# Direct and inverse results in variable Hilbert scales

Peter Mathé [a]

[a] *Weierstraß Institute for Applied Analysis and Stochastics, Mohrenstraße 39,*
*10117 Berlin, Germany*

Bernd Hofmann [b,1]

[b] *Department of Mathematics, Chemnitz University of Technology, 09107*
*Chemnitz, Germany*

---

## Abstract

Variable Hilbert scales are an important tool for the recent analysis of inverse problems in Hilbert spaces, as these constitute a way to describe smoothness of objects other than functions on domains. Previous analysis of such classes of Hilbert spaces focused on interpolation properties, which allows us to vary between such spaces. In the context of discretization of inverse problems, first results on approximation theoretic properties appeared. The present study is the first which aims at presenting such spaces in the context of approximation theory. The authors review and establish direct theorems and also provide inverse theorems, as such are common in approximation theory.

*Key words:* variable Hilbert scales, source conditions, approximability, distance functions, Jackson- and Bernstein-type inequality, inverse theorems, Fenchel duality

---

*Email addresses:* `mathe@wias-berlin.de` (Peter Mathé), `hofmannb@mathematik.tu-chemnitz.de` (Bernd Hofmann).

*URLs:* `http://www.wias-berlin.de/ mathe` (Peter Mathé), `http://www.tu-chemnitz.de/mathematik/inverse_probleme/hofmann/` (Bernd Hofmann).

[1] This author acknowledges support by Deutsche Forschungsgemeinschaft (DFG) under grant HO1454/7-1

*Preprint submitted to Elsevier Science* *4 January 2007*

## 1 Introduction

In recent analysis of ill-posed linear operator equations $Ax = y$ with bounded linear operators $A : X \to Y$ mapping between Hilbert spaces $X$ and $Y$ smoothness in terms of *general source conditions* became attractive, see (14) and the more recent (2; 7). These general source conditions are closely related to classes of Hilbert spaces, which are called *variable Hilbert scales.* Such classes of Hilbert spaces might be also of interest without the context of inverse problems. They constitute analogs and extensions to Sobolev type classes of functions with bounded smoothness. For function classes of this type typical questions arise, and some of those are the objective of the present study. Precisely we shall discuss whether there are *characteristics* for the smoothness of an element $x^\dagger \in X$. Within the classical approximation theory such results are known as direct or *Jackson*-type theorems. Moreover, there exist *inverse* theorems that conclude from the behavior of certain characteristics to the smoothness of $x^\dagger$.

In contrast to the classical approximation theory here we do not deal with the approximation of smooth functions, and we use related concepts of smoothness assigned to elements in Hilbert space. As already mentioned smoothness will be given in terms of general source conditions. As useful characteristics we analyze two functions, one related to the *degree of approximation*, and one measuring the lack of some *benchmark smoothness.* The latter is of interest in the case when some source condition is satisfied only approximately, a situation first studied systematically in (4; 6). This leads to the notion of a distance function, and we shall use the modification as introduced in (7, Sect. 5). Typical direct results assert that smoothness yields approximability as well as a certain decay rate for the related distance function, see e.g., (13, Prop. 2) and (7, Thm. 5.9), respectively.

It is the goal of the present analysis to exhibit some converse results, extending special cases as studied in the literature, in particular (4, Rem. 1).

## 2 Notation and preliminary results

We shall assume that we are given a non-negative self-adjoint operator $H\colon X \to X$, which in addition is compact and injective. Then it admits a singular value decomposition

$$Hx = \sum_{j=1}^{\infty} s_j \langle x, u_j \rangle u_j, \quad x \in X, \tag{1}$$

with (non-increasing) sequence $s_1 \geq s_2 \geq \ldots > 0$, and complete orthonormal system $\{u_j,\ j = 1, 2, \ldots\} \subset X$ . The singular values $s_j$ are obtained as eigenvalues of the mapping $H$, in particular we let $a := \|H\|$.

If such analysis is dealt with linear operator equations $Ax = y$ as in (7) and (15), then we can consider $H = A^*A$ and $\sqrt{s_j}$ are the singular values of $A$. However, here we focus on pure approximation aspects and neither corresponding operator equations nor its regularization are under consideration.

As in (7) we call a function $\varphi\colon [0, a] \to [0, \infty)$ an *index function* if it is continuous and strictly increasing with $\varphi(0) = 0$. An index function $\varphi$ is said to *obey a $\Delta_2$-condition* if there is $C_2 < \infty$ for which $\varphi(2t) \leq C_2\varphi(t)$.

Using spectral calculus we assign each index function $\varphi\colon [0, a] \to [0, \infty)$ the bounded linear operator $\varphi(H)\colon X \to X$ as

$$\varphi(H)x = \sum_{j=1}^{\infty} \varphi(s_j)\langle x, u_j\rangle u_j, \quad x \in X.$$

**Remark 1** *As can be seen from the definition of $\varphi(H)$, we only need information about the index function $\varphi$ on the spectrum of $H$. However, there is good reason to consider this as a function defined on all of $[0, a]$. First, within the context of inverse problems, this notion emerged over a series of papers; it is now well established in its present form. Furthermore, below we shall use concavity and other functional properties which are naturally defined on (compact) intervals.*

### *2.1 Hilbert scales related to general source conditions*

Having fixed the operator $H$ and any index function $\psi$ we assign the *general source set* by

$$H_\psi := \{x \in X, \quad x = \psi(H)v, \text{ for some } v \in X, \ \|v\| \leq 1\}.$$

An element $x^\dagger$ is said to satisfy a *general source condition*, if $x^\dagger \in H_\psi$.

We mention the following

**Lemma 2 ((7, Lemma 2.8))** *If $H\colon X \to X$ is compact and $\psi$ is an index function, then the set $H_\psi$ is compact in $X$.*

As a consequence we may introduce the following scale of Hilbert spaces. We assign any index function $\psi$ the space $X_\psi^H$ which has the source set $H_\psi$ as its unit ball. By Lemma 2 the resulting space is complete and carries a natural scalar product by assigning to any $x, y \in X_\psi^H$ with (unique) source representation $x = \psi(H)u,\ y = \psi(H)v$ the value

$$\langle x, y\rangle_\psi := \langle u, v\rangle.$$

In particular an element $x$ belongs to $X_\psi^H$ if and only if $\sum_{j=1}^\infty |\langle x, u_j\rangle|^2 / \psi^2(s_j) < \infty$. We agree to denote the corresponding norm in $X_\psi^H$ by $\|\cdot\|_\psi$. We mention that in this context Lemma 2 asserts that the spaces $X_\psi^H \subset X$ are densely and compactly embedded.

**Remark 3** *We shall measure smoothness in terms of membership in such Hilbert spaces $X_\psi^H$ with index functions $\psi$. This appears to be natural within the context of regularization of inverse problems in Hilbert spaces. However, there are other ways to do so. The source condition $x^\dagger \in X_\psi^H$ controls the decay of the Fourier coefficients in $l_2$-sense, i.e., $\left(\langle x^\dagger, u_j\rangle/\psi(s_j)\right)_{j\in\mathbb{N}} \in l_2$, and this may be generalized to requiring $\left(\langle x^\dagger, u_j\rangle/\psi(s_j)\right)_{j\in\mathbb{N}} \in l_p$ for $1 \le p \le \infty$. In general the resulting related spaces will not coincide and there will be a gap which depends on the decay of the singular numbers of the underlying mapping $H$. We will delve into this at the end of § 4, below.*

### *2.2 Degree of approximation*

We study approximation by a (nested) sequence $\{X_n\}_{n\in\mathbb{N}}$ of finite dimensional subspaces of $X$, where we normalize to $\dim(X_n) = n$. We agree to call such a sequence an *approximation scheme*. In approximation theory there are various characteristics to describe the quality of an approximation scheme with respect to some smoothness class, and we introduce two of those next.

One way is to ask for the related best approximation of a given $x^\dagger \in X$ by means of elements from $X_n$, i.e., we consider the *degree of approximation*

$$E_n(x^\dagger) := \operatorname{dist}(x^\dagger, X_n) = \|(I - P_n)x^\dagger\|,$$

where $P_n$ denotes the orthogonal projection onto the space $X_n$. If $P_n$ converge point-wise to the identity $I\colon X \to X$ as $n \to \infty$, then $E_n(x^\dagger) \to 0$ for each element $x^\dagger$, and a fortiori for each compact subset $M \subset X$. Convergence may not be uniform for $\|x^\dagger\| \le 1$. But, by Lemma 2, rates of convergence can be expected uniformly for $x^\dagger \in H_\psi$, which will yield a direct Theorem. Results of such type constitute part of classical approximation theory, and we refer the reader to (9, Chapt. 4 and 5).

The approximative power of finite dimensional subspaces with respect to some subset $M \subset X$ may be measured in various ways, and we refer the reader to (21). Here we shall restrict our consideration to ellipsoids, which are obtained as images of some linear mapping in Hilbert spaces: Specifically this holds for the sets $H_\psi$:

In particular, we introduce the $n$-th Kolmogorov widths of the set $H_\psi$ in the spaces $X$ as

$$d_n(H_\psi, X) := \inf_{\dim(Z) \leq n} \sup_{x \in H_\psi} \operatorname{dist}(x, Z), \tag{2}$$

where the infimum is taken over all at most $n$-dimensional subspaces $Z \subset X$. For ellipsoids in Hilbert space these Kolmogorov widths coincide with the linear widths, given as

$$a_n(H_\psi, X) := \inf_{\operatorname{rank}(L) \leq n} \sup_{x \in H_\psi} \|x - Lx\|,$$

this time $L$ ranges among the linear mappings in $X$ with rank at most $n$.

We close with the introduction of another quantity used in classical approximation theory, the *Bernstein widths*, see the formal introduction in (19) and (21). For any $H_\psi \subset X$ we let

$$b_n(H_\psi, X) := \sup_{\dim(Z) \geq n+1} \inf_{0 \neq u \in Z \cap H_\psi} \frac{\|u\|}{\|u\|_\psi}. \tag{3}$$

**Remark 4** *For ellipsoids in Hilbert spaces all $n$-widths coincide, see (21, Chapt. IV) or (20, Chapt. 11). However, in general these $n$-widths may obey different asymptotics, and much effort was undertaken to establish precise asymptotics, and they reflect different aspects of approximation, see e.g. (21; 20; 18). Thus, within the present context, any approximation scheme $\{X_n\}_{n \in \mathbb{N}}$, which is suited for optimal linear approximation provides optimal behavior of the Bernstein widths. We postpone further discussion to § 2.4.*

As mentioned above, the $n$-widths just introduced coincide and agree with the corresponding *eigenvalues* $\lambda_{n+1}(\psi(H))$ of the mapping $\psi(H)$. Thus we state the following well known result.

**Proposition 5 ((1; 20; 21))** *Let $\psi$ be any index function. Then*

$$a_n(H_\psi, X) = d_n(H_\psi, X) = b_n(H_\psi, X) = \psi(s_{n+1}), \quad n = 0, 1, 2, \ldots$$

### *2.3 Approximate source conditions*

As second indicator we use distance functions measuring for an element $x^\dagger \in X$ the violation of a benchmark smoothness characterized by the index function $\varphi$. Having fixed $(H, \varphi)$ we assign any $x^\dagger \in X$ the distance function

$$\varrho_{x^\dagger}(t) = \varrho_{x^\dagger}^{(H,\varphi)}(t) := \operatorname{dist}(tx^\dagger, H_\varphi), \quad t > 0.$$

Of course, if $x^\dagger$ belongs to $\mathcal{R}(\varphi(H))$, the range of the operator $\varphi(H)$, then $\varrho_{x^\dagger}(t) = 0$ in a (right) neighborhood of 0, and this case is not interesting. Therefore we restrict to the complementary case. We recall the following result, similar to (7, Lemma 5.3).

**Lemma 6** *Suppose that* $x^\dagger \notin \mathcal{R}(\varphi(H))$. *Then the mapping* $t \mapsto \varrho_{x^\dagger}(t)$ *is a convex index function. Moreover, also the mapping* $t \mapsto \varrho_{x^\dagger}(t)/t$ *is an index function.*

### *2.4 Bernstein- and Jackson-type inequalities*

The following assumptions are “loosely” related to inequalities of Bernstein- and Jackson-type, where we refer to (9) for the classical context. Given an approximation scheme $\{X_n\}_{n\in\mathbb{N}}$ we agree to denote the *realized* approximability with respect to the operator $H$ by

$$\eta_n := \|H(I - P_n) \colon X \to X\|, \quad n = 1, 2, \ldots, \tag{4}$$

Typically this is known to us (up to constants). In view of the approximation numbers as introduced above we require the following

**Assumption A.1** *There is a constant* $C < \infty$ *such that*

$$\eta_n \le C s_{n+1}, \quad n = 1, 2, \ldots \tag{5}$$

This assumption requires that the subspaces are of optimal order with respect to linear approximation, since by Proposition 5 we have $s_{n+1} \le \eta_n$.

The other assumption is related to the smoothness of the elements from $X_n$, used for approximation. Within the classical context, when using trigonometric polynomials, this results in a norm bound of the derivative in terms of the degree of the polynomial, we refer to (9, Chapt. 3.2) for the Bernstein inequality in its original form. Assumptions of such type are frequently met in the analysis of projection methods for ill-posed problems in Hilbert scales, see (16) and (10), where this is called *inverse property.* Explicitly such assumptions were made in (12).

We start with the following observation. Suppose that $\kappa$ is an index function. If $\{X_n\}_{n\in\mathbb{N}}$ is an approximation scheme with $X_n \subset X_\kappa^H$, then we assign the following measure of injectivity

$$j(H_\kappa, X_n) := \inf_{0\neq u\in X_n} \frac{\|u\|}{\|u\|_\kappa}, \quad n \in \mathbb{N}.$$

By construction of the Bernstein widths from (3) we obtain

$$\kappa(s_n) = s_n(J_\kappa\colon X_\kappa^H \to X) = b_{n-1}(H_\kappa, X) \geq j(H_\kappa, X_n),\ n \in \mathbb{N}. \tag{6}$$

The assumption to be made is that the deviation is only up to a constant.

**Assumption A.2 (Bernstein-type inequality)** *Let $\kappa$ be an index function and $\{X_n\}_{n\in\mathbb{N}}$ be an approximation scheme such that $X_n \subset X_\kappa^H,\ n = 1, 2, \ldots$ The approximation scheme is said to* obey a $(H, \kappa)$-Bernstein inequality *if there is a constant $C_B \geq 1$ such that*

$$j(H_\kappa, X_n) \geq \frac{1}{C_B}\kappa(s_n), \quad n \in \mathbb{N}.$$

We will extend this to "intermediate" smoothness by appropriate interpolation, and we recall the following variant of the interpolation inequality (13, Append. A).

**Proposition 7** *Suppose that $\varphi, \psi$ and $\kappa$ are index functions arranged such that both the functions $\kappa/\varphi$ and $\kappa/\psi$ are such. If the composition*

$$t \longrightarrow \left(\frac{\kappa}{\varphi}\right)^2 \left(\left(\left(\frac{\kappa}{\psi}\right)^2\right)^{-1}(t)\right), \quad 0 < t \leq \frac{\kappa^2(a)}{\psi^2(a)}$$

*is* concave, *then*

$$\left(\frac{\kappa}{\varphi}\right)^{-1}\left(\frac{\|x\|_\varphi}{\|x\|_\kappa}\right) \leq \left(\frac{\kappa}{\psi}\right)^{-1}\left(\frac{\|x\|_\psi}{\|x\|_\kappa}\right), \quad 0 \neq x \in X_\kappa^H. \tag{7}$$

**Corollary 8** *Suppose that $\{X_n\}_{n\in\mathbb{N}}$ is an approximation scheme which obeys the $(H, \kappa)$-Bernstein inequality with constant $C_B$. If $\varphi$ is another index function for which $t \mapsto t/\varphi^2((\kappa^2)^{-1}(t))$ is a concave index function, then $X_\kappa^H \subset$*

$X_{\varphi}^{H}$, *and the approximation scheme also obeys the* $(H,\varphi)$*-Bernstein inequality with constant* $C_B$. *Precisely we have*

$$\|P_n u\|_{\varphi} \leq \frac{C_B}{\varphi(s_n)}\|P_n u\|, \quad u \in X,\ n = 1,2,\dots \tag{8}$$

**Proof:** The interpolation inequality (7) provides us with

$$\left(\frac{\kappa}{\varphi}\right)^{-1}\left(\frac{\|P_n u\|_{\varphi}}{\|P_n u\|_{\kappa}}\right) \leq \kappa^{-1}\left(\frac{\|P_n u\|}{\|P_n u\|_{\kappa}}\right), \quad 0 \neq P_n u \in X_{\kappa}^{H}.$$

Straight calculation yields

$$\frac{\|P_n u\|}{\|P_n u\|_{\varphi}} \geq \varphi\left(\kappa^{-1}\left(\frac{\|P_n u\|}{\|P_n u\|_{\kappa}}\right)\right),$$

and we need to bound the right hand side from below. To this end the assumption that $t \mapsto t/\varphi^2(t)$ is increasing implies for every $0 < c \leq 1$ that $c^2t^2/\varphi^2\left((\kappa^2)^{-1}(c^2t^2)\right) \leq t^2/\varphi^2((\kappa^2)^{-1}(t^2))$, which in turn yields

$$\varphi(\kappa^{-1}(ct)) \geq c\varphi(\kappa^{-1}(t)). \tag{9}$$

Assumption A.2 gives $\|P_n u\|/\|P_n u\|_{\kappa} \geq \kappa(s_n)/C_B$, thus, using (9) we obtain

$$\varphi\left(\kappa^{-1}\left(\frac{\|P_n u\|}{\|P_n u\|_{\kappa}}\right)\right) \geq \varphi\left(\kappa^{-1}\left(\frac{\kappa(s_n)}{C_B}\right)\right) \geq \frac{\varphi(s_n)}{C_B},$$

and the proof is complete. □

**Remark 9** *In case of monomial smoothness* $\varphi(t) = t^{\mu}$, $\kappa(t) := t^{\nu}$ *with* $\mu < \nu$ *the assumption that* $t/\varphi^2((\kappa^2)^{-1}(t))$ *concave, is automatically satisfied.*

## 3 Relating Smoothness and approximability

Clearly, if $x^{\dagger} \in H_{\psi}$ then the degree of approximation $E_n(x^{\dagger})$ of $x^{\dagger}$ by the given scheme $\{X_n\}_{n\in\mathbb{N}}$ is bounded by the supremum over all elements $x \in H_{\psi}$, hence

$$E_n(x^{\dagger}) \leq \sup_{\|x\|_{\psi}\leq 1} \operatorname{dist}(x, X_n).$$

The right-hand side above should be compared to the best possible approximation of elements $x \in H_{\psi} \subset X$, precisely with its $(n+1)$st Kolmogorov width, compare (2). The question arises whether this extends to approximation with respect to the given scheme $\{X_n\}_{n\in\mathbb{N}}$, other than some optimal. Indeed, this holds true for a variety of index functions, and we recall the following direct result from (13, Append. A, Cor. 2).

**Proposition 10** *Suppose that* $x^\dagger \in H_\psi$ *for an index function* $\psi$*, where the function* $t \mapsto \psi^2(\sqrt{t})$ *is assumed to be concave. Moreover let* $\eta_n$ *be as in (4). Then*

$$E_n(x^\dagger) \leq \|I - P_n \colon X_\psi^H \to X\| \leq \psi(\eta_n), \quad n = 1, 2, \dots$$

Therefore, to minimize this, we shall require that the given scheme $\{X_n\}_{n\in\mathbb{N}}$ is almost as good as the best possible accuracy for approximating $H$.

**Corollary 11 (Jackson-type inequality)** *Suppose that the scheme* $\{X_n\}_{n\in\mathbb{N}}$ *obeys Assumption A.1. If the function* $t \mapsto \psi^2(\sqrt{t})$ *is a concave index function and if* $x^\dagger \in H_\psi$ *then*

$$E_n(x^\dagger) \leq C\psi(s_{n+1}), \quad n = 1, 2, \dots \tag{10}$$

**Proof:** This is obtained by simple calculation as follows. Suppose that (5) holds. Then, using the concavity we obtain

$$\psi^2(\eta_n) = \psi^2(\sqrt{\eta_n^2}) \leq \psi^2(\sqrt{C^2 s_{n+1}^2}) \leq C^2\psi^2(\sqrt{s_{n+1}^2}) = C^2\psi^2(s_{n+1}).$$

Taking square roots yields the bound (10) by Proposition 10. □

**Remark 12** *The assumptions in Proposition 10 and Corollary 11 are fulfilled for the functions* $\psi(t) := t^\mu$*, whenever* $0 < \mu \leq 1$*. If this is the case then* $E_n(x^\dagger) \leq C s_{n+1}^\mu$*, provided that* $x^\dagger \in H_\psi$*.*

We turn to discussing an inverse theorem related to the degree of approximation. First we recall the following technical

**Lemma 13 (see e.g., (9, Chapt. 4.4, Lemma 1))** *Suppose that* $f \colon [a, b] \to \mathbb{R}^+$ *is a non-increasing function. Then there is a constant* $M < \infty$ *such that for every sequence* $a \leq u_k \leq u_{k+1} \leq \dots \leq u_l \leq b$ *with* $u_i/u_{i-1} \leq 2$ *it holds true that*

$$\sum_{i=k}^{l} f(u_i) \leq M \sum_{\lfloor \frac{1}{2} u_k \rfloor \leq n < u_l} \frac{f(n)}{n}.$$

The main result in this section is the following

**Theorem 14** *Suppose that $\psi$ is an index function which obeys a $\Delta_2$-condition and is a valid upper bound for the degree of approximation, i.e., $E_n(x^\dagger) \leq \psi(s_{n+1})$. Assume further that the singular values of $H$ are such that there is $1 \leq \gamma < \infty$ for which $s_n/s_{2n} \leq \gamma,\ n \in \mathbb{N}$.*

*If $\varphi$ is any index function such that*

*(1) the scheme $\{X_n\}_{n\in\mathbb{N}}$ obeys the $(H,\varphi)$-Bernstein inequality,*
*(2) the function $\psi/\varphi$ is an index function and*
*(3) the the sum*

$$\sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{\psi}{\varphi}\right)(s_n) < \infty \tag{11}$$

*is convergent,*

*then $x^\dagger \in X_\varphi^H$.*

**Proof:** Suppose that $\psi$ has the properties as stated above. We shall show that the sequence $P_{2^n}x^\dagger$ is a Cauchy sequence is $X_\psi^H$, hence convergent to $x^\dagger$. This in turn ensures $x^\dagger \in X_\varphi^H$.

Since (8) holds true for $\varphi$, we derive for every $m < n$ that

$$\begin{aligned}
\|P_{2^n}x^\dagger - P_{2^m}x^\dagger\|_\varphi &\leq \sum_{k=m}^{n-1} \|P_{2^{k+1}}x^\dagger - P_{2^k}x^\dagger\|_\varphi \\
&\leq C_B \sum_{k=m}^{n-1} \frac{1}{\varphi(s_{2^{k+1}})} \|P_{2^{k+1}}x^\dagger - P_{2^k}x^\dagger\| \\
&\leq 2C_B \sum_{k=m}^{n-1} \frac{1}{\varphi(s_{2^{k+1}})} \|(I - P_{2^k})x^\dagger\| \\
&\leq 2C_B \sum_{k=m}^{n-1} \frac{1}{\varphi(s_{2^{k+1}})} E_{2^k}(x^\dagger) \\
&\leq 2C_B \sum_{k=m}^{n-1} \frac{1}{\varphi(s_{2^{k+1}})} \psi(s_{2^k}) \\
&\leq 2C_B C_\gamma \sum_{k=m}^{n-1} \left(\frac{\psi}{\varphi}\right)(s_{2^{k+1}})
\end{aligned}$$

Now we shall apply Lemma 13 with

$$\phi(k) := \left(\frac{\psi}{\varphi}\right)(s_k), \quad 2^m \leq k \leq 2^{n-1}, \text{ and } u_i := 2^{i+1},\ i = m, \ldots, n-1.$$

This provides us with the following bound

$$\sum_{k=m}^{n-1}\left(\frac{\psi}{\varphi}\right)(s_{2^{k+1}}) \leq M \sum_{2^m \leq n < 2^{n-1}} \frac{1}{n}\left(\frac{\psi}{\varphi}\right)(s_n) \leq M \sum_{n \geq 2^m} \frac{1}{n}\left(\frac{\psi}{\varphi}\right)(s_n) \to 0,$$

by assumption (11), as $m \to \infty$. The proof is complete. $\square$ At a first glance the assumptions formulated in Theorem 14 look rather technical. Therefore it is worth-wile to see them working in the context of monomials.

**Example 15** *Suppose that the singular values of $H$ obey $s_n \asymp n^{-p}$ for some $p > 0$, and that Assumption A.2 holds true for some function $\kappa(t) := t^r$. Then this extends to the validity of (8) for each $\varphi(t) := t^\mu$, whenever $0 \leq \mu \leq r$. If the degree of approximation is bounded for $\psi(t) := t^\nu$ for some $0 < \nu \leq r$, then $x^\dagger \in X^H_{t^\mu}$ for each $0 \leq \mu < \nu$, since in this case*

$$\sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{\psi}{\varphi}\right)(s_n) = \sum_{n=1}^{\infty} n^{-1-p(\nu-\mu)} < \infty,$$

*whenever $\nu - \mu > 0$.*

## 4 Relating Smoothness and distance functions

A major direct result for this indicator was established in (7, Thm. 5.9), see also (8, Proof of Thm. 1). We recall this here as

**Proposition 16** *We suppose that $x^\dagger \in H_\psi$, and that we consider the distance function $\varrho^{(H,\varphi)}_{x^\dagger}(t)$ with respect to the benchmark index function $\varphi$. If the quotient $(\varphi/\psi)(t)$ is an index function for $0 < t \leq a$, then we can estimate*

$$\varrho_{x^\dagger}(t) \leq \varphi\left(\left(\frac{\varphi}{\psi}\right)^{-1}(t)\right) \qquad \text{for all} \quad 0 < t \leq \frac{\varphi(a)}{\psi(a)}.$$

The main inverse result is the following

**Theorem 17** *Let $x^\dagger \in X$. Assume that there is some $\varepsilon > 0$ and an index function $r(t)$, $0 \leq t \leq \varepsilon$, satisfying the inequality*

$$\varrho_{x^\dagger}(t) \leq r(t), \quad 0 \leq t \leq \varepsilon.$$

*Then there is* $j_0 \in \mathbb{N}$ *such that*

$$\left|\langle x^\dagger, u_j\rangle\right| \leq 2\frac{\varphi(s_j)}{r^{-1}(\varphi(s_j))}, \quad j \geq j_0. \tag{12}$$

**Proof:** The proof will be based on tools from convex analysis. Given a convex set $M \subset X$ we assign

$$S(y, M) := \sup\left\{\langle y, z\rangle,\ z \in M\right\},\ y \in X.$$

We recall the following identity, see e.g. (22, Chapt. 2.6, Thm. 1).

$$\operatorname{dist}(x, M) = \sup\left\{\langle x, y\rangle - S(y, M), \quad \|y\| \leq 1\right\}, \quad x \in X. \tag{13}$$

We apply this with $x := tx^\dagger$ and $M := H_\varphi$ and obtain

$$\varrho_{x^\dagger}(t) = \sup\left\{t\langle x^\dagger, y\rangle - S(y, H_\varphi), \quad \|y\| \leq 1\right\}. \tag{14}$$

In particular this yields the inequality (a specific case of the *Fenchel Young Inequality*)

$$t\langle x^\dagger, y\rangle \leq \varrho_{x^\dagger}(t) + S(y, H_\varphi), \quad \|y\| \leq 1,\ t > 0. \tag{15}$$

Since $H_\varphi$ is centrally symmetric this implies

$$\left|\langle x^\dagger, y\rangle\right| \leq \frac{1}{t}\left(\varrho_{x^\dagger}(t) + S(y, H_\varphi)\right), \quad \|y\| \leq 1,\ t > 0.$$

Now, since $r(t) \geq \varrho_{x^\dagger}(t),\ 0 < t \leq \varepsilon$, this extends to

$$\left|\langle x^\dagger, y\rangle\right| \leq \frac{1}{t}\left(r(t) + S(y, H_\varphi)\right), \quad \|y\| \leq 1,\ 0 < t \leq \varepsilon. \tag{16}$$

Let $j_0$ be the smallest index with $\varphi(s_j) \leq r(\varepsilon)$. For any $j \geq j_0$ we use the bound (16) for $y := u_j$, the $j$-th singular function of $H$, to derive

$$\left|\langle x^\dagger, u_j\rangle\right| \leq \frac{1}{t}\left(r(t) + S(u_j, H_\varphi)\right) = \frac{1}{t}\left(r(t) + \varphi(s_j)\right), \quad 0 < t \leq \varepsilon.$$

Balancing this bound with respect to $t$ yields $t^* := r^{-1}(\varphi(s_j))$ and we obtain (12). The proof is complete. □

**Remark 18** *Notice that we used the Fenchel Young inequality from (15), only. The full strength of (13) was not needed. However, the representation (14) proved to be useful in (5), as it allowed to derive lower bounds for the distance function.*

Theorem 17 does not necessarily yield the optimal smoothness of $x^\dagger$ generated by an observed decay rate $r(t) \to 0$ as $t \to 0$ of the distance function $\varrho_{x^\dagger}$. If the function $r$ is such that $t \mapsto r(t)/t$ is an index function, as should be expected by virtue of Lemma 6, then the function $\psi(t) := \varphi(t)/r^{-1}(\varphi(t))$ is an index function and hence the right hand side in (12) tends to zero as $j \to \infty$. In this case Theorem 17 asserts that $\Big(\langle x^\dagger, u_j\rangle/\psi(s_j)\Big)_{j\in\mathbb{N}} \in l_\infty$, which does not imply that $x^\dagger \in X_\psi^H$, for which it would be necessary (and sufficient) to show $\Big(\langle x^\dagger, u_j\rangle/\psi(s_j)\Big)_{j\in\mathbb{N}} \in l_2$. There is a gap, which can be verified rather clear in case of the monomial (power-type) situation as follows.

Let $\varphi(t) = t^\nu$ with some $\nu > 0$ be the benchmark function for the distance function $\varrho_{x^\dagger}$. If $x^\dagger \in X_\psi^H$ for $\psi(t) = t^\mu$ with $\mu < \nu$ then Proposition 16 yields that the distance function can be bounded by $\varrho_{x^\dagger}(t) \le t^{\nu/(\nu-\mu)}$, regardless of the behavior of the singular values of $H$.

On the other hand, if we have $\varrho_{x^\dagger}(t) \le t^{\nu/(\nu-\mu)}$ for $0 < t \le \varepsilon$, then Theorem 17 asserts that $\Big|\langle x^\dagger, u_j\rangle\Big| \le 2 s_j^\mu$ for sufficiently large integers $j$. If we now suppose that the singular values of $H$ behave like $s_j \asymp j^{-p}$ for some $p > 0$, then

$$\sum_{j=1}^{\infty} \frac{\Big|\langle x^\dagger, u_j\rangle\Big|^2}{s_j^{2\alpha}} = \sum_{j=1}^{\infty} \frac{\Big|\langle x^\dagger, u_j\rangle\Big|^2}{s_j^{2\mu}} s_j^{2p(\mu-\alpha)} \le C \sum_{j=1}^{\infty} j^{-2p(\mu-\alpha)} < \infty,$$

only if $2p(\mu - \alpha) > 1$. Thus, in this case Theorem 17 yields $x^\dagger \in X_{t^\alpha}^H$ for all $\alpha < \mu - 1/(2p)$.

However, by recent results on distance functions, see (3; 5), and by well-known converse results from regularization theory, see e.g. (17), we find from $r(t) = C\, t^{\frac{\nu}{\nu-\mu}}$, $0 < \mu < \nu < \infty$ for sufficiently small $t > 0$ and $x^\dagger \notin \mathcal{R}(H^\nu) = X_{t^\nu}^H$ that the solution smoothness obeys $x^\dagger \in X_{t^{\mu-\iota}}^H$ for arbitrarily small $\iota > 0$.

The occurring smoothness gap depends on the decay rate $s_j \asymp j^{-p}$ of the singular values of $H$; it is smaller if $p$ is larger, and it tightens as $p$ tends to infinity.

## 5 Lower bounds for distance functions

Finally we pose the following question: Given $x^\dagger \in X$, can we get information about its distance function $\varrho_{x^\dagger}(t)$ with respect to some benchmark smoothness prescribed by the index function $\varphi$, without knowing the smoothness of $x^\dagger$ relative to $H$? An answer would provide us with a further direct result complementary to Proposition 16.

One attempt would be using the identity (14). To obtain good lower bounds in this way one has to properly design elements $y$, related to $t$ and $x^\dagger$ as well as to the operator $H$. This approach was undertaken in (5), where it could be carried out successfully. However, some smart guess must be made and a careful analysis has to be done.

Here we shall propose a procedure which has its origin in the *a posteriori choice* of regularization parameters in inverse problems, and we refer to (14) for its first use in that context, and to (11) for the most recent formulation of the Lepskiĭ balancing principle. This will result in a lower bound, by just carrying out some iteration of some specific operator equation. As will be seen, by doing so we obtain an increasing function.

To be specific enough we shall exhibit this idea at an approach, related to *Landweber iteration*, because there it is most conveniently explained. So, let us choose some benchmark smoothness $\varphi(t) = t^p$ with $p > 0$ large enough and a parameter $\mu > 0$ such that $\mu\|H\| < 1$. Now fix the element $x^\dagger \in X$ and consider the sequence of $x_k$ $(k = 0, 1, 2, ...)$ of iterates

$$x_0 := \mu H x^\dagger, \tag{17}$$

$$x_k := x_{k-1} + \mu H(x^\dagger - x_{k-1}), \quad k = 1, 2, \dots . \tag{18}$$

This sequence has the following approximative property with respect to $x^\dagger$, see e.g. (7, Thm. 5.5).

$$\|x^\dagger - x_k\| \le \frac{1}{2}\left(\frac{2\gamma_p k^{-p}}{t} + \frac{2\varrho_{x^\dagger}^{(H,t^p)}(t)}{t}\right). \tag{19}$$

**Remark 19** *This bound is obtained for Landweber iteration, since, with the notation from (7), the constant $\gamma_1 = 1$ while $\gamma_p = (p/(\mu e))^p$, is the constant in the qualification of that method, see e.g. (23, Chapt. 2.2).*

Our subsequent analysis uses the terminology and results of (11). We fix $t > 0$. Then we let $\Psi(k) := 2\gamma_p k^{-p}/t,\ k = 1, 2, \dots$ This function is decreasing and it does not depend on (properties of) $x^\dagger$. Moreover, a function $\Phi(k)$ is called admissible, if together with $\Psi(k)$ it is suited for a bound like in (19). For technical reasons it must satisfy $\Phi(1) \leq \Psi(1)$. Hence the constant function $2\varrho_{x^\dagger}^{(H,t^p)}(t)/t$ is admissible, if $0 < t \leq t_0$, where $t_0$ is determined from $t_0\|x^\dagger\| + \varphi(a) \leq \gamma_p$. We now assign the positive integer

$$\bar{j} = \bar{j}(t) := \max\left\{l \in \mathbb{N} : \quad \|x_m - x_l\| \leq 4\gamma_p \frac{m^{-p}}{t}, \text{ for all } m < l\right\}. \tag{20}$$

For this choice of parameter $\bar{j}$ the following bound can be proved.

**Theorem 20** *Let the sequence $x_k$ be as in (17), (18). Given any $0 < t \leq t_0$ determine the corresponding $\bar{j}$ as in (20). Then*

$$\varrho_{x^\dagger}^{(H,t^p)}(t) \geq \gamma_p(\bar{j} + 1)^{-p}. \tag{21}$$

**Proof:** Having fixed any value $t \leq t_0$, we can apply the Lepskiĭ principle. Thus with (11, Prop. 2) (and the notation from there) we let

$$j_{**} := \max\left\{j \in \mathbb{N} : \quad \text{there is admissible } \Phi \text{ for which } \Phi(j) \leq 2\gamma_p j^{-p}/t\right\},$$

and obtain $\bar{j} \geq j_{**}$. Consequently, for the value $\bar{j} + 1$, it holds true that $2\gamma_p(\bar{j} + 1)^{-p}/t \leq \Phi(\bar{j} + 1)$ for every admissible function. In particular this is true for $2\varrho_{x^\dagger}^{(H,t^p)}(t)/t$, which in turn yields (21), and the proof is complete. ☐

The above algorithm can be carried out for any value $0 < t \leq t_0$. If this is done for a decreasing sequence then we obtain a decreasing lower bound.

**Corollary 21** *If $0 < s < t \leq t_0$ then $\bar{j}(t) \leq \bar{j}(s)$.*

**Proof:** This is clear from the construction in (20), since smaller values of $t$ yield less restrictive upper bounds. ☐

**Remark 22** *From Lemma 6 we even know that the function must be convex. So it would be nice to derive related properties for the lower bound.*

As a consequence from Corollary 21 we may proceed as follows. We design any decreasing sequence $t_0 \geq t_1 > t_2 > \ldots > t_m$. For the first value $t_1$ the algorithm yields a choice $n_1 := \bar{j}(t_1) + 1$. Then we continue for $t_2$ by checking (20) starting from $l := n_1$ to obtain $n_2 := \bar{j}(t_2) + 1$, and so forth. In this way we may lower bound the distance function $\varrho_{x^\dagger}$ at any fine grid.

## References

[1] Dz. E. Allahverdiev. On the rate of approximation of completely continuous operators by finite dimensional operators. *Azerbaidzan. Gos. Univ. Ucen. Zap. Ser.Fiz.-Mat.i Him. Nauk*, 2, 1957. Russian.

[2] A. Böttcher, B. Hofmann, U. Tautenhahn, and M. Yamamoto. Convergence rates for Tikhonov regularization from different kinds of smoothness conditions. *Appl. Anal.*, 85(5):555–578, 2006.

[3] D. Düvelmeyer, B. Hofmann, and M. Yamamoto. Range inclusions and approximate source conditions with general benchmark functions. in preparation, 2006.

[4] B. Hofmann, D. Düvelmeyer, and K. Krumbiegel. Approximate source conditions in Tikhonov regularization—new analytical results and some numerical studies. *Math. Model. Anal.*, 11(1):41–56, 2006.

[5] B. Hofmann, M. Schieck, and L. v. Wolfersdorf. On the analyis of distance functions for linear ill-posed problems with an application to the integration operator in $L^2$. submitted, 2006.

[6] B. Hofmann. Approximate source conditions in Tikhonov-Phillips regularization and consequences for inverse problems with multiplication operators. *Math. Methods Appl. Sci.*, 29(3):351–371, 2006.

[7] B. Hofmann and P. Mathé. Analysis of profile functions for general regularization methods. *SIAM J. Numer. Anal.*, 2007. to appear.

[8] B. Hofmann and M. Yamamoto. Convergence rates for Tikhonov regularization based on range inclusions. *Inverse Problems*, 21(3):805–820, 2005.

[9] G. G. Lorentz. *Approximation of functions*. Chelsea Publishing Co., New York, second edition, 1986.

[10] A. K. Louis. *Inverse und schlecht gestellte Probleme*. B. G. Teubner, Stuttgart, 1989.

[11] P. Mathé. The Lepskiĭ principle revisited. *Inverse Problems*, 22(3):L11–L15, 2006.

[12] P. Mathé and S. V. Pereverzev. Optimal discretization of inverse problems in Hilbert scales. Regularization and self-regularization of projection methods. *SIAM J. Numer. Anal.*, 38(6):1999–2021, 2001.

[13] P. Mathé and S. V. Pereverzev. Discretization strategy for linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19(6):1263–1277, 2003.

[14] P. Mathé and S. V. Pereverzev. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19(3):789–803, 2003.

[15] P. Mathé and U. Tautenhahn. Interpolation in variable Hilbert scales with application to inverse problems. *Inverse Problems*, 22(6):2271–2297, 2006.

[16] F. Natterer. Regularisierung schlecht gestellter Probleme durch Projektionsverfahren. *Numer. Math.*, 28(3):329–341, 1977.

[17] A. Neubauer. On converse and saturation results for Tikhonov regularization of linear ill-posed problems. *SIAM J. Numer. Anal.*, 34(2):517–527, 1997.

[18] A. Pietsch. *Eigenvalues and s-Numbers*, volume 43 of *Math. und ihre Anw. in Phys. und Technik*. Geest & Portig, Leipzig, 1987.

[19] A. Pietsch. $s$-numbers of operators in Banach spaces. *Studia Math.*,

51:201–223, 1974.

[20] A. Pietsch. *Operator ideals.* North-Holland Publishing Co., Amsterdam, 1980. Translated from German by the author.

[21] A. Pinkus. *n-Widths in Approximation Theory.* Springer, Berlin, Heidelberg, New York, 1985.

[22] V. M. Tichomirov. *Some Problems in Approximation Theory.* Moscow State University, Moscow, 1976.

[23] G. M. Vaĭnikko and A. Yu. Veretennikov. *Iteratsionnye protsedury v nekorrektnykh zadachakh.* "Nauka", Moscow, 1986.