Chapter 2
Functional Analysis Background of Ill-Posed Problems

The main objective of this chapter is to present some necessary results of functional analysis, frequently used in study of inverse problems. For simplicity, we will derive these results in Hilbert spaces. Let H be a vector space over the field of real ($\mathbb{R}$) or complex ($\mathbb{C}$) numbers. Recall that the mapping $(\cdot, \cdot)_H$ defined on $H \times H$ is called an *inner product* (or *scalar product*) of two elements of H, if the following conditions are satisfied:

$$
\begin{aligned}
&(i1)\ (u_1 + u_2, v)_H = (u_1, v)_H + (u_2, v)_H,\ \forall u_1, u_2, v \in H;\\
&(i2)\ (\alpha u, v)_H = \alpha(u, v)_H,\ \forall \alpha \in \mathbb{C},\ \forall u, v \in H;\\
&(i3)\ (u, v)_H = \overline{(v, u)_H},\ \forall u, v \in H;\\
&(i4)\ (u, u)_H \geq 0,\ \forall u \in H \ \text{ and } (u, u)_H = 0 \ \text{ iff } \ u = 0.
\end{aligned}
$$

Through the following, we will omit the subscript H in the scalar (dot) product and the norm whenever it is clear from the text.

The vector space H together with the inner product is called an *inner product space* or a *pre-Hilbert space*. The norm in a pre-Hilbert space is defined by the above introduced scalar product: $\|u\|_H := (u, u)^{1/2}$, $u \in H$. Hence a pre-Hilbert space H is a normed space. If, in addition, a pre-Hilbert space is complete, it is called a Hilbert space. Thus, a Hilbert space is a Banach space, i.e. complete normed space, with the norm defined via the scalar product.

A Hilbert space is called infinite dimensional (finite dimensional) if the underlying vector space is infinite dimensional (finite dimensional). The basic representative of infinite dimensional space Hilbert in the weak solution theory of PDEs is the space of square integrable functions

$$
L^2(a, b) := \{u : (a, b) \mapsto \mathbb{R} : \int_a^b u^2(x)dx < +\infty\},
$$

with the scalar product

A. Hasanov Hasanoğlu and V.G. Romanov, *Introduction to Inverse Problems for Differential Equations*, DOI 10.1007/978-3-319-62797-7_2

$$(u, v)_{L^2(a,b)} := \int_a^b u(x)v(x)dx,\ u, v \in L^2(a, b).$$

A finite dimensional analogue of this space of square-summable sequences in $\mathbb{R}^n$:

$$l^2 := \{x := (x_1, x_2, ..., x_n) \in \mathbb{R}^n :\ \sum_{k=1}^{n} x_k^2 < +\infty\},$$

with the scalar product

$$(x, y)_{l^2} := \sum_{k=1}^{n} x_k y_k,\ x, y \in \mathbb{R}^n.$$

2.1 Best Approximation and Orthogonal Projection

Let H be an inner product space. The elements $u, v \in H$ are called *orthogonal*, if $(u, v) = 0$. This property is denoted by $u \perp v$. Evidently, $0 \perp v$, for any $v \in H$, since $(0, v) = 0, \forall v \in H$. Let $U \subset H$ be a non-empty subset and $u \in H$ an arbitrary element. If $(u, v) = 0$ for all $v \in U$, then the element $u \in H$ is called orthogonal to the subset U and is denoted by $u \perp U$. The set of all elements $u \in H$ orthogonal to $U \subset H$ is called an *orthogonal complement* of U and is denoted by $U^\perp$:

$$U^\perp := \{u \in H :\ (u, v) = 0,\ \forall v \in U\}. \tag{2.1.1}$$

The subsets $U, V \subset H$ are called orthogonal subsets of H if $(u, v) = 0$ for all $u \in U$ and $v \in V$. This property is denoted by $U \perp V$. Evidently, if $U \perp V$, then $U \cap V = \{0\}$.

Theorem 2.1.1 *Let $U \subset H$ be a subset of an inner product space H. Then an orthogonal complement $U^\perp$ of U is a closed subspace of H. Moreover, the following properties are satisfied*:

(**p1**) $U \cap U^\perp \subset \{0\}$ *and* $U \cap U^\perp = \{0\}$ *iff* U *is a subspace*;
(**p2**) $U \subset \left(U^\perp\right)^\perp =: U^{\perp\perp},\ \left(U^\perp\right)^\perp := \{u \in H : (u, v) = 0,\ \forall v \in U^\perp\}$;
(**p3**) *If* $U_1 \subset U_2 \subset H$, *then* $U_2^\perp \subset U_1^\perp$.

Proof Evidently, $\alpha u + \beta v \in U^\perp$, for all $\alpha,\ \beta \in \mathbb{C}$ and $u, v \in U$, by the above definition of the scalar product: $(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w) = 0$, for all $w \in U$. This implies that $U^\perp$ is a subspace of H. It is easy to prove, by using continuity of the scalar product, that if $\{u_n\} \subset U^\perp$ and $u_n \to u$, as $n \to \infty$, then $u \in U^\perp$. So, $U^\perp$ is closed. To prove (**p1**), we assume that there exists an element $u \in U \cap U^\perp$. Then, by definition (2.1.1) of $U^\perp$, $(u, u) = 0$, which implies $u = 0$. Hence $U \cap U^\perp \subset \{0\}$.

If, in addition, U is a subspace of H, then $0 \in U$ and it yields $U \cap U^{\perp} = \{0\}$. To prove (**p2**) we assume in contrary, that there exists an element $u \in U$ such that $u \notin \left(U^{\perp}\right)^{\perp}$. This means an existence of such an element $v \in U^{\perp}$ that $(u, v) \neq 0$. On the other hand, for all $u \in U$ and $v \in U^{\perp}$, we have $(u, v) = 0$, which is a contradiction. Hence $U \subset \left(U^{\perp}\right)^{\perp}$. Finally, to prove (**p3**), let $v \in U_2^{\perp}$ be an arbitrary element. Then $(u, v) = 0$, for all $u \in U_2$. Since $U_1 \subset U_2$, this holds for all $u \in U_1$ as well. Hence for any element $v \in U_2^{\perp}$, the condition $(u, v) = 0$ holds for all $u \in U_1$. This implies, by the definition, that $v \in U_1^{\perp}$, which completes the proof. □

Definition 2.1.1 (*Best approximation*) Let $U \subset H$ be a subset of an inner product space H and $v \in H$ be a given element. If

$$\|v - u\|_H = \inf_{w \in U} \|v - w\|_H \,, \tag{2.1.2}$$

then $u \in U$ is called the best approximation to the element $v \in H$ with respect to the set $U \subset H$.

The right hand side of (2.1.2) is a distance between a given element $v \in H$ and the set $U \subset H$. Hence, the best approximation is an element with the smallest distance to the set $U \subset H$.

This notion plays a crucial role in inverse problems theory and applications, since a measured output data can only be given with some measurement error. First we will prove that if $U \subset H$ is a closed linear subspace, then the best approximation is determined uniquely. For this aim we will use the main theorem on quadratic variational problems and its consequence, called perpendicular principle [103].

Theorem 2.1.2 *Let $a : H \times H \mapsto \mathbb{R}$ be a symmetric, bounded, strongly positive bilinear form on a real Hilbert space H, and $b : H \mapsto \mathbb{R}$ be a linear bounded functional on H. Then*
(i) The minimum problem

$$J(v) = \min_{w \in H} J(w), \;\; J(w) := \frac{1}{2} a(w, w) - b(w) \tag{2.1.3}$$

has a unique solution $v \in H$.
(ii) This minimum problem is equivalent to the followig variational problem: Find $v \in H$ such that

$$a(v, w) = b(w), \;\; \forall w \in H. \tag{2.1.4}$$

We use this theorem to prove that in a closed linear subspace U of a real or complex Hilbert space H, the best approximation is uniquely determined.

Theorem 2.1.3 *Let U be a closed linear subspace of a Hilbert space H and $v \in H$ be a given element. Then the best approximation problem* (2.1.2) *has a unique solution $u \in U$. Moreover $v - u \in U^{\perp}$.*

Proof We rewrite the norm $\|v-u\|_H$ as follows:

$$\|v-u\|_H := (v,v)-(v,u)-(u,v)+(u,u) \\ = a(v,v)+2\left[\frac{1}{2}a(u,u)-b(u)\right], \tag{2.1.5}$$

where

$$a(u,w) := Re(u,w),\ b(u) := \frac{1}{2}[(v,u)+(u,v)] = Re(v,u)$$

The right hand side of (2.1.5) shows that the best approximation problem (2.1.5) is equivalent to the variational problem (2.1.3) with the above defined bilinear and linear forms. Then it follows from Theorem 2.1.2 that if H is a real Hilbert space, then there exists a unique best approximation $v \in U$ to the element $u \in H$.

If H is a complex Hilbert space, then we can introduce the new scalar product $(v,w)_* := Re(v,w),\ v,w \in H$ and again apply Theorem 2.1.2.

We prove now that if $u \in U$ is the best approximation to the element $v \in H$, then $v-u \perp U$. Indeed, it follows from (2.1.2) that

$$\|v-u\|_H^2 \le \|v-(u+\lambda w)\|_H^2,\ \forall\lambda\in\mathbb{C},\ w\in U.$$

This implies,

$$(v-u,v-u) \le (v-u,v-u)-\overline{\lambda}(v-u,w)-\lambda(w,v-u)+|\lambda|^2(w,w)$$

or

$$0 \le -\overline{\lambda}(v-u,w)-\lambda(w,v-u)+|\lambda|^2(w,w).$$

Assume that $v-u\neq 0,\ w\neq 0$. Then taking $\lambda=(w,v-u)/\|w\|^2$ we obtain: $0\le -|(v-u,w)|^2$, which means that $(v-u,w)=0,\ \forall w\in U$. Note that this remains true also if $v-u=0$. □

Corollary 2.1.1 (Orthogonal decomposition) Let U be a closed linear subspace of a Hilbert space H. Then there exists a unique decomposition of a given arbitrary element $v \in H$ of the form

$$v = u+w,\ u\in U,\ w\in U^{\perp}. \tag{2.1.6}$$

Existence of this orthogonal decomposition follows from Theorem 2.1.3. We prove the uniqueness. Assume, in contrary, that there exists another decomposition of $v \in H$ such that

$$v = u_1+w_1,\ u_1\in U,\ w_1\in U^{\perp}.$$

Since U is a linear subspace of H, we have $u - u_1 \in U$, $w - w_1 \in U^\perp$, and

$$(u - u_1) + (w - w_1) = 0.$$

Multiplying scalarly both sides by $u - u_1$ we get

$$(u - u_1, u - u_1) + (w - w_1, u - u_1) = 0,$$

which implies $u = u_1$, since the second term is zero due to $w - w_1 \in U^\perp$. In a similar way we conclude that $w = w_1$. This completes the proof. □

The orthogonal decomposition (2.1.6) can also be rewritten in terms of the subspaces U and $U^\perp$ as follows:

$$H = U \oplus U^\perp.$$

Example 2.1.1 Let $H := L^2(-1, 1)$, $U := \{u \in L^2(-1, 1) : \ u(-x) = u(x), \text{ a.e. } in\, (-1, 1)\}$ be the set of even functions and $U^\perp := \{u \in L^2(-1, 1) : \ u(-x) = -u(x), \text{ a.e. } in\, (-1, 1)\}$ be the set of odd functions. Then $L^2(-1, 1) = U \oplus U^\perp$. Remark that for any $v \in L^2(-1, 1)$, $v(x) = [v(x) + v(-x)]/2 + [v(x) - v(-x)]/2$.

Corollary 2.1.1 shows that there exists a mapping which uniquely transforms each element $v \in H$ to the element u of a closed linear subspace U of a Hilbert space H. This assertion is called *Orthogonal Projection Theorem*.

Definition 2.1.2 The operator $P : H \mapsto U$, with $Pv = u$, in the decomposition (2.1.6) which maps each element $v \in H$ to the element $u \in U$ is called the projection operator or orthogonal projection.

Using this definition, we may rewrite the best approximation problem (2.1.2) as follows:

$$\|v - Pv\|_H = \inf_{w \in U} \|v - w\|_H.$$

We denote by $\mathcal{N}(P) := \{v \in H : \ Pv = 0\}$ and $\mathcal{R}(P) := \{Pv : \ v \in H\}$ the null-space and the range of the projection operator P, respectively. The theorem below shows that the orthogonal projection $P : H \mapsto U$ is a linear continuous self-adjoint operator.

Theorem 2.1.4 *The orthogonal projection $P : H \mapsto U$ defined from Hilbert space H onto the closed subspace $U \subset H$ is a linear continuous self-adjoint operator with $P^2 = P$ and $\|P\| = 1$, for $U \neq \{0\}$. Conversely, if $P : H \mapsto H$ is a linear continuous self-adjoint operator with $P^2 = P$, then it defines an orthogonal projection from H onto the closed subspace $\mathcal{R}(P)$.*

Proof It follows from (2.1.6) that

$$\|v\|^2 := \|u + w\|^2 = \|u\|^2 + \|w\|^2,$$

since $(u, w) = 0$, by $u \in U$ and $w \in U^{\perp}$. This implies $\|Pv\| \leq \|v\|$, $Pv := u$ for all $v \in H$. In particular, for $v \in U \subset H$ we have $Pv = v$, which means $\|P\| = 1$. We prove now that P is self-adjoint. Let $v_k = u_k + w_k$, $u_k \in U$, $w_k \in U^{\perp}$, $k = 1, 2$. Multiplying both sides of $v_1 = u_1 + w_1$ by u_2 and both sides of $v_2 = u_2 + w_2$ by u_1, then taking into account $(u_k, v_m) = 0$, $k, m = 1, 2$, we conclude

$$(v_1, u_2) = (u_1, u_2), \ (v_2, u_1) = (u_2, u_1).$$

This implies $(u_1, v_2) = (u_1, u_2) = (v_1, u_2)$. Using the definition $Pv_k := u_k$ we deduce:

$$(Pv_1, v_2) := (v_1, Pv_2), \ \forall v_1, v_2 \in H,$$

i.e. the projection operator P is self-adjoint. Assuming now $v = u \in U \subset H$ in (2.1.6) we have: $u = u + 0$, where $0 \in U^{\perp}$. Then $Pu = u$, and for any $v \in H$ we have $P^2 v = P(Pv) = Pu = u = Pv$. Hence $P^2 v = Pv$, for all $v \in H$.

We prove now the second part of the theorem. Evidently, $\mathcal{R}(P) := \{Pv : \ v \in H\}$ is a linear subspace. We prove that the range of the projection operator $\mathcal{R}(P)$ is closed. Indeed, let $\{u_n\} \subset \mathcal{R}(P)$, such that $u_n \to u$, as $n \to \infty$. Then there exists such an element $v_n \in H$ that $u_n = Pv_n$. Together with the property $P^2 v_n = Pv_n$ this implies: $Pu_n = P^2 v_n = Pv_n = u_n$. Hence, $Pu_n = u_n$ for all $u_n \in \mathcal{R}(P)$. Letting to the limit and using the continuity of the operator P, we obtain:

$$u = \lim_{n \to \infty} u_n = \lim_{n \to \infty} Pu_n = Pu,$$

i.e. $Pu = u$, which means $u \in \mathcal{R}(P)$. Thus $\mathcal{R}(P)$ is a linear closed subspace of H and all its elements are fixed points of the operator P, that is, $Pu = u$ for all $u \in \mathcal{R}(P)$. On the other hand, P is a self-adjoint operator with $P^2 = P$, by the asumption. Then for any $v \in H$,

$$(Pv, (I - P)v) = (Pv, v) - (Pv, Pv) = (Pv, v) - (P^2 v, v) = 0.$$

Since $v \in H$ is an arbitrary element, the orthogonality $(Pv, (I - P)v) = 0$ means that $(I - P)v \in \mathcal{R}(P)^{\perp}$. Then the identity

$$v = Pv + (I - P)v, \ \forall v \in H$$

with Corollary 2.1.1 implies that P is a projection operator, since $Pv \in \mathcal{R}(P) \subset H$ and $(I - P)v \in \mathcal{R}(P)^{\perp}$. $\square$

Remark 2.1.1 Based on the properties of the orthogonal projection $P : H \mapsto H$, we conclude that the operator $I - P : H \mapsto H$ is also an orthogonal projection. Indeed, $(I - P)^2 = I - 2P + P^2 = I - P$.

Remark 2.1.2 Let us define the set $\mathcal{M} := \{u \in H : \ Pu = u\}$, i.e. the set of fixed points of the orthogonal projection P. It follows from the proof of Theorem 2.1.4 that $\mathcal{M} = \mathcal{R}(P)$. Similarly, $\mathcal{R}(I - P) = \mathcal{N}(P)$.

Some other useful properties of the projection operator $P : H \mapsto H$ are summarized in the following corollary.

Corollary 2.1.2 *Let $P : H \mapsto H$ be an orthogonal projection defined on a Hilbert space H. Then the following assertions hold:*
(**p1**) *$\mathcal{N}(P)$ and $\mathcal{R}(P)$ are closed linear subspaces of H.*
(**p2**) *Each element $v \in H$ can be written uniquely as the following decomposition:*

$$v = u + w, \ u \in \mathcal{R}(P), \ w \in \mathcal{N}(P). \tag{2.1.7}$$

Moreover,

$$\|v\|^2 = \|u\|^2 + \|w\|^2. \tag{2.1.8}$$

(**p3**) *$\mathcal{N}(P) = \mathcal{R}(P)^{\perp}$ and $\mathcal{R}(P) = \mathcal{N}(P)^{\perp}$.*

Proof The assertions (**p1**)$--$(**p2**) follow from Corollary 2.1.1 and Theorem 2.1.4. We prove (**p3**). Evidently, $\mathcal{N}(P) \perp \mathcal{R}(P)$ and $\mathcal{N}(P) \subset \mathcal{R}(P)^{\perp}$. Hence to prove the first part of the assertion (**p3**) we need to show that $\mathcal{R}(P)^{\perp} \subset \mathcal{N}(P)$. Let $v \in \mathcal{R}(P)^{\perp} \subset H$. Then, there exist such elements $u \in \mathcal{R}(P)$, $w \in \mathcal{N}(P)$ that $v = u + w$, according to (2.1.7). Multiplying both sides by an arbitrary element $\tilde{v} \in \mathcal{R}(P)$ we obtain: $(v, \tilde{v}) = (u, \tilde{v}) + (w, \tilde{v})$. The left hand side is zero, since $v \in \mathcal{R}(P)^{\perp}$. Also, $(w, \tilde{v}) = 0$, by $\mathcal{N}(P) \perp \mathcal{R}(P)$. Thus, $(u, \tilde{v}) = 0$, for all $\tilde{v} \in \mathcal{R}(P)$. But $u \in \mathcal{R}(P)$. This implies $u = 0$, and as a result, $v = 0 + w$, where $w \in \mathcal{N}(P)$. Hence $v \in \mathcal{N}(P)$.

The second part of the assertion (**p3**) can be proved similarly. □

We illustrate the above results in the following example.

Example 2.1.2 Fourier Series and Orthogonal Projection.

Let $\{\varphi_n\}_{n=1}^{\infty}$ be an orthonormal basis of an infinite-dimensional real Hilbert space H. Then any element $u \in H$ can be written uniquely as the following convergent Fourier series:

$$u = \sum_{n=1}^{\infty} (u, \varphi_n)\varphi_n \equiv \sum_{n=1}^{N} (u, \varphi_n)\varphi_n + \sum_{n=N+1}^{\infty} (u, \varphi_n)\varphi_n.$$

Let us consider the finite system $\{\varphi_n\}_{n=1}^{N}$ which forms a basis for the finite-dimensional Hilbert subspace $H_N \subset H$. We define the operator $P : H \mapsto H_N$ as follows:

$$Pu := \sum_{n=1}^{N} (u, \varphi_n)\varphi_n, \ u \in H. \tag{2.1.9}$$

Evidently, P is a linear bounded operator. Moreover, $P^2 = P$. Indeed,

$$P^2 u := P\left(\sum_{n=1}^{N}(u,\varphi_n)\varphi_n\right) = \sum_{m=1}^{N}\left(\sum_{n=1}^{N}(u,\varphi_n)\varphi_n,\varphi_m\right)\varphi_m$$
$$= \sum_{m=1}^{N}(u,\varphi_m)\varphi_m = Pu,$$

by $(\varphi_n,\varphi_m) = \delta_{n,m}$. Thus, $P : H \mapsto H_N$, defined by (2.1.9), is a projection operator from the infinite-dimensional Hilbert space H *onto* the finite-dimensional Hilbert space $H_N \subset H$, with $\mathcal{R}(P) = H_N$. To show the orthogonality of $\mathcal{R}(P)$ and $\mathcal{N}(P)$, let $u \in \mathcal{R}(P)$ and $v \in \mathcal{N}(P)$ be arbitrary elements. Then $Pu = u$ and

$$(v,u) = (v,Pu) = \left(v,\sum_{n=1}^{N}(u,\varphi_n)\varphi_n\right) = \sum_{n=1}^{N}(u,\varphi_n)(v,\varphi_n)$$
$$= \left(\sum_{n=1}^{N}(v,\varphi_n)\varphi_n,u\right) = (Pv,u).$$

But $Pv = 0$, due to $v \in \mathcal{N}(P)$. Hence $(v,u) = 0$, for all $u \in \mathcal{R}(P)$ and $v \in \mathcal{N}(P)$, which implies $\mathcal{R}(P) \perp \mathcal{N}(P)$.

Now we show the projection error, defined as

$$u - Pu := \sum_{n=N+1}^{\infty}(u,\varphi_n)\varphi_n,$$

is orthogonal to H_N. Let $v \in H_N$ be any element. Then

$$(u - Pu, v) := \left(\sum_{n=N+1}^{\infty}(u,\varphi_n)\varphi_n, v\right) = \sum_{n=N+1}^{\infty}(u,\varphi_n)(\varphi_n,v)$$

and the right hand side tends to zero as $N \to \infty$, due to the convergent Fourier series. Thus $u - Pu \perp H_N$ for all $v \in H_N$.

Finally, we use the above result to estimate the approximation error $\|u - v\|_{H_N}$, where $v \in H_N$ is an arbitrary element. We have:

$$\|u - v\|^2_{H_N} := (u - v, u - v) = (u - Pu + Pu - v, u - Pu + Pu - v)$$
$$= \|u - Pu\|^2_{H_N} + \|v - Pu\|^2_{H_N},$$

by (2.1.8). The right hand side has minimum value when $v = Pu$, i.e. $v \in H_N$ is a projection of $u \in H$. In this case we obtain:

$$\inf_{v \in H_N} \|u - v\|_{H_N} = \|u - Pu\|_{H_N},$$

which is the best approximation problem. □

2.2 Range and Null-Space of Adjoint Operators

Relationships between the null-spaces and ranges of a linear operator and its adjoint play an important role in inverse problems. The results given below show that the range of a linear bounded operator can be derived via the null-space of its adjoint. Note that for the linear bounded operator $A : H \mapsto \tilde{H}$, defined between Hilbert spaces H and $\tilde{H}$, the *adjoint operator* $A^* : \tilde{H} \mapsto H$ is defined as follows:

$$(Au, v)_{\tilde{H}} = (u, A^*v)_H, \ \ \forall u \in H, \ v \in \tilde{H}.$$

Theorem 2.2.1 *Let $A : H \mapsto \tilde{H}$ be a linear bounded operator, defined between Hilbert spaces H and $\tilde{H}$, and $A^* : \tilde{H} \mapsto H$ be its adjoint. Then*
(**p1**) $\mathcal{R}(A)^\perp = \mathcal{N}(A^*)$*;*
(**p2**) $\overline{\mathcal{R}(A)} = \mathcal{N}(A^*)^\perp$,
where $\mathcal{R}(A)$ and $\mathcal{N}(A^*)$ are the range and null-space of the operators A and A^*, correspondingly.

Proof Let $v \in \mathcal{N}(A^*)$. Then $A^*v = 0$, and for all $u \in H$ we have:

$$0 = (u, A^*v)_H = (Au, v)_{\tilde{H}}.$$

This implies that $v \in \mathcal{R}(A)^\perp$, i.e. $\mathcal{N}(A^*) \subset \mathcal{R}(A)^\perp$. Now suppose $v \in \mathcal{R}(A)^\perp$. Then $(Au, v)_{\tilde{H}} = 0$, for all $u \in H$. Hence $0 = (Au, v)_{\tilde{H}} = (u, A^*v)_H$, for all $u \in H$, which means $v \in \mathcal{N}(A^*)$. This implies $\mathcal{R}(A)^\perp \subset \mathcal{N}(A^*)$.

To prove (**p2**) let us assume first that $v \in \overline{\mathcal{R}(A)}$ is an arbitrary element. Then there exists such a sequence $\{v_n\} \in \overline{\mathcal{R}(A)}$ that

$$v_n = Au_n \text{ and } \lim_{n\to\infty} v_n = v.$$

Assuming $w \in \mathcal{N}(A^*)$ we conclude that $A^*w = 0$, so

$$(v_n, w)_{\tilde{H}} = (Au_n, w)_{\tilde{H}} = (u_n, A^*w)_H = 0.$$

Hence

$$|(v, w)| \leq |(v - v_n, w)| + |(v_n, w)| \leq \|(v - v_n, w)\| \, \|w\|.$$

The right hand side tends to zero as $n \to \infty$, which implies $(v, w) = 0$, for all $w \in \mathcal{N}(A^*)$, i.e. $v \in \mathcal{N}(A^*)^\perp$. Therefore $\overline{\mathcal{R}(A)} \subset \mathcal{N}(A^*)^\perp$. To prove $\mathcal{N}(A^*)^\perp \subset \overline{\mathcal{R}(A)}$, we need to prove that if $v \notin \overline{\mathcal{R}(A)}$, then $v \notin \mathcal{N}(A^*)^\perp$. Since $\overline{\mathcal{R}(A)}$ is a closed subspace of the Hilbert space H, by Corollary 2.1.1 there exists a unique decomposition of the above defined element $v \in H$:

$$v = v_0 + w_0, \;\; v_0 \in \overline{\mathcal{R}(A)}, \quad w_0 \in \mathcal{R}(A)^\perp,$$

with $v_0 := Pv$. Then $(v, w_0) := (v_0 + w_0, w_0) = \|w_0\|^2 \neq 0$. But by (**p1**), $w_0 \in \mathcal{N}(A^*)$, so $(v, w_0) \neq 0$, which means that $v \notin \mathcal{N}(A^*)^\perp$. This completes the proof. □

Lemma 2.2.1 *Let $A : H \mapsto \tilde{H}$ be a bounded linear operator. Then $(A^*)^* = A$ and $\|A\|^2 = \|A^*\|^2 = \|AA^*\| = \|A^*A\|$.*

Proof We can easily show that $(A^*)^* = A$. Indeed, for all $u \in H$, $v \in \tilde{H}$,

$$(v, \left(A^*\right)^* u)_{\tilde{H}} = (A^*v, u)_H = \overline{(u, A^*v)_H} = \overline{(Au, v)_{\tilde{H}}} = (v, Au)_{\tilde{H}}.$$

Further, it follows from the definition $\|A^*v\|^2 := (A^*v, A^*v)_H$, $v \in \tilde{H}$, that $\|A^*v\|^2 = (AA^*v, v)_{\tilde{H}} \leq \|A\|\|A^*v\|\|v\|$, and hence $\|A^*v\| \leq \|A\|\|v\|$, which implies boundedness of the adjoint operator: $\|A^*\| \leq \|A\|$. In the same way we can deduce $\|A\| \leq \|A^*\|$, interchanging the roles of the operators A and A^*. Therefore, $\|A^*\| = \|A\|$.

To prove the second part of the lemma, we again use the definition $\|Au\|^2 := (Au, Au)_{\tilde{H}}$, $u \in H$. Then, $\|Au\|^2 = (A^*Au, u)_H \leq \|A^*Au\| \;\; \|u\|$, and we get $\|A\|^2 \leq \|A^*A\|$. On the other hand, $\|A^*A\| \leq \|A^*\|\|A\| = \|A\|^2$, since $\|A^*\| = \|A\|$. Thus $\|A\|^2 = \|A^*A\|$. □

Corollary 2.2.1 *Let conditions of Theorem 2.2.1 hold. Then*
(c1) $\mathcal{N}(A^*) = \mathcal{N}(AA^*)$;
(c2) $\overline{\mathcal{R}(A)} = \overline{\mathcal{R}(AA^*)}$.

Proof Let $v \in \mathcal{N}(A^*)$. Then $A^*v = 0$ and hence $AA^*v = 0$, which implies $v \in \mathcal{N}(AA^*)$, i.e. $\mathcal{N}(A^*) \subset \mathcal{N}(AA^*)$. Suppose now $v \in \mathcal{N}(AA^*)$. Then $AA^*v = 0$ and $\|A^*v\|_H^2 := (A^*v, A^*v)_H = (AA^*v, v)_{\tilde{H}} = 0$. This implies $A^*v = 0$, i.e. $v \in \mathcal{N}(A^*)$, which completes the proof of (**c1**). To prove (**c2**) we use the formula $(AA^*)^* := A^{**}A^* = AA^*$ and the second relationship (**p2**) in Theorem 2.2.1, replacing here A by AA^*. We have $\overline{\mathcal{R}(AA^*)} = \mathcal{N}(AA^*)^\perp$. Taking into account here (**c1**) we conclude $\overline{\mathcal{R}(AA^*)} = \mathcal{N}(A^*)^\perp$. With the relationship (**p1**) in Theorem 2.2.1, this completes the proof. □

Remark that if A is a bounded linear operator defined on a Hilbert space H, then AA^* and A^*A are positive.

Now we briefly show a crucial role of adjoint operators in studying the solvability of linear equations. Let $A : H \mapsto \tilde{H}$ be a bounded linear operator. Consider the linear operator equation

$$Au = f, \ u \in H, \ f \in \tilde{H}. \tag{2.2.1}$$

Denote by $v \in \tilde{H}$ a solution of the homogeneous adjoint equation $A^*v = 0$. Then we have:

$$(f, v)_{\tilde{H}} := (Au, v)_{\tilde{H}} = (u, A^*v)_H = 0.$$

Hence $(f, v) = 0$ for all $v \in \mathcal{N}(A^*)$, and by $\tilde{H} = \mathcal{N}(A^*)^\perp \oplus \mathcal{N}(A^*)$, this implies: $f \in \mathcal{N}(A^*)^\perp$. On the other hand, Fredholm alternative asserts that Eq. (2.2.1) has a (non-unique) solution if and only if $f \perp v$ for each solution $v \in \tilde{H}$ of the homogeneous adjoint equation $A^*v = 0$. This leads to the following result.

Proposition 2.2.1 *Let $A : H \mapsto \tilde{H}$ be a bounded linear operator on a Hilbert space H. Then a necessary condition for existence of a solution $u \in H$ of Eq.* (2.2.1) *is the condition*

$$f \in \mathcal{N}(A^*)^\perp. \tag{2.2.2}$$

Using Theorem 2.2.1 we can write H as the orthogonal (direct) sum

$$H = \overline{\mathcal{R}(A)} \oplus \mathcal{N}(A^*). \tag{2.2.3}$$

If the range $\mathcal{R}(A)$ of A is closed in H, that is, if $\overline{\mathcal{R}(A)} = \mathcal{R}(A)$, then using (2.2.3) and Proposition 2.2.1, we obtain the following necessary and sufficient condition for the solvability of Eq. (2.2.1).

Theorem 2.2.2 *Let $A : H \mapsto \tilde{H}$ be a bounded linear operator with closed range. Then Eq.* (2.2.1) *has a solution $u \in H$ if and only if condition* (2.2.2) *holds.*

This theorem provides a useful tool for proving existence of a solution of the closed range operator Eq. (2.2.1) via the null-space of the adjoint operator.

2.3 Moore-Penrose Generalized Inverse

Let $A : H \mapsto \tilde{H}$ be a bounded linear operator between the real Hilbert spaces H and $\tilde{H}$. Consider the operator Eq. (2.2.1). Evidently, a solution of (2.2.1) exists if and only if $f \in \mathcal{R}(A) \subset \tilde{H}$. This means that the first condition (**i1**) of Hadamard's Definition 1.1.1 holds. Assume now that $f \in \tilde{H}$ does not belong to the range $\mathcal{R}(A)$ of the operator A which usually appears in applications. It is natural to extend the notion of solution for this case, looking for an approximate (or generalized) solution of (2.2.1) which satisfies this equation as well as possible. For this aim we introduce the *residual* $\|f - Au\|_{\tilde{H}}$ and then look for an element $u \in H$, as in Definition 2.1.1, which minimizes this norm:

$$\|f - Au\|_{\tilde{H}} = \inf_{v \in H} \|f - Av\|_{\tilde{H}} . \tag{2.3.1}$$

The minimum problem (2.3.1) is called a *least squares problem* and accordingly, the best approximation $u \in H$ is called a *least squares solution* to (2.2.1).

Let us consider first the minimum problem (2.3.1) from differential calculus viewpoint. Introduce the functional

$$\mathcal{J}(u) = \frac{1}{2}\|Au - f\|^2_{\tilde{H}}, \ u \in H.$$

Using the identity

$$\mathcal{J}(u+h) - \mathcal{J}(u) = 2\left(A^*(Au - f), h\right) + \|Au\|^2_H, \ \forall h \in H,$$

we obtain the Fréchet differential

$$\left(\mathcal{J}'(u), h\right) = 2\left(A^*(Au - f), h\right), \ h \in H$$

of this functional. Hence the least squares solution $u \in H$ of the operator Eq. (2.2.1) is defined from the condition $\left(\mathcal{J}'(v), h\right) = 0$, for all $h \in H$, as follows: $(A^*(Au - f), h) = 0$, i.e. as a solution of the equation

$$A^*Au = A^*f. \tag{2.3.2}$$

This shows that least squares problem (2.3.1) is equivalent to Eq. (2.3.2) with the formal solution

$$u = \left(A^*A\right)^{-1} A^*f, \quad f \in \tilde{H}. \tag{2.3.3}$$

Equation (2.3.2) is called the *normal equation*.

The normal equation plays an important role in studying ill-posed problems as we will see in next sessions. First of all, remark that the operator A in (2.2.1) may not be injective, which means non-uniqueness in view of Hadamard's definition. The first important property of the normal Eq. (2.3.2) is that the operator A^*A is injective on the range $\mathcal{R}(A^*)$ of the adjoint operator A^*, even if the bounded linear operator A is not injective. For this reason, the normal equation is the most appropriate one to obtain a least squares solution of an inverse problem.

Lemma 2.3.1 *Let $A : H \mapsto \tilde{H}$ be a bounded linear operator and H, $\tilde{H}$ Hilbert spaces. Then the operator $A^*A : \mathcal{R}(A^*) \subset H \mapsto H$ is injective.*

Proof Note, first of all, that $\overline{\mathcal{R}(A^*)} = \overline{\mathcal{R}(A^*A)}$, as it follows from Corollary 2.2.1 (replacing A by A^*). Let $u \in \mathcal{R}(A^*)$ be such an element that $A^*Au = 0$. Then $Au \in \mathcal{N}(A^*)$, by definition of the null-space. But $\mathcal{N}(A^*) = \mathcal{R}(A)^\perp$, due to Theorem 2.2.1. On the other hand, $Au \in \mathcal{R}(A)$, since Au is the image of the element $u \in H$ under the transformation A. Thus, $Au \in \mathcal{R}(A) \cap \mathcal{R}(A)^\perp$ which implies $Au = 0$. This in turn means that $u \in \mathcal{N}(A) = \mathcal{R}(A^*)^\perp$. With the above assumption $u \in \mathcal{R}(A^*)$, we conclude that $u \in \mathcal{R}(A^*) \cap \mathcal{R}(A^*)^\perp$, i.e. $u = 0$. Therefore $A^*Au = 0$ implies $u = 0$, which proves the injectivity of the operator A^*A. $\square$

Let us explain now an interpretation of the "inverse operator" $(A^*A)^{-1}A^*$ in (2.3.3), in view of the orthogonal projection.

As noted above, $f \in \tilde{H}$ may not belong to the range $\mathcal{R}(A)$ of the operator A. So, we assume that $f \in \tilde{H} \setminus \mathcal{R}(A)$ is an arbitrary element and try *to construct a unique linear extension of an "inverse operator" from* $\mathcal{R}(A)$ *to the subspace* $\mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$. Since the closure $\overline{\mathcal{R}(A)}$ of the range $\mathcal{R}(A)$ is a closed subspace of $\tilde{H}$, by Corollary 2.1.2, $\tilde{H} = \overline{\mathcal{R}(A)} \oplus \mathcal{R}(A)^\perp$ (note that $\overline{\mathcal{R}(A)}^\perp = \overline{\mathcal{R}(A)^\perp} = \mathcal{R}(A)^\perp$). Hence $\mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ is *dense* in $\tilde{H}$. By the same corollary, the projection Pf of the arbitrary element $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ is in $\mathcal{R}(A)$. This means that there exists such an element $u \in H$ that

$$Au = Pf, \quad f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp. \tag{2.3.4}$$

By the Orthogonal Decomposition, the element Au being the projection of f onto $\overline{\mathcal{R}(A)}$, is an element of $\mathcal{R}(A)$. Furthermore, the element $u \in H$ is a least squares solution to (2.2.1), i.e. is a solution of the minimum problem (2.3.1). This implies that a least squares solution $u \in H$ of (2.2.1) exists if and only if f is an element of the *dense* in $\tilde{H}$ subspace $\mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$.

On the other hand, for each element $f \in \overline{\mathcal{R}(A)} \oplus \mathcal{R}(A)^\perp$ the following (unique) decomposition holds:

$$f = Pf + h, \;\; Pf \in \overline{\mathcal{R}(A)}, \;\; h \in \mathcal{R}(A)^\perp. \tag{2.3.5}$$

Hence for each projection $Pf \in \mathcal{R}(A)$ we have $f - Pf \in \mathcal{R}(A)^\perp$. Taking into account (2.3.4), we conclude from (2.3.5) that

$$f - Au \in \mathcal{R}(A)^\perp. \tag{2.3.6}$$

But $\mathcal{R}(A)^\perp = \mathcal{N}(A^*)$, by Theorem 2.2.1. Hence

$$f - Au \in \mathcal{N}(A^*). \tag{2.3.7}$$

By definition of the null-space, (2.3.7) implies that $A^*(f - Au) = 0$. Thus, again we arrive at the same result: $u \in H$ satisfies the normal Eq. (2.3.2).

Therefore we have constructed *a mapping* $A^\dagger$ *from* $\mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ *into* H, *which associates each element* $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ *to the least squares solution* $u \in H$ *of the operator equation* (2.2.1). Furthermore, the domain $\mathcal{D}(A)^\dagger := \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ of this mapping is obtained in a natural way.

This mapping is called the *Moore-Penrose (generalized) inverse* of the bounded linear operator A and is denoted by $A^\dagger$:

$$A^\dagger : \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp \mapsto H. \tag{2.3.8}$$

Evidently, the generalized inverse is a densely defined linear operator, that is, $\overline{\mathcal{D}(A^\dagger)} = \tilde{H}$, since a least squares solution exists only if f is an element of the dense in $\tilde{H}$ subspace $\mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$.

To complete this definition, let us answer the question: the operator $A^\dagger$ is an inverse of which operator? First of all, the normal Eq. (2.3.2) shows that a least squares solution exists if and only if $\mathcal{N}(A^*A) = \{0\}$, or equivalently, $\mathcal{N}(A) = 0$, due to $\mathcal{N}(A^*A) = \mathcal{N}(A)$, by Corollary 2.2.1. Since $A^*A : H \mapsto H$ is a self-adjoint operator, it follows from Theorem 2.2.1 and Corollary 2.2.1 that

$$\begin{aligned} H &:= \overline{\mathcal{R}(A^*A)} \oplus \mathcal{R}(A^*A)^\perp \\ &= \overline{\mathcal{R}(A^*)} \oplus \mathcal{N}(A^*A)^\perp \\ &= \mathcal{N}(A)^\perp \oplus \mathcal{N}(A) \end{aligned} \tag{2.3.9}$$

The last line of decompositions (2.3.9) shows that to ensure the existence, we need to restrict the domain of the linear operator $A : H \mapsto \tilde{H}$ from $\mathcal{D}(A)$ to $\mathcal{N}(A)^\perp$. By this way, we define this operator as follows:

$$\mathring{A} := A|_{\mathcal{N}(A)^\perp}, \quad \mathring{A} : \mathcal{N}(A)^\perp \subset H \mapsto \mathcal{R}(A) \subset \tilde{H}.$$

It follows from this construction that $\mathcal{N}(\mathring{A}) = \{0\}$ and $\mathcal{R}(\mathring{A}) = \mathcal{R}(A)$. Therefore, the inverse operator

$$\mathring{A}^{-1} : \mathcal{R}(A) \subset \tilde{H} \mapsto \mathcal{N}(A)^\perp \subset H \tag{2.3.10}$$

exists. However, the range $\mathcal{R}(\mathring{A}^{-1})$ of this inverse operator is in $\tilde{H}$ and does not contain the elements $f \in \tilde{H} \setminus \mathcal{R}(A)$. For this reason, at the second stage of the above construction, we extended this range from $\mathcal{R}(A)$ to $\mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$, in order include those elements which may not belong to $\mathcal{R}(A)$.

Thus, following to [23], we can define the *Moore-Penrose inverse* $A^\dagger$ as follows.

Definition 2.3.1 The Moore-Penrose (generalized) inverse of a bounded linear operator A is the unique linear extension of the inverse operator (2.3.10) from $\mathcal{R}(A)$ to $\mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$:

$$A^\dagger : \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp \mapsto \mathcal{N}(A)^\perp \subset H, \tag{2.3.11}$$

with

$$\mathcal{N}(A^\dagger) = \mathcal{R}(A)^\perp. \tag{2.3.12}$$

The requirement (2.3.12) in this definition is due to (2.3.6) and (2.3.7). This requirement implies, in particular, that the Moore-Penrose inverse $A^\dagger$ is a linear operator.

Corollary 2.3.1 *Let $f \in \mathcal{D}(A^\dagger)$. Then $u \in H$ is a least squares solution of the operator equation $Au = f$ if and only if it is a solution of the normal Eq.* (2.3.2).

Proof It follows from (2.3.1) that $u \in H$ is a least squares solution of $Au = f$ if and only if Au is the closest element to f in $\mathcal{R}(A)$. The last assertion is equivalent to (2.3.6), i.e. $f - Au \in \mathcal{R}(A)^\perp$. By Theorem 2.2.1, $\mathcal{R}(A)^\perp = \mathcal{N}(A^*)$. Hence $f - Au \in \mathcal{N}(A^*)$, which means $A^*(f - Au) = 0$ or $A^*Au = A^*f$. □

The following theorem shows that the Moore-Penrose inverse of a bounded linear operator is a closed operator. Remark that a linear operator $L : H_1 \mapsto H_2$ is closed if and only if for any sequence $\{u_n\} \subset \mathcal{D}(L)$, satisfying

$$\lim_{n\to\infty} u_n = u, \quad \text{and} \quad \lim_{n\to\infty} Lu_n = v,$$

the conditions hold:

$$u \in \mathcal{D}(L), \quad \text{and} \quad v = Lu. \tag{2.3.13}$$

Theorem 2.3.1 *Let $A : H \mapsto \tilde{H}$ be a linear bounded operator from the Hilbert spaces H into $\tilde{H}$. Then the Moore-Penrose inverse $A^\dagger$, defined by* (2.3.11) *and* (2.3.12)*, is a closed operator.*

Proof Let $\{f_n\} \subset \mathcal{D}(A^\dagger)$, $n = 1, 2, 3, \ldots$, $f_n \to f$, and $A^\dagger f_n \to u$, as $n \to \infty$. Denote by $u_n := A^\dagger f_n$ the unique solution of the normal Eq. (2.3.2) for each n, that is, $A^*Au_n = A^*f_n$, $u_n \in \mathcal{N}(A)^\perp$. Since $\mathcal{N}(A)^\perp$ is closed, $\{u_n\} \subset \mathcal{N}(A)^\perp$ and $u_n \to u$, as $n \to \infty$, which implies $u \in \mathcal{N}(A^\perp)$, i.e. $u \in \mathcal{R}(A^\dagger)$, by (2.3.11). Hence, the first condition of (2.3.13) holds. Now we prove that $u = A^\dagger f$. Due to the continuity of the operators A^*A and A^* we have:

$$A^*Au_n \to A^*Au, \quad \text{and} \quad A^*Au_n = A^*f_n \to A^*f, \quad \text{as} \quad n \to \infty.$$

The left hand sides are equal, so $A^*Au = A^*f$, i.e. $u \in \mathcal{N}(A)^\perp$ is the solution of the normal Eq. (2.3.3). By Corollary 2.3.1, $u = A^\dagger f$. This completes the proof. □

Corollary 2.3.2 *Let conditions of Theorem 2.3.1 hold. Then the Moore-Penrose inverse $A^\dagger$ is continuous if and only if $\mathcal{R}(A)$ is closed.*

Proof Let $A^\dagger$, defined by (2.3.11) and (2.3.12), be a continuous operator. Assume that $\{f_n\} \subset \mathcal{R}(A)$ be a convergent sequence: $f_n \to f$, as $n \to \infty$. We need to prove that $f \in \mathcal{R}(A)$. Denote by $u_n := A^\dagger f_n$. Then $u_n \in \mathcal{N}(A)^\perp$, for all $n = 1, 2, 3, \ldots$. Since $A^\dagger$ is continuous and $\mathcal{N}(A)^\perp$ is closed we conclude:

$$u := \lim_{n\to\infty} u_n = \lim_{n\to\infty} A^\dagger f_n = A^\dagger f, \quad \text{and} \quad u \in \mathcal{N}(A)^\perp.$$

On the other hand, $Au_n = f_n$ and $A : H \mapsto \tilde{H}$ is a continuous operator, so $Au = f$, where f is the above defined limit of the sequence $\{f_n\} \subset \mathcal{R}(A)$. But the element

Au, being the projection of f onto $\overline{\mathcal{R}(A)}$, is an element of $\mathcal{R}(A)$, by (2.3.4). Hence, $f \in \mathcal{R}(A)$, which implies that $\mathcal{R}(A)$ is closed.

To prove the second part of the corollary, we assume now $\mathcal{R}(A)$ is closed. By definition (2.3.11), $\mathcal{D}(A^\dagger) := \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$, which implies $\mathcal{D}(A^\dagger)$ is closed, since the orthogonal complement $\mathcal{R}(A)^\perp$ is a closed subspace. As a consequence, the graph $G_{A^\dagger} := \{(f, A^\dagger f : f \in \mathcal{D}(A^\dagger)\}$ of the operator $A^\dagger$ is closed. Then, as a closed graph linear operator, $A^\dagger$ is continuous. □

Remember that in the case when A is a linear compact operator, the range $\mathcal{R}(A)$ is closed if and only if it is finite-dimensional.

Remark, finally, that the notion of generalized inverse has been introduced by E. H. Moore and R. Penrose [69, 80, 81]. For ill-posed problems this very useful concept has been developed in [23, 31].

2.4 Singular Value Decomposition

As we have seen already in the introduction, inverse problems with compact operators are an challenging case. Most inverse problems related to differential equations are represented by these operators. Indeed, all input-output operators corresponding to these inverse problems, are compact operators. Hence, *compactness of the operator A is a main source of ill-posedness* of the operator equation (2.2.1) and our interest will be directed towards the case $A : H \mapsto \tilde{H}$ in (2.2.1) is a linear compact operator. When A is a self-adjoint compact operator, i.e. for all $u \in H$ and $v \in \tilde{H}$, $(Au, v) = (u, Av)$, we may use the *spectral representation*

$$Au = \sum_{n=1}^{\infty} \lambda_n (u, u_n) u_n, \ \forall u \in H, \tag{2.4.1}$$

where λ_n, $n = 1, 2, 3, \ldots$ are nonzero real eigenvalues (repeated according to its multiplicity) and $\{u_n\} \subset H$ is the complete set of corresponding orthonormal eigenvectors u_n. The set $\{\langle \lambda_n, u_n \rangle\}$, consisting of all pairs of nonzero eigenvalues and corresponding eigenvectors, is defined an *eigensystem* of the self-adjoint operator A. It is also known from the spectral theory of self-adjoint compact operators that $\lambda_n \to 0$, as $n \to \infty$. If $\dim \mathcal{R}(A) = \infty$ then for any $\varepsilon > 0$ the index set $\{n \in \mathbb{N} : \ |\lambda_n| \geq \varepsilon\}$ is finite. Here and below $\mathbb{N}$ is the set of natural numbers. If the range $\mathcal{R}(A)$ of a compact operator is finite, $\lambda_n = 0$, for all $n > \dim \mathcal{R}(A)$.

However, *if A is not-self-adjoint*, there are no eigenvalues, hence no eigensystem. In this case, the notion *singular system* replaces the eigensystem. To describe this system we use the operators A^*A and AA^*. Both $A^*A : H \mapsto H$ and $AA^* : \tilde{H} \mapsto \tilde{H}$ are compact self-adjoint nonnegative operators. We denote the eigensystem of the self-adjoint operator A^*A by $\{\langle \mu_n, u_n \rangle\}$. Then $A^*Au_n = \mu_n u_n$, for all $u_n \in H$, which implies $(A^*Au_n, u_n) = \mu_n (u_n, u_n)$. Hence $\|Au_n\|^2_{\tilde{H}} = \mu_n \|u_n\|^2_H \geq 0$, which means all nonzero eigenvalues are positive: $\mu_n > 0$, $n \in \mathbf{N}$, where $\mathbf{N} := \{n \in \mathbb{N} : \ \mu_n \neq 0\}$

(at most countable) index set of positive eigenvalues. We denote the square roots of the positive eigenvalues μ_n of the self-adjoint operator $A^*A : H \mapsto H$ by $\sigma_n := \sqrt{\mu_n}$, $n \in \mathbf{N}$. Below we will assume that these eigenvalues are ordered as follows: $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_n \geq \ldots > 0$.

Definition 2.4.1 Let $A : H \mapsto \tilde{H}$ be a linear compact operator with adjoint $A^* : \tilde{H} \mapsto H$, H and $\tilde{H}$ be Hilbert spaces. The square root $\sigma_n := \sqrt{\mu_n}$ of the eigenvalue $\mu_n > 0$ of the self-adjoint operator $A^*A : H \mapsto H$ is called the singular value of the operator A.

Using the spectral representation (2.4.1) for the self-adjoint compact operator A^*A we have:

$$A^*Au = \sum_{n=1}^{\infty} \sigma_n^2 (u, u_n) u_n, \ \forall u \in H. \tag{2.4.2}$$

Let us introduce now the orthonormal system $\{v_n\}$ in $\tilde{H}$, via the orthonormal system $\{u_n\} \subset H$ as follows: $v_n := Au_n / \|Au_n\|$. Applying A^* to both sides we have: $A^*v_n = A^*Au_n / \|Au_n\|$. By the above definition $A^*Au_n = \sigma_n^2 u_n$ and $\|Au_n\| = \sigma_n$. This implies:

$$A^*v_n = \sigma_n u_n.$$

Act by the adjoint operator A^* now on both sides of the Fourier representation $v = \sum_{n=1}^{\infty}(v, v_n)v_n$, $v \in \tilde{H}$, where $v_n = Au_n / \|Au_n\|$. Taking into account the definition $A^*v_n = \sigma_n u_n$, we have:

$$A^*v = \sum_{n=1}^{\infty} \sigma_n (v, v_n) u_n, \quad v \in \tilde{H}. \tag{2.4.3}$$

Applying A to both sides of (2.4.3) and using $\|Au_n\| = \sigma_n$ we obtain the spectral representation for the self-adjoint operator AA^*:

$$AA^*v = \sum_{n=1}^{\infty} \sigma_n^2 (v, v_n) v_n, \quad v \in \tilde{H}. \tag{2.4.4}$$

It is seen from (2.4.2) and (2.4.4) that the eigenvalues $\sigma_n^2 > 0$, $n \in \mathbf{N}$, of the self-adjoint operators AA^* and A^*A are the same, as expected.

The representation (2.4.3) is called *singular value expansion* of the adjoint operator $A^* : \tilde{H} \mapsto H$. For the operator $A : H \mapsto \tilde{H}$ this expansion can be derived in a similar way:

$$Au = \sum_{n=1}^{\infty} \sigma_n (u, u_n) v_n, \quad u \in H. \tag{2.4.5}$$

Substituting $v = v_n$ in (2.4.3) and $u = u_n$ in (2.4.5), we obtain the following formulae:

$$Au_n = \sigma_n v_n, \quad A^* v_n = \sigma_n u_n. \tag{2.4.6}$$

The triple $\{\sigma_n, u_n, v_n\}$ is called the *singular system for the non-self-adjoint operator* $A : H \mapsto \tilde{H}$.

As we will see in the next chapter, some input-output operators related to inverse source problems are self-adjoint. If $A : H \mapsto \tilde{H}$ is a self-adjoint operator with eigensystem $\{\langle\lambda_n, u_n\rangle\}$, then $\|Au_n\| = |\lambda_n|$ and

$$v_n := Au_n/\|Au_n\| = \lambda u_n/|\lambda_n|,$$

by the above construction. Therefore, *the singular system for the self-adjoint operator* $A : H \mapsto \tilde{H}$ is defined as the triple $\{\sigma_n, u_n, v_n\}$, with $\sigma = |\lambda_n|$ and $v_n = \lambda u_n/|\lambda_n|$.

Example 2.4.1 Singular values of a self-adjoint integral operator

Assuming $H = L^2(0, \pi)$, we define the non-self-adjoint integral operator $A : H \mapsto H$ as follows:

$$(Au)(x) := \int_x^{\pi} u(\xi)d\xi, \ x \in (0, \pi), \ u \in H. \tag{2.4.7}$$

By using the integration by parts formula and the definition $(Au, v)_{L^2(0,\pi)} = (u, A^*v)_{L^2(0,\pi)}$, $\forall u, v \in H$, we can easily construct the adjoint operator $A^* : H \mapsto H$:

$$(A^*v)(x) = \int_0^{x} v(\xi)d\xi, \ x \in (0, \pi), \ v \in H. \tag{2.4.8}$$

Evidently, both integral operators (2.4.7) and (2.4.8) are compact. Indeed, let $\{u_n\}_{n=1}^{\infty}$ be a bounded sequence in $L^2(0, \pi)$ with $\|u_n\|_{L^2(0,\pi)} \le M$, $M > 0$. Then for any $x_1, x_2 \in [0, \pi]$, (2.4.7) implies:

$$|(Au_n)(x_1) - (Au_n)(x_2)| \le \left| \int_{x_1}^{x_2} u_n(\xi)d\xi \right| \le M|x_1 - x_2|^{1/2}.$$

This shows that $\{(Au)_n\}$ is an equicontinuous family of functions in $C[0, \pi]$. Hence, there exists a subsequence $\{(Au)_m\} \subset \{(Au)_n\}$ that converges uniformly in $C[0, \pi]$ to a continuous function v. Since uniform convergence implies convergence in $L^2[0, \pi]$, we conclude that the subsequence $\{(Tu)_m\}$ converges in $L^2[0, \pi]$. Therefore the integral operator defined by (2.4.7) is compact because the image of a bounded sequence always contains a convergent subsequence.

Now we define the self-adjoint integral operator A^*A:

$$(A^*Au)(x) = \int_0^x \int_\xi^\pi u(\eta)d\eta d\xi,\ u \in H. \tag{2.4.9}$$

To find the nonzero positive eigenvalues $\mu_n > 0$, $n \in \mathbb{N}$, of the self-adjoint integral operator A^*A, defined by (2.4.9), the eigenvalue problem should be solved for the integral equation

$$(A^*Au)(x) = \sigma^2 u(x),\quad x \in (0, \pi). \tag{2.4.10}$$

Differentiating both sides of (2.4.10) twice with respect to $x \in [0, \pi]$ we arrive at the Sturm-Liouville equation: $-u''(x) = \lambda u(x)$, $\lambda = 1/\sigma^2$. To derive the boundary conditions, we first substitute $x = 0$ in (2.4.10). Then we get $u(0) = 0$, by (2.4.9). Differentiating both sides of (2.4.10) and substituting $x = \pi$ we conclude $u'(\pi) = 0$. Hence, problem (2.4.10) is equivalent (in well-known sense) to the eigenvalue problem

$$\begin{cases} -u''(x) = \lambda u(x), \quad \text{a.e. } x \in (0, \pi), \quad \lambda = 1/\sigma^2, \\ u(0) = u'(\pi) = 0, \end{cases} \tag{2.4.11}$$

for the self-adjoint positive-defined differential operator $Au := -u''$. The solution of this problem is in $\mathring{H}^2[0, \pi] := \{u \in H^2(0, \pi) :\ u(0) = 0\}$, where $H^2(0, \pi)$ is the Sobolev space.

Solving the eigenvalue problem (2.4.11) we find the eigenvalues $\lambda_n = (n - 1/2)^2$, $n \in \mathbb{N}$, and the corresponding normalized eigenvectors $u_n(x) = \sqrt{2/\pi}\sin(\sqrt{\lambda_n}x)$. Hence, the eigenvalues $\sigma_n^2 = 1/\lambda_n$ and the corresponding eigenvectors $u_n(x)$ of the self-adjoint integral operator A^*A are

$$\sigma_n^2 = (n - 1/2)^{-2},\ \ u_n(x) = \sqrt{2/\pi}\sin((n - 1/2)x).$$

By Definition 2.4.1, $\sigma_n = (n - 1/2)^{-1}$, $n \in \mathbb{N}$, are the eigenvalues of the non-self-adjoint integral operator A. The corresponding eigenvectors, given in equations (2.4.6) are

$$u_n(x) = \sqrt{2/\pi}\sin((n - 1/2)x),\quad v_n(x) = \sqrt{2/\pi}\cos((n - 1/2)x).$$

Thus, the singular system $\{\sigma_n, u_n, v_n\}$ for the non-self-adjoint integral operator (2.4.7) is defined as follows:

$$\{(n - 1/2)^{-1},\ \sqrt{2/\pi}\sin((n - 1/2)x), \sqrt{2/\pi}\cos((n - 1/2)x)\}. \qquad \square$$

Remark 2.4.1 The above example insights into the degree of ill-posedness of simplest integral equations $Au = f$ and $A^*Au = A^*f$, with the operators A and A^*A, defined by (2.4.7) and (2.4.9). In the first case one needs an operation differentiation (which is an ill-posed procedure) to find $u = A^{-1}f$. As a result, $\sigma_n = \mathcal{O}(n^{-1})$. In the second case the operator A^*A, defined by (2.4.8), contains two integration and hence one needs to differentiate twice to find $u = (A^*A)^{-1}A^*f$, which results in the singular values as $\mathcal{O}(n^{-2})$. We come back to this issue in the next chapter.

The above considerations lead to so-called *Singular Value Decomposition* (or normal form) of compact operators.

Theorem 2.4.1 (Picard) Let H and $\tilde{H}$ be Hilbert spaces and $A : H \mapsto \tilde{H}$ be a linear compact operator with the singular system $\{\sigma, u_n, v_n\}$. Then the equation $Au = f$ has a solution if and only if

$$f \in \mathcal{N}(A^*)^{\perp} \;\text{ and }\; \sum_{n=1}^{\infty} \frac{1}{\sigma_n^2} |(f, v_n)|^2 < +\infty. \tag{2.4.12}$$

In this case

$$u := A^{\dagger} f = \sum_{n=1}^{\infty} \frac{1}{\sigma_n} (f, v_n) u_n \tag{2.4.13}$$

is the solution of the equation $Au = f$.

Proof Let the equation $Au = f$ has a solution. Then f must be in $\overline{\mathcal{R}(A)}$. But by Theorem 2.2.1, $\overline{\mathcal{R}(A)} = \mathcal{N}(A^*)^{\perp}$. Hence $f \in \mathcal{N}(A^*)^{\perp}$ and the first part of (2.4.12) holds. To prove the second part of (2.4.12) we use the relation $A^*v_n = \sigma_n u_n$ in (2.4.6) to get

$$\sigma_n(u, u_n) = (u, A^*v_n) = (Au, v_n) = (f, v_n).$$

Hence, $(u, u_n) = (f, v_n)/\sigma_n$. Using this in

$$u = \sum_{n=1}^{\infty} (u, u_n) u_n, \;\; u \in H \tag{2.4.14}$$

we obtain:

$$u = \sum_{n=1}^{\infty} \frac{1}{\sigma_n} (f, v_n) u_n.$$

But the orthonormal system $\{u_n\}$ is complete, so the Fourier series (2.4.14) is convergent. By the convergence criterion this implies the second condition of (2.4.12):

$$\sum_{n=1}^{\infty} \frac{1}{\sigma_n^2} |(f, v_n)|^2 = \sum_{n=1}^{\infty} |(u, u_n)|^2 < +\infty.$$

To prove the second part of the theorem, we assume now that conditions (2.4.12) hold. Then series (2.4.13) converges. Acting on both sides of this series by the operator A, using $f \in \mathcal{N}(A^*)^\perp = \overline{\mathcal{R}(A)}$ and $Au_n = \sigma_n v_n$ we get:

$$Au = \sum_{n=1}^{\infty} \frac{1}{\sigma_n}(f, v_n)Au_n$$
$$\sum_{n=1}^{\infty}(f, v_n)v_n = f.$$

This completes the proof. $\square$

Since $\mu_n > 0$, $n \in \mathbb{N}$ are eigenvalues of the self-adjoint operator A^*A (as well as AA^*) and $\sigma_n := \sqrt{\mu_n}$, we have: $\sigma_n \to 0$, as $n \to \infty$, if $\dim\mathcal{R}(A) = \infty$. Then it follows from formulae (2.4.12)-(2.4.13) that $A^\dagger$ is an unbounded operator. Indeed, for any fixed eigenvector v_k, with $\|v_k\| = 1$, we have:

$$\|A^\dagger v_k\| = \frac{1}{\sigma_k} \to \infty, \text{ as } n \to \infty.$$

The second condition (2.4.12), called *Picard criterion*, shows that the best approximate solution of the equation $Au = f$ exists if only the Fourier coefficients (f, v_n) of f decay faster than the singular values σ_n. Concrete examples related to this issue will be given in the next chapter.

As noted in Remark 2.4.1, singular value decomposition reflects the ill-posedness of the equation $Au = f$ with a compact operator A between the *infinite dimensional* Hilbert spaces H and $\tilde{H}$. Indeed, the decay rate of the non-increasing sequence $\{\sigma_n\}_{n=1}^{\infty}$ characterizes the *degree of ill-posedness* of an ill-posed problem. In particular, the amplification factors of a measured data errors in nth Fourier component of the series (2.4.13), corresponding to the integral operators (2.4.7) and (2.4.9), increase as n and n^2, respectively, due to the factor $1/\sigma_n$. In terms of corresponding problems $Au = f$ and $A^*Au = A^*f$ this means that the second problem is more ill-posed than the first one. Hence, solving numerically the second ill-posed problem is more difficult than the first one.

These considerations motivate the following definition of ill-posedness of problems governed by compact operators, proposed in [43].

Definition 2.4.2 Let $A : H \mapsto \tilde{H}$ be a linear compact operator between the infinite dimensional Hilbert spaces H and $\tilde{H}$. If there exists a constant $C > 0$ and a real number $s \in (0, \infty)$ such that

$$\sigma_n \ge \frac{C}{n^s}, \text{ for all } n \in \mathbf{N}, \tag{2.4.15}$$

then the equation $Au = f$ is called moderately ill-posed of degree at most s. If for any $\epsilon > 0$, condition (2.4.15) does not hold with s replaced by $s - \epsilon > 0$, then the equation $Au = f$ is called moderately ill-posed of degree s. If no such number

$s \in (0, \infty)$ exists such that condition (2.4.15) holds, then the equation $Au = f$ is called severely ill-posed.

Typical behavior of severe ill-posedness is exponential decay of the singular values of the compact operator A. As we will show in the next chapter, classical backward parabolic problem is severely ill-posed, whereas the final data inverse source problems related to parabolic and hyperbolic equations are only moderately ill-posed. Remark that some authors use more detailed classification, distinguishing between *mildly ill-posedness* ($s \in (0, 1)$) and moderately ill-posedness ($s \in (1, \infty)$).

In applications, to obtain an approximation of $A^\dagger f$ one can truncate the series (2.4.13):

$$u^N := \sum_{n=1}^{N} \frac{1}{\sigma_n}(f, v_n)u_n. \tag{2.4.16}$$

This method of obtaining the *approximate solution* u^N is called the *truncated singular value decomposition* (*TSVD*).

To understand the role of the *cutoff parameter* N, we assume that the right hand side $f \in H$ of the equation $Au = f$ is given with some *measurement error* $\delta > 0$, i.e. $\|f - f^\delta\| \le \delta$, where f^δ is a *noisy data*. Then

$$u^{N,\delta} := \sum_{n=1}^{N} \frac{1}{\sigma_n}(f^\delta, v_n)u_n \tag{2.4.17}$$

is an approximate solution of the equation $Au = f^\delta$ corresponding to the noisy data f^δ. Let us estimate the norm $\|u^N - u^{N,\delta}\|$, i.e. the difference between the approximate solutions corresponding to the noise free (f) and noisy (f^δ) data. From (2.4.16)-(2.4.17) we deduce the estimate:

$$\begin{aligned}\|u^N - u^{N,\delta}\|^2 &= \sum_{n=1}^{N} \frac{1}{\sigma_n^2}\left|(f - f^\delta, v_n)\right|^2 \\ &\le \frac{1}{\sigma_N^2}\sum_{n=1}^{N}\left|(f - f^\delta, v_n)\right|^2 \le \frac{\delta^2}{\sigma_N^2}.\end{aligned}$$

Using this estimate we can find the *accuracy error* $\|u^{N,\delta} - A^\dagger f\|$, i.e. the difference between the best approximate solution $A^\dagger f$, corresponding to the noise free data f, and the approximate solution $u^{N,\delta}$, obtained by TSVD and corresponding to the noisy data f^δ:

$$\begin{aligned}\|u^{N,\delta} - A^\dagger f\| &\le \|u^N - A^\dagger f\| + \|u^N - u^{N,\delta}\| \\ &\le \|u^N - A^\dagger f\| + \frac{\delta}{\sigma_N}.\end{aligned} \tag{2.4.18}$$

The first term $\|u^N - A^\dagger f\|$ on the right hand side of estimate (2.4.18) depends only on the cutoff parameter N and does not depend on the measurement error $\delta > 0$. This term is called the *regularization error*. The second term $\|u^N - u^{N,\delta}\|$ on the right hand side of (2.4.18) depends not only on the cutoff parameter N, but also on the measurement error $\delta > 0$. This term is called the *data error*. This term exhibits some very distinctive features of a solution of the ill-posed problems. Namely, the approximation error $\|u^{N,\delta} - A^\dagger f\|$ decreases with $\delta > 0$, for a fixed value of the cutoff parameter N, on one hand. On the other hand, for a given $\delta > 0$ this error tends to infinity, as $N \to \infty$, since $\sigma_N := \sqrt{\mu_N} \to 0$. Hence, the parameter $N = N(\delta)$ needs to be chosen depending on $\delta > 0$ such that

$$\frac{\delta}{\sigma_{N(\delta)}} \to 0, \quad \text{as} \;\; \delta \to 0. \tag{2.4.19}$$

We use the right hand side of (2.4.17) to introduce the operator $R_{N(\delta)} : \tilde{H} \mapsto H$,

$$R_{N(\delta)} f^\delta := \sum_{n=1}^{N(\delta)} \frac{1}{\sigma_n} (f^\delta, v_n) u_n. \tag{2.4.20}$$

It follows from estimate (2.4.18) that if condition (2.4.19) holds, then

$$\|R_{N(\delta)} f^\delta - A^\dagger f\| \to 0, \;\; \delta \to 0.$$

Hence, the basic idea of TSVD in solving ill-posed problems is finding a finite dimensional approximation of the unbounded operator $A^\dagger$. A class of such finite dimensional operators defined by (2.4.20) and approximating the unbounded operator $A^\dagger$ can be defined as *regularization method*or *regularization strategy*. The cutoff parameter $N(\delta)$ plays role of the parameter of regularization [35].

2.5 Regularization Strategy. Tikhonov Regularization

Let $A : H \mapsto \tilde{H}$ be a linear injective bounded operator between infinite-dimensional real Hilbert spaces H and $\tilde{H}$. Consider the linear ill-posed operator equation

$$Au = f, \;\; u \in H, \;\; f \in \mathcal{R}(A). \tag{2.5.1}$$

By the condition $f \in \mathcal{R}(A)$, the operator Eq. (2.5.1) is ill-posed in the sense that a solution $u \in H$ exists, but doesn't depend continuously on the data $f \in \mathcal{R}(A)$. In practice this data always contains a random noise. We denote by $f^\delta \in \tilde{H}$ the noisy data and assume that

$$\|f^\delta - f\|_{\tilde{H}} \le \delta, \;\; f \in \mathcal{R}(A), \;\; f^\delta \in \tilde{H}, \;\; \delta > 0. \tag{2.5.2}$$

Then the exact equality in the equation $Au = f^\delta$ can not be satisfied due to the noisy data f^δ and we may only consider the minimization problem

$$J(u) = \inf_{v \in H} J(v) \tag{2.5.3}$$

for the *Tikhonov functional*

$$J(u) = \frac{1}{2}\|Au - f^\delta\|^2_{\tilde{H}}, \; u \in H, \; f^\delta \in \tilde{H}, \tag{2.5.4}$$

where $\|Au - f^\delta\|^2_{\tilde{H}} := (Au - f^\delta, Au - f^\delta)_{\tilde{H}}$.

A solution $u \in H$ of the minimization problem (2.5.3)–(2.5.4) is called *quasi-solution or least squares solution* of the ill-posed problem (2.5.1). If, in addition, this solution is defined as the minimum-norm solution, i.e. if

$$\|u\|_H = \inf\{\|w\|_H \; : w \in H \text{ is a least squares solution of } (1.5.1)\},$$

then this solution is called *best approximate solution* of (2.5.1). Note that the concept of quasi-solution has been introduced in [49].

Since H is infinite-dimensional and A is compact, the minimization problem for the functional (2.5.4) is ill-posed, the functional $J(u)$ doesn't depend continuously on the data $f^\delta \in \mathcal{R}(A)$. One of the possible ways of stabilizing the functional is to add the penalty term $\alpha\|u - u^0\|^2_H$, as in Optimal Control Theory, and then consider the minimization problem for the *regularized Tikhonov functional*

$$J_\alpha(u) := \frac{1}{2}\|Au - f^\delta\|^2_{\tilde{H}} + \frac{1}{2}\alpha\|u - u^0\|^2_H, \; u \in H, \; f^\delta \in \tilde{H}. \tag{2.5.5}$$

Here $\alpha > 0$ is the parameter of regularization and $u^0 \in H$ is an initial guess. Usually $u^0 \in H$ is one of the possible good approximations to the exact solution $u \in H$, but if such an initial guess is not known, we may take $u^0 = 0$. Below we assume that $u^0 = 0$.

This approach is defined as *Tikhonov regularization* [95, 96] or *Tikhonov-Phillips regularization* [83].

Theorem 2.5.1 *Let $A : H \mapsto \tilde{H}$ be a linear injective bounded operator between real Hilbert spaces H and $\tilde{H}$. Then the regularized Tikhonov functional (2.5.5) has a unique minimum $u^\delta_\alpha \in H$, for all $\alpha > 0$. This minimum is the solution of the linear equation*

$$\left(A^*A + \alpha I\right) u^\delta_\alpha = A^* f^\delta, \; u^\delta_\alpha \in H, \; f^\delta \in \tilde{H}, \;\; \alpha > 0 \tag{2.5.6}$$

and has the form

$$u^\delta_\alpha = \left(A^*A + \alpha I\right)^{-1} A^* f^\delta. \tag{2.5.7}$$

Moreover, the operator $A^*A + \alpha I$ is boundedly invertible, hence the solution u_α^δ continuously depends on f^δ.

Proof First of all, note that the Fréchet differentiability of the functional (2.5.5) follows from the identity:

$$J_\alpha(u+v) - J_\alpha(u) = \left(A^*(Au - f^\delta) + \alpha u, v\right) + \frac{1}{2}\|Av\|^2 + \frac{1}{2}\alpha\|v\|^2, \ \forall u, v \in H,$$

where $A^* : \tilde{H} \mapsto H$ is the adjoint operator of A. This identity implies:

$$\begin{cases} (J'_\alpha(u), v) = (Au - f^\delta, Av) + \alpha(u, v), \ \forall v \in H; \\ J''_\alpha(u; v, v) = \|Av\|^2 + \alpha\|v\|^2, \ \forall v \in H. \end{cases} \tag{2.5.8}$$

Formula (2.5.8) for the second Fréchet derivative $J''_\alpha(u; v, v)$ shows that the regularized Tikhonov functional $J_\alpha(u)$ defined on a real Hilbert H space is strictly convex, since $\alpha > 0$, and $\lim_{\|u\|\to+\infty} J_\alpha(u) = +\infty$. Then it has a unique minimizer $u_\alpha^\delta \in H$ and this minimum is characterized by the following necessary and sufficient condition

$$(J'_\alpha(u), v) = 0, \ \forall v \in H, \tag{2.5.9}$$

where $J'_\alpha(u)$ is the first Fréchet derivative of the regularized Tikhonov functional. Thus, condition (2.5.9) with formula (2.5.8) implies that the minimum $u_\alpha^\delta \in H$ of the regularized Tikhonov functional is the solution of the linear Eq. (2.5.6). This solution is defined by (2.5.7), since the operator $A^*A + \alpha I$ is boundedly invertible. This follows from the Lax-Milgram lemma and the positive definiteness of the operator $A^*A + \alpha I$:

$$\left((A^*A + \alpha I)v, v\right) = \|Av\|^2 + \alpha\|v\|^2 \geq \alpha\|v\|^2, \ \forall v \in H, \ \alpha > 0.$$

Evidently, the operator $A^*A + \alpha I$ is one-to-one for each positive α. Indeed, multiplication of the homogeneous equation $(A^*A + \alpha I)v = 0$ by $v \in H$ implies: $(A^*Av, v) + \alpha(v, v) = (Av, Av) + \alpha(v, v) = 0$. This holds if and only if $v = 0$. This completes the proof. □

From (2.5.8) we deduce the *gradient formula for the regularized Tikhonov functional.*

Corollary 2.5.1 *For the Fréchet gradient $J'_\alpha(u)$ of the regularized Tikhonov functional* (2.5.4) *the following formula holds:*

$$J'_\alpha(u) = A^*\left(Au - f^\delta\right) + \alpha u, \ u \in H. \tag{2.5.10}$$

The main consequence of the Picard's Theorems 2.4.1 and 2.5.1 is that the solution u_α^δ of the normal Eq. (2.5.6) corresponding to the noisy data f^δ can be represented by the following series:

$$u_\alpha^\delta = \sum_{n=1}^{\infty} \frac{q(\alpha; \sigma_n)}{\sigma_n}(f^\delta, v_n)u_n, \quad \alpha > 0, \tag{2.5.11}$$

where

$$q(\alpha; \sigma) = \frac{\sigma^2}{\sigma^2 + \alpha} \tag{2.5.12}$$

is called the *filter function*.

Corollary 2.5.2 *Let conditions of Theorem 2.4.1 hold and* $f^\delta \in \mathcal{N}(A^*)^\perp$. *Then the unique regularized solution* $u_\alpha^\delta \in H$*, given by* (2.5.7)*, can be represented as the convergent series* (2.5.11).

Proof It follows from the normal Eq. (2.5.6) that $\alpha u_\alpha^\delta = \overline{A^* f^\delta} - A^* A u_\alpha^\delta$. By Corollary 2.2.1, $\overline{\mathcal{R}(A^*A)} = \overline{\mathcal{R}(A^*)}$, so this implies that $u_\alpha^\delta \in \overline{\mathcal{R}(A^*)}$. But due to Theorem 2.2.1, $\overline{\mathcal{R}(A^*)} = \mathcal{N}(A)^\perp$ and we conclude that $u_\alpha^\delta \in \mathcal{N}(A)^\perp$. The orthonormal system $\{u_m\}$ spans $\mathcal{N}(A)^\perp$. Hence

$$u_\alpha^\delta = \sum_{m=1}^{\infty} c_m u_m. \tag{2.5.13}$$

To find the unknown parameters c_m we substitute (2.5.13) into the normal Eq. (2.5.6):

$$\sum_{m=1}^{\infty} (\sigma_m^2 + \alpha) c_m u_m = A^* f^\delta.$$

Multiplying both sides by u_n we get:

$$(\sigma_n^2 + \alpha) c_n = (A^* f^\delta, u_n).$$

But $(A^* f^\delta, u_n) = \left(f^\delta, A u_n\right) = \sigma_n (f^\delta, v_n)$. Therefore

$$(\sigma_n^2 + \alpha) c_n = \sigma_n (f^\delta, v_n)$$

and the unknown parameters are defined as follows:

$$c_m = \frac{\sigma_m}{\sigma_m^2 + \alpha}(f^\delta, v_m)$$

Using this in (2.5.13) we obtain:

$$u_\alpha^\delta = \sum_{m=1}^{\infty} \frac{\sigma_m}{\sigma_m^2 + \alpha} \left(f^\delta, v_m\right) u_n.$$

With formula (2.5.12), this implies (2.5.11). □

Remark that the filter function has the following properties:

$$q(\alpha;\sigma) \le \min\left\{\frac{\sigma}{2\sqrt{\alpha}};\ \frac{\sigma^2}{\alpha}\right\},\ \forall\alpha > 0. \tag{2.5.14}$$

These properties follow from the obvious inequalities:

$$\sigma^2+\alpha \equiv (\sigma-\sqrt{\alpha})^2+2\sigma\sqrt{\alpha} \ge 2\sigma\sqrt{\alpha},$$
$$\frac{\sigma^2}{\sigma^2+\alpha} \equiv \frac{\sigma^2/\alpha}{\sigma^2/\alpha+1} \le \frac{\sigma^2}{\alpha},\ \forall\alpha>0,\ \sigma>0.$$

The linear Eq. (2.5.6) is defined as a *regularized form of the normal equation* $A^*Au = A^*f^\delta$ [97].

If A^*A is invertible, then for $\alpha = 0$ formula (2.5.7) implies that

$$u_0^\delta = (A^*A)^{-1}A^*f^\delta =: A^\dagger f^\delta,\ \alpha = 0 \tag{2.5.15}$$

is the solution of the *normal equation*

$$A^*Au^\delta = A^*f^\delta,\ u^\delta \in H,\ f^\delta \in \tilde{H}. \tag{2.5.16}$$

Evidently $u_0^\delta \in \mathcal{N}(A)^\perp$, by Definition 2.3.1 of Moore-Penrose inverse $A^\dagger$.

It follows from (2.5.15) that Moore-Penrose inverse $A^\dagger := (A^*A)^{-1}A^*$ arises naturally as a result of minimization of Tikhonov functional. Remark that the provisional replacement of the compact operators AA^* or A^*A by the non-singular operators $AA^*+\alpha I$ or $A^*A+\alpha I$, is the main idea of Tikhonov's regularization procedure. This procedure has originally been introduced by Tikhonov in [94] for uniform approximations of solutions of Fredholm's equation of the first kind. However, Tikhonov does not point to the relation of his ideas to Moore-Penrose inverse.

Let us assume now that $\alpha > 0$. The right hand side of (2.5.7) defines family of continuous operators

$$R_\alpha := \left(A^*A+\alpha I\right)^{-1} A^* : \tilde{H} \mapsto H,\ \alpha > 0, \tag{2.5.17}$$

depending on the parameter of regularization $\alpha > 0$. Obviously, when data is noise free, then the regularized solution $R_\alpha f$ should converge (in some sense) to $A^\dagger f$, as $\alpha \to 0$, that is, $R_\alpha f \to A^\dagger f$, for all $f \in \mathcal{D}(A^\dagger)$. In practice this data is always noisy and we may only assume that it is known up to some error $\delta > 0$, that is, $\|f - f^\delta\| \le \delta$. Hence the parameter of regularization $\alpha > 0$ depends on the noise level $\delta > 0$ and the noise data $f^\delta \in \tilde{H}$, and should be chosen appropriately, keeping an error $\|u_\alpha^\delta - u\|$ as small as possible. A strategy of choosing the parameter of regularization $\alpha = \alpha(\delta, f^\delta)$ is called a *parameter choice rule*. These considerations lead to the following definition of *regularization strategy*.

Definition 2.5.1 Let $A : H \mapsto \tilde{H}$ be an injective compact operator. Assume that $f, f^\delta \in \tilde{H}$ be noise free and noisy data respectively, that is, $\|f - f^\delta\| \leq \delta < \|f^\delta\|$, $\delta > 0$. A family $\{R_{\alpha(\delta, f^\delta)}\}_\alpha$ of bounded linear operators is called a regularization strategy or a convergent regularization method if for all $f \in \mathcal{D}(A^\dagger)$,

$$\limsup_{\delta \to 0^+} \left\{ \left\| R_{\alpha(\delta, f^\delta)} f^\delta - A^\dagger f \right\|_H \ : f^\delta \in \tilde{H}, \ \left\| f - f^\delta \right\| \leq \delta \right\} = 0, \quad (2.5.18)$$

with $\alpha : \mathbb{R}_+ \times \tilde{H} \mapsto \mathbb{R}_+$, such that

$$\limsup_{\delta \to 0^+} \left\{ \alpha(\delta, f^\delta) \ : f^\delta \in \tilde{H}, \ \|f - f^\delta\| \leq \delta \right\} = 0. \quad (2.5.19)$$

If the parameter of regularization $\alpha = \alpha(\delta, f^\delta)$ depends only on the noise level $\delta > 0$, the rule $\alpha(\delta, f^\delta)$ is called *a-priori parameter choice rule*. Otherwise, this rule is called *a-posteriori parameter choice rule*. Tikhonov regularization is a typical example of a priori parameter choice rule, since the choice of the parameter of regularization $\alpha > 0$ is made a priori, i.e. before computations, as we will see in Theorem 2.5.2. The Morozov's Discrepancy Principle is a regularization strategy with a-posteriori parameter choice rule, since the choice of the parameter of regularization $\alpha > 0$ is made during the process of computing. Regarding the iterative methods, the number of iterations plays here the role of the regularization parameter. As we will show in the next chapter, Landweber's Method and Conjugate Gradient Method together with appropriate parameter choice rule are also a regularization strategy in sense of Definition 2.5.1. We remark finally that, as in the case of Tikhonov regularization, some widely used regularization operators R_α *are linear*. In particular, the Morozov's Discrepancy Principle and Landweber's Method can be formulated as linear regularization methods. However, the Conjugate Gradient Method is a *nonlinear regularization method*, since the right hand side of the equation $Au = f$ does not depend linearly on the parameter of regularization α.

Above definition, with Theorem 2.5.1, implies that the family of operators $\{R_\alpha\}_{\alpha > 0}$, $R_\alpha : \tilde{H} \mapsto H$, defined as

$$R_\alpha := \left(A^* A + \alpha I\right)^{-1} A^*, \quad \alpha > 0, \quad (2.5.20)$$

is the regularization strategy corresponding to Tikhonov regularization. That is, the regularization operators R_α, $\alpha > 0$, approximate the unbounded inverse $A^\dagger$ of the operator A on $\mathcal{R}(A)$. On the other hand, Corollary 2.5.1, implies that the singular value expansion of this operator is

$$R_\alpha f^\delta := \sum_{n=1}^{\infty} \frac{q(\alpha; \sigma_n)}{\sigma_n} (f^\delta, v_n) u_n, \quad \alpha > 0. \quad (2.5.21)$$

It is easy to prove that the family of operators $\{R_\alpha\}$ are not uniformly bounded, i.e. there exists a sequence $\{\alpha_m\}_{m=1}^{\infty}$, $\alpha_m > 0$, such that $\|R_{\alpha_m}\| \to \infty$, as $\alpha_m \to 0$. Indeed, taking $f_\alpha^\delta = v_m$ in (2.5.21), where v_m is any fixed eigenvector, with $\|v_m\| = 1$, we have:

$$\|R_{\alpha_m} v_m\| = \frac{\sigma_m}{\sigma_m^2 + \alpha_m}.$$

For the sequence $\{\alpha_m\}$ satisfying the conditions

$$\alpha_m \to 0 \text{ and } \alpha_m/\sigma_m \to 1, \quad \text{as } m \to \infty,$$

we conclude:

$$\|R_{\alpha_m} v_m\| = \frac{1}{\sigma_m + \alpha_m/\sigma_m} \to \infty, \text{ as } m \to \infty.$$

The regularization strategy R_α possesses this property in general case as well.

Lemma 2.5.1 *Let R_α be regularization strategy corresponding to the compact operator $A : H \mapsto \tilde{H}$. Then the operators $\{R_\alpha\}_{\alpha>0}$, $R_\alpha : \tilde{H} \mapsto H$, are not uniformly bounded, i.e. there exists a sequence $\{\alpha_m\}_{m=1}^{\infty}$, $\alpha_m > 0$, such that $\|R_{\alpha_m}\| \to \infty$, as $\alpha_m \to 0$.*

Proof Assume, in contrary, that there exists a constant $M > 0$, independent on $\alpha > 0$, such that $\|R_\alpha\| \le M$, for all $\alpha > 0$. Then for any $f \in \mathcal{R}(A) \subset \tilde{H}$ we have:

$$\begin{aligned}\|A^\dagger f\| \le \|A^\dagger f - R_\alpha f\| + \|R_\alpha f\| &\le \|A^\dagger f - R_\alpha f\| + \|R_\alpha\|\|f\| \\ &\le \|A^\dagger f - R_\alpha f\| + M\|f\|.\end{aligned}$$

The first norm on the right-hand side tends to zero as $\alpha \to 0$, by definition (2.5.18). Then passing to the limit we get: $\|A^\dagger f\| \le M\|f\|$, for all $f \in \mathcal{R}(A)$, which implies $\|A^\dagger\| \le M$, i.e. boundedness of the generalized inverse $A^\dagger$. This contradiction completes the proof. □

Therefore, one needs to find a bounded approximation of the unbounded operator $A^\dagger$. This approximation will evidently depends on both, the parameter of regularization $\alpha > 0$ and the noisy data f^δ. So, the main problem of regularization strategy is to find such a bounded approximation, which will be convergent to the best approximate solution $A^\dagger f$ corresponding to the exact data $f \in \mathcal{N}(A^*)^\perp$, as $\alpha \to 0$ and $\delta \to 0$.

The following theorem gives an answer to this issue.

Theorem 2.5.2 *Let $A : H \mapsto \tilde{H}$ be a linear injective bounded operator between Hilbert spaces H and $\tilde{H}$. Assume that conditions of Theorem 2.4.1 hold. Denote by $f^\delta \in \mathcal{N}(A^*)^\perp$ the noisy data: $\|f - f^\delta\|_{\tilde{H}} \le \delta$, $\delta > 0$. Suppose that the conditions hold:*

$$\alpha(\delta) \to 0 \;\; and \;\; \frac{\delta^2}{\alpha(\delta)} \to 0, \;\; as \;\; \delta \to 0. \tag{2.5.22}$$

Then the solution $u_\alpha^\delta := R_{\alpha(\delta)} f^\delta$ *of the regularized form of the normal Eq. (2.5.6) converges to the best approximate solution* $u := A^\dagger f$ *of Eq. (2.5.1) in the norm of the space* H*, that is,*

$$\|R_{\alpha(\delta)} f^\delta - A^\dagger f\|_H \to 0, \;\; \delta \to 0. \tag{2.5.23}$$

Proof Let us estimate the norm $\|R_{\alpha(\delta)} f^\delta - A^\dagger f\|_H$ using (2.4.13) and (2.5.21). We have

$$\begin{aligned}
\|R_{\alpha(\delta)} f^\delta - A^\dagger f\|_H^2 &:= \sum_{n=1}^{\infty} \left\| \left(\frac{q(\alpha(\delta); \sigma_n)}{\sigma_n} (f^\delta, v_n) - \frac{1}{\sigma_n}(f, v_n) \right) u_n \right\|^2 \\
= \sum_{n=1}^{\infty} &\left\| \left(\frac{q(\alpha(\delta); \sigma_n)}{\sigma_n} (f^\delta - f, v_n) + \left(\frac{q(\alpha(\delta); \sigma_n)}{\sigma_n} - \frac{1}{\sigma_n} \right) (f, v_n) \right) u_n \right\|^2 \\
\le 2 \sum_{n=1}^{\infty} &\frac{q^2(\alpha(\delta); \sigma_n)}{\sigma_n^2} |(f^\delta - f, v_n)|^2 + 2 \sum_{n=1}^{\infty} \frac{\alpha^2(\delta)}{\sigma_n^2(\sigma_n^2 + \alpha(\delta))^2} |(f, v_n)|^2.
\end{aligned} \tag{2.5.24}$$

Denote by S_1 and S_2 the first and the second summands on the right-hand side of (2.5.24), respectively.

We use from (2.5.14) the property $q^2(\alpha; \sigma_n)/\sigma_n^2 \le 1/(4\alpha)$ of the filter function and Parseval's identity to estimate the term S_1:

$$S_1 \le \frac{1}{2\alpha(\delta)} \sum_{n=1}^{\infty} |(f^\delta - f, v_n)|^2 \le \frac{1}{2\alpha(\delta)} \|f^\delta - f\|^2 \le \frac{\delta^2}{2\alpha(\delta)} . \tag{2.5.25}$$

For estimating the term S_2 we rewrite it in the following form:

$$\begin{aligned}
S_2 = 2 \sum_{n=1}^{N} \frac{\alpha^2(\delta)}{\sigma_n^2(\sigma_n^2 + \alpha(\delta))^2} |(f, v_n)|^2 + 2 \sum_{n=N+1}^{\infty} \frac{\alpha^2(\delta)}{\sigma_n^2(\sigma_n^2 + \alpha(\delta))^2} |(f, v_n)|^2 \\
=: S_{2N} + R_{2N}.
\end{aligned} \tag{2.5.26}$$

To estimate the Nth partial sum S_{2N} we use the properties $\sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_n \ge \ldots > 0$ and $\sigma_n \to 0$, as $n \to \infty$, of the singular values. For any small value $\alpha = \alpha(\delta) > 0$ of the parameter of regularization, depending on the noise level $\delta > 0$, there exists such a positive integer $N = N(\alpha(\delta)) \equiv N(\delta)$ that

$$\min_{1 \le n \le N(\delta)} \sigma_n^2 = \sigma_{N(\delta)}^2 \ge \sqrt{\alpha(\delta)} > \sigma_{N(\delta)+1}^2, \tag{2.5.27}$$

where $N = N(\alpha(\delta)) \to \infty$, as $\delta \to 0$. For such $N = N(\delta)$ we estimate the Nth partial sum S_{2N} in (2.5.26) as follows:

$$
\begin{aligned}
S_{2N} &\le \frac{2\alpha^2(\delta)}{(\sigma^2_{N(\delta)} + \alpha(\delta))^2} \sum_{n=1}^{N(\delta)} \frac{1}{\sigma_n^2} |(f, v_n)|^2 \\
&= \frac{2\alpha(\delta)}{(\sigma^2_{N(\delta)}/\sqrt{\alpha(\delta)} + \sqrt{\alpha(\delta)})^2} \sum_{n=1}^{N(\delta)} \frac{1}{\sigma_n^2} |(f, v_n)|^2
\end{aligned}
$$

By the condition (2.5.27), $\sigma^2_{N(\delta)}/\sqrt{\alpha(\delta)} \ge 1$. Hence

$$
S_{2N} \le \frac{2\alpha(\delta)}{(1 + \sqrt{\alpha(\delta)})^2} \sum_{n=1}^{N(\delta)} \frac{1}{\sigma_n^2} |(f, v_n)|^2. \tag{2.5.28}
$$

Let us estimate now the series remainder term R_{2N} defined in (2.5.26). We have

$$
\begin{aligned}
R_{2N} = 2 \sum_{n=N(\delta)+1}^{\infty} \frac{1}{(\sigma_n^2/\alpha(\delta) + 1)^2} \frac{1}{\sigma_n^2} |(f, v_n)|^2 \\
\le 2 \sum_{n=N(\delta)+1}^{\infty} \frac{1}{\sigma_n^2} |(f, v_n)|^2.
\end{aligned}
$$

Taking into account this estimate with estimates (2.5.25) and (2.5.28) in (2.5.24) we finally deduce:

$$
\begin{aligned}
\| R_{\alpha(\delta)} f^\delta - A^\dagger f \|_H^2 \le \frac{\delta^2}{2\alpha} + \frac{2\alpha(\delta)}{(1 + \sqrt{\alpha(\delta)})^2} \sum_{n=1}^{N(\delta)} \frac{1}{\sigma_n^2} |(f, v_n)|^2 \\
+ 2 \sum_{n=N(\delta)+1}^{\infty} \frac{1}{\sigma_n^2} |(f, v_n)|^2.
\end{aligned} \tag{2.5.29}
$$

The first right hand side term tends to zero, as $\delta \to 0$, by the second condition of (2.5.22). The factor before the partial sum of the second right hand side term tends to zero, since $\alpha(\delta) \to 0$, as $\delta \to 0$, by the first condition of (2.5.22), and this partial sum is finite due to the convergence condition (2.4.12) of the Picard's Theorem 2.4.1. The third right hand side term also tends to zero, as $\delta \to 0$, since in this case $N(\delta) \to \infty$, and, as a result, the series remainder term tends to zero, by the same convergence condition. This implies (2.5.23). □

Estimate (2.5.29) clearly shows the role of all parameters $\delta > 0$, $\alpha(\delta) > 0$ and $N = N(\delta)$ in the regularization strategy.

As we have seen in the previous section, the number N of first terms in singular value expansion of a compact operator can also be considered as a regularization parameter. As noted above in iterative methods, the number of iterations also plays role of the regularization parameter, that is, $\alpha \sim 1/N$. For more detailed analysis of regularization methods for ill-posed problems we refer to the books [23, 54, 90].

Remark finally that besides the Tikhonov regularization, there are other regularization methods, such as Lavrentiev regularization, asymptotic regularization, local regularization, etc. Since zero is the only accumulation point of the singular values of a compact operator, the underlying idea in all these regularization techniques is modifying the smallest singular values, shifting all singular values by $\alpha > 0$. In other words, the idea is to approximate the compact operator A or A^*A by a family of operators $A + \alpha I$ or $A^*A + \alpha I$. The first one corresponds to *Lavrentiev regularization*. Specifically, while Tikhonov regularization is based on the normal equation $A^*Au = A^*f^\delta$, Lavrentiev regularization is based on the original equation $Au = f^\delta$. The main advantage of the first approach over the second one is that the operator A^*A is always injective, due to Lemma 2.3.1, even if the operator $A :\mapsto \tilde{H}$ is not injective. The second advantage of Tikhonov regularization is that the class of admissible values of the parameter of regularization $\alpha > 0$ for convergent regularization method is larger than the same class in Lavrentiev regularization. Specifically, when $\alpha = \delta$ the condition $\delta^2/\alpha(\delta) \to 0$, as $\delta \to 0$, for convergent regularization strategy in Tikhonov regularization holds, but does not hold in Lavrentiev regularization, as we will see below.

Corollary 2.5.2 can be adopted to the case of *Lavrentiev regularization* when the linear bounded injective operator $A : H \mapsto H$ is self-adjoint and positive semidefinite.

Corollary 2.5.3 *Let conditions of Theorem 2.4.1 hold and $f^\delta \in \mathcal{R}(A)$. Then the unique solution $u_\alpha^\delta \in H$ of the regularized equation*

$$(A + \alpha I)\, u_\alpha^\delta = f^\delta \tag{2.5.30}$$

can be represented as the series

$$u_\alpha^\delta = \sum_{n=1}^{\infty} \frac{\tilde{q}(\alpha; \sigma_n)}{\sigma_n} (f^\delta, u_n) u_n, \quad \alpha > 0, \tag{2.5.31}$$

where

$$\tilde{q}(\alpha; \sigma) = \frac{\sigma}{\sigma + \alpha}. \tag{2.5.32}$$

Proof We use the singular system $\{\sigma_n, u_n, u_n\}$ for the non-self-adjoint operator $A : H \mapsto \tilde{H}$, that is, $Au_n = \sigma_n u_n$ and the representation

$$u_\alpha^\delta = \sum_{m=1}^{\infty} c_m u_m. \tag{2.5.33}$$

Substituting this into the normal Eq. (2.5.30) we obtain:

$$\sum_{m=1}^{\infty}(\sigma_m+\alpha)c_m u_m = f^{\delta}.$$

Multiplying both sides by u_n we find the unknown parameters c_m:

$$c_m = \frac{1}{\sigma_m+\alpha}(f^{\delta}, u_m)$$

Using this in (2.5.33) we find:

$$u_{\alpha}^{\delta} = \sum_{m=1}^{\infty}\frac{1}{\sigma_m+\alpha}\left(f^{\delta}, u_m\right)u_m.$$

By (2.5.32), this is exactly the required expansion (2.5.31). □

Now the question we seek to answer here is that under which conditions it is possible to construct a convergent regularization strategy for Lavrentiev regularization. The following theorem, which is an analogue of Theorem 2.5.2, answers this question.

Theorem 2.5.3 *Let the linear bounded injective operator $A : H \mapsto \tilde{H}$ be a self-adjoint and positive semi-definite. Assume that conditions of Theorem 2.4.1 hold. Denote by $f^{\delta} \in \mathcal{R}(A)$ the noisy data: $\|f - f^{\delta}\|_{\tilde{H}} \le \delta,\ \delta > 0$. Suppose that the following conditions hold:*

$$\alpha(\delta) \to 0 \ \ \textit{and} \ \ \frac{\delta}{\alpha(\delta)} \to 0, \ \ \textit{as} \ \ \delta \to 0. \tag{2.5.34}$$

Then the operator $R_{\alpha(\delta)}$ defined by the right hand side of (2.5.31) is a convergent regularization strategy, that is,

$$R_{\alpha(\delta)} f^{\delta} := \sum_{n=1}^{\infty}\frac{\tilde{q}(\alpha(\delta);\sigma_n)}{\sigma_n}(f^{\delta}, u_n)u_n, \ \ \alpha > 0, \tag{2.5.35}$$

that is, the solution $u_{\alpha}^{\delta} := R_{\alpha(\delta)} f^{\delta}$ of the regularized form of the normal Eq. (2.5.30) converges to the best approximate solution $u := A^{\dagger} f$ of Eq. (2.5.1) in the norm of the space H.

Proof The proof is similar to the proof of Theorem 2.5.2. Here we first use the following property

$$\frac{\sigma}{\sigma+\alpha} \le \frac{\sigma}{\alpha} \tag{2.5.36}$$

of the filter function (2.5.32) in (2.5.24) to obtain the estimate

$$\|R_{\alpha(\delta)}f^\delta - A^\dagger f\|_H^2 := \sum_{n=1}^{\infty}\left\|\left(\frac{\tilde{q}(\alpha(\delta);\sigma_n)}{\sigma_n}(f^\delta,u_n) - \frac{1}{\sigma_n}(f,u_n)\right)u_n\right\|^2$$
$$\le 2\sum_{n=1}^{\infty}\frac{\tilde{q}^2(\alpha(\delta);\sigma_n)}{\sigma_n^2}|(f^\delta - f,u_n)|^2 + 2\sum_{n=1}^{\infty}\frac{\alpha^2(\delta)}{\sigma_n^2(\sigma_n+\alpha(\delta))^2}|(f,u_n)|^2 \quad (2.5.37)$$

Denote by S_1 and S_2 the first and the second right-hand side summands of (2.5.37), respectively. By (2.5.36) we conclude $\tilde{q}^2(\alpha;\sigma_n)/\sigma_n^2 \le 1/\alpha^2$. Using this inequality in (2.5.37) we obtain the estimate

$$S_1 := 2\sum_{n=1}^{\infty}\frac{\tilde{q}^2(\alpha(\delta);\sigma_n)}{\sigma_n^2}|(f^\delta - f,u_n)|^2 \le \frac{1}{\alpha^2(\delta)}\|f^\delta - f\|^2 \le \frac{\delta^2}{\alpha^2(\delta)} \quad (2.5.38)$$

Second, we rewrite the term S_2 in the following form:

$$S_2 = 2\sum_{n=1}^{N}\frac{\alpha^2(\delta)}{\sigma_n^2(\sigma_n+\alpha(\delta))^2}|(f,u_n)|^2 + 2\sum_{n=N+1}^{\infty}\frac{\alpha^2(\delta)}{\sigma_n^2(\sigma_n+\alpha(\delta))^2}|(f,u_n)|^2$$
$$=: S_{2N} + R_{2N}. \quad (2.5.39)$$

For estimating the first right hand side term S_2 in (2.5.39), we use the properties $\sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_n \ge \ldots > 0$, $\sigma_n \to 0$, as $n \to \infty$, of the eigenvalues $\lambda_n = \sigma_n$ of self-adjoint positive semidefinite operator A. For any small values $\alpha = \alpha(\delta) > 0$ of the parameter of regularization, depending on the noise level $\delta > 0$, there exists such a positive integer $N = N(\alpha(\delta)) \equiv N(\delta)$ that

$$\min_{1\le n\le N(\delta)}\sigma_n^2 = \sigma_{N(\delta)}^2 \ge \alpha(\delta) > \sigma_{N(\delta)+1}^2, \quad (2.5.40)$$

where $N = N(\alpha(\delta)) \to \infty$, as $\delta \to 0$. For such $N = N(\delta)$ we get the following estimate for S_{2N} in (2.5.39):

$$S_{2N} \le \frac{2\alpha^2(\delta)}{(\sigma_{N(\delta)}+\alpha(\delta))^2}\sum_{n=1}^{N(\delta)}\frac{1}{\sigma_n^2}|(f,v_n)|^2$$
$$= \frac{2\alpha(\delta)}{(\sigma_{N(\delta)}/\sqrt{\alpha(\delta)}+\sqrt{\alpha(\delta)})^2}\sum_{n=1}^{N(\delta)}\frac{1}{\sigma_n^2}|(f,v_n)|^2$$

By the condition (2.5.40), $\sigma_{N(\delta)}/\sqrt{\alpha(\delta)} \ge 1$, which implies:

$$S_{2N} \le 2\alpha(\delta) \sum_{n=1}^{N(\delta)} \frac{1}{\sigma_n^2} |(f, u_n)|^2. \tag{2.5.41}$$

The factor before the above partial sum, which is finite by the convergence condition (2.4.12) of the Picard's Theorem 2.4.1, tends to zero, as $\alpha(\delta) \to 0$. Hence, $S_{2N} \to 0$, as $\alpha(\delta) \to 0$.

Let us estimate now the series remainder term R_{2N} defined in (2.5.39). We have

$$\begin{aligned} R_{2N} &= 2 \sum_{n=N(\delta)+1}^{\infty} \frac{1}{(\sigma_n/\alpha(\delta) + 1)^2} \frac{1}{\sigma_n^2} |(f, v_n)|^2 \\ &\le 2 \sum_{n=N(\delta)+1}^{\infty} \frac{1}{\sigma_n^2} |(f, v_n)|^2. \end{aligned}$$

Substituting this estimate with (2.5.38) and (2.5.41) into (2.5.37) we conclude:

$$\begin{aligned} &\|R_{\alpha(\delta)} f^\delta - A^\dagger f\|_H^2 \le \\ &\tfrac{\delta^2}{\alpha(\delta)^2} + 2\alpha(\delta) \textstyle\sum_{n=1}^{N(\delta)} \frac{1}{\sigma_n^2} |(f, u_n)|^2 + 2 \sum_{n=N(\delta)+1}^{\infty} \frac{1}{\sigma_n^2} |(f, v_n)|^2 . \end{aligned} \tag{2.5.42}$$

It is easy to verify that under the conditions (2.5.34) all three right hand side terms tend to zero, as $\delta \to 0$. □

Comparing the convergence conditions in Theorems 2.5.2 and 2.5.3 we first observe that, in Tikhonov regularization the class of admissible values of the parameter of regularization $\alpha > 0$ for convergent regularization method is larger than the same class in Lavrentiev regularization. While, for example, in Tikhonov regularization the values $\alpha = \delta$ and $\alpha = \sqrt{\delta}$ of the parameter of regularization are admissible for convergent regularization method, as the second condition of (2.5.22) shows, in Lavrentiev regularization they are not admissible by the condition (2.5.34). Moreover, the dependence on the parameter of regularization $\alpha > 0$ of the numbers $N = N(\alpha(\delta))$, defined in proofs of these theorems and corresponding to these regularizations, are different. Thus, if $\sigma_n = \mathcal{O}(1/n^2)$, then condition (2.5.27) of Theorem 2.5.2 implies that there exists the constants $c_2 > c_1 > 0$ such that

$$\frac{c_2}{N^4} \ge \sqrt{\alpha(\delta)} > \frac{c_1}{(N+1)^4}.$$

For the same $\sigma_n = \mathcal{O}(1/n^2)$, the condition (2.5.40) of Theorem 2.5.3 implies:

$$\frac{\tilde{c}_2}{N^4} \ge \alpha(\delta) > \frac{\tilde{c}_1}{(N+1)^4}, \ \tilde{c}_2 > \tilde{c}_1 > 0.$$

As a result, $N(\alpha(\delta)) = \mathcal{O}\left(\alpha(\delta)^{-1/8}\right)$ in Tikhonov regularization and $N(\alpha(\delta)) = \mathcal{O}\left(\alpha(\delta)^{-1/4}\right)$ in Lavrentiev regularization.

Note, finally that the topic of Tikhonov regularization is very broad and here we described it for only linear inverse problems. We refer the reader to the books [48, 90] on the mathematical theory of regularization methods related to nonlinear inverse problems.

2.6 Morozov's Discrepancy Principle

In applications the right hand side $f \in H$ of the equation $Au = f$ always contains a noise. Instead of this equation one needs to solve the equation $Au^\delta = f^\delta$, where f^δ is a noisy data: $\|f - f^\delta\| \leq \delta$, where $\delta > 0$. This, in particular, implies that in the ideal case the residual or *discrepancy* $\|Au^\delta - f^\delta\|$ can only be at most in the order of δ. On the other hand, in order to construct a bounded approximation $R_{\alpha(\delta)}$ of the unbounded operator $A^\dagger$ one needs to choose the parameter of regularization $\alpha > 0$ depending on the noise level $\delta > 0$. Thus, the parameter of regularization needs to be chosen by a compromise between the residual $\|Au^\delta - f^\delta\|$ and the given bound $\delta > 0$ for the noise level. This is the main criteria of so-called *Morozov's Discrepancy Principle* due to Morozov [67, 68]. This principle is now one of the simplest tools and most widely used regularization method for ill-posed problems.

Definition 2.6.1 Let $f^\delta \in \tilde{H}$ be a noisy data with an arbitrary given noise level $\delta > 0$, that is, $\|f^\delta - f\| \leq \delta$, where $f \in \mathcal{R}(A)$ is a noise free (exact) data. If there exists such a value $\alpha = \alpha(\delta)$ of the parameter of regularization, depending on $\delta > 0$, that the corresponding solution $u^\delta_{\alpha(\delta)} \in H$ of the equation $Au = f^\delta$ satisfies the condition

$$\beta_1 \delta \leq \|Au^\delta_{\alpha(\delta)} - f^\delta\| \leq \beta_2 \delta, \quad \beta_2 \geq \beta_1 \geq 1, \tag{2.6.1}$$

then the parameter of regularization $\alpha = \alpha(\delta)$ is said to be chosen according to Morozov's Discrepancy Principle.

Let us assume that the size $\delta := \|\delta f\| = \|f^\delta - f\| > 0$ of the noise is known (although we do not know the random perturbation δf). Denote by $u \in \mathcal{N}(A)^\perp$ the solution of the equation $Au = f$ with a noise free (exact) data $f \in \mathcal{N}(A^*)^\perp = \overline{\mathcal{R}(A)}$. Then

$$\|Au - f^\delta\| = \|Au - f - \delta f\| = \|\delta f\| =: \delta, \quad \delta > 0.$$

Thus, if u is the exact solution, corresponding to the exact data f, and f^δ is the noisy data, then

$$\|Au - f^\delta\| = \delta. \tag{2.6.2}$$

Now, having the size $\delta = \|f - f^\delta\| > 0$ of noise, we want to use Tikhonov regularization to find the solution u^δ_α defined by (2.5.7) and corresponding to the noisy data $f^\delta := f + \delta f$. Then, as it follows from (2.6.2), *the best that we can require from the parameter of regularization* $\alpha > 0$ *is the residual (or discrepancy)* $\|Au^\delta_{\alpha(\delta)} - f^\delta\| = \delta$.

We prove that there exists such a value of the parameter of regularization $\alpha = \alpha(\delta)$ which satisfies the following conditions:

$$\|Au^\delta_{\alpha(\delta)} - f^\delta\| = \|f - f^\delta\| = \delta, \quad \delta > 0. \tag{2.6.3}$$

Theorem 2.6.1 *Let conditions of Theorem 2.4.1 hold. Denote by $f^\delta \in \tilde{H}$ the noisy data with $\|f^\delta\| > \delta > 0$. Then there exists a unique value $\alpha = \alpha(\delta)$ of the parameter of regularization satisfying conditions* (2.6.3).

Proof By the unique decomposition, for any $f^\delta \in \tilde{H} = \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ we have

$$f^\delta = \sum_{n=1}^{\infty} (f^\delta, v_n)v_n + Pf^\delta, \tag{2.6.4}$$

where $Pf^\delta \in \mathcal{R}(A)^\perp$ is the projection of the noisy data on $\mathcal{R}(A)^\perp$.

Now we rewrite the solution u^δ_α, given by (2.5.11), of the normal Eq. (2.5.6) in the following form:

$$u^\delta_\alpha = \sum_{n=1}^{\infty} \frac{\sigma_n}{\sigma_n^2 + \alpha}(f^\delta, v_n)u_n, \quad \alpha > 0. \tag{2.6.5}$$

Acting on this solution by the operator A and using the relation $Au_n = \sigma_n v_n$ we deduce:

$$Au^\delta_\alpha = \sum_{n=1}^{\infty} \frac{\sigma_n^2}{\sigma_n^2 + \alpha}(f^\delta, v_n)v_n.$$

With (2.6.4) this yields:

$$f^\delta - Au^\delta_\alpha = \sum_{n=1}^{\infty} \frac{\alpha}{\sigma_n^2 + \alpha}(f^\delta, v_n)v_n + Pf^\delta.$$

Then we obtain the discrepancy:

$$\|Au^\delta_\alpha - f^\delta\|^2 = \sum_{n=1}^{\infty} \frac{\alpha^2}{\left(\sigma_n^2 + \alpha\right)^2}|(f^\delta, v_n)|^2 + \|Pf^\delta\|^2. \tag{2.6.6}$$

It can be verified that $g(\alpha) := \|Au^\delta_\alpha - f^\delta\|$ is a monotonically increasing function for $\alpha > 0$. To complete the proof of the theorem we need to show that the equation $g(\alpha) = \delta$, $\alpha \in (0, +\infty)$, has a unique solution. For $\alpha \to 0^+$ we use $Pf = 0$ for the noise free data $f \in \mathcal{R}(A)$ to get

$$\lim_{\alpha \to 0^+} g(\alpha) = \|Pf^\delta\| = \|P(f^\delta - f)\| \le \|f^\delta - f\| = \delta, \tag{2.6.7}$$

by (2.6.3). Note that the sum in (2.6.6) tends to zero, as $\alpha \to 0^+$.

For $\alpha \to +\infty$ we use the following limit

$$\lim_{\alpha \to +\infty} \frac{\alpha^2}{\left(\sigma_n^2 + \alpha\right)^2} = \lim_{\alpha \to +\infty} \frac{1}{\left(\sigma_n^2/\alpha^2 + 1\right)^2} = 1$$

and the identity (2.1.8) to get

$$\lim_{\alpha \to +\infty} g(\alpha) = \left(\sum_{n=1}^{\infty} |(f^\delta, v_n)|^2 + \|Pf^\delta\|^2 \right)^{1/2} =: \|f^\delta\|,$$

by (2.6.6).

By the assumption $\|f^\delta\| > \delta$ we conclude:

$$\lim_{\alpha \to +\infty} g(\alpha) > \delta.$$

This implies with (2.6.7) that the equation $g(\alpha) = \delta$ has a unique solution for $\alpha \in (0, +\infty)$. □

Remark that $\|f^\delta\| > \delta > 0$ is a natural condition in the theorem. Otherwise, i.e. if $\|f^\delta\| < \delta$, then $u^\delta_{\alpha(\delta)} = 0$ can be assigned as a regularized solution.

The theorem below shows that Morozov's Discrepancy Principle provides a convergent regularization strategy. First we need the following notion.

Definition 2.6.2 Let $A : H \mapsto \tilde{H}$ be a linear injective compact operator between Hilbert spaces H and $\tilde{H}$, and $A^* : \tilde{H} \mapsto H$ its adjoint. Denote by $u \in H$ the solution $u = A^\dagger f$ of the equation $Au = f$ with the noise free (exact) data $f \in \mathcal{R}(A)$. If there exists such an element $v \in \tilde{H}$ with $\|v\|_{\tilde{H}} \le M$ that $u = A^* v$, then we say that $u \in H$ satisfies the source condition.

Remark that besides of the classical concept of source condition, in recent years different new concepts, including approximate source conditions, have been developed [44].

Theorem 2.6.2 *Let $A : H \mapsto \tilde{H}$ be a linear injective compact operator between Hilbert spaces H and $\tilde{H}$. Assume that the solution $u = A^\dagger f$ of the equation $Au = f$ with the noise free data $f \in \mathcal{R}(A)$ satisfies the source condition. Suppose that the parameter of regularization $\alpha = \alpha(\delta)$, $\delta > 0$, is defined according to conditions* (2.6.3). *Then the regularization method $R_{\alpha(\delta)}$ is convergent, that is,*

$$\|R_{\alpha(\delta)} f^\delta - A^\dagger f\|_H \to 0, \quad \delta \to 0, \tag{2.6.8}$$

where $f^\delta \in \tilde{H}$ is the noisy data and $R_{\alpha(\delta)} f^\delta =: u^\delta_{\alpha(\delta)}$ is the regularized solution.

Proof $u^\delta_{\alpha(\delta)}$ furnishes a minimum to the regularized Tikhonov functional,

$$J_\alpha(v) := \frac{1}{2}\|Av - f^\delta\|^2_{\tilde{H}} + \frac{1}{2}\alpha\|v\|^2_H,\ v \in H,\ f^\delta \in \tilde{H}. \tag{2.6.9}$$

Hence $J_\alpha(u^\delta_{\alpha(\delta)}) \le J_\alpha(v)$, for all $v \in H$. In particular, for the unique solution $u \in \mathcal{N}(A)^\perp$ of the equation $Au = f$ with the noise free data $f \in \mathcal{R}(A)$ we have:

$$J_\alpha(u^\delta_{\alpha(\delta)}) \le J_\alpha(u). \tag{2.6.10}$$

According to (2.6.3), $\|Au^\delta_{\alpha(\delta)} - f^\delta\| = \delta$ and $\|Au - f^\delta\| = \delta,\ \ \delta > 0$. Taking into account this in (2.6.9) we conclude:

$$J_\alpha(u^\delta_{\alpha(\delta)}) = \frac{1}{2}\delta^2 + \frac{1}{2}\alpha\|u^\delta_{\alpha(\delta)}\|^2_H,$$
$$J_\alpha(u) = \frac{1}{2}\delta^2 + \frac{1}{2}\alpha\|u\|^2_H.$$

This implies that with (2.6.10) that

$$\|u^\delta_{\alpha(\delta)}\|_H \le \|u\|_H,\ \forall\delta > 0,\ \alpha > 0, \tag{2.6.11}$$

i.e. the set $\{u^\delta_{\alpha(\delta)}\}_{\delta>0}$ is uniformly bounded in H by the norm of the solution $u \in \mathcal{N}(A)^\perp$ of the equation $Au = f$.

Having the uniform boundedness of $\{u^\delta_{\alpha(\delta)}\}_{\delta>0}$ we estimate now the difference between the regularized and exact solutions:

$$\begin{aligned}\|u^\delta_{\alpha(\delta)} - u\|^2 &= \|u^\delta_{\alpha(\delta)}\|^2 - 2Re(u^\delta_{\alpha(\delta)}, u) + \|u\|^2 \\ &\le 2\left(\|u\|^2 - Re(u^\delta_{\alpha(\delta)}, u)\right) = 2Re(u - u^\delta_{\alpha(\delta)}, u).\end{aligned}$$

Since $u = A^*v$, we transform this estimate as follows:

$$\begin{aligned}\|u^\delta_{\alpha(\delta)} - u\|^2 &\le 2Re(u - u^\delta_{\alpha(\delta)}, A^*v) = 2Re(Au - Au^\delta_{\alpha(\delta)}, v) \\ = 2Re(f - Au^\delta_{\alpha(\delta)}, v) &\le 2Re(f - f^\delta, v) + 2Re(f^\delta - Au^\delta_{\alpha(\delta)}, v) \\ &\le 2\|v\|\left[\|f - f^\delta\| + \|f^\delta - Au^\delta_{\alpha(\delta)}\|\right].\end{aligned}$$

Using conditions (2.6.3) and $\|v\|_{\tilde{H}} \le M$ on the right hand side we finally get:

$$\|u^\delta_{\alpha(\delta)} - u\| \le 2\sqrt{M\delta}. \tag{2.6.12}$$

The right hand side tends to zero as $\delta \to 0^+$, which is the desired result. □

Remark 2.6.1 Estimate (2.6.12) depends on the norm of the element $v \in \tilde{H}$ which image $u = A^*v$ under the adjoint operator $A^* : \tilde{H} \mapsto H$ is the solution of the equation $Au = f$ with the noise free data $f \in \mathcal{R}(A)$. This element $v \in \tilde{H}$ is called a sourcewise element. In this we say that the exact solution $u = A^\dagger f$ satisfies the source condition $u = A^*v$, $v \in \tilde{H}$.

In applications, it is not necessary to satisfy the condition (2.6.3) exactly. Instead, the relaxed form condition

$$\beta_* \delta \le \|Au^\delta_{\alpha(\delta)} - f^\delta\| \le \beta^* \delta, \quad \beta^* > \beta_* > 0,\ \delta > 0 \tag{2.6.13}$$

can be used.

Therefore, if the noise level $\delta > 0$ is known, then in the iteration algorithm one of the forms of condition (2.6.13) is used as a *stopping rule*, according to Morozov's Discrepancy Principle. Iteration is terminated if

$$\|Au^{\delta,n(\delta)}_{\alpha(\delta)} - f^\delta\| \le \tau_M \delta < \|Au^{\delta,n(\delta)-1}_{\alpha(\delta)} - f^\delta\|, \quad \tau_M > 1,\ \delta > 0, \tag{2.6.14}$$

that is, if for the first time the condition

$$\|Au^{\delta,n(\delta)}_{\alpha(\delta)} - f^\delta\| \le \tau_M \delta \quad \tau_M > 1,\ \delta > 0 \tag{2.6.15}$$

holds. Here $\tau_M > 1$ is a fixed parameter.

http://www.springer.com/978-3-319-62796-0

Introduction to Inverse Problems for Differential Equations
Hasanov Hasanoğlu, A.; Romanov, V.G.
2017, XIII, 261 p. 4 illus., Hardcover
ISBN: 978-3-319-62796-0