Otmar Scherzer

A parameter choice for Tikhonov regularization for solving nonlinear inverse problems leading to optimal convergence rates

Applications of Mathematics, Vol. 38 (1993), No. 6, 479–487

Persistent URL: http://dml.cz/dmlcz/104570

38 (1993) APPLICATIONS OF MATHEMATICS No. 6, 479–487

A PARAMETER CHOICE FOR TIKHONOV REGULARIZATION FOR SOLVING NONLINEAR INVERSE PROBLEMS LEADING TO OPTIMAL CONVERGENCE RATES

Otmar Scherzer,* Linz

Summary. We give a derivation of an a–posteriori strategy for choosing the regularization parameter in Tikhonov regularization for solving nonlinear ill-posed problems, which leads to optimal convergence rates. This strategy requires a special stability estimate for the regularized solutions. A new proof for this stability estimate is given.

We study nonlinear ill-posed problems of the form

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

where $F\colon D(F) \subset X \to Y$ is a nonlinear operator between Hilbert spaces X and Y. As the notion of a "solution" of the equation (1), we choose the concept of an x^*-minimum-norm-solution x_0 (x^*-M.N.S.), i.e.,

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

and

$$\|x_0 - x^*\| = \min_{x \in D(F)} \{\|x - x^*\|\colon F(x) = y_0\}. \tag{3}$$

In the following, we always assume the existence of an x^*-M.N.S. for exact data y_0.

If (1) is ill-posed in the sense of lack of continuity of its solutions with respect to the data, regularization techniques are required. Tikhonov regularization has been investigated in [2], [5], [7] to solve nonlinear ill-posed problems in a stable manner. In Tikhonov regularization, a solution of problem (1) is approximated by a solution of the minimization problem

$$\min_{x \in D(F)} \{\|F(x) - y_\delta\|^2 + \alpha\|x - x^*\|^2\}, \tag{4}$$

* Supported by the Christian-Doppler Society

where $\alpha > 0$ is small parameter and $y_\delta \in Y$ is the available noisy data, for which we have the additional information that

$$\|y_\delta - y_0\| \leqslant \delta. \tag{5}$$

Convergence and rates of convergence results for this method were developed in [2] and [5]. In these papers, the regularization parameter α is chosen in dependence of the noise level δ and smoothness assumptions on x_0. The disadvantage of a-priori strategies as discussed in [2] and [5] is that one cannot check in general whether these smoothness assumptions are fulfilled, and this has the consequence that wrong smoothness assumptions lead to a bad choice of α and consequently to a bad approximation of the x^*-M.N.S. On contrast a-posteriori strategies determine the regularization parameter from quantities that arise during calculations. The most widely used a-posteriori strategy is "Morozov's Discrepancy Principle". In this method the regularization parameter $\alpha(\delta)$ is determined by

$$\|F(x_\alpha^\delta) - y_\delta\|^2 = \delta^2. \tag{6}$$

The disadvantage of Morozov's Discrepancy Principle is, as one sees from the linear case [4], that the regularized solutions converge (beside in trivial cases) at most like $O(\sqrt{\delta})$. We outline the basic idea of an a-posteriori parameter choice strategy, which leads to (quasi-) optimal convergence rates [6]: By x_α^δ, we denote regularized solutions, i.e. any solution of the minimization problem (4), and by x_α, any solution of (4) with the perturbed data y_δ replaced by the exact date y_0. The total error between the regularized solution x_α^δ and the x^*-M.N.S. x_0 can be estimated as follows:

$$\frac{1}{2}\|x_\alpha^\delta - x_0\|^2 \leqslant \|x_\alpha - x_0\|^2 + g^2(\alpha, \delta), \tag{7}$$

where $g(\alpha, \delta)$ is a stability estimate, i.e. a bound for the term $\|x_\alpha^\delta - x_\alpha\|$, which depends only on the regularization parameter α, the noise level δ but not on the special feature of y_δ. The idea of the strategy (proposed in [1], [3] for the case of a linear operator, applied to nonlinear operators) is to choose $\alpha := \alpha(\delta)$ such that the right hand side of (7) is minimized with respect to α. This is achieved by putting the derivative with respect to α equal to 0. Of course, this strategy is not implementable in this form, due to the fact that minimizing the right hand side of (7) would require knowledge of the unknown x^*-M.N.S. x_0 and of the unknown exact data y_0. To make this strategy implementable, the term coming from the differentiation of $\|x_\alpha - x_0\|^2$ has to be approximated by computable terms in such a way that the asymptotic behavior remains unchanged.

A strategy which yields an optimal rate of the regularized solutions requires a choice $g(\alpha, \delta)$ of the form $\sqrt{c\frac{\delta^2}{\alpha}}$, as one sees from the linear case. In the nonlinear case, such a stability estimate seems to be not possible in general.

Throughout this paper it is assumed that F is weakly (sequentially) closed, continuous and Fréchet-differentiable with convex domain $D(F)$. This implies in particular the existence of regularized solutions x_α and x_α^δ. Moreover, we assume that x_0 is an interior point of $D(F)$, i.e., that $\{x \in \|x - x^*\| \leqslant k\|x_0 - x^*\|\} \subseteq D(F)$, with $k > \sqrt{2}$; this assumption guarantees that for $\frac{\delta^2}{\alpha} \leqslant \|x_0 - x^*\|^2$ both x_a^δ and x_α are interior points.

The following assumptions are used in [6] to guarantee that $x_\alpha^\delta \to x_\alpha$ like $O(\frac{\delta}{\sqrt{\alpha}})$.

— There exists a constant K_0 such that for every $(x, z, v) \in D(F) \times D(F) \times X$ there is a $k(x, z, v) \in X$ such that

$$(8) \qquad \big(F'(x) - F'(z)\big)v = F'(z)k(x, z, v),$$

where

$$(9) \qquad \|k(x, z, v)\|_X \leqslant K_0\|v\|_X\|x - z\|_X.$$

— Moreover, there exist constants K_1, K_2 such that for every $(x, z, y) \in D(F) \times D(F) \times Y$ there are $l_1(x, z, y) \in Y$, $l_2\big(x, z, F'(x)^*y\big) \in X$ such that

$$(10) \qquad \big(F'(x)^* - F'(z)^*\big)y = F'(z)^*l_1(x, z, y) + l_2\big(x, z, F'(x)^*y\big),$$

where

$$(11) \qquad \begin{aligned} \|l_1(x, z, y)\|_Y &\leqslant K_1\|y\|_Y\|x - z\|_X, \\ \big\|l_2\big(x, z, F'(x)^*y\big)\big\|_X &\leqslant K_2\|F'(x)^*y\|_X\|x - z\|_X. \end{aligned}$$

Examples illustrating (8)–(11) can be found in [6]. A further assumption in [6] was that

$$(12) \qquad \frac{\delta^2}{\alpha} \leqslant \|x_0 - x^*\|^2.$$

For fixed $\alpha > 0$ this condition is satisfied for all sufficiently small values of δ. For asymptotic results where $\alpha = \alpha(\delta) \to 0$ as $\delta \to 0$, (12) is satisfied if $\frac{\delta^2}{\alpha(\delta)} \to 0$ and if δ is sufficiently small. This is a well-known sufficient condition for convergence of $x_\alpha^\delta \to x_0$ [2], [7]. Condition (12) can be replaced by "$\frac{\delta^2}{\alpha}$ sufficiently small".

In the following we will show that a stability estimate $\|x_\alpha^\delta - x_\alpha\| \leqslant O(\frac{\delta}{\sqrt{\alpha}})$ even holds under weaker assumptions than those proposed in [6]:

Theorem 1 (Stability Estimate). *Let F' be Lipschitz-continuous in x_0, i.e., there exists $L > 0$ such that*

$$\|F'(x_0) - F'(z)\| \leqslant L\|x_0 - z\|, \qquad \textit{for all } z \in D(F).$$

Moreover, let one of the following assumptions (I), (II) *hold:*

(I) *Let* (8), (9) *and* $\|x_0 - x^*\|$ *sufficiently small;*

or

(II) *there exists a* $w \in Y$ *satisfying* $x_0 - x^* = F'(x_0)^* w$, *where* $2L\|w\| < 1$.

Then,

$$\|x_\alpha^\delta - x_\alpha\| \leqslant O\Big(\frac{\delta}{\sqrt{\alpha}}\Big).$$

Proof. In this proof we will make use of the first order necessary optimality conditions for the minima x_α, x_α^δ of $\|F(x)-y_0\|^2+\alpha\|x-x^*\|^2$, $\|F(x)-y_\delta\|^2+\alpha\|x-x^*\|^2$, i.e.,

$$F'(x_\alpha)^*\big(F(x_\alpha)-y_0\big)+\alpha(x_\alpha-x^*)=0,$$
$$F'(x_\alpha^\delta)^*\big(F(x_\alpha^\delta)-y_\delta\big)+\alpha(x_\alpha^\delta-x^*)=0.$$

Since

$$\|F(x_\alpha^\delta)-y_\delta\|^2+\alpha\|x_\alpha^\delta-x^*\|^2 \leqslant \|F(x_\alpha)-y_\delta\|^2+\alpha\|x_\alpha-x^*\|^2,$$

$$\begin{aligned}
(13)\quad & \|F(x_\alpha^\delta)-F(x_\alpha)\|^2+\alpha\|x_\alpha^\delta-x_\alpha\|^2 \\
& \leqslant \|F(x_\alpha^\delta)-y_\delta\|^2+\|F(x_\alpha^\delta)-F(x_\alpha)\|^2-\|F(x_\alpha)-y_\delta\|^2 \\
& \quad +\alpha(\|x_\alpha-x^*\|^2+\|x_\alpha^\delta-x_\alpha\|^2-\|x_\alpha^\delta-x^*\|^2) \\
& = 2\big(F(x_\alpha)-y_\delta, F(x_\alpha)-F(x_\alpha^\delta)\big)+2\alpha(x_\alpha^\delta-x_\alpha, x^*-x_\alpha) \\
& \leqslant 2\big(F(x_\alpha)-y_0, F(x_\alpha)-F(x_\alpha^\delta)\big)+2\big(x_\alpha^\delta-x_\alpha, F'(x_\alpha)^*\big(F(x_\alpha)-y_0\big)\big) \\
& \quad +2\delta\|F(x_\alpha)-F(x_\alpha^\delta)\| \\
& = -2\big(F(x_\alpha)-y_0, F(x_\alpha^\delta)-F(x_\alpha)-F'(x_\alpha)(x_\alpha^\delta-x_\alpha)\big) \\
& \quad +2\delta\|F(x_\alpha)-F(x_\alpha^\delta)\|.
\end{aligned}$$

In the first case we obtain from (8) and (9):

$$\begin{aligned}
& \big|\big(F(x_\alpha)-y_0, F(x_\alpha^\delta)-F(x_\alpha)-F'(x_\alpha)(x_\alpha^\delta-x_\alpha)\big)\big| \\
& = \Big|\Big(F(x_\alpha)-y_0, \int_0^1 \big(F'(x_\alpha+t(x_\alpha^\delta-x_\alpha))-F'(x_\alpha)\big)\,\mathrm{d}t(x_\alpha^\delta-x_\alpha)\Big)\Big| \\
& = \Big|\Big(F(x_\alpha)-y_0, F'(x_\alpha)\int_0^1 k\big(x_\alpha+t(x_\alpha^\delta-x_\alpha), x_\alpha, x_\alpha^\delta-x_\alpha\big)\,\mathrm{d}t\Big)\Big| \\
& \leqslant \alpha\frac{K_0}{2}\|x_\alpha-x^*\|\,\|x_\alpha^\delta-x_\alpha\|^2.
\end{aligned}$$

Together with (13) follows

$$\|F(x_\alpha^\delta)-F(x_\alpha)\|^2+\alpha\|x_\alpha^\delta-x_\alpha\|^2 \leqslant K_0\alpha\|x_\alpha-x^*\|\,\|x_\alpha^\delta-x_\alpha\|^2+2\delta\|F(x_\alpha^\delta)-F(x_\alpha)\|.$$

Therefore,

$$\|F(x_\alpha^\delta)-F(x_\alpha)\|^2+\alpha(1-K_0\|x_\alpha-x^*\|)\|x_\alpha^\delta-x_\alpha\|^2 \leqslant 2\delta\|F(x_\alpha^\delta)-F(x_\alpha)\|.$$

If $q := K_0\|x_0-x^*\| < 1$, then $K_0\|x_\alpha-x^*\| \leqslant q$. Thus

$$\|F(x_\alpha^\delta)-F(x_\alpha)\| \leqslant 2\delta,$$

and consequently

$$\|x_\alpha^\delta-x_\alpha\|^2 \leqslant \frac{4}{1-K_0\|x_0-x^*\|}\,\frac{\delta^2}{\alpha}.$$

In the second case it follows from (13) that

$$\|F(x_\alpha^\delta)-F(x_\alpha)\|^2+\alpha\|x_\alpha^\delta-x_\alpha\|^2 \leqslant L\|F(x_\alpha)-y_0\|\,\|x_\alpha^\delta-x_\alpha\|^2+2\delta\|F(x_\alpha^\delta)-F(x_\alpha)\|.$$

Following the proof of Theorem 2.4 in [2] one finds

$$\|F(x_\alpha)-y_0\| \leqslant 2\|w\|\alpha.$$

Thus

$$\|F(x_\alpha^\delta)-F(x_\alpha)\|^2+(1-2L\|w\|)\|x_\alpha^\delta-x_\alpha\|^2 \leqslant 2\delta\|F(x_\alpha^\delta)-F(x_\alpha)\|.$$

Therefore,

$$\|F(x_\alpha^\delta)-F(x_\alpha)\| \leqslant 2\delta,$$

and thus

$$\|x_\alpha^\delta-x_\alpha\|^2 \leqslant \frac{4}{1-2L\|w\|}\,\frac{\delta^2}{\alpha}.$$

□

Our parameter choice strategy is based on minimizing the right hand side of (7), where we use $g^2(\alpha,\delta) = c\frac{\delta^2}{\alpha}$. Since we do not know the x^*-M.N.S. x_0 and the exact data y_0, we have to approximate the right hand side of this inequality in an appropriate way by computable terms. For the moment, we assume that we have unperturbed data y_0. The effect of perturbed data is taken into account afterwards.

It follows from the first order necessary optimality condition for a minimum of (4) (if we replace the perturbed data y_δ by the unperturbed data y_0) that

$$F'(x_\alpha)^*\big(F(x_\alpha)-y_0\big)+\alpha(x_\alpha-x^*)=0. \tag{14}$$

Formal differentiation of the equation (14) with respect to α yields

$$F'(x_\alpha)^*F'(x_\alpha)\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha}+\big(F'(x_\alpha)^*\big)'\big(F(x_\alpha)-y_0\big)\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha}+\alpha\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha}=-(x_\alpha-x^*), \tag{15}$$

or equivalently,

$$\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha}=-\Big(\alpha I+F'(x_\alpha)^*F'(x_\alpha)+\big(F'(x_\alpha)^*\big)'\big(F(x_\alpha)-y_0\big)\Big)^{-1}(x_\alpha-x^*), \tag{16}$$

if the operator $\big(\alpha I+F'(x_\alpha)^*F'(x_\alpha)+\big(F'(x_\alpha)^*\big)'\big(F(x_\alpha)-y_0\big)\big)$ is positive definite. Taking the weak form of (15), using the special test function $x_\alpha - x_0$ and (14) we obtain

$$\begin{aligned}\big(F(x_\alpha)-y_0,F'(x_\alpha)(x_\alpha-x_0)\big)&=\alpha\Big(\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha}\alpha,x_\alpha-x_0\Big)\\&\quad+\alpha\Big(F'(x_\alpha)\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha},F'(x_\alpha)(x_\alpha-x_0)\Big)\\&\quad+\alpha\Big(F(x_\alpha)-y_0,F''(x_\alpha)\Big(\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha},x_\alpha-x_0\Big)\Big).\end{aligned} \tag{17}$$

The first order necessary optimality condition for the minimum of the right hand side of (7), where we use $g^2(\alpha,\delta)=c\frac{\delta^2}{\alpha}$ is

$$c\delta^2=\Big(\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha}\alpha^2,x_\alpha-x_0\Big). \tag{18}$$

Inserting equation (18) into (17) yields

$$\begin{aligned}\big(F(x_\alpha)-y_0,F'(x_\alpha)(x_\alpha-x_0)\big)&=c\delta^2\\&\quad+\alpha\Big(F'(x_\alpha)\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha},F'(x_\alpha)(x_\alpha-x_0)\Big)\\&\quad+\alpha\Big(F(x_\alpha)-y_0,F''(x_\alpha)\Big(\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha},x_\alpha-x_0\Big)\Big).\end{aligned} \tag{19}$$

If we approximate $F'(x_\alpha)(x_\alpha - x_0)$ by $F(x_\alpha) - F(x_0)$ and neglect the inner product containing the second derivative of F, we obtain from (19)

$$\|F(x_\alpha) - y_0\|^2 - \alpha\Big(F(x_\alpha) - y_0, F'(x_\alpha)\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha}\Big) = c\delta^2. \tag{20}$$

If we replace y_0 by y_δ in (14), (16) and (20) (and consequently x_α by x_α^δ) then we arrive at an implementable strategy: All terms occurring then in the equation (20) are computable.

For reasons of computational complexity, we propose one additional approximation and drop the second derivative term also in the formula (16) for $\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha}$:

$$\frac{\mathrm{d}x_\alpha}{\mathrm{d}\alpha} \sim -\big(\alpha I + F'(x_\alpha)^* F'(x_\alpha)\big)^{-1}(x_\alpha - x^*). \tag{21}$$

Using this further approximation we obtain instead of (20)

$$\|F(x_\alpha)-y_0\|^2+\alpha\big(F(x_\alpha)-y_0, F'(x_\alpha)\big(\alpha I+F'(x_\alpha)^* F'(x_\alpha)\big)^{-1}(x_\alpha-x^*)\big) = c\delta^2. \tag{22}$$

Using (14), we can write (22) also as

$$\begin{aligned}\|F(x_\alpha) - y_0\|^2 - \big(F(x_\alpha) - y_0, \big(\alpha I + F'(x_\alpha)F'(x_\alpha)^*\big)^{-1} \\ \times F'(x_\alpha)F'(x_\alpha)^* \left(F(x_\alpha - y_0)\right) - c\delta^2\end{aligned} \tag{23}$$

or equivalently

$$\alpha\big(F(x_\alpha) - y_0, \big(\alpha I + F'(x_\alpha)F'(x_\alpha)^*\big)^{-1}\big(F(x_\alpha) - y_0\big)\big) = c\delta^2. \tag{24}$$

Replacing y_0 by y_δ and consequently x_α by x_α^δ we obtain the following equation for α:

$$\alpha\big(F(x_\alpha^\delta) - y_0, \big(\alpha I + F'(x_\alpha^\delta)F'(x_\alpha^\delta)^*\big)^{-1}\big(F(x_\alpha^\delta) - y_0\big)\big) = c\delta^2. \tag{25}$$

This is now our proposed parameter choice strategy.

In [6], for the proof that the strategy (25) yields an optimal convergence rate, the problem of solving $F(x) = y_0$ is compared with the problem of solving the linearized problem $F'(x_0)x = F'(x_0)x_0$. It seems to be reasonable that the regularized solutions of the original nonlinear problem and of its linearization have some relation with each other, if the operator F is not "too nonlinear". Assumptions on the operator F which guarantee such a relation are used in [6] to prove the following Theorem:

Theorem 2. *Let c (as in (25)) be chosen appropriately (see [6]) and $\|x_0 - x^*\|$ sufficiently small. If $R(F'(x_0))$ is not closed and $F'(x_0)(x_0 - x^*) \neq 0$, then the* *parameter choice strategy presented in* (25) *is of quasi-optimal order (see* [1] *for the definition of quasi-optimal order). If, in addition, there exists a decreasing sequence* (λ_k) *in* $\sigma\big(F'(x_0)F'(x_0)^*\big)$ *with*

$$\lim_{k\to\infty} \lambda_k = 0 \quad \textit{and} \quad \sup\{\lambda_k\lambda_{k+1}^{-1}\colon k\in\mathbb{N}\} < \infty, \tag{26}$$

then this parameter choice strategy is of optimal order (compare [1] *for the definition of optimal order).*

Example 3. Here we estimate a in the differential equation

$$\begin{aligned} -(au_x)_x &= -e^x, \\ u(0) = 1,&\ u(1) = e. \end{aligned} \tag{27}$$

This problem of estimating a can be formulated via an operator equation

$$F(a_0) = u_0,$$

where F is the "parameter to solution map" of the differential equation (27), i.e.

$$F\colon D(F) := \{a\in H^1]0,1[\colon a(x)\geqslant\gamma>0\}\to L^2.$$
$$a\to u(a)$$

It can be shown that F is weakly closed, continuous, Fréchet-differentiable and that the Fréchet-derivative is Lipschitz-continuous in a neighbourhood of a_0. The Tikhonov functional for this problem reads as follows

$$\|F(a)-u_\delta\|_{L^2}^2 + \alpha\|a-a^*\|_{H^1}^2,$$

where we denote be y_δ given perturbed data of u_0.

If $u_0 = e^x$, then exact parameter is unique and is given by $a_0 = 1$ if a^* is chosen as below. Instead of u_0, we used in our calculations u_δ, where u_δ was a high-frequency) perturbation of u_0 with $\|u_0-u_\delta\|_{L^2}\leqslant\delta$. We used the special perturbation $u_\delta = u_0+\delta\sqrt{2}(10\pi y)$.

Let $a^* = 1+0.05(ay^5+by^4+cy^3+dy^2+ey+f)$, where $a = 1.428065\ldots$, $b = -4.381893049\ldots$, $c = 1.043027472\ldots$, $d = 3.629082416\ldots$, $e = 0$, $f = 1$. The coefficients a, b, c, d, e, f were calculated using MAPLE and are chosen such that $a_0 - a^* \in R\big(F'(a_0)^*F'(a_0)\big)$. In this case the best possible convergence rate for Tikhonov regularization to be expected is $O\big(\delta^{\frac{2}{3}}\big)$ [5].

The regularized solutions a_α^δ were obtained by minimizing the Tikhonov functional on the finite dimensional subspace of piecewise linear splines on a uniform grid with subinterval length $\frac{1}{16}$ with a damped Newton method.

The differential equation (27) was solved approximately with a Galerkin Method on the finite dimensional subspace of piecewise linear splines on a uniform grid with subinterval length $\frac{1}{16}$.

In the Table below one can see the results obtained by using the parameter choice strategy (25) with $c = 1$. The realization of the strategy (25) involves the solution of a nonlinear equation. To minimize the numerical error for this step, a bisection technique on a fine grid was used.

In [5] it has been proved that if $a_0 - a^* \in R(F'(a_0)^* F'(a_0))$ and if $\|a_0 - a^*\|_{H^1}$ is sufficiently small, then the regularized solutions converge to a_0 with a rate $O(\delta^{\frac{2}{3}})$. This convergence rate can be seen in the numerical example.

For more numerical computations and a comparison of Morozov's Discrepancy Principle with the strategy (25) see [6].

Table

δ	α	$e := \|a_\alpha^\delta - a_0\|_{H^1}$	$\delta/\alpha^{\frac{3}{2}}$	$e/\delta^{\frac{2}{3}}$
$0.46e-2$	$0.12e-1$	$0.27e-1$	$0.35e+1$	$0.98e+0$
$0.46e-3$	$0.26e-2$	$0.81e-2$	$0.35e+1$	$0.14e+1$
$0.46e-4$	$0.33e-3$	$0.17e-2$	$0.77e+1$	$0.13e+1$
$0.46e-5$	$0.73e-4$	$0.45e-3$	$0.74e+1$	$0.16e+1$
$0.46e-6$	$0.12e-4$	$0.93e-4$	$0.11e+2$	$0.16e+1$

References

[1] *H. W. Engl, H. Gfrerer*: A posteriori parameter choice for general regularization methods for solving linear ill-posed problems, Appl. Num. Math. *4* (1988), 395–417.

[2] *H. W. Engl, K. Kunisch, A. Neubauer*: Convergence rates for Tikhonov regularization of nonlinears ill-posed problems, Inverse Problems *5* (1989), 523–540.

[3] *H. Gfrerer*: An a-posteriori parameter choice for ordinary and iterated Tikhonov regularization of ill-posed problems leading to optimal convergence rates, Mathematics of Computation *49* (1987), 507–522.

[4] *C. W. Groetsch*: The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind, Pitman, Boston, 1984.

[5] *A. Neubauer*: Tikhonov regularization for non-linear ill-posed problems: optimal convergence rates and finite-dimensional approximation, Inverse Problems *5* (1989), 541–557.

[6] *O. Scherzer, H. W. Engl, K. Kunisch*: Optimal a-posteriori parameter choice for Tikhonov regularization for solving nonlinear ill-posed problems, SIAM J. on Numer. Anal., to appear.

[7] *T. I. Seidman, C. R. Vogel*: Well-posedness and convergence of some regularization methods for nonlinear ill-posed problems, Inverse Problems *5* (1989), 227–238.

Author's address: *Otmar Scherzer*, Institut für Mathematik, Johannes Kepler Universität Linz, A-4040 Linz, Austria.