Operator theory and integral equations
802660S

Lecture Notes
Second printing

Valery Serov
University of Oulu
2012

Edited by Markus Harju

Contents

1 Inner product spaces and Hilbert spaces

A collection of elements is called a complex (real) *vector space* (*linear space*) H if the following axioms are satisfied:

1) To every pair $x, y \in H$ there corresponds a vector $x + y$, called the sum, with the properties:

   a) $x + y = y + x$
   b) $x + (y + z) = (x + y) + z \equiv x + y + z$
   c) there exists unique $0 \in H$ such that $x + 0 = x$
   d) for every $x \in H$ there exists unique $y_1 \in H$ such that $x + y_1 = 0$. We denote $y_1 := -x$.

2) For every $x \in H$ and every $\lambda, \mu \in \mathbb{C}$ there corresponds a vector $\lambda \cdot x$ such that

   a) $\lambda(\mu x) = (\lambda\mu)x \equiv \lambda\mu x$
   b) $(\lambda + \mu)x = \lambda x + \mu x$
   c) $\lambda(x + y) = \lambda x + \lambda y$
   d) $1 \cdot x = x$.

Definition. For a linear space H a mapping $(\cdot, \cdot) : H \times H \to \mathbb{C}$ is called an *inner product* or a *scalar product* if

1) $(x, x) \geq 0$ and $(x, x) = 0$ if and only if $x = 0$
2) $(x, y + z) = (x, y) + (x, z)$
3) $(\lambda x, y) = \lambda(x, y)$
4) $(x, y) = \overline{(y, x)}$

for every $x, y, z \in H$ and $\lambda \in \mathbb{C}$. A linear space equipped with an inner product is called an *inner product space*.

An immediate consequence of this definition is that

$$\begin{aligned} (\lambda x + \mu y, z) &= \lambda(x, z) + \mu(y, z), \\ (x, \lambda y) &= \overline{\lambda}(x, y) \end{aligned}$$

for every $x, y, z \in H$ and $\lambda, \mu \in \mathbb{C}$.

Example 1.1. On the complex Euclidean space $H = \mathbb{C}^n$ the standard inner product is

$$(x, y) = \sum_{j=1}^{n} x_j \overline{y_j},$$

where $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$ and $y = (y_1, \ldots, y_n) \in \mathbb{C}^n$.

Example 1.2. On the linear space $C[a,b]$ of continuous complex-valued functions, the formula
$$(f,g) = \int_a^b f(x)\overline{g(x)}dx$$
defines an inner product.

Definition. Suppose H is an inner product space. We say that

1) $x \in H$ *orthogonal* to $y \in H$ if $(x,y) = 0$.

2) a system $\{x_\alpha\}_{\alpha \in \mathcal{A}} \subset H$ *orthonormal* if $(x_\alpha, x_\beta) = \delta_{\alpha,\beta} = \begin{cases} 1, & \alpha = \beta \\ 0, & \alpha \neq \beta \end{cases}$, where $\mathcal{A}$ is some index set.

3) $\|x\| := \sqrt{(x,x)}$ is called the *length* of $x \in H$.

Exercise 1. Prove the *Theorem of Pythagoras*: If $\{x_j\}_{j=1}^k, k \in \mathbb{N}$ is an orthonormal system in an inner product space H, then
$$\|x\|^2 = \sum_{j=1}^k |(x,x_j)|^2 + \left\| x - \sum_{j=1}^k (x,x_j)x_j \right\|^2$$
for every $x \in H$.

Exercise 2. Prove *Bessel's inequality*: If $\{x_j\}_{j=1}^k, k \leq \infty$ is an orthonormal system then
$$\sum_{j=1}^k |(x,x_j)|^2 \leq \|x\|^2,$$
for every $x \in H$.

Exercise 3. Prove the *Cauchy-Schwarz-Bunjakovskii inequality*:
$$|(x,y)| \leq \|x\| \|y\|, \quad x,y \in H.$$
Prove also that $(\cdot,\cdot)$ is continuous as a map from $H \times H$ to $\mathbb{C}$.

If H is an inner product space, then
$$\|x\| := \sqrt{(x,x)}$$
has the following properties:

1) $\|x\| \geq 0$ for every $x \in H$ and $\|x\| = 0$ if and only if $x = 0$.

2) $\|\lambda x\| = |\lambda| \|x\|$ for every $x \in H$ and $\lambda \in \mathbb{C}$.

3) $\|x+y\| \leq \|x\| + \|y\|$ for every $x,y \in H$. This is the *triangle inequality*.

The function $\|\cdot\| = \sqrt{(\cdot,\cdot)}$ is thus a *norm* on H. It is called the norm *induced by the inner product*.

Every inner product space H is a normed space under the induced norm. The *neighborhood* of $x \in H$ is the *open ball* $B_r(x) = \{y \in H : \|x-y\| < r\}$. This system of neighborhoods defines the *norm topology* on H such that:

1) The addition $x+y$ is a continuous map $H \times H \to H$.

2) The scalar multiplication $\lambda \cdot x$ is a continuous map $\mathbb{C} \times H \to H$.

3) The inner product $(x,y) : H \times H \to \mathbb{C}$ is continuous.

Definition. 1) A sequence $\{x_j\}_{j=1}^{\infty} \subset H$ is called a *Cauchy sequence* if for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $\|x_k - x_j\| < \varepsilon$ for $k, j \geq n_0$.

2) A sequence $\{x_j\}_{j=1}^{\infty} \subset H$ is said to be *convergent* if there exists $x \in H$ such that for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $\|x - x_j\| < \varepsilon$ whenever $j \geq n_0$.

3) The inner product space H is *complete space* if every Cauchy sequence in H converges.

Corollary. *1) Every convergent sequence is a Cauchy sequence.*

2) If $\{x_j\}_{j=1}^{\infty}$ converges to $x \in H$ then

$$\lim_{j\to\infty} \|x_j\| = \|x\|.$$

Definition (J. von Neumann, 1925)**.** A *Hilbert space* is an inner product space which is complete (with respect to its norm topology).

Exercise 4. Prove that in an inner product space the norm induced by this inner product satisfies the *parallelogram law*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Exercise 5. Prove that if in a normed space H the parallelogram law holds, then there is an inner product on H such that $\|x\|^2 = (x,x)$ and that this inner product is defined by the *polarization identity*

$$(x,y) := \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right).$$

Exercise 6. Prove that on $C[a,b]$ the norm

$$\|f\| = \max_{x\in[a,b]} |f(x)|$$

is not induced by an inner product.

Exercise 7. Give an example of an inner product space which is not complete.

Next we list some examples of Hilbert spaces.

1) The Euclidean spaces $\mathbb{R}^n$ and $\mathbb{C}^n$.

2) The matrix space $M_n(\mathbb{C})$ consisting of $n\times n$ -matrices whose elements are complex numbers. For $A, B \in M_n(\mathbb{C})$ the inner product is given by

$$(A, B) = \sum_{k,j=1}^{n} a_{kj}\overline{b_{kj}} = \mathrm{Tr}\,(AB^*),$$

where $B^* = \overline{B}^T$.

3) The *sequence space* $l^2(\mathbb{C})$ defined by

$$l^2(\mathbb{C}) := \left\{ \{x_j\}_{j=1}^{\infty}, x_j \in \mathbb{C} : \sum_{j=1}^{\infty} |x_j|^2 < \infty \right\}.$$

The estimates

$$|x_j + y_j|^2 \le 2\left(|x_j|^2 + |y_j|^2\right), \quad |\lambda x_j|^2 = |\lambda|^2 |x_j|^2$$

and

$$|x_j y_j| \le \frac{1}{2}\left(|x_j|^2 + |y_j|^2\right)$$

imply that $l^2(\mathbb{C})$ is a linear space. Let us define the inner product by

$$(x, y) := \sum_{j=1}^{\infty} x_j \overline{y_j}$$

and prove that $l^2(\mathbb{C})$ is complete. Suppose that $\{x^{(k)}\}_{k=1}^{\infty} \in l^2(\mathbb{C})$ is a Cauchy sequence. Then for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that

$$\left\|x^{(k)} - x^{(m)}\right\|^2 = \sum_{j=1}^{\infty} |x_j^{(k)} - x_j^{(m)}|^2 < \varepsilon^2$$

for $k, m \ge n_0$. It implies that

$$|x_j^{(k)} - x_j^{(m)}| < \varepsilon, \quad j = 1, 2, \ldots$$

or that $\{x_j^{(k)}\}_{k=1}^{\infty}$ is a Cauchy sequence in $\mathbb{C}$ for every $j = 1, 2, \ldots$. Since $\mathbb{C}$ is a complete space then $\{x_j^{(k)}\}_{k=1}^{\infty}$ converges for every fixed $j = 1, 2, \ldots$ i.e. there exists $x_j \in \mathbb{C}$ such that

$$x_j = \lim_{k\to\infty} x_j^{(k)}.$$

This fact and

$$\sum_{j=1}^{l} |x_j^{(k)} - x_j^{(m)}|^2 < \varepsilon^2, \quad l \in \mathbb{N}$$

imply that

$$\lim_{m\to\infty} \sum_{j=1}^{l} |x_j^{(k)} - x_j^{(m)}|^2 = \sum_{j=1}^{l} |x_j^{(k)} - x_j|^2 \leq \varepsilon^2$$

for all $k \geq n_0$ and $l \in \mathbb{N}$. Therefore the sequence

$$s_l := \sum_{j=1}^{l} |x_j^{(k)} - x_j|^2, \quad k \geq n_0$$

is a monotone increasing sequence which is bounded from above by ε^2. Hence this sequence has a limit with the same upper bound i.e.

$$\sum_{j=1}^{\infty} |x_j^{(k)} - x_j|^2 = \lim_{l\to\infty} \sum_{j=1}^{l} |x_j^{(k)} - x_j|^2 \leq \varepsilon^2.$$

That's why we may conclude that

$$\|x\| \leq \left\|x^{(k)}\right\| + \left\|x^{(k)} - x\right\| \leq \left\|x^{(k)}\right\| + \varepsilon$$

and $x \in l^2(\mathbb{C})$.

4) The *Lebesgue space* $L^2(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is an open set. The space $L^2(\Omega)$ consists of all Lebesgue measurable functions f which are square integrable i.e.

$$\int_\Omega |f(x)|^2 dx < \infty.$$

It is a linear space with the inner product

$$(f, g) = \int_\Omega f(x)\overline{g(x)} dx$$

and the Riesz-Fisher theorem reads as: $L^2(\Omega)$ is a Hilbert space.

5) The *Sobolev spaces* $W_2^k(\Omega)$ consisting of functions $f \in L^2(\Omega)$ whose weak or distributional derivatives $D^\alpha f$ also belong to $L^2(\Omega)$ up to order $|\alpha| \leq k, k = 1, 2, \ldots$. On the space $W_2^k(\Omega)$ the natural inner product is

$$(f, g) = \sum_{|\alpha|\leq k} \int_\Omega D^\alpha f(x)\overline{D^\alpha g(x)} dx.$$

Definition. Let H be an inner product space. For any subspace $M \subset H$ the *orthogonal complement* of M is defined as

$$M^{\perp} := \{y \in H : (y,x) = 0, \text{for all}\, x \in M\}.$$

Remark. It is clear that $M^{\perp}$ is a linear subspace of H. Moreover, $M \cap M^{\perp} = \{0\}$ since $0 \in M$ always.

Definition. A *closed subspace* of a Hilbert space H is a linear subspace of H which is closed (i.e. $\overline{M} = M$) with respect to the induced norm.

Remark. The subspace $M^{\perp}$ is closed if M is any subset of a Hilbert space.

Theorem 1 (*Projection theorem*)**.** *Suppose M is a closed subspace of a Hilbert space H. Then every $x \in H$ has the unique representation as*

$$x = u + v,$$

where $u \in M$ and $v \in M^{\perp}$, or equivalently,

$$H = M \oplus M^{\perp}.$$

Moreover, one has that

$$\|v\| = \inf_{y \in M} \|x - y\| := d(x, M).$$

Proof. Let $x \in H$. Then

$$d := d(x, M) \equiv \inf_{y \in M} \|x - y\| \leq \|x - u\|$$

for any $u \in M$. The definition of infimum implies that there exists a sequence $\{u_j\}_{j=1}^{\infty} \subset M$ such that

$$d = \lim_{j \to \infty} \|x - u_j\|.$$

The parallelogram law implies that

$$\begin{aligned} \|u_j - u_k\|^2 &= \|(u_j - x) + (x - u_k)\|^2 \\ &= 2\|u_j - x\|^2 + 2\|x - u_k\|^2 - 4\left\|x - \frac{u_j + u_k}{2}\right\|^2. \end{aligned}$$

Since $(u_j + u_k)/2 \in M$ then

$$\|u_j - u_k\|^2 \leq 2\|u_j - x\|^2 + 2\|x - u_k\|^2 - 4d^2 \to 2d^2 + 2d^2 - 4d^2 = 0$$

as $j, k \to \infty$. Hence $\{u_j\}_{j=1}^{\infty} \subset M$ is a Cauchy sequence in the Hilbert space H. It means that there exists $u \in H$ such that

$$u = \lim_{j \to \infty} u_j.$$

But $M = \overline{M}$ implies that $u \in M$. By construction one has that

$$d = \lim_{j\to\infty} \|x - u_j\| = \|x - u\| .$$

Let us denote $v := x - u$ and show that $v \in M^\perp$. For any $y \in M, y \neq 0$ introduce the number

$$\alpha = -\frac{(v,y)}{\|y\|^2}.$$

Since $u - \alpha y \in M$ we have

$$\begin{aligned} d^2 &\leq \|x - (u - \alpha y)\|^2 = \|v + \alpha y\|^2 = \|v\|^2 + (v, \alpha y) + (\alpha y, v) + |\alpha|^2 \|y\|^2 \\ &= d^2 - \frac{\overline{(v,y)}(v,y)}{\|y\|^2} - \frac{(v,y)(y,v)}{\|y\|^2} + \frac{|(v,y)|^2}{\|y\|^2} = d^2 - \frac{|(v,y)|^2}{\|y\|^2}. \end{aligned}$$

This inequality implies that $(y,v) = 0$. It means that $v \in M^\perp$. In order to prove uniqueness assume that $x = u_1 + v_1 = u_2 + v_2$, where $u_1, u_2 \in M$ and $v_1, v_2 \in M^\perp$. It follows that

$$u_1 - u_2 = v_2 - v_1 \in M \cap M^\perp.$$

But $M \cap M^\perp = \{0\}$ so that $u_1 = u_2$ and $v_1 = v_2$. □

Corollary 1 (*Riesz-Frechet theorem*)**.** *If T is a linear continuous functional on the Hilbert space H then there exists a unique $h \in H$ such that $T(x) = (x,h)$ for all $x \in H$. Moreover, $\|T\|_{H\to\mathbb{C}} = \|h\|$.*

Proof. If $T \equiv 0$ then $h = 0$ will do. If $T \neq 0$ then there exists $v_0 \in H$ such that $T(v_0) \neq 0$. Let

$$M := \{u \in H : T(u) = 0\}.$$

Then $v_0 \in M^\perp, v_0 \neq 0$ and $T(v_0) \neq 0$. Since T is linear and continuous then M is a closed subspace. It follows from Thereom 1 that

$$H = M \oplus M^\perp$$

i.e. every $x \in H$ has the unique representation as $x = \widetilde{u} + \widetilde{v}$. Therefore, for every $x \in H$, we can define

$$u := x - \frac{T(x)}{T(v_0)} v_0.$$

Then $T(u) = 0$ i.e. $u \in M$. It follows that

$$(x, v_0) = (u, v_0) + \frac{T(x)}{T(v_0)} \|v_0\|^2 = \frac{T(x)}{T(v_0)} \|v_0\|^2$$

or

$$T(x) = \frac{T(v_0)}{\|v_0\|^2}(x, v_0) = \left(x, \frac{\overline{T(v_0)}}{\|v_0\|^2} v_0\right),$$

which is of the desired form. The uniqueness of h can be seen as follows. If $T(x) = (x,h) = (x,\widetilde{h})$ then $(x, h-\widetilde{h}) = 0$ for all $x \in H$. In particular $\left\|h-\widetilde{h}\right\|^2 = (h-\widetilde{h}, h-\widetilde{h}) = 0$ i.e. $h = \widetilde{h}$. It remains to prove the statement about the norm $\|T\|_{H\to\mathbb{C}} = \|T\|$. Firstly,

$$\|T\| = \sup_{\|x\|\le 1} |T(x)| = \sup_{\|x\|\le 1} |(x,h)| \le \|h\|.$$

On the other hand $T(h/\|h\|) = \|h\|$ implies that $\|T\| \ge \|h\|$. Thus $\|T\| = \|h\|$. This finishes the proof. □

Corollary 2. *If M is a linear subspace of a Hilbert space H then*

$$M^{\perp\perp} := \left(M^{\perp}\right)^{\perp} = \overline{M}.$$

Proof. It is not so difficult to check that

$$M^{\perp} = \left(\overline{M}\right)^{\perp}.$$

That's why

$$M^{\perp\perp} = \left(\left(\overline{M}\right)^{\perp}\right)^{\perp}$$

and Theorem 1 implies that

$$H = \overline{M} \oplus \left(\overline{M}\right)^{\perp}, \quad H = \left(\overline{M}\right)^{\perp} \oplus M^{\perp\perp}.$$

Uniqueness of this representation guarantees that $M^{\perp\perp} = \overline{M}$. □

Remark. In the frame of this theorem we have that

$$\|x\|^2 = \|u\|^2 + \|v\|^2, \quad \|v\|^2 = (x,v), \quad \|u\|^2 = (x,u).$$

Definition. Let $A \subset H$ be a subset of an inner product space. The subset

$$\operatorname{span} A := \left\{x \in H : x = \sum_{j=1}^{k} \lambda_j x_j, x_j \in A, \lambda_j \in \mathbb{C}\right\}$$

is called the *linear span* of A.

Definition. Let H be a Hilbert space.

1) A subset $B \subset H$ is called a *basis* of H if B is linearly independent in H and

$$\overline{\operatorname{span} B} = H$$

i.e. for every $x \in H$ and every $\varepsilon > 0$ there exist $k \in \mathbb{N}$ and $\{c_j\}_{j=1}^{k} \subset \mathbb{C}$ such that

$$\left\|x - \sum_{j=1}^{k} c_j x_j\right\| < \varepsilon, \quad x_j \in B.$$

2) The Hilbert space is called *separable* if it has a countable or finite basis.

3) An orthonormal system $B = \{x_\alpha\}_{\alpha \in A}$ in H which is a basis is called an *orthonormal basis*.

By the Gram-Schmidt orthonormalization we may conclude that every separable Hilbert space has an orthonormal basis.

Theorem 2 (Characterization of an orthonormal basis)**.** *Let $B = \{x_j\}_{j=1}^{\infty}$ be an orthonormal system in a separable Hilbert space H. Then the following statements are equivalent:*

1) B is maximal i.e. it is not a proper subset of any other orthonormal system.

2) For every $x \in H$ the condition $(x, x_j) = 0, j = 1, 2, \ldots$ implies that $x = 0$.

3) Every $x \in H$ has the Fourier expansion

$$x = \sum_{j=1}^{\infty} (x, x_j) x_j$$

i.e.

$$\left\| x - \sum_{j=1}^{k} (x, x_j) x_j \right\| \to 0, \quad k \to \infty.$$

This means that B is an orthonormal basis.

4) Every pair $x, y \in H$ satisfies the completeness relation

$$(x, y) = \sum_{j=1}^{\infty} (x, x_j) \overline{(y, x_j)}.$$

5) Every $x \in H$ satisfies the Parseval equality

$$\|x\|^2 = \sum_{j=1}^{\infty} |(x, x_j)|^2.$$

Proof. **1)$\Rightarrow$ 2)** Suppose that there is $z \in H, z \neq 0$ such that $(z, x_j) = 0$ for all $j = 1, 2, \ldots$. Then

$$B' := \left\{ \frac{z}{\|z\|}, x_1, x_2, \ldots \right\}$$

is an orthonormal system in H. This fact implies that B is not maximal. It contradicts 1) and proves 2).

2)⇒3) Given $x \in H$ introduce the sequence

$$x^{(k)} = \sum_{j=1}^{k} (x, x_j) x_j.$$

Theorem of Pythagoras and Bessel's inequality (Exercises 1 and 2) imply that

$$\left\| x^{(k)} \right\|^2 = \sum_{j=1}^{k} |(x, x_j)|^2 \leq \|x\|^2 .$$

It follows that

$$\sum_{j=1}^{\infty} |(x, x_j)|^2$$

converges. That's why, for $m < k$,

$$\left\| x^{(k)} - x^{(m)} \right\|^2 = \sum_{j=m+1}^{k} |(x, x_j)|^2 \to 0$$

as $k, m \to \infty$. Hence $x^{(k)}$ is a Cauchy sequence in H. Thus there exists $y \in H$ such that

$$y = \lim_{k \to \infty} x^{(k)} = \sum_{j=1}^{\infty} (x, x_j) x_j.$$

Next, since the inner product is continuous we deduce that

$$(y, x_j) = \lim_{k \to \infty} (x^{(k)}, x_j) = (x, x_j)$$

for any $j = 1, 2, \ldots$. Therefore $(y - x, x_j) = 0$ for any $j = 1, 2, \ldots$. Part 2) implies that $y = x$ and part 3) follows.

3)⇒4) Let $x, y \in H$. We know from part 3) that

$$x = \sum_{j=1}^{\infty} (x, x_j) x_j, \quad y = \sum_{k=1}^{\infty} (y, x_k) x_k.$$

Continuity of the inner product and orthonormality of $\{x_j\}_{j=1}^{\infty}$ allow us to conclude that

$$(x, y) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} (x, x_j) \overline{(y, x_k)} (x_j, x_k) = \sum_{j=1}^{\infty} (x, x_j) \overline{(y, x_j)}.$$

4)⇒5) Take $y = x$ in part 4).

5)⇒1) Suppose that B is not maximal. Then we can add a unit vector $z \in H$ to it which is orthogonal to B. Parseval's equality gives then

$$1 = \|z\|^2 = \sum_{j=1}^{\infty} |(z, x_j)|^2 = 0.$$

This contradiction proves the result.

□

Exercise 8. Let $\{x_j\}_{j=1}^{\infty}$ be an orthonormal system in an inner product space H. Let $x \in H$, $\{c_j\}_{j=1}^{k} \subset \mathbb{C}$ and $k \in \mathbb{N}$. Prove that

$$\left\| x - \sum_{j=1}^{k} (x, x_j) x_j \right\| \leq \left\| x - \sum_{j=1}^{k} c_j x_j \right\|.$$

2 Symmetric operators in the Hilbert space

Assume that H is a Hilbert space. A *linear operator* from H to H is a mapping

$$A : D(A) \subset H \to H,$$

where $D(A)$ is a linear subspace of H and A satisfies the condition

$$A(\lambda x + \mu y) = \lambda Ax + \mu Ay$$

for all $\lambda, \mu \in \mathbb{C}$ and $x, y \in D(A)$. The space $D(A)$ is called the *domain* of A. The space

$$N(A) := \{x \in D(A) : Ax = 0\}$$

is called the *nullspace* (or the *kernel*) of A. The space

$$R(A) := \{y \in H : y = Ax \text{ for some } x \in D(A)\}$$

is called the *range* of A. Both $N(A)$ and $R(A)$ are linear subspaces of H. We say that A is *bounded* if there exists $M > 0$ such that

$$\|Ax\| \le M \|x\|, \quad x \in D(A).$$

We say that A is *densely defined* if $\overline{D(A)} = H$. In such case A can be extended to A_{ex} which will be defined on the whole H with the same norm estimate and we may define

$$\|A\|_{H\to H} := \inf\{M : \|Ax\| \le M \|x\|, x \in D(A)\}$$

or equivalently

$$\|A\|_{H\to H} = \sup_{\|x\|=1} \|Ax\|.$$

Exercise 9 (Hellinger-Toeplitz)**.** Suppose that $D(A) = H$ and

$$(Ax, y) = (x, Ay), \quad x, y \in H.$$

Prove that A is bounded.

Example 2.1 (Integral operator in L^2)**.** Suppose that $K(s,t) \in L^2(\Omega \times \Omega), \Omega \subset \mathbb{R}^n$. Define the integral operator $\widehat{K}$ as

$$\widehat{K}f(s) = \int_\Omega K(s,t)f(t)dt, \quad f \in L^2(\Omega).$$

Let us prove that $\widehat{K}$ is bounded. Indeed,

$$\begin{aligned}
\left\|\widehat{K}f\right\|^2_{L^2(\Omega)} &= \int_\Omega |\widehat{K}f(s)|^2 ds = \int_\Omega \left|\int_\Omega K(s,t)f(t)dt\right|^2 ds \\
&= \int_\Omega \left|(K(s,\cdot), \overline{f})_{L^2}\right|^2 ds \le \int_\Omega \|K(s,\cdot)\|^2_{L^2} \left\|\overline{f}\right\|^2_{L^2} ds \\
&= \int_\Omega \left(\int_\Omega |K(s,t)|^2 dt \int_\Omega |f(t)|^2 dt\right) ds \\
&= \|K\|^2_{L^2(\Omega\times\Omega)} \|f\|^2_{L^2(\Omega)}
\end{aligned}$$

where we have made use of Cauchy-Schwarz-Bunjakovskii inequality. That's why we have

$$\left\|\widehat{K}\right\|_{L^2\to L^2} \le \|K\|_{L^2(\Omega\times\Omega)}.$$

The norm

$$\|K\|_{L^2(\Omega\times\Omega)} := \left\|\widehat{K}\right\|_{HS}$$

is called the *Hilbert-Schmidt norm* of $\widehat{K}$.

Example 2.2 (Schur test)**.** Assume that p and q are positive measurable functions on $\Omega \subset \mathbb{R}^n$ and α and β are positive numbers such that

$$\int_\Omega |K(x,y)|p(y)dy \le \alpha q(x), \quad \text{a.e. in } \Omega$$

and

$$\int_\Omega |K(x,y)|q(x)dx \le \beta p(y), \quad \text{a.e. in } \Omega.$$

Then $\widehat{K}$ is bounded and

$$\left\|\widehat{K}\right\|_{L^2\to L^2} \le \sqrt{\alpha\beta}.$$

Proof. For any $f \in L^2(\Omega)$ we have

$$\begin{aligned}
\int \left(\int |K(x,y)|\cdot|f(y)|dy\right)^2 dx &= \int \left(\int \sqrt{|K(x,y)|}\sqrt{p(y)}\sqrt{\frac{|K(x,y)|}{p(y)}}|f(y)|dy\right)^2 dx \\
&\le \int \left(\int |K(x,y)|p(y)dy\right)\left(\int \frac{|K(x,y)|}{p(y)}|f(y)|^2dy\right)dx \\
&\le \alpha \int \left(\int |K(x,y)|q(x)dx\right)\frac{|f(y)|^2}{p(y)}dy \\
&\le \alpha\beta \int |f(y)|^2dy
\end{aligned}$$

by the Cauchy-Schwarz-Bunjakovskii inequalityand Fubini's theorem. □

Example 2.3 (Differential operator in L^2)**.** Consider the differential operator

$$A := i\frac{d}{dt}$$

of order 1 in $L^2(0,1)$ with the domain

$$D(A) = \left\{f \in C^1[0,1] : f(0) = f(1) = 0\right\}.$$

First of all we have that $\overline{D(A)} = L^2$. Moreover, integration by parts gives

$$(Af,g) = \int_0^1 if'(t)\overline{g(t)}dt = i\left[f\overline{g}|_0^1 - \int_0^1 f(t)\overline{g'(t)}dt\right] = \int_0^1 f(t)\overline{ig'(t)}dt = (f,Ag)$$

for all $f, g \in D(A)$. Let us now consider the sequence

$$u_n(t) := \sin(n\pi t), \quad n = 1, 2, \ldots.$$

Clearly $u_n \in D(A)$ and

$$\|u_n\|_{L^2}^2 = \int_0^1 |\sin(n\pi t)|^2 dt = \frac{1}{2}.$$

But

$$\|Au_n\|_{L^2}^2 = \int_0^1 \left| i\frac{d}{dt}\sin(n\pi t) \right|^2 dt = (n\pi)^2 \int_0^1 |\cos(n\pi t)|^2 dt = (n\pi)^2 \frac{1}{2} = (n\pi)^2 \|u_n\|_{L^2}^2.$$

Therefore A is not bounded. This shows that $D(A) = H$ is an essential assumption in Exercise 9.

Example 2.4 (Differential operator in L^2)**.** Consider the differential operator

$$A := p_0 \frac{d^2}{dt^2} + ip_1 \frac{d}{dt} + p_2$$

of order 2 in $L^2(0, 1)$ with the domain

$$D(A) = \{f \in C^2[0, 1] : f(0) = f(1) = 0\}$$

and with real constant coefficients p_0, p_1 and p_2. The fact $\overline{D(A)} = L^2$ and integration by parts gives

$$\begin{aligned}
(Af, g) &= p_0 \int_0^1 f'' \cdot \overline{g} dt + ip_1 \int_0^1 f' \cdot \overline{g} dt + p_2 \int_0^1 f \cdot \overline{g} dt \\
&= p_0 \left[f'\overline{g}|_0^1 - \int_0^1 f' \cdot \overline{g'} dt \right] + ip_1 \left[f\overline{g}|_0^1 - \int_0^1 f \cdot \overline{g'} dt \right] + p_2 (f, g)_{L^2} \\
&= -p_0 \int_0^1 f' \cdot \overline{g'} dt - ip_1 \int_0^1 f \cdot \overline{g'} dt + (f, p_2 g)_{L^2} \\
&= -p_0 \left[f\overline{g'}|_0^1 - \int_0^1 f \cdot \overline{g''} dt \right] + (f, ip_1 g')_{L^2} + (f, p_2 g)_{L^2} \\
&= p_0 \int_0^1 f \cdot \overline{g''} dt + (f, ip_1 g')_{L^2} + (f, p_2 g)_{L^2} = (f, Ag)_{L^2}
\end{aligned}$$

for all $f, g \in D(A)$. Moreover, for the sequence $u_n(t) = \sin(n\pi t)$ we have (for sufficiently large n) that

$$
\begin{aligned}
\|Au_n\|_{L^2}^2 &= \int_0^1 |p_0(\sin(n\pi t))'' + ip_1(\sin(n\pi t))' + p_2\sin(n\pi t)|^2 dt \\
&= \int_0^1 \left[(p_0(n\pi)^2 - p_2)^2 \sin^2(n\pi t) + (n\pi)^2 p_1^2 \cos^2(n\pi t)\right] dt \\
&\geq \int_0^1 \left[\frac{(n\pi)^4}{2} p_0^2 \sin^2(n\pi t) + (n\pi)^2 p_1^2 \cos^2(n\pi t)\right] dt \\
&\geq (n\pi)^2 p_1^2 \int_0^1 \left(\sin^2(n\pi t) + \cos^2(n\pi t)\right) dt \\
&= 2(n\pi)^2 p_1^2 \frac{1}{2} = 2(n\pi)^2 p_1^2 \|u_n\|_{L^2}^2 .
\end{aligned}
$$

So A is unbounded, since

$$
\|A\|_{L^2 \to L^2} \geq 2(n\pi)^2 p_1^2
$$

for $n \to \infty$.

We assume later on that $\overline{D(A)} = H$ i.e. that A is densely defined in any case.

Definition. The *graph* $\Gamma(A)$ of a linear operator A in the Hilbert space H is defined as

$$
\Gamma(A) := \{(x; y) \in H \times H : x \in D(A) \text{ and } y = Ax\} .
$$

Remark. The graph $\Gamma(A)$ is a linear subspace of the Hilbert space $H \times H$. The inner product in $H \times H$ can be defined as

$$
((x_1; y_1), (x_2; y_2))_{H\times H} := (x_1, x_2)_H + (y_1, y_2)_H
$$

for any $(x_1; y_1), (x_2; y_2) \in H \times H$.

Definition. The operator A is called *closed* if $\Gamma(A) = \overline{\Gamma(A)}$. We denote this fact by $A = \overline{A}$.

By definition, the *criterion for closedness* is that

$$
\begin{cases} x_n \in D(A) \\ x_n \to x \\ Ax_n \to y \end{cases} \Rightarrow \quad \begin{cases} x \in D(A) \\ y = Ax. \end{cases}
$$

The reader is asked to verify that it is also possible to use seemingly weaker, but equivalent, criterion:

$$
\begin{cases} x_n \in D(A) \\ x_n \overset{w}{\to} x \\ Ax_n \overset{w}{\to} y \end{cases} \Rightarrow \quad \begin{cases} x \in D(A) \\ y = Ax, \end{cases}
$$

where $x_n \overset{w}{\to} x$ indicates weak convergence in the sense that

$$(x_n, y) \to (x, y)$$

for all $y \in H$.

Remark. It is important from the point of view of applications (in particular, for numerical procedures) that the closedness of an operator guarantees the convergence of some process to the "correct" (right) result.

Definition. Let A and A_1 be two linear operators in a Hilbert space H. We say that A_1 is an *extension* of A (or A is a *restriction* of A_1) if $D(A) \subset D(A_1)$ and $Ax = A_1 x$ for all $x \in D(A)$. We denote this fact by $A \subset A_1$ and $A = A_1|_{D(A)}$.

Definition. We say that A is *closable* if A has an extension A_1 and $A_1 = \overline{A_1}$. The *closure* of A, denoted by $\overline{A}$, is the smallest closed extension of A if it exists, i.e.

$$\overline{A} = \bigcap_{\substack{A \subset A_1 \\ A_1 = \overline{A_1}}} A_1.$$

Here, by $A_1 \cap \widetilde{A_1}$, we mean the operator whose domain is $D(A_1 \cap \widetilde{A_1}) := D(A_1) \cap D(\widetilde{A_1})$ and

$$(A_1 \cap \widetilde{A_1})x := A_1 x = \widetilde{A_1} x, \quad x \in D(A_1 \cap \widetilde{A_1}),$$

whenever $A \subset A_1 = \overline{A_1}$ and $A \subset \widetilde{A_1} = \overline{\widetilde{A_1}}$.

If A is closable then $\Gamma\left(\overline{A}\right) = \overline{\Gamma(A)}$.

Definition. Consider the subspace

$$D^* := \{v \in H : \text{there exists}\, h \in H\, \text{such that}\, (Ax, v) = (x, h)\, \text{for all}\, x \in D(A)\}.$$

The operator A^* with the domain $D(A^*) := D^*$ and mapping $A^* v = h$ is called the *adjoint operator* of A.

Exercise 10. Prove that A^* exists as unique linear operator.

Remark. The adjoint operator is maximal among all linear operators B (in the sense that $B \subset A^*$) which satisfy

$$(Ax, y) = (x, By)$$

for all $x \in D(A)$ and $y \in D(B)$.

Example 2.5. Consider the operator

$$Af(x) := x^{-\alpha} f(x), \quad \alpha > 0$$

in the Hilbert space $H = L^2(0,1)$. Let us define

$$D(A) := \left\{f \in L^2(0,1) : f(x) = \chi_n(x) g(x), g \in L^2\, \text{for some}\, n \in \mathbb{N}\right\},$$

where

$$\chi_n(x) = \begin{cases} 0, & 0 \le x \le 1/n \\ 1, & 1/n < x \le 1. \end{cases}$$

It is clear that $\overline{D(A)} = L^2(0,1)$. For $v \in D(A^*)$ we have

$$(Af, v) = \int_0^1 x^{-\alpha}\chi_n(x)g(x)\overline{v(x)}dx = \int_0^1 f(x)\overline{x^{-\alpha}v(x)}dx = (f, A^*v).$$

That's why we may conclude that

$$D(A^*) = \left\{v \in L^2 : x^{-\alpha}v \in L^2\right\}.$$

Let us show that A is not closed. To see this take the sequence

$$f_n(x) = \begin{cases} x^{\alpha}, & 1/n < x \le 1 \\ 0, & 0 \le x \le 1/n. \end{cases}$$

Then $f_n \in D(A)$ and

$$Af_n(x) = \begin{cases} 1, & 1/n < x \le 1 \\ 0, & 0 \le x \le 1/n. \end{cases}$$

If we assume that $A = \overline{A}$ then

$$\begin{cases} f_n \in D(A) \\ f_n \to x^{\alpha} \\ Af_n \to 1 \end{cases} \Rightarrow \begin{cases} x^{\alpha} \in D(A) \\ 1 = Ax^{\alpha}. \end{cases}$$

But $x^{\alpha} \notin D(A)$. This contradiction shows us that A is not closed. It is not bounded either since $\alpha > 0$.

Theorem 1. *Let A be linear and densely defined operator. Then*

1) $A^* = \overline{A^*}$.

2) *A is closable if and only if $\overline{D(A^*)} = H$. In this case $A^{**} := (A^*)^* = \overline{A}$.*

3) *If A is closable then $\left(\overline{A}\right)^* = A^*$.*

Proof. 1) Let us define in $H \times H$ the linear and bounded operator V as the mapping

$$V : (u; v) \to (v; -u).$$

It has the property $V^2 = -I$. The equality $(Au, v) = (u, A^*v)$ for $u \in D(A)$ and $v \in D(A^*)$ can be rewritten as

$$(V(u; Au), (v; A^*v))_{H\times H} = 0.$$

It implies that $\Gamma(A^*) \perp V\Gamma(A)$ and $\Gamma(A^*) \perp \overline{V\Gamma(A)}$. It means (see Theorem 1 in Section 1) that $\Gamma(A^*) \subset \left(\overline{V\Gamma(A)}\right)^{\perp}$. Let us check that the criterion for closedness holds i.e.

$$\begin{cases} v_n \in D(A^*) \\ v_n \to v \\ A^* v_n \to y \end{cases} \quad \Rightarrow \quad \begin{cases} v \in D(A^*) \\ y = A^* v. \end{cases}$$

Indeed, for any $u \in D(A)$ we have

$$(Au, v_n) \to (Au, v).$$

On the other hand,

$$(Au, v_n) = (u, A^* v_n) \to (u, y).$$

Hence $(Au, v) = (u, y)$. Thus $v \in D(A^*)$ and $y = A^* v$. This proves 1).

2) Assume $\overline{D(A^*)} = H$. Then we can define $A^{**} := (A^*)^*$ and due to part 1) we may conclude that

$$\Gamma(A^{**}) \perp \overline{V\Gamma(A^*)}.$$

Next, since $\overline{\Gamma(A)}$ is a closed subspace of $H \times H$ we have $\overline{\Gamma(A)} = \left(\overline{\Gamma(A)}\right)^{\perp\perp}$. Since $V^2 = -I$ and V is bounded then

$$\overline{\Gamma(A)} = -\left(V^2 \overline{\Gamma(A)}\right)^{\perp\perp} = -\left(V\left(\overline{V\Gamma(A)}\right)^{\perp}\right)^{\perp} = -\left(V\Gamma(A^*)\right)^{\perp}$$

by 1). Hence

$$\overline{\Gamma(A)} \perp \overline{V\Gamma(A^*)}.$$

It follows that

$$\overline{\Gamma(A)} = \Gamma(A^{**})$$

or

$$\Gamma(\overline{A}) = \Gamma(A^{**})$$

or

$$\overline{A} = A^{**}.$$

This proves 2) in one direction. Let us assume now that A is closable but $\overline{D(A^*)} \neq H$. It is equivalent to the fact that there exists $u_0 \neq 0$ such that $u_0 \perp \overline{D(A^*)}$. In that case for any $v \in D(A^*)$ the element $(u_0; 0) \in H \times H$ is orthogonal to $(v; A^* v) \in \Gamma(A^*) \subset H \times H$. This is equivalent to (see 1)) $(u_0; 0) \in \overline{V\Gamma(A)}$. Since A is closable and V is bounded then $(u_0; 0) \in V\Gamma(\overline{A})$ or

$$-V(u_0; 0) = (0; u_0) \in \Gamma(\overline{A})$$

or $\overline{A}(0) = u_0$. Linearity of $\overline{A}$ implies $u_0 = 0$. This contradiction proves 2).

3) Since A is closable then

$$A^* \overset{1)}{=} \overline{A^*} \overset{2)}{=} (A^*)^{**} = (A)^{***} = (A^{**})^* \overset{2)}{=} \left(\overline{A}\right)^*.$$

This finishes the proof.

□

Example 2.6. Consider the Hilbert space $H = L^2(\mathbb{R})$ and the operator

$$Au(x) = (u, f_0)u_0(x),$$

where $u_0 \not\equiv 0, u_0 \in L^2(\mathbb{R})$ is fixed and $f_0 \neq 0$ is an arbitrary but fixed constant. We consider A on the domain

$$D(A) = \left\{ u \in L^2(\mathbb{R}) : \int_{\mathbb{R}} |f_0 u(x)| dx < \infty \right\} = L^2(\mathbb{R}) \cap L^1(\mathbb{R}).$$

It is known that $\overline{L^2(\mathbb{R}) \cap L^1(\mathbb{R})} = L^2(\mathbb{R})$. Thus A is densely defined. Let v be an element of $D(A^*)$. Then

$$(Au, v) = ((u, f_0)u_0, v) = (u, f_0)(u_0, v) = \left(u, \overline{(u_0, v)} f_0\right) = (u, (v, u_0) f_0).$$

It means that

$$A^* v = (v, u_0) f_0.$$

But $(v, u_0) f_0$ must belong to $L^2(\mathbb{R})$. Since $(v, u_0) f_0$ is a constant and $f_0 \neq 0$ then (v, u_0) must be equal to 0. Thus

$$u_0 \perp D(A^*)$$

which implies that

$$u_0 \perp \overline{D(A^*)}.$$

Since $u_0 \neq 0$ then $\overline{D(A^*)} \neq H$. Thus A^* exists but is not densely defined.

Exercise 11. Assume that A is closable. Prove that $D(\overline{A})$ can be obtained as the closure of $D(A)$ by the norm

$$\left(\|Au\|^2 + \|u\|^2\right)^{1/2}.$$

Definition. Let $A : H \to H$ with $\overline{D(A)} = H$. We say that A is

1) *symmetric* if $A \subset A^*$;

2) *self-adjoint* if $A = A^*$;

3) *essentially self-adjoint* if $\left(\overline{A}\right)^* = \overline{A}$.

Remark. A symmetric operator is always closable and its closure is also symmetric. Indeed, if $A \subset A^*$ then $D(A) \subset D(A^*)$. Hence

$$H = \overline{D(A)} \subset \overline{D(A^*)} \subset H$$

implies that $\overline{D(A^*)} = H$. That's why A is closable. Since $\overline{A}$ is the smallest closed extension of A then

$$A \subset \overline{A} \subset A^* = \left(\overline{A}\right)^*$$

i.e. $\overline{A}$ is also symmetric.

Some properties of symmetric operator A are:

1) $A \subset \overline{A} = A^{**} \subset A^*$
2) $A = \overline{A} = A^{**} \subset A^*$ if A is closed.
3) $A = \overline{A} = A^{**} = A^*$ if A is self-adjoint.
4) $A \subset \overline{A} = A^{**} = A^*$ if A is essentially self-adjoint.

Example 2.7. Consider the operator

$$A := \frac{d^2}{dx^2}$$

in the Hilbert space $H = L^2(0,1)$ with the domain

$$D(A) = \{f \in C^2[0,1] : f(0) = f(1) = f'(0) = f'(1) = 0\}.$$

It is clear that $\overline{D(A)} = L^2(0,1)$ and A is not closed. Moreover, integration by parts gives

$$(Af, g)_{L^2} = (f, Ag)_{L^2}$$

for any $f \in D(A)$ and $g \in W_2^2(0,1)$. That is, A is symmetric such that $A \subset A^*$ and $D(A^*) = W_2^2(0,1)$. As we know, $A^* = \overline{A^*}$ always. Now we will show that $\overline{A}$ is the same differential operator of order 2 with $D(\overline{A}) = \overset{\circ}{W}{}_2^2(0,1)$, where $\overset{\circ}{W}{}_2^2(0,1)$ denotes the closure of $D(A)$ with respect to the norm of Sobolev space $W_2^2(0,1)$. Indeed, for any $f \in D(A)$ we have

$$\|Af\|_{L^2}^2 + \|f\|_{L^2}^2 \le \|f\|_{W_2^2}^2$$

and

$$\begin{aligned} \|f\|_{W_2^2}^2 &= \|Af\|_{L^2}^2 + \|f\|_{L^2}^2 + \int_0^1 |f'|^2 dx \\ &= \|Af\|_{L^2}^2 + \|f\|_{L^2}^2 - \int_0^1 f\overline{f''} dx \le \frac{3}{2}\|Af\|_{L^2}^2 + \frac{3}{2}\|f\|_{L^2}^2 . \end{aligned}$$

It means that

$$\|Af\|_{L^2}^2 + \|f\|_{L^2}^2 \asymp \|f\|_{W_2^2}^2 .$$

Exercise 11 gives now that

$$D(\overline{A}) = \overset{\circ}{W}{}_2^2(0,1).$$

So we have finally

$$D(A) \subsetneq D(\overline{A}) = \overset{\circ}{W}{}_2^2(0,1) = D(A^{**}) \subsetneq W_2^2(0,1).$$

The closure $\overline{A}$ is symmetric, but not self-adjoint since

$$\overset{\circ}{W}{}_2^2(0,1) = D(\overline{A}) \neq D(\overline{A}^*) = D(A^*) = W_2^2(0,1).$$

Theorem 2 (J. von Neumann)**.** *Assume that* $A \subset A^*$*.*

1) If $D(A) = H$ *then* $A = A^*$ *and bounded.*

2) If $R(A) = H$ *then* $A = A^*$ *and* A^{-1} *exists and is bounded.*

3) If A^{-1} *exists then* $A = A^*$ *if and only if* $A^{-1} = (A^{-1})^*$*.*

Proof. 1) Since $A \subset A^*$ then $H = D(A) \subset D(A^*) \subset H$ and hence $D(A) = D(A^*) = H$. Thus $A = A^*$ and the Hellinger-Toeplitz theorem (Exercise 9) says that A is bounded.

2,3) Let us assume that $u_0 \in D(A)$ and $Au_0 = 0$. Then for any $v \in D(A)$ we obtain that

$$0 = (Au_0, v) = (u_0, Av).$$

It means that $u_0 \perp H$ and therefore $u_0 = 0$. It follows that A^{-1} exists and $D(A^{-1}) = R(A) = H$. Hence $(A^{-1})^*$ exists. Let us prove that $(A^*)^{-1}$ exists too and $(A^*)^{-1} = (A^{-1})^*$. Indeed, if $u \in D(A)$ and $v \in D\left((A^{-1})^*\right)$ then

$$(u, v) = (A^{-1}Au, v) = (Au, \left(A^{-1}\right)^* v).$$

This equality implies that

$$\left(A^{-1}\right)^* v \in D(A^*)$$

and

$$A^* \left(A^{-1}\right)^* v = v. \tag{2.1}$$

Similarly, if $u \in D(A^{-1})$ and $v \in D(A^*)$ then

$$(u, v) = (AA^{-1}u, v) = (A^{-1}u, A^*v)$$

and therefore

$$A^*v \in D\left(\left(A^{-1}\right)^*\right)$$

and

$$\left(A^{-1}\right)^* A^*v = v. \tag{2.2}$$

It follows from (2.1) and (2.2) that $(A^*)^{-1}$ exists and $(A^*)^{-1} = (A^{-1})^*$.

Exercise 12. Let A and B be injective operators. Prove that if $A \subset B$ then $A^{-1} \subset B^{-1}$.

Since $A \subset A^*$ we have by Exercise 12 that

$$A^{-1} \subset (A^*)^{-1} = \left(A^{-1}\right)^*$$

i.e. A^{-1} is also symmetric. But $D(A^{-1}) = H$. That's why we may conclude that $H = D(A^{-1}) \subset D\left((A^{-1})^*\right) \subset H$ and hence $D(A^{-1}) = D\left((A^{-1})^*\right) = H$. Thus A^{-1} is self-adjoint and bounded (Hellinger-Toeplitz theorem). Finally,

$$A^{-1} = \left(A^{-1}\right)^* = (A^*)^{-1}$$

if and only if $A = A^*$.

□

Theorem 3 (Basic criterion of self-adjointness)**.** *If $A \subset A^*$ then the following statements are equivalent:*

1) $A = A^*$.

2) $A = \overline{A}$ and $N(A^* \pm iI) = \{0\}$.

3) $R(A \pm iI) = H$.

Proof. **1)⇒2)** Since $A = A^*$ then A is closed. Suppose that $u_0 \in N(A^* - iI)$ i.e. $u_0 \in D(A^*) = D(A)$ and $Au_0 = iu_0$. Then

$$i(u_0, u_0) = (iu_0, u_0) = (Au_0, u_0) = (u_0, Au_0) = (u_0, iu_0) = -i(u_0, u_0).$$

This implies that $u_0 = 0$ i.e. $N(A^* - iI) = \{0\}$. The proof of $N(A^* + iI) = \{0\}$ is left to the reader.

2)⇒3) Since $A = \overline{A}$ and $N(A^* \pm iI) = \{0\}$ then, for example, the equation $A^*u = -iu$ has only the trivial solution $u = 0$. It implies that $\overline{R(A - iI)} = H$. For otherwise there exists $u_0 \neq 0$ such that $u_0 \perp \overline{R(A - iI)}$. It means that for any $u \in D(A)$ we have

$$((A - iI)u, u_0) = 0$$

and therefore $u_0 \in D(A^* + iI)$ and $(A^* + iI)u_0 = 0$ or $A^*u_0 = -iu_0, u_0 \neq 0$. This contradiction proves that $\overline{R(A - iI)} = H$. Next, since A is closed then $\Gamma(A)$ is also closed and due to the fact that A is symmetric we have

$$\begin{aligned} \|(A - iI)u\|^2 &= ((A - iI)u, (A - iI)u) = \|Au\|^2 - i(u, Au) + i(Au, u) + \|u\|^2 \\ &= \|Au\|^2 + \|u\|^2, \quad u \in D(A). \end{aligned}$$

That's why if $(A - iI)u_n \to v_0$ then Au_n and u_n are convergent i.e. $Au_n \to v_0', u_n \to u_0'$ and $u_n \in D(A)$. The closedness of A implies that $u_0' \in D(A)$ and $v_0' = Au_0'$ i.e. $(A - iI)u_n \to Au_0' - iu_0' = v_0$. It means that $R(A - iI)$ is a closed set i.e. $\overline{R(A - iI)} = R(A - iI) = H$. The proof of $R(A + iI) = H$ is left to the reader.

3)⇒1) Assume that $R(A \pm iI) = H$. Since $A \subset A^*$ it suffices to show that $D(A^*) \subset D(A)$. For every $u \in D(A^*)$ we have $(A^* - iI)u \in H$. Part 3) implies that there exists $v_0 \in D(A)$ such that

$$(A - iI)v_0 = (A^* - iI)u.$$

It is clear that $u - v_0 \in D(A^*)$ (since $A \subset A^*$) and

$$\begin{aligned}(A^* - iI)(u - v_0) &= (A^* - iI)u - (A^* - iI)v_0 = (A^* - iI)u - (A - iI)v_0 \\ &= (A - iI)v_0 - (A - iI)v_0 = 0.\end{aligned}$$

Hence $u - v_0 \in N(A^* - iI)$.

Exercise 13. Let A be a linear and densely defined operator in the Hilbert space H. Prove that

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

By this exercise we know that

$$H = N(A^* - iI) \oplus \overline{R(A + iI)}.$$

But in our case $R(A + iI) = H$. Hence $N(A^* - iI) = \{0\}$ and therefore $u = v_0$. Thus $D(A) = D(A^*)$.

□

Example 2.8. Assume that some operator $A = \overline{A}$ is closed and symmetric in the Hilbert space H. Consider the operator A^*A on the domain

$$D(A^*A) = \{f \in D(A) : Af \in D(A^*)\}.$$

This operator is self-adjoint. Indeed,

$$(A^*A)^* = A^*A^{**} = A^*\overline{A} = A^*A.$$

So A^*A is symmetric. At the same time for any $f \in D(A)$ it holds that

$$(A^*Af, f) = (Af, A^{**}f) = (Af, \overline{A}f) = \|Af\|_H^2 .$$

It means that A^*A is positive. This fact leads to $R(A^*A \pm iI) = H$, since $A^*A \pm iI$ is invertible in this case. Thus, Theorem 3 gives us that A^*A is self-adjoint. The same is true for the operator AA^* on the domain

$$D(AA^*) = \{f \in D(A^*) : A^*f \in D(A)\}.$$

It is clear that in general

$$AA^* \neq A^*A.$$

In case the equality holds here the operator A is called *normal*.

Exercise 14. Let $H = L^2(0,1)$ and $A := i\dfrac{d}{dx}$.

1) Prove that A is closed and symmetric on the domain

$$D(A) = \{f \in L^2(0,1) : f' \in L^2(0,1), f(0) = f(1) = 0\} \equiv \overset{\circ}{W}{}_2^1(0,1).$$

2) Prove that A is self-adjoint on the domain

$$D_\gamma(A) = \left\{f \in L^2(0,1) : f' \in L^2(0,1), f(0) = f(1)e^{i\gamma}, \gamma \in \mathbb{R}\right\}.$$

3 J. von Neumann's spectral theorem

Definition. A bounded linear operator P on a Hilbert space H which is self-adjoint and *idempotent* i.e. $P^2 = P$ is called an orthogonal projection operator or a *projector*.

Proposition 1. *Let P be a projector. Then*

1) $\|P\| = 1$ *if* $P \neq 0$.

2) P *is a projector if and only if* $P^{\perp} := I - P$ *is a projector.*

3) $H = R(P) \oplus R(P^{\perp})$, $P|_{R(P)} = I$ *and* $P|_{R(P^{\perp})} = 0$.

4) *There is one-to-one correspondence between projectors on H and closed linear subspaces of H. More precisely, if $M \subset H$ is a closed linear subspace then there exists a projector $P_M : H \to M$ and, conversely, if $P : H \to H$ is a projector then $R(P)$ is a closed linear subspace.*

5) *If* $\{e_j\}_{j=1}^N, N \leq \infty$ *is an orthonormal system then*

$$P_N x := \sum_{j=1}^{N} (x, e_j) e_j, \quad x \in H$$

is a projector.

Proof. 1) Since $P = P^*$ and $P = P^2$ then $P = P^*P$. Hence $\|P\| = \|P^*P\|$. But $\|P^*P\| = \|P\|^2$. Indeed,

$$\|P^*P\| \leq \|P^*\| \, \|P\| \leq \|P\|^2$$

and

$$\begin{aligned} \|P\|^2 &= \sup_{\|x\|=1} \|Px\|^2 = \sup_{\|x\|=1} (Px, Px) = \sup_{\|x\|=1} (P^*Px, x) \leq \sup_{\|x\|=1} \|P^*Px\| \\ &= \|P^*P\| . \end{aligned}$$

Therefore $\|P\| = \|P\|^2$ or $\|P\| = 1$ if $P \neq 0$.

2) Since P is linear and bounded then the same is true about $I - P$. Moreover,

$$(I - P)^* = I - P^* = I - P$$

and

$$(I - P)^2 = (I - P)(I - P) = I - 2P + P^2 = I - P.$$

3) It follows immediately from $I = P + P^{\perp}$ that every $x \in H$ is of the form $u + v$, where $u \in R(P)$ and $v \in R(P^{\perp})$. Let us prove that $R(P) = \left(R(P^{\perp})\right)^{\perp}$. First assume that $w \in \left(R(P^{\perp})\right)^{\perp}$ i.e. $(w, (I - P)x) = 0$ for all $x \in H$. This is equivalent to

$$(w, x) = (w, Px) = (Pw, x), \quad x \in H$$

or $Pw = w$. Hence $w \in R(P)$ and so we have proved that $\left(R(P^{\perp})\right)^{\perp} \subset R(P)$. For the opposite embedding we let $w \in R(P)$. Then there exists $x_w \in H$ such that $w = Px_w$. If $z \in R(P^{\perp})$ then $z = P^{\perp}x_z = (I - P)x_z$ for some $x_z \in H$. Thus

$$(w, z) = (Px_w, (I - P)x_z) = (Px_w, x_z) - (Px_w, Px_z) = 0$$

since P is a projector. Therefore $w \in \left(R(P^{\perp})\right)^{\perp}$ and we may conclude that $R(P) = \left(R(P^{\perp})\right)^{\perp}$. This fact allows us to conclude that $R(P) = \overline{R(P)}$ and $H = R(P) \oplus R(P^{\perp})$. Moreover, it is easy to check by definition that $P|_{R(P)} = I$ and $P|_{R(P^{\perp})} = 0$.

4) If $M \subset H$ is a closed subspace then Theorem 1 in Section 1 implies that $x = u + v \in H$, where $u \in M$ and $v \in M^{\perp}$. In that case let us define $P_M : H \to M$ as

$$P_M x = u.$$

It is clear that $P_M^2 x = P_M u = u = P_M x$ i.e. $P_M^2 = P_M$. Moreover, if $y \in H$ then $y = u_1 + v_1, u_1 \in M, v_1 \in M^{\perp}$ and

$$(P_M x, y) = (u, u_1 + v_1) = (u, u_1) = (u + v, u_1) = (u + v, P_M y) = (x, P_M y)$$

i.e. $P_M^* = P_M$. Hence P_M is a projector. If P is a projector then we know from part 3) that $M := R(P)$ is closed subspace of H.

5) Let us assume that $N = \infty$. Define M as

$$M := \left\{ x \in H : x = \sum_{j=1}^{\infty} c_j e_j, \sum_{j=1}^{\infty} |c_j|^2 < \infty \right\}.$$

Then M is a closed subspace of H. If we define a linear operator P_M as

$$P_M x := \sum_{j=1}^{\infty} (x, e_j) e_j, \quad x \in H$$

then by Bessel's inequality we obtain that $P_M x \in M$ and

$$\|P_M x\| \leq \|x\|.$$

It means that P_M is a bounded linear operator into M. But $P_M e_j = e_j$ and thus $P_M^2 x = P_M x$ for all $x \in H$. Next, for all $x, y \in H$ we have

$$\begin{aligned}(P_M x, y) &= \left(\sum_{j=1}^{\infty}(x, e_j)e_j, y\right) = \sum_{j=1}^{\infty}(x, e_j)\,(e_j, y) = \sum_{j=1}^{\infty}(x, (y, e_j)e_j) \\ &= \left(x, \sum_{j=1}^{\infty}(y, e_j)e_j\right) = (x, P_M y)\end{aligned}$$

i.e. $P_M^* = P_M$. The case of finite N requires no convergence questions and is left to the reader.

□

Definition. A bounded linear operator A on the Hilbert space H is smaller than or equal to a bounded operator B on H if

$$(Ax, x) \le (Bx, x), \quad x \in H.$$

We denote this fact by $A \le B$. We say that A is *non-negative* if $A \ge 0$. We denote $A > 0$, and say that A is *positive*, if $A \ge c_0 I$ for some $c_0 > 0$.

Remark. In the frame of this definition (Ax, x) and (Bx, x) must be real for all $x \in H$.

Proposition 2. *For two projectors P and Q the following statements are equivalent:*

1) $P \le Q$.

2) $\|Px\| \le \|Qx\|$ *for all* $x \in H$.

3) $R(P) \subset R(Q)$.

4) $P = PQ = QP$.

Proof. **1)⇔2)** Follows immediately from $(Px, x) = (P^2 x, x) = (Px, Px) = \|Px\|^2$.

3)⇔4) Assume $R(P) \subset R(Q)$. Then $QPx = Px$ or $QP = P$. Conversely, if $QP = P$ then clearly $R(P) \subset R(Q)$. Finally, $P = QP = P^* = (QP)^* = P^* Q^* = PQ$.

2)⇔4) If 4) holds then $Px = PQx$ and $\|Px\| = \|PQx\| \le \|Qx\|$ for all $x \in H$. Conversely, if $\|Px\| \le \|Qx\|$ then $Px = QPx + Q^{\perp}Px$ implies that

$$\|Px\|^2 = \|QPx\|^2 + \left\|Q^{\perp}Px\right\|^2 \le \|QPx\|^2.$$

Hence

$$\left\|Q^{\perp}Px\right\|^2 = 0$$

i.e. $Q^{\perp}Px = 0$ for all $x \in H$. Hence $P = QP = PQ$.

□

Exercise 15. Let $\{P_j\}_{j=1}^{\infty}$ be a sequence of projectors with $P_j \leq P_{j+1}$ for each $j = 1, 2, \ldots$. Prove that $\lim_{j\to\infty} P_j := P$ exists and that P is a projector.

Definition. Any linear map $A : H \to H$ with the property

$$\|Ax\| = \|x\|, \quad x \in H$$

is called an *isometry*.

Exercise 16. Prove that

1) A is an isometry if and only if $A^*A = I$.

2) Every isometry A has an inverse $A^{-1} : R(A) \to H$ and $A^{-1} = A^*|_{R(A)}$.

3) If A is an isometry then AA^* is a projector on $R(A)$.

Definition. A surjective isometry $U : H \to H$ is called a *unitary operator*.

Remark. It follows that U is unitary if and only if it is surjective and $U^*U = UU^* = I$ i.e. $(Ux, Uy) = (x, y)$ for all $x, y \in H$.

Definition. Let H be a Hilbert space. The family of operators $\{E_\lambda\}_{\lambda=-\infty}^{\infty}$ is called a *spectral family* if the following conditions are satisfied:

1) E_λ is a projector for all $\lambda \in \mathbb{R}$.

2) $E_\lambda \leq E_\mu$ for all $\lambda < \mu$.

3) $\{E_\lambda\}$ is right continuous with respect to the strong operator topology i.e.

$$\lim_{s\to t+0} \|E_s x - E_t x\| = 0$$

for all $x \in H$.

4) $\{E_\lambda\}$ is normalized as follows:

$$\lim_{\lambda\to-\infty} \|E_\lambda x\| = 0, \quad \lim_{\lambda\to+\infty} \|E_\lambda x\| = \|x\|$$

for all $x \in H$. The latter condition can also be formulated as

$$\lim_{\lambda\to+\infty} \|E_\lambda x - x\| = 0.$$

Remark. It follows from the previous definition and Proposition 2 that

$$E_\lambda E_\mu = E_{\min\{\lambda,\mu\}}.$$

Proposition 3. *For every fixed $x, y \in H$, $(E_\lambda x, y)$ is a function of bounded variation with respect to $\lambda \in \mathbb{R}$.*

Proof. Let us define

$$E(\alpha,\beta] := E_\beta - E_\alpha, \quad \alpha < \beta.$$

Then $E(\alpha,\beta]$ is a projector. Indeed,

$$E(\alpha,\beta]^* = E_\beta^* - E_\alpha^* = E_\beta - E_\alpha = E(\alpha,\beta]$$

i.e. $E(\alpha,\beta]$ is self-adjoint. It is also idempotent due to

$$\begin{aligned} (E(\alpha,\beta])^2 &= (E_\beta - E_\alpha)(E_\beta - E_\alpha) = E_\beta^2 - E_\alpha E_\beta - E_\beta E_\alpha + E_\alpha^2 \\ &= E_\beta - E_\alpha - E_\alpha + E_\alpha = E(\alpha,\beta]. \end{aligned}$$

Another property is that

$$E(\alpha_1,\beta_1]x \perp E(\alpha,\beta]y, \quad x, y \in H$$

if $\beta_1 \le \alpha$ or $\beta \le \alpha_1$. To see this for $\beta_1 \le \alpha$ calculate

$$\begin{aligned} (E(\alpha_1,\beta_1]x, E(\alpha,\beta]y) &= (E_{\beta_1}x - E_{\alpha_1}x, E_\beta y - E_\alpha y) \\ &= (E_{\beta_1}x, E_\beta y) - (E_{\alpha_1}x, E_\beta y) - (E_{\beta_1}x, E_\alpha y) + (E_{\alpha_1}x, E_\alpha y) \\ &= (x, E_{\beta_1}y) - (x, E_{\alpha_1}y) - (x, E_{\beta_1}y) + (x, E_{\alpha_1}y) = 0. \end{aligned}$$

Let now

$$\lambda_0 < \lambda_1 < \cdots < \lambda_n.$$

Then

$$\begin{aligned} \sum_{j=1}^n \left|\big(E_{\lambda_j}x, y\big) - \big(E_{\lambda_{j-1}}x, y\big)\right| &= \sum_{j=1}^n |(E(\lambda_{j-1},\lambda_j]x, y)| \\ &= \sum_{j=1}^n |(E(\lambda_{j-1},\lambda_j]x, E(\lambda_{j-1},\lambda_j]y)| \\ &\le \sum_{j=1}^n \|E(\lambda_{j-1},\lambda_j]x\| \, \|E(\lambda_{j-1},\lambda_j]y\| \\ &\le \left(\sum_{j=1}^n \|E(\lambda_{j-1},\lambda_j]x\|^2\right)^{1/2} \left(\sum_{j=1}^n \|E(\lambda_{j-1},\lambda_j]y\|^2\right)^{1/2} \\ &= \left\|\sum_{j=1}^n E(\lambda_{j-1},\lambda_j]x\right\| \left\|\sum_{j=1}^n E(\lambda_{j-1},\lambda_j]y\right\| \\ &= \|E(\lambda_0,\lambda_n]x\| \, \|E(\lambda_0,\lambda_n]y\| \le \|x\| \, \|y\| \, . \end{aligned}$$

Here we have made use of orthogonality, normalization and the Cauchy-Schwarz-Bunjakovskii inequality. □

Due to Proposition 3 we can define a Stieltjes integral. Moreover, for any continuous function $f(\lambda)$ we may conclude that the limit

$$\lim_{\Delta\to 0}\sum_{j=1}^{n} f(\lambda_j^*)\left(E(\lambda_{j-1},\lambda_j]x,y\right) = \lim_{\Delta\to 0}\left(\sum_{j=1}^{n} f(\lambda_j^*)E(\lambda_{j-1},\lambda_j]x,y\right),$$

where $\lambda_j^* \in [\lambda_{j-1},\lambda_j]$, $\alpha=\lambda_0<\lambda_1<\cdots<\lambda_n=\beta$ and $\Delta=\max_{1\le j\le n}|\lambda_{j-1}-\lambda_j|$ exists and by definition this limit is

$$\int_\alpha^\beta f(\lambda)d(E_\lambda x,y), \quad x,y\in H.$$

It can be shown that this is equivalent to the existence of the limit in H

$$\lim_{\Delta\to 0}\sum_{j=1}^{n} f(\lambda_j^*)E(\lambda_{j-1},\lambda_j]x,$$

which we denote by

$$\int_\alpha^\beta f(\lambda)dE_\lambda x.$$

Thus

$$\int_\alpha^\beta f(\lambda)d(E_\lambda x,y) = \left(\int_\alpha^\beta f(\lambda)dE_\lambda x,y\right), \quad x,y\in H.$$

For the spectral representation of self-adjoint operators one needs not only integrals over finite intervals but also over whole line which is naturally defined as the limit

$$\int_{-\infty}^{\infty} f(\lambda)d(E_\lambda x,y) = \lim_{\substack{\alpha\to-\infty\\ \beta\to\infty}}\int_\alpha^\beta f(\lambda)d(E_\lambda x,y) = \left(\int_{-\infty}^{\infty} f(\lambda)dE_\lambda x,y\right)$$

if it exists. Deriving first some basic properties of the integral just defined one can check that

$$\int_{-\infty}^{\infty} f(\lambda)d(E_\lambda E_\beta x,y) = \int_{-\infty}^{\beta} f(\lambda)d(E_\lambda x,y) := \lim_{\alpha\to-\infty}\int_\alpha^\beta f(\lambda)d(E_\lambda x,y), \quad x,y\in H.$$

Theorem 1. *Let $\{E_\lambda\}_{\lambda=-\infty}^{\infty}$ be a spectral family on the Hilbert space H and let f be a real-valued continuous function on the line. Define*

$$D := \left\{x\in H : \int_{-\infty}^{\infty}|f(\lambda)|^2 d(E_\lambda x,x)<\infty\right\}$$

(or $D := \left\{x\in H : \int_{-\infty}^{\infty} f(\lambda)dE_\lambda x \text{ exists}\right\}$). Let us define on this domain an operator A as

$$(Ax,y) = \int_{-\infty}^{\infty} f(\lambda)d(E_\lambda x,y), \quad x\in D(A):=D, y\in H$$

(or $Ax=\int_{-\infty}^{\infty} f(\lambda)dE_\lambda x, x\in D(A)$). Then A is self-adjoint and satisfies

$$E(\alpha,\beta]A \subset AE(\alpha,\beta], \quad \alpha<\beta.$$

Proof. It can be shown that the integral

$$\int_{-\infty}^{\infty} f(\lambda) d\,(E_\lambda x, y)$$

exists for $x \in D$ and $y \in H$. Thus (Ax, y) is well-defined. Let v be any element of H and let $\varepsilon > 0$. Then, by normalization, there exists $\alpha < -R$ and $\beta > R$ with R large enough such that

$$\|v - E(\alpha, \beta]v\| = \|v - E_\beta v + E_\alpha v\| \le \|(I - E_\beta)v\| + \|E_\alpha v\| < \varepsilon.$$

On the other hand,

$$\begin{aligned}
\int_{-\infty}^{\infty} |f(\lambda)|^2 d(E_\lambda E(\alpha, \beta]v, E(\alpha, \beta]v) &= \int_{-\infty}^{\infty} |f(\lambda)|^2 d(E_\lambda E(\alpha, \beta]v, v) \\
&= \int_{-\infty}^{\infty} |f(\lambda)|^2 d(E_\lambda E_\beta v, v) - \int_{-\infty}^{\infty} |f(\lambda)|^2 d(E_\lambda E_\alpha v, v) \\
&= \int_{-\infty}^{\beta} |f(\lambda)|^2 d(E_\lambda v, v) - \int_{-\infty}^{\alpha} |f(\lambda)|^2 d(E_\lambda v, v) \\
&= \int_{\alpha}^{\beta} |f(\lambda)|^2 d(E_\lambda v, v) < \infty.
\end{aligned}$$

These two facts mean that $E(\alpha, \beta]v \in D$ and $\overline{D} = H$. Since $f(\lambda) = \overline{f(\lambda)}$ then A is symmetric. Indeed,

$$\begin{aligned}
(Ax, y) &= \int_{-\infty}^{\infty} f(\lambda) d(E_\lambda x, y) = \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} \int_{\alpha}^{\beta} f(\lambda) d(E_\lambda x, y) \\
&= \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} \int_{\alpha}^{\beta} f(\lambda) d(x, E_\lambda y) = \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} \left(x, \int_{\alpha}^{\beta} f(\lambda) dE_\lambda y \right) \\
&= \left(x, \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} \int_{\alpha}^{\beta} f(\lambda) dE_\lambda y \right) = (x, Ay).
\end{aligned}$$

In order to prove that $A = A^*$ it remains to show that $D(A^*) \subset D(A)$. Let $u \in D(A^*)$. Then

$$(E(\alpha, \beta]z, A^* u) = (AE(\alpha, \beta]z, u) = \int_{\alpha}^{\beta} f(\lambda) d(E_\lambda z, u)$$

for any $z \in H$. This equality implies that

$$\begin{aligned}
(z, A^* u) &= \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} \int_{\alpha}^{\beta} f(\lambda) d(E_\lambda z, u) = \int_{-\infty}^{\infty} f(\lambda) d(E_\lambda z, u) = \int_{-\infty}^{\infty} f(\lambda) d(z, E_\lambda u) \\
&= \overline{\int_{-\infty}^{\infty} f(\lambda) d(E_\lambda u, z)} = \overline{(Au, z)} = (z, Au),
\end{aligned}$$

where the integral exists because (z, A^*u) exists. Hence $u \in D(A)$ and $A^*u = Au$. For the second claim we first calculate

$$\begin{aligned}
E(\alpha,\beta]Ax = (E_\beta - E_\alpha)\,Ax = (E_\beta - E_\alpha)\int_{-\infty}^{\infty} f(\lambda)dE_\lambda x \\
= \int_{-\infty}^{\infty} f(\lambda)dE_\lambda E_\beta x - \int_{-\infty}^{\infty} f(\lambda)dE_\lambda E_\alpha x = \int_{-\infty}^{\beta} f(\lambda)dE_\lambda x - \int_{-\infty}^{\alpha} f(\lambda)dE_\lambda x \\
= \int_{\alpha}^{\beta} f(\lambda)dE_\lambda x = \int_{-\infty}^{\infty} f(\lambda)dE_\lambda\,(E_\beta - E_\alpha)\,x = A\,(E_\beta - E_\alpha)\,x = AE(\alpha,\beta]x
\end{aligned}$$

for any $x \in D(A)$. Since the left hand side is defined on $D(A)$ and the right hand side on all of H then the latter is an extension of the former. □

Exercise 17. Let A be as in Theorem 1. Prove that

$$\|Au\|^2 = \int_{-\infty}^{\infty} |f(\lambda)|^2 d(E_\lambda u, u)$$

if $u \in D(A)$.

Exercise 18. Let $H = L^2(\mathbb{R})$ and $Au(t) = tu(t), t \in \mathbb{R}$. Define $D(A)$ on which $A = A^*$ and evaluate the spectral family $\{E_\lambda\}_{\lambda=-\infty}^{\infty}$.

Theorem 2 (J. von Neumann's spectral theorem)**.** *Every self-adjoint operator A on the Hilbert space H has a unique spectral representation i.e. there is a unique spectral family $\{E_\lambda\}_{\lambda=-\infty}^{\infty}$ such that*

$$Ax = \int_{-\infty}^{\infty} \lambda dE_\lambda x, \quad x \in D(A)$$

(i.e. $(Ax, y) = \int_{-\infty}^{\infty} \lambda d(E_\lambda x, y), x \in D(A), y \in H$), where $D(A)$ is defined as

$$D(A) = \left\{ x \in H : \int_{-\infty}^{\infty} \lambda^2 d(E_\lambda x, x) < \infty \right\}.$$

Proof. At first we assume that this theorem holds when A is bounded, that is, there is a unique spectral family $\{F_\mu\}_{\mu=-\infty}^{\infty}$ such that

$$Au = \int_{-\infty}^{\infty} \mu dF_\mu u, \quad u \in H$$

since $D(A) = H$ in this case. But $F_\mu \equiv 0$ for $\mu < m$ and $F_\mu \equiv I$ for $\mu > M$, where

$$m = \inf_{\|x\|=1} (Ax, x), \quad M = \sup_{\|x\|=1} (Ax, x).$$

That's why the spectral representation has a view

$$Au = \int_{m}^{M} \mu dF_\mu u, \quad u \in H.$$

Let us consider now an unbounded operator which is semibounded from below i.e.

$$(Au, u) \geq m_0(u, u), \quad u \in D(A)$$

with some constant m_0. We assume without loss of generality that $(Au, u) \geq (u, u)$. This condition implies that A^{-1} exists, is defined over whole H and $\|A^{-1}\| \leq 1$. Indeed, A^{-1} exists and is bounded because $Au = 0$ if and only if $u = 0$. The norm estimate follows from

$$(v, A^{-1}v) \geq \left\|A^{-1}v\right\|^2, \quad v \in D(A^{-1}).$$

Since A^{-1} is bounded then $D(A^{-1})$ is a closed subspace in H. But self-adjointness of A means that $A^{-1} = (A^{-1})^*$. That's why A^{-1} is closed and $\overline{D(A^{-1})} = H$ i.e. A^{-1} is densely defined. Therefore $D(A^{-1}) = H$ and $R(A) = H$. Since

$$0 \leq (A^{-1}v, v) \leq \|v\|^2, \quad v \in H$$

we may conclude in this case that $m \geq 0, M \leq 1$ and

$$A^{-1}v = \int_0^1 \mu dF_\mu v, \quad v \in H,$$

where $\{F_\mu\}$ is the spectral family of A^{-1}. Let us note that $F_1 = I$ and $F_0 = 0$. They follow from the spectral theorem and from the fact that $A^{-1}v = 0$ if and only if $v = 0$. Next, let us define the operator $B_\varepsilon, \varepsilon > 0$ as

$$B_\varepsilon u := \int_\varepsilon^1 \frac{1}{\mu} dF_\mu u, \quad u \in D(A).$$

For every $v \in H$ we have

$$\begin{aligned} B_\varepsilon A^{-1}v &= \int_\varepsilon^1 \frac{1}{\mu} dF_\mu (A^{-1}v) = \int_\varepsilon^1 \frac{1}{\mu} dF_\mu \left(\int_0^1 \lambda dF_\lambda v\right) = \int_\varepsilon^1 \frac{1}{\mu} d\left(\int_0^1 \lambda d(F_\mu F_\lambda v)\right) \\ &= \int_\varepsilon^1 \frac{1}{\mu} d\left(\int_\varepsilon^\mu \lambda dF_\lambda v\right) = \int_\varepsilon^1 \frac{1}{\mu}\mu dF_\mu v = \int_\varepsilon^1 dF_\mu v = F_1 v - F_\varepsilon v = v - F_\varepsilon v. \end{aligned}$$

Since every spectral family is right continuous then

$$\lim_{\varepsilon \to 0+0} B_\varepsilon A^{-1} v = v$$

exists. For every $u \in D(A)$ we have similarly,

$$A^{-1}B_\varepsilon u = \int_0^1 \mu dF_\mu (B_\varepsilon u) = \int_\varepsilon^1 \mu d\left(\int_\varepsilon^\mu \frac{1}{\lambda} dF_\lambda u\right) = u - F_\varepsilon u$$

and hence

$$\lim_{\varepsilon \to 0+0} A^{-1}B_\varepsilon u = u$$

exists. These two equalities mean that

$$\lim_{\varepsilon\to 0+0} B_\varepsilon = \left(A^{-1}\right)^{-1} = A$$

exists and the spectral representation

$$A = \int_0^1 \frac{1}{\mu} dF_\mu = \lim_{\varepsilon\to 0+} \int_\varepsilon^1 \frac{1}{\mu} dF_\mu$$

holds. If we define $E_\lambda = I - F_{\frac{1}{\lambda}}, 1 \le \lambda < \infty$ then

$$A = -\int_0^1 \frac{1}{\mu} dE_{\frac{1}{\mu}} = \int_1^\infty \lambda dE_\lambda.$$

Exercise 19. Prove that this $\{E_\lambda\}$ is a spectral family.

Domain $D(A)$ can be characterized as

$$D(A) = \left\{ u \in H : \int_1^\infty \lambda^2 d(E_\lambda u, u) < \infty \right\} = \left\{ u \in H : \int_0^1 \frac{1}{\mu^2} d(F_\mu u, u) < \infty \right\}.$$

This proves the theorem for self-adjoint operators that are semibounded from below. For bounded operators we will only sketch the proof.

Step 1. If $A = A^*$ and bounded then we can define

$$p_N(A) := a_0 I + a_1 A + \cdots + a_N A^N, \quad N \in \mathbb{N},$$

where $a_j \in \mathbb{R}$ for $j = 0, 1, \ldots, N$. Then $p_N(A)$ is also self-adjoint and bounded with

$$\|p_N(A)\| \le \sup_{|t|\le\|A\|} |p_N(t)|.$$

Step 2. For every continuous real-valued function f on $[m, M]$, where m and M are as above we can define $f(A)$ as an approximation by $p_N(A)$ i.e. we can prove that for any $\varepsilon > 0$ there exists $p_N(A)$ such that

$$\|f(A) - p_N(A)\| < \varepsilon.$$

Step 3. For every $u, v \in H$ let us define the functional L as

$$L(f) := (f(A)u, v).$$

Then

$$|L(f)| \le |(f(A)u, v)| \le \|f(A)\| \, \|u\| \, \|v\|$$

that is, $L(f)$ is a bounded linear functional on $C[m, M]$.

Step 4. (Riesz's theorem) Every positive linear continuous functional L on $C_0[a,b]$ can be represented in the form

$$L(f) = \int_a^b f(x)d\nu(x),$$

where ν is a measure that satisfies the conditions

1) $L(f) \geq 0$ for $f \geq 0$
2) $|L(f)| \leq \nu(K)\,\|f\|_K$, where $K \subset [a,b]$ is compact and

$$\|f\|_K = \max_{x\in K} |f(x)|.$$

Step 5. It follows from Step 4 that

$$(Au, v) = \int_m^M \lambda d\nu(\lambda; u, v).$$

Step 6. It is possible to prove that $\nu(\lambda; u, v)$ is a self-adjoint bilinear form. That's why we may conclude that there exists a self-adjoint and bounded operator E_λ such that

$$\nu(\lambda; u, v) = (E_\lambda u, v).$$

This operator is idempotent and we may define $E_\lambda \equiv 0$ for $\lambda < m$ and $E_\lambda \equiv I$ for $\lambda \geq M$. Thus $\{E_\lambda\}_{\lambda=-\infty}^{\infty}$ is the required spectral family and this theorem is proved.

□

Let $A : H \to H$ be a self-adjoint operator in the Hilbert space H. Then by J. von Neumann's spectral theorem we can write

$$Au = \int_{-\infty}^{\infty} \lambda dE_\lambda u, \quad u \in D(A).$$

For every continuous function f we can define

$$D_f := \left\{ u \in H : \int_{-\infty}^{\infty} |f(\lambda)|^2 d(E_\lambda u, u) < \infty \right\}.$$

This set is a linear subspace of H. For every $u \in D_f$ and $v \in H$ let us define the linear functional

$$L(v) := \int_{-\infty}^{\infty} f(\lambda)d(E_\lambda u, v) = \left(\int_{-\infty}^{\infty} f(\lambda)dE_\lambda u, v \right).$$

This functional is continuous because it is bounded. Indeed,

$$|L(v)|^2 \leq \left\| \int_{-\infty}^{\infty} f(\lambda)dE_\lambda u \right\|^2 \|v\|^2 = \int_{-\infty}^{\infty} |f(\lambda)|^2 d(E_\lambda u, u)\,\|v\|^2 = c(u)\,\|v\|^2.$$

By the Riesz-Frechet theorem this functional can be expressed in the form of an inner product i.e. there exists $z \in H$ such that

$$\int_{-\infty}^{\infty} f(\lambda) d(E_\lambda u, v) = (z, v), \quad v \in H.$$

We set

$$z := f(A)u, \quad u \in D_f$$

i.e.

$$(f(A)u, v) = \int_{-\infty}^{\infty} f(\lambda) d(E_\lambda u, v).$$

Remark. Since in general f is not real-valued then $f(A)$ is not a self-adjoint operator in general.

Example 3.1. Consider

$$f(\lambda) = \frac{\lambda - i}{\lambda + i}, \quad \lambda \in \mathbb{R}$$

Denote

$$U_A := f(A) = \int_{-\infty}^{\infty} \frac{\lambda - i}{\lambda + i} dE_\lambda.$$

The operator U_A is called the *Cayley transform*. Since $|f(\lambda)| = 1$ then $D_f = D(U_A) = H$ and

$$\begin{aligned} \|U_A u\|^2 &= \int_{-\infty}^{\infty} |f(\lambda)|^2 d(E_\lambda u, u) = \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} \int_{\alpha}^{\beta} d(E_\lambda u, u) = \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} ((E_\beta u, u) - (E_\alpha u, u)) \\ &= \lim_{\substack{\alpha \to -\infty \\ \beta \to \infty}} \left(\|E_\beta u\|^2 - \|E_\alpha u\|^2 \right) = \|u\|^2 \end{aligned}$$

by normalization of $\{E_\lambda\}$. Hence U_A is an isometry. There is one-to-one correspondence between self-adjoint operators and their Cayley transforms. Indeed,

$$U_A = (A - iI)(A + iI)^{-1}$$

is equivalent to

$$\begin{cases} I - U_A = 2i(A + iI)^{-1} \\ I + U_A = 2A(A + iI)^{-1} \end{cases}$$

or

$$A = i(I + U_A)(I - U_A)^{-1}.$$

Example 3.2. Consider

$$f(\lambda) = \frac{1}{\lambda - z}, \quad \lambda \in \mathbb{R}, z \in \mathbb{C}, \operatorname{Im} z \neq 0.$$

Denote
$$R_z := (A - zI)^{-1} = \int_{-\infty}^{\infty} \frac{1}{\lambda - z} dE_\lambda.$$
The operator R_z is called the *resolvent* of A. Since
$$\left| \frac{1}{\lambda - z} \right| \le \frac{1}{|\operatorname{Im} z|}$$
for all $\lambda \in \mathbb{R}$ then R_z is bounded and defined on whole H.

Example 3.3. Suppose that $K(x, y) \in L^2(\Omega \times \Omega)$. Define the integral operator on $L^2(\Omega)$ as
$$Af(x) = \int_\Omega K(x, y) f(y) dy.$$
Then
$$A^* f(x) = \int_\Omega \overline{K(y, x)} f(y) dy$$
and therefore
$$A^* A f(x) = \int_\Omega \left(\int_\Omega K(y, z) \overline{K(y, x)} dy \right) f(z) dz.$$
As we know from Example 2.8, $A^* A$ is self-adjoint on $L^2(\Omega)$. This fact can also be checked directly, since
$$\int_\Omega K(y, z) \overline{K(y, x)} dy = \overline{\int_\Omega K(y, x) \overline{K(y, z)} dy}.$$
And J. von Neumann's spectral theorem gives us for this operator and for any $s \ge 0$ that
$$(A^* A)^s = \int_0^{\|A\|^2_{L^2 \to L^2}} \lambda^s dE_\lambda,$$
since $A^* A$ is positive and bounded by $\|A\|^2_{L^2 \to L^2}$.

Exercise 20. Let $A = A^*$ with spectral family E_λ. Let $u \in D(f(A))$ and $v \in D(g(A))$. Prove that
$$(f(A)u, g(A)v) = \int_{-\infty}^{\infty} f(\lambda) \overline{g(\lambda)} d(E_\lambda u, v).$$

Exercise 21. Let $A = A^*$ with spectral family E_λ. Let $u \in D(f(A))$. Prove that $f(A)u \in D(g(A))$ if and only if $u \in D((gf)(A))$ and that
$$(gf)(A)u = \int_{-\infty}^{\infty} g(\lambda) f(\lambda) dE_\lambda u.$$

Remark. It follows from Exercise 21 that
$$(gf)(A) = (fg)(A)$$
on the domain $D\left((fg)(A)\right) \cap D\left((gf)(A)\right)$.

4 Spectrum of self-adjoint operators

Definition. Given a linear operator A in the Hilbert space H with domain $D(A)$, $\overline{D(A)} = H$, the set

$$\rho(A) = \left\{ z \in \mathbb{C} : (A - zI)^{-1} \text{ exists as a bounded operator from } H \text{ to } D(A) \right\}$$

is called the *resolvent set* of A. Its complement

$$\sigma(A) = \mathbb{C} \setminus \rho(A)$$

is called the *spectrum* of A.

Theorem 1. *1) If $A = \overline{A}$ then the resolvent set is open and the resolvent operator $R_z := (A - zI)^{-1}$ is an analytic function from $\rho(A)$ to $B(H; H)$, the set of all linear operators in H. Furthermore, the* resolvent identity

$$R_z - R_\xi = (z - \xi) R_z R_\xi, \quad z, \xi \in \rho(A)$$

holds and $R'_z = (R_z)^2$.

2) If $A = A^*$ then $z \in \rho(A)$ if and only if there exists $C_z > 0$ such that

$$\|(A - zI)u\| \geq C_z \|u\|$$

for all $u \in D(A)$.

Proof. 1) Assume that $z_0 \in \rho(A)$. Then R_{z_0} is a bounded linear operator from H to $D(A)$ and thus $r := \|R_{z_0}\|^{-1} > 0$. Let us define for $|z - z_0| < r$ the operator

$$G_{z_0} := (z - z_0) R_{z_0}.$$

Then G_{z_0} is bounded with $\|G_{z_0}\| < 1$. Hence it defines the operator

$$(I - G_{z_0})^{-1} = \sum_{j=0}^{\infty} (G_{z_0})^j$$

because this Neumann series converges. But for $|z - z_0| < r$ we have

$$A - zI = (A - z_0 I)(I - G_{z_0})$$

or

$$(A - zI)^{-1} = (I - G_{z_0})^{-1} R_{z_0}.$$

Hence R_z exists with $D(R_z) = H$ and is bounded. It remains to show that $R(R_z) \subset D(A)$. For $x \in H$ we know that

$$y := (A - zI)^{-1} x \in H.$$

We claim that $y \in D(A)$. Indeed,

$$
\begin{aligned}
y &= (A - zI)^{-1}x = (I - G_{z_0})^{-1}R_{z_0}x = \sum_{j=0}^{\infty}(z - z_0)^j \left(R_{z_0}\right)^{j+1} x \\
&= \lim_{n\to\infty} \sum_{j=0}^{n}(z - z_0)^j \left(R_{z_0}\right)^{j+1} x.
\end{aligned}
$$

It follows from this representation that $R_z = (A - zI)^{-1}$ is an analytic function from $\rho(A)$ to $B(H; H)$. Next we denote

$$
s_n x := \sum_{j=0}^{n}(z - z_0)^j \left(R_{z_0}\right)^{j+1} x.
$$

It is clear that $s_n x \in D(A)$ and that $\lim_{n\to\infty} s_n x = y$. Moreover,

$$
\lim_{n\to\infty}(A - zI)s_n x = x.
$$

Denoting $y_n := s_n x$ we may conclude from the criterion for closedness that

$$
\begin{cases} y_n \in D(A) \\ y_n \to y \\ (A - zI)y_n \to x \end{cases} \Rightarrow \quad \begin{cases} y \in D(A) \\ x = (A - zI)y. \end{cases}
$$

Hence $y = (A - zI)^{-1}x \in D(A)$ and therefore $\rho(A)$ is open. The resolvent identity is proved by straightforward calculation

$$
\begin{aligned}
R_z - R_\xi &= R_z(A - \xi I)R_\xi - R_z(A - zI)R_\xi = R_z\left[(A - \xi I) - (A - zI)\right]R_\xi \\
&= (z - \xi)R_z R_\xi.
\end{aligned}
$$

Finally, the limit

$$
\lim_{z\to\xi}\frac{R_z - R_\xi}{z - \xi} = \lim_{z\to\xi} R_z R_\xi = (R_z)^2
$$

exists and hence $R_z' = (R_z)^2$ exists. It proves this part.

2) Assume that $A = A^*$. If $z \in \rho(A)$ then by definition R_z maps from H to $D(A)$. Hence there exists $M_z > 0$ such that

$$
\|R_z v\| \le M_z \|v\|, \quad v \in H.
$$

Since $u = R_z(A - zI)u$ for any $u \in D(A)$ then we get

$$
\|u\| \le M_z \|(A - zI)u\|, \quad u \in D(A).
$$

This is equivalent to

$$\|(A - zI)u\| \geq \frac{1}{M_z} \|u\|, \quad u \in D(A).$$

Conversely, if there exists $C_z > 0$ such that

$$\|(A - zI)u\| \geq C_z \|u\|, \quad u \in D(A).$$

then $(A - zI)^{-1}$ is bounded. Since A is self-adjoint then $(A - zI)^{-1}$ is defined over whole H. Indeed, if $R(A - zI) \neq H$ then there exists $v_0 \neq 0$ such that $v_0 \perp R(A - zI)$. This means that

$$(v_0, (A - zI)u) = 0, \quad u \in D(A)$$

or

$$(Au, v_0) = (zu, v_0)$$

or

$$(u, A^* v_0) = (u, \overline{z} v_0).$$

Thus $v_0 \in D(A^*)$ and $A^* v_0 = \overline{z} v_0$. Since $A = A^*$ then $v_0 \in D(A)$ and $A v_0 = \overline{z} v_0$ or

$$(A - \overline{z} I)v_0 = 0.$$

It is easy to check that $\|(A - \overline{z}I)u\|^2 = \|(A - zI)u\|^2$ for any $u \in D(A)$. Therefore

$$\|(A - \overline{z}I)v_0\| = \|(A - zI)v_0\| \geq C_z \|v_0\|.$$

Hence $v_0 = 0$ and $D\left((A - zI)^{-1}\right) = R(A - zI) = H$. It means that $z \in \rho(A)$.

□

Corollary 1. *If $A = A^*$ then $\sigma(A) \neq \emptyset, \sigma(A) = \overline{\sigma(A)}$ and $\sigma(A) \subset \mathbb{R}$.*

Proof. If $z = \alpha + i\beta \in \mathbb{C}$ with $\operatorname{Im} z = \beta \neq 0$ then

$$\|(A - zI)x\|^2 = \|(A - \alpha I)x - i\beta x\|^2 = \|(A - \alpha I)x\|^2 + |\beta|^2 \|x\|^2 \geq |\beta|^2 \|x\|^2.$$

It implies (see part 2) of Theorem 1) that $z \in \rho(A)$. It means that $\sigma(A) \subset \mathbb{R}$. Since $A = A^*$ and therefore closed then the spectrum $\sigma(A)$ is closed as a complement of an open set (see part 1) of Theorem 1).

It remains to prove that $\sigma(A) \neq \emptyset$. Assume on the contrary that $\sigma(A) = \emptyset$. Then the resolvent R_z is an entire analytic function. Let us prove that $\|R_z\|$ is uniformly bounded with respect to $z \in \mathbb{C}$. Introduce the functional

$$T_z(y) := (R_z x, y), \quad \|x\| = 1, y \in H.$$

Then $T_z(y)$ is a linear functional on the Hilbert space H. Moreover, since R_z is bounded for any (fixed) $z \in \mathbb{C}$ then

$$|T_z(y)| \leq \|R_z x\| \|y\| \leq \|R_z\| \|y\| = C_z \|y\|.$$

Therefore $T_z(y)$ is continuous i.e. $\{T_z, z \in \mathbb{C}\}$ is a pointwise bounded family of continuous linear functionals. By Banach-Steinhaus theorem we may conclude that

$$\sup_{z\in\mathbb{C}} \|T_z\| = c_0 < \infty.$$

That's why we have

$$|T_z(y)| = |(R_z x, y)| \le c_0 \|y\|, \quad \|x\| = 1, z \in \mathbb{C}.$$

It implies that $\|R_z x\| \le c_0$ i.e. $\|R_z\| \le c_0$. By Liouville theorem we may conclude now that R_z is constant with respect to z. But by J. von Neumann's spectral theorem

$$R_z = \int_{-\infty}^{\infty} \frac{1}{\lambda - z} dE_\lambda,$$

where $\{E_\lambda\}$ is a spectral family of $A = A^*$. Due to the estimate

$$\|R_z\| \le \frac{1}{|\operatorname{Im} z|}$$

we may conclude that $\|R_z\| \to 0$ as $|\operatorname{Im} z| \to \infty$. Hence $R_z \equiv 0$. This contradiction finishes the proof. □

Exercise 22. [Weyl's criterion] Let $A = A^*$. Prove that $\lambda \in \sigma(A)$ if and only if there exists $x_n \in D(A), \|x_n\| = 1$ such that

$$\lim_{n\to\infty} \|(A - \lambda I)x_n\| = 0.$$

Definition. Let us assume that $A = \overline{A}$. The *point spectrum* $\sigma_p(A)$ of A is the set of eigenvalues of A i.e.

$$\sigma_p(A) = \{\lambda \in \sigma(A) : N(A - \lambda I) \neq \{0\}\}.$$

It means that $(A - \lambda I)^{-1}$ does not exist i.e. there exists a non-trivial $u \in D(A)$ such that $Au = \lambda u$. The complement $\sigma(A) \setminus \sigma_p(A)$ is the *continuous spectrum* $\sigma_c(A)$. The *discrete spectrum* is the set

$$\sigma_d(A) = \{\lambda \in \sigma_p(A) : \dim N(A - \lambda I) < \infty \, \text{and} \, \lambda \, \text{is isolated in} \, \sigma(A)\}.$$

The set $\sigma_{\text{ess}}(A) := \sigma(A) \setminus \sigma_d(A)$ is called the *essential spectrum* of A.

In the frame of this definition, the complex plane can be divided into regions according to

$$\mathbb{C} = \rho(A) \cup \sigma(A),$$
$$\sigma(A) = \sigma_p(A) \cup \sigma_c(A)$$

and

$$\sigma(A) = \sigma_d(A) \cup \sigma_{\text{ess}}(A),$$

with all the unions being disjoint.

Remark. If $A = A^*$ then

1) $\lambda \in \sigma_c(A)$ means that $(A - \lambda I)^{-1}$ exists but is not bounded.

2) $\sigma_{\text{ess}}(A) = \overline{\sigma_c(A)} \cup \{\text{eigenvalues of infinite multiplicity and their accumulation points}\} \cup \{\text{accumulation points of } \sigma_d(A)\}$.

Exercise 23. Let $A = A^*$ and $\lambda_1, \lambda_2 \in \sigma_p(A)$. Prove that if $\lambda_1 \neq \lambda_2$ then

$$N(A - \lambda_1 I) \perp N(A - \lambda_2 I).$$

Exercise 24. Let $\{e_j\}_{j=1}^{\infty}$ be an orthonormal basis in H and let $\{s_j\}_{j=1}^{\infty} \subset \mathbb{C}$ be some sequence. Introduce the set

$$D = \left\{ x \in H : \sum_{j=1}^{\infty} |s_j|^2 |(x, e_j)|^2 < \infty \right\}.$$

Define

$$Ax = \sum_{j=1}^{\infty} s_j (x, e_j) e_j, \quad x \in D.$$

Prove that $A = \overline{A}$ and that $\sigma(A) = \overline{\{s_j : j = 1, 2, \ldots\}}$. Prove also that

$$(A - zI)^{-1} x = \sum_{j=1}^{\infty} \frac{1}{s_j - z} (x, e_j) e_j$$

for any $z \in \rho(A)$ and $x \in D$.

Exercise 25. Prove that the spectrum $\sigma(U)$ of a unitary operator U lies on the unit circle in $\mathbb{C}$.

Theorem 2. *Let $A = A^*$ and let $\{E_\lambda\}_{\lambda \in \mathbb{R}}$ be its spectral family. Then*

1) $\mu \in \sigma(A)$ if and only if $E_{\mu+\varepsilon} - E_{\mu-\varepsilon} \neq 0$ for every $\varepsilon > 0$.

2) $\mu \in \sigma_p(A)$ if and only if $E_\mu - E_{\mu-0} \neq 0$. Here $E_{\mu-0} := \lim_{\varepsilon \to 0+} E_{\mu-\varepsilon}$ in the sense of strong operator topology.

Proof. 1) Suppose that $\mu \in \sigma(A)$ but there exists $\varepsilon > 0$ such that $E_{\mu+\varepsilon} - E_{\mu-\varepsilon} = 0$. Then by spectral theorem we obtain for any $x \in D(A)$ that

$$\begin{aligned} \|(A - \mu I)x\|^2 &= \int_{-\infty}^{\infty} (\lambda - \mu)^2 d(E_\lambda x, x) \geq \int_{|\lambda-\mu| \geq \varepsilon} (\lambda - \mu)^2 d(E_\lambda x, x) \\ &\geq \varepsilon^2 \int_{|\lambda-\mu| \geq \varepsilon} d(E_\lambda x, x) = \varepsilon^2 \left[\int_{-\infty}^{\mu-\varepsilon} + \int_{\mu+\varepsilon}^{\infty} \right] d(E_\lambda x, x) \\ &= \varepsilon^2 \left[(E_{\mu-\varepsilon} x, x) + \|x\|^2 - (E_{\mu+\varepsilon} x, x) \right] = \varepsilon^2 \|x\|^2 . \end{aligned}$$

This inequality means (see part 2) of Theorem 1) that $\mu \notin \sigma(A)$ but $\mu \in \rho(A)$. This contradiction proves 1) in one direction. Conversely, if

$$P_n := E_{\mu+\frac{1}{n}} - E_{\mu-\frac{1}{n}} \neq 0$$

for all $n \in \mathbb{N}$ then there is a sequence $\{x_n\}_{n=1}^{\infty}$ such that $x_n \in R(P_n)$ i.e. $x_n = P_n x_n$ i.e. $x_n \in D(A)$ and $\|x_n\| = 1$. For this sequence it is true that

$$\begin{aligned}
\|(A-\mu I)x_n\|^2 &= \int_{-\infty}^{\infty} (\lambda-\mu)^2 d(E_\lambda P_n x_n, P_n x_n) = \int_{|\lambda-\mu|\le 1/n} (\lambda-\mu)^2 d(E_\lambda x_n, x_n) \\
&\le \frac{1}{n^2}\int_{-\infty}^{\infty} d(E_\lambda x_n, x_n) = \frac{1}{n^2}\|x_n\|^2 = \frac{1}{n^2} \to 0
\end{aligned}$$

as $n \to \infty$. Hence, this sequence satisfies Weyl's criterion (see Exercise 22) and therefore $\mu \in \sigma(A)$.

2) Suppose $\mu \in \mathbb{R}$ is an eigenvalue of A. Then there is $x_0 \in D(A), x_0 \neq 0$ such that

$$0 = \|(A-\mu I)x_0\|^2 = \int_{-\infty}^{\infty} (\lambda-\mu)^2 d(E_\lambda x_0, x_0).$$

In particular, for all $n \in \mathbb{N}$ and large enough $\varepsilon > 0$ we have that

$$\begin{aligned}
0 &= \int_{\mu+\varepsilon}^{n} (\lambda-\mu)^2 d(E_\lambda x_0, x_0) \ge \varepsilon^2 \int_{\mu+\varepsilon}^{n} d(E_\lambda x_0, x_0) = \varepsilon^2((E_n - E_{\mu+\varepsilon})x_0, x_0) \\
&= \varepsilon^2 \|(E_n - E_{\mu+\varepsilon})x_0\|^2 .
\end{aligned}$$

Thus we may conclude that

$$0 = E_n x_0 - E_{\mu+\varepsilon} x_0.$$

Similarly we can get that

$$0 = E_{-n} x_0 - E_{\mu-\varepsilon} x_0.$$

Letting $n \to \infty$ and $\varepsilon \to 0$ we obtain

$$x_0 = E_\mu x_0, \quad 0 = E_{\mu-0} x_0.$$

Hence

$$x_0 = (E_\mu - E_{\mu-0})x_0$$

and therefore

$$E_\mu - E_{\mu-0} \neq 0.$$

Conversely, define the projector

$$P := E_\mu - E_{\mu-0}.$$

If $P \neq 0$ then there exists $y \in H, y \neq 0$ such that $y = Py$ (e.g. any $y \in R(P) \neq \{0\}$ will do). For $\lambda > \mu$ it follows that

$$E_\lambda y = E_\lambda P y = E_\lambda E_\mu y - E_\lambda E_{\mu-0} y = Py = y.$$

For $\lambda < \mu$ we have that

$$E_\lambda y = E_\lambda E_\mu y - E_\lambda E_{\mu-0} y = E_\lambda y - E_\lambda y = 0.$$

Hence

$$\|(A - \mu I)y\|^2 = \int_{-\infty}^{\infty} (\lambda - \mu)^2 d(E_\lambda y, y) = \int_{\mu}^{\infty} (\lambda - \mu)^2 d_\lambda (y, y) = 0.$$

That's why $Ay = \mu y$ and $y \in D(A), y \neq 0$ i.e. μ is an eigenvalue of A or $\mu \in \sigma_p(A)$. □

Remark. The statements of Theorem 2 can be reformulated as

1) $\mu \in \sigma_p(A)$ if and only if $E_\mu - E_{\mu-0} \neq 0$.

2) $\mu \in \sigma_c(A)$ if and only if $E_\mu - E_{\mu-0} = 0$.

Definition. Let H and H_1 be two Hilbert spaces. A bounded linear operator $K : H \to H_1$ is called *compact* or *completely continuous* if it maps bounded sets in H into *precompact* sets in H_1 i.e. for every bounded sequence $\{x_n\}_{n=1}^{\infty} \subset H$ the sequence $\{Kx_n\}_{n=1}^{\infty} \subset H_1$ contains a convergent subsequence.

If $K : H \to H_1$ is compact then the following statements hold.

1) K maps every weakly convergent sequence in H into a norm convergent sequence in H_1.

2) If $H = H_1$ is separable then every compact operator is a norm limit of a sequence of operators of *finite rank* (i.e. operators with finite dimensional ranges).

3) The norm limit of a sequence of compact operators is compact.

Let us prove 2). Let K be a compact operator. Since H is separable it has an orthonormal basis $\{e_j\}_{j=1}^{\infty}$. Consider for any $n = 1, 2, \ldots$ the projector

$$P_n x := \sum_{j=1}^{n} (x, e_j) e_j, \quad x \in H.$$

Then $P_n \leq P_{n+1}$ and $\|(I - P_n)x\| \to 0$ as $n \to \infty$. Define

$$d_n := \sup_{\|x\|=1} \|K(I - P_n)x\| \equiv \|K(I - P_n)\|.$$

Since $R(I-P_n) \supset R(I-P_{n+1})$ (see Proposition 2 in Section 3) then $\{d_n\}_{n=1}^{\infty}$ is a monotone decreasing sequence of positive numbers. Hence the limit

$$\lim_{n\to\infty} d_n := d \geq 0$$

exists. Let us choose $y_n \in R(I-P_n)$, $\|y_n\| = 1$ such that

$$\|K(I-P_n)y_n\| = \|Ky_n\| \geq \frac{d}{2}.$$

Then

$$|(y_n, x)| = |((I-P_n)y_n, x)| = |(y_n, (I-P_n)x)| \leq \|y_n\| \, \|(I-P_n)x\| \to 0, \quad n \to \infty$$

for any $x \in H$. It means that $y_n \xrightarrow{w} 0$. Compactness of K implies that $Ky_n \to 0$. Thus $d = 0$. That's why

$$d_n = \|K - KP_n\| \to 0.$$

Since P_n is of finite rank then so is KP_n i.e. K is a norm limit of finite rank operators.

Lemma. *Suppose $A = A^*$ is compact. Then at least one of the two numbers $\pm\|A\|$ is an eigenvalue of A.*

Proof. Since

$$\|A\| = \sup_{\|x\|=1} |(Ax, x)|$$

then there exists a sequence x_n with $\|x_n\| = 1$ such that

$$\|A\| = \lim_{n\to\infty} |(Ax_n, x_n)|.$$

Actually, we can assume that $\lim_{n\to\infty}(Ax_n, x_n)$ exists and equals, say, a. Otherwise we would take a subsequence of $\{x_n\}$. Since $A = A^*$ then a is real and $\|A\| = |a|$. Due to the fact that any bounded set of the Hilbert space is weakly compact (unit ball in our case) we can choose a subsequence of $\{x_n\}$, say, $\{x_{k_n}\}$ which converges weakly i.e. $x_{k_n} \xrightarrow{w} x$. Compactness of A implies that $Ax_{k_n} \to y$. Next we observe that

$$\begin{aligned} \|Ax_{k_n} - ax_{k_n}\|^2 &= \|Ax_{k_n}\|^2 - 2a(Ax_{k_n}, x_{k_n}) + a^2 \leq \|A\|^2 - 2a(Ax_{k_n}, x_{k_n}) + a^2 \\ &= 2a^2 - 2a(Ax_{k_n}, x_{k_n}) \to 2a^2 - 2a^2 = 0, \end{aligned}$$

as $n \to \infty$. Hence

$$\begin{cases} Ax_{k_n} - ax_{k_n} \to 0 \\ Ax_{k_n} \to y \\ x_{k_n} \xrightarrow{w} x \end{cases} \quad \Rightarrow \quad \begin{cases} x_{k_n} \to x \\ Ax = ax. \end{cases}$$

Since $\|x_{k_n}\| = 1$ then $\|x\| = 1$ also. Hence $x \neq 0$ and a is an eigenvalue of A. $\square$

Theorem 3 (Riesz-Schauder)**.** *Suppose $A = A^*$ is compact. Then*

1) *A has a sequence of real eigenvalues $\lambda_j \neq 0$ which can be enumerated in such a way that*
$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_j| \geq \cdots .$$

2) *If there are infinitely many eigenvalues then $\lim_{j\to\infty} \lambda_j = 0$ and 0 is the only accumulation point of $\{\lambda_j\}$.*

3) *The multiplicity of λ_j is finite.*

4) *If e_j is the normalized eigenvector for λ_j then $\{e_j\}_{j=1}^{\infty}$ is an orthonormal system and*
$$Ax = \sum_{j=1}^{\infty} \lambda_j (x, e_j) e_j = \sum_{j=1}^{\infty} (Ax, e_j) e_j, \quad x \in H.$$
It means that $\{e_j\}_{j=1}^{\infty}$ is an orthonormal basis on $R(A)$.

5) *$\sigma(A) = \{0, \lambda_1, \lambda_2, \dots, \lambda_j, \dots\}$ while 0 is not necessarily an eigenvalue of A.*

Proof. Lemma gives the existence of an eigenvalue $\lambda_1 \in \mathbb{R}$ with $|\lambda_1| = \|A\|$ and a normalized eigenvector e_1. Introduce $H_1 = e_1^{\perp}$. Then H_1 is a closed subspace of H and A maps H_1 into itself. Indeed,
$$(Ax, e_1) = (x, Ae_1) = (x, \lambda_1 e_1) = \lambda_1 (x, e_1) = 0$$
for any $x \in H_1$. The restriction of the inner product of H to H_1 makes H_1 a Hilbert space (since H_1 is closed) and the restriction of A to H_1, denoted by $A_1 = A|_{H_1}$, is again a self-adjoint compact operator which is mapping in H_1. Clearly, its norm is bounded by the norm of A i.e. $\|A_1\| \leq \|A\|$. Applying Lemma to A_1 on H_1 we get an eigenvalue λ_2 with $|\lambda_2| = \|A_1\|$ and a normalized eigenvector e_2 with $e_2 \perp e_1$. It is clear that $|\lambda_2| \leq |\lambda_1|$. Next introduce the closed subspace $H_2 = (\operatorname{span}\{e_1, e_2\})^{\perp}$. Again, A leaves H_2 invariant and thus $A_2 := A_1|_{H_2} = A|_{H_2}$ is a self-adjoint compact operator in H_2. Applying Lemma to A_2 on H_2 we obtain λ_3 with $|\lambda_3| = \|A_2\|$ and a normalized eigenvector e_3 with $e_3 \perp e_2$ and $e_3 \perp e_1$. This process in the infinite dimensional Hilbert space leads us to the sequence $\{\lambda_j\}_{j=1}^{\infty}$ such that $|\lambda_{j+1}| \leq |\lambda_j|$ and corresponding normalized eigenvectors. Since $|\lambda_j| > 0$ and monotone decreasing then there is a limit
$$\lim_{j\to\infty} |\lambda_j| = r.$$

Clearly $r \geq 0$. Let us prove that $r = 0$. If $r > 0$ then $|\lambda_j| \geq r > 0$ for each $j = 1, 2, \dots$ or
$$\frac{1}{|\lambda_j|} \leq \frac{1}{r} < \infty.$$

Hence the sequence of vectors
$$y_j := \frac{e_j}{\lambda_j}$$

is bounded and therefore there is a weakly convergent subsequence $y_{j_k} \overset{w}{\to} y$. Compactness of A implies the strong convergence of $Ay_{j_k} \equiv e_{j_k}$. But $\|e_{j_k} - e_{j_m}\| = \sqrt{2}$ for $k \neq m$. This contradiction proves 1) and 2).

Exercise 26. Prove that if H is an infinite dimensional Hilbert space then the identical operator I is not compact and inverse of a compact operator (if it exists) is not bounded.

Exercise 27. Prove part 3) of Theorem 3.

Consider now the projector

$$P_n x := \sum_{j=1}^{n} (x, e_j) e_j, \quad x \in H.$$

Then $I - P_n$ is a projector onto $(\operatorname{span}\{e_1, \ldots, e_n\})^{\perp} \equiv H_n$ and hence

$$\|A(I - P_n)x\| \leq \|A\|_{H_n} \|(I - P_n)x\| \leq |\lambda_{n+1}| \, \|x\| \to 0$$

as $n \to \infty$. Since

$$AP_n x = \sum_{j=1}^{n} (x, e_j) A e_j = \sum_{j=1}^{n} \lambda_j (x, e_j) e_j$$

and

$$\|A(I - P_n)x\| = \|Ax - AP_n x\| \to 0, \quad n \to \infty$$

then

$$Ax = \sum_{j=1}^{\infty} \lambda_j (x, e_j) e_j$$

and part 4) follows. Finally, Exercise 24 gives immediately that

$$\sigma(A) = \{0, \lambda_1, \lambda_2, \ldots, \lambda_j, \ldots\}.$$

This finishes the proof. □

Corollary (Hilbert-Schmidt theorem)**.** *The orthonormal system* $\{e_j\}_{j=1}^{\infty}$ *of eigenvectors of a compact self-adjoint operator* A *in a Hilbert space* H *is an orthonormal basis if and only if* $N(A) = \{0\}$.

Proof. Recall from Exercise 13 that

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

If $N(A) = \{0\}$ then $H = \overline{R(A)}$. It means that for any $x \in H$ and any $\varepsilon > 0$ there exists $y_\varepsilon \in R(A)$ such that

$$\|x - y_\varepsilon\| < \varepsilon/2.$$

But by Riesz-Schauder theorem

$$y_\varepsilon = Ax_\varepsilon = \sum_{j=1}^{\infty} \lambda_j (x_\varepsilon, e_j) e_j.$$

Hence

$$\|x - y_\varepsilon\| = \left\| x - \sum_{j=1}^{\infty} \lambda_j (x_\varepsilon, e_j) e_j \right\| < \varepsilon/2.$$

Making use of the Theorem of Pythagoras, Bessel's inequality and Exercise 8 yields

$$\begin{aligned}
\left\| x - \sum_{j=1}^{n} (x, e_j) e_j \right\| &\le \left\| x - \sum_{j=1}^{n} \lambda_j (x_\varepsilon, e_j) e_j \right\| = \left\| x - \sum_{j=1}^{\infty} \lambda_j (x_\varepsilon, e_j) e_j + \sum_{j=n+1}^{\infty} \lambda_j (x_\varepsilon, e_j) e_j \right\| \\
&< \varepsilon/2 + \left\| \sum_{j=n+1}^{\infty} \lambda_j (x_\varepsilon, e_j) e_j \right\| \\
&\le \varepsilon/2 + \left(\sum_{j=n+1}^{\infty} |\lambda_j|^2 |(x_\varepsilon, e_j)|^2 \right)^{1/2} \\
&\le \varepsilon/2 + |\lambda_{n+1}| \left(\sum_{j=n+1}^{\infty} |(x_\varepsilon, e_j)|^2 \right)^{1/2} \\
&\le \varepsilon/2 + |\lambda_{n+1}| \, \|x_\varepsilon\| < \varepsilon
\end{aligned}$$

for n large enough. It means that $\{e_j\}_{j=1}^{\infty}$ is a basis in H, and moreover, it is an orthonormal basis.

Conversely, if $\{e_j\}_{j=1}^{\infty}$ is complete in H then $\overline{R(A)} = H$ (Riesz-Schauder) and therefore $N(A) = \{0\}$. □

Remark. The condition $N(A) = \{0\}$ means that A^{-1} exists and H must be separable in this case.

Proposition 1 (Riesz)**.** *If A is a compact operator on H and $\mu \in \mathbb{C}$ then the nullspace of $I - \mu A$ is a finite-dimensional subspace.*

Proof. The nullspace $N(I - \mu A)$ is a closed subspace of H since $I - \mu A$ is bounded. Indeed, for each sequence $f_n \to f$ and $f_n - \mu A f_n = 0$ we have that $f - \mu A f = 0$ since A is continuous.

The operator A is compact on H and therefore also compact from $N(I - \mu A)$ onto $N(I - \mu A)$, since $N(I - \mu A)$ is closed. Hence, for any $f \in N(I - \mu A)$ we have

$$If = (I - \mu A) f + \mu A f = \mu A f$$

and I is compact on $N(I - \mu A)$. Thus $N(I - \mu A)$ is finite-dimensional. □

Theorem 4 (Lemma of Riesz)**.** *If A is a compact operator on H and $\mu \in \mathbb{C}$ then $R(I - \mu A)$ is closed in H.*

Proof. If $\mu = 0$ then $R(I-\mu A) = H$. If $\mu \neq 0$ then we assume without loss of generality that $\mu = 1$. Let $f \in \overline{R(I - A)}, f \neq 0$. Then there exists a sequence $\{g_n\} \subset H$ such that

$$f = \lim_{n\to\infty} (I - A)g_n.$$

We will prove that $f \in R(I - A)$ i.e. there exists $g \in H$ such that $f = (I - A)g$. Since $f \neq 0$ then by the decomposition $H = N(I - A) \oplus N(I - A)^{\perp}$ we can assume that $g_n \in N(I - A)^{\perp}$ and $g_n \neq 0$ for all $n \in \mathbb{N}$.

Suppose that g_n is bounded. Then there is a subsequence $\{g_{k_n}\}$ such that

$$g_{k_n} \overset{w}{\to} g.$$

Compactness of A implies that

$$Ag_{k_n} \to h = Ag.$$

Next,

$$g_{k_n} = (I - A)g_{k_n} + Ag_{k_n} \to f + h.$$

Hence $g = f + Ag$ i.e. $f = (I - A)g$.

Suppose that g_n is not bounded. Then we can assume without loss of generality that $\|g_n\| \to \infty$. Introduce a new sequence

$$u_n := \frac{g_n}{\|g_n\|}.$$

Since $\|u_n\| = 1$ then there exists a subsequence $u_{k_n} \overset{w}{\to} u$. Compactness of A gives $Au_{k_n} \to Au$. Since $(I - A)g_n \to f$ then

$$(I - A)u_{k_n} = \frac{1}{\|g_{k_n}\|}(I - A)g_{k_n} \to 0.$$

It means again that

$$u_{k_n} = (I - A)u_{k_n} + Au_{k_n} \to Au$$

and $u = Au$ i.e. $u \in N(I - A)$. But $g_n \in N(I - A)^{\perp}$. Hence $u_{k_n} \in N(I - A)^{\perp}$ and further $u \in N(I - A)^{\perp}$ because $N(I - A)^{\perp}$ is closed. Since $\|u_{k_n}\| = 1$ then $\|u\| = 1$. Therefore $u \neq 0$ while

$$u \in N(I - A) \cap N(I - A)^{\perp}.$$

This contradiction shows that unbounded g_n cannot occur. □

We are now ready to derive the following fundamental result of the Riesz theory.

Theorem 5 (Riesz)**.** *Let $A : H \to H$ be a compact linear operator on the Hilbert space H. Then for any $\mu \in \mathbb{C}$ the operator $I - \mu A$ is* injective *(i.e. $(I - \mu A)^{-1}$ exists) if and only if it is* surjective *(i.e. $R(I - \mu A) = H$). Moreover, in this case the inverse operator $(I - \mu A)^{-1} : H \to H$ is bounded.*

Proof. If $(I - \mu A)^{-1}$ exists then $(I - \overline{\mu} A^*)^{-1}$ exists too and therefore $N(I - \overline{\mu} A^*) = 0$. Then Lemma of Riesz (Theorem 4) and Exercise 13 imply $H = R(I - \mu A)$, i.e. $I - \mu A$ is surjective.

Conversely, if $I - \mu A$ is surjective then $N(I - \overline{\mu} A^*) = 0$ i.e. $I - \overline{\mu} A^*$ injective and so is $I - \mu A$.

It remains to show that $(I - \mu A)^{-1}$ is bounded on H if $I - \mu A$ is injective. Assume that $(I - \mu A)^{-1}$ is not bounded. Then there exists a sequence $f_n \in H$ with $\|f_n\| = 1$ such that

$$\left\|(I - \mu A)^{-1} f_n\right\| \geq n.$$

Define

$$g_n := \frac{f_n}{\|(I - \mu A)^{-1} f_n\|}, \quad \varphi_n := \frac{(I - \mu A)^{-1} f_n}{\|(I - \mu A)^{-1} f_n\|}.$$

Then $g_n \to 0$ as $n \to \infty$ and $\|\varphi_n\| = 1$. Since A is compact we can select a subsequence φ_{k_n} such that $A\varphi_{k_n} \to \varphi$ as $k_n \to \infty$. But

$$\varphi_n - \mu A \varphi_n = g_n$$

and we observe $\varphi_{k_n} \to \mu\varphi$ and $\varphi \in N(I - \mu A)$. Hence $\varphi = 0$ and this contradicts $\|\varphi_n\| = 1$. □

Theorem 6 (Fredholm alternative)**.** *Suppose $A = A^*$ is compact. For given $g \in H$ either the equation*

$$(I - \mu A) f = g$$

has the unique solution ($\mu^{-1} \notin \sigma(A)$) and in this case $f = (I - \mu A)^{-1} g$ or $\mu^{-1} \in \sigma(A)$ and this equation has a solution if and only if $g \in R(I - \mu A)$ i.e. $g \perp N(I - \mu A)$. In this case the general solution of the equation is of the form $f = f_0 + u$, where f_0 is a particular solution and $u \in N(I - \mu A)$ (u is the general solution of the corresponding homogeneous equation) and the set of all solutions is a finite dimensional affine subspace of H.

Proof. Lemma of Riesz (Theorem 4) gives

$$R(I - \mu A) = N(I - \overline{\mu} A)^{\perp}.$$

If $\mu^{-1} \notin \sigma(A)$ then $(\overline{\mu})^{-1} \notin \sigma(A)$ also. Thus

$$R(I - \mu A) = N(I - \overline{\mu} A)^{\perp} = \{0\}^{\perp} = H.$$

Since $A = A^*$ this means that $(I - \mu A)^{-1}$ exists and the unique solution is $f = (I - \mu A)^{-1} g$.

If $\mu^{-1} \in \sigma(A)$ then $R(I - \mu A)$ is a proper subspace of H and the equation $(I - \mu A)f = g$ has a solution if and only if $g \in R(I - \mu A)$. Since the equation is linear then any solution is of the form

$$f = f_0 + u, \quad u \in N(I - \mu A)$$

and the dimension of $N(I - \mu A)$ is finite. □

Exercise 28. Let $A = A^*$ be compact. Prove that $\sigma_p(A) = \sigma_d(A) = \sigma(A) \setminus \{0\}$ and $0 \in \sigma_{\text{ess}}(A)$.

Exercise 29. Consider the Hilbert space $H = l^2(\mathbb{C})$ and

$$A(x_1, x_2, \dots, x_n, \dots) = (0, x_1, \frac{x_2}{2}, \dots, \frac{x_n}{n}, \dots)$$

for $(x_1, x_2, \dots, x_n, \dots) \in l^2(\mathbb{C})$. Show that A is compact and has no eigenvalues (even more, $\sigma(A) = \emptyset$) and is not self-adjoint.

Exercise 30. Consider the Hilbert space $H = L^2(\mathbb{R})$ and

$$(Af)(t) = tf(t).$$

Show that the equation $Af = f$ has no non-trivial solutions and that $(I - A)^{-1}$ does not exist. It means that the Fredholm alternative does not hold for non-compact but self-adjoint operator.

Exercise 31. Let $H = L^2(\mathbb{R}^n)$ and let

$$Af(x) = \int_{\mathbb{R}^n} K(x, y)f(y)dy,$$

where $K(x, y) \in L^2(\mathbb{R}^n \times \mathbb{R}^n)$ is such that $K(x, y) = \overline{K(y, x)}$. Prove that $A = A^*$ and that A is compact.

Theorem 7 (Weyl)**.** *If $A = A^*$ then $\lambda \in \sigma_{ess}(A)$ if and only if there exists an orthonormal system $\{x_n\}_{n=1}^{\infty}$ such that*

$$\|(A - \lambda I)x_n\| \to 0$$

as $n \to \infty$.

Proof. We will provide only a partial proof. Suppose that $\lambda \in \sigma_{\text{ess}}(A)$. If λ is an eigenvalue of infinite multiplicity then there is an infinite orthonormal system of eigenvectors $\{x_n\}_{n=1}^{\infty}$ because $\dim(E_\lambda - E_{\lambda-0})H = \infty$ in this case. Since $(A - \lambda I)x_n \equiv 0$ it is clear that

$$(A - \lambda I)x_n \to 0.$$

Next, suppose that λ is an accumulation point of $\sigma(A)$. It means that $\lambda \in \sigma(A)$ and

$$\lambda = \lim_{n\to\infty} \lambda_n,$$

where $\lambda_n \neq \lambda_m, n \neq m$ and $\lambda_n \in \sigma(A)$. Hence for each $n = 1, 2, \dots$ we have that

$$E_{\lambda_n+\varepsilon} - E_{\lambda_n-\varepsilon} \neq 0$$

for all $\varepsilon > 0$. Therefore there exists a sequence $r_n \to 0$ such that

$$E_{\lambda_n+r_n} - E_{\lambda_n-r_n} \neq 0.$$

That's why we can find a normalized vector $x_n \in R(E_{\lambda_n+r_n} - E_{\lambda_n-r_n})$. Since $\lambda_n \neq \lambda_m$ for $n \neq m$ we can find $\{x_n\}_{n=1}^{\infty}$ as an orthonormal system. By spectral theorem we have

$$\begin{aligned}
\|(A-\lambda I)x_n\|^2 &= \int_{-\infty}^{\infty} (\lambda-\mu)^2 d(E_\mu x_n, x_n) \\
&= \int_{-\infty}^{\infty} (\lambda-\mu)^2 d(E_\mu (E_{\lambda_n+r_n} - E_{\lambda_n-r_n}) x_n, x_n) \\
&= \int_{\lambda_n-r_n}^{\lambda_n+r_n} (\lambda-\mu)^2 d(E_\mu x_n, x_n) \\
&\leq \max_{\lambda_n-r_n\leq\mu\leq\lambda_n+r_n} (\lambda-\mu)^2 \int_{-\infty}^{\infty} d(E_\mu x_n, x_n) \\
&= \max_{\lambda_n-r_n\leq\mu\leq\lambda_n+r_n} (\lambda-\mu)^2 \to 0, \quad n\to\infty.
\end{aligned}$$

□

Theorem 8 (Weyl)**.** *Let A and B be two self-adjoint operators in a Hilbert space. If there is $z \in \rho(A) \cap \rho(B)$ such that*

$$T := (A - zI)^{-1} - (B - zI)^{-1}$$

is a compact operator then $\sigma_{ess}(A) = \sigma_{ess}(B)$.

Proof. We show first that $\sigma_{\text{ess}}(A) \subset \sigma_{\text{ess}}(B)$. Take any $\lambda \in \sigma_{\text{ess}}(A)$. Then there is an orthonormal system $\{x_n\}_{n=1}^{\infty}$ such that

$$\|(A - \lambda I)x_n\| \to 0, \quad n \to \infty.$$

Define the sequence y_n as

$$y_n := (A - zI)x_n \equiv (A - \lambda I)x_n + (\lambda - z)x_n.$$

Due to Bessel's inequality any orthonormal system in the Hilbert space converges weakly to 0. Hence $y_n \xrightarrow{w} 0$. We also have

$$\|y_n\| \geq |\lambda - z| \, \|x_n\| - \|(A - \lambda I)x_n\| = |\lambda - z| - \|(A - \lambda I)x_n\| > \frac{|\lambda - z|}{2} > 0$$

for all $n \geq n_0 >> 1$. Next we take the identity

$$\left[(B - zI)^{-1} - (\lambda - z)^{-1}\right] y_n = -T y_n - (\lambda - z)^{-1}(A - \lambda I)x_n.$$

Since T is compact and $y_n \overset{w}{\to} 0$ we deduce that

$$\left[(B - zI)^{-1} - (\lambda - z)^{-1}\right] y_n \to 0.$$

Introduce

$$z_n := (B - zI)^{-1} y_n.$$

Then

$$z_n - (\lambda - z)^{-1} y_n \to 0$$

or

$$y_n + (z - \lambda) z_n \to 0.$$

This fact and $\|y_n\| > \frac{|\lambda - z|}{2}$ imply that $\|z_n\| \geq \frac{|\lambda - z|}{3}$ for all $n \geq n_0 >> 1$. But

$$(B - \lambda I) z_n \equiv (B - zI) z_n + (z - \lambda) z_n = y_n + (z - \lambda) z_n \to 0.$$

Due to $\|z_n\| \geq \frac{|\lambda - z|}{3} > 0$ the sequence $\{z_n\}_{n=1}^{\infty}$ can be chosen as an orthonormal system. Thus $\lambda \in \sigma_{\text{ess}}(B)$. This proves that $\sigma_{\text{ess}}(A) \subset \sigma_{\text{ess}}(B)$. Finally, since $-T$ is compact too we can interchange the roles of A and B and obtain the opposite embedding. □

5 Quadratic forms. Friedrichs extension.

Definition. Let D be a linear subspace of a Hilbert space H. A function $Q : D \times D \to \mathbb{C}$ is called a *quadratic form* if

1) $Q(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1 Q(x_1, y) + \alpha_2 Q(x_2, y)$

2) $Q(x, \beta_1 y_1 + \beta_2 y_2) = \overline{\beta_1} Q(x, y_1) + \overline{\beta_2} Q(x, y_2)$

for all $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{C}$ and $x_1, x_2, x, y_1, y_2, y \in D$. The space $D(Q) := D$ is called the *domain* of Q. We say that Q is

a) *densely defined* if $\overline{D(Q)} = H$.

b) *symmetric* if $Q(x, y) = \overline{Q(y, x)}$.

c) *semibounded from below* if there exists $\lambda \in \mathbb{R}$ such that $Q(x, x) \geq -\lambda \|x\|^2$ for all $x \in D(Q)$.

d) *closed* (and semibounded) if $D(Q)$ is complete with respect to the norm

$$\|x\|_Q := \sqrt{Q(x, x) + (\lambda + 1) \|x\|^2}.$$

e) *bounded* (continuous) if there exists $M > 0$ such that

$$|Q(x, y)| \leq M \|x\| \|y\|$$

for all $x, y \in D(Q)$.

Exercise 32. Prove that $\|\cdot\|_Q$ is a norm and that

$$(x, y)_Q := Q(x, y) + (\lambda + 1)(x, y)$$

is an inner product.

Theorem 1. *Let Q be a densely defined, closed, semibounded and symmetric quadratic form in a Hilbert space H such that*

$$Q(x, x) \geq -\lambda \|x\|^2, \quad x \in D(Q).$$

Then there exists a unique self-adjoint operator A which is defined by the quadratic form Q as

$$Q(x, y) = (Ax, y), \quad x \in D(A), y \in D(Q),$$

which is semi-bounded from below i.e.

$$(Ax, x) \geq -\lambda \|x\|^2, \quad x \in D(A)$$

and $D(A) \subset D(Q)$.

Proof. Let us introduce an inner product on $D(Q)$ by

$$(x, y)_Q := Q(x, y) + (\lambda + 1)(x, y), \quad x, y \in D(Q)$$

(see Exercise 32). Since Q is closed then $D(Q) = \overline{D(Q)}$ is a closed subspace of H with respect to the norm $\|\cdot\|_Q$. It means that $D(Q)$ with this inner product defines a new Hilbert space H_Q. It is clear also that

$$\|x\|_Q \geq \|x\|$$

for all $x \in H_Q$. Thus, for fixed $x \in H$,

$$L(y) := (y, x), \quad y \in H_Q$$

defines a continuous (bounded) linear functional on the Hilbert space H_Q. Applying the Riesz-Frechet theorem to H_Q we obtain an element $x^* \in H_Q$ ($x^* \in D(Q)$) such that

$$(y, x) \equiv L(y) = (y, x^*)_Q.$$

It is clear that the map

$$H \ni x \mapsto x^* \in H_Q$$

defines a linear operator J such that

$$J : H \to H_Q, \quad Jx = x^*.$$

Hence

$$(y, x) = (y, Jx)_Q, \quad x \in H, y \in H_Q.$$

Next we prove that J is self-adjoint and that it has an inverse operator J^{-1}. For any $x, y \in H$ we have

$$(Jy, x) = (Jy, Jx)_Q = \overline{(Jx, Jy)_Q} = \overline{(Jx, y)} = (y, Jx).$$

Hence $J = J^*$. It is bounded due to Hellinger-Toeplitz theorem (Exercise 9). Suppose that $Jx = 0$. Then

$$(y, x) = (y, Jx)_Q = 0$$

for any $y \in D(Q)$. Since $\overline{D(Q)} = H$ then the last equality implies that $x = 0$ and therefore $N(J) = \{0\}$ and J^{-1} exists. Moreover,

$$H = N(J) \oplus \overline{R(J^*)} = \overline{R(J)}$$

and $R(J) \subset H_Q$. Now we can define a linear operator A on the domain $D(A) \equiv R(J)$ as

$$Ax := J^{-1}x - (\lambda + 1)x, \quad \lambda \in \mathbb{R}.$$

It is clear that A is densely defined and $A = A^*$ (J^{-1} is self-adjoint since J is). If now $x \in D(A)$ and $y \in D(Q) \equiv H_Q$ then

$$Q(x, y) = (x, y)_Q - (\lambda + 1)(x, y) = (J^{-1}x, y) - (\lambda + 1)(x, y) = (Ax, y).$$

The semi-boundedness of A from below follows from that of Q. It remains to prove that this representation for A is unique. Assume that we have two such representations, A_1 and A_2. Then for every $x \in D(A_1) \cap D(A_2)$ and $y \in D(Q)$ we have that

$$Q(x, y) = (A_1 x, y) = (A_2 x, y).$$

It follows that

$$((A_1 - A_2)x, y) = 0.$$

Since $\overline{D(Q)} = H$ then we must have $A_1 x = A_2 x$. This finishes the proof. □

Corollary. *Under the same assumptions as in Theorem 1, there exists* $\sqrt{A + \lambda I}$ *which is self-adjoint on* $D(\sqrt{A + \lambda I}) \equiv D(Q) = H_Q$*. Moreover,*

$$Q(x, y) + \lambda(x, y) = (\sqrt{A + \lambda I}x, \sqrt{A + \lambda I}y)$$

for all $x, y \in D(Q)$.

Proof. Since $A + \lambda I$ is self-adjoint and non-negative there exists a spectral family $\{E_\mu\}_{\mu=0}^{\infty}$ such that

$$A + \lambda I = \int_0^\infty \mu dE_\mu.$$

That's why we can define the operator

$$\sqrt{A + \lambda I} := \int_0^\infty \sqrt{\mu} dE_\mu$$

which is also self-adjoint and non-negative. Then for any $x \in D(A)$ and $y \in D(Q)$ we have that

$$Q(x, y) + \lambda(x, y) = ((A + \lambda I)x, y) = (\sqrt{A + \lambda I}x, \left(\sqrt{A + \lambda I}\right)^* y).$$

This means that $x \in D(\sqrt{A + \lambda I})$ and $y \in D(\left(\sqrt{A + \lambda I}\right)^*)$. But $\sqrt{A + \lambda I}$ is self-adjoint and, therefore,

$$D(\sqrt{A + \lambda I}) = D(\left(\sqrt{A + \lambda I}\right)^*) = D(Q) \equiv H_Q.$$

□

Theorem 2 (Friedrichs extension)**.** *Let A be a non-negative, symmetric linear operator in a Hilbert space H. Then there exists a self-adjoint extension A_F of A which is the smallest among all non-negative self-adjoint extensions of A in the sense that its corresponding quadratic form has the smallest domain. This extension A_F is called the* Friedrichs extension *of A.*

Proof. Let A be a non-negative, symmetric operator with domain $D(A)$ dense in H, $\overline{D(A)} = H$. Its associated quadratic form

$$Q(x,y) := (Ax,y), \quad x,y \in D(Q) \equiv D(A)$$

is densely defined, non-negative and symmetric. Let us define a new inner product

$$(x,y)_Q = Q(x,y) + (x,y), \quad x,y \in D(Q).$$

Then $D(Q)$ becomes an inner product space. This inner product space has a completion H_Q with respect to the norm

$$\|x\|_Q := \sqrt{Q(x,x) + \|x\|^2}.$$

Moreover, the quadratic form $Q(x,y)$ has an extension $Q_1(x,y)$ to this Hilbert space H_Q defined by

$$Q_1(x,y) = \lim_{n\to\infty} Q(x_n,y_n)$$

whenever $x \overset{H_Q}{=} \lim_{n\to\infty} x_n, y \overset{H_Q}{=} \lim_{n\to\infty} y_n, x_n, y_n \in D(Q)$ and these limits exist. The quadratic form Q_1 is densely defined, closed, non-negative and symmetric. That's why Theorem 1, applied to Q_1, gives a unique and non-negative, self-adjoint operator A_F such that

$$Q_1(x,y) = (A_F x,y), \quad x \in D(A_F) \subset H_Q, y \in D(Q_1) \equiv H_Q.$$

Since for $x,y \in D(A)$ one has

$$(Ax,y) = Q(x,y) = Q_1(x,y) = (A_F x,y)$$

then A_F is a self-adjoint extension of A.

It remains to prove that A_F is the smallest non-negative self-adjoint extension of A. Suppose that $B \geq 0, B = B^*$ is such that $A \subset B$. The associated quadratic form $Q_B(x,y) := (Bx,y)$ is an extension of $Q \equiv Q_A$. Hence

$$\overline{Q_B} \supset \overline{Q} = Q_1.$$

This finishes the proof. □

6 Elliptic differential operators

Let Ω be a domain in $\mathbb{R}^n$ i.e. an open and connected set. Introduce the following notation:

1) $x = (x_1, \dots, x_n) \in \Omega$
2) $|x| = \sqrt{x_1^2 + \dots + x_n^2}$
3) $\alpha = (\alpha_1, \dots, \alpha_n)$ is a *multi-index* i.e. $\alpha_j \in \mathbb{N}_0 \equiv \mathbb{N} \cup \{0\}$.
   a) $|\alpha| = \alpha_1 + \dots + \alpha_n$
   b) $\alpha \geq \beta$ if $\alpha_j \geq \beta_j$ for all $j = 1, 2, \dots, n$.
   c) $\alpha + \beta = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)$
   d) $\alpha - \beta = (\alpha_1 - \beta_1, \dots, \alpha_n - \beta_n)$ if $\alpha \geq \beta$
   e) $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ with $0^0 = 1$
   f) $\alpha! = \alpha_1! \cdots \alpha_n!$ with $0! = 1$
4) $D_j = \frac{1}{i}\partial_j = \frac{1}{i}\frac{\partial}{\partial x_j} = -i\partial_j$ and $D^\alpha = D_1^{\alpha_1} \cdots D_n^{\alpha_n} \equiv (-i)^{|\alpha|}\partial^\alpha$

Definition. An *elliptic partial differential operator* $A(x, D)$ of order m on Ω is an operator of the form

$$A(x, D) = \sum_{|\alpha| \leq m} a_\alpha(x) D^\alpha,$$

where $a_\alpha(x) \in C^\infty(\Omega)$ and whose *principal symbol*

$$a(x, \xi) = \sum_{|\alpha| = m} a_\alpha(x) \xi^\alpha, \quad \xi \in \mathbb{R}^n$$

is invertible for all $x \in \Omega$ and $\xi \in \mathbb{R}^n \setminus \{0\}$, that is, $a(x, \xi) \neq 0$ for all $x \in \Omega$ and $\xi \in \mathbb{R}^n \setminus \{0\}$.

Assumption 1. We assume that $a_\alpha(x)$ are real for $|\alpha| = m$.

Under Assumption 1 either $a(x, \xi) > 0$ or $a(x, \xi) < 0$ for all $x \in \Omega$ and $\xi \in \mathbb{R}^n \setminus \{0\}$. Without loss of generality we assume that $a(x, \xi) > 0$. Assumption 1 implies also that m is even and for any compact set $K \subset \Omega$ there exists $C_K > 0$ such that

$$a(x, \xi) \geq C_K |\xi|^m, \quad x \in \Omega, \xi \in \mathbb{R}^n.$$

Assumption 2. We assume that $A(x, D)$ is *formally self-adjoint* i.e.

$$A(x, D) = A^*(x, D) := \sum_{|\alpha| \leq m} (-1)^{|\alpha|} D^\alpha (\overline{a_\alpha(x)} \cdot).$$

Exercise 33. Prove that $A(x, D) = A^*(x, D)$ if and only if

$$a_\alpha(x) = \sum_{\substack{\alpha \le \beta \\ |\beta| \le m}} (-1)^{|\beta|} C_\beta^\alpha D^{\beta-\alpha} \overline{a_\beta(x)},$$

where

$$C_\beta^\alpha = \frac{\beta!}{\alpha!(\beta-\alpha)!}.$$

Hint: Make use of the *generalized Leibniz formula*

$$D^\alpha(fg) = \sum_{\beta \le \alpha} C_\alpha^\beta D^{\alpha-\beta} f D^\beta g.$$

Assumption 3. We assume that $A(x, D)$ has a divergence form

$$A(x, D) \equiv \sum_{|\alpha|=|\beta| \le m/2} D^\alpha(a_{\alpha\beta}(x) D^\beta),$$

where $a_{\alpha\beta} = a_{\beta\alpha}$ and real for all α and β. We assume also the *ellipticity condition*

$$\sum_{|\alpha|=|\beta|=m/2} a_{\alpha\beta}(x) \xi^\alpha \xi^\beta \ge \nu \sum_{|\alpha|=m/2} |\xi^\alpha|^2 = \nu \sum_{|\alpha|=m/2} \xi^{2\alpha},$$

where $\nu > 0$ is called the *constant of ellipticity*. Such operator is called *uniformly elliptic*.

Exercise 34. Prove that

$$\sum_{|\alpha|=m/2} \xi^{2\alpha} \asymp |\xi|^m$$

i.e.

$$c|\xi|^m \le \sum_{|\alpha|=m/2} \xi^{2\alpha} \le C|\xi|^m,$$

where c and C are some constants.

Example 6.1. Let us consider

$$A(x, D) = \sum_{j=1}^n D_j^2 = -\Delta, \quad x \in \Omega \subset \mathbb{R}^n$$

in $H = L^2(\Omega)$ and prove that $A \subset A^*$ with

$$D(A) = C_0^\infty(\Omega) = \left\{ f \in C^\infty(\Omega) : \operatorname{supp} f = \overline{\{x : f(x) \ne 0\}} \text{ is compact in } \Omega \right\}.$$

Let $u, v \in C_0^\infty(\Omega)$. Then

$$\begin{aligned}
(Au, v)_{L^2} &= \int_\Omega \left(\sum_{j=1}^n D_j^2 u \right) \overline{v} dx = -\sum_{j=1}^n \int_\Omega \left(\partial_j^2 u \right) \overline{v} dx \\
&= -\sum_{j=1}^n \int_\Omega \partial_j \left((\partial_j u) \overline{v} \right) dx + \sum_{j=1}^n \int_\Omega (\partial_j u) \left(\overline{\partial_j v} \right) dx \\
&= -\int_{\partial\Omega} (\overline{v} \nabla u, n_x) dx + (\nabla u, \nabla v)_{L^2} = (\nabla u, \nabla v)_{L^2},
\end{aligned}$$

where $\partial\Omega$ is the boundary of Ω and n_x is the unit outward vector at $x \in \partial\Omega$. Here we have made use of the divergence theorem. In a similar fashion we obtain

$$(\nabla u, \nabla v)_{L^2} = -\sum_{j=1}^n \int_\Omega u \partial_j^2 \overline{v} dx = (u, -\Delta v)_{L^2} = (u, Av)_{L^2}.$$

Hence $A \subset A^*$ and A is closable.

Example 6.2. Recall from Example 6.1 that

$$(-\Delta u, v)_{L^2} = (\nabla u, \nabla v)_{L^2}, \quad u, v \in C_0^\infty(\Omega).$$

Hence

$$(-\Delta u, u)_{L^2} = \|\nabla u\|_{L^2}^2 \leq \|u\|_{L^2} \|\Delta u\|_{L^2}, \quad u \in C_0^\infty(\Omega).$$

Therefore,

$$\begin{aligned}
\|u\|_{W_2^2}^2 &= \|u\|_{L^2}^2 + \|\nabla u\|_{L^2}^2 + \|\Delta u\|_{L^2}^2 \\
&\leq \|u\|_{L^2}^2 + \|u\|_{L^2} \|\Delta u\|_{L^2} + \|\Delta u\|_{L^2}^2 \\
&\leq \frac{3}{2} \|u\|_{L^2}^2 + \frac{3}{2} \|\Delta u\|_{L^2}^2 \equiv \frac{3}{2} \|u\|_A^2,
\end{aligned}$$

where $\|\cdot\|_A$ is a norm which corresponds to the operator $A = -\Delta$ as follows:

$$\|u\|_A^2 := \|u\|_{L^2}^2 + \|-\Delta u\|_{L^2}^2.$$

It is also clear that $\|u\|_A \leq \|u\|_{W_2^2}$. Combining these inequalities gives

$$\sqrt{\frac{2}{3}} \|u\|_{W_2^2} \leq \|u\|_A \leq \|u\|_{W_2^2}$$

for all $u \in C_0^\infty(\Omega)$. A completion of $C_0^\infty(\Omega)$ with respect to these norms leads us to the statement:

$$D(\overline{A}) = \overset{\circ}{W}{}_2^2(\Omega).$$

Thus $\overline{A} = -\Delta$ on $D(\overline{A}) = \overset{\circ}{W}{}_2^2(\Omega)$. Let us determine $D(A^*)$ in this case. By the definition of $D(A^*)$ we have

$$D((-\Delta)^*) = \{v \in L^2(\Omega) : \text{there exists}\, v^* \in L^2(\Omega)\, \text{such that} \\ (-\Delta u, v) = (u, v^*)\, \text{for all}\, u \in C_0^\infty(\Omega)\}.$$

If we assume that $v \in W_2^2(\Omega)$ then it is equivalent to

$$(u, (-\Delta)^* v) = (u, v^*)$$

i.e. $(-\Delta)^* v = v^*$ and $D((-\Delta)^*) = W_2^2(\Omega)$. Finally, for $\Omega \subset \mathbb{R}^n$ with $\Omega \neq \mathbb{R}^n$ we obtain that

$$A \subset \overline{A} \subset A^* \equiv (\overline{A})^*$$

and $A \neq \overline{A}$ and $\overline{A} \neq (\overline{A})^*$, that is, the closure of A does not lead us to a self-adjoint operator.

Remark. If $\Omega = \mathbb{R}^n$ then $\overset{\circ}{W}{}_2^2(\mathbb{R}^n) \equiv W_2^2(\mathbb{R}^n)$ and therefore

$$\overline{A} = A^* = (\overline{A})^*.$$

Hence the closure of A is self-adjoint in that case.

Example 6.3. Consider again $A = -\Delta$ on $D(A) = C_0^\infty(\Omega)$ with $\Omega \neq \mathbb{R}^n$. Since

$$(-\Delta u, u)_{L^2} = \|\nabla u\|_{L^2}^2 \geq 0$$

then $-\Delta$ is non-negative with lower bound $\lambda = 0$. That's why

$$Q(u, v) := (\nabla u, \nabla v)_{L^2}$$

is a densely defined and non-negative quadratic form with $D(Q) \equiv D(A) = C_0^\infty(\Omega)$. A new inner product is defined as

$$(u, v)_Q := (\nabla u, \nabla v)_{L^2} + (u, v)_{L^2}$$

and

$$\|u\|_Q^2 \equiv \|u\|_{W_2^1(\Omega)}^2 .$$

If we apply now the procedure from Theorem 2 from Section 5 then we obtain the existence of $Q_1 = \overline{Q}$ with respect to the norm $\|\cdot\|_Q$ which will also be non-negative and closed with $D(Q_1) \equiv \overset{\circ}{W}{}_2^1(\Omega)$. Next step is to get the Friedrichs extension A_F as

$$A_F = J^{-1} - I$$

with $D(A_F) \equiv R(J) \subset \overset{\circ}{W}{}_2^1(\Omega)$. More careful examination of Theorem 1 of Section 5 leads us to the fact

$$D(A_F) = \overset{\circ}{W}{}_2^1(\Omega) \cap D(A^*) = \overset{\circ}{W}{}_2^1(\Omega) \cap W_2^2(\Omega).$$

Remark. In general, for symmetric operator, we have

$$D(A_F) = \{u \in H_Q : Au \in H\}$$

which is equivalent to

$$D(A_F) = \{u \in H_Q : u \in D(A^*)\}.$$

Exercise 35. Let $H = L^2(\Omega)$ and $A(x, D) = -\Delta + q(x)$, where $q(x) = \overline{q(x)}$ and $q(x) \in L^\infty(\Omega)$. Define $\overline{A}, A^*$ and A_F.

Exercise 36. Let $H = L^2(\Omega)$ and

$$A(x, D) = -(\nabla + i\overrightarrow{W}(x))^2 + q(x),$$

where $\overrightarrow{W}$ is an n-dimensional real-valued vector from $W^1_\infty(\Omega)$ and q is a real-valued function from $L^\infty(\Omega)$. Define $\overline{A}, A^*$ and A_F.

Consider now bounded $\Omega \subset \mathbb{R}^n$ and an elliptic operator $A(x, D)$ in Ω of the form

$$A(x, D) = \sum_{|\alpha|=|\beta|\le m/2} D^\alpha(a_{\alpha\beta}(x)D^\beta),$$

where $a_{\alpha\beta}(x) = a_{\beta\alpha}(x)$ are real. Assume that there exists $C_0 > 0$ such that

$$|a_{\alpha\beta}(x)| \le C_0, \quad |\alpha|, |\beta| < \frac{m}{2}$$

for all $x \in \Omega$. Assume also that $A(x, D)$ is elliptic, that is,

$$\sum_{|\alpha|=|\beta|=m/2} a_{\alpha\beta}(x)\xi^\alpha\xi^\beta \ge \nu \sum_{|\alpha|=m/2} |\xi^\alpha|^2, \quad \nu > 0.$$

Theorem 1 (*Gårding's inequality*)**.** *Suppose that $A(x, D)$ is as above. Then for any $\varepsilon > 0$ there is $C_\varepsilon > 0$ such that*

$$(Af, f)_{L^2(\Omega)} \ge (\nu - \varepsilon) \|f\|^2_{W_2^{m/2}(\Omega)} - C_\varepsilon \|f\|^2_{L^2(\Omega)}$$

for any $f \in C_0^\infty(\Omega)$.

Proof. Let $f \in C_0^\infty(\Omega)$. Then integration by parts yields

$$\begin{aligned}
(Af,f)_{L^2(\Omega)} &= \sum_{|\alpha|=|\beta|\le m/2} \int_\Omega D^\alpha(a_{\alpha\beta}(x)D^\beta f)\overline{f}dx \\
&= \sum_{|\alpha|=|\beta|=m/2} \int_\Omega (-1)^{|\alpha|} a_{\alpha\beta}(x) D^\alpha \overline{f} D^\beta f dx \\
&+ \sum_{|\alpha|=|\beta|<m/2} \int_\Omega (-1)^{|\alpha|} a_{\alpha\beta}(x) D^\alpha \overline{f} D^\beta f dx \\
&\ge \nu \sum_{|\alpha|=m/2} \int_\Omega |D^\alpha f|^2 dx - C_0 \sum_{|\alpha|=|\beta|<m/2} \int_\Omega |D^\alpha f||D^\beta f| dx \\
&\ge \nu \sum_{|\alpha|\le m/2} \int_\Omega |D^\alpha f|^2 dx - (C_0+\nu) \sum_{|\alpha|<m/2} \int_\Omega |D^\alpha f|^2 dx \\
&= \nu \|f\|^2_{W_2^{m/2}(\Omega)} - (C_0+\nu)\|f\|^2_{W_2^{m/2-1}(\Omega)}.
\end{aligned}$$

Next we make use of the following

Lemma. *For any $\varepsilon > 0$ and $0 < \delta \le m/2$ there is $C_\varepsilon(\delta) > 0$ such that*

$$(1+|\xi|^2)^{m/2-\delta} \le \varepsilon(1+|\xi|^2)^{m/2} + C_\varepsilon(\delta)$$

for any $\xi \in \mathbb{R}^n$.

Proof. Let $\varepsilon > 0$ and $0 < \delta \le m/2$. If $(1+|\xi|^2)^\delta \ge \frac{1}{\varepsilon}$ then

$$(1+|\xi|^2)^{-\delta} \le \varepsilon.$$

Hence

$$(1+|\xi|^2)^{m/2-\delta} \le \varepsilon(1+|\xi|^2)^{m/2}$$

i.e. the claim holds for any positive constant $C_\varepsilon(\delta)$. For $(1+|\xi|^2)^\delta < \frac{1}{\varepsilon}$ we can get

$$(1+|\xi|^2)^{m/2-\delta} < \left(\frac{1}{\varepsilon}\right)^{\frac{m/2-\delta}{\delta}} \equiv C_\varepsilon(\delta).$$

□

Applying this lemma with $\delta = 1$ to the norm of Sobolev spaces W_2^k we may conclude that

$$\|f\|^2_{W_2^{m/2-1}(\Omega)} \le \varepsilon_1 \|f\|^2_{W_2^{m/2}(\Omega)} + C_{\varepsilon_1}\|f\|^2_{L^2(\Omega)}$$

for any $\varepsilon_1 > 0$. Hence

$$\begin{aligned}
(Af,f)_{L^2(\Omega)} &\ge \nu \|f\|^2_{W_2^{m/2}(\Omega)} - (C_0+\nu)\|f\|^2_{W_2^{m/2-1}(\Omega)} \\
&\ge \nu \|f\|^2_{W_2^{m/2}(\Omega)} - (C_0+\nu)\varepsilon_1 \|f\|^2_{W_2^{m/2}(\Omega)} - (C_0+\nu)C_{\varepsilon_1}\|f\|^2_{L^2(\Omega)} \\
&= (\nu-\varepsilon)\|f\|^2_{W_2^{m/2}(\Omega)} - C_\varepsilon \|f\|^2_{L^2(\Omega)}.
\end{aligned}$$

This proves the theorem. □

Corollary 1. *There exists a self-adjoint Friedrichs extension A_F of A with domain $D(A_F) = \overset{\circ}{W}{}_2^{m/2}(\Omega) \cap W_2^m(\Omega)$.*

Proof. It follows from Gårding's inequality that

$$(Af, f)_{L^2(\Omega)} \quad \geq \quad -C_\varepsilon \, \|f\|^2_{L^2(\Omega)} \,, \qquad f \in D(A).$$

This means that $A_\mu := A + \mu I$ is positive for $\mu > C_\varepsilon$ and therefore Theorem 2 of Section 5 gives us the existence of

$$(A_\mu)_F \equiv (A_F)_\mu = A_F + \mu I$$

with domain

$$D(A_F) = D((A_\mu)_F) = \overset{\circ}{W}{}_2^{m/2}(\Omega) \cap D(A^*),$$

where $\overset{\circ}{W}{}_2^{m/2}(\Omega)$ is the domain of the corresponding closed quadratic form (see Theorem 2). If Ω is bounded with smooth boundary $\partial\Omega$ then it can be proved that

$$D(A^*) = W_2^m(\Omega).$$

□

Gårding's inequality has two more consequences. Firstly,

$$\|(A_F)_\mu f\|_{L^2} \geq C_0 \, \|f\|_{L^2} \,, \quad C_0 > 0$$

so that

$$(A_F)_\mu^{-1} : L^2(\Omega) \to L^2(\Omega).$$

Secondly,

$$\|(A_F)_\mu f\|_{W_2^{-m/2}(\Omega)} \geq C_0' \, \|f\|_{W_2^{m/2}(\Omega)} \,, \quad C_0' > 0$$

so that

$$(A_F)_\mu^{-1} : L^2(\Omega) \to \overset{\circ}{W}{}_2^{m/2}(\Omega).$$

Corollary 2. *The spectrum $\sigma(A_F) = \{\lambda_j\}_{j=1}^\infty$ is the sequence of eigenvalues of finite multiplicity with only one accumulation point at $+\infty$. In short, $\sigma(A_F) = \sigma_d(A_F)$. The corresponding orthonormal system $\{\psi_j\}_{j=1}^\infty$ of eigenfunctions forms an orthonormal basis and*

$$A_F f \overset{L^2}{=} \sum_{j=1}^\infty \lambda_j (f, \psi_j) \psi_j$$

for any $f \in D(A_F)$.

Proof. We begin with a lemma.

Lemma. *The embedding*

$$\overset{\circ}{W}{}_2^{m/2}(\Omega) \hookrightarrow L^2(\Omega)$$

is compact.

Proof. It is enough to show that for any $\{\varphi_k\}_{k=1}^{\infty} \subset \overset{\circ}{W}{}_2^{m/2}(\Omega)$ with $\|\varphi_k\|_{W_2^{m/2}} \le 1$ there exists $\{\varphi_{j_k}\}_{k=1}^{\infty}$ which is a Cauchy sequence in $L^2(\Omega)$. Since Ω is bounded we have

$$|\widehat{\varphi_k}(\xi)| \le \|\varphi_k\|_{L^2} |\Omega|^{1/2}$$

i.e. the Fourier transform $\widehat{\varphi_k}(\xi)$ is uniformly bounded. That's why there exists $\widehat{\varphi_{j_k}}(\xi)$ which converges pointwise in $\mathbb{R}^n$. Next,

$$\begin{aligned}
\|\varphi_{j_k} - \varphi_{j_m}\|_{L^2}^2 &= \int_{\mathbb{R}^n} |\widehat{\varphi_{j_k}}(\xi) - \widehat{\varphi_{j_m}}(\xi)|^2 d\xi \\
&= \int_{|\xi|<r} |\widehat{\varphi_{j_k}}(\xi) - \widehat{\varphi_{j_m}}(\xi)|^2 d\xi + \int_{|\xi|>r} |\widehat{\varphi_{j_k}}(\xi) - \widehat{\varphi_{j_m}}(\xi)|^2 d\xi \\
&\le \int_{|\xi|<r} |\widehat{\varphi_{j_k}}(\xi) - \widehat{\varphi_{j_m}}(\xi)|^2 d\xi \\
&+ \frac{1}{(1+r^2)^{m/2}} \int_{\mathbb{R}^n} (1+|\xi|^2)^{m/2} |\widehat{\varphi_{j_k}}(\xi) - \widehat{\varphi_{j_m}}(\xi)|^2 d\xi \\
&= \int_{|\xi|<r} |\widehat{\varphi_{j_k}}(\xi) - \widehat{\varphi_{j_m}}(\xi)|^2 d\xi + (1+r^2)^{-m/2} \|\varphi_{j_k} - \varphi_{j_m}\|_{W_2^{m/2}}^2 \\
&:= I_1 + I_2.
\end{aligned}$$

The first term I_1 tends to 0 as $k, m \to \infty$ by the dominated convergence theorem of Lebesgue for any fixed $r > 0$. The second term converges to 0 as $r \to \infty$ because $\|\varphi_{j_k} - \varphi_{j_m}\|_{W_2^{m/2}} \le 2$ □

Lemma gives us that

$$(A_\mu)_F^{-1} : L^2(\Omega) \to L^2(\Omega)$$

is a compact operator. Applying Riesz-Schauder and Hilbert-Schmidt theorems we get

1) $\sigma((A_\mu)_F^{-1}) = \{0, \mu_1, \mu_2, \ldots\}$ with $\mu_j \ge \mu_{j+1} > 0$ and $\mu_j \to 0$ as $j \to \infty$.

2) μ_j is of finite multiplicity

3) $(A_\mu)_F^{-1} \psi_j = \mu_j \psi_j$, where $\{\psi_j\}_{j=1}^{\infty}$ is an orthonormal system

4) $\{\psi_j\}_{j=1}^{\infty}$ forms an orthonormal basis in $L^2(\Omega)$.

Since $A_F \psi_j = \lambda_j \psi_j$ with $\lambda_j = \frac{1}{\mu_j} - \mu$ then we may conclude that

$$\sigma(A_F) = \{\lambda_j\}_{j=1}^{\infty}, \quad \lambda_j \le \lambda_{j+1}, \lambda_j \to \infty.$$

Moreover, λ_j has finite multiplicity and ψ_j are the corresponding eigenfunctions. We have also the following representation

$$(A_\mu)_F^{-1} f = \sum_{j=1}^{\infty} \mu_j (f, \psi_j)\psi_j, \quad f \in L^2(\Omega).$$

Exercise 37. Prove that

$$A_F f = \sum_{j=1}^{\infty} \lambda_j (f, \psi_j)\psi_j$$

for any $f \in D(A_F)$.

Now we may conclude that the corollary is proved. $\square$

7 Spectral function

Let us consider a bounded domain $\Omega \subset \mathbb{R}^n$ and an elliptic differential operator $A(x, D)$ in Ω of the form

$$A(x, D) = \sum_{|\alpha|=|\beta|\le m/2} D^\alpha(a_{\alpha\beta}(x)D^\beta),$$

where $a_{\alpha\beta} = a_{\beta\alpha}$ are real, $a_{\alpha\beta} \in C^\infty(\Omega)$ and bounded for all α and β. We assume that

$$\sum_{|\alpha|=|\beta|=m/2} a_{\alpha\beta}(x)\xi^\alpha\xi^\beta \ge \nu|\xi|^m, \quad \nu > 0.$$

As it was proved above there exists at least one self-adjoint extension of A with $D(A) = C_0^\infty(\Omega)$, namely, the Friedrichs extension A_F with

$$D(A_F) = \overset{\circ}{W}{}_2^{m/2}(\Omega) \cap W_2^m(\Omega).$$

Let us consider an arbitrary self-adjoint extension $\widehat{A}$ of A. Without loss of generality we assume that $\widehat{A} \ge 0$. That's why $\widehat{A}$ has the spectral representation

$$\widehat{A} = \int_0^\infty \lambda dE_\lambda$$

with domain

$$D\left(\widehat{A}\right) = \left\{ f \in L^2(\Omega) : \int_0^\infty \lambda^2 d(E_\lambda f, f) < \infty \right\}.$$

In general case we have no such formula for $D\left(\widehat{A}\right)$ as for the Friedrichs extension A_F. But we can say that

$$\overset{\circ}{W}{}_2^m(\Omega) \subset D\left(\widehat{A}\right).$$

Indeed, since $a_{\alpha\beta}(x) \in C^\infty(\Omega)$ and bounded then $A(x, D)$ can be rewritten in the usual form

$$A(x, D) = \sum_{|\gamma|\le m} \widetilde{a_\gamma}(x)D^\gamma$$

with bounded coefficients. Hence

$$\|Af\|_{L^2(\Omega)} \le c \sum_{|\gamma|\le m} \|D^\gamma f\|_{L^2(\Omega)} \equiv c\,\|f\|_{W_2^m(\Omega)}.$$

This proves the embedding.

Theorem 1 (Gårding). *If* $\widehat{A} = \widehat{A}^*$ *then* E_λ *is an integral operator in* $L^2(\Omega)$ *such that*

$$E_\lambda f(x) = \int_\Omega \theta(x, y, \lambda) f(y) dy,$$

where $\theta(x, y, \lambda)$ *is called the* spectral function *and has the properties*

1) $\theta(x,y,\lambda) = \overline{\theta(y,x,\lambda)}$

2)

$$\theta(x,y,\lambda) = \int_\Omega \theta(x,z,\lambda)\theta(z,y,\lambda)dz$$

and

$$\theta(x,x,\lambda) = \int_\Omega |\theta(x,z,\lambda)|^2 dz \geq 0$$

3)

$$\sup_{x\in\Omega_1} \|\theta(x,\cdot,\lambda)\|_{L^2(\Omega)} \leq c_1\lambda^k,$$

where $\overline{\Omega_1} = \Omega_1 \subset \Omega, k \in \mathbb{N}$ *with* $k > \frac{n}{2m}$ *and* $c_1 = c(\Omega_1)$.

Remark. It was proved by L. Hörmander that actually

$$\theta(x,x,\lambda) \leq c_1\lambda^{n/m}.$$

Corollary. *Let* $z \in \rho\left(\widehat{A}\right)$. *Then* $(\widehat{A} - zI)^{-1}$ *is an integral operator whose kernel* $G(x,y,z)$ *is called the* Green's function *corresponding to* $\widehat{A}$ *and which has the properties*

1)

$$G(x,y,z) = \int_0^\infty \frac{d_\lambda\theta(x,y,\lambda)}{\lambda - z}$$

2) $\overline{G(x,y,z)} = G(y,x,\overline{z})$.

Proof. Since $z \in \rho\left(\widehat{A}\right)$ then J. von Neumann's spectral theorem gives us

$$(\widehat{A} - zI)^{-1}f = \int_0^\infty (\lambda - z)^{-1} dE_\lambda f.$$

Next, by Theorem 1 we get

$$\begin{aligned}(\widehat{A} - zI)^{-1}f &= \int_0^\infty (\lambda - z)^{-1} d_\lambda\left(\int_\Omega \theta(x,y,\lambda)f(y)dy\right)\\ &= \int_\Omega\left(\int_0^\infty (\lambda - z)^{-1} d_\lambda\theta(x,y,\lambda)\right)f(y)dy\\ &= \int_\Omega G(x,y,z)f(y)dy,\end{aligned}$$

where $G(x,y,z)$ is as in 1). Since

$$\overline{G(x,y,z)} = \int_0^\infty \frac{d\overline{\theta(x,y,\lambda)}}{\lambda - \overline{z}} = \int_0^\infty \frac{d\theta(y,x,\lambda)}{\lambda - \overline{z}} = G(y,x,\overline{z})$$

then 2) is also proved. □

Exercise 38. Prove that $\theta(x,x,\lambda)$ is a monotone increasing function with respect to λ and

1) $|\theta(x,y,\lambda)|^2 \le \theta(x,x,\lambda)\theta(y,y,\lambda)$
2) $|E_\lambda f(x)| \le \theta(x,x,\lambda)^{1/2}\,\|f\|_{L^2(\Omega)}$.

Exercise 39. Prove that

$$|E_\lambda f(x) - E_\mu f(x)| \le \|E_\lambda f - E_\mu f\|_{L^2(\Omega)}\,|\theta(x,x,\lambda) - \theta(x,x,\mu)|^{1/2}$$

for any $\lambda > 0$ and $\mu > 0$.

Exercise 40. Let us assume that $n < m$. Prove that

$$G(x,y,z) = \int_0^\infty \frac{\theta(x,y,\lambda)d\lambda}{(\lambda - z)^2}$$

and that $G(\cdot,y,z) \in L^2(\Omega)$.

In the case of the Friedrichs extension for bounded domain the spectral function $\theta(x,y,\lambda)$ and the Green's function have special form. We know from Corollary 2 of Theorem 1 of the previous chapter that the spectrum $\sigma(A_F)$ is the sequence $\{\lambda_j\}_{j=1}^\infty$ of eigenvalues with only one accumulation point at $+\infty$ and the corresponding orthonormal system $\{\psi_j\}_{j=1}^\infty$ form an orthonormal basis in $L^2(\Omega)$ such that

$$A_F f = \sum_{j=1}^\infty \lambda_j (f,\psi_j)\psi_j \quad \text{in } L^2.$$

This fact implies that

$$\begin{aligned} E_\lambda f &= \sum_{\lambda_j<\lambda} (f,\psi_j)\psi_j = \sum_{\lambda_j<\lambda} \int_\Omega f(y)\overline{\psi_j(y)}dy\psi_j(x) \\ &= \int_\Omega \left(\sum_{\lambda_j<\lambda} \psi_j(x)\overline{\psi_j(y)}\right) f(y)dy = \int_\Omega \theta(x,y,\lambda)f(y)dy \end{aligned}$$

i.e. the spectral function $\theta(x,y,\lambda)$ has the following form

$$\theta(x,y,\lambda) = \sum_{\lambda_j<\lambda} \psi_j(x)\overline{\psi_j(y)}.$$

Hence (see Corollary of Theorem 1 of this chapter) the Green's function has the form

$$G(x,y,z) = \sum_{j=1}^\infty \frac{\psi_j(x)\overline{\psi_j(y)}}{\lambda_j - z} \quad \text{in } L^2.$$

If we assume now that $n < m$ then we may obtain that the Green's function $G(x, y, z)$ is uniformly bounded in $(x, y) \in \Omega \times \Omega$. Let us assume for simplicity that $z = iz_2$ and $A_F \geq I$. Then applying Hörmander's estimate for the spectral function we obtain

$$
\begin{aligned}
|G(x,y,z)| &\leq \sum_{j=1}^{\infty} \frac{|\psi_j(x)||\psi_j(y)|}{\sqrt{\lambda_j^2 + z_2^2}} = \sum_{k=0}^{\infty} \sum_{2^k \leq \lambda_j < 2^{k+1}} \frac{|\psi_j(x)||\psi_j(y)|}{\sqrt{\lambda_j^2 + z_2^2}} \\
&\leq \sum_{k=0}^{\infty} \frac{1}{(2^{2k} + z_2^2)^{1/2}} \left(\sum_{2^k \leq \lambda_j < 2^{k+1}} |\psi_j(x)|^2 \right)^{1/2} \left(\sum_{2^k \leq \lambda_j < 2^{k+1}} |\psi_j(y)|^2 \right)^{1/2} \\
&\leq \sum_{k=0}^{\infty} \frac{(2^{k+1})^{n/m}}{(2^{2k} + z_2^2)^{1/2}}.
\end{aligned}
$$

Since $n < m$ then this series converges for any z_2.

8 Integral operators with weak singularities. Integral equations of the first and second kind.

Let Ω be a bounded domain in $\mathbb{R}^n$. Then

$$Af(x) = \int_\Omega K(x,y)f(y)dy$$

is an *integral operator* in $L^2(\Omega)$ with *kernel* K.

Definition. Integral operator A is said to be *operator with weak singularity* if its kernel $K(x,y)$ is continuous for all $x,y \in \Omega, x \neq y$ and there are positive constants M and $\alpha \in (0,n]$ such that

$$|K(x,y)| \le M|x-y|^{\alpha-n}, \quad x \neq y.$$

Remark. If $K(x,y)$ is continuous for all $x,y \in \Omega$ and bounded then this integral operator is considered also as an operator with weak singularity.

If we have two integral operators A_1 and A_2 with kernels K_1 and K_2, respectively, then we can consider their composition as follows:

$$\begin{aligned}(A_1 \circ A_2)f(x) &= \int_\Omega K_1(x,y)A_2f(y)dy = \int_\Omega K_1(x,y)\left(\int_\Omega K_2(y,z)f(z)dz\right)dy \\ &= \int_\Omega\left(\int_\Omega K_1(x,y)K_2(y,z)dy\right)f(z)dz\end{aligned}$$

and analogously

$$(A_2 \circ A_1)f(x) = \int_\Omega\left(\int_\Omega K_2(x,y)K_1(y,z)dy\right)f(z)dz$$

assuming that the conditions of Fubini theorem are fulfilled.

So, we may conclude that the compositions $A_1 \circ A_2$ and $A_2 \circ A_1$ are again integral operators with the kernels

$$\begin{aligned}K(x,y) &= \int_\Omega K_1(x,z)K_2(z,y)dz, \\ \widetilde{K}(x,y) &= \int_\Omega K_2(x,z)K_1(z,y)dz,\end{aligned} \tag{8.1}$$

respectively. In general, $K(x,y) \neq \widetilde{K}(x,y)$, that is $A_1 \circ A_2 \neq A_2 \circ A_1$.

Returning to the integral operators with weak singularities we obtain their very important property.

Lemma 1. *If A_1 and A_2 are integral operators with weak singularities then $A_1 \circ A_2$ as well as $A_2 \circ A_1$ are also integral operators with weak singularities. Moreover, if*

$$|K_1(x,y)| \le M_1|x-y|^{\alpha_1-n} \quad \text{and} \quad |K_2(x,y)| \le M_2|x-y|^{\alpha_2-n}, \tag{8.2}$$

then there is $M > 0$ such that

$$|K(x,y)| \le M \begin{cases} |x-y|^{\alpha_1+\alpha_2-n}, & \alpha_1+\alpha_2 < n \\ 1+|\log|x-y||, & \alpha_1+\alpha_2 = n \\ 1, & \alpha_1+\alpha_2 > n. \end{cases} \tag{8.3}$$

The same estimates hold for the kernel $\widetilde{K}(x,y)$.

Proof. Using (8.1) and (8.2) we obtain

$$|K(x,y)| \le M_1M_2 \int_\Omega |x-z|^{\alpha_1-n}|z-y|^{\alpha_2-n}dz.$$

If $\alpha_1+\alpha_2 < n$ then changing the variable $z = y + u|x-y|$ we have

$$x - z = |x-y|(e_0 - u), \quad |e_0| = 1$$

and

$$|K(x,y)| \le M_1M_2|x-y|^{\alpha_1+\alpha_2-n} \int_{\mathbb{R}^n} |u-e_0|^{\alpha_1-n}|u|^{\alpha_2-n}du. \tag{8.4}$$

In order to estimate the latter integral we consider three different cases:

$$|u| \le 1/2, \quad 1/2 \le |u| \le 3/2, \quad |u| \ge 3/2.$$

In the first case

$$|u-e_0| \ge |e_0| - |u| \ge 1 - 1/2 = 1/2$$

and therefore

$$\int_{|u|\le 1/2} |u-e_0|^{\alpha_1-n}|u|^{\alpha_2-n}du \le 2^{n-\alpha_1} \int_{|u|\le 1/2} |u|^{\alpha_2-n}du$$
$$= 2^{n-\alpha_1} \int_0^{1/2} r^{\alpha_2-1}dr \int_{\mathbb{S}^{n-1}} d\theta = \frac{2^{n-\alpha_1-\alpha_2}}{\alpha_2}|\mathbb{S}^{n-1}|, \tag{8.5}$$

where $|\mathbb{S}^{n-1}|$ denotes the area of the unit sphere in $\mathbb{R}^n$.

In the third case

$$|u-e_0| \ge |u| - |e_0| \ge |u| - 1 \ge |u| - \frac{2}{3}|u| = \frac{|u|}{3}$$

and we have analogously

$$\int_{|u|\ge 3/2} |u-e_0|^{\alpha_1-n}|u|^{\alpha_2-n}du \le 3^{n-\alpha_1}|\mathbb{S}^{n-1}| \int_{3/2}^\infty r^{\alpha_1+\alpha_2-n-1}dr = \frac{2^{n-\alpha_1-\alpha_2}3^{\alpha_2}}{n-\alpha_1-\alpha_2}|\mathbb{S}^{n-1}|. \tag{8.6}$$

In the case $1/2 \le |u| \le 3/2$ we have that $|u-e_0| \le 5/2$ and so

$$\int_{1/2\le|u|\le3/2} |u-e_0|^{\alpha_1-n}|u|^{\alpha_2-n}du \le 2^{n-\alpha_2}\int_{|u-e_0|\le5/2}|u-e_0|^{\alpha_1-n}du = \frac{2^{n-\alpha_1-\alpha_2}5^{\alpha_1}}{\alpha_1}|\mathbb{S}^{n-1}|. \tag{8.7}$$

Combining (8.4)-(8.7) we obtain (8.3) for the case $\alpha_1+\alpha_2<n$. It can be mentioned here that the estimate (8.3) in this case holds also in the case of any (not necessarily bounded) domain Ω.

If now $\alpha_1+\alpha_2=n$ then the proof of (8.3) will be a little bit different and holds only for bounded domain Ω. Indeed, for any $z\in\Omega$ and

$$|x-z| \le \frac{|x-y|}{2} \quad \text{or} \quad |z-y| \le \frac{|x-y|}{2}$$

we have in both cases that

$$\begin{aligned} |K(x,y)| &\le M_1M_2 2^{n-\alpha_2}|x-y|^{\alpha_2-n}\int_{\Omega'}|x-z|^{\alpha_1-n}dz \\ &\le M_1M_2 2^{n-\alpha_2}|\mathbb{S}^{n-1}||x-y|^{\alpha_2-n}\int_0^{|x-y|/2} r^{\alpha_1-1}dr \\ &= \frac{M_1M_2}{\alpha_1}|\mathbb{S}^{n-1}| \quad \text{or} \quad \frac{M_1M_2}{\alpha_2}|\mathbb{S}^{n-1}|. \end{aligned} \tag{8.8}$$

If $z\in\Omega$ does not belong to these balls with radius $|x-y|/2$ then we consider two cases: $|z-x|\ge|z-y|$ and $|z-x|\le|z-y|$. In both cases we have

$$|K(x,y)| \le M_1M_2\int_{\Omega\setminus\Omega'}\frac{dz}{|x-z|^n} \le M_1M_2|\mathbb{S}^{n-1}|\int_{|x-y|/2}^d \frac{dr}{r} = M_1M_2|\mathbb{S}^{n-1}|\log\frac{2d}{|x-y|}, \tag{8.9}$$

where $d=\operatorname{diam}\Omega$. The estimates (8.8) and (8.9) give us (8.3) in the case $\alpha_1+\alpha_2=n$.

If finally $\alpha_1+\alpha_2>n$ then since Ω is bounded we can analogously obtain (8.3) in this case. This finishes the proof. □

Remark. In the case $\alpha_1+\alpha_2=n$, since for any $0<t<1$

$$|\log t| \le C_\varepsilon t^{-\varepsilon}, \quad \varepsilon>0$$

instead of logarithmic singularity in (8.3) we may consider weak singularity for the kernel $K(x,y)$ as

$$|K(x,y)| \le M_\varepsilon|x-y|^{-\varepsilon},$$

where $\varepsilon>0$ can be chosen appropriately.

Let A be integral operator in $L^2(\Omega)$ with weak singularity. Then since $0<\alpha\le n$ we have

$$\int_\Omega |x-y|^{\alpha-n}dy \le \beta \quad \text{and} \quad \int_\Omega |x-y|^{\alpha-n}dx \le \beta,$$

where

$$\beta = \sup_{x\in\Omega} \int_\Omega |x-y|^{\alpha-n} dy < \infty.$$

Schur test (see Example 2.2) shows that A is bounded in $L^2(\Omega)$ and

$$\|A\|_{L^2(\Omega)\to L^2(\Omega)} \le M\beta.$$

We can prove even more.

Theorem 1. *The integral operator with weak singularity is compact in $L^2(\Omega)$.*

Proof. Let us introduce the function

$$\chi_\sigma(t) = \begin{cases} 1, & 0 \le t \le \sigma \\ 0, & t > \sigma. \end{cases}$$

Then for any $\sigma > 0$ we may write

$$K(x,y) = \chi_\sigma(|x-y|)K(x,y) + (1-\chi_\sigma(|x-y|))K(x,y) := K_1(x,y) + K_2(x,y).$$

The integral operator with kernel $K_2(x,y)$ is Hilbert-Schmidt operator for any $\sigma > 0$ since

$$\int_\Omega \int_\Omega |K_2(x,y)|^2 dxdy \le M^2 \iint_{\sigma\le|x-y|\le d} |x-y|^{2\alpha-2n} dxdy$$

is finite. That's why it is compact in $L^2(\Omega)$ (see Exercise 31). For the integral operator A_1 with kernel $K_1(x,y)$ we proceed as follows:

$$\|A_1 f\|^2_{L^2(\Omega)} = (A_1 f, A_1 f)_{L^2(\Omega)} = (f, A_1^* \circ A_1 f)_{L^2(\Omega)}, \tag{8.10}$$

where A_1^* is adjoint operator with kernel

$$K_1^*(x,y) = \chi_\sigma(|x-y|)\overline{K(y,x)}$$

which is also with weak singularity. Using Lemma 1 we can estimate the right hand side of (8.10) from above as

$$\int_\Omega \int_\Omega |K_\sigma(x,y)||f(x)||f(y)| dxdy \le \frac{1}{2} \int_\Omega \int_\Omega |K_\sigma(x,y)||f(x)|^2 dxdy + \frac{1}{2} \int_\Omega \int_\Omega |K_\sigma(x,y)||f(y)|^2 dxdy, \tag{8.11}$$

where $K_\sigma(x,y)$ is the kernel of weak singularity i.e.

$$|K_\sigma(x,y)| \le M \begin{cases} |x-y|^{2\alpha-n}, & \alpha < n/2 \\ |x-y|^{-\varepsilon}, & \alpha = n/2 \\ 1, & \alpha > n/2 \end{cases}$$

where $\varepsilon > 0$ can be chosen as small as we want.

Let us note also that the definition of $\chi_\sigma(t)$ implies that $K_\sigma(x,y) = 0$ for $|x-y| > 2\sigma$. Thus (see (8.10) and (8.11)) we have ($\alpha < 2n$)

$$\|A_1 f\|^2_{L^2(\Omega)} \le M \iint_{|x-y|\le 2\sigma} |x-y|^{2\alpha-n}|f(x)|^2 dxdy$$
$$\le M \int_\Omega |f(x)|^2 \int_{|x-y|\le 2\sigma} |x-y|^{2\alpha-n} dydx = M \|f\|^2_{L^2(\Omega)} \frac{(2\sigma)^{2\alpha}}{2\alpha} |\mathbb{S}^{n-1}| \to 0,$$

as $\sigma \to 0$. It means that

$$\|A_1\|_{L^2(\Omega)\to L^2(\Omega)} \to 0, \quad \sigma \to 0.$$

The same fact is valid for the cases $\alpha \ge n/2$. Thus,

$$\|A - A_2\|_{L^2(\Omega)\to L^2(\Omega)} \le \|A_1\|_{L^2(\Omega)\to L^2(\Omega)} \to 0$$

as $\sigma \to 0$. But A_2 is compact for any $\sigma > 0$, therefore A is also compact as the limit of compact operators. This completes the proof. □

We want now to expand the analysis of the integral operators with weak singularity defined on domains in $\mathbb{R}^n$ to the integral operators with weak singularity defined on surfaces of dimension $n-1$.

Assume that $\partial\Omega$ is the boundary of a bounded domain of class C^1. It means, roughly speaking, that at any point $x \in \partial\Omega$ there is a tangent plane with normal vector $\nu(x)$ which is continuous function on $\partial\Omega$ and surface differential $d\sigma(y)$ in the neighborhood of each point $x \in \partial\Omega$ satisfies the inequality (see Kress book)

$$d\sigma(y) \le c_0 \rho^{n-2} d\rho d\theta,$$

where (ρ, θ) are the polar coordinates in the tangent plane with origin x and c_0 is independent of x. According to the dimension $n-1$ of the surface $\partial\Omega$, an integral operator in $L^2(\partial\Omega)$ i.e.

$$Af(x) = \int_{\partial\Omega} K(x,y) f(y) d\sigma(y)$$

is said to be with weak singularity if its kernel $K(x,y)$ is continuous for all $x, y \in \partial\Omega, x \ne y$ and there are constants $M > 0$ and $\alpha \in (0, n-1]$ such that

$$|K(x,y)| \le M|x-y|^{\alpha-(n-1)}, \quad x \ne y.$$

If $K(x,y)$ is continuous everywhere we require that K is bounded on $\partial\Omega \times \partial\Omega$. We can provide now the following theorem.

Theorem 2. *The integral operator with weak singularity is compact in $L^2(\partial\Omega)$.*

Proof. The same as for Theorem 1. □

There is very useful and quite general result for the integral operators with weak singularity both for domains and surfaces in $\mathbb{R}^n$.

Theorem 3. *The integral operator with weak singularity transforms bounded functions into continuous functions.*

Proof. We give the proof for the domains in $\mathbb{R}^n$. The proof for surfaces in $\mathbb{R}^n$ is the same. Let $x, y \in \Omega$ and $|x-y| < \delta$. Then

$$\begin{aligned} |Af(x) - Af(y)| &\le \int_{|x-z|<2\delta} (|K(x,z)| + |K(y,z)|)|f(z)|dz \\ &+ \int_{\Omega\setminus\{|x-z|<2\delta\}} |K(x,z) - K(y,z)||f(z)|dz \\ &\le M \|f\|_{L^\infty(\Omega)} \int_{|x-z|<2\delta} (|x-z|^{\alpha-n} + |y-z|^{\alpha-n})dz \\ &+ \|f\|_{L^\infty(\Omega)} \int_{\Omega\setminus\{|x-z|<2\delta\}} |K(x,z) - K(y,z)|dz := I_1 + I_2. \end{aligned}$$

Since $|z-y| \le |x-z| + |x-y|$ then

$$I_1 \le 2M \|f\|_{L^\infty(\Omega)} |\mathbb{S}^{n-1}| \int_0^{3\delta} r^{\alpha-1}dr = 2M|\mathbb{S}^{n-1}| \|f\|_{L^\infty(\Omega)} \frac{(3\delta)^\alpha}{\alpha} \to 0, \quad \delta \to 0.$$

On the other hand for $|x-y| < \delta$ and $|x-z| \ge 2\delta$ we have that

$$|y-z| \ge |x-z| - |x-y| > 2\delta - \delta = \delta.$$

So the continuity of the kernel K outside of the diagonal implies that

$$K(x,z) - K(y,z) \to 0, \quad \delta \to 0,$$

uniformly in $z \in \Omega \setminus \{|x-z| < 2\delta\}$. Since Ω is bounded we obtain that $I_2 \to 0$ as $\delta \to 0$. This finishes the proof. □

Exercise 41. Prove that if A is as in Theorem 3 then $f(x) + Af(x) \in C(\overline{\Omega})$ for $f \in L^2(\Omega)$ implies $f \in C(\overline{\Omega})$.

We are now in the position to extend the solvability conditions (Fredholm alternative, see Theorem 6 of Chapter 4) to the equations in the Hilbert space with compact but not necessarily self-adjoint operators.

Theorem 4 (Fredholm alternative II)**.** *Suppose $A : H \to H$ is compact. For any $\mu \in \mathbb{C}$ either the equations*

$$(I - \mu A)f = g, \quad (I - \overline{\mu}A^*)f' = g' \tag{8.12}$$

have the unique solutions f and f' for any given g and g' from H or the corresponding homogeneous equations

$$(I-\mu A)f=0, \quad (I-\overline{\mu}A^*)f'=0 \tag{8.13}$$

have nontrivial solutions such that

$$\dim N(I-\mu A)=\dim N(I-\overline{\mu}A^*)<\infty$$

and in this case equations (8.12) have the solutions if and only if

$$\begin{aligned} g\perp N(I-\overline{\mu}A^*) \quad &\Leftrightarrow \quad g\in R(I-\mu A) \\ g'\perp N(I-\mu A) \quad &\Leftrightarrow \quad g'\in R(I-\overline{\mu}A^*) \end{aligned}$$

respectively.

Proof. Lemma of Riesz (see Theorem 4 of Chapter 4) and Exercise 13 give

$$\begin{aligned} R(I-\mu A)&=N(I-\overline{\mu}A^*)^\perp \\ R(I-\overline{\mu}A^*)&=N(I-\mu A)^\perp. \end{aligned}$$

Let us first prove that always

$$\dim N(I-\mu A)=\dim N(I-\overline{\mu}A^*).$$

These two dimensions are finite due to Riesz (see Proposition 1 of Chapter 4). Since every compact operator is a norm limit of a sequence of operators of finite rank (see Chapter 4 for details), for any $\mu\in\mathbb{C},\mu\neq 0$ we have

$$I-\mu A=-\mu A_0+(I-\mu A_1),$$

where A_0 is of finite rank and $\|\mu A_1\|<1$. Then $(I-\mu A_1)^{-1}$ exists and

$$(I-\mu A_1)^{-1}(I-\mu A)=I-\mu(I-\mu A_1)^{-1}A_0:=I-A_2,$$

where A_2 is of finite rank too. Analogously, since $(I-\overline{\mu}A_1^*)^{-1}$ exists then

$$(I-\overline{\mu}A^*)(I-\overline{\mu}A_1^*)^{-1}=I-\overline{\mu}A_0^*(I-\overline{\mu}A_1^*)^{-1}:=I-A_2^*,$$

where A_2^* is adjoint to A_2 and is of finite rank too. These representations allow us to conclude that

$$\begin{aligned} g\in N(I-\mu A) &\Leftrightarrow g\in N(I-A_2) \\ g'\in N(I-\mu A_2^*) &\Leftrightarrow (I-\overline{\mu}A_1^*)^{-1}g'\in N(I-\overline{\mu}A^*) \end{aligned}$$

Thus, it suffices to show that the number of independent solutions of the equations

$$g=A_2g, \quad g'=A_2^*g'$$

are equal.

Since we know that the ranks of A_2 and A_2^* are finite we may represent the mappings of the operators $I - A_2$ and $I - A_2^*$ as the mappings of matrices $I - M_2$ and $I - M_2^*$ with adjoint matrices M_2 and M_2^*. But the ranks of the adjoint matrices are equal and therefore the number of independent solutions of the equations $g = A_2 g$ and $g' = A_2^* g'$ are equal.

The next step is: if $R(I - \mu A) = H$ then $N(I - \overline{\mu} A^*) = \{0\}$ and consequently $N(I - \mu A) = \{0\}$ and $R(I - \overline{\mu} A^*) = H$ (see Exercise 13). This means that both $(I - \mu A)^{-1}$ and $(I - \overline{\mu} A^*)^{-1}$ exist and the unique solutions of (8.12) are given by

$$f = (I - \mu A)^{-1} g, \quad f' = (I - \overline{\mu} A^*)^{-1} g'.$$

If $N(I - \mu A)$ and $N(I - \overline{\mu} A^*)$ are not zero then $R(I - \mu A)$ and $R(I - \overline{\mu} A^*)$ are proper subspaces of H and equations (8.12) have the solutions if and only if

$$g \in R(I - \mu A), \quad g' \in R(I - \overline{\mu} A^*).$$

It is equivalent (see Exercise 13) to

$$g \perp N(I - \overline{\mu} A^*), \quad g' \perp N(I - \mu A).$$

□

We will demonstrate now this Fredholm alternative for the integral operators. Let $\Omega \subset \mathbb{R}^n$ be any domain and let

$$Af(x) = \int_\Omega K(x,y) f(y) dy$$

be a compact integral operator in $L^2(\Omega)$. Then its adjoint is defined as

$$A^* f(x) = \int_\Omega \overline{K(y,x)} f(y) dy.$$

Hence, the Fredholm alternative for these operators reads as: either the equations

$$\begin{aligned} f(x) - \mu \int_\Omega K(x,y) f(y) dy &= g(x) \\ f'(x) - \overline{\mu} \int_\Omega \overline{K(y,x)} f'(y) dy &= g'(x) \end{aligned} \tag{8.14}$$

are uniquely solvable for any g and g' from $L^2(\Omega)$ or the equations

$$\begin{aligned} f(x) &= \mu \int_\Omega K(x,y) f(y) dy \\ f'(x) &= \overline{\mu} \int_\Omega \overline{K(y,x)} f'(y) dy \end{aligned} \tag{8.15}$$

have the same (finite) number of linearly independent solutions. And in this case equations (8.14) are solvable if and only if g and g' are orthogonal to any solution f and f' of the equations (8.15), respectively.

Definition. Equations (8.14) and (8.15) are called the *integral equations of second and first kind*, respectively.

Exercise 42. Consider in $L^2(a,b)$ the integral equation

$$\varphi(x) - \int_a^b e^{x-y}\varphi(y)dy = f(x), \quad x \in [a,b],$$

where $f \in L^2(a,b)$. Solve this equation and formulate the Fredholm alternative for it.

Example 8.1 (Boundary value problems)**.** Consider the ordinary differential equation of the second order

$$a_0(x)u''(x) + a_1(x)u'(x) + a_2(x)u(x) = f(x)$$

on the interval $[0,1]$ with coefficients $f, a_2 \in L^2(0,1), a_1 \in W_2^1(0,1)$ and with smooth $a_0(x) \geq c_0 > 0$ subject to the boundary conditions

$$u(0) = u_0, \quad u(1) = u_1.$$

Dividing this equation by $a_0(x)$ we may consider the boundary value problem in the form

$$u'' + a_1(x)u' + a_2(x)u = f, \quad u(0) = u_0, u(1) = u_1.$$

Using Green's function $G(x,y)$ of the form

$$G(x,y) = \begin{cases} y(1-x), & 0 \leq y \leq x \leq 1 \\ x(1-y), & 0 \leq x \leq y \leq 1 \end{cases}$$

we can rewrite this boundary value problem as

$$u(x) = \varphi_0(x) + \int_0^1 G(x,y)(a_1(y)u' + a_2(y)u - f(y))dy,$$

where $\varphi_0(x) = u_0(1-x) + u_1 x$. Integration by parts implies

$$u(x) = \varphi_0(x) - \int_0^1 G(x,y)f(y)dy + G(x,y)a_1(y)u(y)|_0^1$$
$$- \int_0^1 \left[\partial_y G(x,y)a_1(y) + G(x,y)a_1'(y)\right] u(y)dy + \int_0^1 G(x,y)a_2(y)u(y)dy.$$

Since $G(x,1) = G(x,0) = 0$ then this equation can be rewritten as

$$u(x) = \widetilde{\varphi_0}(x) - \int_0^1 K(x,y)u(y)dy,$$

where

$$\widetilde{\varphi_0}(x) = \varphi_0(x) - \int_0^1 G(x,y)f(y)dy$$

and

$$K(x,y) = \partial_y G(x,y)a_1(y) + G(x,y)a_1'(y) - G(x,y)a_2(y).$$

Exercise 43. 1) Prove that $K(x,y)$ is a Hilbert-Schmidt kernel on $[0,1] \times [0,1]$.

2) Prove that the boundary value problem and this integral equation of the second kind are equivalent.

3) Formulate the solvability condition for the boundary value problem using the Fredholm alternative for this integral operator.

9 Volterra and singular integral equations

In this chapter we consider integral equations of special types on the finite interval $[a,b]$. We consider normed space $L^\infty(a,b)$, the Lebesgue space, and $C^\alpha[a,b]$, the Hölder space (which are not Hilbert spaces) instead of the Hilbert space L^2. The norms of the spaces $L^\infty(a,b)$ and $C^\alpha[a,b]$ are defined as follows:

$$\|f\|_{L^\infty(a,b)} = \inf\{M : |f(x)| \le M \text{ a.e. on } (a,b)\} \tag{9.1}$$

$$\|f\|_{C^\alpha[a,b]} = \|f\|_{L^\infty(a,b)} + \sup_{x,y\in[a,b]} \frac{|f(x)-f(y)|}{|x-y|^\alpha}, \tag{9.2}$$

where $0 < \alpha \le 1$.

The fact that f belongs to the Hölder space $C^\alpha[a,b]$ is equivalent to the fact that $f \in L^\infty(a,b)$ and there is constant $c_0 > 0$ such that for all h (small enough)

$$|f(x+h) - f(x)| \le c_0|h|^\alpha,$$

where $x, x+h \in [a,b]$.

Definition. Integral equations in $L^\infty(a,b)$ of the form

$$f(x) = \int_a^x K(x,y)\varphi(y)dy$$

and

$$\varphi(x) = f(x) + \int_a^x K(x,y)\varphi(y)dy, \tag{9.3}$$

where $x \in [a,b]$ and $\sup_{x,y\in[a,b]} |K(x.y)| < \infty$, are called *Volterra integral equations of the first and second kind*, respectively.

Theorem 1. *For each $f \in L^\infty(a,b)$ the Volterra integral equation of the second kind has a unique solution $\varphi \in L^\infty(a,b)$ such that*

$$|\varphi(x)| \le e^{M(x-a)} \|f\|_{L^\infty(a,b)} \tag{9.4}$$

for any $x \in [a,b]$ and

$$\|\varphi\|_{L^\infty(a,b)} \le \|f\|_{L^\infty(a,b)} e^{M(b-a)}, \tag{9.5}$$

where $M = \sup_{x,y\in[a,b]} |K(x,y)|$.

Proof. We introduce the iterations of the equation (9.3) by

$$\varphi_{j+1}(x) := \int_a^x K(x,y)\varphi_j(y)dy, \quad j = 0,1,2,\ldots$$

with $\varphi_0 = f$. Let us prove by induction that

$$|\varphi_j(x)| \le \frac{(M(x-a))^j}{j!} \|f\|_{L^\infty(a,b)}, \quad j = 0,1,\ldots. \tag{9.6}$$

Indeed, this estimate clearly holds for $j = 0$. Assume that (9.6) is proved for some $j \geq 0$. Then

$$|\varphi_{j+1}(x)| \leq \int_a^x |K(x,y)||\varphi_j(y)|dy \leq M \int_a^x \frac{(M(y-a))^j}{j!} \|f\|_{L^\infty(a,b)}\, dy$$
$$= M^{j+1} \|f\|_{L^\infty(a,b)} \int_a^x \frac{(y-a)^j}{j!} dy = M^{j+1} \|f\|_{L^\infty(a,b)} \frac{(x-a)^{j+1}}{(j+1)!}.$$

This proves (9.6).

Let us introduce the function

$$\varphi(x) := \sum_{j=0}^{\infty} \varphi_j(x). \tag{9.7}$$

Then, from (9.6) we obtain for all $x \in [a, b]$ that

$$|\varphi(x)| \leq \|f\|_{L^\infty(a,b)} \sum_{j=0}^{\infty} \frac{(M(x-a))^j}{j!} = \|f\|_{L^\infty(a,b)}\, e^{M(x-a)}.$$

Thus, the function $\varphi(x)$ is well-defined by the series (9.7) since this series is uniformly convergent with respect to $x \in [a, b]$.

It remains now to show that this $\varphi(x)$ solves (9.3). Since the series (9.7) converges uniformly we may integrate it term by term and obtain

$$\int_a^x K(x,y)\varphi(y)dy = \sum_{j=0}^{\infty} \int_a^x K(x,y)\varphi_j(y)dy = \sum_{j=0}^{\infty} \varphi_{j+1}(x)$$
$$= \sum_{j=1}^{\infty} \varphi_j(x) + \varphi_0(x) - f(x) = \varphi(x) - f(x).$$

So (9.3) holds with this φ. The estimate (9.5) then follows from (9.4) immediately. Finally, the uniqueness of this solution follows from (9.5) too. □

Corollary. *The homogeneous equation*

$$\varphi(x) = \int_a^x K(x,y)\varphi(y)dy$$

has only the trivial solution in $L^\infty(a, b)$.

Proof. Follows from (9.5). □

In general, integral equations of the first kind are more delicate with respect to the solvability than equations of the second kind. However, in some cases Volterra integral equations of the first kind can be treated by reducing them to equations of the second kind. Indeed, consider for $x \in [a, b]$

$$\int_a^x K(x,y)\varphi(y)dy = f(x) \tag{9.8}$$

and assume that the derivatives $\frac{\partial K}{\partial x}(x, y)$ and $f'(x)$ exist and are bounded and that $K(x, x) \neq 0$ for all $x \in [a, b]$. Then, differentiating with respect to x reduces (9.8) to

$$\varphi(x)K(x,x) + \int_a^x \frac{\partial K}{\partial x}(x,y)\varphi(y)dy = f'(x)$$

or

$$\varphi(x) = \frac{f'(x)}{K(x,x)} - \int_a^x \frac{\frac{\partial K}{\partial x}(x,y)}{K(x,x)}\varphi(y)dy. \tag{9.9}$$

Exercise 44. Show that (9.8) and (9.9) are equivalent if $f(a) = 0$.

The second possibility occurs if we assume that $\frac{\partial K}{\partial y}(x, y)$ exists and is bounded and that $K(x, x) \neq 0$ for all $x \in [a, b]$. In this case, setting

$$\psi(x) := \int_a^x \varphi(y)dy, \quad \psi' = \varphi$$

and performing integration by parts in (9.8) yields

$$\begin{aligned} f(x) = \int_a^x K(x,y)\psi'(y)dy &= K(x,y)\psi(y)|_a^x - \int_a^x \frac{\partial K}{\partial y}(x,y)\psi(y)dy \\ &= K(x,x)\psi(x) - \int_a^x \frac{\partial K}{\partial y}(x,y)\psi(y)dy \end{aligned}$$

or

$$\psi(x) = \frac{f(x)}{K(x,x)} + \int_a^x \frac{\frac{\partial K}{\partial y}(x,y)}{K(x,x)}\psi(y)dy.$$

There is an interesting generalization of equation (9.3) when the kernel has weak singularities. More precisely, we consider (9.3) in the space $L^\infty(a, b)$ and assume that the kernel $K(x, y)$ satisfies the estimate

$$|K(x,y)| \le M|x-y|^{-\alpha}, \quad x, y \in [a,b], x \neq y$$

with some $0 < \alpha < 1$. If we consider again the iterations

$$\varphi_j(x) := \int_a^x K(x,y)\varphi_{j-1}(y)dy, \quad j = 1, 2, \ldots$$

with $\varphi_0 = f$, then it can be proved by induction that for all $x \in [a, b]$ we have

$$|\varphi_j(x)| \le \left(\frac{M(x-a)^{1-\alpha}}{1-\alpha}\right)^j \|f\|_{L^\infty(a,b)}, \quad j = 0, 1, \ldots. \tag{9.10}$$

Indeed, since this clearly holds for $j = 0$ assume that it is proved for some $j \geq 0$. Then

$$
\begin{aligned}
|\varphi_{j+1}(x)| &\leq \int_a^x |K(x,y)||\varphi_j(y)|dy \leq M \frac{M^j}{(1-\alpha)^j} \int_a^x |x-y|^{-\alpha}((y-a)^{1-\alpha})^j \|f\|_{L^\infty(a,b)} dy \\
&\leq \frac{M^{j+1}}{(1-\alpha)^j}((x-a)^{1-\alpha})^j \|f\|_{L^\infty(a,b)} \int_a^x (x-y)^{-\alpha} dy \\
&\leq \frac{M^{j+1}}{(1-\alpha)^j}((x-a)^{1-\alpha})^j \|f\|_{L^\infty(a,b)} \frac{(x-a)^{1-\alpha}}{1-\alpha} \\
&= \left(\frac{M(x-a)^{1-\alpha}}{1-\alpha}\right)^{j+1} \|f\|_{L^\infty(a,b)}.
\end{aligned}
$$

If we assume now that

$$\frac{M(b-a)^{1-\alpha}}{1-\alpha} < 1$$

then the series

$$\sum_{j=0}^{\infty} \varphi_j(x)$$

converges uniformly on the interval $[a, b]$ and the function φ defined by

$$\varphi(x) := \sum_{j=0}^{\infty} \varphi_j(x)$$

solves therefore the nonhomogeneous integral equation (9.3). Moreover, the following estimates hold

$$|\varphi(x)| \leq \frac{\|f\|_{L^\infty(a,b)}}{1 - \dfrac{M(x-a)^{1-\alpha}}{1-\alpha}}, \quad x \in [a, b]$$

and

$$\|\varphi\|_{L^\infty(a,b)} \leq \frac{\|f\|_{L^\infty(a,b)}}{1 - \dfrac{M(b-a)^{1-\alpha}}{1-\alpha}}.$$

Exercise 45. Show that the Volterra integral equation of the first kind

$$\varphi(x) = \lambda \int_0^x e^{-(x-y)} \varphi(y) dy$$

has, for any λ, only the trivial solution in $L^\infty(a, b)$.

Definition. Let $0 < \alpha < 1, \varphi \in C^\alpha[-a, a]$ and periodic, i.e. $\varphi(-a) = \varphi(a)$. Integral equation in this space of the form

$$\varphi(x) = f(x) + \lambda \text{p.v.} \int_{-a}^{a} \frac{\varphi(x+y)dy}{y}, \quad \lambda \in \mathbb{C} \tag{9.11}$$

is understood in the sense that

$$\text{p.v.} \int_{-a}^{a} \frac{\varphi(x+y)dy}{y} = \lim_{\varepsilon\to 0} \int_{|y|\geq\varepsilon, y\in[-a,a]} \frac{\varphi(x+y)dy}{y} \tag{9.12}$$

and the function φ is extended periodically (with period $2a$) on the whole line.

Due to (9.12) we have that

$$\text{p.v.} \int_{-a}^{a} \frac{dy}{y} = 0.$$

Thus

$$\text{p.v.} \int_{-a}^{a} \frac{\varphi(x+y)dy}{y} = \text{p.v.} \int_{-a}^{a} \frac{\varphi(x+y)-\varphi(x)}{y}dy = \int_{-a}^{a} \frac{\varphi(x+y)-\varphi(x)}{y}dy$$

and the latter integral can be understood in the usual sense for periodic $\varphi \in C^\alpha[-a,a]$ since

$$\left|\int_{-a}^{a} \frac{\varphi(x+y)-\varphi(x)}{y}dy\right| \leq c_0 \int_{-a}^{a} \frac{|y|^\alpha}{|y|}dy = 2c_0 \int_0^a \xi^{\alpha-1}d\xi = 2c_0\frac{a^\alpha}{\alpha}. \tag{9.13}$$

The inequality (9.13) shows us that for any $\varphi \in C^\alpha[-a,a]$ the integral in (9.11) is uniformly bounded and also periodic with period $2a$. But even more is true.

Proposition 1. *For any $2a$-periodic $\varphi \in C^\alpha[-a,a]$ with $0 < \alpha < 1$ the integral in (9.11) defines a $2a$-periodic function of x which belongs to the same Hölder space $C^\alpha[-a,a]$.*

Proof. Let us denote by $g(x)$ the integral in (9.11). For $|h| > 0$ small enough we have

$$\begin{aligned} g(x+h)-g(x) &= \int_{-a}^{a} \frac{\varphi(x+h+y)-\varphi(x+h)}{y}dy - \int_{-a}^{a} \frac{\varphi(x+y)-\varphi(x)}{y}dy \\ &= \int_{|y|\leq 3|h|} \frac{\varphi(x+h+y)-\varphi(x+h)}{y}dy - \int_{|y|\leq 3|h|} \frac{\varphi(x+y)-\varphi(x)}{y}dy \\ &+ \int_{|y|\geq 3|h|} \frac{\varphi(x+h+y)-\varphi(x)}{y}dy - \int_{|y|\geq 3|h|} \frac{\varphi(x+y)-\varphi(x)}{y}dy \\ &:= I_1 + I_2. \end{aligned}$$

For the first integral I_1 we have

$$\begin{aligned} |I_1| &\leq \int_{|y|\leq 3|h|} \frac{|\varphi(x+y+h)-\varphi(x+h)|}{|y|}dy + \int_{|y|\leq 3|h|} \frac{|\varphi(x+y)-\varphi(x)|}{|y|}dy \\ &\leq c_0 \int_{|y|\leq 3|h|} \frac{|y|^\alpha}{|y|}dy + c_0 \int_{|y|\leq 3|h|} \frac{|y|^\alpha}{|y|}dy \\ &\leq 4c_0 \int_0^{3|h|} \xi^{\alpha-1}d\xi = 4c_0\frac{(3|h|)^\alpha}{\alpha} = \frac{4c_0 3^\alpha}{\alpha}|h|^\alpha. \end{aligned} \tag{9.14}$$

For the estimation of I_2 we first rewrite it as (we change variables in the first integral)

$$\begin{aligned} I_2 &= \int_{|z-h|\geq 3|h|} \frac{\varphi(z+x)-\varphi(x)}{z-h}dz - \int_{|z|\geq 3|h|} \frac{\varphi(z+x)-\varphi(x)}{z}dz \\ &= \int_{|z|\geq 3|h|} (\varphi(z+x)-\varphi(x)) \left[\frac{1}{z-h} - \frac{1}{z}\right] dz \\ &- \int_{\{|z-h|\geq 3|h|\}\setminus\{|z|\geq 3|h|\}} \frac{\varphi(z+x)-\varphi(x)}{z-h}dz. \end{aligned}$$

Then we have

$$\begin{aligned} |I_2| &\leq \int_{|z|\geq 3|h|, z\in[-a,a]} |\varphi(z+x)-\varphi(x)| \frac{|h|dz}{|z|\cdot|z-h|} + \int_{2|h|\leq|z|\leq 3|h|} \frac{|\varphi(z+x)-\varphi(x)|}{|z-h|}dz \\ &\leq c_0|h| \int_{a\geq|z|\geq 3|h|} \frac{|z|^\alpha}{|z|\cdot 2|z|/3}dz + c_0 \int_{2|h|\leq|z|\leq 3|h|} \frac{|z|^\alpha}{|z|/2}dz \\ &= 2\cdot\frac{3c_0}{2}|h| \int_{3|h|}^{a} \xi^{\alpha-2}d\xi + 4c_0 \int_0^{3|h|} \xi^{\alpha-1}d\xi \\ &= 3c_0|h| \left.\frac{\xi^{\alpha-1}}{\alpha-1}\right|_{3|h|}^{a} + 4c_0\frac{(3|h|)^\alpha}{\alpha} = 3c_0|h|\left(\frac{(3|h|)^{\alpha-1}}{1-\alpha} - \frac{a^{\alpha-1}}{1-\alpha}\right) + \frac{4c_0 3^\alpha}{\alpha}|h|^\alpha \\ &< \frac{3^\alpha c_0}{1-\alpha}|h|^\alpha + \frac{4c_0 3^\alpha}{\alpha}|h|^\alpha = c_0 3^\alpha\left(\frac{1}{1-\alpha}+\frac{4}{\alpha}\right)|h|^\alpha, \end{aligned} \tag{9.15}$$

since $0 < \alpha < 1$. Estimates (9.14)-(9.15) show that this Proposition is completely proved. □

If we denote by

$$A\varphi(x) := \text{p.v.} \int_{-a}^{a} \frac{\varphi(x+y)dy}{y} \tag{9.16}$$

a linear operator on periodic $C^\alpha[-a,a], 0 < \alpha < 1$ then Proposition 1 gives that A is bounded in this space. But this operator is not compact there. Nevertheless the following holds.

Corollary. *There is $\lambda_0 > 0$ such that for all $\lambda \in \mathbb{C}, |\lambda| < \lambda_0$ and for any periodic $f \in C^\alpha[-a,a], 0 < \alpha < 1$ the integral equation (9.11) has a unique solution in $C^\alpha[-a,a], 0 < \alpha < 1$.*

Proof. Since operator A from (9.16) is a bounded linear operator in the space $C^\alpha[-a,a]$ then

$$\|A\|_{C^\alpha\to C^\alpha} \leq c_0$$

with some constant $c_0 > 0$. If we choose now $\lambda_0 = 1/c_0$ then for all $\lambda \in \mathbb{C}, |\lambda| < \lambda_0$, operator $I - \lambda A$ will be invertible in the space $C^\alpha[-a,a]$ since

$$\|\lambda A\|_{C^\alpha\to C^\alpha} < 1.$$

This fact implies that the integral equation (9.11) can be solved uniquely in this space and the unique solution φ can be obtained as

$$\varphi = (I - \lambda A)^{-1} f.$$

It is equivalent to the fact that (9.11) can be solved by iterations. $\square$

10 Approximate methods

In this chapter we will study approximate solution methods for equations in the Hilbert space H of the form

$$A\varphi = f, \quad (I - A)\varphi = f \tag{10.1}$$

with bounded or compact operator A. The fundamental concept for approximately solving equations (10.1) is to replace them by the equations

$$A_n\varphi_n = f_n, \quad (I - A_n)\varphi_n = f_n, \tag{10.2}$$

respectively. For practical purposes, the approximating equations (10.2) will be chosen so that they can be reduced to solving a finite-dimensional linear system.

We will start with some general results which are the basis of our considerations.

Theorem 1. *Let $A : H \to H$ be a bounded linear operator with bounded inverse A^{-1}. Assume that the sequence $A_n : H \to H$ of bounded linear operators is norm convergent to A i.e.*

$$\|A_n - A\| \to 0, \quad n \to \infty.$$

Then for all n such that

$$\big\|A^{-1}(A_n - A)\big\| < 1$$

the inverse operators A_n^{-1} exist and

$$\big\|A_n^{-1}\big\| \le \frac{\|A^{-1}\|}{1 - \|A^{-1}(A_n - A)\|}.$$

Moreover, the solutions of (10.1) and (10.2) satisfy the error estimate

$$\|\varphi_n - \varphi\| \le \frac{\|A^{-1}\|}{1 - \|A^{-1}(A_n - A)\|}\Big(\|(A_n - A)\varphi\| + \|f_n - f\|\Big).$$

Proof. Since A^{-1} exists we may write

$$A^{-1}A_n = I - A^{-1}(A - A_n).$$

Since

$$\big\|A^{-1}(A_n - A)\big\| < 1$$

for n large enough, for these values of n we have that

$$\big(I - A^{-1}(A - A_n)\big)^{-1}$$

exists by the Neumann series. Thus,

$$\big(A^{-1}A_n\big)^{-1} = \big(I - A^{-1}(A - A_n)\big)^{-1}$$

or

$$A_n^{-1}A = \big(I - A^{-1}(A - A_n)\big)^{-1}$$

or

$$A_n^{-1} = \left(I - A^{-1}(A - A_n)\right)^{-1} A^{-1}.$$

The error estimate follows immediately from the representation

$$\varphi_n - \varphi = A_n^{-1}(A - A_n)\varphi + A_n^{-1}(f_n - f).$$

□

Theorem 2. *Assume that $A_n^{-1} : H \to H$ exist for all $n \geq n_0$ and their norms are uniformly bounded for such n. Let $\|A_n - A\| \to 0$ as $n \to \infty$. Then the inverse operator A^{-1} exists and*

$$\left\|A^{-1}\right\| \leq \frac{\|A_n^{-1}\|}{1 - \|A_n^{-1}(A_n - A)\|}$$

for all $n \geq n_0$ with $\|A_n^{-1}(A_n - A)\| < 1$.

Exercise 46. Prove Theorem 2 and obtain the error estimate in this case.

Definition. A sequence $\{A_n\}_{n=1}^{\infty}$ of compact operators in the Hilbert space H is called *collectively compact* if for any bounded set $U \subset H$ the image

$$J = \{A_n\varphi : \varphi \in U, n = 1, 2, \ldots\}$$

is *relatively compact* i.e. every sequence from J contains a convergent subsequence.

Exercise 47. Assume that a sequence of compact operators $\{A_n\}_{n=1}^{\infty}$ is collectively compact and converges pointwise to A in H i.e.

$$\lim_{n\to\infty} A_n\varphi = A\varphi, \quad \varphi \in H.$$

Prove that the limit operator A is compact.

Exercise 48. Under the same assumptions for $\{A_n\}_{n=1}^{\infty}$ as in Exercise 47 prove that

$$\|(A_n - A)A\| \to 0, \quad \|(A_n - A)A_n\| \to 0$$

as $n \to \infty$.

Theorem 3. *Let $A : H \to H$ be a compact operator and let $I - A$ be injective. Assume that the sequence $A_n : H \to H$ is collectively compact and pointwise convergent to A. Then for all n such that*

$$\left\|(I - A)^{-1}(A_n - A)A_n\right\| < 1$$

the inverse operators $(I - A_n)^{-1}$ exist and the solutions of (10.1) and (10.2) satisfy the error estimate

$$\|\varphi_n - \varphi\| \leq \frac{1 + \|(I - A)^{-1}A_n\|}{1 - \|(I - A)^{-1}(A_n - A)A_n\|}\Big(\|(A_n - A)\varphi\| + \|f_n - f\|\Big).$$

Proof. By the Riesz theorem (see Theorem 5 of Chapter 4) the inverse operator $(I - A)^{-1}$ exists and is bounded. Due to Exercise 48

$$\|(A_n - A)A_n\| \to 0, \quad n \to \infty.$$

That's why for n large enough we have

$$\left\|(I - A)^{-1}(A_n - A)A_n\right\| < 1.$$

This fact allows us to conclude (as in Theorem 1) that $(I - A_n)^{-1}$ exists and

$$\left\|(I - A_n)^{-1}\right\| \le \frac{1 + \|(I - A)^{-1}A_n\|}{1 - \|(I - A)^{-1}(A_n - A)A_n\|}.$$

The error esimate follows from this inequality and the representation

$$\varphi_n - \varphi = (I - A_n)^{-1}\big((A_n - A)\varphi + f_n - f\big).$$

□

Corollary. *Let A_n be as in Theorem 3. Assume that the inverse operators $(I - A_n)^{-1}$ exist and are uniformly bounded for all $n \ge n_0$. Then the inverse $(I - A)^{-1}$ exists if*

$$\left\|(I - A_n)^{-1}(A_n - A)A\right\| < 1.$$

The solutions of (10.1) and (10.2) satisfy the error estimate

$$\|\varphi_n - \varphi\| \le \frac{1 + \|(I - A_n)^{-1}A\|}{1 - \|(I - A_n)^{-1}(A_n - A)A\|}\Big(\|(A_n - A)\varphi\| + \|f_n - f\|\Big).$$

Theorem 4. *Let $A : H \to H$ be a bounded linear operator with $\|A\| < 1$. Then the* successive approximations

$$\varphi_{n+1} := A\varphi_n + f, \quad n = 0, 1, \ldots \tag{10.3}$$

converge for each $f \in H$ and each $\varphi_0 \in H$ to the unique solution of (10.1).

Proof. The condition $\|A\| < 1$ implies the existence and boundedness of the inverse operator $(I - A)^{-1}$ and the existence of the unique solution of (10.1) as

$$\varphi = (I - A)^{-1}f.$$

It remains only to show that the successive approximations converge to φ for any $\varphi_0 \in H$. The definition (10.3) implies

$$\|\varphi_{n+1} - \varphi_n\| \le \|A\| \, \|\varphi_n - \varphi_{n-1}\| \le \cdots \le \|A\|^n \, \|\varphi_1 - \varphi_0\| .$$

Hence for each $p \in \mathbb{N}$ we have

$$\begin{aligned}\|\varphi_{n+p} - \varphi_n\| &\le \|\varphi_{n+p} - \varphi_{n+p-1}\| + \cdots + \|\varphi_{n+1} - \varphi_n\| \\ &\le \left(\|A\|^{n+p-1} + \|A\|^{n+p-2} + \cdots + \|A\|^n\right) \|\varphi_1 - \varphi_0\| \\ &\le \frac{\|A\|^n}{1 - \|A\|} \|\varphi_1 - \varphi_0\| \to 0\end{aligned}$$

as $n \to \infty$ uniformly in $p \in \mathbb{N}$. It means that $\{\varphi_n\}$ is a Cauchy sequence in the Hilbert space H. Therefore, there exists unique limit

$$\varphi = \lim_{n\to\infty} \varphi_n.$$

Evidently this φ solves (10.1) uniquely. □

We will return to the integral operators

$$Af(x) = \int_\Omega K(x,y) f(y) dy, \tag{10.4}$$

where $K(x,y)$ is assumed to be in $L^2(\Omega \times \Omega)$. In that case, as we know, A is compact in $L^2(\Omega)$.

Definition. A function $K_n(x,y) \in L^2(\Omega \times \Omega)$ is said to be a *degenerate kernel* if

$$K_n(x,y) = \sum_{j=1}^n a_j(x) b_j(y), \tag{10.5}$$

with some functions $a_j, b_j \in L^2(\Omega)$.

We consider the integral equation of second kind with degenerate kernel $K_n(x,y)$ i.e.

$$\varphi_n(x) - \int_\Omega \sum_{j=1}^n a_j(x) b_j(y) \varphi_n(y) dy = f(x) \tag{10.6}$$

in the form

$$\varphi_n(x) - \sum_{j=1}^n \gamma_j a_j(x) = f(x),$$

where $\gamma_j = (\varphi_n, \overline{b_j})_{L^2(\Omega)}$. It means that the solution φ_n of (10.6) is necessarily represented as

$$\varphi_n(x) = f + \sum_{j=1}^n \gamma_j a_j \tag{10.7}$$

such that the coefficients γ_j (which are to be determined) satisfy the linear system

$$\gamma_j - \sum_{k=1}^n \gamma_k (a_k, \overline{b_j})_{L^2(\Omega)} = (f, \overline{b_j})_{L^2(\Omega)} = f_j, \quad j = 1, 2, \ldots, n. \tag{10.8}$$

Hence, the solution φ_n of (10.6) (see also (10.7)) can be obtained whenever we can solve the linear system (10.8) uniquely with respect to γ_j.

Let us consider now the integral equation of second kind with compact self-adjoint operator (10.4) i.e.

$$\varphi(x) - A\varphi(x) = f(x). \tag{10.9}$$

The main idea is to approximate the kernel $K(x, y)$ from (10.9) by the degenerate kernel $K_n(x, y)$ from (10.6) such that

$$\|K(x, y) - K_n(x, y)\|_{L^2(\Omega\times\Omega)} \to 0 \tag{10.10}$$

as $n \to \infty$ and, in addition, the inverse operators $(I - A_n)^{-1}$ exist and are uniformly bounded in n.

In that case the system (10.8) is uniquely solvable and we obtain an approximate solution φ_n such that

$$\|\varphi - \varphi_n\|_{L^2(\Omega)} \to 0, \quad n \to \infty.$$

Indeed, equations (10.6) and (10.9) imply

$$(\varphi - \varphi_n) - A_n(\varphi - \varphi_n) = (A - A_n)\varphi.$$

Since $(I - A_n)^{-1}$ exist and are uniformly bounded we have

$$\|\varphi - \varphi_n\| \le \left\|(I - A_n)^{-1}\right\| \|A - A_n\| \|\varphi\| \to 0$$

as $n \to \infty$ by (10.10). The unique solvability of (10.8) (or the uniqueness of φ_n) follows from the existence of the inverse operators $(I - A_n)^{-1}$.

We may justify this choice of the degenerate kernel $K_n(x, y)$ by the following considerations. Let $\{e_j\}_{j=1}^{\infty}$ be an orthonormal basis in $L^2(\Omega)$. Then $K(x, y) \in L^2(\Omega \times \Omega)$ as a function of $x \in \Omega$ (with parameter $y \in \Omega$) can be represented by

$$K(x, y) = \sum_{j=1}^{\infty} (K(\cdot, y), e_j)_{L^2} e_j(x).$$

Then

$$\left\| K(x, y) - \sum_{j=1}^{n} (K(\cdot, y), e_j)_{L^2} e_j \right\|_{L^2(\Omega\times\Omega)} \to 0$$

as $n \to \infty$, and we may consider the degenerate kernel $K_n(x, y)$ in the form

$$K_n(x, y) = \sum_{j=1}^{n} e_j(x) b_j(y),$$

where $b_j(y) = (K(\cdot, y), e_j)_{L^2}$. The system (10.8) transforms in this case to

$$\gamma_j - \sum_{k=1}^{n} \gamma_k (e_k, (e_j, K(\cdot, y))_{L^2})_{L^2(\Omega)} = f_j.$$

If, for example, e_j are the normalized eigenfunctions of the operator A with corresponding eigenvalues λ_j then the latter system can be rewritten as

$$\gamma_j - \lambda_j \gamma_j = f_j, \quad j = 1, 2, \ldots, n.$$

We assume that $\lambda_j \neq 1$ so that γ_j can be uniquely determined as

$$\gamma_j = \frac{f_j}{1 - \lambda_j}$$

and therefore φ_n is equal to

$$\varphi(x) = f(x) + \sum_{j=1}^{n} \frac{f_j}{1 - \lambda_j} e_j(x).$$

A different method goes back to Nyström. Let us consider instead of integral operator A with kernel $K(x, y)$ the sequence of numerical integration operators

$$A_n \varphi(x) = \sum_{j=1}^{n} \alpha_j^{(n)} K(x, x_j^{(n)}) \varphi(x_j^{(n)}). \tag{10.11}$$

We assume that the points $x_j^{(n)}$ and the weights $\alpha_j^{(n)}$ are chosen so that

$$\|A\varphi - A_n\varphi\|_{L^2}^2 = \int_\Omega \left| \int_\Omega K(x, y)\varphi(y) - \frac{1}{|\Omega|} \sum_{j=1}^{n} \alpha_j^{(n)} K(x, x_j^{(n)}) \varphi(x_j^{(n)}) dy \right|^2 dx \to 0$$

as $n \to \infty$. The main problem here is to choose the weights $\alpha_j^{(n)}$ and the points $x_j^{(n)}$ with this approximation property. The original Nyström method was constructed for continuous kernels $K(x, y)$.

In Hilbert spaces it is more natural to consider projection methods.

Definition. Let $A : H \to H$ be injective bounded linear operator. Let $P_n : H \to H_n$ be projection operators such that $\dim H_n = n$. For given $f \in H$ the *projection method* generated by H_n and P_n approximates the equation $A\varphi = f$ by the projection equation

$$P_n A \varphi_n = P_n f, \quad \varphi_n \in H. \tag{10.12}$$

This projection method is called convergent if there is $n_0 \in \mathbb{N}$ such that for each $f \in H$ the approximating equation (10.12) has a unique solution $\varphi_n \in H_n$ for all $n \geq n_0$ and

$$\varphi_n \to \varphi, \quad n \to \infty,$$

where φ is the unique solution of the equation $A\varphi = f$.

Theorem 5. *A projection method converges if and only if there exist $n_0 \in \mathbb{N}$ and $M > 0$ such that for all $n \geq n_0$ the operators*

$$P_n A : H \to H$$

are invertible and the operators $(P_n A)^{-1} P_n A : H \to H$ are uniformly bounded i.e.

$$\left\|(P_n A)^{-1} P_n A\right\| \leq M, \quad n \geq n_0.$$

In case of convergence we have the error estimate

$$\|\varphi_n - \varphi\| \leq (1 + M) \inf_{\psi \in H_n} \|\psi - \varphi\| .$$

Proof. If a projection method converges then, by definition, $P_n A$ are invertible and the uniform boundedness follows from the Banach–Steinhaus theorem.

Conversely, under the assumptions of the theorem

$$\varphi_n - \varphi = ((P_n A)^{-1} P_n A - I)\varphi.$$

Since for all $\psi \in H_n$ we have trivially $(P_n A)^{-1} P_n A \psi = \psi$ then

$$\varphi_n - \varphi = ((P_n A)^{-1} P_n A - I)(\varphi - \psi)$$

and the error esimate follows. □

Remark. Projection methods make sense and we can expect convergence only if the subspaces H_n possess the *denseness property*

$$\inf_{\psi \in H_n} \|\psi - \varphi\| \to 0, \quad n \to \infty.$$

Theorem 6. *Assume that $A : H \to H$ is compact, $I - A$ is injective and the projection operators $P_n : H \to H_n$ converge pointwise i.e. $P_n \varphi \to \varphi, n \to \infty$ for each $\varphi \in H$. Then the projection method for $I - A$ converges.*

Proof. By the Riesz theorem (see Theorem 5 of Chapter 4) operator $I - A$ has bounded inverse. Since $P_n \varphi \to \varphi$ as $n \to \infty$ we have $P_n A \varphi \to A\varphi$ as $n \to \infty$ too. At the same time the sequence $P_n A$ is collectively compact since A is compact and P_n is of finite rank. Thus, due to Exercise 48 we have

$$\|(P_n A - A) P_n A\| \to 0, \quad n \to \infty. \tag{10.13}$$

Then the operators $(I - P_n A)^{-1}$ exist and are uniformly bounded. Indeed, denoting

$$B_n := I + (I - A)^{-1} P_n A$$

we obtain

$$\begin{aligned} B_n(I - P_n A) &= (I - P_n A) + (I - A)^{-1} P_n A (I - P_n A) \\ &= I - (I - A)^{-1} (P_n A - A) P_n A \\ &= I - S_n. \end{aligned}$$

But it is easy to see from (10.13) that

$$\|S_n\| \to 0, \quad n \to \infty.$$

Hence, both $I - P_nA$ and B_n are injective. Since P_nA is compact then $(I - P_nA)^{-1}$ is bounded. As a consequence of this fact we have that

$$(I - P_nA)^{-1} = (I - S_n)^{-1}B_n.$$

Definition of B_n implies

$$\|B_n\| \leq 1 + \left\|(I - A)^{-1}\right\| \|A\|.$$

Therefore $\|(I - P_nA)^{-1}\|$ is uniformly bounded in n. The exact equation $\varphi - A\varphi = f$ and (10.12) with operator $I - A$ lead to

$$(I - P_nA)(\varphi_n - \varphi) = P_nA\varphi - A\varphi + P_nf - f,$$

which implies also the error estimate

$$\|\varphi_n - \varphi\| \leq \left\|(I - P_nA)^{-1}\right\| \Big(\|P_nA\varphi - A\varphi\| + \|P_nf - f\|\Big).$$

□

Corollary. *Under the assumptions of Theorem 6 and provided additionally that*

$$\|P_nA - A\| \to 0, \quad n \to \infty$$

the approximate equation (10.12) with $I - A$ is uniquely solvable for each $f \in H$ and we have the error estimate

$$\|\varphi_n - \varphi\| \leq M \|P_n\varphi - \varphi\|,$$

where M is an upper bound for the norm $\|(I - P_nA)^{-1}\|$.

Proof. The existence of the inverse operators $(I - P_nA)^{-1}$ and their uniform boundedness follows from

$$I - P_nA = (I - A) - (P_nA - A) = (I - A)\left[I - (I - A)^{-1}(P_nA - A)\right],$$

$$(I - P_nA)^{-1} = \left[I - (I - A)^{-1}(P_nA - A)\right]^{-1}(I - A)^{-1}$$

and

$$\left\|(I - P_nA)^{-1}\right\| \leq \frac{\|(I - A)^{-1}\|}{1 - \|(I - A)^{-1}\| \|(P_nA - A)\|}.$$

The error estimate is the consequence of

$$(\varphi - \varphi_n) - P_nA(\varphi - \varphi_n) = \varphi - P_n\varphi$$

and

$$\|\varphi_n - \varphi\| \leq \left\|(I - P_nA)^{-1}\right\| \|P_n\varphi - \varphi\|.$$

□

Let us return back to the projection equation (10.12). It can be rewritten equivalently as

$$(A\varphi_n - f, g) = 0, \quad g \in H_n. \tag{10.14}$$

Indeed, if $g \in H_n$ then $g = P_n g$, $P_n^* = P_n$ and hence

$$0 = (A\varphi_n - f, g) = (A\varphi_n - f, P_n g) = (P_n(A\varphi_n - f), g)$$

or

$$P_n(A\varphi_n - f) = 0,$$

since H_n is considered here as a Hilbert space. This is the basis for the Galerkin projection method.

Assume that $\{e_j\}_{j=1}^{\infty}$ is an orthonormal basis in the Hilbert space H. Considering

$$H_n = \overline{\text{span}(e_1, \ldots, e_n)}$$

we have for the solution φ_n of the projection equation (10.12) the representation

$$\varphi_n(x) = \sum_{j=1}^{n} \gamma_j e_j. \tag{10.15}$$

The task here is to find (if possible uniquely) the coefficients γ_j such that φ_n from (10.15) solves (10.12). Since (10.14) is equivalent to (10.12) we have from (10.15) that

$$(A\varphi_n, g) = (f, g), \quad g \in H_n$$

or

$$(A\varphi_n, e_k) = (f, e_k) = f_k, \quad k = 1, 2, \ldots, n$$

or

$$\sum_{j=1}^{n} \gamma_j (Ae_j, e_k) = f_k$$

or

$$M\vec{\gamma} = \vec{f}, \tag{10.16}$$

where $\vec{\gamma} = (\gamma_1, \ldots, \gamma_n)$, $\vec{f} = (f_1, \ldots, f_n)$ and $M = \{a_{jk}\}_{n \times n}$ with $a_{jk} = (Ae_j, e_k)$. If operator A is invertible then the matrix M is invertible too and $\vec{\gamma}$ can be obtained uniquely as

$$\vec{\gamma} = M^{-1}\vec{f}.$$

As a result of this consideration we obtain $\varphi_n(x)$ uniquely from (10.15). It remains only to check that this φ_n converges to the solution of the exact equation $A\varphi = f$. In order to verify this fact it is enough to apply Theorem 5.

We apply now this projection method to the equation (10.9) with compact operator A.

Theorem 7. *Let $A : H \to H$ be compact and let $I - A$ be injective. Then the Galerkin projection method converges.*

Proof. By the Riesz theorem, the operator $I - A$ has bounded inverse. That's why φ_n from (10.15) is uniquely defined with γ_j that satisfies the equation (10.16) with matrix $M = \{a_{jk}\}_{n\times n}, a_{jk} = ((I - A)e_j, e_k)$. Since

$$\|P_n\varphi - \varphi\|_H^2 = \sum_{j=n+1}^{\infty} |(\varphi, e_j)|^2 \to 0, \quad n \to \infty$$

then we may apply Theorem 6 and conclude this theorem. □

Index