REGULARIZATION OF EXPONENTIALLY ILL-POSED PROBLEMS

Thorsten Hohage

Institute for Industrial Mathematics/SFB F013
Altenberger Str. 69, A-4040 Linz, Austria

Abstract. Linear and nonlinear inverse problems which are exponentially ill-posed arise in heat conduction, satellite gradiometry, potential theory and scattering theory. For these problems logarithmic source conditions have natural interpretations whereas standard Hölder-type source conditions are far too restrictive. This paper provides a systematic study of convergence rates of regularization methods under logarithmic source conditions including the case that the operator is given only approximately. We also extend previous convergence results for the iteratively regularized Gauß-Newton method to operator approximations.

Key words. regularization, ill-posed problems, logarithmic source conditions, modelling errors, backwards heat equation, sideways heat equation

AMS subject classifications. 65J20, 45L05, 35K05, 65M30

1. Introduction.

Let X, Y be Hilbert spaces and $T : X \to Y$ a bounded linear operator. Our aim is to find the best approximate solution of

$$Tx = y, \qquad y \in \mathrm{ran}(T) \tag{1}$$

i.e. the unique element of $x^\dagger \in X$ that has minimal norm among all minimizers of the residual $\|Tx - y\|$. It is well known (cf., e.g., [6]) that $x^\dagger$ is given by $x^\dagger = T^\dagger y$ where $T^\dagger$ denotes the Moore-Penrose generalized inverse of T. If $\mathrm{ran}(T)$ is not closed in Y, then (1) is ill-posed in the sense that $T^\dagger$ is not continuous. We assume that instead of y only noisy data $y^\delta \in Y$, $\delta \geq 0$ are available satisfying

$$\|y^\delta - y\| \leq \delta. \tag{2}$$

In order to obtain stable approximations to $x^\dagger$, the discontinuous operator $T^\dagger$ has to be approximated by a continuous operator. Usually one considers a familiy of continuous operators converging pointwise to $T^\dagger$. Examples of such *regularization methods* include Tikhonov regularization, Showalter's method, Lardy's method, Landweber iteration, and ν-methods. These methods will be discussed in more detail in Section 3 (cf. also [6, 14, 16] as general references).

Of course, we would like to estimate the error in approximate solutions to (1) computed by regularization methods. It is well known (cf. [6, Proposition 3.11]) that for arbitrary $x^\dagger$ the error may converge to 0 arbitrarily slowly as $\delta \to 0$. Faster convergence can be obtained if $x^\dagger$ satisfies a so-called *source condition.* Standard source conditions have the form

$$x^\dagger = (T^*T)^\mu w, \qquad \|w\| \le \rho \tag{3}$$

with some $\mu > 0$ and are refered to as *Hölder-type source conditions.* For exponentially ill-posed problems these conditions are too restrictive and should be replaced by so-called *logarithmic source conditions*

$$x^\dagger = f_p(T^*T)w, \qquad \|w\| \le \rho \tag{4a}$$

with a parameter $p > 0$ and

$$f_p(\lambda) := \begin{cases} (-\ln\lambda)^{-p}, & 0 < \lambda \le \exp(-1) \\ 0, & \lambda = 0 \end{cases} .$$

In order to avoid the singularity of f_p at $\lambda = 1$, we always assume in this context that the norm in Y is scaled such that

$$\|T^*T\| = \|T\|^2 \le \exp(-1). \tag{4b}$$

In Section 8 we give examples of important inverse problems in heat conduction, satellite gradiometry, and scattering theory where (4) has a natural interpretation as a smoothness condition in terms of Sobolev spaces.

Because of the relevance of the above applications in science and industry, it is desirable to have a convergence theory of regularization methods under source conditions of the form (4). It turns out that there are only a few results available on this topic up to now (cf. [17, 25]). It is the purpose of this paper to provide a systematic study of regularization methods under source conditions of the form (4). The material is organized as follows:

Section 2 collects some known facts about optimality (cf. [17, 23, 25]). It is possible to derive a formula for the best-possible error bounds under a general class of source conditions. A regularization method is called *optimal* under a given source condition if it achieves the best-possible worst-case error estimate. It is called *order optimal* if it is optimal up to a multiplicative constant and *asymptotically optimal* if it is optimal up to a factor tending to 1 as $\delta \to 0$.

Vainikko has constructed optimal regularization methods for Hölder-type source conditions (cf. [16, 27]). Similar results for other source conditions have been established by Tautenhahn ([24, 25]; cf also [23]). Unfortunately, these methods require exact knowledge of all parameters in the source condition. In the simplest case this amounts to knowing the exact smoothness of the true solution and its size in a corresponding Sobolev norm. However, such information is rarely available in practice. Hence, a lot of work has been done to develop order-optimal methods for Hölder-type source conditions that do not require knowledge of the parameters μ and ρ in (3) (cf. [5, 6, 8, 19, 21] and references therein).

It turns out that the situation is different with logarithmic source conditions. In Section 3 we show that a simple a-priori parameter choice rule yields order-optimal convergence for the source conditions (4) *for all* $p, \rho > 0$. This result is valid for most commonly used regularization methods and generalizes a result for Tikhonov regularization which has been derived by Mair [17] using different techniques. In Section 4 we propose a new variant of Morozov's discrepancy principle which leads to asymptotically optimal convergence under logarithmic source conditions. As in Section 3 we do not need a-priori knowledge of the parameters p and ρ in (4).

In Section 5 we prove a converse result stating that the source condition (4), which is sufficient for a logarithmic convergence rate in case of noise-free data, is also "almost" necessary.

Section 6 deals with the situation that the operator T is given only approximately by an operator T_h such that

$$\|T - T_h\| \leq h. \tag{5}$$

The error in the operator T_h may be due to one or a combination of the following sources:

- A physical process is modeled approximately by a bounded linear operator T_h. Our analysis includes the important case that small nonlinearity effects are neglected in the mathematical model.
- T_h is a numerical approximation of T. E.g. we may choose projections $P_h : X \to X_h$ and $Q_h : Y \to Y_h$ onto finite dimensional subspaces and define $T_h := Q_h T P_h$.
- T_h is an integral operator the kernel of which is computed from measured data.

We establish similar convergence results as in Sections 3 and 4 if h tends to 0 sufficiently fast.

In Section 7 we use results from Section 6 to extend previous convergence results for a Newton-type method under logarithmic source conditions (cf. [9]) to the case of approximately given operators.

As mentioned above, Section 8 describes some applications of the theory developed in this paper. Numerical experiments have been reported in [9, 10, 13].

2. Optimality. Assume we want to find the best approximate solution $x^\dagger = T^\dagger y$ to eq. (1), and at our disposal are noisy data y^δ satisfying (2) and the a-priori information that $x^\dagger$ is in some subset $A \subset X$. Let $R : Y \to X$ be an arbitrary mapping to approximately recover $x^\dagger$ from y^δ. Then the *worst case error for* R in this situation is

$$\Delta_R(\delta, A, T) := \sup\{\|Ry^\delta - x^\dagger\| : x^\dagger \in A \wedge \|Tx^\dagger - y^\delta\| \leq \delta\}.$$

The *best possible error bound* is defined as the infimum over all mappings $R : Y \to X$:

$$\Delta(\delta, A, T) := \inf_R \Delta_R(\delta, A, T) \tag{6}$$

We are particularly interested in the case where A is a set of elements which satisfy some source condition, i.e. $A = M_{f,\rho}$ with

$$M_{f,\rho} := \{f(T^*T)w : w \in X \wedge \|w\| \leq \rho\}.$$

Here $f \in C(\sigma(T^*T))$ and $f(0) = 0$. In this case it can be shown (cf. [16, 18]) that the infimum in (6) is actually attained and that

$$\Delta(\delta, M_{f,\rho}, T) = \omega(\delta, M_{f,\rho}, T)$$

with the modulus of continuity defined by

$$\omega(\delta, M_{f,\rho}, T) := \sup\{\|x\| : x \in M_{f,\rho} \wedge \|Tx\| \leq \delta\}.$$

An optimal method is given by a variant of Tikhonov regularization. Unfortunately, the optimal choice of the regularization parameter in this method is very hard to implement in general, and it requires exact knowledge of f and ρ.

The next step is to derive a formula for the modulus of continuity. The following stability estimate is important on its own right and can be established using Jensen's inequality (cf. [17]).

Proposition 1. *Assume that $f \in C([0,\tau])$, $\tau \geq \|T\|^2$, is strictly monotonically increasing with $f(0) = 0$ and that $x^\dagger \in M_{f,\rho}$. Moreover, assume that there exists a convex function $\phi : [0, f(\tau)^2] \to [0, \tau f(\tau)^2]$ satisfying*

$$\phi(\xi) = \xi \cdot (f \cdot f)^{-1}(\xi) \tag{7}$$

for all $\xi \in \{f(\lambda)^2 : \lambda \in \sigma(T^*T) \cup \{0\}\}$. Then $x^\dagger$ satisfies the stability estimate

$$\|x^\dagger\|^2 \leq \rho^2 \phi^{-1}\left(\|Tx^\dagger\|^2/\rho^2\right). \tag{8}$$

With this result we can prove the following formula for the modulus of continuity:

Proposition 2. *Let f satisfy the assumptions of Proposition 1, and let $\tilde{\phi} : [0, f(\tau)^2] \to [0, \tau f(\tau)^2]$ be the greatest convex function satisfying (7) for all $\xi \in \{f(\lambda)^2 : \lambda \in \sigma(T^*T) \cup \{0\}\}$. (Note that $\tilde{\phi}$ is linear on intervals where it is not defined by (7).) Then*

$$\omega(\delta, M_{f,\rho}, T) = \rho\sqrt{\tilde{\phi}^{-1}(\delta^2/\rho^2)} \tag{9}$$

for $\delta \leq \rho\|T\| f(\|T\|^2)$, and $\omega(\delta, M_{f,\rho}, T) = \rho f(\|T\|^2)$ for $\delta > \rho\|T\| f(\|T\|^2)$.

Proof. If $\delta > \rho\|T\| f(\|T\|^2)$, then $\|Tx\| \leq \delta$ holds for all $x \in M_{f,\rho}$, so we have $\omega(\delta, M_{f,\rho}, T) = \sup\{\|x\| : x \in M_{f,\rho}\}$. For $\delta \leq \rho\|T\| f(\|T\|^2)$ the estimate (8) yields "$\leq$" in (9). A proof that this estimate is sharp is given in [11]. ■

For the usual choice $f(\lambda) = \lambda^\mu$, $\mu > 0$, eq. (7) yields the convex function $\phi(\xi) = \xi^{\frac{1+2\mu}{2\mu}}$. This leads to the well-known estimates

$$\|x^\dagger\| \leq \rho^{\frac{1}{1+2\mu}} \|Tx\|^{\frac{2\mu}{1+2\mu}} \qquad \text{and}$$
$$\omega(\delta, M_{\lambda^\mu,\rho}, T) \leq \rho^{\frac{1}{2\mu+1}} \delta^{\frac{2\mu}{1+2\mu}},$$

which are usually proved by interpolation.

The case $f = f_p$ (cf. (4)) has been considered by Mair [17]. He proved that the corresponding functions $\phi_p(\xi) := \xi \exp(-\xi^{-1/2p})$ are convex on $[0,1]$ and that the inverse functions obey

$$\sqrt{\phi_p^{-1}(\lambda)} = f_p(\lambda)(1 + \mathrm{o}(1)), \qquad \lambda \to 0. \tag{10}$$

Consequently, the estimates

$$\|x^\dagger\| \leq \rho f_p\left(\|Tx^\dagger\|^2/\rho^2\right)(1 + \mathrm{o}(1)) \qquad \text{and} \tag{11}$$
$$\omega(\delta, M_{f_p,\rho}, T) \leq \rho f_p\left(\delta^2/\rho^2\right)(1 + \mathrm{o}(1)) \tag{12}$$

hold if the logarithmic source condition (4) is satisfied.

DEFINITION 3. *Let the assumptions of Proposition 1 be satisfied, and let $\phi(\xi)$ be given by (7) for* all $\xi \in [0, f(\tau)^2]$. *Moreover, let $R(\delta) : Y \to X$ be a continuous operator depending on the noise level δ. Convergence of the method $R(\delta)$ on the source set $M_{f,\rho}$ is called*

- optimal *if* $\Delta_{R(\delta)}(\delta, M_{f,\rho}, T) \leq \rho\sqrt{\phi^{-1}(\delta^2/\rho^2)}$
- asymptotically optimal *if* $\Delta_{R(\delta)}(\delta, M_{f,\rho}, T) = \rho\sqrt{\phi^{-1}(\delta^2/\rho^2)}\,(1 + \mathrm{o}(1)), \qquad \delta \to 0$
- of optimal order *if there is a constant $C \geq 1$ such that* $\Delta_{R(\delta)}(\delta, M_{f,\rho}, T) \leq C\rho\sqrt{\phi^{-1}(\delta^2/\rho^2)}$

for δ/ρ sufficiently small.

Note that we have followed the convention to define optimality using the estimate $\rho\sqrt{\phi^{-1}(\delta^2/\rho^2)}$ on $\Delta(\delta, M_{f,\rho}, T)$, although there exists a method R^* described in [18] that is optimal in the strict sense that $\Delta_{R^*} = \Delta$. The term "optimal" is justified in the sense that $\rho\sqrt{\phi^{-1}(\delta^2/\rho^2)}$ is the best-possible estimate on $\Delta(\delta, M_{f,\rho}, T)$ *independent of T*.

3. A-priori parameter choice rules. Approximate solutions to (1) can be constructed by formulae of the form

$$x_\alpha^\delta := g_\alpha(T^*T)T^*y^\delta. \tag{13}$$

Here g_α is a family of piecewise continuous functions parametrized by a *regularization parameter* $\alpha > 0$ such that

$$\lim_{\alpha \searrow 0} g_\alpha(\lambda) = \frac{1}{\lambda} \qquad \text{for all } \lambda > 0 \text{ and} \tag{14}$$

$$\|g_\alpha\|_\infty \leq \frac{C_{\mathrm{g}}}{\alpha}. \tag{15}$$

$g_\alpha(T^*T)$ denotes the functional calculus of T^*T at g_α. The approximate solution for exact data is denoted by $x_\alpha := g_\alpha(T^*T)T^*y$. Most estimates of the total error $\|x^\dagger - x_\alpha^\delta\|$ use the splitting

$$\|x^\dagger - x_\alpha^\delta\| \leq \|x^\dagger - x_\alpha\| + \|x_\alpha - x_\alpha^\delta\| \tag{16}$$

into an *approximation error* $\|x^\dagger - x_\alpha^\delta\|$ and a *propagated data noise error* $\|x_\alpha - x_\alpha^\delta\|$. Obviously, the approximation error is given by

$$x^\dagger - x_\alpha = r_\alpha(T^*T)x^\dagger \tag{17}$$

with $r_\alpha(\lambda) := 1 - \lambda g_\alpha(\lambda)$. We impose the further condition that there exists a number $\bar{\mu} \geq 0$ called the *qualification* of the method such that for all $0 \leq \mu \leq \bar{\mu}$ there exists a constant $C(\mu) > 0$ such that for all $\lambda \in [0, \|T\|^2]$ and $\alpha > 0$

$$\lambda^\mu |r_\alpha(\lambda)| \leq C(\mu)\alpha^\mu \qquad \text{and} \qquad C(0) = 1. \tag{18}$$

Examples of regularization methods of the form (13) satisfying the conditions (14), (15), and (18) include Tikhonov regularization ($r_\alpha(\lambda) = \alpha/(\alpha+\lambda)$; $\bar{\mu} = 1$), truncated singular value decomposition ($r_\alpha(\lambda) = 1$ for $\lambda \leq \alpha$, 0 else; $\bar{\mu} = \infty$) Showalter's method ($r_\alpha(\lambda) = \exp(-\lambda/\alpha)$; $\bar{\mu} = \infty$), Lardy's method ($r_n(\lambda) = (1+\lambda)^{-n}$, $n \in \mathbb{N}$, $\bar{\mu} = \infty$), Landweber iteration ($r_n(\lambda) = (1-\beta\lambda)^n$ with $\beta \geq \|T\|^{-2}$; $\bar{\mu} = \infty$) and ν-methods where the functions r_n are shifted copies of Jacobi polynomials of degree n and $\bar{\mu} = \nu$ ($\nu > 0$) . For Lardy's method and Landweber iteration we set $\alpha = (n+1)^{-1}$, and for ν-methods $\alpha = (n+1)^{-2}$ (cf. [6]).

In this section we consider an *a-priori* parameter choice rule where the regularization parameter α is chosen as a function of δ before computations begin. In Section 4 we will consider an *a-posteriori* parameter choice rule $\alpha = \alpha(\delta, y^\delta)$ which uses quantities arising in the computations to deterimine α. In the terminology of Definition 3 we have $R(\delta)y^\delta = g_{\alpha(\delta,y^\delta)}(T^*T)T^*y^\delta$.

For the propagated data noise error, the assumptions (15), (18) with $\mu = 0$ and Cauchy's inequality imply the frequently used estimate

$$\|x_\alpha - x_\alpha^\delta\|^2 \leq \|TT^* g_\alpha(TT^*)\| \cdot \|g_\alpha(TT^*)\|\delta^2 \leq \frac{2C_{\mathrm{g}}}{\alpha}\delta^2. \tag{19}$$

To estimate the approximation error and its image under T for logarithmic source conditions, we need the following analogues to (18):

$$f_p(\lambda)|r_\alpha(\lambda)| \leq \gamma_1(p) f_p(\alpha) \tag{20}$$

$$\sqrt{\lambda} f_p(\lambda)|r_\alpha(\lambda)| \leq \gamma_2(p)\sqrt{\alpha} f_p(\alpha) \tag{21}$$

for all $\alpha, \lambda \in [0, \exp(-1)]$. The next lemma simplifies and generalizes corresponding results for Tikhonov regularization in [9] and for Landweber iteration in [4].

LEMMA 4. *Assume that $x^\dagger$ satisfies (4).*

1. *Condition (18) with $\bar{\mu} > 0$ implies (20) and*

$$\|x_\alpha - x^\dagger\| \leq \gamma_1(p) f_p(\alpha)\rho \qquad \textit{for all } p > 0. \tag{22}$$

2. *Condition (18) with $\bar{\mu} > 1/2$ implies (21) and*

$$\|Tx_\alpha - Tx^\dagger\| \leq \gamma_2(p)\sqrt{\alpha} f_p(\alpha)\rho. \qquad \textit{for all } p > 0. \tag{23}$$

Proof. We make the substitution $q = \alpha/\lambda$ and write $\tilde{r}(\lambda, q) := r_{q\lambda}(\lambda)$ with $q \in [0, 1/(\mathrm{e}\lambda)]$. Then (18), (20) and (21) transform to

$$\tilde{r}(\lambda, q) \leq C(\mu) q^{\mu}, \tag{24}$$

$$\tilde{r}(\lambda, q) \leq \gamma_1(p) \frac{f_p(q\lambda)}{f_p(\lambda)}, \tag{25}$$

$$\tilde{r}(\lambda, q) \leq \gamma_2(p) \sqrt{q} \frac{f_p(q\lambda)}{f_p(\lambda)}, \tag{26}$$

rsp. Suppose, that (24) holds. Then there exists a constant $C > 0$ depending on p such that for $0 < q \leq 1$ and $0 < \lambda \leq 1/\mathrm{e}$ the estimate

$$\tilde{r}(\lambda, q) \leq C_{\bar{\mu}} q^{\bar{\mu}} \leq C(-\ln q + 1)^{-p} \leq C \left(\frac{\ln q}{\ln \lambda} + 1 \right)^{-p} = C \frac{f_p(q\lambda)}{f_p(\lambda)}$$

holds. For $1 < q \leq 1/(\mathrm{e}\lambda)$, condition (18) with $\mu = 0$ implies that

$$\frac{f_p(q\lambda)}{f_p(\lambda)} > 1 \geq \tilde{r}(\lambda, q).$$

This proves (20). (22) follows from $x^\dagger - x_\alpha = r_\alpha(T^*T) f_p(T^*T) w$ and the isometry of the functional calculus. The proof of (21) is analogous to that of (20). (23) can be derived using the identity $\|Tz\| = \|(T^*T)^{1/2} z\|$. ■

The main result of this section follows easily from the previous lemma.

THEOREM 5. *Assume that (2), (4), (15) and (18) are satisfied with $\bar{\mu} > 0$ and that $\alpha(\delta) = \delta/\rho$. Then the order-optimal error estimate*

$$\|x^\dagger - x_\alpha^\delta\| \leq C(p) \rho f_p(\delta^2/\rho^2)$$

holds for $\delta/\rho \leq \exp(-1)$ with a constant $C(p) \geq 1$.

Proof. Using (16), (19), and (22) we obtain

$$\begin{aligned} \|x^\dagger - x_\alpha^\delta\| &\leq \gamma_1(p) \rho f_p(\alpha) + \sqrt{2C_{\mathrm{g}}} \frac{\delta}{\sqrt{\alpha}} \\ &\leq \gamma_1(p) \rho f_p(\delta/\rho) + \sqrt{2C_{\mathrm{g}}} \rho \sqrt{\delta/\rho} \ \leq \ C(p) \rho f_p \left(\delta^2/\rho^2 \right) \end{aligned}$$

since $f_p(\delta/\rho) = 2^p f_p(\delta^2/\rho^2)$. ■

An inspection of the proof of Theorem 5 reveals that order-optimal convergence also holds for the choice $\alpha = \delta$, but with a constant depending also on ρ. For Tikhonov regularization this has been observed by Mair [17].

4. An a-posteriori rule leading to asymptotically optimal convergence. In this section we consider a posteriori parameter choice rules $\alpha = \alpha(\delta, y^\delta)$. We impose the following conditions:

$$\|Tx_\alpha^\delta - y^\delta\| \le 2\delta \max(-\ln\delta, 1) \tag{27a}$$
$$\alpha \le \exp(-1) \tag{27b}$$
$$\alpha < \exp(-1) \;\Rightarrow\; \exists \alpha' \in [\alpha, 2\alpha] \; \|Tx_{\alpha'}^\delta - y^\delta\| \ge 2\delta \max(-\ln\delta, 1) \tag{27c}$$

As an alternative to (27c) we also consider

$$\alpha < \exp(-1) \;\Rightarrow\; \|Tx_\alpha^\delta - y^\delta\| \ge \frac{3}{2}\delta \max(-\ln\delta, 1). \tag{27d}$$

These conditions are related to Morozov's discrepancy principle where the term $2\delta\max(-\ln\delta, 1)$ is replaced by $\tau\delta$ with some constant $\tau \ge 1$. The next lemma describes a simple algorithm which yields a number $\alpha > 0$ satisfying the conditions (27).

LEMMA 6. *Assume that (2), (14), and (18) hold with $\bar{\mu} \ge 0$ and $\delta > 0$. If (α_n) is some sequence of positive numbers converging monotonically decreasing to 0 with $\alpha_0 = \exp(-1)$ and $\alpha_n/\alpha_{n+1} \le 2$, then the algorithm*

$$\begin{aligned}
&n := 0\\
&\text{while } \|Tx_{\alpha_n}^\delta - y^\delta\| > 2\delta\max(-\ln\delta, 1)\\
&\qquad n := n + 1
\end{aligned}$$

terminates after a finite number N of steps, and $\alpha = \alpha_N$ satisfies (27a)-(27c) with $\alpha' = \alpha_{N-1}$.

Proof. Obviously, it suffices to prove that $\lim_{\alpha\to 0} \|Tx_\alpha^\delta - y^\delta\| < 2\delta$. Note that the identity

$$y^\delta - Tx_\beta^\delta = (I - Tg_\beta(T^*T)T^*)y^\delta = r_\beta(TT^*)y^\delta \tag{28}$$

holds for all $\beta > 0$ since $g_\beta(T^*T)T^* = T^*g_\beta(TT^*)$ (cf. [6, eq. (2.43)]). Let $\{F_\lambda\}$ denote the spectral family of TT^*. Lebesgue's Dominated Convergence Theorem with (14) and (18) for $\mu = 0$ implies

$$\lim_{\alpha\to\infty} \|Tx_\alpha^\delta - y^\delta\|^2 = \lim_{\alpha\to\infty} \int_0^{\|T\|^2+} |r_\alpha(\lambda)|^2 \,\mathrm{d}\|F_\lambda y^\delta\|^2 = \|P_{N(T^*)}y^\delta\|^2$$

where $P_{N(T^*)}$ is the orthogonal projection onto $N(T^*) = \mathrm{ran}(T)^\perp$. Since $y \in \mathrm{ran}(T)$, we have $\|P_{N(T^*)}y^\delta\| = \|P_{N(T^*)}(y^\delta - y)\| \le \delta$. This completes the proof. ∎

For iterative methods, a natural choice of α_n is $\alpha_n = \exp(-1)/(n+1)$ where n is the iteration number. For continuous regularization methods, a parameter α satisfying (27a)-(27c) can be found by a simple bisection algorithm, i.e $\alpha_n = \exp(-1)\cdot 2^{-n}$. For Tikhonov regularization faster convergence is achieved by a Newton method applied to the function $1/\alpha \mapsto \|Tx_\alpha^\delta - y^\delta\|$ (cf. [6, Proposition 9.8]). Here (27d) is more appropriate.

Theorem 7. *Assume (2), (4), (14), (15), (18), (27a), (27b), and either (27c) or (27d) are satisfied with $\bar{\mu} > \frac{1}{2}$. Then the asymptotically optimal error estimate*

$$\|x_\alpha^\delta - x^\dagger\| \leq \rho f_p(\delta^2/\rho^2)(1 + \mathrm{o}\,(1))$$

holds as $\delta \to 0$.

Proof. W.l.o.g we may assume that $\delta \leq \exp(-1)$. Then $\max(-\ln\delta, 1) = -\ln\delta$. We use (16) and treat both terms on the right hand side separately.

1.) To estimate the *approximation error*, we apply Proposition 1 to $x^\dagger - x_\alpha = f_p(T^*T) r_\alpha(T^*T) w$ and obtain

$$\|x^\dagger - x_\alpha\| \leq \rho \sqrt{\phi_p^{-1}\left(\frac{\|y - Tx_\alpha\|^2}{\rho^2}\right)} \tag{29}$$

since $\|r_\alpha(T^*T)w\| \leq C(0)\|w\| \leq \rho$. (27a) and the estimate

$$\|(y - Tx_\alpha) - (y^\delta - Tx_\alpha^\delta)\| = \|r_\alpha(TT^*)(y - y^\delta)\| \leq \delta, \tag{30}$$

which follows from (2), (28), and $\|r_\alpha(TT^*)\| = C(0) = 1$, imply $\|y - Tx_\alpha\| \leq (1 - 2\ln\delta)\delta$. Plugging this into (29) and using (10) yields the upper bound

$$\begin{aligned}\|x^\dagger - x_\alpha\| &\leq \rho \sqrt{\phi_p^{-1}\left(\frac{(1-2\ln\delta)^2\delta^2}{\rho^2}\right)} \\ &= \rho f_p\left((1 - \ln\delta^2)^2 \frac{\delta^2}{\rho^2}\right)(1 + \mathrm{o}\,(1)).\end{aligned}$$

Since

$$\begin{aligned} f_p\left((1-\ln\delta^2)^2\frac{\delta^2}{\rho^2}\right) &= f_p(\delta^2/\rho^2)\left(1 + \frac{-\ln(1-\ln\delta^2)^2}{-\ln\delta^2/\rho^2}\right)^{-p} \\ &= f_p(\delta^2/\rho^2)(1 + \mathrm{o}\,(1)),\end{aligned} \tag{31}$$

we obtain

$$\|x^\dagger - x_\alpha\| \leq \rho f_p(\delta^2/\rho^2)(1 + \mathrm{o}\,(1)).$$

2.) To estimate the *propagated data noise error*, let us first assume that $\alpha < \exp(-1)$. Then, by (27c) and (30),

$$\|Tx_{\alpha'} - y\| \geq \|Tx^\delta_{\alpha'} - y^\delta\| - \delta \geq \delta(-\ln\delta)$$

holds. This together with (23) yields

$$\delta(-\ln\delta) \leq \|Tx_{\alpha'} - y\| \leq \gamma_2(p)\rho\sqrt{\alpha'} f_p(\alpha') \leq \gamma_2(p)\rho\chi_p(\sqrt{\alpha'}) \qquad (32)$$

where $\chi_p(\lambda) := \lambda\, f_p(\lambda)$. Obviously, (32) also holds for $\alpha = \exp(-1)$ if δ is sufficiently small. (19), (32), and the monotonicity of χ_p imply

$$\|x_\alpha - x^\delta_\alpha\| \leq \sqrt{2C_{\mathrm{g}}}\frac{\delta}{\sqrt{\alpha}} \leq 2\sqrt{C_{\mathrm{g}}}\frac{\delta}{\sqrt{\alpha'}} \leq 2\sqrt{C_{\mathrm{g}}}\frac{\delta}{\chi_p^{-1}(\delta(-\ln\delta)/(\gamma_2(p)\rho))}.$$

A very similar calculation gives the same result with (27d) instead of (27c). The functions χ_p^{-1} have the asymptotic behavior $\chi_p^{-1}(\xi) = \xi(-\ln\xi)^p(1+\mathrm{o}\,(1))$ as $\xi \to 0$ (cf. [25]). Hence, a computation similar to (31) yields

$$\begin{aligned}\|x_\alpha - x^\delta_\alpha\| &\leq 2\sqrt{C_{\mathrm{g}}}\gamma_2(p)\rho(-\ln\delta)^{-1} f_p\left(\frac{\delta(-\ln\delta)}{\gamma_2(p)\rho}\right)(1+\mathrm{o}\,(1))\\ &= \rho f_p(\delta^2/\rho^2)\,\mathrm{o}\,(1)\,.\end{aligned}$$

Combining this with the result of the first part gives the assertion. ∎

If $2\delta\max(-\ln\delta, 1)$ in (27a) and (27c) is replaced by $\tau\delta$ with some constant $\tau \geq 1$, then at least order-optimal convergence can be shown.

5. A converse result. The following theorem states that the source condition (4) is not only sufficient for the rate (22) in Lemma 4, but also *almost* necessary.

THEOREM 8. *Assume that (4b), (14), (15), and (18) with $\bar{\mu} \geq 0$ hold. Then the estimate*

$$\|x^\dagger - x_\alpha\| = \mathrm{O}\left(f_q(\alpha)\right), \qquad \alpha \searrow 0 \qquad (33)$$

for some $q > 0$ implies that (4a) holds for all $0 < p < q$.

Proof. The proof is modeled after the proofs of Lemma 4.12. and Proposition 4.13. in [6]. We use (15) in order to obtain the lower bound

$$|r_\alpha(\lambda)| \geq 1 - \lambda|g_\alpha(\lambda)| \geq 1 - \frac{C_{\mathrm{g}}\lambda}{\alpha} \geq \frac{1}{2}$$

for $\lambda \in [0, \frac{\alpha}{2C_{\mathrm{g}}}]$. Then, by (17), we have

$$\begin{aligned}\|x^\dagger - x_\alpha\|^2 &\geq \int_0^{\alpha/2C_{\mathrm{g}}} |r_\alpha(\lambda)|^2\,\mathrm{d}\|E_\lambda x^\dagger\|^2\\ &\geq \frac{1}{4}\int_0^{\alpha/2C_{\mathrm{g}}} \mathrm{d}\|E_\lambda x^\dagger\|^2 = \frac{1}{4}\|E_{\alpha/2C_{\mathrm{g}}} x^\dagger\|^2,\end{aligned}$$

and (33) implies

$$\|E_\epsilon x^\dagger\| = \mathrm{O}\left(f_q(2C_{\mathrm{g}}\epsilon)\right) = \mathrm{O}\left(f_q(\epsilon)\right), \qquad \epsilon \searrow 0. \tag{34}$$

This and a partial integration yield

$$\begin{aligned}\int_\epsilon^{\|T\|^2+} f_p(\lambda)^{-2}\,\mathrm{d}\|E_\lambda x^\dagger\|^2 &= f_p(\|T\|^2)^{-2}\|x^\dagger\|^2 - f_p(\epsilon)^{-2}\|E_\epsilon x^\dagger\|^2 \\ &\quad + \int_\epsilon^{\|T\|^2+} \frac{2p}{\lambda}(-\ln\lambda)^{2p-1}\|E_\lambda x^\dagger\|^2\,\mathrm{d}\lambda \\ &\le C\left(1 + \int_\epsilon^{\|T\|^2+}\frac{2p}{\lambda}(-\ln\lambda)^{2p-1-2q}\,\mathrm{d}\lambda\right) \\ &= C\left(1 + \frac{p}{p-q}(-\ln\lambda)^{2p-2q}\right)\Bigg|_{\lambda=\epsilon}^{\lambda=\|T\|^2}.\end{aligned}$$

for some constant $C > 0$. As $p < q$, we have $\int_0^{\|T\|^2+} f_p(\lambda)^{-2}\,\mathrm{d}\|E_\lambda x^\dagger\|^2 < \infty$, and this implies (4). ∎

6. Operator approximations. In this section we consider an operator equation $F(x) = y$ with an injective, not necessarily linear operator $F : X \supset D(F) \to Y$ that is given only approximately by a linear operator $T_h \in L(X, Y)$, $h > 0$ such that

$$\|F(x) - T_h x\| \le h\|x\| \tag{35}$$

holds for all $x \in X$. Let $x^\dagger \in D(F)$ be the exact solution, $y := F(x^\dagger)$ the exact data and $y^\delta \in Y$ the measured data satisfying (2). Moreover, we assume that $x^\dagger$ fulfils a logarithmic source condition (4) with an operator $T \in L(X, Y)$ satisfying (5) (e.g. $T = F'[x^\dagger]$).

In order to prove convergence rates under logarithmic source conditions with noisy operators, we need an analogue to the estimate

$$\frac{\pi}{4}\|S^\mu - \tilde{S}^\mu\| \le \|S - \tilde{S}\|^\mu, \tag{36}$$

which holds for positive, bounded, linear operators $S, \tilde{S}$, and $0 \le \mu \le 1$ (cf., e.g., [6, §5.2]). This is provided by the following lemma:

Lemma 9. *Let $p > 0$ and $S, \tilde{S}$ positive linear operators in a Hilbert space X with $C := \sqrt{\|S\| + \|\tilde{S}\|} < 1$. Then the estimate*

$$\frac{\pi}{4}\|f_p(S) - f_p(\tilde{S})\| \le f_p(\|S - \tilde{S}\|) + f_p(\sqrt{C})\sqrt{\|S - \tilde{S}\|}$$

holds. Consequently, if $\|\tilde{S} - S\|$ is sufficiently small, then

$$\|f_p(S) - f_p(\tilde{S})\| \leq 2f_p\left(\|S - \tilde{S}\|\right). \tag{37}$$

Proof. Choose $n \in \mathbb{N}_0$ such that $0 < p < n+1$. Then the identity

$$\begin{aligned} A^{-p} &= C(p,n)\int_0^\infty t^{n-p}(tI + A)^{-n-1}\,\mathrm{d}t, \qquad (38) \\ C(p,n) &:= \frac{\sin p\pi}{\pi}\frac{n!}{(1-p)(2-p)\cdots(n-p)}, \end{aligned}$$

holds for any (unbounded) self-adjoint linear operator A with $\sigma(A) \subset [\sigma_0, \infty)$, $\sigma_0 > 0$ (cf. [15, §14]). By the spectral mapping theorem, (38) in particular holds for $A = -\ln S_{|X'}$. Here $X' := N(S)^\perp$, and $S(X') = X'$ since S is self-adjoint. For fixed $t \geq 0$, we have monotonically decreasing convergence of the functions

$$\left(t + \frac{1-\lambda^\mu}{\mu}\right)^{-n-1} \longrightarrow \begin{cases} (t - \ln\lambda)^{-n-1}, & 0 < \lambda \leq \|S\| \\ 0, & \lambda = 0 \end{cases}$$

as $\mu \searrow 0$. (To see this, note that $x \mapsto \frac{1-e^x}{x}$ is monotonically decreasing for $x \in \mathbb{R}$.) By Dini's theorem, convergence is uniform in λ. Hence it follows from the isometry of the functional calculus that

$$\begin{aligned} &\|(tI + \frac{1}{\mu}(I - S_{|X'}{}^\mu))^{-n-1} - (tI - \ln S_{|X'})^{-n-1}\| \\ &= \sup_{\lambda\in\sigma(S_{|X'})}\left|\left(t + \frac{1-\lambda^\mu}{\mu}\right)^{-n-1} - (t-\ln\lambda)^{-n-1}\right| \overset{\mu\searrow 0}{\to} 0. \end{aligned}$$

From this relation and the identity

$$(1-\lambda)^{-n-1} = \sum_{j=0}^\infty \binom{n+j}{n}\lambda^j$$

(with uniform convergence for $0 \leq \lambda \leq \|S\| < 1$), we obtain the representation

$$\begin{aligned} (tI - \ln S_{|X'})^{-n-1} &= \lim_{\mu\searrow 0}\left\{\left(\frac{t\mu+1}{\mu}I - \frac{1}{\mu}S_{|X'}{}^\mu\right)^{-n-1}\right\} \\ &= \lim_{\mu\searrow 0}\left\{\left(\frac{\mu}{t\mu+1}\right)^{n+1}\sum_{j=0}^\infty\binom{n+j}{n}\left(\frac{1}{t\mu+1}\right)^j S_{|X'}{}^{\mu j}\right\}. \end{aligned}$$

Convergence occurs in the operator norm. Plugging the last equation into (38) yields the identity

$$f_p(S)x = C(p,n) \times \int_0^\infty t^{n-p} \lim_{\mu\searrow 0} \left\{ \left(\frac{\mu}{t\mu+1}\right)^{n+1} \sum_{j=0}^\infty \binom{n+j}{n} \left(\frac{1}{t\mu+1}\right)^j S^{\mu j} x \right\} \mathrm{d}t$$

for $x \in X'$. Since this formula is trivially valid for $x \in N(S)$, it holds for all $x \in X$. Subtracting the corresponding formula for $\tilde{S}$, we obtain the estimate

$$\|f_p(S) - f_p(\tilde{S})\| \le C(p,n) \times \int_0^\infty t^{n-p} \lim_{\mu\searrow 0} \left\{ \left(\frac{\mu}{t\mu+1}\right)^{n+1} \sum_{j=0}^\infty \binom{n+j}{n} \left(\frac{1}{t\mu+1}\right)^j \left\| S^{\mu j} - \tilde{S}^{\mu j} \right\| \right\} \mathrm{d}t.$$

We first estimate the terms with $\mu j < 1$ on the right hand side, using (36):

$$\begin{aligned} &\frac{\pi}{4} \left(\frac{\mu}{t\mu+1}\right)^{n+1} \sum_{j<\mu^{-1}} \binom{n+j}{n} \left(\frac{1}{t\mu+1}\right)^j \left\| S^{\mu j} - \tilde{S}^{\mu j} \right\| \\ &\le \left(\frac{\mu}{t\mu+1}\right)^{n+1} \sum_{j=0}^\infty \binom{n+j}{n} \left(\frac{1}{t\mu+1}\right)^j \|S - \tilde{S}\|^{\mu j} \\ &= \left(t + \frac{1 - \|S-\tilde{S}\|^\mu}{\mu}\right)^{-n-1} \\ &\overset{\mu\searrow 0}{\to} \left(t - \ln\|S - \tilde{S}\|\right)^{-n-1}. \end{aligned}$$

To estimate $\|S^{\mu j} - \tilde{S}^{\mu j}\|$ for $\mu j \ge 1$, we choose $M = 2^m$, $m \in \mathbb{N}$, such that $\frac{1}{2} \le \frac{\mu j}{M} < 1$. It can be seen by induction on m that

$$\|S^M - \tilde{S}^M\| \le (\|S\| + \|\tilde{S}\|)^{M-1} \|S - \tilde{S}\|.$$

Hence, by (36),

$$\frac{\pi}{4} \|S^{\mu j} - \tilde{S}^{\mu j}\| \le \|S^M - \tilde{S}^M\|^{\frac{\mu j}{M}} \le C^{\frac{M-1}{M}\mu j} \|S - \tilde{S}\|^{\frac{\mu j}{M}} \le \sqrt{C}^{\mu j} \sqrt{\|S - \tilde{S}\|},$$

and we get

$$
\begin{aligned}
&\frac{\pi}{4}\left(\frac{\mu}{t\mu+1}\right)^{n+1}\sum_{j\geq\mu^{-1}}\binom{n+j}{n}\left(\frac{1}{t\mu+1}\right)^{j}\left\|S^{\mu j}-\tilde{S}^{\mu j}\right\| \\
&\leq \left(\frac{\mu}{t\mu+1}\right)^{n+1}\sum_{j=0}^{\infty}\binom{n+j}{n}\left(\frac{\sqrt{C}^{\mu}}{t\mu+1}\right)^{j}\sqrt{\|S-\tilde{S}\|} \\
&= \left(\frac{\mu}{t\mu+1}\right)^{n+1}\left(1-\frac{\sqrt{C}^{\mu}}{t\mu+1}\right)^{-n-1}\sqrt{\|S-\tilde{S}\|} \\
&\overset{\mu\searrow 0}{\to} \left(t-\ln\sqrt{C}\right)^{-n-1}\sqrt{\|S-\tilde{S}\|}.
\end{aligned}
$$

Putting both estimates together finally gives

$$
\begin{aligned}
\frac{\pi}{4}\|f_p(S)-f_p(\tilde{S})\| \leq C(p,n)\Bigg(&\int_0^{\infty}\frac{t^{n-p}\,\mathrm{d}t}{(t-\ln\|S-\tilde{S}\|)^{n+1}}+ \\
&+\int_0^{\infty}\frac{t^{n-p}\,\mathrm{d}t}{(t-\ln\sqrt{C})^{n+1}}\sqrt{\|S-\tilde{S}\|}\Bigg) \\
=&\ f_p(\|S-\tilde{S}\|)+f_p(C)\sqrt{\|S-\tilde{S}\|}.
\end{aligned}
$$

■

We immediately obtain the following convergence result with an *a-priori* parameter choice rule.

PROPOSITION 10. *Assume that the conditions (4), (5), (15), (18) with $\bar{\mu}>0$ and (35) are satisfied. Then the regularized solutions*

$$
x_\alpha^{\delta,h} := g_\alpha(T_h^*T_h)T_h^*y^\delta \tag{39}
$$

satisfy the error estimate

$$
\|x_\alpha^{\delta,h}-x^\dagger\| \leq C\left(\rho f_p(h)+\rho f_p(\alpha)+\frac{\delta+h\|x^\dagger\|}{\sqrt{\alpha}}\right)
$$

with a constant C independent of h and δ. In particular, for the choice $\alpha = h+\delta/\rho$ the estimate

$$
\|x_\alpha^{\delta,h}-x^\dagger\| \leq C\rho f_p\big((h+\delta/\rho)^2\big)
$$

holds for $h+\delta/\rho$ sufficiently small.

Proof. With $x_\alpha^h := g_\alpha(T_h^* T_h) T_h^* y$ we have

$$\begin{aligned} \|x^\dagger - x_\alpha^{\delta,h}\| &\le \|r_\alpha(T_h^*T_h)x^\dagger\| + \|g_\alpha(T_h^*T_h)T_h^*T_h x^\dagger - x_\alpha^{\delta,h}\| \\ &\le \|r_\alpha(T_h^*T_h)\| \cdot \|f_p(T^*T)w - f_p(T_h^*T_h)w\| \\ &\quad + \|r_\alpha(T_h^*T_h)f_p(T_h^*T_h)w\| + \|g_\alpha(T_h^*T_h)T_h^*\| \cdot \|T_h x^\dagger - y^\delta\|. \end{aligned}$$

On the right hand side of this inequality, the first term is bounded by $C\rho f_p(h)$ due to (18), (37) and the estimate

$$\|T_h^*T_h - T^*T\| \le \|T_h^*(T_h - T)\| + \|(T_h^* - T^*)T\| \le \frac{2}{\mathrm{e}} h < h. \tag{40}$$

The second term is bounded by $\gamma_1(p)\rho f_p(\alpha)$ due to (20). Finally, we have

$$\|T_h x^\dagger - y^\delta\| \le \|T_h - F(x^\dagger)\| + \|y - y^\delta\| \le h\|x^\dagger\| + \delta, \tag{41}$$

due to (2) and (35), so by (19) the last term can be estimated by $C\alpha^{-1/2}(\delta + h\|x^\dagger\|)$. This gives the assertion. ∎

We now consider a modification of an *a-posteriori* parameter choice rule suggested by Plato and Vainikko [20]. Let $Q_h : Y \to \overline{\mathrm{ran}(T_h)}$ be the orthogonal projection onto $\overline{\mathrm{ran}(T_h)}$. We choose the regularization parameter α such that that following conditions are satisfied:

$$\|T_h x_\alpha^{\delta,h} - Q_h y^\delta\| \le 2\delta \max(-\ln\delta, 1) \quad \text{or} \quad \alpha = h \tag{42a}$$

$$h \le \alpha \le \exp(-1) \tag{42b}$$

$$\alpha < \exp(-1) \;\Rightarrow\; \begin{array}{l} \exists \alpha' \in [\alpha, 2\alpha] : \\ \|T_h x_{\alpha'}^{\delta,h} - Q_h y^\delta\| \ge 2\delta \max(-\ln\delta, 1) \end{array} \tag{42c}$$

We also consider the following alternative to (42c):

$$\alpha < \exp(-1) \;\Rightarrow\; \|T_h x_\alpha^{\delta,h} - Q_h y^\delta\| \ge \frac{3}{2}\delta \max(-\ln\delta, 1). \tag{42d}$$

Under the assumptions of Lemma 6 it can be shown that the algorithm

$$\begin{array}{l} n := 0, \qquad \alpha := \exp(-1) \\ \text{while } \big(\|T_h x_{\alpha_n}^{\delta,h} - Q_h y^\delta\| > 2\delta\max(-\ln\delta, 1) \text{ and } \alpha_n > h\big) \\ \qquad n := n + 1 \\ \qquad \alpha := \max(h, \alpha_n) \end{array}$$

terminates after a finite number N of steps and yields a number $\alpha = \alpha_N$ satisfying (42a)-(42c) with $\alpha' = \alpha_{N-1}$ if $N > 0$.

To prove convergence rates with this parameter choice rule we need the following lemma.

LEMMA 11. *For operators $T, T_h \in L(X,Y)$ with $\|T\|, \|T_h\| \le \exp(-1)$ and $\|T - T_h\| \le h$ the estimate*

$$\|Tf_p(T^*T) - T_h f_p(T_h^* T_h)\| \le C\sqrt{h}$$

holds for some constant $C > 0$.

Proof. We use the polar decomposition $T = \tilde{T}|T|$. Here $|T| := (T^*T)^{1/2}$ is self-adjoint and $\tilde{T}$ is a partial isometry. Analogously, we have $T_h = \tilde{T}_h|T_h|$. By the triangle inequality, $\|\tilde{T}_h\| = 1$ (if $T_h \neq 0$), and $\|f_p(|T|^2)\| = \|f_p\|_{\infty,\sigma(|T|^2)} \le 1$ we obtain

$$\begin{aligned}
&\|Tf_p(T^*T) - T_h f_p(T_h^*T_h)\| = \|\tilde{T}|T|f_p(|T|^2) - \tilde{T}_h|T_h|f_p(|T_h|^2)\| \\
&\le \|(\tilde{T} - \tilde{T}_h)|T|\,\| \cdot \|f_p(|T|^2)\| + \|\tilde{T}_h\| \cdot \|\,|T|f_p(|T|) - |T_h|f_p(|T_h|^2)\| \\
&\le \|\tilde{T}|T| - \tilde{T}_h|T_h|\| + \|\tilde{T}_h\big(|T_h| - |T|\big)\| + \|\,|T|f_p(|T|^2) - |T_h|f_p(|T_h|^2)\|.
\end{aligned}$$

By assumption, the first term on the right hand side of this inequality is bounded by h, and due to the normalization of T and T_h we have $h < 1$, so $h < \sqrt{h}$. To estimate the second term, we use (36) and obtain

$$\frac{\pi}{4}\|\,|T| - |T_h|\,\| \le \sqrt{\|T^*T - T_h T_h^*\|} \le \sqrt{h}.$$

To estimate the last term, we proceed as in the proof of Lemma 9 with $S := T^*T$ and $\tilde{S} := T_h^* T_h$ to establish the estimate

$$\|S^{1/2}f_p(S) - \tilde{S}^{1/2}f_p(\tilde{S})\| \le C(p,n)\ \times$$
$$\int_0^\infty t^{n-p} \lim_{\mu\searrow 0}\left(\frac{\mu}{t\mu+1}\right)^{n+1} \sum_{j=0}^{\infty} \binom{n+j}{n} \left(\frac{1}{t\mu+1}\right)^j \left\|S^{\mu j+\frac{1}{2}} - \tilde{S}^{\mu j+\frac{1}{2}}\right\| \,\mathrm{d}t.$$

If $\mu j + \frac{1}{2} \le 1$, we have

$$\frac{\pi}{4}\left\|S^{\mu j+\frac{1}{2}} - \tilde{S}^{\mu j+\frac{1}{2}}\right\| \le \|S - \tilde{S}\|^{\mu j}\sqrt{\|S - \tilde{S}\|} \le C_1^{\mu j}\sqrt{\|S - \tilde{S}\|}$$

with $C_1 = \|S\| + \|\tilde{S}\| \le 2\exp(-1)$ due to (36). If $\mu j + \frac{1}{2} > 1$, we choose $M \in \{2, 2^2, 2^3, \dots\}$ such that $\frac{1}{2} \le \frac{\mu j+1/2}{M} < 1$ to get

$$\begin{aligned}
\frac{\pi}{4}\left\|S^{\mu j+\frac{1}{2}} - \tilde{S}^{\mu j+\frac{1}{2}}\right\| &\le \|S^M - \tilde{S}^M\|^{\frac{\mu j+1/2}{M}} \le C_1^{\frac{M-1}{M}(\mu j+\frac{1}{2})}\|S - \tilde{S}\|^{\frac{\mu j+1/2}{M}} \\
&\le \sqrt{C_1}\sqrt{C_1}^{\mu j}\sqrt{\|S - \tilde{S}\|} \le \sqrt{C_1}^{\mu j}\sqrt{\|S - \tilde{S}\|}.
\end{aligned}$$

Proceeding as in the proof of Lemma 9, we obtain the desired estimate

$$\frac{\pi}{4}\|S^{1/2}f_p(S) - \tilde{S}^{1/2}f_p(\tilde{S})\| \le f_p(\sqrt{C_1})\sqrt{h}.$$

This finishes the proof. ∎

The proof of the next theorem uses ideas from [20].

THEOREM 12. *Assume that the functions r_α and g_α satisfy (14), (15), (18) with $\bar{\mu} > 1/2$, that $x^\dagger$ satisfies a logarithmic source condition (4), that (35) and (5) hold, and that $\alpha = \alpha(\delta, h, y^\delta)$ is chosen such that (42a), (42b), and either (42c) or (42d) are satisfied. Then the estimate*

$$\|x_\alpha^{\delta,h} - x^\dagger\| \le C\rho f_p(h) + \rho f_p\left((\sqrt{h} + \delta/\rho)^2\right)(1 + \mathrm{o}\,(1)). \tag{43}$$

holds for $\sqrt{h} + \delta/\rho$ sufficiently small with a constant C depending only on p and the regularization method.

Proof. Again, we assume w.l.o.g. that $\delta \le \exp(-1)$. As in the proof of Proposition 10 we get the estimate

$$\|x^\dagger - x_\alpha^{\delta,h}\| \le 2\rho f_p(h) + \|f_p(T_h^*T_h)r_\alpha(T_h^*T_h)w\| + C\frac{\delta + h\|x^\dagger\|}{\sqrt{\alpha}} \tag{44}$$

Here and in the following C is a generic constant independent of h and δ.

1.) If $\alpha = h$, then the second term in (44) is bounded by $\gamma_1(p)\rho f_p(h)$ due to (20) which is a sufficient estimate. Let us assume that the other alternative in (42a) holds true. Proposition 1 with w replaced by $r_\alpha(T_h^*T_h)w$ yields

$$\|f_p(T_h^*T_h)r_\alpha(T_h^*T_h)w\| \le \rho\sqrt{\phi_p^{-1}\left(\frac{\|T_h r_\alpha(T_h^*T_h)f_p(T_h^*T_h)w\|^2}{\rho^2}\right)}. \tag{45}$$

From $C(0) = 1$, (4), (5), (41) and Lemma 11 we obtain

$$\begin{aligned}
&\|T_h r_\alpha(T_h^*T_h)f_p(T_h^*T_h)w - Q_h(y^\delta - T_h x_\alpha^{\delta,h})\| \\
&= \|r_\alpha(T_hT_h^*)(T_h f_p(T_h^*T_h)w - Q_h y^\delta)\| \\
&\le \|r_\alpha(T_hT_h^*)\| \cdot \|Q_h(T_h f_p(T_h^*T_h)w - y^\delta)\| \\
&\le \|T_h f_p(T_h^*T_h)w - y^\delta\| \\
&\le \|T_h f_p(T_h^*T_h)w - T f_p(T^*T)w\| + \|Tx^\dagger - T_h x^\dagger\| + \|T_h x^\dagger - y^\delta\| \\
&\le C\sqrt{h}\rho + 2h\|x^\dagger\| + \delta \le C\sqrt{h}\rho + \delta.
\end{aligned}$$

Together with (42a) we get

$$\|T_h r_\alpha(T_h^*T_h)f_p(T_h^*T_h)w\| \le (1 - 2\ln\delta)\delta + C\sqrt{h}\rho.$$

Plugging this into (45) yields

$$\|f_p(T_h^*T_h)r_\alpha(T_h^*T_h)w\| = \rho f_p\big((\sqrt{h}+\delta/\rho)^2\big)(1+\mathrm{o}\,(1))$$

after a computation similar to (31).

2.) We first assume that (42c) holds and show that the last term in (44) is of order $\mathrm{O}\,(f_{p+1}(\delta^2)+f_p(h))$. If $\sqrt{\alpha}\geq \delta/f_{p+1}(\delta^2)$, we immediately get the sufficient estimate $\mathrm{O}\left(f_{p+1}(\delta^2)+\sqrt{h}\right)$ from (42b). Hence, we may assume that $\sqrt{\alpha}<\delta/f_{p+1}(\delta^2)$. Then, for δ sufficiently small, we obtain from (42c) that

$$\begin{aligned}\|r_{\alpha'}(T_hT_h^*)T_hx^\dagger\| &\geq \|r_{\alpha'}(T_hT_h^*)Q_hy^\delta\| - \|Q_h(y^\delta - T_hx^\dagger)\|\\ &\geq \|Q_hy^\delta - T_hx^\delta_{\alpha',h}\| - \delta - h\|x^\dagger\|\\ &\geq \delta(-2\ln\delta - 1) - h\|x^\dagger\|.\end{aligned}$$

Multiplying this inequality by $(\alpha')^{-1/2}$ and using Lemma 4, Lemma 9, (18) with $\mu = 1/2$, and (42b) yields

$$\begin{aligned}\frac{\delta}{\sqrt{\alpha'}}(-\ln\delta^2 - 1) &\leq \frac{1}{\sqrt{\alpha'}}\|r_{\alpha'}(T_hT_h^*)T_hx^\dagger\| + \frac{h}{\sqrt{\alpha'}}\|x^\dagger\|\\ &\leq \frac{1}{\sqrt{\alpha'}}\|r_{\alpha'}(T_hT_h^*)T_h\big(f_p(T^*T) - f_p(T_h^*T_h)w\big)\|\\ &\quad+\frac{1}{\sqrt{\alpha'}}\|r_{\alpha'}(T_hT_h^*)T_hf_p(T_h^*T_h)w\| + \sqrt{2h}\|x^\dagger\|\\ &\leq 2C(\frac{1}{2})f_p(h)\rho + \gamma_2(p)f_p(\alpha')\rho + \sqrt{2h}\|x^\dagger\|.\end{aligned}$$

Now we divide by the last inequality by $(-\ln\delta^2 - 1)$ and use our assumption on α to show that $f_p(\alpha')\leq f_p(\delta^2)(1+\mathrm{o}\,(1))$ (cf. (31)). This yields $\delta/\sqrt{\alpha}\leq \sqrt{2}\delta/\sqrt{\alpha'} = \mathrm{O}\,(f_p(h)f_1(\delta^2) + f_{p+1}(\delta^2))$ and gives the desired estimate on the last term in (44). The proof with (42d) instead of (42c) is very similar. ∎

7. A Newton-type method for nonlinear problems. In this section we consider nonlinear problems of the form

$$F(x) = y, \qquad y\in \operatorname{ran}(F) \tag{46}$$

where $F : X\supset D(F)\to Y$ is a nonlinear operator which is continuously differentiable on its domain $D(F)$. As before, X and Y are Hilbert spaces and y^δ are noisy data satisfying (2). By $x^\dagger$ we denote an exact solution to (46) which is assumed to satisfy the source condition

$$x_0 - x^\dagger = f_p(F'[x^\dagger]^*F'[x^\dagger])w, \qquad \|w\|\leq\rho. \tag{47a}$$

Here x_0 is some initial guess which may incorporate a-priori information about $x^\dagger$. As in the linear case we assume that F is scaled such that

$$\|F'[x^\dagger]\|^2 \leq \exp(-1). \tag{47b}$$

An attractive method to solve (46) is the iteratively regularized Gauß-Newton method (IRGNM) (cf. [1, 2, 4, 9]). The nth iteration step of this method ($n \in \mathbb{N}$) consists in finding the unique element $h_n \in X$ which minimizes the Tikhonov-like functional

$$\|F[x_n^\delta]h + F(x_n^\delta) - y^\delta\|^2 + \alpha_n\|h + x_n^\delta - x_0\|^2 \tag{48}$$

and setting $x_n^\delta := x_{n-1}^\delta + h_n$. The regularization parameters are chosen such that

$$\lim_{n\to\infty} \alpha_n = 0 \quad \text{and} \quad 1 \leq \frac{\alpha_n}{\alpha_{n+1}} \leq \gamma. \tag{49}$$

A convergence analysis of the IRGNM under the logarithmic source condition (47) has been provided in [9] (cf. also [4]). With Lemma 9 these results can easily be extended to the realistic case that in the nth step the operator F is approximated numerically by an differentiable operator $F^{(n)}$ with $D(F^{(n)}) \supset D(F)$, and that $F'[x]$ is approximated by an operator $A_x^{(n)} \in L(X,Y)$. We assume that the following error estimates are satisfied for some $h_n \geq 0$:

$$\begin{aligned} \|F^{(n)}(x_n^\delta) - F(x_n^\delta)\| &\leq f_p(h_n)h_n \\ \max(\|A_x^{(n)} - F'[x]\|, \|A_x^{(n)} - F^{(n)\prime}[x]\|) &\leq h_n \end{aligned} \tag{50}$$

The next theorem also provides a rule how to refine the approximations $F^{(n)}$ and $A_x^{(n)}$ during the iteration such that the order of convergence is not destroyed. This may reduce the total computation time considerably. E.g., for an inverse obstacle scattering problem we found that the computation time was reduced by roughly two thirds with an adaptive choice of regularization parameters (cf. [11] and §8.4).

THEOREM 13. *Let (2), (46), (47), (49) and (50) hold. Moreover, assume that the nonlinearity condition*

$$\begin{aligned} &A_{\bar{x}}^{(n)} = R^{(n)}(\bar{x}, x)A_x^{(n)} + Q^{(n)}(\bar{x}, x) \\ &\|I - R^{(n)}(\bar{x}, x)\| \leq C_R, \quad \|Q^{(n)}(\bar{x}, x)\| \leq C_Q\|A_\dagger^{(n)}(\bar{x} - x)\| \end{aligned} \tag{51}$$

is satisfied for all $x, \bar{x} \in B(x^\dagger, E)$. Finally assume that the stopping index $N \leq \infty$ is determined by

$$\eta\alpha_{N+1} < \delta \leq \eta\alpha_N \tag{52}$$

and that C_R, C_Q, γ, ρ and η are sufficiently small constants. Then the Gauß-Newton iterates x_n^δ, $0 \le n \le N$ are well defined and satisfy the following error estimates:

$$\begin{aligned} \|x_n^\delta - x^\dagger\| &\le E f_p(\alpha_n), & 0 \le n \le N, \\ \|x_N^\delta - x^\dagger\| &\le \mathrm{O}\left(f_p(\delta)\right), & \delta \to 0. \end{aligned}$$

This result can be proved by the same method as [9, Theorem 2.3]. The additional error terms are estimated using Lemma 9. A full proof of Theorem 13 is given in [11, Theorem 4.7]. The smallness conditions in the theorem can be formulated explicitely.

8. Applications. In order to apply the convergence results of this paper to specific problems, it remains to interpret the logarithmic source condition (4) (or (47) for nonlinear problems).

8.1. Backwards Heat Equation.

PROBLEM 1. *Let $\Omega \subset \mathbb{R}^m$ be a smooth (not necessarily bounded) domain, and let $u : \Omega \times (0, \bar{t}) \to \mathbb{R}$ satisfy the heat equation $\frac{\partial u}{\partial t} = \Delta u$ in $\Omega \times (0, \bar{t})$ and the initial condition $u(\cdot, 0) = f$ for some $f \in L^2(\Omega)$. If $\Omega \ne \mathbb{R}^m$, we additionally impose some boundary condition on $\partial\Omega \times (0, \bar{t}]$, e.g. $u = 0$ or $\frac{\partial u}{\partial \nu} = 0$. The problem is to identify the initial values f from measurements of the final values $g := u(\cdot, \bar{t})$.*

It can be shown that the Laplace operator Δ is self-adjoint (i.e. $\langle \Delta x, y\rangle = \langle x, \Delta y\rangle$ for all $x, y \in \mathcal{D}(\Delta)$ and $\mathcal{D}(\Delta) = \mathcal{D}(\Delta^*)$), on both the domains

$$\mathcal{D}(\Delta) = \{v \in H^2(\Omega) : v = 0 \text{ on } \partial\Omega\} \qquad \text{and} \tag{53}$$

$$\mathcal{D}(\Delta) = \{v \in H^2(\Omega) : \frac{\partial v}{\partial \nu} = 0 \text{ on } \partial\Omega\} \tag{54}$$

(cf., e.g., [26, §8.2]). Of course, the boundary conditions have to be understood in the sense of the trace operator. For $\Omega = \mathbb{R}^m$, Δ is self-adjoint on the domain $\mathcal{D}(\Delta) = H^2(\mathbb{R}^m)$.

We interpret the heat equation as an ordinary differential equation for the vector-valued function $t \mapsto u(\cdot, t)$, $[0, \bar{t}] \to L^2(\Omega)$ with an initial condition at $t = 0$ and require that

$$u(\cdot, t) \in \mathcal{D}(\Delta), \qquad 0 < t \le \bar{t}. \tag{55}$$

With (53) or (54), this takes care of the boundary condition, and the heat equation makes sense as an ODE for a vector-valued function. The solution is easily found to be $u(\cdot, t) = \exp(t\Delta) f$, $t \in [0, \bar{t}]$. We interpret this formula in the sense of the functional calculus for unbounded self-adjoint operators. Note that (55) is satisfied since $\Delta \le 0$ and $\sup_{\lambda \le 0} |\lambda \exp(t\lambda)| < \infty$ for $t > 0$.

We can now formulate Problem 1 as an operator equation of the form (1) with $X = Y = L^2(\Omega)$ and $T = T_{\mathrm{BH}}$ given by $T_{\mathrm{BH}} := \frac{1}{\sqrt{\mathrm{e}}}\exp(\bar{t}\Delta)$. The scaling factor $1/\sqrt{\mathrm{e}}$ has been introduced to meet (4b).

PROPOSITION 14. ran $\Big(f_p(T_{\mathrm{BH}}^* T_{\mathrm{BH}})\Big) = \mathcal{D}\Big((I-\Delta)^p\Big)$.

Proof. We have $f_p(T_{\mathrm{BH}}^* T_{\mathrm{BH}}) = \big(f_p \circ \frac{1}{\mathrm{e}}\exp 2\bar{t}\cdot\big)(\Delta) = (I - 2\bar{t}\Delta)^{-p}$. Recall that $\mathcal{D}\Big((I-2\bar{t}\Delta)^p\Big) = \big\{v \in L^2(\Omega) : \int_0^\infty (1+2\bar{t}\lambda)^{2p}\,\mathrm{d}\|E_\lambda v\|^2 < \infty\big\}$ where $\{E_\lambda\}$ is the spectral family of $-\Delta$. Since there are constants $c, C > 0$ such that $c(1+\lambda)^p \le (1+2\bar{t}\lambda)^p \le C(1+\lambda)^p$ for all $\lambda \ge 0$, the assertion follows. ■

REMARK 15. *If $\Omega = \mathbb{R}^m$, then $\mathcal{D}((I-\Delta)^p) = H^{2p}(\Omega)$. If $\partial\Omega \neq 0$ and $p \in \mathbb{N}$, we have*

$$\mathcal{D}\Big((I-\Delta)^p\Big) = \{v \in H^{2p}(\Omega) : \Delta^q v = 0 \text{ on } \partial\Omega \text{ for } q = 0,1,\dots,p-1\}$$
$$\mathcal{D}\Big((I-\Delta)^p\Big) = \{v \in H^{2p}(\Omega) : \frac{\partial \Delta^q v}{\partial \nu} = 0 \text{ on } \partial\Omega \text{ for } q = 0,1,\dots,p-1\}$$

in case of Dirichlet an Neumann boundary conditions, rsp.

We also refer to [25] for related results.

8.2. An inverse problem in satellite gradiometry. The following problem has been studied, e.g., in [7, 22].

PROBLEM 2. *Let u satisfy the Laplace equation $\Delta u = 0$ in $\{x \in \mathbb{R}^{m+1} : |x| > 1\}$. For $m = 2$, u describes the gravitational potential of the earth in a spherical framework if the radius of the earth is normalized to 1. The behavior of u at infinity is described by $|u(x)| = \mathrm{O}\left(|x|^{1-m}\right)$ as $|x| \to \infty$. The problem is to determine the potential $f = u\big|_{S^m}$ at the surface $S^m := \{x \in \mathbb{R}^{m+1} : |x| = 1\}$ of the earth from satellite measurements of $g = \frac{\partial^2 u}{\partial r^2}\big|_{RS^m}$ ($r = |x|$) at $RS^m = \{x \in \mathbb{R}^{m+1} : |x| = R\}$, $R > 1$. Here $-\nabla u$ describes the gravitational force and $-\frac{\partial^2 u}{\partial r^2}$ the rate of change of the gravitational force in radial direction.*

Let v be the solution to the interior Laplace equation with Dirichlet boundary values f. With $r = |x|$ and $\hat{x} = x/|x|$, v is given by the Poisson formula

$$v(r\hat{x}) = v(\hat{x}, r) = \frac{1-r^2}{\gamma_m}\int_{S^m} \frac{f(\hat{y})\,\mathrm{d}s(y)}{|r\hat{x} - \hat{y}|^{m+1}}, \tag{56}$$

where $\gamma_m = 2\pi^{(m+1)/2}/\Gamma((m+1)/2)$ is the surface area of S^m. An elementary calculation using the formula

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{m+1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\Delta_{S^m} \tag{57}$$

yields $u(\hat{x},r) = r^{1-m}v(\hat{x},\frac{1}{r})$. Thus, we may formulate Problem 2 as a first kind integral equation of the form (1) with $T = T_{\rm SG} : L^2(S^m) \to L^2(S^m)$ given by

$$(T_{\rm SG}f)(\hat{x}) := \frac{c}{\gamma_m}\int_{S^m}\frac{\partial^2}{\partial R^2}\left\{R^{1-m}\frac{1-R^{-2}}{|R^{-1}\hat{x}-\hat{y}|^{m+1}}\right\}f(\hat{y})\,\mathrm{d}s(\hat{y}).$$

Here the scaling constant c is chosen such that (4b) is fulfilled. Since $T_{\rm SG}$ has an infinitely smooth kernel, Problem 2 is severely ill-posed. The next proposition shows that again logarithmic sources conditions are smoothness conditions in terms of Sobolev spaces.

PROPOSITION 16. $\operatorname{ran}\Big(f_p(T_{\rm SG}^*T_{\rm SG})\Big) = H^p(S^m)$.

Proof. In analogy to the analysis in §8.1 we interpret the Laplace equation with formula (57) for the Laplacian as a second order ODE in r with an initial condition at $r=1$. Introducing $\Lambda := \left(-\Delta_{S^m} + \left(\frac{m-1}{2}\right)\right)^{1/2}$, we obtain two solutions, one defined for $r \geq 1$, the other one for $0 < r \leq 1$:

$$u(\cdot,r) = r^{-(m-1)/2}r^{-\Lambda}f, \qquad v(\cdot,r) = r^{-(m-1)/2}r^{\Lambda}f,$$

This implies $\frac{\partial^2}{\partial r^2}u(\cdot,r) = \left(-\frac{m-1}{2}-\Lambda\right)\left(-\frac{m+1}{2}-\Lambda\right)r^{-\frac{m+3}{2}-\Lambda}f$, i.e. $T_{\rm SG}^*T_{\rm SG} = \chi(\Lambda)$ with $\chi(\lambda) := c^2\left(-\frac{m-1}{2}-\lambda\right)^2\left(-\frac{m+1}{2}-\lambda\right)^2R^{-(m+3)-2\lambda}$. An easy calculation shows that $\frac{1}{f_p(\chi(\lambda))} = \Big(-\ln(\chi(\lambda))\Big)^p = (\ln R^2)\lambda^p + \Xi(\lambda)$ where $\Xi(\lambda)$ is bounded for $\lambda \in [\frac{m-1}{2},\infty) \supset \sigma(\Lambda)$. Hence, as in the proof of Proposition 14,

$$\operatorname{ran}\Big(f_p(T_{\rm SG}^*T_{\rm SG})\Big) = \operatorname{ran}\Big((f_p\circ\chi)(\Lambda)\Big) = \mathcal{D}\Big((\ln R^2)\Lambda^p + \Xi(\Lambda)\Big) = \mathcal{D}(\Lambda^p),$$

and the assertion follows. ∎

8.3. Sideways Heat Equation.

PROBLEM 3. *Let $u : [0,1]\times\mathbb{R}\to\mathbb{C}$ satisfy the heat equation $u_t(t,x) = u_{xx}(t,x)$. The interval $[0,1]$ may represent some heat conducting medium (e.g. the wall of a furnace) where one side ($x=0$) is accessible and the other one ($x=1$) is not. We assume the accessible side to be insulated, i.e. $u_x(0,t) = 0$ for $t\in\mathbb{R}$. The problem is to determine the temperature $g = u(1,\cdot)$ at the inaccessible side from measurements of the temperature at the accessible side $f = u(0,\cdot)$.*

We make a Fourier transform in the time variable, i.e. we consider the functions $(\mathcal{F}u)(x,\omega) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}\langle\omega,t\rangle}u(x,t)\,\mathrm{d}t$. Under appropriate smoothness assumptions we have $\mathcal{F}u_{xx} = (\mathcal{F}u)_{xx}$ and $\mathcal{F}u_t = \mathrm{i}\omega\mathcal{F}u$. Hence, for each $\omega\in\mathbb{R}$ the function $\mathcal{F}u(\cdot,\omega)$ is characterized by the differential equation $(\mathcal{F}u)_{xx} = \mathrm{i}\omega\mathcal{F}u$ and the boundary conditions $(\mathcal{F}u)_x(0,\omega) = 0$ and

$(\mathcal{F}u)(1,\omega) = \mathcal{F}f(\omega)$. The solution to these boundary value problems is $(\mathcal{F}u)(x,\omega) = \frac{\cosh\sqrt{\mathrm{i}\omega}x}{\cosh\sqrt{\mathrm{i}\omega}}(\mathcal{F}f)(\omega)$ with $\sqrt{\mathrm{i}\omega} = \sqrt{\frac{\omega}{2}} + \mathrm{i}\sqrt{\frac{\omega}{2}}$. Therefore, we describe Problem 3 by an operator equation (1) with $X = Y = L^2(\mathbb{R})$ and $T_{\mathrm{SH}} := \mathcal{F}^{-1}c(\cosh\sqrt{\mathrm{i}\omega})^{-1}\mathcal{F}$. Here $c(\cosh\sqrt{\mathrm{i}\omega})^{-1}$ denotes a multiplication operator, and the scaling constant $c := \exp(-1/2)\inf_\omega|\cosh\sqrt{\mathrm{i}\omega}| > 0$ has been introduced in order to meet (4b).

PROPOSITION 17. $\mathrm{ran}\left(f_p(T_{\mathrm{SH}}^*T_{\mathrm{SH}})\right) = H^{p/2}(\mathbb{R})$.

Proof. If $A : X \to X$ is self-adjoint and $W : Y \to X$ is unitary, then $f(W^{-1}AW) = W^{-1}f(A)W$. By Plancherel's Theorem and $T_{\mathrm{SH}}^*T_{\mathrm{SH}} = \mathcal{F}^{-1}c^2|\cosh\sqrt{\mathrm{i}\omega}|^{-2}\mathcal{F}$, this implies that $\mathrm{ran}\left(f_p(T_{\mathrm{SH}}^*T_{\mathrm{SH}})\right)$ consists of those functions $\varphi \in L^2(\mathbb{R})$ for which $\left(f_p(c^2|\cosh\sqrt{\mathrm{i}\omega}|^{-2})\right)^{-1}\mathcal{F}\varphi \in L^2(\mathbb{R})$. From the identity $|\cosh\sqrt{\mathrm{i}\omega}|^2 = \sinh^2\sqrt{\frac{\omega}{2}} + \cos^2\sqrt{\frac{\omega}{2}}$ it is easily seen that there are constants $c, C > 0$ such that $c(1+\omega^2)^{p/4} \le \left(f_p((c^2|\cosh\sqrt{\mathrm{i}\omega}|)^{-2})\right)^{-1} \le C(1+\omega^2)^{p/4}$ for all $\omega \in \mathbb{R}$. Now the assertion follows from the relation (or definition) $H^{p/2}(\mathbb{R}) = \left\{\varphi \in L^2(\mathbb{R}) : (1+|\omega|^2)^{p/4}\mathcal{F}\varphi \in L^2(\mathbb{R})\right\}$. ∎

A related, but mildly ill-posed problem has been investigated by Tautenhahn [24]. A general class of deconvolution problems was considered by Mair [17].

8.4. Further applications. A linear control problem for the heat equation has been considered by Carthel, Glowinski and Lions [3]. The aim is to find a suitable control at the boundary, given zero initial temperature, such that the final temperature approximates some desired function. By a rather sophisticated argument using Müntz approximation, Kindermann [13] has shown that logarithmic source conditions for this problem can be interpreted as smoothness conditions in Sobolev spaces. Numerical experiments in [13] nicely exhibit logarithmic convergence rates.

The problem of Laplace transform inversion has been considered in [17] using weighted Sobolev spaces.

Nonlinear exponentially ill-posed problems arise in inverse obstacle problems. E.g., an important problem is to identify the shape of a homogeneous heat source from measurements of the heat flux (cf. [9]). In inverse scattering problems, one tries to find the shape of a scattering obstacle from far-field measurements of the scattered field corresponding to one incident wave in the resonance region (cf. [10, 11, 12]). Again, logarithmic source conditions are essentially equivalent to smoothness conditions in Sobolev spaces, and numerical experiments with the iteratively regularized Gauß-Newton method exhibit logarithmic convergence rates.

Acknowledgments. This work has been supported by the Austrian Science Fund (FWF) under the project SFB F013 "Numerical and Symbolic Scientific Computing".

REFERENCES

[1] A. B. Bakushinskii, *The problem of the convergence of the iteratively regularized Gauss-Newton method*, Comput. Maths. Math. Phys. **32** (1992), 1353–1359.

[2] B. Blaschke, A. Neubauer, and O. Scherzer, *On convergence rates for the iteratively regularized Gauß-Newton method*, IMA Journal of Numerical Analysis **17** (1997), 421–436.

[3] C. Carthel, R. Glowinski, and J. P. Lions, *On exact and approximate boundary controllabilities for the heat equation: A numerical approach*, JOTA **82** (1994), 429–484.

[4] P. Deuflhard, H. Engl, and O. Scherzer, *A convergence analysis of iterative methods for the solution of nonlinear ill-posed problems under affinely invariant conditions*, Inverse Problems **14** (1998), 1081 – 1106.

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

[6] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, Kluwer Academic Publisher, Dordrecht Boston London, 1996.

[7] W. Freeden and F. Schneider, *Regularization wavelets and multiresolution*, Inverse Problems **14** (19998), 225–243.

[8] M. Hanke and H. W. Engl, *An optimal stopping rule for the ν-method for solving ill-posed problems using christoffel functions*, J. Approx. Theory **79** (1994), 89–108.

[9] T. Hohage, *Logarithmic convergence rates of the iteratively regularized Gauß-Newton method for an inverse potential and an inverse scattering problem*, Inverse Problems **13** (1997), 1279–1299.

[10] ———, *Convergence rates of a regularized Newton method in sound-hard inverse scattering*, SIAM J. Numer. Anal. **36** (1998), 125–142.

[11] ———, *Iterative methods in inverse obstacle scattering: Regularization theory of linear and nonlinear exponentially ill-posed problems*, Ph.D. thesis, University of Linz, 1999.

[12] T. Hohage and C. Schormann, *A Newton-type method for a transmission problem in inverse scattering*, Inverse Problems **14** (1998), 1207–1227.

[13] S. Kindermann, *Convergence rates of the Hilbert uniqueness method via Tikhonov regularization*, JOTA **103** (1999), 657–671.

[14] A. Kirsch, *An introduction to the mathematical theory of inverse problems*, Springer, New York, Berlin, Heidelberg, 1996.

[15] M. A. Krasnoselskii, P. P. Zabreiko, E. I. Pustylnik, and P. E. Sbolevskii, *Integral operators in spaces of summable functions*, Noordhoff International Publishing, Leyden, The Netherlands, 1976.

[16] A. K. Louis, *Inverse und schlecht gestellte probleme*, Teubner Verlag, Stuttgart, 1989.

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

[18] A. A. Melkman and C. A. Micchelli, *Optimal estimation of linear operators in Hilbert spaces from inaccurate data*, SIAM J. Numer. Anal. **16** (1980), 87–105.

[19] V. A. Morozov, *Methods for solving incorrectly posed problems*, Springer Verlag, New York, Berlin, Heidelberg, 1984.

[20] R. Plato and G. Vainikko, *On the regularization of projection methods for solving ill-posed problems*, Numer. Math. **57** (1990), 63–79.

[21] T. Raus, *On the discrepancy principle for the solution of ill-posed problems with non-selfadjoint operator*, Acta et Comment. Univ. Tartuensis **715** (1985), 12–20.

[22] R. Rummel and O. L. Colombo, *Gravity field determination from satellite gradiometry*, Bull. Geod. **59** (1985), 233–246.

[23] T. I. Seidman, *'Optimal Filtering' for some ill-posed problems*, Wave Propagation and Inversion (W. Witzgibbon and M. Wheeler, eds.), SIAM, Philadelphia, 1992, pp. 108–123.

[24] U. Tautenhahn, *Optimal stable approximations for the sideways heat equation*, J. Inv. Ill-Posed Problems **5** (1997), 287–307.

[25] ______, *Optimality for linear ill-posed problems under general source conditions*, Numer. Funct. Anal. Optim. **19** (1998), 377–398.

[26] M. Taylor, *Partial differential equations*, vol. 2, Springer Verlag, New York, 1996.

[27] G. M. Vainikko, *On the optimality of regularization methods*, Inverse and Ill-Posed Problems (H. W. Engl and C. W. Groetsch, eds.), Academic Press, Boston, 1986, pp. 77–95.