

IPAM 2003

Inverse Problems

Heinz W.Engl

Industrial Mathematics Institute

Johannes–Kepler–Universität

Linz, Austria

and

Johann Radon Institute for

Computational and Applied Mathematics

Austrian Academy of Sciences

Two problems “**inverse to each other**” if the formulation of one problem involves the other one.

“**Direct problem**”, the other one “**inverse problem**”. Possible “inverse problems” are the determination of the present state of the system from future observations (i.e., the calculation of the “evolution” of the system backwards in time) or the identification of physical parameters from observations of the evolution of the system (“parameter identification”).

Two different motivations: First, one wants to **know** past states or parameters of a physical system. Second, one wants to find out how to influence a system via its present state or via parameters in order to **steer** it to a desired state in the future.

Thus, one might say the **inverse problems are concerned with determining causes for a desired or an observed effect.**

Example 1 (Differentiation): Let $f \in C^1[0, 1]$ be any function, $\delta \in]0, 1[$, $n \in \mathbb{N}$ ($n \geq 2$) be arbitrary, and define

$$f_{\delta,n}(x) := f(x) + \delta \sin\left(\frac{nx}{\delta}\right) \quad (x \in [0, 1]). \quad (1)$$

Then

$$f'_{\delta,n}(x) = f'(x) + n \cos\left(\frac{nx}{\delta}\right) \quad (x \in [0, 1]). \quad (2)$$

Now, in the uniform norm,

$$\|f - f_{\delta,n}\|_{\infty} = \delta, \quad (3)$$

but

$$\|f' - f'_{\delta,n}\|_{\infty} = n. \quad (4)$$

The derivative does not depend continuously on the data w.r.t the uniform norm.

f' solves the simple integral equation of the first kind

$$(Sx)(s) := \int_0^t x(s) ds = f(t) - f(0) \quad (5)$$

solvable on $C[0, 1]$ only if $f \in C^1[0, 1]$.

Direct problem: compute f from x , i.e., integration, integration smoothes, highly oscillatory errors in x are damped out.

Why can we differentiate a function?

We have to be able to exclude data errors of arbitrarily high frequency; can be done if we know a bound for f'' .

S is a continuous linear injective operator, whose inverse is unbounded. If we restrict S to the set $\{x \in C^1[0, 1] / \|x\|_\infty + \|x'\|_\infty \leq C\}$, which is compact, then the inverse of this restricted operator is continuous. Thus, we can “restore stability” by knowing an a-priori bound e.g. for f' and f'' .

Let f be the function we want to differentiate, f_δ its noisy version with

$$\|f - f_\delta\|_\infty \leq \delta. \quad (6)$$

Want to use the central difference quotient with step size h . If $f \in C^2[0, 1]$, Taylor expansion yields

$$\frac{f(x+h) - f(x-h)}{2h} = f'(x) + O(h), \quad (7)$$

while for $f \in C^3[0, 1]$,

$$\frac{f(x+h) - f(x-h)}{2h} = f'(x) + O(h^2). \quad (8)$$

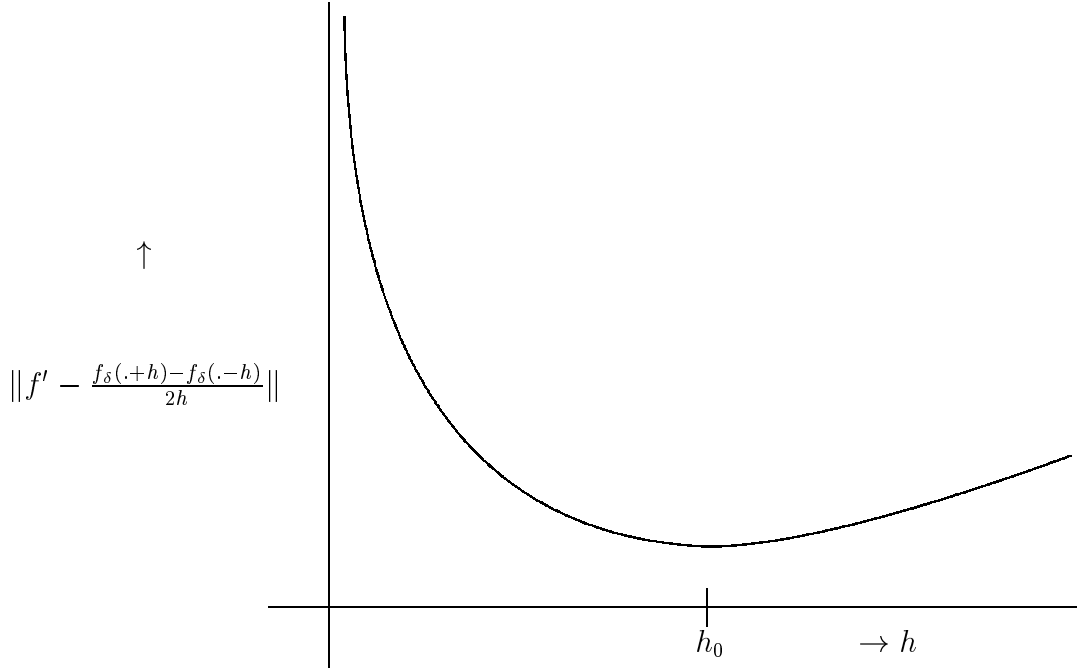
Thus, the accuracy of the central difference quotient depends on the smoothness of the exact data. Instead of f' , we are actually computing

$$\frac{f_\delta(x+h) - f_\delta(x-h)}{2h} \sim \frac{f(x+h) - f(x-h)}{2h} + \frac{\delta}{h}.$$

Thus, the total error behaves like

$$O(h^\nu) + \frac{\delta}{h}, \quad (9)$$

where $\nu = 1$ or 2 , if $f \in C^2[0, 1]$ or $f \in C^3[0, 1]$, respectively. For a fixed error level δ , it looks as follows:



If h becomes too small, total error increases due to the error term $\frac{\delta}{h}$, if h is too large, then the approximation error becomes too large. There is an “optimal” discretization parameter h_0 , estimate the asymptotic behaviour of h_0 :

If h is chosen as a power of δ , i.e.,

$$h \sim \delta^k, \tag{10}$$

then one can minimize (9) by taking $k = \frac{1}{2}$ or $k = \frac{1}{3}$, which results in a behaviour of the total error as $O(\sqrt{\delta})$ or $O(\delta^{\frac{2}{3}})$ for $f \in C^2[0, 1]$ or $f \in C^3[0, 1]$, respectively.

Thus, even in the best possible case, $\nu > 2$ is obviously not possible in (9) and for an optimal choice of h , we obtain only a convergence rate $O(\delta^{\frac{2}{3}})$, where δ denotes the data error, i.e., there is an intrinsic loss of information.

In this example, we saw some effects that are typical for ill-posed problems:

- amplification of high-frequency error
- restoration of stability by using a-priori information
- two error terms of different nature, one for the approximation error, the other one for the propagation of the data error
- the appearance of an optimal discretization parameter, whose choice depends on a-priori information
- loss of information even under optimal circumstances.

Example 2 In physical or technical applications, one often has the situation that the physical laws governing the process are known, but quantitative information about physical parameters are unknown. E.g., if one considers heat conduction in a material occupying a three-dimensional domain Ω whose temperature is kept at 0 at the boundary, the temperature distribution u after sufficiently long time can be modelled by

$$\left. \begin{aligned} -\operatorname{div} (a(x, y, z) \operatorname{grad} u) &= f(x, y, z) & (x, y, z) \in \Omega \\ u &= 0 & \text{on } \partial\Omega, \end{aligned} \right\} \quad (11)$$

where f denotes internal heat sources and a is the (spatially varying) heat conductivity.

The question is then how to determine a from internal measurements of the temperature u or from measurements of the heat flux $a \cdot \frac{\partial u}{\partial n}$ at the boundary $\partial\Omega$.

Consider the one–dimensional version of (11) of determining the coefficient $a = a(x)$ in

$$-(a(x)u_x)_x = f(x) \quad x \in [0, 1] \quad (12)$$

(with appropriate boundary conditions) from measurements of u . If u_x vanishes nowhere, a can be computed explicitly as

$$a(x) = \frac{1}{u_x(x)} \cdot \left[a(0)u_x(0) - \int_0^x f(s)ds \right] \quad (13)$$

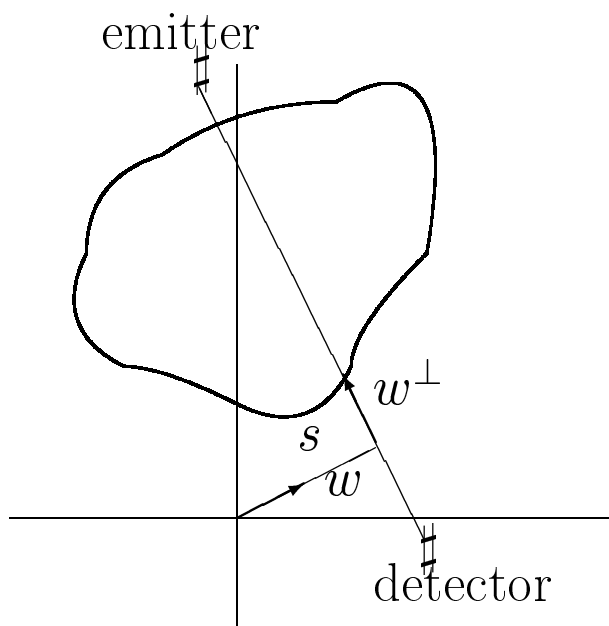
and is therefore uniquely determined. Thus, in order to compute a , one has to differentiate the data u , which is an ill–posed problem as explained in Example 1. In addition, there is another effect of instability coming from the division by u_x in (13):

In regions where u_x is small, errors e.g. in f are amplified, which is not surprising since where u_x vanishes, a cannot be determined at all, so that some instability has to be expected where u_x is small.

This is a nonlinear effect, while the ill-posedness involved with differentiation of the data comes from the fact that also the linearized problem is ill-posed.

Example 3 An inverse problem that was widely studied recently because of its importance in e.g. medical applications arises in “Computerized Tomography (CT)”. We consider the two-dimensional situation:

$D \subseteq \mathbb{R}^2$ compact domain with a spatially varying density f . The aim is to recover the density f from X-ray measurements in the plane where D lies.



If one assumes that the decay $-\Delta I$ of an X-ray beam along a distance Δt is proportional to the intensity I , the density f and to Δt , we obtain

$$\Delta I(sw + tw^\perp) = -I(sw + tw^\perp)f(sw + tw^\perp) \cdot \Delta t, \quad (14)$$

where w^\perp is a unit vector orthogonal to w . By letting Δt tend to 0 in (14), we obtain

$$\frac{d}{dt}I(sw + tw^\perp) = -I(sw + tw^\perp)f(sw + tw^\perp). \quad (15)$$

$I_L(s, w)$ and $I_0(s, w)$: intensity of the X-ray beam measured at the detector and emitted at the emitter.

(15) has the solution

$$\ln I_L(s, w) - \ln I_0(s, w) = - \int_{\mathbb{R}} f(sw + tw^\perp) dt, \quad (16)$$

so that the density f is related to the measured quantities I_L and I_0 via the integral equation of the first kind

$$\begin{aligned} (Rf)(s, w) &:= \int_{\mathbb{R}} f(sw + tw^\perp) dt & (17) \\ &= -\ln \frac{I_L(s, w)}{I_0(s, w)} \quad (w \in \mathbb{R}^2, \|w\| = 1, s > 0). \end{aligned}$$

The integral operator R is called the “Radon transform” after the Austrian mathematician Johann Radon. Assumption:

$$f(s, w) = F(s) \quad (0 < s \leq R, \|w\| = 1) \quad (18)$$

$$g(s) := -\ln \frac{I_L(s, w_0)}{I_0(s, w_0)}. \quad (19)$$

We obtain an “Abel integral equation of the first kind”

$$\int_s^R \frac{r F(r)}{\sqrt{r^2 - s^2}} dr = \frac{g(s)}{2} \quad (0 < s \leq R) \quad (20)$$

$$F(r) = -\frac{1}{\pi r} \int_r^R \frac{g'(s)}{\sqrt{s^2 - r^2}} ds. \quad (21)$$

This inversion formula involves g' , i.e., the data have to be differentiated!

Example 4 A classical inverse problem in connection with heat conduction is the “backwards heat equation”.

Let $\Omega \subseteq \mathbb{R}^3$ be the region that a body, whose boundary is kept at constant temperature 0, occupies.

Putting all physical parameters to 1, the temperature $u = u(x, t)$ obeys the linear heat equation

$$\Delta u = \frac{\partial u}{\partial t} \quad \text{in } D \times [0, T] \quad (22)$$

with Dirichlet boundary condition

$$u = 0 \quad \text{on } \partial D \times [0, T]. \quad (23)$$

We look for the initial temperature $u(x, 0)$ ($x \in D$) from measurements of

$$f(x) := u(x, T) \quad \text{for } x \in D. \quad (24)$$

No solution of this inverse problem exists unless f is analytic! Again, the operator representing the “direct problem” (mapping $u(\cdot, 0)$ onto f) has strong smoothing properties.

Let $\lambda_1, \lambda_2, \lambda_3, \dots$ be the eigenvalues for the Dirichlet problem for Ω , $\phi_1, \phi_2, \dots \in L^2(\Omega)$ the corresponding normalized eigenvalues, i.e., for all $k \in \mathbb{N}$, $\|\phi_k\|_{L^2(\Omega)} = 1$ and

$$\Delta\phi_k + \lambda_k\phi_k = 0 \text{ in } \Omega \tag{25}$$

$$\phi_k = 0 \text{ on } \partial\Omega.$$

Let

$$u_k(x, t) := \frac{1}{\lambda_k}\phi_k(x) \cdot \exp(\lambda_k(T-t)) \quad (x \in \Omega, t \in [0, T]).$$

Then

$$\begin{aligned} (\Delta u_k)(x, t) &= \frac{1}{\lambda_k}(\Delta\phi_k)(x) \cdot \exp(\lambda_k(T-t)) = \\ &= -\phi_k(x) \exp(\lambda_k(T-t)) = \frac{\partial}{\partial t}u_k(x, t), \end{aligned}$$

With $f_k := \frac{\phi_k}{\lambda_k}$, u_k solves (22)–(24). Since $(\lambda_k) \rightarrow +\infty$, $\lim_{k \rightarrow \infty} \|f_k\|_{L^2(\Omega)} = 0$, but $\lim_{k \rightarrow \infty} \|u_k(\cdot, 0)\|_{L^2(\Omega)} = \lim_{k \rightarrow \infty} \left(\frac{1}{\lambda_k} \cdot \exp(\lambda_k T) \right) = +\infty$. Thus, if we consider f_k

as perturbations of $f = 0$ with L^2 -error $\frac{1}{\lambda_k}$, the corresponding error in the solution of the inverse problem is amplified exponentially, namely by the factor $\exp(\lambda_k T)$. Especially, the solution of the inverse problem does not depend continuously on the data.

The exponential amplification of the error appears already for arbitrarily short time $T > 0$, but becomes worse as T increases.

Again, inverse problems of this type can also be represented by an integral equation of the first kind. E.g., for the one-dimensional unbounded case (i.e., $\Omega = \mathbb{R}$), the final temperature $u(., T)$ is related to the initial temperature $u(., 0)$ via

$$\frac{1}{2\sqrt{\pi T}} \int_{-\infty}^{+\infty} u(s, 0) \exp\left(-\frac{(x-s)^2}{4T}\right) ds = u(x, T). \quad (26)$$

Example 5 Let $\Omega \subset \mathbb{R}^3$ symbolize a part of the earth, ρ the density. The gravitation potential at $x \notin \Omega$ is given by

$$\phi(x) := \frac{G}{4\pi} \int_{\Omega} \frac{\rho(y)}{\|x - y\|} dy, \quad (27)$$

where G is the gravitation constant. In geophysics, one is interested in finding ρ (or, at least, obtaining some information about ρ) from gravity measurements; ϕ itself cannot be measured, but derivatives of differences of ϕ can. Via (27), this leads to an integral equation of the first kind for ρ .

Example 6 Practically important class of inverse problems: “inverse scattering problems”, where information about an unknown object (e.g., a body, an inhomogeneity in a material, a potential) is to be recovered from measurements of waves or fields scattered by this object.

Consider the problem of identifying a spatially varying “acoustic profile” (the reciprocal of the sound speed), which is given by a function $n = n(x)$ which equals 1 outside some compact set. We send in a time harmonic wave U^I given by

$$U^I(x, t) = e^{ikt} u^I(x). \quad (28)$$

u^I is the “velocity potential”, whose gradient with respect to x represents the speed of the wave motion. Since the “incident wave” is the one we would observe if the inhomogeneity in n were not there, i.e., if $n \equiv 1$ everywhere, U^I obeys the wave equation

$$\frac{\partial^2 U^I}{\partial t^2} = \Delta U^I, \quad (29)$$

which (by (28)) yields the “reduced wave equation” or “Helmholtz equation”

$$\Delta u^I + k^2 u^I = 0 \quad (30)$$

for the spatial component of the incident wave. Due to the inhomogeneity, the truly observed wave $U(x, t) = e^{ikt} u(x)$ obeys

$$\frac{\partial^2 U}{\partial t^2} = \frac{1}{n^2} \Delta U \quad (31)$$

and hence

$$\Delta u + k^2 n^2 u = 0. \quad (32)$$

We know u^I and observe u , so that it makes sense to define the “scattered wave” by

$$u^s := u - u^I. \quad (33)$$

With

$$f := 1 - n^2, \quad (34)$$

where this unknown function f has compact support, we obtain from (30), (32) and (33)

$$\Delta u^s + k^2 u^s = k^2 f(u^I + u^s). \quad (35)$$

The “inverse scattering problem” is now to compute the unknown function f in (35) from u^I and u^s . The problem of determining f in (35) from u^I and u^s is nonlinear, even though the governing equations are linear. In many applications it is reasonable to assume that the scattering wave is much smaller than the incident wave and to replace (35) by the “Born–(Rytov–)approximation”

$$\Delta u^s + k^2 u^s = k^2 f u^I; \quad (36)$$

the problem of determining f from (36) is now linear and leads to an integral equation of the first kind for f : Since the fundamental solution of the Helmholtz equation is given by

$$G(x, y) := -\frac{1}{4\pi} \frac{e^{ik|x-y|}}{|x-y|} \quad (x \neq y \in \mathbb{R}^3) \quad (37)$$

u^s as given by (36) can be represented as

$$u^s(x) = -\frac{k^2}{4\pi} \int_{\text{supp } f} \frac{e^{ik|x-y|}}{|x-y|} u^I(y) f(y) dy. \quad (38)$$

Inverse problems frequently do not fulfill Hadamard's definition of **well-posedness**, i.e., they do not possess the following properties:

For all admissible data, a solution exists. (39)

For all admissible data, the solution is unique. (40)

The solution depends continuously on the data. (41)

The question of uniqueness is relevant in inverse problems where one looks for a cause for an **observed** effect; if one just wants to find a cause for a **desired** effect, then one is usually content or even happy with having a variety of possible solutions, since then, one can try to pick one which fulfills some additional criteria.

Recall that a linear operator $K : X \rightarrow Y$ between Hilbert spaces X and Y is called “compact” iff for all bounded sets $B \subseteq X$, $\overline{K(B)}$ is compact. An important class of compact operators are integral operators

$$\begin{aligned}
 K : L^2(G) &\rightarrow L^2(G) \\
 x &\mapsto (Kx)(s) := \int_G k(s, t)x(t) dt,
 \end{aligned}
 \tag{42}$$

where $G \subseteq \mathbb{R}^n$ is compact and Jordan-measurable with positive measure, if the kernel k is in $L^2(G \times G)$ (especially, if k is continuous). If the kernel k is weakly singular, i.e., if for all $s \neq t \in G$

$$|k(s, t)| \leq \frac{M}{|s - t|^{n-\alpha}}
 \tag{43}$$

with $M > 0, \alpha > 0$, then K is defined as operator on $C(G)$ and compact there. Note that the kernel in the Abels equation (20) is weakly singular (for $n = 1$).

For any compact operator K between Hilbert spaces X and Y , there exists a “singular system” $(\sigma_i; u_i, v_i)_{i \in \mathbb{N}}$, which is defined as follows:

If $K^* : Y \rightarrow X$ denotes the adjoint of K (defined via the requirement that for all $x \in X$ and $y \in Y$, $\langle Kx, y \rangle = \langle x, K^*y \rangle$ holds), then the $(\sigma_i^2)_{i \in \mathbb{N}}$ are the non-zero eigenvalues of the self-adjoint operator K^*K (and also of KK^*), written down in decreasing order with multiplicity, $\sigma_i > 0$, the $(u_i)_{i \in \mathbb{N}}$ are a corresponding complete orthonormal system of eigenvectors of K^*K (which spans $\overline{R(K^*)} = \overline{R(K^*K)}$), and the $(v_i)_{i \in \mathbb{N}}$ are defined via

$$v_i := \frac{Ku_i}{\|Ku_i\|}. \quad (44)$$

As in the finite-dimensional situation (\rightarrow singular value decomposition of a matrix), the $(v_i)_{i \in \mathbb{N}}$ are a complete orthonormal system of eigenvalues of KK^* and span $\overline{R(K)} = \overline{R(KK^*)}$. Analogous formulas hold, namely

$$Ku_i = \sigma_i v_i \quad (45)$$

$$K^*v_i = \sigma_i u_i \quad (46)$$

$$Kx = \sum_{i=1}^{\infty} \sigma_i \langle x, u_i \rangle v_i \quad (x \in X) \quad (47)$$

$$K^*y = \sum_{i=1}^{\infty} \sigma_i \langle y, v_i \rangle u_i \quad (y \in Y), \quad (48)$$

where these infinite series converge in the Hilbert space norms of X and Y , respectively; (47) and (48) are called “singular value expansion (SVE)” and are the infinite-dimensional analogues of the SVD.

If (and only if) K has a finite-dimensional range, K has only finitely many singular values, so that all infinite series involving singular values degenerate to finite sums. If K is an integral operator of the form (42) with an L^2 -kernel k , this happens iff the kernel k is “degenerate”, i.e., has the form

$$k(s, t) = \sum_{i=1}^n \varphi_i(s) \psi_i(t) \quad (s, t \in G) \quad (49)$$

with $n \in \mathbf{N}$ and $\varphi_i, \psi_i \in L^2(G)$.

If there are infinitely many singular values, they accumulate (only) at 0, i.e.,

$$\lim_{i \rightarrow \infty} \sigma_i = 0, \quad (50)$$

which will be a crucial fact (and prove to be the inherent reason for the ill-posedness of integral equations of the first kind).

The range $R(K)$ is closed if and only if it is finite-dimensional, so that in the (generic) case of infinitely many singular values, $R(K)$ is non-closed. We will see that the closedness of the range is equivalent to the well-posedness of the best-approximate solution of a first-kind equation

$$Kx = y. \quad (51)$$

A prototype for (51) is an integral equation of the first kind

$$\int_G k(s, t)x(t) dt = y(s) \quad (s \in G) \quad (52)$$

with $k \in L^2(G \times G)$, $y \in L^2(G)$, as it comes up in many inverse problems.

Definition 7 Let $T : X \rightarrow Y$ be a bounded linear operator.

a) $x \in X$ is called “least-squares-solution” of (51), if

$$\|Tx - y\| = \inf\{\|Tz - y\|/z \in X\}. \quad (53)$$

b) $x \in X$ is called “best-approximate solution” of (51), if x is least-squares-solution of (51) and

$$\|x\| = \inf\{\|z\|/z \text{ is least-squares-solution of (51)}\} \quad (54)$$

holds.

Definition 8 The “Moore-Penrose (generalized) inverse” of $T \in L(X, Y)$, symbol T^\dagger , is defined as the

unique linear extension of \tilde{T}^{-1} to

$$D(T^\dagger) := R(T) \dot{+} R(T)^\perp \quad (55)$$

with

$$N(T^\dagger) = R(T)^\perp, \quad (56)$$

where

$$\tilde{T} := T|_{N(T)^\perp} : N(T)^\perp \rightarrow R(T). \quad (57)$$

Let now (and below) P and Q be the orthogonal projectors onto $N(T)$ and $\overline{R(T)}$, respectively. Then:

$$TT^\dagger T = T \quad (58)$$

$$T^\dagger TT^\dagger = T^\dagger \quad (59)$$

$$T^\dagger T = I - P \quad (60)$$

$$TT^\dagger = Q|_{D(T^\dagger)}. \quad (61)$$

Proposition 9 a) T^\dagger has a closed graph.

b) The following statements are equivalent:

b1) T^\dagger is bounded, i.e., continuous.

b2) $R(T)$ is closed.

Proposition 10 Let $K : X \rightarrow Y$ be compact, $\dim R(K) = \infty$. Then K^\dagger is a densely defined unbounded operator with closed graph.

Theorem 11 Let $y \in D(T^\dagger)$. Then, (51) has a unique best-approximate solution, which is given by $T^\dagger y$. The set of all least-squares solutions is $T^\dagger y + N(T)$.

Theorem 12 Let $y \in D(T^\dagger)$. Then $x \in X$ is a least-squares-solution of (51) if and only if the “normal equation”

$$T^*Tx = T^*y \tag{62}$$

holds.

Theorem 13 Let $(\sigma_n; u_n, v_n)$ be a singular system for $K; y \in Y$. Then we have:

a) $y \in D(K^\dagger) \Leftrightarrow \sum_{n=1}^{\infty} \frac{|\langle y, v_n \rangle|^2}{\sigma_n^2} < \infty$.

b) For $y \in D(K^\dagger)$,

$$K^\dagger y = \sum_{n=1}^{\infty} \frac{\langle y, v_n \rangle}{\sigma_n} u_n. \tag{63}$$

The condition for the existence of a best- approximate solution is called the “Picard criterion”.

If $\dim R(K) = \infty$, (50) holds, so that data errors of a fixed size can be amplified arbitrarily much, namely by the factor $\frac{1}{\sigma_n}$, which increases without bound. E.g., if $y_{\delta,n} := y + \delta \cdot v_n$, then $\|y_{\delta,n} - y\| = \delta$, but $K^\dagger y - K^\dagger y_{\delta,n} = \frac{\langle \delta v_n, v_n \rangle}{\sigma_n} u_n$ by (63) and hence $\|K^\dagger y - K^\dagger y_{\delta,n}\| = \frac{\delta}{\sigma_n} \rightarrow \infty$ as $n \rightarrow \infty$. A special case of this effect appeared already in Example 1 for the simple integral equation (5).

Of course, the instability in (63) (and the solvability condition of Picard’s criterion) becomes the more severe the faster the singular value decay. This makes it possible to quantify the “degree of ill-posedness” of (51): Usually, a problem is called “mildly (modestly) ill-posed”, if $\sigma = O(n^{-\alpha})$ for some $\alpha \in \mathbb{R}^+$, and “severely ill-posed”, if $\sigma_n = O(e^{-n})$ holds.

For numerical differentiation, i.e., for the integral equation (5), $\sigma_n := O(n^{-1})$, while for the Abel equation (20), which modeled the axisymmetric case of CT , $\sigma_n = O(n^{-\frac{1}{2}})$; roughly that rate also holds in general for CT in two dimensions, so that these problems are mildly ill-posed. However, “incomplete data problems” in CT , where e.g. X-ray measurements are available only for some directions, are severely ill-posed.

Example 14 We consider the one-dimensional version of the backwards heat equation.

We consider the heat equation

$$\frac{\partial u}{\partial t}(x, t) = \frac{\partial^2 u}{\partial x^2}(x, t) \quad x \in [0, \pi], t \geq 0 \quad (64)$$

with homogenous Dirichlet boundary condition

$$u(0, t) = u(\pi, t) = 0, \quad t \geq 0 \quad (65)$$

and assume that the final temperature

$$f(x) := u(x, 1) \quad (x \in [0, \pi]) \quad (66)$$

is given with $f(0) = f(\pi) = 0$; we want to determine the initial temperature

$$u_o(x) := u(x, 0) \quad (x \in [0, \pi]) \quad (67)$$

The functions $\varphi_n(x) := \sqrt{\frac{2}{\pi}} \cdot \sin(nx)$ are a complete orthonormal system in $L^2[0, \pi]$ and eigenfunctions of $\frac{d^2}{dx^2}$ on $[0, \pi]$ with homogenous Dirichlet boundary conditions. Thus, $u_o \in L^2[0, \pi]$ can be expanded as

$$u_o(x) = \sum_{n=1}^{\infty} c_n \varphi_n(x) \quad (x \in [0, \pi]) \quad (68)$$

with $c_n = \sqrt{\frac{2}{\pi}} \int_0^{\pi} u_o(\tau) \sin(n\tau) d\tau$.

Ansatz by separation of variables:

$$u(x, t) := \sum_{n=1}^{\infty} a_n(t) \varphi_n(x) \quad (x \in [0, \pi], t \geq 0) \quad (69)$$

$\sum_{n=1}^{\infty} a'_n(t) \varphi_n(x) = - \sum_{n=1}^{\infty} a_n(t) \varphi_n''(x)$ and hence since
 $(\varphi_n'' = -n^2 \varphi_n)$

$$\sum_{n=1}^{\infty} a'_n(t) \varphi_n(x) = - \sum_{n=1}^{\infty} n^2 a_n(t) \varphi_n(x). \quad (70)$$

The a_n have to solve the initial value problems

$$\left. \begin{aligned} a'_n(t) &= -n^2 a_n(t) \quad t \geq 0 \\ a_n(0) &= c_n \end{aligned} \right\} \quad (71)$$

Hence, for $n \in \mathbb{N}$,

$$a_n(t) = c_n \cdot e^{-n^2 t} \quad (t \geq 0). \quad (72)$$

Thus, $u(x, t) = \sum_{n=1}^{\infty} c_n e^{-n^2 t} \varphi_n(x)$, so that by (66),

$$\begin{aligned} f(x) &= \sum_{n=1}^{\infty} c_n e^{-n^2} \varphi_n(x) = \\ &= \frac{2}{\pi} \sum_{n=1}^{\infty} \int_0^{\pi} u_0(\tau) \sin(n\tau) d\tau e^{n^2} \sin(nx) \end{aligned}$$

With

$$k(x, \tau) := \frac{2}{\pi} \sum_{n=1}^{\infty} e^{-n^2} \sin(n\tau) \sin(nx), \quad (73)$$

we thus have

$$\int_0^{\pi} k(x, \tau) u_0(\tau) d\tau = f(x). \quad (74)$$

Thus, the inverse problem is equivalent to solving the integral equation of the first kind (74).

Note that a singular system for the integral operator in (74) is given by $(e^{-n^2}; \sqrt{\frac{2}{\pi}} \sin(nx), \sqrt{\frac{2}{\pi}} \sin(nx))$.

Since the singular functions are complete in $L^2[0, \pi]$, we have that $N(K) = N(K^*) = \{0\}$, and that $D(K^\dagger) = R(K)$ is dense in $L^2[0, \pi]$. Thus, it follows from Theorem 13 that (74) and hence our inverse problem is (uniquely) solvable if and only if

$$\sum_{n=1}^{\infty} e^{2n^2} |f_n|^2 < \infty \quad (75)$$

holds, where the

$$f_n := \sqrt{\frac{2}{\pi}} \int_0^\pi f(\tau) \sin(n\tau) d\tau \quad (76)$$

are the (classical) Fourier coefficients of f . In this case, the solution is given by

$$u_0(x) = \sqrt{\frac{2}{\pi}} \sum_{n=1}^{\infty} e^{n^2} f_n \cdot \sin(nx). \quad (77)$$

(75) and (77) show that this inverse problem is extremely ill-posed:

A solution exists only for such f for which the Fourier coefficients (f_n) decay extremely rapidly (much faster than (e^{-n^2})), i.e., for very smooth f . A small error in the n -th Fourier coefficient is amplified by the factor e^{n^2} ! Thus, already an error of, say 10^{-8} in the 5th

Fourier coefficient of the data leads to an error of about 720 in the initial temperature. Thus, one can consider at most about 3 degrees of freedom in the data and has to filter out everything else. The problem is just very severely ill-posed!

Now, we turn to methods that can cope with these problems, namely **“regularization methods”**:

In general terms, “regularization” is the approximation of an ill-posed problem by a (parameter-dependent) family of neighbouring well-posed problems. There are various ways of finding such approximations. First, we describe a method based on spectral theory.

We start by recalling that (σ_n^2, u_n) is an eigensystem for the self-adjoint compact operator K^*K and that

$$K^*Kx = \sum_{i=1}^{\infty} \sigma_i^2 \langle x, u_i \rangle u_i \quad (78)$$

holds, which will be written below as

$$K^*Kx = \int \lambda dE_\lambda x. \quad (79)$$

We define

$$\int_{-\infty}^{+\infty} f(\lambda) dE_\lambda x := \sum_{n=1}^{\infty} f(\sigma_n^2) \langle x, u_n \rangle u_n, \quad (80)$$

$$\int_{-\infty}^{+\infty} f(\lambda) d\langle E_\lambda x, y \rangle := \sum_{n=1}^{\infty} f(\sigma_n^2) \langle x, u_n \rangle \langle y, v_n \rangle, \quad (81)$$

$$\int_{-\infty}^{+\infty} f(\lambda) d\|E_\lambda x\|^2 := \sum_{n=1}^{\infty} f(\sigma_n^2) |\langle x, u_n \rangle|^2 \quad (82)$$

(for non-compact operators defined via functional calculus). Now, we will present a quite general spectral- theoretic method for constructing regularization methods. Heuristic motivation:

$$“K^\dagger y = \left(\int \frac{1}{\lambda} dE_\lambda \right) K^* y “. \quad (83)$$

However, since 0, where the integrand has a pole, is in the spectrum of K , the integral in (83) does not make sense. Now, the idea is to replace the integrand $\frac{1}{\lambda}$ by a parameter- dependent family of functions $\{(U(\alpha, \lambda), \alpha > 0\}$ which are (piecewise) continuous on an interval containing $[0, \sigma_1^2]$ and which converge to $\frac{1}{\lambda}$ as $\alpha \rightarrow 0$;

$$x_\alpha := U(\alpha, K^* K) K^* y = \left(\int U(\alpha, \lambda) dE_\lambda \right) K^* y; \quad (84)$$

for perturbed data y_δ fulfilling

$$\|y - y_\delta\| \leq \delta \quad (85)$$

we define

$$x_\alpha^\delta := U(\alpha, K^*K)K^*y_\delta \quad (86)$$

and call $x_\alpha^{(\delta)}$ “regularized solutions”.

Theorem 15 Let, for an $\varepsilon > 0$, $U : \mathbb{R}^+ \times [0, \sigma_1^2 + \varepsilon] \rightarrow \mathbb{R}$ fulfill the following assumptions

$$\text{for all } \alpha > 0, U(\alpha, \cdot) \text{ is piecewise cont.;} \quad (87)$$

$$\text{there is a } C > 0 \text{ such that for all} \quad (88)$$

$$(\alpha, \lambda), |\lambda \cdot U(\alpha, \lambda)| \leq C \text{ holds}$$

$$\text{for all } \lambda \neq 0, \lim_{\alpha \rightarrow 0} U(\alpha, \lambda) = \frac{1}{\lambda}. \quad (89)$$

Then, for all $y \in D(K^\dagger)$,

$$\lim_{\alpha \rightarrow 0} U(\alpha, K^*K)K^*y = K^\dagger y. \quad (90)$$

holds.

Theorem 16 Let U be as in Theorem 15, x_α and x_α^δ be defined by (84)–(86). For $\alpha > 0$, let

$$g_U(\alpha) := \sup\{|U(\alpha, \lambda)|/\lambda \in [0, \sigma_1^2]\}. \quad (91)$$

Then

$$\|x_\alpha - x_\alpha^\delta\| \leq \delta \cdot \sqrt{Cg_U(\alpha)} \quad (92)$$

holds.

The convergence of x_α to $K^\dagger y$ can be arbitrarily slow without additional assumptions which is typical for the approximation of ill-posed problems. A-priori assumptions that enable us to estimate the convergence rate are so-called “source conditions”

$$K^\dagger y \in R((K^* K)^\nu) \quad (93)$$

for a $\nu > 0$, which can be interpreted in various ways: First, (93) means that there exists a $w \in X$ such that

$$K^\dagger y = (K^* K)^\nu w = \sum_{i=1}^{\infty} \sigma_i^{2\nu} \langle w, u_i \rangle u_i;$$

because of (63), this is equivalent to

$$\frac{\langle y, v_i \rangle}{\sigma_i} = \sigma_i^{2\nu} \langle w, u_i \rangle \text{ for } i \in \mathbb{N}. \quad (94)$$

Since $(\langle w, v_i \rangle) \in l^2$, this in turn is equivalent to the condition

$$\sum_{i=1}^{\infty} \frac{|\langle y, v_i \rangle|^2}{\sigma_i^{2+4\nu}} < \infty, \quad (95)$$

which reduces to the Picard criterion (cf. Theorem 13a)) for $\nu = 0$.

Theorem 17 Let U be as in Theorem 15, x_α be defined by (84). Then, if (93) holds for a $\nu \in \mathbb{R}^+$,

$$\|x_\alpha - K^\dagger y\| = O(w(\alpha, \nu)). \quad (96)$$

where w is such that for all λ in an open interval containing $[0, \sigma_1^2]$,

$$\lambda^\nu |1 - \lambda U(\alpha, \lambda)| \leq w(\alpha, \nu) \quad (97)$$

holds. Then

$$\|x_\alpha^\delta - K^\dagger y\| \leq w(\alpha, \nu) + \delta \cdot \sqrt{C \cdot g_U(\alpha)}. \quad (98)$$

One has to choose α such that

$$\lim_{\delta \rightarrow 0} (\delta^2 \cdot g_U(\alpha(\delta))) = 0 \quad (99)$$

in order to obtain convergence; this condition is not only sufficient, but also necessary. $\alpha = \alpha(\delta)$ is not allowed to go to 0 too quickly, since otherwise, stability is lost in the sense that the data error is amplified too

much. Of course, also

$$\lim_{\delta \rightarrow 0} \alpha(\delta) = 0, \quad (100)$$

has to hold in order to achieve that the first error term goes to 0.

Now, one has at least the following possibilities for a parameter choice:

- a) “A–priori parameter choice”: Here, $\alpha = \alpha(\delta)$ is chosen such that the total estimate on the right–hand side of (98) goes to 0 as fast as possible. In order to optimize this estimate, however, knowledge of ν is necessary, since ν appears in the error term $w(\alpha, \nu)$. Thus, such an a–priori choice can only be optimal under the assumption of having the “correct” a–priori information (93).
- b) “A–posteriori choice”: Here, $\alpha = \alpha(\delta, y_\delta)$ is chosen during the calculations already using quantities as they appear in the calculations, e.g., certain residuals. Quite surprisingly, in this way one can find

parameter choice strategies that give rise to optimal convergence rates without having to use any a-priori knowledge of the form (93).

We specialize these results to concrete methods now:

Example 18 The simplest choice of U fulfilling the requirements of Theorem 15 is

$$U(\alpha, \lambda) := \frac{1}{\alpha + \lambda}; \quad (101)$$

for this choice, $C = 1$, $g_U(\alpha) \left(= \sup \left\{ \left| \frac{1}{\alpha + \lambda} \right| / \lambda \in [0, \sigma_1^2] \right\} \right) = \frac{1}{\alpha}$.

Hence, (99) reduces to the requirement that $\frac{\delta^2}{\alpha(\delta)} \rightarrow 0$.

We can take $w(\alpha, \nu) = \frac{\lambda^\nu \alpha}{\alpha + \lambda}$ and obtain by calculating the supremum of this function that

$$w(\alpha, \nu) = \begin{cases} o(\alpha^\nu) & 0 < \nu < 1 \\ O(\alpha) & \nu \geq 1. \end{cases} \quad (102)$$

Note that for $\nu > 1$, the rate with which $w(\alpha, \nu)$ goes to 0 does not improve as compared to $\nu = 1$. Thus, a-priori information (93) can effectively be used only for $\nu \leq 1$. This is different for other methods, as we

will see!

Now, (98) implies that

$$\|x_\alpha^\delta - K^\dagger y\| = \begin{cases} o(\alpha^\nu) + \frac{\delta}{\sqrt{\alpha}} & (0 < \nu < 1) \\ O(\alpha) + \frac{\delta}{\sqrt{\alpha}} & (\nu \geq 1) \end{cases} \quad (103)$$

under the a-priori assumption (93). If we want to choose $\alpha \sim \delta^k$ and want to optimize w.r.t. k , we immediately find that the choice minimizing the right-hand side of (103) is $k = \frac{2}{1+2\nu}$, i.e., the “optimal” a-priori choice of α is

$$\alpha(\delta) \sim \delta^{\frac{2}{1+2\nu}} \quad (104)$$

and leads to the rate

$$\|x_\alpha^\delta - K^\dagger y\| = O(\delta^{\frac{2\nu}{1+2\nu}}). \quad (105)$$

What is actually the method considered? It follows from (84) that

$$x_\alpha^\delta = (\alpha I + K^* K)^{-1} K^* y_\delta, \quad (106)$$

which is called “Tikhonov regularization”. We see that

$$x_\alpha^\delta = \sum_{n=1}^{\infty} \frac{\sigma_n}{\sigma_n^2 + \alpha} \langle y, v_n \rangle u_n, \quad (107)$$

Tikhonov regularization also has a “variational” characterization: x_α^δ is the unique minimizer of the functional

$$x \rightarrow \|Kx - y_\delta\|^2 + \alpha\|x\|^2. \quad (108)$$

Another motivation: if we have to solve (51) with perturbed data y_δ fulfilling (85), it does not make sense to look for an x fulfilling $Kx = y_\delta$. Any x fulfilling

$$\|Kx - y_\delta\| \leq \delta \quad (109)$$

is sufficient. Since we are looking for a least-squares solution of (51), it makes sense to pick (out of the many solutions of (109)) the solution of (109) with minimum norm. Thus, we solve $\|x\| \rightarrow \min.$ under the constraint (109). Via Lagrange multiplier theory, this leads to (108). The minimum of $\|x\|$ is assumed at the boundary of the feasible set (109), which gives a nonlinear equation for the Lagrange multiplier, i.e., for the regularization parameter, namely

$$\|Kx_{\alpha(\delta)}^\delta - y_\delta\| = \delta. \quad (110)$$

This parameter choice rule is called the “Morozov discrepancy principle”.

Example 19 An extension of Tikhonov regularization is “iterated Tikhonov regularization (of order n)”, where the regularized solution $x_{\alpha,n}^{\delta}$ is defined iteratively via

$$\begin{aligned}
 x_{\alpha,0}^{\delta} &:= 0 \\
 (\alpha I + K^*K)x_{\alpha,i}^{\delta} &= K^*y_{\delta} + \alpha x_{\alpha,i-1}^{\delta} \quad (i \in \{1, \dots, n\})
 \end{aligned}
 \tag{111}$$

Example 20 Another easy choice of U that fulfills the assumption of Theorem 15 is

$$U(\alpha, \lambda) = \begin{cases} \frac{1}{\lambda} & \lambda \geq \alpha \\ 0 & \lambda < \alpha. \end{cases}
 \tag{112}$$

By (86), this results in the “truncated singular value expansion”

$$x_\alpha^\delta = \sum_{\substack{n=1 \\ \sigma_n^2 \geq \alpha}}^{\infty} fty \frac{\langle y_\delta, v_n \rangle}{\sigma_n} u_n \quad (113)$$

i.e., the small singular values (“high frequencies”) are filtered out by a “low-pass filter”. By specializing the results obtained in this Chapter, we obtain that convergence is obtained if the regularization parameter (“cutoff frequency”) $\alpha = \alpha(\delta)$ is chosen such that $\lim_{\delta \rightarrow 0} \alpha(\delta) = 0$ and $\lim_{\delta \rightarrow 0} \frac{\delta^2}{\alpha(\delta)} = 0$

Also iterative regularization methods play a role. There, one use an iterative method that would converge to $K^\dagger y$ for exact data y and uses the iteration index as regularization parameter, i.e., one stops the iteration according to a “stopping rule” where the stopping index is chosen as a function of the noise level δ .

$$x_n^{(\delta)} = U(n, K^* K) K^* y_{(\delta)}, \quad (114)$$

Example 21 Let

$$U(n, \lambda) := \sum_{i=1}^n (1 + \lambda)^{-i}; \quad (115)$$

i.e., $x_n^\delta = \sum_{i=1}^n (I + K^*K)^{-i} K^* y_\delta$, which can also be written in iterative form as

$$x_0^\delta := 0 \quad (116)$$

$$(I + K^*K)x_n^\delta = K^*y_\delta + x_{n-1}^\delta \quad (n \in \mathbb{N}).$$

This method is called ‘‘Lardy’s method’’.

Example 22 A more widely used method is ‘‘Landweber iteration’’, which is defined as follows:

For $\beta \in]0, \frac{2}{\sigma_1^2}[$, let

$$x_0^\delta := \beta K^* y_\delta \quad (117)$$

$$x_n^\delta := (I - \beta K^*K)x_{n-1}^\delta + \beta K^* y_\delta.$$

As opposed to (116), where the operator $I + K^*K$ has to be inverted (which can be done in a very stable way, since all eigenvalues exceed 1), (117) is an explicit iteration.

The CG–method is iteratively defined as follows:

Let $x_0 \in N(K)^\perp$, $r_0 = -d_0 := K^*(Kx_0 - y)$, and let

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

$$\left\{ \begin{array}{l} \alpha_n \quad := \frac{\|r_n\|^2}{\|Kd_n\|^2}, \\ x_{n+1} \quad := x_n + \alpha_n d_n, \\ r_{n+1} \quad := K^*(Kx_{n+1} - y), \\ \beta_n \quad := \frac{\|r_{n+1}\|^2}{\|r_n\|^2}, \\ d_{n+1} \quad := -r_{n+1} + \beta_n d_n. \end{array} \right. \quad (118)$$

Convergence rate under (93) with $\nu > 0$:

$$\|x_n - K^\dagger y\| = O(n^{-2\nu}), \quad (119)$$

For compact K with singular values (σ_n) , the finer result

$$\|x_n - K^\dagger y\| = o(\sigma_{n+1}^{2\nu}), \quad (120)$$

holds under the same condition.

The theory developed in this chapter so far is still infinite–dimensional. For numerical computations, one still has to combine any regularization method

with some projection into a finite-dimensional space (e.g, discretization, collocation, quadrature, Galerkin methods,...). This involves an additional error, which can essentially be treated in an analogous way as the data error. Before we discuss this in more detail, we turn to “regularization by projection”:

The first idea would be to use a sequence of finite-dimensional subspaces $X_1 \subseteq X_2 \subseteq \dots \subseteq X$ with $\overline{\bigcup_{n=1}^{\infty} X_n} = X$ and use $T_n^\dagger y$ (with $T_n := T|_{X_n}$) as approximation of $T^\dagger y$. This approach, however does not always lead to convergence even for exact data:

Another approach is to construct finite-dimensional approximations via the range space:

Let $\{\psi_1, \psi_2, \psi_3, \dots\} \subseteq \overline{R(K)}$ linearly independent with linear span dense in $\overline{R(K)}$. As n -the approximation for $K^\dagger y$ ($y \in D(K^\dagger)$) we use the solution x_n

of the following problem:

$$\langle Kx, \psi_j \rangle = \langle y, \psi_j \rangle \quad j \in \{1, \dots, n\} \quad (121)$$

$$x \in X_n := \text{lin} \{K^*\psi_1, \dots, K^*\psi_n\}.$$

x_n can be computed as follows: Let $\alpha_1, \dots, \alpha_n$ be the unknown coefficients in $x_n = \sum_{i=1}^n \alpha_i K^*\psi_i$; then (121) implies for $j \in \{1, \dots, n\}$:

$$\langle y, \psi_j \rangle = \sum_{i=1}^n \alpha_i \langle KK^*\psi_i, \psi_j \rangle = \sum_{i=1}^n \alpha_i \langle K^*\psi_i, K^*\psi_j \rangle.$$

The coefficients $\alpha_1, \dots, \alpha_n$ are thus solutions of the linear system

$$Q_n \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = y_n \quad (122)$$

with

$$Q_n = (\langle K^*\psi_i, K^*\psi_j \rangle)_{1 \leq i, j \leq n}, \quad (123)$$

$$y_n = \begin{pmatrix} \langle y, \psi_1 \rangle \\ \vdots \\ \langle y, \psi_n \rangle \end{pmatrix} =: R_n y. \quad (124)$$

Since the $\psi_i \in \overline{R(K)} = N(K^*)^\perp$, the $K^*\psi_i$ are linearly independent, which implies the regularity of Q_n .

Hence, the $\alpha_1, \dots, \alpha_n$ and thus

$$x_n = \sum_{j=1}^n \alpha_j K^* \psi_j \quad (125)$$

are uniquely determined.

We have the following convergence result for exact data:

Theorem 23 For $y \in D(K^\dagger)$,

$$x_n \rightarrow K^\dagger y, \quad (126)$$

where x_n is determined by (121), or, equivalently, by (122)–(124).

The next question concerns of course the influence of perturbed data on x_n as defined by (121). As data, we take $R_n y$ (defined by (124)), which is the quantity actually used for computations, and assume that we have perturbed data fulfilling

$$\|\tilde{y}_n - R_n y\|_{\mathbb{R}^n} \leq \delta_n. \quad (127)$$

The approximate solution \tilde{x}_n is then defined by

$$\left. \begin{aligned} (\langle Kx, \psi_j \rangle)_{1 \leq j \leq n} &= \tilde{y}_n \\ x &\in X_n, \end{aligned} \right\} \quad (128)$$

We obtain the (sharp!) estimate

$$\|x_n - \tilde{x}_n\| \leq \frac{\delta_n}{\sqrt{\lambda_n}}. \quad (129)$$

Theorem 24 Let \tilde{x}_n be defined by (128) and let $(\delta_n) \rightarrow 0$ and

$$\frac{\delta_n}{\sqrt{\lambda_n}} \rightarrow 0 \text{ as } n \rightarrow \infty, \quad (130)$$

Then

$$\lim_{n \rightarrow \infty} \|\tilde{x}_n - K^\dagger y\| = 0. \quad (131)$$

Thus, λ_n plays the role of a regularization parameter; it depends on the dimension n of the approximating subspaces X_n , on the operator K and on the choice of the basis functions ψ_i .

So far, we were looking for best-approximate solutions of (51), i.e., for least-squares solutions of minimal norm (cf. Definition 7). In some applications, it will

not be appropriate to minimize the norm among all least-squares solutions; e.g., one might like to find a function fulfilling some integral equation that minimizes $\int_0^1 (x''(s))^2 ds$ among all its solutions. In order to incorporate this, one defines the “ L -best approximate solution” of (51) as the least-squares solution such that (54) is replaced by

$$\|Lx\| = \inf \{ \|Lz\| / z \text{ is least-squares solution of (51)} \} \quad (132)$$

and denotes it by $K_L^\dagger y$.

Theory can be carried over from case $L = I$ by considering the Hilbert space $D(L)$ with the inner product

$$[x_1, x_2] := \langle Kx_1, Kx_2 \rangle + \langle Lx_1, Lx_2 \rangle$$

(under some conditions on K and L).

Tikhonov regularization in this structure can also be written as

$$(K^*K + \alpha L^*L)x = K^*y_\delta. \quad (133)$$