INSTITUTE OF PHYSICS PUBLISHING INVERSE PROBLEMS

Inverse Problems **21** (2005) 805–820 doi:10.1088/0266-5611/21/3/002

Convergence rates for Tikhonov regularization based on range inclusions

Bernd Hofmann[1] **and Masahiro Yamamoto**[2]

[1] Faculty of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany
[2] Department of Mathematical Sciences, The University of Tokyo, 3-8-1 Komaba Meguro, Tokyo 153, Japan

E-mail: hofmannb@mathematik.tu-chemnitz.de and myama@ms.u-tokyo.ac.jp

Received 21 December 2004, in final form 9 February 2005
Published 4 March 2005
Online at stacks.iop.org/IP/21/805

Abstract
This paper provides some new *a priori* choice strategy for regularization parameters in order to obtain convergence rates in Tikhonov regularization for solving ill-posed problems $Af_0 = g_0$, $f_0 \in X$, $g_0 \in Y$, with a linear operator A mapping in Hilbert spaces X and Y. Our choice requires only that the range of the adjoint operator A^* includes a member of some variable Hilbert scale and is, in principle, applicable in the case of general f_0 without source conditions imposed otherwise in the existing papers. For testing our strategies, we apply them to the determination of a wave source, to the Abel integral equation, to a backward heat equation and to the determination of initial temperature by boundary observation.

1. Introduction

Let X and Y be infinite-dimensional separable Hilbert spaces over $\mathbb{R}$. We consider a bounded injective linear operator A from X to Y and we will discuss the operator equation

$$Af = g, \qquad f \in X, \quad g \in Y. \tag{1.1}$$

We are mainly concerned with the case of a non-closed range $\mathcal{R}(A) \neq \overline{\mathcal{R}(A)}$, and so $A^{-1} : \mathcal{R}(A) \subset Y \to X$ is not continuous with respect to the norms in X and Y, which describes a general linear ill-posed problem. Then equation (1.1) is unstable and the stable approximate solution of the uniquely determined solution $f_0 \in X$ of (1.1) for the exact right-hand side $g_0 \in \mathcal{R}(A)$ requires some regularization technique whenever noisy data $g_\delta \in Y$ with known noise level $\delta > 0$ satisfying the estimate

$$\|g_0 - g_\delta\|_Y \leqslant \delta, \tag{1.2}$$

are available instead of g_0. We discuss the classical Tikhonov regularization

$$\text{Minimize}\|Af - g_\delta\|_Y^2 + \alpha\|f\|_X^2 \qquad \text{over} \quad f \in X, \tag{1.3}$$

where $\alpha > 0$ denotes the regularization parameter and

$$f_{\alpha,\delta} = (A^*A + \alpha I)^{-1}A^*g_\delta \tag{1.4}$$

is the uniquely determined minimizer of (1.3), which is called a regularized solution. In particular, we are concerned with an *a priori* choice strategy of α realizing an optimal (or quasi-optimal) rate of the convergence $\lim_{\delta\to 0} f_{\alpha,\delta} = f_0$ in the norm of X. There are many articles concerning *a priori* assumptions on the exact solution f_0 to be reconstructed, which guarantee such a convergence rate: as monographs, see for example, Baumeister [3], Colton and Kress [5], Engl, Hanke and Neubauer [6], Groetsch [9], Hofmann [11], Kirsch [16], Tikhonov and Arsenin [26], Tikhonov, Goncharsky, Stepanov and Yagola [27], Vasin and Ageev [28], and moreover we can refer to Hegland [10], Hohage [15], Mair [17], Mathé and Pereverzev [18, 19], Neubauer [21, 22], Tautenhahn [25] as related papers, for instance.

In the majority of books and papers mentioned above, the authors require so-called *source conditions* in a more or less generalized form which assume that f_0 either belongs to one of the ranges of A^* or a fractional power $(A^*A)^\gamma$ or belongs to the range of an increasing nonnegative index function ρ applied to the operator A^*A. For practical inverse problems for partial differential equations, in general, it is very difficult to characterize such range spaces. Moreover, even though we can characterize $\mathcal{R}(A^*)$, $\mathcal{R}((A^*A)^\gamma)$ or $\mathcal{R}(\rho(A^*A))$, if f_0 is not in those ranges, then the existing strategies do not give any information on convergence rates. Although there are works on adaptation of source conditions (e.g., section 6 in [18]), the existing *a priori* choice strategies do not work for actual inverse problems such as the following example.

Example 1 (inverse wave source problem)**.** Let $\Omega \subset \mathbb{R}^r$ be a bounded domain whose boundary $\partial\Omega$ is of C^2-class. Let $u(f) = u(f)(x,t) \in C\big([0,T]; H_0^1(\Omega)\cap H^2(\Omega)\big)\cap C^1\big([0,T]; H_0^1(\Omega)\big)$ satisfy

$$\begin{cases} \partial_t^2 u(x,t) = \Delta u(x,t) + \lambda(t)f(x), & x\in\Omega, \quad 0<t<T,\\ u(x,t) = 0, & x\in\partial\Omega, \quad 0<t<T,\\ u(x,0) = \partial_t u(x,0) = 0, & x\in\Omega, \end{cases} \tag{1.5}$$

where $\lambda \in C^1[0,\infty)$ is a given function and we assume that $\lambda(0) \neq 0$. Then our inverse wave source problem is the determination of $f \in L^2(\Omega)$ from the boundary observation $\frac{\partial u}{\partial\nu}\big|_{\partial\Omega\times(0,T)}$. This inverse problem is discussed, for example, in Yamamoto [29].

Let $X = L^2(\Omega)$ and $Y = L^2(\partial\Omega\times(0,T))$, and let us define an operator $A : X \longrightarrow Y$ by

$$Af = \frac{\partial u}{\partial\nu}\bigg|_{\partial\Omega\times(0,T)}.$$

Then A is injective whenever $T > \frac{1}{2}\sup_{x,x'\in\Omega}|x-x'|$ [29].

Let us discuss the Tikhonov regularization for this inverse problem:

$$\text{Minimize}\|Af - g_\delta\|^2_{L^2(\partial\Omega\times(0,T))} + \alpha\|f\|^2_{L^2(\Omega)} \qquad \text{over} \quad f\in L^2(\Omega),$$

where $g_\delta \in L^2(\partial\Omega\times(0,T))$ is our available data such that

$$\left\|\frac{\partial u(f_0)}{\partial\nu} - g_\delta\right\|_{L^2(\partial\Omega\times(0,T))} \leqslant \delta.$$

We can prove (e.g., [29]) that there exists a unique minimizer $f_{\alpha,\delta}$ for a given $\alpha > 0$ and that $\mathcal{R}(A^*) \supset H_0^1(\Omega)$ which implies that if $f_0 \in H_0^1(\Omega)$ and $\alpha \sim \delta$ as $\delta \to 0$, then $\|f_{\alpha,\delta} - f_0\|_{L^2(\Omega)} = O(\sqrt{\delta})$ as $\delta\to 0$. Here and henceforth $\alpha\sim\delta$ means that $\alpha = O(\delta)$ and $\delta = O(\alpha)$.

In this example, we can incidentally give a sufficiently large subset of $\mathcal{R}(A^*)$, namely $H_0^1(\Omega)$, but it is extremely difficult to do so for $\mathcal{R}((A^*A)^\gamma)$, because A is not explicitly described and, for example, the spectral properties of A^*A are quite complicated in order to characterize $\mathcal{R}((A^*A)^\gamma)$. Furthermore, we have had no information of the convergence rates in the case of $f_0 \notin H_0^1(\Omega)$, which must be considered if we have to reconstruct a characteristic function $f_0 = \chi_D$ of a unknown subdomain $D \subset \Omega$. Note that $\chi_D \notin H_0^1(\Omega)$.

The purpose of this paper is to give an *a priori* choice strategy for α under more applicable *a priori* information of the exact solution f_0 which is preferably described by means of conventional function spaces, so that we can apply it, for example, to the reconstruction of $f_0 = \chi_D$.

Remark 1. Let us consider a different regularization where we choose a regularizing term with stronger norm than in X:

$$\text{Minimize } \|Af - g_\delta\|_X^2 + \alpha\|f\|_Z^2 \qquad \text{over} \quad f \in Z,$$

where the embedding $Z \subset X$ is continuous (usually compact). If we have a conditional stability estimate $\|f\|_X \leqslant \omega(\|Af\|_Y)$ for any f in a bounded subset of Z, where $\omega = \omega(s) > 0$ is a continuous monotone increasing function such that $\omega(0) = 0$, then Cheng and Yamamoto [4] give an *a priori* choice strategy for α. As for other *a priori* strategy based on conditional stability, see section 3 of chapter 6 in Baumeister [3] for example. In our strategy (1.3), we take a regularizing term $\alpha\|f\|_X^2$ with the same norm as in X, and we do not require any conditional stability with rate function ω.

2. Main result

Henceforth, $\|\cdot\|_X$ and $(\cdot,\cdot)_X$ denote the norm and the scalar product in a Hilbert space X, and $\mathcal{D}(L)$ is the domain of an operator L.

We set

$$\mathcal{I} = \{\rho : [0,\infty) \longrightarrow \mathbb{R};\ \rho \text{ is continuous and increasing and } \rho(0) = 0\}$$

and make use of variable Hilbert scales $\{X_\rho(G)\}_{\rho\in\mathcal{I}}$ as introduced by Hegland [10] (see also [18]) which are generated by an injective compact positive self-adjoint linear operator G in X with an orthonormal basis $\{\varphi_j\}_{j\in\mathbb{N}}$ of its eigenvectors and ordered positive eigenvalues

$$\sigma_1 \geqslant \sigma_2 \geqslant \sigma_3 \geqslant \cdots \longrightarrow 0$$

satisfying $G\varphi_j = \sigma_j\varphi_j$, $j \in \mathbb{N}$. We consider $\rho \in \mathcal{I}$ as an index function. Then the Hilbert space $X_\rho(G)$, $\rho \in \mathcal{I}$, is the completion of

$$\left\{\sum_{j=1}^N c_j\varphi_j;\ N \in \mathbb{N}, c_1, \ldots, c_N \in \mathbb{R}\right\}$$

with respect to the norm

$$\left\|\sum_{j=1}^N c_j\varphi_j\right\|_{X_\rho(G)} = \left(\sum_{j=1}^N \frac{c_j^2}{\rho(\sigma_j)^2}\right)^{\frac{1}{2}}.$$

Note that we can also write $X_\rho(G) = \mathcal{R}(\rho(G))$. Namely, the Hilbert space $X_\rho(G)$ contains just those elements of X which belong to the range of the operator $\rho(G)$ defined by

$$\rho(G)x = \sum_{j=1}^\infty \rho(\sigma_j)(x,\varphi_j)_X\varphi_j, \qquad x \in \mathcal{D}(\rho(G)).$$

Standing assumption. Throughout this paper we assume there exist $\rho_1, \rho_2 \in \mathcal{I}$ such that

$$X_{\rho_1}(G) \subset \mathcal{R}(A^*) = \mathcal{R}\big((A^*A)^{\frac{1}{2}}\big), \tag{2.1}$$

$$f_0 \in X_{\rho_2}(G), \tag{2.2}$$

and

$$\begin{cases} \text{there exists } t_1 \in (0, \sigma_1] \text{ such that} \\ \frac{\rho_2(t)}{\rho_1(t)} \text{ is strictly monotone decreasing in } 0 < t \leqslant t_1, \\ \lim_{t\to 0} \frac{\rho_2(t)}{\rho_1(t)} = \infty, \text{ and there exists a constant } C_1 \geqslant 1 \text{ such that} \\ \max_{t_1 \leqslant t \leqslant \sigma_1} \left(\frac{\rho_2}{\rho_1}\right)(t) \leqslant C_1 \left(\frac{\rho_2}{\rho_1}\right)(t_1). \end{cases} \tag{2.3}$$

In the context of (2.3) we denote by $\left(\frac{\rho_2}{\rho_1}\right)^{-1}$ the inverse function to $\frac{\rho_2}{\rho_1}$, where

$$\left(\frac{\rho_2}{\rho_1}\right)(t) = \eta \qquad \text{for} \quad 0 < t \leqslant t_1 \quad \text{if and only if} \quad t = \left(\frac{\rho_2}{\rho_1}\right)^{-1}(\eta).$$

Moreover, we set

$$[R_1, \infty) = \left\{ \left(\frac{\rho_2}{\rho_1}\right)(t); 0 < t \leqslant t_1 \right\}.$$

Let us remark that the Hilbert spaces $X_{\rho_1}(G)$ and $X_{\rho_2}(G)$ generated by G can be taken rather independently of the forward operator A of equation (1.1) or its spectral properties. In the case where A is compact, for the singular system of A and the eigensystem of G, we require a loose relation (2.1) which is merely an algebraic inclusion. The verification of (2.1) should be done according to a concrete ill-posed problem under consideration. Now we are ready to state.

Theorem 1. *Let us hold standing assumptions (2.1) through (2.3) and denote by $f_{\alpha,\delta}$ the Tikhonov-regularized solution (1.4). For $\delta > 0, \alpha > 0$ and $R \geqslant R_1$, we set*

$$\Psi(R, \alpha; \delta) = \rho_2\left(\left(\frac{\rho_2}{\rho_1}\right)^{-1}(R)\right) + \sqrt{\alpha}R + \frac{\delta}{\sqrt{\alpha}}, \qquad R \geqslant R_1 \tag{2.4}$$

and we assume that, for a given $\delta > 0$, at $R = R(\delta)$ and $\alpha = \alpha(\delta)$, a function Ψ in R and α gains the minimum:

$$\Psi_0(\delta) \equiv \Psi(R(\delta), \alpha(\delta); \delta).$$

Moreover, we assume that $\alpha(\delta) > 0$, and

$$\lim_{\delta\to 0} \alpha(\delta) = 0, \qquad \lim_{\delta\to 0} R(\delta) = \infty. \tag{2.5}$$

Then, setting $\alpha = \alpha(\delta)$, we have

$$\|f_{\alpha,\delta} - f_0\|_X = O(\Psi_0(\delta)) \qquad \text{as} \quad \delta \longrightarrow 0.$$

If our choice guarantees $\lim_{\delta\to 0} \Psi_0(\delta) = 0$, then the conclusion gives a convergence rate of $f_{\alpha,\delta}$ to f_0 as $\delta \longrightarrow 0$.

Although the choices of ρ_1 and ρ_2 are possible only by detailed study of the original ill-posed problem and such studies are not trivial for concrete ill-posed problems, our main theorem can give a flexible strategy for given *a priori* information on f_0:

(1) Find an operator G and $\rho_2 \in \mathcal{I}$ such that (2.2) is satisfied.
(2) Next find $\rho_1 \in \mathcal{I}$ such that (2.1) and (2.3) are satisfied.

In the case where f_0 is assumed to be in a Sobolev space (that is, we assume some finite smoothness *a priori* information), in the step (1), we can usually take the inverse operator to $-\Delta$ with a suitable boundary condition and $\rho_2(t) = t$ or t^μ with $\mu > 0$ (see theorem 2 and examples 3 and 4 in section 5). Then the choice of ρ_1 in step (2) is an essential and difficult part where we need detailed analysis for (1.1). On the other hand, in the existing papers (e.g., [15, 18]), we should first pose that f_0 satisfies a condition called a source condition, and it is frequently difficult to adapt when f_0 is given in an arbitrary *a priori* bounded set. From a strategic viewpoint, we need no such adaptation for f_0, but the choice of ρ_1 can be done after the choice of ρ_2 for any given f_0. In contrast, in the existing strategies, the main issue is to first find the adaptation of a source condition for f_0 and after suitable adaptation, the derivation of a concrete convergence rate is automatic. In sections 4 and 5, we will explain the choices of ρ_1 and ρ_2 in four ill-posed problems.

The assertion of theorem 1 is essentially based on lemma 1 which was presented by Baumeister in [3] as theorem 6.8 on pp 97–98. We set

$$f_\alpha = (A^*A + \alpha I)^{-1} A^* g_0,$$

where we recall that $Af_0 = g_0$. In other words, f_α is the regularized solution for the exact data g_0. Then, we can formulate the key lemma:

Lemma 1. *Set*

$$d_R = \inf\{\|f_0 - A^*g\|_X;\ \|g\|_Y \leqslant R\}.$$

Then,

$$\|f_\alpha - f_0\|_X \leqslant \left(d_R^2 + \alpha R^2\right)^{\frac{1}{2}} \tag{2.6}$$

for all $\alpha > 0$ and $R > 0$.

For completeness we will repeat the proof of lemma 1 in the appendix. Some more discussion concerning the distance function d_R is presented in [12].

3. Proof of theorem 1

First step. First we will estimate $\left(d_R^2 + \alpha R^2\right)^{\frac{1}{2}}$. For this, we show

Lemma 2. *For any $R > 0$, there exists $C_2 > 0$ such that*

$$\left\{w \in X_{\rho_1}(G);\ \|w\|_{X_{\rho_1}(G)} \leqslant C_2 R\right\} \subset \overline{\{A^*g;\ \|g\|_Y \leqslant R\}}. \tag{3.1}$$

Proof of lemma 2. By assumption (2.1), we have

$$\begin{aligned}\left\{w;\ \|w\|_{X_{\rho_1}(G)} \leqslant 1\right\} &= \bigcup_{n=1}^{\infty} \left\{A^*g;\ \|g\|_Y \leqslant n\right\} \cap \left\{w;\ \|w\|_{X_{\rho_1}(G)} \leqslant 1\right\} \\ &\subset \bigcup_{n=1}^{\infty} \overline{\{A^*g;\ \|g\|_Y \leqslant n\} \cap \left\{w;\ \|w\|_{X_{\rho_1}(G)} \leqslant 1\right\}}^{X_{\rho_1}(G)}.\end{aligned}$$

In contrast to the closure $\overline{\{\cdot\}}$ with respect to the norm in X used in formula (3.1) we denote by $\overline{\{\cdot\}}^{X_{\rho_1}(G)}$ the closure with respect to the norm in $X_{\rho_1}(G)$. Then by means of Baire's category theorem (e.g., [30]), there exist $w_0 \in X_{\rho_1}(G)$, $\varepsilon_0 > 0$ and $n_0 \in \mathbb{N}$ such that

$$\begin{aligned}\left\{w;\ \|w - w_0\|_{X_{\rho_1}(G)} \leqslant \varepsilon_0\right\} &\subset \overline{\{A^*g;\ \|g\|_Y \leqslant n_0\} \cap \left\{w;\ \|w\|_{X_{\rho_1}(G)} \leqslant 1\right\}}^{X_{\rho_1}(G)} \\ &\subset \overline{\{A^*g;\ \|g\|_Y \leqslant n_0\}}.\end{aligned} \tag{3.2}$$

Here since $\lim_{n\to\infty}\sigma_n = 0$ and ρ_1 is increasing, we note that $\overline{\mathcal{U}}^{X_{\rho_1}(G)} \subset \overline{\mathcal{U}}$. Then, we can prove that

$$\left\{w;\ \|w\|_{X_{\rho_1}(G)} \leqslant \varepsilon_0\right\} \subset \overline{\{A^*g;\ \|g\|_Y \leqslant 2n_0\}}. \tag{3.3}$$

In fact, since $w_0 \in \overline{\{A^*g;\ \|g\|_Y \leqslant n_0\}}$, there exist $g_m, m \in \mathbb{N}$, such that $\|g_m\|_Y \leqslant n_0$ and $\lim_{m\to\infty}\|A^*g_m - w_0\|_X = 0$. Let $v \in X_{\rho_1}(G)$ be an arbitrary element satisfying $\|v\|_{X_{\rho_1}(G)} \leqslant \varepsilon_0$. Therefore by (3.2), we can choose $\widetilde{g_m}, m \in \mathbb{N}$, such that $\|\widetilde{g_m}\|_Y \leqslant n_0$ and $\lim_{m\to\infty}\|A^*\widetilde{g_m} - (w_0 + v)\|_X = 0$. Therefore, we have chosen $z_m = \widetilde{g_m} - g_m, m \in \mathbb{N}$, such that $\lim_{m\to\infty}\|A^*z_m - v\|_X = 0$ and $\|z_m\|_Y \leqslant \|\widetilde{g_m}\|_Y + \|g_m\|_Y \leqslant 2n_0$. This means that $v \in \overline{\{A^*g;\ \|g\|_Y \leqslant 2n_0\}}$. Since $v \in \left\{w;\ \|w\|_{X_{\rho_1}(G)} \leqslant \varepsilon_0\right\}$ is arbitrary, inclusion (3.3) is valid.

To complete the proof of lemma 2, we set $C_2 = \frac{\varepsilon_0}{2n_0}$. Let $\|w\|_{X_{\rho_1}(G)} \leqslant C_2R$. For $\widetilde{w} = \frac{\varepsilon_0}{C_2R}w$, we then have $\|\widetilde{w}\|_{X_{\rho_1}(G)} \leqslant \varepsilon_0$. Hence, (3.3) yields

$$\widetilde{w} = \frac{\varepsilon_0}{C_2R}w \in \overline{\{A^*g;\ \|g\|_Y \leqslant 2n_0\}},$$

that is,

$$w \in \overline{\left\{A^*\left(\frac{C_2R}{\varepsilon_0}g\right);\ \|g\|_Y \leqslant 2n_0\right\}} = \overline{\{A^*h;\ \|h\|_Y \leqslant R\}}.$$

Thus the proof of lemma 2 is complete. □

Second step. In this step, we estimate from above the error $\|f_0 - f_\alpha\|_X$. Since $f_0 = 0$ implies $g_0 = 0$, $f_\alpha = 0$ and $\|f_0 - f_\alpha\|_X = 0$, we can assume here that $f_0 \neq 0$. We will separately discuss the two cases:

Case 1. $0 < \|f_0\|_{X_{\rho_2}(G)} \leqslant \frac{C_2}{C_1}$.
Case 2. $\|f_0\|_{X_{\rho_2}(G)} \geqslant \frac{C_2}{C_1}$.

Case 1. We will estimate from above

$$\inf_{\|w\|_{X_{\rho_1}(G)}\leqslant C_2R}\|f_0 - w\|_X.$$

Let $t \in (0, t_1)$ be arbitrarily given. Then, we can determine $N \in \mathbb{N}$ such that $\sigma_{N+1} \leqslant t < \sigma_N < t_1$. We set $w = \sum_{n=1}^N (f_0, \varphi_n)\varphi_n$, where $(\cdot,\cdot)$ denotes the scalar product in X. Then, by (2.3), we have

$$\|w\|^2_{X_{\rho_1}(G)} = \sum_{n=1}^N \frac{|(f_0,\varphi_n)|^2}{\rho_1(\sigma_n)^2} = \sum_{n=1}^N \frac{|(f_0,\varphi_n)|^2}{\rho_2(\sigma_n)^2}\left(\frac{\rho_2(\sigma_n)}{\rho_1(\sigma_n)}\right)^2 \leqslant C_1^2\left(\frac{\rho_2(\sigma_N)}{\rho_1(\sigma_N)}\right)^2\|f_0\|^2_{X_{\rho_2}(G)}.$$

Therefore, since ρ_2 is increasing, we obtain

$$\begin{aligned}\|f_0 - w\|_X^2 &= \sum_{n=N+1}^\infty \rho_2(\sigma_n)^2\rho_2(\sigma_n)^{-2}|(f_0,\varphi_n)|^2\\ &\leqslant \rho_2(\sigma_{N+1})^2\sum_{n=N+1}^\infty \frac{|(f_0,\varphi_n)|^2}{\rho_2(\sigma_n)^2} \leqslant \rho_2(\sigma_{N+1})^2\|f_0\|^2_{X_{\rho_2}(G)},\end{aligned}$$

that is,

$$\inf\left\{\|f_0 - w\|_X;\ \|w\|_{X_{\rho_1}(G)} \leqslant C_1\left(\frac{\rho_2}{\rho_1}\right)(\sigma_N)\|f_0\|_{X_{\rho_2}(G)}\right\} \leqslant \rho_2(\sigma_{N+1})\|f_0\|_{X_{\rho_2}(G)}. \tag{3.4}$$

Since $\frac{\rho_2}{\rho_1}$ is decreasing and ρ_2 is increasing in $(0, t_1]$, we have $\left(\frac{\rho_2}{\rho_1}\right)(\sigma_N) < \left(\frac{\rho_2}{\rho_1}\right)(t)$ and $\rho_2(\sigma_{N+1}) \leqslant \rho_2(t)$ for any $t \in [\sigma_{N+1}, \sigma_N)$. Since

$$\left\{w;\ \|w\|_{X_{\rho_1}(G)} \leqslant C_1\left(\frac{\rho_2}{\rho_1}\right)(\sigma_N)\|f_0\|_{X_{\rho_2}(G)}\right\} \subset \left\{w;\ \|w\|_{X_{\rho_1}(G)} \leqslant C_1\left(\frac{\rho_2}{\rho_1}\right)(t)\|f_0\|_{X_{\rho_2}(G)}\right\},$$

by (3.4) we have

$$\inf\left\{\|f_0 - w\|_X;\ \|w\|_{X_{\rho_1}(G)} \leqslant C_1\left(\frac{\rho_2}{\rho_1}\right)(t)\|f_0\|_{X_{\rho_2}(G)}\right\}$$
$$\leqslant \inf\left\{\|f_0 - w\|_X;\ \|w\|_{X_{\rho_1}(G)} \leqslant C_1\left(\frac{\rho_2}{\rho_1}\right)(\sigma_N)\|f_0\|_{X_{\rho_2}(G)}\right\}$$
$$\leqslant \rho_2(\sigma_{N+1})\|f_0\|_{X_{\rho_2}(G)} \leqslant \rho_2(t)\|f_0\|_{X_{\rho_2}(G)}.$$

By means of $C_1\|f_0\|_{X_{\rho_2}} \leqslant C_2$, setting $R = \left(\frac{\rho_2}{\rho_1}\right)(t)$, we have

$$\inf_{\|w\|_{X_{\rho_1}(G)}\leqslant C_2 R} \|f_0 - w\|_X \leqslant \rho_2\left(\left(\frac{\rho_2}{\rho_1}\right)^{-1}(R)\right)\|f_0\|_{X_{\rho_2}(G)}, \qquad R \geqslant R_1.$$

Hence lemma 2 yields

$$\inf_{\|g\|_Y\leqslant R} \|f_0 - A^*g\|_X \leqslant \rho_2\left(\left(\frac{\rho_2}{\rho_1}\right)^{-1}(R)\right)\|f_0\|_{X_{\rho_2}(G)}, \qquad R \geqslant R_1.$$

Thus, by estimate (2.6) of lemma 1, we obtain in this case

$$\|f_0 - f_\alpha\|_X \leqslant \rho_2\left(\left(\frac{\rho_2}{\rho_1}\right)^{-1}(R)\right)\|f_0\|_{X_{\rho_2}(G)} + \sqrt{\alpha}R, \qquad R \geqslant R_1. \tag{3.5}$$

Case 2. Let $C_3 = \frac{1}{2\|f_0\|_{X_{\rho_2}(G)}}\frac{C_2}{C_1}$. Then, $\|C_3 f_0\|_{X_{\rho_2}(G)} \leqslant \frac{C_2}{C_1}$. Therefore, noting that $f_\alpha = (A^*A + \alpha I)^{-1}A^*g_0$, inequality (3.5) of case 1 yields

$$\|f_0 - f_\alpha\|_X \leqslant \rho_2\left(\left(\frac{\rho_2}{\rho_1}\right)^{-1}(R)\right)\|f_0\|_{X_{\rho_2}(G)} + \frac{1}{C_3}\sqrt{\alpha}R, \qquad R \geqslant R_1. \tag{3.6}$$

Third step. In this step, we will complete the proof of theorem 1. We have $f_{\alpha,\delta} = (A^*A + \alpha I)^{-1}A^*g_\delta$ and by the spectral theory $\|(A^*A + \alpha I)^{-1}A^*\|_{\mathcal{L}(Y,X)} \leqslant \frac{1}{2\sqrt{\alpha}}$ as a consequence of $\frac{\sqrt{\lambda}}{\lambda+\alpha} \leqslant \frac{1}{2\sqrt{\alpha}}$ for all $\lambda \geqslant 0$ and $\alpha > 0$ (cf, e.g., formula (2.48) on p 45 in Engl *et al* [6] or, for compact A, theorem 4.13 in Colton and Kress [5]). From (3.5) and (3.6), we then obtain for $\delta > 0$, $\alpha > 0$ and $R \geqslant R_1$

$$\|f_0 - f_\alpha\|_X \leqslant C\left\{\rho_2\left(\left(\frac{\rho_2}{\rho_1}\right)^{-1}(R)\right) + \sqrt{\alpha}R\right\} \tag{3.7}$$

and

$$\|f_{\alpha,\delta} - f_0\|_X \leqslant \|f_\alpha - f_0\|_X + \|f_{\alpha,\delta} - f_\alpha\|_X$$
$$\leqslant \|f_\alpha - f_0\|_X + \|(A^*A + \alpha I)^{-1}A^*\|_{\mathcal{L}(Y,X)}\|g_\delta - g_0\|_Y$$
$$\leqslant C\left\{\rho_2\left(\left(\frac{\rho_2}{\rho_1}\right)^{-1}(R)\right) + \sqrt{\alpha}R + \frac{\delta}{\sqrt{\alpha}}\right\} \leqslant C\Psi(R, \alpha; \delta), \qquad R \geqslant R_1 \tag{3.8}$$

with a constant $C = \max\left\{\|f_0\|_{X_{\rho_2}(G)}, C_3^{-1}, 1\right\}$. This estimate ensures the assertion of theorem 1 and completes the proof. □

In the following sections we will discuss some consequences of theorem 1 with specific choices of ρ_1 and ρ_2 and compare them with the former results in the regularization theory.

4. Hölder-type index functions

In this section, we consider the case where the index functions in formulae (2.1) and (2.2) of the standing assumption are of the form

$$\rho_1(t) = t^\nu, \qquad \rho_2(t) = t^\mu \qquad (0 \leqslant t \leqslant \sigma_1) \quad \text{with fixed exponents} \quad 0 < \mu \leqslant \nu. \tag{4.1}$$

Then $X_{\rho_1}(G)$ and $X_{\rho_2}(G)$ are two elements of a conventional Hilbert scale $\{\widetilde{X}_s(G)\}_{s\in[0,\infty)}$ generated by the operator G with $\widetilde{X}_s(G) = \mathcal{R}(G^s)$, $\widetilde{X}_0(G) = X$ and $\|f\|_{\widetilde{X}_s(G)} = \|G^{-s}f\|_X$. This case is of particular interest if the operator A is *finitely smoothing* in the sense of Mair [17], i.e., if the ordered singular values $\sigma_n(A)$ of the compact operator A decay to zero not faster than a power n^{-p} with a finite exponent $p > 0$ as $n \to \infty$. Then, we can formulate

Theorem 2. *Let us hold with* $0 < \mu \leqslant \nu$

$$\mathcal{R}(G^\nu) \subset \mathcal{R}(A^*) = \mathcal{R}\big((A^*A)^{\frac{1}{2}}\big) \qquad \textit{and} \qquad f_0 \in \mathcal{R}(G^\mu). \tag{4.2}$$

If we denote by $f_{\alpha,\delta}$ *the Tikhonov-regularized solution (1.4), then for the* a priori *regularization parameter choice*

$$\alpha = c_0 \delta^{\frac{2\nu}{\nu+\mu}} \tag{4.3}$$

with some constant $c_0 > 0$ *we obtain the convergence rate*

$$\|f_{\alpha,\delta} - f_0\|_X = O\big(\delta^{\frac{\mu}{\nu+\mu}}\big) \qquad \text{as} \quad \delta \to 0. \tag{4.4}$$

Proof. Note that (4.2) coincides with (2.1)–(2.2) in the standing assumption. Now we distinguish case 1 with $\mu < \nu$, where theorem 2 is a corollary of theorem 1, and case 2 with $\mu = \nu$, where the result is well known (see, e.g., corollary 3.1.3 in Groetsch [9]).

Case 1 ($\mu < \nu$). In this case, the index functions (4.1) satisfy conditions (2.3) with $C_1 = 1$, since $\big(\frac{\rho_2}{\rho_1}\big)(t) = t^{\mu-\nu}$, $t > 0$, is strictly monotone decreasing with $\lim_{t\to 0} t^{\mu-\nu} = \infty$. Then, inequality (2.4) attains the form

$$\|f_{\alpha,\delta} - f_0\|_X \leqslant C\left(R^{\frac{\mu}{\mu-\nu}} + \sqrt{\alpha}R + \frac{\delta}{\sqrt{\alpha}}\right). \tag{4.5}$$

By equating the first and the second terms in the sum of the right-hand side of formula (4.5), we obtain $R = \alpha^{\frac{\mu-\nu}{2\nu}}$. This ansatz for $R = R(\alpha)$ need not be optimal, but implies the error estimate

$$\|f_{\alpha,\delta} - f_0\|_X \leqslant C\left(2\alpha^{\frac{\mu}{2\nu}} + \frac{\delta}{\sqrt{\alpha}}\right),$$

and with *a priori* choice (4.3) for $\alpha = \alpha(\delta)$, we can obtain convergence rate (4.4).

Case 2 ($\mu = \nu$). Here lemma 1 directly applies with $d_R \equiv 0$ for all $R > 0$. This yields $\|f_\alpha - f_0\|_X \leqslant \sqrt{\alpha}R$ and (4.4) whenever α is chosen by (4.3). □

Remark 2. We should note that theorem 2 can be proven alternatively based on the conclusion

$$\mathcal{R}(G^\nu) \subset \mathcal{R}\big((A^*A)^{\frac{1}{2}}\big) \quad \Longrightarrow \quad \mathcal{R}(G^\mu) = \mathcal{R}\big((G^\nu)^{\frac{\mu}{\nu}}\big) \subset \mathcal{R}\big((A^*A)^{\frac{\mu}{2\nu}}\big) \tag{4.6}$$

which is, for $0 < \mu \leqslant \nu$, an immediate consequence of the Heinz–Kato inequality (see, e.g., the corollary of theorem 2.3.3 on p 45 in Tanabe [24] or proposition 8.21 in Engl *et al* [6]) taking into account that, for $s > 0$, the range $\mathcal{R}(G^s)$ of the injective compact operator G^s and the domain $\mathcal{D}(G^{-s})$ coincide. Namely, under assumption (4.2) we obtain from (4.6) a source condition

$$f_0 \in \mathcal{R}((A^*A)^\gamma) \tag{4.7}$$

with $\gamma = \frac{\mu}{2\nu}$. As is well known (see, e.g., corollary 3.1.1 in Groetsch [9]), condition (4.7) provides for any $0 < \gamma \leqslant 1$ an error estimate

$$\|f_\alpha - f_0\|_X \leqslant \widetilde{C}\alpha^\gamma$$

of Tikhonov regularization with a constant $\widetilde{C}$ depending on γ. Similarly, by (3.8), this implies (4.4) if α is chosen according to (4.3).

Now, we return to the inverse wave source problem in $\Omega \subset \mathbb{R}^2$ introduced in example 1 in section 1 and consider (1.5) under the assumption that

$$T > \frac{1}{2} \sup_{x,x'\in\Omega} |x - x'|.$$

Let us recall that we define the linear operator $A : L^2(\Omega) \longrightarrow L^2(\partial\Omega \times (0, T))$ by $Af = \frac{\partial u(f)}{\partial\nu}\big|_{\partial\Omega\times(0,T)}$. We set $X = L^2(\Omega)$ and $Y = L^2(\partial\Omega\times(0, T))$. Let $(Lu)(x) = -\Delta u(x)$, $x \in \Omega$ with $\mathcal{D}(L) = H^2(\Omega) \cap H_0^1(\Omega)$. Then the fractional power L^s, $s > 0$, is defined (e.g., [24]), and $\mathcal{R}(L) = L^2(\Omega)$, $G = L^{-1}$, is a compact and positive self-adjoint operator. Moreover, $G^s = L^{-s}$ and $\mathcal{R}(G^s) = H^{2s}(\Omega)$ if $0 \leqslant s < \frac{1}{4}$, and $\mathcal{R}(G^s) = H_0^{2s}(\Omega)$ if $\frac{1}{4} < s < \frac{3}{4}$ (e.g., [7]). Since theorem 3 in Yamamoto [29] shows that

$$H_0^1(\Omega) \subset \mathcal{R}(A^*),$$

the condition $\mathcal{R}(G^\nu) \subset \mathcal{R}(A^*)$ in (4.2) holds here with $\nu = \frac{1}{2}$. If $f_0 \in H_0^1(\Omega)$, then our theorem 2 recovers theorem 4 in [29]. A more interesting situation is $f_0 = \chi_D$, where χ_D denotes the characteristic function of a smooth subdomain $D \subset \Omega$. Then, by the definition of Sobolev spaces of fractional orders (e.g., [1]), we can verify that $\chi_D \in H^{2\mu}(\Omega) = \mathcal{R}(G^\mu)$ if $\mu < \frac{1}{4}$. Thus, our strategy applies to the reconstruction of a source term concentrating in D. The choice $\alpha = c_0\delta^{\frac{2}{1+2\mu}}$, with $0 < \mu < \frac{1}{4}$, yields

$$\|f_{\alpha,\delta} - f_0\|_{L^2(\Omega)} = O\big(\delta^{\frac{2\mu}{1+2\mu}}\big) \qquad \text{as} \quad \delta \longrightarrow 0.$$

Another approach (see [20]) also yielding convergence rate (4.4) for the Tikhonov regularization with $f_0 \in \mathcal{R}(G^\mu)$ and *a priori* choice (4.3) of the regularization parameter is based on a given *degree of ill-posedness* $\nu > 0$ for the operator A determined by estimates of the form

$$\widehat{C}^{-1}\|f\|_{\mathcal{R}(G^{-\nu})} \leqslant \|Af\|_Y \leqslant \widehat{C}\|f\|_{\mathcal{R}(G^{-\nu})} \qquad \text{for all} \quad f \in X \tag{4.8}$$

with $\|f\|_{\mathcal{R}(G^{-\nu})} = \|G^\nu f\|_X$ and a fixed constant $\widehat{C} > 0$. Taking the dual, we see that (4.8) implies $\mathcal{R}(G^\nu) \subset \mathcal{R}(A^*)$ such that theorem 2 is applicable for $0 < \mu \leqslant \nu$. Example 2 below presents such a situation. However, we should note that requirement (4.8) because of the right inequality can be essentially stronger than the purely algebraic inclusion $\mathcal{R}(G^\nu) \subset \mathcal{R}(A^*)$ in theorem 2.[3]

Example 2 (Abel integral equation)**.** Let $X = Y = L^2(0, 1)$, $0 < \nu \leqslant 1$ and let us consider a linear Abel integral operator $A : X \longrightarrow X$ defined by

$$(Af)(t) = \frac{1}{\Gamma(\nu)} \int_0^t (t - \xi)^{\nu-1} K(t, \xi) f(\xi)\, d\xi, \qquad 0 \leqslant t \leqslant 1.$$

Here $\Gamma(\nu)$ is the gamma function, and $K = K(t, \xi)$ is assumed to satisfy the conditions:

$$\begin{cases} K \in C(\{(t, \xi); 0 \leqslant \xi \leqslant t \leqslant 1\}), \qquad K(t, t) = 1, \quad 0 \leqslant t \leqslant 1, \\ \text{there exists a decreasing function } k \in L^2(0, 1) \text{ such that} \\ \left|\frac{\partial K}{\partial\xi}(t, \xi)\right| \leqslant k(\xi), \qquad 0 \leqslant \xi \leqslant t \leqslant 1. \end{cases}$$

[3] Recently, the authors realized that $\mathcal{R}(G^\nu) \subset \mathcal{R}(A^*)$ implies the left inequality of (4.8) for some constant $\hat{C} > 0$. Details and consequences concerning this fact will be discussed in a forthcoming paper with A Böttcher and U Tautenhahn.

We introduce a Hilbert scale (see [8])

$$\widetilde{X}_s(G) = \begin{cases} H^s(0,1), & 0 \leqslant s < \frac{1}{2}, \\ \left\{u \in H^{\frac{1}{2}}(0,1);\ \int_0^1 (1-t)^{-1}|u(t)|^2\,\mathrm{d}t < \infty\right\}, & s = \frac{1}{2}, \\ \{u \in H^s(0,1);\ u(1) = 0\}, & \frac{1}{2} < s \leqslant 1. \end{cases} \tag{4.9}$$

Then, for our example, we can prove (see theorem 1 in [8]) an inequality chain of form (4.8) and by taking the dual, theorem 2 is applicable.

Remark 3. In the context of formula (4.8) for $X = L^2(0,1)$ and elements $\widetilde{X}_s(G)$ as (4.9), Hilbert scales occur if the operator G corresponds with fractional powers J^β, $\beta > 0$ of the operator $(Jf)(t) = \int_0^t f(\xi)\,\mathrm{d}\xi$, $0 \leqslant t \leqslant 1$ (see, e.g., [8] or [14]). These scales are appropriate for compact integral operators A. In such a case, the exponent $\mu > 0$ in theorem 2 expresses the smoothness of f_0 measured by using a Sobolev scale. On the other hand, a study on non-compact multiplication operators A in section 4 of Hofmann and Fleischer [13] shows that convergence rates of Tikhonov regularization only depend on the smoothing properties of A whenever $f_0 \in L^\infty(0,1)$. For that situation, theorem 2 does not apply.

5. Strictly convex index function and logarithmic convergence rates

In this section, we consider the case where the index functions in formulae (2.1) and (2.2) of the standing assumption are of the form

$$\begin{cases} \rho_1 \in C^2[0,\sigma_1], & \rho_1 \text{ is strictly increasing and is strictly convex in } 0 \leqslant t \leqslant t_1 \leqslant \sigma_1, \\ \rho_2(t) = t, & 0 \leqslant t \leqslant \sigma_1, \qquad \lim_{t\to 0} \frac{t}{\rho_1(t)} = \infty. \end{cases} \tag{5.1}$$

Then, we have $\left(\frac{\rho_2}{\rho_1}\right)'(t) = \frac{\rho_1(t) - t\rho_1'(t)}{\rho_1^2(t)}$ and $(\rho_1 - t\rho_1')' = -t\rho_1'' < 0$ in $0 < t \leqslant t_1$. Hence $(\rho_1 - t\rho_1')(t) < (\rho_1 - t\rho_1')(0) = 0$, which means that $\left(\frac{\rho_2}{\rho_1}\right)' < 0$. Therefore, condition (2.3) is also satisfied. Note that as a consequence of (5.1), the inverse function ρ_1^{-1} is strictly concave in a right neighbourhood of zero with $\lim_{t\to 0} \frac{\rho_1^{-1}(t)}{t} = \infty$. In the following, we focus on situations such that we moreover have

$$\lim_{t\to 0} \frac{t^\kappa}{\rho_1(t)} = \infty \qquad \text{for all exponents} \quad \kappa > 0. \tag{5.2}$$

This case is in particular of interest if A is *infinitely smoothing* in the sense of [17], i.e., for *severely ill-posed* problems (1.1), where the requirements for conventional source conditions $f_0 \in \mathcal{R}((A^*A)^\gamma)$ for some $0 < \gamma < 1$ are rather hard to satisfy (see also [15]). Then we can formulate

Theorem 3. *Let us hold*

$$\mathcal{R}(\rho_1(G)) \subset \mathcal{R}(A^*) = \mathcal{R}\left((A^*A)^{\frac{1}{2}}\right) \qquad \textit{and} \qquad f_0 \in \mathcal{R}(G). \tag{5.3}$$

By $f_{\alpha,\delta}$ we denote the Tikhonov-regularized solution (1.4), and we set

$$\Theta(s) = \rho_1^{-1}(\sqrt{s})\sqrt{s}, \qquad 0 \leqslant s \leqslant s_1. \tag{5.4}$$

Then for the a priori *regularization parameter choice*

$$\alpha = \Theta^{-1}(c_1\delta), \tag{5.5}$$

with some constant $c_1 > 0$, we obtain the convergence rate

$$\|f_{\alpha,\delta} - f_0\|_X = O\left(\rho_1^{-1}\left(\sqrt{\Theta^{-1}(\delta)}\right)\right) \qquad \textit{as} \quad \delta \to 0. \tag{5.6}$$

Proof. This theorem is derived from theorem 1. We equate both terms in the right-hand side of formula (3.7) and have an equation $\frac{\rho_2}{\rho_1}(\sqrt{\alpha}R) = R$, that is, $\frac{\rho_1^{-1}(\sqrt{\alpha})}{\sqrt{\alpha}} = R$. By setting $R(\alpha) = \frac{\rho_1^{-1}(\sqrt{\alpha})}{\sqrt{\alpha}}$ we have $\lim_{\alpha\to 0} R(\alpha) = \infty$, since $\lim_{t\to 0} \frac{\rho_1^{-1}(t)}{t} = \infty$. For this choice, according to (2.4) we can write

$$\Psi(R(\alpha), \alpha; \delta) = 2\rho_1^{-1}(\sqrt{\alpha}) + \frac{\delta}{\sqrt{\alpha}}. \tag{5.7}$$

We note that $\Theta(t)$ and $\rho_1^{-1}(\sqrt{\Theta^{-1}(t)})$ are strictly increasing index functions. Then the parameter choice (5.5) is well defined for sufficiently small $\delta > 0$ and we easily derive the convergence rate

$$\|f_{\alpha,\delta} - f_0\|_X = O\big(\rho_1^{-1}\big(\sqrt{\Theta^{-1}(c_1\delta)}\big)\big) \qquad \text{as} \quad \delta \to 0 \tag{5.8}$$

from formula (5.7). This, however, immediately implies the convergence rate (5.6) to be proven. Namely, we have $\rho_1^{-1}(\sqrt{\Theta^{-1}(c_1\delta)}) \leqslant \max(c_1, 1)\rho_1^{-1}(\sqrt{\Theta^{-1}(\delta)})$ for sufficiently small $\delta > 0$ as a consequence of the monotonicity of $\rho_1^{-1}(\sqrt{\Theta^{-1}(\delta)})$ for $c_1 \leqslant 1$ and as a consequence of $\frac{\rho_1^{-1}(\sqrt{\Theta^{-1}(c_1\delta)})}{c_1\delta} \leqslant \frac{\rho_1^{-1}(\sqrt{\Theta^{-1}(\delta)})}{\delta}$ for $c_1 > 1$. By setting $s = \rho_1^{-1}(\sqrt{\Theta^{-1}(t)})$ it holds $\frac{\rho_1^{-1}(\sqrt{\Theta^{-1}(t)})}{t} = \frac{s}{s\rho_1(s)} = \frac{1}{\rho_1(s)}$ and we easily see that the function $\frac{\rho_1^{-1}(\sqrt{\Theta^{-1}(t)})}{t}$ is decreasing for sufficiently small t. Hence the proof of theorem 3 is complete. □

It should be mentioned that a convergence rate of form (5.6) is *order optimal* and is valid (see, e.g., the remarks in Mathé and Pereverzev [19, p 1265]) if a general source condition

$$f_0 \in \mathcal{R}(\rho_o(A^*A)) \tag{5.9}$$

with the *concave* index function $\rho_0(t) = \rho_1^{-1}(\sqrt{t})\ (0 \leqslant t \leqslant \bar{t})$ is assumed. The interplay between this fact and theorem 3 would be completely evident if we could prove the implication

$$\mathcal{R}(\rho_1(G)) \subset \mathcal{R}\big((A^*A)^{\frac{1}{2}}\big) \quad \Longrightarrow \quad \mathcal{R}(G) \subset \mathcal{R}(\rho_0(A^*A)) \tag{5.10}$$

for every strictly convex index function ρ_1. This would be an essential generalization of the Heinz–Kato inequality, and to our best knowledge, (5.10) is an open problem for general index functions ρ_1.

Example 3 (backward heat equation)**.** Let $\Omega \subset \mathbb{R}^r$ be a bounded domain whose boundary $\partial\Omega$ is of C^2-class. We consider

$$\begin{cases} \partial_t u(x,t) = \Delta u(x,t), & x \in \Omega, \quad 0 < t < T, \\ u(x,t) = 0, & x \in \partial\Omega, \quad 0 < t < T, \\ u(x,0) = f_0(x), & x \in \Omega. \end{cases} \tag{5.11}$$

Let $T > 0$ be arbitrarily fixed and let us discuss the determination of an initial value $f_0(x), x \in \Omega$, by $u(x,T), x \in \Omega$. This is a classical severely ill-posed problem and there are many papers on its analysis and regularization (for example, Ames and Straughan [2], Baumeister [3, chapter 11]). Let $X = L^2(\Omega)$ be a usual real L^2-space, and let $(\cdot,\cdot)$ and $\|\cdot\|$ denote the scalar product and the norm in X, respectively. Let us number the eigenvalues of $-\Delta$ with the homogeneous Dirichlet boundary condition repeatedly according to their multiplicities:

$$0 < \lambda_1 \leqslant \lambda_2 \leqslant \lambda_3 \leqslant \cdots \longrightarrow \infty.$$

Let $\{\varphi_n\}_{n\in\mathbb{N}}$ be corresponding eigenfunctions such that $(\varphi_n, \varphi_n) = 1$. Then, it is known that $\{\varphi_n\}_{n\in\mathbb{N}}$ is an orthonormal basis in X. Moreover, we can represent the solution to (5.11) by

$$u(x,t) = \sum_{n=1}^{\infty} \mathrm{e}^{-\lambda_n t}(f, \varphi_n)\varphi_n(x), \qquad x \in \Omega, t > 0.$$

Therefore, our operator $A : X \longrightarrow X$ is defined by

$$(Af)(x) = \sum_{n=1}^{\infty} \mathrm{e}^{-\lambda_n T}(f, \varphi_n)\varphi_n(x), \qquad x \in \Omega. \tag{5.12}$$

Then by theorem 1 we will derive an *a priori* choice strategy of regularizing parameters in reconstructing f_0 under an *a priori* condition $f_0 \in H_0^1(\Omega) \cap H^2(\Omega)$. We can easily verify

$$(A^*g)(x) = (Ag)(x) = \sum_{n=1}^{\infty} \mathrm{e}^{-\lambda_n T}(g, \varphi_n)\varphi_n(x), \qquad x \in \Omega. \tag{5.13}$$

We choose G as the inverse of the operator $-\Delta$ with the homogeneous Dirichlet boundary condition and set

$$\rho_1(t) = \mathrm{e}^{-\frac{T}{t}}, \qquad \rho_2(t) = t, \quad 0 < t < T. \tag{5.14}$$

Then ρ_1 and ρ_2 satisfy conditions (2.3) and (5.1) with $t_1 = \frac{T}{2}$. Moreover, we can easily see that

$$X_{\rho_1}(G) = \mathcal{R}(A^*), \qquad X_{\rho_2}(G) = H_0^1(\Omega) \cap H^2(\Omega). \tag{5.15}$$

Note that $\sigma_n = \frac{1}{\lambda_n}, n \in \mathbb{N}$, are all the eigenvalues of G, and the norm $\|f\|_{H_0^1(\Omega)\cap H^2(\Omega)}$ is equivalent to $\left(\sum_{n=1}^{\infty} \lambda_n^2 (f, \varphi_n)^2\right)^{\frac{1}{2}}$.

First, we will apply theorem 3. Equation (5.5) is equivalent to

$$\frac{T\sqrt{\alpha}}{\log \frac{1}{\sqrt{\alpha}}} = c_1\delta, \tag{5.16}$$

and so under choice (5.16) of α, we have

$$\|f_{\alpha,\delta} - f_0\| = O\left(\frac{1}{\log \frac{1}{\alpha}}\right)$$

by theorem 3. Since (5.16) is not solved in α explicitly, we will consider a quasi-minimum of Ψ defined by (2.4). As in the proof of theorem 3 we set

$$\left(\frac{\rho_2}{\rho_1}\right)^{-1}(R) = \sqrt{\alpha}R,$$

that is, $1 = \sqrt{\alpha}\,\mathrm{e}^{\frac{T}{\sqrt{\alpha}R}}$. Therefore, we have $R = \frac{2T}{\sqrt{\alpha}\,\log\frac{1}{\alpha}}$. Without loss of generality, we may assume that $\alpha > 0$ is small, so that $R \geqslant R_1$. Then,

$$\Psi\left(\frac{2T}{\sqrt{\alpha}\ \log\frac{1}{\alpha}}, \alpha; \delta\right) = \frac{4T}{\log\frac{1}{\alpha}} + \frac{\delta}{\sqrt{\alpha}}.$$

Let us determine α in the form of

$$\alpha = c_2\delta^{\kappa}, \qquad c_2 > 0, \quad 0 < \kappa < 2. \tag{5.17}$$

Then,

$$\Psi_0(\delta) \equiv \min_{\alpha>0, R\geqslant R_1} \Psi(R,\alpha;\delta) \leqslant \Psi\left(\frac{2T}{c_2^{\frac{1}{2}}\delta^{\frac{\kappa}{2}}\log\frac{1}{c_2\delta^{\kappa}}}, c_2\delta^{\kappa};\delta\right)$$

$$= \frac{4T}{\log\frac{1}{c_2}+\kappa\log\frac{1}{\delta}} + c_2^{-\frac{1}{2}}\delta^{1-\frac{\kappa}{2}} = O\left(\frac{1}{\log\frac{1}{\delta}}\right) \qquad \text{as} \quad \delta \longrightarrow 0.$$

Consequently, we can state one *a priori* strategy for α:

Proposition 1. *In (5.11), we assume that* $f_0 \equiv u(\cdot,0) \in H_0^1(\Omega)\cap H^2(\Omega)$, $g_0 = Af_0$ *and* $\|g_\delta - g_0\|_{L^2(\Omega)} \leqslant \delta$. *Let* $f_{\alpha,\delta}$ *be the minimizer to* $\|Af-g_\delta\|^2_{L^2(\Omega)} + \alpha\|f\|^2_{L^2(\Omega)}$ *over* $f\in L^2(\Omega)$. *If* α *is chosen according to (5.17) for the noise level* δ, *then*

$$\|f_{\alpha,\delta} - f_0\|_{L^2(\Omega)} = O\left(\frac{1}{\log\frac{1}{\delta}}\right) \qquad \textit{as} \quad \delta\longrightarrow 0. \tag{5.18}$$

This proposition realizes the convergence shown in the existing papers by means of the source condition (e.g., theorem 5 and proposition 14 in [15]). In the case where $f_0 \in H_0^{\mu}(\Omega)$ with some $\mu > 0$, we can similarly argue and establish the same convergence rate with the same choice of α. We can expect only the conditional stability of logarithmic type, even if f_0 is *a priori* assumed to be in a Sobolev space of higher order. Thus, this convergence rate of regularized solutions is acceptable and extremely difficult to be improved for general $f_0 \in H_0^{\mu}(\Omega)$. Moreover, the exponent $\kappa \in (0,2)$ in the choice of α for the noise level δ does not influence the convergence rate. Here we consider a simple heat equation only for convenience, but our treatment is the same for a general backward parabolic equation with variable coefficients, and for our strategy, we need not know exact values of the eigenvalues λ_n (cf section 4 of chapter 11 in Baumeister [3]).

We note that the parameter choice (5.17) in proposition 1 is completely different from choice (5.5) in more general theorem 3. More precisely, (5.17) *oversmooths* with respect to (5.5) under assumptions (5.1) and (5.2). On the other hand, choice (5.17) has the advantage that it does not depend on ρ_1. From our standing assumption, (5.1) and (5.2), we derive

$$\|f_{\alpha,\delta}-f_0\|_X \leqslant \widehat{C}\rho_1^{-1}\left(\sqrt{c_2}\delta^{\frac{\kappa}{2}}\right), \qquad \widehat{C} = \widehat{C}(\kappa) > 0 \tag{5.19}$$

with *a priori* choice (5.17) of α, formula (5.7) and an inequality

$$\rho_0(t) = \rho_1^{-1}(\sqrt{t}) \geqslant \widehat{c}t^{\xi}, \qquad \widehat{c} = \widehat{c}(\xi) > 0, \tag{5.20}$$

which is valid for all $\xi > 0$ and sufficiently small $t > 0$ and follows directly from assumption (5.2). Moreover, from (5.19) and (5.20) we obtain the order optimal convergence rate (5.6) also for the *a priori* choice (5.17). It is well known as an intrinsic advantage of the method of Tikhonov regularization that the *a priori* parameter choice (5.17) yields order optimal convergence rates for logarithmic source conditions with index functions $\rho_0(t) = (\log(1/t))^{-\eta}$ in (5.9) *uniformly* for all $\eta > 0$ (e.g., [18, p 802] and for the special case $\kappa = 1$ [15, 17]). More generally, order optimal convergence rates based on (5.9) and (5.17) occur if the twice differentiable and concave index function ρ_0 satisfies limit conditions $\lim_{t\to 0}\frac{\rho_0(t)}{t^{\zeta}} = \infty$ for all $\zeta > 0$. Such requirements are just fulfilled whenever $\rho_0(t) = \rho_1^{-1}(\sqrt{t})$ meets (5.1) and (5.2).

Example 4 (determination of initial temperature by boundary observation)**.** Let $\Omega \subset \mathbb{R}^r$ be a bounded domain whose boundary $\partial\Omega$ is of C^2-class. We consider (5.11). Here $\nu = \nu(x)$

denotes the unit outward normal vector to $\partial\Omega$ at x and we set $\frac{\partial u}{\partial \nu} = \nabla u \cdot \nu$. We discuss the determination of $f_0(x)$, $x \in \Omega$ by boundary observation $\frac{\partial u}{\partial \nu}\big|_{\partial\Omega\times(0,T)}$.

Let us recall that $(Lu)(x) = -\Delta u(x)$, $x \in \Omega$ with $\mathcal{D}(L) = H^2(\Omega) \cap H_0^1(\Omega)$ and let us number the eigenvalues of L according to the multiplicities: $0 < \lambda_1 < \lambda_2 \leqslant \lambda_3 \leqslant \cdots \longrightarrow \infty$. By $\{\varphi_n\}_{n\in\mathbb{N}}$ we denote the corresponding eigenvectors such that $(\varphi_n, \varphi_n) = 1$. Let $X = L^2(\Omega)$ and $Y = L^2(\partial\Omega \times (0, T))$. Then we can represent $A : X \longrightarrow Y$ as

$$(Af)(x, t) = \frac{\partial u}{\partial \nu}(x, t), \qquad x \in \partial\Omega, \quad 0 < t < T.$$

Then our problem is described by (1.1). By using the operator theory [7, 24] and the trace theorem [1], we can see that $A : X \longrightarrow Y$ is bounded. Now, we will determine A^*. We introduce

$$\begin{cases} \partial_t v(x, t) = -\Delta v(x, t), & x \in \Omega, \quad 0 < t < T, \\ v(x, t) = g(x, t), & x \in \partial\Omega, \quad 0 < t < T, \\ u(x, T) = 0, & x \in \Omega. \end{cases} \tag{5.21}$$

For $g \in C_0^\infty(\partial\Omega \times (0, T))$ and $f_0 \in C_0^\infty(\Omega)$, we see that the solutions u and v to (5.11) and (5.21) are sufficiently smooth, so that we can calculate $\int_0^T \int_\Omega (\partial_t u) v \,\mathrm{d}x\,\mathrm{d}t$ by the Green theorem and integration by parts in t. Then, we have

$$(Af_0, g)_{L^2(\partial\Omega\times(0,T))} = \left(\frac{\partial u}{\partial \nu}, v\right)_{L^2(\partial\Omega\times(0,T))} = -(f_0, v(\cdot, 0))_{L^2(\Omega)},$$

which implies

$$A^* g = -v(\cdot, 0), \qquad g \in C_0^\infty(\partial\Omega \times (0, T)). \tag{5.22}$$

In order to verify (2.1), we have to characterize $\mathcal{R}(A^*) = \{A^* g;\ g \in Y\}$. By theorem 2.3 in Russell [23], we know that $\mathcal{R}(A^*) \supset \mathcal{D}\big(\exp\big(C_4 L^{\frac{1}{2}}\big)\big)$, where $C_4 > 0$ is a constant. Since a system $\{\varphi_n\}_{n\in\mathbb{N}}$ of the eigenfunctions is an orthonormal basis in X, we have

$$\exp\big(C_4 L^{\frac{1}{2}}\big) a = \sum_{n=1}^{\infty} \exp\big(C_4 \lambda_n^{\frac{1}{2}}\big)(a, \varphi_n)\varphi_n \tag{5.23}$$

for $a \in \mathcal{D}\big(\exp\big(C_4 L^{\frac{1}{2}}\big)\big)$. Let us choose $G = L^{-1}$ and

$$\rho_1(t) = \exp\left(\frac{-C_4}{\sqrt{t}}\right), \qquad \rho_2(t) = t, \quad t > 0. \tag{5.24}$$

Since $\sigma_n = \frac{1}{\lambda_n}$, $n \in \mathbb{N}$, if $a \in X_{\rho_1}(G)$, then

$$\sum_{n=1}^{\infty} \exp\big(2C_4 \lambda_n^{\frac{1}{2}}\big)(a, \varphi_n)^2 < \infty$$

by the definition of $\|\cdot\|_{X_{\rho_1}(G)}$. Therefore, (5.23) yields $a \in \mathcal{D}\big(\exp\big(C_4 L^{\frac{1}{2}}\big)\big)$ with choice (5.24). We can argue similarly to example 3, so that choice (5.17) of α implies

$$\|f_{\alpha,\delta} - f_0\|_{L^2(\Omega)} = O\left(\frac{1}{\big(\log\frac{1}{\delta}\big)^2}\right) \qquad \text{as} \quad \delta \longrightarrow 0,$$

for $f_0 \in H_0^1(\Omega) \cap H^2(\Omega)$.

Acknowledgments

The main part of this paper was written during the stay of the second author in September 2003 at Chemnitz and during the stay of the first author in January 2004 at Tokyo. The authors thank the Faculty of Mathematics of Chemnitz University of Technology and the Graduate School of Mathematical Sciences of the University of Tokyo for this kind support. The second author is partly supported by grant 15340027 from the Japan Society for the Promotion of Science and grant 15654015 from the Ministry of Education, Culture, Sport and Technology.

Appendix. Proof of lemma 1

First, we note

$$(A^*A + \alpha I)f_\alpha = A^*g_0, \qquad A^*Af_0 = A^*g_0. \tag{A.1}$$

Let $g \in Y$ and $\|g\|_Y \leqslant R$. Then, by (A.1) and the Cauchy–Schwarz inequality, we have

$$\begin{aligned}
\|Af_\alpha - Af_0\|_Y^2 &= (Af_\alpha - Af_0, Af_\alpha - Af_0) = -\alpha\|f_\alpha - f_0\|_X^2 - \alpha(f_0, f_\alpha - f_0)\\
&= -\alpha\|f_\alpha - f_0\|_X^2 - \alpha(f_0 - A^*g, f_\alpha - f_0) - \alpha(g, A(f_\alpha - f_0))\\
&\leqslant -\alpha\|f_\alpha - f_0\|_X^2 + \alpha\|f_0 - A^*g\|_X\|f_\alpha - f_0\|_X + \alpha\|g\|_Y\|Af_\alpha - Af_0\|_Y\\
&\leqslant -\alpha\|f_\alpha - f_0\|_X^2 + \alpha\|f_0 - A^*g\|_X\|f_\alpha - f_0\|_X + \alpha R\|Af_\alpha - Af_0\|_Y.
\end{aligned}$$

Taking the infimum in g, we obtain

$$\|Af_\alpha - Af_0\|_Y^2 \leqslant -\alpha\|f_\alpha - f_0\|_X^2 + \alpha d_R\|f_\alpha - f_0\|_X + \alpha R\|Af_\alpha - Af_0\|_Y.$$

Therefore,

$$\|Af_\alpha - Af_0\|_Y^2 \leqslant -\alpha\|f_\alpha - f_0\|_X^2 + \alpha\left(\tfrac{1}{2}d_R^2 + \tfrac{1}{2}\|f_\alpha - f_0\|_X^2\right) + \tfrac{1}{2}\alpha^2R^2 + \tfrac{1}{2}\|Af_\alpha - Af_0\|_Y^2,$$

so that

$$\frac{\alpha}{2}\|f_\alpha - f_0\|_X^2 + \frac{1}{2}\|Af_\alpha - Af_0\|_Y^2 \leqslant \frac{\alpha}{2}d_R^2 + \frac{1}{2}\alpha^2R^2,$$

which completes the proof of lemma 1.

References

[1] Adams R A 1975 *Sobolev Spaces* (New York: Academic)

[2] Ames K A and Straughan B 1997 *Non-Standard and Improperly Posed Problems* (San Diego, CA: Academic)

[3] Baumeister J 1987 *Stable Solution of Inverse Problems* (Braunschweig: Vieweg)

[4] Cheng J and Yamamoto M 2000 One new strategy for *a priori* choice of regularizing parameters in Tikhonov's regularization *Inverse Problems* **16** L31–8

[5] Colton D and Kress R 1992 *Inverse Acoustic and Electromagnetic Scattering Theory* (Berlin: Springer)

[6] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)

[7] Fujiwara D 1967 Concrete characterization of the domains of fractional powers of some elliptic differential operators of the second order *Proc. Japan Acad.* **43** 82–6

[8] Gorenflo R and Yamamoto M 1999 Operator theoretic treatment of linear Abel integral equations of first kind *Japan. J. Ind. Appl. Math.* **16** 137–61

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

[10] Hegland M 1995 Variable Hilbert scales and their interpolation inequalities with applications to Tikhonov regularization *Appl. Anal.* **59** 207–23

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

[12] Hofmann B 2004 The potential for ill-posedness of multiplication operators occurring in inverse problems *Preprint* 2004-17, Faculty of Mathematics, Chemnitz University of Technology, Chemnitz

[13] Hofmann B and Fleischer G 1999 Stability rates for linear ill-posed problems with compact and non-compact operators *J. Anal. Appl.* **18** 267–86
[14] Hofmann B and Tautenhahn U 1997 On ill-posedness measures and space change in Sobolev scales *J. Anal. Appl.* **16** 979–1000
[15] Hohage T 2000 Regularization of exponentially ill-posed problems *Numer. Funct. Anal. Optim.* **21** 439–64
[16] Kirsch A 1996 *An Introduction to the Mathematical Theory of Inverse Problems* (New York: Springer)
[17] Mair B A 1994 Tikhonov regularization for finitely and infinitely smoothing smoothing operators *SIAM J. Math. Anal.* **25** 135–47
[18] Mathé P and Pereverzev S V 2003 Geometry of linear ill-posed problems in variable Hilbert scales *Inverse Problems* **19** 789–803
[19] Mathé P and Pereverzev S V 2003 Discretization strategy for linear ill-posed problems in variable Hilbert scales *Inverse Problems* **19** 1263–77
[20] Natterer F 1984 Error bounds for Tikhonov regularization in Hilbert scales *Appl. Anal.* **18** 29–37
[21] Neubauer A 1987 Numerical realization of an optimal discrepancy principle for Tikhonov regularization in Hilbert scales *Computing* **39** 43–55
[22] Neubauer A 1988 When do Sobolev spaces form a Hilbert scale? *Proc. Am. Math. Soc.* **102** 587–92
[23] Russell D L 1973 A unified boundary controllability theory for hyperbolic and parabolic partial differential equations *Stud. Appl. Math.* **52** 189–211
[24] Tanabe H 1979 *Equations of Evolution* (London: Pitman)
[25] Tautenhahn U 1998 Optimality for ill-posed problems under general source conditions *Numer. Funct. Anal. Opt.* **19** 377–98
[26] Tikhonov A N and Arsenin V Y 1977 *Solutions of Ill-Posed Problems* (New York: Wiley)
[27] Tikhonov A N, Goncharsky A, Stepanov V and Yagola A 1995 *Numerical Methods for the Solution of Ill-posed Problems* (Dordrecht: Kluwer)
[28] Vasin V V and Ageev A L 1995 *Ill-Posed Problems with A Priori Information* (Utrecht: VSP)
[29] Yamamoto M 1995 Stability, reconstruction formula and regularization for an inverse source hyperbolic problem by a control method *Inverse Problems* **11** 481–96
[30] Yosida K 1980 *Functional Analysis* (New York: Springer)