lecture notes on

Inverse Problems

University of Göttingen
Summer 2002

Thorsten Hohage

Contents

1. Introduction

In [Kel76] Keller formulated the following very general definition of inverse problems, which is often cited in the literature:

> We call two problems *inverses* of one another if the formulation of each involves all or part of the solution of the other. Often, for historical reasons, one of the two problems has been studied extensively for some time, while the other is newer and not so well understood. In such cases, the former problem is called the *direct problem*, while the latter is called the *inverse problem*.

In many cases one of the two problems is not well-posed in the following sense:

Definition 1.1. (Hadamard) A problem is called *well-posed* if

1. there exists a solution to the problem (existence),
2. there is at most one solution to the problem (uniqueness),
3. the solution depends continuously on the data (stability).

A problem which is not well-posed is called *ill-posed*. If one of two problems which are inverse to each other is ill-posed, we call it the *inverse problem* and the other one the *direct problem*. All inverse problems we will consider in the following are ill-posed.

If the data space is defined as set of solutions to the direct problem, *existence* of a solution to the inverse problem is clear. However, a solution may fail to exist if the data are perturbed by noise. This problem will be addressed below. *Uniqueness* of a solution to an inverse problem is often not easy to show. Obviously, it is an important issue. If uniqueness is not guaranteed by the given data, then either additional data have to be observed or the set of admissible solutions has to be restricted using a-priori information on the solution. In other words, a remedy against non-uniqueness can be a reformulation of the problem.

Among the three Hadamard criteria, a failure to meet the third one is most delicate to deal with. In this case inevitable measurement and round-off errors can be amplified by an arbitrarily large factor and make a computed solution completely useless. Until the beginning of the last century it was generally believed that for natural problems the solution will always depend continuously on the data ('natura non facit salti'). If this was not the case, the mathematical model of the problem was believed to be inadequate. Therefore, these problems were called ill- or badly posed. Only in the second half of the last century it was realized that a huge number of problems aring in science and technology are ill-posed in any reasonable mathematical setting. This initiated a large amount of research in stable and accurate methods for the numerical solution of ill-posed problems. Today inverse and ill-posed problems are still an active area of research. This is reflected in a large number of journals ("Inverse Problems", "Inverse and Ill-Posed Problems", "Inverse Problems in Engeneering") and monographs ([BG95, Bau87, EHN96, Gla84, Gro84, Hof86, Isa98, Kir96, Kre89, Lav67, Lou89, Mor84, Mor93, TA77, TGSY95]).

Let us consider some typical examples of inverse problems. Many more can be found in the references above.

Example 1.2. (Numerical differentiation)
Differentiation and integration are two problems, which are inverse to each other. Although symbolic differentiation is much simpler than symbolic integration, we will call differentiation the inverse problem since it is ill-posed in the setting considered below. For this reason differentiation turns out to be the more delicate problem from a numerical point of view.

We define the direct problem to be the evalutation of the integral

$$(T_{\mathrm{D}}\varphi)(x) := \int_0^x \varphi(t)\,\mathrm{d}t \qquad \text{for } x \in [0,1]$$

for a given $\varphi \in C([0,1])$. The inverse problem consists in solving the equation

$$T_{\mathrm{D}}\varphi = g \tag{1.1}$$

for a given $g \in C([0,1])$ satisfying $g(0) = 0$, or equivalently computing $\varphi = g'$.

Obviously, (1.1) has a solution φ in $C([0,1])$ if and only if $g \in C^1([0,1])$. The inverse problem (1.1) would be well-posed if the data g were measured with the C^1-norm. Although this is certainly a natural setting, it is of no use if the given data contain measurement or round-off errors which can only be estimated with respect to the supremum norm.

Let us assume that we are given noisy data $g^\delta \in C([0,1])$ satisfying

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

with noise level $0 < \delta < 1$. The functions

$$g_n^\delta(x) := g(x) + \delta \sin \frac{nx}{\delta}, \qquad x \in [0,1],$$

$n = 2, 3, 4, \ldots$ satisfy this error bound, but for the derivates

$$(g_n^\delta)'(x) = g'(x) + n \cos \frac{nx}{\delta}, \qquad x \in [0,1]$$

we find that

$$\|(g_n^\delta)' - g'\|_\infty = n.$$

Hence, the error in the solutions tends blows up without bound as $n \to \infty$ although the error in the data is bounded by δ. This shows that (1.1) is ill-posed with respect to the supremum norm.

Let us look at the approximate solution of (1.1) by the central difference quotients

$$(R_h g)(x) := \frac{g(x+h) - g(x-h)}{2h}, \qquad x \in [0,1]$$

for $h > 0$. To make $(R_h g)(x)$ well defined for x near the boundaries, we assume for simplicity that g is periodic with period 1. A Taylor expansion of g yields

$$\begin{aligned} \|g' - R_h g\|_\infty &\leq \frac{h}{2}\|g''\|_\infty, \\ \|g' - R_h g\|_\infty &\leq \frac{h^2}{6}\|g'''\|_\infty \end{aligned}$$

if $g \in C^2([0,1])$ or $g \in C^3([0,1])$, respectively. For noisy data the total error can be estimated by

$$\begin{aligned} \|g' - R_h g^\delta\|_\infty &\leq \|g' - R_h g\|_\infty + \|R_h g - R_h g^\delta\|_\infty \\ &\leq \frac{h}{2}\|g''\|_\infty + \frac{\delta}{h} \end{aligned} \tag{1.2}$$

if $g \in C^2([0,1])$. In this estimate we have split the total error into an *approximation error* $\|g' - R_h g\|_\infty$, which tends to 0 as $h \to 0$, and a *data noise error* $R_h g^\delta - R_h g$, which explodes as $h \to 0$. To get a good approximation we have to balance these two error terms by a good choice of the discretization parameter h. The minimum of the right hand side of the error estimate is attained at $h = (2/\|g''\|)^{1/2}\delta^{1/2}$. With this choice of h the total error is of the order

$$\|g' - R_h g^\delta\|_\infty = \mathrm{O}\left(\delta^{1/2}\right). \tag{1.3a}$$

If $g \in C^3([0,1])$, a similar computation shows that for $h = (3/\|g'''\|)^{1/3}\delta^{1/3}$ we get the better convergence rate

$$\|g' - R_h g^\delta\|_\infty = \mathrm{O}\left(\delta^{2/3}\right). \tag{1.3b}$$

More regularity of g does not improve the order $\delta^{2/3}$ for the central difference quotient R_h. It can be shown that even for higher order difference schemes the convergence rate is always smaller than $\mathrm{O}(\delta)$. This order can only be achieved for well-posed problems.

The convergence rate (1.3a) reflects the fact that stability can be restored to the inverse problem (1.1) if the *a-priori information*

$$\|g''\|_\infty \leq E \tag{1.4}$$

is given. In other words, the restriction of the differentiaton operator to the set $W_E := \{g : \|g''\| \leq E\}$ is continuous with respect to the maximum norm. This follows from the estimate

$$\begin{aligned}
\|g_1' - g_2'\|_\infty &\leq \|(\frac{d}{dx} - R_h)(g_1 - g_2)\|_\infty + \|R_h(g_1 - g_2)\|_\infty \\
&\leq \frac{h}{2}\|g_1'' - g_2''\|_\infty + \frac{\|g_1 - g_2\|_\infty}{h} \\
&\leq 2\sqrt{E\|g_1 - g_2\|_\infty},
\end{aligned}$$

which holds for $g_1, g_2 \in W_E$ with the choice $h = \sqrt{\|g_1 - g_2\|_\infty / E}$.

In studying this first example we have seen a number of typical properties of ill-posed problems:

- amplification of high frequency errors
- dependence of ill-posedness on the choice of norms, which is often determined by practical needs
- restoration of stability by a-priori information
- a trade-off between accuracy and stability in the choice of the discretization parameter
- dependence of the optimal choice of the discretization parameter and the convergence rate on the smoothness of the solution

Example 1.3. (Backwards heat equation)

Direct Problem: Given $\varphi \in L^2([0,1])$ find $g(x) = u(x,T)$ $(T > 0)$ where $u : [0,1] \times [0,T] \to \mathbb{R}$ satisfies

$$\frac{\partial}{\partial t}u(x,t) = \Delta u(x,t), \qquad x \in (0,1), t \in (0,T), \tag{1.5a}$$

$$u(0,t) = u(1,t) = 0, \qquad t \in (0,T], \tag{1.5b}$$

$$u(x,0) = \varphi(x), \qquad x \in [0,1]. \tag{1.5c}$$

φ may describe a temperature profile at time $t = 0$. On the boundaries of the interval $[0,1]$ the temperature is kept at 0. The task is to find the temperature at time $t = T$. The inverse problem consists in finding the initial temperature given the temperature at time $t = T$.

Inverse Problem: Given $g \in L^2([0,1])$, find $\varphi \in L^2([0,1])$ such that $u(\cdot, T) = g$ and u satisfies (1.5).

Let $\varphi_n := \sqrt{2}\int_0^1 \sin(\pi n x)\varphi(x)\,\mathrm{d}x$ denote the Fourier coefficients of φ with respect to the complete orthonormal system $\{\sqrt{2}\sin(\pi n \cdot) : n = 1, 2, \dots\}$ of $L^2([0,1])$. A separation of variables leads to the formal solution

$$u(x,t) = \sqrt{2}\sum_{n=1}^{\infty} \varphi_n e^{-\pi^2 n^2 t}\sin(n\pi x) \tag{1.6}$$

of (1.5). It is a straightforward exercise to show that u defined by (1.6) belongs to $C^\infty([0,1]\times(0,T])$ and satisfies (1.5a) and (1.5b). Moreover, the initial condition (1.5c) is satisfied in the L^2 sense, i.e. $u(\cdot,t)\to\varphi$ as $t\to 0$ in $L^2([0,1])$.

Introducing the operator $T_{\mathrm{BH}}: L^2([0,1])\to L^2([0,1])$ by

$$(T_{\mathrm{BH}}\varphi)(x) := \int 2\sum_{n=1}^{\infty}\left(e^{-\pi^2n^2T}\sin(n\pi x)\sin(n\pi y)\right)\varphi(y)\,\mathrm{d}y, \tag{1.7}$$

we may formulate the inverse problem as an integral equation of the first kind

$$T_{\mathrm{BH}}\varphi = g. \tag{1.8}$$

Note that the direct solution operator T_{BH} damps out high frequency components with an exponentially descreasing factor $e^{-\pi^2n^2T}$. Therefore, in the inverse problem a data error in the nth Fourier component of g is amplified by the factor $e^{\pi^2n^2T}$! This shows that the inverse problem is severely ill-posed.

Also note that the inverse problem does not have a solution for arbitrary $g\in L^2([0,1])$. For more information on inverse problems in diffusion processes we refer to [ER95].

Example 1.4. (Sideways heat equation and deconvolution problems)
We again consider the heat equation on the interval $[0,1]$, but this time with an infinite time interval:

$$u_t(x,t) = u_{xx}(x,t), x\in(0,1), t\in\mathbb{R} \tag{1.9a}$$

The interval $[0,1]$ may describe some heat conducting medium, e.g. the wall of a furnace. We assume that the exterior side of the wall is insulated,

$$u_x(0,t) = 0, \qquad t\in\mathbb{R}, \tag{1.9b}$$

and that the interior side has the temperature of the interior of the furnace,

$$u(1,t) = \varphi(t), \qquad t\in\mathbb{R}. \tag{1.9c}$$

Direct Problem: Given $\varphi\in L^2(\mathbb{R})$ find $u:[0,1]\times\mathbb{R}\to\mathbb{R}$ satisfying (1.9).

In practice the interior of a furnace is not accessible. Temperature measurements can only be taken at the exterior side of the wall. This leads to the

Inverse Problem: Given $g\in L^2(\mathbb{R})$, find $\varphi\in L^2(\mathbb{R})$ such that the solution to (1.9) satisfies

$$u(0,t) = g(t), \qquad t\in\mathbb{R}. \tag{1.10}$$

The Fourier transform of u with respect to the time variable,

$$(\mathcal{F}u)(x,\omega) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\omega t}u(x,t)\,\mathrm{d}t$$

formally satisfies

$$\mathcal{F}u_{xx} = (\mathcal{F}u)_{xx} \qquad \text{and} \qquad \mathcal{F}u_t = i\omega\mathcal{F}u$$

for $0 < x < 1$. Hence, $\mathcal{F}u$ is characterized by the equations

$$(\mathcal{F}u)_{xx} = i\omega\mathcal{F}u, \qquad (\mathcal{F}u)_x(0,\cdot) = 0, \qquad (\mathcal{F}u)(1,\cdot) = \mathcal{F}\varphi,$$

which have the explicit solution

$$(\mathcal{F}u)(x,\omega) = \frac{\cosh\sqrt{i\omega}x}{\cosh\sqrt{i\omega}}(\mathcal{F}\varphi)(\omega)$$

where $\sqrt{i\omega} = \sqrt{\frac{\omega}{2}} + i\sqrt{\frac{\omega}{2}}$. It is easily checked that the inverse Fourier transform of the right hand side of this equation really defines a solution to (1.9). Therefore, we can reformulate our problem as an operator equation $T_{\mathrm{SH}}\varphi = g$ with the operator $T_{\mathrm{SH}}\varphi : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ defined by

$$T_{\mathrm{SH}}\varphi := \mathcal{F}^{-1}(\cosh\sqrt{i\omega})^{-1}\mathcal{F}\varphi. \tag{1.11}$$

By the Convolution Theorem T_{SH} can be written as a convolution operator

$$(T_{\mathrm{SH}}\varphi)(t) = \int_{-\infty}^{\infty} k(t-s)\varphi(s)\,\mathrm{d}s$$

with kernel $k = \sqrt{2\pi}\mathcal{F}^{-1}((\cosh\sqrt{i\omega})^{-1})$.

Integral equations of convolution type also arise in other areas of the applied sciences, e.g. in deblurring of images.

Example 1.5. (Deblurring of images)
In early 1990 the Hubble Space Telescope was launched into the low-earth orbit outside of the disturbing atmosphere in order to provide images with a unprecedented spatial resolution. Unfortunately, soon after launch a manufacturing error in the main mirror was detected, causing severe spherical aberrations in the images. Therefore, before the space shuttle Endeavour visited the telescope in 1993 to fix the error, astronomers employed inverse problem techniques to improve the blurred images (cf. [Ado95, LB91]).

The true image φ and the blurred image g are related by a first kind integral equation

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k(x,y;x',y')\varphi(x',y')\,\mathrm{d}x'\,\mathrm{d}y' = g(x,y) \tag{1.12}$$

where k is the blurring function. $k(\cdot;x_0,y_0)$ describes the blurred image of a point source at (x_0,y_0). It is usually assumed that k is spatially invariant, i.e.

$$k(x,y;x',y') = h(x-x',y-y'), \qquad x,x',y,y' \in \mathbb{R}. \tag{1.13}$$

h is called *point spread function.* Under the assumption (1.13) the direct problem is described by the convolution operator

$$(T_{DB}\varphi)(x,y) := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x-x',y-y')\varphi(x',y')\,\mathrm{d}x'\,\mathrm{d}y'.$$

The exact solution to the problem

$$T_{DB}\varphi = g \tag{1.14}$$

can in principle be computed by Fourier transformation as for the sideways heat equation. We get $\varphi = (2\pi)^{-1}\mathcal{F}^{-1}(1/\hat{h})\mathcal{F}g$. Again, the multiplication by $1/\hat{h}$ is unstable since $\hat{h} := \mathcal{F}h$ vanishes asymptotically for large arguments. Therefore the *inverse problem* to determine $\varphi \in L^2(\mathbb{R}^2)$ given $g \in L^2(\mathbb{R}^2)$ is ill-posed.

In many image restoration applications, a crucial problem is to find a suitable point spread function. For the Hubble Space Telescope it turned out that the blurring function was to some extent spatially varying, i.e. that assumption (1.13) was violated. Moreover the functions $k(\cdot; x_0, y_0)$ had a comparatively large support.

Example 1.6. (Computerized tomography)
Computerized tomography is used in medical imaging and other applications and has been studied intensively for a long time (cf. [Hel80, Nat86, RK96]). In medical X-ray tomography one tries to determine the density φ of a two-dimensional cross section of the human body by measuring the attenuation of X-rays. We assume that the support of φ is contained in the disc of radius 1. Let $I(t) = I_{\vartheta,s}(t)$ denote the intensity of an X-ray traveling in direction $\vartheta^\perp$ along the line $t \mapsto s\vartheta + t\vartheta^\perp$ where $\vartheta \in S^1 := \{x \in \mathbb{R}^2 : |x| = 1\}$ and $\vartheta^\perp := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\vartheta$ (cf. Fig. 1.1). Then I satisfies the differential equation

$$I'(t) = -\varphi(s\vartheta + t\vartheta^\perp)I(t)$$

with the explicit solution

$$I(t) = I(-\infty)\exp\left(-\int_{-\infty}^{t} \varphi(s\vartheta + \tilde{t}\vartheta^\perp)\,\mathrm{d}\tilde{t}\right).$$

If φ has compact support, then $I(t)$ is constant for $|t|$ sufficiently large. The asymptotic values $I(\pm\infty)$ of $I(t)$ are the given data. We have $-\ln\frac{I_{\vartheta,s}(\infty)}{I_{\vartheta,s}(-\infty)} = (R\varphi)(\vartheta, s)$ where the *Radon transform* is defined by

$$(R\varphi)(\vartheta, s) := \int_{-\infty}^{\infty} \varphi(s\vartheta + t\vartheta^\perp)\,\mathrm{d}t.$$

The *direct problem* consists in evaluating the Radon transform $(R\varphi)(\vartheta, s)$ for a given density distribution $\varphi \in L^2(B_1)$, $B_1 := \{x \in \mathbb{R}^2 : |x| \le 1\}$ at all $\vartheta \in S^1$ and $s \in \mathbb{R}$. The *inverse problem* is to find $\varphi \in L^2(B_1)$ given $g = R\varphi \in L^2(S^1 \times \mathbb{R})$.

In the special case that φ is radially symmetric we only need rays from one direction since $(R\varphi)(\vartheta, s)$ is constant with respect to the direction ϑ. If $\varphi(x) = \Phi(|x|^2)$, we obtain

$$(R\varphi)(\vartheta, s) = \int_{-\sqrt{1-s^2}}^{\sqrt{1-s^2}} \varphi(s\vartheta + t\vartheta^\perp)\,\mathrm{d}t = 2\int_0^{\sqrt{1-s^2}} \Phi(t^2 + s^2)\,\mathrm{d}t = \int_\sigma^1 \frac{\Phi(\tau)}{\sqrt{\tau - \sigma}}\,\mathrm{d}\tau$$

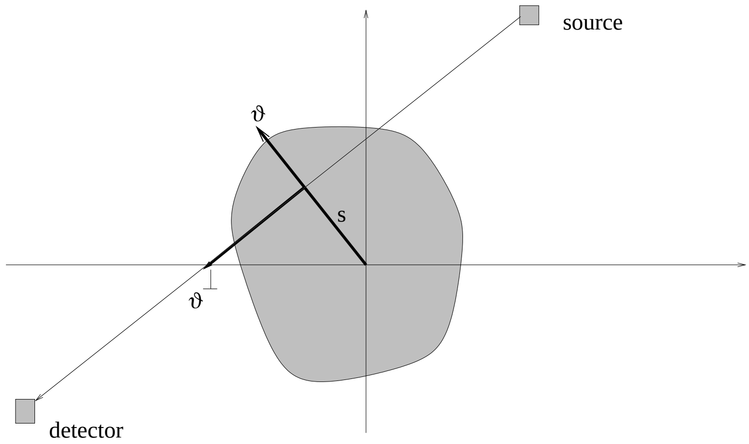

Fig. 1.1: The Radon transform

where we have used the substitutions $\tau = t^2 + s^2$ and $\sigma = s^2$. Hence our problem reduces to the *Abel integral equation*

$$(T_{\mathrm{CT}}\Phi)(\sigma) = g(\sigma), \qquad \sigma \in [0,1] \tag{1.15}$$

with the Volterra integral operator

$$(T_{\mathrm{CT}}\Phi)(\sigma) := \int_\sigma^1 \frac{\Phi(\tau)}{\sqrt{\tau - \sigma}}\,\mathrm{d}\tau. \tag{1.16}$$

Example 1.7. (Electrical Impedance Tomography, EIT)
Let $D \subset \mathbb{R}^d$, $d = 2, 3$ describe an electrically conducting medium with spatially varying conductivity $\sigma(x)$. We denote the voltage by u and assume that the electric field $E = -\operatorname{grad} u$ is stationary, i.e. $\partial_t E = 0$. By Ohm's law, the current density j satisfies $j = \sigma E = -\sigma \operatorname{grad} u$. Applying the div operator to the Maxwell equation $\operatorname{rot} H = j + \partial_t(\epsilon E)$ yields

$$\operatorname{div} \sigma \operatorname{grad} u = 0, \qquad \text{in } D. \tag{1.17a}$$

since $\operatorname{div}\operatorname{rot} = 0$. On the boundary a current distribution

$$\sigma \frac{\partial u}{\partial \nu} = I, \qquad \text{on } \partial D \tag{1.17b}$$

is prescribed. By Gauss' law (or the conservation of charge) it must satisfy

$$\int_{\partial D} I \,\mathrm{d}s = 0. \tag{1.17c}$$

Since the voltage is only determined up to a constant, we can normalize u by

$$\int_{\partial D} u \,\mathrm{d}s = 0. \tag{1.17d}$$

Direct problem: Given σ, determine the voltage $u|_{\partial D}$ on the boundary for all current distributions I satisfying (1.17c) by solving the elliptic boundary value problem (1.17a), (1.17b), (1.17d). In other words, determine the *Neumann-to-Dirichlet map* $\Lambda_\sigma : H^{-1/2}(\partial D) \to H^{1/2}(\partial D)$ defined by $\Lambda_\sigma I := u|_{\partial D}$.

Inverse problem: Given measurements of the voltage distributation $u|_{\partial D}$ for all current distributations I, i.e. given the Neumann-to-Dirichlet map Λ_σ, reconstruct σ.

This problem has been studied intensively due to numereous applications in medical imaging and nondestructive testing. Note that even though the underlying differential equation (1.17a) is linear, the mapping $\sigma \mapsto \Lambda_\sigma$ is nonlinear. Some important papers on theoretical aspects of this problem are [Ale88, Nac88, Nac95, SU87].

For $d = 1$ the Neumann-to-Dirichlet map is determined by one real number, which is certainly not enough information to reconstruct the function $\sigma(x)$. If we assume the interior measurements of u are available, i.e. that $u(x)$ is known for say $x \in [0, 1]$ in addition to $a(0)u'(0)$, then σ can be computed explicitely. Since $(\sigma u')' = 0$, we have $\sigma(x)u'(x) - \sigma(0)u'(0) = \int_0^x (\sigma u')' \, \mathrm{d}\tilde{x} = 0$, and hence

$$\sigma(x) = \frac{\sigma(0)u'(0)}{u'(x)}$$

provided $u'(x)$ vanishes nowhere. This problem is (mildly) ill-posed as it involves differentiation of the given data u (see example 1.2). If u' is small in some areas, this may cause additional instability.

We have introduced two problems where a coefficient in a partial differential equation is to be determined from (partial) knowledge of the solution of the equation. Such problems are refered to as *parameter identification problems* (cf. [BK89, Isa90, Isa98]).

Example 1.8. (Inverse obstacle scattering problems)
Another particularly important class of inverse problems are inverse scattering problems. Such problems arise in acoustics, electromagnetics, elasticity, and quantum theory. The aim is to identify properties of inaccessible objects by measuring waves scattered by them. For simplicity we only look at the acoustic case. Let $U(x, t) = \mathrm{Re}(u(x)e^{-i\omega t})$ describe the velocity potential of a time harmonic wave with frequency ω and space-dependent part u propagating through a homogeneous medium. Then the wave equation $\frac{1}{c^2}\frac{\partial^2}{\partial t^2}U = \Delta U$ reduces to the *Helmholtz equation*

$$\Delta u + k^2 u = 0 \qquad \text{in } \mathbb{R}^d \backslash K. \tag{1.18a}$$

Here $k = \omega/c$ is the wave number, and K describes an impenetrable, smooth, bounded obstacle K. The boundary condition on ∂K depends on the surface properties. For sound-soft obstacles we have the Dirichlet condition

$$u = 0 \qquad \text{on } \partial K. \tag{1.18b}$$

We assume that the total field $u = u_i + u_s$ is the superposition of an known incident field $u_i(x) = e^{ik\langle x,d\rangle}$ with direction $d \in \{x \in \mathbb{R}^2 : |x| = 1\}$ and a scattered field u_s. The scattered field satisfies the *Sommerfeld's radiation condition*

$$\lim_{r\to\infty} r^{(d-1)/2}\left(\frac{\partial u_s}{\partial r} - iku_s\right) = 0 \qquad r = |x|, \text{uniformly for all } \hat{x} = x/r, \tag{1.18c}$$

which guarantees that asymptotically energy is transported away from the origin. Moreover, this condition implies that the scattered field behaves asymptotically like an outgoing wave:

$$u_s(x) = \frac{e^{ikr}}{r^{(d-1)/2}}\left(u_\infty(\hat{x}) + \mathrm{O}\left(\frac{1}{|x|}\right)\right), \qquad |x| \to \infty \tag{1.19}$$

The function $u_\infty : S^{d-1} \to \mathbb{C}$ defined on the sphere $S^{d-1} := \{\hat{x} \in \mathbb{R}^d : |\hat{x}| = 1\}$ is called *far field pattern* or *scattering amplitude* of u_s.

> *Direct Problem:* Given a smooth bounded obstacle K and an incident wave u_i find the far-field pattern $u_\infty \in L^2(S^{d-1})$ of the scattered field satisfying (1.18).
>
> *Inverse Problem:* Given the far-field pattern $u_\infty \in L^2(S^{d-1})$ and the incident wave u_i find the obstacle K (e.g. a parametrization of its boundary ∂K).

For more information on inverse scattering problems we refer to the monographs [CK83, CK97b, Ram86] and the literature therein.

All the examples we have considered can be formulated as operator equations

$$F(\varphi) = g \tag{1.20}$$

for operators $F : X \to Y$ between normed spaces X and Y. For the linear problems 1.2, 1.3, 1.4, 1.5, and 1.6, F is given by a linear integral operator

$$(F(\varphi))(x) = (T\varphi)(x) = \int k(x,y)\varphi(y)\,\mathrm{d}y$$

with some kernel k, i.e. these problems can be formulated as *integral equations of the first kind.* For operator equations of the form (1.20) Hadamard's criteria of well-posedness in Definition 1.1 are

1. $F(X) = Y$.
2. F is one-to-one.
3. F^{-1} is continuous.

The third criterium cannot be satisfied if F is a compact operator (i.e. it maps bounded sets to relatively compact sets) and if $\dim X = \infty$. Otherwise the unit ball $B = F^{-1}(F(B))$ in X would be relatively compact, which is false.

If g in equation (1.20) does not describe *measured* data, but a *desired* state and if φ describes control parameter, one speaks of an *optimal control problem.* Such problems are related to inverse problems, but the point of view is different in some respects. First of all, one is mainly interested in convergence in the image space Y rather than the pre-image space X. Often one is not even care about uniqueness of a solution φ to (1.20). Whereas in inverse problems it is assumed that g (not g^δ!) belongs to the range of F, this is not true for optimal control problems. If the desired state g is not attainable, one tries to get as close as possible to g, i.e. (1.20) is replaced by the minimization problem

$$\|F(\varphi) - g\| = \min! \qquad \varphi \in X.$$

2. Algorithms for the solution of linear inverse problems

In this chapter we introduce methods for the stable solution of linear ill-posed operator equations

$$T\varphi = g. \tag{2.1}$$

Here $T : X \to Y$ is a bounded linear injective operator between Hilbert spaces X and Y, and $g \in R(T)$. The algorithms we are going to discuss have to take care of the fact that the solution of (2.1) does not depend continuously on the data, i.e. that $\|T^{-1}\| = \infty$ and that the data may be perturbed by noise. We assume that only noisy data g^δ are at our disposal and that

$$\|g^\delta - g\| \leq \delta. \tag{2.2}$$

Tikhonov regularization

The problem to solve (2.1) with noise data g^δ may be equivalently reformulated as finding the minimum of the functional $\varphi \mapsto \|T\varphi - g^\delta\|^2$ in X. Of course, the solution to this minimization problem again does not depend continuously on the data. One possibility to restore stability is to add a penalty term to the functional involving the distance of φ to some initial guess φ_0:

$$J_\alpha(\varphi) := \|T\varphi - g^\delta\|^2 + \alpha\|\varphi - \varphi_0\|^2$$

The parameter $\alpha > 0$ is called *regularization parameter*. If no initial guess is known, we take $\varphi_0 = 0$.

Theorem 2.1. *The Tikhonov functional J_α has a unique minimum φ_α^δ in X for all $\alpha > 0$, $g^\delta \in Y$, and $\varphi_0 \in X$. This minimum is given by*

$$\varphi_\alpha^\delta = (T^*T + \alpha I)^{-1}(T^*g^\delta + \alpha\varphi_0). \tag{2.3}$$

The operator $T^*T + \alpha I$ is boundedly invertible, so φ_α^δ depends continuously on g^δ.

To prove this theorem we need some preparations.

Definition 2.2. Let X, Y be normed spaces, and let U be an open subset of X. A mapping $F : U \to Y$ is called *Fréchet differentiable* at $\varphi \in U$ if there exists a bounded linear operator $F'[\varphi] : X \to Y$ such that

$$\lim_{h\to 0} \frac{1}{\|h\|}\left(F(\varphi + h) - F(\varphi) - F'[\varphi]h\right) = 0. \tag{2.4}$$

$F'[\varphi]$ is called the *Fréchet derivative* of F at φ. F is called *Fréchet* differentiable if it is Fréchet differentiable at every point $\varphi \in U$.

Lemma 2.3. *Let U be an open subset of a normed space X. If $F : U \to \mathbb{R}$ is Fréchet differentiable at φ and if $\varphi \in U$ is a local minimum of F, then $F'[\varphi] = 0$.*

Proof. Assume on the contrary that $F'[\varphi]h \neq 0$. After possibly changing the sign of h we may assume that $F'[\varphi]h < 0$. Then

$$\lim_{\epsilon \searrow 0} \frac{F(\varphi + \epsilon h) - F(\varphi)}{\epsilon} = F'[\varphi]h < 0.$$

This contradicts the assumption that φ is a local minimum of F. □

Lemma 2.4. *The Tikhonov functional is Fréchet differentiable for every $\alpha \geq 0$, and the Fréchet derivative is given by*

$$J'_\alpha[\varphi]h = 2 \operatorname{Re} \left\langle T^*(T\varphi - g^\delta) + \alpha(\varphi - \varphi_0), h \right\rangle .$$

Proof. The assertion follows from the identity

$$J_\alpha(\varphi + h) - J_\alpha(\varphi) - J'_\alpha[\varphi]h = \|Th\|^2 + \alpha\|h\|^2.$$

□

Proof of Theorem 2.1. Assume that φ_α^δ minimizes the Tikhonov functional J_α. Then $J'[\varphi_\alpha^\delta]h = 0$ for all $h \in X$ by Lemma 2.3 and 2.4. The choice $h = T^*(T\varphi - g^\delta) + \alpha(\varphi - \varphi_0)$ implies that

$$(T^*T + \alpha I)\varphi_\alpha^\delta = T^*g^\delta + \alpha\varphi_0.$$

The bounded invertibility of $T^*T + \alpha I$ follows from the Lax-Milgrim lemma and the inequality

$$\operatorname{Re} \langle (T^*T + \alpha I)\varphi, \varphi\rangle = \|T\varphi\|^2 + \alpha\|\varphi\|^2 \geq \alpha\|\varphi\|^2.$$

To show that φ_α^δ defined by (2.3) minimizes J_α, note that for all $h \in X \setminus \{0\}$ the function $\psi(t) := J_\alpha(\varphi_\alpha^\delta + th)$ is a polynomial of degree 2 with $\psi \geq 0$ and $\psi'(0) = 0$. Hence, $\psi(t) \geq \psi(0)$ for all $t \in \mathbb{R}$ with equality only for $t = 0$. □

Landweber iteration

Another idea is to minimize the functional $J_0(\varphi) = \|T\varphi - g^\delta\|^2$ by the steepest decent method. According to Lemma 2.4 the direction of steepest decent is $h = -T^*(T\varphi - g^\delta)$. This leads to the recursion formula

$$\begin{aligned} \varphi_0 &= 0 && (2.5a) \\ \varphi_{n+1} &= \varphi_n - \mu T^*(T\varphi_n - g^\delta), \qquad n \geq 0, && (2.5b) \end{aligned}$$

known as *Landweber iteration*. We will see later that the step size parameter μ has be chosen such that $\mu\|T^*T\| \leq 1$. It can be shown by induction that the nth Landweber iterate is given by

$$\varphi_n = \sum_{j=0}^{n-1} (I - \mu T^*T)^j \mu T^* g^\delta. \tag{2.6}$$

In fact, the equation is obviously correct for $n = 0$, and if (2.6) holds true for n, then

$$\varphi_{n+1} = (I - \mu T^*T)\varphi_n + \mu T^* g^\delta = \sum_{j=0}^{n} (I - \mu T^*T)^j \mu T^* g^\delta.$$

If some initial guess φ_0 to the solution is known, the iteration should be started at φ_0. This case can be reduced to (2.5) by introducing the data $\tilde{g}^\delta = g^\delta - T\varphi_0$. The Landweber iterates $\tilde{\varphi}_n$ corresponding to these data are related to φ_n by $\tilde{\varphi}_n = \varphi_n - \varphi_0$, $n \geq 0$.

By possibly rescaling the norm in Y, we may also assume for theoretical purposes that

$$\|T\| \leq 1. \tag{2.7}$$

Then we may set $\mu = 1$, and the formulas above become a bit simpler. In practice it is necessary to estimate the norm of $\|T^*T\|$ since the speed of convergence depends sensitively on the value $\mu\|T^*T\|$, which should not be much smaller than 1. This can be done by a few steps of the power method $\psi_{n+1} := T^*T\psi_n / \|T^*T\psi_n\|$. Usually after about 5 steps with a random starting vector ψ_0 the norm of $\|T^*T\psi_n\|$ is sufficiently close to $\|T^*T\|$.

Discrepancy principle

A crucial problem concerning Tikhonov regularization as well as other regularization methods is the choice of the regularization parameter. For Landweber iteration the number of iterations plays the role of the regularization parameter. For Tikhonov regularization with $\varphi_0 = 0$ the total error $\varphi - \varphi_\alpha^\delta$ can be decomposed as follows:

$$\begin{aligned} \varphi - \varphi_\alpha^\delta &= \varphi - (\alpha I + T^*T)^{-1} T^*T\varphi + (T^*T + \alpha I)^{-1} T^*(g - g^\delta) \\ &= \alpha(\alpha I + T^*T)^{-1}\varphi + (T^*T + \alpha I)^{-1} T^*(g - g^\delta) \end{aligned}$$

The first term on the right hand side of this identity is called the *approximation error*, and the second term is called the *data noise error*. Formally (at least if T is a non-zero multiple of the identity operator) the approximation error tends to 0 as $\alpha \to 0$. We will show later that this is true for arbitrary T and $\varphi \in N(T)^\perp$. On the other hand, the operator $(T^*T + \alpha I)^{-1}T^*$ formally tends to the unbounded operator T^{-1}, so we expect that the propagated data noise error explods as $\alpha \to 0$. This situation is illustrated in Figure 2.1. We have a trade-off between accuracy and stability: If α is too large, we get a poor approximation of the exact solution even for exact data, and if α is too small, the reconstruction becomes unstable. The optimal value of α depends both on the data noise level δ and the exact solution φ.

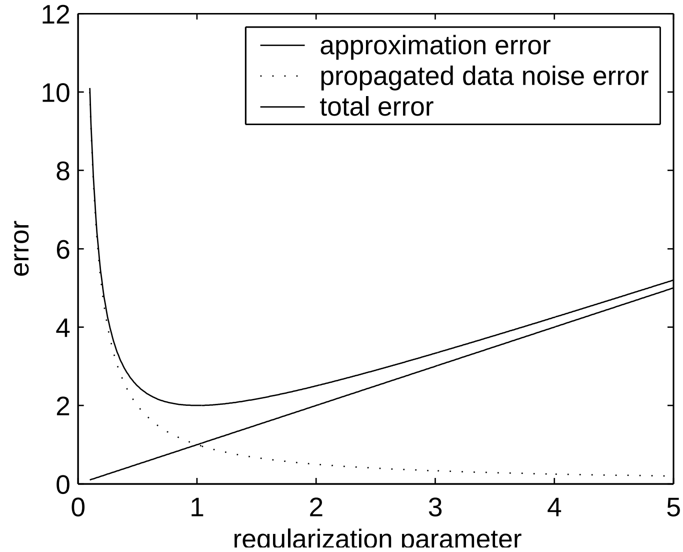

Fig. 2.1: Dependence of error terms on regularization parameter

There exist a variety of strategies for choosing the regularization parameter. The most well-known is *Morozov's discrepancy principle*. It implies that one should not try to satisfy the operator equation more accurately than the data noise error. More precisely, it consists in taking the largest regularization parameter $\alpha = \alpha(\delta, g^\delta)$ such that the residual $\|T\varphi_\alpha^\delta - g^\delta\|$ is lower or equal $\tau\delta$, where $\tau \geq 1$ is some fixed parameter:

$$\alpha(\delta, g^\delta) := \sup\{\alpha > 0 : \|T\varphi_\alpha^\delta - g^\delta\| \leq \tau\delta\} \tag{2.8}$$

We will prove later that for Tikhonov regularization and most other regularization methods the function $\alpha \mapsto \|T\varphi_\alpha^\delta - g^\delta\|$ is monotonely increasing. Usually it is sufficient to find α such that $\tau_1\delta \leq \|T\varphi_\alpha^\delta - g^\delta\| \leq \tau_2\delta$ for given constants $1 \leq \tau_1 < \tau_2$. This can be done by a simple bisection algorithm. Faster convergence can be achieved by Newton's method applied to $1/\alpha$ as unknown.

For iterative methods such as Landweber iteration, the discrepancy principle consists in stopping the iteration at the first index N for which $\|T\varphi_N^\delta - g^\delta\| \leq \tau\delta$.

Other implicit methods

Once we have computed the Tikhonov solution φ_α^δ defined by (2.3) we may find a better approximation by applying Tikhonov regularization again using φ_α^δ as initial guess φ_0. This leads to *iterated Tikhonov regularization*:

$$\varphi_{\alpha,0}^\delta := 0 \tag{2.9a}$$

$$\varphi_{\alpha,n+1}^\delta := (T^*T + \alpha I)^{-1}(T^*g^\delta + \alpha\varphi_{\alpha,n}^\delta), \qquad n \geq 0 \tag{2.9b}$$

Note that only one operator $T^*T + \alpha I$ has to be inverted to compute $\varphi^\delta_{\alpha,n}$ for any $n \in \mathbb{N}$. If we use, e.g., the LU factorization to apply $(T^*T + \alpha I)^{-1}$, the computation of $\varphi^\delta_{\alpha,n}$ for $n \geq 2$ is not much more expensive than the computation of $\varphi^\delta_{\alpha,1}$.

The following expression for $\varphi^\delta_{\alpha,n}$ is easily derived by induction:

$$\varphi^\delta_{\alpha,n} := (\alpha I + T^*T)^{-n}(T^*T)^{-1}\left((\alpha I + T^*T)^n - \alpha^n I\right) T^* g^\delta \tag{2.10}$$

Whereas in iterated Tikhonov regularization n is fixed and α is determined by some parameter choice rule (e.g. the discrepancy principle), we may also hold $\alpha > 0$ fixed and iterpret (2.9) as an iterative method. In this case one speaks of *Lardy's method*.

In his original paper on integral equations of the first kind Tikhonov [Tik63] suggested to include a derivative operator L in the penalty term in order to damp out high frequencies, i.e. to replace J_α by the functional

$$\|T\varphi - g^\delta\|^2 + \alpha\|L(\varphi - \varphi_0)\|^2. \tag{2.11}$$

Since a minimum φ^δ_α of this functional must belong to the domain of the differential operator, and since $\|L\varphi^\delta_\alpha\|$ must not be too large, the incooperation of L has a smoothing effect. In many cases this leads to a considerable improvement of the results. If $N(L) = \{0\}$, the situation can usually be reduced to the situation considered before by introducing a Hilbert space $X_L := \mathcal{D}(L) \subset X$ with the norm $\|\varphi\|_{X_L} := \|L\varphi\|_X$ [1] (cf. [Kre89, Section 16.5]). Typically, X_L is a Sobolev space. If $0 < \dim N(L) < \infty$ and $\|T\varphi\|^2 + \|L\varphi\|^2 \geq \gamma\|\varphi\|^2$ for all $\varphi \in \mathcal{D}(L)$ and some constant $\gamma > 0$, then the regularization (2.11) still works, but technical complications occur (cf. [EHN96, Chapter 8])

Other penalty terms have been investigated for special situations. E.g. if jumps of the unknown solution φ have to be detected, the smoothing effect of (2.11) is most undesirable. Much better results are obtained with the bounded variation norm, which is defined by $\|\varphi\|_{BV} = \|\varphi\|_{L^1} + \|\operatorname{grad}\varphi\|_{L^1}$ for smooth functions φ. In general, if $\varphi \in L^1(\Omega)$, $\Omega \subset \mathbb{R}^d$ is not smooth, the weak formulation

$$\|\varphi\|_{BV} := \|\varphi\|_{L^1} + \sup\left\{\int_\Omega \varphi \operatorname{div} f \,\mathrm{d}x : f \in C^1_0(\Omega), |f| \leq 1\right\} \tag{2.12}$$

has to be used. For more information on functions of bounded variation we refer to [Giu84]. The non-differentiability of the bounded variation norm causes difficulties both in the convergence analysis and the numerical implementation of this method (cf. [CK97a, NS98, VO96]).

Maximum entropy regularization is defined by

$$\|T\varphi - g^\delta\|^2 + \alpha \int_a^b \varphi(x) \log\frac{\varphi(x)}{\varphi_0(x)}\,\mathrm{d}x \tag{2.13}$$

where $\varphi_0 > 0$ is some initial guess. It is assumed that φ_0 and φ are nonnegative funtions satisfying $\int \varphi\,\mathrm{d}x = \int \varphi_0\,\mathrm{d}x = 1$ (cf. [EHN96, SG85]). In other words, φ and φ_0 can be interpreted as probability densities. Note that the penalty term is nonlinear here. For a convergence analysis of the maximum entropy method we refer to [EL93].

[1] It can be shown by results from spectral theory of unbounded operators that X_L is complete if L is self-adjoint and there exists a constant $\gamma > 0$ such that $\|L\varphi\| \geq \gamma\|\varphi\|$ for all $\varphi \in \mathcal{D}(L)$.

Other explicit methods

Implicit methods are characterized by the fact that some operator involving T has to be inverted. For many problems the application of the operator T to a vector $\varphi \in X$ can be implemented without the need to set up a matrix corresponding to T. For example, the application of convolution operators can be implemented by FFT using $O(N \log N)$ flops where N is the number of unknowns. This is much faster then a matrix-vector multiplication with $O(N^2)$ flops. Other examples include the application of inverses of elliptic differential operator by multigrid method and the application of boundary integral operators by multipole or panel-clustering methods. For these problems explicit schemes such as Landweber iteration are very attractive since they do not require setting up a matrix for T. To apply such schemes we only need routines implementing the application of the operators T and T^* to given vectors.

Note that the nth Landweber iterate belongs the *Krylov subspace* defined by

$$\mathcal{K}_n(T^*T, T^*g^\delta) := \operatorname{span}\{(T^*T)^j T^* g^\delta : j = 1, \ldots, n\}$$

Since the computation of any element of $\mathcal{K}_n(T^*T, T^*g^\delta)$ requires only (at most) n applications of T^*T, one may try to look for better approximations in the Krylov subspace $\mathcal{K}_n(T^*T, T^*g^\delta)$.

The *conjugate gradient method* applied to the normal equation $T^*T\varphi = T^*y$ is characterized by the optimality condition

$$\|T\varphi_n^\delta - g^\delta\| = \min_{\varphi \in \mathcal{K}_n(T^*T, T^*g^\delta)} \|T\varphi - g^\delta\|.$$

The algorithm is defined as follows:

$$\varphi_0^\delta = 0; \quad d_0 = g^\delta; \quad p_1 = s_0 = T^* d_0 \tag{2.14a}$$

for $k = 1, 2, \ldots,$ unless $s_{k-1} = 0$

$$q_k = T p_k \tag{2.14b}$$

$$\alpha_k = \|s_{k-1}\|^2 / \|q_k\|^2 \tag{2.14c}$$

$$\varphi_k^\delta = \varphi_{k-1}^\delta + \alpha_k p_k \tag{2.14d}$$

$$d_k = d_{k-1} - \alpha_k q_k \tag{2.14e}$$

$$s_k = T^* d_k \tag{2.14f}$$

$$\beta_k = \|s_k\|^2 / \|s_{k-1}\|^2 \tag{2.14g}$$

$$p_{k+1} = s_k + \beta_k p_k \tag{2.14h}$$

Note that φ_n^δ depends nonlinearly on g^δ. Moreover, it can be shown that the function $g^\delta \mapsto \varphi_n^\delta$ is discontinuous. Both of these facts make the analysis of the conjugate gradient method more difficult than the analysis of other methods.

Quasi solutions

A different method for the solution of general (linear and nonlinear) ill-posed problems is based on the following topological result:

Theorem 2.5. *Let K be a compact topological space, and let Y be topological space satisfying the Hausdorff separation axiom. Moreover, let $f : K \to Y$ be continuous and injective. Then $f^{-1} : f(K) \to K$ is continuous with respect to the topology of $f(K)$ induced by Y.*

Proof. We first show that f maps closed sets to closed sets. Let $A \subset K$ be closed. Since K is compact, A is compact as well. Due to the continuity of f, $f(A)$ is compact. Since Y satisfies the Hausdorff axiom, $f(A)$ is closed.

Now let $U \subset K$ be open. Then $f(K \backslash U)$ is closed by our previous argument. It follows that $f(U)$ is open in $f(K)$ since $f(U) \cup f(K \setminus U) = f(K)$ and $f(U) \cap f(K \setminus U) = \emptyset$ due to the injectivity of f. Hence, f^{-1} is continuous. □

By Theorem 2.5 the restriction of any injective, continuous operator to a compact set is boundedly invertible on its range.

Definition 2.6. Let $F : X \to Y$ be a continuous (not necessarily linear) operator and let $K \subset X$. $\varphi_K^\delta \in K$ is called a *quasi-solution* of $F(\varphi) = g^\delta$ with contraint K if

$$\|F(\varphi_K^\delta) - g^\delta\| = \inf\left\{\|F(\varphi) - g^\delta\| : \varphi \in K\right\}. \tag{2.15}$$

Obviously, a quasi-solution exists if and only if there exists a best-approximation $Q_{F(K)}g^\delta$ to g^δ in $F(K)$. If $F(K)$ is convex, then $Q_{F(K)}g^\delta$ exists for all $g^\delta \in Y$, and $Q_{F(K)}$ is continuous. If, moreover, $K \subset X$ is compact, then $\varphi_K^\delta = (F|_K)^{-1}Q_{F(X)}g^\delta$ depends continuously on g^δ by Theorem 2.5. Note that $F(K)$ is convex if F is linear and K is convex.

Since the embedding operators of function spaces with higher regularity to function spaces with lower regularity are typically compact, convex and compact subsets K of X are given by imposing a bound on the corresponding stronger norm. E.g., if $X = L^2([0,1])$, then the sets $\{\varphi \in C^1([0,1]) : \|\varphi\|_{C^1} \le R\}$ and $\{\varphi \in H_0^1([0,1]) : \|\varphi'\|_{L^2} \le R\}$ are compact in X for all $R > 0$. Note the in the latter case (2.11) with $L = d/dx$ can be interpreted as a penalty method for solving the constraint optimization problem (2.15) with $F = T$. In the former case the constraint optimization problem (2.15) is more difficult to solve.

Regularization by discretization

We have seen in Example 1.2 that inverse problems can be regularized by discretization. In fact, the restriction of T to any finite dimensional subspace $X_n \subset X$ yields a well-posed problem since linear operators on finite dimensional spaces are always bounded. However, the condition number of the finite dimensional problem may be very large unless X_n is chosen properly. Whereas for differentiation and some other important problems it is well understood how to choose finite dimensional subspaces such that the condition number can be controlled, appropriate finite dimensional subspaces are not always known a-priori, and the numerical computation of such spaces is usually too expensive.

In regularization by discretization the size of the finite dimensional subspace acts as a regularization parameter. Therefore, asymptotically the solution becomes less reliable as the discretization gets finer. We refer to [Nat77], [EHN96, Section 3.3], and [Kre89, Chapter 17] as general references.

Of course, if other regularization methods such as Tikhonov regularization are used, the regularized problems have to be discretized for a numerical implementation, and the effects of this discretization have to be investigated ([Neu88, PV90, Hoh00]). This is refered to as regularization *with* discretization as opposed to regularization *by* discretization discussed before. In the former case the choice of the discretization scheme is not as critical, and for any reasonable method the quality of the solution does not deteriorate as the discretization gets finer. In this lecture we will neither consider regularization by nor regularization with discretization.

3. Regularization methods

Let X and Y be Hilbert spaces, and let $L(X, Y)$ denote the space of linear bounded operators from X to Y. We define $L(X) := L(X, X)$. For $T \in L(X, Y)$, the null-space and the range of T are denoted by $N(T) := \{\varphi \in X : T\varphi = 0\}$ and $R(T) := T(X)$.

Orthogonal projections

Theorem 3.1. *Let U be a closed linear subspace of X. Then for each $\varphi \in X$ there exists a unique vector $\psi \in U$ satisfying*

$$\|\psi - \varphi\| = \inf_{u \in U} \|u - \varphi\|. \tag{3.1}$$

ψ is called the best approximation *to φ in U. ψ is the unique element of U satisfying*

$$\langle \varphi - \psi, u \rangle = 0 \qquad \textit{for all } u \in U. \tag{3.2}$$

Proof. We abbreviate the right hand side of (3.1) by d and choose a sequence $(u_n) \subset U$ such that

$$\|\varphi - u_n\|^2 \leq d^2 + \frac{1}{n}, \qquad n \in N. \tag{3.3}$$

Then

$$\begin{aligned} & \|(\varphi - u_n) + (\varphi - u_m)\|^2 + \|u_n - u_m\|^2 \\ = \; & 2\|\varphi - u_n\|^2 + 2\|\varphi - u_m\|^2 \leq 4d^2 + \frac{2}{n} + \frac{2}{m} \end{aligned}$$

for all $n, m \in \mathbb{N}$. Since $\frac{1}{2}(u_n + u_m) \in U$ it follows that

$$\|u_n - u_m\|^2 \leq 4d^2 + \frac{2}{n} + \frac{2}{m} - 4\left\|\varphi - \frac{1}{2}(u_n + u_m)\right\| \leq \frac{2}{n} + \frac{2}{m}.$$

This shows that (u_n) is a Cauchy sequence. Since U is complete, it has a unique limit $\psi \in U$. Passing to the limit $n \to \infty$ in (3.3) shows that ψ is a best approximation to φ. From the equality

$$\|(\varphi - \psi) + tu\|^2 = \|\varphi - \psi\|^2 + 2t \operatorname{Re} \langle \varphi - \psi, u \rangle + t^2 \|u\|^2, \tag{3.4}$$

which holds for all $u \in U$ and all $t \in \mathbb{R}$, it follows that $2|\operatorname{Re} \langle \varphi - \psi, u \rangle| \leq t\|u\|^2$ for all $t > 0$. This implies (3.2). Going back to (3.4) shows that ψ is the only best approximation to φ in U. $\square$

Remark 3.2. An inspection of the proof shows that Theorem 3.1 remains valid if U is a closed convex subset of X. In this case (3.2) has to be replaced by

$$\operatorname{Re}\langle\varphi-\psi, u-\psi\rangle \leq 0 \qquad \text{for all } u \in U. \tag{3.5}$$

Theorem 3.3. *Let $U \neq \{0\}$ be a closed linear subspace of X and let $P : X \to U$ denote the* orthogonal projection onto U*, which maps a vector $\varphi \in X$ to its best approximation in U. Then P is a linear operator with $\|P\| = 1$ satisfying*

$$P^2 = P \qquad \textit{and} \qquad P^* = P. \tag{3.6}$$

$I - P$ is the orthogonal projection onto the closed subspace $U^\perp := \{v \in X : \langle v, u\rangle = 0$ for all $u \in U\}$. Moreover,

$$X = U \oplus U^\perp \qquad \textit{and} \qquad U^{\perp\perp} = U. \tag{3.7}$$

Proof. Linearity of P follows from the characterization (3.2). Since $P\varphi = \varphi$ for $\varphi \in U$, we have $P^2 = P$ and $\|P\| \geq 1$. (3.2) with $u = P\varphi$ implies $\|\varphi\|^2 = \|P\varphi\|^2 + \|(I-P)\varphi\|^2$. Hence, $\|P\| \leq 1$. Since

$$\langle P\varphi, \psi\rangle = \langle P\varphi, P\psi + (I-P)\psi\rangle = \langle P\varphi, P\psi\rangle$$

and analogously $\langle \varphi, P\psi\rangle = \langle P\varphi, P\psi\rangle$, the operator P is self-adjoint.

To show that $U^\perp$ is closed, let (φ_n) be a convergent subsequence in $U^\perp$ with limit $\varphi \in X$. Then $\langle\varphi, u\rangle = \lim_{n\to\infty}\langle\varphi_n, u\rangle = 0$, so $\varphi \in U^\perp$. It follows from (3.2) that $(I-P)\varphi \in U^\perp$ for all $\varphi \in X$. Moreover, $\langle\varphi - (I-P)\varphi, v\rangle = \langle P\varphi, v\rangle = 0$ for all $v \in U^\perp$. By Theorem 3.1 this implies that $(I-P)\varphi$ is the best approximation to φ in $U^\perp$.

It follows immediately from the definition of $U^\perp$ that $U \cap U^\perp = \{0\}$. Moreover, $U + U^\perp = X$ since $\varphi = P\varphi + (I-P)\varphi$ for all $\varphi \in X$ with $P\varphi \in U$ and $(I-P)\varphi \in U^\perp$. Finally, $U^{\perp\perp} = R(I - (I-P)) = R(P) = U$. □

Theorem 3.4. *If $T \in L(X, Y)$ then*

$$N(T) = R(T^*)^\perp \qquad \textit{and} \qquad \overline{R(T)} = N(T^*)^\perp. \tag{3.8}$$

Proof. If $\varphi \in N(T)$, then $\langle\varphi, T^*\psi\rangle = \langle T\varphi, \psi\rangle = 0$ for all $\psi \in Y$, so $\varphi \in R(T^*)^\perp$. Hence, $N(T) \subset R(T^*)^\perp$. If $\varphi \in R(T^*)^\perp$, then $0 = \langle\varphi, T^*\psi\rangle = \langle T\varphi, \psi\rangle$ for all $\psi \in Y$. Hence $T\varphi = 0$, i.e. $\varphi \in N(T)$. This shows that $R(T)^\perp \subset N(T)$ and completes the proof of the first equality in (3.8). Interchanging the roles of T and T^* gives $N(T^*) = R(T)^\perp$. Hence, $N(T^*)^\perp = R(T)^{\perp\perp} = \overline{R(T)}$. The last equality follows from (3.7) since $R(T)^\perp = \overline{R(T)}^\perp$. □

The Moore-Penrose generalized inverse

We consider an operator equation

$$T\varphi = g \tag{3.9}$$

with an operator $T \in L(X, Y)$. Since we also want to consider optimal control problems, we neither assume that T is injective nor that $g \in R(T)$. First we have to define what is meant by a solution of (3.9) in this case. This leads to a generalization of the notion of the inverse of T.

Definition 3.5. φ is called a *least-squares solution* of (3.9) if

$$\|T\varphi - g\| = \inf\{\|T\psi - g\| : \psi \in X\}. \tag{3.10}$$

$\varphi \in X$ is called a *best approximate solution* of (3.9) if φ is a least-squares solution of (3.9) and if

$$\|\varphi\| = \inf\{\|\psi\| : \psi \text{ is least-squares solution of } T\psi = g\}. \tag{3.11}$$

Theorem 3.6. *Let $Q : Y \to \overline{R(T)}$ denote the orthogonal projection onto $\overline{R(T)}$. Then the following three statements are equivalent:*

$$\varphi \text{ is a least-squares solution to } T\varphi = g. \tag{3.12}$$

$$T\varphi = Qg \tag{3.13}$$

$$T^*T\varphi = T^*g \tag{3.14}$$

Proof. Since $\langle T\varphi - Qg, (I - Q)g\rangle = 0$ by Theorem 3.3, we have

$$\|T\varphi - g\|^2 = \|T\varphi - Qg\|^2 + \|Qg - g\|^2.$$

This shows that (3.13) implies (3.12). Vice versa, if φ_0 is a least-squares solution, the last equation shows that φ_0 is a minimum of the functional $\varphi \mapsto \|T\varphi - Qg\|$. Since $\inf_{\varphi\in X} \|T\varphi - Qg\| = 0$ by the definition of Q, φ must satisfy (3.13).

Since $N(T^*) = R(T)^\perp = R(I-Q)$ by Theorems 3.3 and 3.4, the identity $T^*(I-Q) = 0$ holds true. Hence, (3.13) implies (3.14). Vice versa, assume that (3.14) holds true. Then $T\varphi - g \in N(T^*) = R(T)^\perp$. Hence, $0 = Q(T\varphi - g) = T\varphi - Qg$. □

Equation (3.14) is called the *normal equation* of (3.9). Note that a least-squares solution may not exist since the infimum in (3.10) may not be attained. It follows from Theorem 3.6 that a least-squares solution of (3.9) exists if and only if $g \in R(T) + R(T)^\perp$.

Let $P : X \to N(T)$ denote the orthogonal projection onto the null-space $N(T)$ of T. If φ_0 is a least-squares solution to (3.9) then the set of all least-squares solution to (3.9) is given by $\{\varphi_0 + u : u \in N(T)\}$. Since

$$\|\varphi_0 + u\|^2 = \|(I - P)(\varphi_0 + u)\|^2 + \|P(\varphi_0 + u)\|^2 = \|(I - P)\varphi_0\|^2 + \|P\varphi_0 + u\|^2,$$

the best-approximate solution of (3.9) is given by $(I - P)\varphi_0$. In particular, *a best-approximate solution of (3.9) is unique* if it exists.

After these preparations we introduce the Moore-Penrose inverse as follows:

Definition 3.7. The *Moore-Penrose (generalized) inverse* $T^\dagger : \mathcal{D}(T^\dagger) \to X$ of T defined on $\mathcal{D}(T^\dagger) := R(T) + R(T)^\perp$ maps $g \in \mathcal{D}(T^\dagger)$ to the best-approximate solution of (3.9).

Obviously, $T^\dagger = T^{-1}$ if $R(T)^\perp = 0$ and $N(T) = 0$. Note that $T^\dagger g$ is not defined for all $g \in Y$ if $R(T)$ is not closed.

Let $\tilde{T} : N(T)^\perp \to R(T)$, $\tilde{T}\varphi := T\varphi$ denote the restriction of T to $N(T)^\perp$. Since $N(\tilde{T}) = N(T) \cap N(T)^\perp = \{0\}$ and $R(\tilde{T}) = R(T)$, the operator $\tilde{T}$ is invertible. By the remarks above, the Moore-Penrose inverse can be written as

$$T^\dagger g = \tilde{T}^{-1} Q g \qquad \text{for all } g \in \mathcal{D}(T^\dagger). \tag{3.15}$$

Definition and properties of regularization methods

Definition 3.8. Let $R_\alpha : Y \to X$ be a family of continuous (not necessarily linear) operators defined for α in some index set A, and let $\overline{\alpha} : (0, \infty) \times Y \to A$ be a parameter choice rule. For a given noisy data $g^\delta \in Y$ and noise level $\delta > 0$ such that $\|g^\delta - g\| \le \delta$ the exact solution $T^\dagger g$ is approximated by $R_{\overline{\alpha}(\delta, g^\delta)} g^\delta$. The pair $(R, \overline{\alpha})$ is called a *regularization method* for (3.9) if

$$\lim_{\delta \to 0} \sup \left\{ \left\| R_{\overline{\alpha}(\delta, g^\delta)} g^\delta - T^\dagger g \right\| : g^\delta \in Y, \|g^\delta - g\| \le \delta \right\} = 0 \tag{3.16}$$

for all $g \in \mathcal{D}(T^\dagger)$. $\overline{\alpha}$ is called an *a-priori parameter choice rule* if $\overline{\alpha}(\delta, g^\delta)$ depends only on δ. Otherwise $\overline{\alpha}$ is called an *a-posteriori parameter choice rule.*

In most cases the operators R_α are linear. E.g., for Tikhonov regularization we have $A = (0, \infty)$ and $R_\alpha = (\alpha I + T^*T)^{-1} T^*$. The discrepancy principle is an example of an a-posteriori parameter choice rule since the regularization parameter depends on quantities arising in the computations, which in turn depend on the observed data g^δ. An example of an a-priori parameter choice rule for Tikhonov regularization is $\overline{\alpha}(\delta, g^\delta) = \delta$. For iterative methods the number of iterations plays the role of the regularization parameter, i.e. we have $A = \mathbb{N}$. We will show later that Tikhonov regularization, Landweber iteration and other method introduced in the previous chapter together with appropriate parameter choice rules are regularization methods in the sense of Definition 3.8.

The number

$$\sup \left\{ \left\| R_{\overline{\alpha}(\delta, g^\delta)} g^\delta - T^\dagger g \right\| : g^\delta \in Y, \|g^\delta - g\| \le \delta \right\}$$

in (3.16) is the *worst case error* of the regularization method $(R, \overline{\alpha})$ for the exact data $g \in \mathcal{D}(T^\dagger)$ and the noise level δ. We require that the worst case error tends to 0 as the noise level tends to 0.

If regularization with discretization is considered, the regularization parameter is of the form $\alpha = (\tilde{\alpha}, h)$, where $\tilde{\alpha}$ is the regularization parameter of the infinite dimensional method and h is a discretization parameter.

Note that an arbitrary reconstruction method $T^\dagger g \approx S(\delta, g^\delta)$ with a (not necessarily continuous) mapping $S : (0, \infty) \times Y \to X$ can be written in the form of Definition 3.8 by setting $A = X$ and defining $R_\alpha g := \alpha$ for all $g \in Y$. This is important concerning the negative results in Theorem 3.9 and 3.11 below.

Since it is sometimes hard or impossible to obtain information on the size of the noise level in practice, parameter choice rules $\overline{\alpha}$ have been devised, which only require knowledge of the measured data g^δ, but not of the noise level δ (cf. [EHN96, Section 4.5] and the references therein). Although such so-called *error-free* parameter choice rules give good results in many cases, the following result by Bakushinskii shows that for an ill-posed problem one cannot get convergence in the sense of Definition 3.8.

Theorem 3.9. *Assume there exists a regularization method $(R_\alpha, \overline{\alpha})$ for (3.9) with a parameter choice rule $\overline{\alpha}(\delta, g^\delta)$, which depends only on g^δ, but not on δ. Then $T^\dagger$ is continuous.*

Proof. Choosing $g^\delta = g$ in (3.16) shows that $R_{\overline{\alpha}(g)}g = T^\dagger g$ for all $g \in \mathcal{D}(T^\dagger)$. Let $(g_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{D}(T^\dagger)$ which converges to $g \in \mathcal{D}(T^\dagger)$ as $n \to \infty$. Then $R_{\overline{\alpha}(g_n)}g_n = T^\dagger g_n$ for all $n \in \mathbb{N}$. Using (3.16) again with $g^\delta = g_n$ gives $0 = \lim_{n\to\infty} \|R_{\overline{\alpha}(g_n)}g_n - T^\dagger g\| = \lim_{n\to\infty} \|T^\dagger g_n - T^\dagger g\|$. This proves the assertion. □

Whereas we are considering the error as a *deterministic* quantity, it may be more appropriate in some applications to consider it as a *probabalistic* quantity. The standard method to determine the regularization parameter in this setting is *generalized cross validation* (cf. [Wah90]). Usually it is assumed in this setting that the image space Y is finite dimensional and that some information on the distribution of the error is known.

Let us consider regularization methods $(R_\alpha, \overline{\alpha})$ which satisfy the following assumption:

$$\begin{aligned} &R_\alpha : Y \to X,\ \alpha \in A \subset (0,\infty) \text{ is a family of } \textit{linear} \text{ operators and} \\ &\lim_{\delta\to 0} \sup\left\{\overline{\alpha}(\delta, g^\delta) : g^\delta \in Y, \|g^\delta - g\| \le \delta\right\} = 0. \end{aligned} \tag{3.17}$$

It follows from (3.16) and (3.17) with $g^\delta = g$ that R_α converges *pointwise* to $T^\dagger$:

$$\lim_{\alpha\to 0} R_\alpha g = T^\dagger g \qquad \text{for all } g \in \mathcal{D}(T^\dagger). \tag{3.18}$$

Theorem 3.10. *Assume that (3.17) holds true and $T^\dagger$ is unbounded. Then the operators R_α cannot be uniformly bounded with respect to α, and the operators $R_\alpha T$ cannot be norm convergent as $\alpha \to 0$.*

Proof. For the first statement, assume on the contrary that $\|R_\alpha\| \le C$ for all $\alpha \in A$. Then (3.18) implies $\|T^\dagger\| \le C$ which contradicts our assumption that $\|T^\dagger\| = \infty$.

For the second statement, assume that we have norm convergence. Then there exists $\alpha \in A$ such that $\|R_\alpha T - I\| \le 1/2$. It follows that

$$\|T^\dagger g\| \le \|T^\dagger g - R_\alpha T T^\dagger g\| + \|R_\alpha Q g\| \le \frac{1}{2}\|T^\dagger g\| + \|R_\alpha\|\|g\|$$

for all $g \in \mathcal{D}(T^\dagger)$, which implies $\|T^\dagger g\| \le 2\|R_\alpha\|\|g\|$. Again, this contradicts our assumption. □

We have required that (3.16) holds true for all $y \in \mathcal{D}(T^\dagger)$. Our next result shows that it is not possible to get a uniform convergence rate in (3.16) for all $y \in \mathcal{D}(T^\dagger)$.

Theorem 3.11. *Assume that there exist a regularization method $(R_\alpha, \overline{\alpha})$ for (3.9) and a continuous function $f : [0,\infty) \to [0,\infty)$ with $f(0) = 0$ such that*

$$\sup\left\{\left\|R_{\overline{\alpha}(\delta,g^\delta)}g^\delta - T^\dagger g\right\| : g^\delta \in Y, \|g^\delta - g\| \le \delta\right\} \le f(\delta) \tag{3.19}$$

for all $g \in \mathcal{D}(T^\dagger)$ with $\|g\| \le 1$ and all $\delta > 0$. Then $T^\dagger$ is continuous.

Proof. Let $(g_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{D}(T^\dagger)$ converging to $\tilde{g} \in \mathcal{D}(T^\dagger)$ as $n \to \infty$ such that $\|g_n\| \le 1$ for all n. With $\delta_n := \|g_n - \tilde{g}\|$ we obtain

$$\left\|T^\dagger g_n - T^\dagger \tilde{g}\right\| \le \left\|T^\dagger g_n - R_{\overline{\alpha}(\delta_n, g_n)} g_n\right\| + \left\|R_{\overline{\alpha}(\delta_n, g_n)} g_n - T^\dagger \tilde{g}\right\|.$$

The first term on the right hand side of this inequality can be estimated by setting $g = g^\delta = g_n$ in (3.19), and the second term by setting $g = \tilde{g}$ and $g^\delta = g_n$. We obtain

$$\left\|T^\dagger g_n - T^\dagger \tilde{g}\right\| \le 2f(\delta_n).$$

Since f is continuous and $f(0) = 0$, this implies that $T^\dagger g_n \to T^\dagger \tilde{g}$ as $n \to \infty$. Therefore, $T^\dagger$ is continuous at all points $\tilde{g}$ with $\|\tilde{g}\| \le 1$. This implies that $T^\dagger$ is continuous everywhere. □

Theorem 3.11 shows that for any regularization method for an ill-posed problem *convergence can be arbitrarily slow.* However, we have seen in Example 1.2 for the central difference quotient that convergence rates $f(\delta) = C\sqrt{\delta}$ or $f(\delta) = C\delta^{2/3}$ can be shown if smoothness properties of the exact solution $\varphi^\dagger = T^\dagger g$ are known a-priori. Later we will get to know generalized smoothness conditions which yield convergence rates in the general setting considered here.

4. Spectral theory

Compact operators in Hilbert spaces

Lemma 4.1. *Let $A \in L(X)$ be self-adjoint and assume that $X \neq \{0\}$. Then*

$$\|A\| = \sup_{\|\varphi\|=1} |\langle A\varphi, \varphi\rangle|. \tag{4.1}$$

Proof. Let a denote the right hand side of equation (4.1). It follows from the Cauchy-Schwarz inequality that $\|A\| \geq a$. To show that $\|A\| \leq a$ first note that

$$\|A\| = \sup_{\|\varphi\|=1} \|A\varphi\| \leq \sup_{\|\varphi\|=\|\psi\|=1} |\langle A\varphi, \psi\rangle|$$

since we may choose $\psi = A\varphi/\|A\varphi\|$. Let $\varphi, \psi \in X$ with $\|\varphi\| = \|\psi\| = 1$. We choose $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ such that $|\langle A\varphi, \psi\rangle| = \langle A\varphi, \alpha\psi\rangle$ and set $\tilde{\psi} := \alpha\psi$. Using the polarization identity and the fact that $\langle A\chi, \chi\rangle \in \mathbb{R}$ for all $\chi \in X$ we obtain

$$\begin{aligned}
|\langle A\varphi, \psi\rangle| &= \frac{1}{4}\sum_{k=0}^{3} i^k \left\langle A(\varphi + i^k\tilde{\psi}), \varphi + i^k\tilde{\psi}\right\rangle \\
&= \frac{1}{4}\left\langle A(\varphi + \tilde{\psi}), \varphi + \tilde{\psi}\right\rangle - \frac{1}{4}\left\langle A(\varphi - \tilde{\psi}), \varphi - \tilde{\psi}\right\rangle \\
&\leq \frac{a}{4}\left(\|\varphi + \tilde{\psi}\|^2 + \|\varphi - \tilde{\psi}\|^2\right) \\
&= \frac{a}{4}\left(2\|\varphi\|^2 + 2\|\tilde{\psi}\|^2\right) \\
&= a.
\end{aligned}$$

This shows that $\|A\| \leq a$. $\square$

Lemma 4.2. *Let $A \in L(X)$ be compact and self-adjoint and assume that $X \neq \{0\}$. Then there exists an eigenvalue λ of A such that $\|A\| = |\lambda|$.*

Proof. By virtue of Lemma 4.1 there exists a sequence (φ_n) with $\|\varphi_n\| = 1$ for all n such that $\langle A\varphi_n, \varphi_n\rangle \to \lambda$, $n \to \infty$ where $\lambda \in \{\|A\|, -\|A\|\}$. To prove that λ is an eigenvalue, we note that

$$\begin{aligned}
0 &\leq \|A\varphi_n - \lambda\varphi_n\| = \|A\varphi_n\|^2 - 2\lambda\langle A\varphi_n, \varphi_n\rangle + \lambda^2\|\varphi_n\|^2 \\
&\leq \|A\|^2 - 2\lambda\langle A\varphi_n, \varphi_n\rangle + \lambda^2 = 2\lambda(\lambda - \langle A\varphi_n, \varphi_n\rangle) \to 0, \qquad n \to 0,
\end{aligned}$$

so

$$A\varphi_n \to \lambda\varphi_n, \qquad n \to \infty. \tag{4.2}$$

Since A is compact, there exists a subsequence $(\varphi_{n(k)})$ such that $A\varphi_{n(k)} \to \psi$, $k \to \infty$ for some $\psi \in X$. We may assume that $A \neq 0$ since for $A = 0$ the statement of the theorem is trivial. It follows from (4.2) that $\lambda\varphi_{n(k)} \to \psi$, $k \to \infty$. Therefore, $\varphi_{n(k)}$ converges to $\varphi := \psi/\lambda$ as $k \to \infty$, and $A\varphi = \lambda\varphi$. □

Theorem 4.3. (Spectral theorem for compact self-adjoint operators) *Let $A \in L(X)$ be compact and self-adjoint. Then there exists a complete orthonormal system $E = \{\varphi_j : j \in I\}$ of X consisting of eigenvectors of A. Here I is some index set, and $A\varphi_j = \lambda_j\varphi_j$ for $j \in I$. The set $J = \{j \in I : \lambda_j \neq 0\}$ is countable, and*

$$A\varphi = \sum_{j\in J} \lambda_j \langle \varphi, \varphi_j \rangle \varphi_j \tag{4.3}$$

for all $\varphi \in X$. Moreover, for any $\epsilon > 0$ the set $J_\epsilon := \{j \in I : |\lambda_j| \geq \epsilon\}$ is finite.

Proof. By Zorn's lemma there exists a maximal orthonormal set E of eigenfunctions of A. Let U denote the closed linear span of E. Obviously, $A(U) \subset U$. Moreover, $A(U^\perp) \subset U^\perp$ since $\langle Au, \varphi\rangle = \langle u, A\varphi\rangle = 0$ for all $u \in U^\perp$ and all $\varphi \in U$. As $U^\perp$ is closed, $A|_{U^\perp}$ is compact, and of course $A|_{U^\perp}$ is self-adjoint. Hence, if $U^\perp \neq \{0\}$, there exists an eigenvector $\psi \in U^\perp$ of A due to Lemma 4.2. Since this contradicts the maximality of E, we conclude that $U^\perp = \{0\}$. Therefore $U = X$, i.e. the orthonormal system E is complete.

To show (4.3) we apply A to the Fourier representation

$$\varphi = \sum_{j\in I} \langle \varphi, \varphi_j \rangle \varphi_j \tag{4.4}$$

with respect to the Hilbert basis E.[1]

Assume that J_ϵ is infinite for some $\epsilon > 0$. Then there exists a sequence $(\varphi_n)_{n\in\mathbb{N}}$ of orthonormal eigenvectors such that $|\lambda_n| \geq \epsilon$ for all $n \in \mathbb{N}$. Since A is compact, there exists a subsequence $(\varphi_{n(k)})$ of (φ_n) such that $(A\varphi_{n(k)}) = (\lambda_{n(k)}\varphi_{n(k)})$ is a Cauchy-sequence. This is a contradiction since $\|\lambda_n\varphi_n - \lambda_m\varphi_m\|^2 = \lambda_n^2 + \lambda_m^2 \geq 2\epsilon^2$ for $m \neq n$ due to the orthonormality of the vectors φ_n. □

Theorem and Definition 4.4. (Singular value decomposition) *Let $T \in L(X, Y)$ be compact, and let $P \in L(X)$ denote the orthogonal projection onto $N(T)$. Then there exist singular values $\sigma_0 \geq \sigma_1 \geq \cdots > 0$ and orthonormal systems $\{\varphi_0, \varphi_1, \ldots\} \subset X$ and $\{g_0, g_1, \ldots\} \subset Y$ such that*

$$\varphi = \sum_{n=0}^{\infty} \langle \varphi, \varphi_n \rangle \varphi_n + P\varphi \tag{4.5}$$

[1]Recall that only a countable number of terms in (4.4) can be non-zero even if E is not countable. By Bessel's inequality, we have $\sum_{\varphi_j\in E} |\langle \varphi, \varphi_j\rangle|^2 \leq \|\varphi\|^2 < \infty$ where $\sum_{\varphi_j\in E} |\langle \varphi, \varphi_j\rangle|^2 := \sup\{\sum_{\varphi_j\in G} |\langle \varphi, \varphi_j\rangle|^2 : G \subset E, \#G < \infty\}$. Therefore, for each $n \in \mathbb{N}$ the set $S_n := \{\varphi_j \in E : |\langle \varphi, \varphi_j\rangle| \in (\frac{\|\varphi\|}{n+1}, \frac{\|\varphi\|}{n}]\}$ is finite. Hence, $S = \bigcup_{n\in\mathbb{N}} S_n = \{\varphi_j \in E : |\langle \varphi, \varphi_j\rangle| > 0\}$ is countable.

and

$$T\varphi = \sum_{n=0}^{\infty} \sigma_n \langle \varphi, \varphi_n \rangle g_n. \tag{4.6}$$

for all $\varphi \in X$*. A system* $\{(\sigma_n, \varphi_n, g_n)\}$ *with these properties is called a* singular system of T*. If* $\dim R(T) < \infty$*, the series in (4.5) and (4.6) degenerate to finite sums. The singular values* $\sigma_n = \sigma_n(T)$ *are uniquely determined by* T *and satisfy*

$$\sigma_n(T) \to 0 \qquad \textit{as } n \to \infty. \tag{4.7}$$

If $\dim R(T) < \infty$ *and* $n \geq \dim R(T)$*, we set* $\sigma_n(T) := 0$*.*

Proof. We use the notation of Theorem 4.3 with $A = T^*T$. Then $\lambda_n = \langle \lambda_n \varphi_n, \varphi_n \rangle = \|T\varphi_n\|^2 > 0$ for all n. Set

$$\sigma_n := \sqrt{\lambda_n} \qquad \text{and} \qquad g_n := T\varphi_n / \|T\varphi_n\|. \tag{4.8}$$

The vectors g_n are orthogonal since $\langle T\varphi_n, T\varphi_m \rangle = \langle T^*T\varphi_n, \varphi_m \rangle = \lambda_n \langle \varphi_n, \varphi_m \rangle = 0$ for $n \neq m$. (4.5) follows from (4.4) since $P\varphi = \sum_{j \in I \setminus J} \langle \varphi, \varphi_j \rangle \varphi_j$. Applying T to both sides of (4.5) gives (4.6) by the continuity of the scalar product. Now let $\{(\sigma_n, \varphi_n, g_n)\}$ be any singular system of T. It follows from (4.6) with $\varphi = \varphi_n$ that $T\varphi_n = \sigma_n g_n$ for all n. Moreover, since $T^* g_n \in N(T)^\perp$ and $\langle T^* g_n, \varphi_m \rangle = \langle g_n, T\varphi_m \rangle = \sigma_m \langle g_n, g_m \rangle = \sigma_n \delta_{n,m}$ for all n, m, we have

$$T^* g_n = \sigma_n \varphi_n. \tag{4.9}$$

Hence, $T^*T\varphi_n = \sigma_n^2 \varphi_n$. This shows that the singular values $\sigma_n(T)$ are uniquely determined as positive square roots of the eigenvalues of $A = T^*T$. (4.7) follows from the fact that 0 is the only possible accumulation point of the eigenvectors of A. ☐

Example 4.5. For the operator T_{BH} defined in Example 1.3 a singular system is given by

$$\begin{aligned} \sigma_n(T_{\mathrm{BH}}) &= \sqrt{2} \exp(-\pi^2 (n+1)^2 T), \\ \varphi_n(x) = g_n(x) &= \sin(\pi(n+1)x), \qquad x \in [0,1] \end{aligned}$$

for $n \in \mathbb{N}_0$.

Theorem 4.6. *The singular values of a compact operator* $T \in L(X, Y)$ *satisfy*

$$\sigma_n(T) = \inf \{ \|T - F\| : F \in L(X, Y), \dim R(F) \leq n \}, \qquad n \in \mathbb{N}_0. \tag{4.10}$$

Proof. Let $\alpha_n(T)$ denote the right hand side of (4.10). To show that $\alpha_n(T) \leq \sigma_n(T)$, we choose $F\varphi := \sum_{j=0}^{n-1} \sigma_n \langle \varphi, \varphi_n \rangle g_n$, where $(\sigma_n, \varphi_n, g_n)$ is a singular system of T. Then by Parseval's equality

$$\|T\varphi - F\varphi\|^2 = \sum_{j=n}^{\infty} \sigma_n^2 |\langle \varphi, \varphi_j \rangle|^2 \leq \sigma_n^2 \|\varphi\|^2.$$

This implies $\alpha_n(T) \leq \sigma_n(T)$. To prove the reverse inequality, let $F \in L(X,Y)$ with $\dim R(F) \leq n$ be given. Then the restriction of F to $\operatorname{span}\{\varphi_0, \ldots, \varphi_n\}$ has a non-trivial null-space. Hence, there exists $\varphi = \sum_{j=0}^{n} \alpha_n \varphi_n$ with $\|\varphi\| = 1$ and $F\varphi = 0$. It follows that

$$\|T - F\|^2 \geq \|(T-F)\varphi\|^2 = \|T\varphi\|^2 = \left\| \sum_{j=0}^{n} \sigma_j \alpha_j g_j \right\| = \sum_{j=0}^{n} \sigma_j^2 |\alpha_j|^2 \geq \sigma_n^2.$$

Hence, $\alpha_n(T) \geq \sigma_n(T)$. □

Theorem 4.7. (Picard) *Let $T \in L(X,Y)$ be compact, and let $\{(\sigma_n, \varphi_n, g_n)\}$ be a singular system of T. Then the equation*

$$T\varphi = g \tag{4.11}$$

is solvable if and only if $g \in N(T^*)^\perp$ and if the Picard criterion

$$\sum_{n=0}^{\infty} \frac{1}{\sigma_n^2} \left|\langle g, g_n \rangle\right|^2 < \infty \tag{4.12}$$

is satisfied. Then the solution is given by

$$\varphi = \sum_{n=0}^{\infty} \frac{1}{\sigma_n} \langle g, g_n \rangle \varphi_n. \tag{4.13}$$

Proof. Assume that $g \in N(T^*)^\perp$ and that (4.12) holds true. Then $g = \sum_{n=0}^{\infty} \langle g, g_n \rangle g_n$, and the series in (4.13) converges. Applying T to both sides of (4.13) gives (4.11). Vice versa, assume that (4.11) holds true. Then $g \in N(T^*)^\perp$ by Theorem 3.4. Since $\langle g, g_n \rangle = \langle T\varphi, g_n \rangle = \langle \varphi, T^* g_n \rangle = \sigma_n \langle \varphi, \varphi_n \rangle$ by (4.9), it follows that

$$\sum_{n=0}^{\infty} \frac{1}{\sigma_n^2} \left|\langle g, g_n \rangle\right|^2 = \sum_{n=0}^{\infty} \left|\langle \varphi, \varphi_n \rangle\right|^2 = \|\varphi\|^2 < \infty.$$

□

The solution formula (4.13) nicely illustrates the ill-posedness of linear operator equations with a compact operator: Since $1/\sigma_n \to \infty$ by (4.7), the large Fourier modes are amplified without bound. Typically, as in Example 4.5 the large Fourier modes correspond to high frequencies. Obviously, the faster the decay of the singular values, the more severe is the ill-posedness of the problem.

We say that (4.11) is *mildly ill-posed* if the singular values decay to 0 at a polynomial rate, i.e. if there exist constants $C, p > 0$ such that $\sigma_n \geq C n^{-p}$ for all $n \in \mathbb{N}$. Otherwise (4.11) is called *severely ill-posed.* If there exist constants $C, p > 0$ such that $\sigma_n \leq C \exp(-n^p)$, we call the problem (4.11) *exponentially ill-posed.*

One possibility to restore stability in (4.13) is to truncate the series, i.e. to compute

$$R_\alpha g := \sum_{\{n : \sigma_n \geq \alpha\}} \frac{1}{\sigma_n} \langle g, g_n \rangle \varphi_n. \tag{4.14}$$

for some regularization parameter $\alpha > 0$. This is called *truncated singular value decomposition.* A modified version is

$$R_\alpha g := \sum_{\{n:\sigma_n \geq \alpha\}} \frac{1}{\sigma_n} \langle g, g_n \rangle \varphi_n. + \sum_{\{n:\sigma_n < \alpha\}} \frac{1}{\alpha} \langle g, g_n \rangle \varphi_n. \tag{4.15}$$

This method is usually only efficient if a singular system of the operator is known explicitely since a numerical computation of the singular values and vectors is too expensive for large problems.

The spectral theorem for bounded self-adjoint operators

Our next aim is to find a generalization of Theorem 4.3 for bounded self-adjoint operators. This theorem may be reformulated as follows: We define the operator $W : l^2(I) \to X$ by $W(f) := \sum_{j\in I} f(j)\varphi_j$. Here $l^2(I)$ is the Hilbert space of all functions $f : I \to \mathbb{C}$ with the norm $\|f\|^2 := \sum_{j\in I} |f(j)|^2$. By Parseval's equality, W is a unitary operator. Its inverse is given by $(W^{-1}\varphi)(j) = \langle \varphi, \varphi_j \rangle$ for $j \in J$. Eq. (4.3) is equivalent to

$$W^*AW = M_\lambda$$

where $M_\lambda \in L(l^2(I))$ is defined by $(M_(\lambda)f)(j) = \lambda_j f(j)$, $j \in I$. In other words, there exists a unitary map W such that A is transformed to the multiplication operator M_λ on $l^2(I)$.

An important class of operators, which occured in the examples 1.4 and 1.5, are convolution operators. Let $k \in L^1(\mathbb{R}^d)$ satisfy $k(x) = \overline{k(-x)}$ for $x \in \mathbb{R}^d$. Then

$$(A\varphi)(x) := \int_{\mathbb{R}^d} k(x-y)\varphi(y)\,\mathrm{d}y$$

defines a self-adjoint operator $A \in L(X)$. Recall that the Fourier transform

$$(\mathcal{F}\varphi)(\omega) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{-i\langle\omega,x\rangle}\varphi(x)\,\mathrm{d}x, \qquad \omega \in \mathbb{R}^d$$

is unitary on $L^2(\mathbb{R}^d)$ and that the inverse operator is given by

$$(\mathcal{F}^{-1}f)(x) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{i\langle\omega,x\rangle} f(\omega)\,\mathrm{d}\omega, \qquad x \in \mathbb{R}^d.$$

Due to the assumed symmetry of k, the function $\mathcal{F}k$ is real-valued, and it is bounded since $k \in L^1(\mathbb{R}^d)$. Introducing the function $\lambda := (2\pi)^{d/2}\mathcal{F}k$ and the multiplication operator $M_\lambda \in L(L^2(\mathbb{R}^d))$, $(M_\lambda f)(\omega) := \lambda(\omega)f(\omega)$, the Convolution Theorem implies that

$$\mathcal{F}A\mathcal{F}^{-1} = M_\lambda.$$

Thus we have again found a unitary map which transforms A to a multiplication operator. The following theorem shows that this is always possible for a bounded self-adjoint operator.

Theorem 4.8 (spectral theorem for bounded self-adjoint operators)**.** *Let $A \in L(X)$ be self-adjoint. Then there exist a locally compact space Ω, a positive Borel measure μ on Ω, a unitary map*

$$W : L^2(\Omega, d\mu) \longrightarrow X, \tag{4.16}$$

and a real-valued function $\lambda \in C(\Omega)$ such that

$$W^* A W = M_\lambda, \tag{4.17}$$

where $M_\lambda \in L(L^2(\Omega, d\mu))$ is the multiplication operator defined by $(M_\lambda f)(\omega) := \lambda(\omega) f(\omega)$ for $f \in L^2(\Omega, d\mu)$ and $\omega \in \Omega$.

Note that we have already proved this theorem for all the linear examples in Chapter (1) by the remarks above. Our proof follows the presentation in [Tay96, Section 8.1] and makes use of the following special case of the representation theorem of Riesz, the proof of which can be found e.g. in [Bau90, MV92]. For other versions of the spectral theorem we refer to [HS71, Rud73, Wei76].

Theorem 4.9 (Riesz)**.** *Let Ω be a locally compact space, and let $C_0(\Omega)$ denote the space of continuous, compactly supported functions on Ω. Let $L : C_0(\Omega) \to \mathbb{R}$ be a positive linear functional, i.e. $L(f) \geq 0$ for all $f \geq 0$. Then there exists a positive Borel measure μ on Ω such that for all $f \in C_0(\Omega)$*

$$L(f) = \int f \, d\mu.$$

Lemma 4.10. *Let $A \in L(X)$ be self-adjoint. Then the initial value problem*

$$\frac{d}{dt} U(t) = iAU(t), \qquad U(0) = I \tag{4.18}$$

has a unique solution $U \in C^1(\mathbb{R}, L(X))$ given by

$$U(t) = \sum_{n=0}^{\infty} \frac{1}{n!} (itA)^n. \tag{4.19}$$

$\{U(t) : t \in \mathbb{R}\}$ is a group of unitary operators with the multiplication

$$U(s+t) = U(s)U(t). \tag{4.20}$$

Proof. The series (4.19) together with its term by term derivative converges uniformly on bounded intervals $t \in [a, b]$ with respect to the operator norm since $\sum_{n=0}^{\infty} 1/n! \, \|itA\|^n = \exp(t\|A\|) < \infty$ and $\sum_{n=1}^{\infty} n/n! \, \|iA\| \cdot \|itA\|^{n-1} = \|A\| \exp(t\|A\|) < \infty$. Hence, $t \mapsto U(t)$ is differentiable, and

$$U'(t) = iA \sum_{n=1}^{\infty} \frac{n}{n!} (itA)^{n-1} = iAU(t).$$

If $\tilde{U} \in C^1(\mathbb{R}, L(X))$ is another solution to (4.18), then $Z = U - \tilde{U}$ satisfies $Z'(t) = iAZ(t)$ and $Z(0) = 0$. Since A is self-adjoint

$$\frac{d}{dt}\|Z(t)\varphi\|^2 = 2\operatorname{Re}\langle Z'(t)\varphi, Z(t)\varphi\rangle = 2\operatorname{Re}\langle iAZ(t)\varphi, Z(t)\varphi\rangle = 0$$

for all $\varphi \in X$. Hence $Z(t) = 0$ and $U(t) = \tilde{U}(t)$ for all t.

To establish (4.20) we observe that $Z(s) := U(s+t) - U(s)U(t)$ satisfies $Z'(s) = iAZ(s)$ and $Z(0) = 0$. It follows from the argument above that $Z(s) = 0$ for all s. By the definition (4.19), $U(t)^* = U(-t)$ for all $t \in \mathbb{R}$. Now (4.20) implies that $U(t)^*U(t) = U(t)U(t)^* = I$, i.e. $U(t)$ is unitary. $\square$

For $\varphi \in X$ we call

$$X_\varphi := \overline{\operatorname{span}\{U(t)\varphi : t \in \mathbb{R}\}} \tag{4.21}$$

the *cyclic subspace generated by* φ. We say that $\varphi \in X$ is a cyclic vector of X if $X_\varphi = X$.

Lemma 4.11. *If $\{U(t) : t \in \mathbb{R}\}$ is a unitary group on a Hilbert space X, then X is an orthogonal direct sum of cyclic subspaces.*

Proof. An easy application of Zorn's lemma shows that there exists a maximal set $\{\varphi_j : j \in I\}$ of vectors $\varphi \in X$ such that the cyclic subspaces X_{φ_j} are pairwise orthogonal. Let $V := \bigoplus_{j\in I} X_{\varphi_j}$. Obviously, $U(t)V \subset V$ for all $t \in \mathbb{R}$. Assume that there exists a vector $\psi \in V^\perp$ with $\psi \neq 0$. Since for all $t \in \mathbb{R}$ and all $\varphi \in V$

$$\langle U(t)\psi, \varphi\rangle = \langle \psi, U(t)^*\varphi\rangle = \langle \psi, U(-t)\varphi\rangle = 0,$$

it follows that $X_\psi \subset V^\perp$. This contradicts the maximality of the set $\{\varphi_j : j \in I\}$. It follows that $V^\perp = \{0\}$ and $X = V$. $\square$

Lemma 4.12. *If $\{U(t) : t \in \mathbb{R}\}$ is a continuous unitary group of operators on a Hilbert space X, having a cyclic vector φ, then there exist a positive Borel measure μ on $\mathbb{R}$ and a unitary map $W : L^2(\mathbb{R}, d\mu) \to X$ such that*

$$W^*U(t)Wf(\omega) = e^{it\omega}f(\omega), \qquad \omega \in \mathbb{R} \tag{4.22}$$

for all $f \in L^2(\mathbb{R}, d\mu)$ and $t \in \mathbb{R}$.

Proof. We consider the function

$$\zeta(t) := (2\pi)^{-1/2}\langle U(t)\varphi, \varphi\rangle, \qquad t \in \mathbb{R}$$

and define the linear functional

$$L(f) := \int_{-\infty}^{\infty} \zeta(-t)\mathcal{F}f(t)\,\mathrm{d}t$$

for $f \in C_0^2(\mathbb{R})$, where $\hat{f} = \mathcal{F}f$ denotes the Fourier transform of f. $L(f)$ is well-defined since $|\zeta(t)| \le \|\varphi\|^2 < \infty$ and

$$\hat{f}(t) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-i\omega t} f(\omega)\,\mathrm{d}\omega = (2\pi)^{-1/2}(-it)^{-2} \int_{-\infty}^{\infty} e^{-i\omega t} f''(\omega)\,\mathrm{d}\omega$$

for $t \neq 0$ by partial integration, so $|(\mathcal{F}f)(t)| = \mathrm{O}\left(|t|^{-2}\right)$ as $|t| \to \infty$. Therefore, Lf is well defined for $f \in C_0^2(\mathbb{R})$. Moreover,

$$W(f) := (2\pi)^{-1/2} \int_{-\infty}^{\infty} (\mathcal{F}f)(t) U(t)\varphi\,\mathrm{d}t$$

is well defined, and

$$\begin{aligned}
\|W(f)\|^2 &= (2\pi)^{-1} \left\langle \int_{-\infty}^{\infty} \hat{f}(s) U(s)\varphi\,\mathrm{d}s, \int_{-\infty}^{\infty} \hat{f}(t) U(t)\varphi\,\mathrm{d}t \right\rangle \\
&= (2\pi)^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{f}(s)\overline{\hat{f}(t)} \left\langle U(s-t)\varphi, \varphi \right\rangle \,\mathrm{d}t\,\mathrm{d}s \\
&= (2\pi)^{-1/2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{f}(s)\overline{\hat{f}(t-s)}\,\mathrm{d}s\,\zeta(-t)\,\mathrm{d}t \qquad (4.23) \\
&= \int_{-\infty}^{\infty} \mathcal{F}(f\overline{f})(t)\,\zeta(-t)\,\mathrm{d}t \\
&= L(|f|^2).
\end{aligned}$$

This shows that $L(|f|^2) \ge 0$ for $f \in C_0^2(\mathbb{R})$.

Our next aim is to show that for each interval $[a, b] \subset \mathbb{R}$, L has a unique extension to a bounded, positive linear functional on $C_0([a, b])$. For this end we choose a smooth cut-off function $\chi \in C_0^2(\mathbb{R})$ satisfying $0 \le \chi \le 1$ and $\chi(\omega) = 1$ for $\omega \in [a, b]$. Let $g \in C_0^2([a, b])$ with $\|g\|_\infty < 1$. Then $\sqrt{\chi^2 - g} \in C_0^2(\mathbb{R})$. Hence, $L(\chi^2 - g) \ge 0$ by (4.23). Analogously, $L(\chi^2 + g) \ge 0$. Both inequalities and the linearity of L imply that $|L(g)| \le L(\chi^2)$. Therefore, L is continuous on $C^2([a, b])$ with respect to the maximum norm. To show that $L(g) \ge 0$ for $g \in C_0^2(\mathbb{R})$ with $g \ge 0$, we introduce $f_\epsilon := g/\sqrt{g + \epsilon}$ for all $\epsilon > 0$. Note that $f_\epsilon \in C_0^2(\mathbb{R})$ and that

$$\|g - f_\epsilon^2\|_\infty = \left\| \frac{g^2 + g\epsilon - g^2}{g + \epsilon} \right\|_\infty \le \epsilon.$$

Hence, $L(g) = \lim_{\epsilon \to 0} L(f_\epsilon^2) \ge 0$. As $C_0^2([a, b])$ is dense in $C_0([a, b])$, L has a unique extension to a bounded, positive linear functional on $C_0([a, b])$.

Now Theorem 4.9 implies that $L(f) = \int f\,\mathrm{d}\mu$, $f \in C_0(\mathbb{R})$ for some positive Borel measure μ on $\mathbb{R}$. Eq. (4.23) shows that W can be extended to an isometry $W : L^2(\mathbb{R}, \mu) \to X$. Assume that $\psi \in R(W)^\perp$. Then $0 = \langle \psi, W(f) \rangle = (2\pi)^{-1/2} \int \hat{f}(t) \langle \psi, U(t)\varphi \rangle \,\mathrm{d}t$ for all $f \in C_0^2(\mathbb{R})$. Hence $\langle \psi, U(t)\varphi \rangle = 0$ for all $t \in \mathbb{R}$. Since φ is assumed to be cyclic, $\psi = 0$. Therefore, $R(W)^\perp = \{0\}$, i.e. W is unitary.

For $s \in \mathbb{R}$ and $f \in C_0^2(\mathbb{R})$ we have

$$\begin{aligned} W(e^{is\cdot}f) &= \int_{-\infty}^{\infty} \hat{f}(t-s)U(t)\varphi\,\mathrm{d}t = \int_{-\infty}^{\infty} \hat{f}(t)U(t+s)\varphi\,\mathrm{d}t \\ &= U(s)\int_{-\infty}^{\infty} \hat{f}(t)U(t)\varphi\,\mathrm{d}t = U(s)W(f). \end{aligned}$$

Applying W^* to both sides of this equation gives (4.22) for $f \in C_0^2(\mathbb{R})$. Then a density argument shows that (4.22) holds true for all $f \in L^2(\mathbb{R}, \mu)$. ☐

Proof of Theorem 4.8. If X is cyclic, then we apply $-id/dt$ to (4.22) with U defined in Lemma 4.10 and obtain (4.17) with $\lambda(\omega) = \omega$ and $\Omega = \mathbb{R}$.

In general, by Lemma 4.11 we can decompose X into an orthogonal sum of cyclic subspaces, $X = \bigoplus_{j\in I} X_{\varphi_j}$. Hence, we obtain a family of Borel measures $\{\mu_j : j \in I\}$ defined on $\Omega_j = \mathbb{R}$, a family of continuous function $\lambda_j \in C(\Omega_j)$, and a familiy of unitary mappings $W_j : L^2(\Omega_j, d\mu_j) \to X_{\varphi_j}$ such that $W_j^* A W_j = M_{\lambda_j}$ for all $j \in I$. Let μ denote the sum of the measures $\{\mu_j : j \in I\}$ defined on the topological sum Ω of the spaces Ω_j. Obviously, Ω is locally compact. Moreover, $L^2(\Omega, \mu)$ can be identified with $\bigoplus_{j\in I} L^2(\Omega_j, d\mu_j)$. Defining $W : L^2(\Omega, d\mu) \to X$ as the sum of the operators W_j, and $M_\lambda \in L(L^2(\Omega, d\mu))$ as sum of the operators M_{λ_j} completes the proof. ☐

Since we have shown that all bounded self-adjoint operators can be transformed to multiplication operators, we will now look at this class of operators more closely.

Lemma 4.13. *Let μ be a Borel measure on a locally compact space Ω, and let $f \in C(\Omega)$. Then the norm of $M_f \in L(L^2(\Omega, d\mu))$ defined by $(M_f g)(\omega) := f(\omega)g(\omega)$ for $g \in L^2(\Omega, d\mu)$ and $\omega \in \Omega$ is given by*

$$\|M_f\| = \|f\|_{\infty,\,\mathrm{supp}\,\mu} \tag{4.24}$$

where $\mathrm{supp}\,\mu = \Omega \setminus \bigcup_{S\,\mathrm{open},\,\mu(S)=0} S$.

Proof. Since

$$\|M_f g\|^2 = \int_\Omega |fg|^2\,d\mu = \int_{\mathrm{supp}\,\mu} |fg|^2\,d\mu \le \|f\|^2_{\infty,\mathrm{supp}\,\mu} \int_{\mathrm{supp}\,\mu} |g|^2 d\mu = \|f\|^2_{\infty,\mathrm{supp}\,\mu} \|g\|^2$$

we have $\|M_f\| \le \|f\|_{\infty,\,\mathrm{supp}\,\mu}$. To show the reverse inequality, we may assume that $\|f\|_{\infty,\mathrm{supp}\,\mu} > 0$ and choose $0 < \epsilon < \|f\|_{\infty,\mathrm{supp}\,\mu}$. By continuity of f, $\Omega_\epsilon := (f|_{\mathrm{supp}\,\mu})^{-1}(\{\omega : |\omega| > \|f\|_{\infty,\,\mathrm{supp}\,\mu} - \epsilon\})$ is open, and $|f(\omega)| \ge \|f\|_{\infty,\,\mathrm{supp}\,\mu} - \epsilon$ for all $\omega \in \overline{\Omega_\epsilon}$. Since Ω is locally compact, there exists a compact neighborhood K to any point in Ω_ϵ, and $\mu(K) < \infty$ since μ is a Borel measure. Moreover, $\mu(K \cap \Omega_\epsilon) > 0$ as $\Omega_\epsilon \subset \mathrm{supp}\,\mu$ and the interior of $K \cap \Omega_\epsilon$ is not empty. Hence, the characteristic function g of $K \cap \Omega_\epsilon$ belongs to $L^2(\Omega, d\mu)$, $g \ne 0$, and $\|M_f g\| \ge (\|f\|_{\infty\ \mathrm{supp}\,\mu} - \epsilon)\|g\|$. Therefore, $\|M_f\| \ge \|f\|_{\infty,\,\mathrm{supp}\,\mu} - \epsilon$ for any $\epsilon > 0$. This completes the proof of (4.24). ☐

Next we define the spectrum of an operator, which is a generalization of the set of eigenvalues of a matrix.

Definition 4.14. Let X be a Banach space and $A \in L(X)$. The *resolvent set* $\rho(A)$ of A is the set of all $\tilde{\lambda} \in \mathbb{C}$ for which $N(\tilde{\lambda}I - A) = \{0\}$, $R(\tilde{\lambda}I - A) = X$, and $(\tilde{\lambda}I - A)$ is boundedly invertible. The *spectrum of* A is defined as $\sigma(A) := \mathbb{C} \setminus \rho(A)$.

It follows immediately from the definition that the spectrum is invariant under unitary transformations. In particular, $\sigma(A) = \sigma(M_\lambda)$ with the notation of Theorem 4.8. By (4.24), $(\tilde{\lambda}I - M_\lambda)^{-1}$ exists and is bounded if and only if the function $\omega \mapsto (\tilde{\lambda} - \lambda(\omega))^{-1}$ is well-defined for μ-almost all $\omega \in \Omega$ and if $\|(\tilde{\lambda} - \lambda)^{-1}\|_{\infty, \operatorname{supp} \mu} < \infty$. Since λ is continuous, the latter condition is equivalent to $\tilde{\lambda} \notin \overline{\lambda(\operatorname{supp} \mu)}$. In other words, $\rho(M_\lambda) = \mathbb{C} \setminus \overline{\lambda(\operatorname{supp} \mu)}$ or

$$\sigma(A) = \overline{\lambda(\operatorname{supp} \mu)}. \tag{4.25}$$

It follows from (4.24) and (4.25) that $\sigma(A)$ is closed and bounded and hence compact.

If $p(\lambda) = \sum_{j=0}^{n} p_j \lambda^j$ is a polynomial, it is natural to define

$$p(A) := \sum_{j=0}^{n} p_j A^j. \tag{4.26}$$

The next theorem generalizes this definition to continuous functions on $\sigma(A)$.

Theorem 4.15 (functional calculus)**.** *With the notation of Theorem 4.8 define*

$$f(A) := W M_{f \circ \lambda} W^* \tag{4.27}$$

for a real-valued function $f \in C(\sigma(A))$. Here $(f \circ \lambda)(\omega) := f(\lambda(\omega))$. Then $f(A) \in L(X)$ is self-adjoint and satisfies (4.26) if f is a polynomial. The mapping $f \mapsto f(A)$, which is called the functional calculus at A*, is an isometric algebra homomorphism from $C(\sigma(A))$ to $L(X)$, i.e. for $f, g \in C(\sigma(A))$ and $\alpha, \beta \in \mathbb{R}$ we have*

$$\begin{aligned} (\alpha f + \beta g)(A) &= \alpha f(A) + \beta g(A), && (4.28a)\\ (fg)(A) &= f(A) g(A), && (4.28b)\\ \|f(A)\| &= \|f\|_\infty. && (4.28c) \end{aligned}$$

The functional calculus is uniquely determined by the properties (4.26) and (4.28).

Proof. By (4.24) and (4.25), $f(A)$ is well-defined and bounded. It is self-adjoint since f is real-valued. For a polynomial p we have $p(A) = W p(M_\lambda) W^*$ with the definition (4.26). Since $p(M_\lambda) = M_{p \circ \lambda}$, this coincides with definition (4.27). The proof of (4.28a) is straightforward. To show (4.28b), we write

$$f(A) g(A) = W M_{f \circ \lambda} W^* W M_{g \circ \lambda} W^* = W M_{(fg) \circ \lambda} W^* = (fg)(A).$$

Finally, using (4.24), (4.25), and the continuity of f,

$$\|f(A)\| = \|M_{f \circ \lambda}\| = \|f \circ \lambda\|_{\infty, \operatorname{supp} \mu} = \|f\|_{\infty, \sigma(A)}.$$

Let $\Phi_A : C(\sigma(A)) \to L(X)$ be any isometric algebra homomorphism satisfying $\Phi_A(p) = p(A)$ for all polynomials. By the Weierstraß Approximation Theorem, for any $f \in C(\sigma(A))$ there exists a sequence of polynomials (p_n) such that $\|f - p_n\|_{\infty,\sigma(A)} \to 0$ as $n \to \infty$. Using (4.28c) we obtain $\Phi_A(f) = \lim_{n\to\infty} \Phi_A(p_n) = \lim_{n\to\infty} p_n(A) = f(A)$. Therefore, the functional calculus is uniquely determined by (4.26) and (4.28). □

Theorem 4.15 will be a powerful tool for the convergence analysis in the next section as it allows to reduce the estimation of operator norms to the estimation of functions defined on an interval. We will also need the following

Lemma 4.16. *If $T \in L(X,Y)$ and $f \in C([0, \|T^*T\|])$, then*

$$Tf(T^*T) = f(TT^*)T. \tag{4.29}$$

Proof. It is obvious that (4.29) holds true if f is a polynomial. By the Weierstraß Approximation Theorem, for any $f \in C([0, \|T^*T\|])$ there exists a sequence (p_n) of polynomials such that $\|f - p_n\|_{\infty,[0,\|T^*T\|]} \to 0$ as $n \to \infty$. Hence, $Tf(T^*T) = \lim_{n\to\infty} Tp_n(T^*T) = \lim_{n\to\infty} p_n(TT^*)T = f(TT^*)T$ by virtue of (4.28c). □

Finally we show that the functional calculus can be extended to the algebra $\mathcal{M}(\sigma(A))$ of bounded, Borel-measureable functions on $\sigma(A)$ with the norm $\|f\|_\infty := \sup_{t\in\sigma(A)} |f(t)|$.

Theorem 4.17. *The mapping $f \mapsto f(A) := WM_{f\circ\lambda}W^*$ is a norm-decreasing algebra homomorphism from $\mathcal{M}(\sigma(A))$ to $L(X)$, i.e. for $f, g \in \mathcal{M}(\sigma(A))$ and $\alpha, \beta \in R$ we have*

$$(\alpha f + \beta g)(A) = \alpha f(A) + \beta g(A), \tag{4.30a}$$

$$(fg)(A) = f(A)g(A), \tag{4.30b}$$

$$\|f(A)\| \leq \|f\|_\infty. \tag{4.30c}$$

If (f_n) is a sequence in $\mathcal{M}(\sigma(A))$ converging pointwise to a function $f \in \mathcal{M}(\sigma(A))$ such that $\sup_{n\in\mathbb{N}} \|f_n\| < \infty$*, then*

$$\|f_n(A)\varphi - f(A)\varphi\| \to 0 \qquad \text{as } n \to \infty \tag{4.31}$$

for all $\varphi \in X$.

Proof. (4.30a) and (4.30b) are shown in the same way as (4.28a) and (4.28b). For the proof of (4.30c) we estimate $\|f(A)\| = \|M_{f\circ\lambda}\| \leq \|f\|_\infty$. To prove (4.31), we define $g := W^*\varphi$ and $C := \sup_{n\in\mathbb{N}} \|f_n\|$. Since the functions $|f_n \circ \lambda - f \circ \lambda|^2|g|^2$ converge pointwise to 0 and since they are dominated by the integrable function $4C^2|g|^2$, Lebesgue's Dominated Convergence Theorem implies that

$$\|f_n(A)\varphi - f(A)\varphi\|^2 = \|M_{f_n\circ\lambda}g - M_{f\circ\lambda}g\|^2 = \int |f_n(\lambda(\omega)) - f(\lambda(\omega))|^2|g(\omega)|^2 d\mu \to 0$$

as $n \to \infty$. □

5. Convergence analysis of linear regularization methods

We consider an operator equation

$$T\varphi = g \tag{5.1}$$

where $T \in L(X,Y)$ and X and Y are Hilbert spaces. We assume that $g \in \mathcal{D}(T^\dagger)$ and that the data $g^\delta \in Y$ satisfy

$$\|g - g^\delta\| \leq \delta. \tag{5.2}$$

In this chapter we are going to consider the convergence of regularization methods of the form

$$R_\alpha g^\delta := q_\alpha(T^*T)T^*g^\delta \tag{5.3}$$

with some functions $q_\alpha \in C([0, \|T^*T\|]))$ depending on some regularization parameter $\alpha > 0$. We denote the reconstructions for exact and noisy data by $\varphi_\alpha := R_\alpha g$ and $\varphi_\alpha^\delta := R_\alpha g^\delta$, respectively and use the symbol $\varphi^\dagger := T^\dagger g$ for the exact solution. Since $T^*g = T^*Qg = T^*T\varphi^\dagger$, the reconstruction error for exact data is given by

$$\varphi^\dagger - \varphi_\alpha = (I - q_\alpha(T^*T)T^*T)\varphi^\dagger = r_\alpha(T^*T)\varphi^\dagger \tag{5.4}$$

with

$$r_\alpha(\lambda) := 1 - \lambda q_\alpha(\lambda), \qquad \lambda \in [0, \|T^*T\|]. \tag{5.5}$$

The following table lists the functions q_α and r_α for some regularization methods.

Tikhonov regularization	$q_\alpha(\lambda) = \frac{1}{\lambda+\alpha}$	$r_\alpha(\lambda) = \frac{\alpha}{\lambda+\alpha}$
iterated Tikhonov regularization (2.9), $n \geq 1$	$q_\alpha(\lambda) = \frac{(\lambda+\alpha)^n - \alpha^n}{\lambda(\lambda+\alpha)^n}$	$r_\alpha(\lambda) = \left(\frac{\alpha}{\lambda+\alpha}\right)^n$
truncated singular value decomposition (4.14)	$q_\alpha(\lambda) = \begin{cases} \lambda^{-1}, & \lambda \geq \alpha \\ 0, & \lambda < \alpha \end{cases}$	$r_\alpha(\lambda) = \begin{cases} 0, & \lambda \geq \alpha \\ 1, & \lambda < \alpha \end{cases}$
truncated singular value decomposition (4.15)	$q_\alpha(\lambda) = \begin{cases} \lambda^{-1}, & \lambda \geq \alpha \\ \alpha^{-1}, & \lambda < \alpha \end{cases}$	$r_\alpha(\lambda) = \begin{cases} 0, & \lambda \geq \alpha \\ 1 - \lambda/\alpha, & \lambda < \alpha \end{cases}$
Landweber iteration with $\mu = 1$	$q_n(\lambda) = \sum_{j=0}^{n-1}(1-\lambda)^j$	$r_n(\lambda) = (1-\lambda)^n$

In all these cases the functions r_α satisfy

$$\lim_{\alpha\to 0} r_\alpha(\lambda) = \begin{cases} 0, & \lambda > 0 \\ 1, & \lambda = 0 \end{cases} \tag{5.6}$$

$$|r_\alpha(\lambda)| \leq C_{\mathrm{r}} \qquad \text{for } \lambda \in [0, \|T^*T\|]. \tag{5.7}$$

with some constant $C_{\mathrm{r}} > 0$. The limit function defined by the right hand side of (5.6) is denoted by $r_0(\lambda)$. For Landweber iteration we set $\alpha = 1/n$ and assume that the normalization condition (2.7) holds true. Note that (5.6) is equivalent to $\lim_{\alpha\to 0} q_\alpha(\lambda) = 1/\lambda$ for all $\lambda > 0$. Hence, q_α explodes near 0. For all methods listed in the table above this growth is bounded by

$$|q_\alpha(\lambda)| \leq \frac{C_{\mathrm{q}}}{\alpha} \qquad \text{for } \lambda \in [0, \|T^*T\|] \tag{5.8}$$

with some constant $C_{\mathrm{q}} > 0$.

Theorem 5.1. *If (5.6) and (5.7) hold true, then the operators R_α defined by (5.3) converge pointwise to $T^\dagger$ on $\mathcal{D}(T^\dagger)$ as $\alpha \to 0$. With the additional assumption (5.8) the norm of the regularization operators can be estimated by*

$$\|R_\alpha\| \leq \sqrt{\frac{(C_{\mathrm{r}}+1)C_{\mathrm{q}}}{\alpha}}. \tag{5.9}$$

If $\overline{\alpha}(\delta, g^\delta)$ is a parameter choice rule satisfying

$$\overline{\alpha}(\delta, g^\delta) \to 0, \qquad \textit{and} \qquad \delta/\sqrt{\overline{\alpha}(\delta, g^\delta)} \to 0 \qquad \textit{as } \delta \to 0, \tag{5.10}$$

then $(R_\alpha, \overline{\alpha})$ is a regularization method in the sense of Definition 3.8.

Proof. We first aim to show the pointwise convergence $R_\alpha \to T^\dagger$. Let $g \in \mathcal{D}(T^\dagger)$, $\varphi^\dagger = T^\dagger g$, and $A := T^*T$. Recall from (5.4) that $T^\dagger g - R_\alpha g = r_\alpha(A)\varphi^\dagger$. Using the boundedness condition (5.7), it follows from (4.31) in Theorem 4.17 that $\lim_{\alpha\to 0} r_\alpha(A)\varphi^\dagger = r_0(A)\varphi^\dagger$. Since r_0 is real-valued and $r_0^2 = r_0$, the operator $r_0(A)$ is an orthogonal projection. Moreover, $R(r_0(A)) \subset N(A)$ since $\lambda r_0(\lambda) = 0$ for all λ and hence $Ar_0(A) = 0$. By (3.8) we have $N(T) = N(A)$. Hence, $\|r_0(A)\varphi^\dagger\|^2 = \left\langle r_0(A)\varphi^\dagger, \varphi^\dagger \right\rangle = 0$ as $\varphi^\dagger \in N(T)^\perp = N(A)^\perp$. This shows that $\|R_\alpha g - T^\dagger g\| \to 0$ as $\alpha \to 0$.

Using Lemma 4.16 and the Cauchy-Schwarz inequality we obtain that

$$\|R_\alpha \psi\|^2 = \langle TT^* q_\alpha(TT^*)\psi, q_\alpha(TT^*)\psi \rangle \leq \|\lambda q_\alpha(\lambda)\|_\infty \|q_\alpha(\lambda)\|_\infty \|\psi\|^2$$

for $\psi \in Y$. Now (5.9) follows from the assumptions (5.7) and (5.8).

To prove that $(R, \overline{\alpha})$ is a regularization method, we estimate

$$\|\varphi^\dagger - \varphi_\alpha^\delta\| \leq \|\varphi^\dagger - \varphi_\alpha\| + \|\varphi_\alpha - \varphi_\alpha^\delta\|. \tag{5.11}$$

The approximation error $\|\varphi^\dagger - \varphi_\alpha\|$ tends to 0 due to the pointwise convergence of R_α and the first assumption in (5.10). The propagated data noise error $\|\varphi_\alpha - \varphi_\alpha^\delta\| = \|R_{\overline{\alpha}(\delta)}(g - g^\delta)\|$ vanishes asymptotically as $\delta \to 0$ by (5.2), (5.9), and the second assumption in (5.10). □

Source conditions

We have seen in Theorem 3.11 that the convergence of any regularization method can be arbitrarily slow in general. On the other hand, we have seen in Example 1.2 that estimates on the rate of convergence as the noise level δ tends to 0 can be obtained under a-priori smoothness assumptions on the solution. In a general Hilbert space setting such conditions have the form

$$\varphi^\dagger = f(T^*T)w, \qquad w \in X, \|w\| \leq \rho \tag{5.12}$$

with a continuous function f satisfying $f(0) = 0$. (5.12) is called a *source condition*. The most common choice $f(\lambda) = \lambda^\mu$ with $\mu > 0$ leads to source conditions of Hölder type,

$$\varphi^\dagger = (T^*T)^\mu w, \qquad w \in X, \|w\| \leq \rho. \tag{5.13}$$

Since T is typically a smoothing operator, (5.12) and (5.13) can be seen as abstract smoothness conditions. In the next chapter we will show for some important problems that source conditions can be interpreted as classical smoothness conditions in terms of Sobolev spaces. In (5.13) the case $\mu = 1/2$ is of special importance, since

$$R((T^*T)^{1/2}) = R(T^*) \tag{5.14}$$

as shown in the exercises. To take advantage of the source condition (5.13) we assume that there exist constants $0 \leq \mu_0 \leq \infty$ and $C_\mu > 0$ such that

$$\sup_{\lambda \in [0, \|T^*T\|]} |\lambda^\mu r_\alpha(\lambda)| \leq C_\mu \alpha^\mu \qquad \text{for } 0 \leq \mu \leq \mu_0. \tag{5.15}$$

The constant μ_0 is called the *qualification* of the family of regularization operators (R_α) defined by (5.3). A straightforward computation shows that the qualification of (iterated) Tikhonov regularization $\mu_0 = 1$ (or $\mu_0 = n$, respectively), and that the qualification of Landweber iteration and the truncated singular value decomposition is $\mu_0 = \infty$. By the following theorem, μ_0 is a measure of the maximal degree of smoothness, for which the method converges of optimal order.

Theorem 5.2. *Assume that (5.13) and (5.15) hold. Then the approximation error and its image under T satisfy*

$$\begin{aligned} \|\varphi^\dagger - \varphi_\alpha\| &\leq C_\mu \alpha^\mu \rho, && \text{for } 0 \leq \mu \leq \mu_0, && (5.16)\\ \|T\varphi^\dagger - T\varphi_\alpha\| &\leq C_{\mu+1/2} \alpha^{\mu+1/2} \rho, && \text{for } 0 \leq \mu \leq \mu_0 - \frac{1}{2}. && (5.17) \end{aligned}$$

Under the additional assumptions (5.2) and (5.8) and with the a-priori parameter choice rule $\alpha = c\delta^{\frac{2}{2\mu+1}}$, $c > 0$, the total error is bounded by

$$\|\varphi^\dagger - \varphi_\alpha^\delta\| \leq c_\mu \delta^{\frac{2\mu}{2\mu+1}} \tag{5.18}$$

with some constant $c_\mu > 0$ independent of g^δ. For the parameter choice rule $\alpha = c(\delta/\rho)^{\frac{2}{2\mu+1}}$, $c > 0$ we have

$$\|\varphi^\dagger - \varphi_\alpha^\delta\| \leq c_\mu \rho^{\frac{1}{2\mu+1}} \delta^{\frac{2\mu}{2\mu+1}} \tag{5.19}$$

with c_μ independent of g^δ and ρ.

Proof. Using (5.4), (5.13), (5.15), and the isometry of the functional calculus (4.28c), we obtain

$$\|\varphi^\dagger - \varphi_\alpha\| = \|r_\alpha(T^*T)\varphi^\dagger\| = \|r_\alpha(T^*T)(T^*T)^\mu w\| \le \|\lambda^\mu r_\alpha(\lambda)\|_\infty \rho \le C_\mu \alpha^\mu \rho.$$

(5.17) follows from (5.16) together with the identify

$$\|T\psi\|^2 = \langle T\psi, T\psi\rangle = \langle T^*T\psi, \psi\rangle = \left\langle (T^*T)^{1/2}\psi, (T^*T)^{1/2}\psi\right\rangle = \|(T^*T)^{1/2}\psi\|^2. \qquad (5.20)$$

Due to the assumption (5.2) and (5.9), this gives the estimate

$$\|\varphi^\dagger - \varphi_\alpha^\delta\| \le \|\varphi^\dagger - \varphi_\alpha\| + \|\varphi_\alpha - \varphi_\alpha^\delta\| \le C_\mu \alpha^\mu \rho + \sqrt{(C_\mathrm{r}+1)C_\mathrm{q}}\,\alpha^{-1/2}\delta.$$

This shows (5.18) and (5.19). □

The choice $f(\lambda) = \lambda^\mu$ in (5.12) is not always appropriate. Whereas Hölder-type source conditions (5.13) have natural interpretations for many mildly ill-posed problems, they are far too restrictive for most exponentially ill-posed problems. (5.13) often implies that $\varphi^\dagger$ in an analytic function in such situations. As we will see, an appropriate choice of f for most exponentially ill-posed problems is

$$f_p(\lambda) := \begin{cases} (-\ln\lambda)^{-p}, & 0 < \lambda \le \exp(-1) \\ 0, & \lambda = 0 \end{cases} \qquad (5.21)$$

with $p > 0$, i.e.

$$\varphi^\dagger = f_p(T^*T)w, \qquad \|w\| \le \rho. \qquad (5.22a)$$

We call (5.22) a *logarithmic source condition*. In order to avoid the singularity of $f_p(\lambda)$ at $\lambda = 1$, we always assume in this context that the norm in Y is scaled such that

$$\|T^*T\| = \|T\|^2 \le \exp(-1). \qquad (5.22b)$$

The equality in (5.22b) is a consequence of Lemma 4.1. Of course, scaling the norm of Y has the same effect as scaling the operator T.

Theorem 5.3. *Assume that (5.22) with $p > 0$ and (5.15) with $\mu_0 > 0$ hold true. Then there exist constants $\gamma_p > 0$ such that the approximation error is bounded by*

$$\|\varphi^\dagger - \varphi_\alpha\| \le \gamma_p f_p(\alpha)\rho \qquad (5.23)$$

for all $\alpha \in [0, \exp(-1)]$. Under the additional assumptions (5.2) and (5.8) and with the a-priori parameter choice rule $\alpha = \delta$, the total error is bounded by

$$\|\varphi^\dagger - \varphi_\alpha^\delta\| \le c_p f_p(\delta) \qquad \textit{for } \delta \le \exp(-1) \qquad (5.24)$$

with some constant $c_p > 0$ independent of g^δ. If $\alpha = \delta/\rho$, then

$$\|\varphi^\dagger - \varphi_\alpha^\delta\| \le c_p \rho f_p(\delta/\rho), \qquad \textit{for } \delta/\rho \le \exp(-1) \qquad (5.25)$$

with some constant $c_p > 0$ independent of g^δ and ρ.

Proof. It follows from (5.4), (5.22), and (4.28c) that

$$\|\varphi^\dagger - \varphi_\alpha\| = \|r_\alpha(T^*T)\varphi^\dagger\| = \|r_\alpha(T^*T)f_p(T^*T)w\| \le \rho\|r_\alpha f_p\|_{\infty,[0,\exp(-1)]}.$$

If we can show that (5.15) implies

$$|r_\alpha(\lambda)|f_p(\lambda) \le \gamma_p f_p(\alpha) \tag{5.26}$$

for all $\lambda, \alpha \in [0, \exp(-1)]$, we have proved (5.23). To show (5.26) we make the substitution $\zeta = \alpha/\lambda$ and write $\tilde{r}(\lambda,\zeta) := r_{\zeta\lambda}(\lambda)$ with $\zeta \in [0, 1/(e\lambda)]$. (5.15) and (5.26) are equivalent to

$$\begin{aligned} \tilde{r}(\lambda,\zeta) &\le C_\mu \zeta^\mu, & (5.27)\\ \tilde{r}(\lambda,\zeta) &\le \gamma_p \frac{f_p(\zeta\lambda)}{f_p(\lambda)}, & (5.28) \end{aligned}$$

respectively. Suppose that (5.27) holds true, and define $\tilde{\gamma}_p := \sup_{0<\zeta\le 1} C_{\overline{\mu}}\zeta^{\overline{\mu}}(-\ln\zeta + 1)^p < \infty$. Then

$$\tilde{r}(\lambda,\zeta) \le C_{\bar{\mu}}\zeta^{\bar{\mu}} \le \tilde{\gamma}_p(-\ln\zeta + 1)^{-p} \le \tilde{\gamma}_p\left(\frac{\ln\zeta}{\ln\lambda} + 1\right)^{-p} = \tilde{\gamma}_p\frac{f_p(\zeta\lambda)}{f_p(\lambda)}$$

for $0 < \zeta \le 1$ and $0 < \lambda \le \exp(-1)$ since $\ln\lambda \le -1$. For $1 < \zeta \le 1/(e\lambda)$, condition (5.15) with $\mu = 0$ implies that

$$C_0\frac{f_p(\zeta\lambda)}{f_p(\lambda)} > C_0 \ge \tilde{r}(\lambda,\zeta).$$

This proves (5.26) with $\gamma_p = \max(C_0, \tilde{\gamma}_p)$.

Using (5.11), (5.9), and (5.23) we obtain

$$\|\varphi^\dagger - \varphi_\alpha^\delta\| \le \gamma_p\rho f_p(\alpha) + \sqrt{(1+C_0)C_{\mathrm{q}}}\frac{\delta}{\sqrt{\alpha}}.$$

This implies (5.24) with $c_p = \gamma_p\rho + \sqrt{(1+C_0)C_{\mathrm{q}}}\sup_{0<\lambda\le\exp(-1)}\sqrt{\lambda}/f_p(\lambda)$ and (5.25) with $c_p = \gamma_p + \sqrt{(1+C_0)C_{\mathrm{q}}}\sup_{0<\lambda\le\exp(-1)}\sqrt{\lambda}/f_p(\lambda)$. ☐

Optimality and an abstract stability result

Assume we want to find the best approximate solution $\varphi^\dagger = T^\dagger g$ to (5.1), and at our disposal are noisy data g^δ satisfying (5.2) and the a-priori information that $\varphi^\dagger$ satisfies (5.12). Define

$$M_{f,\rho} = \{\varphi^\dagger \in X : \varphi^\dagger \text{ satisfies (5.12)}\},$$

and let $R: Y \to X$ be an arbitrary mapping to approximately recover $\varphi^\dagger$ from g^δ. Then the *worst case error of the method* R is

$$\Delta_R(\delta, M_{f,\rho}, T) := \sup\{\|Rg^\delta - \varphi^\dagger\| : \varphi^\dagger \in M_{f,\rho},\ g^\delta \in Y,\ \|T\varphi^\dagger - Qg^\delta\| \le \delta\}.$$

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

$$\Delta(\delta, M_{f,\rho}, T) := \inf_R \Delta_R(\delta, M_{f,\rho}, T) \tag{5.29}$$

Theorem 5.4. *Let* $\omega(\delta, M_{f,\rho}, T) := \sup\{\|\varphi^\dagger\| : \varphi^\dagger \in M_{f,\rho}, \|T\varphi^\dagger\| \leq \delta\}$ *denote the modulus of continuity of* $(T|_{M_{f,\rho}})^{-1}$*. Then*

$$\Delta(\delta, M_{f,\rho}, T) \geq \omega(\delta, M_{f,\rho}, T). \tag{5.30}$$

Proof. Let $R : Y \to X$ be an arbitrary mapping, and let $\varphi^\dagger \in M_{f,\rho}$ such that $\|T\varphi^\dagger\| \leq \delta$. Choosing $g^\delta = 0$ in the definition of Δ_R gives

$$\Delta_R(\delta, M_{f,\rho}, T) \geq \|R(0) - \varphi^\dagger\|.$$

Since also $-\varphi^\dagger \in M_{f,\rho}$ and $\| - T\varphi^\dagger\| \leq \delta$, we get

$$\Delta_R(\delta, M_{f,\rho}, T) \geq \|R(0) + \varphi^\dagger\|,$$

and hence

$$2\|\varphi^\dagger\| \leq \|R(0) - \varphi^\dagger\| + \|R(0) + \varphi^\dagger\| \leq 2\Delta_R(\delta, M_{f,\rho}, T).$$

This implies $\omega(\delta, M_{f,\rho}, T) \leq \Delta_R(\delta, M_{f,\rho}, T)$ after taking the supremum over all $\varphi^\dagger \in M_{f,\rho}$ with $\|T\varphi^\dagger\| \leq \delta$. Since R was arbitrary, we have proved (5.30). □

It can be shown that actually equality holds in (5.30). The proof of the reverse inequality, which is much more difficult than the proof of Theorem 5.4, proceeds by the construction of an optimal method for the set $M_{f,\rho}$ (cf. [MM79, Lou89, GP95]). These methods require the a-priori knowledge of both f and ρ.

Our next aim is to find a computable expression for $\omega(\delta, M_{f,\rho}, T)$ which allows us to compare error bounds for practical regularization methods to the best possible error bound. The proof of the next theorem is based on

Lemma 5.5. (Jensen's inequality) *Assume that* $\phi \in C^2([\alpha, \beta])$ *with* $\alpha, \beta \in \mathbb{R} \cup \{\pm\infty\}$ *is convex, and let* μ *be a finite measure on some measure space* Ω*. Then*

$$\phi\left(\frac{\int \chi \, d\mu}{\int d\mu}\right) \leq \frac{\int \phi \circ \chi \, d\mu}{\int d\mu} \tag{5.31}$$

holds for all $\chi \in L^1(\Omega, d\mu)$ *satisfying* $\alpha \leq \chi \leq \beta$ *almost everywhere* $d\mu$*. The right hand side may be infinite if* $\alpha = -\infty$ *or* $\beta = \infty$*.*

Proof. W.l.o.g. we may assume that $\int d\mu = 1$. Let $M := \int \chi d\mu$, and consider the Taylor expansion $\phi(\xi) = \phi(M) + \phi'(M)(\xi - M) + \phi''(\eta) \cdot (\xi - M)^2/2$ for some $\eta \in (\alpha, \beta)$. Since $\phi''(\eta) \geq 0$, we have $\phi(M) + \phi'(M)(\xi - M) \leq \phi(\xi)$ for all $\xi \in [\alpha, \beta]$. Hence, with $\xi = \chi(x)$,

$$\phi(M) + \phi'(M)(\chi - M) \leq \phi \circ \chi$$

for almost all $x \in \Omega$. An integration $d\mu$ yields (5.31). □

For the special case $\phi(t) = t^p$, $p > 1$, Jensen's inequality becomes

$$\int \chi d\mu \leq \left(\int \chi^p d\mu\right)^{\frac{1}{p}} \left(\int d\mu\right)^{\frac{1}{q}}$$

with $q = \frac{p}{p-1}$. From this form, we easily obtain Hölder's inequality

$$\int |a||b| d\tilde{\mu} \leq \left(\int |a|^p d\tilde{\mu} \right)^{\frac{1}{p}} \left(\int |b|^q d\tilde{\mu} \right)^{\frac{1}{q}}$$

for positive measures $\tilde{\mu}$ on Ω, $a \in L^p(\tilde{\mu})$, and $b \in L^q(\tilde{\mu})$ by setting $\mu = |b|^q \tilde{\mu}$ and $\chi = |a||b|^{-\frac{1}{p-1}}$.

Theorem 5.6. *Let (5.12) hold, and assume that $f \in C([0,\tau])$, $\tau = \|T\|^2$, is strictly monotonically increasing with $f(0) = 0$. Moreover, assume that the function $\phi : [0, f(\tau)^2] \to [0, \tau f(\tau)^2]$ defined by*

$$\phi(\xi) := \xi \cdot (f \cdot f)^{-1}(\xi) \tag{5.32}$$

is convex and twice continuously differentiable. Then the stability estimate

$$\|\varphi^\dagger\|^2 \leq \rho^2 \phi^{-1} \left(\frac{\|T\varphi^\dagger\|^2}{\rho^2} \right). \tag{5.33}$$

holds. Consequently, for $\delta \leq \rho\sqrt{\tau} f(\tau)$,

$$\omega(\delta, M_{f,\rho}, T) \leq \rho \sqrt{\phi^{-1}\left(\delta^2/\rho^2\right)}. \tag{5.34}$$

Proof. By linearity, we may assume that $\rho = 1$. With the notation of Theorem 4.8 let $\mu_w := |W^{-1}w|^2 \mu$. Then (5.31) and (5.32) yield

$$\begin{aligned} \phi\left(\frac{\|\varphi^\dagger\|^2}{\|w\|^2} \right) &= \phi\left(\frac{\|W^{-1}\varphi^\dagger\|^2_{L^2(\Omega,d\mu)}}{\|W^{-1}w\|^2_{L^2(\Omega,d\mu)}} \right) = \phi\left(\frac{\int f(\lambda)^2 d\mu_w}{\int d\mu_w} \right) \\ &\leq \frac{\int \phi(f(\lambda)^2) d\mu_w}{\int d\mu_w} = \frac{\int \lambda f(\lambda)^2 d\mu_w}{\|w\|^2} = \frac{\|(T^*T)^{1/2} f(T^*T) w\|^2}{\|w\|^2} = \frac{\|T\varphi^\dagger\|^2}{\|w\|^2}. \end{aligned}$$

By the convexity of ϕ, the fact that $\phi(0) = 0$, and $\|w\| \leq 1$, this estimate implies

$$\phi(\|\varphi^\dagger\|^2) \leq \|T\varphi^\dagger\|^2. \tag{5.35}$$

Since f is strictly increasing, so are $f \cdot f$, $(f \cdot f)^{-1}$, ϕ, and ϕ^{-1}. Hence, applying ϕ^{-1} to (5.35) yields (5.33). (5.34) follows from (5.33) and the definition. $\square$

The estimate (5.33) is a *stability estimate* for (5.1) based on the a-priori information that $\varphi^\dagger \in M_{f,\rho}$. It corresponds to the stability estimate derived in Example 1.2, which was based on an a-priori bound on the second derivative of the solution.

Remark 5.7. We discuss when equality holds in (5.33) and (5.34). Let w be an eigenvector of T^*T such that $\|w\| = \rho = 1$ and $T^*Tw = \tilde{\lambda} w$. Then $\varphi^\dagger = f(\tilde{\lambda}) w$ and

$$\|\varphi^\dagger\|^2 = f(\tilde{\lambda})^2 = \phi^{-1}(\tilde{\lambda} f(\tilde{\lambda})^2) = \phi^{-1}(\|T\varphi^\dagger\|^2).$$

In the second equality we have used the definition (5.32) with $\xi = f(\tilde{\lambda})^2$, and in the last equality the identity (5.20). Hence, (5.33) is sharp in this case. Moreover, equality holds in (5.34) if $(\delta/\rho)^2$ is an eigenvalue of $T^*Tf(T^*T)^2$. An exact expression for $\omega(\delta, M_{f,\rho}, T)$ for all δ is derived in [Hoh99].

Definition 5.8. Let $(R_\alpha, \overline{\alpha})$ be a regularization method for (5.1), and let the assumptions of Theorem 5.6 be satisfied. Convergence on the source sets $M_{f,\rho}$ is said to be

- *optimal* if
$$\Delta_{R_\alpha}(\delta, M_{f,\rho}, T) \le \rho\sqrt{\phi^{-1}\left(\delta^2/\rho^2\right)}$$
- *asymptotically optimal* if
$$\Delta_{R_\alpha}(\delta, M_{f,\rho}, T) = \rho\sqrt{\phi^{-1}\left(\delta^2/\rho^2\right)}\,(1 + \mathrm{o}\,(1)), \qquad \delta \to 0$$
- *of optimal order* if there is a constant $C \ge 1$ such that
$$\Delta_{R_\alpha}(\delta, M_{f,\rho}, T) \le C\rho\sqrt{\phi^{-1}\left(\delta^2/\rho^2\right)}$$
for δ/ρ sufficiently small.

For $f(\lambda) = \lambda^\mu$, $\mu > 0$, the assumptions of Theorem 5.6 are satisfied, and $\phi(\xi) = \xi^{\frac{1+2\mu}{2\mu}}$. We get

Corollary 5.9. *(5.13) implies*
$$\|\varphi^\dagger\| \le \rho^{\frac{1}{1+2\mu}}\|T\varphi^\dagger\|^{\frac{2\mu}{1+2\mu}}.$$
Moreover,
$$\omega(\delta, M_{\lambda^\mu,\rho}, T) \le \rho^{\frac{1}{2\mu+1}}\delta^{\frac{2\mu}{1+2\mu}}.$$

Corollary 5.9 implies that the method in Theorem 5.2 is of optimal order. Note, however, that this method requires knowledge of the parameter μ, i.e. the degree of smoothness of the unknown solution. We will see below that a-posteriori parameter choice rules can lead to order-optimal methods, which do not require such a-priori knowledge.

Usually, Corollary 5.9 is proved by interpolation instead of Jensen's inequality (cf.,e.g.,[EHN96, Lou89]). We have chosen the latter approach since it also allows to treat logarithmic source conditions (cf. [Mai94]):

Corollary 5.10. *The assumptions of Theorem 5.6 are satisfied for $f = f_p$ (see (5.21)), and the inverses of the corresponding functions ϕ_p have the asymptotic behavior*
$$\sqrt{\phi_p^{-1}(\lambda)} = f_p(\lambda)(1 + \mathrm{o}\,(1)), \qquad \lambda \to 0. \tag{5.36}$$
Consequently,
$$\|\varphi^\dagger\| \le \rho f_p\left(\frac{\|T\varphi^\dagger\|^2}{\rho^2}\right)(1 + \mathrm{o}\,(1)), \qquad \|T\varphi^\dagger\|/\rho \to 0, \tag{5.37}$$
$$\omega(\delta, M_{f_p,\rho}, T) \le \rho f_p\left(\delta^2/\rho^2\right)(1 + \mathrm{o}\,(1)), \qquad \delta/\rho \to 0. \tag{5.38}$$

Proof. By (5.22b), we have $\tau = \exp(-1)$. It is obvious that f_p is continuous on $[0,\tau]$ and strictly monotonically increasing. $\phi_p : [0,1] \to [0, \exp(-1)]$ is given by

$$\phi_p(\xi) = \xi \exp(-\xi^{-\frac{1}{2p}}).$$

From

$$\phi_p''(\xi) = \exp(-\xi^{-1/2p}) \frac{\xi^{-1-\frac{1}{2p}}}{(2p)^2} (2p - 1 + \xi^{-\frac{1}{2p}})$$

it is easily seen that $\phi_p''(\xi) \geq 0$ for $\xi \in [0,1]$, i.e. ϕ_p is convex.

To prove the estimate on $\sqrt{\phi_p^{-1}(\lambda)}$, first note that $\xi = \phi_p^{-1}(\lambda)$ implies

$$\ln \lambda = \ln \xi - \xi^{-\frac{1}{2p}}.$$

Therefore,

$$\begin{aligned} \xi &= (\ln \xi - \ln \lambda)^{-2p} \\ &= (-\ln \lambda)^{-2p} \left(1 - \frac{\ln \xi}{\ln \lambda}\right)^{-2p} \\ &= f_p(\lambda)^2 \left(1 - \frac{\ln \xi}{\ln \xi - \xi^{-\frac{1}{2p}}}\right)^{-2p} \end{aligned}$$

Since $\lim_{\xi \to 0} \frac{\ln \xi}{\ln \xi - \xi^{-1/2p}} = 0$ and $\lim_{\lambda \to 0} \xi = \lim_{\lambda \to 0} \phi_p^{-1}(\lambda) = 0$, the assertion follows. □

As $f_p(\delta^2/\rho^2) = 2^{-p} f_p(\delta/\rho)$, Corollary 5.10 implies that the method in Theorem 5.3 is of optimal order.

The discrepancy principle

We assume in this section the exact data g are attainable, i.e.

$$g \in R(T). \tag{5.39}$$

Assume that (5.7) holds true and choose

$$\tau > C_{\mathrm{r}}. \tag{5.40}$$

Instead of (2.8) we consider a version of the discrepancy principle which allows some sloppyness in the choice of the regularization paramter. If $\|g^\delta\| > \tau\delta$, i.e. if the *signal-to-noise ration* $\|g^\delta\|/\delta$ is greater than τ, we assume that $\alpha = \overline{\alpha}(\delta, g^\delta)$ is chosen such that

$$\|g^\delta - T\varphi_\alpha^\delta\| \leq \tau\delta \leq \|g^\delta - T\varphi_{\alpha'}^\delta\| \tag{5.41a}$$

for some $\alpha' \in [\alpha, 2\alpha]$. Otherwise, if $\|g^\delta\| \leq \tau\delta$, we set

$$\overline{\alpha}(\delta, g^\delta) = \infty \tag{5.41b}$$

and $\varphi_\infty^\delta := 0$. In the latter case, when there is more noise than data, there may not exist α' such that (5.41a) is satisfied. Note that under assumption (5.8),

$$\|\varphi_\alpha^\delta\| = \|q_\alpha(T^*T)T^*g^\delta\| \leq \frac{C_{\mathrm{q}}}{\alpha}\|T^*g^\delta\| \to 0 \qquad \text{as } \alpha \to \infty, \tag{5.42}$$

which explains the notation $\varphi_\infty^\delta := 0$. Finally note that

$$g^\delta - T\varphi_\beta^\delta = (I - Tq_\beta(T^*T)T^*)g^\delta = r_\beta(TT^*)g^\delta \tag{5.43}$$

for all $\beta > 0$ by virtue of Lemma 4.16. Hence, $\lim_{\alpha\to 0}\|g^\delta - T\varphi_\alpha^\delta\| = \|r_0(TT^*)g^\delta\| \leq \|g^\delta\| \leq \delta$ by Theorem 4.17. Together with (5.42) this shows that (5.41a) can be satisfied if $\|g^\delta\| \geq \tau\delta$.

Theorem 5.11. *Let r_α and q_α satisfy (5.7), (5.8), and (5.15) with $\mu_0 > 1/2$, and let $\overline{\alpha}(\delta, g^\delta)$ be a parameter choice rule satisfying (5.40) and (5.41). Then $(R_\alpha, \overline{\alpha})$ is a regularization method in the sense of Defintion 3.8.*

Proof. We will repeatedly use the inequalities

$$\begin{aligned} \|g - T\varphi_\alpha\| &\leq (\tau + C_{\mathrm{r}})\delta & (5.44\mathrm{a}) \\ (\tau - C_{\mathrm{r}})\delta &\leq \|g - T\varphi_{\alpha'}\| & (5.44\mathrm{b}) \end{aligned}$$

which follow from (5.41a), (5.40), and the estimate

$$\|(g - T\varphi_\beta) - (g^\delta - T\varphi_\beta^\delta)\| = \|r_\beta(T^*T)(g - g^\delta)\| \leq C_{\mathrm{r}}\delta. \tag{5.45}$$

Eq. (5.45), which holds for all $\beta > 0$, is a consequence of (5.2) and the identity (5.43).

The proof is again based on the splitting (5.11) of the total error into an approximation error and a propagated data error. Let $(g^{\delta_n})_{n\in\mathbb{N}}$ be a sequence in Y such that $\|g - g^{\delta_n}\| \leq \delta_n$ and $\delta_n \to 0$ as $n \to \infty$, and let $\alpha_n := \overline{\alpha}(\delta_n, g^{\delta_n})$. Assume that the approximation error does not tend to 0. Then, after passing to a subsequence, we may assume that there exists $\epsilon > 0$ such that

$$\|\varphi^\dagger - \varphi_{\alpha_n}\| \geq \epsilon \qquad \text{for all } n. \tag{5.46}$$

Using the notation of Theorem 4.8 with $T^*T = A = WM_\lambda W^{-1}$, let $\psi^\dagger := W^{-1}\varphi^\dagger$. It follows from (5.4) and (5.20), and (5.44a) that

$$\int \lambda |r_{\alpha_n} \circ \lambda|^2 |\psi^\dagger|^2 \, d\mu = \|\sqrt{A}\, r_{\alpha_n}(A)\varphi^\dagger\|^2 = \|T(\varphi^\dagger - \varphi_{\alpha_n})\|^2 \to 0$$

as $n \to \infty$. Since $0 \leq \lambda \leq \|A\|$, it follows from the Riesz-Fischer Theorem that there exists a subsequence of $n(k)$, such that $\lambda|r_{\alpha_{n(k)}} \circ \lambda|^2|\psi^\dagger|^2 \to 0$ pointwise almost everywhere $d\mu$ as $k \to \infty$. Hence, $|r_{\alpha_{n(k)}} \circ \lambda - r_0 \circ \lambda|^2|\psi^\dagger|^2 \to 0$ pointwise almost everywhere $d\mu$ as $k \to \infty$, where r_0 denotes the limit function defined by the right-hand side of (5.6). With assumption (5.7), it follows from Lebesque's Dominated Convergence Theorem that

$\int |r_{\alpha_{n(k)}} \circ \lambda - r_0 \circ \lambda|^2 |\psi^\dagger|^2 \, \mathrm{d}\mu \to 0$ as $k \to \infty$. Since $r_0(A)\varphi^\dagger = 0$ as shown in the proof of Theorem 5.1, this implies that

$$\|\varphi^\dagger - \varphi_{\alpha_{n(k)}}\| = \|r_{\alpha_{n(k)}}(A)\varphi^\dagger\| = \|r_{\alpha_{n(k)}}(A)\varphi^\dagger - r_0(A)\varphi^\dagger\| \to 0$$

as $k \to \infty$. This contradicts (5.46). Hence, $\|\varphi^\dagger - \varphi_{\alpha_n}\| \to 0$ as $n \to \infty$.

Now assume that the propagated data noise error does not tend to 0. In analogy to (5.46) we may assume that there exists $\epsilon > 0$ such that

$$\|\varphi_{\alpha_n} - \varphi_{\alpha_n}^{\delta_n}\| \geq \epsilon \qquad \text{for all } n. \tag{5.47}$$

It follows from (5.9) that

$$\|\varphi_{\alpha_n} - \varphi_{\alpha_n}^{\delta}\| \leq C\frac{\delta_n}{\sqrt{\alpha_n}} \leq C\frac{\delta_n}{\sqrt{\alpha_n'}} \tag{5.48}$$

with a generic constant C independent of n. This implies that (5.48) $\alpha_n' \to 0$ as $n \to \infty$. Now it follows from (5.44b), (5.17), and the assumption $\mu_0 > 1/2$ that

$$\|\varphi_{\alpha_n} - \varphi_{\alpha_n}^{\delta}\| \leq C\frac{\delta_n}{\sqrt{\alpha_n'}} \leq C\frac{\|T(\varphi^\dagger - \varphi_{\alpha_n'})\|}{\sqrt{\alpha_n'}} \leq C(\alpha_n')^{\mu_0 - 1/2} \to 0 \tag{5.49}$$

as $n \to \infty$ with a generic constant C. This contradicts (5.47). Hence, both the approximation error and the propagated data noise error tend to 0, and the proof is complete. □

The next theorem shows that the discrepancy principle leads to order optimal convergence rates.

Theorem 5.12. *Under the assumptions of Theorem 5.11 let $\varphi^\dagger$ satisfy the Hölder source condition (5.13) with $0 < \mu \leq \mu_0 - 1/2$. Then there exists a constant $c_\mu > 0$ independent of ρ, δ, and $\varphi^\dagger$ such that*

$$\|\varphi^\dagger - \varphi_{\overline{\alpha}(\delta, g^\delta)}\| \leq c_\mu \rho^{\frac{1}{2\mu+1}} \delta^{\frac{2\mu}{2\mu+1}}. \tag{5.50}$$

Proof. To estimate the approximation error we use Corollary 5.9 with $\varphi^\dagger$ replaced by $\varphi^\dagger - \varphi_\alpha = r_\alpha(T^*T)\varphi^\dagger = (T^*T)^\mu r_\alpha(T^*T)w$. Since $\|r_\alpha(T^*T)w\| \leq C_{\mathrm{r}}\rho$, this gives

$$\|\varphi^\dagger - \varphi_\alpha\| \leq (\rho C_{\mathrm{r}})^{\frac{1}{1+2\mu}} \|T(\varphi^\dagger - \varphi_\alpha)\|^{\frac{2\mu}{2\mu+1}} \leq (\rho C_{\mathrm{r}})^{\frac{1}{1+2\mu}} ((\tau + C_{\mathrm{r}})\delta)^{\frac{2\mu}{1+2\mu}}. \tag{5.51}$$

Here the second inequality is a consequence of (5.44a). To estimate the propagated data noise error, note that (5.17) and (5.44b) imply

$$(\tau - C_{\mathrm{r}})\delta \leq \|g - T\varphi_{\alpha'}\| \leq C_{\mu+1/2}(\alpha')^{\mu+1/2}\rho.$$

Together with (5.2) and (5.9) it follows that

$$\|\varphi_\alpha - \varphi_\alpha^\delta\| \leq \sqrt{\frac{(C_{\mathrm{r}}+1)C_{\mathrm{q}}}{\alpha}}\delta \leq \sqrt{\frac{2(C_{\mathrm{r}}+1)C_{\mathrm{q}}}{\alpha'}}\delta \leq c_\mu \rho^{\frac{1}{2\mu+1}} \delta^{\frac{2\mu}{2\mu+1}}$$

with a constant c_μ independent of ρ, δ, and $\varphi^\dagger$. This together with (5.51) shows (5.50). $\square$

By Theorem 5.12, the discrepancy principle leads to order-optimal convergence only for $\mu \leq \mu_0 - 1/2$, whereas the a-priori rule in Theorem 5.2 yields order-optimal convergence for all $\mu \leq \mu_0$. It can be shown that (5.50) cannot be extended to $\mu \in (\mu_0 - 1/2, \mu_0]$ (cf. [Gro83] for Tikhonov regularization). This fact has initiated a considerable amount of research in improved a-posteriori rule, which yield order-optimal convergence for all $\mu \leq \mu_0$ without a-priori knowledge of μ (cf. [EG88, Gfr87] and references in [EHN96]).

For logarithmic source conditions, the discrepancy principle also leads to order-optimal convergence. With a modified version of the discrepancy principle even asymptotically optimal convergence rates can be achieved without a-priori knowledge of p (cf. [Hoh00]).

6. Interpretation of source conditions

In the previous chapter we have established convergence rates of regularization methods for ill-posed operator equations $T\varphi = g$ if the exact solution $\varphi^\dagger$ satisfies a source condition

$$\varphi^\dagger = f(T^*T)w, \qquad \|w\| \le \rho$$

with a continuous function $f : [0, \|T^*T\|] \to [0, \infty)$ satisfying $f(0) = 0$. It is usually not obvious what such a condition means for a specific inverse problem. The aim of this chapter is to interpret source condition for some important inverse problems.

Sobolev spaces of periodic functions

Let

$$f_n(x) := \frac{1}{\sqrt{2\pi}} \exp(inx), \qquad n \in \mathbb{Z}$$

denote the standard orthonormal basis of $L^2([0, 2\pi])$ with respect to the inner product $\langle \varphi, \psi \rangle := \int_0^{2\pi} \varphi(x)\overline{\psi(x)}\,\mathrm{d}x$. For $\varphi \in L^2([0, 2\pi])$ we denote the Fourier coefficients of φ by

$$\hat{\varphi}(n) := \langle \varphi, f_n \rangle, \qquad n \in \mathbb{Z}.$$

Then $\|\varphi\|_{L^2}^2 = \sum_{n\in\mathbb{Z}} |\hat{\varphi}(n)|^2$ by Parseval's equality, and $\varphi = \sum_{n\in\mathbb{Z}} \hat{\varphi}(n) f_n$.

Definition 6.1. For $0 \le s < \infty$ we define

$$\|\varphi\|_{H^s} := \left(\sum_{n\in\mathbb{Z}} (1+n^2)^s |\hat{\varphi}(n)|^2 \right)^{1/2} \tag{6.1}$$

and

$$H^s([0, 2\pi]) := \{\varphi \in L^2([0, 2\pi]) : \|\varphi\|_{H^s} < \infty\}.$$

$H^s([0, 2\pi])$ is called a *Sobolev space* of index s. Note that $H^0([0, 2\pi]) = L^2([0, 2\pi])$.

Theorem 6.2. *$H^s([0, 2\pi])$ is a Hilbert space for $s \ge 0$.*

Proof. It is easy to show that $H^s([0, 2\pi])$ is a linear space and that

$$\langle \varphi, \psi \rangle_{H^s} := \sum_{n\in\mathbb{Z}} (1+n^2)^s \hat{\varphi}(n)\overline{\hat{\psi}(n)}$$

is an inner product on $H^s([0,2\pi])$ satisfying $\|\varphi\|_{H^s}^2 = \langle\varphi,\varphi\rangle_{H^s}$ for all $\varphi \in H^s([0,2\pi])$. To show completeness, let $(\varphi_k)_{k\in\mathbb{N}}$ be a Cauchy sequence in $H^s([0,2\pi])$. Then $(\hat{\varphi}_k(n))_{k\in\mathbb{N}}$ is a Cauchy sequence for each $n \in \mathbb{Z}$, and $\hat{\varphi}(n) := \lim_{k\to\infty}\hat{\varphi}_k(n)$ is well defined. We have to show that $\varphi := \sum_{n\in\mathbb{Z}} f_n\hat{\varphi}(n)$ belongs to $H^s([0,2\pi])$ and that $\|\varphi-\varphi_k\|_{H^s} \to 0$ as $k\to\infty$. Given $\epsilon > 0$ there exists $k > 0$ such that $\|\varphi_k - \varphi_l\|_{H^s}^2 \leq \epsilon$ for all $l \geq k$. Hence,

$$\begin{aligned}\sum_{n=-N}^{N}(1+n^2)^s|\hat{\varphi}(n)-\hat{\varphi}_k(n)|^2 &= \lim_{l\to\infty}\sum_{n=-N}^{N}(1+n^2)^s|\hat{\varphi}_l(n)-\hat{\varphi}_k(n)|^2\\ &\leq \sup_{l\geq k}\|\varphi_l-\varphi_k\|_{H^s}^2 \leq \epsilon\end{aligned}$$

for all $N > 0$. Taking the supremum over $N \in \mathbb{N}$ shows that $\|\varphi - \varphi_k\|_{H^s} \leq \epsilon$. This implies that $\|\varphi\|_{H^s} < \infty$ and $\varphi_k \to \varphi$ in H^s as $k \to \infty$. □

Remark 6.3. Another definition of Sobolev spaces of integral index $s \in \mathbb{N}$ uses the concept of weak derivatives. A 2π-periodic function $\varphi \in L^2([0,2\pi])$ is said to have a *weak (or distributional) derivative* $D^k\varphi \in L^2([0,2\pi])$ of order $k \in \mathbb{N}$ if the periodic continuation of $D^k\varphi$ satisfies

$$\int_{\mathbb{R}} \varphi\chi^{(k)}\,\mathrm{d}x = (-1)^k\int_{\mathbb{R}}(D^k\varphi)\chi\,\mathrm{d}x$$

for all $\chi \in C^\infty(\mathbb{R})$ with compact support. It can be shown that $H^k([0,2\pi])$ is the set of all functions in $L^2([0,2\pi])$ which have weak derivatives in $L^2([0,2\pi])$ of order $\leq k$ (see exercises). This alternative definition can be used to introduce Sobolev spaces on more general domains.

For more information on Sobolev spaces of periodic functions, we refer to [Kre89, Chapter 8].

Hölder-type source conditions

Our first example is numerical differention. Since differentiation annihilates constant function we define the Hilbert space

$$L_0^2([0,2\pi]) := \left\{\varphi \in L^2([0,2\pi]) : \int_0^{2\pi}\varphi(x)\,\mathrm{d}x = 0\right\}$$

of L^2 function with zero mean. The inverse of the differentiation operator on $L_0^2([0,2\pi])$ is given by

$$(T_{\mathrm{D}}\varphi)(x) := \int_0^x \varphi(t)\,\mathrm{d}t + c(\varphi), \qquad x \in [0,2\pi]$$

where $c(\varphi) := -\frac{1}{2\pi}\int_0^{2\pi}\int_0^x\varphi(t)\,\mathrm{d}t\,\mathrm{d}x$ is defined such that $T_{\mathrm{D}}\varphi \in L_0^2([0,2\pi])$ for all $\varphi \in L_0^2([0,2\pi])$, i.e. $T_{\mathrm{D}} \in L(L_0^2([0,2\pi]))$.

Theorem 6.4. *$R((T_D^*T_D)^\mu) = H^{2\mu}([0,2\pi]) \cap L_0^2([0,2\pi])$ for all $\mu \geq 0$. Moreover, there exist constants $c, C > 0$ such that $c\|w\|_{L^2} \leq \|(T_D^*T_D)^\mu w\|_{H^{2\mu}} \leq C\|w\|_{L^2}$ for all $w \in L^2([0,2\pi])$, i.e. $\|(T_D^*T_D)^\mu w\|_{H^{2\mu}} \sim \|w\|_{L^2}$.*

Proof. Note that $L_0^2([0,2\pi]) = \overline{\mathrm{span}\{f_n : n \in \mathbb{Z}, n \neq 0\}}$. A straightforward computation shows that $T_D f_n = \frac{1}{in} f_n$ for $n \in \mathbb{Z} \setminus \{0\}$. It follows that $T_D^* f_n = -\frac{1}{in} f_n$ since $\langle T\varphi, f_n\rangle = \frac{1}{in}\hat{\varphi}(n) = \left\langle \varphi, \frac{-1}{in} f_n\right\rangle$. Hence,

$$T_D^* T_D f_n = \frac{1}{n^2} f_n, \qquad n \in \mathbb{Z} \setminus \{0\},$$

i.e. $\{f_n : n \in \mathbb{Z} \setminus \{0\}\}$ is complete orthonormal system of eigenfunctions of $T_D^*T_D$. It follows from Theorem 4.15 with $W : l^2(\mathbb{Z} \setminus \{0\}) \to L_0^2([0,2\pi])$, $W\hat{\varphi} := \sum_{n\neq 0} \hat{\varphi}(n) f_n$ and $\lambda(n) = n^{-2}$ that

$$(T_D^*T_D)^\mu w = \sum_{n\neq 0} \frac{1}{n^{2\mu}} \hat{w}(n) f_n$$

for all $w \in L_0^2([0,2\pi])$. Therefore,

$$\|(T_D^*T_D)^\mu w\|_{H^{2\mu}}^2 = \sum_{n\neq 0} \left(\frac{1+n^2}{n^2}\right)^{2\mu} |\hat{w}(n)|^2 \leq 2^{2\mu} \|\hat{w}(n)\|_{L^2}^2 < \infty$$

which implies $R((T_D^*T_D)^\mu) \subset H^{2\mu}([0,2\pi]) \cap L^0([0,2\pi])$. To prove the reverse inclusion, let $\varphi \in H^{2\mu}([0,2\pi])$ and define $w := \sum_{n\neq 0} n^{2\mu}\hat{\varphi}(n) f_n$. Then $w \in L_0^2([0,2\pi])$ is well defined since $\|w\|_{L^2}^2 = \sum_{n\neq 0} (n^{2\mu})^2 |\hat{\varphi}(n)|^2 \leq \|\varphi\|_{H^{2\mu}}^2 < \infty$, and $(T_D^*T_D)^\mu w = \varphi$. □

Theorem 6.4 implies that if $\varphi^\dagger \in H^{2\mu}([0,2\pi])$ and if the regularization method satisfies the assumptions of either Theorem 5.2 or Theorem 5.12, then the error can be estimated by

$$\|\varphi^\dagger - \varphi_{\overline{\alpha}(\delta, g^\delta)}\|_{L^2} \leq c_\mu \rho^{\frac{1}{2\mu+1}} \delta^{\frac{2\mu}{2\mu+1}}.$$

with $\rho = \|\varphi^\dagger\|_{H^{2\mu}}$. Replacing the Sobolev spaces $H^{2\mu}([0,2\pi])$ by the classical function spaces $C^{2\mu}([0,2\pi])$ with $\mu \in \{0, 1/2, 1\}$, this corresponds to the convergence rates (1.3) obtained for the central difference quotient.

Our second example is related to solution of the boundary value problem

$$\Delta u = 0 \quad \text{in } \Omega, \qquad u = f \quad \text{on } \partial\Omega \tag{6.2}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded, simply connected region with smooth boundary $\partial\Omega$ and $f \in C(\partial\Omega)$. The single layer potential

$$u(x) = -\frac{1}{\pi} \int_{\partial\Omega} \psi(y) \ln|x-y| \,\mathrm{d}s(y), \qquad x \in \Omega$$

solves (6.2) if and only if the density $\psi \in C(\partial\Omega)$ solves *Symm's equation*

$$-\frac{1}{\pi} \int_{\partial\Omega} \psi(y) \ln|x-y| \,\mathrm{d}s(y) = f(y), \qquad x \in \partial\Omega$$

(cf. [Kre89]). It is a prototype of a one-dimensional first-kind integral equation with a logarithmic singularity. We consider the special case that $\partial\Omega$ is a circle of radius $a > 0$ parametrized by $z(t) = a(\cos t, \sin t)$, $t \in [0, 2\pi]$. Then $|z(t) - z(\tau)|^2 = 4a^2 \sin^2 \frac{t-\tau}{2}$, and setting $\varphi(t) := a\psi(z(t))$ and $g(t) := f(z(t))$ we get

$$-\frac{1}{\pi}\int_0^{2\pi} \varphi(\tau)\left(\ln a + \frac{1}{2}\ln\left(4\sin^2\frac{t-\tau}{2}\right)\right)\mathrm{d}\tau = g(t), \qquad t \in [0, 2\pi]. \tag{6.3}$$

The left hand side of this equation is denoted by $(T_{\mathrm{Sy}}\varphi)(t)$. We now consider (6.3) as an integral equation in $L^2([0, 2\pi])$. Since the integral operator T_{Sy} has a square integrable kernel, it is a compact operator in $L(L^2([0, 2\pi]))$. Therefore, (6.3) is ill-posed.

Theorem 6.5. *If $a \neq 1$, then $R((T^*_{\mathrm{Sy}}T_{\mathrm{Sy}})^\mu) = H^{2\mu}([0, 2\pi])$ for all $\mu \geq 0$. Moreover, $\|(T^*_{\mathrm{Sy}}T_{\mathrm{Sy}})^\mu w\|_{H^{2\mu}} \sim \|w\|_{L^2}$ for $w \in L^2([0, 2\pi])$.*

Proof. Using the integrals

$$\frac{1}{2\pi}\int_0^{2\pi} e^{int}\ln\left(4\sin^2\frac{t}{2}\right)\mathrm{d}t = \begin{cases} -1/|n|, & n \in \mathbb{Z}\setminus\{0\} \\ 0, & n = 0 \end{cases}$$

(cf. [Kir96, Lemma 3.17]) we obtain

$$T_{\mathrm{Sy}} f_n = \begin{cases} 1/|n| f_n, & n \in \mathbb{Z}\setminus\{0\} \\ 2\ln a f_0, & n = 0 \end{cases}$$

As $T^*_{\mathrm{Sy}} = T_{\mathrm{Sy}}$ it follows that

$$(T^*_{\mathrm{Sy}}T_{\mathrm{Sy}}) f_n = \begin{cases} 1/|n|^2 f_n, & n \in \mathbb{Z}\setminus\{0\} \\ (2\ln a)^2 f_0, & n = 0 \end{cases}$$

Now the proof is almost identical to that of Theorem 6.4. □

Logarithmic source conditions

We consider the initial value problem for the heat equation with periodic boundary conditions:

$$\frac{\partial}{\partial t}u(x,t) = \frac{\partial^2}{\partial x^2}u(x,t), \qquad x \in (0, 2\pi), t \in (0, T] \tag{6.4a}$$

$$u(0,t) = u(2\pi, t), \qquad t \in (0, T) \tag{6.4b}$$

$$u(x, 0) = \varphi(x), \qquad x \in [0, 2\pi]. \tag{6.4c}$$

This is equivalent to the heat equation on a circle. In analogy to (1.6), the solution to this evolution problem is given by

$$u(x,t) = \sum_{n\in\mathbb{Z}} \exp(-n^2 t)\hat{\varphi}(n) f_n(x)$$

We define the operator $T_{\mathrm{BH}} \in L(X,Y)$ with $X = Y = L^2([0,2\pi])$ by $(T_{\mathrm{BH}}\varphi)(x) := u(x,T)$ where u satisfies (6.4), i.e.

$$T_{\mathrm{BH}}\varphi = \sum_{n\in\mathbb{Z}} \exp(-n^2T)\hat{\varphi}(n)f_n. \tag{6.5}$$

In order to meat the normalization condition (5.22b), we define the norm in Y to be $\|\psi\|_Y := \exp(-1/2)\|\psi\|_{L^2}$.

Theorem 6.6. *$R(f_p(T_{\mathrm{BH}}^*T_{\mathrm{BH}})) = H^{2p}([0,2\pi])$ for all $p > 0$, and $\|f_p(T_{\mathrm{BH}}^*T_{\mathrm{BH}})w\|_{H^{2p}} \sim \|w\|_{L^2}$.*

Proof. It follows from (6.5) and the definition of the norm in Y that

$$T_{\mathrm{BH}}^*T_{\mathrm{BH}}\varphi = \sum_{n\in\mathbb{Z}} \exp(-1)\exp(-2Tn^2)\hat{\varphi}(n)f_n.$$

Hence,

$$f_p(T_{\mathrm{BH}}^*T_{\mathrm{BH}})w = \sum_{n\in\mathbb{Z}} f_p(\exp(-1)\exp(-2Tn^2))\hat{w}(n)f_n = \sum_{n\in\mathbb{Z}}(1+2Tn^2)^{-p}\hat{w}(n)f_n$$

for $w \in L^2([0,2\pi])$ and

$$\|f_p(T_{\mathrm{BH}}^*T_{\mathrm{BH}})w\|_{H^{2p}}^2 = \sum_{n\in\mathbb{Z}}\left(\frac{1+n^2}{1+2Tn^2}\right)^{2p}|\hat{w}(n)|^2 \le T^{-2p}\|w\|_{L^2}^2.$$

Therefore, $R(f_p(T_{\mathrm{BH}}^*T_{\mathrm{BH}})) \subset H^{2p}([0,2\pi])$. Vice versa, let $\varphi \in H^{2p}([0,2\pi])$, and define $w := \sum_{n\in\mathbb{Z}}(1+2Tn^2)^p\hat{\varphi}(n)f_n$ such that $f_p(T_{\mathrm{BH}}^*T_{\mathrm{BH}})w = \varphi$. It follows from $\|w\|_{L^2} = \sum_{n\in\mathbb{Z}}(1+2Tn^2)^{2p}|\hat{\varphi}(n)|^2 \le (2T)^{2p}\|\varphi\|_{H^{2p}}^2 < \infty$ that $w \in L^2([0,2\pi])$. Hence, $H^{2p}([0,2\pi]) \subset R(f_p(T_{\mathrm{BH}}^*T_{\mathrm{BH}}))$. □

Theorem 6.6 shows that if $\varphi^\dagger \in H^{2p}([0,2\pi])$ and the regularization method satisfies the assumptions of Theorem 5.3, then the error can be estimated by

$$\|\varphi^\dagger - \varphi_\alpha^\delta\| \le c_p\rho f_p(\delta/\rho)$$

with $\rho = \|\varphi^\dagger\|_{H^{2p}}$.

Analogous results for the heat equation in more general domains, for the sideways heat equation, and for some other problems can be found in [Hoh00].

7. Nonlinear operators

The Fréchet derivative

Most regularization methods for nonlinear problems are based on a linearization of the operator equation. Recall the definition of the Fréchet derivative from Chapter 2:

Definition 7.1. Let X, Y be normed spaces, and let U be an open subset of X. A mapping $F : U \to Y$ is called *Fréchet differentiable* at $\varphi \in U$ if there exists a bounded linear operator $F'[\varphi] : X \to Y$ such that

$$\|F(\varphi + h) - F(\varphi) - F'[\varphi]h\| = \mathrm{o}\left(\|h\|\right) \tag{7.1}$$

uniformly as $\|h\| \to 0$. $F'[\varphi]$ is called the *Fréchet derivative* of F at φ. F is called *Fréchet* differentiable if it is Fréchet differentiable at every point $\varphi \in U$. F is called *continuously differentiable* if F is differentiable and if $F' : U \to L(X, Y)$ is continuous.

We collect some basic properties of the Fréchet derivative.

Theorem 7.2. *Let $F : U \subset X \to Y$ be Fréchet differentiable, and let Z be a normed space.*

1. *The Fréchet derivative of F is uniquely determined.*

2. *If $G : U \to Y$ is Fréchet differentiable, then $\alpha F + \beta G$ is differentiable for all $\alpha, \beta \in \mathbb{R}$ (or $\mathbb{C}$) and*
$$(\alpha F + \beta G)'[\varphi] = \alpha F'[\varphi] + \beta G'[\varphi], \qquad \varphi \in U. \tag{7.2}$$

3. **(Chain rule)** *If $G : Y \to Z$ is Fréchet differentiable, the $G \circ F : U \to Z$ if Fréchet differentiable, and*
$$(G \circ F)'[\varphi] = G'[F(\varphi)]F', \qquad \varphi \in U. \tag{7.3}$$

4. **(Product rule)** *A bounded bilinear mapping $b : X \times Y \to Z$ is Fréchet differentiable, and*
$$b'[(\varphi_1, \varphi_2)](h_1, h_2) = b(\varphi_1, h_2) + b(h_1, \varphi_2) \tag{7.4}$$
for all $\varphi_1, h_1 \in X$ and $\varphi_2, h_2 \in Y$.

5. **(Derivative of the operator inverse)** *Let X, Y be Banach spaces, and assume that the set $U \subset L(X, Y)$ of operators which have a bounded inverse is not empty. Then the mapping* $\mathrm{inv} : U \to L(Y, X)$ *defined by* $\mathrm{inv}(T) := T^{-1}$. *is Fréchet differentiable, and*
$$\mathrm{inv}'[T]H = -T^{-1}HT^{-1} \tag{7.5}$$
for $T \in U$ and $H \in L(X, Y)$.

Remark 7.3. If Q is a normed space and $F_1 : Q \to X$ and $F_2 \to Y$ are Fréchet differentiable, then the product rule and the chain rule with $G = b$ and $F = (F_1, F_2) : Q \to X \times Y$ imply that $b(F_1, F_2)$ is Fréchet differentiable with

$$b(F_1, F_2)'[q]h = b(F_1(q), F_2'[q]h) + b(F_1'[q]h, F_2(q)).$$

If b is the product of real numbers, this is the ordinary product rule. Similarly, part 5 and the chain rule imply that for a mapping $T : Q \to U$ the function $\mathrm{inv} \circ T$ is Fréchet differentiable with

$$(\mathrm{inv} \circ T)'[q]h = -T(q)^{-1} \, (T'[q]h) \, T(q)^{-1}.$$

If $X = Y = \mathbb{R}$, this is the quotient rule.

Proof of Theorem 7.2 1) Let $\tilde{F}'$ be another Fréchet derivative of F. Then for all $\varphi \in U$, $h \in X$ and $\epsilon > 0$ we have

$$\begin{aligned} \|F'[\varphi]\epsilon h - \tilde{F}'[\varphi]\epsilon h\| \;\leq\; & \|F(\varphi + \epsilon h) - F(\varphi) - F'[\varphi]\epsilon h\| \\ & + \| - F(\varphi + \epsilon h) - F(\varphi) - \tilde{F}'[\varphi]\epsilon h\| = \mathrm{o}\,(\epsilon) \end{aligned}$$

as $\epsilon \to 0$. Dividing by ϵ shows that $F'[\varphi]h = \tilde{F}'[\varphi]h$.

2) This is obvious.

3) We have

$$\begin{aligned} & \|G(F(\varphi + h)) - G(F(\varphi)) - G'[F(\varphi)]F'[\varphi]h\| \\ \leq\; & \|G(F(\varphi + h)) - G(F(\varphi)) - G'[F(\varphi)](F(\varphi + h) - F(\varphi))\| \\ & + \|G'[\varphi]\| \, \|F(\varphi + h) - F(\varphi) - F'[\varphi]h\|. \end{aligned}$$

It follows from the Fréchet differentiability of F that $\|F(\varphi + h) - F(\varphi)\| = \|F'[\varphi]h\| + \mathrm{o}\,(\|h\|) = \mathrm{O}\,(\|h\|)$ as $\|h\| \to 0$. Since G is Fréchet differentiable, the first term on the right hand side is of order $\mathrm{o}\,(\|F(\varphi + h) - F(\varphi)\|) = \mathrm{o}\,(\|h\|)$. Due to the boundedness of $\|G'[\varphi]\|$ and the Fréchet differentiability of F, the second term is also of order $\mathrm{o}\,(\|h\|)$.

4) This follows from the identity

$$b(\varphi_1 + h_1, \varphi_2 + h_2) - b(\varphi_1, \varphi_2) - b(\varphi_1, h_2) - b(h_1, \varphi_2) = b(h_1, h_2)$$

and the boundedness of b.

5) By the Neumann series

$$(T + H)^{-1} = T^{-1}(I + HT^{-1})^{-1} = T^{-1} - T^{-1}HT^{-1} + R$$

for $\|H\| < \|T^{-1}\|$ with $R = \sum_{j=2}^{\infty} T^{-1}(-HT^{-1})^j$. As $\|R\| = \mathrm{O}\,(\|T^{-1}(-HT^{-1})^2\|) = \mathrm{O}\,(\|H\|^2)$, this shows (7.5). ☐

Next we want to define the integral over a continuous function $G : [a, b] \to X$ defined on a bounded interval $[a, b] \subset \mathbb{R}$ with values in a Hilbert space X. The functional $L : X \to \mathbb{R}$ (or $L : X \to \mathbb{C}$ if X is a complex Hilbert space) given by

$$L(\psi) := \int_a^b \langle G(t), \psi \rangle \; \mathrm{d}t$$

is (anti)linear and bounded since the $t \mapsto \|G(t)\|$ is continuous on the compact interval $[a,b]$ and hence bounded. Using the Riesz representation theorem, we define $\int_0^1 G(t)\,\mathrm{d}t$ to be the unique element in X satisfying

$$\left\langle \int_a^b G(t)\,\mathrm{d}t, \psi \right\rangle = L(\psi)$$

for all $\psi \in X$. It follows immediately from this definition that

$$\| \int_a^b G(t)\,\mathrm{d}t\| \leq \int_a^b \|G(t)\|\,\mathrm{d}t. \tag{7.6}$$

Lemma 7.4. *Let X, Y be Hilbert spaces and $U \subset X$ open. Moreover, let $\varphi \in U$ and $h \in X$ such that $\varphi + th \in U$ for all $0 \leq t \leq 1$. If $F : U \to Y$ is Fréchet differentiable, then*

$$F(\varphi + h) - F(\varphi) = \int_0^1 F'[\varphi + th]h\,\mathrm{d}t. \tag{7.7}$$

Proof. For a given $\psi \in X$ consider the function

$$f(t) := \langle F(\varphi + th), \psi \rangle\,, \qquad 0 \leq t \leq 1.$$

By the chain rule, f is differentiable and

$$f'(t) = \langle F'[\varphi + th]h, \psi \rangle\,.$$

Hence, by the Fundamental Theorem of Calculus, $f(1) - f(0) = \int_0^1 f'(t)\,\mathrm{d}t$ or

$$\langle F(\varphi + h) - F(\varphi), \psi \rangle = \left\langle \int_0^1 F'[\varphi + th]h\,\mathrm{d}t, \psi \right\rangle.$$

Since this holds true for all $\psi \in X$, we have proved the assertion. □

Lemma 7.5. *Let X, Y be Hilbert spaces and $U \subset X$ open. Let $F : U \to Y$ be Fréchet differentiable and assume that there exists a Lipschitz constant L such that*

$$\|F'[\varphi] - F'[\psi]\| \leq L\|\varphi - \psi\| \tag{7.8}$$

for all $\varphi, \psi \in U$. If $\varphi + th \in U$ for all $t \in [0,1]$, then (7.1) can be improved to

$$\|F(\varphi + h) - F(\varphi) - F'[\varphi]h\| \leq \frac{L}{2}\|h\|^2.$$

Proof. Using Lemma 7.4, (7.6) and (7.8) we get

$$\begin{aligned}
\|F(\varphi + h) - F(\varphi) - F'[\varphi]h\| &= \left\| \int_0^1 (F'[\varphi + th]h - F'[\varphi]h)\,\mathrm{d}t \right\| \\
&\leq \int_0^1 \|F'[\varphi + th]h - F'[\varphi]h\|\,\mathrm{d}t \\
&\leq \int_0^1 Lt\|h\|^2\,\mathrm{d}t = \frac{L}{2}\|h\|^2.
\end{aligned}$$

□

Compactness

Definition 7.6. Let X, Y be normed spaces, and let U be a subset of X. An operator $F : U \to Y$ is called *compact* if it maps bounded sets to relatively compact sets. It is called *completely continuous* if it is continuous and compact.

Unlike in the linear case, a nonlinear compact operator is not necessarily continuous (see exercises). We mention that the terms compact and complete continuous are sometimes used differently in the literature.

Let Z be another normed space, and let $F = H \circ G$ with $G : U \to Z$ and $H : Z \to Y$. It follows immediately from the definition that F is compact if G is compact and H is continuous, or if G maps bounded sets to bounded sets and H is compact. This facts can often be used to prove the compactness of nonlinear operators.

Theorem 7.7. *Let X, Y be normed spaces, and let $U \subset X$ be open. If $F : U \to Y$ is compact and X is infinite dimensional, then F^{-1} cannot be continuous, i.e. the equation $F(\varphi) = g$ is ill-posed.*

Proof. If F^{-1} does not exist, there is nothing to show. Assume that F^{-1} exists as a continuous operator. Then F^{-1} maps relatively compact sets to relatively compact sets. Hence every ball $B = \{\varphi \in X : \|\varphi - \varphi_0\| < \epsilon\}$ contained in U is relatively compact since $F(B)$ is relatively compact and $B = F^{-1}(F(B))$. This is not possible since $\dim X = \infty$. $\square$

The idea of Newton-type methods, which will be discussed in the next section, is to replace $F(\varphi) = g$ by the linearized equation

$$F'[\varphi_n]h_n = g - F(\varphi_n) \tag{7.9}$$

in each step and update φ_n by $\varphi_{n+1} := \varphi_n + h_n$. The following result implies that (7.9) inherits the ill-posedness of the original equation if F is completely continuous.

Theorem 7.8. *Let X be a normed space, Y a Banach space, and let $U \subset X$ be open. If $F : U \to Y$ is completely continuous and Fréchet diffentiable, then $F'[\varphi]$ is compact for all $\varphi \in U$.*

Proof. The proof relies on the fact that a subset K of the Banach space Y is relatively compact if and only if it is totally bounded, i.e. if for any $\epsilon > 0$ there exist $\psi_1, \dots, \psi_n \in Y$ such that $\min_{1 \le j \le n} \|\psi - \psi_j\| \le \epsilon$ for all $\psi \in K$ (cf. e.g. [Kre89, Section 1.4]).

We have to show that

$$K := \{F'[\varphi]h : h \in X, \|h\| \le 1\}$$

is relatively compact. Let $\epsilon > 0$. By the definition of the Fréchet derivative there exists $\delta > 0$ such that $\varphi + h \in U$ and

$$\|F(\varphi + h) - F(\varphi) - F'[\varphi]h\| \le \frac{\epsilon}{3}\|h\| \tag{7.10}$$

for all $\|h\| \leq \delta$. Since F is compact, the set $\{F(\varphi + \delta h) : h \in X, \|h\| \leq 1\}$ is relatively compact and hence totally bounded, i.e. there exist $h_1, \dots, h_n \in X$ with $\|h_j\| \leq 1$, $j = 1, \dots, n$ such that for all $h \in X$ satisfying $\|h\| \leq 1$ there exists and index $j \in \{1, \dots, n\}$ such that

$$\|F(\varphi + \delta h) - F(\varphi + \delta h_j)\| \leq \frac{\epsilon\delta}{3}. \tag{7.11}$$

Using (7.10) and (7.11) we obtain

$$\begin{aligned} \delta\|F'[\varphi]h - F'[\varphi]h_j\| &\leq \|F(\varphi + \delta h) - F(\varphi + \delta h_j)\| \\ +\| - F(\varphi + \delta h) + F(\varphi) + F'[\varphi]\delta h\| &+ \|F(\varphi + \delta h_j) - F(\varphi) - F'[\varphi]\delta h_j\| \leq \delta\epsilon, \end{aligned}$$

i.e. K is totally bounded. □

Apart from Theorem 7.10 the connection between the ill-posedness of a nonlinear problem and its linearization is less close than one may expect as counter-examples in [EKN89, Sch02] show.

The inhomogeneous medium scattering problem

A classical inverse problem is to determine the refractive index of a medium from measurements of far field patterns of scattered time-harmonic acoustic waves in this medium. Let u_{i} be an incident field satisfying the Helmholtz equation $\Delta u_{\mathrm{i}} + k^2 u_{\mathrm{i}} = 0$, e.g. $u_{\mathrm{i}}(x) = \exp(-ikx \cdot \theta)$ with $\theta \in S^{d-1} := \{x \in \mathbb{R}^d : |x| = 1\}$. The direct scattering problem is described by the system of equations

$$\Delta u + k^2 nu = 0, \qquad x \in \mathbb{R}^d, \tag{7.12a}$$

$$u_{\mathrm{i}} + u_{\mathrm{s}} = 0, \tag{7.12b}$$

$$\lim_{r\to\infty} r^{(d-1)/2}\left(\frac{\partial u_{\mathrm{s}}}{\partial r} - iku_{\mathrm{s}}\right) = 0, \qquad \text{uniformly for all } \hat{x} = \tfrac{x}{|x|}. \tag{7.12c}$$

Here $k > 0$ denotes the wave number, n is the refractive index of the medium, u_{s} is the scattered field, and u is the total field. Absorbing media are modelled by complex-valued refractive indices n. We assume that $\operatorname{Re} n \geq 0$, $\operatorname{Im} n \geq 0$. Moreover, we assume that n is constant and equal to 1 outside of the ball $B_\rho := \{x \in \mathbb{R}^3 : |x| \leq \rho\}$, $\rho > 0$, i.e.

$$n = 1 - a$$

with $\operatorname{supp} a \subset B_\rho$. The fundamental solution to the Helmholtz equation is given by

$$\Phi(x, y) = \frac{1}{4\pi}\frac{\exp(k|x-y|)}{|x-y|}, \qquad \text{for } d = 3, \tag{7.13}$$

$$\Phi(x, y) = \frac{i}{4}H_0^{(1)}(k|x-y|), \qquad \text{for } d = 2 \tag{7.14}$$

for $x \neq y$ where $H_0^{(1)}$ denotes the Hankel function of the first kind of order 0.

A basic tool for the analysis of the direct medium scattering problem is a reformulation of the system (7.12) as an integral equation of the second kind:

$$u(x) + k^2 \int_{B_\rho} \Phi(x-y)a(y)u(y)\,\mathrm{d}y = u_{\mathrm{i}}(x), \qquad x \in \mathbb{R}^d. \tag{7.15}$$

(7.15) is called the *Lippmann-Schwinger equation.*

Theorem 7.9. *Assume that $a \in C^1(\mathbb{R}^d)$ satisfies $\operatorname{supp} a \subset B_\rho$. Then any solution $u \in C^2(\mathbb{R}^d)$ to (7.12) satisfies (7.15). Vice versa, let $u \in C(B_\rho)$ be a solution to (7.15). Then*

$$u_s(x) := -k^2 \int_{B_\rho} \Phi(x-y)a(y)u(y)\,\mathrm{d}y, \qquad x \in \mathbb{R}^d \tag{7.16}$$

belongs to $C^2(\mathbb{R}^d)$ and satisfies (7.12).

To prove Theorem 7.9 we need a few preparations. All function $v \in C^2(B_\rho) \cap C^1(\overline{B_\rho})$ satisfy *Green's representation formula*

$$\begin{aligned} v(x) &= \int_{\partial B_\rho} \left\{ \frac{\partial v}{\partial \nu}(y)\Phi(x,y) - v(y)\frac{\partial \Phi(x,y)}{\partial \nu(y)} \right\} \mathrm{d}s(y) \\ &\quad - \int_{B_\rho} \left\{ \Delta v(y) + k^2 v(y) \right\} \Phi(x,y)\,\mathrm{d}y, \qquad x \in B_\rho \end{aligned} \tag{7.17}$$

where ν denotes the outer normal vector on ∂B_ρ (cf. [CK97b, Theorem 2.1]).

Moreover, we need the following properties of the volume potential

$$(V\varphi)(x) := \int_{\mathbb{R}^d} \Phi(x,y)\varphi(y)\,\mathrm{d}y, \qquad x \in \mathbb{R}^d. \tag{7.18}$$

Theorem 7.10. *Let $\operatorname{supp}\varphi \subset B_\rho$. If $\varphi \in C(\mathbb{R}^d)$, then $V\varphi \in C^1(\mathbb{R}^d)$, and if $\varphi \in C^1(\mathbb{R}^d)$, then $V\varphi \in C^2(\mathbb{R}^d)$. In the latter case*

$$(\Delta + k^2)(V\varphi) = -\varphi \tag{7.19}$$

holds true.

The mapping properties of V stated in this theorem are crude, but sufficient for our purposes. In terms of Hölder and Sobolev spaces, V is a smoothing operator of order 2.

Proof of Theorem 7.9. Let $u \in C^2(\mathbb{R}^d)$ be a solution to (7.12), and let $x \in B_\rho$. It follows from Green's representation formula (7.17) with $v = u$ and $\Delta u + k^2 u = k^2 a u$ that

$$u(x) = \int_{\partial B_\rho} \left\{ \frac{\partial u}{\partial \nu}(y)\Phi(x,y) - u(y)\frac{\partial \Phi(x,y)}{\partial \nu(y)} \right\} \mathrm{d}s(y) - k^2 \int_{B_\rho} \Phi(x,y)a(y)u(y)\,\mathrm{d}y \tag{7.20}$$

where ν denotes the outer normal vector on ∂B_ρ. Green's formula (7.17) applied to u_{i} gives

$$u_{\mathrm{i}}(x) = \int_{B_\rho} \left\{ \frac{\partial u_{\mathrm{i}}}{\partial \nu}\Phi(x,y) - u_{\mathrm{i}}(y)\frac{\partial \Phi(x,y)}{\partial \nu(y)} \right\} \mathrm{d}s(y). \tag{7.21}$$

Finally, we choose $R > \rho$ and apply Green's second theorem to $\Phi(x, \cdot)$ and u_s in $B_R \setminus B_\rho$

$$\begin{aligned}
&\int_{\partial B_\rho} \left\{ \frac{\partial u_\mathrm{s}}{\partial \nu}(y)\Phi(x,y) - u_\mathrm{s}(y)\frac{\partial \Phi(x,y)}{\partial \nu(y)} \right\} \mathrm{d}s(y) \\
=\ & \int_{\partial B_R} \left\{ \frac{\partial u_\mathrm{s}}{\partial \nu}(y)\Phi(x,y) - u_\mathrm{s}(y)\frac{\partial \Phi(x,y)}{\partial \nu(y)} \right\} \mathrm{d}s(y) \\
=\ & \int_{\partial B_R} \left\{ \frac{\partial u_\mathrm{s}}{\partial \nu}(y) - iku_\mathrm{s}(y) \right\} \Phi(x,y)\,\mathrm{d}s(y) + \int_{\partial B_R} u_\mathrm{s}(y) \left\{ ik\Phi(x,y) - \frac{\partial \Phi(x,y)}{\partial \nu(y)} \right\} \mathrm{d}s(y)
\end{aligned}$$

where ν on ∂B_R is the outer normal vector. Since both u_s and $\Phi(x,\cdot)$ satisfy the Sommerfeld radiation condition and since (7.12c) implies that $|u_\mathrm{s}(y)| = \mathrm{O}\left(|y|^{-(d-1)/2}\right)$ as $|y| \to \infty$, the right hand side of the last equation tends to 0 as $R \to \infty$. Combining this result with (7.20) and (7.21) and using (7.12b) shows that (7.15) is satisfied.

Vice versa, let $u \in C(B_\rho)$ be a solution to (7.15). Since $\Phi(\cdot, y)$ satisfies the Sommerfeld radiation condition uniformly for $y \in B_\rho$, u_s defined by (7.16) satisfies (7.12c). Moreover, $u_\mathrm{s} \in C^1(\mathbb{R}^d)$ by Theorem 7.10. Now the second statement in Theorem 7.10 and the assumption $a \in C^1_c(\mathbb{R}^d)$ imply that $u_\mathrm{s} \in C^2(\mathbb{R}^d)$ and that

$$\Delta u_\mathrm{s} + k^2 u_\mathrm{s} - k^2 a u = 0.$$

Since $\Delta u_\mathrm{i} + k^2 u_\mathrm{i} = 0$, $u = u_\mathrm{s} + u_\mathrm{i}$ satisfies (7.12a). □

Theorem 7.11. *The Lippmann-Schwinger equation has a unique solution $u \in C(B_\rho)$ if $\|a\|_\infty < (k^2\|V\|_\infty)^{-1}$.*

Proof. Under the given assumption we have $\|k^2 V M_a\|_\infty < 1$ where $M_a : C(B_\rho) \to C(B_\rho)$ is defined by $M_a v := av$. Therefore, the operator $I + k^2 V M_a$ has a bounded inverse given by the Neumann series. □

Using Riesz theory it can be shown that the smallness assumption on $\|a\|_\infty$ in Theorem 7.11 can be dropped. However, it is not trivial to show uniqueness of a solution to (7.15) or equivalently, a solution to (7.12) (cf. [CK97b, Häh98, Kir96]).

Recall the definition of the far-field pattern from (1.19). As straightforward computation shows that for $d = 3$ the far field pattern of $\Phi(\cdot, y)$ is given by $\Phi_\infty(\hat{x}, y) = \gamma_3 \exp(-ik\hat{x}\cdot y)$, $\gamma_3 = 1/(4\pi)$ and that $\Phi(\cdot, y)$ satisfies (1.19) uniformly with respect to $y \in B_\rho$. By the asymptotic behavior of the Hankel functions for large arguments (cf. [CK97b, Leb65]), the same holds true for $d = 2$ with $\gamma_2 = e^{i\pi/4}/\sqrt{8\pi k}$. Hence, for both $d = 2$ and $d = 3$ the far field pattern of u_s is given by

$$u_\infty(\hat{x}) = -k^2\gamma_d \int_{B_\rho} \exp(-ik\hat{x}\cdot y) a(y) u(y)\,\mathrm{d}y, \qquad \hat{x} \in S^{d-1}. \tag{7.22}$$

The right hand side of this equation defines a linear integral operator $E : C(B_\rho) \to L^2(S^{d-1})$ with $u_\infty = E(au)$.

Now we turn to the inverse problem to recover a from measurements of u_∞. Since for $d = 3$ the far field pattern u_∞ is a function of two variables whereas the unknown coefficient a is a function of three variables, we cannot expect to be able to reconstruct a from far-field measurements corresponding to just one incident field. A similar argument holds for $d = 2$. Therefore, we consider incident fields $u_\mathrm{i}(x) = u_\mathrm{i}(x, \theta) = \exp(ik\theta \cdot x)$ from all directions $\theta \in S^{d-1}$ and denote the corresponding solutions to the direct problem by $u(x, \theta), u_\mathrm{s}(x, \theta)$, and $u_\infty(x, \theta)$. The direct solution operator $F_\mathrm{IM} : D(F_\mathrm{IM}) \subset C_c^1(B_\rho) \to L^2(S^{d-1} \times S^{d-1})$ is defined by

$$(F_\mathrm{IM}(a))(\hat{x}, \theta) := u_\infty(\hat{x}, \theta), \qquad \hat{x}, \theta \in S^{d-1} \tag{7.23}$$

where $u_\infty(\cdot, \theta)$ is the far field pattern corresponding to the solution $u_\mathrm{s}(\cdot, \theta)$ of (7.12) with $n = 1-a$ and $u_\mathrm{i} = u_\mathrm{i}(\cdot, \theta)$. Here $C_c^1(B_\rho)$ is the space of all continuously differentiable function on B_ρ with $\operatorname{supp} a \subset B_\rho$, and the domain of F_IM incorporates the physical restrictions on $n = 1 - a$:

$$D(F_\mathrm{IM}) := \{a \in C_c^1(B_\rho) : \mathrm{Re}(1 - a) \geq 0, \mathrm{Im}(a) \leq 0\}.$$

It can be shown that F_IM is one-to-one, i.e. that the far-field patterns $u_\infty(\cdot, \theta)$ for all directions $\theta \in S^2$ of the incident wave determine the refractive index $n = 1 - a$ uniquely (cf. [CK97b, Häh98, Kir96]).

Theorem 7.12. *1. The operator $G : D(F_\mathrm{IM}) \to C(\mathbb{R}^d \times S^{d-1})$ is Fréchet differentiable with respect to the supremum norm on $D(F_\mathrm{IM})$, and $u' := G'[a]h$ satisfies the integral equation*

$$u' + k^2 V(au') = -k^2 V(hu) \tag{7.24}$$

for all $a \in D(F_\mathrm{IM})$ and $h \in C_c^1(B_\rho)$ with $u := G(a)$.

2. F_IM is Fréchet differentiable with respect to the maximum norm on $D(F_\mathrm{IM})$.

Proof. 1) The mapping $D(F_\mathrm{IM}) \to L(C(B_\rho \times S^{d-1}))$, $a \mapsto M_a$ where $(M_a v)(x, \theta) := a(x)v(x, \theta)$ is linear and bounded with respect to the supremum norm and hence Fréchet differentiable. It follows from Theorem 7.2, Parts 3 and 5 that the the mapping $D(F_\mathrm{IM}) \to L(C(B_\rho \times S^{d-1}))$ $a \mapsto (I + k^2 V M_a)^{-1}$ is Fréchet-differentiable. An application of the chain rule yields the Fréchet differentiability of $G(a) = (I + k^2 V M_a)^{-1} u_\mathrm{i}$. Now (7.24) can be derived by differentiating both sides of the Lippmann-Schwinger equation (7.15) with respect to a using the product rule.

2) Since $F(a) = E(M_a G(a))$, this follows from the first part, the product and the chain rule. $\square$

8. Nonlinear Tikhonov regularization

Let X, Y be Hilbert spaces and $F : D(F) \subset X \to Y$ a continuous operator. We want to solve the operator equation

$$F(\varphi) = g \tag{8.1}$$

given noisy data $g^\delta \in Y$ satisfying $\|g^\delta - g\| \le \delta$. Let $\varphi^\dagger$ denote the exact solution. We assume that the solution to (8.1) with exact data $g = F(\varphi^\dagger)$ is unique, i.e. that

$$F(\varphi) = g \qquad \Rightarrow \qquad \varphi = \varphi^\dagger \tag{8.2}$$

although many of the results below can be obtained in a modified form without this assumption.

The straightforward generalization of linear Tikhonov regularization leads to the minimization problem

$$\|F(\varphi) - g^\delta\|^2 + \alpha\|\varphi - \varphi_0\|^2 = \min! \tag{8.3}$$

over $\varphi \in D(F)$ where φ_0 denotes some initial guess of $\varphi^\dagger$. The mapping $D(F) \to \mathbb{R}$, $\varphi \mapsto \|F(\varphi) - g^\delta\|^2 + \alpha\|\varphi - \varphi_0\|^2$ is called *(nonlinear) Tikhonov functional.* Note that as opposed to the linear case, the element $0 \in X$ does not have a special role any more. Therefore, $\varphi_0 = 0$ is as good as any other initial guess.

As opposed to the linear case it is not clear under the given assumptions if the minimization problem (8.3) has a solution. We will have to impose additional assumptions on F to ensure existence. Moreover, even if (8.3) has a unique solution for $\alpha = 0$ there may be more than one global minimizer of (8.3) for $\alpha > 0$.

As in the linear case, it is sometimes useful to consider other penalty terms in (8.3) than $\alpha\|\varphi - \varphi_0\|^2$.

Weak convergence in Hilbert spaces

Definition 8.1. We say that a sequence (φ_n) in a Hilbert space X *converges weakly to* $\varphi \in X$ and write $\varphi_n \rightharpoonup \varphi$ as $n \to \infty$ if

$$\langle \varphi_n, \psi \rangle \to \langle \varphi, \psi \rangle\,, \qquad n \to \infty \tag{8.4}$$

for all $\psi \in X$.

If $\tilde\varphi$ is another weak limit of the sequence (φ_n), then $\langle \varphi - \tilde\varphi, \psi \rangle = 0$ for all $\psi \in X$. Choosing $\psi = \varphi - \tilde\varphi$ shows that $\varphi = \tilde\varphi$, i.e. *a weak limit of a sequence is uniquely determined.* It follows from the Cauchy-Schwarz inequality that strong convergence implies weak convergence. The following example shows that the reverse implication is not true in general.

Example 8.2. Let $\{\varphi_n : n \in \mathbb{N}\}$ be an orthonormal system in X. Then $\|\varphi_n - \varphi_m\| = \sqrt{2}$ for $m \neq n$, i.e. the sequence (φ_n) cannot be strongly convergent. On the other hand, it follows from Bessel's inequality

$$\sum_{n=1}^{\infty} |\langle \varphi_n, \psi \rangle|^2 \leq \|\psi\|^2$$

for $\psi \in X$ that $|\langle \varphi_n, \psi \rangle|^2 \to 0$ as $n \to \infty$. Therefore, $\varphi_n \rightharpoonup 0$.

Lemma 8.3. *If $T \in L(X, Y)$, then T is weakly continuous, i.e. $\varphi_n \rightharpoonup \varphi$ implies $T\varphi_n \rightharpoonup T\varphi$ as $n \to \infty$.*

Proof. Let $\varphi_n \rightharpoonup \varphi$. Then for any $\psi \in Y$ we have

$$\langle T\varphi_n, \psi \rangle = \langle \varphi_n, T^*\psi \rangle \to \langle \varphi, T^*\psi \rangle = \langle T\varphi, \psi \rangle .$$

□

Lemma 8.4. *If $\varphi_n \rightharpoonup \varphi$, then $\limsup_{n\to\infty} \|\varphi_n\| \geq \|\varphi\|$, i.e. the norm is weakly lower semicontinuous.*

Proof. We have $\langle \varphi_n, \varphi \rangle \to \|\varphi\|^2$ as $n \to \infty$. It follows from the Cauchy-Schwarz inequality that $\limsup_{n\to\infty} \|\varphi_n\|\|\varphi\| \geq \|\varphi\|^2$. This implies the assertion. □

Theorem 8.5. *Every bounded sequence has a weakly convergent subsequence.*

Proof. Let $(\varphi_n)_{n\in\mathbb{N}}$ be some sequence in X such that $\|\varphi_n\| \leq 1$ for all $n \in \mathbb{N}$, and let $\{e_j : j \in \mathbb{N}\}$ be a complete orthonormal system in $\tilde{X} := \overline{\mathrm{span}\{\varphi_n : n \in \mathbb{N}\}}$. Since $\langle \varphi_n, e_1 \rangle$ is a bounded sequence of complex numbers, there exists a convergent subsequence $\left\langle \varphi_{n_1(k)}, e_1 \right\rangle$. Since $\left\langle \varphi_{n_1(k)}, e_2 \right\rangle$ is bounded, there exists a subsequence $n_2(k)$ of $n_1(k)$ such that $\left\langle \varphi_{n_2(k)}, e_2 \right\rangle$ is convergent. By continuing this process, we obtain subsequences $n_l(k)$ for all $l \in \mathbb{N}$ such that $\left\langle \varphi_{n_l(k)}, e_l \right\rangle$ is convergent and $n_{l+1}(k)$ is a subsequence of $n_l(k)$. The diagonal sequence $(\varphi_{n_l(l)})_{l\in\mathbb{N}}$ has the property that $\left\langle \varphi_{n_l(l)}, e_k \right\rangle$ converges to some $\xi_k \in \mathbb{C}$ as $l \to \infty$ for all $k \in \mathbb{N}$. Then $\varphi := \sum_{k\in\mathbb{N}} \xi_k e_k$ defines an element of $\tilde{X}$ with $\|\varphi\| \leq 1$ since for all $K \in \mathbb{N}$ we have

$$\sum_{k=1}^{K} |\xi|^2 = \lim_{l\to\infty} \sum_{k=1}^{K} |\left\langle \varphi_{n_l(l)}, e_k \right\rangle|^2 \leq \limsup_{l\to\infty} \|\varphi_{n_l(l)}\|^2 \leq 1.$$

We have to show that $\left\langle \varphi_{n_l(l)}, \psi \right\rangle \to \langle \varphi, \psi \rangle$ for all $\psi \in X$. It suffices to consider $\psi \in \tilde{X}$ since $\varphi_{n_l(l)}, \varphi \in \tilde{X}$. Let $\epsilon > 0$ and choose $K \in \mathbb{N}$ such that $\sum_{k=K+1}^{\infty} |\langle \psi, e_k \rangle|^2 \leq (\epsilon/4)^2$. There exists $L > 0$ such that

$$\left| \langle \varphi - \varphi_{n_l(l)}, \sum_{k=1}^{K} \langle \psi, e_k \rangle e_k \rangle \right| \leq \frac{\epsilon}{2}$$

for $l \geq L$. Now it follows from the triangle inequality and the Cauchy-Schwarz inequality that $|\left\langle \varphi - \varphi_{n_l(l)}, \psi \right\rangle| \leq \epsilon$ for $l \geq L$. $\square$

Definition 8.6. A subset K of a Hilbert space X is called *weakly closed* if it contains the weak limits of all weakly convergent sequences contained in K.
An operator $F : \mathcal{D}(F) \subset X \to Y$ is called *weakly closed* if its graph $\operatorname{gr} F := \{(\varphi, F(\varphi)) : \varphi \in D(F)\}$ is weakly closed in $X \times Y$, i.e. if $\varphi_n \rightharpoonup \varphi$ and $F(\varphi_n) \rightharpoonup g$ imply that $\varphi \in D(F)$ and $F(\varphi) = g$.

Note that if F is weakly continuous and if $D(F)$ is weakly closed, then F is weakly closed. The following result gives sufficient conditions for the weak closedness of $D(F)$.

Theorem 8.7. *If $K \subset X$ is convex and closed, then K is weakly closed.*

Proof. Let (φ_n) be some sequence in K converging weakly to some $\varphi \in X$. By Remark 3.2 there exists a best approximation $\psi \in K$ to φ in K satisfying

$$\operatorname{Re} \langle u - \psi, \varphi - \psi \rangle \leq 0$$

for all $u \in K$. Substituting $u = \varphi_n$ and taking the limit $n \to \infty$ shows that $\|\varphi - \psi\|^2 \leq 0$. Hence, $\psi = \varphi$, and in particular $\varphi \in K$. $\square$

Convergence analysis

Theorem 8.8. *Assume that F is weakly closed. Then the Tikhonov functional (8.3) has a global minimum for all $\alpha > 0$.*

Proof. Let $I := \inf_{\varphi \in D(F)} \|F(\varphi) - g^\delta\|^2 + \alpha\|\varphi - \varphi_0\|^2$ denote the infimum of the Tikhonov functional and choose a sequence (φ_n) in $D(F)$ such that

$$\|F(\varphi_n) - g^\delta\|^2 + \alpha\|\varphi_n - \varphi_0\|^2 \leq I + \frac{1}{n}. \tag{8.5}$$

Since $\alpha > 0$, φ_n is bounded. Hence, by Theorem 8.5 there exists a weakly convergent subsequence $\varphi_{n(k)}$ with a weak limit $\varphi \in X$. Moreover, it follows from (8.5) that $F(\varphi_{n(k)})$ is bounded. Therefore, there exists a further subsequence such that $F(\varphi_{n(k(l))})$ is weakly convergent. Now the weak closedness of F implies that $\varphi \in D(F)$ and that $F(\varphi_{n(k(l))}) \rightharpoonup F(\varphi)$ as $l \to \infty$. It follows from Lemma 8.4 that

$$\|F(\varphi) - g^\delta\|^2 + \alpha\|\varphi - \varphi_0\|^2 \leq \limsup_{n\to\infty} \left\{\|F(\varphi_n) - g^\delta\|^2 + \alpha\|\varphi_n - \varphi_0\|^2\right\} \leq I.$$

Hence, φ is a global minimum of the Tikhonov functional. $\square$

We do not know if a solution to (8.3) is unique. Nevertheless, it can be shown that an arbitrary sequence of minimizers converges to the exact solution $\varphi^\dagger$ as the noise level δ tends to 0. In analogy to Definition 3.8 this means that nonlinear Tikhonov regularization is a regularization method.

Theorem 8.9. *Assume that F is weakly closed and that (8.2) holds true. Let $\alpha = \overline{\alpha}(\delta)$ be chosen such that*

$$\overline{\alpha}(\delta) \to 0 \quad \text{and} \quad \delta^2/\overline{\alpha}(\delta) \to 0 \qquad \text{as } \delta \to 0. \tag{8.6}$$

If g^{δ_k} is some sequence in Y such that $\|g^{\delta_k} - g\| \le \delta_k$ and $\delta_k \to 0$ as $k \to \infty$, and if $\varphi_{\alpha_k}^{\delta_k}$ denotes a solution to (8.3) with $g^\delta = g^{\delta_k}$ and $\alpha = \alpha_k = \overline{\alpha}(\delta_k)$, then $\|\varphi_{\alpha_k}^{\delta_k} - \varphi^\dagger\| \to 0$ as $k \to \infty$.

Proof. Since $\varphi_{\alpha_k}^{\delta_k}$ minimizes the Tikhonov functional, we have

$$\begin{aligned}\|F(\varphi_{\alpha_k}^{\delta_k}) - g^{\delta_k}\|^2 + \alpha_k\|\varphi_{\alpha_k}^{\delta_k} - \varphi_0\|^2 &\le \|F(\varphi^\dagger) - g^{\delta_k}\|^2 + \alpha_k\|\varphi^\dagger - \varphi_0\|^2 \\ &\le \delta_k^2 + \alpha_k\|\varphi^\dagger - \varphi_0\|^2.\end{aligned}$$

The assumptions $\delta_k \to 0$ and $\alpha_k \to 0$ imply that

$$\lim_{k\to\infty} F(\varphi_{\alpha_k}^{\delta_k}) = g, \tag{8.7}$$

and the assumption $\delta_k^2/\alpha_k \to 0$ yields

$$\limsup_{k\to\infty} \|\varphi_{\alpha_k}^{\delta_k} - \varphi_0\|^2 \le \limsup_{k\to\infty} \left\{\delta_k^2/\alpha_k + \|\varphi^\dagger - \varphi_0\|^2\right\} = \|\varphi^\dagger - \varphi_0\|^2. \tag{8.8}$$

It follows from (8.8) and Theorem 8.5 that there exists a weakly convergent subsequence of $\varphi_{\alpha_k}^{\delta_k}$ with some weak limit $\varphi \in X$. By virtue of the weak closedness of F we have $\varphi \in D(F)$ and $F(\varphi) = g$, so $\varphi = \varphi^\dagger$ by (8.2).

It remains to show that $\|\varphi_{\alpha_k}^{\delta_k} - \varphi^\dagger\| \to 0$. Assume on the contrary that there exists $\epsilon > 0$ such that

$$\|\varphi_{\alpha_k}^{\delta_k} - \varphi^\dagger\| \ge \epsilon \tag{8.9}$$

for some subsequence of $(\varphi_{\alpha_k}^{\delta_k})$ which may be assumed to be identical to $(\varphi_{\alpha_k}^{\delta_k})$ without loss of generality. By the argument above, we may further assume that $\varphi_{\alpha_k}^{\delta_k} \rightharpoonup \varphi^\dagger$. Since

$$\|\varphi_{\alpha_k}^{\delta_k} - \varphi^\dagger\|^2 = \|\varphi_{\alpha_k}^{\delta_k} - \varphi_0\|^2 + \|\varphi_0 - \varphi^\dagger\|^2 + 2\,\mathrm{Re}\left\langle \varphi_{\alpha_k}^{\delta_k} - \varphi_0, \varphi_0 - \varphi^\dagger\right\rangle,$$

it follows from (8.8) that

$$\limsup_{k\to\infty} \|\varphi_{\alpha_k}^{\delta_k} - \varphi^\dagger\|^2 \le 2\|\varphi^\dagger - \varphi_0\|^2 + 2\,\mathrm{Re}\left\langle \varphi^\dagger - \varphi_0, \varphi_0 - \varphi^\dagger\right\rangle = 0.$$

This contradicts (8.9). □

As in the linear case we need a source condition to establish estimates on the rate of convergence as $\delta \to 0$. Source conditions for nonlinear problems usually involve $\varphi^\dagger - \varphi_0$ instead of $\varphi^\dagger$ because of the loss of the special role of $0 \in X$.

Theorem 8.10. *Assume that F is weakly closed and Fréchet differentiable, that $D(F)$ is convex, and that there exists a Lipschitz constant $L > 0$ such that*

$$\|F'[\varphi] - F'[\psi]\| \le L\|\varphi - \psi\| \tag{8.10}$$

for all $\varphi, \psi \in D(F)$. Moreover, assume that the source condition

$$\begin{aligned} \varphi^\dagger - \varphi_0 &= F'[\varphi^\dagger]^* w, && (8.11a)\\ L\|w\| &< 1 && (8.11b) \end{aligned}$$

is satisfied for some $w \in Y$ and that a parameter choice rule $\alpha = c\delta$ with some $c > 0$ is used. Then there exists a constant $C > 0$ independent of δ such that every global minimum φ_α^δ of the Tikhonov functional satisfies the estimates

$$\begin{aligned} \|\varphi_\alpha^\delta - \varphi^\dagger\| &\leq C\sqrt{\delta}, && (8.12a)\\ \|F(\varphi_\alpha^\delta) - g\| &\leq C\delta. && (8.12b) \end{aligned}$$

Proof. As in the proof of the Theorem 8.9 we use the inequality

$$\|F(\varphi_\alpha^\delta) - g^\delta\|^2 + \alpha\|\varphi_\alpha^\delta - \varphi_0\|^2 \leq \delta^2 + \alpha\|\varphi^\dagger - \varphi_0\|^2$$

for global minimum φ_α^δ of the Tikhonov functional. Since this estimate would not give the optimal rate for $\|F(\varphi_\alpha^\delta) - g^\delta\|^2$, we add $\alpha\|\varphi_\alpha^\delta - \varphi^\dagger\|^2 - \alpha\|\varphi_\alpha^\delta - \varphi_0\|^2$ on both sides to obtain

$$\begin{aligned} \|F(\varphi_\alpha^\delta) - g^\delta\|^2 + \alpha\|\varphi_\alpha^\delta - \varphi^\dagger\|^2 &\leq \delta^2 + 2\alpha \operatorname{Re}\left\langle \varphi^\dagger - \varphi_0, \varphi^\dagger - \varphi_\alpha^\delta \right\rangle\\ &= \delta^2 + 2\alpha \operatorname{Re}\left\langle w, F'[\varphi^\dagger](\varphi^\dagger - \varphi_\alpha^\delta) \right\rangle. \end{aligned}$$

Here (8.11a) has been used in the second line. Using the Cauchy-Schwarz inequality and inserting the inequality

$$\begin{aligned} \|F'[\varphi^\dagger](\varphi^\dagger - \varphi_\alpha)\| &\leq \frac{L}{2}\|\varphi_\alpha^\delta - \varphi^\dagger\|^2 + \|F(\varphi_\alpha^\delta) - F(\varphi^\dagger)\|\\ &\leq \frac{L}{2}\|\varphi_\alpha^\delta - \varphi^\dagger\|^2 + \|F(\varphi_\alpha^\delta) - g^\delta\| + \delta, \end{aligned}$$

which follows from Lemma 7.5, yields

$$\|F(\varphi_\alpha^\delta) - g^\delta\|^2 + \alpha\|\varphi_\alpha^\delta - \varphi^\dagger\|^2 \leq \delta^2 + 2\alpha\delta\|w\| + 2\alpha\|w\|\,\|F(\varphi_\alpha^\delta) - g^\delta\| + \alpha L\|w\|\|\varphi_\alpha^\delta - \varphi^\dagger\|^2,$$

and hence

$$\left(\|F(\varphi_\alpha^\delta) - g^\delta\| - \alpha\|w\|\right)^2 + \alpha(1 - L\|w\|)\|\varphi_\alpha^\delta - \varphi^\dagger\|^2 \leq (\delta + \alpha\|w\|)^2.$$

Therefore,

$$\|F(\varphi_\alpha^\delta) - g^\delta\| \leq \delta + 2\alpha\|w\|$$

and, due to (8.11b),

$$\|\varphi_\alpha^\delta - \varphi^\dagger\| \leq \frac{\delta + \alpha\|w\|}{\sqrt{\alpha(1 - L\|w\|)}}.$$

With the parameter choice rule $\alpha = c\delta$, this yields (8.12). □

By virtue of (5.14), condition (8.11a) is equivalent to

$$\varphi^\dagger - \varphi_0 = (F'[\varphi^\dagger]^* F'[\varphi^\dagger])^{1/2}\tilde{w}, \qquad L\|\tilde{w}\| < 1$$

for some $\tilde{w} \in X$. Hence, for linear problems Theorem 8.10 reduces to a special case of Theorem 5.2.

Theorem 8.10 was obtained by Engl, Kunisch and Neubauer [EKN89]. Other Hölder-type source conditions with a-priori parameter choice rules are treated in [Neu89], and a posteriori parameter choice rules were investigated in [SEK93]. For further references and results we refer to [EHN96, Chapter 10].

9. Iterative regularization methods

Let X, Y be Hilbert spaces, $D(F) \subset X$ open, and let $F : D(F) \to Y$ be a continuously Fréchet differentiable operator. We consider the nonlinear operator equation

$$F(\varphi) = g \tag{9.1}$$

and assume that for the given exact data g there exists a unique solution $\varphi^\dagger \in D(F)$ to (9.1). Moreover, we assume that some initial guess $\varphi_0 \in D(F)$ to $\varphi^\dagger$ is given, which will serve as starting point of the iteration. Finally, as before, $g^\delta \in Y$ denote noisy data satisfying

$$\|g - g^\delta\| \leq \delta. \tag{9.2}$$

Examples

Since the gradient of the cost functional $I(\varphi) := \|F(\varphi) - g^\delta\|^2$ is given by $I'[\varphi]h = 2\,\mathrm{Re}\left\langle F'[\varphi]^*(F(\varphi) - g^\delta), h\right\rangle$, the minimization of I by moving in the direction of steepest descent leads to the iteration formula

$$\varphi_{n+1}^\delta = \varphi_n^\delta - \mu F'[\varphi_n^\delta]^*(F(\varphi_n^\delta) - g^\delta), \qquad n = 0, 1, \ldots \tag{9.3}$$

known as *nonlinear Landweber iteration*. The step size parameter $\mu > 0$ should be chosen such that $\mu\|F'[\varphi_n^\delta]^* F'[\varphi_n^\delta]\| \leq 1$ for all n.

It is well known that the convergence of Landweber iteration is very slow. We may expect faster convergence from Newton-type methods. Recall that Newton's method consists in replacing (9.1) in the nth step by the linearized equation

$$F'[\varphi_n^\delta]h_n = y^\delta - F(\varphi_n^\delta) \tag{9.4}$$

and updating φ_n^δ by

$$\varphi_{n+1}^\delta = \varphi_n^\delta + h_n. \tag{9.5}$$

If (9.1) is ill-posed then in general $F'[\varphi_n^\delta]^{-1}$ will not be bounded and the range of $F'[\varphi_n^\delta]$ will not be closed (cf. Theorem 7.8). Therefore, the standard Newton method may not be well defined even for exact data since we cannot guarantee that $y - F(\varphi_n^\delta) \in R(F'[\varphi_n^\delta])$ in each step. Even if φ_n^δ is well defined for $n \geq 1$, it does not depend continuously on the data. Therefore, some sort of regularization has to be employed. In principle any regularization method for linear ill-posed problems can be used to compute a stable solution to (9.4). Tikhonov regularization with regularization parameter $\alpha_n > 0$ leads to the iteration formula

$$\varphi_{n+1}^\delta = \varphi_n^\delta + (\alpha_n I + F'[\varphi_n^\delta]^* F'[\varphi_n^\delta])^{-1} F'[\varphi_n^\delta]^* (g^\delta - F(\varphi_n^\delta)). \tag{9.6}$$

By Theorem 2.1 this is equivalent to solving the minimization problem

$$h_n = \mathrm{argmin}_{h \in X} \left(\|F'[\varphi_n^\delta] h + F(\varphi_n^\delta) - g^\delta\|^2 + \alpha_n \|h\|^2 \right) \tag{9.7}$$

using the update formula (9.5). This method is known as the *Levenberg-Marquardt algorithm*. The original idea of the Levenberg-Marquardt algorithm is to minimize $\|F(\varphi) - g^\delta\|^2$ within a *trust region* $\{\varphi \in X : \|\varphi - \varphi_n^\delta\| \leq \rho_n\}$ in which the approximation $F(x) \approx F(\varphi_n^\delta) + F'[\varphi_n^\delta](x - \varphi_n^\delta)$ is assumed to be valid. Here α_n plays the rôle of a Lagrange parameter. Depending on the agreement of the actual residual $\|F(\varphi_{n+1}^\delta) - g^\delta\|$ and the predicted residual $\|F(\varphi_n^\delta) + F'[\varphi_n^\delta](x - \varphi_n^\delta) - g^\delta\|$ the trust region radius ρ_n is enlarged or reduced. If $\|\varphi_0 - \varphi^\dagger\|$ is sufficiently small, then a simple parameter choice rule of the form

$$\alpha_n = \alpha_0 \left(\frac{1}{2}\right)^n \tag{9.8}$$

is usually sufficient and computationally more effective for ill-posed problems (cf. [Han97a, Hoh99]).

In [Bak92] Bakushinskii suggested a related method called *iteratively regularized Gauss-Newton method* where (9.7) is replaced by

$$h_n = \mathrm{argmin}_{h \in X} \left(\|F'[\varphi_n^\delta] h + F(\varphi_n^\delta) - g^\delta\|^2 + \alpha_n \|h + \varphi_n^\delta - \varphi_0\|^2 \right) \tag{9.9}$$

The penalty term in (9.9) involves the distance of the new iterate $\varphi_{n+1}^\delta = h_n + \varphi_n^\delta$ from the initial guess φ_0. This yields additional stability since it is not possible that noise components sum up over the iteration. Of course, to achieve convergence, the regularization parameters must be chosen such that $\alpha_n \to 0$ as $n \to \infty$. By Theorem 2.1, the unique solution to (9.9) is given by

$$h_n = (\alpha_n I + F'[\varphi_n^\delta]^* F'[\varphi_n^\delta])^{-1} \left\{ F'[\varphi_n^\delta]^* (g^\delta - F(\varphi_n^\delta)) + \alpha_n (\varphi_0 - \varphi_n^\delta) \right\}. \tag{9.10}$$

Let us compare the Levenberg-Marquardt algorithm (9.6) and the iteratively regularized Gauss-Newton method (9.10) if $F = T$ is linear. In this case performing n steps with constant $\alpha_n = \alpha$ is equivalent to applying iterated Tikhonov regularization with n iterations because (9.7) reduces to

$$\varphi_{n+1}^\delta = \mathrm{argmin}_{\varphi \in X} \left(\|T\varphi - g^\delta\|^2 + \alpha \|\varphi - \varphi_n^\delta\|^2 \right).$$

Since we let n tend to infinity, the Levenberg-Marquardt algorithm becomes Lardy's method. The iteratively regularized Gauss-Newton method (9.9) applied to linear problems is

$$\varphi_{n+1}^\delta = \mathrm{argmin}_{\varphi \in X} \left(\|T\varphi - g^\delta\|^2 + \alpha_n \|\varphi - \varphi_0\|^2 \right),$$

i.e. it reduces to ordinary Tikhonov regularization with initial guess φ_0 and regularization parameter α_n. In particular, φ_{n+1}^δ is independent of all the previous iterates.

The convergence of Newton-type methods is much faster than the convergence of Landweber iteration. It can be shown that the number of Landweber steps which is necessary to achieve an accuracy comparable to n Newton steps increases exponentially with n (cf. [DES98]). On the other hand, a disadvantage of the Levenberg-Marquardt algorithm and the iteratively regularized Gauss-Newton method is that the (typically full) matrix corresponding to the operator $F'[\varphi_n^\delta] \in L(X,Y)$ has to be computed, and $\alpha_n I + F'[\varphi_n^\delta]^* F'[\varphi_n^\delta]$ has to be inverted in each step. For large scale problems, e.g. the inhomogeneous medium scattering problem discussed in Chapter 7, this can be prohibitively expensive since the computation of one column $F'[\varphi_n^\delta]h$ of this matrix, is typically as expensive as an evaluation of the operator F.

Setting up a matrix for $F'[\varphi_n^\delta]$ can be avoided by using an iterative method to solve (9.4). Then only the application of $F'[\varphi_n^\delta]$ and $F'[\varphi_n^\delta]^*$ to a given vector are needed. For many problems highly efficient implementations of the applications of $F'[\varphi_n^\delta]$ and $F'[\varphi_n^\delta]^*$ are available, e.g. by FFT, fast multipole or finite element techniques. The regularization of (9.4) is then achieved by an early stopping of the inner iteration. Using Landweber iteration, a ν-method or the conjugate gradient method for the normal equation as inner iteration leads to the *Newton-Landweber method*, the *Newton-ν-method* or the *Newton-CG method*, respectively (cf. [Han97b, Kal97, Rie99]). In all cases we we have a choice of using either 0 or $\varphi_0 - \varphi_n^\delta$ as starting point of the inner iteration, corresponding to either a Levenberg-Marquardt or an iteratively regularized Gauss-Newton scheme.

Another possibility is to solve the regularized equation

$$(\alpha_n I + F'[\varphi_n^\delta]^* F'[\varphi_n^\delta])h_n = r_n \tag{9.11}$$

by some iterative method, e.g. the conjugate gradient method. For exponentially ill-posed problems it is possible to construct special preconditioners for this inner iteration which increase the speed of convergence significantly (cf. [Hoh01]).

Convergence

As in the linear case, the quality of the iterates deteriorates in the presence of noise as the number of iterations tends to infinity. Therefore, the iteration has to be stopped before the propagated data noise error becomes too large. An iteration scheme has to be accompanied by an appropriate stopping rule to give a regulization method. The most commonly used stopping rule is the discrepancy principle which requires to stop the iteration at the first iteration index $N = N(\delta, g^\delta)$ for which

$$\|F(\varphi_N^\delta) - g^\delta\| \le \tau\delta \tag{9.12}$$

with some fixed constant $\tau \ge 1$.

Definition 9.1. An iterative method $\varphi_{n+1}^\delta := \Phi_n(\varphi_n^\delta, \dots, \varphi_0, g^\delta)$ together with a stopping rule $N(\delta, g^\delta)$ is called an *iterative regularization method* for F if for all $\varphi^\dagger \in D(F)$, all

g^δ satisfying (9.2) with $g := F(\varphi^\dagger)$ and all initial guesses φ_0 sufficiently close to $\varphi^\dagger$ the following conditions hold:

1. (Well-definedness) φ_n^δ is well defined for $n = 1, \dots, N(\delta, g^\delta)$, and $N(\delta, g^\delta) < \infty$ for $\delta > 0$.

2. (Convergence for exact data) For exact data ($\delta = 0$) either $N = N(0, g) < \infty$ and $\varphi_N^\delta = \varphi^\dagger$ or $N = \infty$ and $\|\varphi_n - \varphi^\dagger\| \to 0$ for $n \to \infty$.

3. (Convergence for noisy data) The following regularization property holds:

$$\sup\{\|\varphi^\delta_{N(\delta,g^\delta)} - \varphi^\dagger\| : g^\delta \in Y, \|g^\delta - g\| \leq \delta\} \to 0 \qquad \text{as } \delta \to 0. \tag{9.13}$$

The last property in Definition 9.1 can often be established with the help of the following theorem.

Theorem 9.2. *Assume that the first two properties in Definition 9.1 are satisfied and that the discrepancy principle (9.12) is used as stopping rule. Moreover, assume that the* monotonicity *condition*

$$\|\varphi_n^\delta - \varphi^\dagger\| \leq \|\varphi_{n-1}^\delta - \varphi^\dagger\| \tag{9.14}$$

and the stability *condition*

$$\|\varphi_n^\delta - \varphi_n\| \to 0 \qquad \text{as } \delta \to 0 \tag{9.15}$$

are satisfies for $1 \leq n \leq N(\delta, g^\delta)$. *Then the regularizing property (9.13) holds true.*

Proof. Let (g^{δ_k}) be a sequence in Y such that $\|g^{\delta_k} - g\| \leq \delta_k$ and $\delta_k \to 0$ as $k \to \infty$, and define $N_k := N(\delta_k, g^{\delta_k})$.

We first assume that $\overline{N}$ is a finite accumulation point of N_k. Without loss of generality we may assume that $N_k = \overline{N}$ for all k. The stability assumption (9.15) implies that

$$\varphi_{\overline{N}}^{\delta_k} \to \varphi_{\overline{N}} \qquad \text{as } k \to \infty. \tag{9.16}$$

It follows from (9.12) that $\|g^{\delta_k} - F(\varphi_{\overline{N}}^{\delta_k})\| \leq \tau\delta_k$ for all k. Taking the limit $k \to \infty$ shows that $F(\varphi_{\overline{N}}) = g$, i.e. $\varphi_{\overline{N}}$ is a solution to (9.1). Since we assume the solution to (9.1) to be unique, this implies (9.13).

It remains to consider the case that $N_k \to \infty$ as $k \to \infty$. Without loss of generality we may assume that N_k increases monotonically with k. Then (9.14) yields

$$\|\varphi_{N_k}^{\delta_k} - \varphi^\dagger\| \leq \|\varphi_{N_l}^{\delta_k} - \varphi^\dagger\| \leq \|\varphi_{N_l}^{\delta_k} - \varphi_{N_l}\| + \|\varphi_{N_l} - \varphi^\dagger\| \tag{9.17}$$

for $l \leq k$. Given $\epsilon > 0$, it follows from the second property in Definition 9.1 that the second term on the right hand side of (9.17) is $\leq \epsilon/2$ for some $l = l(\epsilon)$. Moreover, it follows from (9.15) that there exists $K \geq l(\epsilon)$ such that the first term on the right hand side of (9.17) is $\leq \epsilon/2$ for $k \geq K$. This shows (9.13). $\square$

All known proofs establishing the properties in Definition 9.1 for some method require some condition restricting the degree of nonlinearity of the operator F. A commonly used condition is

$$\|F(\varphi) - F(\psi) - F'[\varphi](\varphi - \psi)\| \leq \eta \|F(\varphi) - F(\psi)\|, \qquad \eta < \frac{1}{2} \tag{9.18}$$

for $\|\varphi - \psi\|$ sufficiently small. At a first glance, (9.18) may look like a weak condition since the Taylor remainder can be estimated by

$$\|F(\varphi) - F(\psi) - F'[\varphi](\varphi - \psi)\| \leq \frac{L}{2}\|\varphi - \psi\|^2 \tag{9.19}$$

under the Lipschitz condition (8.10) (cf. Lemma 7.5), and the right hand side of (9.19) involves a second power in the distance between φ and ψ whereas the right hand side of (9.18) only involves a first power of the distance of the images under F. However, if (9.1) is ill-posed, then $\|F(\varphi) - F(\psi)\|$ may be much smaller than $\|\varphi - \psi\|$. In this sense (9.18) is more restrictive than (9.19).

Whereas (9.19) could be shown for a number of parameter identification problems in partitial differential equation involving distributed measurements, it has not been possible to show (9.18) for other interesting problem such as inverse scattering problems or impedance tomography, yet.

Our next theorem shows how (9.18) can be used to establish the monotonicity condition (9.14) for Landweber iteration:

Theorem 9.3. *Assume that (9.18) is satisfied in some ball $B_\rho := \{\varphi : \|\varphi - \varphi^\dagger\| \leq \rho\}$ contained in $D(F)$ and that $\|F'[\varphi]\| \leq 1$ for all $\varphi \in B_\rho$. Then the Landweber iteration with $\mu = 1$ together with the discrepancy principle (9.12) with*

$$\tau = 2\frac{1+\eta}{1-2\eta}$$

is well-defined and satisfies (9.14).

Proof. Assume that $\varphi_n^\delta \in B_\rho$ for some $n < N(\delta, g^\delta)$. Then

$$\begin{aligned}
&\|\varphi^\dagger - \varphi_{n+1}^\delta\|^2 - \|\varphi^\dagger - \varphi_n^\delta\|^2 \\
&= 2\,\mathrm{Re}\left\langle \varphi_n^\delta - \varphi^\dagger, \varphi_{n+1}^\delta - \varphi_n^\delta \right\rangle + \|\varphi_{n+1}^\delta - \varphi_n^\delta\|^2 \\
&= 2\,\mathrm{Re}\left\langle F'[\varphi_n^\delta](\varphi_n^\delta - \varphi^\dagger), g^\delta - F(\varphi_n^\delta) \right\rangle + \mathrm{Re}\left\langle g^\delta - F(\varphi_n^\delta), F'[\varphi_n^\delta]F'[\varphi_n^\delta]^*(g^\delta - F(\varphi_n^\delta)\right\rangle \\
&= 2\,\mathrm{Re}\left\langle g^\delta - F(\varphi_n^\delta) + F'[\varphi_n^\delta](\varphi_n^\delta - \varphi^\dagger), g^\delta - F(\varphi_n^\delta) \right\rangle \\
&\quad - \mathrm{Re}\left\langle g^\delta - F(\varphi_n^\delta), (I - F'[\varphi_n^\delta]F'[\varphi_n^\delta]^*)(g^\delta - F(\varphi_n^\delta)\right\rangle - \|g^\delta - F(\varphi_n^\delta)\|^2 \\
&\leq 2\,\mathrm{Re}\left\langle g^\delta - F(\varphi_n^\delta) + F'[\varphi_n^\delta](\varphi_n^\delta - \varphi^\dagger), g^\delta - F(\varphi_n^\delta) \right\rangle - \|g^\delta - F(\varphi_n^\delta)\|^2
\end{aligned}$$

since $I - F'[\varphi_n^\delta]F'[\varphi_n^\delta]^*)$ is positive semidefinite by the assuption $\|F'[\cdot]\| \leq 1$. Using (9.18) and the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}
&\|\varphi^\dagger - \varphi_{n+1}^\delta\|^2 - \|\varphi^\dagger - \varphi_n^\delta\|^2 \\
&\quad\leq\ \|g^\delta - F(\varphi_n^\delta)\| \left(2\eta\|g - F(\varphi_n^\delta)\| + 2\delta - \|g^\delta - F(\varphi_n^\delta)\|\right) \\
&\quad\leq\ \|g^\delta - F(\varphi_n^\delta)\| \left((2\eta - 1)\|g^\delta - F(\varphi_n^\delta)\| + 2(1+\eta)\delta\right) \\
&\quad\leq\ 0
\end{aligned}$$

Here the last inequality follows from $n < N(\delta, g^\delta)$, the definition (9.12) of the discrepancy principle, and the definition of τ. In particular, $\varphi_{n+1}^\delta \in B_\rho$. The assertion now follows by induction on n. □

The stability condition (9.15) is obvious from the continuity of F and $F'[\cdot]^*$. Hence, to prove that Landweber iteration is an iterative regularization method in the sense of definition 9.1, it remains to show that it terminates after a finite number of steps for $\delta > 0$, i.e. $N(\delta, g^\delta) < \infty$ and that it converges for exact data. For this we refer to the originial paper [HNS95] or the monography [EHN96].

Based on Theorem 9.2 it has been shown by Hanke [Han97a, Han97b] that the Levenberg-Marquardt algorithm and the Newton-CG method are iterative regularization methods in the sense of Definition 9.1 if the operator F satisfies the nonlinearity condition

$$\|F(\varphi) - F(\psi) - F'[\varphi](\varphi - \psi)\| \leq c\|\varphi - \psi\|\,\|F(\varphi) - F(\psi)\| \tag{9.20}$$

for $\|\varphi - \psi\|$ sufficiently small. Although (9.20) is formally stronger than (9.18), for most problems where (9.18) can be shown, (9.20) can be established as well.

Convergence rates

Let

$$e_n := \varphi_n^\delta - \varphi^\dagger, \qquad n = 0, 1, \dots,$$

denote the error of the nth iterate of the iteratively regularized Gauss-Newton method. It follows from (9.10) that

$$\begin{aligned}
e_{n+1} &= e_n + (T_n^*T_n + \alpha_n I)^{-1}\left\{T_n^*(g^\delta - F(\varphi_n^\delta)) + \alpha_n(e_0 - e_n)\right\} \\
&= (T_n^*T_n + \alpha_n I)^{-1}\left\{\alpha_n e_0 + T_n^*(g^\delta - F(\varphi_n^\delta) + T_n e_n)\right\}
\end{aligned}$$

with $T_n := F'[\varphi_n^\delta]$. After a few further manipulations we see that e_{n+1} is the sum of an approximation error e_{n+1}^{app}, a data noise error e_{n+1}^{noi} and a nonlinearity error e_{n+1}^{nl} given by

$$\begin{aligned}
e_{n+1}^{\mathrm{app}} &= \alpha_n(T^*T + \alpha_n I)^{-1}e_0, \\
e_{n+1}^{\mathrm{noi}} &= (T_n^*T_n + \alpha_n I)^{-1}T_n^*(g^\delta - g), \\
e_{n+1}^{\mathrm{nl}} &= (T_n^*T_n + \alpha_n I)^{-1}T_n^*(F(\varphi^\dagger) - F(\varphi_n^\delta) + T_n e_n) \\
&\quad + \alpha_n(T_n^*T_n + \alpha_n I)^{-1}(T_n^*T_n - T^*T)(T^*T + \alpha_n I)^{-1}e_0
\end{aligned}$$

where $T := F'[\varphi^\dagger]$. If F is linear, then $e^{\mathrm{nl}}_{n+1} = 0$ and we can use the error analysis in Chapter 5. The error e^{nl}_{n+1} caused by the nonlineariy of the operator can be estimated separately. This possiblity makes the iteratively regularized Gauss-Newton method easier to analyze than other methods.

Lemma 9.4. *Assume that the source condition*

$$\varphi^\dagger - \varphi_0 = (F'[\varphi^\dagger]^* F'[\varphi^\dagger])^\mu w, \qquad \|w\| \le \rho \tag{9.21}$$

is satisfied for some $\frac{1}{2} \le \mu \le 1$ *and* $w \in X$ *and that there exists a Lipschitz constant* L *such that*

$$\|F'[\varphi] - F'[\psi]\| \le L\|\varphi - \psi\| \tag{9.22}$$

for all $\varphi, \psi \in D(F)$. *Moreover, assume that the stopping index* N *is chosen such that*

$$\alpha_N^{\mu+\frac{1}{2}} < \eta\delta \le \alpha_n^{\mu+\frac{1}{2}}, \qquad 0 \le n < N \tag{9.23}$$

with some constant $\eta > 0$, *and that* φ_n^δ *is well-defined. Then*

$$\|e_{n+1}\| \le \left(\rho + \frac{1}{2\eta}\right)\alpha_n^\mu + L\rho\left(\frac{\alpha_n^{\mu-1/2}}{2} + \|T^*T\|^{\mu+1/2}\right)\|e_n\| + \frac{L}{4\sqrt{\alpha_n}}\|e_n\|^2. \tag{9.24}$$

Proof. It follows from (5.20) that

$$\|(T_n^*T_n + \alpha_n I)^{-1}T_n^*\| \le \sup_{\lambda \ge 0}\left|\frac{\sqrt{\lambda}}{\lambda + \alpha}\right| = \frac{1}{2\sqrt{\alpha}}. \tag{9.25}$$

By virtue of (9.21) and Theorem 5.2 we have $\|e^{\mathrm{app}}_{n+1}\| \le \alpha_n^\mu \rho$ as $\mu \le 1$. It follows from (9.25) and (9.23) that $\|e^{\mathrm{noi}}_{n+1}\| \le \alpha_n^{-1/2}\delta/2 \le \alpha_n^\mu/(2\eta)$ for $n < N$. The first term in e^{nl}_{n+1} can be estimated by $\alpha_n^{-1/2}(L/4)\|e_n\|^2$ due to Lemma 7.5 and (9.25). For the second term in e^{nl}_{n+1} we get

$$\begin{aligned}
& \|\alpha_n(T_n^*T_n + \alpha_n I)^{-1}(T_n^*T_n - T^*T)(T^*T + \alpha_n I)^{-1}e_0\| \\
= \;& \|\alpha_n(T_n^*T_n + \alpha_n I)^{-1}\left\{T_n^*(T - T_n) + (T^* - T_n^*)T\right\}(T^*T + \alpha_n I)^{-1}e_0\| \\
\le \;& L\|e_n\|\rho\left(\frac{\alpha_n^{\mu-1/2}}{2} + \|T^*T\|^{\mu+1/2}\right)
\end{aligned}$$

using (9.22), (9.25), (9.21) and Theorem 5.2. Putting all these estimates together gives the assertion. $\square$

Theorem 9.5. *Let the assumptions of Lemma 9.4 be satisfied, and assume that*

$$1 \le \frac{\alpha_n}{\alpha_{n+1}} \le r, \qquad \lim_{n\to\infty} \alpha_n \to 0, \qquad \alpha_n > 0 \tag{9.26}$$

with $\alpha_0 \le 1$ and $r > 1$. Moreover, assume that η is sufficiently large and ρ is sufficiently small. Then for exact data the error satisfies

$$\|\varphi_n - \varphi^\dagger\| = \mathrm{O}\left(\alpha_n^\mu\right), \qquad \text{as } n \to \infty. \tag{9.27}$$

For noisy data we have

$$\|\varphi_{N(\delta,g^\delta)} - \varphi^\dagger\| = \mathrm{O}\left(\delta^{\frac{2\mu}{2\mu+1}}\right) \qquad \text{as } \delta \to 0. \tag{9.28}$$

Proof. Under the assumptions of Lemma 9.4 the quantities $\Theta_n := \alpha_n^{-\mu}\|e_n\|$ fulfil the inequality

$$\Theta_{n+1} \le a + b\Theta_n + c\Theta_n^2, \tag{9.29}$$

where the coefficients are defined by $a := r^\mu(\rho+1/(2\eta))$, $b := r^\mu L\rho\left(\alpha_n^{\mu-1/2}/2 + \|T^*T\|^{\mu+1/2}\right)$, and $c := r^\mu L/4$. Here we have used that $\mu \ge 1/2$, and hence $\alpha_n^{-1/2} \le \alpha_n^{-\mu}$. Let t_1 and t_2 be the solutions to the fixed point equation $a + bt + ct^2 = t$, i.e.

$$t_1 := \frac{2a}{1-b+\sqrt{(1-b)^2-4ac}}, \qquad t_2 := \frac{1-b+\sqrt{(1-b)^2-4ac}}{2c},$$

let the stopping index $N \le \infty$ be given by (9.23) and define $C_\Theta := \max(\Theta_0, t_1)$. We will show by induction that

$$\Theta_n \quad \le \quad C_\Theta \tag{9.30}$$

for $0 \le n \le N$ if

$$b + 2\sqrt{ac} < 1, \tag{9.31a}$$
$$\Theta_0 \le t_2, \tag{9.31b}$$
$$\{\varphi \in X : \|\varphi - \varphi^\dagger\| \le \alpha_0^\mu C_\Theta\} \subset D(F). \tag{9.31c}$$

It is easy to see that the conditions (9.31a) are satisfied if η is suffiently large and ρ sufficiently small. For $n = 0$, (9.30) is true by the definition of C_Θ. Assume that (9.30) is true for some $k < N$. Then (9.31c) implies that $\varphi_n^\delta \in D(F)$, i.e. $\varphi_{n+1}^\delta \in X$ is well defined, and (9.29) is true for $n = k$. By (9.31a) we see that $t_1, t_2 \in \mathbb{R}$ and $t_1 < t_2$, and by (9.30) we have $0 \le \Theta_k \le t_1$ or $t_1 < \Theta_k \le \Theta_0$. In the first case, since $a, b, c \ge 0$, we conclude that

$$\Theta_{k+1} \le a + b\Theta_k + c\Theta_k^2 \le a + bt_1 + ct_1^2 = t_1,$$

and in the second case we use (9.31b) and the fact that $a + (b-1)t + ct^2 \le 0$ for $t_1 \le t \le t_2$ to show that

$$\Theta_{k+1} \le a + b\Theta_k + c\Theta_k^2 \le \Theta_k \le \Theta_0.$$

Thus in both cases there holds $\Theta_{k+1} \le C_\Theta$. This completes the proof of (9.30).

(9.30) immediately implies (9.27). The convergence rate (9.28) follows from the first inequality in (9.23) since

$$\|\varphi_N - \varphi^\dagger\| \leq C_\theta \alpha_N^\mu \leq C_\theta (\eta\delta)^{\frac{2\mu}{2\mu+1}}.$$

□

In [BNS97] it has been shown that the iteratively regularized Gauß-Newton method is an iterative regularization method in the sense of Definition 9.1 if the operator F satisfies the following nonlinearity condition: For $\varphi, \bar{\varphi}$ in a neighborhood of $\varphi^\dagger$ there exist operators $R(\bar{\varphi}, \varphi) \in L(Y)$ and $Q(\bar{\varphi}, \varphi) \in L(X, Y)$ and constants $C_Q, C_R > 0$ such that

$$\begin{aligned} &F'[\bar{\varphi}] = R(\bar{\varphi}, \varphi) F'[\varphi] + Q(\bar{\varphi}, \varphi) \\ &\|I - R(\bar{\varphi}, \varphi)\| \leq C_R, \quad \|Q(\bar{\varphi}, \varphi)\| \leq C_Q \|F'[\varphi^\dagger](\bar{\varphi} - \varphi)\| \end{aligned} \tag{9.32}$$

Under this assumption order optimal convergence rates have been shown for $\mu < \frac{1}{2}$ using the discrepancy principle. Convergence rates under logarithmic source conditions

$$\varphi^\dagger - \varphi_0 = f_p(F'[\varphi^\dagger]^* F'[\varphi^\dagger]) w, \qquad \|w\| \leq \rho,$$

(cf. (5.21)) have been obtained in [Hoh97]. Such conditions are appropriate for exponentially ill-posed problems such as inverse obstacle scattering problems and can often be roughly interpreted as smoothness conditions in terms of Sobolev spaces.

Bibliography

[Ado95] H.-M. Adorf. Hubble space telescope image restoration in its fourth year. *Inverse Problems*, 11:639–653, 1995.

[Ale88] G. Alessandrini. Stable determination of conductivities by boundary measurements. *Appl. Anal.*, 27:153–172, 1988.

[Bak92] A. B. Bakushinsky. The problem of the convergence of the iteratively regularized Gauss-Newton method. *Comput. Maths. Math. Phys.*, 32:1353–1359, 1992.

[Bau87] J. Baumeister. *Stable Solution of Inverse Problems.* Vieweg, Braunschweig, 1987.

[Bau90] H. Bauer. *Maß- und Integrationstheorie.* de Gruyter, Berlin, New York, 1990.

[BG95] A. B. Bakushinsky and A. Goncharsky. *ill-posed Problems: Theory and Applications.* Kluwer, Dordrecht, 1995.

[BK89] H. Banks and K. Kunisch. *Parameter Estimation Techniques for Distributed Systems.* Birkhäuser, Boston, 1989.

[BNS97] B. Blaschke, A. Neubauer, and O. Scherzer. On convergence rates for the iteratively regularized Gauß-Newton method. *IMA Journal of Numerical Analysis*, 17:421–436, 1997.

[CK83] D. Colton and R. Kreß. *Integral Equation Methods in Scattering Theory.* Wiley, New York, 1983.

[CK97a] G. Chavent and K. Kunisch. Regularization of linear least squares problems by total bounded variation. *ESAIM: Control, Optimisation and Calculus of Variations*, 2:359–376, 1997.

[CK97b] D. Colton and R. Kreß. *Inverse Acoustic and Electromagnetic Scattering Theory.* Springer Verlag, Berlin, Heidelberg, New York, second edition, 1997.

[DES98] P. Deuflhard, H. Engl, and O. Scherzer. A convergence analysis of iterative methods for the solution of nonlinear ill-posed problems under affinely invariant conditions. *Inverse Problems*, 14:1081–1106, 1998.

[EG88] H. W. Engl and H. Gfrerer. A posteriori parameter choice for general regularization methods for solving linear ill-posed problems. *Appl. Numer. Math.*, 4:395–417, 1988.

[EHN96] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems.* Kluwer Academic Publisher, Dordrecht, Boston, London, 1996.

[EKN89] H. W. Engl, K. Kunisch, and A. Neubauer. Convergence rates for Tikhonov regularization of nonlinear ill-posed problems. *Inverse Problems*, 5:523–540, 1989.

[EL93] H. W. Engl and G. Landl. Convergence rates for maximum entropy regularization. *SIAM J. Numer.oo Anal.*, 30:1509–1536, 1993.

[ER95] H. W. Engl and W. Rundell, eds. *Inverse Problems in diffusion processes.* SIAM, Philadelphia, 1995.

[Gfr87] H. Gfrerer. An a-posteriori parameter choice for ordinary and iterated Tikhonov regularization of ill-posed problems leading to optimal convergence rates. *Math. Comp.*, 49:507–522 and S5–S12, 1987.

[Giu84] E. Giusti. *Minimal surfaces and functions of bounded variation.* Birkhäuser, Boston, 1984.

[Gla84] V. B. Glasko. *Inverse Problems of Mathematical Physics.* American Institute of Physics, New York, 1984.

[GP95] R. D. Grigorieff and R. Plato. On a minimax equality for seminorms. *Linear Algebra Appl.*, 221:227–243, 1995.

[Gro83] C. W. Groetsch. Comments on Morozov's discrepancy principle. In: G. Hämmerlin and K. H. Hoffmann, eds, *Improperly Posed Problems and Their Numerical Treatment*, pp. 97–104. Birkhäuser, Basel, 1983.

[Gro84] C. W. Groetsch. *The Theory of Tikhonov regularization for Fredholm equations of the first kind.* Pitman, Boston, 1984.

[Häh98] P. Hähner. On acoustic, electromagnetic, and elastic scattering problems in inhomogeneous media, 1998. Habilitation thesis, Göttingen.

[Han97a] M. Hanke. A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems. *Inverse Problems*, 13:79–95, 1997.

[Han97b] M. Hanke. Regularizing properties of a truncated Newton-CG algorithm for nonlinear inverse problems. *Numer. Funct. Anal. Optim.*, 18:971–993, 1997.

[Hel80] S. Helgason. *The Radon Transform.* Birkhäuser Verlag, Boston, 1980.

[HNS95] M. Hanke, A. Neubauer, and O. Scherzer. A convergence analysis of the Landweber iteration for nonlinear ill-posed problems. *Numer. Math.*, 72:21–37, 1995.

[Hof86] B. Hofmann. *Regularization of Applied Inverse and ill-posed Problems.* Teubner, Leipzig, 1986.

[Hoh97] T. Hohage. Logarithmic convergence rates of the iteratively regularized Gauß-Newton method for an inverse potential and an inverse scattering problem. *Inverse Problems*, 13:1279–1299, 1997.

[Hoh99] T. Hohage. *Iterative Methods in Inverse Obstacle Scattering: Regularization Theory of Linear and Nonlinear Exponentially ill-posed Problems.* PhD thesis, University of Linz, 1999.

[Hoh00] T. Hohage. Regularization of exponentially ill-posed problems. *Numer. Funct. Anal. Optim.*, 21:439–464, 2000.

[Hoh01] T. Hohage. On the numerical solution of a three-dimensional inverse medium scattering problem. *Inverse Problems*, 17:1743–1763, 2001.

[HS71] F. Hirzebruch and W. Scharlau. *Einführung in die Funktionalanalysis.* Spektrum Akademischer Verlag, Heidelberg, Berlin, Oxford, 1971.

[Isa90] V. Isakov. *Inverse Source Problems*, Vol. 34 of *Mathematical Surveys and Monographs.* Amer. Math. Soc., Providence, 1990.

[Isa98] V. Isakov. *Inverse Problems for Partial Differential Equations.* Springer, Berlin, Heidelberg, New York, 1998.

[Kal97] B. Kaltenbacher. Some Newton-type methods for the regularization of nonlinear ill-posed problems. *Inverse Problems*, 13:729–753, 1997.

[Kel76] J. B. Keller. Inverse problems. *Am. Math. Mon.*, 83:107–118, 1976.

[Kir96] A. Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems.* Springer, New York, Berlin, Heidelberg, 1996.

[Kre89] R. Kreß. *Linear Integral Equations.* Springer Verlag, Berlin, Heidelberg, New York, 1989.

[Lav67] M. Lavrentiev. *Some Improperly Posed Problems of Mathematical Physics.* Springer, New York, 1967.

[LB91] R. L. Lagendijk and J. Biemond. *Iterative Identification and Restoration of Images.* Kluwer, Dordrecht,London, 1991.

[Leb65] N. N. Lebedev. *Special Functions and Their Applications.* Prentice-Hall, Inc., Englewood, N.J., 1965.

[Lou89] A. K. Louis. *Inverse und schlecht gestellte Probleme*. Teubner Verlag, Stuttgart, 1989.

[Mai94] B. A. Mair. Tikhonov regularization for finitely and infinitely smoothing operators. *SIAM J. Math. Anal.*, 25:135–147, 1994.

[MM79] A. A. Melkman and C. A. Micchelli. Optimal estimation of linear operators in Hilbert spaces from inaccurate data. *SIAM J. Numer. Anal.*, 16:87–105, 1979.

[Mor84] V. A. Morozov. *Methods for Solving Incorrectly Posed Problems*. Springer Verlag, New York, Berlin, Heidelberg, 1984.

[Mor93] V. A. Morozov. *Methods for Solving Incorrectly Posed Problems*. CRC Press, Boca Raton, CA, 1993.

[MV92] R. Meise and D. Vogel. *Einführung in die Funktionalanalysis*. Vieweg Verlag, Braunschweig Wiesbaden, 1992.

[Nac88] A. Nachman. Reconstructions from boundary measurements. *Annals of Math.*, 128:531–576, 1988.

[Nac95] A. Nachman. Global uniqueness for the two-dimensional inverse boundary-value problem. *Ann. of Math.*, 142:71–96, 1995.

[Nat77] F. Natterer. Regularisierung schlecht gestellter Probleme durch Projektionsverfahren. *Numer. Math.*, 28:329–341, 1977.

[Nat86] F. Natterer. *The Mathematics of Computerized Tomography*. Teubner, Stuttgart, 1986.

[Neu88] A. Neubauer. An a posteriori parameter choice for Tikhonov regularization in the presence of modelling error. *Appl.Numer.Math.*, 4:507–519, 1988.

[Neu89] A. Neubauer. Tikhonov regularization for nonlinear ill-posed problems: optimal convergence and finite-dimensional approximation. *Inverse Problems*, 5:541–557, 1989.

[NS98] M. Z. Nashed and O. Scherzer. Least squares and bounded variation regularization with nondifferentiable functionals. *Numer. Funct. Anal. and Optimiz.*, 19:873–901, 1998.

[PV90] R. Plato and G. Vainikko. On the regularization of projection methods for solving ill-posed problems. *Numer. Math.*, 57:63–79, 1990.

[Ram86] A. Ramm. *Scattering by Obstacles*. Reidel, Dordrecht, 1986.

[Rie99] A. Rieder. On the regularization of nonlinear ill-posed problems via inexact Newton iterations. *Inverse Problems*, 15:309–327, 1999.

[RK96] A. G. Ramm and A. I. Katsevich. *The Radon transform and local tomography.* CRC Press, Boca Raton, CA, 1996.

[Rud73] W. Rudin. *Functional Analysis.* McGraw-Hill, New York, 1973.

[Sch02] E. Schock. Non-linear ill-posed problems: counter-examples. *Inverse Problems*, 18:715–717, 2002.

[SEK93] O. Scherzer, H. W. Engl, and K. Kunisch. Optimal a-posteriori parameter choice for Tikhonov regularization for solving nonlinear ill-posed problems. *SIAM J. Numer. Anal.*, 30:1796–1838, 1993.

[SG85] C. R. Smith and W. T. Grandy, eds. *Maximum-Entropy and Bayesian Methods in Inverse Problems.* Fundamental Theories of Physics. Reidel, Dordrecht, 1985.

[SU87] J. Sylvester and G. Uhlmann. A global uniqueness theorem for an inverse boundary value problem. *Annals of Math.*, 125:153–169, 1987.

[TA77] A. N. Tikhonov and V. Arsenin. *Solutions of ill-posed Problems.* Wiley, New York, 1977.

[Tay96] M. Taylor. *Partial Differential Equations: Qualitative Studies of Linear Equations*, Vol. 2. Springer Verlag, New York, 1996.

[TGSY95] A. N. Tikhonov, A. V. Goncharsky, V. V. Stephanov, and A. G. Yagola. *Numerical Methods for the Solution of ill-posed Problems.* Kluwer Academic Press, Dordrecht, 1995.

[Tik63] A. N. Tikhonov. On the solution of incorrectly formulated problems and the regularization method. *Soviet Math. Doklady*, 4:1035–1038, 1963. Englisch translation.

[VO96] C. Vogel and M. Oman. Iterative methods for total variation denoising. *SIAM J. Sci. Comput.*, 17:227–238, 1996.

[Wah90] G. Wahba. *Spline Models for Observational Data.* SIAM, Philadelphia, 1990.

[Wei76] J. Weidmann. *Lineare Operatoren in Hilberträumen.* Teubner, Stuttgart, 1976.