinitely many n_m assume the same value. We consider the second case first and assume without loss of generality $n_m = \tilde{n}$ for all m. Again, $\{|c_{\tilde{n}}^{\beta_m}|\}_{m=1}^{\infty}$ is a bounded set and by Lemma 6. 2(a) any weakly convergent subsequence converges to a solution $c_{\tilde{n}}^{\beta^*}$ of $(\mathrm{ROLS})_{\beta^*}^{w_{\tilde{n}}^0}$. Thus taking the limit in (6. 11) we obtain

$$2\beta^* - k_1^{-2}|u(c_{\tilde{n}}^{\beta^*}) - z_{\tilde{n}}|^2 \leqq \frac{\varepsilon}{4},$$

which contradicts (6. 8). On the other hand, if $\lim n_m = \infty$, then it is simple to see that $c_{n_m}^{\beta_m}$ contains a subsequence converging weakly to a solution $c_0^{\beta^*}$ of $(\mathrm{ROLS})_{\beta^*}^{w_0^0}$. Again we take the limit in (6. 11) and arrive at

$$\beta^* - k_1^{-2}|u(c_0^{\beta^*}) - z_0^0|^2 \leqq \frac{\beta^*}{4},$$

which contradicts (6. 7). Thus (6. 10) is verified and the theorem is proved.

7. Numerical experiments

In this section we briefly discuss some features of experiments that were carried out to solve (OLS) and (ROLS) numerically. The approach that we took is rather simple. The differential equation was approximated by a Galerkin approximation using linear spline functions with equidistant knots at $\frac{i}{N}$, $i = 0, \ldots, N$. For the unknown function c we used approximation by linear spline functions with equidistant knots at $\frac{j}{NP}$, $j = 0, \ldots, NP$; thus we searched for α_i^{NP} in $c^{NP} = \sum_{i=0}^{NP} \alpha_i^{NP} B_i^{NP}$, where B_i^{NP} are the linear spline basis functions. As "observations" we took the values of the true solution evaluated at $\frac{j}{N}$, $j=0,\ldots,N$. The minimization problem was solved by the Levenberg-Marquardt routine, available from the IMSL-library. Two classes of examples (E.1) and (E.2) were tested:

(E.1)	$a \equiv 1, \quad c(x) = e^x$,	Neumann boundary conditions,
	$f(x) = \cos k\pi x(k^2\pi^2 + e^x)$,	for various values of k,
(E.2)	$a \equiv 1, \quad c(x) = 1 + \cos n\pi x$,	Neumann boundary conditions,
	$f(x) = \cos k\pi x(k^2\pi^2 + \cos n\pi x + 1)$,	for various values of k and n.

The solution of (E.1) and (E.2) is given by $u(x) = \cos k\pi x$.

It was generally observed that:

(i) The converged values for the optimal parameters were rather independent of any reasonable start-up value for the minimization routine.

(ii) Taking fewer observations did not lead to reasonable results.

(iii) The criterion that terminated the minimization algorithm was generally the gradient criterion.

(iv) Taking the H^0- or the H^1-norm in the fit-to-data criterion, did not lead to significantly different results.

(v) Taking $N = NP$ did not lead to good results unless a regularization-term was used; best results were obtained with $\beta = 10^{-6}$ or 10^{-5}.

(vi) We do not have a criterion to calculate the "best" β-value. Sometimes $\beta = 0$ gave best values and any $\beta > 0$ introduced too much damping in the oscillation of the coefficient $c(x) = 1 + \cos n\pi x$ (see (E.2)). Using a small regularization term, however, we got a qualitatively correct result in all but one of our calculations. For $\beta = 0$ excessive oscillations can occur.

For the numerical calculations we thank Dipl.-Ing. G. Moyschewitz.

Acknowledgement. We thank Prof. F. Lempio for bringing the paper [1] by Dr. W. Alt to our attention.

References

[1] *W. Alt*, Lipschitzian perturbations of infinite optimization problems; in: Mathematical Programming with Data Perturbations, ed. by A. V. Fiacco, New York 1983, 7—21.

[2] *J. P. Aubin*, Approximation of Elliptic Boundary Value Problems, London 1972.

[3] *A. Bamberger*, *G. Chavent* and *P. Lailly*, About the stability of the inverse problem in the 1-D wave equation — Application to the interpretation of seismic profiles, Appl. Math. Opt. **5** (1979), 1—47.

[4] *H. T. Banks*, A survey of some problems and recent results for parameter estimation and optimal control in delay and distributed parameter systems, in: Volterra and Functional Differential Equations, ed. by K. B. Hannsgen, et al., New York 1982, 3—24.

[5] *H. T. Banks* and *K. A. Murphy*, Estimation of coefficients and boundary parameters in hyperbolic systems, to appear in SIAM J. Contr. Opt.

[6] *H. Bauer*, Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie, Berlin 1968.

[7] *G. Chavent*, Analyse fonctionnelle et identification de coefficients repartis dans les equations aux derivees partielles, Thèse, Paris 1971.

[8] *G. Chavent*, Local stability of the output least square parameter estimation technique, Mat. Apl. Comput. **2** (1983), 3—22.
[9] *E. Deuflhard* and *E. Hairer* (eds.), Numerical Treatment of Inverse Problems in Differential and Integral Equations, Stuttgart 1983.
[10] *G. Hellwig*, Differential Operators of Mathematical Physics, London 1964.
[11] *C. Kravaris*, Identification of spatially varying parameters indistributed systems, Thesis, Los Angeles 1984.
[12] *C. Kravaris* and *J. H. Seinfeld*, Identification of parameters in distributed parameter systems by regularization, SIAM J. Contr. Opt. **23** (1985), 217—241.
[13] *K. Kunisch* and *L. W. White*, Parameter estimation, regularity and the penalty method for a class of elliptic problems, Nonl. Anal. TMA **10** (1986), 121—146.
[14] *H. Maurer* and *J. Zowe*, First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems, Math. Progr. **16** (1979), 98—110.
[15] *S. M. Robinson*, Stability theory for systems of inequalities. II, differentiable nonlinear systems, SIAM J. Num. Anal. **13** (1976), 487—513.
[16] *M. Schechter*, Negative norms and boundary problems, Anal. Math. **72** (1960), 581—593.
[17] *A. N. Tikhonov* and *V. Ya. Arsenin*, Solutions of Illposed Problems, New York 1977.
[18] *L. W. White*, Identification of friction parameters in a first order linear hyperbolic system, Proceedings of the 22nd CDC-Conference, San Antonio 1983.
[19] *A. Wouk*, A Course of Applied Functional Analysis, New York 1979.
[20] *J. Zowe* and *S. Kurcyusz*, Regularity and Stability for the mathematical programming problem in Banach Spaces, Appl. Math. Opt. **5** (1979), 49—62.

Institut für Dynamische Systeme, Universität Bremen, Postfach 330440, D-2800 Bremen 33

Institut für Mathematik, Technische Universität Graz, Kopernikusgasse 24, A-8010 Graz

Eingegangen 16. November 1984