Proc. Estonian Acad. Sci. Phys. Math., 2004, **53**, 2, 116–123

On order optimal regularization under general source conditions

Ulrich Tautenhahn

Department of Mathematics, University of Applied Sciences Zittau/Görlitz, P.O. Box 1454, 02754 Zittau, Germany; u.tautenhahn@hs-zigr.de

Received 20 November 2003

Abstract. We study the problem of solving ill-posed problems with linear operators acting between Hilbert spaces, where instead of exact data noisy data with a known noise level are given. Regularized approximations are obtained by a general regularization scheme. Assuming the unknown solution belongs to some general source set, we prove that the regularized approximations are order optimal on this set provided the regularization parameter is chosen either *a priori* or *a posteriori* by the Raus–Gfrerer rule or the monotone error rule. Our results cover the special cases of finitely and infinitely smoothing operators.

Key words: ill-posed problems, regularization, *a priori* parameter choice, *a posteriori* rules, order optimal error bounds, general source conditions.

1. INTRODUCTION

In this paper we are interested in the minimum-norm solution $x^\dagger \in X$ of the ill-posed problem

$$Ax = y, \tag{1}$$

where $A \in \mathcal{L}(X,Y)$ is a linear, injective, and bounded operator with non-closed range $R(A)$ of A, and X, Y are Hilbert spaces. Throughout the paper we assume that $y^\delta \in Y$ are the available noisy data with

$$\|y - y^\delta\| \le \delta \tag{2}$$

and a known noise level δ. The numerical treatment of ill-posed problems (1), (2) requires the application of special regularization methods. A large class of such methods fit into a *general regularization scheme* in which a regularized approximation x_α^δ is given by

$$x_\alpha^\delta = g_\alpha(A^*A)A^*y^\delta \tag{3}$$

with some properly chosen operator function g_α. Note that the choice $g_\alpha(\lambda) = \frac{1}{\lambda+\alpha}$ in (3) leads to the method of Tikhonov regularization [1].

The paper is organized as follows. In Section 2 we discuss some facts on optimality and order optimality. In Sections 3 and 4 we prove that under certain conditions concerning g_α the regularized approximations (3) are order optimal on some general source set M provided the regularization parameter α is chosen either *a priori* or *a posteriori* by the Raus–Gfrerer rule or the monotone error rule.

2. OPTIMALITY AND ORDER OPTIMALITY

In order to guarantee convergence rates for $\|x_\alpha^\delta - x^\dagger\|$, source conditions are necessary. In this paper we are interested in order optimality results under *general source conditions* $x^\dagger \in M_{\varphi,E}$ with $M_{\varphi,E}$ given by

$$M_{\varphi,E} = \left\{ x \in X \,\middle|\, x = [\varphi(A^*A)]^{1/2} v\,,\ \|v\| \le E \right\} \tag{4}$$

and *general source functions* φ satisfying

Assumption A1. $\varphi : (0, a] \to (0, \infty)$ with $\|A^*A\| \le a$ is continuous and satisfies

(i) φ is strictly monotonically increasing with $\lim_{\lambda\to 0} \varphi(\lambda) = 0,$

(ii) ρ, implicitly defined by $\rho(\varphi(\lambda)) = \lambda\varphi(\lambda)$, is convex.

In (4) the function φ is well defined via $\varphi(A^*A) = \int_0^a \varphi(\lambda)\,\mathrm{d}E_\lambda$, where $A^*A = \int_0^a \lambda\,\mathrm{d}E_\lambda$ is the spectral resolution of A^*A and $\|A^*A\| \le a$. Typically, for Eqs. (1) with finitely smoothing operators, *polynomial source conditions* with $\varphi(\lambda) = \lambda^p$ and $p > 0$ are exploited (see [2−6]). For infinitely smoothing operators, polynomial source conditions are too restrictive. In this case it is natural to use *logarithmic source conditions* with $\varphi(\lambda) = \left[\ln \frac{1}{\lambda}\right]^{-p}$ and $p > 0$, see [7−9].

Any operator $R : Y \to X$ can be considered as a special method for solving (1). The approximate solution to (1) is then given by Ry^δ. Let us consider the *worst case error* of the method R defined by

$$\Delta(\delta, R) = \sup\left\{ \|Ry^\delta - x^\dagger\| \,\middle|\, x^\dagger \in M_{\varphi,E}\,,\ y^\delta \in Y\,,\ \|y - y^\delta\| \le \delta \right\}.$$

This worst case error characterizes the maximal error of the method R if the minimum-norm solution $x^\dagger$ of problem (1) varies in the set $M_{\varphi,E}$. An optimal method $R_{\rm opt}$ is characterized by $\Delta(\delta, R_{\rm opt}) = \inf_R \Delta(\delta, R)$. It can easily be realized that $\inf_R \Delta(\delta, R) \geq \omega(\delta, M_{\varphi,E})$ with

$$\omega(\delta, M_{\varphi,E}) = \sup\left\{ \|x\| \,\Big|\, x \in M_{\varphi,E}\,,\ \|Ax\| \leq \delta \right\} .$$

Under Assumption A1, the *modulus of continuity* $\omega(\delta, M_{\varphi,E})$ of the inverse operator A^{-1} on the set $M_{\varphi,E}$ can be estimated as follows.

Theorem 1. *Let $M_{\varphi,E}$ be given by* (4) *and let Assumption* A1 *be satisfied. If δ/E is sufficiently small such that $\delta^2/E^2 \leq a\varphi(a)$, then*

$$\omega(\delta, M_{\varphi,E}) \leq E\,\sqrt{\rho^{-1}(\delta^2/E^2)}\,. \tag{5}$$

If $\delta^2/E^2 \in \sigma(A^*A\varphi(A^*A))$, then there holds equality in (5).

In Theorem 1, $\sigma(A^*A)$ denotes the spectrum of the operator A^*A. For the proof of Theorem 1 in the case of compact operators A see [10], in the general case where the operator A is not necessarily compact see [9], and in the special case of source sets (4) with $\varphi(\lambda) = \lambda^p$ see [6]. Due to Theorem 1, any regularized approximation $x_\alpha^\delta = R_\alpha^\delta y^\delta$ for problem (1), (2) is called

(i) *optimal on the set* $M_{\varphi,E}$ if $\|x_\alpha^\delta - x^\dagger\| \leq E\,\sqrt{\rho^{-1}(\delta^2/E^2)}$,

(ii) *order optimal on the set* $M_{\varphi,E}$ if $\|x_\alpha^\delta - x^\dagger\| \leq c\,E\,\sqrt{\rho^{-1}(\delta^2/E^2)}$,

where $c \geq 1$ is a constant. Note that for polynomial source sets with $\varphi(\lambda) = \lambda^p$ estimate (5) attains the form $\omega(\delta, M_{\varphi,E}) \leq E^{1/(p+1)}\delta^{p/(p+1)}$ and that in the case of logarithmic source sets with $\varphi(\lambda) = \left[\ln\frac{1}{\lambda}\right]^{-p}$ estimate (5) provides $\omega(\delta, M_{\varphi,E}) \leq E\left[\ln\frac{E^2}{\delta^2}\right]^{-p/2}(1+o(1))$ for $\delta \to 0$.

3. *A PRIORI* PARAMETER CHOICE

We assume in this section that $g_\alpha : [0,a] \to \mathbb{R}$ with $\|A^*A\| \leq a$ is piecewise continuous with $\lim_{\alpha\to 0} g_\alpha(\lambda) = 1/\lambda$. For studying scheme (3), besides A1 the following additional assumption is required, which has been exploited in [6] for polynomial source sets, in [7] for logarithmic source sets, and in [11,12] for general source sets.

Assumption A2. There exist constants γ_1 and γ_2 such that

(i) $\displaystyle\sup_{\lambda\geq 0}\left|\sqrt{\lambda}g_\alpha(\lambda)\right| \leq \frac{\gamma_1}{\sqrt{\alpha}}$,

(ii) $\displaystyle\sup_{\lambda\geq 0}\left|1-\lambda g_\alpha(\lambda)\right|\sqrt{\varphi(\lambda)} \leq \gamma_2\sqrt{\varphi(\alpha)}\,.$

Our next theorem shows that the regularized approximations (3) provide order optimal error bounds provided α is chosen *a priori*.

Theorem 2. *Let $x^\dagger \in M_{\varphi,E}$ with φ satisfying Assumption A*1. *Let x_α^δ be the regularized approximation* (3) *satisfying Assumption A*2 *and let α be chosen by*

$$\alpha = \varphi^{-1}(\rho^{-1}(\delta^2/E^2)) . \tag{6}$$

Then,

$$\|x_\alpha^\delta - x^\dagger\| \leq (\gamma_1 + \gamma_2) E \sqrt{\rho^{-1}(\delta^2/E^2)} . \tag{7}$$

Proof. Due to the definition of ρ according to Assumption A1 (ii) and due to the choice of α according to (6), we have

$$\delta/\sqrt{\alpha} = E\sqrt{\rho^{-1}\left(\delta^2/E^2\right)} . \tag{8}$$

Let $x_\alpha = g_\alpha(A^*A)A^*y$. Exploiting Assumption A2 (i) and (8), we obtain

$$\|x_\alpha^\delta - x_\alpha\| \leq \delta \sup_{\lambda \geq 0} \left| \sqrt{\lambda} g_\alpha(\lambda) \right| \leq \gamma_1 \frac{\delta}{\sqrt{\alpha}} = E\gamma_1 \sqrt{\rho^{-1}(\delta^2/E^2)} . \tag{9}$$

Since $x^\dagger \in M_{\varphi,E}$, we have due to Assumption A2 (ii) and (6) that

$$\|x_\alpha - x^\dagger\| \leq E \sup_{\lambda \geq 0} \left| 1 - \lambda g_\alpha(\lambda) \right| \sqrt{\varphi(\lambda)} \leq E\gamma_2 \sqrt{\rho^{-1}(\delta^2/E^2)} . \tag{10}$$

Now (7) follows from (9), (10) and the triangle inequality. □

Let us discuss Assumption A2 for special regularization methods that fit into the framework of methods (3).

Example 3. (*Tikhonov regularization of order* m). These methods are characterized by (3) with $g_\alpha(\lambda) = \frac{1}{\lambda}\left[1 - \left(\frac{\alpha}{\lambda+\alpha}\right)^m\right]$. In this method, $x_\alpha^\delta := x_{\alpha,m}^\delta$ can be obtained by solving the m operator equations

$$(A^*A + \alpha I)x_{\alpha,k}^\delta = A^*y^\delta + \alpha x_{\alpha,k-1}^\delta , \quad k = 1, ..., m , \quad x_{\alpha,0}^\delta = 0 . \tag{11}$$

Assumption A2 (i) is satisfied with constant $\gamma_1 = \frac{1}{2}$ for $m = 1$ and $\gamma_1 = \sqrt{m}$ for $m \geq 2$ (see [6]). It can be shown that Assumption A2 (ii) holds true with $\gamma_2 = 1$ provided $\varphi^{1/(2m)}$ is concave. For polynomial source sets (4) with $\varphi(\lambda) = \lambda^p$, the function $\varphi^{1/(2m)}$ is concave for $p \leq 2m$.

Example 4. (*Spectral methods*). Consider two methods (3) with

$$g_\alpha(\lambda) = \begin{cases} 1/\lambda & \text{for } \lambda \geq \alpha \\ 1/\alpha & \text{for } \lambda < \alpha \end{cases} \quad \text{and} \quad g_\alpha(\lambda) = \begin{cases} 1/\lambda & \text{for } \lambda \geq \alpha \\ 0 & \text{for } \lambda < \alpha . \end{cases}$$

For both methods, Assumption A2 (i) is satisfied with $\gamma_1 = 1$ and Assumption A2 (ii) is satisfied with $\gamma_2 = 1$ provided φ is increasing. For polynomial source sets (4) with $\varphi(\lambda) = \lambda^p$, the function φ is increasing for arbitrary $p < \infty$.

4. *A POSTERIORI* RULES FOR TIKHONOV REGULARIZATION

The *a priori* parameter choice (6) requires the knowledge of the function φ, which is generally unknown. One prominent *a posteriori* rule for choosing α, which does not require to know φ, is Morozov's discrepancy principle (see [6,13,14]). In this principle α is chosen as the solution of the nonlinear equation $\|Ax_\alpha^\delta - y^\delta\| = C\delta$ with $C \geq 1$, and order optimality on the source set (4) with $\varphi(\lambda) = \lambda^p$ has been established for the range $p \in (0, 1]$ (see [6,13,14]). For order optimality results of the discrepancy principle on general source sets (4) with concave source functions φ see [12].

In this section we restrict our studies to the method of Tikhonov regularization of order m discussed in Example 3 and consider the Raus–Gfrerer rule (see [15,16]) and the monotone error rule (see [3]). As Morozov's discrepancy principle, these rules do not require to know the function φ which characterizes the source set $M_{\varphi,E}$ given by (4).

Rule R1 (*Raus–Gfrerer rule*). Let $R_\alpha = \alpha(AA^* + \alpha I)^{-1}$. For a given constant $C \geq 1$, choose $\alpha = \alpha_{\mathrm{RG}}$ from the equation

$$d_{\mathrm{RG}}(\alpha) := \|R_\alpha^{1/2}(Ax_\alpha^\delta - y^\delta)\| = C\delta\,. \tag{12}$$

Rule R2 (*Monotone-error rule*). Let $R_\alpha = \alpha(AA^* + \alpha I)^{-1}$. For a given constant $C \geq 1$, choose $\alpha = \alpha_{\mathrm{ME}}$ from the equation

$$d_{\mathrm{ME}}(\alpha) := \frac{\|R_\alpha^{1/2}(Ax_\alpha^\delta - y^\delta)\|^2}{\|R_\alpha(Ax_\alpha^\delta - y^\delta)\|} = C\delta\,. \tag{13}$$

The nonlinear scalar equations (12) and (13) possess unique solutions provided $\|Py^\delta\| < C\delta < \|y^\delta\|$, where P is the orthoprojection of Y onto $N(A^*) = \overline{R(A)}^{\perp}$ (see, e.g., [3]). Note that Rule R2 always provides more accurate regularized approximations than Rule R1 and that the choice $C = 1$ in Rules R1 and R2 is the best possible (see [3]).

In our subsequent considerations we prove that for concave functions $\varphi^{1/(2m)}$ the method of Tikhonov regularization of order m combined with Rule R1 or Rule R2 is *order optimal* on the source set $M_{\varphi,E}$ given by (4). We start our studies with providing some monotonicity property for Rules R1 and R2 which may be found in [3].

Proposition 5. *Let $x_\alpha^\delta := x_{\alpha,m}^\delta$ be defined by* (11) *and let α_{RG} and α_{ME} be chosen by Rule R*1 *and Rule R*2*, respectively. Then,*

$$\alpha_{\mathrm{RG}} \geq \alpha_{\mathrm{ME}} \quad \textit{and} \quad \frac{\mathrm{d}}{\mathrm{d}\alpha}\|x_\alpha^\delta - x^\dagger\|^2 \geq 0 \quad \textit{for} \quad \alpha \geq \alpha_{\mathrm{ME}}\,.$$

Now we estimate $\|x_\alpha - x^\dagger\|$, where x_α is the approximation (11) with exact data and α is chosen either by Rule R1 or by Rule R2.

Proposition 6. *Let $x^\dagger \in M_{\varphi,E}$ with φ satisfying Assumption A*1*. Let $x_\alpha^\delta := x_{\alpha,m}^\delta$ be the regularized approximation* (11) *and x_α the corresponding regularized approximation with y^δ replaced by y. Let α be chosen either by Rule R*1 *or by Rule R*2*. If $\varphi^{1/(2m)}$ is concave, then*

$$\|x_\alpha - x^\dagger\| \le (C+1)E\sqrt{\rho^{-1}(\delta^2/E^2)}\,. \tag{14}$$

Proof. Due to Proposition 5 it is sufficient to prove (14) for $\alpha = \alpha_{\mathrm{RG}}$. Let us use the notations $R_\alpha = \alpha(AA^* + \alpha I)^{-1}$ and $r_\alpha(\lambda) = \frac{\alpha}{\lambda+\alpha}$. Since $\varphi^{1/(2m)}$ is concave with $\lim_{\lambda\to 0}\varphi(\lambda) = 0$, we have $t\varphi^{1/(2m)}(\lambda) \le \varphi^{1/(2m)}(t\lambda)$ for $t \in [0,1]$, or equivalently, $\varphi^{-1}(t^{2m}\varphi(\lambda)) \le t\lambda$. We multiply by $t^{2m}\varphi(\lambda)$ and obtain $\rho(t^{2m}\varphi(\lambda)) \le t^{2m+1}\lambda\varphi(\lambda)$. Choosing $t = r_\alpha(\lambda)$ yields

$$\rho\left(r_\alpha^{2m}(\lambda)\varphi(\lambda)\right) \le \lambda r_\alpha^{2m+1}(\lambda)\varphi(\lambda)\,. \tag{15}$$

Let α be chosen by Rule R1. Since $d_{\mathrm{RG}}(\alpha) = \|R_\alpha^{m+1/2}y^\delta\|$, we obtain

$$\|R_\alpha^{m+1/2}y\| \le \|R_\alpha^{m+1/2}y^\delta\| + \|R_\alpha^{m+1/2}(y-y^\delta)\| \le (C+1)\delta\,. \tag{16}$$

We use Assumption A1, (15), (16) and obtain by exploiting Jensen's inequality

$$\rho\left(\frac{\|x_\alpha - x^\dagger\|^2}{\|v\|^2}\right) \le \frac{\int_0^a \rho(r_\alpha^{2m}(\lambda)\varphi(\lambda))\,\mathrm{d}\|E_\lambda v\|^2}{\int_0^a \mathrm{d}\|E_\lambda v\|^2} \le \frac{(C+1)^2\delta^2}{\|v\|^2}\,. \tag{17}$$

From the monotonicity of φ^{-1}, the definition of ρ and (17) we obtain

$$\varphi^{-1}\left(\frac{\|x_\alpha - x^\dagger\|^2}{(C+1)^2E^2}\right) \le \varphi^{-1}\left(\frac{\|x_\alpha - x^\dagger\|^2}{\|v\|^2}\right) \le \frac{(C+1)^2\delta^2}{\|x_\alpha - x^\dagger\|^2}\,.$$

Due to $\rho(\lambda) = \lambda\varphi^{-1}(\lambda)$, this estimate provides (14). □

Exploiting Proposition 6, we obtain order optimal error bounds provided α is chosen *a posteriori* either by Rule R1 or by Rule R2.

Theorem 7. *Let $x^\dagger \in M_{\varphi,E}$ with φ satisfying Assumption A*1*. Let $x_\alpha^\delta := x_{\alpha,m}^\delta$ be the regularized approximation* (11) *and let α be chosen either by Rule R*1 *or by Rule R*2*. If $\varphi^{1/(2m)}$ is concave, then*

$$\|x_\alpha^\delta - x^\dagger\| \le c_0E\sqrt{\rho^{-1}(\delta^2/E^2)} \tag{18}$$

with $c_0 = C+1+\gamma_1$, $\gamma_1 = \frac{1}{2}$ for $m = 1$, and $\gamma_1 = \sqrt{m}$ for $m \ge 2$.

Proof. Due to Proposition 5, it is sufficient to prove (18) for $\alpha = \alpha_{\mathrm{RG}}$. We use the decomposition $x_\alpha^\delta - x^\dagger = (x_\alpha - x^\dagger) + (x_\alpha^\delta - x_\alpha)$ and obtain from (14) and the first part of (9) that

$$\|x_{\alpha_{\mathrm{RG}}}^\delta - x^\dagger\| \le (C+1)E\sqrt{\rho^{-1}(\delta^2/E^2)} + \frac{\gamma_1 \delta}{\sqrt{\alpha_{\mathrm{RG}}}}\,. \tag{19}$$

We distinguish a *first case* $\alpha_{\mathrm{RG}} \ge \varphi^{-1}\left(\rho^{-1}(\delta^2/E^2)\right) := \alpha_0$ and a *second case* $\alpha_{\mathrm{RG}} \le \alpha_0$. In the *first case* we obtain from (19) and (8) that

$$\|x_{\alpha_{\mathrm{RG}}}^\delta - x^\dagger\| \le (C+1+\gamma_1)E\sqrt{\rho^{-1}(\delta^2/E^2)}\,. \tag{20}$$

In the *second case* we use the estimate $\|x_{\alpha_{\mathrm{RG}}}^\delta - x^\dagger\| \le \|x_{\alpha_0}^\delta - x^\dagger\|$ which follows from Proposition 5, exploit Theorem 2 and obtain

$$\|x_{\alpha_{\mathrm{RG}}}^\delta - x^\dagger\| \le \|x_{\alpha_0}^\delta - x^\dagger\| \le (1+\gamma_1)E\sqrt{\rho^{-1}(\delta^2/E^2)}\,. \tag{21}$$

Now (18) follows from (20) and (21). □

ACKNOWLEDGEMENTS

Thanks are due to the organizers of the 2nd International Conference on Approximation Methods and Orthogonal Expansions, Kääriku, Estonia, 12–14 September 2003, and especially for the opportunity to present this paper.

REFERENCES

1. Tikhonov, A. N. and Arsenin, V. Y. *Solution of Ill-Posed Problems*. New York, Wiley, 1977.
2. Engl, H. W., Hanke, M. and Neubauer, A. *Regularization of Inverse Problems*. Dordrecht, Kluwer, 1993.
3. Tautenhahn, U. and Hämarik, U. The use of monotonicity for choosing the regularization parameter in ill-posed problems. *Inverse Probl.*, 1999, **15**, 1487–1505.
4. Vainikko, G. On the optimality of methods for ill-posed problems. *J. Anal. Anw.*, 1987, **6**, 351–362.
5. Vainikko, G. On the optimality of regularization methods. In *Inverse and Ill-Posed Problems* (Engl, H. W. et al., eds.). Academic Press, New York, 1987, 77–95.
6. Vainikko, G. M. and Veretennikov, A. Y. *Iteration Procedures in Ill-Posed Problems*. Nauka, Moscow, 1986 (in Russian).
7. Hohage, T. Regularization of exponentially ill-posed problems. *Numer. Funct. Anal. and Optimiz.*, 2000, **21**, 439–464.
8. Mair, B. A. Tikhonov regularization for finitely and infinitely smoothing operators. *SIAM J. Math. Anal.*, 1994, **25**, 135–147.
9. Tautenhahn, U. Optimality for ill-posed problems under general source conditions. *Numer. Funct. Anal. and Optimiz.*, 1998, **19**, 377–398.

10. Ivanov, V. and Korolyuk, T. Error estimates for solutions of ill-posed problems. *USSR Comput. Math. Math. Phys.*, 1969, **9**, 35–49.
11. Mathé, P. and Pereverzev, S. V. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Probl.*, 2003, **19**, 789–803.
12. Nair, M. T., Schock, E. and Tautenhahn, U. Morozov's discrepancy principle under general source conditions. *J. Anal. Anw.*, 2003, **22**, 199–214.
13. Vainikko, G. M. The discrepancy principle for a class of regularization methods. *USSR Comput. Math. Math. Phys.*, 1982, **22**, 1–19.
14. Vainikko, G. M. The critical level of discrepancy in regularization methods. *USSR Comput. Math. Math. Phys.*, 1983, **23**, 1–9.
15. Gfrerer, H. An a posteriori parameter choice for ordinary and iterated Tikhonov regularization of ill-posed problems leading to optimal convergence rates. *Math. Comput.*, 1987, **49**, 507–522.
16. Raus, T. On the discrepancy principle for the solution of ill-posed problems with non-selfadjoint operators. *Acta Comment. Univ. Tartuensis*, 1985, **715**, 12–20 (in Russian).

Optimaalse järguga regulariseerimisest üldise allikatingimuse korral

Ulrich Tautenhahn

On vaadeldud regulariseerimismeetodite klassi mittekorrektse ülesande lahendamiseks lähteandmete teadaoleva veataseme korral. Eeldatakse, et täpsel lahendil on üldistatud allikataoline esitus. Näidatakse, et regulariseerimisparameetri valik Rausi–Gfrereri reegli või monotoonse vea reegli kohaselt tagab lähislahendile optimaalse järguga veahinnangu.