ERRORS OF REGULARISATION UNDER RANGE INCLUSIONS USING VARIABLE HILBERT SCALES

MARKUS HEGLAND
Centre for Mathematics and its Applications
The Australian National University
Canberra ACT, 0200, Australia

BERND HOFMANN
Department of Mathematics
Chemnitz University of Technology
09107 Chemnitz, Germany.

Abstract. Based on the variable Hilbert scale interpolation inequality bounds for the error of regularisation methods are derived under range inclusions. In this context, new formulae for the modulus of continuity of the inverse of bounded operators with non-closed range are given. Even if one can show the equivalence of this approach to the version used previously in the literature, the new formulae and corresponding conditions are simpler than the former ones. Several examples from image processing and spectral enhancement illustrate how the new error bounds can be applied.

1. Introduction

Let $X$ and $Y$ be infinite dimensional separable Hilbert spaces with norms $\| \cdot \|$ and scalar products $(\cdot, \cdot)$. We study linear inverse problems in form of ill-posed operator equations

$$Af = g, \quad f \in X, \ g \in Y,$$

characterised by an injective bounded linear forward operator $A : X \to Y$ for which the range range$(A)$ is a non-closed subset of $Y$. Then equation (1) is unstable in the sense that the inverse operator $A^{-1} : \text{range}(A) \subseteq Y \to X$ is unbounded and hence the use of perturbed data $g^\delta$ instead of the exact right-hand side $g$ with

$$\|g - g^\delta\| \leq \delta$$

and noise level $\delta > 0$ may lead to arbitrarily large errors in the solution of (1) even if the noise level is extremely small. As a consequence of this ill-posedness phenomenon regularisation methods are required for the stable approximate solution of the inverse problem. Their basic idea consists in finding approximations to the exact solution $f$ in form of solutions $f_\alpha = f_\alpha(g^\delta)$ to stable auxiliary problems neighbouring (1). Those solutions are obtained by using the noisy data $g^\delta$. The degree of neighbourhood of the exploited auxiliary problems is controlled by a regularisation parameter $\alpha > 0$. In this context, small $\alpha$ express closeness to (1) in combination with a low level of stability, whereas larger $\alpha$ ensure better stability,
however combined with a low level of approximation. For the success of any regularisation method an appropriate trade-off between stability and approximation has to be aspired when choosing the regularisation parameter.

As already outlined and summarised in the monograph [9] by Engl, Hanke and Neubauer a successful way for doing regularisation for linear ill-posed problems in Hilbert spaces including convergence and convergence rates of constructed methods requires some knowledge on the impact of smoothness on the regularised solutions. Smoothness should be understood there in a very generalised sense as both solution smoothness and smoothing properties of the forward operator. In [9] such smoothness fitting is focused on Hölder type source conditions yielding Hölder type convergence rates when the regularisation method has a sufficiently high level of qualification. This theory is closely connected with associated (classical) Hilbert scales, where we refer to the seminal paper by Natterer [27]. An extension of that theory to generalised source conditions implying also more general convergence rates was performed rather independently by two different approaches. The first approach initiated and established by Hegland (see [11, 12]) introduced variable Hilbert scales with positive index functions the behaviour of which is in particular of interest for large arguments covering the spectrum of an injective unbounded linear operator with bounded inverse. Results in this approach are based on interpolation inequalities. An alternative second approach was developed and published by Mathé and Pereverzev (see [24, 25, 26]) and complimented by ideas of Hofmann and other co-workers (see, e.g., [15, 16, 17]). This approach, in principle, also exploits variable Hilbert scales, but the index functions occurring there are more specific and their behaviour is of interest just for small positive arguments covering on the spectrum of $A^*A$. The index functions of the second approach are monotonically increasing and tend to zero as the positive arguments tend to zero. Here we call them rate functions abbreviated by over-lined small Greek letters, because they also express the convergence rate of approximate solutions. One of the main goals of this paper is to compare both approaches, their results, required conditions and their natural interplay. Moreover, some more consequences and new convergence rates results of Hegland’s approach shall be formulated and proven in the sequel.

In our study we use variable Hilbert scales and corresponding interpolation inequalities in order to obtain bounds from above for the error $\|f - f_\alpha\|$ of regularised solutions under conditions imposed on $f$. As is well-known the convergence of regularised solutions $f_\alpha \to f$ in $X$, even if $\delta \to +0$ and $\alpha = \alpha(\delta, g^\delta) \to +0$ is chosen in an appropriate manner, can be arbitrarily slow for solutions $f$ which are in some sense non-smooth with respect to the forward operator $A$. It is of essential interest in regularisation theory to obtain convergence rates

$$\|f - f_\alpha\| = \mathcal{O}(\bar{\eta}(\delta)) \quad \text{as} \quad \delta \to +0$$

with appropriate rate functions $\bar{\eta}$.

For a non-closed range of $A$ convergence rates require that general source conditions are satisfied which attain in the standard case the form

$$f = \tilde{\psi}(A^*A)v$$

with source element $v \in X$ and with some rate function $\tilde{\psi}$ defined at least on the interval $(0, \|A\|^2]$ covering the spectrum of $A^*A$. Here, $\psi(A^*A)$ is well-defined as an injective bounded positive self-adjoint linear operator by spectral calculus (see, e.g., [9, Section 2.3]). If the regularisation method has a sufficiently high qualification (see, e.g., [15, Definition 2.6]), then an asymptotically fast decay of
\[ \tilde{\psi}(t) \to 0 \text{ as } t \to +0 \] corresponds with high order convergence rates (3) of regularised solutions and vice versa. Note that the tool of general source conditions with rate functions \( \psi \) different from monomials was early applied to regularisation theory by Tautenhahn in [33]. Later Hohage (see [19, 20]) studied in detail the case of logarithmic functions \( \tilde{\psi} \) in (4) and corresponding convergence rates.

Using a non-standard source condition
\[ (5) \quad f = Gw \]
with source element \( w \in X \), the priori information about the smoothness of the solution \( f \) can be expressed by an injective bounded self-adjoint positive definite linear operator \( G : X \to X \), where \( \text{range}(G) \) is a non-closed subset of \( X \). In most cases the character of the operator \( G \) is rather independent of the forward operator \( A \) and hence \( G \) need not be a function of \( A^*A \). However, in order to make conclusions to convergence rates of regularised solutions the assumption \( f \in \text{range}(G) \) of (5) and the assumption \( f \in \text{range}(\tilde{\psi}(A^*A)) \) of (4) have to be connected anyway. In the framework of Böttcher et al. [6] conditions for such connections and their interplay are discussed in a comprehensive manner. In [15] they are called link conditions. Besides the simplest link type \( G = \tilde{\varphi}(A^*A) \), which is typical for commuting \( G \) and \( A^*A \), one of the most convincing class of link conditions represent range inclusions introduced with Yamamoto in [18] to regularisation theory. The favourite form of such inclusion is
\[ (6) \quad \text{range}(G) \subseteq \text{range}(\tilde{\psi}(A^*A)) \]
with some rate function \( \tilde{\psi} \). Evidently, under (5) a range inclusion (6) immediately implies a source condition (4). The higher the rate expressed by \( \tilde{\psi} \) is, i.e. the faster the decay \( \tilde{\psi}(t) \to 0 \) as \( t \to +0 \) goes on, the smaller the set \( \text{range}(\tilde{\psi}(A^*A)) \) becomes. Hence the condition (6) is a strong one for higher rates \( \tilde{\psi} \) and vice versa.

An alternative link condition is
\[ (7) \quad \| \tilde{g}(G)x \| \leq C \| Ax \| \quad \text{for all } x \in X \]
for some constant \( C > 0 \), where the rate function \( \tilde{g} \) acts as a benchmark for the degree of ill-posedness of equation (1) with respect to the a priori information (5). From Proposition 2.1 in [6] we know that a range inclusion \( \text{range}(G_1) \subseteq \text{range}(G_2) \) and a condition of the form \( \| G_1x \| \leq C \| G_2x \| \) for all \( x \in X \) and some \( C > 0 \) are equivalent. Consequently, with Proposition 2.18 in [9] the condition (7) is equivalent to the range inclusion
\[ (8) \quad \text{range}(\tilde{g}(G)) \subseteq \text{range}(A^*) = \text{range}((A^*A)^{1/2}) \]
taking into account the identity \( \| Ax \| = \|(A^*A)^{1/2}x\| \) for all \( x \in X \).

As exploited in [6, § 4] one can reduce (8) to the form (6) with \( \tilde{\psi}(t) = \tilde{g}^{-1}(\sqrt{t}) \) if \( [\tilde{g}^{-1}(\sqrt{t})]^2 \) is an operator monotone function (cf. [5]). An important special case of that implication, namely for \( g(t) = t^{\mu/2} \) with \( 0 < \mu \leq 1/2 \), is well-known as Heinz-Kato inequality (see, e.g., [6, Proposition 8.21] or the corollary of Theorem 2.3.3 in [32, p. 45]). In that special case, (8) yields (6) with \( \tilde{\psi}(t) = t^\mu \) for exponents \( 0 < \mu \leq 1/2 \).

In the next section we review the definition and some properties of index functions and variable Hilbert scales. The fundamental interpolation inequality is given with a short proof together with an application to a general regularisation method. We then show how the variable Hilbert scales provide natural source conditions. In the third section bounds for the modulus of continuity are given in a variable Hilbert
scale setting. An important part of this section compares the new bounds on the modulus of continuity with some obtained earlier and shows how the new results have a substantially simpler structure. The fourth section analyses linear regularisation methods and parameter choices using the variable Hilbert scale approach. In section 5 we consider several examples from image processing and spectral enhancement and the paper finishes with some final remarks.

2. Interpolation inequalities and consequences

The main tool used here to derive error bounds for regularised solutions is an extension of interpolation inequalities to variable Hilbert scales. For classical Hilbert scales \( \{X_r\}_{r \in \mathbb{R}} \) with real numbers as scale index \( r \), interpolation inequalities are well-established. These interpolation inequalities were initially applied to the treatment of linear ill-posed problems (1) by Natterer in [27] (see also the monograph by Engl, Hanke and Neubauer [9]). For variable Hilbert scales, new interpolation inequalities have to be formulated. Here the scale index \( r \) is replaced by a wide class of index functions defined as:

**Definition 2.1.** We call a real function \( \theta \) defined on the open interval \((0, \infty)\) an *index function* if it is continuous and positive. The *index set* denoted by \( \mathcal{I} \) is then the set of all such index functions.

We call an index function \( \hat{\theta} \in \mathcal{I} \) a *rate function* if it is monotonically increasing and if it satisfies the limit condition \( \lim_{t \to +0} \hat{\theta}(t) = 0 \).

Note that any monotonically increasing continuous function \( \theta(t) \) defined on a finite interval \([0, t_0]\) satisfying \( \lim_{t \to +0} \theta(t) = 0 \) can be extended to a rate function \( \hat{\theta} \) such that \( \hat{\theta}(t) = \theta(t) \) for \( t \in (0, t_0] \). Furthermore, the index functions corresponding to the classical Hilbert scales \( X_r \) can be seen to be power functions \( \theta(\lambda) = \lambda^r \) for real \( r \). Rate functions are obtained for this case if \( r > 0 \).

The set of index functions \( \mathcal{I} \) includes the positive constant functions and all power functions but not the zero function. We denote the pointwise operations by \( \phi + \psi, \phi \psi \) and \( \phi / \psi \), respectively. As usual, multiplication by a constant \( \gamma \) is denoted by \( \gamma \phi \). The composition is denoted by \( \phi \circ \psi \) where \( (\phi \circ \psi)(\lambda) = \phi(\psi(\lambda)) \). The pointwise maximum of two index functions is \( \phi \lor \psi \) defined by \( (\phi \lor \psi)(\lambda) = \max(\phi(\lambda), \psi(\lambda)) \) and the pointwise minimum is \( \phi \land \psi \). One verifies that the index set \( \mathcal{I} \) from Definition 2.1 is closed under

- pointwise addition, multiplication and division,
- multiplication with positive constants,
- pointwise maximum and minimum and
- composition.

If an index function is injective and surjective, the inverse denoted by \( \phi^{-1} \) is also an index function. Not every index function is invertible, however. The reciprocal function of \( \phi \) (with values \( 1/\phi(\lambda) \)) is denoted by \( 1/\phi \).

The variable Hilbert scales are then families of Hilbert spaces indexed by \( \mathcal{I} \).

**Definition 2.2.** For a given injective self-adjoint positive definite linear operator \( T \) densely defined on a Hilbert space \( X \) we define the *variable Hilbert scale* \( \{X_\theta\}_{\theta \in \mathcal{I}} \) as a family of Hilbert spaces \( X_\theta \) indexed by functions \( \theta \) from the set \( \mathcal{I} \) of index functions in the sense of Definition 2.1. Every Hilbert space \( X_\theta \) with \( \theta \in \mathcal{I} \) is then
the closure of the domain of the quadratic form

\[ \|f\|_0^2 = (f, \theta(T) f). \tag{9} \]

In such a way, variable Hilbert scales were introduced by Hegland in [11] for the special case of \( T \) being the inverse of a compact operator and in [12] for more general \( T \). The choice of the operator \( T \) determines the Hilbert scale. In the simplest case where both \( T \) and its inverse \( T^{-1} \) are bounded all the Hilbert spaces \( X_\theta \) are isomorphic to \( X \) because of the continuity of the index functions \( \theta \). In this paper, in the context of ill-posed problems (1) we often assume that \( T \) is unbounded but has a bounded inverse, i.e. the spectrum of \( T \) is contained in the interval \([\|T^{-1}\|^{-1}, \infty)\) and has \(+\infty\) as an accumulation point. As the function \( 1/\lambda \) is an index function and the set of index functions is closed under composition, the inverse \( T^{-1} \) generates the same Hilbert scale as \( T \). It is thus not necessary to consider variable Hilbert scales generated by invertible \( T \) and bounded \( T^{-1} \) separately. The more general case where both \( T \) and the inverse \( T^{-1} \) are unbounded is only considered for the negative Laplacian \( T = -\Delta \) and in particular \( T = -d^2/dx^2 \). For the more general case where also \( A \) is unbounded we refer to the recent paper [17]. To get a link with (1), a particular \( T \) is suggested either by the forward operator \( A \), by the operator \( G \) of condition (6) or based on a combination of both. A common choice is \( T = (A^*A)^{-1} \) for injective operators \( A \) with a non-closed range. It follows that \( A^*A = \theta(T) \) if \( \theta(\lambda) = 1/\lambda \). For classes of problems connected with deconvolution, however, \( T = -d^2/dx^2 \) on \( L_2(\mathbb{R}) \) is the canonical choice as \( T \) is the generator of symmetric convolutions. An index function \( \theta \) such that \( A^*A = \theta(T) \) is then found using Fourier transforms. More generally, for problems where the source conditions relate to smoothness, \( T = -\Delta \) can be chosen. In such a case \( \Delta \) denotes the Laplacian on \( L_2(\Omega) \) for some domain \( \Omega \subseteq \mathbb{R}^d \).

It was shown in [12] that there exists a continuous embedding \( X_\psi \hookrightarrow X_\psi \) if and only if \( \phi \leq \gamma \psi \) for some constant \( \gamma > 0 \). If two different index functions \( \theta_1 \) and \( \theta_2 \) are identical on the spectrum of \( T \) they define the same norms and hence the same space \( X_{\theta_1} = X_{\theta_2} \). If they differ on the spectrum they do define different norms, however, these norms may be equivalent and thus the Hilbert spaces \( X_{\theta_1} \) and \( X_{\theta_1} \) as elements of the variable Hilbert scale \( \{X_\theta\}_{\theta \in \mathcal{I}} \) are indistinguishable.

The most important connection between the norms of different spaces \( X_\theta \) is the interpolation inequality for variable Hilbert scales.

**Lemma 2.3** (Interpolation inequality). Let \( T \) be an unbounded injective self-adjoint positive definite linear operator densely defined on the Hilbert space \( X \) with bounded inverse \( T^{-1} : X \rightarrow X \). Moreover let \( \phi, \psi, \theta \) and \( \Psi \) be index functions such that \( \Psi \) is concave and

\[ \phi(\lambda) \leq \Psi(\psi(\lambda)), \quad \text{for} \quad \|T^{-1}\|^{-1} \leq \lambda < \infty. \tag{10} \]

Then for any element \( 0 \neq f \in X_\theta \cap X_\psi \) one gets \( f \in X_{\phi} \) and

\[ \frac{\|f\|_{\phi}^2}{\|f\|_{\phi}^2} \leq \Psi \left( \frac{\|f\|_{\psi}^2}{\|f\|_{\psi}^2} \right). \tag{11} \]

**Proof.** Let in the following the measure \( \nu \) be defined by

\[ d\nu(\lambda) = \|f\|_{\phi}^2 \theta(\lambda) d(f, E(\lambda)f) \quad \text{for} \quad 0 \neq f \in X_\theta \]
where \( E(\lambda) \) is the spectral family or resolution of the identity defined by \( T \). By definition, because \( \theta \) is positive and the integral of \( d\nu \) equals 1, \( \nu \) is a probability measure. Taking into account that \( f \in X_{\psi \theta} \) and that \( \Psi \) is concave we obtain from the inverse Jensen inequality that 
\[
\int_{0}^{\infty} \Psi(\psi(\lambda)) \, d\nu(\lambda) \leq \Psi \left( \int_{0}^{\infty} \psi(\lambda) \, d\nu(\lambda) \right) < \infty.
\]
Because integration is monotone and the inequality (10) holds one gets 
\[
\int_{0}^{\infty} \psi(\lambda) \, d\nu(\lambda) < \infty
\]
and hence \( f \in X_{\psi \theta} \). Summarising the results we arrive at the inequality
\[
\int_{0}^{\infty} \phi(\lambda) \, d\nu(\lambda) \leq \Psi \left( \int_{0}^{\infty} \psi(\lambda) \, d\nu(\lambda) \right)
\]
which provides us with the required inequality (11) by replacing \( d\nu(\lambda) \) by its definition.

The concavity of \( \Psi \) is the key property which enables us to use Jensen’s inequality. The Lemma 2.5 below shows that this property has only to be established for large arguments. We can focus on large arguments, if the spectrum of \( T \) under consideration contains only sufficiently large values and has \( +\infty \) as an accumulation point. We need some auxiliary result:

**Lemma 2.4.** If \( \theta : [t_0, \infty) \to (0, \infty) \) is concave for some \( t_0 > 0 \) then \( \theta \) is monotonically increasing. If moreover \( \lim_{t \to \infty} \theta(t) = \infty \), then \( \theta \) is even strictly increasing.

**Proof.** We show the contraposition. Assume that \( \theta : [t_0, \infty) \to (0, \infty) \) is not monotonically increasing. Then there exist \( t_0 < t_1 < t_2 \) such that \( \theta(t_1) > \theta(t_2) \). Let
\[
l(t) = \frac{t - t_1}{t_2 - t_1} \theta(t_2) + \frac{t_2 - t}{t_2 - t_1} \theta(t_1)
\]
be the linear interpolant of \( \theta \) in \([t_1, t_2]\). As the slope of \( l(t) \) is \((\theta(t_2) - \theta(t_1))/(t_2 - t_1) < 0\) one has \( l(t) \to -\infty \) for \( t \to \infty \). As \( \theta(t) \geq 0 \) there exists a \( t_3 > t_2 \) such that \( \theta(t_3) > l(t_3) \). By rearranging this inequality one gets
\[
\theta(t_2) \leq \frac{t_2 - t_3}{t_3 - t_1} \theta(t_3) + \frac{t_3 - t_2}{t_3 - t_1} \theta(t_1)
\]
and so \( \theta \) is not concave. The strict monotonicity for \( \lim_{t \to \infty} \theta(t) = \infty \) follows immediately from the fact that the hypograph of a concave function is a convex set.

Now we can replace index functions which are concave for large arguments by such which are globally concave in the following way:

**Lemma 2.5.** Let \( \theta(\lambda) \) be an index function which is concave and hence by Lemma 2.4 increasing on the interval \( 0 < \lambda_0 \leq \lambda < \infty \). Then there exists an index function \( \Psi(\lambda) \) which is concave for all \( 0 < \lambda < \infty \) such that with some \( \lambda_1 > \lambda_0 \) one has
\[
\Psi(\lambda) = \theta(\lambda), \quad \text{for } \lambda_1 \leq \lambda < \infty,
\]
\[
\Psi(\lambda) = \lambda \Psi(\lambda_1)/\lambda_1, \quad \text{for } 0 < \lambda \leq \lambda_1.
\]

**Proof.** To obtain the assertion of this lemma we consider the set of real numbers
\[
\{ \alpha \mid \alpha \lambda \geq \theta(\lambda), \lambda_0 \leq \lambda < \infty \}.
\]
As \( \theta(\lambda) \) is concave for \( \lambda_0 \leq \lambda < \infty \) this set is not empty and it is bounded below by zero. Thus it does have a greatest lower bound \( \alpha_0 \geq 0 \) such that
- \( \alpha_0 \lambda \geq \theta(\lambda) \) for \( \lambda_0 \leq \lambda < \infty \),
- there is a \( \lambda_1 \) such that \( \alpha_0 \lambda_1 = \theta(\lambda_1) \) if not, \( \alpha_0 \) would not be the greatest lower bound.
Hence, knowing from Lemma 2.4 that an index function Ψ which is concave for all $0 < \lambda < \infty$ is always increasing, the function $\Psi(\lambda)$ can be composed of a linear function growing from zero to $\theta(\lambda_1)$ in the interval $(0, \lambda_1]$ and coinciding with $\theta$ for greater arguments.

The interpolation inequality is the main tool to obtain error bounds for solvers of linear ill-posed problems. However, taking into account Lemma 2.5 by inspection it becomes clear that rate results derived from Lemma 2.3 are only based on the behaviour of $\Psi(\lambda)$ for large $\lambda \geq \lambda_1$. Without loss of generality $\Psi$ can be amended for $0 < \lambda \leq \lambda_1$ by the linear function $\Psi(\lambda) = \Psi(\lambda_1) \lambda / \lambda_1$ for $0 < \lambda < \lambda_1$.

Three typical choices for $\Psi(\lambda)$ being concave at least for sufficiently large $\lambda$ are

- $\Psi(\lambda) = \lambda^\kappa$ where $\kappa \in (0, 1)$
- $\Psi(\lambda) = \lambda / \log(\lambda)$
- $\Psi(\lambda) = \log(\lambda)$

For all three choices we have the limit condition

$$\lim_{\lambda \to \infty} \frac{\Psi(\lambda)}{\lambda} = 0$$

and one gets the following versions of interpolation inequalities from Lemma 2.3:

- For $\Psi(\lambda) = \lambda^\kappa$ one gets
  $$\|f\|_{\phi_\theta} \leq \|f\|_{\theta}^{1-\kappa} \|f\|_{\psi_\theta}^\kappa,$$
- for $\Psi(\lambda) = \lambda / \log(\lambda)$ one gets
  $$\|f\|_{\phi_\theta} \leq \frac{\|f\|_{\psi_\theta}}{\sqrt{2 \log(\|f\|_{\psi_\theta} / \|f\|_{\theta})}},$$
- and for $\Psi(\lambda) = \log(\lambda)$ one has
  $$\|f\|_{\phi_\theta} \leq \|f\|_{\theta} \sqrt{2 \log(\|f\|_{\psi_\theta} / \|f\|_{\theta})}.$$

Asymptotically, i.e. for $\|f\|_{\theta} \to 0$, the interpolation inequality allows us to find error bounds in the application to the error estimation for the solution of equation (1). One aims to get bounds for the norm $\|f\|$ in $X$ using values of the image norm $\|Af\|$ in $Y$ and values of the norm $\|f\|_{\psi_\theta}$ which expresses the specific additional smoothness of $f$. The terms in the interpolation inequality (11) are then

$$\|f\|_{\phi_\theta} = \|f\|_{\phi_\theta}$$

for $f \in X_{\phi_\theta}$ and $\|f\|_{\theta} = \|Af\|$ for $f \in X_\theta$.

The first condition leads to $\phi(\lambda) \theta(\lambda) = 1$ for all $\lambda$ and the second condition gives $\theta(T) = A^* A$ and with $\theta(\lambda) := 1 / \lambda$ the relations $T = (A^* A)^{-1}$ and $\phi(\lambda) = \lambda$. We are still free to choose the index functions $\psi$ and do it in the form $\psi(\lambda) := \chi(\lambda) \lambda$ with an appropriate index function $\chi$.

For later use we add here some observations about convex functions which are stated as a lemma:

**Lemma 2.6.** Let $\Psi : (0, \infty) \to (0, \infty)$ be a concave function. Then we have the following properties:

(a) The function $\Xi : (0, \infty) \to (0, \infty)$ defined by

$$\Xi(\lambda) := \frac{\Psi(\lambda)}{\lambda}, \quad 0 < \lambda < \infty$$

is monotonically decreasing.
The function \( \Phi : (0, \infty) \to (0, \infty) \) defined by

\[
\Phi(\mu) := \mu \Psi \left( \frac{1}{\mu} \right), \quad 0 < \mu < \infty
\]

is concave and hence monotonically increasing.

Proof. (a) Let \( 0 < \lambda_0 < \lambda_1 < \lambda_2 \). As \( \Psi \) is concave and positive one has

\[
\Psi(\lambda_1) \geq \frac{\lambda_1 - \lambda_0}{\lambda_2 - \lambda_0} \Psi(\lambda_2) + \frac{\lambda_2 - \lambda_1}{\lambda_2 - \lambda_0} \Psi(\lambda_0)
\]

\[
\geq \frac{\lambda_1 - \lambda_0}{\lambda_2 - \lambda_0} \Psi(\lambda_2).
\]

As this holds for arbitrarily small \( \lambda_0 > 0 \) on has

\[
\Psi(\lambda_1) \geq \frac{\lambda_1}{\lambda_2} \Psi(\lambda_2)
\]

and consequently \( \Xi(\lambda_1) \geq \Xi(\lambda_2) \). This proves assertion (a) of the lemma.

(b) Let \( 0 < \mu_0 < \mu_1 < \mu_2 \) and \( \lambda_i = 1/\mu_i \). Then one has \( 0 < \lambda_2 < \lambda_1 < \lambda_0 \) and by the concavity of \( \Psi \) and some simple algebraic manipulations one gets

\[
\frac{\mu_1 - \mu_0}{\mu_2 - \mu_0} \Phi(\mu_2) + \frac{\mu_2 - \mu_1}{\mu_2 - \mu_0} \Phi(\mu_0) = \frac{\lambda_1 - \lambda_0}{\lambda_2 - \lambda_0} \Psi(\lambda_2) + \frac{\lambda_2 - \lambda_1}{\lambda_2 - \lambda_0} \Psi(\lambda_0)
\]

\[
= \frac{1}{\lambda_1} \left( \frac{\lambda_0 - \lambda_1}{\lambda_0 - \lambda_2} \Psi(\lambda_2) + \frac{\lambda_1 - \lambda_2}{\lambda_0 - \lambda_2} \Psi(\lambda_0) \right)
\]

\[
\leq \frac{1}{\lambda_1} \Psi(\lambda_1) = \Phi(\mu_1).
\]

It follows that \( \Phi \) is concave and hence by Lemma 2.4 also increasing. This completes the proof of the lemma.

Remark 1. We note here that the transformation \( S : \Psi \in \mathcal{I} \mapsto \Phi \in \mathcal{I} \) according to formula (14), applicable to every index function and preserving concavity, is an involution, that means \( S^{-1} = S \) and hence \( S \) is bijective. If the concave index function \( \Psi \) satisfies \( \lim_{\lambda \to \infty} \Psi(\lambda) = \infty \), then by Lemma 2.4 the function is even strictly increasing and if, in addition, \( \Psi \) is a rate function, i.e., it satisfies the additional limit condition \( \lim_{\lambda \to 0^+} \Psi(\lambda) = 0 \) (which is also motivated by Lemma 2.5), the inverse function \( \Psi^{-1} \) is a well-defined and convex index function. If, on the other hand, the limit condition (12) holds, then we have

\[
\lim_{\mu \to +0} \Phi(\mu) = \lim_{\mu \to +0} \mu \Psi(1/\mu) = \lim_{\lambda \to \infty} \Psi(\lambda)/\lambda = 0
\]

and taking into account Lemma 2.6 (a) and (b) one sees that \( \Phi = S(\Psi) \) is a concave rate function. Vice versa we have that \( \Psi = S(\Phi) \) satisfies (12) whenever \( \Phi \) is a rate function.

By inspection of the proof of Lemma 2.6 one can also see the following facts: If \( \Psi(\lambda) \) is only concave for \( \lambda \in [\lambda_0, \infty) \), then \( \Phi(\mu) = [S(\Psi)](\mu) \) is concave for \( \mu \in (0, \mu_0) \) with \( \mu_0 = 1/\lambda_0 \). The involution \( S \) preserves also the convexity of an index function and if the concavity or convexity is strict, then the strictness carries over to the transformed function.

Now we are ready to draw conclusions from Lemma 2.3. A first, abstract version of bounds for errors of regularised solutions is given in the following corollary.

We will denote by \( f_\alpha \) an approximation of a solution \( f \) to equation (1) which is
computed from an approximate right-hand side using a regularisation method and a regularisation parameter \(\alpha > 0\).

**Corollary 2.7.** Let \(A : X \to Y\) be an injective bounded linear operator with non-closed range mapping between the two Hilbert spaces \(X\) and \(Y\). Furthermore let the variable Hilbert scale \(\{X_\nu\}_{\nu \in I}\) be generated by \(T = (A^\top A)^{-1}\) such that any scale element \(X_\nu\) has a norm denoted by \(\| \cdot \|_\nu\). Moreover let \(\chi\) and \(\Psi\) be index functions and \(\Psi\) be concave such that

\[
\Psi (\chi (\lambda) \lambda) \geq \lambda \quad \text{for all} \quad \| T^{-1} \|^{-1} \leq \lambda < \infty.
\]

If the solution \(f\) to \((1)\) in addition satisfies the condition \(f \in X_\chi\) and if \(f_\alpha \in X_\chi\) is such that

\[
\| f_\alpha - f \|_\chi = \zeta > 0
\]

\[
\| Af_\alpha - g \| = \epsilon > 0
\]

then

\[
\| f - f_\alpha \| \leq \epsilon \sqrt{\Psi(\zeta^2/\epsilon^2)}.
\]

**Proof.** By Lemma 2.3 with \(\theta(\lambda) = 1/\lambda, \phi(\lambda) = \lambda, \psi(\lambda) = \chi(\lambda) \lambda\) and for a concave index function \(\Psi\) satisfying \((15)\) one has for all \(0 \neq h \in X_\chi\)

\[
\| h \|_\chi^2 \leq \| Ah \|_\chi^2 \Psi \left( \frac{\| h \|_\chi^2}{\| Ah \|_\chi^2} \right).
\]

Setting \(h := f_\alpha - f\) this yields the estimate

\[
\| f_\alpha - f \|_\chi^2 \leq \| Af_\alpha - f \|_\chi^2 \Psi \left( \frac{\| f_\alpha - f \|_\chi^2}{\| Af_\alpha - Af \|_\chi^2} \right) = \epsilon^2 \Psi((\zeta/\epsilon)^2)
\]

and proves the assertion of the corollary.

Results similar to those of Corollary 2.7 can be found for other choices of \(T\), see for example Corollary 5.3 where \(T = -d^2/dt^2\). The Corollary 2.7 can be interpreted as an instance of an abstract Lax theorem [1] where the condition \((16)\) is a stability condition and the bound \((17)\) relates to consistency.

Note that the error estimate \((18)\) of Corollary 2.7 requires the essential conditions \(f \in X_\chi\) and \(f_\alpha \in X_\chi\), i.e. the approximate solutions \(f_\alpha\) are constructed such that they obtain the same smoothness level with respect to \(T\) as the exact solution \(f\). A next step for drawing conclusions of Lemma 2.3 will be formulated in Corollary 2.8 by assuming that \(f\) belongs to ball

\[
B_\chi (R) := \{ h \in X_\chi : \| h \|_\chi \leq R \}
\]

in \(X_\chi\) with positive radius \(R = R_1\) and that the approximate solutions \(f_\alpha\) for all \(\alpha > 0\) under consideration belong to another such ball with radius \(R = R_2\). Moreover, we consider for data \(g^\delta\) satisfying \((2)\) the limit process \(\delta \to +0\) in correspondence with associated regularized solutions \(f_\alpha\), where the regularisation parameter \(\alpha > 0\) is chosen either a priori as \(\alpha = \alpha(\delta)\) or a posteriori as \(\alpha = \alpha(\delta, g^\delta)\).

**Corollary 2.8.** Under the setting of Corollary 2.7 let the limit condition \((12)\) be satisfied and let \(f \in B_\chi (R_1), R_1 > 0\). Moreover with prescribed \(\delta_{\text{max}} > 0\) let \(f_\alpha \in B_\chi (R_2), R_2 > 0\), for all \(\alpha\) attributed to \(\delta \in (0, \delta_{\text{max}}]\) and \(g^\delta\) satisfying \((2)\) such that

\[
\| Af_\alpha - g \| \leq \tilde{C} \xi(\delta), \quad 0 < \delta \leq \delta_{\text{max}},
\]
for some rate function $\xi$ and some constant $\bar{C} > 0$. Then we have
\begin{equation}
\|f - f_\alpha\| \leq \bar{C}\xi(\delta) \sqrt{\frac{R_1 + R_2}{C\xi(\delta)^2}}, \quad 0 < \delta \leq \delta_{\text{max}},
\end{equation}
where the upper bound in (22) is a rate function, i.e., it tends to zero as $\delta \to 0$.

**Proof.** Since $\Psi$ is concave by Lemma 2.4 the error norm $\epsilon \sqrt{\Psi(\zeta^2/\epsilon^2)}$ obtained from (18) is increasing in $\zeta > 0$ for fixed $\epsilon > 0$ and as a consequence of Lemma 2.6 (a) this upper bound $\epsilon \sqrt{\Psi(\zeta^2/\epsilon^2)} = \zeta \sqrt{\frac{\Psi(\zeta^2/\epsilon^2)}{(\zeta/\epsilon)^2}}$ is increasing in $\epsilon > 0$ for fixed $\zeta > 0$. Moreover, due to (12) in the limit process $\epsilon \to 0$ for fixed $\zeta > 0$ implying $\zeta/\epsilon \to \infty$ this bound and hence the error norm in $X$ even tends to zero. For the mentioned kinds of monotonicity we obtain formula (22) by $\epsilon \leq \bar{C}\xi(\delta)$ and $\zeta = \|f - f_\alpha\|_X \leq \|f\|_X + \|f_\alpha\|_X \leq R_1 + R_2$. The upper bound in (22) is a rate function declining to zero as $\delta \to 0$ because $\xi$ is a rate function. \hfill \Box

**Remark 2.** As a special case for the situation of Corollary 2.8 we can consider an a posteriori choice $\alpha_{\text{dis}} = \alpha_{\text{dis}}(\delta, g^\delta)$ for the regularisation parameter realised by a discrepancy principle
\begin{equation}
\|Af_{\alpha_{\text{dis}}} - g^\delta\| = C_{\text{dis}}\delta
\end{equation}
with some prescribed $C_{\text{dis}} > 0$. Then by using the triangle inequality we obtain with (2) as noise model
\begin{equation}
\|Af_{\alpha_{\text{dis}}} - g\| \leq \|Af_{\alpha_{\text{dis}}} - g^\delta\| + \|g^\delta - g\| \leq (C_{\text{dis}} + 1)\delta = \bar{C}\delta.
\end{equation}
Then for such $\alpha = \alpha_{\text{dis}}$ under (12) the regularisation method converges strongly in $X$ with the convergence rate
\begin{equation}
\|f - f_\alpha\| = \mathcal{O}\left(\delta \sqrt{\Psi(K/\delta^2)}\right) \quad \text{as} \quad \delta \to 0
\end{equation}
for some constant $K > 0$. Note that beside the assumption $f \in B_X(R_1)$ on the solution smoothness for that result the strong condition $f_{\alpha(\delta, g^\delta)} \in B_X(R_2)$ for all $\delta \in (0, \delta_{\text{max}}]$ and all associated $g^\delta$ satisfying (2) is required.

The convergence rate in (24) depends only on the asymptotic behaviour of $\Psi(\lambda)$ as $\lambda \to \infty$. Thus the alteration of $\Psi(\lambda)$ for small $\lambda$ in the sense of Lemma 2.5 has no influence on that rate. For the class of functions $\Psi(\lambda) = \lambda^\kappa$ with $0 < \kappa < 1$ rate functions proportional to $\delta^{1-\kappa}$ occur in (24). All those error rates are lower than the rate $\delta$ which is typical for well-posed problems. It should be mentioned that $\Psi(\lambda) = \lambda$ fails to satisfy the condition (12) and used in Corollary 2.7 the inequality (19) does not yield a convergence rate.

To get a feeling for the role of the solution smoothness $f \in X_\chi$ we can study consequences of the inequality (15) as a hypothesis of Corollary 2.7 taking into account Lemma 2.4. One consequence of (15) is the limit condition $\lim_{\lambda \to \infty} \Psi(\lambda) = \infty$ for the function $\Psi$ which is because of its concavity then strictly increasing and invertible with convex $\Psi^{-1}(\lambda)$ also tending to infinity as $\lambda \to \infty$. Then (15) implies $\chi(\lambda) \geq \frac{\Psi^{-1}(\lambda)}{\lambda}$ for large $\lambda$. Under that condition (12) is equivalent to $\lim_{\lambda \to \infty} \Psi^{-1}(\lambda)/\lambda = \infty$. Hence, the index function $\chi(\lambda)$ tends to infinity for $\lambda \to \infty$ provided that (12) holds true.

When setting $\phi(\lambda) := \lambda$, $\theta(\lambda) := 1/\lambda$ and $\psi(\lambda) := \Psi^{-1}(\lambda)$ in the interpolation inequality (11) then the corresponding regularity condition $f \in X_{\psi\theta}$ is equivalent
to a source condition (4) which expresses the specific smoothness of the solution $f$ with respect to the forward operator $A$ of equation (1).

**Proposition 2.9.** Let $\Psi(\lambda)$, for $0 < \lambda < \infty$, be a concave and strictly increasing index function satisfying the limit conditions $\lim_{\lambda \to +0} \Psi(\lambda) = 0$, $\lim_{\lambda \to \infty} \Psi(\lambda) = \infty$ and (12). Moreover let $T = (A^*A)^{-1}$ and set $\phi(\lambda) := \lambda$, $\theta(\lambda) := 1/\lambda$ as well as $\psi(\lambda) := \Psi^{-1}(\lambda)$ for $0 < \lambda < \infty$. Then we have $f \in X_{\psi\theta}$ if and only if $f$ satisfies a source condition (4) with the function

$$\tilde{\psi}(t) = \frac{1}{\sqrt{t\Psi^{-1}(1/t)}}, \quad 0 < t < \infty,$$

which is then a rate function.

**Proof.** Under the stated assumptions the function $\tilde{\psi}$ is well-defined and a rate function. Namely, we can write $\frac{1}{\sqrt{t\Psi^{-1}(1/t)}} = \sqrt{\frac{\Psi(u)}{u}}$ when using the substitution $u := \Psi^{-1}(1/t)$. The variable $u > 0$ is strictly decreasing with respect to $t > 0$ such that $u \to \infty$ corresponds with $t \to +0$ and vice versa $t \to \infty$ corresponds with $u \to +0$, because $\Psi^{-1}$ is also strictly increasing and we have $\lim_{\lambda \to \infty} \Psi^{-1}(\lambda) = \infty$ and $\lim_{\lambda \to +0} \Psi^{-1}(\lambda) = 0$ for the functions $\Psi$ under consideration. Now by (12) we have $\lim_{u \to \infty} \frac{\Psi(u)}{u} = 0$ and with Lemma 2.6 (a) the quotient $\frac{\Psi(u)}{u}$ is monotonically decreasing in $u > 0$. This, however, implies that $\psi(t)$ is monotonically increasing for $t > 0$ with limit condition $\lim_{t \to +0} \psi(t) = 0$. Hence, $\psi$ is a rate function.

Moreover, we have

$$f \in X_{\psi\theta} \iff (f, \Psi^{-1}(T)T^{-1}f) < \infty$$

and

$$f = \tilde{\psi}(A^*A)v, \text{ for } v \in X \iff (\tilde{\psi}(A^*A))^{-1}f, (\tilde{\psi}(A^*A))^{-1}f) < \infty.$$

One has equivalence if and only if

$$\Psi^{-1}((A^*A)^{-1})(A^*A) = (\tilde{\psi}(A^*A))^{-2}$$

and the claim follows. This proves the proposition. \qed

After the millennium Mathé and Pereverzev with coauthors seized, reused and extended Hegland’s ideas and concepts of variable Hilbert scales and corresponding interpolation inequalities from [11, 12] for linear ill-posed problems and their regularisation (cf. [24, 25]) and combined it (cf. [15, 26]) with the concept of approximate source conditions (cf. [8, 14]). The comprehensive theory developed therein considers only rate functions as index functions. Such an approach leads in general to different formulae compared with the results based on the concept of Lemma 2.3, but as we will outline in the sequel clear cross-connections and sometimes even equivalences of the assertions obtained characterise the two different ways.
3. Modulus of Continuity of $A^{-1}$

The *modulus of continuity* of $A^{-1}$ restricted to the set $AM$ with $M \subseteq X$ is

$$\omega(M, \delta) = \sup \{ \|x\| : x \in M, \|Ax\| \leq \delta \}.$$

The impact of the modulus of continuity on error bounds in regularisation has recently been discussed in the paper [16, §4]. It is well-known that the worst case error

$$e(\hat{f}, M, \delta) := \sup_{f \in M} e(\hat{f}, f, \delta)$$

for

$$e(\hat{f}, f, \delta) := \sup_{g^\delta \in Y : \|Af - g^\delta\| \leq \delta} \|\hat{f}(g^\delta) - f\|$$

of linear and nonlinear reconstruction methods $\hat{f} : g^\delta \in Y \mapsto \hat{f}(g^\delta) \in X$ has an infimum

$$e(M, \delta) := \inf_{\hat{f} : Y \to X} e(\hat{f}, M, \delta)$$

which satisfies for centrally symmetric and convex sets $M$ the inequalities

$$\omega(M, \delta) \leq e(M, \delta) \leq \omega(M, 2\delta) \leq 2\omega(M, \delta).$$

Hence the modulus of continuity $\omega(M, \delta)$ serves as benchmark for the reconstruction error of $\hat{f}$ when $f, \hat{f} \in M$ can be assumed. For example, from [16, Lemma 4.2] one can find a minimax-expression for the modulus of continuity in the case of centrally symmetric and convex source sets $M$ of the form

$$M = G[\bar{B}(R)] := \{ x \in X : x = Gv, v \in X, \|v\| \leq R \}.$$

corresponding to condition (5). This expression gets an explicit bound from above for the special case $G = \bar{\psi}(A^*A)$ and

$$M = \bar{\psi}(A^*A)[B(R)] := \{ x \in X : x = \bar{\psi}(A^*A)v, v \in X, \|v\| \leq R \}$$

associated with the source condition (4). Note that the rate function $\bar{\psi}(t)$ is only of interest here for arguments $0 < t \leq \|A\|^2$, but without loss of generality (cf. [17, Theorem 1 (b)]) we can extend $\bar{\psi}$ to be a monotonically increasing index function defined on $(0, \infty)$. Then by using the strictly increasing auxiliary function

$$\Theta(t) := \sqrt{t} \bar{\psi}(t), \quad 0 < t < \infty$$

satisfying the limits conditions $\lim_{t \to +0} \Theta(t) = 0$ and $\lim_{t \to \infty} \Theta(t) = \infty$ one obtains for

$$M = \bar{\psi}(A^*A)[B(R)]$$

$$\omega(M, \delta) \leq R \bar{\psi} \left( \Theta^{-1} \left( \frac{\delta}{R} \right) \right), \quad \delta > 0$$

provided that

$$\bar{\psi}^2((\Theta^2)^{-1}(t))$$

is concave for $0 < t < \infty$.

This result can be derived from Corollary 3.7 and Theorem 2.1(c) in [16] (see also Theorem 1 in the earlier paper [25]). A similar assertion was already mentioned in a rudimentary form in a paper by Ivanov and Korolyuk in 1969 [21].

The following proposition also yields an upper bound for the modulus of continuity based on a variable Hilbert scale interpolation inequality using Lemma 2.3 or Corollary 2.7. For the proof we use Lemma 2.6 (b).
Proposition 3.1. Let \( \Psi(\lambda) \), for \( 0 < \lambda < \infty \), be a concave and strictly increasing index function satisfying the limit conditions \( \lim_{\lambda \to 0^+} \Psi(\lambda) = 0 \), \( \lim_{\lambda \to \infty} \Psi(\lambda) = \infty \) and (12), for which an index function \( \chi \) exists that satisfies
\[
\chi(\lambda) \geq \Psi^{-1}(\lambda)/\lambda, \quad 0 < \lambda < \infty.
\]
Furthermore let \( X_\chi \) be an element of a Hilbert scale generated by \( T = (A^*A)^{-1} \) where \( A \) is injective. Then
\[
(30) \quad \omega(M, \delta) \leq \delta \left\{ \sqrt{\Psi \left( \frac{M^2}{\delta^2} \right)} \right\}, \quad \delta > 0
\]
for \( M = B_\chi(R) \).

Proof. Under the assumptions stated on \( \Psi \) and \( \chi \) Corollary 2.7 applies. Then from formula (19) we can conclude that \( \|h\| \leq \|Ah\| \sqrt{\Psi \left( \frac{M^2}{\delta^2} \right)} \) for all \( h \in X_\chi \).

As \( \Psi \) is monotonically increasing one then gets \( \|h\| \leq \|Ah\| \sqrt{\Psi \left( \frac{M^2}{\delta^2} \right)} \). Now by Lemma 2.6 (a) the function \( \Xi(\zeta) = \Psi(\zeta)/\zeta \) is monotonically increasing and so \( \Xi(\zeta_1) \geq \Xi(\zeta_2) \) for \( 0 < \zeta_1 \leq \zeta_2 < \infty \). This gives with \( \zeta_1 := \frac{R^2}{\delta^2} \) and \( \zeta_2 := \frac{R^2}{\|Ah\|^2} \) the estimate \( \|h\| \leq \delta \sqrt{\Psi \left( \frac{R^2}{\delta^2} \right)} \) for all \( h \in B_\chi(R) \) satisfying the additional condition \( \|Ah\| \leq \delta \). Thus the proposition is proven. \( \square \)

We note that for centrally symmetric and convex sets \( M, f \in M \), and regularised solutions \( f_{\text{dis}} \in M \) obtained from the discrepancy principle of form (23) mentioned in Remark 2 we easily derive along the lines of [16, Lemma 2.2] that
\[
(31) \quad \|f - f_{\text{dis}}\| \leq \omega(2M, (C_{\text{dis}} + 1)\delta)
\]
with \( 2M := \{u \in X : u = 2v, v \in M\} \). In the case \( M = B_\chi(R) \) with \( 2M = B_\chi(2R) \) the estimate (31) yields with (30) a convergence rate of the form (24) with constant \( K = 4R^2/(C_{\text{dis}} + 1)^2 \). With more generality such rates were verified above directly from Corollary 2.7.

Under weak additional assumptions (see [16, Corollary 3.7]) there is also a constant \( C > 0 \) such that
\[
\omega(B_\chi(R), \delta) \geq C\delta \sqrt{\Psi \left( \frac{R^2}{\delta^2} \right)}, \quad \delta > 0.
\]

Then a convergence rate of the form (24) is order optimal independent of the constant \( K > 0 \) because of \( \sqrt{\Psi \left( \frac{C_{\text{dis}}R^2}{\delta^2} \right)} \leq \max\{C, 1\} \sqrt{\Psi \left( \frac{R^2}{\delta^2} \right)} \) for all \( C > 0 \). On the other hand, Corollary 2.8 yields an error estimate of best order just for \( \xi(\delta) \sim \delta \), hence the discrepancy principle is order optimal in that sense.

Evidently, under the assumptions of Proposition 2.9 with the additional setting
\[
(32) \quad \chi(\lambda) := \frac{\Psi^{-1}(\lambda)}{\lambda} = \frac{1}{\psi(1/\lambda)^2}, \quad 0 < \lambda < \infty
\]
one has
\[
\bar{\psi}(A^*A)[B(R)] = B_\chi(R)
\]
where \( B_\chi(R) \) denotes the ball (20) of radius \( R \) in \( X_\chi \), an element of the Hilbert scale generated by \( T = (A^*A)^{-1} \) expressed through the index function \( \chi \). We emphasise
that the upper bound in (30) for the modulus of continuity from Proposition 3.1 needing only one function $\Psi$ has a much simpler structure than the nested upper bound in (28) composing the functions $\psi$ and $\Theta^{-1}$. Also the required concavity of $\Psi$ for obtaining (30) looks much simpler than the needed concavity of the composite function

$$\bar{\psi}^2((\Theta^2)^{-1}(t)) = \bar{\psi}^2(\Theta^{-1}(\sqrt{t})), \quad 0 < t < \infty,$$

for obtaining (28).

Owing to the correspondence (25) between the concave index function $\Psi$ and the rate function $\psi$ it is of some interest to compare the quality of the estimates (28) and (30) as well as the strength of conditions which have to imposed in order to ensure those bounds coincide, i.e., we have

$$\lim_{\lambda \to 0^+} \Psi(\lambda) = 0, \quad \lim_{\lambda \to \infty} \Psi(\lambda) = \infty$$

and (12). Then for the rate function $\bar{\psi}(t) := 1/\sqrt{R\Psi^{-1}(1/t)}$ (cf. (25)) and by setting $\Theta(t) := \sqrt{R}\bar{\psi}(t)$, $0 < t < \infty$, we have the following assertions: The error bounds in (30) and in (28) and the corresponding concavity conditions required for obtaining those bounds coincide, i.e., we have

$$\delta \sqrt{\Psi \left( \frac{R^2}{\delta^2} \right)} = R\bar{\psi} \left( \Theta^{-1} \left( \frac{\delta}{R} \right) \right), \quad R > 0, \quad \delta > 0.$$

Moreover, the function $\bar{\psi}^2((\Theta^2)^{-1}(t))$ is concave for all $0 < t < \infty$.

Vice versa, any rate function $\bar{\psi}(t)$, $0 < t < \infty$, determines by equation (25) in a unique manner a strictly increasing index function $\Psi(\lambda)$, $0 < \lambda < \infty$, satisfying the limit conditions $\lim_{\lambda \to 0^+} \Psi(\lambda) = 0, \lim_{\lambda \to \infty} \Psi(\lambda) = \infty$ and (12) which is concave for all $0 < \lambda < \infty$ if $\bar{\psi}^2((\Theta^2)^{-1}(t))$ is concave for all $0 < t < \infty$ which again implies the coincidence (33) of the error bounds.

**Proof.** First we find from Proposition 2.9 that $\bar{\psi}(t)$, $t > 0$, is a rate function if $\Psi(\lambda)$, $0 < \lambda < \infty$ is a concave and strictly increasing index function satisfying the limit conditions $\lim_{\lambda \to 0^+} \Psi(\lambda) = 0, \lim_{\lambda \to \infty} \Psi(\lambda) = \infty$ and (12). Then from the right equation in (32) (cf. (25)) we have $\Psi^{-1}(\lambda) = \lambda/\bar{\psi}^2(1/\lambda)$. By using the bijective substitution $u = 1/\lambda$ in $(0, \infty)$ this yields $\Psi^{-1}(1/u) = 1/\bar{\psi}^2(u)$ and $1/\bar{\psi}^2(u) = \Psi \left( \frac{1}{\bar{\psi}^2(u)} \right)$ for $0 < u < \infty$. Multiplying the last equation by the factor $u \bar{\psi}^2(u)$ we derive

$$\bar{\psi}^2(u) = u \bar{\psi}^2(u) \Psi \left( \frac{1}{\Theta^2(u)} \right) = \Theta^2(u) \Psi \left( \frac{1}{\Theta^2(u)} \right)$$

and $\bar{\psi}(u) = \Theta(u) \sqrt{\Psi \left( \frac{1}{\Theta^2(u)} \right)}$. By exploiting the bijection $t = \Theta(u)$ of $(0, \infty)$ into itself this provides us with the equation $\bar{\psi} \left( \Theta^{-1}(t) \right) = t \sqrt{\Psi \left( \frac{1}{t} \right)}$ which implies the required identity (33) by inserting $t := \delta/R$ and multiplying the arising equation by $R$.

In a second step we note that by using the monotonically increasing bijection $s = \Theta^2(u)$ between $s \in (0, \infty)$ and $u \in (0, \infty)$ and once more by exploiting the right equation in (32) we can write as follows for all $s > 0$:

$$\bar{\psi}^2((\Theta^2)^{-1}(s)) = \bar{\psi}^2(u) = \frac{\Theta^2(u)}{u} = \Theta^2(u) \Psi \left( \frac{1}{\Theta^2(u)} \right) = s \Psi \left( \frac{1}{s} \right) = [S(\Psi)](s).$$
Hence, by Lemma 2.6 (b) we immediately see that as required \( \tilde{\psi}^2((\Theta^2)^{-1}(s)), s > 0 \), is concave if \( \Psi(\lambda), \lambda > 0 \), is concave.

Since the involution \( S \) (cf. Remark 1) preserves concavity, the reverse assertion formulated in Proposition 3.2 becomes immediately clear, since (25) represents a one-to-one correspondence between index functions \( \tilde{\psi} \) and strictly increasing functions \( \Psi \) with the limit conditions under consideration. \( \square \)

We now investigate the concavity condition for the function \( \tilde{\psi}^2((\Theta^2)^{-1}(s)) \) in more detail. For this a characterisation of the concavity of index functions is given in terms of the monotonicity of certain divided differences.

**Lemma 3.3.** Let \( \psi \) be an index function. Then the three following statements are equivalent:

1. \( \psi \) is concave
2. \( (\psi(s_0 + s) - \psi(s_0))/s \) is a decreasing index function for all \( s_0 > 0 \)
3. \( (\psi(s_0) - \psi(s_0 - s))/s \) is an increasing continuous function \( (0, s_0) \to \mathbb{R}_+ \) for all \( s_0 > 0 \).

**Proof.** If \( \psi \) is a concave index function then by Lemma 2.4 \( \psi \) is increasing and so both \( (\psi(s_0 + s) - \psi(s_0))/s \) and \( (\psi(s_0) - \psi(s_0 - s))/s \) are positive continuous functions for \( s > 0 \) and \( s \in (0, s_0) \), respectively. Furthermore by definition

\[
(t_2 - t_0)\psi(t_1) \geq (t_2 - t_1)\psi(t_0) + (t_1 - t_0)\psi(t_2)
\]

and by simple algebraic manipulations and the right choice of \( t_0 < t_1 < t_2 \) one gets the second and third statement from the first.

Conversely, if \( (\psi(s_0 + s) - \psi(s_0))/s \) is a decreasing index function for all \( s_0 > 0 \) one has for all \( t_0 < t_1 < t_2 \)

\[
\frac{\psi(t_1) - \psi(t_0)}{t_1 - t_0} \geq \frac{\psi(t_2) - \psi(t_0)}{t_2 - t_0}
\]

and thus \( \psi \) is concave. A similar argument shows that \( \psi \) is concave if the third statement holds. \( \square \)

A direct consequence of this lemma is that for concave rate functions \( \tilde{\psi} \) one has

\[
\frac{\tilde{\psi}(s_0) - \tilde{\psi}(s_0 - s)}{s} \leq \frac{\tilde{\psi}(s_0)}{s_0}
\]

as \( \lim_{s \to 0} \tilde{\psi}(s) = 0 \). Another consequence is

**Proposition 3.4.** If \( \psi(t) \) is a concave rate function then so is \( \psi(\sqrt{t})^2 \).

**Proof.** By lemma 3.3 we have to show that for all \( t_0 > 0 \) the function \( (\psi(\sqrt{t} + t_0)^2 - \psi(t-0)^2)/t \) is a decreasing index function. As the mapping \( s \to (s+s_0)^2 \) is monotone it is sufficient to show that

\[
\omega(s) = \frac{\psi(s + s_0)^2 - \psi(s_0)^2}{(s + s_0)^2 - s_0^2}
\]

is monotonically decreasing.

As \( \psi \) is assumed to be concave, Lemma 3.3 implies that

\[
\sigma(s) = \frac{\psi(s + s_0) - \psi(s_0)}{s}
\]
is monotonically decreasing. Furthermore
\[ \omega(s) = \sigma(s) \left( \frac{\psi(s + s_0) + \psi(s_0)}{s + 2s_0} \right) = \sigma(s) \frac{s\sigma(s) + 2\psi(s_0)}{s + 2s_0}. \]

Now let \( s_2 < s_2 \). As \( \sigma(s) \) is monotonically decreasing on has
\[ \omega(s_1) \geq \sigma(s_2) \frac{s_1\sigma(s_2) + 2\psi(s_0)}{s_1 + 2s_0} = \sigma(s_2) \frac{s_1 + 2\psi(s_0)/\sigma(s_2)}{s_1 + 2s_0}. \]

The right-hand side is a decreasing function of \( s_1 \) if \( 2s_0 \leq 2\psi(s_0)/\sigma(s_2) \), i.e., \( \sigma(s - 2) \leq \psi(s_0) \). This is a consequence of Lemma 3.3 as stated in the remark after the lemma. Replacing \( s_1 \) by \( s_2 \) thus gives a lower bound for \( \omega(s_1) \) and thus
\[ \omega(s_1) \geq \sigma(s_2) \frac{2s_2 + 2\psi(s_0)/\sigma(s_2)}{s_2 + 2s_0} = \omega(s_2). \]

It follows that \( \omega \) is monotonically decreasing. \( \square \)

A consequence of this lemma is that for the concavity of the function \( \bar{\psi}^2((\Theta^2)^{-1}(s)) = \bar{\psi}^2(\Theta^{-1}(\sqrt{s})) \) it is thus sufficient to show that \( \bar{\psi} \circ \Theta^{-1} \) is concave.

Finally we conjecture that a similar result to the proposition above also holds more generally, i.e., that a sufficient condition for concavity of \( g \circ \bar{\psi} \circ g^{-1} \) is the concavity of \( \bar{\psi} \) where \( g \) belongs to a class of suitably chosen functions.

4. Linear Regularisation Approaches

Our goal in this section is to draw conclusions from Corollary 2.7 for linear regularisation methods. Taking into account the setting of Corollary 2.7 we assume throughout this section that the index function \( \Psi(\lambda) \) is concave and strictly increasing for all \( 0 < \lambda < \infty \) satisfying the limit conditions \( \lim_{\lambda \to +0} \Psi(\lambda) = 0 \), \( \lim_{\lambda \to \infty} \Psi(\lambda) = \infty \) (cf. Lemma 2.5), and (12). Moreover, we set
\[ \chi(\lambda) := \frac{\Psi^{-1}(\lambda)}{\lambda}, \quad 0 < \lambda < \infty, \quad \text{and} \quad \bar{\psi}(t) := \frac{1}{\sqrt{\chi(t^2)}}, \quad 0 < t < \infty. \]

Then \( \chi \) is an increasing index function with \( \lim_{\lambda \to \infty} \chi(\lambda) = \infty \) and \( \bar{\psi} \) is an increasing index function with \( \lim_{t \to +0} \bar{\psi}(t) = 0 \), hence a rate function. As outlined in section 3 under these assumptions we have \( \bar{\psi}(A^*A)[B(R)] = B_{\chi}(R) \) and the best case for regularised solutions \( f_\alpha \) approximating the exact solution \( f \in X_\chi \) based on data \( g^\delta \) satisfying (2) by using an a priori choice \( \alpha = \alpha(\delta) \) or a posteriori choice \( \alpha = \alpha(\delta, g^\delta) \) is to achieve the order optimal convergence rate (24). It is a specific consequence of interpolation theory and can be seen easily by inspection of Corollary 2.7 that a successful use requires the focus on regularisation methods which yield regularised solutions of appropriate smoothness. Precisely, there must be a ball \( B_{\chi}(R) \) to which the elements \( f_\alpha \) belong for all \( \alpha > 0 \) attributed to sufficiently small \( \delta > 0 \) and \( g^\delta \) satisfying (2).

4.1. General linear regularisation schemata. In a first approach we are going to consider linear regularisation schemes as described in many textbooks on linear regularisation theory (see, e.g., [9, Chap. 4], [10, Chap. 2] and [2, 3, 22, 23, 29]). We consider approximate solutions
\[ f_\alpha := h_\alpha(A^*A)A^*g^\delta. \]
to $f$ based on a family of piecewise continuous real functions $h_\alpha(t), \ 0 < t \leq \|A\|^2$, to which we assign bias functions

$$r_\alpha(t) := t h_\alpha(t) - 1, \ 0 < t \leq \|A\|^2.$$  

These functions depend on a regularisation parameter $\alpha \in (0, \alpha_{\text{max}}]$, where $\alpha_{\text{max}}$ may be a finite real number or $\infty$. Small $\alpha > 0$ characterise good approximation of the original problem (1), whereas larger values $\alpha$ are connected with more stability. Hence, an appropriate trade-off between the two conflicting goals approximation and stability can be controlled by the choice of $\alpha$. We say that such a function $h_\alpha$ describes a linear regularisation method if the properties

$$\lim_{\alpha \rightarrow +0} r_\alpha(t) = 0, \ 0 < t \leq \|A\|^2,$$

and

$$\sup_{0 < \alpha \leq \alpha_{\text{max}}} \sup_{0 < t \leq \|A\|^2} t |h_\alpha(t)| \leq C_1$$

with a constant $C_1 > 0$ hold. Because of (37) we have another constant $C_2 > 0$ such that

$$\sup_{0 < \alpha \leq \alpha_{\text{max}}} \sup_{0 < t \leq \|A\|^2} |r_\alpha(t)| \leq C_2$$

and hence for all $0 < \alpha \leq \alpha_{\text{max}}$ the estimate

$$\|A f_\alpha - g^\delta\| = \|(Ah_\alpha(A^*A)A^* - I)g^\delta\| \leq \sup_{0 < t \leq \|A\|^2} |r_\alpha(t)| \|g^\delta\| \leq C_2 \|g^\delta\|.$$  

This implies the limit condition $\lim_{\alpha \rightarrow +0} \|A f_\alpha - g^\delta\| = 0$ for all data $g^\delta \in Y$. As a consequence we have that there is always a parameter choice $\alpha = \alpha(\delta, g^\delta), \ 0 < \delta \leq \delta_{\text{max}},$ such that

$$\|A f_{\alpha(\delta, g^\delta)} - g^\delta\| \leq C_{\text{dis}} \delta \quad (0 < \delta \leq \delta_{\text{max}})$$

for some prescribed constant $C_{\text{dis}} > 0$. If the mapping $\alpha \mapsto \|A f_\alpha - g^\delta\|$ is even continuous, then the discrepancy principle can be realised by a parameter choice $\alpha_{\text{dis}} = \alpha_{\text{dis}}(\delta, g^\delta)$ satisfying the equation (33).

Here we call a rate function $\tilde{\varphi}$ a "qualification" of the regularisation method generated by $h_\alpha$ if there is a constant $C_{\text{quali}} > 0$ such that

$$\sup_{0 < t \leq \|A\|^2} |r_\alpha(t)| \tilde{\varphi}(t) \leq C_{\text{quali}} \tilde{\varphi}(\alpha), \ 0 < \alpha \leq \alpha_{\text{max}}.$$  

Now we are going to study under what conditions the inequality (21) in Corollary 2.8 can be fulfilled here with $\tilde{\xi}(\delta) = \delta$. First we obtain

$$\|A f_\alpha - g\| = \|Ar_\alpha(A^*A)f + Ah_\alpha(A^*A)A^*(g^\delta - Af)\| \leq \|Ar_\alpha(A^*A)f\| + C_1 \delta.$$  

In order to apply that corollary for obtaining a convergence rate (24) we assume $f \in B_\psi(R_1) = \psi(A^*A)[B(R_1)]$ taking into account the cross-connection (34). So let $f = \tilde{\psi}(A^*A)v, \ |v| \leq R_1$. Provided that $\Theta(t) := \sqrt{\tilde{\psi}(t)}$ is a qualification of the method with constant $C_{\text{quali}} > 0$ this gives with (39)

$$\|A f_\alpha - g\| \leq \sup_{0 < t \leq \|A\|^2} r_\alpha(t) \Theta(t) R_1 + C_1 \delta \leq C_{\text{quali}} R_1 \Theta(\alpha) + C_1 \delta$$

and hence an estimate of type (21) is fulfilled with $\tilde{\xi}(\delta) = \delta$ when an a priori parameter choice $\alpha = \Theta^{-1}(\delta)$ is used.
Next we will check whether \( f_\alpha \in B_\chi(R_2) \) for some \( 0 < R_2 < \infty \). We have
\[
f_\alpha = h_\alpha(A^*A)(g^\delta - Af) + h_\alpha(A^*A)A^*Af
\]
and after some reformulation
\[
f_\alpha = \tilde{\psi}(A^*A) \left[ h_\alpha(A^*A)(\psi(A^*A))^{-1}(A^*A)^{\frac{1}{2}}\tilde{g} + h_\alpha(A^*A)A^*Av \right]
\]
with \( \|\tilde{g}\| \leq \delta \), since the different functions of \( A^*A \) are commutable. Now let the interplay of the regularisation method expressed by \( h_\alpha(t) \) and the parameter choice \( \alpha = \alpha(\delta, g^\delta) \) be such that there is a constant \( C_{\text{para}} > 0 \) with
\[
\sup_{0 < t \leq \|A\|^2} \frac{\sqrt{t}|h_\alpha(t)|}{\psi(t)} \delta \leq C_{\text{para}}, \quad 0 < \delta \leq \delta_{\text{max}}.
\] (41)
The upper bound \( C_{\text{para}} \) in (41) must hold for all data \( g^\delta \in Y \) associated with the noise level \( \delta > 0 \) and satisfying (2), where the case of an a priori parameter choice \( \alpha = \alpha(\delta) \) should be included as a special case. Under (41) we have with (37)
\[
\|h_\alpha(\delta, g^\delta)(A^*A)(\psi(A^*A))^{-1}(A^*A)^{\frac{1}{2}}\tilde{g} + h_\alpha(\delta, g^\delta)(A^*A)A^*Av\| \leq R_2 := C_{\text{para}} + C_1 R_1,
\]
in other terms \( f_\alpha(\delta, g^\delta) \in \tilde{\psi}(A^*A)[B(R_2)] = B_\chi(R_2) \).

If there is a function \( \Gamma(\alpha) \) satisfying for sufficiently small \( \alpha > 0 \) the inequality
\[
\sup_{0 < t \leq \|A\|^2} \frac{\sqrt{t}|h_\alpha(t)|}{\psi(t)} \leq \Gamma(\alpha)
\] (42)
such that
\[
\delta \Gamma(\alpha(\delta, g^\delta)) \leq C_{\text{para}}, \quad 0 < \delta \leq \delta_{\text{max}},
\] (43)
this represents a sufficient condition for (41). In particular, if moreover the a priori parameter choice \( \alpha(\delta, g^\delta) := \Theta^{-1}(\delta) \) satisfies (43) we have an estimate of type (21) with \( \xi(\delta) = \delta \) for that a priori parameter choice whenever \( \Theta \) is a qualification of the regularisation method under consideration.

Hence the considerations above gave a sketch of the proof for the following proposition as a consequence of Corollary 2.8:

**Proposition 4.1.** Under the standing assumptions of this section including (34) let \( f \in X_\chi = \text{range}(\tilde{\psi}(A^*A)) \) and consider regularised solutions (35) with a generator function \( h_\alpha \) that determines the regularisation method and satisfies (36) – (37) as well as (42) with some function \( \Gamma \) such that \( \Theta(t) := \sqrt{t}\tilde{\psi}(t) \) satisfies (43) with some constant \( C_{\text{para}} > 0 \) and is a qualification of the method (cf. (38)). Then for the a priori regularisation parameter choice \( \alpha = \alpha(\delta) := \Theta^{-1}(\delta) \rightarrow +0 \) as \( \delta \rightarrow +0 \) we have the convergence rate
\[
\|f - f_\alpha\| = \mathcal{O} \left( \delta \sqrt{\Psi(K/\delta^2)} \right) \quad \text{as} \quad \delta \rightarrow +0
\] (44)
with some constant \( K > 0 \).

Note that in Proposition 4.1 the rate (44) also holds for any other parameter choice \( \alpha = \alpha(\delta, g^\delta) \) that fulfils the inequalities (43) and
\[
\|A_f(\alpha(\delta, g^\delta)) - g\| \leq C\delta, \quad 0 < \delta \leq \delta_{\text{max}},
\] (45)
with some constant \( \hat{C} > 0 \).
Example 1. The most prominent example of a linear regularisation method (35) is the Tikhonov regularisation with the generator function $h_\alpha(t) = \frac{1}{t + \alpha}$ and with the bias function $r_\alpha(t) = \frac{1}{t^{\alpha}}$, where the requirements (36) and (37) are satisfied for the constants $C_1 = C_2 = 1$. It is well known that all concave rate functions $\bar{\varphi}$ are qualifications of the method satisfying (38) with the constant $C_{\text{quali}} = 1$. From that class we consider the monomials $\bar{\varphi}(t) = t^{\nu}$ for exponents $0 < \nu \leq 1$. Then $\Theta(t) = \sqrt{\bar{\varphi}}(t)$ is a qualification with the same constant for the Tikhonov regularisation in case of a rate function $\bar{\varphi}(t) = t^{\nu}$ with $0 < \mu \leq 1/2$. Taking into account (34) this rate function is associated with $\chi(\lambda) = \lambda^{2\mu}$ and the strictly concave function $\Psi(\lambda) = \lambda^{\frac{\mu}{2}}$. By the estimate (40) we have then (45) with $\tilde{C} = R_1 + 1$ for $f = (A^*A)^{\mu}v$, $\|v\| \leq R_1$ and for the a priori parameter choice

$$\alpha = \Theta^{-1}(\delta) = \delta^{\frac{1}{2\mu}}. $$

To derive a function $\Gamma$ such that (42) is valid, we exploit the inequality

$$\frac{t^{\kappa}}{t + \alpha} \leq (1 - \kappa)^{1-\kappa} \kappa^{\kappa} \alpha^{\kappa-1},$$

which holds for all $t > 0$, $\alpha > 0$ and $0 < \kappa < 1$. In the limit case $\kappa = 0$ we also have the inequality $1/(t+\alpha) \leq 1/\alpha$. Thus there is a constant $\hat{c}$ depending on $\kappa \in (0,1)$ such that $\frac{\hat{c}}{t^{\kappa}(t+\alpha)} \leq \frac{\hat{c}}{\alpha^{\kappa}}$. By setting $\kappa := 1/2 - \mu$ we obtain for $\psi(t) = t^{\mu}$, $0 < \mu \leq 1/2$ the inequality (42) with the function

$$\Gamma(\alpha) = \frac{\hat{c}}{\alpha^{\mu+\frac{1}{2}}},$$

Then one easily verifies that $\delta \Gamma(\Theta^{-1}(\delta)) \leq \frac{\delta \hat{c}}{(\delta^{\frac{1}{2\mu}})^{\mu+\frac{1}{2}}} = \hat{c}$ and that (43) is fulfilled

with $C_{\text{para}} = \hat{c}$. Hence Proposition 4.1 applies and we obtain for the parameter choice (46) and all $0 < \mu \leq 1/2$ the optimal convergence rate

$$\|f - f_\alpha\| = O\left(\delta^{\frac{1}{2\mu}}\right) \quad \text{as} \quad \delta \to +0.$$ 

The best possible rate obtained in that way is $\|f - f_\alpha\| = O\left(\sqrt{\delta}\right)$ for $\mu = 1/2$. For $\mu > 1/2$ the function $\Psi$ remains strictly concave, but a finite function $\Gamma(\alpha)$ in (42) fails to exist, since we have $\sup_{0 < t \leq \|A\|^2} \frac{\sqrt{\delta}}{\psi(t)(t+\alpha)} = +\infty$. The limitation of Proposition 4.1 to lower Hölder rates than the saturation of Tikhonov’s method admits seems to be a consequence of the fact that our approach based on Corollary 2.8 and the construction (35) do not interact good enough in case of higher smoothness of $f$. In order to overcome that effect, we will consider another approach in the following subsection.

4.2. Regularisation with unbounded operators and range inclusions. In a second approach, under a non-standard source condition (5) characterising the available a priori knowledge on the solution smoothness, we exploit a variant of the Tikhonov regularisation with regularised solutions

$$f_\alpha := G(GA^*AG + \alpha I)^{-1}GA^*g^\delta,$$

where $G : X \to X$ is an injective bounded self-adjoint positive definite linear operator $G : X \to X$ with non-closed range, i.e., zero is an accumulation point of the spectrum $\text{spec}(G)$ of the operator $G$. Since the unbounded linear operator with $B = G^{-1} : \text{range}(G) \subseteq X \to X$ is frequently a differential operator, this approach is
sometimes called *regularisation with differential operators*. Precisely, by construction the element \( f_\alpha \in \text{range}(G) \) is well-defined for all \( \alpha > 0 \) as the minimiser of the extremal problem
\[
T_\alpha(\tilde{f}) := \| A\tilde{f} - g^\delta \|^2 + \alpha \| B\tilde{f} \|^2 \to \min, \quad \text{subject to} \quad \tilde{f} \in \text{range}(G),
\]
and then the penalty term in \( T_\alpha \) contains derivatives of the function \( \tilde{f} \).

To apply Corollary 2.8 under our setting (34) we assume \( f \in G[B(R_1)] \), with \( G[B(R)] \) from (26), and a link condition
\[
\text{range}(G) \subseteq X_\chi = \text{range}(\tilde{\psi}(A^*A)),
\]
which is equivalent to
\[
\| Gw \| \leq C \| \tilde{\psi}(A^*A)w \|, \quad \text{for all} \quad w \in X,
\]
with some \( C > 0 \). Then from [15, Lemma 6.2] we obtain that \( f \in G[B(R_1)] \) implies \( f \in B_\chi(CR_1) = \tilde{\psi}(A^*A)[B(CR_1)] \).

Along the lines of the paper [7] by Cheng and Yamamoto we consider an a priori parameter choice \( \alpha = \alpha(\delta) \) as
\[
\xi \delta^2 \leq \alpha(\delta) \leq \tau \delta^2, \quad 0 < \xi \leq \tau < \infty,
\]
with constants \( 0 < \xi \leq \tau < \infty \), for which we obtain from \( T_\alpha(f_\alpha) \leq T_\alpha(f) \) the inequalities
\[
\| Af_{\alpha(\delta)} - g^\delta \|^2 + \alpha(\delta) \| G^{-1}f_{\alpha(\delta)} \|^2 \leq \| Af - g^\delta \|^2 + \alpha(\delta) \| G^{-1}f \|^2 \leq \delta^2 + \tau \delta^2 R_1^2.
\]
Now we have
\[
\| Af_{\alpha(\delta)} - g \| \leq \bar{C} \delta, \quad \text{with} \quad \bar{C} = \sqrt{1 + \tau R_1^2} + 1
\]
satisfying condition (21) with \( \xi(\delta) = \delta \) and
\[
\| G^{-1}f_{\alpha(\delta)} \| \leq \sqrt{\frac{\delta^2}{\alpha(\delta)}} + \| G^{-1}f \| \leq \sqrt{\frac{1}{\xi} + R_1^2} =: R_2.
\]
This yields \( f_{\alpha(\delta)} \in G[B(R_2)] \), thus \( f_{\alpha(\delta)} \in B_\chi(CR_2) = \tilde{\psi}(A^*A)[B(CR_2)] \) and consequently an estimate of type (22) with \( \xi(\delta) = \delta \) and \( CR_1, CR_2 \) instead of \( R_1, R_2 \).

With the above considerations we have shown the convergence rate result of the following proposition again as a consequence of Corollary 2.8:

**Proposition 4.2.** Under the standing assumptions of this section including (34) let \( f \) satisfy (5), where the link condition (48) is valid. Then for the a priori regularisation parameter choice (50) we have the convergence rate (44) with some constant \( K > 0 \).

Due to [16, Corollary 4.5] for all concave \( \Psi \) fulfilling the standing assumptions of this section the rate (44) is even order optimal in the sense of
\[
\mathcal{O} \left( \delta \sqrt{\Psi(K/\delta^2)} \right) = \mathcal{O} \left( \omega(G[B(R_1)], \delta) \right) \quad \text{as} \quad \delta \to +0.
\]

As already discussed in the introduction the requirement (48) gets stronger for higher rates in (44). In many applications (see as an illustration the examples in [18]) one can only verify range inclusions of the form (8) with some rate function \( \tilde{\rho} \). Under operator monotonicity of the function \( [\tilde{\rho}^{-1}(\sqrt{t})]^2 \) (8) implies (48) with \( \tilde{\psi}(t) = \tilde{\rho}^{-1}(\sqrt{t}) \) and \( \chi(\lambda) = \left[ \tilde{\rho}^{-1} \left( 1/\sqrt{\lambda} \right) \right]^2 \).
In order to verify in general for what index functions $\chi$ a range inclusion (48) with $\text{range}(G) \subseteq X_\chi$ is fulfilled, one can use the spectral theorem for unbounded self-adjoint operators $T$ (see [34, Chapter VII.3] and also [28, Chapter VIII]). In the Hilbert space $X$, the injective, densely defined, self-adjoint, positive definite, and unbounded linear operator $T$ is unitarily invariant to a multiplication operator $\mathcal{M}$ expressed by a real multiplier function $m$. This means that there are a measure space $(\Sigma, \mathcal{A}, \mu)$ with finite measure $\mu$, a unitary operator $U: X \rightarrow L^2(\Sigma, \mathcal{A}, \mu)$ and a real measurable function $m(t)$, $t \in \Sigma$, such that $[\mathcal{M} h](t) := m(t) h(t)$ a.e., where $\mathcal{M}$ maps in $L^2(\Sigma, \mathcal{A}, \mu)$, and

$$U T U^* h = \mathcal{M} h = m \cdot h$$

for all $h$ from the domain of $\mathcal{M}$. We note that the closure of the range $\text{range}(m)$ of the multiplier function $m$ and the spectrum $\text{spec}(T) \subseteq [||T||^{-1}, \infty)$ of the operator $T$, possessing $+\infty$ as an accumulation point, coincide. Moreover, we have for index functions $\psi \in \mathcal{I}$ and $h$ from the domain of $\psi(\mathcal{M})$

$$U \psi(T) U^* h = \psi(\mathcal{M}) h = \psi(m) \cdot h.$$

Then by using the notations $\hat{f} := U f \in L^2(\Sigma, \mathcal{A}, \mu)$ and $(\widehat{Gw}) := U G w \in L^2(\Sigma, \mathcal{A}, \mu)$ by definition we immediately find that $\text{range}(G) \subseteq X_\chi$ is equivalent to the condition that

$$\int_{\Sigma} \chi(m(t))(\widehat{Gw})(t)^2 dt < \infty$$

holds for all $w \in X$. In Example 2 with background in imaging (cf. [30]) we will consider the special case that $U$ denotes the two-dimensional Fourier transform and that the corresponding measure space is $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu)$ with the associated Borel $\sigma$-algebra and measure. In that example, $T$ and $G$ are commuting operators, both non-compact with a non-closed range.

On the other hand, in Example 3 we will exploit the one-dimensional Fourier transform to formulate sufficient conditions such that classical source conditions are satisfied for linear compact integral operators.

5. Examples

In the remaining examples we illustrate the theory. All the occurring operators $A$ are linear integral operators. First the Example 2 refers to convolution operators $A$ which occur, for example, when the deblurring of noisy images is under consideration. Then the Example 3 illustrates the low rate case where an integral equation with a smooth kernel is solved and it is known that the solution is in a Sobolev space. The situation here is similar as in the case of elliptic partial differential equations and has been discussed in [6]. In contrast to the PDE situation here convergence rates are low, typically of the form $O(|\log(\delta)|^{-k})$. The final Example 4 illustrates the high rate case where a derivative of data in the range of an integral operator with smooth kernel is considered. The high convergence rates are here of the form $O(\delta|\log(\delta)|^k)$.

In the examples we consider functions over $\mathbb{R}^d$ ($d = 1, 2$) and Sobolev spaces $H^l(\mathbb{R}^d)$ ($l = 1, 2, ...$) of Hilbert type will be used with norms $|| \cdot ||_l$ defined by

$$||x||_l^2 = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} (1 + |\omega|^2 + \cdots + |\omega|^2) |\hat{x}|^2 d\omega,$$
where \( \hat{x} = \hat{x}(\omega) \), \( \omega \in \mathbb{R}^d \), is the Fourier transform of \( x \). Now let \( E_1 : H^1(\mathbb{R}^d) \to L^2(\mathbb{R}^d) \) denote the embedding and \( E_1^* \) the adjoint of \( E_1 \). Then \( E_1 E_1^* : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d) \) is an integral operator and

\[
E_1 E_1^* y(\omega) = \frac{\hat{y}(\omega)}{1 + |\omega|^2 + \cdots + |\omega|^{2l}}.
\]

**Example 2.** In this example with \( X = Y = L^2(\mathbb{R}^2) \) we are interested in deblurring, that means in finding a true picture which is characterised by a function \( f = f(t) \in L^2(\mathbb{R}^2) \), \( t = (t_1, t_2)^T \), that satisfies a linear operator equation (1) of convolution type

\[
Af(s) = \int_{\mathbb{R}^2} k(s - t)f(t) \, dt = g(s), \quad s = (s_1, s_2)^T \in \mathbb{R}^2,
\]

where \( g \in L^2(\mathbb{R}^2) \) is a blurred image of \( f \) which is additionally contaminated with noise such that only the noisy blurred image \( g^\delta \in L^2(\mathbb{R}^2) \) satisfying (2) available as data. Following [4, Chapter 3] the kernel function \( k(\tau), \tau = (\tau_1, \tau_2)^T \in \mathbb{R}^2 \), is called point spread function of a space invariant imaging system under consideration. We assume that the kernel is such that its Fourier transform \( \hat{k} = \hat{k}(\omega), \omega = (\omega_1, \omega_2)^T \), called transfer function is bounded. Different variants of such deblurring problems are presented and analysed in [4]. As a reference situation we exploit for illustration a variant of an out-of-focus blur for which

\[
\hat{k}(\omega) = \frac{2 J_1(D|\omega|)}{D|\omega|}
\]

where \( J_1 \) is the Bessel function of order one and \( D \) is the radius of the circle of confusion (cf. [4, formula (3.25) on p.60]). The linear convolution operator \( A : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2) \) in this example has a non-closed range but it is non-compact and the kernel is not square integrable.

In order to apply our theory to this example one needs to find an index function \( \theta \) and a symmetric positive definite operator \( T \) such that \( A^* A = \theta(T) \). A natural choice in this context is \( T = -\Delta \) and in this case \( \theta \) needs to satisfy \( |\hat{k}(\omega)|^2 = \theta(|\omega|^2) \).

This, however, is not possible, as \( \hat{k}(\omega) \) is zero for some finite \( \omega \) but an index function has to satisfy \( \theta(\lambda) > 0 \) for all \( \lambda > 0 \) and it can only be zero asymptotically at zero or infinity. It is thus not possible to get error bounds for the deblurring problem using the variable Hilbert scale theory and \( T = -\Delta \).

One does not have this problem if one chooses \( T = (A^* A)^{-1} \). Let us define the solution smoothness as \( f \in H^1(\mathbb{R}^2) \). Then we have the operator \( G = E_1^* E_1 \) in (5) characterising the associated non-standard source condition. To find index functions \( \chi \) that satisfy the link condition (48) we can make use of formula (51) taking into account that \( m(\omega) = 1/|\hat{k}(\omega)|^2 \) and

\[
|\hat{G}w(\omega)|^2 = (1 + |\omega|^2 + \cdots + |\omega|^{2l})^{-1} |\hat{w}(\omega)|^2.
\]

Then the range inclusion \( \text{range}(G) \subseteq X_\chi \) takes the form

\[
\chi \left( \frac{1}{|\hat{k}(\omega)|^2} \right) \frac{1}{(1 + |\omega|^2 + \cdots + |\omega|^{2l})} \leq C < \infty \quad \text{for all} \quad \omega \in \mathbb{R}^2.
\]

This range condition can only be satisfied if \( \chi \) is bounded, i.e., i.e. \( \chi(\lambda) \leq C < \infty \) for all \( \lambda > 0 \). This is again a consequence of the existence of zeros of \( \hat{k}(\omega) \) for finite
$\omega$. A consequence of the finiteness of $\chi$ is $L^2(\mathbb{R}^2) \subset H_\chi$ and it follows that the “source condition” reduces to $f \in L^2(\mathbb{R})$ which does not lead to an error bound.

The failure of the above attempts to get error bounds clearly illustrates the need to extend the variable Hilbert scale theory to be able to cope with the deblurring problem. One can, however, deal with a partial deblurring problem. Observe that one has the asymptotics

$$2 \left| \frac{D[l(|\omega|)]}{D[|\omega|]} \right| \asymp |\omega|^{-3/2}$$

for large $|\omega|$ (cf. [4, formula (3.29) on p.60]). It follows that $\hat{k}(\omega) = |\omega|^{-3/2} \kappa(\omega)$ for some bounded $\kappa(\omega)$. The first factor $|\omega|^{3/2}$ relates to a “smoothing component” of the out-of-focus blur situation. We now consider inversion of this smoothing component only. For this we introduce an integral operator $A$ with kernel $k$ which satisfies

$$(54) \quad \hat{k}(\omega) = |\omega|^{3/2}. \tag{54}$$

For the “partial” out-of-focus blur situation (54) and monomials $\chi(\lambda) = \lambda^\kappa$, $\kappa > 0$, we have (53) if and only if $\kappa \leq \frac{2}{3}$. With the relation $\chi(\lambda) = \Psi^{-1}(\lambda)/\lambda$ this corresponds with $\Psi(\lambda) \leq \lambda^{3/(2\kappa+3)}$. Hence based on Proposition 4.2 for the situation (54) and under $f \in H^1(\mathbb{R}^2)$ a best possible convergence rate

$$\|f - f_\alpha\|_{L^2(\mathbb{R}^2)} = O(\delta^{2\kappa/\kappa+3}) \quad \text{as} \quad \delta \to +0$$

can be obtained by Tikhonov regularisation with $H^1$-penalty term.

**Example 3.** In this example we consider compact forward operators $A$ in equation (1) with $X = Y = L^2(\mathbb{R})$ in form of linear operators $A : L^2(\mathbb{R}) \to L^2(\mathbb{R})$, for which the range of the operator $K := \tilde{\psi}(A^* A)$ is a subset of $H_\phi$ with some index function $\phi$ and some rate function $\psi$. That means, we have range($K$) $\subseteq H_\phi$ and a classical source conditions (4) is valid for $f \in H_\phi$ implying the corresponding convergence rates in regularisation. In this context, let $K$ be a linear Fredholm integral operator of Hilbert-Schmidt type. For such operators one can provide conditions on the kernel which guarantee this range condition.

**Lemma 5.1.** Let $K : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ be a Hilbert-Schmidt operator with kernel $k(t, s) \in L^2(\mathbb{R}^2)$. Furthermore, let $\tilde{K} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ be an integral operator with kernel $\hat{k}(\omega, s) = \int_R e^{-i\omega t} k(t, s) \, dt$. Then $\tilde{K}$ is a Hilbert-Schmidt operator and

$$\tilde{K}x = \overline{Kx}, \quad x \in L^2(\mathbb{R}).$$

**Proof.** The adjoint operator $K^*$ of $K$ is an integral operator with kernel $k^*(s, t) = \overline{k(t, s)}$ as a consequence of the theorem of Fubini. By Plancherel’s theorem one has

$$K^* u = \frac{1}{2\pi} \tilde{K}^* \hat{u}.$$

An application of Parseval’s identity several times gives for $u, v \in L^2(\mathbb{R})$:

$$\frac{1}{2\pi} (\hat{u}, \tilde{K}v) = (u, Kv) = (K^* u, v) = \frac{1}{2\pi} (\hat{u}, \tilde{K}v) = \frac{1}{2\pi} (\hat{u}, \hat{K}v).$$

**Proof.**
Proposition 5.2. Let $K : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ be a Hilbert-Schmidt operator where the Fourier transform $\hat{k}(\omega, s) = \int_{\mathbb{R}} e^{-i\omega t} k(t, s) \, dt$ of the kernel of $K$ satisfies
\[
\int_{\mathbb{R}} \int_{\mathbb{R}} \phi(\omega^2) |\hat{k}(\omega, s)|^2 \, ds \, d\omega < \infty
\]
for some index function $\phi$. Then $\text{range}(K) \subseteq H_\phi$.

Proof. By Lemma 5.1 one has
\[
\hat{y}(\omega) = \int_{\mathbb{R}} \hat{k}(\omega, s) x(s) \, ds
\]
which we insert into the following bound, obtained from the Cauchy-Schwarz inequality:
\[
\left| \int_{\mathbb{R}} \hat{k}(\omega, s) x(s) \, ds \right|^2 \leq \int_{\mathbb{R}} |\hat{k}(\omega, s)|^2 \, ds \|x\|^2.
\]
It follows that for $y = Kx$ with $x \in L^2(\mathbb{R})$ one has
\[
\|y\|^2_\phi = \frac{1}{2\pi} \int_{\mathbb{R}} \phi(\omega^2) |\hat{y}(\omega)|^2 \, d\omega
\leq \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \phi(\omega^2) |\hat{k}(\omega, s)|^2 \, ds \, d\omega \|x\|^2
\]
and consequently $y \in H_\phi$. □

Example 4. As a concrete application example we consider a problem from derivative spectroscopy [31]. Here numerical derivatives are used to enhance the resolution of measured spectra in order to separate close peaks. An instance is the Eddington correction formula. The approach determines
\[
f = Lg := g - \frac{g^{(2)}}{2}
\]
from observed $g_\delta$ where $g^{(2)}$ is the second derivative of $g$. We now apply the theory developed so far to determine how well $f = Lg$ can be determined from spectral data $g_\delta$.

For $f \in H^2(\mathbb{R})$ and $f = Lg$ the Fourier transforms $\hat{f}$ and $\hat{g}$ satisfy
\[
\hat{f}(\omega) = (1 + \omega^2/2) \hat{g}(\omega), \quad \text{a.e.}
\]
Using Plancherel’s theorem, one obtains from this the bounds
\[
\frac{1}{2} \|f\|_2 \leq \|Lf\| \leq \|f\|_2, \quad f \in H^2(\mathbb{R})
\]
which means in particular that $\|Lf\|$ is an equivalent norm for $H^2(\mathbb{R})$. Using standard arguments, one can then show that $L : H^2(\mathbb{R}) \to L^2(\mathbb{R})$ is a Hilbert space isomorphism. Using the convolution theorem one sees that $A = E_2 L^{-1} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is an integral operator with
\[
Af(t) = \frac{1}{\sqrt{2}} \int_{\mathbb{R}} \exp(-\sqrt{2}|t-s|) f(s) \, ds \quad t \in \mathbb{R}
\]
where $E_2$ denotes the embedding $H^2(\mathbb{R}) \to L^2(\mathbb{R})$. As $L^{-1}$ maps $L^2(\mathbb{R})$ onto $H^2(\mathbb{R})$ the range of $A$ can be identified with $H^2(\mathbb{R})$. 
In addition to the Sobolev spaces, which form a classical Hilbert scale, we will use a variable Hilbert scale $H_\phi$ with norms $\| \cdot \|_\phi$ defined by

$$\| x \|_\phi^2 = \frac{1}{(2\pi)^k} \int_{-\infty}^{\infty} \phi(\omega^2) |\hat{x}(\omega)|^2 d\omega$$

where $\phi$ are index functions. Note that we have here $H_\phi = X_\phi$ for $X = L^2(\mathbb{R})$ (see [11]) and the generating operator is the second order differential operator $T = -d^2/dt^2$. The index functions

$$\nu_k(\lambda) = 1 + \lambda + \cdots + \lambda^k = \frac{\lambda^k - 1}{\lambda - 1}$$

define the Sobolev spaces, in particular, one has $H_{\nu_k} = H^k(\mathbb{R})$ and furthermore, the Sobolev norm is equal to the norm of the corresponding variable Hilbert scale:

$$\| f \|_{k} = \| f \|_{\nu_k}, \quad f \in H_{\nu_k}.$$

In this framework, we now get error bounds analogue to the ones in Corollary 2.7 which are again a consequence of Lemma 2.3.

**Corollary 5.3.** Let $H_\nu$ be the Hilbert scales generated by $T = -d^2/dt^2$ from $L^2(\mathbb{R})$. Furthermore, let $A : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ be a (convolution) operator satisfying

$$A^* A = \theta(T)$$

for some bounded index function $\theta$. Moreover, let $\phi, \psi$ and $\Psi$ be index functions and $\Psi$ be concave such that

$$\phi(\lambda) \leq \Psi(\phi(\lambda)), \quad \phi(\lambda) \theta(\lambda) = 1, \quad \text{for} \quad t > 0.$$

If $Af = g \in H_\psi$ and if $f_\alpha$ is such that $Af_\alpha \in H_\psi$, and

$$\| Af_\alpha - g \|_\psi = \zeta$$

then

$$\| f - f_\alpha \| \leq \epsilon \sqrt{\Psi(\zeta^2/\epsilon^2)}.$$

**Proof.** Note that $\theta(T)$ is well defined by the Fourier transform and as $\theta$ is bounded, so is $A$. Furthermore it follows from the condition $A^* A = \theta(T)$ that $\| h \|_\alpha = \| Ah \|$ for all $h \in L^2(\mathbb{R})$ and $\| h \|_{\psi \theta} = \| Ah \|_\psi$, for all $h$ with $Ah \in H_\psi$.

By the variable Hilbert scale interpolation inequality (Lemma 2.3) one has

$$\| f - f_\alpha \|_{\psi \theta}^2 \leq \| f - f_\alpha \|_\phi^2 \Psi \left( \frac{\| f - f_\alpha \|_{\psi \theta}^2}{\| f - f_\alpha \|_\phi^2} \right).$$

Now $\phi(\lambda) \theta(\lambda) = 1$ and $\| f_\alpha - f \|_\alpha = \| Af_\alpha - g \|$ as $f_\alpha - f \in L^2(\mathbb{R})$. Furthermore, $\| f_\alpha - f \|_{\psi \theta} = \| Af_\alpha - g \|_\psi$. By Lemma 2.6 one then has

$$\| f_\alpha - f \| \leq \| Af_\alpha - g \| \sqrt{\Psi(\| Af_\alpha - g \|_{\psi \theta}^2/\| Af_\alpha - g \|^2)}.$$

The bound follows by another application of Lemma 2.6. \hfill \Box

In comparison with Corollary 2.7 this corollary uses an operator $T$ which is not necessarily equal to $(A^* A)^{-1}$ but more importantly, the source condition is here not given as a property of the solution $f$ but of the data $g$.

For the application of this corollary to the case of the Eddington correction formula one chooses $\theta(\lambda) = 1/(1 + \lambda/2)$ and so $\phi(\lambda) = 1 + \lambda/2$. 
In contrast to the usual case, where the source condition is stated as a condition on \( f \), here the source condition is stated as a condition on (the original spectrum) \( g \). This source condition results from physical models for the spectrum, and, in particular for the so-called spectral broadening. A variety of models are used, the most common ones are the Gaussian, Lorenz and Voigt spectra where a Voigt spectrum is a combination of a Lorenz and a Gaussian spectrum. Here we consider Gaussian spectra defined by

\[
g(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{(t-s)^2}{2}\right)v(s)\, ds
\]

for some \( v \in L^2(\mathbb{R}) \). For a different discussion and more background on the problem, the reader may consult the paper by Hegland [13].

It follows that \( g \in H_\psi \) with \( \psi(\lambda) = \exp(\lambda) \). The concave function \( \Psi \) can then be chosen as

\[
\Psi(\lambda) = \begin{cases} 
\lambda, & \text{for } \lambda \leq 1 \\
(1 + \log(\lambda)/2)^2, & \text{for } \lambda \geq 1.
\end{cases}
\]

It follows that \( \Psi \) is concave and that \( \phi(\lambda) \leq \Psi(\psi(\lambda)) \). As a consequence one gets the error bounds

\[
\|f - f_\alpha\| \leq \delta(1 + \log(\eta/\delta))
\]

for \( \delta < \eta \) and \( \|f - f_\alpha\| \leq \eta \) if \( \delta \geq \epsilon \). The stabilisation guarantees that even if the errors are very large, the error of the approximation does not grow to infinity. In fact, the solution \( f_\alpha = 0 \) would probably be a good choice for the large data error case.

References

[1] K. Atkinson and W. Han, *Theoretical Numerical Analysis – A Functional Analysis Framework*, 3rd ed., Texts in Applied Mathematics, vol. 39, Springer, Dordrecht, 2009.
[2] A. B. Bakushinsky and M. Yu. Kokurin, *Iterative Methods for Approximate Solution of Inverse Problems*, Springer, Dordrecht, 2004.
[3] J. Baumeister, *Stable Solution of Inverse Problems*, Friedr. Vieweg & Sohn, Braunschweig, 1987.
[4] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*, Institute of Physics Publishing, Bristol, 1998.
[5] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.
[6] A. Böttcher, B. Hofmann, U. Tautenhahn, and M. Yamamoto, *Convergence rates for Tikhonov regularization from different smoothness conditions*, Appl. Anal. **85** (2006), no. 5, 555–578.
[7] J. Cheng and M. Yamamoto, *One new strategy for a priori choice of regularizing parameters in Tikhonov’s regularization*, Inverse Problems **16** (2000), no. 4.
[8] D. Düvelmeyer, B. Hofmann, and M. Yamamoto, *Range inclusions and approximate source conditions with general benchmark functions*, Numer. Funct. Anal. Optim. **28** (2007), no. 11-12, 1245–1261.
[9] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, 1996.
[10] C. W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman, Boston, MA, 1984.
[11] M. Hegland, *An optimal order regularization method which does not use additional smoothness assumptions*, SIAM J. Numer. Anal. **29** (1992), no. 5, 1446–1461.
[12] , *Variable Hilbert scales and their interpolation inequalities with applications to Tikhonov regularization*, Appl. Anal. **59** (1995), no. 1-4, 207–223.
[13] , *Error bounds for spectral enhancement which are based on variable Hilbert scale inequalities*, 2009, http://www.citebase.org/abstract?id=oai:arXiv.org:0911.2695.
[14] B. Hofmann, *Approximate source conditions in Tikhonov-Phillips regularization and consequences for inverse problems with multiplication operators*, Mathematical Methods in the Applied Sciences 29 (2006), 351–371.

[15] B. Hofmann and P. Mathé, *Analysis of profile functions for general linear regularization methods*, SIAM J. Numer. Anal. 45 (2007), no. 3, 1122–1141.

[16] B. Hofmann, P. Mathé, and M. Schieck, *Modulus of continuity for conditionally stable ill-posed problems in Hilbert space*, J. Inverse Ill-Posed Probl. 16 (2008), no. 6, 567–585.

[17] B. Hofmann, P. Mathé, and H. von Weizsäcker, *Regularization in Hilbert space under unbounded operators and general source conditions*, Inverse Problems 25 (2009), no. 11, 115013.

[18] B. Hofmann and M. Yamamoto, *Convergence rates for Tikhonov regularization based on range inclusions*, Inverse Problems 21 (2005), no. 3, 805–820.

[19] 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), no. 5, 1279–1299.

[20] ———. *Regularization of exponentially ill-posed problems*, Numer. Funct. Anal. Optim. 21 (2000), no. 3-4, 439–464.

[21] V. K. Ivanov and T. I. Koroljuk, *The estimation of errors in the solution of linear ill-posed problems*, Ž. Vyčisl. Mat. i Mat. Fiz. 9 (1969), 30–41.

[22] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer-Verlag, New York, 1996.

[23] P. Mathé and S.V. Pereverzev, *Discretization strategy for linear ill-posed problems in variable Hilbert scales*, Inverse Problems 19 (2003), no. 6, 1263–1277.

[24] B. Hofmann and U. Tautenhahn, *Interpolation in variable Hilbert scales with application to inverse problems*, Inverse Problems 22 (2006), no. 6, 2271–2297.

[25] ———. *Geometry of linear ill-posed problems in variable Hilbert scales*, Inverse Problems 19 (2003), no. 3, 789–803.

[26] P. Mathé and U. Tautenhahn, *Interpolation in variable Hilbert scales with application to inverse problems*, Inverse Problems 22 (2006), no. 6, 2271–2297.

[27] F. Natterer, *Error bounds for Tikhonov regularization in Hilbert scales*, Appl. Anal. 18 (1984), no. 1-2, 29–37.

[28] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, second ed., Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980.

[29] A. Rieder, *Keine Probleme mit inversen Problemen*, Friedr. Vieweg & Sohn, Braunschweig, 2003.

[30] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational Methods in Imaging*, Applied Mathematical Sciences, vol. 167, Springer, New York, 2009.

[31] F.R. Stauffer and H. Sakai, *Derivative spectroscopy*, Applied Optics 7 (1968), no. 1, 61.

[32] H. Tanabe, *Equations of evolution*, Monographs and Studies in Mathematics, vol. 6, Pitman, Boston, Mass., 1979, Translated from the Japanese.

[33] U. Tautenhahn, *Error estimates for regularization methods in Hilbert scales*, SIAM J. Numer. Anal. 33 (1996), no. 6, 2120–2130.

[34] D. Werner, *Funktionalanalysis*, extended ed., Springer-Verlag, Berlin, 2000.

E-mail address: markus.hegland@anu.edu.au

E-mail address: bernd.hofmann@mathematik.tu-chemnitz.de