Regularization of Nonlinear Illposed Problems with Closed Operators

Karl Kunisch* and Wolfgang Ring*
Institut für Mathematik
Technische Universität Graz
Kopernikusgasse 24, A-8010 Graz, Austria

July 1992

Abstract

In this paper Tikhonov regularization for nonlinear illposed problems is investigated. The regularization term is characterized by a closed linear operator, permitting seminorm regularization in applications. Results for existence, stability, convergence and convergence rates of the solution of the regularized problem in terms of the noise level are given. An illustrating example involving parameter estimation for a one dimensional stationary heat equation is given.

1 Introduction

The subject of the present paper is the problem

$$F(x_0) = y_0, \tag{1}$$

where $F : \text{dom } F \subset X \to Y$ is a not necessarily linear operator between Hilbert spaces X and Y. We are interested in the situation where (1) is illposed in the sense of lack of uniqueness and continuous dependence on the data y_0 of the solutions of (1). We shall also require a linear operator

$$D : \text{dom } D \subset X \to Z,$$

with Z a Hilbert space, which will allow to define a seminorm $|\cdot|$ on dom F by means of $|x| = \|Dx\|_Z$. The following assumptions on (F, D) shall hold throughout the paper:

(A1) dom $F \subset$ dom D

and

(A2) $(F, D) : \text{dom } F \subset X \to Y \times Z$ is weakly sequentially closed, i.e. for any sequence $\{x_n\} \subset \text{dom } F$ with $x_n \rightharpoonup x$ and $(F(x_n), Dx_n) \rightharpoonup (y, z)$ it follows that $x \in \text{dom } F$ and $(F(x), Dx) = (y, z)$. (Here '$\rightharpoonup$' denotes weak convergence.)

*Research supported by Fonds zur Förderung der wissenschaftlichen Forschung, Austria, under P 7869-Phy

Under the hypothesis that (A1) holds, the assumption (A2) is fulfilled if, for instance, F is weakly sequentially closed and D is a closed linear operator. The weak closedness of D is then an immediate consequence of Mazur's theorem on weakly closed sets in a Banach space.

We choose a generalized concept of 'solution' for problem (1). An element x_0 is called *'x^*-minimum-seminorm solution'* (x^*-MSS) of problem (1) for given $y_0 \in Y$ and $x^* \in \text{dom } D$ if

$$F(x_0) = y_0 \tag{2}$$

and

$$\|Dx_0 - Dx^*\|_Z = \min\{\|Dx - Dx^*\|_Z \mid F(x) = y_0\}. \tag{3}$$

A solution of (2) and (3) need not exist, and even if it exists, it need not be unique partly because F is nonlinear and partly because in condition (3) — which selects from all solutions of (2) the one(s) which is (or which are) closest to the a-priori guess x^* — the distance is measured by the *seminorm* $|\cdot|$.

To stabilize problem (1), we use the Tikhonov regularization method with the seminorm $|\cdot|$ in the regularization term, i.e. we seek to approximate a solution of (1) by a solution of the problem

$$\min_{x \in \text{dom } F} (\|F(x) - y_0\|_Y^2 + \alpha\|Dx - Dx^*\|_Z^2), \tag{4}$$

where $y_0 \in Y$, $x^* \in \text{dom } D$ and $\alpha > 0$.

In the following section we prove an existence result for problem (4) and we show that solutions of (4) (for fixed $\alpha > 0$ and $x^* \in \text{dom } D$) are stable with respect to perturbations of the data y_0. A theorem which ensures the convergence of solutions of the regularized problem (4), with y_0 replaced by a perturbation y_δ, towards an x^*-MSS of (1) if $\|y_\delta - y_0\|_Y$ and the regularization parameter α go to zero, is presented at the end of section 2. In section 3 we present a result on convergence rates for solutions of the regularized problem (4). In that section we use the concepts provided in [1]. In that work the regularization term was restricted to be constructed from the norm on X, i.e. $D =$ identity. The practical advantage of allowing for regularization with an operator D is given by the fact that one can realize seminorm regularization terms, which penalize undesired oscillations in the numerical solution without significantly effecting its low modes. Numerical experiments also confirm that the range for the regularization parameter within which good numerical results for certain nonlinear illposed problems are obtained is increased by regularizing with a seminorm as opposed to using a full norm regularization term. We end with an example, where we apply our results to a parameter estimation problem to identify the heat conductivity for a one dimensional stationary heat equation.

2 Existence, Stability and Convergence

We give conditions under which the minimisation problem (4) has a solution which depends continuously on the data $y \in Y$ and which converges to an x^*-MSS of the problem (1) for $\alpha \to 0$. We shall need two additional assumptions on the operators F and D.

(A3) Let $\{x_n\}$ be a sequence in dom F. If $\{(F(x_n), Dx_n)\}$ is bounded in $Y \times Z$ then $\{x_n\}$ is bounded.

(A4) If $x_n \rightharpoonup x$ and $\|Dx_n\|_Z \to \|Dx\|_Z$ for a sequence $\{x_n\}$ in dom D, then $x_n \to x$.

In case D is a closed linear operator we have the following equivalence.

Lemma 1 *Let D : dom $D \subset X \to Z$ be a closed linear operator between Hilbert spaces X and Z. Then (A4) is equivalent to*

- $\mathrm{ker}D$ *is finite dimensional and* $\mathrm{rg}D$ *is closed.*

PROOF

Suppose that (A4) holds. Then $\mathrm{ker}D$ cannot have infinite dimension since an infinite orthonormal basis in $\mathrm{ker}D$ would violate (A4).

Next it will be shown that $\mathrm{rg}D$ is closed. The kernel $\mathrm{ker}D$ is closed and one can use the decomposition $X = \mathrm{ker}D \oplus (\mathrm{ker}D)^\perp$. The operator $\bar{D} = D|_{(\mathrm{ker}D)^\perp}$ is closed and injective with $\mathrm{rg}\bar{D} = \mathrm{rg}D$. We can therefore assume that D is injective, as otherwise one could factor $\mathrm{ker}D$ and deal with $\bar{D}$.

Suppose now that $Dx_n \to z$ in Z. We first show that $\{x_n\}$ is bounded in X. Else we can assume that $\|x_n\|_X \to \infty$ and then for $w_n = \frac{1}{\|x_n\|_X} x_n$ we have

$$Dw_n = \frac{1}{\|x_n\|_X} Dx_n \to 0. \tag{5}$$

The sequence $\{w_n\}$ is bounded in X, hence we can find a subsequence $w_{n_k} \rightharpoonup w$. From (5) we get $Dw_{n_k} \rightharpoonup 0$ and since D is weakly closed we obtain that $w \in$ dom D and $Dw = 0$. Since D is injective, it follows that $w = 0$. We thus have $w_{n_k} \rightharpoonup 0$ and $\|Dw_{n_k}\|_Z \to 0$ so (A4) implies that $w_{n_k} \to 0$, which is a contradiction to $\|w_{n_k}\|_X = 1$ for all $k \in \mathbb{N}$.

We can therefore conclude that $\{x_n\}$ must be bounded. Hence there exists a weakly convergent subsequence $x_{n_k} \rightharpoonup x$. From $Dx_{n_k} \rightharpoonup z$ and from the weak closedness of D it follows that $x \in$ dom D and $Dx = z$. This proves that $\mathrm{rg}D$ is closed in Z.

On the other hand suppose that $x_n \rightharpoonup x$ and $\|Dx_n\|_Z \to \|Dx\|_Z$. Let $P : X \to \mathrm{ker}D$ and $Q : X \to (\mathrm{ker}D)^\perp$ denote the orthogonal projections on $\mathrm{ker}D$ and $(\mathrm{ker}D)^\perp$ respectively. The projections are weakly continuous and therefore $Px_n \rightharpoonup Px$. In the finite dimensional subspace $\mathrm{ker}D$ weak- and norm convergence coincide, so we have $Px_n \to x$.

We next show that $Qx_{n_k} \to Qx$ for a subsequence $\{x_{n_k}\}$ of $\{x_n\}$. The sequence $\{Dx_n\}$ is bounded hence there exists a weakly convergent subsequence $Dx_{n_k} \rightharpoonup z$. The weak closedness of D gives $z = Dx$. It is a well known fact that in a Hilbert space $Dx_{n_k} \rightharpoonup Dx$ and $\|Dx_{n_k}\|_Z \to \|Dx\|_Z$ implies $Dx_{n_k} \to Dx$. From the closedness of $\mathrm{rg}D$ it follows that the generalized inverse $D^\dagger : Z \to (\mathrm{ker}D)^\perp$ is a continuous operator on Z (c.f. Groetsch [2, pages 36–46]). We thus find

$$Qx_{n_k} = D^\dagger Dx_{n_k} \to D^\dagger Dx = Qx$$

and consequently

$$x_{n_k} = Px_{n_k} + Qx_{n_k} \to Px + Qx = x.$$

Finally the weak convergence of x_n to x and the norm convergence of a subsequence imply $x_n \to x$. ☺

We prove now an existence result for the minimization problem (4).

Theorem 2 (Existence) *Under the assumptions (A1)–(A3), the problem*

$$\min_{x \in \mathrm{dom}\ F} \left(\|F(x) - y_0\|^2 + \alpha \|Dx - Dx^*\|^2 \right)$$

has at least one solution $x_0 \in \mathrm{dom}\ F$ *for given* $y_0 \in Y$, $x^* \in \mathrm{dom}\ D$ *and* $\alpha > 0$.

Proof

Let $\{x_n\}$ be a minimizing sequence for (4). Then $\{(F(x_n), Dx_n)\}$ is bounded in $Y \times Z$ and therefore (due to (A3)) $\{(x_n, F(x_n), Dx_n)\}$ is bounded in $X \times Y \times Z$. Taking a weakly convergent subsequence $(x_{n_k}, F(x_{n_k}), Dx_{n_k}) \rightharpoonup (x_0, y, z)$, the weak closedness of (F, D) implies that $x_0 \in \mathrm{dom}\ F \subset \mathrm{dom}\ D$ and $(y, z) = (F(x_0), Dx_0)$. From the weak lower semicontinuity (w.l.s.c.) of the norm it follows that x_0 is a solution of (4). ☺

Let us introduce

$$(x, y)_D = (x, y)_X + (Dx, Dy)_Z \tag{6}$$

for $x \in \mathrm{dom}\ D$. The set dom D endowed with the graph norm $\|x\|_D = (x, x)_D^{1/2}$ is denoted by $(\mathrm{dom}\ D, \|\cdot\|_D)$. It is a Hilbert space. The next theorem establishes the stability of solutions of (4) in $(\mathrm{dom}\ D, \|\cdot\|_D)$ with respect to perturbations of y_0 for fixed $\alpha > 0$.

Theorem 3 (Stability) *Assume that (A1)–(A4) hold. Let* $\{y_n\} \subset Y$ *with* $y_n \to y$ *and let* $\alpha > 0$ *and* $x^* \in \mathrm{dom}\ D$ *be fixed. Further let* $\{x_n\}$ *be a sequence of solutions for*

$$\min_{x \in \mathrm{dom}\ F} \Big(\|F(x) - y_n\|_Y^2 + \alpha \|Dx - Dx^*\|_Z^2 \Big). \tag{7}$$

Then there exists a subsequence $\{x_{n_k}\}$ *of* $\{x_n\}$ *convergent in* $(\mathrm{dom}\ D, \|\cdot\|_D)$. *Every* X*-weakly convergent subsequence* $\{x_{n_k}\}$ *of* x_n *converges in* $(\mathrm{dom}\ D, \|\cdot\|_D)$ *to a solution of*

$$\min_{x \in \mathrm{dom}\ F} \Big(\|F(x) - y\|_Y^2 + \alpha \|Dx - Dx^*\|_Z^2 \Big).$$

Proof

From the fact that x_n is a solution of (7) it easily follows that $\{(F(x_n), Dx_n)\}$ is bounded in $Y \times Z$. Furthermore (A3) implies that $\{x_n\}$ is bounded in X and obviously $\{\|Dx_n - Dx^*\|_Z\}$ is bounded in $\mathbb{R}$. We can therefore find a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that

$$x_{n_k} \rightharpoonup x; \quad F(x_{n_k}) \rightharpoonup v; \quad Dx_{n_k} \rightharpoonup z; \quad \text{and} \quad \|Dx_{n_k} - Dx^*\|_Z \to c \geq 0.$$

Due to the weak closedness of (F, D) we have $x \in \text{dom } F \subset \text{dom } D$, $F(x) = v$ and $Dx = z$. Let $\tilde{x} \in \text{dom } F$. The following inequalites

$$\begin{aligned}
\|F(x) - y\|_y^2 + \alpha\|Dx - Dx^*\|_Z^2 &\leq \liminf_{k\to\infty} \left(\|F(x_{n_k}) - y_{n_k}\|_Y^2 + \alpha\|Dx_{n_k} - Dx^*\|_Z^2\right) \\
&\leq \lim_{k\to\infty} \left(\|F(\tilde{x}) - y_{n_k}\|_Y^2 + \alpha\|D\tilde{x} - Dx^*\|_Z^2\right) \\
&= \|F(\tilde{x}) - y\|_Y^2 + \alpha\|D\tilde{x} - Dx^*\|_Z^2
\end{aligned}$$

show that x is a solution of

$$\min_{x\in\text{dom } F} \|F(x) - y\|_Y^2 + \alpha\|Dx - Dx^*\|_Z^2.$$

It remains to prove that $\{x_{n_k}\}$ converges to x not only in the weak topology but also in $(\text{dom } D, \|\cdot\|_D)$. We claim that $\|Dx_{n_k} - Dx^*\|_Z \to \|Dx - Dx^*\|_Z$. From $Dx_{n_k} \rightharpoonup Dx$ and from the w.l.s.c. of the norm it follows that

$$\|Dx - Dx^*\|_Z \leq \lim_{k\to\infty} \|Dx_{n_k} - Dx^*\|_Z = c.$$

Suppose that there exists $\rho > 0$ with

$$\|Dx - Dx^*\|_Z^2 \leq c^2 - \rho. \tag{8}$$

We choose $\epsilon \in (0, \alpha\rho)$. By the w.l.s.c. of the norm we have $\|F(x) - y\|_Y^2 \leq \liminf_{k\to\infty} \|F(x_{n_k}) - y_{n_k}\|_Y^2$, so we can find an integer $N_0 \in \mathbb{N}$ such that

$$\|F(x) - y\|_Y^2 \leq \|F(x_{n_k}) - y_{n_k}\|_Y^2 + \frac{\epsilon}{4} \tag{9}$$

for all $k \geq N_0$. Moreover $y_{n_k} \to y$ implies $\|F(x) - y_{n_k}\|_Y^2 \to \|F(x) - y\|_Y^2$ for $k \to \infty$. Hence we can assume — possibly with a larger N_0 — that

$$\|F(x) - y_{n_k}\|_Y^2 \leq \|F(x) - y\|_Y^2 + \frac{\epsilon}{4} \tag{10}$$

for all $k \geq N_0$. If we assume further that N_0 is large enough such that

$$\alpha c^2 \leq \alpha\|Dx_{n_k} - Dx^*\|_Z^2 + \frac{\epsilon}{2} \tag{11}$$

for all $k \geq N_0$, then (8)–(11) imply

$$\begin{aligned}
\|F(x) - y_{n_k}\|_Y^2 + \alpha\|Dx - Dx^*\|_Z^2 &\leq \|F(x_{n_k}) - y_{n_k}\|_Y^2 + \alpha(c^2 - \rho) + \frac{\epsilon}{2} \\
&\leq \|F(x_{n_k}) - y_{n_k}\|_Y^2 + \alpha\|Dx_{n_k} - Dx^*\|_Z^2 - \alpha\rho + \epsilon,
\end{aligned}$$

for $k \geq N_0$. Since $\epsilon \in (0, \alpha\rho)$ it follows that

$$\|F(x) - y_{n_k}\|_Y^2 + \alpha\|Dx - Dx^*\|_Z^2 < \|F(x_{n_k}) - y_{n_k}\|_Y^2 + \alpha\|Dx_{n_k} - Dx\|_Z^2,$$

which is a contradiction to the fact that x_{n_k} is a solution of the minimisation problem (7).

We thus have $\|Dx_{n_k} - Dx^*\|_Z \to \|Dx - Dx^*\|_Z$ and $x_{n_k} - x^* \rightharpoonup x - x^*$. It follows from (A4) that $x_{n_k} - x^* \to x - x^*$ in X and since $Dx_{n_k} - Dx^* \rightharpoonup Dx - Dx^*$ we obtain $x_{n_k} \to x$ in $(\mathrm{dom}\ D, \|\cdot\|_D)$. ☺

The following theorem shows that the solutions of (4) with y_0 replaced by an inaccurate measurement y_δ converge towards an x^*-MSS of (1) if the error δ between y_δ and y_0 and the regularization parameter α go to 0 appropriately. We also take into account that the minimisation problem (4) can only be solved approximately in practical applications. We assume that

$$y_\delta \in Y \quad \text{with} \quad \|y_\delta - y_0\|_Y \le \delta \tag{12}$$

and that $x_\alpha^{\delta,\eta} \in \mathrm{dom}\ F$ satisfies

$$\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 + \alpha\|Dx_\alpha^{\delta,\eta} - Dx^*\|_Z^2 \le \|F(x) - y_\delta\|_Y^2 + \alpha\|Dx - Dx^*\|_Z^2 + \eta \tag{13}$$

for all $x \in \mathrm{dom}\ F$, where $\eta > 0$ is a given (small) error bound for an approximate solution of (4).

Theorem 4 (Convergence) *Assume that (A1)–(A4) hold. For a given $y_0 \in Y$ let*

- *$x_0 \in \mathrm{dom}\ F$ with $F(x_0) = y_0$, i.e. y_0 is attainable.*
- *y_δ and $x_\alpha^{\delta,\eta}$ as in (12) and in (13) respectively.*

Assume further that the regularization parameter $\alpha = \alpha(\delta, \eta)$ is choosen such that

$$\alpha(\delta,\eta) \to 0; \qquad \frac{\delta^2}{\alpha(\delta,\eta)} \to 0; \qquad \frac{\eta}{\alpha(\delta,\eta)} \to 0$$

for $(\delta,\eta) \to (0,0)$.

Then for $\delta_k \to 0$, $\eta_k \to 0$ and $\alpha_k = \alpha(\delta_k,\eta_k)$ we have:

1. *$\{x_{\alpha_k}^{\delta_k,\eta_k}\}$ has a subsequence convergent in $(\mathrm{dom}\ D, \|\cdot\|_D)$. Every X-weakly convergent subsequence of $\{x_{\alpha_k}^{\delta_k,\eta_k}\}$ converges in $(\mathrm{dom}\ D, \|\cdot\|_D)$ and its limit $\hat{x}$ is an x^*-MSS of (1), i.e.*

$$F(\hat{x}) = y_0 \quad \text{and} \quad \|D\hat{x} - Dx^*\|_Z = \min\{\|D\tilde{x} - Dx^*\|_Z \,|\, F(\tilde{x}) = y_0\}.$$

2. *If only one x^*-MSS $\hat{x}$ of (1) exists then we have*

$$\lim_{k\to\infty} x_{\alpha_k}^{\delta_k,\eta_k} = \hat{x}$$

in $(\mathrm{dom}\ D, \|\cdot\|_D)$.

PROOF

We make the following abbreviations: $x_k = x_{\alpha_k}^{\delta_k,\eta_k}$ and $y_k = y_{\delta_k}$.

From (13) it follows that $\{F(x_k)\}$ is a bounded sequence. We further find

$$\|Dx_k - Dx^*\|_Z^2 \le \frac{\delta_k^2}{\alpha_k} + \frac{\eta_k}{\alpha_k} + \|Dx_0 - Dx^*\|_Z^2,$$

and due to the assumption on α_k it follows that $\{Dx_k\}$ is bounded. (A3) implies that $\{x_k\}$ is bounded as well and hence the sequence $\{(x_k, F(x_k), Dx_k)\}$ is bounded in $X \times Y \times Z$. Since (F, D) is weakly closed, we can choose a weakly convergent subsequence such that

$$x_{k_m} \rightharpoonup \hat{x}; \quad F(x_{k_m}) \rightharpoonup F(\hat{x}); \quad Dx_{k_m} \rightharpoonup D\hat{x}.$$

We have

$$\begin{aligned}
\|F(\hat{x}) - y_0\|_Y^2 &\le \liminf_{m\to\infty} \left(\|F(x_{k_m}) - y_{k_m}\|_Y^2 + \alpha_{k_m}\|Dx_{k_m} - Dx^*\|_Z^2\right) \\
&\le \liminf_{m\to\infty} \left(\|F(x_0) - y_{k_m}\|_Y^2 + \alpha_{k_m}\|Dx_0 - Dx^*\|_Z^2 + \eta_{k_m}\right) \\
&= \|F(x_0) - y_0\|_Y^2 = 0,
\end{aligned}$$

and consequently $F(\hat{x}) = y_0$.

Let $\tilde{x} \in \operatorname{dom} F$ with $F(\tilde{x}) = y_0$ be arbitrary. We find

$$\begin{aligned}
\|D\hat{x} - Dx^*\|_Z^2 &\le \liminf_{m\to\infty} \|Dx_{k_m} - Dx^*\|_Z^2 \\
&\le \lim_{m\to\infty} \alpha_{k_m}^{-1}\left(\|F(\tilde{x}) - y_{k_m}\|_Y^2 + \alpha_{k_m}\|D\tilde{x} - Dx^*\|_Z^2 + \eta_{k_m}\right) \\
&\le \lim_{m\to\infty} \left(\frac{\delta_{k_m}^2}{\alpha_{k_m}} + \frac{\eta_{k_m}}{\alpha_{k_m}} + \|D\tilde{x} - Dx^*\|_Z^2\right) \\
&= \|D\tilde{x} - Dx^*\|_Z^2,
\end{aligned}$$

so $\|D\hat{x} - Dx^*\|_Z = \min\{\|D\tilde{x} - Dx^*\|_Z \,|\, F(\tilde{x}) = y_0\}$. This proves that $\hat{x}$ is an x^*-MSS of problem (1).

We still have to show that $x_{k_m} \to \hat{x}$ in $(\operatorname{dom} D, \|\cdot\|_D)$. We first prove that $x_{k_m} \to \hat{x}$ in X. Due to (A4) and the fact that $x_{m_k} - x^* \rightharpoonup \hat{x} - x^*$ it is sufficient to prove that $\|Dx_{k_m} - Dx^*\|_Z \to \|D\hat{x} - Dx^*\|_Z$. We have

$$\begin{aligned}
\|D\hat{x} - Dx^*\|_Z^2 &\le \liminf_{m\to\infty} \|Dx_{k_m} - Dx^*\|_Z^2 \le \limsup_{m\to\infty} \|Dx_{k_m} - Dx^*\|_Z^2 \\
&\le \lim_{m\to\infty} \left(\frac{\delta_{k_m}^2}{\alpha_{k_m}} + \frac{\eta_{k_m}}{\alpha_{k_m}} + \|D\hat{x} - Dx^*\|_Z^2\right) = \|D\hat{x} - Dx^*\|_Z^2,
\end{aligned}$$

so $\|Dx_{k_m} - Dx^*\| \to \|D\hat{x} - Dx^*\|$ and $x_{k_m} \to \hat{x}$ in X. Moreover $\|Dx_{k_m} - Dx^*\| \to \|D\hat{x} - Dx^*\|$ and $Dx_{n_k} \rightharpoonup D\hat{x}$ immediately implies that $Dx_{n_k} \to D\hat{x}$ in Z, i.e. $x_{n_k} \to \hat{x}$ in $(\text{dom D}, \|\cdot\|_D)$. This proves the first claim of the theorem. The second is an immediate consequence. ☺

3 Rate of Convergence

In this section we shall derive a result on the rate of convergence for the solutions $x_\alpha^{\delta,\eta}$ of the regularized problem (13) as $\delta \to 0^+$.

We shall make frequent use of the the following inequality:

$$\|a-b\|_H^2 \geq \mu\|a\|_H^2 - \frac{\mu}{1-\mu}\|b\|_H^2, \tag{14}$$

for every $a, b \in H$ and $\mu \in (0,1)$, H being a Hilbert space with norm $\|\cdot\|_H$.

In the following proposition a norm equivalent to $\|\cdot\|_D$ is constructed.

Proposition 5 *Let X, Y and Z be Hilbert spaces and let $D : \mathrm{dom}\ D \subset X \to Z$ be a closed linear operator. Moreover let $A : (\mathrm{dom}\ D, \|\cdot\|_D) \to Y$ be a bounded linear Operator and assume further that*

- $\ker D \cap \ker A = \{0\}$
- $\mathrm{rg} D$ *is closed*
- *there exists a number $\beta > 0$ such that*

$$\|Aw\|_Y \geq \beta\|w\|_X \quad \textit{for all} \quad w \in \ker D.$$

Then

$$(x,y)_A = (Ax, Ay)_Y + (Dx, Dy)_Z \tag{15}$$

is an inner product on $\mathrm{dom}\ D$ *and the norms $\|x\|_D$ and $\|x\|_A = (x,x)_A^{1/2}$ are equivalent on* $\mathrm{dom} D$*. In particular, there exists a number $\kappa > 0$ such that*

$$\|x\|_A \geq \kappa\|x\|_D \quad \textit{for all} \quad x \in \mathrm{dom}\ D. \tag{16}$$

PROOF

Since $(x,x)_A = 0$ implies that $x \in \ker A \cap \ker D = \{0\}$ it follows that the symmetric bilinear form $(\cdot,\cdot)_A$ is positive definite and that $(\cdot,\cdot)_A$ is an inner product on dom D.

For $w \in \mathrm{dom}\ D$ we have

$$\|Aw\|_Y^2 + \|Dw\|_Z^2 \leq (1+\|A\|^2)(\|w\|_X^2 + \|Dw\|_Z^2).$$

To prove the equivalence of the two norms we still need to verify (16). First we observe that any element $w \in$ dom D can be decomposed as

$$w = w_1 + w_2 \quad \text{with} \quad w_1 \in \ker D,\ w_2 \in (\ker D)^\perp \cap \mathrm{dom}\ D.$$

Since $\mathrm{rg} D$ is closed in Z the generalized inverse $D^\dagger$ of D is a bounded linear operator from Z into $(\ker D)^\perp$. For $w_2 \in (\ker D)^\perp \cap \mathrm{dom}\ D$ we have $w_2 = D^\dagger D w_2$, and consequently

$$\|w_2\|_X \leq \|D^\dagger\|\,\|Dw_2\|_Z.$$

Using (14) with some $0 < \mu < 1$ we find for $w \in \text{dom } D$

$$
\begin{aligned}
&\|Aw\|_Y^2 + \|Dw\|_Z^2 \\
&\geq \mu \|Aw_1\|_Y^2 - \frac{\mu}{1-\mu}\|Aw_2\|_Y^2 + \|Dw_2\|_Z^2 \\
&\geq \mu\beta^2\|w_1\|_X^2 - \frac{\mu}{1-\mu}\|A\|^2(\|w_2\|_X^2 + \|Dw_2\|_Z^2) + \|Dw_2\|_Z^2 \\
&\geq \mu\beta^2\|w_1\|_X^2 + \Big((1 - \frac{\mu}{1-\mu}\|A\|^2)\|D^\dagger\|^{-2} - \frac{\mu}{1-\mu}\|A\|^2\Big)\|w_2\|_X^2 \geq \kappa'\|w\|_X^2
\end{aligned}
$$

where

$$\kappa' = \min\Big(\mu\beta^2, \|D^\dagger\|^{-2} - \frac{\mu}{1-\mu}C\Big)$$

with $C = \|A\|^2(1 + \|D^\dagger\|^{-2})$. The function $\kappa'(\mu)$ achieves its maximum $\kappa' = \beta\mu_0$ at $\mu_0 \in (0,1)$ given by

$$\mu_0 = \frac{1}{2\beta^2}\Big(\|D^\dagger\|^{-2} + C + \beta^2 - \sqrt{(\|D^\dagger\|^{-2} + C - \beta^2)^2 + 4\beta^2 C}\Big).$$

We thus find

$$\|w\|_A^2 \geq \frac{\kappa'}{1+\kappa'}\|w\|_D^2,$$

which implies (16) with $\kappa = \sqrt{\frac{\kappa'}{1+\kappa'}}$. This proves that the norms are equivalent and that dom D is a Hilbert space with respect to $\|\cdot\|_D$ and $\|\cdot\|_A$. ☺

A similar result but with A a bounded linear operator from X to Y is proved with a different technique in Locker and Prenter [3].

We present now the convergence result which we announced above. Conditions for 7. bellow to hold were given in Theorem 4.

Theorem 6 *Let X, Y and Z be Hilbert spaces and let F : dom $F \subset X \to Y$ and let D : dom $D \subset X \to Z$ be a closed, densely defined linear operator with* dom $F \subset$ dom D. *Further let y_δ and $x_\alpha^{\delta,\eta}$ satisfy (12) and in (13) respectively and let x_0 be an x^*-MSS of problem (1). Assume that*

1. rg$D$ *is closed,*

2. *F is Fréchet differentiable as an operator on a convex neighbourhood $V(x_0)$ of x_0 in* dom F *endowed with the $\|\cdot\|_D$-norm to X,*

3. ker$D\cap$ker$F'(x_0) = \{0\}$ *and there exists a number $\beta > 0$ such that $\|F'(x_0)w\|_Y \geq \beta\|w\|_X$ for all $w \in$* ker$D$*,*

4. *there exists $L > 0$ such that $\|F'(x_0) - F'(z)\| \leq L\|x_0 - z\|_D$ for all $z \in V(x_0)$,*

5. *$x_0 - x^* \in$* dom D^*D *and there exist an $\omega \in Y^*$ such that*

$$D^*D(x_0 - x^*) = \big(F'(x_0)\big)^*\omega,$$

6. $L\|\omega\|_{Y^*} < \kappa^2$, *where* κ *is defined as in the preceeding proposition with* A *replaced by* $F'(x_0)$, *i.e.*

$$\left(\|F'(x_0)x\|_Y^2 + \|Dx\|_X^2\right)^{\frac{1}{2}} \geq \kappa\|x\|_D \quad \textit{for all} \quad x \in \textit{dom } D.$$

7. $\delta \to 0$; $\alpha \sim \delta$; $\eta \sim \delta^2$ *and* $x_\alpha^{\delta,\eta} \to x_0$ *in* $(\text{dom } D, \|\cdot\|_D)$ *for* $\delta \to 0$.

Then

$$\|x_\alpha^{\delta,\eta} - x_0\|_D = O(\sqrt{\delta})$$

and

$$\|F(x_\alpha^{\delta,\eta}) - y_0\|_Y = O(\delta).$$

Proof

From (13) it follows that

$$\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 + \alpha\|Dx_\alpha^{\delta,\eta} - Dx^*\|_Z^2 \leq \delta^2 + \eta + \alpha\|Dx_0 - Dx^*\|_Z^2,$$

and using the assumption that $x_0 - x^* \in \text{dom } D^*D$ we find

$$\begin{aligned}
&\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 + \alpha\|Dx_\alpha^{\delta,\eta} - Dx_0\|_Z^2 \\
&\leq \delta^2 + \eta + \alpha(\|Dx_0 - Dx^*\|_Z^2 - \|Dx_\alpha^{\delta,\eta} - Dx^*\|_Z^2 + \|Dx_\alpha^{\delta,\eta} - Dx_0\|_Z^2) \\
&= \delta^2 + \eta + 2\alpha\big(D(x_0 - x^*), D(x_0 - x_\alpha^{\delta,\eta})\big)_Z \\
&= \delta^2 + \eta + 2\alpha\langle D^*D(x_0 - x^*), x_0 - x_\alpha^{\delta,\eta}\rangle_{X^*,X}.
\end{aligned}$$

Due to 7. one concludes that $x_\alpha^{\delta,\eta} \to x_0$ in $(\text{dom } D, \|\cdot\|_D)$ as $\delta \to 0$. Assumption 5. further implies that

$$\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 + \alpha\|Dx_\alpha^{\delta,\eta} - Dx_0\|_Z^2 \leq \delta^2 + \eta + 2\alpha\langle\omega, F'(x_0)(x_0 - x_\alpha^{\delta,\eta})\rangle_{Y^*,Y}. \tag{17}$$

From the differentiablity of F at x_0 and condition 4. it follows that for all δ sufficiently small, as will be assumed henceforth,

$$F(x_\alpha^{\delta,\eta}) = F(x_0) + F'(x_0)(x_\alpha^{\delta,\eta} - x_0) + r_\alpha^{\delta,\eta} \tag{18}$$

with

$$\|r_\alpha^{\delta,\eta}\|_Y \leq \frac{1}{2}L\|x_\alpha^{\delta,\eta} - x_0\|_D^2. \tag{19}$$

Using (17)–(19) we obtain

$$\begin{aligned}
&\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 + \alpha\|Dx_\alpha^{\delta,\eta} - Dx_0\|_Z^2 \\
&\leq \delta^2 + \eta + 2\alpha\langle\omega, F(x_0) - y_\delta + y_\delta - F(x_\alpha^{\delta,\eta}) + r_\alpha^{\delta,\eta}\rangle_{Y^*,Y} \\
&\leq \delta^2 + \eta + 2\alpha\delta\|\omega\|_{Y^*} + 2\alpha\|\omega\|_{Y^*}\,\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y + \alpha\|\omega\|_{Y^*}L\|x_0 - x_\alpha^{\delta,\eta}\|_D^2
\end{aligned} \tag{20}$$

Let $0 < \mu < 1$. By (14) (18) and (19) we obtain the estimate

$$\begin{aligned}
\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 &= \frac{1}{2}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 + \frac{1}{2}\|F(x_\alpha^{\delta,\eta}) - F(x_0) + F(x_0) - y_\delta\|_Y^2 \\
&\geq \frac{1}{2}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 + \frac{\mu}{2}\|F(x_\alpha^{\delta,\eta}) - F(x_0)\|_Y^2 - \frac{\mu}{2(1-\mu)}\delta^2 \\
&\geq \frac{1}{2}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 + \frac{\mu}{4}\|F'(x_0)(x_\alpha^{\delta,\eta} - x_0)\|_Y^2 \\
&\quad -\frac{\mu}{2}\|r_\alpha^{\delta,\eta}\|_Y^2 - \frac{\mu}{2(1-\mu)}\delta^2 \\
&\geq \frac{1}{2}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 + \frac{\mu}{4}\|F'(x_0)(x_\alpha^{\delta,\eta} - x_0)\|_Y^2 \\
&\quad -\frac{\mu L^2}{8}\|x_0 - x_\alpha^{\delta,\eta}\|_D^4 - \frac{\mu}{2(1-\mu)}\delta^2. \qquad (21)
\end{aligned}$$

Combining (20) and (21) gives

$$\begin{aligned}
&\frac{1}{2}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 + \frac{\mu}{4}\|F'(x_0)(x_\alpha^{\delta,\eta} - x_0)\|_Y^2 + \alpha\|Dx_\alpha^{\delta,\eta} - Dx_0\|_Z^2 \\
&\leq \left(1 + \frac{\mu}{2(1-\mu)}\right)\delta^2 + \eta + 2\alpha\delta\|\omega\|_{Y^*} + 2\alpha\|\omega\|_{Y^*}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y \\
&\quad +\Big(\alpha\|\omega\|_{Y^*}L + \frac{\mu L^2}{8}\|x_0 - x_\alpha^{\delta,\eta}\|_D^2\Big)\|x_0 - x_\alpha^{\delta,\eta}\|_D^2. \qquad (22)
\end{aligned}$$

If we choose $\mu = 4\alpha$ and δ sufficiently small such that $\mu \in (0,1)$, then

$$\begin{aligned}
&\frac{1}{2}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 + \alpha\|x_\alpha^{\delta,\eta} - x_0\|_A^2 \\
&\leq \frac{1-2\alpha}{1-4\alpha}\delta^2 + \eta + 2\alpha\delta\|\omega\|_{Y^*} + 2\alpha\|\omega\|_{Y^*}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y \\
&\quad +\alpha\Big(\|\omega\|_{Y^*}L + \frac{L^2}{2}\|x_0 - x_\alpha^{\delta,\eta}\|_D^2\Big)\|x_0 - x_\alpha^{\delta,\eta}\|_D^2, \qquad (23)
\end{aligned}$$

where $A = F'(x_0)$. With κ as in 6. we find

$$\begin{aligned}
&\frac{1}{2}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 + \alpha\kappa^2\|x_\alpha^{\delta,\eta} - x_0\|_D^2 \\
&\leq \frac{1-2\alpha}{1-4\alpha}\delta^2 + \eta + 2\alpha\delta\|\omega\|_{Y^*} + 2\alpha\|\omega\|_{Y^*}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y \\
&\quad +\alpha\Big(\|\omega\|_{Y^*}L + \frac{L^2}{2}\|x_0 - x_\alpha^{\delta,\eta}\|_D^2\Big)\|x_0 - x_\alpha^{\delta,\eta}\|_D^2. \qquad (24)
\end{aligned}$$

Since $x_\alpha^{\delta,\eta} \to x_0$ in $(\mathrm{dom}\, D, \|\cdot\|_D)$ and due to conditions 6. and 7. we can choose M and δ_0 sufficiently small such that

$$0 < M \leq \kappa^2 - L\|\omega\|_{Y^*} - \frac{L^2}{2}\|x_0 - x_\alpha^{\delta,\eta}\|_D^2 \qquad (25)$$

for all $\delta \in (0, \delta_0]$. This together with (24) implies that

$$\frac{1}{2}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y^2 \leq \frac{1-2\alpha}{1-4\alpha}\delta^2 + \eta + 2\alpha\delta\|\omega\|_{Y^*} + 2\alpha\|\omega\|_{Y^*}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y$$

for all $\delta \in (0, \delta_0]$. The implication

$$(a, b, c \geq 0;\ a^2 \leq ab + c^2) \Rightarrow a \leq b + c$$

yields

$$\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y \leq \left(\frac{2(1-2\alpha)}{1-4\alpha}\delta^2 + 2\eta + 4\alpha\delta\|\omega\|_{Y^*}\right)^{\frac{1}{2}} + 4\alpha\|\omega\|_{Y^*} \tag{26}$$

for $\delta \in (0, \delta_0]$ so, for the convergence rate of the outputs $F(x_\alpha^{\delta,\eta})$ towards y_0 as $\delta \to 0$, we find

$$\|F(x_\alpha^{\delta,\eta}) - y_0\|_Y \leq \|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y + \delta = O(\delta). \tag{27}$$

Using (24) once again we obtain

$$\alpha M\|x_\alpha^{\delta,\eta} - x_0\|_D^2 \leq \frac{1-2\alpha}{1-4\alpha}\delta^2 + \eta + 2\alpha\delta\|\omega\|_{Y^*} + 2\alpha\|\omega\|_{Y^*}\ \|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y$$

and consequently

$$\|x_\alpha^{\delta,\eta} - x_0\|_D \leq \left(\frac{\frac{1-2\alpha}{1-4\alpha}\delta^2 + \eta + 2\alpha\delta\|\omega\|_{Y^*} + 2\alpha\|\omega\|_{Y^*}\|F(x_\alpha^{\delta,\eta}) - y_\delta\|_Y}{M\alpha}\right)^{\frac{1}{2}} = O(\sqrt{\delta}) \tag{28}$$

due to condition 7. ☺

4 An Illustrating Example

Consider the two-point boundary value problem

$$\begin{cases} -(au_x)_x + cu = f \text{ in } (0,1) \\ u(0) = u(1) = 0 \end{cases} \tag{29}$$

with $f \in L^2$, $f \neq 0$ and $c \in L^2$, $c(x) \geq 0$ almost everywhere. Unless specified otherwise all function spaces are considered over the interval $(0, 1)$. The problem consists in determinating the coefficient a from noisy data u_δ of the state variable u of (29). We introduce the nonlinear parameter-to-solution mapping

$$F : \text{dom } F \to L^2,$$

where

$$\text{dom } F = \{a \in H^1 \mid 0 < k \leq a(x) \leq K, \text{ for all } x \in [0, 1]\},$$

with k and K positive constants and

$$F(a) = u(a)$$

where $u(a)$ denotes the solution to (29) as a function of a. Here we use the fact that for every $a \in \text{dom } F$, equation (29) admits a unique solution $u(a) \in H_0^1 \cap H^2$. It is assumed that the noisefree observation u_0 is attainable, i.e. there exists $a_0 \in \text{dom } F$ with

$$F(a_0) = u(a_0) = u_0,$$

and that for $\delta > 0$ the noisy data u_δ satisfy

$$\|u_\delta - u_0\|_{L^2} \leq \delta.$$

The determination of a_0 from data $u_\delta \in L^2$ is illposed (see e.g. [1]) when the distance in the parameter space is measured by the H^1- or the L^2-norm for example, and hence a formulation as least-squares problem is appropriate:

$$\min_{a \in \text{dom } F} \|F(a) - u_\delta\|_{L^2}^2 + \alpha \|Da - Da^*\|_{L^2}^2,$$

where $a^* \in H^1$ and D is given by

$$Da = a_x.$$

We shall express this problem in two different ways within the framework of the analysis of the previous sections. It will be convenient to summarize some technicalities. The space H^1 is endowed with $\langle \varphi, \psi \rangle_{H^1} = \langle \varphi, \psi \rangle_{L^2} + \langle \varphi_x, \psi_x \rangle_{L^2}$ as inner product and associated norm. The inner product on H_0^1 is taken to be $\langle \varphi, \psi \rangle_{H_0^1} = \langle \varphi_x, \psi_x \rangle_{L^2}$. For every $a \in \text{dom } F$, an operator $A(a) : H_0^1 \cap H^2 \to L^2$ is defined by

$$A(a)u = -(au_x)_x + cu.$$

It is well known that there exists a constant $c_1 > 0$ depending only on k in the definition of dom F, such that

$$\|v\|_{H^1}^2 \leq c_1 \langle A(a)v, v \rangle_{L^2} \tag{30}$$

for all $v \in H_0^1 \cap H^2$ and $a \in \text{dom } F$. In pariticular this implies that

$$\|u(a)\|_{H^1} \leq c_1 \|f\|_{H^{-1}} \quad \text{for all } a \in \text{dom } F. \tag{31}$$

For a_1 and a_2 in dom F one obtains

$$A(a_1)\Big(u(a_1) - u(a_2)\Big) = ((a_1 - a_2)u(a_2)_x)_x,$$

and by (30)

$$\|u(a_1) - u(a_2)\|_{H^1}^2 \leq c_1 \|a_1 - a_2\|_{L^\infty} \|u(a_2)_x\|_{L^2} \|u(a_1) - a(a_2)\|_{H^1}.$$

Employing (31) one has

$$\|u(a_1) - u(a_2)\|_{H^1} \leq \sqrt{2} c_1^2 \|a_1 - a_2\|_{H^1} \|f\|_{H^{-1}} \tag{32}$$

and the global Lipschitz continuity of F from dom $F \subset H^1$ to H^1 follows.

4.1

We shall first demonstrate the applicability of the results of Sections 2 and 3 for the choice

$$X = H^1; \quad Y = Z = L^2.$$

In this case D is a bounded linear operator and (A1) trivially holds.

ad (A2). It suffices to argue the weak sequential closedness of F. Let $a_n \rightharpoonup a$ in H^1 and $u(a_n) \rightharpoonup u$ in L^2 with $a_n \in \text{dom } F$. It follows that for some subsequence of $\{a_n\}$ again denoted by the same symbol, $a_n \to a$ in the maximum norm and $u(a_n) \rightharpoonup u$ in H^1 by (31). Consequently $a \in \text{dom } F$ and, using the weak form of (29) one obtains $u(a) = u$. This gives the weak closure of F.

ad (A3). This is an obvious consequence of the properties of dom F and D.

ad (A4). The operator D is bounded, surjective, with kernel consisting of all constant functions on $(0,1)$. Due to Lemma 1 (A4) follows.

We turn to conditions 1.–7. of Theorem 6 and commence with

ad condition 7. It must be assumed that $\delta \to 0$, $\alpha \sim \delta$ and $\eta \sim \delta^2$. If in addition $F(a) = u_0$ has a unique solution a_0, then Theorem 4 implies $a_\alpha^{\delta,\eta} \to a_0$ in H^1 as $\delta \to 0$. Otherwise we restrict our attention to a solution a_0 of $F(a) = u_0$ and a sequence $a_\alpha^{\delta,\eta}$ with $a_\alpha^{\delta,\eta} \to a_0$ as $\delta \to 0$.

ad conditions 2. and 4. The Fréchet derivative of F can be characterized with standard techniques, see [1] for example. It is found to be

$$F'(a)h = A^{-1}(a)(hu(a)_x)_x, \quad \text{for } a \in \text{dom } F \text{ and all } h \in H^1.$$

Next let a_1 and a_2 be in dom F and let $h \in H^1$. Then

$$\begin{aligned}
\|(F'(a_1) - F'(a_2))h\|_{L^2} &\le \|[A^{-1}(a_1) - A^{-1}(a_2)](hu(a_1)_x)_x\|_{L^2} \\
&\quad + \|A^{-1}(a_2)\Big(hu(a_1)_x - hu(a_2)_x\Big)_x\|_{L^2} \\
&\le \sqrt{2}c_1^2\|a_1 - a_2\|_{H^1}\|hu(a_1)_x\|_{L^2} \\
&\quad + c_1\|h(u(a_1)_x - u(a_2)_x)\|_{L^2},
\end{aligned}$$

where we used (32) and (31). Employing (32) and (31) once again gives

$$\|F'(a_1)h - F'(a_2)h\|_{L^2} \le 4c_1^3\|a_1 - a_2\|_{H^1}\|h\|_{H^1}\|f\|_{H^{-1}} \tag{33}$$

and condition 4. follows.

ad condition 3. Suppose that $g \in \ker D \cap \ker F'(a_0)$. This implies that g is constant and $gA^{-1}(a_0)(u(a_0)_{xx}) = 0$. If $g \neq 0$ then this implies that $u(a_0)_{xx} = 0$, which can happen only if $f = 0$, which is excluded. Since $\dim(\ker D) = 1$, there clearly exists $\beta > 0$ with $\|F'(a_0)g\|_{L^2} \ge \beta\|g\|_{H^1}$ for all $g \in \ker D$.

ad condition 5. This condition requires the existence of $\omega \in L^2$ such that

$$D^*D(a_0 - a^*) = F'(a_0)^*\omega. \tag{34}$$

To interpret (34) we first observe that

$$F'(a_0)^*\omega = -\mathcal{N}^{-1}\Big(u(a_0)_x(A(a_0)^{-1}\omega)_x\Big), \tag{35}$$

where $\mathcal{N}^{-1} : (H^1)^* \to H^1$ is the solution operator of the Neumann problem given by $\mathcal{N}^{-1}\psi = \varphi$ with φ the unique variational solution of

$$\langle \varphi_x, v_x \rangle_{L^2} + \langle \varphi, v \rangle_{L^2} = \langle \psi, v \rangle_{L^2} \quad \text{for all } v \in H^1$$

The element $b = D^*D(a_0 - a^*) \in H^1$ on the left hand side of (34) is caracterized as the unique solution of the variational equation

$$\langle b_x, v_x \rangle_{L^2} + \langle b, v \rangle_{L^2} = \langle (a_0 - a^*)_x, v_x \rangle_{L^2} \quad \text{for all } v \in H^1.$$

Thus $b = D^*D(a_0 - a^*)$ is the variational solution of

$$\begin{aligned} &-b_{xx} + b = -(a_0 - a^*)_{xx} \\ &b_x(0) = (a_0 - a^*)_x(0), \quad b_x(1) = (a_0 - a^*)_x(1). \end{aligned}$$

The operator $D^*D : H^1 \to H^1$ does not regularize $a_0 - a^*$. On the other hand, since $u(a_0)_x(A(a_0)^{-1}\omega)_x \in H^1$, the element on the right hand side of (35) is in H^3, and hence (34) represents a regularity condition on $a_0 - a^*$. We point out that the inverse Neumann operators on the left and the right hand side of (34) do not cancel out since they involve different variational boundary conditions.

ad condition 6. This is a smallness condition on $a_0 - a^*$ in H^1. Alternatively, in view of (33), this condition holds if $\|f\|_{H^{-1}}$ is sufficiently small or k is sufficiently large.

Remark 8.

The upper bound on the elements $a \in \text{dom } F$ was used only to ascertain (A3). We now argue that the applicapility of Theorem 6 remains valid for dom $F = \{a \in H^1 \,|\, 0 < k \leq a(x),\, x \in [0,1]\}$. In fact, (A3) was used for Theorem 6 in the application of Theorem 4, which provides convergence of $a_\alpha^{\delta,\eta} \to a_0$. Returning to the proof of Theorem 4, one notices that (A3) was employed to argue the boundedness of $a_k = a_{\alpha_k}^{\delta_k,\eta_k}$ in H^1. Every a_k can be decomposed as $a_k = a_k^{(1)} + a_k^{(2)}$, where $a_k^{(1)} = \int_0^1 a_k\, dx$ and $\int_0^1 a_k^{(2)} dx = 0$. As in the proof of Theorem 4 we find that $\{Da_k\} = \{Da_k^{(2)}\}$ is bounded in L^2, hence $\{a_k^{(2)}\}$ is bounded in H^1. If $\{a_k^{(1)}\}$ were unbounded, then $u(a_k) \to 0$ in H^1. This would imply that $\|y_0\|_{L^2}^2 = \|u(a_0)\|_{L^2}^2 = \lim_{k\to\infty} \|u(a_k) - y_k\|_{L^2}^2 \leq 0$, by (13). Since $u(a_0) \neq 0$, as a consequence of $f \neq 0$ this is a contradiction. The remaining part of the proof of Theorem 4 remains unchanged.

4.2

We now choose

$$X = Y = Z = L^2.$$

In this case $D : H^1 \subset L^2 \to L^2$ is a closed, densely defined, surjective linear operator. As before the kernel consists of all constant functions. (A1)–(A4) as well as the conditions of Theorem 6 are also satisfied for this choice of the spaces X, Y and Z. We restrict ourselves to discuss only the changes that need to be made when compared to the arguments that we gave in 4.1.

ad (A2). Let $a_n \rightharpoonup a$ in L^2 and $(u(a_n), Da_n) \rightharpoonup (u, b)$ in $L^2 \times L^2$. Since D is closed in L^2 it follows that $a \in H^1$, $Da = b$ and $a_n \rightharpoonup a$ in H^1. It can now be verified as in 4.1 that $a \in \operatorname{dom} F$ and $u(a) = u$. We point out that F itself is not weakly sequentially closed.

ad (A3). Here Remark 8 applies.

ad condition 5. This condition requires the existence of $\omega \in L^2$ such that

$$D^*D(a_0 - a^*) = F'(a_0)\omega. \tag{36}$$

The adjoint D^* of D is given by $\operatorname{dom} D^* = \{a \in H^1 \,|\, a(0) = a(1) = 0\}$, and $D^*a = -a_x$. Defining the operator N in L^2 by $\operatorname{dom} N = \{a \in H^2 \,|\, a_x(0) = a_x(1) = 0\}$ and $Na = -a_{xx}$, we find that (36) requires that $a_0 - a^* \in \operatorname{dom} N$, and ω satisfies (36) if and only if

$$N(a_0 - a^*) = -u(a_0)_x \Big(A(a_0)^{-1}\omega\Big)_x. \tag{37}$$

For (37) to hold with $\omega \in L^2$ it is necessary that $a_0 - a^* \in H^3$ and that $\Big\langle u(a_0)_x \Big(A(a_0)^{-1}\omega\Big)_x, 1\Big\rangle_{L^2} = 0$.

Comparing the conclusions that can be drawn from Theorem 6 for the choice of spaces in 4.1 and 4.2 we observe that different conditions for ω were obtained under which a convergence rate result in H^1 holds. The necessary regularity condition is $a_0 - a^* \in H^3$ in either case. On a technical level the extra work due to considering unbounded operators in Sections 2 and 3 is rewarded by a simpler characterization of the adjoint operators in Section 4.2.

References

[1] H.W. Engl, K. Kunisch, and A. Neubauer. Convergence rates for Tikhonov regularisation of non-linear ill-posed problems. *Inverse Problems*, 5:523–540, 1989.

[2] C.W. Groetsch. *Generalized Inverses of Linear Operators.* Volume 37 of *Pure and Applied Mathematics*, Marcel Dekker, New York, 1977.

[3] J. Locker and P. M. Prenter. Regularization with differential operators I. general theory. *Journal of Mathematical Analysis and Applications*, 74:504–529, 1980.