Inverse Problems in Imaging

Martin Benning and Matthias J. Ehrhardt

Last updated on: November 29, 2016

Lecture Notes
Michaelmas Term 2016

Contents

Complementary material to the lecture notes can be found in the following books and lecture notes:

(a) Heinz Werner Engl, Martin Hanke, and Andreas Neubauer. *Regularization of Inverse Problems.* Vol. 375. Springer Science & Business Media, 1996.

(b) Martin Burger. *Inverse Problems.* Lecture notes winter 2007/2008.

`http://www.math.uni-muenster.de/num/Vorlesungen/IP_WS07/skript.pdf`

(c) Andreas Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems.* Vol. 120. Springer Science & Business Media, 1996.

(d) Kazufumi Ito and Bangti Jin. *Inverse Problems: Tikhonov Theory and Algorithms.* World Scientific, 2014.

(e) Per Christian Hansen. *Discrete Inverse Problems: Insight and Algorithms.* Fundamentals of Algorithms, SIAM Philadelphia, 2010.

(f) Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier and Frank Lenzen. *Variational Methods in Imaging.* Applied Mathematical Sciences, Springer New York, 2008.

(g) Jennifer L. Mueller and Samuli Siltanen. *Linear and Nonlinear Inverse Problems with Practical Applications.* Vol. 10. SIAM, 2012.

(h) Andreas Rieder. *Keine Probleme mit Inversen Problemen (in German).* Vieweg+Teubner Verlag. 2003.

(i) Christian Clason. *Inverse Probleme (in German)*, Lecture notes winter 2014/2015

`https://www.uni-due.de/~adf040p/teaching/inverse_14/InverseSkript.pdf`

The lecture notes are under constant redevelopment and will likely contain errors and mistakes. We very much appreciate the finding and reporting of those (to mb941@cam.ac.uk or m.j.ehrhardt@damtp.cam.ac.uk). Thanks!

Chapter 1

Introduction to inverse problems

Solving an inverse problem is the task of computing an unknown physical quantity that is related to given, indirect measurements via a forward model. Inverse problems appear in a vast majority of applications, including imaging (Computed Tomography (CT), Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI), Electron Tomography (ET), microscopic imaging, geophysical imaging), signal- and image-processing, computer vision, machine learning and (big) data analysis in general, and many more.

Mathematically, an inverse problem can be described as the solution of the operator equation

$$Ku = f \tag{1.1}$$

with given measurement data f for the unknown quantity u. Here, $K\colon \mathcal{U} \to \mathcal{V}$ denotes an operator mapping from the Banach space $\mathcal{U}$ to the Banach space $\mathcal{V}$. For the better part of this lecture, we are going to restrict ourselves to linear and bounded operators though.

Inverting a forward model however is not straightforward in most relevant applications, for two basic reasons: either a (unique) inverse model simply does not exist, or existing inverse models heavily amplify small measurement errors. In the sense of Hadamard the problem (1.1) is called *well-posed* if

- for all input data there exists a solution of the problem, i.e. for all $f \in \mathcal{V}$ there exists a $u \in \mathcal{U}$ with $Ku = f$.
- for all input data this solution is unique, i.e. $u \neq v$ implies $Kv \neq f$.
- the solution of the problem depends continuously on the input datum, i.e. for all $\{u_k\}_{k\in\mathbb{N}}$ with $Ku_k \to f$ we have $u_k \to u$.

If any of these conditions is violated, problem (1.1) is called *ill-posed*. In the following we are going to see that most practically relevant inverse problems are ill-posed or approximately ill-posed.[1]

1.1 Examples

In the following we are going to present various examples of inverse problems and highlight the challenges of dealing with them.

[1] In fact the name ill-posed problems may be a more suitable name for this lecture, as the real challenge is to deal with the ill-posedness of the inverse problems. However, the name inverse problems became more widely accepted for this area of mathematics.

1.1.1 Matrix inversion

One of the most simple (class of) inverse problems that arises from (numerical) linear algebra is the solution of linear systems. These can be written in the form of (1.1) with $u \in \mathbb{R}^n$ and $f \in \mathbb{R}^n$ being n-dimensional vectors with real entries and $K \in \mathbb{R}^{n\times n}$ being a matrix with real entries. We further assume K to be a symmetric, positive definite matrix. In that case we know from the spectral theory of symmetric matrices that there exist eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0$ and corresponding eigenvectors $k_j \in \mathbb{R}^n$ for $j \in \{1, \ldots, n\}$ such that K can be written as

$$K = \sum_{j=1}^{n} \lambda_j k_j k_j^T \, . \tag{1.2}$$

It is well known from numerical linear algebra that the condition number $\kappa = \lambda_1/\lambda_n$ is a measure of how stable (1.1) can be solved which we will illustrate in the following.

We assume that we observe f^δ instead of f, with $\|f - f^\delta\|_2 \leq \delta\|K\| = \delta\lambda_1$, where $\|\cdot\|_2$ denotes the Euclidean norm and $\|K\|$ the operator norm of K (largest singular value of K). Then, if we further denote with u^δ the solution of $Ku^\delta = f^\delta$, the difference between u^δ and the solution u of (1.1) reads as

$$u - u^\delta = \sum_{j=1}^{n} \lambda_j^{-1} k_j k_j^T \left(f - f^\delta\right) .$$

Therefore we can estimate

$$\left\|u - u^\delta\right\|_2^2 = \sum_{j=1}^{n} \lambda_j^{-2} \underbrace{\|k_j\|_2^2}_{=1} \left|k_j^T \left(f - f^\delta\right)\right|^2 \leq \lambda_n^{-2} \left\|f - f^\delta\right\|_2^2 ,$$

due to $\lambda_n \leq \lambda_j$ for $j \neq 1$ and the orthogonality of the eigenvectors. Thus, taking the square root yields the estimate

$$\left\|u - u^\delta\right\|_2 \leq \lambda_n^{-1} \left\|f - f^\delta\right\|_2 \leq \kappa\delta \, .$$

Hence, we observe that in the worst case an error δ in the data y is amplified by the condition number κ of the matrix K. A matrix with large κ is therefore called *ill-conditioned*. We want to demonstrate the effect of this error amplification with a small example.

Example 1.1. Let us consider the matrix

$$K = \begin{pmatrix} 1 & 1 \\ 1 & \frac{1001}{1000} \end{pmatrix}$$

and the vector $f = (1, 1)^T$. Then the solution of $Ku = f$ is simply given via $u = (1, 0)^T$. If we, however, consider the perturbed data $f^\delta = (99/100, 101/100)^T$ instead of f, the solution u^δ of $Ku^\delta = f^\delta$ is (exactly) $u^\delta = (-19.01, 20)^T$. The eigenvalues in this example are $\lambda_{1/2} = 1 + \frac{1}{2000} \pm \sqrt{1 + \frac{1}{2000^2}}$ which leads to the condition number $\kappa \approx 4002 \gg 1$ and operator norm $\|K\| \approx 2$. The condition number in this example reflects nicely the amplication of the noise: error in data $n = (-0.01, 0.01), \delta = \|n\|/\|K\| \approx \sqrt{2}/200)$, error in reconstruction $e = (-20.01, 20)$, $\|e\| \approx 20\sqrt{2}$, noise amplification $\|e\|/\delta \approx 4000$.

1.1.2 Differentiation

Another classic inverse problem is differentiation. Assume we are given a function f with $f(0) = 0$ for which we want to compute $u = f'$. For f smooth enough, these conditions are satisfied if and only if u and f satisfy the operator equation

$$f(y) = \int_0^y u(x)\, dx \,,$$

which can be written as the operator equation $Ku = f$ with the linear operator $(K\cdot)(y) := \int_0^y \cdot(x)\, dx$. As in the previous section, we assume that instead of f we observe a noisy version f^δ for which we further assume that the perturbation is additive, i.e. $f^\delta = f + n^\delta$ with $f \in C^1([0,1])$ and $n^\delta \in L^\infty([0,1])$.

It is obvious that the derivative u exists if the noise n^δ is differentiable. However, even in the (unrealistic) case n^δ is differentiable the error in the derivative can become arbitrarily large. Consider the sequence of noise functions $n^\delta \in C^1([0,1]) \hookrightarrow L^\infty([0,1])$ with

$$n^\delta(x) := \delta \sin\left(\frac{kx}{\delta}\right) \,, \tag{1.3}$$

for a fixed but arbitrary number k. We on the one hand observe $\left\|n^\delta\right\|_{L^\infty([0,1])} = \delta \to 0$, but on the other hand have

$$u^\delta(x) = f'(x) + k \cos\left(\frac{kx}{\delta}\right) \,,$$

and therefore obtain the estimate

$$\left\|u - u^\delta\right\|_{L^\infty([0,1])} = \left\|(n^\delta)'\right\|_{L^\infty([0,1])} = k \,.$$

Thus, despite the noise in the data becoming arbitrarily small, the error in the derivative can become arbitrarily big (depending on k). In any case for $k > 0$ we observe that the solution does not depend continuously on the data.

Note that considering a decreasing error in the norm of the Banach space $C^1([0,1])$ will yield a different result. If we have a sequence of noise functions (other than those defined in equation (1.3)) with $\left\|n^\delta\right\|_{C^1([0,1])} \leq \delta \to 0$ instead, we can conclude

$$\left\|u - u^\delta\right\|_{L^\infty([0,1])} = \left\|(n^\delta)'\right\|_{L^\infty([0,1])} \leq \left\|n^\delta\right\|_{C^1([0,1])} \to 0 \,,$$

due to $C^1([0,1])$ being embedded in $L^\infty([0,1])$. In contrast to the previous example the sequence of functions $n^\delta(x) := \delta \sin(kx)$ for instance satisfies

$$\left\|n^\delta\right\|_{C^1([0,1])} = \sup_{x\in[0,1]} \left|n^\delta(x)\right| + \sup_{x\in[0,1]} \left|(n^\delta)'(x)\right| = (1+k)\delta \to 0 \,.$$

However, for a fixed δ the bound on $\left\|u - u^\delta\right\|_{L^\infty([0,1])}$ can obviously still become fairly large compared to δ, depending on how large k is.

1.1.3 Deconvolution

An interesting problem that occurs in many imaging, image- and signal processing applications is the deblurring or *deconvolution* of signals from a known, linear degradation. Deconvolution of a signal can be modelled as solving the inverse problem of the convolution, which reads as

$$f(y) = (Ku)(y) := \int_{\mathbb{R}^n} u(x)g(y-x)\,dx\,. \tag{1.4}$$

Here f denotes the blurry image, u is the (unknown) true image and g is the function that models the degradation. Due to the Fourier convolution theorem we can rewrite (1.4) to

$$f = (2\pi)^{\frac{n}{2}}\mathcal{F}^{-1}\left(\mathcal{F}(u)\mathcal{F}(g)\right)\,. \tag{1.5}$$

with $\mathcal{F}$ denoting the Fourier transform

$$\mathcal{F}(u)(\xi) := (2\pi)^{-\frac{n}{2}}\int_{\mathbb{R}^n} u(x)\exp(-ix\cdot\xi)\,dx \tag{1.6}$$

and $\mathcal{F}^{-1}$ being the inverse Fourier transform

$$\mathcal{F}^{-1}(f)(x) := (2\pi)^{-\frac{n}{2}}\int_{\mathbb{R}^n} f(\xi)\exp(ix\cdot\xi)\,d\xi \tag{1.7}$$

It is important to note that the inverse Fourier transform is indeed the unique, inverse operator of the Fourier transform in the Hilbert-space L^2 due to the theorem of Plancherel. If we rearrange (1.5) to solve it for u we obtain

$$u = (2\pi)^{-\frac{n}{2}}\mathcal{F}^{-1}\left(\frac{\mathcal{F}(f)}{\mathcal{F}(g)}\right)\,, \tag{1.8}$$

and hence, we allegedly can recover u by simple division in the Fourier domain. However, we are quickly going to discover that this inverse problem is ill-posed and the division will lead to heavy amplifications of small errors.

Let u denote the image that satisfies (1.4). Further we assume that instead of the blurry image f we observe $f^\delta = f + n^\delta$ instead, and that u^δ is the solution of (1.8) with input datum f^δ. Hence, we observe

$$(2\pi)^{\frac{n}{2}}\left|u-u^\delta\right| = \left|\mathcal{F}^{-1}\left(\frac{\mathcal{F}(f-f^\delta)}{\mathcal{F}(g)}\right)\right| = \left|\mathcal{F}^{-1}\left(\frac{\mathcal{F}(n^\delta)}{\mathcal{F}(g)}\right)\right|\,. \tag{1.9}$$

As the convolution kernel g usually has compact support, $\mathcal{F}(g)$ will tend to zero for high frequencies. Hence, the denominator of (1.9) becomes fairly small, whereas the numerator will be non-zero as the noise is of high frequency. Thus, in the limit the solution will not depend continuously on the data and the convolution problem therefore be ill-posed.

1.1.4 Tomography

In almost any tomography application the underlying inverse problem is either the inversion of the Radon transform or of the X-ray transform in dimensions higher than two. For $u \in C_0^\infty(\mathbb{R}^n)$,

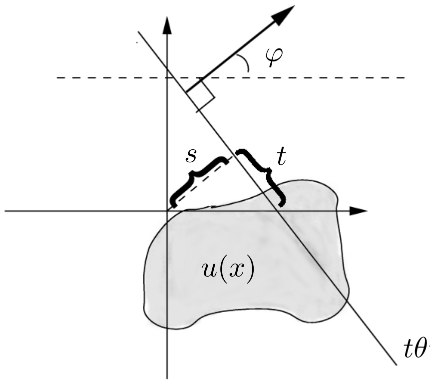

Figure 1.1: Visualization[3] of the Radon transform in 2D (which conincides with the X-ray transform). The function u is integrated over the ray parametrized by φ and s.

$s \in \mathbb{R}$ and $\theta \in S^{n-1}$, the Radon transform[2] $R : C_0^\infty(\mathbb{R}^n) \to C^\infty(S^{n-1} \times \mathbb{R})$ can be defined as the integral operator

$$f(\theta, s) = (\mathcal{R}u)(\theta, s) = \int_{x\cdot\theta=s} u(x)\, dx = \int_{\theta^\perp} u(s\theta + y)\, dy\,, \tag{1.10}$$

which for $n = 2$ coincides with the X-ray transform

$$f(\theta, s) = (\mathcal{P}u)(\theta, s) = \int_{\mathbb{R}} u(s\theta + t\theta^\perp)\, dt\,,$$

for $\theta \in S^{n-1}$ and $x \in \theta^\perp$. Hence, the X-ray transform (and therefore also the Radon transform in two dimensions) integrates the function u over lines in $\mathbb{R}^n$.

Example 1.2. Let $n = 2$. Then S^{n-1} is simply the unit sphere $S^1 = \{\theta \in \mathbb{R}^2 \mid \|\theta\|_2 = 1\}$. We can choose for instance $\theta = (\cos(\varphi), \sin(\varphi))^T$, $\varphi \in [0, 2\pi[$, and parametrise the Radon transform in terms of φ and s, i.e.

$$f(\varphi, s) = (\mathcal{R}u)(\varphi, s) = \int_{\mathbb{R}} u(s\cos(\varphi) - t\sin(\varphi), s\sin(\varphi) + t\cos(\varphi))\, dt\,. \tag{1.11}$$

Note that—with respect to the origin of the reference coordinate system—φ determines the angle of the line along one wants to integrate, while s is the offset of that line to the centre of the coordinate system.

X-ray Computed Tomography (CT)

In X-ray computed tomography (CT), the unknown quantity u represents a spatially varying density that is exposed to X-radiation from different angles, and that absorbs the radiation according to its material or biological properties.

[2]Named after the Austrian mathematician Johann Karl August Radon (16 December 1887 – 25 May 1956)
[3]Figure adapted from wikipedia `https://commons.wikimedia.org/w/index.php?curid=3001440`

The basic modelling assumption for the intensity decay of an X-ray beam is that on a small distance Δt it is proportional to the intensity itself, the density and the distance, i.e.

$$\frac{I(x + (t + \Delta t)\theta) - I(x + t\theta)}{\Delta t} = -I(x + t\theta)u(x + t\theta)\,,$$

for $x \in \theta^\perp$. By taking the limit $\Delta t \to 0$ we end up with the ordinary differential equation

$$\frac{d}{dt} I(x + t\theta) = -I(x + t\theta)u(x + t\theta)\,. \tag{1.12}$$

We now integrate (1.12) from $t = -\sqrt{R^2 - \|x\|_2^2}$, the position of the emitter, to $t = \sqrt{R^2 - \|x\|_2^2}$, the position of the detector, to obtain

$$\int_{-\sqrt{R^2-\|x\|_2^2}}^{\sqrt{R^2-\|x\|_2^2}} \frac{\frac{d}{dt}I(x + t\theta)}{I(x + t\theta)}\,dt = -\int_{-\sqrt{R^2-\|x\|_2^2}}^{\sqrt{R^2-\|x\|_2^2}} u(x + t\theta)\,dt\,.$$

Note that due to $d/dx \log(f(x)) = f'(x)/f(x)$ the left hand side in the above equation simplifies to

$$\int_{-\sqrt{R^2-\|x\|_2^2}}^{\sqrt{R^2-\|x\|_2^2}} \frac{\frac{d}{dt}I(x + t\theta)}{I(x + t\theta)}\,dt = \log\left(I\left(x + \sqrt{R^2 - \|x\|_2^2}\theta\right)\right) - \log\left(I\left(x - \sqrt{R^2 - \|x\|_2^2}\theta\right)\right)\,.$$

As we know the radiation intensity at both the emitter and the detector, we therefore know $f(x, \theta) := \log(I(x - \theta\sqrt{R^2 - \|x\|_2^2})) - \log(I(x + \theta\sqrt{R^2 - \|x\|_2^2}))$ and we can write the estimation of the unknown density u as the inverse problem of the X-ray transform (1.11) (if we further assume that u can be continuously extended to zero outside of the circle of radius R).

Positron Emission Tomography (PET)

In Positron Emission Tomography (PET) a so-called radioactive tracer (a positron emitting radionuclide on a biologically active molecule) is injected into a patient (or subject). The emitted positrons of the tracer will interact with the subjects' electrons after travelling a short distance (usually less than 1mm), causing the annihilation of both the positron and the electron, which results in a pair of gamma rays moving into (approximately) opposite directions. This pair of photons is detected by the scanner detectors, and an intensity $f(\varphi, s)$ can be associated with the number of annihilations detected at the detector pair that forms the line with offset s and angle φ (with respect to the reference coordinate system). Thus, we can consider the problem of recovering the unknown tracer density u as a solution of the inverse problem (1.10) again. The line of integration is determined by the position of the detector pairs and the geometry of the scanner.

1.1.5 Magnetic Resonance Imaging (MRI)

Magnetic resonance imaging (MRI) is an imaging technique that allows to visualise the chemical composition of patients or materials. MRI scanners use strong magnetic fields and radio waves to excite subatomic particles (like protons) that subsequently emit radio frequency signals which can be measured by the radio frequency coils. In the following we want to briefly outline the mathematics of the acquisition process. Subsequently we are going to see that finding the unknown spin proton density basically leads to solving the inverse problem of the Fourier transform (1.6).

The magnetisation of a so-called spin isochromat can be described by the Bloch equations[4]

$$\frac{d}{dt}\begin{pmatrix} M_x(t) \\ M_y(t) \\ M_z(t) \end{pmatrix} = \begin{pmatrix} -\frac{1}{T_2} & \gamma B_z(t) & -\gamma B_y(t) \\ -\gamma B_z(t) & -\frac{1}{T_2} & \gamma B_x(t) \\ \gamma B_y(t) & -\gamma B_x(t) & -\frac{1}{T_1} \end{pmatrix} \begin{pmatrix} M_x(t) \\ M_y(t) \\ M_z(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \frac{M_0}{T_1} \end{pmatrix}. \tag{1.13}$$

Here $M(t) = (M_x(t), M_y(t), M_z(t))$ is the nuclear magnetisation (of the spin isochromat), γ is the gyromagnetic ratio, $B(t) = (B_x(t), B_y(t), B_z(t))$ denotes the magnetic field experienced by the nuclei, T_1 is the longitudinal and T_2 the transverse relaxation time and M_0 the magnetisation in thermal equilibrium. If we define $M_{xy}(t) = M_x(t) + iM_y(t)$ and $B_{xy}(t) = B_x(t) + iB_y(t)$, we can rewrite (1.13) to

$$\frac{d}{dt}M_{xy}(t) = -i\gamma\left(M_{xy}(t)B_z(t) - M_z(t)B_{xy}(t)\right) - \frac{M_{xy}(t)}{T_2} \tag{1.14a}$$

$$\frac{d}{dt}M_z(t) = i\frac{\gamma}{2}\left(M_{xy}(t)\overline{B_{xy}}(t) - \overline{M_{xy}}(t)B_{xy}(t)\right) - \frac{M_z(t) - M_0}{T_1} \tag{1.14b}$$

with $\bar{\cdot}$ denoting the complex conjugate of $\cdot$.

If we assume for instance that $B = (0, 0, B_0)$ is just a constant magnetic field in z-direction, (1.14) reduces to the decoupled equations

$$\frac{d}{dt}M_{xy}(t) = -i\gamma B_0 M_{xy}(t) - \frac{M_{xy}(t)}{T_2}, \tag{1.15a}$$

$$\frac{d}{dt}M_z(t) = -\frac{M_z(t) - M_0}{T_1}. \tag{1.15b}$$

It is easy to see that this system of equations (1.15) has the unique solution

$$M_{xy}(t) = e^{-t(i\omega_0 + 1/T_2)}M_{xy}(0) \tag{1.16a}$$

$$M_z(t) = M_z(0)e^{-\frac{t}{T_1}} + M_0\left(1 - e^{-\frac{t}{T_1}}\right) \tag{1.16b}$$

for $\omega_0 := \gamma B_0$ denoting the Lamor frequency, and $M_{xy}(0)$, $M_z(0)$ being the initial magnetisations at time $t = 0$.

Rotating frame

Thus, for a constant magnetic background field in z-direction, B_0, M_{xy} basically rotates around the z-axis in clockwise direction with frequency ω_0 (if we ignore the T_2 decay for a moment). Rotating the x- and y-coordinate axes with the same frequency yields the representation of the Bloch equations in the so-called rotating frame. If we substitute $M^r_{xy}(t) := e^{i\omega_0 t}M_{xy}(t)$, $B^r_{xy}(t) := B_{xy}(t)e^{i\omega_0 t}$, $M^r_z(t) := M_z(t)$ and $B^r_z(t) := B_z(t)$, we obtain

$$\frac{d}{dt}M^r_{xy}(t) = -i\gamma\left(M^r_{xy}(t)(B^r_z(t) - B_0) - M^r_z(t)B^r_{xy}(t)\right) - \frac{M^r_{xy}(t)}{T_2} \tag{1.17a}$$

$$\frac{d}{dt}M^r_z(t) = i\frac{\gamma}{2}\left(M^r_{xy}(t)\overline{B^r_{xy}}(t) - \overline{M^r_{xy}}(t)B^r_{xy}(t)\right) - \frac{M^r_z(t) - M_0}{T_1} \tag{1.17b}$$

instead of (1.14).

[4]Named after the Swiss born American physicist Felix Bloch (23 October 1905 - 10 September 1983)

Thus, if we assume the magnetic field to be constant with magnitude B_0 in z-direction within the rotating frame, i.e. $B^r(t) = (B^r_x(t), B^r_y(t), B_0)$, (1.17a) simplifies to

$$\frac{d}{dt} M^r_{xy}(t) = i\gamma M^r_z(t) B^r_{xy}(t) - \frac{M^r_{xy}(t)}{T_2} \,. \tag{1.18}$$

90° pulse

Now we assume that $B^r_x(t) = c$, c constant, and $B^r_y(t) = 0$ for $t \in [0, \tau]$, and $\tau \ll T_1$ and $\tau \ll T_2$. Then we can basically ignore the effect of $M^r_{xy}(t)/T_2$ and $(M^r_z(t) - M_0)/T_1$, and the Bloch equations in the rotating frame simplify to

$$\frac{d}{dt} \begin{pmatrix} M^r_x(t) \\ M^r_y(t) \\ M^r_z(t) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \omega \\ 0 & -\omega & 0 \end{pmatrix} \begin{pmatrix} M^r_x(t) \\ M^r_y(t) \\ M^r_z(t) \end{pmatrix} \tag{1.19}$$

with $\omega := \gamma c$, in matrix form with separate components. Assuming the initial magnetisations in the rotating frame to be zero in the x-y plane, i.e. $M^r_x(0) = 0$ and $M^r_y(0) = 0$, and constant in the z-plane with value $M^r_z(0)$, the solution of (1.19) can be written as

$$\begin{pmatrix} M^r_x(t) \\ M^r_y(t) \\ M^r_z(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\omega t) & \sin(\omega t) \\ 0 & -\sin(\omega t) & \cos(\omega t) \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ M^r_z(0) \end{pmatrix} . \tag{1.20}$$

Thus, equation (1.20) rotates the initial z-magnetisation around the x-axis by the angle $\theta := \omega t$. Note that if c and τ are chosen such that $\theta = \pi$, all magnetisation is rotated from the z-axis to the y-axis, i.e. $M^r_y(\tau) = M^r_z(0)$. In analogy, choosing $B_x(t) = 0$ and $B_y(t) = c$, all magnetisation can be shifted from the z- to the x-axis.

Signal acquisition

If the radio-frequency (RF) pulse is turned off and thus, $B^r_x(t) = 0$ and $B^r_y(t) = 0$ for $t > \tau$, the same coils that have been used to induce the RF pulse can be used to measure the x-y magnetisation. Since we measure a volume of the whole x-y net-magnetisation, the acquired signal equals

$$y(t) = \int_{\mathbb{R}^3} M(x,t)\, dx = \int_{\mathbb{R}^2} e^{-i\omega_0(x)t} M^r(x,t)\, dx \tag{1.21}$$

with $M(x,t)$ denoting $M_{xy}(t)$ for a specific spatial coordinate $x \in \mathbb{R}^3$ ($M^r(x,t)$ respectively). Using (1.16a) and assuming $\tau < t \ll T_2$, this yields

$$y(t) = \int_{\mathbb{R}^3} M_\tau(x) e^{-i\omega_0(x)t}\, dx \,, \tag{1.22}$$

with M_τ denoting the x-y-magnetisation at spatial location $x \in \mathbb{R}^3$ and time $t = \tau$. Note that $M_\tau = 0$ without any RF pulse applied in advance.

Signal recovery

The basic clue to allow for spatially resolving nuclear magnetic resonance spectrometry is to add a magnetic field $\hat{B}(t)$ to the constant magnetic field B_0 in z-direction that varies spatially over time. Then, (1.15a) changes to

$$\frac{d}{dt}M_{xy}(t) = -i\gamma(B_0 + \hat{B}(t))M_{xy}(t) - \frac{M_{xy}(t)}{T_2}\,,$$

which, for initial value $M_{xy}(0)$, has the unique solution

$$M_{xy}(t) = e^{-i\gamma\left(B_0 t + \int_0^t \hat{B}(\tau)\,d\tau\right)} e^{-\frac{t}{T_2}} M_{xy}(0) \tag{1.23}$$

if we ensure $\hat{B}(0) = 0$. If now $x(t)$ denotes the spatial location of a considered spin isochromat at time t, we can write $\hat{B}(t)$ as $\hat{B}(t) = x(t) \cdot G(t)$, with a function G that describes the influence of the magnetic field gradient over time.
Based on the considerations that lead to (1.22) we therefore measure

$$y(t) = \int_{\mathbb{R}^3} M_\tau(x) e^{-i\gamma\left(B_0(x)t + \int_0^t x(\tau)\cdot G(\tau)\,d\tau\right)}\,dx$$

in an NMR experiment. Assuming that B_0 is also constant in space, we can consider the equation in the rotating frame (see Section 1.1.5) by eliminating this term and by writing the signal acquisition as

$$e^{i\gamma B_0 t} y(t) = \int_{\mathbb{R}^3} M_\tau(x) e^{-i\gamma \int_0^t x(\tau)\cdot G(\tau)\,d\tau}\,dx\,. \tag{1.24}$$

In the following we assume that $x(t)$ can be approximated reasonably well via its zero-order Taylor approximation around $t_0 = 0$, i.e.

$$\int_0^t x(\tau)\cdot G(\tau)\,d\tau \approx x(0)\cdot \int_0^t G(\tau)\,d\tau\,. \tag{1.25}$$

and hence, the inverse problem of finding the unknown spin-proton density M_τ for given measurements y is equivalent to solving the inverse problem of the Fourier transform

$$f(t) = (KM_\tau)(t) = \int_{\mathbb{R}^3} M_\tau(x) e^{-ix\cdot\xi(t)}\,dx\,, \tag{1.26}$$

with $f(t) := e^{i\gamma B_0 t} y(t)$ and $\xi(t) := \gamma \int_0^t G(\tau)\,d\tau$.

Chapter 2

Linear inverse problems

Throughout this lecture we deal with functional analytic operators. For the sake of brevity, we cannot recall all basic concepts of functional analysis but refer to popular textbooks that deal with this subject, like [4, 15]. Nevertheless, we want to recall a few important properties that are going to be important for the further course of this lecture. In particular, we are going to focus mainly on inverse problems with *bounded, linear operators* K only, i.e. $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ with

$$\|K\|_{\mathcal{L}(\mathcal{U},\mathcal{V})} := \sup_{u \in \mathcal{U} \setminus \{0\}} \frac{\|Ku\|_{\mathcal{V}}}{\|u\|_{\mathcal{U}}} = \sup_{\|u\|_{\mathcal{U}} \leq 1} \|Ku\|_{\mathcal{V}} < \infty \,.$$

For $K : \mathcal{U} \to \mathcal{V}$ we further want to denote with

(a) $\mathcal{D}(K) := \mathcal{U}$ the domain

(b) $\mathcal{N}(K) := \{u \in \mathcal{U} \mid Ku = 0\}$ the kernel

(c) $\mathcal{R}(K) := \{f \in \mathcal{V} \mid f = Ku, u \in \mathcal{U}\}$ the range

of K. We say that K is continuous in $u \in \mathcal{U}$ if there exists a $\delta > 0$ for all $\varepsilon > 0$ with

$$\|Ku - Kv\|_{\mathcal{V}} \leq \varepsilon \text{ for all } v \in \mathcal{U} \text{ with } \|u - v\|_{\mathcal{U}} \leq \delta.$$

For linear K it can be shown that continuity is equivalent to the existence of a positive constant C such that

$$\|Ku\|_{\mathcal{V}} \leq C\|u\|_{\mathcal{U}}$$

for all $u \in \mathcal{U}$. Note that this constant C actually equals the operator norm $\|K\|_{\mathcal{L}(\mathcal{U},\mathcal{V})}$.

For the first part of the lecture we only consider $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ with $\mathcal{U}$ and $\mathcal{V}$ being Hilbert spaces. From functional calculus we know that every Hilbert space is equipped with a *scalar product*, which we are going to denote by $\langle \cdot, \cdot \rangle_{\mathcal{U}}$ (if $\mathcal{U}$ denotes the corresponding Hilbertspace). In analogy to the transpose of a matrix, this scalar product structure together with the theorem of Fréchet-Riesz [15, Section 2.10, Theorem 2.E] allows us to define the (unique) *adjoint operator* of K, denoted with K^*, as follows:

$$\langle Ku, v \rangle_{\mathcal{V}} = \langle u, K^* v \rangle_{\mathcal{U}} \,.$$

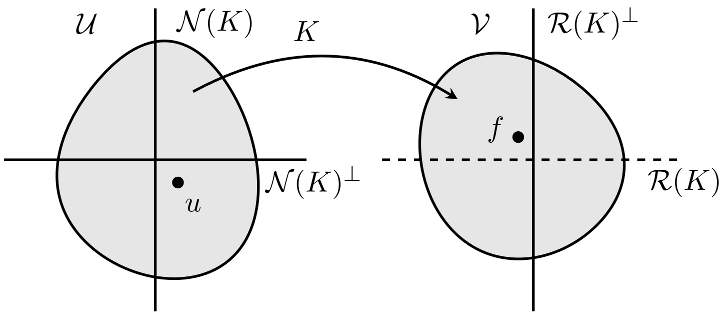

Figure 2.1: Visualization of the setting for linear inverse problems where we want to solve the inverse problem (1.1). The operator K is a linear mapping between $\mathcal{U}$ and $\mathcal{V}$. The kernel $\mathcal{N}(\mathcal{U})$ and range $\mathcal{R}(K)$ are used to analyse solutions to the inverse problem.

In addition to that, a scalar product allows to have a notion of orthogonality. Two functions $u, v \in \mathcal{U}$ are said to be *orthogonal* if $\langle u, v\rangle_{\mathcal{U}} = 0$. For a subset $\mathcal{X} \subset \mathcal{U}$ the *orthogonal complement* of $\mathcal{X}$ in $\mathcal{U}$ is defined as

$$\mathcal{X}^\perp := \{u \in \mathcal{U} \mid \langle u, v\rangle_{\mathcal{U}} = 0 \text{ for all } v \in \mathcal{X}\} \, .$$

From this definition we immediately observe that $\mathcal{X}^\perp$ is a closed subspace. Further we have $\mathcal{U}^\perp = \{0\}$. Moreover, we have $\mathcal{X} \subset (\mathcal{X}^\perp)^\perp$. If $\mathcal{X}$ is a closed subspace we even have $\mathcal{X} = (\mathcal{X}^\perp)^\perp$. In this case there exists the *orthogonal decomposition*

$$\mathcal{U} = \mathcal{X} \oplus \mathcal{X}^\perp \, ,$$

which means that every element $u \in \mathcal{U}$ can uniquely be represented as

$$u = x + x^\perp \text{ with } x \in \mathcal{X}, x^\perp \in \mathcal{X}^\perp \, ,$$

see for instance [15, Section 2.9, Corollary 1]. The mapping $u \mapsto x$ defines a linear operator $P_{\mathcal{X}} \in \mathcal{L}(\mathcal{U}, \mathcal{U})$ that is called *orthogonal projection* on $\mathcal{X}$.

Lemma 2.1 (cf. [10, Section 5.16])**.** *Let $\mathcal{X} \subset \mathcal{U}$ be a closed subspace. The orthogonal projection onto $\mathcal{X}$ satisfies the following conditions:*

(a) $P_{\mathcal{X}}$ is self-adjoint, i.e. $P_{\mathcal{X}}^* = P_{\mathcal{X}}$;

(b) $\|P_{\mathcal{X}}\|_{\mathcal{L}(\mathcal{U},\mathcal{U})} = 1$ (if $\mathcal{X} \neq \{0\}$);

(c) $I - P_{\mathcal{X}} = P_{\mathcal{X}^\perp}$;

(d) $\|u - P_{\mathcal{X}} u\|_{\mathcal{U}} \leq \|u - v\|_{\mathcal{U}}$ for all $v \in \mathcal{X}$;

(e) $x = P_{\mathcal{X}} u$ if and only if $x \in \mathcal{X}$ and $u - x \in \mathcal{X}^\perp$.

Remark 2.1. Note that for a non-closed subspace $\mathcal{X}$ we only have $(\mathcal{X}^\perp)^\perp = \overline{\mathcal{X}}$. For $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ we therefore have $\mathcal{R}(K)^\perp = \mathcal{N}(K^*), \mathcal{R}(K^*)^\perp = \mathcal{N}(K)$ and thus $\mathcal{N}(K^*)^\perp = \overline{\mathcal{R}(K)}$ and $\mathcal{N}(K)^\perp = \overline{\mathcal{R}(K^*)}$. Hence, we can conclude the orthogonal decompositions

$$\mathcal{U} = \mathcal{N}(K) \oplus \overline{\mathcal{R}(K^*)} \text{ and } \mathcal{V} = \mathcal{N}(K^*) \oplus \overline{\mathcal{R}(K)} \, .$$

In the following we want to investigate the concept of generalised inverses of bounded linear operators, before we will identify compactness of operators as the major source of ill-posedness. Subsequently we are going to discuss this in more detail by analysing compact operators in terms of their singular value decomposition.

2.1 Generalised solutions

In order to overcome the issues of non-existence or non-uniqueness we want to generalise the concept of least squares solutions to linear operators in Hilbert spaces.

If we consider the generic inverse problem (1.1) again, we know that there does not exist a solution of the inverse problem if $f \notin \mathcal{R}(K)$. In that case it seems reasonable to find an element $u \in \mathcal{U}$ for which $\|Ku - f\|_{\mathcal{V}}$ gets minimal instead. If $\mathcal{V} = L^2$ then u minimizes the squared error and thus motivates the name least squares solution.

However, for $\mathcal{N}(K) \neq \{0\}$ there are infinitely many solutions that minimise $\|Ku - f\|_{\mathcal{V}}$ of which we have to pick one. Picking the one with minimal norm $\|u\|_{\mathcal{U}}$ brings us to the definition of the minimal norm solution.

Definition 2.1. *We call $u \in \mathcal{U}$ a* least squares solution *of the inverse problem* (1.1)*, if*

$$\|Ku - f\|_{\mathcal{V}} \leq \|Kv - f\|_{\mathcal{V}} \quad \text{for all } v \in \mathcal{U}. \tag{2.1}$$

Furthermore, we call $u^\dagger \in \mathcal{U}$ a minimal norm solution *of the inverse problem* (1.1)*, if*

$$\|u^\dagger\|_{\mathcal{U}} \leq \|v\|_{\mathcal{U}} \quad \text{for all least squares solutions } v. \tag{2.2}$$

Remark 2.2. Let u be a least squares solution to $Ku = f$. It is easy to see that each $v \in \{u\} + \mathcal{N}(K)$ is a least squares solution as well.

Moreover, let $u^\dagger$ be a minimal norm solution, then $u^\dagger \in \mathcal{N}(K)^\perp$. If this was not the case, then there exists $x^\perp \in \mathcal{N}(K)^\perp$ and $x \in \mathcal{N}(K)$ with $\|x\|_{\mathcal{U}} > 0$ such that $u^\dagger = x + x^\perp$. It is obvious that $x^\perp$ is a least squares solution and has smaller norm than $u^\dagger$ which contradicts that $u^\dagger$ is of minimal norm, thus $u^\dagger \in \mathcal{N}(K)^\perp$.

In numerical linear algebra it is a well known fact that the normal equations can be considered to compute least squares solutions. The same is true in the continuous case.

Theorem 2.1. *Let $f \in \mathcal{V}$ and $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$. The following three assertions are equivalent.*

(a) $u \in \mathcal{U}$ satisfies $Ku = P_{\overline{\mathcal{R}(K)}} f$

(b) u is a least squares solution of the inverse problem (1.1)*.*

(c) u solves the normal equation

$$K^* K u = K^* f\,. \tag{2.3}$$

Remark 2.3. The name normal equation is derived from the fact that for any solution u its residual $Ku - f$ is orthogonal (normal) to $\mathcal{R}(K)$. This can be readily seen, as we have for any $v \in \mathcal{U}$ that

$$0 = \langle v, K^*(Ku - f)\rangle_{\mathcal{U}} = \langle Kv, Ku - f\rangle_{\mathcal{V}}$$

which shows $Ku - f \in \mathcal{R}(K)^\perp$.

Proof of Theorem 2.1. For (a) $\Rightarrow$ (b): Let $u \in \mathcal{U}$ such that $Ku = P_{\overline{\mathcal{R}(K)}} f$ and arbitrary $v \in \mathcal{U}$. With the basic properties of the orthogonal projection (Lemma 2.1 (d)), we have

$$\|Ku - f\|_{\mathcal{V}}^2 = \|(I - P_{\overline{\mathcal{R}(K)}})f\|_{\mathcal{V}}^2 \leq \inf_{g \in \overline{\mathcal{R}(K)}} \|g - f\|_{\mathcal{V}}^2 \leq \inf_{v \in \mathcal{U}} \|Kv - f\|_{\mathcal{V}}^2$$

which shows that u is a least squares solution.

For (b) $\Rightarrow$ (c): Let $u \in \mathcal{U}$ be a least squares solution $v \in \mathcal{U}$ an arbitrary element and define the quadratic polynom $F\colon \mathbb{R} \to \mathbb{R}$,

$$F(\lambda) := \|K(u + \lambda v) - f\|_{\mathcal{V}}^2 = \lambda^2 \|Kv\|_{\mathcal{V}}^2 - 2\lambda \langle Kv, f - Ku\rangle_{\mathcal{V}} + \|f - Ku\|_{\mathcal{V}}^2 .$$

A necessary condition for $u \in \mathcal{U}$ to be a least squares solution is $F'(0) = 0$ which leads to $\langle v, K^*(f - Ku)\rangle_{\mathcal{V}} = 0$. As v was arbitrary, it follows that the normal equation holds.

For (c) $\Rightarrow$ (a): From the normal equation it follows that $K^*(f - Ku) = 0$ which is equivalent to $f - Ku \in \mathcal{R}(K)^{\perp}$, see Remark 2.3. As $\mathcal{R}(K)^{\perp} = \overline{\mathcal{R}(K)}^{\perp}$ and $Ku \in \mathcal{R}(K) \subset \overline{\mathcal{R}(K)}$, the assertion follows from the basic properties of the orthogonal projection, i.e. Lemma 2.1 (e). □

Lemma 2.2. *Let $f \in \mathcal{V}$ and $\mathbb{L}$ be the set of least squares solutions to the inverse problem* (1.1). *Then $\mathbb{L}$ is non-empty if and only if $f \in \mathcal{R}(K) \oplus \mathcal{R}(K)^{\perp}$.*

Proof. Let $u \in \mathbb{L}$. It is easy to see that $f = Ku + (f - Ku) \in \mathcal{R}(K) \oplus \mathcal{R}(K)^{\perp}$ as the normal equations are equivalent to $f - Ku \in \mathcal{R}(K)^{\perp}$.

Consider now $f \in \mathcal{R}(K) \oplus \mathcal{R}(K)^{\perp}$. Then there exists $u \in \mathcal{U}$ and $g \in \mathcal{R}(K)^{\perp} = \overline{\mathcal{R}(K)}^{\perp}$ such that $f = Ku + g$ and thus $P_{\overline{\mathcal{R}(K)}} f = P_{\overline{\mathcal{R}(K)}} Ku + P_{\overline{\mathcal{R}(K)}} g = Ku$ and the assertion follows from Theorem 2.1. □

Remark 2.4. If the dimensions of $\mathcal{U}$ and $\mathcal{R}(K)$ are finite, then $\mathcal{R}(K)$ is closed, i.e. $\overline{\mathcal{R}(K)} = \mathcal{R}(K)$. Thus, in a finite dimensional setting, there always exists a least squares solution.

It is natural to ask whether there are always least squares solutions. From the remark above it is clear that we have to look for an example in infinite dimensional spaces. The answer is negative as we see from the following counter example.

Example 2.1. Let $\mathcal{U} = \ell^2, \mathcal{V} = \ell^2$, where the space ℓ^2 is the space of all square summable sequences, i.e.

$$\ell^2 := \left\{ \{x_j\}_{j\in\mathbb{N}} \;\middle|\; x_j \in \mathbb{R},\ \sum_{j=1}^{\infty} x_j^2 < \infty \right\} .$$

It is a Hilbert space with inner product and norm given by

$$\langle x, y\rangle_{\ell^2} := \sum_{j=1}^{\infty} x_j y_j \quad \text{and} \quad \|x\|_{\ell^2} := \left(\sum_{j=1}^{\infty} x_j^2 \right)^{1/2} .$$

For more information, consult for instance [4].

Consider the inverse problem $Kx = f$, where the linear operator $K\colon \ell^2 \to \ell^2$ is defined by

$$(Kx)_j := \frac{x_j}{j} .$$

and the data by $f_j := j^{-1}$. It is easy to see that K is linear and bounded (what is the operator norm of K?), i.e. $K \in \mathcal{L}(\ell^2, \ell^2)$ and $f \in \ell^2$.

We will show that $f \in \overline{\mathcal{R}(K)} \setminus \mathcal{R}(K)$ and thus $f \notin \mathcal{R}(K) \oplus \mathcal{R}(K)^{\perp}$. With Lemma 2.2 it follows then that there are no least squares solutions.

First we show that $f \notin \mathcal{R}(K)$ by contradiction. Assume that $f \in \mathcal{R}(K)$, then there exists $x \in \ell^2$ such that $Kx = f$ and thus $j^{-1}x_j = j^{-1}$ for all $j \in \mathbb{N}$. Therefore, we have $x_j = 1$ and $x \notin \ell^2$.

Next we show that $f \in \overline{\mathcal{R}(K)}$. Let $\{x^k\}_{k\in\mathbb{N}} \subset \ell^2$ be a sequence in ℓ^2 (each element is a sequence as well), with

$$(x^k)_j := \begin{cases} 1, & j \leq k \\ 0, & j > k \end{cases} .$$

It is easy to see that $x^k \in \ell^2$ as it has only finitely many non-negative components. In addition, we have

$$f^k := Kx^k, \quad (f^k)_j = \begin{cases} \frac{1}{j}, & j \leq k \\ 0, & j > k \end{cases}$$

and therefore

$$\|f - f^k\|_{\ell^2}^2 = \sum_{j=k+1}^{\infty} f_j^2 = \sum_{j=1}^{\infty} f_j^2 - \sum_{j=1}^{k} f_j^2 \to 0$$

by definition of a convergent series. Therefore $f^k \to f$ in ℓ^2 and thus $f \in \overline{\mathcal{R}(K)}$.

Theorem 2.2. *Let $f \in \mathcal{R}(K) \oplus \mathcal{R}(K)^\perp$. Then there exists a unique minimal norm solution $u^\dagger$ to the inverse problem* (1.1) *and all least squares solutions are given by $\{u^\dagger\} + \mathcal{N}(K)$.*

Proof. From Lemma 2.2 we know that there exist least squares solutions and denote any arbitrary two of them (not necessarily different) by $u, v \in \mathcal{U}$. Then there exist $\varphi, \psi \in \mathcal{N}(K)^\perp$ and $x, y \in \mathcal{N}(K)$ such that $u = \varphi + x$ and $v = \psi + y$. As we noted in Remark 2.2 φ and ψ are least squares solutions as well. With Theorem 2.1 we conclude

$$K(\varphi - \psi) = K\varphi - K\psi = P_{\overline{\mathcal{R}(K)}}f - P_{\overline{\mathcal{R}(K)}}f = 0 \tag{2.4}$$

which shows that $\varphi - \psi \in \mathcal{N}(K)$. But as $\varphi - \psi \in \mathcal{N}(K)^\perp$ and $\mathcal{N}(K) \cap \mathcal{N}(K)^\perp = \{0\}$ we see that $\varphi = \psi$. Therefore all least squares solutions are of the form $\{\varphi\} + \mathcal{N}(K)$.

Moreover, we know that $u^\dagger$ is a least squares solution and that $u^\dagger \in \mathcal{N}(K)^\perp$, see Remark 2.2. Thus we have that $u^\dagger = \varphi$ which completes the proof. □

Corollary 2.1. *The minimal norm solution is the unique solution of the normal equation in $\mathcal{N}(K)^\perp$.*

2.2 Generalised inverse

We have seen that for arbitrary f a least squares solution does not need to exist if $\mathcal{R}(K)$ is not closed. If, however, a least squares solution exists, then we have shown that the minimum norm solution is unique. We will see in the following that the minimum norm solution can be computed via the *Moore–Penrose* generalised inverse.

Definition 2.2. *Let $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ and denotes the restriction of K to $\mathcal{N}(K)^\perp$ by*

$$\tilde{K} := K|_{\mathcal{N}(K)^\perp} : \mathcal{N}(K)^\perp \to \mathcal{R}(K)\,.$$

The Moore–Penrose inverse $K^\dagger\colon \mathcal{D}(K^\dagger) \to \mathcal{U}$ is defined as the unique linear extension of $\tilde{K}^{-1}$ to $\mathcal{D}(K^\dagger) := \mathcal{R}(K) \oplus \mathcal{R}(K)^\perp$ with $K^\dagger f = 0$ for $f \in \mathcal{R}(K)^\perp$, i.e. $\mathcal{N}(K^\dagger) = \mathcal{R}(K)^\perp$.

Remark 2.5. Note that $\tilde{K}$ is injective due to the restriction to $\mathcal{N}(K)^\perp$, and surjective due to the restriction to $\mathcal{R}(K)$. Hence, $\tilde{K}^{-1}$ exists, and—as a consequence—$K^\dagger$ is well-defined on $\mathcal{R}(K)$. Due to the orthogonal decomposition $\mathcal{D}(K^\dagger) = \mathcal{R}(K) \oplus \mathcal{R}(K)^\perp$, for arbitrary $f \in \mathcal{D}(K^\dagger)$ there exist $f_1 \in \mathcal{R}(K)$ and $f_2 \in \mathcal{R}(K)^\perp$ with $f = f_1 + f_2$. Hence, we have

$$K^\dagger f = K^\dagger f_1 + K^\dagger f_2 = K^\dagger f_1 = \tilde{K}^{-1} f_1 = \tilde{K}^{-1} P_{\overline{\mathcal{R}(K)}} f\,, \tag{2.5}$$

where we used that $f_2 \in \mathcal{R}(K)^\perp = \mathcal{N}(K^\dagger)$. Thus, $K^\dagger$ is well-defined on the whole of $\mathcal{D}(K^\dagger)$.

Example 2.2. To illustrate the definition of the Moore–Penrose inverse we consider a simple example in finite dimensions. Let the linear operator $K\colon \mathbb{R}^3 \to \mathbb{R}^2$ be given by

$$Kx = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 0 \end{pmatrix}$$

and consider the right hand side $\tilde{f} = (8, 1)^T$. It is easy to see that $\mathcal{R}(K) = \{f \in \mathbb{R}^2 \mid f_2 = 0\}$ and $\mathcal{N}(K) = \{x \in \mathbb{R}^3 \mid x_1 = 0\}$, thus $\mathcal{N}(K)^\perp = \{x \in \mathbb{R}^3 \mid x_2, x_3 = 0\}$. Therefore, $\tilde{K}\colon \mathcal{N}(K)^\perp \to \mathcal{R}(K), x \mapsto (2x_1, 0)^T$ which is bijective and can be easily inverted: $\tilde{K}^{-1}\colon \mathcal{R}(K) \to \mathcal{N}(K)^\perp, f \mapsto (f_1/2, 0, 0)^T$. As the orthogonal projection onto $\mathcal{R}(K)$ is given by $f = (f_1, f_2) \mapsto (f_1, 0)$, the Moore–Penrose inverse of K is $K^\dagger\colon \mathbb{R}^2 \to \mathbb{R}^3$,

$$K^\dagger f = \begin{pmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} f_1/2 \\ 0 \\ 0 \end{pmatrix}\,,$$

and thus $K^\dagger \tilde{f} = K^\dagger (8, 0)^T = (4, 0, 0)^T$.

It can be shown that $K^\dagger$ can be characterized by the Moore–Penrose equations.

Lemma 2.3. *The Moore–Penrose inverse $K^\dagger$ is linear and $\mathcal{R}(K^\dagger) = \mathcal{N}(K)^\perp$. Moreover, it satisfies the Moore–Penrose equations*

(a) $KK^\dagger K = K$,

(b) $K^\dagger K K^\dagger = K^\dagger$,

(c) $K^\dagger K = I - P_{\mathcal{N}(K)}$,

(d) $KK^\dagger = P_{\overline{\mathcal{R}(K)}}\Big|_{\mathcal{D}(K^\dagger)}$,

where $P_{\mathcal{N}(K)}$ and $P_{\overline{\mathcal{R}(K)}}$ denote the orthogonal projections on $\mathcal{N}(K)$ and $\overline{\mathcal{R}(K)}$, respectively.

Proof. First of all we note that by definition the Moore–Penrose inverse is linear.

Let us now prove $\mathcal{R}(K^\dagger) = \mathcal{N}(K)^\perp$. Let $u \in \mathcal{R}(K^\dagger)$, then there exists a $f \in \mathcal{D}(K^\dagger)$ with $u = K^\dagger f$ and according to (2.5) we observe that $u = K^\dagger f = \tilde{K}^{-1} P_{\overline{\mathcal{R}(K)}} f$. Hence, $u \in \mathcal{R}(\tilde{K}^{-1}) = \mathcal{N}(K)^\perp$ and therefore $\mathcal{R}(K^\dagger) \subset \mathcal{N}(K)^\perp$. To prove $\mathcal{N}(K)^\perp \subset \mathcal{R}(K^\dagger)$, let $u \in \mathcal{N}(K)^\perp$ and it holds $u = \tilde{K}^{-1}\tilde{K}u = K^\dagger K u$, thus $u \in \mathcal{R}(K^\dagger)$.

It remains to prove the Moore–Penrose equations. We begin with (d): For $f \in \mathcal{D}(K^\dagger)$ it follows from (2.5) and $K = \tilde{K}$ on $\mathcal{N}(K)^\perp$ that

$$KK^\dagger f = K\tilde{K}^{-1} P_{\overline{\mathcal{R}(K)}} f = \tilde{K}\tilde{K}^{-1} P_{\overline{\mathcal{R}(K)}} f = P_{\overline{\mathcal{R}(K)}} f \, .$$

(c): According to the definition of $K^\dagger$ we have $K^\dagger K u = \tilde{K}^{-1} K u$ for all $u \in \mathcal{U}$ and thus

$$K^\dagger K u = \tilde{K}^{-1} \underbrace{K P_{\mathcal{N}(K)} u}_{=0} + \tilde{K}^{-1} K \underbrace{(I - P_{\mathcal{N}(K)})}_{=P_{\mathcal{N}(K)^\perp}} u = (I - P_{\mathcal{N}(K)}) u \, .$$

(b): Inserting (d) into (2.5) yields

$$K^\dagger f = K^\dagger P_{\overline{\mathcal{R}(K)}} f = K^\dagger K K^\dagger f \, .$$

(a): With (c) we have

$$KK^\dagger K = K(I - P_{\mathcal{N}(K)}) = K - KP_{\mathcal{N}(K)} = K \, .$$

□

The following theorem states that minimum norm solutions can be computed via the generalised inverse.

Theorem 2.3. *For each $f \in \mathcal{D}(K^\dagger)$ the minimal norm solution $u^\dagger$ to the inverse problem* (1.1) *is given via*

$$u^\dagger = K^\dagger f \, .$$

Proof. As $f \in \mathcal{D}(K^\dagger)$, we know from Theorem 2.2 that the minimal norm solution $u^\dagger$ exists and is unique. With $u^\dagger \in \mathcal{N}(K)^\perp$, Lemma 2.3 and Theorem 2.1 we conclude that

$$u^\dagger = (I - P_{\mathcal{N}(K)}) u^\dagger = K^\dagger K u^\dagger = K^\dagger P_{\overline{\mathcal{R}(K)}} f = K^\dagger K K^\dagger f = K^\dagger f \, .$$

□

As a direct consequence from Theorem 2.3 and Theorem 2.1 we obtain

$$K^\dagger f = (K^* K)^\dagger K^* f \, ,$$

and hence, in order to approximate $K^\dagger f$ we may also compute an approximation via the normal equations (2.3) instead.

At the end of this section we further want to analyse the domain of the generalised inverse in more detail. Due to the construction of the Moore–Penrose inverse we have $\mathcal{D}(K^\dagger) = \mathcal{R}(K) \oplus \mathcal{R}(K)^\perp$. As orthogonal complements are always closed we can conclude

$$\overline{\mathcal{D}(K^\dagger)} = \overline{\mathcal{R}(K)} \oplus \mathcal{R}(K)^\perp = \mathcal{V} \, ,$$

and hence, $\mathcal{D}(K^\dagger)$ is dense in $\mathcal{V}$. Thus, if $\mathcal{R}(K)$ is closed it follows that $\mathcal{D}(K^\dagger) = \mathcal{V}$ and vice versa, $\mathcal{D}(K^\dagger) = \mathcal{U}$ implies $\mathcal{R}(K)$ to be closed. Moreover, for $f \in \mathcal{R}(K)^\perp = \mathcal{N}(K^\dagger)$ the minimum norm solution is $u^\dagger = 0$. Therefore, for given $f \in \overline{\mathcal{R}(K)}$ the important question to address is when f also satisfies $f \in \mathcal{R}(K)$. If this is the case, $K^\dagger$ has to be continuous. However, the existence of a single element $f \in \overline{\mathcal{R}(K)} \setminus \mathcal{R}(K)$ is enough already to prove that $K^\dagger$ is discontinuous.

Definition 2.3. *Let $\mathcal{V}$ and $\mathcal{U}$ be Hilbert spaces and consider $A\colon \mathcal{V} \to \mathcal{U}$. We call the graph of A*

$$\mathrm{gr}(A) := \{(f,u) \in \mathcal{V} \times \mathcal{U} \mid Af = u\}$$

closed, if for any sequence $\{(f_j, u_j)\}_{j\in\mathbb{N}}$ with $u_j = Af_j$, $f_j \to f \in \mathcal{V}$ and $u_j \to u \in \mathcal{U}$ we have $Af = u$.

Theorem 2.4 (Closed graph theorem [13, Proposition 2.14 and Theorem 2.15])**.** *Let $\mathcal{V}$ and $\mathcal{U}$ be Hilbert spaces, $A\colon \mathcal{V} \to \mathcal{U}$ a linear mapping and with a closed graph. Then $A \in \mathcal{L}(\mathcal{V},\mathcal{U})$.*

Theorem 2.5. *Let $K \in \mathcal{L}(\mathcal{U},\mathcal{V})$. Then $K^\dagger$ is continuous, i.e. $K^\dagger \in \mathcal{L}(\mathcal{D}(K^\dagger),\mathcal{U})$, if and only if $\mathcal{R}(K)$ is closed.*

Proof. We will show first that the graph of the Moore–Penrose inverse is closed. Let $\{(f_j, u_j)\}_{j\in\mathbb{N}} \subset \mathrm{gr}(K^\dagger)$ be a sequence in the graph of the Moore–Penrose inverse, i.e. $u_j = K^\dagger f_j$, and $f_j \to f$ and $u_j \to u$. Then because of Lemma 2.1 and the continuity of the orthogonal projection and K, we have

$$Ku = \lim_{j\to\infty} Ku_j = \lim_{j\to\infty} KK^\dagger f_j = \lim_{j\to\infty} P_{\overline{\mathcal{R}(K)}} f_j = P_{\overline{\mathcal{R}(K)}} f\,,$$

thus u is a least squares solution. As $K^\dagger f_j \in \mathcal{N}(K)^\perp$ and $\mathcal{N}(K)^\perp$ is closed, we have $u \in \mathcal{N}(K)^\perp$ and it follows from the uniqueness of the minimal norm solution that $u = K^\dagger f$. This shows that the graph of $K^\dagger$ is closed.

For the proof of the actual theorem, assume first that $\mathcal{R}(K)$ is closed so that $\mathcal{D}(K^\dagger) = \mathcal{V}$. Therefore, $K^\dagger$ is bounded by the Closed graph theorem (Theorem 2.4).

Conversely, let $K^\dagger$ be continuous. As $\mathcal{D}(K^\dagger)$ is dense in $\mathcal{V}$, there is a unique continuous extension A of $K^\dagger$ to $\mathcal{V}$. From Lemma 2.3 (d), $KK^\dagger = P_{\overline{\mathcal{R}(K)}}|_{\mathcal{D}(K^\dagger)}$, and the continuity of K we conclude that $KA = P_{\overline{\mathcal{R}(K)}}$ so that for $f \in \overline{\mathcal{R}(K)}$, $f = P_{\overline{\mathcal{R}(K)}} f = KAf \in \mathcal{R}(K)$. Thus, $\overline{\mathcal{R}(K)} \subset \mathcal{R}(K)$, so that $\mathcal{R}(K)$ is closed. □

In the next section we are going to discover that the class of compact operators is a class for which the Moore–Penrose inverses are discontinuous.

2.3 Compact operators

Compact operators are very common in inverse problems; in fact, almost all (linear) inverse problems involve the inversion of compact operators. Compact operators are defined as follows.

Definition 2.4. *Let $K \in \mathcal{L}(\mathcal{U},\mathcal{V})$. Then K is said to be* compact *if the image of a bounded sequence $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}$ contains a convergent subsequence $\{Ku_{j_k}\}_{k\in\mathbb{N}} \subset \mathcal{V}$. We denote the space of compact operators by $\mathcal{K}(\mathcal{U},\mathcal{V})$.*

Remark 2.6. We can equivalently define an operator K to be compact if and only if for any bounded set B, the closure of its image $\overline{K(B)}$ is compact.

Example 2.3 (Follows from e.g. [16, p. 49])**.** Let $I\colon \mathcal{U} \to \mathcal{U}$ be the identity operator on $\mathcal{U}$, i.e. $u \mapsto u$. Then I is compact if and only if the dimension of I is finite.

Example 2.4 (e.g. [16, p. 286, Proposition 5] or [4, p. 186])**.** Let $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$. If the range of K is finite dimensional, then K is compact.

Example 2.5 ([8, p. 230])**.** The operator $K\colon \ell^2 \to \ell^2, (Kx)_j = j^{-1}x_j$ from Example 2.1 is compact.

Example 2.6 ([8, p. 231])**.** Let $\emptyset \neq \Omega \subset \mathbb{R}^n$ be compact. Let $k \in L^2(\Omega \times \Omega)$ and define the integral operator $K\colon L^2(\Omega) \to L^2(\Omega)$ with

$$(Ku)(x) = \int_\Omega k(x,y)u(x)\,dy\,.$$

Then K is compact.

Example 2.7 ([11, p. 38])**.** Let $B := \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$ denote the unit ball in $\mathbb{R}^2$ and $Z := [-1,1] \times [0,\pi)$. Moreover, let $\theta(\varphi) := (\cos(\varphi), \sin(\varphi))^T$, $\theta^\perp(\varphi) := (\sin(\varphi), -\cos(\varphi))^T$ be the unit vectors pointing in the direction described by φ and orthogonal to it. Then the Radon transform / X-ray transform is defined as the operator $R\colon L^2(B) \to L^2(Z)$ with

$$(Ru)(s,\varphi) := \int_{-\sqrt{1-s^2}}^{\sqrt{1-s^2}} u\Big(s\theta(\varphi) + t\theta^\perp(\varphi)\Big)\,dt\,.$$

It can be shown that the Radon transform is linear and continuous, i.e. $R \in \mathcal{L}(L^2(B), L^2(Z))$, and even compact, i.e. $R \in \mathcal{K}(L^2(B), L^2(Z))$.

Compact operators can be seen as the infinite dimensional analogue to ill-conditioned matrices. Indeed it can be seen that compactness is a main source of ill-posedness in infinite dimensions, confirmed by the following result.

Theorem 2.6. *Let $K \in \mathcal{K}(\mathcal{U}, \mathcal{V})$ with an infinite dimensional range. Then, the Moore–Penrose inverse of K is discontinuous.*

Proof. As the range $\mathcal{R}(K)$ is of infinite dimension, we can conclude that $\mathcal{U}$ and $\mathcal{N}(K)^\perp$ are also infinite dimensional. We can therefore find a sequence $\{u_j\}_{j\in\mathbb{N}}$ with $u_j \in \mathcal{N}(K)^\perp$, $\|u_j\|_\mathcal{U} = 1$ and $\langle u_j, u_k\rangle_\mathcal{U} = 0$ for $j \neq k$. Since K is a compact operator the sequence $f_j = Ku_j$ has a convergent subsequence, hence, for all $\delta > 0$ we can find j, k such that $\|f_j - f_k\|_\mathcal{V} < \delta$. However, we also obtain

$$\begin{aligned}\|K^\dagger f_j - K^\dagger f_k\|_\mathcal{U}^2 &= \|K^\dagger K u_j - K^\dagger K u_k\|_\mathcal{U}^2 \\ &= \|u_j - u_k\|_\mathcal{U}^2 = \|u_j\|_\mathcal{U}^2 - 2\langle u_j, u_k\rangle_\mathcal{U} + \|u_k\|_\mathcal{U}^2 = 2\,,\end{aligned}$$

which shows that $K^\dagger$ is discontinuous. □

To have a better understanding of when we have $f \in \overline{\mathcal{R}(K)} \setminus \mathcal{R}(K)$ for compact operators K, we want to consider the singular value decomposition of compact operators.

2.4 Singular value decomposition of compact operators

We want to characterise the Moore–Penrose inverse of compact operators in terms of a spectral decomposition. Like in the finite dimensional case of matrices, we can only expect a spectral decomposition to exist for self-adjoint operators.

Theorem 2.7 ([8, p. 225, Theorem 9.16]). *Let $\mathcal{U}$ be a Hilbert space and $K \in \mathcal{K}(\mathcal{U},\mathcal{U})$ be self-adjoint. Then there exists an orthonormal basis $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}$ of $\overline{\mathcal{R}(K)}$ and a sequence of Eigenvalues $\{\lambda_j\}_{j\in\mathbb{N}} \subset \mathbb{R}$ with $|\lambda_1| \geq |\lambda_2| \geq \ldots > 0$ such that for all $u \in \mathcal{U}$ we have*

$$Ku = \sum_{j=1}^{\infty} \lambda_j \langle u, u_j \rangle_{\mathcal{U}} u_j \,.$$

The sequence $\{\lambda_j\}_{j\in\mathbb{N}}$ is either finite or we have $\lambda_j \to 0$.

Remark 2.7. The notation in the theorem above only makes sense if the sequence $\{\lambda_j\}_{j\in\mathbb{N}}$ is infinite. For the case that there are only finitely many λ_j the sum has to be interpreted as a finite sum.

Moreover, as the absolute value of the Eigenvalues $|\lambda_j|$ are sorted, we have $\|K\|_{\mathcal{L}(\mathcal{U},\mathcal{U})} = |\lambda_1|$.

Due to Theorem 2.1 we can consider K^*K instead of K, which brings us to the singular value decomposition of linear, compact operators.

Theorem 2.8. *Let $K \in \mathcal{K}(\mathcal{U},\mathcal{V})$. Then there exists i) a not-necessarily infinite null sequence $\{\sigma_j\}_{j\in\mathbb{N}}$ with $\sigma_1 \geq \sigma_2 \geq \ldots > 0$, ii) an orthonormal basis $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}$ of $\mathcal{N}(K)^\perp$ and iii) an orthonormal basis $\{v_j\}_{j\in\mathbb{N}} \subset \mathcal{V}$ of $\overline{\mathcal{R}(K)}$, with*

$$Ku_j = \sigma_j v_j, \quad K^* v_j = \sigma_j u_j \quad \text{for all } j \in \mathbb{N} \tag{2.6}$$

and for all $w \in \mathcal{U}$ we have the representation

$$Kw = \sum_{j=1}^{\infty} \sigma_j \langle w, u_j \rangle_{\mathcal{U}} \, v_j \,. \tag{2.7}$$

The sequence $\{(\sigma_j, u_j, v_j)\}$ is called singular system *or* singular value decomposition (SVD) *of K.*

Proof. As $K^*K \colon \mathcal{U} \to \mathcal{U}$ is compact and self-adjoint, by Theorem 2.7 there exists a decreasing (in terms of absolute values) null sequence $\{\lambda_j\}_{j\in\mathbb{N}} \subset \mathbb{R} \setminus \{0\}$ and an orthonormal basis $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}$ of $\overline{\mathcal{R}(K^*K)}$ with $K^*Ku = \sum_{j=1}^{\infty} \lambda_j \langle u, u_j \rangle_{\mathcal{U}} u_j$ for all $u \in \mathcal{U}$.

Due to

$$\lambda_j = \lambda_j \|u_j\|_{\mathcal{U}}^2 = \langle \lambda_j u_j, u_j \rangle_{\mathcal{U}} = \langle K^*Ku_j, u_j \rangle_{\mathcal{U}} = \langle Ku_j, Ku_j \rangle_{\mathcal{V}} = \|Ku_j\|_{\mathcal{V}}^2 > 0$$

we can define $\sigma_j := \sqrt{\lambda_j}$ and $v_j := (Ku_j)/\sigma_j \in \mathcal{V}$ for all $j \in \mathbb{N}$. Further, we observe

$$K^* v_j = \frac{1}{\sigma_j} K^*Ku_j = \frac{\lambda_j}{\sigma_j} u_j = \sigma_j u_j$$

which proves Equation (2.6).

We also obverse that $\{v_j\}_{j\in\mathbb{N}}$ form an orthonormal system due to

$$\langle v_i, v_j\rangle_{\mathcal{V}} = \frac{1}{\sigma_i\sigma_j}\langle Ku_i, Ku_j\rangle_{\mathcal{V}} = \frac{1}{\sigma_i\sigma_j}\langle K^*Ku_i, u_j\rangle_{\mathcal{U}} = \frac{\lambda_i}{\sigma_i\sigma_j}\langle u_i, u_j\rangle_{\mathcal{U}} = \begin{cases} 1 & i=j \\ 0 & \text{else} \end{cases}.$$

We know that $\{u_j\}_{j\in\mathbb{N}}$ is an orthonormal basis of $\overline{\mathcal{R}(K^*K)}$ and we want to show that it is also an orthonormal basis of $\mathcal{N}(K)^\perp$. As $\overline{\mathcal{R}(K^*)} = \mathcal{N}(K)^\perp$ it is sufficient to show that $\overline{\mathcal{R}(K^*K)} = \overline{\mathcal{R}(K^*)}$. It is clear that $\overline{\mathcal{R}(K^*K)} = \overline{\mathcal{R}(K^*|_{\mathcal{R}(K)})} \subset \overline{\mathcal{R}(K^*)}$, such that we are left to prove $\overline{\mathcal{R}(K^*)} \subset \overline{\mathcal{R}(K^*K)}$. Let $u \in \overline{\mathcal{R}(K^*)}$ and $\varepsilon > 0$. Then there exists $f \in \mathcal{N}(K^*)^\perp$ with $\|K^*f - u\|_{\mathcal{U}} < \varepsilon/2$. As $\mathcal{N}(K^*)^\perp = \overline{\mathcal{R}(K)}$, there exists $x \in \mathcal{U}$ such that $\|Kx - f\|_{\mathcal{V}} < \varepsilon/(2\|K\|)$. Putting these together we have

$$\begin{aligned}\|K^*Kx - u\|_{\mathcal{U}} &\le \|K^*Kx - K^*f\|_{\mathcal{U}} + \|K^*f - u\|_{\mathcal{U}} \\ &\le \underbrace{\|K^*\|\|Kx - f\|_{\mathcal{V}}}_{<\varepsilon/2} + \underbrace{\|K^*f - u\|_{\mathcal{U}}}_{<\varepsilon/2} < \varepsilon\end{aligned}$$

which shows that $u \in \overline{\mathcal{R}(K^*K)}$.

To show the basis representation of Kw let

$$w_N := \sum_{j=1}^N \langle w, u_j\rangle_{\mathcal{U}} u_j$$

be a finite approximation of the basis representation of $w \in \mathcal{N}(K)^\perp$. Then it is easy to see that

$$\begin{aligned}Kw_N &= \sum_{j=1}^N \langle w, u_j\rangle_{\mathcal{U}} Ku_j = \sum_{j=1}^N \sigma_j\langle w, u_j\rangle_{\mathcal{U}} v_j \\ &= \sum_{j=1}^N \langle w, \sigma_j u_j\rangle_{\mathcal{U}} v_j = \sum_{j=1}^N \langle w, K^*v_j\rangle_{\mathcal{U}} v_j = \sum_{j=1}^N \langle Kw, v_j\rangle_{\mathcal{V}} v_j .\end{aligned}$$

With $w_N \to w$ and the continuity of K we have that

$$Kw = \lim_{N\to\infty} Kw_N = \lim_{N\to\infty} \sum_{j=1}^N \sigma_j\langle w_N, u_j\rangle_{\mathcal{U}} v_j = \sum_{j=1}^\infty \sigma_j\langle w_N, u_j\rangle_{\mathcal{U}} v_j$$

which shows (2.7). Moreover, we can also conclude that

$$Kw = \sum_{j=1}^\infty \langle Kw, u_j\rangle_{\mathcal{U}} v_j$$

which shows that $\{v_j\}_{j\in\mathbb{N}}$ is an orthonormal basis of $\mathcal{R}(K)$ and thus also of $\overline{\mathcal{R}(K)}$. □

Remark 2.8. Since Eigenvalues of K^*K with Eigenvectors u_j are also Eigenvalues of KK^* with Eigenvectors v_j, we further obtain a singular value decomposition of K^*, i.e.

$$K^*z = \sum_{j=1}^\infty \sigma_j\langle z, v_j\rangle_{\mathcal{V}}\, u_j .$$

A singular system allows us to characterize elements in the range of the operator.

Theorem 2.9. *Let* $K \in \mathcal{K}(\mathcal{U}, \mathcal{V})$ *with singular system* $\{(\sigma_j, u_j, v_j)\}_{j \in \mathbb{N}}$, *and* $f \in \overline{\mathcal{R}(K)}$. *Then* $f \in \mathcal{R}(K)$ *if and only if the* Picard criterion

$$\sum_{j=1}^{\infty} \frac{|\langle f, v_j \rangle_{\mathcal{V}}|^2}{\sigma_j^2} < \infty \tag{2.8}$$

is met.

Proof. Let $f \in \mathcal{R}(K)$, thus there is a $u \in \mathcal{U}$ such that $Ku = f$. It is easy to see that we have

$$\langle f, v_j \rangle_{\mathcal{V}} = \langle Ku, v_j \rangle_{\mathcal{V}} = \langle u, K^* v_j \rangle_{\mathcal{U}} = \sigma_j \langle u, u_j \rangle_{\mathcal{U}}$$

and therefore

$$\sum_{j=1}^{\infty} \sigma_j^{-2} |\langle f, v_j \rangle_{\mathcal{V}}|^2 = \sum_{j=1}^{\infty} |\langle u, u_j \rangle_{\mathcal{U}}|^2 \leq \|u\|_{\mathcal{U}}^2 < \infty \,.$$

Now let the Picard criterion (2.8) hold and define $u := \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f, v_j \rangle_{\mathcal{V}} u_j \in \mathcal{U}$. It is well-defined by the Picard criterion (2.8) and we conclude

$$Ku = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f, v_j \rangle_{\mathcal{V}} K u_j = \sum_{j=1}^{\infty} \langle f, v_j \rangle_{\mathcal{V}} v_j = P_{\overline{\mathcal{R}(K)}} f = f \,,$$

which shows $f \in \mathcal{R}(K)$. □

Remark 2.9. The Picard criterion is a condition on the decay of the coefficents $\langle f, v_j \rangle_{\mathcal{V}}$. As the singular values σ_j decay to zero as $j \to \infty$, the Picard criterion is only met if the coefficients $\langle f, v_j \rangle_{\mathcal{V}}$ decay sufficiently fast.

In case the singular system is given by the Fourier basis, then the coefficents $\langle f, v_j \rangle_{\mathcal{V}}$ are just the Fourier coefficents of f. Therefore, the Picard criterion is a condition on the decay of the Fourier coefficients which is equivalent to the smoothness of f.

We can now derive a representation of the Moore–Penrose inverse in terms of the singular value decomposition.

Theorem 2.10. *Let* $K \in \mathcal{K}(\mathcal{U}, \mathcal{V})$ *with singular system* $\{(\sigma_j, v_j, u_j)\}_{j \in \mathbb{N}}$ *and* $f \in \mathcal{D}(K^\dagger)$. *Then the Moore–Penrose inverse of* K *can be written as*

$$K^\dagger f = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f, v_j \rangle_{\mathcal{V}} u_j \,. \tag{2.9}$$

Proof. As $f \in \mathcal{R}(K) \oplus \mathcal{R}(K)^\perp$ there exist $u \in \mathcal{N}(K)^\perp$ and $g \in \mathcal{R}(K)^\perp$ such that $f = Ku + g$. As $\{u_j\}_{j \in \mathbb{N}}$ is an orthonormal system of $\mathcal{N}(K)^\perp$ we have that

$$\begin{aligned} u &= \sum_{j=1}^{\infty} \langle u, u_j \rangle_{\mathcal{U}} u_j = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle u, \sigma_j u_j \rangle_{\mathcal{U}} u_j = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle u, K^* v_j \rangle_{\mathcal{U}} u_j \\ &= \sum_{j=1}^{\infty} \sigma_j^{-1} \langle Ku, v_j \rangle_{\mathcal{U}} u_j = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f - g, v_j \rangle_{\mathcal{U}} u_j = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f, v_j \rangle_{\mathcal{U}} u_j \end{aligned}$$

where we used for the last equality that $g \in \mathcal{R}(K)^\perp$ and $v_j \in \overline{\mathcal{R}(K)}$.

Moreover, in addition to $u \in \mathcal{N}(K)^\perp$ we have that u satisfies the normal equation

$$K^*Ku = \sum_{j=1}^{\infty} \sigma_j^2 \sigma_j^{-1} \langle f, v_j \rangle_{\mathcal{U}} u_j = \sum_{j=1}^{\infty} \sigma_j \langle f, v_j \rangle_{\mathcal{U}} u_j = K^* f$$

and is therefore the minimal norm solution to the inverse problem $Ku = f$ (1.1). With Theorem 2.3 we conclude that $u = K^\dagger f$. □

From representation (2.9) we can see what happens in case of noisy measurements. Assume we are given $f^\delta = f + \delta v_j$ and denote by $u^\dagger$ and $u_\delta^\dagger$ the minimal norm solutions of $Ku = f$ and $Ku = f^\delta$. Then we observe

$$\|u^\dagger - u_\delta^\dagger\|_{\mathcal{U}} = \|K^\dagger f - K^\dagger f^\delta\|_{\mathcal{U}} = \delta \|K^\dagger v_j\|_{\mathcal{U}} = \frac{\delta}{\sigma_j} \to \infty \text{ for } j \to \infty\,.$$

For static j we see that the amplification of the error δ depends on how small σ_j is. Hence, the faster the singular values decay, the stronger the amplification of errors. For that reason, one distinguishes between three classes of ill-posed problems:

Definition 2.5. *We say that an ill-posed inverse problem* (1.1) *is* severely ill-posed *if the singular values decay as* $\sigma_j = \mathcal{O}(\exp(-j))$*, where the "Big-O-notation" means that there exists* j_0 *and* $c > 0$ *such that for all* $j \geq j_0$ *there is* $\sigma_j \leq c\exp(-j)$*. We call the ill-posed inverse problem* mildly ill-posed *if it is not severely ill-posed.*

Example 2.8. Let us consider the example of differentiation again, as introduced in Section 1.1.2. The operator $K\colon L^2([0,1] \to L^2([0,1]$ of the inverse problem (1.1) of differentiation is given as

$$(Ku)(y) = \int_0^y u(x)\, dx = \int_0^1 k(x,y) u(x)\, dx\,,$$

with $k\colon [0,1] \times [0,1] \to \mathbb{R}$ defined as

$$k(x,y) := \begin{cases} 1 & x \leq y \\ 0 & \text{else} \end{cases}.$$

This is a special case of the integral operators as introduced in Example 2.6 due to its kernel k being square integrable and thus K is compact.

In order to compute the singular value decomposition of K we compute its adjoint K^* first, which is characterised via

$$\langle Ku, v \rangle_{L^2([0,1])} = \langle u, K^* v \rangle_{L^2([0,1])}\,.$$

Hence, we obtain

$$\langle Ku, v \rangle_{L^2([0,1])} = \int_0^1 \int_0^1 k(x,y) u(x)\, dx\, v(y)\, dy = \int_0^1 u(x) \int_0^1 k(x,y) v(y)\, dy\, dx\,.$$

Hence, the adjoint operator K^* is given via

$$(K^* v)(x) = \int_0^1 k(x,y) v(y)\, dy = \int_x^1 v(y)\, dy\,. \tag{2.10}$$

Now we want to compute the Eigenvalues and Eigenvectors of K^*K, i.e. we look for $\lambda > 0$ and $u \in L^2([0,1])$ with

$$\lambda u(x) = (K^*Ku)(x) = \int_x^1 \int_0^y u(z)\, dz\, dy\,.$$

We immediately observe $u(1) = 0$ and further

$$\lambda u'(x) = \frac{d}{dx} \int_x^1 \int_0^y u(z)\, dz\, dy = -\int_0^x u(z)\, dz\,,$$

from which we conclude $u'(0) = 0$. Taking the derivative another time thus yields the ordinary differential equation

$$\lambda u''(x) + u(x) = 0\,,$$

for which solutions are of the form

$$u(x) = c_1 \sin(\sigma^{-1}x) + c_2 \cos(\sigma^{-1}x)\,,$$

with $\sigma := \sqrt{\lambda}$ and constants c_1, c_2. In order to satisfy the boundary conditions $u(1) = c_1 \sin(\sigma^{-1}) + c_2 \cos(\sigma^{-1}) = 0$ and $u'(0) = c_1 = 0$, we chose $c_1 = 0$ and σ such that $\cos(\sigma^{-1}) = 0$. Hence, we have

$$\sigma_j = \frac{2}{(2j-1)\pi} \text{ for } j \in \mathbb{N}\,,$$

and by choosing $c_2 = \sqrt{2}$ we obtain the following normalised representation of u_j:

$$u_j(x) = \sqrt{2} \cos\left(\left(j - \frac{1}{2}\right)\pi x\right)\,.$$

According to (2.6) we further obtain

$$v_j(x) = \sigma_j^{-1}(Ku_j)(x) = \left(j - \frac{1}{2}\right)\pi \int_0^x \sqrt{2} \cos\left(\left(j - \frac{1}{2}\right)\pi y\right)\, dy = \sqrt{2} \sin\left(\left(j - \frac{1}{2}\right)\pi x\right)\,,$$

and hence, for $f \in L^2([0,1])$ the Picard criterion becomes

$$2 \sum_{j=1}^{\infty} \sigma_j^{-2} \left(\int_0^1 f(x) \sin\left(\sigma_j^{-1}x\right)\, dx\right)^2 < \infty\,.$$

Thus, the Picard criterion holds if f is differentiable and $f' \in L^2([0,1])$.

From the decay of the singular values we see that this inverse problem is mildly ill-posed.

Chapter 3

Regularisation

We have seen in the previous section that the major source of ill-posedness of inverse problems of the type (1.1) is a fast decay of the singular values of K. An idea to overcome this issue is to define approximations of $K^\dagger$ in the following fashion. Consider the family of operators

$$R_\alpha f := \sum_{j=1}^{\infty} g_\alpha(\sigma_j) \langle f, v_j \rangle_{\mathcal{V}} \, u_j \,, \tag{3.1}$$

with functions $g_\alpha : \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$ that converge to $1/\sigma_j$ as α converges to zero. We are going to see that such an operator R_α is what is called a *regularisation* (of $K^\dagger$), if g_α is bounded, i.e.

$$g_\alpha(\sigma) \leq C_\alpha \text{ for all } \sigma \in \mathbb{R}_{>0}. \tag{3.2}$$

In case (3.2) holds true, we immediately observe

$$\|R_\alpha f\|_{\mathcal{U}}^2 = \sum_{j=1}^{\infty} g_\alpha(\sigma_j)^2 |\langle f, v_j \rangle_{\mathcal{V}}|^2 \leq C_\alpha^2 \sum_{j=1}^{\infty} |\langle f, v_j \rangle_{\mathcal{V}}|^2 \leq C_\alpha^2 \|f\|_{\mathcal{V}}^2 \,,$$

which means that C_α is a bound for the norm of R_α and thus $R_\alpha \in \mathcal{L}(\mathcal{V}, \mathcal{U})$.

Example 3.1 (Truncated singular value decomposition)**.** As a first example for a spectral regularisation of the form (3.1) we want to consider the so-called *truncated singular value decomposition.* As the name suggests, the idea is to discard all singular values below a certain threshold α

$$g_\alpha(\sigma) = \begin{cases} \frac{1}{\sigma} & \sigma \geq \alpha \\ 0 & \sigma < \alpha \end{cases} . \tag{3.3}$$

Note that for all $\sigma > 0$ we naturally obtain $\lim_{\alpha \to 0} g_\alpha(\sigma) = 1/\sigma$. Equation (3.1) then reads as

$$R_\alpha f = \sum_{\sigma_j \geq \alpha} \frac{1}{\sigma_j} \langle f, v_j \rangle_{\mathcal{V}} \, u_j \,, \tag{3.4}$$

for all $f \in \mathcal{V}$. Note that (3.4) is always well-defined (i.e. finite) for $\alpha > 0$ as zero is the only accumulation point of singular vectors of compact operators. From (3.3) we immediately observe $g_\alpha(\sigma) \leq 1/\alpha$ so that $\|R_\alpha\|_{\mathcal{L}(\mathcal{U},\mathcal{V})} \leq 1/\alpha$.

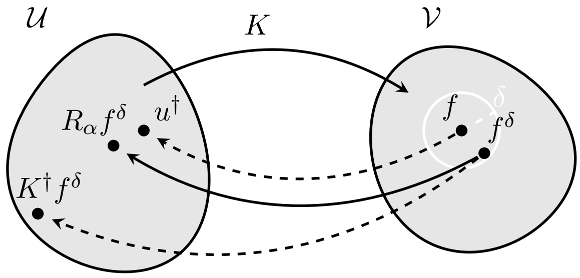

Figure 3.1: Visualization of reconstruction from noisy data. While the Moore–Penrose inverse reconstructs optimally from noiseless data, its noise amplification renders it useless when small errors are present in the data. A regularisation operator gives a robust solution while still approximating the Moore–Penrose inverse.

Example 3.2 (Tikhonov regularisation)**.** The main idea behind Tikhonov regularisation[1] is to shift the singular values of K^*K by a constant factor, which will be associated with the regularisation parameter α. This shift can be realised via the function

$$g_\alpha(\sigma) = \frac{\sigma}{\sigma^2 + \alpha}\,. \tag{3.5}$$

Again, we immediately observe that for all $\sigma > 0$ we have $\lim_{\alpha\to 0} g_\alpha(\sigma) = 1/\sigma$. Further, we can estimate $g_\alpha(\sigma) \leq 1/(2\sqrt{\alpha})$ due to $\sigma^2 + \alpha \geq 2\sqrt{\alpha}\sigma$. The corresponding Tikhonov regularisation (3.1) reads as

$$R_\alpha f = \sum_{j=1}^{\infty} \frac{\sigma_j}{\sigma_j^2 + \alpha} \langle f, v_j\rangle_{\mathcal{V}}\, u_j\,. \tag{3.6}$$

After getting an intuition about regularisation of the form (3.1) via examples, we want to define what a regularisation actually is, and what properties come along with it.

Definition 3.1. *Let $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ be a bounded operator. A family $\{R_\alpha\}_{\alpha>0}$ of continuous operators is called* regularisation *(or* regularisation operator*) of $K^\dagger$ if*

$$R_\alpha f \to K^\dagger f = u^\dagger$$

for all $f \in \mathcal{D}(K^\dagger)$ as $\alpha \to 0$.

Definition 3.2. *We further call $\{R_\alpha\}_{\alpha>0}$ a* linear regularisation*, if Definition 3.1 is satisfied together with the additional assumption*

$$R_\alpha \in \mathcal{L}(\mathcal{V},\mathcal{U})\,,$$

for all $\alpha \in \mathbb{R}_{>0}$.

Hence, a regularisation is a pointwise approximation of the Moore–Penrose inverse with continuous operators, see Figure 3.1 for an illustration. As in the interesting cases the Moore–Penrose inverse may not be continuous we cannot expect that the norms of a regularisation stay bounded as $\alpha \to 0$. This is confirmed by the following results.

[1]Named after the Russian mathematician Andrey Nikolayevich Tikhonov (30 October 1906 - 7 October 1993)

Theorem 3.1 (Banach–Steinhaus e.g. [4, p. 78], [16, p. 173])**.** *Let $\mathcal{U}, \mathcal{V}$ be Hilbert spaces and $\{K_j\}_{j\in\mathbb{N}} \subset \mathcal{L}(\mathcal{U},\mathcal{V})$ a family of point-wise bounded operators, i.e. for all $u \in \mathcal{U}$ there exists a constant $C(u) > 0$ with $\sup_{j\in\mathbb{N}} \|K_j u\|_{\mathcal{V}} \leq C(u)$. Then*

$$\sup_{j\in\mathbb{N}} \|K_j\|_{\mathcal{L}(\mathcal{U},\mathcal{V})} < \infty\,.$$

Corollary 3.1 ([16, p. 174])**.** *Let $\mathcal{U}, \mathcal{V}$ be Hilbert spaces and $\{K_j\}_{j\in\mathbb{N}} \subset \mathcal{L}(\mathcal{U},\mathcal{V})$. Then the following two conditions are equivalent:*

(a) There exists $K \in \mathcal{L}(\mathcal{U},\mathcal{V})$ such that

$$Ku = \lim_{j\to\infty} K_j u \quad \text{for all } u \in \mathcal{U}.$$

(b) There is a dense subset $\mathcal{X} \subset \mathcal{U}$ such that $\lim_{j\to\infty} K_j u$ exists for all $u \in \mathcal{X}$ and

$$\sup_{j\in\mathbb{N}} \|K_j\|_{\mathcal{L}(\mathcal{U},\mathcal{V})} < \infty\,.$$

Theorem 3.2. *Let $\mathcal{U}$, $\mathcal{V}$ be Hilbert spaces, $K \in \mathcal{L}(\mathcal{U},\mathcal{V})$ and $\{R_\alpha\}_{\alpha>0}$ a liner regularisation as defined in Definition 3.2. If $K^\dagger$ is not continuous, $\{R_\alpha\}_{\alpha>0}$ cannot be uniformly bounded. In particular this implies the existence of an element $f \in \mathcal{V}$ with $\|R_\alpha f\|_{\mathcal{U}} \to \infty$ for $\alpha \to 0$.*

Proof. We prove the theorem by contradiction and assume that $\{R_\alpha\}_{\alpha>0}$ is uniformly bounded. Hence, there exists a constant C with $\|R_\alpha\|_{\mathcal{L}(\mathcal{V},\mathcal{U})} \leq C$ for all $\alpha > 0$. Due to Definition 3.1, we have $R_\alpha \to K^\dagger$ on $\mathcal{D}(K^\dagger)$. Corollary 3.1 then already implies $K^\dagger \in \mathcal{L}(\mathcal{V},\mathcal{U})$, which is a contradiction to the assumption that $K^\dagger$ is not continuous.

It remains to show the existence of the element $f \in \mathcal{V}$ with $\|R_\alpha f\|_{\mathcal{U}} \to \infty$ for $\alpha \to 0$. If such an element would not exist, we could conclude $\{R_\alpha\}_{\alpha>0} \subset \mathcal{L}(\mathcal{V},\mathcal{U})$. However, Theorem 3.1 then implies that $\{R_\alpha\}_{\alpha>0}$ has to be uniformly bounded, which contradicts the first part of the proof. □

With the additional assumption that $\|KR_\alpha\|_{\mathcal{L}(\mathcal{V},\mathcal{V})}$ is bounded, we can even show that $R_\alpha f$ diverges for all $f \notin \mathcal{D}(K^\dagger)$.

Theorem 3.3. *Let $K \in \mathcal{L}(\mathcal{U},\mathcal{V})$ and $\{R_\alpha\}_{\alpha>0}$ be a linear regularisation of $K^\dagger$, and define $u_\alpha := R_\alpha f$. If*

$$\sup_{\alpha>0} \|KR_\alpha\|_{\mathcal{L}(\mathcal{V},\mathcal{V})} < \infty\,,$$

then $\|u_\alpha\|_{\mathcal{U}} \to \infty$ for $f \notin \mathcal{D}(K^\dagger)$.

Proof. The convergence in case of $f \in \mathcal{D}(K^\dagger)$ simply follows from Definition 3.1. We therefore only need to consider the case $f \notin \mathcal{D}(K^\dagger)$. We assume that there exists a sequence $\alpha_k \to 0$ such that $\|u_{\alpha_k}\|_{\mathcal{U}}$ is uniformly bounded. Then there exists a weakly convergent subsequence $u_{\alpha_{k_l}}$ with some limit $u \in \mathcal{U}$, cf. [7, Section 2.2, Theorem 2.1]. As continuous linear operators are also weakly continuous, we further have $Ku_{\alpha_{k_l}} \rightharpoonup Ku$. However, as KR_α are uniformly bounded operators, we also conclude $Ku_{\alpha_{k_l}} = KR_{\alpha_{k_l}} f \rightharpoonup P_{\overline{\mathcal{R}(K)}} f$ for all $f \in \mathcal{V}$ (and not just $f \in \mathcal{D}(K^\dagger)$), because of Corollary 3.1. Hence, we further conclude $f \in \mathcal{R}(K)$ and therefore $f \in \mathcal{D}(K^\dagger)$ in contradiction to the assumption $f \notin \mathcal{D}(K^\dagger)$. □

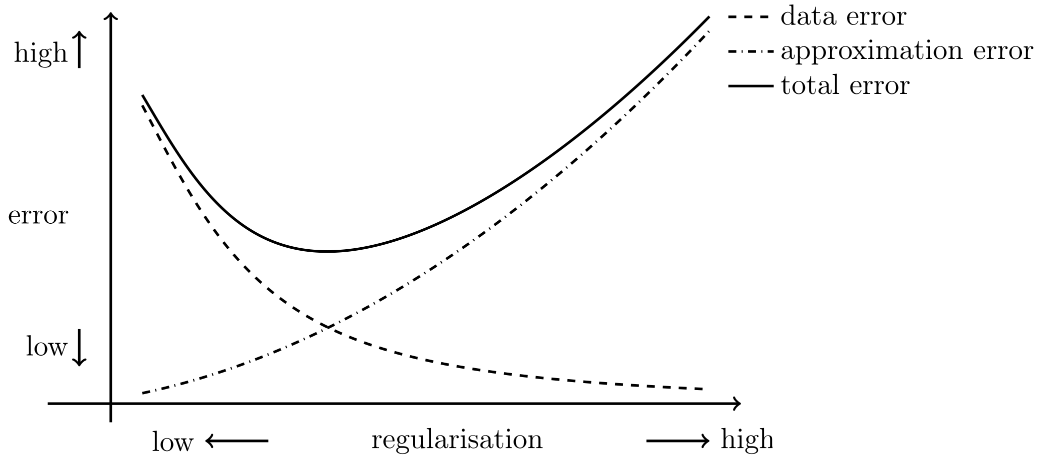

Figure 3.2: The *total error* between a regularised solution and the minimal norm solution decomposes into the *data error* and the *approximation error*. These two errors have opposing trends: For a small regularisation parameter α the error in the data gets amplified through the ill-posedness of the problem and for large α the operator R_α is a poor approximation of the Moore–Penrose inverse.

Usually we cannot expect $f \in \mathcal{D}(K^\dagger)$ for most applications, due to measurement and modelling errors. However, we assume that there exists $f \in \mathcal{D}(K^\dagger)$ such that we have

$$\left\| f - f^\delta \right\|_{\mathcal{V}} \leq \delta$$

for measured data $f^\delta \in \mathcal{V}$. For linear regularisations we can split the *total error* between the regularised solution of the noisy problem $R_\alpha f^\delta$ and the minimal norm solution of the noise-free problem $u^\dagger = K^\dagger f$ as

$$\begin{aligned} \|R_\alpha f^\delta - u^\dagger\|_{\mathcal{U}} &\leq \|R_\alpha f^\delta - R_\alpha f\|_{\mathcal{U}} + \|R_\alpha f - u^\dagger\|_{\mathcal{U}} \\ &\leq \underbrace{\delta \|R_\alpha\|_{\mathcal{L}(\mathcal{V},\mathcal{U})}}_{\text{data error}} + \underbrace{\|R_\alpha f - K^\dagger f\|_{\mathcal{U}}}_{\text{approximation error}} . \end{aligned} \tag{3.7}$$

The first term of (3.7) is the *data error*; this term unfortunately does not stay bounded for $\alpha \to 0$, which we can conclude from Theorem 3.2. The second term, known as the *approximation error*, however vanishes for $\alpha \to 0$, due to the pointwise convergence of R_α to $K^\dagger$. Hence it becomes evident from (3.7) that a good choice of α depends on δ, and needs to be chosen such that the approximation error becomes as small as possible, whilst the data error is being kept at bay. See Figure 3.2 for a visualisation of this situation. In the following we are going to discuss typical strategies for choosing α appropriately.

3.1 Parameter-choice strategies

In this section we want to discuss three standard rules for the choice of the regularisation parameter α and whether they lead to (convergent) regularisation methods.

Definition 3.3. *A function* $\alpha\colon \mathbb{R}_{>0}\times\mathcal{V}\to\mathbb{R}_{>0}$*,* $(\delta, f^\delta)\mapsto\alpha(\delta, f^\delta)$ *is called* parameter choice rule. *We distinguish between*

(a) a-priori parameter choice rules, if they depend on δ *only;*

(b) a-posteriori parameter choice rules, if they depend on δ *and* f^δ*;*

(c) heuristic parameter choice rules, if they depend on f^δ *only.*

In case of (a) or (c) we would simply write $\alpha(\delta)$, respectively $\alpha(f^\delta)$, instead of $\alpha(\delta, f^\delta)$.

Definition 3.4. *If* $\{R_\alpha\}_{\alpha>0}$ *is a regularisation of* $K^\dagger$ *and* α *is a parameter choice rule, then the pair* (R_α,α) *is called* convergent regularisation*, if for all* $f\in\mathcal{D}(K^\dagger)$ *there exists a parameter choice rule* $\alpha:\mathbb{R}_{>0}\times\mathcal{V}\to\mathbb{R}_{>0}$ *such that*

$$\lim_{\delta\to 0}\sup\left\{\left\|R_\alpha f^\delta - K^\dagger f\right\|_{\mathcal{U}}\ \middle|\ f^\delta\in\mathcal{V}, \left\|f-f^\delta\right\|_{\mathcal{V}}\le\delta\right\}=0 \tag{3.8}$$

and

$$\lim_{\delta\to 0}\sup\ \left\{\alpha(\delta, f^\delta)\ \middle|\ f^\delta\in\mathcal{V}, \left\|f-f^\delta\right\|_{\mathcal{V}}\le\delta\right\}=0 \tag{3.9}$$

are guaranteed.

3.1.1 A-priori parameter choice rules

First of all we want to discuss a-priori parameter choice rules in more detail. In fact, it can be shown that for every regularisation an a-priori parameter choice rule, and thus, a convergent regularisation, exists.

Theorem 3.4. *Let* $\{R_\alpha\}_{\alpha>0}$ *be a regularisation of* $K^\dagger$*, for* $K\in\mathcal{L}(\mathcal{U},\mathcal{V})$*. Then there exists an a-priori parameter choice rule, such that* (R_α,α) *is a convergent regularisation.*

Proof. Let $f\in\mathcal{D}(K^\dagger)$ be arbitrary but fixed. We can find a monotone increasing function $\gamma:\mathbb{R}_{>0}\to\mathbb{R}_{>0}$ with $\lim_{\varepsilon\to0}\gamma(\varepsilon)=0$ such that for every $\varepsilon>0$ we have

$$\left\|R_{\gamma(\varepsilon)}f-K^\dagger f\right\|_{\mathcal{U}}\le\frac{\varepsilon}{2},$$

due to the pointwise convergence $R_\alpha\to K^\dagger$.
As the operator $R_{\gamma(\varepsilon)}$ is continuous for fixed ε, there exists $\rho(\varepsilon)>0$ with

$$\left\|R_{\gamma(\varepsilon)}g-R_{\gamma(\varepsilon)}f\right\|_{\mathcal{U}}\le\frac{\varepsilon}{2}\ \text{for all}\ g\in\mathcal{V}\ \text{with}\ \|g-f\|_{\mathcal{V}}\le\rho(\varepsilon).$$

Without loss of generality we can assume ρ to be a continuous, strictly monotone increasing function with $\lim_{\varepsilon\to0}\rho(\varepsilon)=0$. Then, due to the inverse function theorem there exists a strictly monotone and continuous function ρ^{-1} on $\mathcal{R}(\rho)$ with $\lim_{\delta\to0}\rho^{-1}(\delta)=0$. We continuously extend ρ^{-1} on $\mathbb{R}_{>0}$ and define our a-priori strategy as

$$\alpha:\mathbb{R}_{>0}\to\mathbb{R}_{>0},\quad \delta\to\gamma(\rho^{-1}(\delta)).$$

Then $\lim_{\delta\to0}\alpha(\delta)=0$ follows. Furthermore, there exists $\delta:=\rho(\varepsilon)$ for all $\varepsilon>0$, such that with $\alpha(\delta)=\gamma(\varepsilon)$

$$\left\|R_{\alpha(\delta)}f^\delta-K^\dagger f\right\|_{\mathcal{U}}\le\left\|R_{\gamma(\varepsilon)}f^\delta-R_{\gamma(\varepsilon)}f\right\|_{\mathcal{U}}+\left\|R_{\gamma(\varepsilon)}f-K^\dagger f\right\|_{\mathcal{U}}\le\varepsilon$$

follows for all $f^\delta\in\mathcal{V}$ with $\|f-f^\delta\|_{\mathcal{V}}\le\delta$. Thus, (R_α,α) is a convergent regularisation method. ☐

For linear regularisations we can characterise a-priori parameter choice strategies that lead to convergent regularisation methods via the following theorem.

Theorem 3.5. *Let $\{R_\alpha\}_{\alpha>0}$ be a linear regularisation, and $\alpha : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ an a-priori parameter choice rule. Then (R_α, α) is a convergent regularisation method if and only if*

(a) $\lim_{\delta\to 0} \alpha(\delta) = 0$

(b) $\lim_{\delta\to 0} \delta \|R_{\alpha(\delta)}\|_{\mathcal{L}(\mathcal{V},\mathcal{U})} = 0$

Proof. $\Leftarrow$: Let condition a) and b) be fulfilled. From (3.7) we then observe

$$\left\| R_{\alpha(\delta)} f^\delta - K^\dagger f \right\|_{\mathcal{U}} \to 0 \text{ for } \delta \to 0.$$

Hence, (R_α, α) is a convergent regularisation method.
$\Rightarrow$: Now let (R_α, α) be a convergent regularisation method. We prove that conditions 1 and 2 have to follow from this by showing that violation of either one of them leads to a contradiction to (R_α, α) being a convergent regularisation method. If condition a) is violated, (3.9) is violated and hence, (R_α, α) is not a convergent regularisation method. If condition a) is fulfilled but condition b) is violated, there exists a null sequence $\{\delta_k\}_{k\in\mathbb{N}}$ with $\delta_k \|R_{\alpha(\delta_k)}\|_{\mathcal{L}(\mathcal{V},\mathcal{U})} \geq C > 0$, and hence, we can find a sequence $\{g_k\}_{k\in\mathbb{N}} \subset \mathcal{V}$ with $\|g_k\|_{\mathcal{V}} = 1$ and $\delta_k \|R_{\alpha(\delta_k)} g_k\|_{\mathcal{U}} \geq \tilde{C}$ for some $\tilde{C}$. Let $f \in \mathcal{D}(K^\dagger)$ be arbitrary and define $f_k := f + \delta_k g_k$. Then we have on the one hand $\|f - f_k\|_{\mathcal{V}} \leq \delta_k$, but on the other hand the norm of

$$R_{\alpha(\delta_k)} f_k - K^\dagger f = R_{\alpha(\delta_k)} f - K^\dagger f + \delta_k R_{\alpha(\delta_k)} g_k$$

cannot converge to zero, as the second term $\delta_k R_{\alpha(\delta_k)} g_k$ is bounded from below by construction. Hence, (3.8) is violated for $f^\delta = g_k$ and thus, (R_α, α) is not a convergent regularisation method. □

3.1.2 A-posteriori parameter choice rules

In the following sections we are going to see that Theorem 3.5 basically means that $\alpha(\delta)$ cannot converge too quickly to zero in relation to δ; typical parameter choice strategies will be of the form $\alpha(\delta) = \delta^p$. However, finding an optimal choice of p often requires additional information about $u^\dagger$, for instance in terms of source conditions that we are going to discuss in Section 3.2.4. A-posteriori parameter choice rules have the advantage that they do not require this additional information. The basic idea is as follows. We again have $f \in \mathcal{D}(K^\dagger)$ and f^δ with $\|f - f^\delta\|_{\mathcal{V}} \leq \delta$, and now consider the *residual* between f^δ and $u_\alpha := R_\alpha f^\delta$, i.e.

$$\|K u_\alpha - f^\delta\|_{\mathcal{V}} .$$

If we assume that $u^\dagger$ is the minimal norm solution and f is given via $f = K u^\dagger$, we immediately observe that $u^\dagger$ satisfies

$$\|K u^\dagger - f^\delta\|_{\mathcal{V}} = \|f - f^\delta\|_{\mathcal{V}} = \delta .$$

Hence, it appears not to be useful to choose $\alpha(\delta, f^\delta)$ with $\|K u_{\alpha(\delta, f^\delta)} - f^\delta\|_{\mathcal{V}} < \delta$, which motivates Morozov's discrepancy principle.

Definition 3.5 (Morozov's discrepancy principle). *Let $\alpha(\delta, f^\delta)$ be chosen such that*

$$\|Ku_{\alpha(\delta,f^\delta)} - f^\delta\|_{\mathcal{V}} \leq \eta\delta \tag{3.10}$$

is satisfied, for given δ, f^δ, and a fixed constant $\eta > 1$. Then $u_{\alpha(\delta,f^\delta)} = R_{\alpha(\delta,f^\delta)} f^\delta$ is said to satisfy Morozov's discrepancy principle.

Remark 3.1. It is important to point out that (3.10) may never be fulfilled, as is the case for $f \in \mathcal{R}(K)^\perp$. Following Lemma 2.3 (d), even for exact data $f^\delta = f$ we observe

$$\|Ku^\dagger - f\|_{\mathcal{V}} = \|KK^\dagger f - f\|_{\mathcal{V}} = \|P_{\overline{\mathcal{R}(K)}} f - f\|_{\mathcal{V}} = \|f\|_{\mathcal{V}} > \delta$$

in this case, for δ being small enough. In order to avoid this scenario, we ideally ensure that $\mathcal{R}(K)$ is dense in $\mathcal{V}$, as this already implies $\mathcal{R}(K)^\perp = \{0\}$ due to Remark 2.1.

Practical a-posteriori regularisation strategies are usually designed as follows. We pick a null sequence $\{\alpha_j\}_{j\in\mathbb{N}}$ and iteratively compute $u_{\alpha_j} = R_{\alpha_j} f^\delta$ for $j \in \{1, \ldots, j^*\}$, $j^* \in \mathbb{N}$, until $u_{\alpha_{j^*}}$ satisfies (3.10). This procedure is justified by the following theorem.

Theorem 3.6. *Let $\{R_\alpha\}_{\alpha>0}$ be a regularisation of $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, and let $\mathcal{R}(K)$ be dense in $\mathcal{V}$. Further, let $\{\alpha_j\}_{j\in\mathbb{N}}$ be a strictly monotonically decreasing null sequence, and let $\eta > 1$. If the family of operators $\{KR_\alpha\}_{\alpha>0}$ is uniformly bounded, there exists a finite index $j^* \in \mathbb{N}$ such that for all $f \in \mathcal{D}(K^\dagger)$ and f^δ with $\|f - f^\delta\|_{\mathcal{V}} \leq \delta$ the inequalities*

$$\|Ku_{\alpha_{j^*}} - f^\delta\|_{\mathcal{V}} \leq \mu\delta < \|Ku_{\alpha_j} - f^\delta\|_{\mathcal{V}}$$

are satisfied for all $j < j^*$.

Proof. We know that KR_α converges pointwise to $KK^\dagger = P_{\overline{\mathcal{R}(K)}}$ in $\mathcal{D}(K^\dagger)$, which together with the uniform boundedness assumption already implies pointwise convergence in $\mathcal{V}$, as we have already shown in the proof of Theorem 3.2. Hence, for all $f \in \mathcal{D}(K^\dagger)$ and $f^\delta \in \mathcal{V}$ with $\|f - f^\delta\|_{\mathcal{V}} \leq \delta$ we can conclude

$$\begin{aligned}\lim_{j\to\infty} \|Ku_{\alpha_j} - f^\delta\|_{\mathcal{V}} &= \lim_{j\to\infty} \|KR_{\alpha_j} f^\delta - f^\delta\|_{\mathcal{V}} = \left\| P_{\overline{\mathcal{R}(K)}} f^\delta - f^\delta \right\|_{\mathcal{V}} \\ &= \inf_{g\in\overline{\mathcal{R}(K)}} \|g - f^\delta\|_{\mathcal{V}} \leq \|f - f^\delta\|_{\mathcal{V}} \leq \delta\,.\end{aligned}$$

□

We are going to demonstrate later that (3.10) in combination with specific regularisations is indeed a regularisation method. Before we do so, we want to conclude the discussion of parameter choice strategies by investigating heuristic regularisation methods.

3.1.3 Heuristic parameter choice rules

Heuristic parameter choice rules do not require knowledge of the noise level δ, which makes them popular strategies in practice. In the following we give three examples of popular heuristic parameter choice rules.

Quasi-optimality principle For the first n elements of a null sequence, i.e. $\{\alpha_j\}_{j\in\{1,\ldots,n\}}$, we choose $\alpha(f^\delta) = \alpha_{j^*}$ with

$$\alpha_{j^*} = \operatorname*{arg\,min}_{1\leq j<n} \|u_{\alpha_{j+1}} - u_{\alpha_j}\|_{\mathcal{U}}\,.$$

Hanke-Raus rule The parameter $\alpha(f^\delta)$ is chosen via

$$\alpha(f^\delta) = \operatorname*{arg\,min}_{\alpha>0} \frac{1}{\sqrt{\alpha}} \|Ku_\alpha - f^\delta\|_{\mathcal{V}} \,.$$

L-curve method The parameter $\alpha(f^\delta)$ is chosen via

$$\alpha(f^\delta) = \operatorname*{arg\,min}_{\alpha>0} \|u_\alpha\|_{\mathcal{U}} \|Ku_\alpha - f^\delta\|_{\mathcal{V}} \,.$$

Despite their popularity and the fact that they do not require any knowledge about δ, heuristic parameter choice rules have one significant theoretical disadvantage. While any regularisation can be equipped with an a-priori parameter choice rule to form a convergent regularisation as seen in Theorem 3.4, heuristic parameter choice rules cannot lead to convergent regularisations, a result that has become famous as the so-called Bakushinskiĭ veto [2].

Theorem 3.7. *Let $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ with $\mathcal{R}(K) \neq \overline{\mathcal{R}(K)}$. Then for any regularisation $\{R_\alpha\}_{\alpha>0}$ and any heuristic parameter choice rule $\alpha(f^\delta)$ the pair $(\{R_\alpha\}, \alpha)$ is not a convergent regularisation.*

Proof. Assume that $(\{R_\alpha\}, \alpha)$ is a convergent regularisation method and that the parameter choice rule is heurstic, i.e. $\alpha = \alpha(f^\delta)$. Then it follows from (3.8) that

$$\lim_{\delta\to 0} \sup \left\{ \left\| R_{\alpha(f^\delta)} f^\delta - K^\dagger f \right\|_{\mathcal{U}} \;\middle|\; f^\delta \in \mathcal{V}, \left\| f - f^\delta \right\|_{\mathcal{V}} \leq \delta \right\} = 0$$

and in particular $R_{\alpha(f)} f = K^\dagger f$ for all $f \in \mathcal{D}(K^\dagger)$. Thus, for any sequence $\{f_j\}_{j\in\mathbb{N}} \subset \mathcal{D}(K^\dagger)$ which converges to $f \in \mathcal{D}(K^\dagger)$ we have that

$$\lim_{j\to\infty} K^\dagger f_j = \lim_{j\to\infty} R_{\alpha(f_j)} f_j = K^\dagger f$$

which shows that $K^\dagger$ is continuous. It follows from Theorem 2.5 that the range of K is closed, which contradicts the assumption. □

Remark 3.2. We want to point out that Theorem 3.7 does not automatically make any heuristic parameter choice rule useless, for two reasons. Firstly, because Theorem 3.7 applies to infinite dimensional problems. Hence, discretised, ill-conditioned problems can still benefit from heuristic parameter choice rules. Secondly, the proof of Theorem 3.7 explicitly uses perturbed data $f_j \in \mathcal{D}(K^\dagger)$ to show the contradiction. For actual perturbed data f^δ however, it is quite unusual that they will satisfy $f^\delta \in \mathcal{D}(K^\dagger)$. It can indeed be shown that, under the additional assumption $f^\delta \notin \mathcal{D}(K^\dagger)$, a lot of regularisation strategies together with a whole class of heuristic parameter choice strategies can be turned into convergent regularisations.

3.2 Spectral regularisation methods

Now we revisit (3.1) and finally prove that these methods are regularisation methods for piecewise continuous functions g_α satisfying (3.2).

Theorem 3.8. *Let $g_\alpha : \mathbb{R}_{>0} \to \mathbb{R}$ be a piecewise continuous function satisfying* (3.2), $\lim_{\alpha\to 0} g_\alpha(\sigma) = \frac{1}{\sigma}$ *and*

$$\sup_{\alpha,\sigma} \sigma g_\alpha(\sigma) \leq \gamma \,, \tag{3.11}$$

for some constant $\gamma > 0$. If R_α is defined as in (3.1), we have

$$R_\alpha f \to K^\dagger f \text{ as } \alpha \to 0,$$

for all $f \in \mathcal{D}(K^\dagger)$.

Proof. From the singular value decomposition of $K^\dagger$ and the definition of R_α we obtain

$$R_\alpha f - K^\dagger f = \sum_{j=1}^\infty \left(g_\alpha(\sigma_j) - \frac{1}{\sigma_j} \right) \langle f, v_j \rangle_{\mathcal{V}} \, u_j = \sum_{j=1}^\infty \left(\sigma_j g_\alpha(\sigma_j) - 1 \right) \langle u^\dagger, u_j \rangle_{\mathcal{U}} \, u_j \, .$$

From (3.11) we can conclude

$$\left| \left(\sigma_j g_\alpha(\sigma_j) - 1 \right) \langle u^\dagger, u_j \rangle_{\mathcal{U}} \right| \leq (1 + \gamma) \| u^\dagger \|_{\mathcal{U}} \, ,$$

and hence, each element of the sum stays bounded. Thus, we can also estimate

$$\begin{aligned} \| R_\alpha f - K^\dagger f \|_{\mathcal{U}}^2 &= \sum_{j=1}^\infty | \sigma_j g_\alpha(\sigma_j) - 1 |^2 \left| \langle u^\dagger, u_j \rangle_{\mathcal{U}} \right|^2 \leq (1 + \gamma)^2 \sum_{j=1}^\infty \left| \langle u^\dagger, u_j \rangle_{\mathcal{U}} \right|^2 \\ &= (1 + \gamma)^2 \| u^\dagger \|_{\mathcal{U}}^2 < \infty \end{aligned}$$

and conclude that $\| R_\alpha f - K^\dagger f \|_{\mathcal{U}}$ is bounded from above. This allows the application of the reverse Fatou lemma, which yields the estimate

$$\begin{aligned} \limsup_{\alpha \to 0} \left\| R_\alpha f - K^\dagger f \right\|_{\mathcal{U}}^2 &\leq \limsup_{\alpha \to 0} \sum_{j=1}^\infty | \sigma_j g_\alpha(\sigma_j) - 1 |^2 \left| \langle u^\dagger, u_j \rangle_{\mathcal{U}} \right|^2 \\ &\leq \sum_{j=1}^\infty \left| \lim_{\alpha \to 0} \sigma_j g_\alpha(\sigma_j) - 1 \right|^2 \left| \langle u^\dagger, u_j \rangle_{\mathcal{U}} \right|^2 . \end{aligned}$$

Due to the pointwise convergence of $g_\alpha(\sigma_j)$ to $1/\sigma_j$ we obtain $\lim_{\alpha \to 0} \sigma_j g_\alpha(\sigma_j) - 1 = 0$. Hence, we have $\left\| R_\alpha f - K^\dagger f \right\|_{\mathcal{U}} \to 0$ for $\alpha \to 0$ for all $f \in \mathcal{D}(K^\dagger)$. ☐

Proposition 3.1. *Let the same assumptions hold as in Theorem 3.8. Further, let α be an a-priori parameter choice rule. Then $(R_{\alpha(\delta)}, \alpha(\delta))$ is a convergent regularisation method if*

$$\lim_{\delta \to 0} \delta C_{\alpha(\delta)} = 0$$

is guaranteed.

Proof. The result follows immediately from $\| R_{\alpha(\delta)} \|_{\mathcal{L}(\mathcal{V}, \mathcal{U})} \leq C_{\alpha(\delta)}$ and Theorem 3.5. ☐

3.2.1 Convergence rates

Knowing that spectral regularisation methods of the form (3.1) together with (3.2) represent convergent regularisation methods, we now want to understand how the error in the data propagates to the error in the reconstruction.

Theorem 3.9. *Let the same assumptions hold for g_α as in Theorem 3.8. If we define $u_\alpha := R_\alpha f$ and $u_\alpha^\delta := R_\alpha f^\delta$, with $f \in \mathcal{D}(K^\dagger)$, $f^\delta \in \mathcal{V}$ and $\|f - f^\delta\|_\mathcal{V} \leq \delta$, then*

$$\|Ku_\alpha - Ku_\alpha^\delta\|_\mathcal{V} \leq \gamma\delta\,, \tag{3.12}$$

and

$$\|u_\alpha - u_\alpha^\delta\|_\mathcal{U} \leq C_\alpha\delta \tag{3.13}$$

hold true.

Proof. From the singular value decomposition we can estimate

$$\begin{aligned}\|Ku_\alpha - Ku_\alpha^\delta\|_\mathcal{V}^2 &\leq \sum_{j=1}^\infty \sigma_j^2 g_\alpha(\sigma_j)^2 |\langle f - f^\delta, v_j\rangle_\mathcal{V}|^2 \\ &\leq \gamma^2 \sum_{j=1}^\infty |\langle f - f^\delta, v_j\rangle_\mathcal{V}|^2 = \gamma^2\|f - f^\delta\|_\mathcal{V}^2 \leq \gamma^2\delta^2\,,\end{aligned}$$

which yields (3.12). In the same fashion we can estimate

$$\begin{aligned}\|u_\alpha - u_\alpha^\delta\|_\mathcal{U}^2 &\leq \sum_{j=1}^\infty g_\alpha(\sigma_j)^2 |\langle f - f^\delta, v_j\rangle_\mathcal{V}|^2 \\ &\leq C_\alpha^2 \sum_{j=1}^\infty |\langle f - f^\delta, v_j\rangle_\mathcal{V}|^2 = C_\alpha^2\|f - f^\delta\|_\mathcal{V}^2 \leq C_\alpha^2\delta^2\,,\end{aligned}$$

to obtain (3.13). □

Remark 3.3. At first glance (3.13) gives the impression as if the error in the reconstruction is also of order δ. This, however, is not the case, as C_α also depends on δ, as we have seen in Proposition 3.1. The condition $\lim_{\delta\to 0}\delta C_\alpha = 0$ will in particular force C_α to decay more quickly than δ. Hence, $C_\alpha\delta$ will be of order δ^ν, with $0 < \nu < 1$.

Combining the assertions of Theorem 3.8, Proposition 3.1 and Theorem 3.9, we obtain the following convergence results of the regularised solutions.

Proposition 3.2. *Let the assumptions of Theorem 3.8, Proposition 3.1 and Theorem 3.9 hold true. Then,*

$$u_{\alpha(\delta)} \to u^\dagger$$

is guaranteed as $\delta \to 0$.

3.2.2 Truncated singular value decomposition

As a first example for a spectral regularisation of the form (3.1) we have considered the so-called truncated singular value decomposition in Example 3.1. From (3.3) we immediately observe $g_\alpha(\sigma) \leq C_\alpha = 1/\alpha$. Thus, according to Proposition 3.1 the truncated singular value decomposition, together with an a-priori parameter choice strategy satisfying $\lim_{\delta\to 0}\alpha(\delta) = 0$, is a convergent regularisation method if $\lim_{\delta\to 0}\delta/\alpha(\delta) = 0$.

Moreover, we observe $\sup_{\sigma,\alpha} \sigma g_\alpha(\sigma) = \gamma = 1$ and hence, we obtain the error estimates $\|Ku_\alpha - Ku_\alpha^\delta\|_\mathcal{V} \leq \delta$ and $\|u_\alpha - u_\alpha^\delta\|_\mathcal{U} \leq \delta/\alpha(\delta)$ as a consequence of Theorem 3.9.

Let $K \in \mathcal{K}(\mathcal{U}, \mathcal{V})$ with singular system $\{\sigma_j, u_j, v_j\}_{j\in\mathbb{N}}$, and choose for $\delta > 0$ an index function $j^* : \mathbb{R}_{>0} \to \mathbb{N}$ with $j^*(\delta) \to \infty$ for $\delta \to 0$ and $\lim_{\delta\to 0} \delta/\sigma_{j^*(\delta)} = 0$. We can then choose $\alpha(\delta) = \sigma_{j^*(\delta)}$ as our a-priori parameter choice rule to obtain a convergent regularisation.

Note that in practice a larger δ implies that more and more singular values have to be cut off in order to guarantee a stable recovery that successfully suppresses the data error.

3.2.3 Tikhonov regularisation

The second example we were considering was Tikhonov regularisation in Example 3.2, where we have shifted the singular values of K^*K by a constant factor, which will be associated with the regularisation parameter α.

In case of g_α as defined in (3.5) we observe $\lim_{\alpha\to 0} g_\alpha(\sigma) = 1/\sigma$ for $\sigma > 0$. Further, we can estimate $g_\alpha(\sigma) \leq 1/(2\sqrt{\alpha}) = C_\alpha$ due to $\sigma^2 + \alpha \geq 2\sqrt{\alpha}\sigma$. Moreover, we discover $\sigma g_\alpha(\sigma) = \sigma^2/(\sigma^2+\alpha) < 1 =: \gamma$ for $\alpha > 0$. Consequently, we have to ensure $\delta/(2\sqrt{\alpha(\delta)}) \to 0$ for $\delta \to 0$ to obtain a convergent regularisation, and in that case get the estimates $\|Ku_\alpha - Ku_\alpha^\delta\|_\mathcal{V} \leq \delta$ and $\|u_\alpha - u_\alpha^\delta\|_\mathcal{U} \leq \delta/(2\sqrt{\alpha(\delta)})$. Thus, equipping $R_{\alpha(\delta)}$ for instance with the a-priori parameter choice rule $\alpha(\delta) = \delta/4$ will lead to a convergent regularisation for which we have $\|u_\alpha - u_\alpha^\delta\|_\mathcal{U} = \mathcal{O}(\sqrt{\delta})$.

Note that Tikhonov regularisation can be computed without knowledge of the singular system. Considering the equation $(K^*K + \alpha I)u_\alpha$ in terms of the singular value decomposition, we observe

$$\sum_{j=1}^\infty \frac{\sigma_j}{\sigma_j^2 + \alpha} \langle f, v_j\rangle_\mathcal{V} \underbrace{K^* \underbrace{K u_j}_{=\sigma_j v_j}}_{=\sigma_j^2 u_j} + \sum_{j=1}^\infty \frac{\alpha\sigma_j}{\sigma_j^2+\alpha}\langle f, v_j\rangle_\mathcal{V}\, u_j$$

$$= \sum_{j=1}^\infty \frac{\sigma_j(\sigma_j^2+\alpha)}{\sigma_j^2+\alpha}\langle f, v_j\rangle_\mathcal{V}\, u_j = \sum_{j=1}^\infty \sigma_j \langle f, v_j\rangle_\mathcal{V}\, u_j = K^* f\,.$$

Hence, the Tikhonov-regularised solution u_α can be obtained by solving

$$(K^*K + \alpha I)u_\alpha = K^* f \tag{3.14}$$

for u_α. The advantage in computing u_α via (3.14) is that its computation does not require the singular value decomposition of K, but only involves the inversion of a linear, well-posed operator equation with a symmetric, positive definite operator.

3.2.4 Source-conditions

Before we continue to investigate other examples of regularisations we want to briefly address the question of the convergence speed of a regularisation method. From Theorem 3.9 we have already obtained a convergence rate result; however, with additional regularity assumptions on the (unknown) minimal norm solution we are able to improve those. The regularity assumptions that we want to consider are known as *source conditions*, and are of the form

$$\exists w \in \mathcal{U} \;:\; u^\dagger = (K^*K)^\mu w\,. \tag{3.15}$$

The power $\mu > 0$ of the operator is understood in the sense of the consider the μ-th power of the singular values of the operator K^*K, i.e.

$$(K^*K)^\mu w = \sum_{j=1}^{\infty} \sigma_j^{2\mu} \langle w, u_j \rangle_{\mathcal{U}} u_j \,.$$

Example 3.3 (Differentiation)**.** We want to take a look at what (3.15) actually means in the case of a specific example. We therefore again consider the inverse problem of differentiation, i.e.

$$(Ku)(y) = \int_0^y u(x)\, dx \,.$$

In case of $\mu = 1$ (3.15) reads as

$$u^\dagger(x) = \int_x^1 \int_0^y w(z)\, dz\, dy \,.$$

due to (2.10). Hence, (3.15) does simply imply that $u^\dagger$ has to be twice weakly differentiable. It becomes even more obvious if we look at twice differentiable $u^\dagger$. In that case applying the Leibniz differentiation rule for parameter integrals leaves us with

$$(u^\dagger)''(x) = -w(x) \,.$$

Hence, any twice differentiable $u^\dagger$ automatically satisfies the source condition (3.15) for $\mu = 1$. Similar results follow for different choices of $\mu \in \mathbb{N}$.

The rate of convergence of a regularisation scheme to the minimal norm solution now depends on the specific choice of g_α. We assume that g_α satisfies

$$\sigma^{2\mu} |\sigma g_\alpha(\sigma) - 1| \leq \omega_\mu(\alpha) \,,$$

for all $\sigma > 0$. In case of the truncated singular value decomposition we would for instance have $\omega_\mu(\alpha) = \alpha^{2\mu}$. With this additional assumption, we can improve the estimate in Theorem 3.8 as follows:

$$\begin{aligned}
\|R_\alpha f - K^\dagger f\|_{\mathcal{V}}^2 &\leq \sum_{j=1}^{\infty} |\sigma_j g_\alpha(\sigma_j) - 1|^2 |\langle u^\dagger, u_j \rangle_{\mathcal{U}}|^2 \\
&= \sum_{j=1}^{\infty} |\sigma_j g_\alpha(\sigma_j) - 1|^2 \sigma_j^{4\mu} |\langle w, u_j \rangle_{\mathcal{U}}|^2 \\
&\leq \omega_\mu(\alpha)^2 \|w\|_{\mathcal{U}}^2
\end{aligned}$$

Hence, we have obtained the estimate

$$\|u_\alpha - u^\dagger\|_{\mathcal{U}} \leq \omega_\mu(\alpha) \|w\|_{\mathcal{U}} \,.$$

Together with (3.7) we can further estimate

$$\|u_{\alpha(\delta)} - u^\dagger\|_{\mathcal{U}} \leq \omega_\mu(\alpha) \|w\|_{\mathcal{U}} + C_\alpha \delta \,. \tag{3.16}$$

Example 3.4. In case of the truncated singular value decomposition we know from Section 3.2.2 that $C_\alpha = 1/\alpha$, and we can further conclude $\omega_\mu(\alpha) = \alpha^{2\mu}$. Hence, (3.16) simplifies to

$$\|u_{\alpha(\delta)} - u^\dagger\|_{\mathcal{U}} \leq \alpha^{2\mu}\|w\|_{\mathcal{U}} + \delta\alpha^{-1} \tag{3.17}$$

in this case. In order to make the right-hand-side of (3.17) as small as possible, we have to choose α such that

$$\alpha = \left(\frac{\delta}{2\mu\|w\|_{\mathcal{U}}}\right)^{\frac{1}{2\mu+1}} .$$

With this choice of α we estimate

$$\begin{aligned}\|u_{\alpha(\delta)} - u^\dagger\|_{\mathcal{U}} &\leq \underbrace{2^{\frac{1-2\mu}{1+2\mu}}}_{\leq 2} \underbrace{\mu^{\frac{1-2\mu}{1+2\mu}}}_{\leq 1} \delta^{\frac{2\mu}{2\mu+1}} \|w\|_{\mathcal{U}}^{\frac{1}{2\mu+1}} \\ &\leq 2\delta^{\frac{2\mu}{2\mu+1}} \|w\|_{\mathcal{U}}^{\frac{1}{2\mu+1}} .\end{aligned}$$

It is important to note that no matter how large μ is, the rate of convergence $\delta^{\frac{2\mu}{2\mu+1}}$ will always be slower than δ, due to the ill-posedness of the inversion of K.

3.2.5 Asymptotic regularisation

Another form of regularisation is asymptotic regularisation of the form

$$\begin{aligned}\partial_t u(t) &= K^*\left(f - Ku(t)\right) \\ u(0) &= 0\end{aligned} . \tag{3.18}$$

As the linear operator K does not change with respect to the time t, we can make the Ansatz of writing $u(t)$ in terms of the singular value decomposition of K as

$$u(t) = \sum_{j=1}^\infty \gamma_j(t) u_j \, , \tag{3.19}$$

for some function $\gamma : \mathbb{R} \to \mathbb{R}$. From the initial conditions we immediately observe $\gamma(0) = 0$. From the singular value decomposition and (3.18) we further see

$$\sum_{j=1}^\infty \gamma_j'(t) u_j = \sum_{j=1}^\infty \sigma_j \left(\langle f, v_j\rangle_{\mathcal{V}} - \sigma_j \gamma(t) \underbrace{\langle u_j, u_j\rangle_{\mathcal{U}}}_{=\|u_j\|_{\mathcal{U}}^2=1} \right) u_j \, .$$

Hence, by equating the coefficients we get

$$\gamma_j'(t) = \sigma_j \langle f, v_j\rangle_{\mathcal{V}} - \sigma_j^2 \gamma_j(t) \, ,$$

and together with $\gamma_j(0)$ we obtain

$$\gamma_j(t) = \left(1 - e^{-\sigma_j^2 t}\right) \frac{1}{\sigma_j} \langle f, v_j\rangle_{\mathcal{V}}$$

as a solution for all j and hence, (3.19) reads as

$$u(t) = \sum_{j=1}^{\infty} \left(1 - e^{-\sigma_j^2 t}\right) \frac{1}{\sigma_j} \langle f, v_j \rangle_{\mathcal{V}} u_j \, .$$

If we substitute $t = 1/\alpha$, we obtain the regularisation

$$u_\alpha = \sum_{j=1}^{\infty} \left(1 - e^{-\frac{\sigma_j^2}{\alpha}}\right) \frac{1}{\sigma_j} \langle f, v_j \rangle_{\mathcal{V}} u_j$$

with $g_\alpha(\sigma) = \left(1 - e^{-\frac{\sigma^2}{\alpha}}\right) \frac{1}{\sigma}$. We immediately see that $g_\alpha(\sigma)\sigma \leq 1 =: \gamma$, and due to $e^x \geq 1 + x$ we further observe $1 - e^{-\frac{\sigma^2}{\alpha}} \leq \sigma^2/\alpha$ and therefore $(1 - e^{-\frac{\sigma^2}{\alpha}})/\sigma \leq \max_j \sigma_j/\alpha = \sigma_1/\alpha = \|K\|_{\mathcal{L}(\mathcal{U},\mathcal{V})}/\alpha =: C_\alpha$.

3.2.6 Landweber iteration

If we approximate (3.18) via a forward finite-difference discretisation, we end up with the iterative procedure

$$\begin{aligned} \frac{u^{k+1} - u^k}{\tau} &= K^* \left(f - K u^k\right) \, , && (3.20) \\ \Leftrightarrow \qquad u^{k+1} &= u^k + \tau K^* \left(f - K u^k\right) \, , \\ \Leftrightarrow \qquad u^{k+1} &= (I - \tau K^* K) u^k + \tau K^* f \, , \end{aligned}$$

for some $\tau > 0$ and $u^0 \equiv 0$. Iteration (3.20) is known as the so-called Landweber iteration. We assume $f \in \mathcal{D}(K^\dagger)$ first, and with the singular value decomposition of K and K^* we obtain

$$\sum_{j=1}^{\infty} \langle u^{k+1}, u_j \rangle_{\mathcal{U}} u_j = \sum_{j=1}^{\infty} \left(\left(1 - \tau\sigma_j^2\right) \langle u^k, u_j \rangle_{\mathcal{U}} + \tau\sigma_j \langle f, v_j \rangle_{\mathcal{V}} \right) u_j \, , \tag{3.21}$$

and hence, by equating the individual summands

$$\langle u^{k+1}, u_j \rangle_{\mathcal{U}} = \left(1 - \tau\sigma_j^2\right) \langle u^k, u_j \rangle_{\mathcal{U}} + \tau\sigma_j \langle f, v_j \rangle_{\mathcal{V}} \, . \tag{3.22}$$

Assuming $u^0 \equiv 0$, summing up equation (3.22) yields

$$\langle u^k, u_j \rangle_{\mathcal{U}} = \tau\sigma_j \langle f, v_j \rangle_{\mathcal{V}} \sum_{i=1}^{k} (1 - \tau\sigma_j^2)^{k-i} \, . \tag{3.23}$$

The following Lemma will help us simplifying (3.23).

Lemma 3.1. *For $k \in \mathbb{N} \setminus \{1\}$ we have*

$$\sum_{i=1}^{k} (1 - \tau\sigma^2)^{k-i} = \frac{1 - \left(1 - \tau\sigma^2\right)^k}{\tau\sigma^2} \, . \tag{3.24}$$

Proof. Equation (3.24) can simply be verified via induction. We immediately see that

$$\sum_{i=1}^{2}(1-\tau\sigma^2)^{2-i} = 1 + (1-\tau\sigma^2) = \frac{1-(1-2\tau\sigma^2+\tau^2\sigma^4)}{\tau\sigma^2} = \frac{1-\left(1-\tau\sigma^2\right)^2}{\tau\sigma^2}$$

serves as as our induction base. Considering $k \to k+1$, we observe

$$\begin{aligned}
\sum_{i=1}^{k+1}(1-\tau\sigma^2)^{k+1-i} &= 1 + \sum_{i=1}^{k}(1-\tau\sigma^2)^{k+1-i} \\
&= 1 + (1-\tau\sigma^2)\sum_{i=1}^{k}(1-\tau\sigma^2)^{k-i} \\
&= 1 + (1-\tau\sigma^2)\frac{1-\left(1-\tau\sigma^2\right)^k}{\tau\sigma^2} \\
&= \frac{1-\left(1-\tau\sigma^2\right)^{k+1}}{\tau\sigma^2}\,,
\end{aligned}$$

and we are done. □

If we now insert (3.24) into (3.23) we therefore obtain

$$\langle u^k, u_j\rangle_{\mathcal{U}} = \left(1-(1-\tau\sigma_j^2)^k\right)\frac{1}{\sigma_j}\langle f, v_j\rangle_{\mathcal{V}}. \tag{3.25}$$

The important consequence of Equation (3.25) is that we now immediately see that $\langle u^k, u_j\rangle_{\mathcal{U}} \to \langle u^\dagger, u_j\rangle_{\mathcal{U}}$ if we ensure $(1-\tau\sigma_j^2)^k \to 0$. In other words, we need to choose τ such that $|1-\tau\sigma_j^2| < 1$ (respectively $0 < \tau\sigma_j < 2$) for all j. As in the case of asymptotic regularisation we exploit that $\sigma_1 = \|K\|_{\mathcal{L}(\mathcal{U},\mathcal{V})} > \sigma_j$ for all j and select τ such that

$$0 < \tau < \frac{2}{\|K\|^2_{\mathcal{L}(\mathcal{U},\mathcal{V})}} \tag{3.26}$$

is satisfied. If we interpret the iteration number as the regularisation parameter $\alpha := 1/k$, we obtain the regularisation method

$$u_\alpha = R_\alpha f = \sum_{j=1}^{\infty}\left(1-\left(1-\tau\sigma_j^2\right)^{\frac{1}{\alpha}}\right)\frac{1}{\sigma_j}\langle f, v_j\rangle_{\mathcal{V}}$$

with $g_\alpha(\sigma) = \left(1-(1-\tau\sigma^2)^{\frac{1}{\alpha}}\right)/\sigma$.

Landweber Iteration & the discrepancy principle

To conclude this section on the Landweber iteration we want to prove convergence rates given $u^\dagger$ satisfies a source condition. We further want to demonstrate that Landweber iteration in combination with the a-posteriori parameter choice rule defined in Definition 3.5 is a sensible strategy that ensures $u^k \to u^\dagger$ as long as the discrepancy principle is violated. Following the introduction of the source condition in Section 3.2.4, we want to assume a source condition similar (3.15) for $\mu = 1/2$, i.e. there exists a $w \in \mathcal{V}$ such that

$$u^\dagger = K^* w \tag{3.27}$$

is satisfied. Under that additional assumption we can conclude the following convergence rate in the case of noise-free data $f^\delta = f$.

Lemma 3.2. *Let* (3.27) *be satisfied. Then the Landweber iterates* (3.20) *satisfy*

$$\|u^k - u^\dagger\|_{\mathcal{U}} = \mathcal{O}\left(\frac{1}{\sqrt{k}}\right) = \mathcal{O}\left(\sqrt{\alpha}\right) \,,$$

for $f = Ku^\dagger$.

Proof. We start proving this statement by showing that the inner product of $u^k - u^\dagger$ with a singular vector u_j simplifies to

$$\begin{aligned}
\langle u^k - u^\dagger, u_j\rangle_{\mathcal{U}} &= \langle u^k, u_j\rangle_{\mathcal{U}} - \langle u^\dagger, u_j\rangle_{\mathcal{U}} \\
&= \left(1 - (1 - \tau\sigma_j^2)^k\right) \langle u^\dagger, u_j\rangle_{\mathcal{U}} - \langle u^\dagger, u_j\rangle_{\mathcal{U}} \\
&= (1 - \tau\sigma_j^2)^k \langle u^\dagger, u_j\rangle_{\mathcal{U}} \\
&= \underbrace{\sigma_j (1 - \tau\sigma_j^2)^k}_{=: r(\sigma_j)} \langle w, u_j\rangle_{\mathcal{U}} \,,
\end{aligned}$$

with the second equality following from Equation (3.25). As our next step, we want to find an upper bound for $r(\sigma_j)$. We therefore analyse the concave function $r(\sigma) = \sigma(1 - \tau\sigma^2)^k$ by computing its first derivative, setting it to zero and inserting the resulting argument that maximises r. This yields

$$\max_\sigma r(\sigma) = \frac{\left(\frac{2k}{2k+1}\right)^k}{\sqrt{\tau(2k+1)}} \leq \frac{1}{\sqrt{\tau(2k+1)}}$$

for $k \in \mathbb{N}$. Hence, we obtain the estimate

$$\left|\langle u^k - u^\dagger, u_j\rangle_{\mathcal{U}}\right| \leq \frac{|\langle w, u_j\rangle_{\mathcal{U}}|}{\sqrt{\tau(2k+1)}} \,,$$

and consequently

$$\|u^k - u^\dagger\|_{\mathcal{U}} = \sqrt{\sum_{j=1}^\infty |\langle u^k - u^\dagger, u_j\rangle_{\mathcal{U}}|^2} \leq \frac{1}{\sqrt{\tau(2k+1)}} \sqrt{\sum_{j=1}^\infty |\langle w, u_j\rangle_{\mathcal{U}}|^2} = \frac{\|w\|_{\mathcal{U}}}{\sqrt{\tau(2k+1)}} \,.$$

□

Together with the stepsize-constraint (3.26) we can further conclude convergence of the iterates to a least squares solution.

Lemma 3.3. *Let* (3.26) *be satisfied. Then the iterates* (3.20) *satisfy*

$$\|Ku^{k+1} - f\|_{\mathcal{V}} \leq \|Ku^k - f\|_{\mathcal{V}} \,,$$

for $f = Ku^\dagger$ *and all* $k \in \mathbb{N}$*, where equality only holds if* u^k *already satisfies the normal equation* (2.3)*.*

Proof. We easily estimate

$$\begin{aligned}
\|Ku^{k+1} - f\|_{\mathcal{V}}^2 &= \|K(I - \tau K^* K)u^k - (I - \tau K^*)f\|_{\mathcal{V}}^2 \\
&= \|Ku^k - f - \tau K K^*(Ku^k - f)\|_{\mathcal{V}}^2 \\
&= \|Ku^k - f\|_{\mathcal{V}}^2 - 2\tau\langle K^*(Ku^k - f), K^*(Ku^k - f)\rangle_{\mathcal{U}} + \tau^2\|KK^*(Ku^k - f)\|_{\mathcal{V}}^2 \\
&= \|Ku^k - f\|_{\mathcal{V}}^2 + \tau\left(\tau\|KK^*(Ku^k - f)\|_{\mathcal{V}}^2 - 2\|K^*(Ku^k - f)\|_{\mathcal{U}}^2\right) \\
&\leq \|Ku^k - f\|_{\mathcal{V}}^2 + \tau\|K^*(Ku^k - f)\|_{\mathcal{U}}^2 \underbrace{\left(\tau\|K\|_{\mathcal{L}(\mathcal{U},\mathcal{V})}^2 - 2\right)}_{<0} \\
&\leq \|Ku^k - f\|_{\mathcal{V}}^2 \,,
\end{aligned}$$

which proves the statement. □

Lemma 3.2 and Lemma 3.3 allow us to conclude the following proposition.

Proposition 3.3. *The Landweber iteration is a linear regularisation in the sense of Definition 3.2.*

In order to show that the Landweber iteration (3.20) in combination with the discrepancy principle (3.10) is also a convergent regularisation, we obviously have to look at the case of noisy data f^δ with $\|f^\delta - f\|_{\mathcal{V}} \leq \delta$ for $f = Ku^\dagger$. We denote the solution of (3.20) in case of noisy data f^δ as u_δ^k for all $k \in \mathbb{N}$ and obtain the following estimate for the norm between u_δ^k and $u^\dagger$.

Lemma 3.4. *Let* (3.27) *be satisfied. Then the Landweber iterates* (3.20) *satisfy*

$$\|u_\delta^k - u^\dagger\|_{\mathcal{U}} \leq \tau k \delta \|K\|_{\mathcal{L}(\mathcal{U},\mathcal{V})} + \frac{\|w\|_{\mathcal{U}}}{\sqrt{\tau(2k-1)}} \tag{3.28}$$

for $k \in \mathbb{N} \setminus \{1\}$, $f = Ku^\dagger$, $f^\delta \in \mathcal{V}$ *and* $\|f^\delta - f\|_{\mathcal{V}} \leq \delta$.

Proof. Similar to the proof of Lemma 3.2 we consider the inner product between $u_\delta^k - u^\dagger$ and a singular vector u_j, which yields

$$\begin{aligned}
\langle u_\delta^k - u^\dagger, u_j\rangle_{\mathcal{U}} &= \frac{1}{\sigma_j}\left(\left(1 - (1 - \tau\sigma_j^2)^k\right)\langle f^\delta, v_j\rangle_{\mathcal{V}} - \langle f, v_j\rangle_{\mathcal{V}}\right) \\
&= \frac{1}{\sigma_j}\left(1 - (1 - \tau\sigma_j^2)^k\right)\langle f^\delta - f, v_j\rangle_{\mathcal{V}} - \sigma_j(1 - \tau\sigma_j^2)^k\langle w, v_j\rangle_{\mathcal{V}} \,.
\end{aligned}$$

Hence, for $k > 1$ we can use (3.24) to estimate

$$\begin{aligned}
\frac{1}{\sigma_j}\left(1 - (1 - \tau\sigma_j^2)^k\right)\left|\langle f^\delta - f, v_j\rangle_{\mathcal{V}}\right| &= \tau\sigma_j\sum_{j=1}^{k}(1 - \tau\sigma_j^2)^{k-j}\left|\langle f^\delta - f, v_j\rangle_{\mathcal{V}}\right| \\
&\leq \tau k\sigma_j\left|\langle f^\delta - f, v_j\rangle_{\mathcal{V}}\right| \leq \tau k\sigma_1\left|\langle f^\delta - f, v_j\rangle_{\mathcal{V}}\right| \,.
\end{aligned}$$

Together with the result from Lemma 3.2 we conclude

$$\|u_\delta^k - u^\dagger\|_{\mathcal{U}} \leq \tau k \delta \|K\|_{\mathcal{L}(\mathcal{U},\mathcal{V})} + \frac{\|w\|_{\mathcal{U}}}{\sqrt{\tau(2k-1)}} \,.$$

□

Note that the decrease of the residual in Lemma 3.3 holds true for all $f \in \mathcal{V}$. As we obviously do not want to iterate until infinity – this would blow up the data error in (3.28) – this decrease together with the stepsize-constraint (3.26) motivates the use of (3.10) as a stopping criterion. The following lemma shows that with (3.20) we indeed minimise the difference between u_δ^k and $u^\dagger$ (in terms of the $\mathcal{U}$ norm) as long as (3.10) is violated.

Lemma 3.5. *Let* (3.26) *be satisfied. Then the iterates of* (3.20) *satisfy*

$$\|u_\delta^{k+1} - u^\dagger\|_{\mathcal{U}} \leq \|u_\delta^k - u^\dagger\|_{\mathcal{U}}$$

for $k \leq k^*$, $f = Ku^\dagger$ and $f^\delta \in \mathcal{V}$ with $\|f^\delta - f\|_{\mathcal{V}} \leq \delta$. Here, k^* satisfies the discrepancy principle (3.10) *for $\eta = 2/(2 - \tau\|K\|^2_{\mathcal{L}(\mathcal{U},\mathcal{V})}) > 1$. Moreover, equality can only be attained for $\delta = 0$ and u_δ^k satisfying the normal equation* (2.3).

Proof. We prove the statement by showing that $\|u_\delta^{k+1} - u^\dagger\|^2_{\mathcal{U}} - \|u_\delta^k - u^\dagger\|^2_{\mathcal{U}}$ is negative whilst the discrepancy principle is not violated. We estimate

$$\begin{aligned}
\|u_\delta^{k+1} - u^\dagger\|^2_{\mathcal{U}} - \|u_\delta^k - u^\dagger\|^2_{\mathcal{U}} &= \|u_\delta^k - \tau K^*(Ku_\delta^k - f^\delta) - u^\dagger\|^2_{\mathcal{U}} - \|u_\delta^k - u^\dagger\|^2_{\mathcal{U}} \\
&= \tau^2\|K^*(Ku_\delta^k - f^\delta)\|^2_{\mathcal{U}} - 2\tau\langle Ku_\delta^k - f^\delta, Ku_\delta^k - f\rangle_{\mathcal{V}} \\
&\leq \tau^2\|K\|^2_{\mathcal{L}(\mathcal{U},\mathcal{V})}\|Ku_\delta^k - f^\delta\|^2_{\mathcal{V}} - 2\tau \underbrace{\langle Ku_\delta^k - f^\delta, Ku_\delta^k - f + f^\delta - f^\delta\rangle_{\mathcal{V}}}_{=\|Ku_\delta^k - f^\delta\|^2_{\mathcal{V}} + \langle Ku_\delta^k - f^\delta, f^\delta - f\rangle_{\mathcal{V}}} \\
&= \tau\left(\tau\|K\|^2_{\mathcal{L}(\mathcal{U},\mathcal{V})} - 2\right)\|Ku_\delta^k - f^\delta\|^2_{\mathcal{V}} + 2\tau\langle f - f^\delta, Ku_\delta^k - f^\delta\rangle_{\mathcal{V}} \\
&\leq \tau\left(\tau\|K\|^2_{\mathcal{L}(\mathcal{U},\mathcal{V})} - 2\right)\|Ku_\delta^k - f^\delta\|^2_{\mathcal{V}} + 2\tau\delta\|Ku_\delta^k - f^\delta\|_{\mathcal{V}} \\
&= -\tau\|Ku_\delta^k - f^\delta\|_{\mathcal{V}}\left(\left(2 - \tau\|K\|^2_{\mathcal{L}(\mathcal{U},\mathcal{V})}\right)\|Ku_\delta^k - f^\delta\|_{\mathcal{V}} - 2\delta\right) \\
&= -\frac{2\tau}{\eta}\|Ku_\delta^k - f^\delta\|_{\mathcal{V}}\left(\|Ku_\delta^k - f^\delta\|_{\mathcal{V}} - \eta\delta\right) .
\end{aligned}$$

Hence, for $k \leq k_*$ we conclude $\|u_\delta^{k+1} - u^\dagger\|_{\mathcal{U}} \leq \|u_\delta^k - u^\dagger\|_{\mathcal{U}}$. □

3.3 Tikhonov regularisation revisited

We conclude this chapter by showing that Tikhonov regularisation can not just be interpreted as the spectral regularisation (3.6) and the solution of the well-posed operator equation (3.14), but also as the minimiser of a functional.

Theorem 3.10. *For $f \in \mathcal{V}$ the Tikhonov-regularised solution $u_\alpha = R_\alpha f$ with R_α as defined in* (3.6) *is uniquely determined as the global minimiser of the Tikhonov-functional*

$$T_\alpha(u) := \frac{1}{2}\|Ku - f\|^2_{\mathcal{V}} + \frac{\alpha}{2}\|u\|^2_{\mathcal{U}} . \tag{3.29}$$

Proof. $\Rightarrow$: Let u_α be the Tikhonov-regularised solution and we show that it is also a global minimiser. A global minimiser $\hat{u} \in \mathcal{U}$ of $T_\alpha(\hat{u})$ is characterised via $T_\alpha(\hat{u}) \leq T_\alpha(u)$ for all $u \in \mathcal{U}$.

Hence, it follows from

$$\begin{aligned}
T_\alpha(u) - T_\alpha(u_\alpha) &= \frac{1}{2}\left\|Ku - f\right\|_{\mathcal{V}}^2 + \frac{\alpha}{2}\left\|u\right\|_{\mathcal{U}}^2 - \frac{1}{2}\left\|Ku_\alpha - f\right\|_{\mathcal{V}}^2 - \frac{\alpha}{2}\left\|u_\alpha\right\|_{\mathcal{U}}^2 \\
&= \frac{1}{2}\|Ku\|_{\mathcal{V}}^2 - \langle Ku, f\rangle + \frac{\alpha}{2}\|u\|_{\mathcal{U}}^2 - \frac{1}{2}\|Ku_\alpha\|_{\mathcal{V}}^2 + \langle Ku_\alpha, f\rangle - \frac{\alpha}{2}\|u_\alpha\|_{\mathcal{U}}^2 \\
&\quad + \underbrace{\langle (K^*K + \alpha I)u_\alpha - K^*f, u_\alpha - u\rangle}_{=0} \\
&= \frac{1}{2}\left\|Ku - Ku_\alpha\right\|_{\mathcal{V}}^2 + \frac{\alpha}{2}\left\|u - u_\alpha\right\|_{\mathcal{U}}^2 \\
&\geq 0
\end{aligned}$$

that u_α is a global minimiser of T_α.
$\Leftarrow$: Let now $\hat{u}$ be a global minimiser. If we have $T_\alpha(\hat{u}) \leq T_\alpha(u)$ (for all $u \in \mathcal{U}$), it follows with $u = \hat{u} + \tau v$ for arbitrary $\tau > 0$ and fixed $v \in \mathcal{U}$ that

$$0 \leq T_\alpha(u) - T_\alpha(\hat{u}) = \frac{\tau^2}{2}\|Kv\|_{\mathcal{V}}^2 + \frac{\tau^2\alpha}{2}\|v\|_{\mathcal{U}}^2 + \tau\langle (K^*K + \alpha I)\hat{u} - K^*f, v\rangle_{\mathcal{U}}$$

holds true. Dividing by τ and subsequent consideration of the limit $\tau \downarrow 0$ thus yields

$$\langle (K^*K + \alpha I)\hat{u} - K^*f, v\rangle_{\mathcal{U}} \geq 0\,, \text{ for all } v \in \mathcal{U}.$$

Thus $(K^*K + \alpha I)\hat{u} - K^*f = 0$ and we conclude $\hat{u} = u_\alpha$, i.e. a global minimiser is the Tikhonov-regularised solution. This also shows that the global minimiser of the Tikhonov functional (3.29) is unique. $\square$

This result paves the way for a generalisation of Tikhonov regularisation to a much broader class of regularisation methods that we want to discuss in the following chapter.

Chapter 4

Variational regularisation for linear inverse problems

At the end of the last chapter we have seen that Tikhonov regularisation[1] $R_\alpha f^\delta$ can be characterised as the solution of the minimisation problem

$$R_\alpha f^\delta = \arg\min_{u\in\mathcal{U}} \Big\{ \frac{1}{2}\|Ku - f^\delta\|_{\mathcal{V}}^2 + \frac{\alpha}{2}\|u\|_{\mathcal{U}}^2 \Big\}\,.$$

It is well known that the solution to an unconstrained minimisation problem has a vanishing derivative. In function spaces, the (Gâteaux-) derivative is also called the "first variation" such that minimisation problems are also called *variational problems* and methods that rely on minimising a functional *variational methods*. In this section we want to investigate variational methods for regularisation of linear inverse problems. To do so we will generalise Tikhonov regularisation by choosing different regularisation functionals $J\colon \mathcal{U} \to \mathbb{R}$ and compute regularised solutions by minimising the functional

$$\Phi_{\alpha,f^\delta} := \frac{1}{2}\|Ku - f^\delta\|_{\mathcal{V}}^2 + \alpha J(u)\,.$$

Regularisation of this form is sometimes called *Tikhonov-type regularisation* but we will refer to this as *variational regularisation*. Before we have a look at the theory behind variational regularisation such as the existence and uniqueness of minimisers we will discuss several examples of *regularisation functionals* J.

Example 4.1 (Tikhonov-Philipps regularisation)**.** The easiest way to extend classical Tikhonov regularisation to a more general regularisation method is to replace $\frac{1}{2}\|u\|_{\mathcal{U}}^2$ by $\frac{1}{2}\|Du\|_{\mathcal{Z}}^2$ where $D\colon \mathcal{U} \to \mathcal{Z}$ is a linear (not necessarily bounded) operator and we thus minimise

$$\frac{1}{2}\|Ku - f^\delta\|_{\mathcal{V}}^2 + \frac{\alpha}{2}\|Du\|_{\mathcal{Z}}^2\,,$$

which became known as *Tikhonov-Philipps* regularisation. While Tikhonov regularisation penalises the norm of u, in Tikhonov-Philipps regularisation only certain features of u (depending on the choice of D) are penalised. The most frequent used operator D in imaging applications is the gradient operator ∇ such that the regulariser J corresponds to the semi-norm on $H^1(\Omega)$ which is the Sobolev space of functions $u \in L^2(\Omega)$ such that the weak derivative ∇u exists and $\nabla u \in$

[1] This regularisation is called *ridge regression* in the statistical literature.

$L^2(\Omega, \mathbb{R}^n)$. By using this regulariser, only the variations in u but not the actual intensities are penalised which helps to control noise without a bias of the intensities towards zero.

If the operator D is given by $Du = (u, \nabla u)$ and $\mathcal{Z} = L^2(\Omega) \times L^2(\Omega, \mathbb{R}^n)$ is equipped with the natural inner product for product spaces, then

$$J(u) = \frac{1}{2}\|Du\|_{\mathcal{Z}}^2 = \frac{1}{2}\|u\|_{L^2}^2 + \frac{1}{2}\|\nabla u\|_{L^2}^2$$

is the norm on $H^1(\Omega)$ and it corresponds to classical Tikhonov regularisation on $H^1(\Omega)$.

Example 4.2 (Maximum-entropy regularisation)**.** Maximum-entropy regularisation is of particular interest if solutions of the inverse problem are assumed to be probability density functions (pdf), i.e. functions in the set

$$\mathrm{PDF}(\Omega) := \left\{ u \in L^1(\Omega) \;\middle|\; \int_\Omega u(x)\, dx = 1,\; u \geq 0 \right\} .$$

The set $\mathrm{PDF}(\Omega)$ is a convex subset but it is not a subspace as differences of pdfs are not necessarily pdfs. The (differential) entropy used in physics and information theory is defined as the functional $\mathrm{PDF}(\Omega) \to \mathbb{R}$ with

$$u \mapsto -\int_\Omega u(x) \log(u(x))\, dx ,$$

and the convention $0 \log(0) := 0$. The corresponding regularisation with the negative entropy reads as

$$\min_{u \in \mathrm{PDF}(\Omega)} \left\{ \frac{1}{2} \|Ku - f\|_{\mathcal{V}}^2 + \alpha \int_\Omega u(x) \log(u(x))\, dx \right\} ,$$

for operators $K \in \mathcal{L}(L^1(\Omega), \mathcal{V})$.

Example 4.3 (ℓ^1-regularisation)**.** When it comes to non-injective operators $K \in \mathcal{L}(\ell^1, \ell^2)$ between sequence spaces, the ℓ^1-norm, i.e. $\|u\|_{\ell^1} := \sum_{j=1}^\infty |u_j|$ is often used as a regulariser, in order to enforce sparse solutions, see example in Figure 4.1. The corresponding minimisation problem[2] reads as

$$\min_{u \in \ell^1} \left\{ \frac{1}{2} \|Ku - f\|_{\ell^2}^2 + \alpha \sum_{j=1}^\infty |u_j| \right\} .$$

Example 4.4 (Elastic net)**.** Another regularisation method from statistics is the *elastic net*, where the regulariser is the weighted sum of the ℓ^1-norm and the squared ℓ^2-norm:

$$J(u) = \|u\|_{\ell^1} + \frac{\beta}{2}\|u\|_{\ell^2}^2 .$$

Here the idea is to combine two favourable models in order to get sparse solutions with more stability. As $\ell^1 \subset \ell^2$ we could either consider the elastic net on the Banach space ℓ^1 or on the Hilbert space ℓ^2. In case we decide to do the latter, we can extend the elastic net such that

$$J(u) = \begin{cases} \|u\|_{\ell^1} + \frac{\beta}{2}\|u\|_{\ell^2}^2 & \text{if } u \in \ell^1 \\ \infty & \text{if } u \in \ell^2 \setminus \ell^1 \end{cases} .$$

Intuitively, the value ∞ makes sure that a minimiser will never be in $\ell^2 \setminus \ell^1$ but we will discuss this aspect in more detail later.

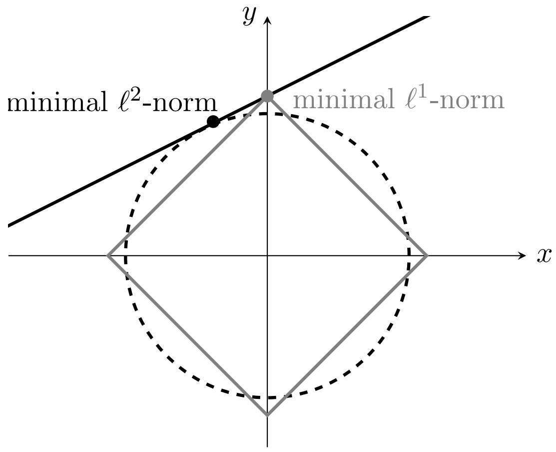

Figure 4.1: Non-injective operators have a non-trivial kernel such that the inverse problem has more than one solution and the solutions form an affine subspace visualised by the solid line. Different regularisation functionals favour different solutions. The circle and the diamond indicate all points with constant ℓ^2-norm, respectively ℓ^1-norm, and the minimal ℓ^2-norm and ℓ^1-norm solutions are the intersections of the line with the circle, respectively the diamond. As it can be seen, the minimal ℓ^2-norm solution has two non-zero components while the minimal ℓ^1-norm solution has only one non-zero component and thus is *sparser*.

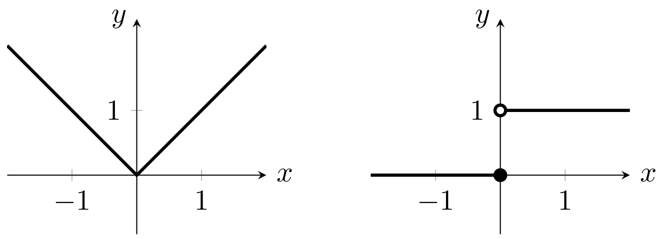

Figure 4.2: The absolute value function on the left is in $H^{1,1}(\mathbb{R})$ while the Heaviside function on the right is not. The solid dot at a jump indicates the value that the function takes. However, the Heaviside function is in $\mathrm{BV}(\mathbb{R})$ which shows that $\mathrm{BV}(\Omega)$ is larger than $H^{1,1}(\mathbb{R})$. Moreover, it shows that $\mathrm{BV}(\mathbb{R})$ includes function with discontinuities which is a favourable model for images with sharp edges.

Example 4.5 (Total variation)**.** Total variation as a regulariser has originally been introduced for image-denoising and -restoration applications with the goal to preserve edges in images, respectively discontinuities in signals [12]. For smooth signals $u \in H^{1,1}(\Omega)$, i.e. $u \in L^1(\Omega)$ and has a weak derivative $\nabla u \in L^1(\Omega, \mathbb{R}^n)$, the total variation is simply defined as the semi-norm on the Sobolev space $H^{1,1}(\Omega)$

$$\mathrm{TV}(u) := \int_\Omega \|\nabla u(x)\|_2 \, dx \, .$$

However, functions in $H^{1,1}(\Omega)$ may not allow discontinuities which are useful in imaging applications to model images with sharp edges.

To allow discontinuities while still preserving some regularity (otherwise we could model images in $L^1(\Omega)$ for instance) we generalise the definition of the total variation. It is well-known (e.g. Cauchy–Schwarz inequality) that for $x, v \in \mathbb{R}^n$ with $\|v\|_2 \leq 1$ we have that $\langle v, x \rangle \leq \|x\|_2$. Thus,

[2] This is called *lasso* in the statistical literature.

for any *test function* $\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)$ with

$$\mathcal{D}(\Omega, \mathbb{R}^n) := \left\{ \varphi \in C_0^\infty(\Omega; \mathbb{R}^n) \,\middle|\, \|\varphi(x)\|_2 \leq 1 \right\}$$

we have that

$$\mathrm{TV}(u) = \int_\Omega \|\nabla u(x)\|_2 \, dx \geq \int_\Omega \langle \nabla u(x), \varphi(x) \rangle \, dx = -\int_\Omega \langle u(x), \operatorname{div} \varphi(x) \rangle \, dx$$

where the last equality is due to partial integration (Gauss' divergence theorem). In fact one can show that

$$\mathrm{TV}(u) = \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \int_\Omega \langle u(x), \operatorname{div} \varphi(x) \rangle \, dx \,,$$

which gives rise to the definition of *functions of bounded variation.*

$$\mathrm{BV}(\Omega) := \left\{ u \in L^1(\Omega) \,\middle|\, \|u\|_{\mathrm{BV}} := \|u\|_{L^1} + \mathrm{TV}(u) < \infty \right\}$$

It can be shown that $\mathrm{BV}(\Omega)$ is much larger than $H^{1,1}(\Omega)$ and contains functions with discontinuities, see examples in Figure 4.2.

The total variation regularisation can then be written as

$$\min_{u \in \mathrm{BV}(\Omega)} \left\{ \frac{1}{2} \|Ku - f\|_{\mathcal{V}}^2 + \alpha \, \mathrm{TV}(u) \right\} \,, \tag{4.1}$$

for $K \in \mathcal{L}(\mathrm{BV}(\Omega), \mathcal{V})$.

To summarise the introduction, variational regularisation aims at finding approximations to the solution of the inverse problem (1.1) by minimising appropriate functionals of the form

$$\Phi_{\alpha, f^\delta}(u) := \frac{1}{2} \|Ku - f^\delta\|_{\mathcal{V}}^2 + \alpha J(u) \,, \tag{4.2}$$

where $J : \mathcal{U} \to \mathbb{R} \cup \{+\infty\}$ represents a functional over the Banach space $\mathcal{U}$, $\mathcal{V}$ is a Hilbert space and $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ a linear and continuous operator, and $\alpha > 0$ is a real, positive constant. The term $D(u) := \frac{1}{2} \|Ku - f^\delta\|_{\mathcal{V}}^2$ is usually named *fidelity* or *data term*, as it measures the deviation between the measured data f^δ and the forward model Ku. The functional J is the *regularisation term* or *regulariser* as it will impose certain regularity conditions on the unknown u. The *regularisation parameter* will balance between both terms. Next, we will study some general theory on variational methods that will tell us under which conditions we can expect existence and uniqueness of solutions to those minimisation problems.

4.1 Variational methods

4.1.1 Background

Banach spaces and weak convergence

To cover all the examples of the beginning of this chapter we have to extend our setting to include *Banach spaces.* These are complete, normed vector spaces (as Hilbert spaces) but they may not have an inner product. For every Banach space $\mathcal{U}$, we can define the space of linear and continuous functionals which is called the *dual space* $\mathcal{U}^*$ of $\mathcal{U}$, i.e. $\mathcal{U}^* := \mathcal{L}(\mathcal{U}, \mathbb{R})$. Let $u \in \mathcal{U}$ and $p \in \mathcal{U}^*$, then we usually write the *dual product* $\langle p, u\rangle$ instead of $p(u)$. Obviously, the dual product is not symmetric (in contrast to the inner product of Hilbert spaces). Moreover, for any $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ there exists a unique operator $K^* \colon \mathcal{V}^* \to \mathcal{U}^*$, called the *adjoint* of K such that for all $u \in \mathcal{U}$ and $p \in \mathcal{V}^*$ we have

$$\langle K^* p, u\rangle = \langle p, Ku\rangle \,.$$

It is easy to see that either sides of the equation are well-defined, e.g. $K^*p \in \mathcal{U}^*$ and $u \in \mathcal{U}$.

As the dual space is a Banach space as well, it has a dual space as well which we will call the bi-dual space of $\mathcal{U}$ and denote it with $\mathcal{U}^{**} := (\mathcal{U}^*)^*$. As every $u \in \mathcal{U}$ defines a continuous and linear mapping on the dual space $\mathcal{U}^*$ by $\langle E(u), u^*\rangle := \langle u^*, u\rangle$, the mapping $E \colon \mathcal{U} \to \mathcal{U}^{**}$ is well-defined. It can be shown that E is a linear and continuous isometry (and thus injective). In the special case when E is surjective, we call $\mathcal{U}$ *reflexiv*. Examples of reflexive Banach spaces include Hilbert spaces and L^p, ℓ^p spaces with $1 < p < \infty$. We call the space $\mathcal{U}$ *separable* if there exists a set $\mathcal{X} \subset \mathcal{U}$ of at most countable cardinality such that $\overline{\mathcal{X}} = \mathcal{U}$.

A problem in infinite dimensional spaces is that bounded sequences may fail to have convergent subsequences. An example is for instance in ℓ^2 the sequence $\{u^k\}_{k\in\mathbb{N}} \subset \ell^2, u^k_j = 1$ if $k = j$ and 0 otherwise. It is easy to see that $\|u^k\|_{\ell^2} = 1$ and that there is no $u \in \ell^2$ such that $u^k \to u$. To circumvent this problem, we define a new weaker topology on $\mathcal{U}$. We say that $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}$ *converges weakly* to $u \in \mathcal{U}$ if and only if for all $p \in \mathcal{U}^*$ the sequence of real numbers $\{\langle p, u_j\rangle\}_{j\in\mathbb{N}}$ converges and

$$\langle p, u_j\rangle \to \langle p, u\rangle \,.$$

We will denote weak convergence by $u_j \rightharpoonup u$. On a dual space $\mathcal{U}^*$ we could define another topology (in addition to the strong topology induced by the norm and the weak topology as the dual space is a Banach space as well). We say a sequence $\{p_j\}_{j\in\mathbb{N}} \subset \mathcal{U}^*$ *converges in weak-$*$* to $p \in \mathcal{U}^*$ if and only if

$$\langle p_j, u\rangle \to \langle p, u\rangle \quad \text{for all } u \in \mathcal{U}$$

and we denote weak-$*$ convergence by $p_j \stackrel{*}{\rightharpoonup} p$. Similarly, for any topology τ on $\mathcal{U}$ we denote the convergence in that topology by $u_j \stackrel{\tau}{\to} u$.

With these two new notions of convergence, we can solve the problem of bounded sequences:

Theorem 4.1 (Sequential Banach-Alaoglu Theorem, e.g. [13, p. 70] or [14, p. 141])**.** *Let $\mathcal{U}$ be a separable normed vector space. Then every bounded sequence $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}^*$ has a weak-* convergent subsequence.*

Corollary 4.1 ([16, p. 64])**.** *Each bounded sequence $\{u_j\}_{j\in\mathbb{N}}$ in a reflexive Banach space $\mathcal{U}$ has a weakly convergent subsequence.*

Infinity calculus

We will look at functionals $E : \mathcal{U} \to \mathbb{R}_\infty$ whose range is modelled to be the *extended real line* $\mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$ where the symbol ∞ denotes an element that is not part of the real line that is by definition larger than any other element of the reels, i.e. $x < \infty$ for all $x \in \mathbb{R}$. This is useful to model constraints: For instance, if we were trying to minimise $E : [-1, \infty) \to \mathbb{R}, x \mapsto x^2$ we could remodel this minimisation problem by $\tilde{E} : \mathbb{R} \to \mathbb{R}_\infty$

$$\tilde{E}(x) = \begin{cases} x^2 & \text{if } x > -1 \\ \infty & \text{else} \end{cases} .$$

Obviously both functionals have the same minimiser but $\tilde{E}$ is defined on a vector space and not only on a subset. This has two important features: On the on hand, it makes many theoretical arguments easier as we do not need to worry whether $E(x + y)$ is defined or not. On the other hand, it makes practical implementations easier as we are dealing with unconstrained optimisation instead of constrained optimisation. This comes at a cost that some algorithms are not applicable anymore, e.g. the function $\tilde{E}$ is not differentiable everywhere whereas E is (in the interior of its domain).

It is useful to note that one can calculate on the extended real line $\mathbb{R}_\infty$ as we are used to on the real line $\mathbb{R}$ but the operations with ∞ need yet to be defined. As ∞ is larger than any other element it makes sense that it dominates any other calculation, i.e. for all $x \in \mathbb{R}$ and $\lambda > 0$, we have

$$x + \infty := \infty + x := \infty, \quad \lambda \cdot \infty := \infty \cdot \lambda := \infty, \quad x/\infty := 0.$$

However, care needs to be taken as some calculations are not defined, e.g. $\infty - \infty$.

Definition 4.1. *Let $\mathcal{C} \subset \mathcal{U}$ be a subset of a vector space $\mathcal{U}$ and $E\colon \mathcal{C} \to \mathbb{R}_\infty$ a functional. Then the* effective domain *of E is*

$$\operatorname{dom}(E) := \{u \in \mathcal{C} \mid E(u) < \infty\} .$$

Convex calculus

A property of fundamental importance of sets and functions is convexity. Simply said, a set (or function) is convex if the shape is regular. More precisely it is defined as follows.

Definition 4.2. *Let $\mathcal{U}$ be a vector space. A subset $\mathcal{C} \subset \mathcal{U}$ is called* convex*, if $\lambda u + (1 - \lambda)v \in \mathcal{C}$ for all $\lambda \in (0, 1)$ and all $u, v \in \mathcal{C}$.*

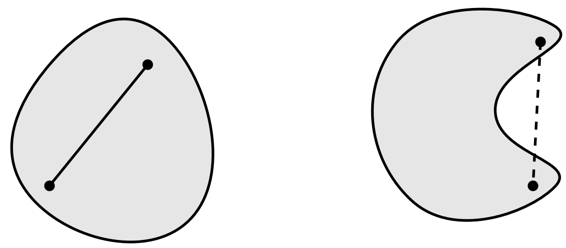

Figure 4.3: Example of a convex set (left) and non-convex set (right).

In analogy we can define convex functionals.

Definition 4.3. *Let $\mathcal{C} \subset \mathcal{U}$ be a convex set. A functional $E\colon \mathcal{C} \to \mathbb{R}_\infty$ is called* convex*, if*

$$E(\lambda u + (1 - \lambda)v) \leq \lambda E(u) + (1 - \lambda)E(v)$$

for all $\lambda \in (0, 1)$ and all $u, v \in \operatorname{dom}(E)$ with $u \neq v$. It is called strictly *convex if the inequality is strict.*

Example 4.6. The absolute value function $\mathbb{R} \to \mathbb{R}, x \mapsto |x|$ is convex but not strictly convex while the quadratic function $x \mapsto x^2$ is stricly convex. For other examples, see Figure 4.4.

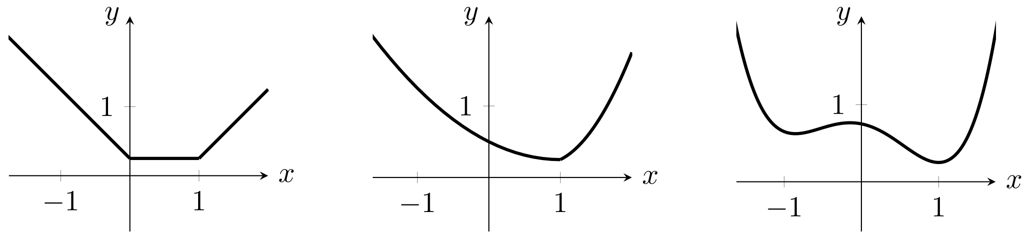

Figure 4.4: Example of a convex function (left), a strictly convex function (middle) and a non-convex function (right).

Example 4.7. Let $\mathcal{C} \subset \mathcal{U}$ be a set. Then the *characteristic function* $\chi_{\mathcal{C}} \colon \mathcal{U} \to \mathbb{R}_\infty$ with

$$\chi_{\mathcal{C}}(u) := \begin{cases} 0 & u \in \mathcal{C} \\ +\infty & u \in \mathcal{U} \setminus \mathcal{C} \end{cases} \tag{4.3}$$

is convex if and only if $\mathcal{C} \subset \mathcal{U}$ is a convex subset. To see the convexity, if both u and v are in $\mathcal{C}$, then by the convexity of $\mathcal{C}$ the convex combination $\lambda u + (1-\lambda)v$ is as well in $\mathcal{C}$ and both the left and the right hand side of the desired inequality are zero.

Lemma 4.1. *Let $\alpha, \beta \geq 0$ and $E, F \colon \mathcal{U} \to \mathbb{R}_\infty$ be two convex functions. Then $\alpha E + \beta F \colon \mathcal{U} \to \mathbb{R}_\infty$ is convex. Furthermore, if $\beta > 0$ and F strictly convex, then $\alpha E + \beta F$ is strictly convex.*

Proof. The proof shall be done as an exercise. □

Definition 4.4. *Let $\mathcal{U}$ be a Banach space and $E \colon \mathcal{U} \to \mathbb{R}_\infty$ a functional. Then, E is called* subdifferentiable *at $u \in \mathcal{U}$, if there exists an element $p \in \mathcal{U}^*$ such that*

$$E(v) \geq E(u) + \langle p, v - u \rangle$$

holds, for all $v \in \mathcal{U}$. Furthermore, we call p a subgradient *at position u. The collection of all subgradients at position u, i.e.*

$$\partial E(u) := \{ p \in \mathcal{U}^* \mid E(v) \geq E(u) + \langle p, v - u \rangle, \forall v \in \mathcal{U} \} ,$$

is called subdifferential *of E at u.*

Remark 4.1. Let $E \colon \mathcal{C} \to \mathbb{R}_\infty$ be a convex functional. Then the subdifferential is non-empty at all $u \in \mathrm{dom}(E)$.

For non-differentiable functionals the subdifferential is multivalued; we want to consider the subdifferential of the absolute value function as an illustrative example.

Example 4.8. Let $\mathcal{U} = \mathbb{R}$, and let $E \colon \mathbb{R} \to \mathbb{R}$ be the absolute value function $E(x) = |x|$. Then, the subdifferential of E at x is given by

$$\partial E(x) = \mathrm{sign}(x) := \begin{cases} \{1\} & \text{for } x > 0 \\ [-1, 1] & \text{for } x = 0 \\ \{-1\} & \text{for } x < 0 \end{cases} ,$$

which you will prove as an exercise. A visual explanation is given in Figure 4.5.

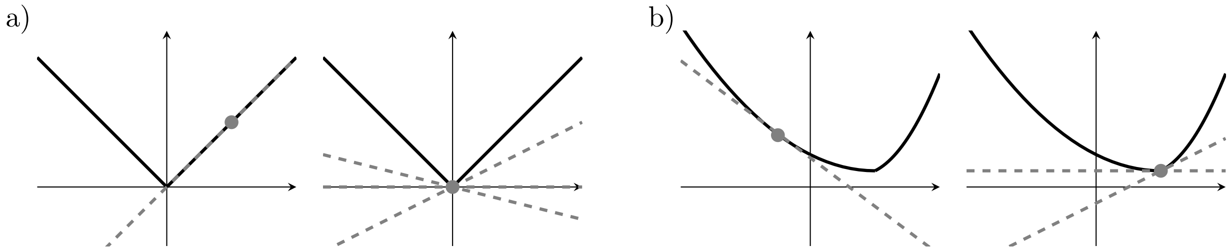

Figure 4.5: Visualisation of the subdifferential. Linear approximations of the functional have to lie completely underneath the function. For points where the function is not differentiable there may be more than one such approximation.

4.1.2 Minimisers

Definition 4.5. *Let $\mathcal{C} \subset \mathcal{U}$ be a set and $E\colon \mathcal{C} \to \mathbb{R}_\infty$ a functional. We say that $u^* \in \mathcal{C}$ solves the minimisation problem*

$$\min_{u \in \mathcal{C}} E(u)$$

if and only if $E(u^*) < \infty$ and $E(u^*) \le E(v)$, for all $v \in \mathcal{C}$. We call u^* a minimiser *of E.*

We will now review two properties that are necessary for the well-definedness of a minimisation problem.

Definition 4.6. *A functional $E\colon \mathcal{U} \to \mathbb{R}_\infty$ is called* proper*, if the effective domain* dom(E *isis not empty.*

Definition 4.7. *A functional $E\colon \mathcal{U} \to \mathbb{R}_\infty$ is called* bounded from below *if there exists a constant $C > -\infty$ such that for all $u \in \mathcal{U}$ we have $E(u) \ge C$.*

This condition is obviously necessary for the existence of the infimum $\inf_{u \in \mathcal{U}} E(u)$.
Finally we characterise minimisers of convex functionals.

Theorem 4.2. *Let $E\colon \mathcal{U} \to \mathbb{R}_\infty$ be a proper, convex functional. An element $u \in \mathcal{U}$ is a minimiser of E if and only if $0 \in \partial E(u)$.*

Proof. By definition, $0 \in \partial E(u)$ if and only if for all $v \in \mathcal{U}$ it holds

$$E(v) \ge E(u) + \langle 0, v - u \rangle = E(u)\,,$$

which is by definition the case if and only if u is a minimiser of E. □

4.1.3 Existence

If all minimising sequences (that converge to the infimum assuming it exists) are unbounded, then there cannot exist a minimiser. A sufficient condition to avoid such a scenario is *coercivity*.

Definition 4.8. *A functional $E\colon \mathcal{U} \to \mathbb{R}_\infty$ is called* coercive*, if for all $\{u_j\}_{j \in \mathbb{N}}$ with $\|u_j\|_{\mathcal{U}} \to \infty$ we have $E(u_j) \to \infty$.*

Remark 4.2. Coercivity is equivalent to that if the function values $\{E(u_j)\}_{j \in \mathbb{N}} \subset \mathbb{R}$ are bounded, so is the sequence $\{u_j\}_{j \in \mathbb{N}} \subset \mathcal{U}$.

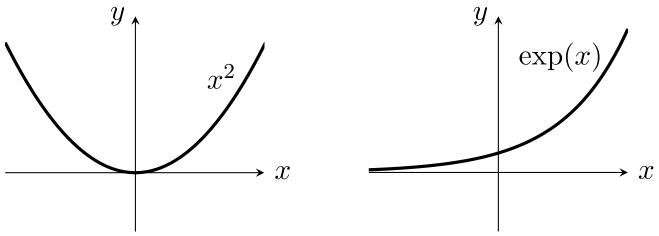

Figure 4.6: While the coercive function on the left has a minimiser, it is easy to see that the non-coercive function on the right does not have a minimiser.

Although coercivity is not strictly speaking necessary, it is sufficient that all minimising sequences are bounded.

Lemma 4.2. *Let $E\colon \mathcal{U} \to \mathbb{R}_\infty$ be a proper, coercive functional and bounded from below. Then the infimum $\inf_{u\in\mathcal{U}} E(u)$ is finite. Then, there are minimising sequences, i.e. $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}$ with $E(u_j) < \infty$ and $E(u_j) \to \inf_{u\in\mathcal{U}} E(u)$. Moreover, all minimising sequences are bounded, in the sense that there exists a constant $C > 0$ such that $\|u_j\|_\mathcal{U} < C$ for all $j \in \mathbb{N}$.*

Proof. As E is proper and bounded from below, there exists a $C_1 > 0$ such that we have $-\infty < -C_1 < \inf_u E(u) < \infty$ which also guarantees the existence of a minimising sequence. Let $\{u_j\}_{j\in\mathbb{N}}$ be any minimising sequence, i.e. $E(u_j) \to \inf_u E(u)$. Then there exists a $j_0 \in \mathbb{N}$ such that for all $j > j_0$ we have

$$E(u_j) \leq \underbrace{\inf_u E(u) + 1}_{=:C_2} < \infty\,.$$

With $C_3 := \max_{1\leq j\leq j_0} E(u_j)$ and $C := \max(C_1, C_2, C_3)$ we get that $|E(u_j)| < C$ for all $j \in \mathbb{N}$. From the coercivity it follows that $\{u_j\}_{j\in\mathbb{N}}$ is bounded, see Remark 4.2. □

More importantly we are going to need that functionals are sequentially lower semi-continuous. Roughly speaking this means that the functional values for arguments near an argument u are either close to $E(u)$ or greater than $E(u)$.

Definition 4.9. *Let $\mathcal{U}$ be a Banach space with topology $\tau_\mathcal{U}$. The functional $E\colon \mathcal{U} \to \mathbb{R}_\infty$ is said to be* sequentially lower semi-continuous *with respect to $\tau_\mathcal{U}$ at $u \in \mathcal{U}$ if*

$$E(u) \leq \liminf_{j\to\infty} E(u_j)$$

for all sequences $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}$ with $u_j \to u$ in the topology $\tau_\mathcal{U}$ of $\mathcal{U}$.

Remark 4.3. For topologies that are not induced by a metric we have to differ between a topological property and its sequential version, e.g. continuous and sequentially continous. If the topology is induced by a metric, then these two are the same. However, for instance the weak and weak-$*$ topology are generally not induced by a metric.

Example 4.9. The functional $\|\cdot\|_1 : \ell^2 \to \mathbb{R}_\infty$ with

$$\|u\|_1 = \begin{cases} \sum_{j=1}^\infty |u_j| & \text{if } u \in \ell^1 \\ \infty & \text{else} \end{cases}$$

is lower semi-continuous with respect to ℓ^2.

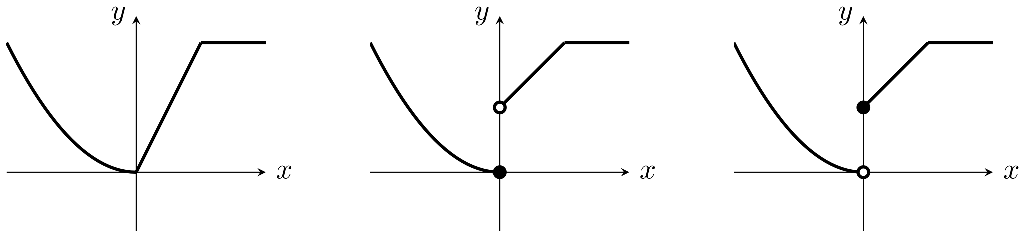

Figure 4.7: Visualisation of lower semi-continuity. The solid dot at a jump indicates the value that the function takes. The function on the left is continuous and thus lower semi-continuous. The functions in the middle and on the right are discontinuous. While the function in the middle is lower semi-continuous, the function on the right is not (due to the limit from the left at the discontinuity).

Proof. Let $\{u^j\}_{j\in\mathbb{N}} \subset \ell^2$ be a squence with $u^j \to u \in \ell^2$. As strong convergence implies weak converge, which implies convergence of the components (the functionals $\delta_i : \ell^2 \to \mathbb{R}, \delta_i(u) = u_i$ are linear and continuous), we have that for all $k \in \mathbb{N}$ that $u_k^j \to u_k$. The assertion follows then with Fatou's lemma

$$\|u\|_1 = \sum_{k=1}^{\infty} |u_k| = \sum_{k=1}^{\infty} \lim_{j\to\infty} |u_k^j| \leq \liminf_{j\to\infty} \sum_{k=1}^{\infty} |u_k^j| = \liminf_{j\to\infty} \|u^j\|_1 \,.$$

Note that it is note clear whether the right hand side is finite. The left hand side certainly is. □

Example 4.10. Let $\Omega \subset \mathbb{R}^n$ be open and bounded. Then, the total variation is lower semi-continuous with respect to L^1.

Proof. Recall that the total variation was defined by means of the test functions

$$\mathcal{D}(\Omega, \mathbb{R}^n) := \left\{ \varphi \in C_0^\infty(\Omega; \mathbb{R}^n) \,\middle|\, \|\varphi(x)\|_2 \leq 1 \right\}$$

as

$$\mathrm{TV}(u) = \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \int_\Omega \langle u(x), \operatorname{div} \varphi(x) \rangle \, dx \,.$$

Let $\{u_j\}_{j\in\mathbb{N}} \subset \mathrm{BV}(\Omega)$ be a sequence converging in $L^1(\Omega)$ with $u_j \to u$ in $L^1(\Omega)$. Then for any test function $\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)$

$$\int_\Omega [u(x) - u_j(x)] \operatorname{div} \varphi(x) dx \leq \underbrace{\int_\Omega |u(x) - u_j(x)| dx}_{\to 0} \sup_{x\in\Omega} |\operatorname{div} \varphi(x)| \to 0$$

and thus

$$\int_\Omega u(x) \operatorname{div} \varphi(x) dx = \lim_{j\to\infty} \int_\Omega u_j(x) \operatorname{div} \varphi(x) dx \leq \liminf_{j\to\infty} \mathrm{TV}(u_j) \,.$$

Taking the supremum over all test functions shows the assertion. Note that the right hand side may not be finite. □

This leads to the "direct method" or "fundamental theorem of optimisation".

Theorem 4.3 ("Direct method", David Hilbert, around 1900)**.** *Let $\mathcal{U}$ be a Banach space and $\tau_{\mathcal{U}}$ a topology (not necessarily the one induced by the norm) on $\mathcal{U}$ such that bounded sequences have $\tau_{\mathcal{U}}$-convergent subsequences. Let $E\colon \mathcal{U} \to \mathbb{R}_\infty$ be proper, bounded from below, coercive and sequentially lower semi-continuous with respect to $\tau_{\mathcal{U}}$. Then there exists a minimiser u^* of E.*

Proof. From Lemma 4.2 we know that $\inf_{u\in\mathcal{U}} E(u)$ is finite, minimising sequences exist and that they are bounded. Let $\{u_j\}_{j\in\mathbb{N}} \in \mathcal{U}$ be a minimising sequence. Thus, from the assumption on the topology $\tau_{\mathcal{U}}$ there exists a subsequence $\{u_{j_k}\}_{k\in\mathbb{N}}$ and $u^* \in \mathcal{U}$ with $u_{j_k} \overset{\tau_{\mathcal{U}}}{\to} u^*$ for $k \to \infty$. From the sequential lower semi-continuity of E we obtain

$$E(u^*) \le \liminf_{k\to\infty} E(u_{j_k}) = \lim_{j\to\infty} E(u_j) = \inf_{u\in\mathcal{U}} E(u) < \infty\,,$$

consequently u^* minimises E. □

The above theorem is very general but its conditions are hard to verify but the situation is a easier in *reflexive* Banach spaces.

Corollary 4.2. *Let $\mathcal{U}$ be a reflexive Banach space and $E\colon \mathcal{U} \to \mathbb{R}_\infty$ be a proper, bounded from below, coercive and sequentially lower semi-continuous functional with respect to the weak topology. Then there exists a minimiser of E.*

Proof. The statement follows from the direct method, Theorem 4.3, as in reflexive Banach spaces bounded sequences have weakly convergent subsequences, see Corollary 4.1. □

Remark 4.4. For convex functions on reflexive Banach spaces, the situation is even easier. It can be shown that a convex function is sequentially lower semi-continuous with respect to the weak topology if and only if it is lower semi-continuous with respect to the strong topology (see e.g. [3, p. 149] for Hilbert spaces).

Remark 4.5. It is easy to see that the key ingredient for the existence of minimisers is that bounded sequences have a convergent subsequence which is difficult to prove in practical situations. Another option is to change the space and consider a space in which $\mathcal{U}$ is compactly embedded in, i.e. the mapping $\mathcal{U} \to \mathcal{V}, u \mapsto u$ is compact. Then (by definition) every bounded sequence in $\mathcal{U}$ has a convergent subsequence in $\mathcal{V}$.

4.1.4 Uniqueness

Theorem 4.4. *Assume that the functional $E\colon \mathcal{U} \to \mathbb{R}_\infty$ has at least one minimiser and let E be strictly convex. Then the minimiser is unique.*

Proof. Let u and v be two minimisers of E. Assume that they are different, i.e. $u \neq v$. Then it follows from the minimising properties of u and v as well as the strict convexity of E that

$$E(u) \le E(\tfrac{1}{2}u + \tfrac{1}{2}v) < \frac{1}{2}E(u) + \frac{1}{2}\underbrace{E(v)}_{\le E(u)} \le E(u)$$

which is a contradiction. Thus, $u = v$ and the assertion is proven. □

Example 4.11. Convex (but not strictly convex) functions may have have more than one minimiser, examples include constant and trapezoidal functions, see Figure 4.8. On the other hand, convex (and even non-convex) functions may have a unique minimiser, see Figure 4.8.

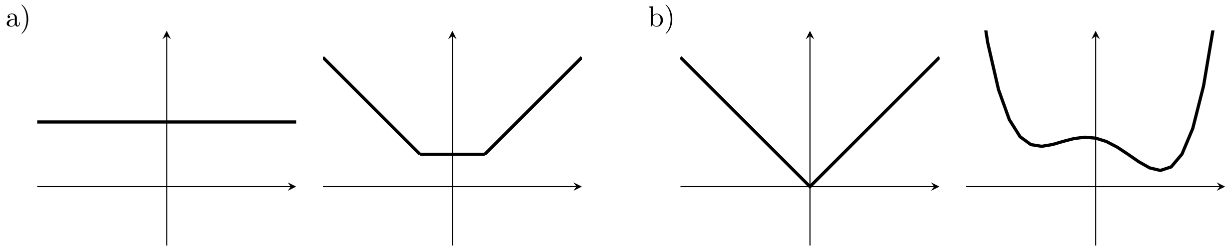

Figure 4.8: a) Convex functions may not have a unique minimiser. b) Neither strict convexity nor convexity is necessary for the uniqueness of a minimiser.

4.2 Variational regularisation

The aim of this section is to have a detailed look at the model $R_\alpha \colon \mathcal{V} \to \mathcal{U}$ with

$$R_\alpha f := u_\alpha := \arg\min_u \left\{ \Phi_{\alpha,f}(u) := \frac{1}{2}\|Ku - f\|_{\mathcal{V}}^2 + \alpha J(u) \right\}. \tag{4.4}$$

We will establish conditions on the spaces $\mathcal{U}, \mathcal{V}$, the functional J and the operator K under which the minimiser exists and is unique and therefore the mapping R_α is well-defined. We will analyse the continuity of the mapping R_α which means that the solution depends continuous on the data and thus can handle small variations due to noise. We also show that there are parameter choice rules that make R_α a convergent regularisation in a modified sense (that we will define later) and prove convergence rates under a source condition.

4.2.1 Existence and uniqueness

Existence

Lemma 4.3. *Let $\mathcal{U}$ be a Banach space and $\tau_{\mathcal{U}}$ a topology on it. Let $E \colon \mathcal{U} \to \mathbb{R}$ and $F \colon \mathcal{U} \to \mathbb{R}_\infty$ be proper functionals that are both sequentially lower semi-continuous with respect to the topology $\tau_{\mathcal{U}}$ and bounded from below. Then $E + F \colon \mathcal{U} \to \mathbb{R}_\infty$ is proper, sequentially lower semi-continuous with respect to the topology $\tau_{\mathcal{U}}$ and bounded from below.*

Proof. First of all, as F is proper, there exists $u \in \mathcal{U}$ such that $F(u) < \infty$ and as $E(u) < \infty$ it is clear that $(E + F)(u) < \infty$ which shows that $E + F$ is proper.

Second, for all $u \in \mathcal{U}$ we have from the boundedness from below of E and F that $E(u) \geq C_1$ and $F(u) \geq C_2$ and thus,

$$(E + F)(u) = E(u) + F(u) \geq C_1 + C_2 > -\infty\,.$$

Finally, let $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}$ be a sequence and $u \in \mathcal{U}$ with $u_j \to u$ in $\tau_{\mathcal{U}}$. Then by the sequential lower semi-continuity with respect to $\tau_{\mathcal{U}}$ we have that

$$\begin{aligned}(E + F)(u) &\leq \liminf_{j\to\infty} E(u_j) + \liminf_{j\to\infty} F(u_j) \\ &\leq \liminf_{j\to\infty} \left(E(u_j) + F(u_j)\right) = \liminf_{j\to\infty} (E + F)(u_j)\end{aligned}$$

which shows that $E + F$ is sequentially lower semi-continuous with respect to $\tau_{\mathcal{U}}$ and all assertions are proven. □

Lemma 4.4. *Let $\mathcal{U}$ be a Banach space and $E, F\colon \mathcal{U} \to \mathbb{R}_\infty$ be functionals. Let E be coercive and F be bounded from below, then $E + F$ is coercive.*

Proof. From the boundedness from below of F, there exists a constant $C \geq -\infty$ such that $F(u) > C$ for all $u \in \mathcal{U}$. Thus we see that

$$(E + F)(u) = E(u) + F(u) \geq E(u) + C \to \infty$$

as $\|u\|_{\mathcal{U}} \to \infty$ which proves that $E + F$ is coercive. □

In many situations of interest, the lemma above does not apply because the coercivity comes jointly from the data term and the prior as we will see in the following example.

Example 4.12. Let $\Omega \subset \mathbb{R}^n$ be a bounded. Consider the space $\mathcal{U} = \mathrm{BV}(\Omega)$ and the regularisation functional $J = \mathrm{TV}$. One can easily see (e.g. integration by parts) that $\mathrm{TV}(u+c) = \mathrm{TV}(u)$, for all $c \in \mathbb{R}$, $u \in \mathrm{BV}(\Omega)$ such that constant functions have zero total variation. Notice that this implies that J is not coercive on the whole space $\mathcal{U}$ as $u_j(x) = j/|\Omega|, |\Omega| := \int_\Omega 1\,dx$ defines a sequence such that $\|u_j\|_{L^1} = j$ and $\mathrm{TV}(u_j) = 0$. However, we can make use of a Poincaré–Wirtinger inequality for BV.

Proposition 4.1 ([6, p. 24])**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain (non-empty, open, connected and bounded) with Lipschitz boundary. There exists a constant $C > 0$ such that for all $u \in \mathrm{BV}(\Omega)$ the* Poincaré–Wirtinger *inequality is satisfied*

$$\|u - u_\Omega\|_{L^1} \leq C\,\mathrm{TV}(u)$$

where $u_\Omega := \frac{1}{|\Omega|}\int_\Omega u(x)dx$ is the mean-value of u over Ω.

Continuation of Example 4.12. Let Ω now fulfill the conditions of Proposition 4.1. Furthermore, let $p_0 \in \mathcal{U}^*$ with

$$\langle p_0, u\rangle = \frac{1}{|\Omega|}\int_\Omega u(x)dx$$

and denote the space of zero-mean functions by $\mathcal{U}_0 := \{u \in \mathcal{U} \,|\, u \in \mathcal{N}(p_0)\}$. By the Poincaré–Wirtinger inequality it is clear that the total variation is coercive on $\mathcal{U}_0$ and the data term has to make sure that the whole functional $\Phi_{\alpha,f}$ is coercive on the whole space $\mathcal{U}$. As we will see in the next, Lemma 4.5, the condition $1 \notin \mathcal{N}(K)$ is sufficient to guarantee coercivity in this scenario.

Lemma 4.5. *Let $\mathcal{U}, \mathcal{V}$ be Banach spaces, $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, $J\colon \mathcal{U} \to [0,\infty]$ and $f \in \mathcal{V}$. Let $p_0 \in \mathcal{U}^*, u_0 \in \mathcal{U}, \langle p_0, u_0\rangle = 1$,*

$$\mathcal{U}_0 := \{u \in \mathcal{U} \mid u \in \mathcal{N}(p_0)\}$$

so that $u_0 \notin \mathcal{N}(K)$ and J is coercive on $\mathcal{U}_0$ in the sense that

$$\|u - \langle p_0, u\rangle u_0\|_{\mathcal{U}} \to \infty \quad \textit{implies} \quad J(u) \to \infty\,.$$

Then the variational regularisation functional $\Phi_{\alpha,f}$ defined by (4.4) is coercive.

Proof. Any $u \in \mathcal{U}$ can be decomposed into $u = v + w$ where $v := u - \langle p_0, u\rangle u_0 \in \mathcal{U}_0$ and $w := \langle p_0, u\rangle u_0 \in \mathrm{span}(u_0)$. Now, let $\{u^j\}_{j\in\mathbb{N}} \subset \mathcal{U}$ be a sequence with $\|u^j\|_{\mathcal{U}} \to \infty$. On the one hand, if $\|v^j\|_{\mathcal{U}} \to \infty$, then by the coercivity of J on $\mathcal{U}_0$ and the boundedness from below of the data term, we have that $\Phi_{\alpha,f}(u^j) \to \infty$. On the other hand, if $\|v^j\|_{\mathcal{U}} < C$ for some $C > 0$, then from

$$\|u^j\|_{\mathcal{U}} \leq \|v^j\|_{\mathcal{U}} + \|\langle p_0, u^j\rangle u_0\|_{\mathcal{U}} < C + |\langle p_0, u^j\rangle| \|u_0\|_{\mathcal{U}}$$

it follows that $|\langle p_0, u^j\rangle| \to \infty$. Therefore,

$$\begin{aligned} \|Ku^j - f\|_{\mathcal{V}} &= \|K(v^j + w^j) - f\|_{\mathcal{V}} \\ &\geq \|Kw^j\|_{\mathcal{V}} - \|f - Kv^j\|_{\mathcal{V}} \\ &> \underbrace{\|Ku_0\|_{\mathcal{V}}}_{>0} \underbrace{|\langle p_0, u^j\rangle|}_{\to\infty} \underbrace{-\|f\|_{\mathcal{V}} - \|K\|C}_{\text{bounded from below}} \to \infty \end{aligned}$$

and thus $\Phi_{\alpha,f}(u^j) \to \infty$ as the regularisation functional J is bounded from below. □

Remark 4.6. A natural question here is whether the coercivity can also come completely from the data term $\frac{1}{2}\|Ku - f\|_{\mathcal{V}}^2$. On the one hand if K is not injective, then the kernel is non-trivial, thus we cannot expect the data term to be coercive. On the other hand, even if K was injective we cannot expect coercivity. Assume that the data term was coercive, $\mathcal{U}$ a Hilbert space, the topologies $\tau_{\mathcal{U}}$ and $\tau_{\mathcal{V}}$ the weak topologies on $\mathcal{U}$ and $\mathcal{V}$ and $f \in \overline{\mathcal{R}(K)} \setminus \mathcal{R}(K)$. Then we can apply the direct method on the data term only and we get the existence of a minimiser which is by definition a least squares solution, see Chapter 2. This is a contradiction to Lemma 2.2.

The remark will be illustrated by the following example.

Example 4.13. Let us consider the Example 2.1 again where the operator was $K\colon \ell^2 \to \ell^2, (Ku)_j := u_j/j$ and the data $f \in \ell^2$ with $f_j := 1/j$. Then the $\{u^k\}_{k\in\mathbb{N}} \subset \ell^2$ with

$$u_j^k := \begin{cases} 1 & j \leq k \\ 0 & \text{else} \end{cases}$$

defines a sequence $\{Ku^k\}_{k\in\mathbb{N}}$ which is in the range of K and $Ku^k \to f$ in ℓ^2 but $f \notin \mathcal{R}(K)$.

In addition to the observations in Example 2.1, we see K is injective and that $\{u^k\}_{k\in\mathbb{N}}$ is a minimising sequence of the data term, i.e. $\|Ku^k - f\|_{\ell^2}^2 \to 0$ but there is no minimiser as $f \notin \mathcal{R}(K)$. However, as $\|u^k\|_{\ell^2} = k$ the sequence $\{u^k\}_{k\in\mathbb{N}}$ is unbounded and thus $u \mapsto \|Ku - f\|_{\ell^2}^2$ cannot be coercive.

Lemma 4.6. *Let $\mathcal{U}$ and $\mathcal{V}$ be Banach spaces with topologies $\tau_{\mathcal{U}}$ and $\tau_{\mathcal{V}}$. Moreover, let the norm on $\mathcal{V}$ be sequentially lower semi-continuous with respect to $\tau_{\mathcal{V}}$, the operator $K\colon \mathcal{U} \to \mathcal{V}$ be sequentially continuous with respect to the topologies $\tau_{\mathcal{U}}$ and $\tau_{\mathcal{V}}$ and let $\{f_j\}_{j\in\mathbb{N}} \subset \mathcal{V}$ be convergent in $\tau_{\mathcal{V}}$ with $f_j \to f \in \mathcal{V}$. Then for any $\tau_{\mathcal{U}}$-convergent sequence $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}$ with $u_j \to u \in \mathcal{U}$, we have*

$$\frac{1}{2}\|Ku - f\|_{\mathcal{V}}^2 \leq \liminf_{j\to\infty} \frac{1}{2}\|Ku_j - f_j\|_{\mathcal{V}}^2 .$$

In particular, if $f_j = f$, then $D\colon \mathcal{U} \to \mathbb{R}, u \mapsto \frac{1}{2}\|Ku - f\|_{\mathcal{V}}^2$ is sequentially lower semi-continuous with respect to $\tau_{\mathcal{U}}$.

Proof. Let $\{u_j\}_{j\in\mathbb{N}}$ be a $\tau_{\mathcal{U}}$-convergent sequence and denote its limit by $u \in \mathcal{U}$, i.e. $u_j \to u$ in $\tau_{\mathcal{U}}$. Because K is continuous with respect to $\tau_{\mathcal{U}}$ and $\tau_{\mathcal{V}}$ we have that $Ku_j \to Ku$ in $\tau_{\mathcal{V}}$ and thus $Ku_j - f_j \to Ku - f$ in $\tau_{\mathcal{V}}$. Thus, the assertion follows from the sequential lower semi-continuity of the norm with respect to $\tau_{\mathcal{V}}$. □

Remark 4.7. If the topologies $\tau_\mathcal{U}$ and $\tau_\mathcal{V}$ are the weak topologies, then the situation is much simpler as continuity in the strong topologies implies continuity in the weak topologies. Thus the assumptions of Lemma 4.6 are met if K is continuous.

Now we are in a position to state sufficient assumptions for the existence of minimisers.

Assumption 4.1. *Sufficient assumptions for the existence of minimisers of $\Phi_{\alpha,f}$ are:*

(a) The Banach space $\mathcal{U}$ and Hilbert space $\mathcal{V}$ are associated with the topologies $\tau_\mathcal{U}$ and $\tau_\mathcal{V}$. The pair $(\mathcal{U}, \tau_\mathcal{U})$ has the property that bounded sequences have $\tau_\mathcal{U}$-convergent subsequences. Moreover, the norm on $\mathcal{V}$ is sequentially lower semi-continuous with respect to $\tau_\mathcal{V}$.

(b) The operator $K\colon \mathcal{U} \to \mathcal{V}$ is linear and sequentially continuous with respect to the topologies $\tau_\mathcal{U}$ and $\tau_\mathcal{V}$.

(c) The functional $J\colon \mathcal{U} \to [0, \infty]$ is proper and sequentially lower semi-continuous with respect to $\tau_\mathcal{U}$.

(d) Either J is coercive or the pair (K, J) fulfill the assumptions of Lemma 4.5.

Theorem 4.5. *Let the Assumptions 4.1 hold and let $f \in \mathcal{V}, \alpha > 0$. Then the variational regularisation functional $\Phi_{\alpha,f}$ defined by (4.4) has a minimiser.*

Proof. It follows from the assumptions by Lemmata 4.3 and 4.6 that $\Phi_{\alpha,f}$ is proper, sequentially lower semi-continuous with respect to $\tau_\mathcal{U}$ and bounded from below. Moreover, from Lemmata 4.4 or 4.5 (depending on assumption 4.1 (d)) $\Phi_{\alpha,f}$ is coercive. Then from the direct method, Theorem 4.3, it follows that there exists a minimiser. ☐

Uniqueness

Lemma 4.7. *Let $\mathcal{U}$ be a Banach space and and $\mathcal{V}$ a Hilbert space. Furthermore, let $K \in \mathcal{L}(\mathcal{U}, \mathcal{V}), f \in \mathcal{V}$ and $D\colon \mathcal{V} \to \mathbb{R}_\infty$ be defined as $D(u) := \frac{1}{2}\|Ku - f\|_\mathcal{V}^2$. Then E is convex. Furthermore, D is strictly convex if and only if K is injective.*

Proof. The proof is left as an exercise. ☐

Remark 4.8. The lemma is not true if in general if $\mathcal{U}$ is a Banach space, consider for instance the examples ℓ^1 and ℓ^∞.

Example 4.14. Let $\mathcal{U}$ be continuously embedded into the Hilbert space $\mathcal{Z}$ (in symbols $\mathcal{U} \hookrightarrow \mathcal{Z}$), i.e. there exists a constant $C > 0$ such that for all $u \in \mathcal{U}$ there is $\|u\|_\mathcal{Z} \leq C\|u\|_\mathcal{U}$. Furthermore, let $\beta > 0$. Then the functional $\Phi_{\alpha,f} + \frac{\beta}{2}\|\cdot\|_\mathcal{Z}^2$ is always strictly convex independent of K.

Consider the product space $\mathcal{V} \times \mathcal{Z}$ which is a Hilbert space with the inner product

$$\langle (v_1, z_1), (v_2, z_2)\rangle_{\mathcal{V}\times\mathcal{Z}} := \langle v_1, v_2\rangle_\mathcal{V} + \langle z_1, z_2\rangle_\mathcal{Z} \,.$$

Then we can rewrite $\frac{1}{2}\|Ku - f\|_\mathcal{V}^2 + \frac{\beta}{2}\|u\|_\mathcal{Z}^2$ as

$$\frac{1}{2}\left\|\begin{pmatrix} K \\ \sqrt{\beta} I\end{pmatrix} u - \begin{pmatrix} f \\ 0\end{pmatrix}\right\|_{\mathcal{V}\times\mathcal{Z}}^2 = \frac{1}{2}\|\tilde{K}u - \tilde{f}\|_\mathcal{V}^2$$

where the modified operator $\tilde{K}$ is injective. Therefore, adding the term $\frac{\beta}{2}\|u\|_\mathcal{Z}^2$ can be seen as a regularisation of the linear operator K directly.

Theorem 4.6. *Let the Assumptions 4.1 are met and let J be convex. Moreover, let either K be injective or J be strictly convex. Then for any $f \in \mathcal{U}$ and $\alpha > 0$ the variational regularisation model is well-defined in the sense that there exists a unique mimimiser of the functional $\Phi_{\alpha,f}$ defined by (4.4).*

Proof. Existence follows immediately from Theorem 4.5. For the uniqueness, notice that both $\frac{1}{2}\|Ku - f\|_{\mathcal{V}}^2$ and αJ are convex and either of them is strictly convex by assumption, see Lemma 4.7. Thus by Lemma 4.1 the whole functional $\Phi_{\alpha,f}$ is strictly convex and therefore the minimiser is unique, see Theorem 4.4. □

Example 4.15. Let $\alpha > 0, \eta \geq 0, K \in \mathcal{L}(\ell^2, \ell^2)$ and consider the elastic net variational regularisation model $\frac{1}{2}\|Ku - f\|_{\ell^2}^2 + \alpha J(u)$ with

$$J(u) = \begin{cases} \eta\|u\|_1 + \frac{1}{2}\|u\|_2^2 & \text{if } u \in \ell^1 \\ \infty & \text{else} \end{cases} .$$

As ℓ^2 is a Hilbert space we will employ Corollary 4.2. We choose the topologies $\tau_{\mathcal{U}}, \tau_{\mathcal{V}}$ to be the weak topology on ℓ^2. By Remark 4.7 the Assumptions 4.1 (a) and (b) are fulfilled and we can show lower semi-continuity in the strong topology rather than the weak one. It is easy to see that the prior J is strictly convex, proper and coercive. It remains to show that J is lower semi-continuous with respect to ℓ^2. The squared ℓ^2-norm is continuous, thus lower semi-continuous and the lower semi-continuity of the ℓ^1-norm has been proven in Example 4.9 such that the whole prior is lower semi-continuous by Lemma 4.3.

For the example of the total variation we need to have some knowledge about compact embeddings of BV.

Theorem 4.7 (Rellich-Kandrachov, [1, p. 168])**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with Lipschitz boundary and either*

$$\begin{aligned} & n > mp \quad \textit{and} \quad p^* := np/(n - mp) \\ \textit{or} \quad & n \leq mp \quad \textit{and} \quad p^* := \infty . \end{aligned}$$

Then the embedding $H^{m,p}(\Omega) \to L^q(\Omega)$ is continous if $1 \leq q \leq p^*$ and compact if $1 \leq q < p^*$.

Due to approximations of $u \in \mathrm{BV}(\Omega)$ by smooth functions this gives us compactness.

Corollary 4.3 ([6, p. 17])**.** *Let $\Omega \subset \mathbb{R}^n$ be bounded with Lipschitz boundary, and let $p^* := n/(n-1)$ if $n > 1$ or $p^* := \infty$ else. Then the embedding $\mathrm{BV}(\Omega) \to L^q(\Omega)$ is continous if $1 \leq q \leq p^*$ and compact if $1 \leq q < p^*$.*

Example 4.16. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with Lipschitz boundary and let $1 \leq q \leq n/(n-1)$. Let $\alpha > 0, K \in \mathcal{L}(L^q(\Omega), L^2(\Omega))$ and $K1 \notin \mathcal{N}(K)$ be injective and consider the TV-variational regularisation model $\Phi_{\alpha,f} \colon \mathrm{BV}(\Omega) \to \mathbb{R}_\infty$, with

$$\Phi_{\alpha,f}(u) = \frac{1}{2}\|Ku - f\|_{L^2}^2 + \alpha \, \mathrm{TV}(u) .$$

This time we are neither in a Hilbert nor reflexive Banach space setting but from Corollary 4.3 we see that $\mathrm{BV}(\Omega)$ is compactly embedded in $L^1(\Omega)$. Thus every sequence bounded in $\mathrm{BV}(\Omega)$ has a convergent subsequence in $L^1(\Omega)$. Let $\tau_{\mathcal{U}}$ and $\tau_{\mathcal{V}}$ be the topologies induced by the L^q-norm, respectively L^2-norm. It is clear that the assumptions on the spaces and topologies are met. The lower semi-continuity of TV with respect to L^1 was shown in Example 4.10. Moreover, it can be shown that TV is proper and convex. From Example 4.12 and $1 \notin \mathcal{N}(K)$ it follows that $\Phi_{\alpha,f}$ is coercive. Thus, a minimiser exists. The injectivity of the operator K guarantees the uniqueness of the minimiser.

4.2.2 Continuity

We have seen that under some assumptions the variational regularisation R_α is well-defined (solutions exists and are unique). In this section we show that variational regularisation is continuous with respect to the data, i.e. small variations in the data do not lead to arbitrary large distortions in the solution. To establish the main result we have to prove auxiliary lemmata.

Lemma 4.8. *Let $\mathcal{V}$ be a normed space. For all $f, g \in \mathcal{V}$ there is*

$$\|f+g\|_{\mathcal{V}}^2 \leq 2\|f\|_{\mathcal{V}}^2 + 2\|g\|_{\mathcal{V}}^2 \,.$$

Proof. For any $f, g \in \mathcal{V}$ we have with $2ab \leq a^2 + b^2, a, b \in \mathbb{R}$ so that

$$\begin{aligned} \|f+g\|_{\mathcal{V}}^2 &\leq \Big(\|f\|_{\mathcal{V}} + \|g\|_{\mathcal{V}}\Big)^2 \\ &= \|f\|_{\mathcal{V}}^2 + 2\|f\|_{\mathcal{V}}\|g\|_{\mathcal{V}} + \|g\|_{\mathcal{V}}^2 \leq 2\|f\|_{\mathcal{V}}^2 + 2\|g\|_{\mathcal{V}}^2 \,. \end{aligned}$$

□

Lemma 4.9. *Let $\mathcal{U}, \mathcal{V}$ be Banach spaces. For all $u \in \mathcal{U}$ and $f, g \in \mathcal{V}$ there is*

$$\Phi_{\alpha,f}(u) \leq 2\Phi_{\alpha,g}(u) + \|f-g\|_{\mathcal{V}}^2 \,.$$

Proof. Using Lemma 4.8 and $J(u) \geq 0$, we have

$$\begin{aligned} \Phi_{\alpha,f}(u) = \frac{1}{2}\|Ku-f\|_{\mathcal{V}}^2 + \alpha J(u) &\leq \|Ku-g\|_{\mathcal{V}}^2 + \|g-f\|_{\mathcal{V}}^2 + 2\alpha J(u) \\ &= 2\left(\frac{1}{2}\|Ku-g\|_{\mathcal{U}}^2 + \alpha J(u)\right) + \|f-g\|_{\mathcal{V}}^2 \\ &= 2\Phi_{\alpha,g}(u) + \|f-g\|_{\mathcal{V}}^2 \,. \end{aligned}$$

□

Theorem 4.8 (Continuity)**.** *Assume the setting of Theorem 4.6 that guarantees the existance and uniqueness of minimisiers of $\Phi_{\alpha,f}(u) := \frac{1}{2}\|Ku-f\|_{\mathcal{V}}^2 + \alpha J(u)$ for any $f \in \mathcal{V}$ and $\alpha > 0$. Moreover, let the topology $\tau_{\mathcal{V}}$ on $\mathcal{V}$ be weaker than the norm topology in the sense that convergence in norm implies convergence in $\tau_{\mathcal{V}}$. Then, the mapping $R_\alpha \colon \mathcal{V} \to \mathcal{U}, R_\alpha f := \arg\min_{u\in\mathcal{U}} \Phi_{\alpha,f}(u)$ is sequentially strong-$\tau_{\mathcal{U}}$ continuous, i.e. for all sequences $\{f_j\}_{j\in\mathbb{N}} \subset \mathcal{V}$ with $f_j \to f$ we have*

$$R_\alpha f_j \overset{\tau_{\mathcal{U}}}{\to} R_\alpha f \,.$$

Moreover, we have that $J(R_\alpha f_j) \to J(R_\alpha f)$.

Proof. Let $\{f_j\}_{j\in\mathbb{N}} \subset \mathcal{V}$ be a convergent sequence with $f_j \to f$ and let $u_j := R_\alpha f_j$ be the minimiser of Φ_{α,f_j} and $u := R_\alpha f$ the minimiser of $\Phi_{\alpha,f}$.

We first show that $\{\Phi_{\alpha,f}(u_j)\}_{j\in\mathbb{N}} \subset \mathbb{R}$ is bounded. To see this, as J is proper, there exists $\tilde{u} \in \mathcal{U}$ such that $J(\tilde{u}) < \infty$ and we denote $C := 2\|K\tilde{u}\|_{\mathcal{V}}^2 + 2\alpha J(\tilde{u})$. With Lemmata 4.8 and 4.9 and the minimising property of u_j we have that

$$
\begin{aligned}
\Phi_{\alpha,f}(u_j) &\leq 2\underbrace{\Phi_{\alpha,f_j}(u_j)}_{\leq \Phi_{\alpha,f_j}(\tilde{u})} + \|f - f_j\|_{\mathcal{V}}^2 \\
&\leq \|K\tilde{u} - f_j\|_{\mathcal{V}}^2 + 2\alpha J(\tilde{u}) + \|f - f_j\|_{\mathcal{V}}^2 \\
&\leq 2\|K\tilde{u}\|_{\mathcal{V}}^2 + 2\|f_j\|_{\mathcal{V}}^2 + 2\alpha J(\tilde{u}) + \|f - f_j\|_{\mathcal{V}}^2 = 2\|f_j\|_{\mathcal{U}}^2 + \|f - f_j\|_{\mathcal{V}}^2 + C\,.
\end{aligned}
$$

As f_j converges to f, there exists a $j_0 \in \mathbb{N}$ such that for all $j > j_0$ there is

$$
\begin{aligned}
\Phi_{\alpha,f}(u_j) &\leq 2\underbrace{\|f_j\|_{\mathcal{V}}^2}_{\|f\|_{\mathcal{V}}^2+1} + \underbrace{\|f - f_j\|_{\mathcal{V}}^2}_{\leq 1} + C \\
&\leq 2\|f\|_{\mathcal{V}}^2 + C + 3 < \infty\,.
\end{aligned}
$$

By the coercivity of $\Phi_{\alpha,f}$ we know that the sequence $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}$ is bounded, see Remark 4.2. Thus there exist $\tau_{\mathcal{U}}$-convergent subsequences and let $\{u_{j_k}\}_{k\in\mathbb{N}} \subset \mathcal{U}$ be any one of those. We denote its limit by $\hat{u} \in \mathcal{U}$, i.e. $u_{j_k} \to \hat{u}$ in $\tau_{\mathcal{U}}$.

From Lemma 4.6 and the sequential lower semi-continuity of J we have that

$$
\frac{1}{2}\|K\hat{u} - f\|_{\mathcal{V}}^2 \leq \liminf_{k\to\infty} \frac{1}{2}\|Ku_{j_k} - f_{j_k}\|_{\mathcal{V}}^2 \quad \text{and} \quad J(\hat{u}) \leq \liminf_{k\to\infty} J(u_{j_k})\,. \tag{4.5}
$$

Thus, we conclude with (4.5) that

$$
\begin{aligned}
\Phi_{\alpha,f}(\hat{u}) &= \frac{1}{2}\|K\hat{u} - f\|_{\mathcal{V}}^2 + \alpha J(\hat{u}) \\
&\leq \liminf_{k\to\infty} \frac{1}{2}\|Ku_{j_k} - f_{j_k}\|_{\mathcal{V}}^2 + \alpha \liminf_{k\to\infty} J(u_{j_k}) \\
&\leq \liminf_{k\to\infty} \left(\frac{1}{2}\|Ku_{j_k} - f_{j_k}\|_{\mathcal{V}}^2 + \alpha J(u_{j_k})\right) = \liminf_{k\to\infty} \Phi_{\alpha,f_{j_k}}(u_{j_k}) \\
&\leq \liminf_{k\to\infty} \Phi_{\alpha,f_{j_k}}(u) = \lim_{k\to\infty} \Phi_{\alpha,f_{j_k}}(u) = \Phi_{\alpha,f}(u)\,.
\end{aligned}
\tag{4.6}
$$

Thus, as the minimiser of $\Phi_{\alpha,f}$ is unique, we have that $\hat{u} = u$. Repeating the same arguments as above for any subsequence of $\{u_j\}_{j\in\mathbb{N}}$ instead of $\{u_j\}_{j\in\mathbb{N}}$, we see that every subsequence has a convergent subsequence that converges to u in $\tau_{\mathcal{U}}$. Thus, $\{u_j\}_{j\in\mathbb{N}}$ is convergent in $\tau_{\mathcal{U}}$ and we have $u_j \to u$ in $\tau_{\mathcal{U}}$ and the first assertion is proven.

Moreover, also from Equation (4.6) with the convergence of $\{u_j\}_{j\in\mathbb{N}}$ it follows that

$$
\lim_{j\to\infty} \Phi_{\alpha,f_j}(u_j) = \Phi_{\alpha,f}(u)\,.
$$

Thus with the sequential lower semi-continuity of J with respect to $\tau_{\mathcal{U}}$ we arrive at the second assertion as

$$\begin{aligned}\limsup_{j\to\infty} \alpha J(u_j) &= \limsup_{j\to\infty} \left(\frac{1}{2}\|Ku_j - f_j\|_{\mathcal{V}}^2 + \alpha J(u_j) - \frac{1}{2}\|Ku_j - f_j\|_{\mathcal{V}}^2\right) \\ &\leq \underbrace{\limsup_{j\to\infty} \left(\frac{1}{2}\|Ku_j - f_j\|_{\mathcal{V}}^2 + \alpha J(u_j)\right)}_{=\lim_{j\to\infty}\Phi_{\alpha,f_j}(u_j)=\Phi_{\alpha,f}(u)} + \limsup_{j\to\infty} \left(-\frac{1}{2}\|Ku_j - f_j\|_{\mathcal{V}}^2\right) \\ &= \frac{1}{2}\|Ku - f\|_{\mathcal{V}}^2 + \alpha J(u) \underbrace{- \liminf_{j\to\infty} \frac{1}{2}\|Ku_j - f_j\|_{\mathcal{V}}^2}_{\leq -\frac{1}{2}\|Ku-f\|_{\mathcal{V}}^2 \text{ by (4.5)}} \\ &\leq \alpha J(u) \leq \liminf_{j\to\infty} \alpha J(u_j)\,.\end{aligned}$$

□

Remark 4.9. In the theorem above we could only prove convergence in $\tau_{\mathcal{U}}$. If J statisfies the *Radon-Riesz property* with respect to the topology $\tau_{\mathcal{U}}$, i.e. $u_j \to u$ in $\tau_{\mathcal{U}}$ and $J(u_j) \to J(u)$ imply $\|u_j - u\|_{\mathcal{U}} \to 0$, then the convergence is in the norm topology. An example of a functional satisfying the Radon-Riesz property is $\|\cdot\|_{L^p}^p$ / $\|\cdot\|_{\ell^p}^p$ with $1 < p < \infty$ if the underlying space is L^p / ℓ^p and $\tau_{\mathcal{U}}$ is the weak topology.

4.2.3 Convergent regularisation

Note that variational regularisation for general J is not necessarily a regularisation in the sense of Definition 3.1, as we cannot expect $R_\alpha f = u_\alpha = \arg\min_{u\in\mathcal{U}} \Phi_{\alpha,f}(u) \to u^\dagger$ for $\alpha \to 0$ where $u^\dagger$ is the minimal norm solution. However, we can generalise Definition 2.1 of a minimal norm solution (and a least squares solution) to justify calling R_α a regularisation.

Definition 4.10. *Let $\mathcal{U}$ and $\mathcal{V}$ be Banach spaces and $f \in \mathcal{V}$. We call $u \in \mathcal{U}$ a* least squares solution *of the inverse problem* (1.1)*, if*

$$u \in \arg\min_{v\in\mathcal{U}} \|Kv - f\|_{\mathcal{V}} \tag{4.7}$$

As in the case of Hilbert spaces, we denote by $\mathbb{L}$ the set of all least squares solutions (it might be empty). Furthermore, we call $u^\dagger \in \mathcal{U}$ a J-minimising solution *of the inverse problem* (1.1)*, if*

$$u^\dagger \in \arg\min_{v\in\mathbb{L}} J(v)\,. \tag{4.8}$$

Remark 4.10. If $\mathcal{V}$ is a Hilbert space (as in our setting in Assumption 4.1), then most of the statements from Chapter 2 about least squares solutions still hold. In particular Lemma 2.2, which states that $\mathbb{L} \neq \emptyset$ if and only if $f \in \mathcal{R}(K) \oplus \mathcal{R}(K)^\perp$. However, the minimal norm solution (J-minimising solution for $J = \|\cdot\|_{\mathcal{U}}$) may not be unique anymore.

Lemma 4.10. *Let $\mathcal{U}$ and $\mathcal{V}$ be Banach spaces, $f \in \mathcal{V}$ and $K\colon \mathcal{U} \to \mathcal{V}$ be linear. Then the set of least squares solutions $\mathbb{L}$ is convex. Moreover, $\mathbb{L}$ is at most a singleton if K is injective.*

Proof. Let $u, v \in \mathbb{L}, u \neq v$ and $\lambda \in (0,1)$. Then for any $w \in \mathcal{U}$ we have

$$\begin{aligned}\|K(\lambda u + (1-\lambda)v) - f\|_{\mathcal{V}} &= \|\lambda(Ku - f) + (1-\lambda)(Kv - f)\|_{\mathcal{V}} \\ &\leq \lambda\|Ku - f\|_{\mathcal{V}} + (1-\lambda)\|Kv - f\|_{\mathcal{V}} \\ &\leq \lambda\|Kw - f\|_{\mathcal{V}} + (1-\lambda)\|Kw - f\|_{\mathcal{V}} = \|Kw - f\|_{\mathcal{V}}\end{aligned}$$

which shows that $\lambda u + (1-\lambda)v \in \mathbb{L}$ and thus $\mathbb{L}$ is convex.

For the second part, assume that K is injective and that least squares solutions exist which are equivalently characterised by

$$u \in \arg\min_{v\in\mathcal{U}} \left\{ \Psi(v) := \frac{1}{2}\|Kv - f\|_{\mathcal{V}}^2 \right\}. \tag{4.9}$$

From Lemma 4.7 we know that Ψ is strictly convex and thus by Theorem 4.4 the minimiser is unique. □

Proposition 4.2. *Let the assumptions of Theorem 4.6 hold and $f \in \mathcal{R}(K) \oplus \mathcal{R}(K)^\perp$. Then a J-minimising solution exists and is unique.*

Proof. The condition $f \in \mathcal{R}(K) \oplus \mathcal{R}(K)^\perp$ guarantees the existence of least squares solutions, i.e. $\mathbb{L} \neq \emptyset$. For the existence of J-minimising solutions via the direct method, we see that only the coercivity on $\mathbb{L}$ may not be guaranteed by the assumptions. If J is coercive, then it is obviously also coercive on $\mathbb{L}$. If J is only coercive on $\mathcal{U}_0$, see Lemma 4.5 for a definition, then a similar calculation as in the proof of Lemma 4.5 shows that for any sequence $\{u_j\}_{j\in\mathbb{N}} \subset \mathbb{L}$ we have

$$\begin{aligned}
\|f\|_{\mathcal{V}} = \|K0 - f\|_{\mathcal{V}} &\geq \|Ku^j - f\|_{\mathcal{V}} \\
&= \|K(v^j + w^j) - f\|_{\mathcal{V}} \\
&\geq \|Kw^j\|_{\mathcal{V}} - \|f - Kv^j\|_{\mathcal{V}} \\
&> \|Ku_0\|_{\mathcal{V}} |\langle p_0, u^j\rangle| - \|f\|_{\mathcal{V}} - \|K\|\|v^j\|_{\mathcal{V}} \,,
\end{aligned}$$

thus if $\|v^j\|_{\mathcal{V}}$ is bounded, then $|\langle p_0, u^j\rangle|$ is also bounded. Therefore for $\{u_j\}_{j\in\mathbb{N}} \subset \mathbb{L}$ with $\|u_j\|_{\mathcal{U}} \to \infty$ we have that $\|u_j - \langle p_0, u_j\rangle\|_{\mathcal{U}} \to \infty$ and thus $J(u_j) \to \infty$ by the coercivity on $\mathcal{U}_0$.

For the uniqueness, either J is strictly convex (and thus a minimiser is unique) or K is injective and only one least squares solution exists. □

Definition 4.11 (Regularisation)**.** *Let $\mathcal{U}, \mathcal{V}$ be Banach spaces, $\tau_{\mathcal{U}}$ a topology on $\mathcal{U}$, $f \in \mathcal{V}$ and $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$. Moreover, let $u^\dagger$ be the J-minimising solution (assuming it exists and is unique). We call the family of operators $\{R_\alpha\}_{\alpha>0}, R_\alpha \colon \mathcal{V} \to \mathcal{U}$ a* regularisation *(with respect to $\tau_{\mathcal{U}}$) of the inverse problem* (1.1)*, if R_α is sequentially strong-$\tau_{\mathcal{U}}$ continuous for all $\alpha > 0$ and*

$$R_\alpha f \overset{\tau_{\mathcal{U}}}{\to} u^\dagger \quad \text{as} \quad \alpha \to 0\,.$$

Theorem 4.9 (Convergent regularisation)**.** *Let the assumptions of Theorem 4.6 hold and assume (for simplicity) that the clean data is in the range, i.e. $f \in \mathcal{R}(K)$, thus the J-minimising solution $u^\dagger$ exists and is unique. Moreover, assume that the topology $\tau_{\mathcal{V}}$ is weaker than the norm topology on $\mathcal{V}$. Let $\alpha : (0,\infty) \to (0,\infty)$ be a parameter choice rule with*

$$\alpha(\delta) \to 0, \quad \text{and} \quad \frac{\delta^2}{\alpha(\delta)} \to 0 \quad \text{as} \quad \delta \to 0\,.$$

Let $\{\delta_j\}_{j\in\mathbb{N}} \subset [0,\infty)$ be a sequence of noise levels with $\delta_j \to 0$ and $\{f_j\}_{j\in\mathbb{N}} \subset \mathcal{V}$ be a sequence of noisy observations with $\|f - f_j\|_{\mathcal{V}} \leq \delta_j$. Set $\alpha_j := \alpha(\delta_j)$ and let $\{u_j\}_{j\in\mathbb{N}}$ be the sequence of minimisers of Φ_{α_j, f_j}, i.e. $u_{\alpha_j} := R_{\alpha_j} f_j$.

Then $\{u_j\}_{j\in\mathbb{N}}$ converges in $\tau_{\mathcal{U}}$ and $u_j \overset{\tau_{\mathcal{U}}}{\to} u^\dagger$. Moreover, we have $J(u_j) \to J(u^\dagger)$. In particular, (as it implies pointwise convergence for exact data) R_α is a regularisation.

Proof. From the definition of u_j (minimising property) it follows that

$$\begin{aligned} 0 &\leq \frac{1}{2}\|Ku_j - f_j\|_{\mathcal{V}}^2 + \alpha_j J(u_j) \\ &\leq \frac{1}{2}\|Ku^\dagger - f_j\|_{\mathcal{V}}^2 + \alpha_j J(u^\dagger) \leq \frac{\delta_j^2}{2} + \alpha_j J(u^\dagger) \to 0 \end{aligned} \tag{4.10}$$

as $\delta_j, \alpha_j \to 0$. Thus $\lim_{j\to\infty} \|Ku_j - f_j\|_{\mathcal{V}} = 0$, and then

$$\|Ku_j - f\|_{\mathcal{V}} \leq \|Ku_j - f_j\|_{\mathcal{V}} + \|f_j - f\|_{\mathcal{V}} \leq \|Ku_j - f_j\|_{\mathcal{V}} + \delta_j \to 0\,. \tag{4.11}$$

Similarly, we see from (4.10) that

$$\limsup_{j\to\infty} J(u_j) \leq \limsup_{j\to\infty} \frac{\delta_j^2}{2\alpha_j} + J(u^\dagger) = J(u^\dagger)\,. \tag{4.12}$$

Let $\alpha^+ := \max_{j\in\mathbb{N}} \alpha_j$ be the largest regularisation parameter (which exists as $\alpha_j \to 0$), then

$$\begin{aligned} \limsup_{j\to\infty} \Phi_{\alpha^+,f}(u_j) &= \limsup_{j\to\infty} \left(\frac{1}{2}\|Ku_j - f\|_{\mathcal{V}}^2 + \alpha^+ J(u_j)\right) \\ &\leq \underbrace{\limsup_{j\to\infty} \frac{1}{2}\|Ku_j - f\|_{\mathcal{V}}^2}_{=0} + \underbrace{\limsup_{j\to\infty} \alpha^+ J(u_j)}_{\leq \alpha^+ J(u^\dagger)} \leq \alpha^+ J(u^\dagger) =: C < \infty \end{aligned}$$

This shows that there exists a $j_0 \in \mathbb{N}$ such that for all $j \geq j_0$ we have that $\Phi_{\alpha^+,f}(u_j) \leq C + 1$. From the coercivity and the assumptions on the topology, it follows that $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{U}$ has a $\tau_{\mathcal{U}}$-convergent subsequence $\{u_{j_k}\}_{k\in\mathbb{N}}$ with $u_{j_k} \to \hat{u}$ with respect to $\tau_{\mathcal{U}}$. By the continuity of K with respect to $\tau_{\mathcal{U}}$ and $\tau_{\mathcal{V}}$ we have that $Ku_{j_k} \to K\hat{u}$ with respect to $\tau_{\mathcal{V}}$ and with (4.11) and the assumptions on $\tau_{\mathcal{V}}$ it follows that $Ku_{j_k} \to f$ with respect to $\tau_{\mathcal{V}}$, thus $K\hat{u} = f$. From the sequential lower semi-continuity of J and (4.12) we have that

$$J(\hat{u}) \leq \liminf_{k\to\infty} J(u_{j_k}) \leq \limsup_{k\to\infty} J(u_{j_k}) \leq J(u^\dagger)\,. \tag{4.13}$$

Thus, $\hat{u}$ is a J-minimising solution, which implies by its uniqueness that $\hat{u} = u^\dagger$. Moreover, from (4.13) and $\hat{u} = u^\dagger$ we can deduce that $J(u_{j_k}) \to J(u^\dagger)$.

As in the proof of Theorem 4.8, all arguments can be applied to any subsequence of $\{u_j\}_{j\in\mathbb{N}}$, which shows that $u_j \to u^\dagger$ in $\tau_{\mathcal{U}}$ and $J(u_j) \to J(u^\dagger)$. □

Remark 4.11. Similar to the stability we can get strong convergence if J satisfies the Radon-Riesz property.

4.2.4 Convergence rates

In the last section we have proven convergence of the regularisation method in the topology $\tau_{\mathcal{U}}$ and not in the norm as in Chapter 3. Thus, we cannot expect to prove convergence rates in the norm. However, it turns out we can prove convergence rates in the Bregman distance.

Definition 4.12. *Let $\mathcal{U}$ be a Banach space and $E : \mathcal{U} \to \mathbb{R}_\infty$ be a proper and convex functional. Moreover, let $u, v \in \mathcal{U}, E(v) < \infty$ and $p \in \partial E(v)$. Then the* (generalised) Bregman distance *of E is defined as*

$$D_E^p(u, v) := E(u) - E(v) - \langle p, u - v\rangle\,. \tag{4.14}$$

Remark 4.12. It is easy to check that a Bregman distance somewhat resembles a metric as for all $u, v \in \mathcal{U}, p \in \partial E(v)$ there is $D_E^p(u,v) \geq 0$ and $D_E^p(v,v) = 0$. There are functionals where the Bregman distance (up to a square root) is actually a metric: Let $\mathcal{U}$ be a Hilbert space and let $E(u) := \frac{1}{2}\|u\|_{\mathcal{U}}^2$. Then $D_J^p(u,v) = \frac{1}{2}\|u-v\|_{\mathcal{U}}^2$. However, there are functionals E where $D_E^p(u,v) = 0$ does not imply $u = v$, see the third example sheet for examples.

Theorem 4.10. *Assume the setting of Theorem 4.6 that guarantees that the mapping R_α is well-defined. Let $f \in \mathcal{R}(K)$ be clean data and $u^\dagger$ be a solution of the inverse problem, i.e. $f = Ku^\dagger$, and consider noisy data $f^\delta \in \mathcal{V}$ with $\|f - f^\delta\|_{\mathcal{V}} \leq \delta$. Moreover, let $u^\dagger$ satisfy the* source condition

$$p = K^* w \in \partial J(u^\dagger)$$

and denote $u_\alpha^\delta := R_\alpha f^\delta$. Then,

(a) $$D_J^p(u_\alpha^\delta, u^\dagger) \leq \frac{\delta^2}{2\alpha} + \alpha\|w\|_{\mathcal{V}^*}\delta + \frac{\alpha^2\|w\|_{\mathcal{V}^*}^2}{2}\,,$$

(b) $$\frac{1}{2}\|Ku_\alpha^\delta - f^\delta\|_{\mathcal{V}}^2 \leq \delta^2 + 2\alpha\|w\|_{\mathcal{V}^*}\delta + 2\alpha^2\|w\|_{\mathcal{V}^*}^2\,,$$ *and*

(c) $$J(u_\alpha^\delta) \leq \frac{\delta^2}{2\alpha} + J(u^\dagger)\,.$$

Moreover, for the a-priori parameter choice rule $\alpha(\delta) = \delta$ we have

$$D_J^p(u_\alpha^\delta, u^\dagger) = \mathcal{O}(\delta), \quad \|Ku_\alpha^\delta - f^\delta\|_{\mathcal{V}} = \mathcal{O}(\delta), \quad \text{and} \quad J(u_\alpha^\delta) \leq J(u^\dagger) + \mathcal{O}(\delta)\,.$$

Proof. From the minising property of u_α^δ and $Ku^\dagger = f$ it follows that

$$\frac{1}{2}\|Ku_\alpha^\delta - f^\delta\|_{\mathcal{V}}^2 + \alpha J(u_\alpha^\delta) \leq \frac{1}{2}\|Ku^\dagger - f^\delta\|_{\mathcal{V}}^2 + \alpha J(u^\dagger) \leq \frac{\delta^2}{2} + \alpha J(u^\dagger)\,. \tag{4.15}$$

From the non-negativity of the data term and (4.15) we derive assertion (c) as

$$J(u_\alpha^\delta) \leq \frac{1}{\alpha}\left(\frac{1}{2}\|Ku_\alpha^\delta - f^\delta\|_{\mathcal{V}}^2 + \alpha J(u_\alpha^\delta)\right) \leq \frac{\delta^2}{2\alpha} + J(u^\dagger)\,.$$

Moreover, by reordering the terms of (4.15) and completing Bregman distance, we get

$$\frac{1}{2}\|Ku_\alpha^\delta - f^\delta\|_{\mathcal{V}}^2 + \alpha D_J^p(u_\alpha^\delta, u^\dagger) \leq \frac{\delta^2}{2} - \alpha\langle p, u_\alpha^\delta - u^\dagger\rangle$$

where we can further estimate

$$\begin{aligned} -\langle p, u_\alpha^\delta - u^\dagger\rangle &= -\langle w, K(u_\alpha^\delta - u^\dagger)\rangle = -\langle w, Ku_\alpha^\delta - f\rangle \\ &\leq \|w\|_{\mathcal{V}^*}\|Ku_\alpha^\delta - f\|_{\mathcal{V}} \leq \|w\|_{\mathcal{V}^*}\left(\|Ku_\alpha^\delta - f^\delta\|_{\mathcal{V}} + \delta\right) \end{aligned}$$

Combining the two yields

$$\begin{aligned} \frac{1}{2}\|Ku_\alpha^\delta - f^\delta\|_{\mathcal{V}}^2 + \alpha D_J^p(u_\alpha^\delta, u^\dagger) &\leq \frac{\delta^2}{2} + \alpha\|w\|_{\mathcal{V}^*}\delta + \alpha\|w\|_{\mathcal{V}^*}\|Ku_\alpha^\delta - f^\delta\|_{\mathcal{V}} \\ &\leq \frac{\delta^2}{2} + \alpha\|w\|_{\mathcal{V}^*}\delta + \frac{\alpha^2\|w\|_{\mathcal{V}^*}^2}{2\gamma} + \frac{\gamma}{2}\|Ku_\alpha^\delta - f^\delta\|_{\mathcal{V}}^2 \end{aligned}$$

where we used $ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$ for the second inequality. Thus, we derive

$$(1-\gamma)\frac{1}{2}\|Ku_\alpha^\delta - f^\delta\|_{\mathcal{V}}^2 + \alpha D_J^p(u_\alpha^\delta, u^\dagger) \leq \frac{\delta^2}{2} + \alpha\|w\|_{\mathcal{V}^*}\delta + \frac{\alpha^2\|w\|_{\mathcal{V}^*}^2}{2\gamma}$$

Choosing $\gamma = 1$ and $\gamma = 1/2$ yields the assertions (a) and (b). □

Remark 4.13. Note that we did not use the source condition for the assertion (c), thus it is true for all solutions $u^\dagger$ of the inverse problem, i.e. $Ku^\dagger = f$.

Remark 4.14. We did not assume that $u^\dagger$ is a J-minimising solution. However, let $\mathcal{U}$ be a Hilbert space and $J(u) = \frac{1}{2}\|u\|_{\mathcal{U}}^2$. Then the source condition is equivalent to $K^*w = u^\dagger$ which is in turn equivalent to $u^\dagger \in \mathcal{R}(K^*) = \mathcal{N}(K)^\perp$. Thus, $u^\dagger$ is the minimial norm solution.

4.3 Numerical implementation

4.3.1 Saddle point problems

In order to compute a solution to an inverse problem with variational regularisation, we have to solve the minimisation problem

$$\min_{u\in\mathcal{U}}\left\{\frac{1}{2}\|Ku - f\|_{\mathcal{V}}^2 + \alpha J(u)\right\}.$$

However, for total variation regularisation for instance, the minimisation problem actually becomes a *saddle point problem*

$$\min_{u\in\mathcal{U}} \sup_{\varphi\in\mathcal{D}(\Omega,\mathbb{R}^n)} \left\{\frac{1}{2}\|Ku - f\|_{\mathcal{V}}^2 + \alpha \int_\Omega u(x)\mathrm{div}\varphi(x)\, dx\right\}.$$

where we have to minimise with respect to u but to maximise with respect to φ.

While for the total variation the minimisation problem is intrinsically a saddle point problem, we can rewrite many other variational regularisation models as a saddle point problem by means of the Fenchel conjugate.

Definition 4.13. *Let $\mathcal{U}$ be a Banach space and let $E\colon \mathcal{U} \to \mathbb{R}_\infty$ be proper, lower semi-continuous and convex. Then the* Fenchel conjugate *or* convex conjugate *of E is defined to be the mapping $E^*\colon \mathcal{U}^* \to \mathbb{R}_\infty$ with*

$$E^*(v) := \sup_{u\in\mathcal{U}}\left\{\langle v, u\rangle - E(u)\right\}.$$

Remark 4.15. It can be shown that in a Hilbert space the following identity holds: $E^{**} := (E^*)^* = E$. Thus in a Hilbert space, we can always reformulate our regularisation functional as

$$E(u) = \sup_{v\in\mathcal{U}^*}\left\{\langle v, u\rangle - E^*(v)\right\}.$$

Example 4.17. Let $\Omega \subset \mathbb{R}^n, 1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Moreover, let $E\colon L^p(\Omega) \to \mathbb{R}$ with $E(u) := \frac{1}{p}\|u\|_p^p$. Then the Fenchel conjugate of E is given by $E^*\colon L^q(\Omega) \to \mathbb{R}$ with $E^*(v) = \frac{1}{q}\|v\|_q^q$. In particular, the idenitity $(\frac{1}{2}\|\cdot\|_2^2)^* = \frac{1}{2}\|\cdot\|_2^2$ holds.

Definition 4.14. *Let $\mathcal{U}, \mathcal{V}$ be Banach spaces. Let $E\colon \mathcal{U} \to \mathbb{R}_\infty, F^*\colon \mathcal{V}^* \to \mathbb{R}_\infty$ be proper, convex and lower semi-continuous functionals where F^* is the Fenchel conjugate of a functional $F\colon \mathcal{V} \to \mathbb{R}_\infty$ and $D\colon \mathcal{U} \to \mathcal{V}^*$ be a linear and continuous operator. We can associate to any minimisation problem*

$$\min_{u\in\mathcal{U}}\left\{E(u) + F^*(Du)\right\} \tag{4.16}$$

a corresponding saddle point problem

$$\min_{u\in\mathcal{U}}\sup_{v\in\mathcal{V}}\Big\{\Psi(u,v) := E(u) + \langle Du, v\rangle - F(v)\Big\}, \tag{4.17}$$

the solution of which we call a saddle point*.*

Remark 4.16. If (u^*, v^*) is a saddle point, then u^* solves the minimisation problem (4.16).

4.3.2 Optimality condition for saddle point problems

One can show that for any saddle point (u^*, v^*) we have for all $(u, v) \in \mathcal{U} \times \mathcal{V}$ that

$$\Psi(u^*, v) \le \Psi(u^*, v^*) \le \Psi(u, v^*)$$

which shows that

$$u^* \in \arg\min_{u\in\mathcal{U}} \Psi(u, v^*) \quad \text{and} \quad v^* \in \underbrace{\arg\max_{v\in\mathcal{V}} \Psi(u^*, v)}_{=\arg\min_{v\in\mathcal{V}}[-\Psi(u^*,v)]} .$$

Thus, as $\Psi(u, v)$ is convex in u and $-\Psi(u, v)$ is convex in v, we see that necessary and sufficient optimality conditions for saddle point problems (4.17) are

$$0 \in \partial_u\Psi(u^*, v^*) \quad \text{and} \quad 0 \in \partial_v[-\Psi(u^*, v^*)] \tag{4.18}$$

In order to make better sense out of the optimality conditions, we have to discuss some more properties of the subdifferential.

Definition 4.15. *Let $E : \mathcal{U} \to \mathbb{R}$ be a mapping from the Banach space $\mathcal{U}$ and $u \in \mathcal{U}$. If there exists a $A \in \mathcal{L}(\mathcal{U}, \mathbb{R})$ that*

$$\lim_{h\to 0} \frac{|E(u + h) - E(u) - Ah|}{\|h\|_{\mathcal{U}}} = 0,$$

holds true, then E is called Fréchet differentiable *in x and $E'(u) := A$ the* Fréchet derivative *in u. If the Fréchet derivative exists for all $u \in \mathcal{U}$, the operator $E' : \mathcal{U} \to \mathcal{U}^*$ is called Fréchet derivative.*

Example 4.18. Let $\mathcal{U}$ be a Banach space and $p \in \mathcal{U}^*$. Then the Fréchet derivative of p is given by $p' = p$.

Example 4.19. Let $\mathcal{U}, \mathcal{V}$ be Hilbert spaces, $K \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, $f \in \mathcal{V}$ and $E : \mathcal{U} \to \mathbb{R}$ be defined as $E(u) := \frac{1}{2}\|Ku - f\|^2_{\mathcal{U}}$. Then the Fréchet derivative of E is given by $E' : \mathcal{U} \to \mathcal{U}^*$ with

$$E'(u) = \langle K^*(Ku - f), \cdot\rangle_{\mathcal{U}},$$

thus by the Riesz representation theorem can be identified with $K^*(Ku - f)$.

Proof. For any $u \in \mathcal{U}$, an easy calculation shows that

$$\frac{1}{2}\|K(u + h) - f\|^2_{\mathcal{U}} - \frac{1}{2}\|Ku - f\|^2_{\mathcal{U}} = \langle Ku - f, Kh\rangle_{\mathcal{U}} + \frac{1}{2}\|Kh\|^2_{\mathcal{U}}$$

and thus with $Ah := \langle K^*(Ku-f), h\rangle_{\mathcal{U}}$ we have that

$$\begin{aligned}\frac{|E(u+h)-E(u)-Ah|}{\|h\|_{\mathcal{U}}} &= \frac{|\frac{1}{2}\|K(u+h)-f\|_{\mathcal{U}}^2 - \frac{1}{2}\|Ku-f\|_{\mathcal{U}}^2 - \langle K^*(Ku-f), h\rangle_{\mathcal{U}}|}{\|h\|_{\mathcal{U}}} \\ &= \frac{|\langle Ku-f, Kh\rangle_{\mathcal{U}} + \frac{1}{2}\|Kh\|_{\mathcal{U}}^2 - \langle Ku-f, Kh\rangle_{\mathcal{U}}|}{\|h\|_{\mathcal{U}}} \\ &= \frac{1}{2}\|Kh\|_{\mathcal{U}} \leq \frac{1}{2}\|K\|\|h\|_{\mathcal{U}} \to 0\end{aligned}$$

as $\|h\|_{\mathcal{U}} \to 0$. □

Proposition 4.3. *Let $\mathcal{U}$ be a Banach space and $E : \mathcal{U} \to \mathbb{R}$ be a convex functional that is Fréchet differentiable in $u \in \mathcal{U}$. Then*

$$\partial E(u) = \{E'(u)\}\,.$$

Proof. The proof is left as an exercise. □

Proposition 4.4. *Let $\mathcal{U}$ be a Banach space, $F\colon \mathcal{U} \to \mathbb{R}_\infty$ convex, $E\colon \mathcal{U} \to \mathbb{R}$ E be convex and* Fréchet differentiable *and $G\colon \mathcal{U} \to \mathbb{R}_\infty, G(u) := E(u) + F(u)$. Then for all $u \in \mathrm{dom}(G) = \mathrm{dom}(F)$ it holds*

$$\partial G(u) = E'(u) + \partial F(u)\,.$$

Proof. It is trivial to see that for any $E, F\colon \mathcal{U} \to \mathbb{R}_\infty$ we have that

$$\partial E + \partial F \subset \partial(E+F)$$

and thus it remains to show that

$$\partial G(u) \subset E'(u) + \partial F(u)\,.$$

For any $u \in \mathrm{dom}(G)$, let $p \in \partial G(u)$ and denote $q := p - E'(u)$. We will show that $q \in \partial F(u)$. For any $v \in \mathcal{U}$ it holds from the subgradient condition for p that

$$\begin{aligned}F(v) &= F(v) + E(v) - E(v) = G(v) - E(v) \\ &\geq G(u) + \langle p, v-u\rangle - E(v) \\ &= F(u) + \langle q, v-u\rangle + \underbrace{E(u) - E(v) + \langle E'(u), v-u\rangle}_{\geq 0,\text{ as } E'(u)\in\partial E(u)} \\ &\geq F(u) + \langle q, v-u\rangle\,,\end{aligned}$$

and the assertion is proven. □

With the help of Proposition 4.4 we can rewrite the optimality conditions (4.18) as

$$\begin{aligned}&0 \in \partial_u \Psi(u^*, v^*) = \partial_u\Big\{E(u^*) + \langle Du^*, v^*\rangle - F(v^*)\Big\} = \partial E(u^*) + D^* v^* \\ &0 \in \partial_v[-\Psi(u^*, v^*)] = -Du^* + \partial F(v^*)\end{aligned}$$

which simplify to

$$-D^*v^* \in \partial E(u^*) \quad \text{and} \quad Du^* \in \partial F(v^*)\,. \tag{4.19}$$

4.3.3 Proximal operators

Definition 4.16. *Let $\mathcal{U}$ be a Hilbert space and let $E\colon \mathcal{U} \to \mathbb{R}_\infty$ be proper, lower semi-continuous and convex. Then we define the* proximal operator $\operatorname{prox}_E\colon \mathcal{U} \to \mathcal{U}$ *by*

$$\operatorname{prox}_E(z) := \arg\min_{u\in\mathcal{U}} \left\{ \frac{1}{2}\|u - z\|_{\mathcal{U}}^2 + E(u) \right\}. \tag{4.20}$$

Remark 4.17. It can be proven with tools from convex analysis, that the function to be minimised in (4.20) is bounded from below and coercive. Intuitively, this is the case as convex functions are not allow to decrease faster than linear functions. Therefore, the quadratic term dominates and these two properties hold. Then the existence follows from the direct method and the uniqueness from the strict convexity of the squared Hilbert space norm.

Remark 4.18. With the help of Proposition 4.4 we see that $x = \operatorname{prox}_E(z)$ if and only if

$$0 \in \partial \left(\frac{1}{2}\|x - z\|_{\mathcal{U}}^2 + E(x) \right) = x - z + \partial E(x) = (I + \partial E)(x) - z$$

which in turn is the case if and only if $x = (I + \partial E)^{-1}(z)$. It is interesting to see that despite the mapping $I + \partial E$ being multi-valued, its inverse is single-valued and can be computed by the proximal operator.

Example 4.20. Let $\mathcal{U}, \mathcal{V}$ be Hilbert spaces, $K \in \mathcal{L}(\mathcal{U}, \mathcal{V}), w \in \mathcal{U}, f \in \mathcal{V}$ and $\tau \geq 0$. Then the proximal operator for $\tau E(u) := \frac{\tau}{2}\|Ku - f\|_{\mathcal{V}}^2$ is given by

$$\operatorname{prox}_{\tau E}(z) = (I + \tau K^* K)^{-1}(z + \tau K^* f).$$

Proof. From Example 4.19 we see that the necessary and sufficient condition for the minimiser of $\frac{1}{2}\| \cdot - z\|_{\mathcal{U}}^2 + \tau E$ is given by

$$u - z + \tau K^*(Ku - f) = 0$$

and thus we derive

$$u = (I + \tau K^* K)^{-1}(z + \tau K^* f).$$

□

Example 4.21. Let $\Omega \subset \mathbb{R}^n$ and $\mathcal{U} = L^2(\Omega, \mathbb{R}^n)$. Moreover, let

$$C := \Big\{ u \in \mathcal{U} \,|\, \|u(x)\|_2 \leq 1 \text{ for almost every } x \in \Omega \Big\}.$$

Then the proximal operator for $E\colon \mathcal{U} \to \mathbb{R}_\infty, E = \chi_C$, i.e.

$$E(u) = \begin{cases} 0 & \text{if } u \in C \\ \infty & \text{else} \end{cases}$$

is given by the orthogonal projection of u onto C, i.e.

$$[\operatorname{prox}_E(u)](x) = \frac{u(x)}{\max(1, \|u(x)\|_2)} \quad \text{for almost every } x \in \Omega.$$

4.3.4 Primal-dual hybrid gradient method

In this section we discuss how to solve a saddle point problem (4.17) numerically. Let $(u^*, v^*) \in \mathcal{U} \times \mathcal{V}$ be a saddle point, then the optimality conditions (4.19) are equivalent to the existence of subgradients $p^* \in \partial E(u^*)$ and $q^* \in \partial F(v^*)$ so that

$$p^* + D^* v^* = 0 \quad \text{and} \quad q^* - D u^* = 0\,. \tag{4.21}$$

We approach finding a saddle point of (4.17) via the iterative approach

$$\begin{pmatrix} p^{k+1} + D^* v^{k+1} \\ q^{k+1} - D u^{k+1} \end{pmatrix} + M \begin{pmatrix} u^{k+1} - u^k \\ v^{k+1} - v^k \end{pmatrix} = 0\,, \tag{4.22}$$

for $p^{k+1} \in \partial E(u^{k+1})$ and $q^{k+1} \in \partial F(v^{k+1})$, and a 2×2 operator-matrix M.

Remark 4.19. We observe that if (u^{k+1}, v^{k+1}) is a fixed point of the iterations, i.e. $(u^{k+1}, v^{k+1}) = (u^k, v^k)$, then (u^{k+1}, v^{k+1}) is a saddle point of (4.17).

An important question is how to choose M such that (4.22) results in a convergent algorithm with relatively simple update steps for u^{k+1} and v^{k+1}? A naïve choice for M could simply be the identity; however, in that case the two updates for u^{k+1} and v^{k+1} are coupled which makes it difficult to solve. Alternatively, we propose to use

$$M := \begin{pmatrix} \frac{1}{\tau} I & -D^* \\ -D & \frac{1}{\sigma} I \end{pmatrix} \tag{4.23}$$

instead, with τ and σ being positive scalars. For this choice, (4.22) simplifies to the equations

$$\begin{aligned} p^{k+1} + D^* v^{k+1} + \tau^{-1}(u^{k+1} - u^k) - D^*(v^{k+1} - v^k) &= 0 \\ q^{k+1} - D u^{k+1} - D(u^{k+1} - u^k) + \sigma^{-1}(v^{k+1} - v^k) &= 0 \end{aligned}$$

which are equivalent to

$$\begin{aligned} u^{k+1} + \tau p^{k+1} &= u^k - \tau D^* v^k \\ v^{k+1} + \sigma q^{k+1} &= v^k + \sigma D(2u^{k+1} - u^k) \end{aligned}\,. \tag{4.24}$$

Due to $p^{k+1} \in \partial E(u^{k+1})$ and $q^{k+1} \in \partial F(u^{k+1})$, Equations (4.24) can be rewritten as

$$\begin{aligned} u^{k+1} &= (I + \tau \partial E)^{-1}(u^k - \tau D^* v^k)\,, \\ \overline{u}^{k+1} &= 2u^{k+1} - u^k \\ v^{k+1} &= (I + \sigma \partial F)^{-1}(v^k + \sigma D \overline{u}^{k+1})\,, \end{aligned} \tag{4.25}$$

The iterates (4.25) are known as the *primal-dual hybrid gradient method (PDHGM)*. Before we prove actual convergence of those iterates to a saddle point of (4.17), we want to highlight what makes PDHGM so useful. First of all, the particular choice of M (4.23) decouples u^{k+1} and v^{k+1} in the update for u^{k+1}, which makes updating u^{k+1} and v^{k+1} in an alternating fashion possible. Secondly, PDHGM now only requires basic arithmetic operations, operator and adjoint applications, and the evaluation of the proximal operations with respect to E and F. If these are simple, the overall PDHGM is simple.

Algorithm 1 Primal-Dual Hybrid Gradient Method.

Initialise: $(u^0, v^0) \in \mathcal{U}^* \times \mathcal{V}$, step sizes: $\tau, \sigma > 0$, with $\sigma\tau\|D\|^2 < 1$, number of iterations: $K \in \mathbb{N}$
Iterate:

1: **for** $k = 0, \dots, K-1$ **do**
2: $\quad u^{k+1} = \operatorname{prox}_{\tau E}(u^k + \tau D^* v^k)$
3: $\quad \overline{u}^{k+1} = 2u^{k+1} - u^k$
4: $\quad v^{k+1} = \operatorname{prox}_{\sigma F}(v^k - \sigma D\overline{u}^{k+1})$
5: **end for**

4.3.5 Convergence of PDHGM

Before we prove convergence of the iterates, we want to prove that M as defined in (4.23) is positive definite under suitable conditions on τ and σ.

Lemma 4.11. *Let $\tau, \sigma > 0$ with $\tau\sigma\|D\|^2 < 1$. Moreover, let $\mathcal{U}, \mathcal{V}$ be Hilbert spaces and denote the Hilbert space $W := \mathcal{U} \times \mathcal{V}$ with inner product $\langle (u_1, v_1), (w_1, w_2) \rangle := \langle u_1, u_2 \rangle_{\mathcal{U}} + \langle v_1, v_2 \rangle_{\mathcal{V}}$. Consider the self-adjoint operator $M\colon W \to W$ as defined in (4.23). Then, M is positive definite.*

Proof. Let $(u, v) \in \mathcal{U} \times \mathcal{V}, (u, v) \neq 0$ and denote $\hat{u} := \tau^{-1/2} u$ and $\hat{v} := \sigma^{-1/2} v$. Without loss of generality, let $u \neq 0$. Then

$$\begin{aligned}
\langle M(u,v), (u,v) \rangle &= \langle (\tau^{-1}u - D^*v, -Du + \sigma^{-1}v), (u,v) \rangle \\
&= \frac{1}{\tau}\|u\|_{\mathcal{U}}^2 + \frac{1}{\sigma}\|v\|_{\mathcal{V}}^2 - 2\langle Du, v \rangle_{\mathcal{V}} \\
&= \|\hat{u}\|_{\mathcal{U}}^2 + \|\hat{v}\|_{\mathcal{V}}^2 - 2\underbrace{\langle D\tau^{1/2}\hat{u}, \sigma^{1/2}\hat{v} \rangle_{\mathcal{V}}}_{=\langle D\tau^{1/2}\sigma^{1/2}\hat{u}, \hat{v} \rangle_{\mathcal{V}}} \\
&\geq \|\hat{u}\|_{\mathcal{U}}^2 + \|\hat{v}\|_{\mathcal{V}}^2 - \|D\|^2\tau\sigma\|\hat{u}\|_{\mathcal{U}}^2 - \|\hat{v}\|_{\mathcal{V}}^2 = \frac{(1 - \|D\|^2\tau\sigma)}{\tau}\|u\|_{\mathcal{U}}^2 > 0
\end{aligned}$$

□

Lemma 4.12. *Let W be a Hilbert space and $M\colon W \to W$ be linear and self-adjoint. Then, for all $w^{k+1}, w^k, w^* \in W$ we have that*

$$\langle M(w^{k+1} - w^k), w^{k+1} - w^* \rangle = \frac{1}{2}\left(\|w^{k+1} - w^k\|_M^2 - \|w^k - w^*\|_M^2 + \|w^{k+1} - w^*\|_M^2 \right),$$

where $\|w\|_M^2 := \langle Mw, w \rangle$.

Proof. Long but straight forward calculations lead to

$$\begin{aligned}
&\|w^{k+1} - w^k\|_M^2 - \|w^k - w^*\|_M^2 + \|w^{k+1} - w^*\|_M^2 \\
&= \langle M(w^{k+1} - w^k), w^{k+1} - w^k \rangle - \langle M(w^k - w^*), w^k - w^* \rangle + \langle M(w^{k+1} - w^*), w^{k+1} - w^* \rangle \\
&= \langle M(w^{k+1} - w^k), w^{k+1} \rangle - \langle M(w^{k+1} - w^k), w^k \rangle \\
&\quad - \langle M(w^k - w^*), w^k \rangle + \langle M(w^k - w^*), w^* \rangle \\
&\quad + \langle M(w^{k+1} - w^*), w^{k+1} \rangle - \langle M(w^{k+1} - w^*), w^* \rangle \\
&= \langle M(w^{k+1} - w^k), w^{k+1} \rangle - \langle Mw^{k+1}, w^k \rangle \\
&\quad + \langle Mw^*, w^k \rangle + \langle Mw^k, w^* \rangle
\end{aligned}$$

$$
\begin{aligned}
&+ \langle M(w^{k+1} - w^*), w^{k+1}\rangle - \langle Mw^{k+1}, w^*\rangle \\
&= 2\Big[\langle Mw^{k+1}, w^{k+1}\rangle - \langle Mw^{k+1}, w^k\rangle + \langle Mw^*, w^k\rangle - \langle Mw^{k+1}, w^*\rangle\Big] \\
&= 2\Big[\langle Mw^{k+1}, w^{k+1} - w^*\rangle - \langle w^{k+1}, Mw^k\rangle + \langle w^*, Mw^k\rangle\Big] \\
&= 2\Big[\langle Mw^{k+1}, w^{k+1} - w^*\rangle - \langle w^{k+1} - w^*, Mw^k\rangle\Big] \\
&= 2\langle M(w^{k+1} - w^k), w^{k+1} - w^*\rangle\,.
\end{aligned}
$$

□

Similar to the regularisation results in the last section, we will prove the convergence of PDHGM in a weaker Bregman distance setting. However, here we are able to prove the results in the symmetric Bregman distance.

Definition 4.17. *Let $\mathcal{U}$ be a Banach space and $E\colon \mathcal{U} \to \mathbb{R}_\infty$ be a functional defined on it. Let $u_1, u_2 \in \operatorname{dom}(E), p_i \in \partial E(u_i)$, $i = 1, 2$. Then the* (generalised) symmetric Bregman distance *is defined as*

$$
D_E^{symm}(x_1, x_2) := D_E^{p_1}(x_2, x_1) + D_E^{p_2}(x_1, x_2) = \langle x_1 - x_2, p_1 - p_2\rangle\,.
$$

Note that we have omitted p_1 and p_2 in the notation of D_E^{symm} only for the sake of brevity.

Theorem 4.11. *Let $\mathcal{U}, \mathcal{V}$ be Hilbert spaces and $W := \mathcal{U} \times \mathcal{V}$. Let $\tau, \sigma > 0$, with $\tau\sigma\|D\|^2 < 1$, $M\colon W \to W$ defined via* (4.23)*, $w^0 := (u^0, v^0) \in W, \overline{u}^0 = u^0$ and the sequence $\{w^k\}_{k\in\mathbb{N}} := \{u^k, v^k\}_{k\in\mathbb{N}}$ be defined via Algorithm 1. Let $w^* := (u^*, v^*) \in W$ be any saddle point of* (4.17)*. Then the following assertions hold.*

(a) The sequence $\{w^k\}_{k\in\mathbb{N}}$ is bounded and $\|w^{k+1} - w^k\| \to 0$.

(b) The M-distance of w^k to w^* is not increasing, i.e.

$$
\|w^{k+1} - w^*\|_M \leq \|w^k - w^*\|_M\,.
$$

(c) The sequence $(u^k, v^k)_{k\in\mathbb{N}}$ converges to (u^*, v^*) in a Bregman sense, i.e.

$$
\lim_{k\to\infty} D_E^{symm}(u^k, u^*) = 0 \qquad \text{and} \qquad \lim_{k\to\infty} D_F^{symm}(v^k, v^*) = 0\,.
$$

If in addition, $\dim W < \infty$ (e.g. after discretisation), then $\{w^k\}_{k\in\mathbb{N}}$ is convergent in norm and its limit is a saddle point of (4.17)*.*

Proof. Note that we can combine (4.21) and (4.22) to

$$
\begin{aligned}
0 &= \begin{pmatrix} p^{k+1} + D^* v^{k+1} \\ q^{k+1} - Du^{k+1} \end{pmatrix} + M \begin{pmatrix} u^{k+1} - u^k \\ v^{k+1} - v^k \end{pmatrix} - \begin{pmatrix} p^* + D^* v^* \\ q^* - Du^* \end{pmatrix} \\
&= \begin{pmatrix} p^{k+1} - p^* + D^*(v^{k+1} - v^*) \\ q^{k+1} - q^* - D(u^{k+1} - u^*) \end{pmatrix} + M \begin{pmatrix} u^{k+1} - u^k \\ v^{k+1} - v^k \end{pmatrix}
\end{aligned}
\tag{4.26}
$$

Taking the inner product the first term with $(u^{k+1}-u^*, v^{k+1}-v^*)$ yields

$$\begin{aligned}
&\left\langle \begin{pmatrix} p^{k+1}-p^*+D^*(v^{k+1}-v^*) \\ q^{k+1}-q^*-D(u^{k+1}-u^*) \end{pmatrix}, \begin{pmatrix} u^{k+1}-u^* \\ v^{k+1}-v^* \end{pmatrix} \right\rangle \\
&= \langle p^{k+1}-p^*+D^*(v^{k+1}-v^*), u^{k+1}-u^* \rangle + \langle q^{k+1}-q^*-D(u^{k+1}-u^*), v^{k+1}-v^* \rangle \\
&= \langle p^{k+1}-p^*, u^{k+1}-u^* \rangle + \langle q^{k+1}-q^*, v^{k+1}-v^* \rangle \\
&= D_E^{\text{symm}}(u^{k+1}, u^*) + D_F^{\text{symm}}(v^{k+1}, v^*)\,.
\end{aligned}$$

Furthermore, taking the inner product of the second term in (4.26) with $w^{k+1}-w^*$ (change the notation to $w := (u,v)$) is non-positive and we can conclude with Lemma 4.12 that

$$\langle M(w^{k+1}-w^k), w^{k+1}-w^* \rangle = \frac{1}{2}\left(\|w^{k+1}-w^k\|_M^2 - \|w^k-w^*\|_M^2 + \|w^{k+1}-w^*\|_M^2 \right),$$

thus we have

$$\begin{aligned}
D_E^{\text{symm}}&(u^{k+1}, u^*) + D_F^{\text{symm}}(v^{k+1}, v^*) \\
&+ \frac{1}{2}\left(\|w^{k+1}-w^k\|_M^2 - \|w^k-w^*\|_M^2 + \|w^{k+1}-w^*\|_M^2 \right) = 0 \qquad (4.27)
\end{aligned}$$

As symmetric Bregman distances are non-negative, this yields

$$\|w^k-w^*\|_M^2 \geq \|w^{k+1}-w^*\|_M^2 + \|w^{k+1}-w^k\|_M^2\,.$$

This has two implications. First, we have that

$$\|w^{k+1}-w^*\|_M \leq \|w^k-w^*\|_M$$

which shows (b). Moreover, we have that $\|w^k-w^*\|_M \leq \|w^0-w^*\|_M$ such that

$$\|w^k\|_M \leq \|w^0-w^*\|_M - \|w^*\|_M < \infty\,.$$

As $\|\cdot\|_M$ defines an equivalent norm on W this shows that $\{w^k\}_{k\in\mathbb{N}}$ is bounded.

Summing up (4.27) from $k=0,\ldots,K-1$ yields

$$\begin{aligned}
&\sum_{k=0}^{K-1} 2\left(D_E^{\text{symm}}(u^{k+1}, u^*) + D_F^{\text{symm}}(v^{k+1}, v^*) \right) + \sum_{k=0}^{K-1} \|w^{k+1}-w^k\|_M^2 \\
&= \sum_{k=0}^{K-1} \left(\|w^k-w^*\|_M^2 - \|w^{k+1}-w^*\|_M^2 \right) \\
&= \|w^0-w^*\|_M^2 - \|w^K-w^*\|_M^2 \leq \|w^0-w^*\|_M^2\,,
\end{aligned}$$

thus taking the limit $K\to\infty$

$$\sum_{k=0}^{\infty} 2\left(D_E^{\text{symm}}(u^{k+1}, u^*) + D_F^{\text{symm}}(v^{k+1}, v^*) \right) + \sum_{k=0}^{\infty} \|w^{k+1}-w^k\|_M^2 \leq \|w^0-w^*\|_M^2 < \infty\,.$$

This implies that

$$D_E^{\text{symm}}(u^{k+1}, u^*) \to 0, \quad D_F^{\text{symm}}(v^{k+1}, v^*) \to 0, \quad \text{and} \quad \|w^{k+1}-w^k\|_M \to 0\,,$$

and thus proves (a) and (c).

Let now W be finite dimensional. Then the boundedness of $\{w^k\}_{k\in\mathbb{N}}$ implies that there exists a convergent subsequence $\{w^{k_j}\}_{j\in\mathbb{N}}$ and $w^\infty \in W$ with $w^{k_j} \to w^\infty$. As $w^{k+1} - w^k \to 0$ we also have that $w^{k_j+1} \to w^\infty$ which shows that w^∞ is a fixed point of the iteration and thus a saddle point of (4.17), see Remark 4.19. Let $\varepsilon > 0$. By the convergence of the subsequence there exists a $j_0 \in \mathbb{N}$ so that $\|w^{k_{j_0}} - w^\infty\|_M < \varepsilon$. But as w^∞ is a saddle point, this means by assertion (b) that for all $k > k_{j_0}$ we have that

$$\|w^k - w^\infty\|_M \leq \|w^{k_{j_0}} - w^\infty\|_M < \varepsilon$$

which shows that $w^k \to w^\infty$. □

Remark 4.20. For certain functionals, the convergence of the symmetric Bregman distance implies convergence in norm. E.g. $E(u) = \frac{1}{2}\|u\|_{\mathcal{U}}^2$.

4.3.6 Deconvolution with total variation regularisation

An example for total variation regularisation of the inverse problem of image convolution is given as Exercise 2 on Exercisesheet 3.

Example 4.22. Hence, PDHGM reads as

$$u^{k+1} = (I + \tau K^* K)^{-1}(u^k + \tau(\mathrm{div} v^k + K^* f)) \tag{4.28}$$

$$\overline{u}^{k+1} = 2u^{k+1} - u^k \tag{4.29}$$

$$v_j^{k+1} = \frac{v_j^k + \sigma \nabla_j \overline{u}^{k+1}}{\max(1, \|v^k + \sigma \nabla \overline{u}^{k+1}\|_2)} \qquad \text{for all } j \in \{1, \dots, n\} \tag{4.30}$$

in case of total variation regularisation, due to $\nabla^* = -\mathrm{div}$. Note that in case of ROF-denoising as proposed in [12], Equation (4.28) simplifies to

$$u^{k+1} = \frac{u^k + \tau(\mathrm{div} v^k + f)}{1 + \tau} .$$

Note that for real world applications one obviously has to find appropriate discretisations of K and ∇.

Chapter 5

Inverse problems with non-linear forward operator

To conclude this lecture we want to look into inverse problems with non-linear forward operator. For simplicity, we stick to the Hilbert space setting, and consider inverse problems of the form

$$F(u) = f^\delta , \tag{5.1}$$

for a non-linear operator $F : \mathcal{U} \to \mathcal{H}$, where both $\mathcal{U}$ and $\mathcal{H}$ are Hilbert spaces. Typical examples for non-linear inverse problems are parameter identification problems of partial differential equations.

Example 5.1 (Groundwater filtration)**.** The problem of groundwater filtration can be modelled as the inverse problem (5.1) for which F is the operator that maps the hydraulic permittivity u to the solution of the partial differential equation (PDE)

$$\mathrm{div}(u\nabla f) = g \,,$$

under suitable boundary conditions. Here f represents the unknown solution of the PDE, and g is a given source.

Example 5.2 (Impedance tomography)**.** An inverse problem very relevant in imaging and closely related to the above parameter identification problem of ground water filtration is impedance tomography. In its simplest form, the mathematical description of the forward process can be described as the solution of the elliptic partial differential equation

$$\mathrm{div}(u\nabla g) = 0 \qquad \text{in} \quad \Omega \tag{5.2}$$

where g is the electric potential and u the conductivity, both modelled as functions in some function space over a spatial domain Ω. On the boundary $\partial\Omega$, the electric potential g directly relates to the voltages h applied to the system, i.e. we have

$$g = h \qquad \text{on} \quad \partial\Omega \,.$$

The measured currents over the boundary for a specific voltage function h are then given as

$$f_h = u\frac{\partial g}{\partial n} \qquad \text{on} \quad \partial\Omega \,.$$

Here $\frac{\partial g}{\partial n}$ denotes the normal derivative of g. Hence, if the functions u and h are given (in suitable function spaces), the forward problem consists of applying the so-called Dirichlet-to-Neumann map

$$\Lambda_u : g \mapsto f_h \,,$$

which is a linear operator due to the linearity of (5.2) for given u. The inverse problem of impedance tomography, also known as the inverse conductivity problem, aims at reconstructing the conductivity u on the whole domain Ω based on a known voltage function h and measured currents f_h on $\partial\Omega$. Due to the multiplication of u with g in (5.2) the inverse problem is automatically non-linear.

For the setup (5.1), all regularisation approaches discussed in the previous part of this lecture are useless, due to the non-linearity of F. However, we can pick a specific regularisation strategy and try to adapt it to the non-linear case, similar to the generalisations of Tikhonov regularisation in Section 4. Due to its explicit nature, we are going to pick the Landweber iteration as introduced in Section 3.2.6, and define the non-linear Landweber iteration as follows:

$$u^{k+1} = u^k - \tau^k (F'(u^k))^* \left(F(u^k) - f^\delta \right) \,, \tag{5.3}$$

where $\{\tau_k\}_{k\in\mathbb{N}}$ is a sequence of positive parameters, and $(F'(u^k))^*$ is the adjoint operator of the Fréchet derivative of F at u^k. It is straight forward to see that (5.3) is nothing else but gradient descent applied to the minimisation of the (generally non-convex) functional

$$E(u) := \frac{1}{2}\|F(u) - f^\delta\|_{\mathcal{V}}^2 \,.$$

In the following we want to prove that (5.3) does converge to a critical point $\hat{u}$ of E, i.e. $0 = E'(\hat{u}) = (F'(\hat{u}))^*(F(\hat{u}) - f^\delta)$. In order to do so, we need to make some assumptions on E (respectively on F) first. The first assumption that is fairly standard is that the Fréchet derivative of the functional E is Lipschitz continuous, i.e. there exists a constant $0 < L < \infty$ such that the inequality

$$\|E'(u) - E'(v)\|_{\mathcal{U}} \leq L\|u - v\|_{\mathcal{U}} \tag{5.4}$$

is satisfied for all $u, v \in \mathcal{U}$.

Given that the domain of E is convex, which is certainly true if the domain is a Hilbert space, we can conclude the following useful result.

Lemma 5.1. *Let $\mathcal{U}$ be a Hilbert space, $E : \mathcal{U} \to \mathbb{R}$ be a Fréchet-differentiable functional, and let E' be Lipschitz continuous with constant L. Then the functional $G(u) := \frac{L}{2}\|u\|_{\mathcal{U}}^2 - E(u)$ is convex.*

Proof. From the Cauchy-Schwarz inequality and (5.4) we obtain

$$\langle E'(u) - E'(v), u - v\rangle \leq L\|u - v\|_{\mathcal{U}}^2 \,,$$

which we can rewrite to

$$0 \leq \langle Lu - E'(u) - (Lv - E'(v)), u - v\rangle = D_G^{\text{symm}}(u, v) \,,$$

due to the convexity of $\mathcal{U}$, $\text{dom}(G) = \mathcal{U}$ and $G'(u) = Lu - E'(u)$ for all $u \in \mathcal{U}$. Given that D_G^{symm} is non-negative it makes it a symmetric Bregman distance, which further implies convexity of G. □

Note that convexity of $G(u) = \frac{L}{2}\|u\|_{\mathcal{U}} - E(u)$ implies the very useful Lipschitz estimate

$$E(u) \leq E(v) + \langle E'(v), u - v\rangle + \frac{L}{2}\|u - v\|_{\mathcal{U}}^2 \,, \tag{5.5}$$

for all $u, v \in \mathcal{U}$.

Next we want to assume that E satisfies what is known as the Kurdyka-Łojasiewicz (KL) property. Let $\eta \in]0, \infty]$. We consider functions $\varphi : [0, \eta[\to \mathbb{R}_{\geq 0}$ of the class of all concave and continuous functions that satisfy

(a) $\varphi(0) = 0$.

(b) φ is C^1 on $]0, \eta[$ and continuous at 0.

(c) $\varphi'(s) > 0$ for all $s \in]0, \eta[$.

Then the KL property is defined as follows.

Definition 5.1 (Kurdyka-Łojasiewicz property)**.** *Let $E : \mathcal{U} \to \mathbb{R}$ be a Fréchet-differentiable functional with its Fréchet derivative E' being well-defined for all $u \in \mathcal{U}$.*

(a) The functional E fulfills the KL property at point $\overline{u} \in \mathcal{U}$ if there exists $\eta \in]0, \infty[$, a neighbourhood U of $\overline{u}$ and a function φ satisfying the conditions above, such that for all

$$u \in U \cap \{u \mid E(\overline{u}) < E(u) < E(\overline{u}) + \eta\}$$

the inequality

$$\varphi'(E(u) - E(\overline{u}))\|E'(u)\|_{\mathcal{U}} \geq 1 \tag{5.6}$$

is satisfied.

(b) If E satisfies the KL property at each point $\overline{u} \in \mathcal{U}$, E is called a KL functional.

Example 5.3. In order to clarify what (5.6) means we want to consider the classic example from Stanisław Łojasiewicz who considered functions φ of the form $\varphi(x) = \frac{1}{1-\theta}|x|^{1-\theta}$ for $0 < \theta < 1$. It is easy to see that this choice of φ satisfies the above conditions, and that (5.6) in this case transforms to

$$|E(u) - E(\overline{u})|^{\theta} \leq \|E'(u)\|_{\mathcal{U}} \,.$$

Hence, we obtain that the Fréchet-derivative of E is bounded from below by its functional values, which is a very useful estimate as it implies that the closer we get to a critical point, the closer we also get to the functional evaluation of this critical point.

Example 5.4 (KL functions)**.** If we assume for a moment that E is simply a function, it can be shown that surprisingly many functions E already satisfy (5.6). Functions of the following classes are known to satisfy (5.6) (see [9]):

(a) Semi-algebraic, e.g. $E(x, y) = \sqrt{x^4 + y^4}$.

(b) Globally subanalytic, e.g. $E(x, y) = \frac{y}{\sin x}$, for $x \in]0, \pi[$.

(c) $(\mathbb{R}, \exp)$-definable, e.g. $f(x,y) = x^2 \exp\left(-\frac{y^2}{x^4+y^2}\right) \ln x$.

(d) $(\mathbb{R}_{an}, \exp)$-definable, e.g. $E(x,y) = x^{\sqrt{2}} \ln(\sin y)$.

(e) $(\mathbb{R}_{an}^{\mathbb{R}})$-definable, e.g. $E(x,y) = x^{\sqrt{2}} \exp\left(\frac{x}{y}\right)$, $0 < x < y < 1$.

Bolte et al. have further transferred the KL condition concept from functions to function spaces (cf. [5]), which is why we can also apply the concept to functionals.

Assuming that E is a KL functional with Lipschitz continuous Fréchet-derivative E' with Lipschitz constant L, we now want to prove that (5.3) satisfies the three following properties:

(a) We can find a positive constant ρ_1 such that the *sufficient decrease property*

$$\rho_1 \|u^{k+1} - u^k\|_{\mathcal{U}}^2 \leq E(u^k) - E(u^{k+1}) \qquad \forall k = 0, 1, \ldots$$

is satisfied.

(b) Assume that the sequence (5.3) is bounded. Then we can find a positive constant ρ_2 such that the *gradient lower bound for the iterates gap*, i.e.

$$\|\nabla E(u^k)\|_{\mathcal{U}} \leq \rho_2 \|u^{k+1} - u^k\|_{\mathcal{U}} \qquad \forall k = 0, 1, \ldots\,,$$

holds true.

(c) Together with the *KL property* we then show that the generated sequence $\{u^k\}_{k \in \mathbb{N}}$ is a Cauchy sequence.

These three properties will be sufficient to prove global convergence, i.e. regardless of how we start, the iterates of (5.3) will always converge to a critical point of E.

Theorem 5.1 (Sufficient decrease property). *Let $E : \mathcal{U} \to \mathbb{R}$ be a Fréchet-differentiable functional with locally Lipschitz continuous Fréchet-derivative E' with Lipschitz constant $L > 0$. If we choose $0 < \tau^k < 2/L$ such that*

$$\rho_1 \|u^{k+1} - u^k\|_{\mathcal{U}} \leq \left(\frac{1}{\tau^k} - \frac{L}{2}\right) \|u^{k+1} - u^k\|_{\mathcal{U}}^2 \tag{5.7}$$

holds true for all k and a fixed constant $0 < \rho < \infty$, then the iterates of the non-linear Landweber iteration (5.3) *satisfy the descent estimate*

$$E(u^{k+1}) + \rho_1 \|u^{k+1} - u^k\|_{\mathcal{U}}^2 \leq E(u^k)\,. \tag{5.8}$$

In addition, we observe

$$\lim_{k \to \infty} \|u^{k+1} - u^k\|_{\mathcal{U}} = 0\,.$$

Proof. First of all, we rewrite the Landweber iteration (5.3) in terms of E', i.e

$$\tau^k E'(u^k) + u^{k+1} - u^k = 0\,,$$

and then take the Hilbert space inner product with $u^{k+1} - u^k$, which yields

$$\tau^k \langle E'(u^k), u^{k+1} - u^k \rangle + \|u^{k+1} - u^k\|_{\mathcal{U}}^2 = 0\,.$$

Thus, we obtain

$$\langle E'(u^k), u^{k+1} - u^k \rangle = -\frac{1}{\tau^k} \|u^{k+1} - u^k\|_{\mathcal{U}}^2 \,. \tag{5.9}$$

Due to the local Lipschitz-continuity of E' we can use (5.5) to further estimate

$$E(u^{k+1}) \leq E(u^k) + \langle u^{k+1} - u^k, E'(u^k) \rangle + \frac{L}{2} \|u^{k+1} - u^k\|_{\mathcal{U}}^2 \,.$$

Together with (5.9) we therefore obtain the estimate

$$E(u^{k+1}) + \left(\frac{1}{\tau^k} - \frac{L}{2} \right) \|u^{k+1} - u^k\|_{\mathcal{U}}^2 \leq E(u^k) \,.$$

Using (5.7) then allows us to conclude

$$0 \leq \rho_1 \|u^{k+1} - u^k\|_{\mathcal{U}}^2 \leq E(u^k) - E(u^{k+1}) \,;$$

hence, summing up over all N iterates and telescoping yields

$$\begin{aligned} \sum_{k=0}^{N} \rho_1 \|u^{k+1} - u^k\|_{\mathcal{U}}^2 &\leq \sum_{k=0}^{N} E(u^k) - E(u^{k+1}) \,, \\ &= E(u^0) - E(u^{N+1}) \,, \\ &\leq E(u^0) - \overline{E} < \infty \,. \end{aligned}$$

Taking the limit $N \to \infty$ therefore implies

$$\sum_{k=0}^{\infty} \rho_1 \|u^{k+1} - u^k\|_{\mathcal{U}}^2 < \infty \,,$$

and thus, we have $\lim_{k\to\infty} \|u^{k+1} - u^k\|_{\mathcal{U}}^2 = 0$ due to $\rho_1 > 0$. □

The next step is to show that $E'(u^k)$ is bounded from above for every k by some positive multiple of $\|u^{k+1} - u^k\|_{\mathcal{U}}$, as we then can also conclude $\|E'(u^k)\|_{\mathcal{U}} \to 0$.

Theorem 5.2 (A gradient lower bound for the iterates gap)**.** *The iterates of the non-linear Landweber iteration* (5.3) *satisfy*

$$\|E'(u^k)\|_{\mathcal{U}} \leq \rho_2 \|u^{k+1} - u^k\|_{\mathcal{U}} \,, \tag{5.10}$$

for $\rho_2 := 1/\overline{\tau}$ *and* $0 < \overline{\tau} := \min_k \tau^k$.

Proof. From (5.3) we trivially observe

$$\|E'(u^k)\|_{\mathcal{U}} = \frac{1}{\tau^k} \|u^{k+1} - u^k\|_{\mathcal{U}} \leq \rho_2 \|u^{k+1} - u^k\|_{\mathcal{U}} \,,$$

for ρ_2 as defined above. □

Together with the KL property we now have everything that is necessary in order to prove global convergence of (5.3).

Theorem 5.3 (Global convergence)**.** *Suppose that E is a weakly continuous KL functional in the sense of Definition 5.1, with Lipschitz continuous Fréchet-derivative E' with Lipschitz constant $L > 0$. Let $\{u^k\}_{k\in\mathbb{N}}$ be a sequence generated by* (5.3)*, which is further assumed to be bounded. Then the sequence $\{u^k\}_{k\in\mathbb{N}}$ has a strongly convergent subsequence that converges to a critical point $\hat{u}$ with $E'(\hat{u}) = 0$.*

Proof. Given that $\{u^k\}_{k\in\mathbb{N}}$ is assumed to be bounded, we know that there exists a weakly convergent subsequence $\{u^{k_j}\}_{j\in\mathbb{N}}$ with $u^{k_j} \rightharpoonup \overline{u}$. Since E is assumed to be weakly continuous we also have

$$\lim_{j\to\infty} E(u^{k_j}) = E(\overline{u})\,. \tag{5.11}$$

If there exists an index q such that $E(u^{k_q}) = E(\overline{u})$, then (5.8) already implies $u^{k_{q+1}} = u^{k_q}$, and we can show via induction that the sequence $\{u^{k_j}\}_{j\in\mathbb{N}}$ is stationary, trivially implying finite length and convergence to a critical point.

If such an index does not exist, we know from (5.8) that $\{E(u^{k_j})\}_{j\in\mathbb{N}}$ is a non-increasing sequence, therefore (5.11) implies $E(\overline{u}) < E(u^{k_j})$ for all $j > 0$. We also know from (5.11) that there exist $j_1 \in \mathbb{N}$ and $\eta \in]0,\infty[$ such that $E(u^{k_j}) < E(\overline{u})+\eta$ for all $j > j_1$. Due to the convergence of the sub-sequence $\{u^{k_j}\}_{j\in\mathbb{N}}$ there also exists a j_2 such that $\|u^{k_j} - \overline{u}\|_{\mathcal{U}} < \varepsilon$ for all $j > j_2$. Hence, the sequence u^{k_j} belongs to the intersection $\{u \mid \|u - \overline{u}\|_{\mathcal{U}} < \varepsilon\} \cap \{u \mid E(\overline{u}) < E(u) < E(\overline{u}) + \eta\}$ for all $j > l := \max(j_1, j_2)$, which implies (5.6) for all $j > l$.

(a) Due to the previous considerations we have for any $j > l$, for some $l \in \mathbb{N}$, the estimate

$$\varphi'(E(u^{k_j}) - E(\overline{u}))\|E'(u^{k_j})\|_{\mathcal{U}} \geq 1\,,$$

which makes sense due to $E(u^{k_j}) > E(\overline{u})$ for all $j > l$. We then obtain from (5.10) the estimate

$$\varphi'(E(u^{k_j}) - E(\overline{u})) \geq \rho_2^{-1}\|u^{k_{j+1}} - u^{k_j}\|_{\mathcal{U}}^{-1}\,.$$

From the concavity of φ we also estimate

$$\frac{\varphi(E(u^{k_j}) - E(\overline{u})) - \varphi(E(u^{k_{j+1}}) - E(\overline{u}))}{E(u^{k_j}) - E(u^{k_{j+1}})} \geq \varphi'(E(u^{k_j}) - E(\overline{u}))\,.$$

Together with the previous estimate, this yields

$$\frac{\varphi(E(u^{k_j}) - E(\overline{u})) - \varphi(E(u^{k+1}) - E(\overline{u}))}{E(u^{k_j}) - E(u^{k_{j+1}})} \geq \rho_2^{-1}\|u^{k_{j+1}} - u^{k_j}\|_{\mathcal{U}}^{-1}\,.$$

With Theorem 5.1 we can therefore conclude

$$\frac{\varphi(E(u^{k_j}) - E(\overline{u})) - \varphi(E(u^{k_{j+1}}) - E(\overline{u}))}{\rho_1\|u^{k_{j+1}} - u^{k_j}\|_{\mathcal{U}}^2} \geq \rho_2^{-1}\|u^{k_{j+1}} - u^{k_j}\|_{\mathcal{U}}^{-1}\,,$$

respectively

$$\frac{\rho_1}{\rho_2}\|u^{k_{j+1}} - u^{k_j}\|_{\mathcal{U}} \leq \varphi(E(u^{k_j}) - E(\overline{u})) - \varphi(E(u^{k_{j+1}}) - E(\overline{u}))\,.$$

Summing up from $j = 0$ to $j = N$ then yields

$$\sum_{j=0}^{N} \|u^{k_{j+1}} - u^{k_j}\|_{\mathcal{U}} \leq \frac{\rho_1}{\rho_2} \Big(\varphi(E(u^{k_0}) - E(\overline{u})) - \varphi(E(u^{k_{N+1}}) - E(\overline{u})) \Big)$$
$$\leq \frac{\rho_1}{\rho_2} \varphi(E(u^{k_0}) - E(\overline{u})) < \infty \,.$$

Hence, we can conclude $\sum_{j=0}^{\infty} \|u^{k_{j+1}} - u^{k_j}\|_{\mathcal{U}} < \infty$ by taking the limit $N \to \infty$.

(b) The property $\sum_{j=0}^{\infty} \|u^{k_{j+1}} - u^{k_j}\|_{\mathcal{U}} < \infty$ implies that each u^{k_r} and u^{k_s} with $s > r > l$ are bounded w.r.t the $\mathcal{U}$ norm. This follows from

$$\|u^{k_r} - u^{k_s}\|_{\mathcal{U}} = \left\| \sum_{j=r}^{s-1} u^{k_{j+1}} - u^{k_j} \right\|_{\mathcal{U}} \leq \sum_{j=r}^{s-1} \|u^{k_{j+1}} - u^{k_j}\|_{\mathcal{U}}$$

Since $\sum_{j=0}^{\infty} \|u^{k_{j+1}} - u^{k_j}\|_{\mathcal{U}} < \infty$ implies $\lim_{l\to\infty} \sum_{j=l+1}^{\infty} \|u^{k_{j+1}} - u^{k_j}\|_{\mathcal{U}} = 0$, we can conclude that $\{u^{k_j}\}_{j\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{U}$. As the set of limit points of a Fréchet-differentiable functional is non-empty this concludes the proof.

□

Remark 5.1. For $\dim(\mathcal{U}) < \infty$ Theorem 5.3 can be extended to a global (strong) convergence result for all $\{u^k\}_{k\in\mathbb{N}}$. In that case it is obviously also sufficient to just assume that E is Fréchet-differentiable and therefore continuous, and not just weakly continuous.

As we know from the case of linear F, converging to a critical point is obviously only desirable if $f^\delta \in \mathcal{D}(F^\dagger)$. For non-linear F it is not even clear how to extent the concept of generalised inverses to make sense of an expression such as $f^\delta \in \mathcal{D}(F^\dagger)$. But even if we were, it is quite unlikely that f^δ would satisfy such a smoothness condition, which is why the iteration has to be stopped after a finite number of iterations. Thanks to Theorem 5.1 it seems reasonable to use Morozov's discrepancy principle (3.10) as a stopping criterion, similar to Chapter 3.2.6. With additional restrictions on the non-linearity of F we can show that (5.3) together with (3.10) satisfies a descent result in analogy to Lemma 3.5. In order to do so, we need to establish the definition of a neighbourhood first.

Definition 5.2. *A* neighbourhood $B_r(u_0)$ *of* $u_0 \in \mathcal{U}$ *is defined as the set*

$$B_r(u_0) := \{u \in \mathcal{U} \mid \|u_0 - u\|_{\mathcal{U}} \leq r\} \subset \mathcal{U}\,,$$

for a positive constant $r > 0$*.*

Now we can start proving a result similar to Lemma 3.5 under the additional assumption of the tangential cone condition.

Lemma 5.2. *Let* $F : \mathcal{U} \to \mathcal{H}$ *be smooth and continuous. Assume that for* $u_0 \in \mathcal{U}$ *there exists* $r > 0$ *with* $B_{2r}(u_0) \subset \mathcal{U}$ *such that a solution* $u^\dagger \in B_r(u_0)$ *to* $F(u^\dagger) = f$ *with* $\|f - f^\delta\|_{\mathcal{H}}$ *exists, and that for all* $u, \tilde{u} \in B_{2r}(u_0)$ *the conditions*

$$\|F'(u)\|_{\mathcal{L}(\mathcal{U},\mathcal{H})} \leq 1\,, \tag{5.12}$$

$$\|F(u) - F(\tilde{u}) - F'(\tilde{u})(u - \tilde{u})\|_{\mathcal{H}} \leq \mu \|F(u) - F(\tilde{u})\|_{\mathcal{H}} \qquad \text{for } 0 < \mu < \frac{1}{2}, \tag{5.13}$$

are met. If u_k^δ *as an iterate of* (5.3) *satisfies* $u_k^\delta \in B_r(u^\dagger)$ *for* $\delta \geq 0$ *and*

$$\|F(u_k^\delta) - f^\delta\|_{\mathcal{H}} \geq 2\frac{1+\mu}{1-2\mu}\delta\,, \tag{5.14}$$

then this immediately implies

$$\|u_{k+1}^\delta - u^\dagger\|_{\mathcal{U}} \leq \|u_k^\delta - u^\dagger\|_{\mathcal{U}}\,,$$

and consequently $u_{k+1}^\delta \in B_r(u^\dagger) \subset B_{2r}(u_0)$.

Proof. Using (5.3) we obtain

$$\begin{aligned}
\|u_{k+1}^\delta - u^\dagger\|_{\mathcal{U}}^2 - \|u_k^\delta - u^\dagger\|_{\mathcal{U}}^2 &= 2\langle u_{k+1}^\delta - u_k^\delta, u_k^\delta - u^\dagger\rangle + \|u_{k+1}^\delta - u_k^\delta\|_{\mathcal{U}}^2 \\
&= 2\langle F'(u_k^\delta)^*(f^\delta - F(u_k^\delta)), u_k^\delta - u^\dagger\rangle + \|F'(u_k^\delta)^*(f^\delta - F(u_k^\delta))\|_{\mathcal{U}}^2 \\
&\leq 2\langle f^\delta - F(u_k^\delta), F'(u_k^\delta)(u_k^\delta - u^\dagger)\rangle + \|f^\delta - F(u_k^\delta)\|_{\mathcal{H}}^2 \\
&= 2\langle f^\delta - F(u_k^\delta), f^\delta - F(u_k^\delta) + F'(u_k^\delta)(u_k^\delta - u^\dagger)\rangle - \|f^\delta - F(u_k^\delta)\|_{\mathcal{H}}^2 \\
&\leq \|f^\delta - F(u_k^\delta)\|_{\mathcal{H}} \Big(2\|f^\delta - F(u_k^\delta) + F'(u_k^\delta)(u_k^\delta - u^\dagger)\|_{\mathcal{H}} \\
&\qquad - \|f^\delta - F(u_k^\delta)\|_{\mathcal{H}}\Big)\,,
\end{aligned}$$

where we have made use of (5.12). Applying the triangular inequality to the first term in the bracket yields

$$\begin{aligned}
\|f^\delta - F(u_k^\delta) + F'(u_k^\delta)(u_k^\delta - u^\dagger)\|_{\mathcal{H}} &\leq \delta + \|F'(u_k^\delta)(u_k^\delta - u^\dagger)\|_{\mathcal{H}}\,, \\
&\leq \delta + \mu\|F(u_k^\delta) - F(u^\dagger)\|_{\mathcal{H}}\,, \\
&\leq (1+\mu)\delta + \mu\|F(u_k^\delta) - f^\delta\|_{\mathcal{H}}\,,
\end{aligned}$$

thanks to $F(u^\dagger) = f$ and condition (5.13). Hence, we obtain the overall estimate

$$\|u_{k+1}^\delta - u^\dagger\|_{\mathcal{U}}^2 - \|u_k^\delta - u^\dagger\|_{\mathcal{U}}^2 \leq \|f^\delta - F(u_k^\delta)\|_{\mathcal{H}} \Big(2(1+\mu)\delta - (1-2\mu)\|f^\delta - F(x_k^\delta)\|_{\mathcal{H}}\Big) \leq 0\,,$$

due to (5.14). □

Theorem 5.4. *Let the same assumptions hold true as in Lemma 5.2. We further assume that the stopping index* $k^*(\delta, f^\delta)$ *is chosen according to the discrepancy principle* (3.10) *with* η *satisfying*

$$2 < 2\frac{1+\mu}{1-2\mu} < \eta\,.$$

Then we have

$$k^*(\delta, f^\delta) < C\delta^{-2}\,,$$

for a constant $C > 0$.

Proof. The initial $u_0^\delta = u_0 \in B_{2r}(u_0)$ and the choice of η allow us to apply Lemma 5.2; we particularly obtain the estimate

$$\|u_{k+1}^\delta - u^\dagger\|_{\mathcal{U}}^2 - \|u_k^\delta - u^\dagger\|_{\mathcal{U}}^2 < \|f^\delta - F(u_k^\delta)\|_{\mathcal{H}} \left(\frac{2}{\mu}(1+\mu)\delta - (1-2\mu)\right)$$

for all $k < k^*(\delta, f^\delta)$. Summing up over all $k \in \{0, \dots, k^*(\delta, f^\delta)\}$ therefore yields

$$\left(1 - 2\mu - \frac{2}{\eta}(1+\mu)\right) \sum_{k=0}^{k^*(\delta,f^\delta)-1} \|F(u_k^\delta) - f^\delta\|_{\mathcal{H}}^2 < \|u_0 - u^\dagger\|_{\mathcal{U}}^2 - \|u_{k^*(\delta,f^\delta)}^\delta - u^\dagger\|_{\mathcal{U}}^2 .$$

Since we have $\|F(u_k^\delta) - f^\delta\|_{\mathcal{H}} > \eta\delta$ for all $k < k^*(\delta, f^\delta)$, we can further estimate

$$k^*(\delta, f^\delta)\eta^2\delta^2 < \sum_{k=0}^{k^*(\delta,f^\delta)-1} \|F(u_k^\delta) - f^\delta\|_{\mathcal{H}}^2 < \left(1 - 2\mu - \frac{2}{\eta}(1+\mu)\right)^{-1} \|u_0 - u^\dagger\|_{\mathcal{U}}^2 ,$$

which yields the desired estimate for $C := 1/((1-2\mu)\eta^2 - 2(1+\mu)\eta)\|u_0 - u^\dagger\|_{\mathcal{U}}^2 > 0$. □

In order to show that the non-linear Landweber iteration in combination with Morozov's discrepancy principle is behaving like the equivalent of convergent regularisation methods for non-linear operators we need to verify the following lemma.

Lemma 5.3. *Let $k^*(\delta, f^\delta)$ be chosen according to the discrepancy principle* (3.10)*. If an iterative method satisfies*

$$\tilde{k} := \tilde{k}(0, f) < \infty, \quad u_{\tilde{k}} = u^\dagger \qquad \textit{or} \qquad \tilde{k} = \infty, \quad u_k \to u^\dagger \textit{ for } k \to \infty, \tag{5.15}$$

and the two conditions

$$\|u_k^\delta - u^\dagger\|_{\mathcal{U}} \leq \|u_{k-1}^\delta - u^\dagger\|_{\mathcal{U}} , \tag{5.16}$$

$$\lim_{\delta \to 0} \|u_k^\delta - u_k\|_{\mathcal{U}} = 0 , \tag{5.17}$$

for all $k \in \{1, \dots, k^*(\delta, f^\delta)\}$, then it also satisfies

$$\limsup_{\delta \to 0} \left\{ \left\| u_{k^*(\delta,f^\delta)} - u^\dagger \right\|_{\mathcal{U}} \mid f^\delta \in \mathcal{H}, \|f^\delta - f\|_{\mathcal{H}} \leq \delta \right\} . \tag{5.18}$$

Proof. We know by assumption that $F : \mathcal{U} \to \mathcal{H}$ is continuous. Let $\{f^{\delta_j}\}_{j \in \mathbb{N}} \subset \mathcal{H}$ with $\|f - f^{\delta_j}\}_{\mathcal{H}} \leq \delta_j$ and $\delta_j \to 0$ for $j \to \infty$, and define $k_j^* := k^*(\delta_j, f^{\delta_j})$.

We first investigate the case for which $\{k_j^*\}_{j \in \mathbb{N}}$ has a finite limiting point $k^* < \infty$. With a subsequence argument we can argue $k_j^* = k^*$ for all $j \in \mathbb{N}$, and therefore $u_{k^*}^{\delta_j} \to u_{k^*}$ for $j \to \infty$ due to (5.17). Since each k_j^* are chosen according to the discrepancy principle (3.10), we further know

$$\|F(u_{k^*}^{\delta_j}) - f^{\delta_j}\|_{\mathcal{H}} \leq \eta\delta_j$$

for all $j \in \mathbb{N}$. Taking the limit $j \to \infty$ on both sides yields $F(u_{k^*}) = f$, due to the continuity of F, which already implies (5.18).

Otherwise we have $k_j^* \to \infty$. With another subsequence argument we can assume that k_j^* is monotonically increasing. From (5.16) we therefore obtain

$$\|u_{k_j^*}^{\delta_j} - u^\dagger\|_{\mathcal{U}} \leq \|u_{k_i^*}^{\delta_j} - u^\dagger\|_{\mathcal{U}} \leq \|u_{k_i^*}^{\delta_j} - u_{k_i^*}\|_{\mathcal{U}} + \|u_{k_i^*} - u^\dagger\|_{\mathcal{U}}$$

for all $i \leq j$. Let $\varepsilon > 0$ be arbitrary, then there exists $M > 0$ such that $\|u_{k_M^*} - u^\dagger\|_{\mathcal{U}} \leq \frac{\varepsilon}{2}$, due to (5.15). On the other hand we can also conclude the existence of an index $N > 0$ such that $\|u_{k_M^*}^{\delta_j} - u_{k_M^*}\|_{\mathcal{U}} \leq \frac{\varepsilon}{2}$ is satisfied for all $j \geq N$, due to (5.17) for $k = k_M^*$. This also implies (5.18), which concludes the proof. □

Proposition 5.1. *Without proof we want to state that for noise-free data and the assumption of the conditions* (5.12) *and* (5.13) *we are not just able to show strong convergence of sub-sequences as in Theorem 5.3, but strong convergence for the entire sequences.*

Theorem 5.5. *Let the same assumptions hold true as in Lemma 5.2. Then* $u_{k^*(\delta,f^\delta)} \to u^\dagger$ *for* $F(u^\dagger) = f$ *and* $\delta \to 0$.

Proof. We apply Lemma 5.3. Condition (5.15) follows from Proposition 5.1. Since F and F' are assumed to be continuous, we further know that the right-hand-side of (5.3) depends continuously on u_k, for fixed $k \in \mathbb{N}$. Hence, for $\delta \to 0$ the right-hand-side of (5.3) for u_{j+1}^δ converges to the right-hand-side of (5.3) for u_{j+1}, for all $j \leq k$, which implies $u_{j+1}^\delta \to u_{j+1}$ and consequently (5.17). As the monotonicity (5.16) follows from Lemma 5.2, we can further conclude (5.18). □

Bibliography

[1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces.* Elsevier Science, Singapore, 2003.

[2] Anatolii Borisovich Bakushinskii. Remarks on the choice of regularization parameter from quasioptimality and relation tests. *Zhurnal Vychislitel'noï Matematiki i Matematicheskoï Fiziki*, 24(8):1258–1259, 1984.

[3] Heinz H Bauschke and Patrick L Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* 2011.

[4] Béla Bollobás. *Linear Analysis: An Introductory Course.* Cambridge University Press, Cambridge, second edi edition, 1999.

[5] Jérôme Bolte, Aris Daniilidis, Olivier Ley, and Laurent Mazet. Characterizations of łojasiewicz inequalities: subgradient flows, talweg, convexity. *Transactions of the American Mathematical Society*, 362(6):3319–3363, 2010.

[6] Enrico Giusti. *Minimal Surfaces and Functions of Bounded Variation.* Birkhaeuser, Basel, Boston, Stuttgart, 1984.

[7] Charles W Groetsch. *Stable approximate evaluation of unbounded operators.* Springer, 2006.

[8] John Hunter and Bruno Nachtergaele. *Applied Analysis.* World Scientific Publishing Company Incorporated, 2001.

[9] Krzysztof Kurdyka. On gradients of functions definable in o-minimal structures. In *Annales de l'institut Fourier*, volume 48, pages 769–783, 1998.

[10] Arch W Naylor and George R Sell. *Linear Operator Theory in Engineering and Science.* Springer Science & Business Media, 2000.

[11] A. Rieder. *Keine Probleme mit Inversen Problemen: Eine Einführung in ihre stabile Lösung.* Vieweg+Teubner Verlag, 2003.

[12] Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[13] W. Rudin. *Functional Analysis.* International series in pure and applied mathematics. McGraw-Hill, 1991.

[14] Terence Tao. *Epsilon of Room, One*, volume 1. American Mathematical Soc., 2010.

[15] Eberhard Zeidler. *Applied Functional Analysis: Applications to Mathematical Physics*, volume 108 of *Applied Mathematical Sciences Series.* Springer, 1995.

[16] Eberhard Zeidler. *Applied Functional Analysis: Main Principles and Their Applications*, volume 109 of *Applied Mathematical Sciences Series*. Springer, 1995.