Morozov's Discrepancy Principle Under General Source Conditions

M. Thamban Nair[1], Eberhard Schock[2] and Ulrich Tautenhahn[3]

Abstract. In this paper we study linear ill-posed problems $Ax = y$ in a Hilbert space setting where instead of exact data y noisy data y^δ are given satisfying $\|y - y^\delta\| \le \delta$ with known noise level δ. Regularized approximations are obtained by a general regularization scheme where the regularization parameter is chosen from Morozov's discrepancy principle. Assuming the unknown solution belongs to some general source set M we prove that the regularized approximation provides order optimal error bounds on the set M. Our results cover the special case of finitely smoothing operators A and extends recent results for infinitely smoothing operators.

1 Introduction

Ill-posed problems arise in several contexts and have important applications in science and engineering (see, e.g., [1, 3, 4, 6, 10, 13]). In this paper we are interested in the minimum-norm-solution $x^\dagger \in X$ of the ill-posed problem

$$Ax = y, \tag{1.1}$$

where $A : X \to Y$ is a linear bounded operator between infinite dimensional Hilbert spaces X and Y with non-closed range $R(A)$ of A, and $y \in R(A)$. Throughout this paper we assume that $y^\delta \in Y$ are the available noisy data with $\|y - y^\delta\| \le \delta$ and known noise level $\delta > 0$. For the stable approximate solution of problem (1.1) some regularization technique has to be applied, which provides regularized approximations $x_\alpha^\delta = R_\alpha^\delta y^\delta$ with property $x_\alpha^\delta \to x^\dagger \quad \text{as} \quad \delta \to 0$ where the regularization parameter $\alpha := \alpha(\delta, y^\delta)$ has to be chosen properly. Hence, regularized solutions x_α^δ depend continuously on the data. However, the convergence of x_α^δ to $x^\dagger$ can be arbitrarily slow without assuming additional quantitative a priori restrictions on the unknown solution $x^\dagger$, which

[1]Department of Mathematics, University of Kaiserslautern, P.O.Box 3049, 67653 Kaiserslautern, Germany; Home Institution: Department of Mathematics, Indian Institute of Technology Madras, Chennai 600 036, India; E.Mail: mtnair@iitm.ac.in

[2]Department of Mathematics, University of Kaiserslautern, P.O.Box 3049, 67653 Kaiserslautern, Germany; E-Mail: schock@mathematik.uni-kl.de

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

is typical for ill-posed problems, see [12]. In order to guarantee certain convergence rates the set of solutions has to be restricted to certain source sets. Typically, for operator equations (1.1) with finitely smoothing operators A, source conditions of the type $x^\dagger \in M_{p,E}$ with

$$M_{p,E} = \Big\{ x \in X \mid x = (A^*A)^{p/2} v\,,\ \|v\| \le E \Big\}\,, \quad p > 0, \tag{1.2}$$

are exploited, see [1, 2, 6, 15]. For infinitely smoothing operators A source conditions of the type (1.2) are generally too restrictive. In this case it is natural to assume that $x^\dagger \in M_{p,E}^{\log}$ with

$$M_{p,E}^{\log} = \Big\{ x \in X \mid x = \ln^{-p/2}(A^*A)^{-1} v\,,\ \|v\| \le E, \Big\}\,, \quad p > 0, \tag{1.3}$$

see [5, 7, 11, 13]. 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

$$M_{\varphi,E} = \Big\{ x \in X \mid x = [\varphi(A^*A)]^{1/2} v\,,\ \|v\| \le E \Big\}\ . \tag{1.4}$$

In (1.4) the operator function $\varphi(A^*A)$ is well defined via spectral representation

$$\varphi(A^*A) = \int_0^a \varphi(\lambda)\, \mathrm{d}E_\lambda$$

where $A^*A = \int_0^a \lambda\, \mathrm{d}E_\lambda$ is the spectral representation of A^*A, $\{E_\lambda : 0 \le \lambda \le a\}$ is the spectral family of A^*A with $a > 0$ satisfying $\sigma(A^*A) \subseteq [0, a]$ with $\|A\|^2 \le a$, where $\sigma(A^*A)$ denotes the spectrum of the operator A^*A.

Throughout this paper we assume that $x^\dagger \in M_{\varphi,E}$ so that

$$x^\dagger = [\varphi(A^*A)]^{1/2} v \quad \text{with} \quad \|v\| \le E,$$

where the function φ satisfies

Assumption 1.1 *The function $\varphi : (0, a] \to (0, \infty)$, where $a > 0$ with $\|A^*A\| \le a$, is continuous and satisfies the following:*

(i) $\lim_{\lambda \to 0} \varphi(\lambda) = 0$,

(ii) *φ is strictly monotonically increasing on $(0, a]$* ,

(iii) *$\rho : (0, \varphi(a)] \to (0, a\varphi(a)]$ defined by $\rho(\lambda) := \lambda \varphi^{-1}(\lambda), \lambda \in (0, \varphi(a)]$, is convex.*

2 Optimality and Order Optimality

Any operator $R : Y \to X$ can be considered as a special method for solving (1.1). The approximate solution to (1.1) is then given by Ry^δ. Let us consider the *worst case error* for identifying the minimum-norm-solution $x^\dagger$ of problem (1.1) from $y^\delta \in Y$ under the assumptions $\|y - y^\delta\| \le \delta$ and $x^\dagger \in M_{\varphi,E}$ defined by

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

This worst case error characterizes the maximal error of the method R if the minimum-norm-solution $x^\dagger$ of problem (1.1) varies in the set $M_{\varphi,E}$. An optimal method R_0 would be the one which satisfies

$$\Delta(\delta, R_0) = \inf_R \Delta(\delta, R).$$

It is easy to see that

$$\inf_R \Delta(\delta, R) \geq \omega(\delta, M_{\varphi,E}),$$

where

$$\omega(\delta, M_{\varphi,E}) = \sup\{\|x\| \mid x \in M_{\varphi,E}\,,\ \|Ax\| \leq \delta\}\,. \tag{2.1}$$

In view of the above relation, it is important to have some estimates for the quantity $\omega(\delta, M_{\varphi,E})$. In this regard we have the following result from Tautenhahn [13].

Theorem 2.1 *Let $M_{\varphi,E}$ be given by (1.4) and let Assumption 1.1 be satisfied. Then,*

$$\omega(\delta, M_{\varphi,E}) \leq E\,\sqrt{\rho^{-1}(\delta^2/E^2)}\,. \tag{2.2}$$

If $\delta^2/E^2 \in \sigma(A^*A\varphi(A^*A))$, then there holds equality in (2.2).

Due to Theorem 2.1, the following definition makes sense.

Definition 2.2 *Consider the set $M_{\varphi,E}$ of (1.4) and let Assumption 1.1 be satisfied. Then, any regularization method $R = R_\alpha^\delta$, or any regularized approximation $x_\alpha^\delta = R_\alpha^\delta y^\delta$ 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

$$\|x_\alpha^\delta - x^\dagger\| \leq c\,E\,\sqrt{\rho^{-1}(\delta^2/E^2)} \quad \textit{with} \quad c \geq 1\,.$$

In Tautenhahn [13] it has been proved that there exist special regularization methods that are optimal on the set $M_{\varphi,E}$. One of these methods is the the method of generalized Tikhonov regularization. In this method a regularized approximation x_α^δ is determined by solving the minimization problem

$$\min_{x\in X} J_\alpha(x)\,;\ \ J_\alpha(x) = \|Ax - y^\delta\|^2 + \alpha\|[\varphi(A^*A)]^{-1/2}x\|^2, \quad x \in X, \tag{2.3}$$

or equivalently, by solving the Euler equation $(A^*A + \alpha[\varphi(A^*A)]^{-1})x_\alpha^\delta = A^*y^\delta$ of Tikhonov's functional $J_\alpha(x)$. This method is *optimal* on the set $M_{\varphi,E}$ given by (1.4) provided the regularization parameter α is chosen properly. In fact, we have the following result from [13].

Theorem 2.3 *Let $M_{\varphi,E}$ be given by (1.4), let Assumption 1.1 be satisfied, $\varphi(\lambda)$: $(0,a] \to \mathbb{R}$ be twice differentiable, $\rho(\lambda)$ be strictly convex on $(0,\varphi(a)]$ and $\delta^2/E^2 \leq a\varphi(a)$. Then the Tikhonov regularized solution x_α^δ defined by (2.3) is optimal on $M_{\varphi,E}$ provided the regularization parameter α is chosen by*

$$\alpha = \frac{\lambda}{\varphi^{-1}(\lambda)\varphi'(\varphi^{-1}(\lambda))}\left(\frac{\delta}{E}\right)^2 \quad \textit{with} \quad \lambda = \rho^{-1}(\delta^2/E^2)\,. \tag{2.4}$$

It is to be mentioned that, for the method (2.3), in a general setting of an unbounded operator L in place of $\varphi(A^*A)]^{-1/2}$, *order optimal* results are obtained by Mair [7] under an apriori choice of the parameter, and by Nair [9] by using the Morozov discrepancy principle.

3 The General Regularization Scheme

The construction of regularized approximations that are *optimal* on the source set $M_{\varphi,E}$ given in (1.4), such as the ones considered in [7], [9] and [13], requires the knowledge of the function φ. In practice, however, the smoothness properties of the unknown solution $x^\dagger$ of problem (1.1) are generally unknown. Hence, there arises the question if there are regularization methods which do not require the knowledge of the function φ and which are *order optimal* on the set $M_{\varphi,E}$. We will prove in this section that if the function φ is concave, then the classical regularization methods such as Tikhonov regularization, iterated Tikhonov regularization, asymptotical regularization, regularized singular value decomposition and others, provide, combined with Morozov's discrepancy principle regularized approximations x_α^δ which are *order optimal* on the set $M_{\varphi,E}$.

It is well known that in the classical regularization methods the regularized approximations x_α^δ can be represented in the general form

$$x_\alpha^\delta = g_\alpha(A^*A)A^*y^\delta \quad \text{with} \quad 0 \leq 1 - \lambda g_\alpha(\lambda) \leq 1, \quad \lambda \in \sigma(A^*A). \tag{3.1}$$

This representation appears to be useful in the theoretical study of regularization methods. In Morozov's discrepancy principle (cf. [8]), the regularization parameter α is chosen as the solution of the nonlinear scalar equation

$$\|Ax_\alpha^\delta - y^\delta\| = C\delta \tag{3.2}$$

with some constant $C \geq 1$. Let x_α the regularized solution (3.1) with y^δ replaced by the exact data y, that is,

$$x_\alpha = g_\alpha(A^*A)A^*y\,. \tag{3.3}$$

In our next theorem we provide order optimal error bounds for $\|x_\alpha^\delta - x^\dagger\|$. We shall make use of the relation

$$\|x_\alpha^\delta - x^\dagger\|^2 \leq (x^\dagger - x_\alpha, x^\dagger) = (R_\alpha x^\dagger, x^\dagger) \tag{3.4}$$

established in [15].

Theorem 3.1 *Let $M_{\varphi,E}$ be given by (1.4) and let Assumption 1.1 be satisfied. Let x_α^δ be the regularized solution obtained by the general regularization scheme (3.1) and let α be chosen by Morozov's discrepancy principle (3.2). If the function φ is concave, then x_α^δ is order optimal on the set $M_{\varphi,E}$. In fact,*

$$\|x_\alpha^\delta - x^\dagger\| \le (C+1)E\sqrt{\rho^{-1}(\delta^2/E^2)}\,. \tag{3.5}$$

Proof. Let us use the notations

$$R_\alpha = I - A^*Ag_\alpha(A^*A)\,, \quad \widehat{R}_\alpha = I - AA^*g_\alpha(AA^*) \quad \text{and} \quad r_\alpha(\lambda) = 1 - \lambda g_\alpha(\lambda)\,.$$

Let $\alpha = \alpha_D$ be the regularization parameter chosen by (3.2). By the relation (3.4), we have

$$\|x_\alpha^\delta - x^\dagger\| \le \|R_\alpha^{1/2}x^\dagger\|\,. \tag{3.6}$$

Since φ is a concave function with $\lim_{\lambda\to 0}\varphi(\lambda) = 0$ we have $t\varphi(\lambda) \le \varphi(t\lambda)$ for $t \in [0,1]$, or equivalently, $\varphi^{-1}(t\varphi(\lambda)) \le \lambda t$, which, due to $\rho(\lambda) := \lambda\varphi^{-1}(\lambda)$ provides the relation $\rho(t\varphi(\lambda)) \le t^2\lambda\varphi(\lambda)$. We apply this estimate with $t = r_\alpha(\lambda) := 1 - \lambda g_\alpha(\lambda)$ and obtain

$$\rho(r_\alpha(\lambda)\varphi(\lambda)) \le \lambda r_\alpha^2(\lambda)\varphi(\lambda)\,. \tag{3.7}$$

Since α is the solution of (3.2), we have

$$\begin{aligned}
\|Ax_\alpha - Ax^\dagger\| &= \|AR_\alpha x^\dagger\| \\
&= \|\widehat{R}_\alpha Ax^\dagger\| \\
&\le \|\widehat{R}_\alpha y^\delta\| + \|\widehat{R}_\alpha(Ax^\dagger - y^\delta)\| \\
&= \|Ax_\alpha^\delta - y^\delta\| + \|\widehat{R}_\alpha(Ax^\dagger - y^\delta)\| \\
&\le (C+1)\delta\,.
\end{aligned} \tag{3.8}$$

Now, we exploit (1.4), the convexity of ρ, (3.7) and (3.8) and the Jensen's inequality. Thus,

$$\begin{aligned}
\rho\left(\frac{\|R_\alpha^{1/2}x^\dagger\|^2}{\|v\|^2}\right) &= \rho\left(\frac{\int_0^a r_\alpha(\lambda)\varphi(\lambda)\,\mathrm{d}\|E_\lambda v\|^2}{\int_0^a \mathrm{d}\|E_\lambda v\|^2}\right) \\
&\le \frac{\int_0^a \rho(r_\alpha(\lambda)\varphi(\lambda))\,\mathrm{d}\|E_\lambda v\|^2}{\int_0^a \mathrm{d}\|E_\lambda v\|^2} \\
&\le \frac{\int_0^a \lambda r_\alpha^2(\lambda)\varphi(\lambda)\,\mathrm{d}\|E_\lambda v\|^2}{\int_0^a \mathrm{d}\|E_\lambda v\|^2} \\
&= \frac{\|(A^*A)^{1/2}R_\alpha[\varphi(A^*A)]^{1/2}v\|^2}{\|v\|^2} \\
&= \frac{\|Ax_\alpha - Ax^\dagger\|^2}{\|v\|^2} \\
&\le \frac{(C+1)^2\delta^2}{\|v\|^2}\,.
\end{aligned} \tag{3.9}$$

Using the monotonicity of φ^{-1} and the realtions $\varphi^{-1}(\lambda) = \frac{1}{\lambda}\rho(\lambda)$ and (3.9), we obtain

$$\begin{aligned}
\varphi^{-1}\left(\frac{\|R_\alpha^{1/2}x^\dagger\|^2}{(C+1)^2E^2}\right) &\leq \varphi^{-1}\left(\frac{\|R_\alpha^{1/2}x^\dagger\|^2}{\|v\|^2}\right) \\
&= \frac{\|v\|^2}{\|R_\alpha^{1/2}x^\dagger\|^2}\rho\left(\frac{\|R_\alpha^{1/2}x^\dagger\|^2}{\|v\|^2}\right) \\
&\leq \frac{(C+1)^2\delta^2}{\|R_\alpha^{1/2}x^\dagger\|^2},
\end{aligned}$$

or equivalently,

$$\rho\left(\frac{\|R_\alpha^{1/2}x^\dagger\|^2}{(C+1)^2E^2}\right) \leq \frac{\delta^2}{E^2}.$$

This estimate together with (3.6) leads to (3.5). □

Let us discuss two special cases. In our first special case we consider operator equations with finitely smoothing operators A and assume that $x^\dagger \in M_{p,E}$ with $M_{p,E}$ given in (1.2). This set has the form (1.4) with $\varphi(\lambda) = \lambda^p$. For this function Asumption 1.1 is satisfied. Since $\varphi^{-1}(\lambda) = \lambda^{1/p}$ we obtain $\rho(\lambda) := \lambda\varphi^{-1}(\lambda) = \lambda^{(p+1)/p}$, consequently, $\rho^{-1}(\lambda) = \lambda^{p/(p+1)}$. Since $\varphi(\lambda) = \lambda^p$ is concave for $p \leq 1$ the following result is a special case of Theorem 3.1.

Corollary 3.2 *Let x_α^δ be the regularized solution obtained by the general regularization scheme (3.1) and let α be chosen by Morozov's discrepancy principle (3.2). If $p \leq 1$, then x_α^δ is order optimal on the set $M_{p,E}$ given in (1.2), and*

$$\|x_\alpha^\delta - x^\dagger\| \leq (C+1)E^{\frac{1}{p+1}}\,\delta^{\frac{p}{p+1}}.$$

Although this classical result of Corollary 3.2 is not new (see [15]) we realize that it can be obtained as a special case of our more general Theorem 3.1.

In our second special case we consider operator equations with infinitely smoothing operators A and assume that $x^\dagger \in M_{p,E}^{\log}$ with $M_{p,E}^{\log}$ given in (1.3). This set has the form (1.4) with $\varphi(\lambda) = \left[\ln\frac{1}{\lambda}\right]^{-p}$. For this function Asumption 1.1 is satisfied. Since $\varphi^{-1}(\lambda) = e^{-1/\lambda^{1/p}}$ we obtain $\rho(\lambda) := \lambda\varphi^{-1}(\lambda) = \lambda e^{-1/\lambda^{1/p}}$, consequently,

$$\rho^{-1}(\lambda) = \left[\ln\frac{1}{\lambda}\right]^{-p}(1+o(1)) \quad \text{for} \quad \lambda \to 0,$$

see, e.g., [7]. Since $\varphi(\lambda)$ is concave for $\lambda \leq 1/(p+1)^p$, from Theorem 3.1, we obtain the following result.

Corollary 3.3 *Let x_α^δ the regularized solution obtained by the general regularization scheme (3.1) and let α be chosen by Morozov's discrepancy principle (3.2). If $\|A\| \leq 1/(p+1)^{p/2}$, then x_α^δ is order optimal on the set $M_{p,E}^{\log}$ given in (1.3), and*

$$\|x_\alpha^\delta - x^\dagger\| \leq (C+1)E\left[\ln\frac{E^2}{\delta^2}\right]^{-p/2}(1+o(1)) \quad \textit{for} \quad \delta \to 0.$$

Remarks. For Tikhonov regularization, a result analogous to Corollary 3.3 has been proved by Pereverzev and Schock [11] under a modified form of Morozov's discrepancy principle. For the general regularization of the form (3.1), Hohage [5] proved a result of the form Corollary 3.3 under a stronger assumption on g_α, namely, $\lambda^\mu[1-\lambda g_\alpha(\lambda)] \le c_\mu \alpha^\mu$ for $0 \le \mu \le \mu_0$ for some $\mu_0 > 0$. We deduced Corollary 3.3 from a more general result, Theorem 3.1, whose proof also seems to be much simpler than the procedures adopted in [11] and [5].

While proving Theorem 3.1 we made use of the relation (3.4) proved in [15]. For the special case of Tikhonove regularization, we shall provide, in Section 5, an alternate proof which does not require (3.4).

4 The General Regularization Scheme Revisited

Theorem 3.1 requires the assumption that φ is concave. This assumption is too strong for certain special regularization methods and certain special source sets. Therefore we reconsider the general regularization scheme under some other assumption which appears to be weaker in certain circumstances. However, let us start our studies with some general order optimality result for the regularization error $\|x_\alpha - x^\dagger\|$.

Theorem 4.1 *Let $M_{\varphi,E}$ be given by (1.4). Let x_α be the regularized solution (3.3) and let α be chosen by Morozov's discrepancy principle (3.2). If $x^\dagger \in M_{\varphi,E}$, then*

$$\|x_\alpha - x^\dagger\| \le (C+1)\,\omega(\delta, M_{\varphi,E}) \tag{4.1}$$

with $\omega(\delta, M_{\varphi,E})$ given by (2.1). If in addition Assumption (1.1) is satisfied, then

$$\|x_\alpha - x^\dagger\| \le (C+1)E\,\sqrt{\rho^{-1}(\delta^2/E^2)}\,. \tag{4.2}$$

Proof. We define the element $z_\alpha = \frac{1}{C+1}(x_\alpha - x^\dagger)$ with α chosen by (3.2). Due to

$$x_\alpha - x^\dagger = [g_\alpha(A^*A)A^*A - I][\varphi(A^*A)]^{1/2}v$$

we have $z_\alpha = [\varphi(A^*A)]^{1/2}w$, where $w := [g_\alpha(A^*A)A^*A - I]v$ satisfies

$$\|w\| = \frac{1}{C+1}\|[I - g_\alpha(A^*A)A^*A]v\| \le \frac{1}{C+1}\|v\| \le E\,.$$

Hence, $z_\alpha \in M_{\varphi,E}$. In addition, due to (3.8), $\|Az_\alpha\| \le \delta$. Hence, due to (2.1) we obtain $\|z_\alpha\| \le \omega(\delta, M_{\varphi,E})$, or equivalently, (4.1). Now (4.2) follows from (4.1) and (2.2). □

Our next theorem provides an order optimality result under the condition that $\alpha_D \ge \alpha_0$ with α_D as the solution of (3.2) and α_0 given by

$$\alpha_0 = \varphi^{-1}(\rho^{-1}(\delta^2/E^2)), \tag{4.3}$$

under an additional assumption on g_α that there exists some constant c_0 such that

$$\sup_{\lambda\in[0,a]} \left|\sqrt{\lambda}\, g_\alpha(\lambda)\right| \le \frac{c_0}{\sqrt{\alpha}}\,. \tag{4.4}$$

Theorem 4.2 *Let $M_{\varphi,E}$ be given by (1.4) and let Assumption 1.1 be satisfied. Assume further that (4.4) is satisfied. Let x_α^δ be the regularized solution (3.1), α_D be the solution of (3.2), and α_0 be given by (4.3). If $\alpha_D \ge \alpha_0$, then $x_{\alpha_D}^\delta$ is order optimal on the set $M_{\varphi,E}$, and*

$$\|x_{\alpha_D}^\delta - x^\dagger\| \le (C + c_0 + 1)E\sqrt{\rho^{-1}(\delta^2/E^2)}\,. \tag{4.5}$$

Proof. Due to (4.4) we have

$$\|x_\alpha^\delta - x_\alpha\| = \|g_\alpha(A^*A)A^*(y^\delta - y)\| \le \frac{c_0\delta}{\sqrt{\alpha}}\,. \tag{4.6}$$

From (4.6) and $\alpha_D \ge \alpha_0$ we obtain

$$\|x_{\alpha_D}^\delta - x_{\alpha_D}\| \le \frac{c_0\delta}{\sqrt{\alpha_D}} \le \frac{c_0\delta}{\sqrt{\alpha_0}} = c_0 E\sqrt{\rho^{-1}(\delta^2/E^2)}\,.$$

Consequently, (4.5) follows from (4.2) and the triangle inequality. □

Now it remains to study the case $\alpha_D \le \alpha_0$. In this case, however, we need the additional condition that there exists some constant c_φ such that

$$\sup_{0<\lambda\le a} |[1 - \lambda g_\alpha(\lambda)]\varphi(\lambda)| \le c_\varphi \varphi(\alpha)\,. \tag{4.7}$$

Theorem 4.3 *Let $M_{\varphi,E}$ be given by (1.4) and let Assumption 1.1 be satisfied. Let x_α^δ be the regularized solution (3.1). Let α_D be the solution of (3.2). If assumptions (4.4) and (4.7) are satisfied, then $x_{\alpha_D}^\delta$ is order optimal on the set $M_{\varphi,E}$, and*

$$\|x_{\alpha_D}^\delta - x^\dagger\| \le c\,E\sqrt{\rho^{-1}(\delta^2/E^2)}, \tag{4.8}$$

where $c = \max\{C + c_0 + 1, \sqrt{c_\varphi}\}$.

Proof. Let us distinguish two cases. In the first case of $\alpha_D \ge \alpha_0$ with α_0 given by (4.3), the proof of this theorem follows from Theorem 4.2. It remains to consider the second case $\alpha_D \le \alpha_0$. Using the relation (3.4), and exploiting (1.4) and (4.7) we obtain

$$\begin{aligned}
\|x_{\alpha_D}^\delta - x^\dagger\|^2 &\le ([I - g_{\alpha_D}(A^*A)A^*A]\varphi(A^*A)v, v) \\
&\le E^2 \sup_{0<\lambda\le a} |[1 - \lambda g_{\alpha_D}(\lambda)]\varphi(\lambda)| \\
&\le E^2 c_\varphi \varphi(\alpha_D)\,.
\end{aligned}$$

Hence, since $\alpha_D \le \alpha_0$ and φ is monotonically increasing, we have

$$\|x_{\alpha_D}^\delta - x^\dagger\|^2 \le E^2 c_\varphi \varphi(\alpha_0) = E^2 c_\varphi \rho^{-1}(\delta^2/E^2)\,.$$

This completes the proof. □

The proof of the following corollary, a companion result to Corollary 3.2, is immediate from Theorem 4.3.

Corollary 4.4 *Let x_α^δ be the regularized solution obtained by the general regularization scheme (3.1) and let α be chosen by Morozov's discrepancy principle (3.2). Suppose, in addtion, that (4.4) is satisfied and there exists $p_0 > 0$ with the property that for each $p \in (0, p_0]$, there exists $c_p > 0$ such that*

$$\sup_{0<\lambda\leq a} \lambda^p[1-\lambda g_\alpha(\lambda)] \leq c_p \alpha^p \tag{4.9}$$

for $0 < p \leq p_0$. Then for each $p \in (0, p_0]$, x_α^δ is order optimal on the set $M_{p,E}$ given in (1.2), and

$$\|x_\alpha^\delta - x^\dagger\| \leq c\, E^{\frac{1}{p+1}}\, \delta^{\frac{p}{p+1}},$$

where $c = \max\{C + c_0 + 1, \sqrt{c_p}\}$.

Our next proposition shows that assumption (4.7) of Theorem 4.3 is indeed weaker than the concavity condition for the function φ in Theorem 3.1, for many regularization methods.

Proposition 4.5 *Let Assumption 1.1 be satisfied. Assume further that*

$$h(\lambda) := 1 - \alpha\lambda g_\alpha(\alpha\lambda), \quad 0 \leq \lambda < \infty. \tag{4.10}$$

is independend of α, and $0 \leq h(\lambda) \leq 1$ and $\lambda h(\lambda) \leq 1$. If φ is a concave function, then assumption (4.7) holds true with $c_\varphi = 1$.

Proof. We substitute $\lambda = \alpha s$ into (4.7) and obtain from (4.10)

$$\sup_{\lambda\in[0,a]} |[1-\lambda g_\alpha(\lambda)]\,\varphi(\lambda)| \leq \sup_{s\geq 0} |[1-\alpha s g_\alpha(\alpha s)]\,\varphi(\alpha s)| = \sup_{s\geq 0} |h(s)\varphi(\alpha s)|\,. \tag{4.11}$$

Since φ is a concave function with $\lim_{\lambda\to 0}\varphi(\lambda) = 0$ we have $t\varphi(s) \leq \varphi(ts)$ for $t \in [0,1]$. We apply this estimate with $t = h(s)$ and obtain

$$h(s)\varphi(\alpha s) \leq \varphi(\alpha s h(s)).$$

Since φ is monotonically increasing and $sh(s) \leq 1$ there follows

$$h(s)\varphi(\alpha s) \leq \varphi(\alpha s h(s)) \leq \varphi(\alpha).$$

This estimate and (4.11) provide the result of the proposition. □

Examples for regularization methods (3.1) with h in (4.10) satisfying $0 \leq h(\lambda) \leq 1$ and $\lambda h(\lambda) \leq 1$ include

(a) the method of ordinary Tikhonov regularization with $h(\lambda) = \frac{1}{\lambda+1}$,

(b) the method of asymptotical regularization with $h(\lambda) = e^{-\lambda}$ and

(c) the method of truncated singular value decomposition with $h(\lambda) = 0$ for $\lambda \geq 1$ and $h(\lambda) = 1 - \lambda$ for $\lambda \leq 1$.

5 Discretization in Tikhonov's Method

In the method of ordinary Tikhonov regularization the regularized approximation x_α^δ is determined by solving the minimization problem

$$\min_{x\in X} J_\alpha(x)\,; \quad J_\alpha(x) = \|Ax - y^\delta\|^2 + \alpha\|x\|^2, \quad x \in X,$$

or equivalently, x_α^δ is the unique solution of the Euler equation

$$(A^*A + \alpha I)x_\alpha^\delta = A^*y^\delta$$

of Tikhonov's functional $J_\alpha(x)$. Hence, x_α^δ has the form (3.1) with $g_\alpha(\lambda) = 1/(\lambda+\alpha)$.

The computation of regularized approximations x_α^δ requires the numerical realization of (3.1) and (3.2) in finite dimensional spaces. One may take a finite dimensional subspace X_n, of dimension n, and define a finite dimensional regularized approximation $x_{n,\alpha}^\delta$ by solving the finite dimensional minimization problem

$$\min_{x\in X_n} J_\alpha(x)\,; \quad J_\alpha(x) = \|Ax - y^\delta\|^2 + \alpha\|x\|^2\,. \tag{5.1}$$

If $\{\phi_1, ..., \phi_n\}$ is a basis of X_n, then it is clear that the uniquely determined solution of problem (5.1) has the form

$$x_{n,\alpha}^\delta = \sum_{j=1}^{n} c_j\phi_j \tag{5.2}$$

with $\mathbf{c} = (c_1, ..., c_n)^T \in \mathbb{R}^n$ as the unique solution of the linear system of equations

$$[\mathbf{A} + \alpha\mathbf{\Phi}]\mathbf{c} = \mathbf{d}$$

where $\mathbf{A}$ and $\mathbf{\Phi}$ are $n\times n$ matrices with their ij-th entries given by $(A\phi_j, A\phi_i)$ and (ϕ_j, ϕ_j) respectively, and $\mathbf{d} = (d_1, ..., d_n)^T \in \mathbb{R}^n$ with $d_i = (y^\delta, A\phi_i)$.

For choosing the regularization parameter we consider a discretized discrepancy principle, that is, we choose α as the solution of the scalar nonlinear equation

$$\|Ax_{n,\alpha}^\delta - y^\delta\| = (1+\varepsilon_n)\delta \tag{5.3}$$

with some constant $\varepsilon_n \geq 0$. In our next theorem we show that our finite dimensional regularized approximation (5.2) with α chosen by (5.3) is order optimal on the general source set (1.4) provided φ is concave and the dimension number n is chosen sufficiently large. In fact, we state and prove the theorem in the following more general setting :

Let (X_n) be a sequence of subspaces of X, and for each n, let $P_n : X \to X$ be the orthogonal projection onto X_n such that $\|x - P_nx\| \to 0$ as $n\to\infty$ for every $x \in X$. (This condition is satisfied if, for example, $X_n := \text{span}\{\phi_1, ..., \phi_n\}$, where $\{\phi_1, \phi_2, \ldots\}$ is an orthonormal basis of X.) Let $x_{n,\alpha}^\delta \in X_n$ be the unique solution of (5.1); equivalently, $x_{n,\alpha}^\delta$ satisfies the equation

$$(A_n^*A_n + \alpha I)x_{n,\alpha}^\delta = A_n^*y^\delta, \tag{5.4}$$

where $A_n := AP_n$.

Theorem 5.1 *Let $M_{\varphi,E}$ be given by (1.4) and let Assumption 1.1 be satisfied. Let $x_{n,\alpha}^\delta$ the unique solution of (5.4) and let α be chosen by the discretized discrepancy principle (5.3). If P_n satisfies*

$$\|A(I-P_n)x^\dagger\| \le \varepsilon_n \delta, \tag{5.5}$$

and if φ is concave, then $x_{n,\alpha}^\delta$ is order optimal on the set $M_{\varphi,E}$, and

$$\|x_{n,\alpha}^\delta - x^\dagger\| \le (2+\varepsilon_n)E\sqrt{\rho^{-1}(\delta^2/E^2)}\,. \tag{5.6}$$

Proof. Since $x_{n,\alpha}^\delta$ satisfies the equation $\alpha x_{n,\alpha}^\delta = P_n A^*[y^\delta - Ax_{n,\alpha}^\delta]$ we obtain due to $AP_n x_{n,\alpha}^\delta = Ax_{n,\alpha}^\delta$ that

$$\begin{aligned}
\|x_{n,\alpha}^\delta - x^\dagger\|^2 &= (x^\dagger - x_{n,\alpha}^\delta, x^\dagger) - (x^\dagger - x_{n,\alpha}^\delta, x_{n,\alpha}^\delta)\\
&= (x^\dagger - x_{n,\alpha}^\delta, x^\dagger) - \frac{1}{\alpha}\left(x^\dagger - x_{n,\alpha}^\delta, P_nA^*[y^\delta - Ax_{n,\alpha}^\delta]\right)\\
&= (x^\dagger - x_{n,\alpha}^\delta, x^\dagger) - \frac{1}{\alpha}\left(AP_nx^\dagger - Ax_{n,\alpha}^\delta, y^\delta - Ax_{n,\alpha}^\delta\right)\\
&= (x^\dagger - x_{n,\alpha}^\delta, x^\dagger) + \frac{1}{2\alpha}\Big\{\|AP_nx^\dagger - y^\delta\|^2 - \|Ax_{n,\alpha}^\delta - y^\delta\|^2\\
&\qquad\qquad - \|AP_nx^\dagger - Ax_{n,\alpha}^\delta\|^2\Big\}\,.
\end{aligned} \tag{5.7}$$

From (5.5) we have $\|AP_nx^\dagger - y^\delta\| \le \|Ax^\dagger - y^\delta\| + \|A(I-P_n)x^\dagger\| \le (1+\varepsilon_n)\delta$. Consequently, due to (5.7) and (5.3) we obtain

$$\|x_{n,\alpha}^\delta - x^\dagger\|^2 \le (x^\dagger - x_{n,\alpha}^\delta, x^\dagger)\,.$$

Exploiting the source condition $x^\dagger \in M_{\varphi,E}$ with $M_{\varphi,E}$ given in (1.4) yields

$$\|x_{n,\alpha}^\delta - x^\dagger\|^2 \le E\|[\varphi(A^*A)]^{1/2}(x_{n,\alpha}^\delta - x^\dagger)\|\,,$$

hence,

$$\frac{\|x_{n,\alpha}^\delta - x^\dagger\|^2}{(2+\varepsilon_n)^2E^2} \le \frac{\|[\varphi(A^*A)]^{1/2}(x_{n,\alpha}^\delta - x^\dagger)\|^2}{\|x_{n,\alpha}^\delta - x^\dagger\|^2}\,.$$

Since φ is monotone and concave we obtain that φ^{-1} is monotone and convex. We use the monotonicity of φ^{-1}, apply Jensen's inequality and obtain

$$\begin{aligned}
\varphi^{-1}\left(\frac{\|x_{n,\alpha}^\delta - x^\dagger\|^2}{(2+\varepsilon_n)^2E^2}\right) &\le \varphi^{-1}\left(\frac{\|[\varphi(A^*A)]^{1/2}(x_{n,\alpha}^\delta - x^\dagger)\|^2}{\|x_{n,\alpha}^\delta - x^\dagger\|^2}\right)\\
&= \varphi^{-1}\left(\frac{\int_0^a \varphi(\lambda)\,\mathrm{d}\|E_\lambda(x_{n,\alpha}^\delta - x^\dagger)\|^2}{\int_0^a \mathrm{d}\|E_\lambda(x_{n,\alpha}^\delta - x^\dagger)\|^2}\right)\\
&\le \frac{\int_0^a \lambda\,\mathrm{d}\|E_\lambda(x_{n,\alpha}^\delta - x^\dagger)\|^2}{\int_0^a \mathrm{d}\|E_\lambda(x_{n,\alpha}^\delta - x^\dagger)\|^2}\\
&= \frac{\|Ax_{n,\alpha}^\delta - Ax^\dagger\|^2}{\|x_{n,\alpha}^\delta - x^\dagger\|^2}\,.
\end{aligned}$$

Due to $\lambda\varphi^{-1}(\lambda) = \rho(\lambda)$ and $\|Ax_{n,\alpha}^\delta - Ax^\dagger\| \le \|Ax_{n,\alpha}^\delta - y^\delta\| + \|Ax^\dagger - y^\delta\| \le (2+\varepsilon_n)\delta$ we obtain

$$\rho\left(\frac{\|x_{n,\alpha}^\delta - x^\dagger\|^2}{(2+\varepsilon_n)^2E^2}\right) \le \frac{\delta^2}{E^2}\,,$$

which provides (5.6). □

Note that for $P_n = I$, $X_n = X$, $\varepsilon_n = 0$, and $1+\varepsilon_n = C$ the results of Theorem 5.1 and Theorem 3.1 with $g_\alpha(\lambda) = 1/(\lambda+\alpha)$ coincide.

6 Asymptotical Regularization

In the method of asymptotical regularization the regularized approximation x_α^δ is given by $x_\alpha^\delta = x(1/\alpha)$ where $x(t)$ is obtained by solving the initial value problem problem

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) + A^*Ax(t) = A^*y^\delta \quad \text{for} \quad 0 < t \le 1/\alpha\,, \quad x(0) = 0\,. \tag{6.1}$$

It is well known that x_α^δ has the representation

$$x_\alpha^\delta = g_\alpha(A^*A)A^*y^\delta \quad \text{with} \quad g_\alpha(\lambda) = (1-e^{-\lambda/\alpha})/\lambda\,. \tag{6.2}$$

Obviously, this method fits into the framework of the general regularization method (3.1). In this section we will give, compared with Theorem 4.3, some other proof for order optimal error bounds which, instead of assumption (4.7) requires some weaker assumption. We start our studies with providing some monotonicity property.

Theorem 6.1 *Let x_α^δ the regularized solution obtained by the method of asymptotical regularization (6.2) and let α_D the regularization parameter obtained by the Morozov's discrepancy principle (3.2). Then,*

$$\|x_{\alpha_D}^\delta - x^\dagger\| \le \|x_\alpha^\delta - x^\dagger\| \quad \textit{for} \quad \alpha_D \le \alpha\,. \tag{6.3}$$

Proof. From $g_\alpha(\lambda) = (1-e^{-\lambda/\alpha})/\lambda$ we have

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(\lambda) = -\frac{1}{\alpha^2}e^{-\lambda/\alpha} = \frac{1}{\alpha^2}(\lambda g_\alpha(\lambda) - 1)\,.$$

Consequently,

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(AA^*)y^\delta = \frac{1}{\alpha^2}[AA^*g_\alpha(AA^*) - I]y^\delta = \frac{1}{\alpha^2}[Ax_\alpha^\delta - y^\delta]\,.$$

This identity provides the estimate

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\alpha}\|x_\alpha^\delta - x^\dagger\|^2 &= \Big(x_\alpha^\delta - x^\dagger, A^*\frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(AA^*)y^\delta\Big) \\
&= \Big(Ax_\alpha^\delta - y^\delta + (y^\delta - y), \frac{\mathrm{d}}{\mathrm{d}\alpha}g_\alpha(AA^*)y^\delta\Big) \\
&= \frac{1}{\alpha^2}\Big(Ax_\alpha^\delta - y^\delta + (y^\delta - y), Ax_\alpha^\delta - y^\delta\Big) \\
&\ge \frac{\|Ax_\alpha^\delta - y^\delta\|}{\alpha^2}\Big\{\|Ax_\alpha^\delta - y^\delta\| - \delta\Big\}\,.
\end{aligned}$$

From this estimate and the monotonicity of the norm $\|Ax_\alpha^\delta - y^\delta\|$ as a function of α, it follows that for all $\alpha \geq \alpha_D$,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\alpha}\|x_\alpha^\delta - x^\dagger\|^2 \geq \frac{C\delta}{\alpha_D^2}(C-1)\delta \geq 0.$$

Thus, we obtain the required property (6.3). $\square$

In order to derive order optimal error bounds for $\|x_{\alpha_D}^\delta - x^\dagger\|$ we distinguish as in Section 4 two cases, a first case with $\alpha_D \geq \alpha_0$ and a second case with $\alpha_D \leq \alpha_0$ where α_0 is given by (4.3). In the first case we will exploit Theorem 4.2, and in the second case Theorem 6.1. The second case requires, instead of assumption (4.7), the condition that there exists some constant d_φ such that

$$\sup_{0<\lambda\leq a}\left|[1-\lambda g_\alpha(\lambda)]\sqrt{\varphi(\lambda)}\right| \leq d_\varphi\sqrt{\varphi(\alpha)}\,. \tag{6.4}$$

Note that, due to the condition in (3.1) on $g_\alpha(\lambda)$, the relation (4.7) implies (6.4).

Theorem 6.2 *Let $M_{\varphi,E}$ be given by (1.4) and let Assumption 1.1 be satisfied. Let x_α^δ the regularized solution (6.2). Let α_D the solution of (3.2). If assumption (4.4) and (6.4) are satisfied, then $x_{\alpha_D}^\delta$ is order optimal on the set $M_{\varphi,E}$, and*

$$\|x_{\alpha_D}^\delta - x^\dagger\| \leq c\,E\,\sqrt{\rho^{-1}(\delta^2/E^2)} \tag{6.5}$$

with $c = \max\{C + c_0 + 1,\ d_\varphi + 1\}$.

Proof. For the case of $\alpha_D \geq \alpha_0$ with α_0 given by (4.3) the proof of this theorem follows from Theorem 4.2. It remains to consider the second case $\alpha_D \leq \alpha_0$. From Theorem 6.1, the triangle inequality, (4.5) and the identity $\frac{\delta}{\sqrt{\alpha_0}} = E\sqrt{\rho^{-1}(\delta^2/E^2)}$ we obtain

$$\begin{aligned}
\|x_{\alpha_D}^\delta - x^\dagger\| &\leq \|x_{\alpha_0}^\delta - x^\dagger\| \\
&\leq \|x_{\alpha_0} - x^\dagger\| + \|x_{\alpha_0}^\delta - x_{\alpha_0}\| \\
&\leq \|x_{\alpha_0} - x^\dagger\| + E\sqrt{\rho^{-1}(\delta^2/E^2)}\,.
\end{aligned} \tag{6.6}$$

Exploiting (6.2), (1.4), (6.4) as well as (4.3), we obtain

$$\begin{aligned}
\|x_{\alpha_0} - x^\dagger\| &= \|[g_{\alpha_0}(A^*A)A^*A - I][\varphi(A^*A)]^{1/2}v\| \\
&\leq d_\varphi E\sqrt{\varphi(\alpha_0)} \\
&= d_\varphi E\,\sqrt{\rho^{-1}(\delta^2/E^2)}\,.
\end{aligned} \tag{6.7}$$

Now (6.5) follows from (6.6) and (6.7). $\square$

Acknowledgement: M. Thamban Nair gratefully acknowledges the support received from German Academic Exchange Service (DAAD) to visit the Department of Mathematics, University of Kaiserslautern, during April-July 2002, under the Guest Professor's Programme (DAAD).

References

[1] Engl, H. W., Hanke, M. and Neubauer, A (1993): *Regularization of Inverse Problems*, Dordrecht: Kluwer.

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

[3] Hanke, M. and Hansen, P. C. (1993): Regularization methods for large-scale problems, *Surv. Math. Ind.* **3**, 253–315.

[4] Hofmann, B. (1986): *Regularization for Applied Inverse and Ill-Posed Problems*, Leipzig: Teubner.

[5] Hohage, T. (2000): Regularization of exponentially ill-posed problems, *Numer. Funct. Anal. and Optimiz.* **21**, 439–464.

[6] Louis, A. K. (1989): *Inverse und schlecht gestellte Probleme*, Stuttgart: Teubner.

[7] Mair, B. A. (1994): Tikhonov regularization for finitely and infinitely smoothing operators, *SIAM J. Math. Anal.* **25**, 135–147.

[8] Morozov, V. A. (1966): On the solution of functional equationsby the method of regularization, *Soviet Math. Dokl.* **7**, 414–417.

[9] Nair, M. T. (1999): On Morozov's method for Tikhonov regularization as an optimal order yielding algorithm, *Journal for Analysis and its Applications* **18**, 37–46.

[10] Payne, L. E. (1975): *Improperly Posed Problems in Partial Differential Equations*, Philadelphia: SIAM.

[11] Pereverzev, S. and Schock, E. (2000): Morozov's discrepancy principle for Tikhonov regularization of severely ill-posed problems, *Numer. Funct. Anal. and Optimiz.* **21**, 901–920.

[12] Schock, E. (1985): Approximate solution of ill-posed equations: arbitrary slow convergence vs. superconvergence, In: G. Hämmerlin and K. H. Hofmann, eds., *Constructive Methods for the Practical Treatment of Integral Equations*, Basel: Birkhäuser, 234–243.

[13] Tautenhahn, U. (1998): Optimality for ill-posed problems under general source conditions, *Numer. Funct. Anal. and Optimiz.* **19**, 377–398.

[14] Tikhonov, A. N. and Arsenin, V. Y. (1977): *Solution of Ill-Posed Problems*, New York: Wiley.

[15] Vainikko, G. M. and Veretennikov, A. Y. (1986): *Iteration Procedures in Ill-Posed Problems*, Moscow: Nauka (In Russian).