Functional analysis (MS-E1460, -E1461, -E1462)
Ville Turunen, Aalto University, 27.3.2020

Welcome to the introducture lecture course on functional analysis. We shall deal mostly with Hilbert spaces, but starting first with more general Banach spaces. There will be 12 lectures (without the prequel of Lecture 0), each taking 180 minutes of time in the class room. A remark: the text is organized in the fashion that enables starting also from the second half (the Hilbert space lectures), only later coming back to the first half of the functional analysis notes. Let us quickly list the topics of the lectures:

0. *Functional analysis* $\approx$ **vector spaces** + **topology**
1. Banach spaces
2. Bounded operators
3. Fruits of completeness: corollaries of Baire
4. Duality in Banach spaces
5. Compact operators
6. Spectral properties in Banach spaces
7. Hilbert spaces
8. Orthogonality
9. Duality in Hilbert spaces
10. Operators in Hilbert spaces
11. Spectral properties in Hilbert spaces
12. Singular value decomposition (SVD)

0. *Functional analysis* $\approx$ vector spaces + topology

It could be said that a course in freshman linear algebra is "finite-dimensional functional analysis". We could describe functional analysis to be the discipline that combines linear algebra with metric and topological structures, enabling us to study also the infinite-dimensional vector spaces. Why should we bother? Such vector spaces and related linear operators live everywhere in mathematics and its applications: e.g. when studying partial differential equations, integral transforms, differential geometry, probability, classical mechanics, quantum mechanics, signal processing etc.

In the sequel, we write the scalar field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ (real or complex). In this course, metric space topology will be enough. The reader should already know the basic linear algebra concepts: vector spaces and linear mappings, norms and inner products; moreover, we assume familiarity with continuity and compactness in metric spaces, and also Cauchy sequences and metric completeness.

In the sequel, in each class of vector spaces, please pay attention to the following properties:

(*.1) Basic vector structures.
(*.2) Operators related to such structures.
(*.3) Related functionals (Duality).

Moreover, also spectral properties of linear operators (e.g. eigenvalues and eigenvectors) are important. Let us reflect on the properties that we listed above:

Informal example. Forgetting technicalities, think about:

(*.1) Vector spaces $X(M)$ consisting of nice-enough functions

$$u : M \to \mathbb{K}$$

(here spaces M, N might have some interesting extra structure).

(*.2) Linear mappings $A : X(N) \to X(M)$, for instance integral transform

$$Av(x) := \int_N K_A(x, y)\, v(y)\, \mathrm{d}y$$

(where the integral kernel $K_A : N \times M \to \mathbb{K}$ is nice-enough).

(*.3) Linear functionals $\varphi : X(M) \to \mathbb{K}$, e.g. given by a weighted integral

$$\varphi(u) := \int_M u(x)\, w_\varphi(x)\, \mathrm{d}x$$

(where the weight $w_\varphi : M \to \mathbb{K}$ is nice-enough).

Please keep in mind that this text is ought to be an introduction to functional analysis: in the first reading, try to get just the main ideas, and only afterwards fill in the logical details — do not check your every step if you want to run fast... ;)

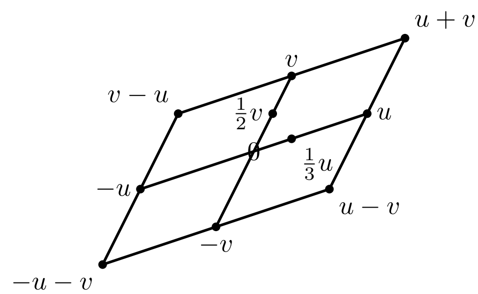


Figure 1: Schematic picture of vector operations.

Remark. Let $0 \leq \lambda_\alpha < \infty$ for each $\alpha \in J$. Define (potentially infinite) sum

$$\sum_{\alpha \in J} \lambda_\alpha := \sup \left\{ \sum_{\alpha \in S} \lambda_\alpha : \ S \subset J \text{ finite}, \ S \neq \emptyset \right\} \quad \in \quad [0, \infty]. \tag{1}$$

Here it is easy to show that if $\sum_{\alpha \in J} \lambda_\alpha < \infty$ then $\lambda_\alpha > 0$ only for at most countably many indices $\alpha \in J$.

0.1 Vector spaces

As the reader already should now, vectors can be added, and multiplied by scalars. For convenience, here is the precise description:

Definition. A $\mathbb{K}$-*vector space* (or just *vector space*, if $\mathbb{K}$ is known) is a set V of *vectors* $u \in V$, endowed with mappings

$$((u, v) \mapsto u + v) : V \times V \to V,$$
$$((\lambda, u) \mapsto \lambda u) : \mathbb{K} \times V \to V$$

with *origin* $0 \in V$, such that for all $u, v, w \in V$ and $\lambda, \mu \in \mathbb{K}$:

$$\begin{aligned}
&(u+v)+w = u+(v+w),\\
&u+v = v+u,\\
&u+0 = u,\\
&u+(-1)u = 0,\\
&1u = u,\\
&\lambda(\mu u) = (\lambda\mu)u,\\
&\lambda(u+v) = \lambda u + \lambda v,\\
&(\lambda+\mu)u = \lambda u + \mu u.
\end{aligned}$$

Write $u + v + w := (u + v) + w = u + (v + w)$ and $-u := (-1)u$.

Example. As a set, the *n-dimensional Euclidean vector space* $\mathbb{K}^n$ consists of points $u = (u_k)_{k=1}^n = (u_1, \cdots, u_n)$, where the kth coordinate of $u \in \mathbb{K}^n$ is $u_k \in \mathbb{K}$. Then $\mathbb{K}^n$ is a vector space over $\mathbb{K}$, with vector operations given by

$$\begin{aligned} (u_k)_{k=1}^n + (v_k)_{k=1}^n &:= (u_k + v_k)_{k=1}^n, \\ \lambda (u_k)_{k=1}^n &:= (\lambda u_k)_{k=1}^n. \end{aligned}$$

Especially, vector space $\mathbb{K}^1$ can be identified with the scalar field $\mathbb{K}$.

Definition. The *direct sum* of a non-empty family $\{V_\alpha\}_{\alpha \in J}$ of $\mathbb{K}$-vector spaces is the vector space

$$W := \bigoplus_{\alpha \in J} V_\alpha,$$

which as a set is the Cartesian product of spaces V_α. Here $u = (u_\alpha)_{\alpha \in J} \in W$ is a function on J such that $u_\alpha := u(\alpha) \in V_\alpha$ is the α-coordinate of u, and

$$\begin{aligned} u + v &:= (u_\alpha + v_\alpha)_{\alpha \in J}, \\ \lambda u &:= (\lambda u_\alpha)_{\alpha \in J}. \end{aligned}$$

Often in the literature, the following notations are used for the same direct sum:

$$V_1 \oplus V_2, \qquad \bigoplus_{\alpha \in \{1,2\}} V_\alpha, \qquad V_1 \times V_2.$$

Above, $\mathbb{K}^n$ can be identified with a direct sum of n copies of $\mathbb{K}$.

Example. Let V be a vector space and M a non-empty set. Let $V_x := V$ for each $x \in M$. Vector space

$$V^M := \bigoplus_{x \in M} V_x$$

of all functions $u, v : M \to V$ has vector operations defined by

$$\begin{aligned} (u + v)(x) &:= u(x) + v(x), \\ (\lambda u)(x) &:= \lambda\, u(x). \end{aligned}$$

Notice that if set M has n points, then $\mathbb{K}^M$ can be identified with $\mathbb{K}^n$.

0.2 Subspaces, quotient spaces

Definition. Subset $Z \subset V$ is a *subspace* of vector space V if it is a vector space with respect to the restrictions of vector space operations. Vector space V has always *trivial subspaces* $\{0\}$ and V.

Example. The vector space $\mathbb{K}^{\mathbb{R}}$ of all functions $u : \mathbb{R} \to \mathbb{K}$ has vector subspaces $C^k(\mathbb{R})$ consisting of k times continuously differentiable functions. These spaces have the subspace

$$C^\infty(\mathbb{R}) = \bigcap_{k=0}^{\infty} C^k(\mathbb{R})$$

of infinitely smooth functions. This subspace contains the subspace consisting of all polynomials $u : \mathbb{R} \to \mathbb{K}$,

$$u(x) = \sum_{m=0}^{N} a_m \, x^m.$$

Exercise. Let V_1, V_2 be vector subspaces of V. Show that

$$V_1 + V_2 := \{u_1 + u_2 \mid u_1 \in V_1, \; u_2 \in V_2\}$$

is a vector subspace of V.

Definition. If V_1, V_2 are vector subspaces of V such that

$$V_1 + V_2 = V \quad \text{and} \quad V_1 \cap V_2 = \{0\}, \tag{2}$$

we say that V is a *direct sum* of V_1, V_2, and we write $V = V_1 \oplus V_2$.

Exercise. Let $V_\alpha \subset V$ be a vector subspace for each $\alpha \in J$, where J is any non-empty index set. Show that $\bigcap_{\alpha \in J} V_\alpha$ is a vector subspace.

Example. Let M be a non-empty set. For $K \subset M$, vector space V^M of functions $u : M \to V$ has a subspace

$$Z(K) := \{u : M \to V \mid \forall x \in K : \; u(x) = 0\}.$$

Here $Z(M) = \{0\} \subset V^M$ and $Z(\emptyset) = V^M$. Also structures on set M (e.g. topology, smoothness, symmetries, measure...) give ideas for nice subspaces of V^M: for instance, think of vector spaces of continuous functions, smooth functions, measurable functions, etc.

Definition. Let $Z \subset V$ be a vector subspace. Then

$$u \sim v \overset{\text{definition}}{\iff} u - v \in Z$$

gives an equivalence relation on V, with equivalence classes

$$[u] = \{v \in V : \; u \sim v\} = \{u + z : \; z \in Z\} =: u + Z.$$

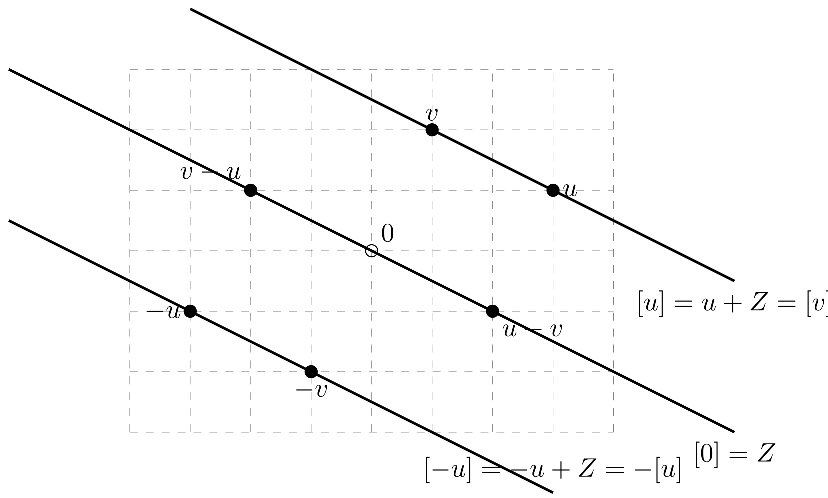


Figure 2: Some equivalence classes for a quotient vector space

Quotient vector space $V/Z := \{u + Z \mid u \in V\}$ has operations

$$\begin{aligned} (([u],[v]) \mapsto [u+v]) &\quad : \quad V/Z \times V/Z \to V/Z, \\ (\lambda,[v])) \mapsto [\lambda v]) &\quad : \quad \mathbb{K} \times V/Z \to V/Z. \end{aligned}$$

Notice that $[0] = Z \subset V$ is the origin of the quotient space V/Z.

Exercise. Verify that quotient space V/Z defined above is indeed a vector space.

0.3 Linear operators

Linear operators (or linear functions) are the mappings between vector spaces that preserve the natural vector operations:

Definition. Function $A : V \to W$ in $\mathbb{K}$-vector spaces V, W is *linear* (a linear operator or a morphism), denoted by $A \in \mathrm{Hom}(V,W) = \mathrm{Mor}(V,W)$, if

$$\begin{cases} A(u+v) = A(u) + A(v), \\ A(\lambda v) = \lambda\, A(v) \end{cases}$$

for all $u, v \in V$ and $\lambda \in \mathbb{K}$; write $Av := A(v)$. Let $\mathrm{Hom}(V) := \mathrm{Hom}(V,V)$. The *null space* or *kernel* $A^{-1}\{0\} = \mathcal{N}(A) = \ker(A) \subset V$ of A is

$$A^{-1}\{0\} = \mathcal{N}(A) = \ker(A) := \{u \in V : Au = 0\}.$$

The *range* $A(V) = \mathcal{R}(A) = \mathrm{ran}(A) \subset W$ of A is defined by

$$A(V) = \mathcal{R}(A) = \mathrm{ran}(A) := \{Au : u \in V\}.$$

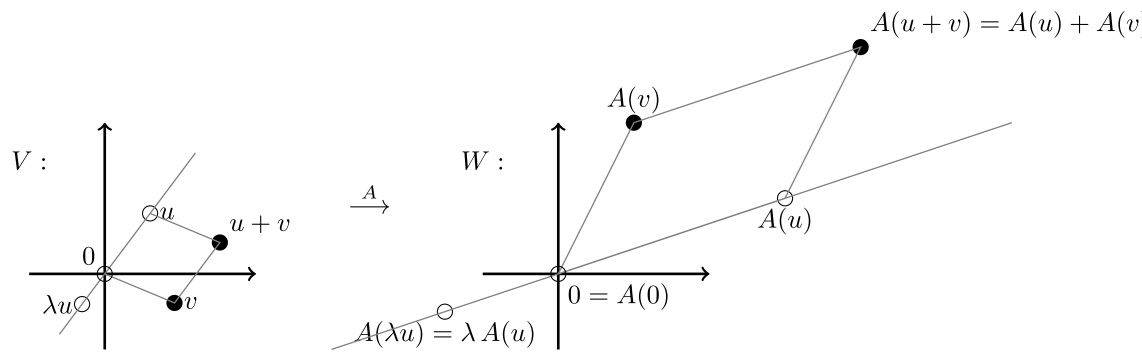


Figure 3: Idea of a linear mapping $A : V \to W$.

Notice that the kernel subspace $\ker(A)$ has nothing to do with the integral kernel K_A of an integral operator A, where $Au(x) = \displaystyle\int K_A(x,y)\,u(y)\,\mathrm{d}y$.

Example. Identity operator $I = (u \mapsto u) : V \to V$ is clearly linear, as well as $\lambda I = (u \mapsto \lambda u) : V \to V$ for all $\lambda \in \mathbb{K}$. Mappings $(\lambda \mapsto \lambda u) : \mathbb{K} \to V$ are linear for all $u \in V$. If $((u,v) \mapsto \langle u,v\rangle) : V \times V \to \mathbb{K}$ is an inner product of V then $(u \mapsto \langle u,v\rangle) : V \to \mathbb{K}$ is linear for all $v \in V$.

Exercise. Let $A : V \to W$ be linear. Show that $\ker(A)$ is a vector subspace of V and that $\mathrm{ran}(A)$ is a vector subspace of W.

Exercise. Let $A : V \to W$ be linear. Prove the following claims:
(a) Linear operator A is injective if and only if $\ker(A) = \{0\}$.
(b) $\widetilde{A} : V/\ker(A) \to \mathrm{ran}(A)$ is a linear bijection, where $\widetilde{A}[u] := Au$, when $[u] := u + \ker(A) \in V/\ker(A)$.

Example. If Z is a vector subspace of V then $A : V \to V/Z$ is linear, where $Au := u + Z$. Then A is surjective and $\ker(A) = Z$.

Example. Linear $A : \mathbb{K}^m \to \mathbb{K}^n$ has a matrix $[A] = [A_{jk}] \in \mathbb{K}^{n\times m}$,

$$(Au)_j = \sum_{k=1}^{m} A_{jk}\,u_k \in \mathbb{K}.$$

In this case, we may identify linear operator A with its matrix $[A]$.

Informal example. Suppose we could integrate on set N. Let $V(M), V(N)$ be vector subspaces of suitable functions on M, N, respectively. Then formula

$$Av(x) = \int_N K_A(x,y)\,v(y)\,\mathrm{d}y$$

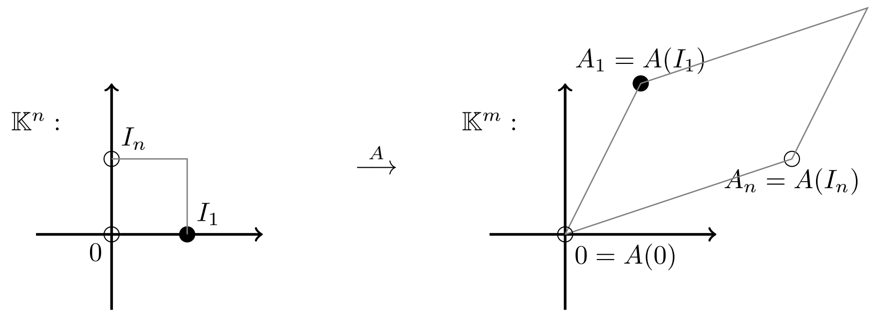


Figure 4: Matrix $[A] = [A_1 \cdots A_n] \in \mathbb{K}^{m\times n}$ of linear mapping $A : \mathbb{K}^n \to \mathbb{K}^m$.

might define linear operator $A : V(N) \to V(M)$, with its "Schwartz integral kernel" $K_A : M \times N \to \mathbb{K}$ analogous to matrix presentation of A.

Example. For smooth functions $u : \mathbb{R} \to \mathbb{K}$ we have

$$u(x) = u(0) + \int_0^x u'(t)\,\mathrm{d}t$$

by the Fundamental Theorem of Calculus. Thereby the linear mapping $A = (u \mapsto u') : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ is surjective but not injective.

Exercise. Let $A : V \to W$ and $B : W \to X$ be linear. Show that the composite mapping $BA := B \circ A : V \to X$ is linear.

Powers of operators. For linear $A : V \to V$ and $k \in \mathbb{Z}^+$, we define the kth power $A^k : V \to V$ by $A^k := A^{k-1}A$, with convention that $A^0 = I : V \to V$ is the identity operator. If A is invertible, then define $(A^{-1})^k = (A^k)^{-1} =: A^{-k}$.

Definition. Linear $P : X \to X$ on a vector space X is a *projection* if $P^2 = P$. Then we have the direct sum

$$X = \mathrm{ran}(P) \oplus \ker(P). \tag{3}$$

Notice that if $X = V \oplus W$ for some vector subspaces $V, W \subset X$, then $(v+w) \mapsto v$ (where naturally $v \in V$ and $w \in W$) is a projection with range V and kernel W. Notice also that $I - P$ is a projection if P is a projection.

0.4 Convexity

Definition. Let S be a subset of a $\mathbb{K}$-vector space X.
(a) S is *absorbing* if for all $u \in X$ there is $r_u > 0$ so that $u \in rS$ when $r \geq r_u$.
(b) S is *balanced* if $\lambda u \in S$ when $u \in S$ and $|\lambda| \leq 1$.
(c) S is *convex* if $tx + (1-t)y \in S$ for every $x, y \in S$ when $0 < t < 1$.

Exercise. Let $C \subset V$ be convex and $A \in \mathrm{Hom}(V, W)$. Show that $A(C) \subset W$ is convex.

Definition. The *convex hull* of a subset S of a vector space X is the intersection of all convex sets that contain S. (Notice that at least X is a convex set containing S.)

Exercise. Show that the convex hull of S is the smallest convex set that contains S.

Exercise. Show that $v \in X$ belongs to the convex hull of S if and only if

$$v = \sum_{k=1}^{n} t_k \, v_k$$

for some $n \in \mathbb{Z}^+$, where the vectors $v_k \in S$, and $t_k > 0$ are such that $\sum_{k=1}^{n} t_k = 1$.

0.5 Duality

Definition. *Linear functionals* $\varphi \in \mathrm{Hom}(V, \mathbb{K})$ form *algebraic dual* $V^t = \mathrm{Hom}(V, \mathbb{K})$ ("$t = transpose$"). For $v \in V$ and $\varphi \in V^t$, write

$$\langle v, \varphi \rangle := \varphi(v).$$

Example. Inner product $\langle u, v \rangle$ gives linear functionals $u \mapsto \langle u, v \rangle$.

Informal example. Let $V(M)$ consist of nice $u : M \to \mathbb{K}$. Then

$$\langle u, \varphi \rangle = \varphi(u) := \int_M u(x) \, w_\varphi(x) \, \mathrm{d}x$$

may be a linear functional $\varphi : V(M) \to \mathbb{K}$.

Definition. Function $A^t : W^t \to V^t$ is *transpose* of $A \in \mathrm{Hom}(V, W)$ if

$$\langle v, A^t(w^t) \rangle = \langle Av, w^t \rangle$$

for all $v \in V$ and $w^t \in W^t$. Then $A^t \in \mathrm{Hom}(W^t, V^t)$.

Definition. The *span* of a non-empty subset S of a vector space V is

$$\mathrm{span}(S) := \left\{ \sum_{u \in S} \lambda(u) u \;\middle|\; \lambda : S \to \mathbb{K} \text{ finitely supported} \right\}.$$

Thus $\mathrm{span}(S) \subset V$ is the smallest vector subspace containing $S \subset V$, consisting of all the linear combinations of vectors of S.

Definition. A non-empty subset S of vector space V is *linearly dependent*, if it is not linearly independent: $S \neq \emptyset$ is *linearly independent* if

$$\sum_{u \in S} \lambda(u) u = 0 \quad \Rightarrow \quad \lambda \equiv 0$$

whenever $\lambda : S \to \mathbb{K}$ is finitely supported. A subset $S \subset V$ is called an *algebraic basis* (or a *Hamel basis*) of V if S is linearly independent and $V = \mathrm{span}(S)$.

Example. The canonical algebraic basis for $\mathbb{K}^n$ is $\{e_k\}_{k=1}^n$, where $e_k = (\delta_{jk})_{j=1}^n$ and $\delta_{jk} \in \{0,1\}$ is the *Kronecker delta*: that is, $\delta_{kk} = 1$ and $\delta_{jk} = 0$ if $j \neq k$.

Remark: Let $\mathcal{B}$ be an algebraic basis for V. Then there exists a unique set of linear functionals $(u \mapsto \langle u, b \rangle_{\mathcal{B}}) : V \to \mathbb{K}$ such that

$$u = \sum_{b \in \mathcal{B}} \langle u, b \rangle_{\mathcal{B}}\, b$$

for all $u \in V$. Clearly, $\langle u, b \rangle_{\mathcal{B}} \neq 0$ for at most finitely many $b \in \mathcal{B}$. The following Basis Lemma tells us that there are a plenty of non-zero linear functionals on $V \neq \{0\}$:

Basis Lemma ($\iff$ Zorn's Lemma $\iff$ Axiom of Choice). *Any vector space $V \neq \{0\}$ has an algebraic basis, which has a definite cardinality (called the* dimension *of the vector space).*

Proof. Let $\mathcal{F}$ be the family of all linearly independent subsets of V. Now $\mathcal{F} \neq \emptyset$, because $\{u\} \in \mathcal{F}$ for every $u \in V \setminus \{0\}$. Endow $\mathcal{F}$ with a partial order by inclusion. Chain $\mathcal{C} \subset \mathcal{F}$ clearly has an upper bound $F := \bigcup \mathcal{C} \in \mathcal{F}$. By Zorn's Lemma, there is a maximal element $M \in \mathcal{F}$. Obviously, M is an algebraic basis for V.

Let $\mathcal{A}, \mathcal{B}$ be algebraic bases for V. By induction, $\mathrm{card}(\mathcal{A}) = \mathrm{card}(\mathcal{B})$ when $\mathcal{A}$ is finite. So suppose $\mathrm{card}(\mathcal{A}) \leq \mathrm{card}(\mathcal{B})$, where $\mathcal{A}$ is infinite. Now $\mathrm{card}(\mathcal{A}) = \mathrm{card}(S)$, where

$$S := \{(a,b) \in \mathcal{A} \times \mathcal{B} : \ \langle a, b \rangle_{\mathcal{B}} \neq 0\}.$$

Assume $\mathrm{card}(\mathcal{A}) < \mathrm{card}(\mathcal{B})$. Thus

$$\exists b_0 \in \mathcal{B} \ \forall a \in \mathcal{A} : \ \langle a, b_0 \rangle_{\mathcal{B}} = 0.$$

But then

$$\begin{aligned}
b_0 &= \sum_{a\in\mathcal{A}} \langle b_0, a\rangle_{\mathcal{A}}\, a \\
&= \sum_{a\in\mathcal{A}} \langle b_0, a\rangle_{\mathcal{A}} \sum_{b\in\mathcal{B}} \langle a, b\rangle_{\mathcal{B}}\, b \\
&= \sum_{b\in\mathcal{B}} \left(\sum_{a\in\mathcal{A}} \langle b_0, a\rangle_{\mathcal{A}}\, \langle a, b\rangle_{\mathcal{B}} \right) b \\
&= \sum_{b\in\mathcal{B}\setminus\{b_0\}} \left(\sum_{a\in\mathcal{A}} \langle b_0, a\rangle_{\mathcal{A}}\, \langle a, b\rangle_{\mathcal{B}} \right) b \\
&\in \operatorname{span}(\mathcal{B}\setminus\{b_0\}),
\end{aligned}$$

contradicting the linear independence of $\mathcal{B}$. Thus $\operatorname{card}(\mathcal{A}) = \operatorname{card}(\mathcal{B})$. □

Definition. By the Basis Lemma, we may define the *algebraic dimension* $\dim_{\mathbb{K}}(V) = \dim(V)$ of a $\mathbb{K}$-vector space V to be the cardinality of its any algebraic basis. The vector space V is said to be *finite-dimensional* if $\dim(V)$ is finite, and *infinite-dimensional* otherwise. The *codimension* $\operatorname{codim}_V(Z)$ of a vector subspace $Z \subset V$ is the algebraic dimension of quotient vector space V/Z.

Remark. Of course, $\dim_{\mathbb{K}}(\mathbb{K}^n) = n$, but beware: Sometimes the same set can be viewed as a vector space over different fields, affecting naturally the algebraic dimension. For instance, $\dim_{\mathbb{C}}(\mathbb{C}) = 1 \neq 2 = \dim_{\mathbb{R}}(\mathbb{C})$, where in the latter case we identified $\mathbb{C}$ with the $\mathbb{R}$-vector space $\mathbb{R}^2$.

Example. Let M be a set. Vector space $\mathbb{K}^M$ of functions $u : M \to \mathbb{K}$ is finite-dimensional if and only if M is finite.

Example. The span of vectors $u_1, \cdots, u_k \in \mathbb{K}^n$ is the vector subspace

$$Z_k = \operatorname{span}\{u_j\}_{j=1}^k := \left\{ \sum_{j=1}^k \lambda_j u_j \in \mathbb{K}^n : \ \lambda_1, \cdots, \lambda_k \in \mathbb{K} \right\}.$$

E.g. $Z_1 = \mathbb{K}u_1 = \{\lambda_1 u_1 : \ \lambda \in \mathbb{K}\} \subset \mathbb{K}^n$ (the line through the points $u_1 \neq 0, 0$, or just $\{0\}$ if $u_1 = 0$), and

$$\{0\} \subset Z_1 \subset Z_2 \subset \cdots \subset Z_{k-1} \subset Z_k \subset \mathbb{K}^n.$$

The dimensions are $\dim(\mathbb{K}^n) = n$ and $\dim(\{0\}) = 0$, and if here $Z_{j+1} \neq Z_j$, then $\dim(Z_{j+1}) = 1 + \dim(Z_j)$. Hence

$$0 \leq \dim(Z_k) \leq k.$$

Vectors $u_1, \cdots, u_k$ are *linearly independent* if $\dim(Z_k) = k$ (otherwise *linearly dependent*).

Example. $\mathrm{codim}_V(\ker(\varphi)) = 1$ if $0 \neq \varphi \in V^t = \mathrm{Hom}(V, \mathbb{K})$.

Exercise. Show that $\dim(V_1 + V_2) + \dim(V_1 \cap V_2) = \dim(V_1) + \dim(V_2)$.

Exercise. Let X be a vector subspace of V. Show that there exists a vector subspace Y such that $V = X + Y$ and $X \cap Y = \{0\}$. In other words, $V \cong X \oplus Y$. (Hint: Zorn's Lemma.)

Exercise. Let Z_k be a vector subspace of V when $1 \leq k \leq n \in \mathbb{Z}^+$. Show that

$$\mathrm{codim}_V(\bigcap_{k=1}^{n} Z_k) \leq \sum_{k=1}^{n} \mathrm{codim}_V(Z_k).$$

Exercise. Let $V = C([0,1])$ and $Z = \{u \in V : \ u(0) = u(1)\}$. Find $\mathrm{codim}_V(Z)$. Find also vector subspace Y of V such that $V = Y + Z$ and $Y \cap Z = \{0\}$.

0.6 Spectral theory

Definition. Let V be a $\mathbb{K}$-vector space, with identity operator

$$I = (u \mapsto u) : V \to V.$$

The *spectrum* of linear $A : V \to V$ is

$$\sigma(A) := \{\lambda \in \mathbb{K} : \ \lambda I - A \text{ is not bijective}\}.$$

If $Au = \lambda u$ where $\lambda \in \mathbb{K}$ and $0 \neq u \in V$ then $\lambda \in \sigma(A)$ is called an *eigenvalue* corresponding to the *eigenvector* u. The subset of the eigenvalues is called the *point spectrum* of A.

Remark. For linear $A : \mathbb{C}^n \to \mathbb{C}^n$, we have $\sigma(A) \neq \emptyset$, and the spectrum consists of eigenvalues only: The *characteristic polynomial* of matrix $[A] \in \mathbb{C}^{n\times n}$ is $p_A : \mathbb{C} \to \mathbb{C}$, where

$$p_A(z) := \det[A - zI],$$

$\det[B]$ being the determinant of matrix $[B] \in \mathbb{C}^{n\times n}$. By the **Fundamental Theorem of Algebra** [Gauss], polynomials split uniquely in $\mathbb{C}$ into product of first order terms; thus

$$p_A(z) = (-1)^n (z - \lambda_1) \cdots (z - \lambda_n),$$

where $\lambda_1, \cdots, \lambda_n \in \mathbb{C}$ are the eigenvalues of A. The *algebraic multiplicity* $d = m_a(\lambda)$ of an eigenvalue $\lambda \in \mathbb{C}$ is the degree d of the factor $(z - \lambda)^d$ in p_A. The *geometric multiplicity* $m_g(\lambda)$ of an eigenvalue $\lambda \in \mathbb{C}$ is the dimension of the vector space spanned by the corresponding eigenvectors. Notice that always

$$1 \leq m_g(\lambda) \leq m_a(\lambda) \leq n.$$

Actually,

$$p_A(z) = (-1)^n \left(z^n - \operatorname{tr}[A]z^{n-1} + \cdots + (-1)^n \det[A]\right),$$

where $\operatorname{tr}[A] \in \mathbb{C}$ is the *trace* of $[A] \in \mathbb{C}^{n\times n}$, satisfying

$$\operatorname{tr}[A] := \sum_{k=1}^{n} A_{kk} = \sum_{k=1}^{n} \lambda_k;$$

typically here $A_{kk} \neq \lambda_k$.

Exercise. Let $V = C([0,1])$ be the infinite-dimensional vector space of continuous functions on the closed interval $[0,1]$. Define a linear mapping $A : V \to V$ by

$$Au(x) := \int_0^x u(t)\,\mathrm{d}t.$$

Find $\ker(A)$ and $\operatorname{ran}(A)$. Show that A does not have any eigenvalues. Especially, $0 \in \sigma(A)$ is not an eigenvalue here!

1 Banach spaces

Loosely speaking, normed spaces are those vector spaces where we can measure distances between points in a manner that respects the usual vector operations. Banach spaces are metrically complete normed spaces, and later there will be a canonical procedure to complete any normed space.

1.1 Seminorms and norms

Definition. *Seminorm* on a $\mathbb{K}$-vector space X is function $s : X \to \mathbb{R}$, where

$$\begin{cases} s(u+v) \leq s(u) + s(v) & \text{(subadditivity)}, \\ s(\lambda u) = |\lambda|\, s(u) \end{cases}$$

for all $u, v \in X$ and for all $\lambda \in \mathbb{K}$. Seminorm s is a *norm* if $s(u) > 0$ whenever $u \neq 0$. A norm on X is written as $u \mapsto \|u\|_X$ or simply $\|u\|$. A vector space with a norm is called a *normed space*. Subset S of a normed space X is *bounded* if $\|u\| < c$ for all $u \in S$, where $c < \infty$ is a constant.

Remark. For a seminorm s, we clearly have $s(u) \leq s(u-v) + s(-v)$, yielding

$$|s(u) - s(v)| \leq s(u-v). \tag{4}$$

Especially, a seminorm cannot have negative values.

Example. For $1 \leq p < \infty$, norm

$$u \mapsto \|u\|_p := \left(\sum_{k=1}^{d} |u_k|^p \right)^{1/p}$$

on $\mathbb{K}^d$, where $u \mapsto |u_k|$ is a seminorm for each $k \in \{1, \cdots, d\}$. Norm $\|u\|_p$ is called the p-norm, and especially: *Taxicab norm* $\|u\|_1$ is the sum of these seminorms $u \mapsto |u_k|$. *Euclidean norm* $\|u\|_2$ is the "usual Pythagorean norm". *Maximum norm* satisfies $\|u\|_\infty := \max\{|u_1|, \cdots, |u_d|\} = \lim_{p\to\infty} \|u\|_p$.

Exercise. Let $s : X \to \mathbb{R}$ a be seminorm, and for $r > 0$ define the respective "open" and "closed" *r-semiballs* $\{s < r\}, \{s \leq r\}$ by

$$\{s < r\} \quad := \quad \{u \in X : \; s(u) < r\}, \tag{5}$$

$$\{s \leq r\} \quad := \quad \{u \in X : \; s(u) \leq r\}. \tag{6}$$

Prove that these r-semiballs are
(**a**) **a**bsorbing,
(**b**) **b**alanced,
(**c**) **c**onvex.

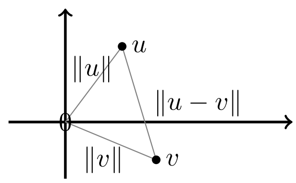


Figure 5: Norms, distances.

Exercise. Let X be a vector space, and $D \subset X$ **a**bsorbing **b**alanced **c**onvex.
(a) Show that the *Minkowski functional* $f_D : X \to \mathbb{R}$ is a seminorm, where

$$f_D(u) := \inf \{ r > 0 : \; u/r \in D \} . \tag{7}$$

(b) Let $O := \{f_D < 1\}$ and $C := \{f_D \leq 1\}$. Show that

$$O \subset D \subset C, \qquad f_O = f_D = f_C .$$

Exercise. Let s be a seminorm on a $\mathbb{K}$-vector space X, and define

$$u \sim v \quad \overset{\text{definition}}{\Longleftrightarrow} \quad s(u - v) = 0.$$

Let $[u] := \{v \in X : \; u \sim v\}$. Prove the following claims:
(a) $\sim$ is an equivalence relation on X.
(b) $L := X/s^{-1}(0) = \{[u] : \; u \in X\}$ is a normed space with

$$[u] + [v] := [u + v], \quad \lambda[u] := [\lambda u], \quad \text{quotient norm} \quad [u] \mapsto s(u).$$

Definition. The *norm metric*

$$((u, v) \mapsto \|v - u\|) : X \times X \to \mathbb{R} \tag{8}$$

yields *norm topology*, where the open balls are of form

$$\mathbb{B}_X(u, r) = \mathbb{B}(u, r) := \{v \in X : \; \|v - u\| < r\} . \tag{9}$$

In other words, $\|v - u\|$ is the *distance* between the points $u, v \in X$. We may write

$$\mathbb{B}_r = r\mathbb{B} = \mathbb{B}(0, r) = \{\|v\| < r\} , \tag{10}$$

where $\mathbb{B}_1 = \mathbb{B} = \mathbb{B}(0, 1) = \{\|v\| < 1\}$ is the open unit ball centered at the origin. Then $\overline{\mathbb{B}_r} = \{v \in X : \; \|v\| \leq r\} = \{\|v\| \leq r\}$ is the closed ball of radius r centered at the origin. *Banach space* is a normed space with complete norm metric: i.e. all the Cauchy sequences converge there.

Exercise. In a normed space, prove the following claims:
(a) Finite-dimensional subspaces are closed.
(b) Closures and interiors of convex sets are convex.

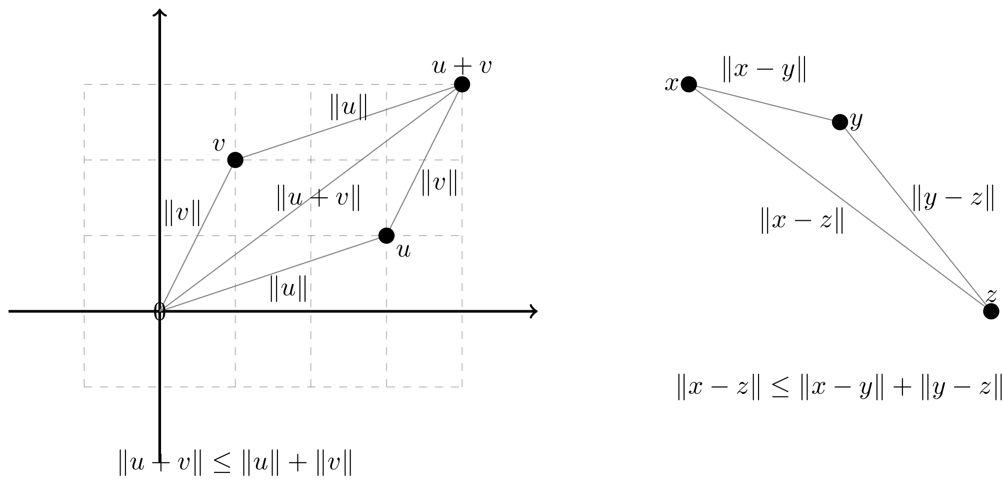


Figure 6: Triangle inequality of vectors.

1.2 p-summability, ℓ^p spaces

Definition. For an index set $M \neq \emptyset$ and for $u \in \mathbb{K}^M$, define the *little ℓ^p-norm*

$$\|u\|_{\ell^p} := \begin{cases} \left(\sum_{x\in M} |u(x)|^p\right)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \sup_{x\in M} |u(x)|, & \text{if } p = \infty. \end{cases}$$

Let $\ell^p = \ell^p(M) := \left\{u \in \mathbb{K}^M : \|u\|_{\ell^p} < \infty\right\}$ be the corresponding *little ℓ^p-space.*

Theorem. *$\ell^\infty = \ell^\infty(M)$ is a Banach space.*

Proof. The reader can check that ℓ^∞ is a vector space, and that $u \mapsto \|u\|_{\ell^\infty}$ is indeed a norm. We need to show that the norm metric is complete. Take a Cauchy sequence $(u_k)_{k=1}^\infty$ in ℓ^∞. For each $x \in M$, numbers $u_k(x) \in \mathbb{K}$ form a Cauchy sequence, because

$$|u_j(x) - u_k(x)| \leq \|u_j - u_k\|_{\ell^\infty} \overset{j,k\to\infty}{\longrightarrow} 0.$$

Hence due to the completeness of $\mathbb{K}$, we can define function $u : M \to \mathbb{K}$ by

$$u(x) := \lim_{k\to\infty} u_k(x).$$

Cauchy sequences in metric spaces are always bounded, so that $\|u_k\|_{\ell^\infty} \leq c$ for a constant $c < \infty$, for all $k \in \mathbb{Z}^+$. Then also $|u(x)| = \lim_k |u_k(x)| \leq c$ for all $x \in M$, yielding $\|u\|_{\ell^\infty} \leq c$. Thus $u \in \ell^\infty$. But does $\|u - u_k\|_{\ell^\infty} \to 0$ as $k \to \infty$? Fix $\varepsilon > 0$. Since $(u_k)_{k=1}^\infty$ is a Cauchy sequence, there is $N_\varepsilon \in \mathbb{Z}^+$ such that

$\|u_j - u_k\|_{\ell^\infty} < \varepsilon$ when $j,k > N_\varepsilon$. Hence if $j,k > N_\varepsilon$ then for all $x \in M$ we have

$$\begin{aligned} |u(x) - u_k(x)| &\leq |u(x) - u_j(x)| + |u_j(x) - u_k(x)| \\ &\leq |u(x) - u_j(x)| + \|u_j - u_k\|_{\ell^\infty} \\ &< |u(x) - u_j(x)| + \varepsilon \\ &\stackrel{j\to\infty}{\longrightarrow} \varepsilon. \end{aligned}$$

So, $\|u - u_k\|_{\ell^\infty} \leq \varepsilon$ whenever $k > N_\varepsilon$. Therefore $\lim\limits_{k\to\infty} u_k = u$ in space ℓ^∞. $\square$

Exercise. Let $1 \leq p < \infty$. Show that $\ell^p := \ell^p(M)$ is a Banach space with respect to the norm $u \mapsto \|u\|_{\ell^p}$. You may assume here *Minkowski's inequality*

$$\|u + v\|_{\ell^p} \leq \|u\|_{\ell^p} + \|v\|_{\ell^p}. \tag{11}$$

(Hint: Given a Cauchy sequence $(u_k)_{k=1}^\infty$ in ℓ^p, show that $(u_k(x))_{k=1}^\infty$ is a Cauchy sequence in Banach space $\mathbb{K}$, for each $x \in M$. Hence $u_k(x) \to u(x) \in \mathbb{K}$ defines $u : M \to \mathbb{K}$. If $\|u_k\|_{\ell^p} \leq c$, then $\sum_{x\in S} |u_k(x)|^p \leq c^p$ for all finite subsets $S \subset M$. Use this to deduce that $u \in \ell^p$. Finally, modify this deduction to show that $\|u - u_k\|_{\ell^p} \to 0$, as $k \to \infty$.)

Example. $V := \{u \in \ell^\infty(\mathbb{Z}) : \ \{x \in \mathbb{Z} : \ u(x) \neq 0\} \text{ finite}\}$ is not a dense subspace of $\ell^\infty(\mathbb{Z})$. The closure of V in $\ell^\infty(\mathbb{Z})$ is

$$c_0(\mathbb{Z}) := \left\{ u : \mathbb{Z} \to \mathbb{C} : \ \lim_{x\to\pm\infty} u(x) = 0 \right\}.$$

Exercise. Show that

$$V := \{u \in \ell^p(M) : \ \{x \in M : \ u(x) \neq 0\} \text{ finite}\}$$

is a dense subspace of $\ell^p(M)$ for $1 \leq p < \infty$.

Exercise. Norms s_1, s_2 on X are called *equivalent* if

$$a^{-1}\, s_1(u) \leq s_2(u) \leq a\, s_1(u)$$

for all $u \in X$, for a constant $a \geq 1$. Show that all norms on a finite-dimensional $\mathbb{K}$-vector space are equivalent. Consequently, a finite-dimensional normed space over field $\mathbb{K}$ is a Banach space, regardless of the chosen norm.
(Hint: It is enough to consider norms s_1, s_2 on $\mathbb{K}^n$, and take s_2 to be the ℓ^2-norm (why?). Recall that the closed ball in the Euclidean space is a compact set.)

1.3 p-integrability: Lebesgue's L^p spaces

Spaces $\ell^p = \ell^p(M)$ defined above were special cases of Lebesgue's L^p spaces, when M is endowed with the counting measure.

Informal example. Let M be a space with a positive measure. Let X^p consist of all measurable functions $u : M \to \mathbb{K}$ for which $s_p(u) < \infty$, where

$$s_p(u) := \begin{cases} \left[\int_M |u(x)|^p \,\mathrm{d}x\right]^{1/p}, & \text{if } 1 \le p < \infty, \\ \operatorname{ess\,sup}_{x \in M} |u(x)|, & \text{if } p = \infty. \end{cases}$$

Then $s_p : X^p \to \mathbb{R}$ is a seminorm. Let

$$Z^p := \{u \in X^p : \; s_p(u) = 0\}.$$

Define $L^p = L^p(M) := X^p/Z^p$, with the quotient norm

$$\|u\|_{L^p} := s_p(u).$$

Here we write simply $u \in L^p(M)$ instead of the correct notation

$$[u] = u + Z^p \in L^p(M).$$

Inequality $\|u + v\|_{L^p} \le \|u\|_{L^p} + \|v\|_{L^p}$ is called *Minkowski's inequality.*

1.4 Banach space of continuous functions

Let (M, d) be a compact metric space. Then

$$C(M) \subset \ell^\infty(M) \subset \mathbb{K}^M,$$

where $C(M)$ is the vector subspace of the continuous functions $u : M \to \mathbb{K}$. Endow $C(M)$ with the norm $u \mapsto \|u\|_{C(M)}$, where

$$\|u\|_{C(M)} := \max_{x \in M} |u(x)| \tag{12}$$

(remember: continuous $|u|$ has maximum in compact set M). Clearly, here $\|u\|_{C(M)} = \|u\|_{\ell^\infty}$ for all $u \in C(M)$. We already know that $\ell^\infty(M)$ is a Banach space. What about $C(M)$? We start investigating this question by an auxiliary result:

Lemma. *On compact metric spaces, continuity is uniform continuity.*

Proof. Let $u \in C(M)$, where (M, d) is a compact metric space. Continuity of u at point $x \in M$ means that for all $\varepsilon > 0$ there exists $\delta_{\varepsilon x} > 0$ such that

$$\forall x, y \in M : \quad d(x, y) < \delta_{\varepsilon x} \quad \Longrightarrow \quad |u(x) - u(y)| < \varepsilon. \tag{13}$$

Let $\mathbb{B}_d(x, r) = \{y \in M : d(x, y) < r\}$ be the open ball of radius $r > 0$ with center at $x \in M$. Then $\{\mathbb{B}_d(x, \delta_{\varepsilon x}) : \; x \in M\}$ is an open cover of the compact space M, thus having a subcover $\{\mathbb{B}_d(x, \delta_{\varepsilon x}) : \; x \in S\}$, where $S \subset M$ is a finite set. Defining

$$\delta_\varepsilon := \min_{x \in S} \delta_{\varepsilon x} > 0,$$

we obtain

$$\forall x, y \in M : \quad d(x,y) < \delta_\varepsilon \quad \Longrightarrow \quad |u(x) - u(y)| < \varepsilon, \tag{14}$$

i.e. u is uniformly continuous. $\square$

Theorem. *$C(M)$ is a Banach space when metric space (M, d) is compact.*

Proof. Suppose $(u_k)_{k=1}^\infty$ is a Cauchy sequence in $C(M)$. Then $(u_k)_{k=1}^\infty$ is also a Cauchy sequence in $\ell^\infty(M)$, which is a Banach space. We already know that

$$u \in \ell^\infty(M) \quad \text{for} \quad u(x) := \lim_{k\to\infty} u_k(x), \quad \text{and} \quad \lim_{k\to\infty} \|u - u_k\|_{\ell^\infty} = 0.$$

Is this limit function $u \in \ell^\infty(M)$ continuous? On compact metric spaces, continuity is uniform continuity. Take $\varepsilon > 0$. Take k such that $\|u - u_k\|_{\ell^\infty} < \varepsilon$. Take $\delta > 0$ such that $|u_k(x) - u_k(y)| < \varepsilon$ whenever $d(x,y) < \delta$ (by the uniform continuity of u_k). If $d(x,y) < \delta$, then

$$\begin{aligned} |u(x) - u(y)| &\leq |u(x) - u_k(x)| + |u_k(x) - u_k(y)| + |u_k(y) - u(y)| \\ &\leq \|u - u_k\|_{\ell^\infty} \quad + \quad \varepsilon \quad + \quad \|u_k - u\|_{\ell^\infty} \\ &\leq 3\varepsilon. \end{aligned}$$

Thus u is uniformly continuous. Thereby $C(M)$ is a Banach space. $\square$

Exercise. Let $a, b \in \mathbb{R}$ with $a < b$. For $M = [a,b]$ and for $u \in C(M)$ define $\|u\|$ as in (12), and let

$$\|u\|_1 := \int_a^b |u(x)| \, \mathrm{d}x.$$

Show that $C(M)$ is not complete with respect to the norm $u \mapsto \|u\|_1$.

1.5 Higher smoothness

Example. Let $k \in \mathbb{Z}^+ \cup \{0\}$. Let $X = C^k([a,b])$, the vector space of k times continuously differentiable functions $u : [a,b] \to \mathbb{K}$ (think of the derivatives at the end-points $a, b \in [a,b]$ as one-sided limits). Natural seminorms $s_0, \cdots, s_k$ on X would be given by

$$s_j(u) := \max_{x\in[a,b]} |u^{(j)}(x)|.$$

Here s_0 would actually be a norm. In a similar fashion, we could consider space

$$X = C^\infty([a,b]) = \bigcap_{k=0}^\infty C^k([a,b])$$

of infinitely smooth functions: there we would have infinitely many natural seminorms $s_0, s_1, s_2, s_3, \cdots$.

Remark. Let $s_1, \cdots, s_n : X \to \mathbb{R}$ be seminorms. Suppose that for all $u \in X$ there exists $k \in \{1, \cdots, n\}$ such that $s_k(u) > 0$. Then $s := s_1 + \cdots + s_n$ is a norm on X.

Exercise. Let $p_j : V \to \mathbb{K}$ is a seminorm for each $j \in \mathbb{Z}^+$. Suppose $p_j(u) = 0$ for all $j \in \mathbb{Z}^+$ only if $u = 0$. Construct a metric $d : V \times V \to \mathbb{R}$ such that each p_j is continuous in the metric topology and such that d is translation-invariant in the sense that $d(x + z, y + z) = d(x, y)$ for every $x, y, z \in V$.

Exercise. Let $\Omega \subset \mathbb{C}$ be open and non-empty. Endow the space of analytic functions $f : \Omega \to \mathbb{C}$ with semi-norms which give it a complete metric topology.

1.6 Semi-normed non-normable spaces

Example. Let us define the *Schwartz space* $\mathscr{S}(\mathbb{R})$ of test functions: $u \in \mathscr{S}(\mathbb{R})$ if $u : \mathbb{R} \to \mathbb{C}$ is infinitely smooth and

$$\lim_{|t| \to \infty} t^n \, u^{(m)}(t) = 0 \tag{15}$$

for all $m, n \in \mathbb{N} = \{0, 1, 2, 3, 4, 5, \cdots\}$. Natural seminorms would be $s_{m,n}$ given by

$$s_{m,n}(u) := \sup_{t \in \mathbb{R}} |t^n \, u^{(m)}(t)|.$$

The reader may easily imagine the seminorms for the Schwartz space $\mathscr{S}(\mathbb{R}^d)$ of multi-dimensional test functions $u : \mathbb{R}^d \to \mathbb{C}$.

Remark. Let $s_1, s_2, s_3, \cdots : X \to \mathbb{R}$ be seminorms. Suppose for all $u \in X$ there exists $k \in \mathbb{Z}^+$ such that $s_k(u) \neq 0$. Then X can be endowed with a natural metric, but not necessarily with a norm. For instance, the Schwartz space $\mathscr{S}(\mathbb{R}^d)$ of test functions $u : \mathbb{R}^d \to \mathbb{C}$ is not a normed space, but it has a natural countably infinite collection of seminorms.

1.7 Polynomial approximation of continuous functions

Continuous functions on compact intervals can be approximated by polynomials:

Weierstrass' Theorem (1885). *Polynomials are dense in* $C([a, b])$.

Proof. Evidently, it is enough to consider the case $[a, b] = [0, 1]$. Let $u \in C([0, 1])$, and let $v(x) = u(x) - (u(0) + (u(1) - u(0))x)$; then $v \in C(\mathbb{R})$ if we define $v(x) = 0$ for $x \in \mathbb{R} \setminus [0, 1]$. For $n \in \mathbb{N}$ let us define $k_n : \mathbb{R} \to [0, \infty)$ by

$$k_n(x) := \begin{cases} \frac{(1-x^2)^n}{\int_{-1}^{1} (1-t^2)^n \, \mathrm{d}t}, & \text{when } |x| < 1, \\ 0, & \text{when } |x| \geq 1. \end{cases}$$

Then define $p_n := v * k_n$ (*convolution* of v and k_n), that is

$$\begin{aligned} p_n(x) &= \int_{-\infty}^{\infty} v(x-t)\, k_n(t)\, \mathrm{d}t \\ &= \int_0^1 v(x-t)\, k_n(t)\, \mathrm{d}t, \end{aligned}$$

and from this last expression we see that p_n is a polynomial on $[0,1]$. Notice that p_n is real-valued if u is real-valued. Take any $\varepsilon > 0$. The function v is uniformly continuous, so that there exists $\delta > 0$ such that

$$\forall x, y \in \mathbb{R} : \ |x-y| < \delta \Rightarrow |v(x) - v(y)| < \varepsilon.$$

Let $\|v\| = \max_{t\in[0,1]} |v(t)|$. Take $x \in [0,1]$. Then

$$\begin{aligned} & |p_n(x) - v(x)| \\ = & \left| \int_{-\infty}^{\infty} v(x-t)\, k_n(t)\, \mathrm{d}t - v(x) \int_{-\infty}^{\infty} k_n(t)\, \mathrm{d}t \right| \\ = & \left| \int_{-1}^{1} (v(x-t) - v(x))\ k_n(t)\, \mathrm{d}t \right| \\ \leq & \int_{-1}^{1} |v(x-t) - v(x)|\ k_n(t)\, \mathrm{d}t \\ \leq & \int_{-1}^{-\delta} 2\|v\|\, k_n(t)\, \mathrm{d}t + \int_{-\delta}^{\delta} \varepsilon\, k_n(t)\, \mathrm{d}t + \int_{\delta}^{1} 2\|v\|\, k_n(t)\, \mathrm{d}t \\ \leq & 4\|v\| \int_{\delta}^{1} k_n(t)\, \mathrm{d}t + \varepsilon. \end{aligned}$$

The reader may verify that $\int_{\delta}^{1} k_n(t)\, \mathrm{d}t \to 0$ as $n \to \infty$, for every $\delta > 0$. Hence

$$\lim_{n\to\infty} \|q_n - u\| = 0,$$

where $q_n(x) = p_n(x) + u(0) + (u(1) - u(0))x$. □

Exercise. Why $\lim_{n\to\infty} \int_{\delta}^{1} k_n(t)\, \mathrm{d}t = 0$ in the proof of Weierstrass' Theorem?

2 Bounded operators

In vector spaces, linear operators are the most natural mappings. For them in normed spaces, continuity is equivalent to so-called boundedness. Related important concept is the completion of normed spaces to Banach spaces, with essentially unique opportunity to extend the relevant bounded operators.

Definition. Linear operator $A : X \to Y$ between normed spaces X, Y is called *bounded*, denoted by $A \in \mathscr{B}(X,Y)$, if

$$\|Au\| \leq \textit{constant}\, \|u\|$$

for all $u \in X$, for *constant* $< \infty$. Then such a minimal constant is the *norm* (or the *operator norm*)

$$\|A\| = \|A\|_{X\to Y} := \sup_{u\in X:\ \|u\|\leq 1} \|Au\|.$$

We often abbreviate $\mathscr{B}(X) := \mathscr{B}(X,X)$.

Proposition. *Let X, Y be normed spaces and $A : X \to Y$ a linear operator. Then the following conditions* (a,b,c) *are equivalent:*
(a) *A is bounded.*
(b) *A is continuous.*
(c) *A is continuous at $0 \in X$.*

Proof. Suppose A is bounded. Then it is (even Lipschitz) continuous, because

$$\|Au - Av\| = \|A(u-v)\| \leq \|A\|\, \|u-v\|.$$

Thus (a) implies (b). Condition (b) trivially implies (c). Finally, assume (c). This means that

$$\forall \varepsilon > 0\ \exists \delta_\varepsilon > 0 : \quad \|u - 0\| \leq \delta_\varepsilon \ \Rightarrow\ \|Au - A0\| \leq \varepsilon.$$

Especially, $\|Au\| \leq 1/\delta_1$ whenever $\|u\| \leq 1$. Hence $\|A\| \leq 1/\delta_1$. □

Remark. For normed spaces X, Y, the space $\mathscr{B}(X,Y)$ of bounded operators is a normed space: First, $\|A+B\| \leq \|A\| + \|B\|$, because

$$\|(A+B)u\| = \|Au + Bu\| \leq \|Au\| + \|Bu\| \leq \|A\|\|u\| + \|B\|\|u\|.$$

Notice that $\|\lambda A\| = |\lambda|\, \|A\|$, because $\|\lambda Au\| = |\lambda|\, \|Au\|$ for all $u \in X$. Moreover, $\|A\| = 0$ means $\|Au\| = 0$ for all $u \in X$, so $Au = 0$, i.e. $A = 0$.

Exercise. Show that $\mathscr{B}(X,Y)$ is a Banach space if Y is Banach.

Example. If $A \in \mathscr{B}(X,Y)$, $B \in \mathscr{B}(Y,Z)$, then $BA \in \mathscr{B}(X,Z)$. Why so? First, $BA : X \to Z$ is linear, because

$$BA(u+v) = B(Au+Av) = BAu + BAv,$$
$$BA(\lambda u) = B(\lambda Au) = \lambda\, BAu.$$

Finally, $\|BA\| \le \|B\|\|A\|$, since

$$\|BAu\| \le \|B\|\,\|Au\| \le \|B\|\,\|A\|\,\|u\|.$$

Example. Let $A : X \to Y$ be bijective and linear. Then $B := A^{-1} : Y \to X$ is linear, because if $v_k = Au_k$ then

$$B(v_1+v_2) = B(Au_1+Au_2) = BA(u_1+u_2) = u_1+u_2 = Bv_1+Bv_2,$$
$$B(\lambda v_1) = B(\lambda Au_1) = BA(\lambda u_1) = \lambda u_1 = \lambda Bv_1.$$

Here $B = A^{-1}$ is bounded if

$$\|Au\| \ge c\|u\|$$

for all $u \in X$, where $c > 0$ is a constant. Then $\|B(Au)\| = \|u\| \le \|Au\|/c$, so that $\|B\| \le 1/c$.

Exercise. Let X be a Banach space and $u_k \in X$ such that $\displaystyle\sum_{k=1}^{\infty} \|u_k\| < \infty$. Show that the vectors

$$v_N := \sum_{k=1}^{N} u_k \in X$$

form a Cauchy sequence $(v_N)_{N=1}^{\infty}$, thus converging to

$$v = \lim_{N\to\infty} v_N =: \sum_{k=1}^{\infty} u_k \in X.$$

Moreover, when Y is another Banach space and $A \in \mathscr{B}(X,Y)$, show that here

$$A\sum_{k=1}^{\infty} u_k = \sum_{k=1}^{\infty} Au_k.$$

Example. Let X be a Banach space. Let $Q \in \mathscr{B}(X)$ such that $\|Q\| < 1$. Then the so-called *Neumann series* $\displaystyle\sum_{k=0}^{\infty} Q^k$ converges, and

$$(I-Q)\sum_{k=0}^{\infty} Q^k = I = \left(\sum_{k=0}^{\infty} Q^k\right)(I-Q).$$

Thereby if $\|Q\| < 1$ then

$$(I-Q)^{-1} = \sum_{k=0}^{\infty} Q^k, \quad \left\|(I-Q)^{-1}\right\| \leq \sum_{k=0}^{\infty} \|Q\|^k = \frac{1}{1-\|Q\|}. \tag{16}$$

The Neumann series is a generalization of the geometric series of numbers.

Example. Regardless of convergence issues, let us write

$$(Au)_j := \sum_{k\in\mathbb{Z}} A_{jk} u_k,$$

where $u, Au : \mathbb{Z} \to \mathbb{C}$, $u_k = u(k)$, $(Au)_j = Au(j)$, and $A_{jk} \in \mathbb{C}$ for all $j,k \in \mathbb{Z}$. At least, the series converges if u here is finitely supported. It is easy to show that here $\|Au\|_{\ell^\infty} \leq \sup_{j,k\in\mathbb{Z}} |A_{jk}|\, \|u\|_{\ell^1}$.

Informal example. Let $X(M), X(N)$ be vector subspaces of nice-enough functions on M, N, respectively. Then formula

$$Av(x) = \int_N K(x,y)\, v(y)\, \mathrm{d}y$$

may define linear $A : X(M) \to X(N)$. If $K \in L^\infty(M \times N)$ then

$$|Av(x)| \leq \int_N |K(x,y)|\, |v(y)|\, \mathrm{d}y \leq \|K\|_{L^\infty} \int_N |v(y)|\, \mathrm{d}y,$$

so that we may think that we have a bounded linear operator

$$A : L^1(N) \to L^\infty(M),$$

with $\|Av\|_{L^\infty(M)} \leq \|K\|_{L^\infty}\, \|v\|_{L^1(N)}$.

Example. As a special case of the previous informal example, for $u \in L^1(\mathbb{R})$, the *Fourier transform* $\widehat{u} = \mathcal{F}u : \mathbb{R} \to \mathbb{C}$ is

$$\mathcal{F}u(\eta) := \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}2\pi y\cdot\eta}\, u(y)\, \mathrm{d}y.$$

Then $\|\mathcal{F}u\|_{L^\infty} \leq \|u\|_{L^1}$.

Example. For $u \in \ell^1(\mathbb{Z})$, the *Fourier transform* $\widehat{u} = \mathcal{F}u : \mathbb{R} \to \mathbb{C}$ is

$$\mathcal{F}u(\eta) := \sum_{y\in\mathbb{Z}} \mathrm{e}^{-\mathrm{i}2\pi y\cdot\eta}\, u(y).$$

Actually, here $\mathcal{F}u(\eta-1) = \mathcal{F}u(\eta)$; that is, function $\mathcal{F}u$ is 1-periodic, denoted by $\mathcal{F}u : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$. Then $\mathcal{F} : \ell^1(\mathbb{Z}) \to L^\infty(\mathbb{R}/\mathbb{Z})$ is a bounded linear operator, as $|\mathcal{F}u(\eta)| \leq \sum_{y\in\mathbb{Z}} |u(y)| = \|u\|_{\ell^1}$: this shows that $\|\mathcal{F}\|_{\ell^1\to L^\infty} \leq 1$. Actually, $\|\mathcal{F}\|_{\ell^1\to L^\infty} = 1$, because if $v(0) = 1$ and $v(x) = 0$ for $x \neq 0$ then $\mathcal{F}v(\eta) = 1$ for all η, so that then $\|\mathcal{F}v\|_{L^\infty} = 1 = \|v\|_{\ell^1}$.

Remark. If $A \in \mathscr{B}(X, Y)$ then

$$\ker(A) = \{u \in X : \ Au = 0\} = A^{-1}(\{0\})$$

is a closed vector subspace (because it is the inverse image of a closed set under a continuous map A). However, there is no guarantee that the vector subspace

$$\operatorname{ran}(A) = A(X) = \{Au \in Y : \ u \in X\}$$

would be closed.

Example. Let X be a Banach space and $P = P^2 \in \mathscr{B}(X)$ (that is, P is a bounded *projection*). Then $\operatorname{ran}(P)$ is closed, as $Pu = u \iff (I - P)u = 0$ implies here $\operatorname{ran}(P) = \ker(I - P)$, where naturally $I - P \in \mathscr{B}(X)$.

Exercise. Let $\|v\|_{C^k}$ denote the natural Banach space norm of $v \in C^k([a, b])$. Show that the differentiation $u \mapsto u'$ defines a bounded linear operator

$$A_k : C^{k+1}([a, b]) \to C^k([a, b]).$$

Exercise. Let $X = \{u \in C^2([0, 1]) : \ u(0) = 0 = u'(0)\}$. Show that the linear mapping $A = (u \mapsto u'') : X \to C([0, 1])$ is bounded and bijective, and that

$$A^{-1}v(x) = \int_0^x (x - y)\, v(y)\, \mathrm{d}y.$$

Exercise. Let $A : \mathbb{K}^n \to \mathbb{K}^m$ be a linear mapping defined by

$$(Au)_j := \sum_{k=1}^{n} A_{jk} u_k.$$

Show that

$$\begin{aligned}
\|A\|_{\ell^1 \to \ell^1} &= \max_{k \in \{1, \cdots, n\}} \sum_{j=1}^{m} |A_{jk}|, \\
\|A\|_{\ell^\infty \to \ell^\infty} &= \max_{j \in \{1, \cdots, m\}} \sum_{k=1}^{n} |A_{jk}|, \\
\|A\|_{\ell^2 \to \ell^2} &\leq \left(\sum_{j=1}^{m} \sum_{k=1}^{n} |A_{jk}|^2 \right)^{1/2}.
\end{aligned}$$

2.1 Banach space completion; extension of operators

We should not worry too much about non-complete normed spaces: in the following exercise it turns out that there always is a canonical way of "fixing the holes or the fringes", yielding a nice essentially unique Banach space.

Definition. Mapping $\kappa : Q \to R$ is a *completion* of metric space (Q, d_Q) if $\kappa(Q)$ is dense in complete metric space $(\mathbb{R}, d_R)$ and κ is an isometry: that is,

$$d_R(\kappa(x), \kappa(y)) = d_Q(x, y)$$

for all $x, y \in Q$.

Finding a completion. A completion (R, d_R) of a metric space (Q, d_Q) can be constructed as follows. For Cauchy sequences $u, v : \mathbb{Z}^+ \to Q$ define an equivalence relation by

$$u \sim v \iff \lim_{k\to\infty} d_Q(u(k), v(k)) = 0. \tag{17}$$

Let R be the set of the corresponding equivalence classes $[u] := \{v : \ u \sim v\}$, and equip it with the metric given by

$$d_R([u], [v]) := \lim_{k\to\infty} d_Q(u(k), v(k)). \tag{18}$$

Then $x \in Q$ can be identified with $\kappa(x) = [u_x] \in R$, where $u_x(k) := x$ for all $k \in \mathbb{Z}^+$. In other words, the mapping $\kappa = (x \mapsto [u_x]) : Q \to R$ embeds the metric space Q isometrically to a subset of R. It is customary to identify Q with the $\kappa(Q) \subset R$, often even writing simply $Q \subset R$.

Example. Think of completing the space $\mathbb{Q}$ of rational numbers to the space $\mathbb{R}$ of real numbers with respect to the absolute value metric $(x, y) \mapsto |x - y|$.

Exercise. Show that (R, d_R) in the example above is a complete metric space, and that $\kappa(Q) \subset R$ is dense.

Exercise. Show that a metric completion is unique in the following sense: If $\iota : Q \to S$ and $\kappa : Q \to R$ are metric completions, then there is a bijective isometry $\psi : S \to R$ with commuting diagram

$$\begin{array}{ccc} S & \xrightarrow{\psi} & R \\ \uparrow \iota & & \uparrow \kappa \\ Q & \equiv & Q \end{array}$$

Also, if Q is normed, show that S has a natural Banach structure, and that ι is linear: this is the *canonical Banach space completion* of a normed space.

Example. By Weierstrass' Approximation Theorem, polynomials are dense in $C([a, b])$. That is, for any continuous $u : [a, b] \to \mathbb{K}$ and for all $\varepsilon > 0$ there exists a polynomial $p : [a, b] \to \mathbb{K}$ such that

$$\|u - p\| := \max_{x\in[a,b]} |u(x) - p(x)| < \varepsilon.$$

Example. Endowing *Schwartz space* $X := \mathscr{S}(\mathbb{R})$ with the L^p-norm, we get $\widetilde{X} = L^p(\mathbb{R})$ when $1 \leq p < \infty$. However, $\mathscr{S}(\mathbb{R})$ is not dense in $L^\infty(\mathbb{R})$.

Above, we completed normed spaces to Banach spaces. The same procedure applies also to bounded linear operators between normed spaces, extending them to essentially uniquely to bounded linear operators between the respective Banach space completions. This is very useful, as it can be practically simpler to define operators on nice dense subspaces.

Exercise. Let $A \in \mathscr{B}(X,Y)$, where normed spaces X, Y have respective Banach completions $\iota_X : X \to \widetilde{X}$ and $\iota_Y : Y \to \widetilde{Y}$. Show that there is unique $\widetilde{A} \in \mathscr{B}(\widetilde{X}, \widetilde{Y})$ so that $\widetilde{A} \circ \iota_X = \iota_Y \circ A$, satisfying automatically $\|\widetilde{A}\| = \|A\|$. In other words, the following diagram commutes:

$$\begin{array}{ccc} \widetilde{X} & \xrightarrow{\widetilde{A}} & \widetilde{Y} \\ \uparrow \iota_X & & \uparrow \iota_Y \\ X & \xrightarrow{A} & Y \end{array}$$

To simplify notation, it is customary to write $\widetilde{A} = A$.

Example. The Fourier transform $(u \mapsto \widehat{u}) : \mathscr{S}(\mathbb{R}) \to \mathscr{S}(\mathbb{R})$ defined by

$$\widehat{u}(\eta) := \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}2\pi y \cdot \eta}\, u(y)\, \mathrm{d}y$$

is bijective, where $\|\widehat{u}\|_{L^\infty} \leq \|u\|_{L^1}$ and $\|\widehat{u}\|_{L^2} = \|u\|_{L^2}$. Thus this uniquely extends to a bounded linear mapping $L^1(\mathbb{R}) \to L^\infty(\mathbb{R})$ and to a linear isometric bijection $L^2(\mathbb{R}) \to L^2(\mathbb{R})$, where in the first case the Lebesgue integral interpretation is still valid, but in the L^2 case the Lebesgue integral does not make sense: then the extension is to be interpreted just as a limit.

Example. We study the boundedness of the Hilbert transform for square-integrable functions. We build on L. Grafakos' note [10]. For $v \in \ell^2 = \ell^2(\mathbb{Z})$, define the *discrete Hilbert transform* $Dv : \mathbb{Z} \to \mathbb{C}$ by

$$Dv(j) := \sum_{k \in \mathbb{Z} \setminus \{j\}} \frac{v(k)}{\pi(j-k)}. \tag{19}$$

Notice that D maps real-valued functions to real-valued ones. Thereby since $\|v\|^2 = \|\mathrm{Re}(v)\|^2 + \|\mathrm{Im}(v)\|^2$, it is enough to consider just real-valued v. Then

$$\begin{aligned}
\|Dv\|^2 &= \sum_{j\in\mathbb{Z}} \left(\sum_{k\in\mathbb{Z}\setminus\{j\}} \frac{v(k)}{\pi(j-k)} \right)^2 \\
&= \sum_{l\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} v(k)\, v(l) \sum_{j\in\mathbb{Z}\setminus\{k,l\}} \frac{1}{\pi^2(j-k)(j-l)} \\
&\le \sum_{l\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \frac{v(k)^2+v(l)^2}{2} \sum_{j\in\mathbb{Z}\setminus\{k,l\}} \frac{1}{\pi^2(j-k)(j-l)} \\
&= \sum_{l\in\mathbb{Z}} v(l)^2\, \frac{1}{\pi^2} \sum_{k\in\mathbb{Z}} \sum_{j\in\mathbb{Z}\setminus\{0,k\}} \frac{1}{(j-k)j}.
\end{aligned}$$

Here

$$\begin{aligned}
& \sum_{k\in\mathbb{Z}} \sum_{j\in\mathbb{Z}\setminus\{0,k\}} \frac{1}{(j-k)j} \\
= \;& \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{1}{k} \sum_{j\in\mathbb{Z}\setminus\{0,k\}} \left(\frac{1}{j-k} - \frac{1}{j} \right) + \sum_{j\in\mathbb{Z}\setminus\{0\}} \frac{1}{(j-0)j} \\
= \;& \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{1}{k} \lim_{N\to\infty} \left(\sum_{|j|\le N, j\ne k} \frac{1}{j-k} - \frac{1}{0-k} - \sum_{|j|\le N, j\ne 0} \frac{1}{j} + \frac{1}{k} \right) + \sum_{j\in\mathbb{Z}\setminus\{0\}} \frac{1}{j^2} \\
= \;& 6 \sum_{k=1}^{\infty} \frac{1}{k^2} \overset{\text{Euler}}{=} \pi^2.
\end{aligned}$$

Thus $\|Dv\| \le \|v\|$ for all $v \in \ell^2$, that is

$$\sum_{j\in\mathbb{Z}} \left| \sum_{k\in\mathbb{Z}\setminus\{j\}} \frac{v(k)}{\pi(j-k)} \right|^2 \le \sum_{j\in\mathbb{Z}} |v(j)|^2. \tag{20}$$

For compactly supported continuous $u : \mathbb{R} \to \mathbb{C}$, define the *Hilbert transform* $Au : \mathbb{R} \to \mathbb{C}$ by

$$Au(x) := \lim_{0<\varepsilon\to 0} \int_{|x-y|>\varepsilon} \frac{u(y)}{\pi(x-y)}\, \mathrm{d}y. \tag{21}$$

Then

$$\begin{aligned}
\|Au\|^2 &= \int |Au(x)|^2 \,\mathrm{d}x \\
&= \int \left| \lim_{0<\varepsilon\to 0} \int_{|x-y|>\varepsilon} \frac{u(y)}{\pi(x-y)} \,\mathrm{d}y \right|^2 \mathrm{d}x \\
&= \lim_{0<\varepsilon\to 0} \varepsilon \sum_{j\in\mathbb{Z}} \left| \varepsilon \sum_{k\in\mathbb{Z}\setminus\{j\}} \frac{u(k\varepsilon)}{\pi(j\varepsilon - k\varepsilon)} \right|^2 \\
&= \lim_{0<\varepsilon\to 0} \varepsilon \sum_{j\in\mathbb{Z}} \left| \sum_{k\in\mathbb{Z}\setminus\{j\}} \frac{u(k\varepsilon)}{\pi(j-k)} \right|^2 \\
&\overset{(20)}{\leq} \lim_{0<\varepsilon\to 0} \varepsilon \sum_{j\in\mathbb{Z}} |u(j\varepsilon)|^2 \quad = \quad \int |u(x)|^2 \,\mathrm{d}x \quad = \quad \|u\|^2.
\end{aligned}$$

As continuous compactly supported functions are dense in $L^2(\mathbb{R})$, the Hilbert transform extends uniquely to a linear mapping $A : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ for which $\|Au\| \leq \|u\|$ for all $u \in L^2(\mathbb{R})$. Exploiting the Fourier transform, it would actually turn out that $\widehat{Au}(\eta) = -\mathrm{i}\,\mathrm{sgn}(\eta)\,\widehat{u}(\eta)$, so that A would be a bijective isometry.

2.2 Rigidity of real normed spaces

Real normed spaces are rather rigid objects:

Mazur–Ulam Theorem. *Let X, Y be real normed spaces. Let $A : X \to Y$ be a bijective isometry such that $A(0) = 0$. Then A is linear.*

Proof. Take $u, v \in X$. Define their "midpoint set"

$$S_0 := \left\{ x \in X : \ \|x-u\| = \frac{\|u-v\|}{2} = \|x-v\| \right\}.$$

Set $S_0 \subset X$ is *symmetric at* $\frac{u+v}{2} \in S_0$ in the sense that

$$\forall h \in X : \ \frac{u+v}{2} + h \in S_0 \iff \frac{u+v}{2} - h \in S_0.$$

Define recursively sets $S_{n+1} \subset S_n$ such that

$$S_{n+1} := \left\{ x \in S_n : \ \|x - x_n\| \leq 2^{-n} \mathrm{diam}(S_n) \text{ for all } x_n \in S_n \right\},$$

where the *diameter* of $S \subset X$ is defined by $\mathrm{diam}(S) := \sup\{\|a-b\| : \ a, b \in S\}$. Then S_{n+1} is symmetric at $\frac{u+v}{2} \in S_{n+1}$. Clearly, $\bigcap_{n=0}^{\infty} S_n = \left\{ \frac{u+v}{2} \right\}$. Now

$$\begin{aligned}
A(S_{n+1}) &= \{A(x) : \ x \in S_{n+1}\} \\
&= \left\{ y \in A(S_n) : \ \|y - y_n\| \leq 2^{-n} \mathrm{diam}(S_n) \text{ for all } y_n \in A(S_n) \right\}
\end{aligned}$$

is symmetric at $\frac{A(u)+A(v)}{2} \in A(S_{n+1})$. Clearly, $\bigcap_{n=0}^{\infty} A(S_n) = \left\{ \frac{A(u)+A(v)}{2} \right\}$.
So,

$$A(\frac{u+v}{2}) = \frac{A(u)+A(v)}{2}, \tag{22}$$

especially implying

$$A(w/2) \overset{A(0)=0}{=} A(w)/2. \tag{23}$$

So we obtain the additivity

$$A(u+v) \overset{(23)}{=} 2\,A(\frac{u+v}{2}) \overset{(22)}{=} A(u)+A(v). \tag{24}$$

Especially, $0 = A(0) = A(u-u) = A(u)+A(-u)$, so that $A(-u) = A(u)$. Combining this to (23) and (24), we obtain

$$A(\lambda v) = \lambda A(v), \tag{25}$$

whenever $\lambda = k/2^n$ for $k, n \in \mathbb{Z}$. By continuity, (25) extends to all $\lambda \in \mathbb{R}$. Thus A is linear. □

Remark. The Mazur–Ulam Theorem does not hold in complex normed spaces. Think e.g. of $X = \mathbb{C}$, with isometry $A : \mathbb{C} \to \mathbb{C}$ such that $A(0) = 0$. Define $B : \mathbb{C} \to \mathbb{C}$ by $B(u) := A(u)/A(1)$. Now $B(0) = 0$ and $B(1) = 1$, so $B(u) = u$ for all $u \in \mathbb{R} \subset \mathbb{C}$. Then for all $u \in \mathbb{C}$ either $B(u) = u$ or $B(u) = u^*$. The situation becomes more complicated in higher-dimensional complex normed spaces.

2.3 Unbounded operators

Let X be an infinite-dimensional normed space. Let us show that there are unbounded linear operators $A : X \to X$. Let $\mathcal{B}$ be an algebraic basis for vector space X. In other words, there exists a unique set of linear functionals $(u \mapsto \langle u, b\rangle_{\mathcal{B}}) : X \to \mathbb{K}$ such that

$$u = \sum_{b\in\mathcal{B}} \langle u, b\rangle_{\mathcal{B}}\, b$$

for all $u \in X$, where $\langle u, b\rangle_{\mathcal{B}} \neq 0$ for at most finitely many $b \in \mathcal{B}$. Pick a function $f : \mathcal{B} \to \mathbb{K}$ so that

$$\sup_{b\in\mathcal{B}} \|f(b)\, b\| = \infty.$$

Define linear mapping $A : X \to X$ by

$$Au := \sum_{b\in\mathcal{B}} f(b)\, \langle u, b\rangle_{\mathcal{B}}\, b,$$

so that especially $Ab = f(b)\, b$. Such operator $A : X \to X$ is clearly unbounded.

3 Fruits of completeness

Next we study the role of completeness in normed spaces. Remember that a normed space can be embedded into a Banach space in a canonical fashion: thereby the possibly missing completeness is easy to cure in practise.

Baire's Theorem (or Baire's Category Theorem, [3]) deals with density in complete metric spaces. It will be the key to major results in Banach spaces: Zabreiko's Lemma, and Theorems on Uniform Boundedness (Banach–Steinhaus), Open Mapping (Banach–Schauder), and Closed Graph. For Banach–Steinhaus Theorem, we also present an alternative proof avoiding Baire's result [18].

Baire's Theorem. *Let $U_k \subset X$ be dense and open in a complete metric space X for each $k \in \mathbb{Z}^+$. Then $S = \bigcap_{k=1}^{\infty} U_k$ is dense.*

Proof. We show that $S \cap \mathbb{B}(v_0, r_0) \neq \emptyset$ for any $v_0 \in X$ and $r_0 > 0$: Assuming $X \neq \emptyset$, take v_1 and r_1 such that

$$\overline{\mathbb{B}(v_1, r_1)} \subset U_1 \cap \mathbb{B}(v_0, r_0).$$

Inductively, we choose v_{k+1} and $r_{k+1} < 1/k$ so that

$$\overline{\mathbb{B}(v_{k+1}, r_{k+1})} \subset U_{k+1} \cap \mathbb{B}(v_k, r_k).$$

So $(v_k)_{k=1}^{\infty}$ is Cauchy, converging to some $v \in X$ by completeness. By construction, $v \in S \cap \mathbb{B}(v_0, r_0)$. □

Example. As a concrete instance of Baire's Theorem, think of $X = \mathbb{R}$ with the absolute value metric, with dense open sets $U_k := \mathbb{R} \setminus \{u_k\}$, where the rational numbers are enumerated by $\mathbb{Q} = \{u_k : \ k \in \mathbb{Z}^+\}$. Then $S = \bigcap_{k=1}^{\infty} U_k = \mathbb{R} \setminus \mathbb{Q}$, the set of irrational numbers, which is a dense non-open subset of $\mathbb{R}$.

Exercise. Use Baire's Theorem to prove that an algebraic basis of an infinite-dimensional Banach space must be uncountable.

3.1 Zabreiko's Lemma and its consequences

The next intuitive result by Petr Petrovich Zabreiko [23] from 1969 will be invaluable when giving simple proofs to many other major results in Banach spaces:

Zabreiko's Lemma. *A seminorm s on a Banach space X is bounded if it is countably subadditive in the sense that*

$$s\Big(\sum_{n=0}^{\infty} u_n\Big) \leq \sum_{n=0}^{\infty} s(u_n) \quad \text{whenever} \quad \sum_{n=0}^{\infty} \|u_n\| < \infty. \tag{26}$$

In other words, then $s(u) \leq C\|u\|$ for all $u \in X$, where $C < \infty$ is a constant.

Proof. For $r > 0$, let $U_r := X \setminus \overline{\{s \le r\}}$, where we have the semiball

$$\{s \le r\} := \{v \in X : \ s(v) \le r\}.$$

Here $X = \bigcup_{k=1}^{\infty} \{s \le k\} = \bigcup_{k=1}^{\infty} \overline{\{s \le k\}}$, thus $\emptyset = \bigcap_{k=1}^{\infty} (X \setminus \overline{\{s \le k\}}) = \bigcap_{k=1}^{\infty} U_k$. As $U_k = kU_1$, we get $\overline{U_1} \neq X$ by Baire's Theorem. So the interior V of $X \setminus U_1 = \overline{\{s \le 1\}}$ is non-empty. Since $\{s \le 1\} = -\{s \le 1\}$ is convex, $V = -V$ is convex too. So if $v \in V$ then $0 = (-v+v)/2 \in V$. Thus $\varepsilon\mathbb{B} = \mathbb{B}(0,\varepsilon) \subset V$ for some $\varepsilon > 0$. Let $u \in \varepsilon\mathbb{B} \subset V = \mathrm{int}(\overline{\{s \le 1\}}) \subset \overline{\{s \le 1\}}$. Hence we can take $u_0 \in \{s \le 1\}$ such that $\|u - u_0\| < 2^{-1}\varepsilon$, i.e. $u - u_0 \in 2^{-1}\varepsilon\mathbb{B}$. Take inductively $u_n \in \{s \le 2^{-n}\}$ such that

$$\|u - \sum_{n=0}^{N-1} u_n\| < 2^{-N}\varepsilon.$$

Noticing that $u_N = (u - \sum_{n=0}^{N-1} u_n) - (u - \sum_{n=0}^{N} u_n)$, we get

$$\sum_{N=0}^{\infty} \|u_N\| \le \sum_{N=0}^{\infty} \left(2^{-N}\varepsilon + 2^{-N-1}\varepsilon\right) = 3\varepsilon < \infty,$$

so that we have

$$s(u) = s(\sum_{n=0}^{\infty} u_n) \overset{(26)}{\le} \sum_{n=0}^{\infty} s(u_n) \overset{u_n \in \{s \le 2^{-n}\}}{\le} \sum_{n=0}^{\infty} 2^{-n} = 2.$$

Hence $s(w) \le C\|w\|$ for all $w \in X$, where $C = 2/\varepsilon$. □

Immediate corollary. *Let X be a Banach space and V a normed space. Linear mapping $A : X \to V$ is bounded if*

$$\|A \sum_{n=0}^{\infty} u_n\| \le \sum_{n=0}^{\infty} \|Au_n\| \quad \text{whenever} \quad \sum_{n=0}^{\infty} \|u_n\| < \infty. \tag{27}$$

Proof. Just let $s(u) := \|Au\|$ in Zabreiko's Lemma. □

Exercise. Check the induction argument in the proof of Zabreiko's Lemma.

Exercise. Prove the following easy converse to Zabreiko's Lemma:
Bounded seminorms on Banach spaces are countably subadditive.

3.1.1 Uniform Boundedness Principle (Banach–Steinhaus Theorem)

The next result is by Stefan Banach and Hugo Steinhaus [2] from 1927, and it was independently discovered by Hans Hahn. This *Uniform Boundedness Principle* (or *Banach–Steinhaus Theorem*) states that for a family of bounded linear operators, pointwise bounds imply a uniform bound. We first prove this anachronistically by applying Zabreiko's Lemma from 1969, but later we give another independent proof in a more historical fashion. The result can be generalized considerably (e.g. as in [14]), but here we formulate a natural basic version; more precisely:

Uniform Boundedness Principle (Banach–Steinhaus Theorem). *Let X be a Banach space, V a normed space, and $\{A_\alpha\}_{\alpha\in J} \subset \mathscr{B}(X,V)$ such that*

$$\sup_{\alpha\in J} \|A_\alpha u\| < \infty$$

for every $u \in X$. Then $\sup_{\alpha\in J} \|A_\alpha\| < \infty$.

Proof. Let $s(u) := \sup_{\alpha\in J} \|A_\alpha u\|$. Whenever $\sum_{n=0}^{\infty} \|u_n\| < \infty$, we have

$$\begin{aligned} s(\sum_{n=0}^{\infty} u_n) = \sup_{\alpha\in J} \|A_\alpha \sum_{n=0}^{\infty} u_n\| &\overset{A_\alpha\in\mathscr{B}(X,V)}{=} \sup_{\alpha\in J} \|\sum_{n=0}^{\infty} A_\alpha u_n\| \\ &\leq \sup_{\alpha\in J} \sum_{n=0}^{\infty} \|A_\alpha u_n\| \\ &\leq \sum_{n=0}^{\infty} \sup_{\alpha\in J} \|A_\alpha u_n\| = \sum_{n=0}^{\infty} s(u_n). \end{aligned}$$

By Zabreiko's Lemma, $\sup_{\alpha\in J} \|A_\alpha u\| = s(u) \leq C\|u\|$, so that $\sup_{\alpha\in J} \|A_\alpha\| \leq C$. □

Exercise. Using the Uniform Boundedness Principle, the reader may then prove the following Corollary that often enables a nice way to define bounded linear operators:

Corollary. *In Banach spaces, let $A_k \in \mathscr{B}(X,Y)$ satisfy*

$$\lim_{k\to\infty} \|A_k u - Au\|_Y = 0 \quad \text{for all } u \in X. \tag{28}$$

Then $A = (u \mapsto Au) : X \to Y$ is a bounded linear operator.

Remark. Limit process (28) is called *strong convergence*, denoted by

$$A_k \overset{strong}{\longrightarrow} A.$$

This does not necessarily imply *norm convergence* $\|A_k - A\| \to 0$ (simply denoted by $A_k \longrightarrow A$), as discussed in exercises below. However, norm convergence trivially implies strong convergence: thus, the corresponding *norm operator topology* is stronger (i.e. has more open sets) than the *strong operator topology* of $\mathscr{B}(X,Y)$.

Example. Let $X = Y = L^2([0,1])$. Define the *Fourier transform* (or more precisely *Fourier coefficient transform*) $\widehat{u} : \mathbb{Z} \to \mathbb{C}$ of $u \in L^2([0,1])$ by

$$\widehat{u}(\eta) := \int_0^1 \mathrm{e}^{-\mathrm{i}2\pi y\cdot\eta}\, u(y)\,\mathrm{d}y.$$

Let

$$A_k u(x) := \sum_{\eta=-k}^{+k} \mathrm{e}^{\mathrm{i}2\pi x\cdot\eta}\, \widehat{u}(\eta).$$

Then $\|A_k\|_{X\to Y} = 1$. Actually, $A_k u \to Au = u$ in $L^2([0,1])$, as $k \to \infty$.

Exercise. Let $X = \ell^p(\mathbb{Z})$ for $1 \le p < \infty$. Define $P_N \in \mathscr{B}(X)$ by

$$P_N u(k) = \begin{cases} u(k) & \text{if } |k| \le N, \\ 0 & \text{otherwise.} \end{cases}$$

Show that P_N converges strongly to $I \in \mathscr{B}(X)$ as $N \to \infty$, but not in norm.

Exercise. Let $1 \le p < \infty$ and $A_k : L^p(\mathbb{R}) \to L^p(\mathbb{R})$ such that

$$A_k u(x) := u(x + 1/k).$$

Show that $A_k \overset{strong}{\longrightarrow} I$ as $k \to \infty$, but that $\|A_k - I\| \not\to 0$ as $k \to \infty$.

3.1.2 Closed Graph Theorem

Definition. The *graph* of a mapping $f : X \to Y$ is

$$\Gamma(f) \quad := \quad \{(u, f(u)) \in X \times Y \mid u \in X\}.$$

We endow $X \times Y$ with the natural vector space structure. For $(u,v) \in X \times Y$, define the Banach space norm e.g. by $\|(u,v)\| := \|u\| + \|v\|$.

Closed Graph Theorem. *Let X, Y be Banach Banach spaces, $A : X \to Y$ linear. Then A is continuous if and only if its graph is closed. (In other words, linear $A : X \to Y$ is bounded if and only if $\Gamma(A)$ is a Banach space.)*

Proof. First, let A be continuous. Let $((w_k, Aw_k))_{k=1}^{\infty}$ be a Cauchy sequence in $\Gamma(A) \subset X \times Y$. Then $(w_k)_{k=1}^{\infty}$ is a Cauchy sequence, converging to $w \in X$ by completeness. Then $Aw_k \to Aw$ by continuity. Now $\Gamma(A)$ is closed, because

$$(w_k, Aw_k) \to (w, Aw) \in \Gamma(A).$$

For the converse, let $\Gamma(A) \subset X \times Y$ be closed. Let $s(u) := \|Au\|$. When $\sum_{n=0}^{\infty} \|u_n\| < \infty$, we can define $u := \sum_{n=0}^{\infty} u_n$. Assuming $\sum_{n=0}^{\infty} s(u_n) < \infty$, we have

$$A \sum_{n=0}^{N} u_n = \sum_{n=0}^{N} Au_n \overset{N\to\infty}{\longrightarrow} \sum_{n=0}^{\infty} Au_n =: v \in Y.$$

Thus $(u, v) = (u, Au) \in \Gamma(A)$, because the graph is closed. Thereby

$$s(\sum_{n=0}^{\infty} u_n) = \|A \sum_{n=0}^{\infty} u_n\| \overset{Au=v}{=} \|\sum_{n=0}^{\infty} Au_n\| \leq \sum_{n=0}^{\infty} \|Au_n\| = \sum_{n=0}^{\infty} s(u_n).$$

By Zabreiko's Lemma, $\|Au\| = s(u) \leq C\|u\|$ for all $u \in X$. □

Exercise. Let $1 \leq p, q \leq \infty$. Let $B : L^p([0,1]) \to L^q([0,1])$ be a bounded linear operator such that $Bu \in C([0,1])$ for all $u \in C([0,1])$. Show that

$$(u \mapsto Bu) : C([0,1]) \to C([0,1])$$

is bounded.

Corollary. *Projection P is bounded if and only if* $\mathrm{ran}(P), \ker(P)$ *are closed.*

Proof. Let $P = P^2$. Then $I - P$ is also a projection, and $\ker(I-P) = \mathrm{ran}(P)$. If projection P is bounded then $\ker(P) = P^{-1}(\{0\})$ is closed, and then likewise $\mathrm{ran}(P) = \ker(I - P)$ is closed.

For the converse, assume now that $\mathrm{ran}(P), \ker(P) \subset X$ are closed for a projection P. Suppose $u_k \in X$ such that $(u_k, Pu_k) \to (u, v)$ in $X \times X$ as $k \to \infty$. Then $Pv = v$, because $P^2 = P$ and $\mathrm{ran}(P)$ is closed. Moreover,

$$\ker(P) \overset{P^2=P}{\ni} u_k - Pu_k \overset{k\to\infty}{\longrightarrow} u - v \in \ker(P),$$

because $\ker(P)$ is closed. Thereby $Pu = Pv = v$, and so $(u_k, Pu_k) \to (u, Pu)$ as $k \to \infty$. In other words, the graph of P is closed. Hence $P \in \mathscr{B}(X)$ by the Closed Graph Theorem. □

3.1.3 Open Mapping Theorem (Banach–Schauder Theorem)

Next we present the *Open Mapping Theorem* (or *Banach–Schauder Theorem*), and prove it by applying Zabreiko's Lemma. Later we shall give another proof for it. The Open Mapping Theorem speaks about the stability of solving linear equations $Au = v$; here $u \in X$ is unknown, while $v \in Y$ and $A \in \mathscr{B}(X, Y)$ are known:

Open Mapping Theorem (Banach–Schauder Theorem). *Let X, Y be Banach spaces and $A \in \mathscr{B}(X,Y)$ surjective. Then A is open, i.e. maps open sets to open sets. (Especially, if $A \in \mathscr{B}(X,Y)$ is bijective then A^{-1} is continuous.)*

Proof. For the closed vector subspace $Z := \ker(A) \subset X$, write $[u] := u + Z = \{u + z : z \in Z\}$. The the quotient space X/Z is a Banach space with the norm $\|[u]\| := \inf\{\|u+z\| : z \in Z\}$. The quotient map $\pi = (u \mapsto [u]) : X \to X/Z$ is a continuous open linear mapping, and $B = ([u] \mapsto Au) : X/Z \to Y$ is a continuous linear bijection. Here $A = B\pi$. It is enough to show that B is an open map. Define a seminorm s on Y by $s(v) := \|B^{-1}v\|$. Define $v := \sum_{n=0}^{\infty} v_n \in Y$ when $\sum_{n=0}^{\infty} \|v_n\| < \infty$. Assuming $\sum_{n=0}^{\infty} s(v_n) < \infty$, we have

$$s\Big(\sum_{n=0}^{\infty} v_n\Big) = \|B^{-1} \sum_{n=0}^{\infty} BB^{-1} v_n\| \overset{B \in \mathscr{B}(X/Z,Y)}{=} \|\sum_{n=0}^{\infty} B^{-1} v_n\| \leq \sum_{n=0}^{\infty} \|B^{-1} v_n\| = \sum_{n=0}^{\infty} s(v_n).$$

By Zabreiko's Lemma, $B^{-1} \in \mathscr{B}(Y, X/Z)$, so that $B : X/Z \to Y$ is open. $\square$

Remark. Later in 3.2.3, we prove the Closed Graph Theorem as a corollary to the Open Mapping Theorem. Thus by the following Exercise, we may think that the Closed Graph and the Open Mapping Theorems are logically equivalent:

Exercise. Use the Closed Graph Theorem to directly prove the Open Mapping Theorem (without another application of Zabreiko's Lemma).

Remarks. Once more, to emphasize the fact: In the Open Mapping Theorem, notice that the surjectivity of $A \in \mathscr{B}(X,Y)$ means that the linear equation

$$Au = v$$

has a solution $u \in X$ for every $v \in Y$. Then the openness of operator A means stability in finding such solutions! As an application of the Open Mapping Theorem, we learn that inclusion-comparable Banach space norm topologies must be the same, and the norms are essentially the same, too:

Corollary. *Let the same vector space $X = Y$ have Banach norm topologies τ_X, τ_Y such that $\tau_X \supset \tau_Y$. Then $\tau_X = \tau_Y$ and the norms are equivalent.*

Proof. The bijective linear mapping $(u \mapsto u) : X \to Y$ is trivially continuous, because $\tau_Y \subset \tau_X$. Hence the linear mapping $(u \mapsto u) : Y \to X$ is continuous by the Open Mapping Theorem, yielding $\tau_X \subset \tau_Y$. For linear mappings in normed spaces, continuity means boundedness, so we have $\|u\|_X \leq C\|u\|_Y \leq D\|u\|_X$ for some constants $C, D < \infty$ for all $u \in X = Y$. □

3.2 Alternative proofs without Zabreiko's Lemma

Historically, the Uniform Boundedness Principle, the Open Mapping Theorem and the Closed Graph Theorems were proven decades before Zabreiko's Lemma. Consequently, we shall now present alternative proofs for those major results, without relying directly on Zabreiko's Lemma.

3.2.1 Uniform Boundedness Principle (Banach–Steinhaus)

Uniform Boundedness Principle (Banach–Steinhaus Theorem). *Let X be a Banach space, let V be a normed space, and let $\{A_\alpha\}_{\alpha \in J} \subset \mathscr{B}(X, V)$ be such that*

$$\sup_{\alpha \in J} \|A_\alpha u\| < \infty$$

for every $u \in X$. Then $\sup_{\alpha \in J} \|A_\alpha\| < \infty$.

Proof. By *continuity* of A_α, for each $k \in \mathbb{Z}^+$, define *open* sets

$$U_k := \bigcup_{\alpha \in J} \{u \in X : \ \|A_\alpha u\| > k\} = \{u \in X : \ \sup_{\alpha \in J} \|A_\alpha u\| > k\}, \tag{29}$$

where $U_k = kU_1 \subset U_1$ and $\bigcap_{k=1}^{\infty} U_k = \emptyset$. Hence $\overline{U_1} \neq X$ by Baire's Theorem. Take $v \in X$ and $\delta > 0$ such that

$$\mathbb{B}(v, \delta) \subset X \setminus \overline{U_1}. \tag{30}$$

Finally, $\|A_\alpha\| \leq 1/\delta$ for every $\alpha \in J$, because if $\|z\| \leq 1$ then

$$\begin{aligned} 2\delta\|A_\alpha z\| &= \|A_\alpha(v + \delta z) - A_\alpha(v - \delta z)\| \\ &\leq \|A_\alpha(v + \delta z)\| + \|A_\alpha(v - \delta z)\| \overset{(29),(30)}{\leq} 2. \end{aligned}$$

□

Exercise. In the Banach–Steinhaus proof above, show that $C_1 := X \setminus U_1$ is a closed convex set, which is also *symmetric* in the sense that $-v \in C_1$ whenever $v \in C_1$. Furthermore, show that its interior

$$\mathrm{int}(C_1) = X \setminus \overline{U_1}$$

is an open convex symmetric set, containing especially the origin $0 \in X$.

Proof of Banach–Steinhaus (proof without Baire's Theorem, modified from [18]). First, by taking supremum over $u \in \mathbb{B}(0,r)$ in

$$\|A_\alpha u\| \le \frac{\|A_\alpha(u-v)\| + \|A_\alpha(u+v)\|}{2} \le \max\{\|A_\alpha(u-v)\|, \|A_\alpha(u+v)\|\}$$

for any $v \in X$ yields

$$r\,\|A_\alpha\| \le \sup_{w \in \mathbb{B}(v,r)} \|A_\alpha w\|. \tag{31}$$

To get a contradiction, suppose $\sup\limits_{\alpha \in J} \|A_\alpha\| = \infty$. Let $c, d > 1$ and $0 < \delta < 1$ (we refine these later): For $k \in \mathbb{Z}^+$, choose $\alpha_k \in J$ such that $\|A_{\alpha_k}\| \ge c^k$, and inductively choose $u_k \in X$ such that $u_0 = 0$, $\|u_k - u_{k-1}\| \le \delta^k$ and

$$\delta^k \|A_{\alpha_k}\| \le d\,\|A_{\alpha_k} u_k\| \tag{32}$$

by an application of (31). Due to completeness, Cauchy sequence $(u_k)_{k=1}^\infty$ converges to some $u \in X$. By geometric series, $\|u - u_k\| \le \delta^{k+1}/(1-\delta)$, and so

$$\begin{aligned} \|A_{\alpha_k} u\| &= \|A_{\alpha_k} u_k + A_{\alpha_k}(u - u_k)\| \\ &\ge \|A_{\alpha_k} u_k\| - \|A_{\alpha_k}\|\,\|u - u_k\| \\ &\overset{(32)}{\ge} \|A_{\alpha_k}\| \left(\delta^k/d - \delta^{k+1}/(1-\delta)\right) \\ &\ge \frac{1-(d+1)\delta}{d(1-\delta)}\,(c\delta)^k, \end{aligned}$$

which tends to ∞ as k grows (whenever $c\delta > 1 > (d+1)\delta$; so let $0 < \delta < 1/2$, and choose $c > 1/\delta$ and $1 < d < 1/\delta - 1$). But this is a contradiction! Hence we must have $\sup\limits_{\alpha \in J} \|A_\alpha\| < \infty$. □

3.2.2 Open Mapping Theorem (Banach–Schauder)

Now prove the Open Mapping Theorem without the help of Zabreiko's Lemma:

Open Mapping Theorem (Banach–Schauder Theorem). *Let X, Y be Banach spaces and $A \in \mathscr{B}(X,Y)$ surjective. Then A is open, i.e. maps open sets to open sets. (Especially, if $A \in \mathscr{B}(X,Y)$ is bijective then A^{-1} is continuous.)*

Proof. Let $\mathbb{B}_r := \mathbb{B}(0,r)$. Now

$$Y \overset{\text{surjection}}{=} A(X) = A(\bigcup_{k=1}^\infty \mathbb{B}_k).$$

Thus $\bigcap\limits_{k=1}^\infty U_k = \emptyset$ for $U_k := Y \setminus \overline{A(\mathbb{B}_k)}$. Here $U_k = kU_1 \subset U_1$. So $\overline{U_k} \ne Y$ by Baire's Theorem. Especially, $0 \notin \overline{U_k}$ (why?). Take $\varepsilon > 0$ with

$$\mathbb{B}_\varepsilon \subset Y \setminus \overline{U_1}.$$

Let

$$v \in \mathbb{B}_\varepsilon \subset Y \setminus U_1 = \overline{A(\mathbb{B}_1)}.$$

Inductively for $k \geq 0$, take $u_k \in \mathbb{B}_{2^{-k}}$ with $\left\| v - \sum_{j=0}^{k} A u_j \right\| < \frac{\varepsilon}{2^{k+1}}$. Now

$$\sum_{j=0}^{\infty} u_j =: u \in \mathbb{B}_2,$$

since X is complete. Then $v = Au$ by continuity of A. Thus

$$\mathbb{B}_\varepsilon \subset A(\mathbb{B}_2).$$

Hence A is open. □

Exercise. Supply details to the proof, especially the induction argument, and why $0 \notin \overline{U_k}$. Show also that $\mathbb{B}_\varepsilon \subset A(\mathbb{B}_1)$.

3.2.3 Closed Graph Theorem

Let us prove again the Closed Graph Theorem, this time as a corollary to the Open Mapping Theorem:

Closed Graph Theorem. *Let X, Y be Banach, $A : X \to Y$ linear. Then A is continuous if and only if its graph is closed. (Linear $A : X \to Y$ is bounded if and only if $\Gamma(A) \subset X \times Y$ is a Banach space.)*

Proof. First, suppose A is continuous. Let $((u_k, Au_k))_{k=1}^\infty$ be a Cauchy sequence in $\Gamma(A) \subset X \times Y$. Then $(u_k)_{k=1}^\infty$ is a Cauchy sequence, converging to $u \in X$ by completeness. Then $Au_k \to Au$ by continuity. Hence

$$(u_k, Au_k) \to (u, Au) \in \Gamma(A).$$

Thus $\Gamma(A)$ is closed.

Now let $\Gamma(A) \subset X \times Y$ be closed: thus $\Gamma(A)$ is a Banach subspace. Define

$$B := (u \mapsto (u, Au)) : X \to \Gamma(A).$$

This is a linear bijection, and it is continuous by the Open Mapping Theorem. So $A \in \mathscr{B}(X, Y)$. □

4 Duality in Banach spaces

When studying vector spaces, of special importance is duality: to an extent, dual X' might mirror the properties of space X. The dual consists of those bounded linear operators that take the space to its scalar field. Such operators are called *functionals*, and as one may expect, they are of uttermost importance in the field of Functional Analysis.

Definition. A *functional* on a $\mathbb{K}$-normed space X is a mapping $\varphi : X \to \mathbb{K}$. The *Banach dual* of X is

$$X' = \mathscr{B}(X,\mathbb{K}) := \{\varphi : X \to \mathbb{K} \mid \varphi \text{ bounded and linear}\}.$$

We also write

$$\langle u, \varphi\rangle := \varphi(u)$$

for $u \in X$ and $\varphi \in X'$. We equip X' with Banach space norm

$$\|\varphi\| := \sup\{|\langle u,\varphi\rangle| : \ u \in X,\ \|u\| \leq 1\}.$$

Indeed, X' is a Banach space even when the normed space X is not be complete.

Example. Let $X = \mathbb{K}^n$ and $\varphi \in X'$. Let $u \in X$. Then

$$u = (u_k)_{k=1}^n = (u_1, \cdots, u_n) = \sum_{k=1}^n u_k\, e_k \in X,$$

where $(e_k)_{k=1}^n$ is the standard basis of X: kth coordinate $(e_k)_k = 1$, and otherwise the coordinates are $(e_k)_j = 0$ for $j \neq k$. We have

$$\langle u, \varphi\rangle = \varphi(u) = \varphi(\sum_{k=1}^n u_k\, e_k) = \sum_{k=1}^n u_k\, \varphi(e_k).$$

Obviously, we may naturally identify $\varphi \in X'$ with vector $(\varphi(e_k))_{k=1}^n \in X$. In this sense, $X' \cong X$ here: $(\mathbb{K}^n)' \cong \mathbb{K}^n$.

Informal example. If $u \in L^1(M)$ and $w \in L^\infty(M)$ then

$$\langle u, \varphi\rangle = \varphi(u) := \int_M u(x)\, w(x)\, \mathrm{d}x$$

defines $\varphi \in (L^1(M))'$ of norm $\|\varphi\| \leq \|w\|_{L^\infty}$.

Example. Let $C(M)$ be the space of continuous functions $u : M \to \mathbb{K}$ on a compact space M. Then the evaluation at point $p \in M$ given by

$$(u \mapsto u(p)) : C(M) \to \mathbb{K}$$

is a bounded linear functional of norm 1.

Example. Let $L^p = L^p(\mathbb{R})$, where $1 < p < \infty$. Let $p^{-1} + q^{-1} = 1$. For $u \in L^p$ and $v \in L^q$ define

$$\varphi_v(u) = \langle u, \varphi_v \rangle := \int_{\mathbb{R}} u(x)\, v(x)\, \mathrm{d}x.$$

Then $(v \mapsto \varphi_v) : L^q \to (L^p)'$ is a Banach space isomorphism: $\|\varphi_v\|_{(L^p)'} = \|v\|_{L^q}$. Hence $(L^p(\mathbb{R}))' \cong L^q(\mathbb{R})$, especially $(L^2)' \cong L^2$.

Example. Let C_0 consist of continuous functions $u : \mathbb{R} \to \mathbb{K}$ for which $\lim_{|x|\to\infty} u(x) = 0$. Then in the spirit of the previous example, $(C_0)' \cong L^1$ and $(L^1)' \cong L^\infty$, but $C_0 \ncong (L^\infty)' \ncong L^1$.

Example. Let $\ell^p := \ell^p(\mathbb{Z}^+)$ and $c_0 := \{u \in \ell^\infty :\ \lim_{k\to\infty} u(k) = 0\}$. We have $(c_0)' \cong \ell^1$ and $(\ell^1)' \cong \ell^\infty$ but $c_0 \ncong (\ell^\infty)' \ncong \ell^1$. However, $(\ell^p)' \cong \ell^q$ if $1 < p < \infty$ and $p^{-1} + q^{-1} = 1$.

Exercise. Show that $(\ell^p)' \cong \ell^q$ if $1 < p < \infty$ and $p^{-1} + q^{-1} = 1$ by proving *Hölder's inequality*

$$\sum_{k=1}^{\infty} |u(k)\, v(k)| \leq \left(\sum_{j=1}^{\infty} |u(j)|^p \right)^{1/p} \left(\sum_{k=1}^{\infty} |v(k)|^q \right)^{1/q} = \|u\|_{\ell^p}\, \|v\|_{\ell^q}. \tag{33}$$

Example [F. Riesz]. For the space $\mathcal{M}(M)$ of signed measures on a compact metric space M, $(\mu \mapsto \varphi_\mu) : \mathcal{M}(M) \to C(M)'$ is a Banach isomorphism when

$$\varphi_\mu(f) := \int_M f\, \mathrm{d}\mu.$$

4.1 Hahn–Banach Theorem

Hahn–Banach Theorem is about extending bounded linear functionals from a vector subspace to the whole normed space. Notice that we do not need assume completeness here.

Hahn–Banach Theorem. *Let X be a normed $\mathbb{K}$-space and $\varphi : Z_\varphi \to \mathbb{K}$ be bounded and linear on a vector subspace $Z_\varphi \subset X$. Then there exists a bounded linear functional $\Phi : X \to \mathbb{K}$ so that $\Phi|_{Z_\varphi} = \varphi$ and $\|\Phi\| = \|\varphi\|$.*

Proof when $\mathbb{K} = \mathbb{R}$. Notice that $\varphi \in S \neq \emptyset$, where the set

$$S := \left\{ \alpha \in \mathscr{B}(Z_\alpha, \mathbb{R}) :\ Z_\varphi \subset Z_\alpha \subset X,\ \alpha|_{Z_\varphi} = \varphi,\ \|\alpha\| = \|\varphi\| \right\}$$

has partial order

$$\beta \leq \alpha \quad \Longleftrightarrow \quad Z_\beta \subset Z_\alpha :\ \beta = \alpha|_{Z_\beta}.$$

For any chain $(\varphi_\gamma)_{\gamma\in J} \subset S$, let $Z_\alpha := \bigcup_{\gamma\in J} Z_{\varphi_\gamma}$, $\alpha|_{Z_{\varphi_\gamma}} := \varphi_\gamma \leq \alpha$. Hence there exists a maximal element $\Phi \in S$ by Zorn's Lemma. Suppose $u_0 \in X \setminus Z_\Phi$. Let $Z := Z_\Phi + \mathbb{R}u_0$. For $a \in \mathbb{R}$ let

$$\begin{aligned} &\Psi_a : Z \to \mathbb{R}, \\ &\Psi_a(u + tu_0) := \Phi(u) + t\,a, \end{aligned}$$

so that $Z \neq Z_\Phi \subset Z \subset X$, $\Psi_a|_{Z_\Phi} = \Phi$ and $\|\Psi_a\| \geq \|\varphi\|$. Since

$$\begin{aligned} |\Phi(u) - \Phi(v)| &\leq \|\Phi\|\,\|u - v\| \\ &\leq \|\varphi\|\left(\|u - u_0\| + \|u_0 - v\|\right), \end{aligned}$$

there exists $a_0 \in \mathbb{R}$ such that

$$\Phi(u) - \|\varphi\|\,\|u + u_0\| \leq a_0 \leq \Phi(v) + \|\varphi\|\,\|v + u_0\|$$

for every $u, v \in Z_\Phi$ (notice that the assumption on realness was used here). Hence

$$|\Phi(w) - a_0| \leq \|\varphi\|\,\|w - u_0\|$$

for every $w \in Z_\Phi$. Thus

$$\begin{aligned} |\Psi_{a_0}(u - tu_0)| &= |\Phi(u) - t\,a_0| \\ &\overset{u=tw}{\leq} \|\varphi\|\,\|u - tu_0\|, \end{aligned}$$

i.e. $\|\Psi_{a_0}\| \leq \|\varphi\|$, so $\Psi_{a_0} \in S$: a contradiction! So $Z_\Phi = X$. □

Proof idea for $\mathbb{K} = \mathbb{C}$. With a pinch of salt, complex vector spaces can also be considered as real vector spaces (think e.g. of $\mathbb{C}$ and $\mathbb{R}^2$). Reduce the complex scalar case to the real scalar Hahn–Banach Theorem as follows: Let $\varphi = \varphi_1 + \mathrm{i}\varphi_2$, where $\varphi_k : Z_\varphi \to \mathbb{R}$ are bounded linear functionals such that $\|\varphi_k\| = \|\varphi\|$. Applying the real scalar Hahn–Banach Theorem, extend φ_k to $\Phi_k \in \mathscr{B}(X, \mathbb{R})$ such that $\|\Phi_k\| = \|\varphi\|$. Then let $\Phi = \Phi_1 + \mathrm{i}\Phi_2$. Finally, check that $\|\Phi\| = \|\varphi\|$. □

Exercise. Prove the complex scalar Hahn–Banach Theorem.

Remark. In the proof of Hahn–Banach Theorem above, we do not need completeness: the proof works in any normed space! However, in the treatment above we resorted to Zorn's Lemma (equivalent to Axiom of Choice, or to Well-Ordering Principle), but actually strictly weaker methods would have been enough.

Remark. Hahn–Banach Theorem provides lots of bounded linear functionals. The Banach dual X' actually *separates* the points of X: if $u, v \in X$ and $u \neq v$ then $\Phi(u) \neq \Phi(v)$ for some $\Phi \in X'$. We get Φ by the Hahn–Banach extension of $\varphi : Z_\varphi \to \mathbb{K}$, where $Z_\varphi := \{\lambda(u - v) :\ \lambda \in \mathbb{K}\}$ and $\varphi(\lambda(u - v)) := \lambda$.

Example. There exists a linear functional $\Phi : L^\infty(M) \to \mathbb{C}$ such that $\|\Phi\| = 1$ and $\Phi(u) = u(p)$ for all constant functions $u : M \to \mathbb{K}$.

Exercise. Show: if $v \in X \setminus Z$ (where $Z \subset X$ is a closed vector subspace) then there exists $\Phi \in X'$ such that $\Phi(v) \neq 0 = \Phi(u)$ for all $u \in Z$.
(Hint: Let $Z_\varphi := \{\lambda v - u : \ \lambda \in \mathbb{K}, \ u \in Z\}$ and $\varphi(\lambda v - u) := \lambda$.)

Exercise. Let X be a Banach space and $u \in X$. Show that

$$\|u\| = \sup \{|\langle u, \varphi\rangle| : \ \varphi \in X', \ \|\varphi\| \leq 1\}.$$

Remark. For $u \in X$ in Banach space X, define $u'' : X' \to \mathbb{K}$ by

$$u''(\Phi) := \Phi(u).$$

(Warning: here u'' is **not** the derivative of any u'.) By the previous exercise, $u'' \in (X')' = X''$ with $\|u''\| = \|u\|$, so $X \subset X''$. Here X is *reflexive* if $X = X''$. Actually, X is reflexive if and only if its closed balls are weakly compact: the *weak topology* of X is the smallest topology for which all $\Phi \in X'$ are continuous. So, bounded closed sets in a reflexive space are weakly compact.

Remark. Let $x \mapsto \|x\|$ be the norm of a vector space X over a field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. The dual space $X' = \mathscr{B}(X, \mathbb{K})$ of X is set of bounded linear functionals $f : X \to \mathbb{K}$, having a norm

$$\|f\| := \sup_{x \in X: \ \|x\| \leq 1} |f(x)|.$$

This endows X' with a Banach space structure. However, it is often better to use a weaker topology for the dual: let us define $x(f) := f(x)$ for every $x \in X$ and $f \in X'$; this gives the interpretation $X \subset X'' := \mathscr{B}(X', \mathbb{K})$, because

$$|x(f)| = |f(x)| \leq \|f\| \|x\|.$$

So we may treat X as a set of functions $X' \to \mathbb{K}$, and we define the *weak*-topology* of X' to be the X-induced topology of X' (the weakest topology that makes each $f \mapsto x(f)$ continuous).

Example. Let $L^p = L^p(\mathbb{R})$. Spaces L^p are reflexive for $1 < p < \infty$, and $(L^1)' \cong L^\infty$. Yet L^1 and L^∞ are not reflexive. Spaces $C([0,1])$ and $C_0(\mathbb{R})$ are not reflexive.

Exercise. Let $1 < p < \infty$. Show that $\ell^p = \ell^p(\mathbb{Z}^+)$ is reflexive. What about ℓ^1 and ℓ^∞?

Exercise. Show that $C([0,1])$ is not reflexive.

Exercise. Let V be a closed vector subspace of a reflexive Banach space X. Show that V and X/V are reflexive.

Exercise. Show that X is reflexive if and only if X' is reflexive.

Banach–Alaoglu Theorem. *Let X be a Banach space. Then the closed unit ball*

$$K := \overline{\mathbb{B}_{X'}(0,1)} = \{\phi \in X' : \ \|\phi\|_{X'} \leq 1\}$$

of X' is weak-compact.*

Proof. Due to Tihonov's Compactness Theorem,

$$P := \prod_{x \in X} \{\lambda \in \mathbb{C} : \ |\lambda| \leq \|x\|\} = \overline{\mathbb{D}(0, \|x\|)}^X$$

is compact in the product topology τ_P. Any element $f \in P$ is a mapping

$$f : X \to \mathbb{C} \quad \text{such that} \quad f(x) \leq \|x\|.$$

Hence $K = X' \cap P$. Let τ_1 and τ_2 be the relative topologies of K inherited from the weak*-topology $\tau_{X'}$ of X' and the product topology τ_P of P, respectively. We shall prove that $\tau_1 = \tau_2$ and that $K \subset P$ is closed; this would show that K is a compact Hausdorff space.

First, let $\phi \in X'$, $f \in P$, $S \subset X$, and $\delta > 0$. Define

$$\begin{aligned} U(\phi, S, \delta) &:= \{\psi \in X' : \ x \in S \Rightarrow |\psi x - \phi x| < \delta\}, \\ V(f, S, \delta) &:= \{g \in P : \ x \in S \Rightarrow |g(x) - f(x)| < \delta\}. \end{aligned}$$

Then

$$\begin{aligned} \mathcal{U} &:= \{U(\phi, S, \delta) \mid \phi \in X', \ S \subset X \text{ finite}, \ \delta > 0\}, \\ \mathcal{V} &:= \{V(f, S, \delta) \mid f \in P, \ S \subset X \text{ finite}, \ \delta > 0\} \end{aligned}$$

are bases for the topologies $\tau_{X'}$ and τ_P, respectively. Clearly

$$K \cap U(\phi, S, \delta) = K \cap V(\phi, S, \delta),$$

so that the topologies $\tau_{X'}$ and τ_P agree on K, i.e. $\tau_1 = \tau_2$.

Still we have to show that $K \subset P$ is closed. Let $f \in \overline{K} \subset P$. First we show that f is linear. Take $x, y \in X$, $\lambda, \mu \in \mathbb{C}$ and $\delta > 0$. Choose $\phi_\delta \in K$ such that

$$f \in V(\phi_\delta, \{x, y, \lambda x + \mu y\}, \delta).$$

Then

$$\begin{aligned} & |f(\lambda x + \mu y) - (\lambda f(x) + \mu f(y))| \\ \leq\ & |f(\lambda x + \mu y) - \phi_\delta(\lambda x + \mu y)| + |\phi_\delta(\lambda x + \mu y) - (\lambda f(x) + \mu f(y))| \\ =\ & |f(\lambda x + \mu y) - \phi_\delta(\lambda x + \mu y)| + |\lambda(\phi_\delta x - f(x)) + \mu(\phi_\delta y - f(y))| \\ \leq\ & |f(\lambda x + \mu y) - \phi_\delta(\lambda x + \mu y)| + |\lambda| \ |\phi_\delta x - f(x)| + |\mu| \ |\phi_\delta y - f(y)| \\ \leq\ & \delta \ (1 + |\lambda| + |\mu|). \end{aligned}$$

This holds for every $\delta > 0$, so that actually

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y),$$

f is linear! Moreover, $\|f\| \leq 1$, because

$$|f(x)| \leq |f(x) - \phi_\delta x| + |\phi_\delta x| \leq \delta + \|x\|.$$

Hence $f \in K$, K is closed. $\square$

Remark. The Banach–Alaoglu Theorem implies that a bounded weak*-closed subset of the dual space is a compact Hausdorff space in the relative weak*-topology. However, in a normed space norm-closed balls are compact if and only if the dimension is finite (see Riesz's Compactness Theorem in the next Chapter).

Exercise. Let X be a Banach space. Prove that X is reflexive if and only if its closed unit ball is compact in the weak topology. (Hint: Hahn–Banach and Banach–Alaoglu).

Definition. In a Banach space X, sequence $(u_k)_{k=1}^\infty$ *converges weakly* to a point $u \in X$ (denoted by $u_k \overset{weak}{\to} u$) if for every $\varphi \in X'$

$$\lim_{k \to \infty} \langle u_k - u, \varphi \rangle = 0.$$

Exercise. Show that if $\|u_k - u\| \to 0$ then $u_k \overset{weak}{\to} u$.

Exercise. Show that a weakly convergent sequence $(u_k)_{k=1}^\infty$ is uniformly bounded: $\|u_k\| \leq \mathit{constant} < \infty$.

Exercise. Give an example of a sequence converging weakly to 0, but not converging in norm as usual. (Hint: Consider Kronecker deltas in ℓ^2.)

4.2 Finite-dimensional projections

Theorem. *Let V be a finite-dimensional vector subspace of a Banach space X. Then $X = V \oplus W$ for a closed vector subspace $W \subset X$.*

Proof. Let $(v_k)_{k=1}^n$ be an algebraic basis of V. Choose the algebraic basis $(\varphi_j)_{j=1}^n$ for V' such that $\varphi_j(v_k) = \delta_{jk}$, where $\delta_{jk} \in \{0,1\}$ is the Kronecker delta (i.e. $\delta_{kk} = 1$, and $\delta_{jk} = 0$ if $j \neq k$). Then $\varphi_j \in V'$ has extension $\Phi_j \in X'$ by the Hahn–Banach Theorem. Define $P : X \to X$ by

$$Pu := \sum_{k=1}^{n} \langle u, \Phi_k \rangle \, v_k. \tag{34}$$

It is easy to check that P is a bounded linear projection with range V. Then $X = V \oplus W$ for the closed vector subspace $W := \ker(P)$. $\square$

4.3 Banach limits

Definition. Converging sequences $u : \mathbb{Z}^+ \to \mathbb{C}$ form the Banach subspace $c \subset \ell^\infty = \ell^\infty(\mathbb{Z}^+)$, where the functional $\lim : c \to \mathbb{C}$ is defined by

$$\lim(u) := \lim_{k\to\infty} u(k). \tag{35}$$

A bounded linear functional $\phi : \ell^\infty \to \mathbb{C}$ is called a *Banach limit* if it extends $\lim : c \to \mathbb{C}$ so that $\|\phi\| = 1$ and it is both *positive* ($v \geq 0 \Rightarrow \phi(v) \geq 0$) and *shift-invariant* ($\phi(v) = \phi(Lv)$, where $Lv(k) := v(k+1)$ defines the left-shift operator $L : \ell^\infty \to \ell^\infty$). A sequence $v \in \ell^\infty$ *almost converges to* $\lambda \in \mathbb{C}$ if $\phi(v) = \lambda$ for all Banach limits ϕ.

Example. Let $\phi : \ell^\infty \to \mathbb{C}$ be a Banach limit. Define sequences $v, \mathbf{1} \in \ell^\infty$ by $v(k) := (-1)^k$ and $\mathbf{1}(k) = 1$. Then $v \notin c$, so $\lim(v)$ is not defined. However,

$$\phi(v) = \phi(v+\mathbf{1})-1 = \frac{\phi(v+\mathbf{1}) + \phi(L(v+\mathbf{1}))}{2} - 1 = \frac{\phi(\mathbf{1}) + \phi(\mathbf{1})}{2} - 1 = 1-1 = 0.$$

Hence v almost converges to 0. Notice that Banach limits are *non-multiplicative*, as here $1 = \phi(v^2) \neq \phi(v)^2 = 0$.

Exercise. Let ϕ be a Banach limit and $v : \mathbb{Z}^+ \to \mathbb{R}$. Show that $\phi(v) \in \mathbb{R}$, where

$$\liminf_{k\to\infty} v(k) \leq \phi(v) \leq \limsup_{k\to\infty} v(k).$$

Exercise. Check that Banach limits do exist.
(Hint: Consider first the real scalars. Apply the Hahn–Banach Theorem to find bounded linear $\phi : \ell^\infty \to \mathbb{R}$ such that $\phi(\mathbf{1}) = 1$ and $\phi(v - Lv) = 0$ for all $v \in \ell^\infty$, with $\|\phi\| = 1$. Finally, consider the complex scalars.)

4.4 Adjoint operator

Just as dual X' might mirror space X, so may adjoint operator $A' \in \mathscr{B}(Y', X')$ mirror the properties of operator $A \in \mathscr{B}(X, Y)$; this is the message of the next exercise:

Exercise. Show that for $A \in \mathscr{B}(X, Y)$ there is unique $A' \in \mathscr{B}(Y', X')$ so that

$$\langle Av, \psi \rangle = \langle v, A'\psi \rangle$$

for every $v \in X$ and $\psi \in Y'$. Moreover, show:
(a) $\|A'\| = \|A\|$.
(b) $(BA)' = A'B'$ if $B \in \mathscr{B}(Y, Z)$.
(c) $(A^{-1})' = (A')^{-1}$ if A is invertible.

Definition. Let $A \in \mathscr{B}(X, Y)$ be as in the previous exercise. Then $A' \in \mathscr{B}(Y', X')$ is called the (*Banach*) *adjoint* of A.

Example. Linear $A : \mathbb{K}^m \to \mathbb{K}^n$ has a matrix $[A] = [A_{jk}] \in \mathbb{K}^{n\times m}$,

$$(Av)_j = \sum_{k=1}^{m} A_{jk}\, v_k \in \mathbb{K}.$$

Let us identify $(\mathbb{K}^\ell)'$ naturally with $\mathbb{K}^\ell$. Then adjoint $A' : \mathbb{K}^n \to \mathbb{K}^m$ satisfies

$$(A'\psi)_k = \sum_{j=1}^{n} A_{jk}\, \psi_j \in \mathbb{K}.$$

Informal example. Let $A : X(N) \to X(M)$, where

$$Av(x) = \int_N K_A(x, y)\, v(y)\, \mathrm{d}y.$$

The dual operator $A' : X(M)' \to X(N)'$ satisfies

$$\langle v, A'\psi\rangle = \langle Av, \psi\rangle.$$

Suppose $X(M)'$ and $X(N)'$ are isomorphic to spaces of functions on M and N, respectively. In the spirit of earlier informal examples,

$$\begin{aligned}
\int_N v(y)\, A'\psi(y)\, \mathrm{d}y &= \int_M Av(x)\, \psi(x)\, \mathrm{d}x \\
&= \int_M \left(\int_N K_A(x, y)\, v(y)\, \mathrm{d}y \right) \psi(x)\, \mathrm{d}x \\
&= \int_N v(y) \left(\int_M K_A(x, y)\, \psi(x)\, \mathrm{d}x \right) \mathrm{d}y
\end{aligned}$$

suggests that $K_A(x, y) = K_{A'}(y, x)$: that is,

$$A'\psi(y) = \int_M K_A(x, y)\, \psi(x)\, \mathrm{d}x.$$

Definition. Let X be a Banach space. The *annihilators* of $S \subset X$ and $F \subset X'$ are respectively

$$S^\perp := \{\varphi \in X' : \langle u, \varphi\rangle = 0 \text{ for all } u \in S\}, \tag{36}$$

$$^\perp F := \{u \in X : \langle u, \varphi\rangle = 0 \text{ for all } \varphi \in F\}. \tag{37}$$

Exercise. Show that $S^\perp \subset X'$ and $^\perp F \subset X$ are closed vector subspaces. Moreover, show that

$$\begin{aligned}
S &\subset {}^\perp(S^\perp) = \overline{\mathrm{span}(S)}, \\
F &\subset ({}^\perp F)^\perp = \overline{\mathrm{span}(F)}, \\
S^\perp &= ({}^\perp(S^\perp))^\perp, \\
{}^\perp F &= {}^\perp(({}^\perp F)^\perp).
\end{aligned}$$

Exercise. Let X, Y be Banach spaces and $A \in \mathscr{B}(X, Y)$. Show that

$$\begin{aligned}
\operatorname{ran}(A)^{\perp} &= \ker(A'), && (38)\\
\overline{\operatorname{ran}(A)} &= {}^{\perp}\ker(A'), && (39)\\
{}^{\perp}\operatorname{ran}(A') &= \ker(A), && (40)\\
\overline{\operatorname{ran}(A')} &= \ker(A)^{\perp}. && (41)
\end{aligned}$$

5 Compact operators

Compact operators are important special cases of bounded linear operators. For instance, those bounded linear operators that can be approximated by finite-rank operators turn out to be compact, but there might be more examples in Banach spaces (famous result by Per Enflo in 1973, see [6]). Nevertheless, a good initial idea is that compact operators boundedly squeeze the spaces to nearly finite dimensional; in Hilbert spaces, we make this claim precise in Chapter 11.

Compactness in topological spaces. First, for the reader's convenience, we recall some definitions and facts about compactness. First, let (X, τ) be a topological space. An *open cover* of a subset $K \subset X$ is a collection $\mathscr{U} \subset \tau$ of open sets such that $K \subset \bigcup \mathscr{U}$. Such K is called *compact* if its each open cover $\mathscr{U}$ has a *finite subcover*, i.e. a finite subset $\mathscr{V} \subset \mathscr{U}$ which is still an open cover of K. For instance, it is easy to prove that if $f : X \to Y$ is continuous, $K \subset X$ is compact and $C \subset X$ closed, then both $K \cap C \subset X$ and $f(K) \subset Y$ are compact. And if $(u_k)_{k=1}^{\infty}$ is a sequence in X converging to $u \in X$ then $\{u\} \cup \{u_k\}_{k=1}^{\infty} \subset X$ is a compact set.

Compactness in metric spaces. Now let us consider compactness in the metric topology of a metric space (X, d). It turns out that compactness is equivalent to so called sequential compactness: $K \subset X$ is *sequentially compact* if and only if each sequence $(u_j)_{j=1}^{\infty}$ in K has a subsequence $(u_{j_k})_{k=1}^{\infty}$ that converges to some point $v \in K$. How to prove the equivalence of the compactness and the sequential compactness in metric spaces? It is easy to see that in (X, d) the sequential compactness of $K \subset X$ is equivalent to K being both complete and totally bounded: *total boundedness* means that for each $\varepsilon > 0$ the open cover

$$\mathscr{U}_\varepsilon := \{\mathbb{B}(u, \varepsilon) : \ u \in K\}$$

for K has a finite subcover. It is clear that compactness implies the completeness and the total boundedness; so let us show the reverse implication. Let $\mathscr{U}$ be an open cover of a sequentially compact (i.e. complete totally bounded) set $K \subset X$. Total boundedness indicates that there is a dense countable set $S \subset K$: just consider the open covers $\mathscr{U}_\varepsilon$ when $\varepsilon \in \mathbb{Q}^+$, with S consisting of the centers of the balls in the respective finite subcovers. Let us define the countable set

$$J := \left\{(v, \varepsilon) \subset S \times \mathbb{Q}^+ : \ \mathbb{B}(v, \varepsilon) \subset W_{(v,\varepsilon)} \text{ for some } W_{(v,\varepsilon)} \in \mathscr{U}\right\}.$$

Clearly $\{W_{(v,\varepsilon)} : (v, \varepsilon) \in J\} =: \{U_k\}_{k=1}^{\infty}$ is a countable subcover of the totally bounded set K. Suppose that the open cover $\{U_k\}_{k=1}^{\infty} \subset \mathscr{U}$ of K would not have any finite subcover (so that then K would not be compact). Then take

$$u_k \in K \setminus \bigcup_{j=1}^{k} U_j.$$

By the sequential compactness, $(u_k)_{k=1}^{\infty}$ has a subsequence converging to some $v \in K$. But this is a contradiction, since $v \in U_j$ for some $j \in \mathbb{Z}^+$, but $u_k \notin U_j$ whenever $k \geq j$. Hence K must be compact.

Exercise. Let $X = \ell^2(\mathbb{Z}^+)$. Show that the *Hilbert cube*

$$K = \left\{ u \in X : \ |u(k)| \leq \frac{1}{k} \text{ for all } k \in \mathbb{Z}^+ \right\}$$

is a compact subset of X.

Definition. Let X, Y be Banach spaces. Linear mapping $A : X \to Y$ is *compact*, denoted by $A \in \mathscr{K}(X,Y)$, if the closure of $A(\mathbb{B}_X(0,1)) \subset Y$ is compact (so $\mathscr{K}(X,Y) \subset \mathscr{B}(X,Y)$, as compact sets in metric space are bounded). Let us write $\mathscr{K}(X) := \mathscr{K}(X,X)$.

Remark. A closed subset of a compact set is compact. So, $\overline{A(S)} \subset Y$ is compact, if $S \subset X$ is bounded and $A \in \mathscr{K}(X,Y)$.

Example. If $A \in \mathscr{B}(X,Y)$ has finite-dimensional range $A(X) \subset Y$ then $A \in \mathscr{K}(X,Y)$: in $A(X)$, apply the Heine–Borel Theorem (compact if and only if closed and bounded, when dimension is finite).

Example. If $A_1, K_1 \in \mathscr{B}(X,Y)$ and $A_2, K_2 \in \mathscr{B}(Y,Z)$ with compact K_1, K_2, then $K_2A_1, A_2K_1 \in \mathscr{K}(X,Z)$. Why? Let $\mathbb{B} = \mathbb{B}_X(0,1)$. Then $A_1(\mathbb{B})$ is bounded, so $\overline{K_2A_1(\mathbb{B})}$ is compact. Then $A_2(\overline{K_1(\mathbb{B})})$ is compact, as A_2 is continuous and $\overline{K_1(\mathbb{B})}$ is compact. Hence

$$\overline{A_2K_1(\mathbb{B})} \subset \overline{A_2(\overline{K_1(\mathbb{B})})} = A_2(\overline{K_1(\mathbb{B})})$$

is compact.

Remark. Linear $A : X \to Y$ is compact if and only if it maps *bounded sequences to sequences having a convergent subsequence.* I.e. if $(u_k)_{k=1}^{\infty}$ so that $\|u_k\|_X \leq$ *constant* for all k, then $(Au_k)_{k=1}^{\infty}$ has a subsequence $(Au_{k_j})_{j=1}^{\infty}$ which converges in Y. Why so? In metric spaces, compactness **is** *sequential compactness*!

Closure Lemma. *$\mathscr{K}(X,Y) \subset \mathscr{B}(X,Y)$ is closed.*

Proof. Let $\|A_n - A\| \to 0$ as $n \to \infty$, where $A_n \in \mathscr{K}(X,Y)$ and $A \in \mathscr{B}(X,Y)$. Let $\|u_k\|_X \leq 1$ and $u_{0,k} := u_k$. Now $(A_n u_{n-1,k})_{k=1}^{\infty}$ has a converging subsequence $(A_n u_{n,k})_{k=1}^{\infty}$. For $v_n := u_{n,n} \in X$, sequence $(Av_n)_{n=1}^{\infty}$ is Cauchy in Y, as

$$\begin{aligned}\|Av_n - Av_m\| &= \|A_n(v_n - v_m) + (A - A_n)(v_n - v_m)\| \\ &\leq \|A_n(v_n - v_m)\| + \|A - A_n\|(\|v_n\| + \|v_m\|) \\ &\leq \|A_n v_n - A_n v_m\| + 2\,\|A - A_n\| \\ &\overset{m>n\to\infty}{\longrightarrow} 0.\end{aligned}$$

Thus $(Av_n)_{n=1}^{\infty}$ converges in complete space Y, hence $A \in \mathscr{K}(X, Y)$. □

Remark. Hence, by combining earlier results: If $\lim_{n\to\infty} \|A_n - A\| = 0$, where $A_n \in \mathscr{B}(X, Y)$ and $\dim(A_n(X)) < \infty$, then $A \in \mathscr{K}(X, Y)$.

Exercise. Let $A' \in \mathscr{B}(Y', X')$ be the Banach adjoint of $A \in \mathscr{B}(X, Y)$. Show that A' is compact if and only if A is compact.

Exercise. Let $A \in \mathscr{B}(X)$ be compact, where X is a Banach space. Let λ be a non-zero scalar. Show that the range set

$$\operatorname{ran}(\lambda I - A) := (\lambda I - A)(X) = \{\lambda u - Au : \ u \in X\}$$

is closed, $\ker(\lambda I - A) = \{u \in X : Au = \lambda u\}$ is finite-dimensional, and that

$$\begin{aligned} & \dim\left(\ker(\lambda I - A)\right) \\ = \ & \dim\left(\ker(\lambda I - A')\right) \\ = \ & \dim\left(X/((\lambda I - A)(X))\right) \\ = \ & \dim\left(X'/((\lambda I - A')(X'))\right). \end{aligned}$$

5.1 Almost Orthogonality Lemma

Recall that in a metric space (M, d), the distance between non-empty subsets $Y, Z \subset M$ is

$$\operatorname{dist}(Y, Z) := \inf\{d(y, z) : \ y \in Y, \ z \in Z\}. \tag{42}$$

Notice that $u \in M$ is in the closure of Z if and only if $\operatorname{dist}(\{u\}, Z) = 0$.

In general, we may not talk about orthogonality and angles between vectors in Banach spaces; in Hilbert spaces we shall encounter no problems in this respect. However, it is possibly to deal with "almost orthogonality" in Banach spaces: To visualize the following result, think of X as a plane, with Z a line through the origin, with difficulty in exactly projecting orthogonally onto Z:

Almost Orthogonality Lemma [F. Riesz]. *Let X be a normed space with a closed vector subspace $Z \neq X$. Then for each $\varepsilon > 0$ there exists $u_\varepsilon \in X$ such that $\|u_\varepsilon\| = 1$ and $\operatorname{dist}(\{u_\varepsilon\}, Z) \geq 1 - \varepsilon$.*

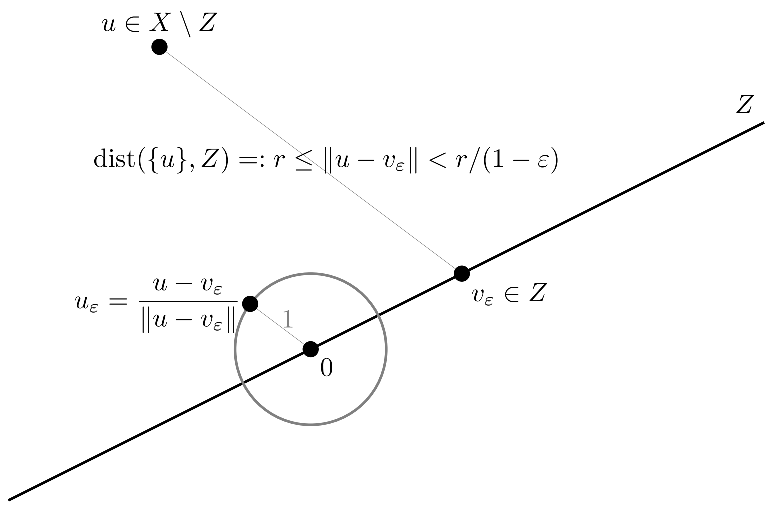


Figure 7: Proof of Riesz's Almost Orthogonality Lemma. (In the picture, in the Euclidean sense, the closest point in Z to $u \in X \setminus Z$ would be $0 \in Z$.)

Proof. Let $u \in X \setminus Z$ and $r := \text{dist}(\{u\}, Z)$. Here $r > 0$, because Z is closed. Let $0 < \varepsilon < 1$. Take $v_\varepsilon \in Z$ such that

$$r \leq \|u - v_\varepsilon\| < r/(1-\varepsilon).$$

Let $u_\varepsilon := (u - v_\varepsilon)/\|u - v_\varepsilon\|$. If $v \in Z$ then

$$\begin{aligned} \|u_\varepsilon - v\| &= \left\| \frac{u - v_\varepsilon}{\|u - v_\varepsilon\|} - v \right\| \\ &= \frac{\|u - (v_\varepsilon + \|u - v_\varepsilon\| v)\|}{\|u - v_\varepsilon\|} \\ &> \frac{r}{r/(1-\varepsilon)} \\ &= 1 - \varepsilon, \end{aligned}$$

showing that $\text{dist}(\{u_\varepsilon\}, Z) \geq 1 - \varepsilon$. □

Remark. As a corollary to the Almost Orthogonality Lemma, we learn that normed spaces with compact balls are of finite dimension:

Riesz's Compactness Theorem. *Let X be a normed space. Then X is finite-dimensional if and only if $\overline{\mathbb{B}(0,1)}$ is compact. (Normed space X is finite-dimensional if and only if the identity operator $I = (u \mapsto u) : X \to X$ is compact.)*

Proof. In finite-dimensional spaces, by the Heine–Borel Theorem, a closed set is compact if and only if bounded. Let X be infinite-dimensional. Let $0 < \varepsilon < 1$. Take $e_1 \in X$ such that $\|e_1\| = 1$. Inductively, let

$$Z_k := \text{span}\{e_j\}_{j=1}^k,$$

where $Z_k \neq X$ due to infinite-dimensionality, the Almost Orthogonality Lemma giving vectors $e_{k+1} \in X$ so that

$$\begin{aligned} \|e_{k+1}\| &= 1, \\ \text{dist}(\{e_{k+1}\}, Z_k) &\geq 1 - \varepsilon. \end{aligned}$$

Then sequence $(e_k)_{k=1}^\infty$ cannot have a converging subsequence. Hence $\overline{\mathbb{B}(0,1)}$ is not compact. $\square$

Remark. Riesz's Compactness Theorem tells us that if $A^{-1} \in \mathscr{B}(X)$ and $A \in \mathscr{K}(X)$ then X is finite-dimensional, because $I = A^{-1}A \in \mathscr{K}(X)$ here.

Corollary. *Let $A \in \mathscr{K}(X)$ bijective in Banach space X. Then X is finite-dimensional.*

Proof. Since $A \in \mathscr{K}(X) \subset \mathscr{B}(X)$ is bijective, here $A^{-1} \in \mathscr{B}(X)$ by the Open Mapping Theorem. But then $I = A^{-1}A \in \mathscr{K}(X)$ as a composition of a bounded operator and a compact operator. Hence X is finite-dimensional by Riesz's Compactness Theorem. $\square$

5.2 Hahn–Banach implies Riesz's Compactness Theorem

Riesz's Compactness Theorem can also be obtained as a follow-up of the Hahn–Banach Theorem, as in [5]:

Corollary (to Hahn–Banach). *Let X be a normed space. Then $\overline{\mathbb{B}(0,1)}$ is compact if and only if X is finite-dimensional.*

Proof. A set in a finite-dimensional normed space is compact if and only if it is bounded, by the Heine–Borel Theorem.

For the converse, suppose $\overline{\mathbb{B}(0,1)}$ is compact. Then the unit sphere $S := \{u \in X : \|u\| = 1\}$ is compact, and

$$\{S \cap \ker(\varphi) : \varphi \in \mathscr{B}(X, \mathbb{K})\}$$

is a family of compact sets, whose intersection is empty since $\mathscr{B}(X, \mathbb{K})$ separates the points of X (due to Hahn–Banach Theorem). Thereby there exists a finite set $\{\varphi_k\}_{k=1}^n \subset \mathscr{B}(X, \mathbb{K})$ such that

$$\bigcap_{k=1}^n S \cap \ker(\varphi_k) = \emptyset, \quad \text{i.e.} \quad \bigcap_{k=1}^n \ker(\varphi_k) = \{0\}.$$

As $\dim(X/\ker(\varphi_k)) \in \{0,1\}$, this implies that

$$\dim(X) = \dim(X/\bigcap_{k=1}^{n} \ker(\varphi_k)) \le \sum_{k=1}^{n} \dim(X/\ker(\varphi_k)) \le n. \qquad \square$$

5.3 Fredholm operators

Fredholm operators $A : X \to Y$ are "almost invertible" in the following sense:

Definition. Let X, Y be Banach spaces. Then $A \in \mathscr{B}(X,Y)$ is a *Fredholm operator* of *index* $\mathrm{ind}(A) \in \mathbb{Z}$ if $\ker(A), Y/\mathrm{ran}(A)$ are finite-dimensional, and

$$\mathrm{ind}(A) \quad := \quad \dim(\ker(A)) - \dim(Y/\mathrm{ran}(A)). \tag{43}$$

Example. Invertible $A \in \mathscr{B}(X,Y)$ is a Fredholm operator with $\mathrm{ind}(A) = 0$.

Example. Let $Z := \ell^p(\mathbb{Z}^+)$, and define the left- and right-shift operators $L, R : Z \to Z$ by

$$Lu(k) := u(k+1), \quad Ru(k+1) := u(k), \quad Ru(1) := 0.$$

Then L, R are Fredholm operators, with indices

$$\mathrm{ind}(L) = 1 - 0 = +1, \quad \mathrm{ind}(R) = 0 - 1 = -1.$$

It is easy to check that $\mathrm{ind}(L^n) = +n$ and that $\mathrm{ind}(R^n) = -n$ for all $n \in \mathbb{Z}^+$.

Exercise. Let $A \in \mathscr{B}(X,Y)$ be a Fredholm operator. Show that $A' \in \mathscr{B}(Y', X')$ is a Fredholm operator such that $\mathrm{ind}(A') = -\mathrm{ind}(A)$.

Definition. Let $(A_k)_{k=0}^n$ be a sequence of linear mappings $A_k : X_k \to X_{k+1}$. In other words,

$$X_0 \xrightarrow{A_0} X_1 \xrightarrow{A_1} X_2 \xrightarrow{A_2} \cdots \xrightarrow{A_{n-2}} X_{n-1} \xrightarrow{A_{n-1}} X_n \xrightarrow{A_n} X_{n+1}. \tag{44}$$

This is called an *exact sequence* if the vector spaces X_k are finite-dimensional such that $X_0 = \{0\} = X_{n+1}$ and $\mathrm{ran}(A_k) = \ker(A_{k+1})$ when $1 \le k < n$.

Remark. In the exact sequence (44), we observe that A_1 is injective as $A_0 = 0$, and A_{n-1} is surjective as $A_n = 0$. Especially in case $n = 2$, we have so-called *short exact sequence*

$$\{0\} \xrightarrow{A_0=0} X_1 \xrightarrow{A_1} X_2 \xrightarrow{A_2=0} \{0\},$$

where the linear mapping A_1 must be bijective.

Exact Lemma. *In the exact sequence* (44), *we have* $\sum_{k=1}^{n}(-1)^k\dim(X_k) = 0.$

Proof. When $n = 1$, we just have the trivial zero exact sequence

$$\{0\} = X_0 \xrightarrow{A_0=0} X_1 = \{0\} \xrightarrow{A_1=0} X_2 = \{0\}.$$

We shall reduce case $n+1$ to case n. Let $(B_k)_{k=0}^{n+1}$ be an exact sequence of operators $B_k : Y_k \to Y_{k+1}$, so that

$$\{0\} \xrightarrow{0} Y_1 \xrightarrow{B_1} Y_2 \xrightarrow{B_2} \cdots \longrightarrow Y_{n-1} \xrightarrow{B_{n-1}} Y_n \xrightarrow{B_n} Y_{n+1} \xrightarrow{0} \{0\},$$

where B_1 is injective and B_n surjective. Define an exact sequence $(A_k)_{k=0}^{n}$ of operators $A_k : X_k \to X_{k+1}$ as follows:

$$\begin{aligned} A_k &:= B_k \quad \text{when} \quad k < n-1, \\ A_{n-1} &:= (u \mapsto B_{n-1}u) : Y_{n-1} \to \ker(B_n). \end{aligned}$$

We have $\sum_{k=1}^{n}(-1)^k\dim(X_k) = 0$ by the induction hypothesis, and so

$$\begin{aligned} & \sum_{k=1}^{n+1}(-1)^k\dim(Y_k) \\ = \ & \sum_{k=1}^{n}(-1)^k\dim(X_k) + (-1)^n\left(\dim(Y_n) - \dim(Y_{n+1}) - \dim(X_n)\right) \\ = \ & (-1)^n\left(\dim(Y_n) - \dim(\operatorname{ran}(B_n)) - \dim(\ker(B_n))\right) \quad = \quad 0. \end{aligned}$$

This completes the proof. □

Index Theorem. *Let* $A \in \mathscr{B}(X,Y)$ *and* $B \in \mathscr{B}(Y,Z)$ *be Fredholm operators. Then* $BA \in \mathscr{B}(X,Z)$ *is a Fredholm operator, and*

$$\operatorname{ind}(BA) = \operatorname{ind}(A) + \operatorname{ind}(B). \tag{45}$$

Proof. First, $BA \in \mathscr{B}(X,Z)$ is a Fredholm operator, because

$$\begin{aligned} \dim(\ker(BA)) &\leq \dim(\ker(A)) + \dim(\ker(B)) &< \infty, \\ \dim(Z/\operatorname{ran}(BA)) &\leq \dim(Y/\operatorname{ran}(A)) + \dim(Z/\operatorname{ran}(B)) &< \infty. \end{aligned}$$

Define a sequence $(A_k)_{k=0}^{6}$ of linear operators $A_k : X_k \to X_{k+1}$, where

$$\begin{aligned} \{0\} &\xrightarrow{A_0} \ker(A) \xrightarrow{A_1} \ker(BA) \xrightarrow{A_2} \ker(B) \\ &\xrightarrow{A_3} Y/\operatorname{ran}(A) \xrightarrow{A_4} Z/\operatorname{ran}(BA) \xrightarrow{A_5} Z/\operatorname{ran}(B) \xrightarrow{A_6} \{0\}, \end{aligned}$$

$$
\begin{aligned}
A_1 u &:= u, \\
A_2 u &:= Au, \\
A_3 v &:= v + \mathrm{ran}(A), \\
A_4(v + \mathrm{ran}(A)) &:= Bv + \mathrm{ran}(BA), \\
A_5(w + \mathrm{ran}(BA)) &:= w + \mathrm{ran}(B).
\end{aligned}
$$

Here $(A_k)_{k=0}^6$ is an exact sequence, and so by the Exact Lemma

$$
\begin{aligned}
0 &= \sum_{k=1}^{6} (-1)^k \dim(X_k) \\
&= -\dim(\ker(A)) + \dim(\ker(BA)) - \dim(\ker(B)) + \\
&\quad +\dim(Y/\mathrm{ran}(A)) - \dim(Z/\mathrm{ran}(BA)) + \dim(Z/\mathrm{ran}(B)) \\
&= -\mathrm{ind}(A) + \mathrm{ind}(BA) - \mathrm{ind}(B).
\end{aligned}
$$

This completes the proof. □

Exercise. Check that $(A_k)_{k=0}^6$ in the proof of the Index Theorem is indeed an exact sequence.

Proposition. *Let $A : X \to Y$ be a Fredholm operator. Then $\mathrm{ran}(A) \subset Y$ is a closed vector subspace.*

Proof. First, $X = W \oplus \ker(A)$ for a closed vector subspace $W \subset X$. Let $(v_k + \mathrm{ran}(A))_{k=1}^n$ be an algebraic basis form $Y/\mathrm{ran}(A)$, where $v_k \in Y \setminus \mathrm{ran}(A)$. Then $Z := \mathrm{span}\{v_k\}_{k=1}^n$ is a vector subspace of $Y = \mathrm{ran}(A) \oplus Z$. Define operator $B : (X/\ker(A)) \oplus Z \to Y$ by

$$
B([u], v) := Au + v,
$$

where $[u] = u + \ker(A)$. Here define the norm of $([u], v) \in (X/\ker(A)) \oplus Z$ by $\|([u], v)\| := \|v\| + \inf\{\|w\| : \ w \in [u]\}$. Then B is a Banach space isomorphism by the Open Mapping Theorem, so that $\mathrm{ran}(A) = B((X/\ker(A)) \oplus \{0\})$ is closed, because it is the B-image of a closed set. □

Fredholm operators can be thought as "invertible modulo compact operators":

Corollary. *$A \in \mathscr{B}(X, Y)$ is a Fredholm operator if and only if there exists $B \in \mathscr{B}(Y, X)$ such that $AB - I$ and $BA - I$ are compact operators.*

Proof. First assume that $A \in \mathscr{B}(X, Y)$ is a Fredholm operator. Using the notation of the previous proof, we may identify Fredholm operator $A \in \mathscr{B}(X, Y)$ with operator $L \oplus 0 : W \oplus \ker(A) \to \mathrm{ran}(A) \oplus Z$, where

$$
L \oplus 0 = \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix}, \quad L : W \to \mathrm{ran}(A), \quad Lu := Au.
$$

Here $L := (u \mapsto Au) : W \to \operatorname{ran}(A)$ is a Banach space isomorphism, hence having the inverse $L^{-1} : \operatorname{ran}(A) \to W$. Define

$$B := L^{-1} \oplus 0 = \begin{bmatrix} L^{-1} & 0 \\ 0 & 0 \end{bmatrix} : \operatorname{ran}(A) \oplus Z \to W \oplus \ker(A).$$

Then $AB - I$ and $BA - I$ have finite-dimensional ranges, so they are compact.

Now suppose $A \in \mathscr{B}(X,Y)$ and $B \in \mathscr{B}(Y,X)$ such that $K := AB - I, BA - I$ are compact. By Riesz's Compactness Theorem, $\ker(BA) \subset X$ is then finite-dimensional (why?). Thereby $\ker(A) \subset \ker(BA)$ is finite-dimensional. On the other hand, $\dim(Y/\operatorname{ran}(A)) = \dim(\operatorname{ran}(A)^\perp) = \dim(\ker(A'))$, where $\ker(A') \subset \ker(B'A')$ is finite-dimensional, as $B'A' - I = (AB - I)' = K'$ is compact. $\square$

Exercise. In the previous proof, explain why Riesz's Compactness Theorem implies that $\ker(BA)$ is finite-dimensional when $BA - I$ is compact.

Exercise. Let $A \in \mathscr{B}(X,Y)$ be a Fredholm operator. Show that there exists $\varepsilon > 0$ such that if $B \in \mathscr{B}(X,Y)$ satisfies $\|A - B\| < \varepsilon$, then B is a Fredholm operator such that $\operatorname{ind}(B) = \operatorname{ind}(A)$.

Exercise. Suppose $A, K \in \mathscr{B}(X,Y)$, where A is Fredholm and K compact. Show that $A + K$ is a Fredholm operator such that $\operatorname{ind}(A + K) = \operatorname{ind}(A)$.

6 Spectral properties in Banach spaces

In this Chapter, we study the invertibility of linear operators on Banach spaces. First, we introduce Banach algebras.

Definition. A vector space $\mathscr{A}$ over the field $\mathbb{C}$ is an *algebra* with *unit* $I = I_{\mathscr{A}} \in \mathscr{A} \setminus \{0\}$, with *multiplication*

$$((A,B) \mapsto AB) : \mathscr{A} \times \mathscr{A} \to \mathscr{A},$$

if the mappings $(A \mapsto AB), (B \mapsto AB)$ are linear such that for all $A, B, C \in \mathscr{A}$

$$A(BC) = (AB)C, \tag{46}$$

$$IA = A = AI. \tag{47}$$

We briefly write $ABC := A(BC)$. If $AB = BA$ for every $A, B \in \mathscr{A}$ then algebra $\mathscr{A}$ is called *commutative*. Element $A \in \mathscr{A}$ is called *invertible* if there exists $A^{-1} \in \mathscr{A}$ such that

$$A^{-1}A = I = AA^{-1}. \tag{48}$$

The *inverse* A^{-1} is unique if it exists, because if $AB = I$ then

$$B = IB = A^{-1}AB = A^{-1}I = A^{-1}.$$

Definition. An algebra $\mathscr{A}$ (with unit I) is called a *(unital) Banach algebra* if it is a Banach space (where $\|I\| = 1$), satisfying for all $A, B \in \mathscr{A}$

$$\|AB\| \leq \|A\| \, \|B\|. \tag{49}$$

Exercise. Let K be a compact space. Show that $C(K)$ is a Banach algebra with the norm $u \mapsto \|u\| = \max_{x \in K} |u(x)|$.

Example. Let X be a Banach space. Then the Banach space $\mathscr{B}(X)$ of bounded linear operators $X \to X$ is a Banach algebra if the multiplication is the composition of operators, since $\|AB\| \leq \|A\| \, \|B\|$ for every $A, B \in \mathscr{B}(X)$. Here the unit is the identity operator $I = (u \mapsto u) : X \to X$. And obviously any (unital) norm-closed subalgebra of $\mathscr{B}(X)$ would a (unital) Banach algebra. Actually, this is not far away from characterising all the Banach algebras:

Theorem (Characterisation of Banach algebras). *Banach algebra $\mathscr{A}$ is isometrically isomorphic to a norm-closed subalgebra of $\mathscr{B}(X)$, where $X := \mathscr{A}$.*

Proof. For $A \in X := \mathscr{A}$, let us define

$$m(A) : X \to X \quad \text{by} \quad m(A)B := AB.$$

Obviously $m(A)$ is a linear mapping, $m(AB) = m(A)\,m(B)$, $m(I_{\mathscr{A}}) = I_X$, and

$$\begin{aligned} \|m(A)\| &= \sup_{B \in X:\ \|B\| \le 1} \|AB\| \\ &\le \sup_{B \in X:\ \|B\| \le 1} (\|A\|\,\|B\|) = \|A\| = \|m(A)\,I_X\| \\ &\le \|m(A)\|\,\|I_X\| = \|m(A)\|; \end{aligned}$$

briefly, $m = (A \mapsto m(A)) \in \mathrm{Hom}(\mathscr{A}, \mathscr{B}(X))$ is isometric. Thereby algebra $m(\mathscr{A}) \subset \mathscr{B}(X)$ is a closed subspace, and hence a Banach algebra. □

Exercise. Let X be a Banach space and $A \in \mathscr{B}(X)$. Let $\mathscr{A} \subset \mathscr{B}(X)$ be the algebra of all polynomials $\sum_{k=0}^{n} c_k A^k$, with $c_k \in \mathbb{C}$ and $n \in \mathbb{N}$ arbitrary. Show that the norm closure of $\mathscr{A}$ is a commutative Banach subalgebra of $\mathscr{B}(X)$.

Exercise. Let $\mathscr{A}$ be a Banach algebra, and $A, B \in \mathscr{A}$ such that $A^2 = A$, $B^2 = B$, $AB = BA$. Show that either $A = B$ or $\|A - B\| \ge 1$.
(Find also a low-dimensional example, where $AB \ne BA$ for non-orthogonal projections such that $\|A - B\| < 1$.)

Lemma (Continuity of inversion). *Let $\mathscr{A}$ be a Banach algebra. Then the set $G_{\mathscr{A}} \subset \mathscr{A}$ of its invertible elements is open. Moreover, the mapping $(A \mapsto A^{-1}) : G_{\mathscr{A}} \to G_{\mathscr{A}}$ is a homeomorphism.*

Proof. Take $A \in G_{\mathscr{A}}$ and $\varepsilon \in \mathscr{A}$. Apply the Neumann series to get

$$(A - \varepsilon)^{-1} = \left(I - A^{-1}\varepsilon\right)^{-1} A^{-1} = \sum_{k=0}^{\infty} (A^{-1}\varepsilon)^k A^{-1},$$

valid if $\|A^{-1}\|\,\|\varepsilon\| < 1$, that is $\|\varepsilon\| < \|A^{-1}\|^{-1}$; thus $G_{\mathscr{A}} \subset \mathscr{A}$ is open. Clearly $(A \mapsto A^{-1}) : G_{\mathscr{A}} \to G_{\mathscr{A}}$ is its own inverse, and thus it is continuous as

$$\begin{aligned} \|(A - \varepsilon)^{-1} - A^{-1}\| &= \|(I - A^{-1}\varepsilon)^{-1} A^{-1} - A^{-1}\| \\ &\le \|(I - A^{-1}\varepsilon)^{-1} - I\|\,\|A^{-1}\| \\ &= \left\| \sum_{k=1}^{\infty} (A^{-1}\varepsilon)^k \right\| \|A^{-1}\| \\ &\le \|\varepsilon\| \left(\sum_{k=1}^{\infty} \|A^{-1}\|^{k+1}\,\|\varepsilon\|^{k-1} \right) \quad \overset{\varepsilon \to 0}{\longrightarrow} \quad 0. \end{aligned}$$

Hence $A \mapsto A^{-1}$ is a homeomorphism. □

Exercise. In a Banach algebra $\mathscr{A}$, element $A \in \mathscr{A}$ is a *topological zero divisor* if there exists a sequence $(B_n)_{n=1}^{\infty}$ in $\mathscr{A}$ such that $\|B_n\| = 1$ for all n and

$$\lim_{n\to\infty} A\,B_n = 0 = \lim_{n\to\infty} B_n\,A.$$

(a) Show that if $G_{\mathscr{A}} \ni A_n \to A \in \partial G_{\mathscr{A}}$ as $n \to \infty$ then $\|A_n^{-1}\| \to \infty$ as $n \to \infty$.
(b) Show that the boundary points of $G_{\mathscr{A}}$ are topological zero divisors.

Definition. Let $\mathscr{A}$ be an algebra. The *spectrum* $\sigma(A) = \sigma_{\mathscr{A}}(A)$ of an element $A \in \mathscr{A}$ is the set

$$\sigma(A) := \{\lambda \in \mathbb{C} : \lambda I - A \text{ is not invertible in } \mathscr{A}\}.$$

Exercise. Let $\mathscr{A}$ be a Banach algebra. The *commutant* of a set $S \subset \mathscr{A}$ is

$$\mathscr{C}(S) := \{A \in \mathscr{A} : \ AB = BA \text{ for all } B \in S\}. \tag{50}$$

Prove the following claims:
a) $\mathscr{C}(S)$ is a Banach (sub)algebra.
b) $S \subset \mathscr{C}(\mathscr{C}(S))$.
c) If $AB = BA$ for all $A, B \in S$, then $\mathscr{B} := \mathscr{C}(\mathscr{C}(S))$ is a commutative Banach algebra such that $\sigma_{\mathscr{B}}(C) = \sigma_{\mathscr{A}}(C)$ for all $C \in \mathscr{B}$.

Exercise. Let $\mathscr{A}$ be an algebra, and $A \in \mathscr{A}$. Give an example, where $\sigma(A) = \emptyset \neq \sigma(A^2)$.

Exercise. Let $\mathscr{A}$ be an algebra, and $A \in \mathscr{A}$. Show that $\sigma(A) = \{0\}$ if A is nilpotent, i.e. if $A^k = 0$ for some $k \in \mathbb{Z}^+$. (Hint: Think of the "geometric sum" (the Neumann sum) $\displaystyle\sum_{j=0}^{k-1}(A/\lambda)^j$.)

Exercise. Let $\mathscr{A}$ be an algebra, and $A, B \in \mathscr{A}$. Show that $\sigma(AB) \cup \{0\} = \sigma(BA) \cup \{0\}$ in general, and that $\sigma(AB) = \sigma(BA)$ if A is invertible.

Example. Recall that as a consequence of the Open Mapping Theorem and Riesz's Compactness Theorem, compact linear bijections exists only in finite-dimensional Banach spaces. So if $A \in \mathscr{K}(X)$ where X is an infinite-dimensional Banach space then $0 \in \sigma(A)$.

Example. Let $X = C([0,1])$, and define $A \in \mathscr{B}(X)$ by

$$Au(x) := \int_0^x u(t)\,\mathrm{d}t.$$

Then $0 \in \sigma(A)$ is not an eigenvalue. Actually, it is easy to show that this A does not have any eigenvalues.

Theorem (Gelfand, 1939). *Let $\mathscr{A}$ be a Banach algebra, $A \in \mathscr{A}$. Then the spectrum $\sigma(A) := \sigma_{\mathscr{A}}(A) \subset \{z \in \mathbb{C} : \ |z| \leq \|A\|\}$ is a non-empty compact set.*

Proof. If $|\mu| > \|A\|$ then $\mu \notin \sigma(A)$ due to the convergent Neumann series

$$(\mu I - A)^{-1} = (I - A/\mu)^{-1}\mu^{-1} = \sum_{k=0}^{\infty} (A/\mu)^k \mu^{-1}.$$

Hence $\sigma(A) \subset \mathbb{C}$ is contained in the 0-centered closed disk of radius $\|A\|$. Also

$$\mathbb{C} \setminus \sigma(A) = g^{-1}(G_{\mathscr{A}}),$$

where $g = (\lambda \mapsto \lambda I - A) : \mathbb{C} \to \mathscr{A}$ is continuous. Thus $\sigma(A) \subset \mathbb{C}$ is closed, as $G_{\mathscr{A}} \subset \mathscr{A}$ is open. By the Heine–Borel Theorem, $\sigma(A) \subset \mathbb{C}$ is compact.

Showing that $\sigma(A) \neq \emptyset$ is more complicated. Let us start with defining the *resolvent mapping* $R : \mathbb{C} \setminus \sigma(A) \to G_{\mathscr{A}}$ by

$$R(z) := (zI - A)^{-1}. \tag{51}$$

Then $f \circ R : \mathbb{C} \setminus \sigma(A) \to \mathbb{C}$ is analytic for all $f \in \mathscr{A}' = \mathscr{B}(\mathscr{A}, \mathbb{C})$, because

$$\begin{aligned}
\frac{f(R(z+h)) - f(R(z))}{h} &= f\left(\frac{R(z+h) - R(z)}{h}\right) \\
&= f\left(\frac{R(z+h)R(z)^{-1} - I}{h}\, R(z)\right) \\
&= f\left(\frac{R(z+h)\left(R(z+h)^{-1} - hI\right) - I}{h}\, R(z)\right) \\
&= f(-R(z+h)R(z)) \\
&\overset{h \to 0}{\longrightarrow} f(-R(z)^2),
\end{aligned}$$

since f and R are continuous. Briefly, this means $(f \circ R)'(z) = f(-R(z)^2)$. Applying the Neumann series and the geometric series, we see that

$$\begin{aligned}
\|R(z)\| &= \left\|(zI - A)^{-1}\right\| = \left\|(I - A/z)^{-1}\right\| |z|^{-1} \\
&= \left\| \sum_{k=0}^{\infty} (A/z)^k \right\| |z|^{-1} \\
&\leq \sum_{k=0}^{\infty} \left(\frac{\|A\|}{|z|}\right)^k |z|^{-1} = \frac{1}{|z| - \|A\|} \\
&\overset{|z| \to \infty}{\longrightarrow} 0.
\end{aligned}$$

Thus $\lim_{|z| \to \infty} (f \circ R)(z) = 0$. What if $\sigma(A) = \emptyset$? Then $f \circ R = 0$ by Liouville's Theorem in complex analysis (for any $f \in \mathscr{A}'$). By the Hahn–Banach Theorem, $R(z) = 0$, which is a contradiction: 0 is not invertible. Hence $\sigma(A) \neq \emptyset$. □

Remark. Above we used Liouville's Theorem, stating the following:
All bounded analytic functions $g : \mathbb{C} \to \mathbb{C}$ are constants.

Proof. If here $g(z) = \sum_{k=0}^{\infty} c_k z^k$ then for all $r > 0$ we have

$$\sum_{k=0}^{\infty} |c_k|^2 \, r^{2k} = \int_0^1 |g(r\, \mathrm{e}^{\mathrm{i}2\pi t})|^2 \, \mathrm{d}t \le \sup_{z \in \mathbb{C}} |g(z)|^2 < \infty,$$

which implies $g(z) = c_0 = g(0)$ for all $z \in \mathbb{C}$. □

Exercise. Let $H = \ell^2(\mathbb{Z}^+)$ and linear $L : H \to H$, where $(Lu)_k := u_{k+1}$, that is for $u_k := u(k)$ we have

$$Lu = L(u_1, u_2, u_3, \cdots) = (u_2, u_3, u_4, \cdots).$$

(a) Is L bounded? Is it compact? Justify your answers!
(b) Find the spectrum $\sigma(L)$.
(Hint: For $\lambda \in \mathbb{C}$, is $\lambda \in \sigma(L)$? What happens if $|\lambda| < 1$? What if $|\lambda| > 1$?)

Exercise. Let $\mathscr{A}$ be a Banach algebra, $A \in \mathscr{A}$, $\Omega \subset \mathbb{C}$ an open set, and $\sigma(A) \subset \Omega$. Show that

$$\exists \delta > 0 \; \forall B \in \mathscr{A} : \; \|B\| < \delta \; \Rightarrow \; \sigma(A + B) \subset \Omega.$$

Corollary (Gelfand–Mazur Theorem). *Suppose 0 is the only non-invertible element of a Banach algebra $\mathscr{A}$. Then $\mathscr{A}$ is isometrically isomorphic to $\mathbb{C}$.*

Proof. Take $A \in \mathscr{A}$, $A \neq 0$. Since $\sigma(A) \neq \emptyset$, pick $\lambda(A) \in \sigma(A)$. Then $\lambda(A)I - A$ is non-invertible, so that it must be 0. Hence $A = \lambda(A)I$. By defining $\lambda(0) := 0$, we have an algebra isomorphism

$$\lambda : \mathscr{A} \to \mathbb{C}.$$

Moreover, $|\lambda(A)| = \|\lambda(A)I\| = \|A\|$. □

Exercise. Let $\mathscr{A}$ be a Banach algebra.
(a) Assume that 0 is the only topological zero divisor. Show that $\mathscr{A} \cong \mathbb{C}$ isometrically. (Hint: modify the Gelfand–Mazur proof.)
(b) Assume that there is a constant $k < \infty$ such that

$$\|A\| \, \|B\| \le k \, \|AB\|$$

for every $A, B \in \mathscr{A}$. Show that $\mathscr{A} \cong \mathbb{C}$ isometrically. (Hint: Apply (a).)

Definition. Let $\mathscr{A}$ be a Banach algebra. The *spectral radius* of $A \in \mathscr{A}$ is

$$\rho(A) := \sup_{\lambda \in \sigma(A)} |\lambda|; \tag{52}$$

this is well-defined, because $\sigma(A) \neq \emptyset$ due to Gelfand's Theorem. In other words, $\overline{\mathbb{D}(0,\rho(A))} = \{\lambda \in \mathbb{C} : \ |\lambda| \leq \rho(A)\}$ is the smallest 0-centered closed disk containing $\sigma(A) \subset \mathbb{C}$. Notice that $\rho(A) \leq \|A\|$, since $\lambda I - A = \lambda(I - A/\lambda)$ is invertible if $|\lambda| > \|A\|$.

Spectral Radius Formula (Beurling, 1938; Gelfand, 1939). *Let $\mathscr{A}$ be a Banach algebra, $A \in \mathscr{A}$. Then*

$$\rho(A) = \lim_{n \to \infty} \|A^n\|^{1/n}. \tag{53}$$

(Notice that trivially $\liminf_{n\to\infty} \|A^n\|^{1/n} \leq \limsup_{n\to\infty} \|A^n\|^{1/n} \leq \|A\| < \infty$.)

Proof. The claim is trivial if $A = 0$, so assume $A \neq 0$. By Gelfand's Theorem, $\sigma(A) \neq \emptyset$. Let $\lambda \in \sigma(A)$ and $n \in \mathbb{Z}^+$. Notice that in an algebra, if both BC and CB are invertible then the elements B, C are invertible. Therefore

$$\lambda^n I - A^n = (\lambda I - A)\left(\sum_{k=0}^{n-1} \lambda^{n-1-k} A^k\right) = \left(\sum_{k=0}^{n-1} \lambda^{n-1-k} A^k\right)(\lambda I - A)$$

implies that $\lambda^n \in \sigma(A^n)$. Thus $|\lambda^n| \leq \|A^n\|$, i.e. $|\lambda| \leq \|A^n\|^{1/n}$, so that

$$\rho(A) = \sup_{\lambda \in \sigma(A)} |\lambda| \leq \liminf_{n\to\infty} \|A^n\|^{1/n}.$$

Let $f \in \mathscr{A}'$ and $z \in \mathbb{C}$ such that $|z| > \|A\|$. By applying the Neumann series,

$$f(R(z)) = f\left((zI - A)^{-1}\right) = f\left(z^{-1}\sum_{n=0}^{\infty} A^n/z^n\right) = z^{-1}\sum_{n=0}^{\infty} f(A^n/z^n).$$

This is true for $|z| > \rho(A)$, as $f \circ R$ is analytic in $\mathbb{C} \setminus \sigma(A) \supset \mathbb{C} \setminus \overline{\mathbb{D}(0,\rho(A))}$. So, applying the Banach–Steinhaus Theorem to $\{B_n\}_{n=0}^{\infty} \subset \mathscr{A}''$, where $B_n(f) := f(A^n/z^n)$, we see that $\|B_n\| \leq M$ for all $n \in \mathbb{N}$ for a constant $M = M_{A/z} < \infty$ whenever $|z| > \rho(A)$. Then

$$\begin{aligned}
\|A^n\|^{1/n} &\overset{\text{Hahn-Banach}}{=} \sup_{f \in \mathscr{A}':\ \|f\| \leq 1} |f(A^n)|^{1/n} \\
&= \sup_{f \in \mathscr{A}':\ \|f\| \leq 1} |B_n(f)|^{1/n}\,|z| \\
&\leq M^{1/n}\,|z| \overset{n\to\infty}{\longrightarrow} |z|
\end{aligned}$$

whenever $|z| > \rho(A)$. Thus

$$\limsup_{n\to\infty} \|A^n\|^{1/n} \leq \rho(A).$$

This completes the proof of the Spectral Radius Formula. □

Remark. The Spectral Radius Formula (53) contains startling information. There was initially no guarantee that the limit of $\|A^n\|^{1/n}$ would exist. The spectral radius $\rho(A)$ is purely an algebraic property (though related to a topological algebra), but the limit of $\|A^n\|^{1/n}$ relies on both algebraic and metric properties. Yet the numbers are equal!

Remark. $\rho(A)^{-1}$ is the radius of convergence of the $\mathscr{A}$-valued power series

$$z \mapsto \sum_{n=0}^{\infty} z^n A^n.$$

Remark. Let $\mathscr{A}$ be a closed subalgebra of a Banach algebra $\mathscr{B}$. Then

$$\sigma_{\mathscr{B}}(A) \subset \sigma_{\mathscr{A}}(A)$$

for each $A \in \mathscr{A}$. This inclusion can be proper, but the spectral radii for both Banach algebras are the same, since

$$\rho_{\mathscr{A}}(A) = \lim_{n\to\infty} \|A^n\|^{1/n} = \rho_{\mathscr{B}}(A).$$

Exercise. Let $\mathscr{A}$ be a Banach algebra, $A, B \in \mathscr{A}$. Show that $\rho(AB) = \rho(BA)$. Show that if $A \in \mathscr{A}$ is *nilpotent* (i.e. $A^k = 0$ for some $k \in \mathbb{N}$) then $\sigma(A) = \{0\}$.

Exercise. Let $\mathscr{A}$ be a Banach algebra and $A, B \in \mathscr{A}$ such that $AB = BA$. Prove that $\rho(AB) \leq \rho(A)\,\rho(B)$.

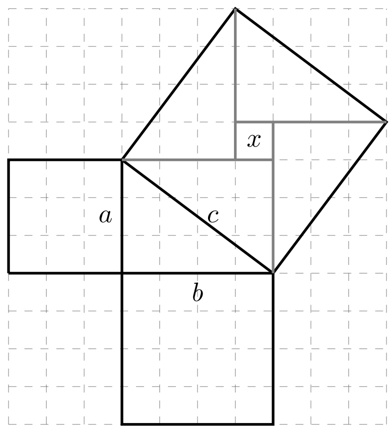


Figure 8: Geometry of Pythagoras' equation $a^2 + b^2 = c^2$, with $x = |a - b|$.

Notice! Hilbert spaces are special cases of Banach spaces. So whatever is true in Banach spaces is also true in Hilbert spaces (but not vice versa). Thus due to this logic, one could study first Banach and then Hilbert spaces (just as we have organized the lecture notes). However, the Hilbert spaces have so pleasant rich geometry that for this reason they would be the natural first introduction to infinite-dimensional vector spaces. Thus we shall present the Hilbert spaces without need for reading the previous chapters on Banach spaces!

7 Hilbert spaces

In order to understand functional analysis in Hilbert spaces, the reader does not actually have to first master the earlier text on Banach spaces, even though those notes would deal with more general topological vector spaces; in the first introduction to functional analysis, Hilbert spaces are actually a natural first place to start investigation. In vector spaces, norm tells us the distances in a uniform fashion. Inner product in tells us more refined information: not only distances, but also "angles" between vectors, especially sharp orthogonality (and not just "almost orthogonality", like in normed spaces). Hilbert spaces are those Banach spaces where this "angle information" is available, and they can be thought as generalizations of the usual Euclidean spaces $\mathbb{K}^n$. Thereby the reader is strongly encouraged to sketch planar pictures that would illuminate the proofs of the results in the sequel.

For instance, in the visualization of Pythagoras' equation, by studying the areas of the right-angled triangles and squares, we see that $c^2 = (a-b)^2 + 4ab/2$, yielding

$$a^2 + b^2 = c^2.$$

7.1 Inner product, Hilbert spaces

As before, the scalar field $\mathbb{K}$ is either the real field $\mathbb{R}$ or the complex field $\mathbb{C}$. The complex conjugate of $\lambda \in \mathbb{C}$ is denoted by $\lambda^* = \overline{\lambda} \in \mathbb{C}$.

Definition. Let H be a $\mathbb{K}$-vector space, where the scalar field is $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. A mapping

$$((u,v) \mapsto \langle u,v \rangle) : H \times H \to \mathbb{K}$$

is an *inner product* if for all $u, v \in H$ and $\lambda \in \mathbb{K}$ we have

$$\begin{aligned}
&\langle u+v, w \rangle = \langle u, w \rangle + \langle v, w \rangle, \\
&\langle \lambda u, v \rangle = \lambda \langle u, v \rangle, \\
&\langle v, u \rangle = \langle u, v \rangle^*, \\
&\langle u, u \rangle \geq 0, \\
&\langle u, u \rangle = 0 \quad \Rightarrow \quad u = 0.
\end{aligned}$$

Then H endowed with the inner product is an *inner product space*, and

$$\|u\| := \langle u, u \rangle^{1/2}$$

is called the *canonical norm* of $u \in H$. If there are several inner products available, we may emphasize the inner product space H by subscripts like $\langle u, v \rangle_H = \langle u, v \rangle$ and $\|u\|_H = \|u\|$.

Example. Why above $\langle v, u \rangle = \langle u, v \rangle^*$ and not $\langle v, u \rangle = \langle u, v \rangle$? If we would change this symmetry axiom this way for $\mathbb{K} = \mathbb{C}$, we would obtain

$$0 \leq \langle \mathrm{i}u, \mathrm{i}u \rangle = \mathrm{i}\langle u, \mathrm{i}u \rangle \overset{\langle u,v \rangle = \langle v,u \rangle}{=} \mathrm{i}\langle \mathrm{i}u, u \rangle = \mathrm{i}^2 \langle u, u \rangle = -\langle u, u \rangle \leq 0,$$

i.e. $\langle u, u \rangle = 0$ for all $u \in H$: thus $H = \{0\}$. So better have $\langle v, u \rangle = \langle u, v \rangle^*$.

Idea: Number $\|u\|^2 = \langle u, u \rangle \geq 0$ can be regarded as the "energy" of vector $u \in H$. Also, $\langle u, v \rangle = \|u\| \|v\| \cos(\alpha)$, where α is the angle between $u, v \in \mathbb{R}^2$.

Remark. Recall that $u \mapsto \|u\|$ is a *norm* on vector space H when

$$\begin{aligned}
\|u+v\| &\leq \|u\| + \|v\|, \\
\|\lambda u\| &= |\lambda| \, \|u\|, \\
u \neq 0 &\implies \|u\| > 0
\end{aligned}$$

for all $u \in H$ and $\lambda \in \mathbb{K}$, and then

$$d = ((u,v) \mapsto \|u - v\|) : H \times H \to \mathbb{R}$$

is a metric called the *norm metric* of H. The natural topology on H is then the *norm topology* given by this metric d. Inner product spaces can be considered as special cases of normed spaces: We shall soon show that the **canonical norm** $u \mapsto \|u\|$ is indeed a **norm** on H; the only non-trivial issue here is to prove the triangle inequality $\|u+v\| \leq \|u\| + \|v\|$, which soon follows from the Cauchy–Schwarz inequality.

Definition. Inner product space H is a *Hilbert space* if it is a complete metric space with respect to its canonical norm metric $d : H \times H \to \mathbb{R}$, where

$$d(u,v) := \|u-v\| = \langle u-v, u-v \rangle^{1/2}.$$

Thus, Hilbert spaces can be considered as special cases of Banach spaces (which are the normed vector spaces with the complete norm metric).

Exercise. For $u \in \mathbb{K}^M$ (that is, for functions $u : M \to \mathbb{K}$), let

$$\|u\|^2 := \sum_{x \in M} |u(x)|^2.$$

Show that

$$\ell^2(M) = \left\{ u \in \mathbb{K}^M : \ \|u\| < \infty \right\}$$

is a Hilbert space, where the inner product given by

$$\langle u, v \rangle = \sum_{x \in M} u(x)\, \overline{v(x)}. \tag{54}$$

Informal example. For a measurable function $u : M \to \mathbb{K}$, let

$$\|u\|^2 := \int_M |u(x)|^2 \, \mathrm{d}x.$$

Then $u \mapsto \|u\|$ defines the canonical norm of the Hilbert space $L^2(M)$ of (equivalence classes of) square-integrable functions on M, where the inner product given by

$$\langle u, v \rangle = \int_M u(x)\, v(x)^* \, \mathrm{d}x. \tag{55}$$

Example. For matrices $A, B \in \mathbb{K}^{d_1 \times d_2}$, define the *Hilbert–Schmidt* inner product $(A,B) \mapsto \langle A, B \rangle_{HS}$ by

$$\begin{aligned}
\langle A, B \rangle_{HS} &:= \mathrm{Tr}(A\,B^*) \\
&= \sum_{j=1}^{d_1} (A\,B^*)_{jj} \\
&= \sum_{j=1}^{d_1} \sum_{k=1}^{d_2} A_{jk} (B^*)_{kj} \\
&= \sum_{j=1}^{d_1} \sum_{k=1}^{d_2} A_{jk}\, (B_{jk})^*.
\end{aligned}$$

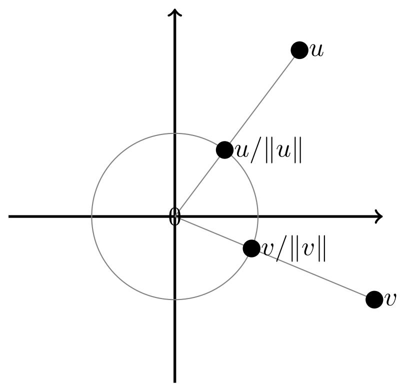


Figure 9: Unit normalizations $u/\|u\|$ and $v/\|v\|$ of vectors $u \neq 0$ and $v \neq 0$.

Then $\mathbb{K}^{d_1 \times d_2}$ is a Hilbert space, where the canonical norm $A \mapsto \|A\|_{HS}$ satisfies

$$\|A\|_{HS}^2 = \langle A, A \rangle_{HS} = \sum_{j=1}^{d_1} \sum_{k=1}^{d_2} |A_{jk}|^2.$$

7.2 Geometric inequalities

The inner product is naturally bounded by the canonical norm:

Cauchy–Schwarz inequality. *Let X be an inner product space. Then for every $u, v \in X$*

$$|\langle u, v \rangle| \leq \|u\| \, \|v\|. \tag{56}$$

Proof. Assume the non-trivial case $0 < \|u\| \, \|v\|$. Notice that then

$$|\langle u, v \rangle| = \|u\| \, \|v\| \left\langle \frac{u}{\|u\|}, \frac{v}{\|v\|} \right\rangle,$$

so that we may assume $\|u\| = 1 = \|v\|$. Let $|\lambda| = 1$. Then

$$\begin{aligned} 0 &\leq \|\lambda u - v\|^2 \\ &= \|\lambda u\|^2 + \|v\|^2 - \langle \lambda u, v \rangle - \langle v, \lambda u \rangle \\ &= 2 - 2\,\mathrm{Re}(\lambda \langle u, v \rangle), \end{aligned}$$

so $\mathrm{Re}(\lambda \langle u, v \rangle) \leq 1$. Especially, $|\langle u, v \rangle| \leq 1$. □

Example. For $u, v \in \ell^2(M)$ with normalization $\|u\| = \|v\|$, we obtain (56) also directly from the triangle inequality $(\star)$ of scalars:

$$\begin{aligned} |\langle u, v\rangle| &= \left|\sum_{x\in M} u(x)\,\overline{v(x)}\right| \\ &\overset{(\star)}{\leq} \sum_{x\in M} |u(x)|\,|v(x)| \\ &\overset{0\leq(|u(x)|-|v(x)|)^2}{\leq} \sum_{x\in M} \frac{|u(x)|^2+|v(x)|^2}{2} \overset{\|u\|=\|v\|}{=} \|u\|\,\|v\|. \end{aligned}$$

A similar reasoning would hold for $u, v \in L^2(M)$, too:

$$\begin{aligned} |\langle u, v\rangle| &= \left|\int_M u(x)\,\overline{v(x)}\,\mathrm{d}x\right| \\ &\leq \int_M |u(x)|\,|v(x)|\,\mathrm{d}x \\ &\leq \int_M \frac{|u(x)|^2+|v(x)|^2}{2}\,\mathrm{d}x \overset{\|u\|=\|v\|}{=} \|u\|\,\|v\|. \end{aligned}$$

Informal example. For $K_A \in L^2(M\times N)$, define linear $A : L^2(N) \to L^2(M)$ by

$$Av(x) = \int_N K_A(x,y)\,v(y)\,\mathrm{d}y.$$

Reasoning as in the previous example (or using the Hölder inequality), we obtain

$$\begin{aligned} \|Av\|^2_{L^2(M)} &= \int_M |Av(x)|^2\,\mathrm{d}x \\ &= \int_M \left|\int_N K_A(x,y)\,v(y)\,\mathrm{d}y\right|^2\,\mathrm{d}x \\ &\leq \int_M \left(\int_N |K_A(x,y)|^2\,\mathrm{d}y\right)\left(\int_M |v(y)|^2\,\mathrm{d}y\right)\,\mathrm{d}x \\ &= \|K_A\|^2_{L^2(M\times N)}\,\|v\|^2_{L^2(M)}. \end{aligned}$$

Such *Hilbert–Schmidt operators* $A : L^2(N) \to L^2(M)$ form a Hilbert space, with inner product

$$\langle A, B\rangle_{HS} := \int_N \int_M K_A(x,y)\,K_B(x,y)^*\,\mathrm{d}x\,\mathrm{d}y,$$

with the canonical norm $A \mapsto \|A\|_{HS}$, where

$$\|A\|^2_{HS} = \int_N \int_M |K_A(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y = \|K_A\|^2_{L^2(M\times N)}.$$

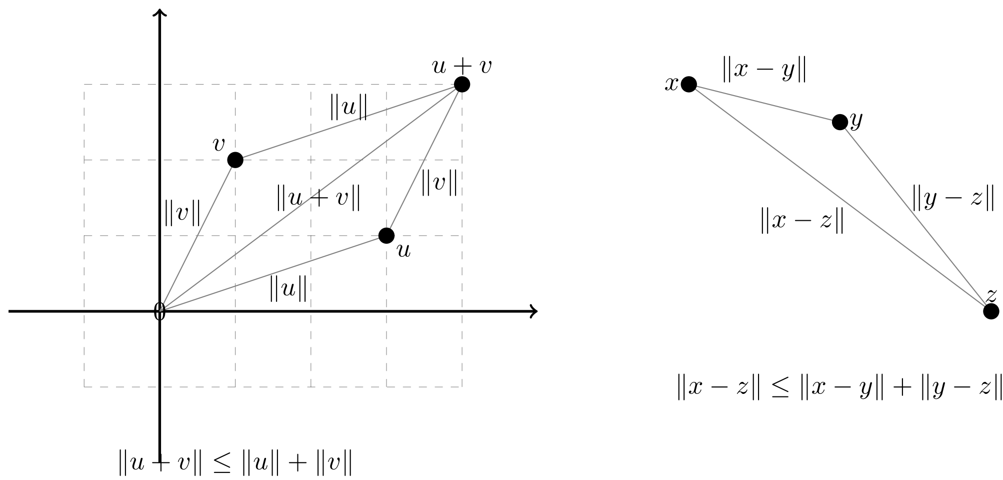


Figure 10: Triangle inequality of vectors.

Remark. As a corollary to the Cauchy–Schwarz inequality, the canonical norm turns out to be really a norm:

Corollary: Triangle inequality. *For all u, v in an inner product space,*

$$\|u + v\| \leq \|u\| + \|v\|. \tag{57}$$

Proof. The claim follows from taking the square roots of

$$\begin{aligned} \|u+v\|^2 &= \langle u+v, u+v\rangle \\ &= \langle u,u\rangle + \langle v,v\rangle + \langle u,v\rangle + \langle v,u\rangle \\ &\overset{(56)}{\leq} \|u\|^2 + \|v\|^2 + 2\,\|u\|\,\|v\| \\ &= (\|u\| + \|v\|)^2 . \end{aligned}$$ □

Remark. The triangle inequality is often in the equivalent form

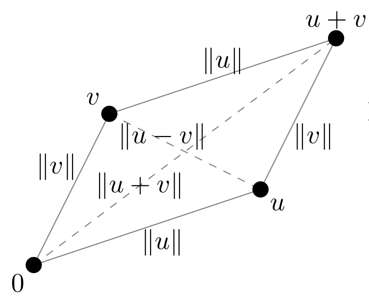
$$\|u - w\| \leq \|u - v\| + \|v - w\|, \tag{58}$$

stating that "Distance from u to w is at most distance from u via v to w." Think e.g. of cities $u = London$, $v = Paris$, $w = Rome$.

7.3 Norm topology

Definition. The *norm topology* τ_n of a normed space H is the topology that coming from the norm metric

$$(u, v) \mapsto d(u, v) = \|u - v\|.$$



For a parallelogram in an inner product space,
the sum of the squares of the diagonals
= the sum of the squares of the edges.

Figure 11: Parallelogram Identity: $\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$.

In other words, the norm topology is generated by the open balls

$$\mathbb{B}(u,r) := \{v \in H : \ \|v-u\| < r\} \tag{59}$$

of centers $u \in H$ and radii $r > 0$. That is, $U \subset H$ belongs to τ_n if and only if for all $u \in U$ there exists $\varepsilon > 0$ such that

$$\mathbb{B}(u,\varepsilon) \subset U.$$

7.4 Geometric identities

Soon we find that the inner product is actually encoded in the norm!

Polarization Identity. The inner product can be recovered from the canonical norm by the **Polarization Identity**

$$\mathrm{Re}\langle u,v\rangle = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2\right). \tag{60}$$

Exercise. Let H be a complex inner product space. Starting with the Polarization Identity (60) for the real part $\mathrm{Re}\langle u,v\rangle$, find the corresponding identity for the imaginary part $\mathrm{Im}\langle u,v\rangle$.

Parallelogram Identity. Also the **Parallelogram Identity** holds in inner product spaces:

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2 \tag{61}$$

(think of distances in a parallelogram with vertices at $0, u, v, u+v$). Actually, the Parallelogram Identity characterizes inner product spaces among normed spaces — this is the message of the following Jordan–von Neumann Theorem:

Jordan–von Neumann Theorem. *Let the Parallelogram Identity* (61) *hold in a normed space X. Then X is an inner product space.*

Proof. It is enough consider $\mathbb{R}$-normed space X (Why?). Define

$$\langle u, v\rangle := \|u+v\|^2/4 - \|u-v\|^2/4.$$

Clearly, $\langle u, u\rangle = \|u\|^2$, and $\langle v, u\rangle = \langle u, v\rangle$. Next,

$$\begin{aligned}
\langle u, w\rangle + \langle v, w\rangle &= \frac{1}{4}\left(\|u+w\|^2 - \|u-w\|^2 + \|v+w\|^2 - \|v-w\|^2\right) \\
&\overset{(61)}{=} \frac{1}{2}\left(\|\frac{u+v}{2} + w\|^2 + \|\frac{u-v}{2}\|^2 - \|\frac{u+v}{2} - w\|^2 - \|\frac{u-v}{2}\|^2\right) \\
&= 2\langle (u+v)/2, w\rangle.
\end{aligned}$$

Thus $\langle x, w\rangle = \langle x, w\rangle + \langle 0, w\rangle = 2\langle x/2, w\rangle$, so

$$\langle u, w\rangle + \langle v, w\rangle = \langle u+v, w\rangle.$$

From this inductively, we get

$$\langle ku, w\rangle = k\langle u, w\rangle$$

for all $k \in \mathbb{Z}^+$. Then also

$$k\langle u/k, w\rangle = \langle u, w\rangle.$$

Clearly

$$\langle -u, w\rangle = -\langle u, w\rangle.$$

Therefore

$$\langle \lambda u, w\rangle = \lambda\langle u, w\rangle$$

for every $\lambda \in \mathbb{Q}$. This extends to all $\lambda \in \mathbb{R}$, as the norm is continuous. □

Exercise. Let $p \in [1, \infty]$. Show that the Banach space

$$\ell^p(M) = \{u \in \mathbb{K}^M : \ \|u\|_{\ell^p} < \infty\}$$

is a Hilbert space if and only if $p = 2$. Recall that if $1 \le p < \infty$ then

$$\begin{aligned}
\|u\|_{\ell^p} &= \left(\sum_{x\in M} |u(x)|^p\right)^{1/p}, \\
\|u\|_{\ell^\infty} &= \sup_{x\in M} |u(x)|.
\end{aligned}$$

Informal example. Lebesgue space $L^p(M)$ is a Hilbert space if and only if $p = 2$. Here if $1 \le p < \infty$ then

$$\begin{aligned}
\|u\|_{L^p(M)} &= \left(\int_M |u(x)|^p \, \mathrm{d}x\right)^{1/p}, \\
\|u\|_{L^\infty(M)} &= \operatorname{ess\,sup}\{|u(x)| : \ x \in M\}.
\end{aligned}$$

Exercise. Let X be an inner product space. Show that the metric completion H of X can be given a natural Hilbert space structure.

7.5 Bounded operators

Even if the reader already would know bounded linear operators, we encourage thinking carefully about the following minimal treatise that prepares the study in Hilbert spaces.

Definition. A linear operator $A : H \to G$ between normed spaces H, G is *bounded* if there is a constant $C < \infty$ such that

$$\|Au\| \leq C\,\|u\|$$

for all $u \in H$ (with natural norm of G, H). Then we denote $A \in \mathscr{B}(H, G)$, and

$$\|A\| := \sup\{\|Au\| : \ u \in H, \ \|u\| \leq 1\}$$

is called the *norm* of A. We shall be especially interested in

$$\mathscr{B}(H) := \mathscr{B}(H, H).$$

Exercise. Let $A : H \to G$ be linear between normed spaces H, G. Show that A is bounded if and only if it is continuous.

Exercise. Let H be a normed space with a dense vector subspace S. Let $A : S \to H$ be linear such that

$$\|Au\| \leq C\|u\|$$

for a constant $C < \infty$, for all $u \in S$. Show that there is unique bounded linear extension $\widetilde{A} \in \mathscr{B}(H)$ such that

$$\widetilde{A}u = Au$$

for all $u \in S$, and that $\|\widetilde{A}\| \leq C$.
(Remark: here it is typical to simply write $\widetilde{A} = A$.)

Exercise. Prove the *Hilbert integral inequality*

$$\int_0^\infty \int_0^\infty \frac{|u(x)|\,|v(y)|}{x+y}\,\mathrm{d}x\,\mathrm{d}y \leq \pi \left(\int_0^\infty |u(x)|^2\,\mathrm{d}x \int_0^\infty |v(y)|^2\,\mathrm{d}y \right)^{1/2} \tag{62}$$

by using "exotic polar coordinates" $(\sqrt{x}, \sqrt{y}) = (r\cos(\varphi), r\sin(\varphi))$, and applying the Cauchy–Schwarz inequality. This approach is by David C. Ullrich [21].)

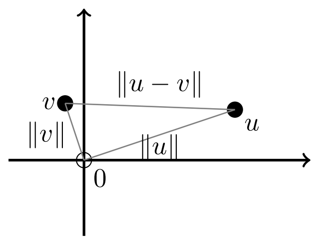


Figure 12: Orthogonality of vectors $u, v \in X$ means $\langle u, v\rangle = 0$. Then the Pythagorean equality $\|u-v\|^2 = \|u\|^2 + \|v\|^2$ holds.

8 Orthogonality

Definition. Let X be an inner product space. Vectors $u, v \in X$ are *orthogonal* if $\langle u, v\rangle = 0$. Subspaces $Z_1, Z_2 \subset X$ are *orthogonal*, denoted by $Z_1 \perp Z_2$, if $\langle u, v\rangle = 0$ for all $u \in Z_1$ and $v \in Z_2$. For $S \subset X$, the *orthogonal complement* is

$$S^{\perp} := \{u \in X \mid \forall v \in S : \langle u, v\rangle = 0\}.$$

Exercise. Let S be a subset of a Hilbert space H. Show that $S^{\perp} \subset H$ is a closed vector subspace, and that $S \subset (S^{\perp})^{\perp}$.

Remark. If $u, v \in X$ are orthogonal then

$$\|u-v\|^2 = \|u\|^2 + \|v\|^2 - 2\,\mathrm{Re}\,\langle u, v\rangle \overset{\langle u,v\rangle=0}{=} \|u\|^2 + \|v\|^2,$$

i.e. we obtained Pythagorean-type equality $\|u-v\|^2 = \|u\|^2 + \|v\|^2$ (think about a "right-angled triangle" with vertices at $0, u, v \in X$). When $\mathbb{K} = \mathbb{C}$, it is easy to check that $\langle u, v\rangle = 0$ if and only if

$$\|u-v\|^2 = \|u\|^2 + \|v\|^2 = \|u - \mathrm{i}v\|^2.$$

More generally, then we have

$$\|\mathrm{e}^{\mathrm{i}\alpha}u - \mathrm{e}^{\mathrm{i}\beta}v\|^2 = \|u\|^2 + \|v\|^2$$

for all $\alpha, \beta \in \mathbb{R}$. Actually, if u is a wavefunction in quantum mechanics (or a complex-valued signal in time-frequency analysis), then instead of the vector u, it is customary to deal with equivalence classes

$$[u] := \left\{\mathrm{e}^{\mathrm{i}\alpha}u : \ \alpha \in \mathbb{R}\right\} = \{\lambda u : \ \lambda \in \mathbb{C}, \ |\lambda| = 1\}.$$

8.1 Orthogonal projections

Orthogonal projections are bounded linear operators that naturally "cast shadows" within Hilbert spaces. Later they will appear useful in understanding more complicated operators. We show first that non-empty closed subsets in Hilbert space have a unique element nearest to the origin. Recall that a subset C of a vector space H is called *convex* if

$$tu + (1-t)v \in C$$

whenever $0 < t < 1$ and $u, v \in C$.

Lemma. *Let C be a non-empty closed convex subset of Hilbert space H. Then there exists a unique point $v \in C$ such that for all $w \in C$*

$$\|v\| \leq \|w\|.$$

Proof. Let

$$r := \inf\{\|w\| : \ w \in C\}.$$

Take $(v_k)_{k=1}^{\infty}$ in C such that $\|v_k\| \overset{k\to\infty}{\longrightarrow} r$. Now $\dfrac{v_j + v_k}{2} \in C$ by convexity, so

$$\begin{aligned}
(2r)^2 + \|v_j - v_k\|^2 &\leq \left(2\left\|\frac{v_j+v_k}{2}\right\|\right)^2 + \|v_j - v_k\|^2 \\
&= \|v_j + v_k\|^2 + \|v_j - v_k\|^2 \\
&\overset{\text{Parallelogram Id.}}{=} 2\|v_j\|^2 + 2\|v_k\|^2 \\
&\underset{j,k\to\infty}{\longrightarrow} (2r)^2,
\end{aligned}$$

i.e. $\|v_j - v_k\| \overset{j,k\to\infty}{\longrightarrow} 0$: So $(v_k)_{k=1}^{\infty}$ is Cauchy, converging to some $v \in C$ in closed set C, such that $\|v\| = r$. If $w \in C$ and $\|w\| = r$ then

$$\begin{aligned}
(2r)^2 + \|v - w\|^2 &\leq \left(2\left\|\frac{v+w}{2}\right\|\right)^2 + \|v - w\|^2 \\
&= \|v + w\|^2 + \|v - w\|^2 \\
&\overset{\text{Parallelogram Id.}}{=} 2\|v\|^2 + 2\|w\|^2 \\
&= (2r)^2,
\end{aligned}$$

so that $\|v - w\| = 0$, that is $v = w$. □

Definition. Let Z be a closed vector subspace of Hilbert space H, and $u \in H$ Then $C := u - Z = \{u - v : \ v \in Z\}$ is a closed non-empty set. By the previous Lemma, there exists a unique $Q(u) \in C$ of minimal norm: for all $v \in Z$ we have

$$\|Q(u)\| \leq \|u - v\|. \tag{63}$$

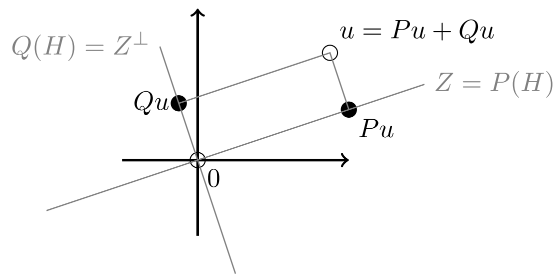


Figure 13: Orthogonal projection $P \in \mathscr{B}(H)$ onto a closed vector subspace $Z \subset H$. Then the linear mapping $Q = I - P$ is the orthogonal projection onto the orthogonal vector subspace $Z^\perp = \{u \in H : \langle u, z\rangle = 0 \text{ for all } z \in Z\}$.

Then $P(u) := u - Q(u)$ is the unique closest point in Z to $u \in H$: for all $v \in Z$ we have

$$\|u - P(u)\| \leq \|u - v\|. \tag{64}$$

This defines the *orthogonal projection* $P = P_Z : H \to H$ onto the closed vector subspace $Z \subset H$. It turns out that P is linear and bounded, and more:

Orthogonal Projection Theorem. *Let Z be a closed vector subspace of Hilbert space H. Let $P = P_Z : H \to H$, let $Q = I - P$ and $u \in H$. Then $P, Q \in \mathscr{B}(H)$, $\langle Pu, Qu\rangle = 0$, $\|u\|^2 = \|Pu\|^2 + \|Qu\|^2$, and $Q = P_{Z^\perp}$.*

Proof. In subspace Z, point $P(u)$ is closest to u, so that point $P(u) - P(u) = 0$ is closest to $u - P(u) = Q(u)$. Thus for all $t \in \mathbb{R}$ and $\lambda \in \mathbb{K}$,

$$\begin{aligned} 0 &\leq \|t\lambda P(u) - Q(u)\|^2 - \|Q(u)\|^2 \\ &= t^2|\lambda|^2\|P(u)\|^2 - 2t\mathrm{Re}(\lambda\langle P(u), Q(u)\rangle), \end{aligned}$$

Viewing at $t \approx 0$, we get $\langle P(u), Q(u)\rangle = 0$. Thus $Q(u) \in Z^\perp$, as $P(H) = Z$. Since

$$\begin{aligned} \lambda u &= \lambda\left(P(u) + Q(u)\right), \\ \mu v &= \mu\left(P(v) + Q(v)\right), \\ \lambda u + \mu v &= P(\lambda u + \mu v) + Q(\lambda u + \mu v), \end{aligned}$$

for every $u, v \in H$ and $\lambda, \mu \in \mathbb{K}$, we get

$$\begin{aligned} Z &\ni P(\lambda u + \mu v) - \lambda P(u) - \mu P(v) \\ &= \lambda Q(u) + \mu Q(v) - Q(\lambda u + \mu v) \\ &\in Z^\perp. \end{aligned}$$

Hence P and Q are linear, because $Z \cap Z^{\perp} = \{0\}$. Finally,

$$\begin{aligned} \|u\|^2 &= \|Pu + Qu\|^2 \\ &= \|Pu\|^2 + \|Qu\|^2 + 2\mathrm{Re}\langle Pu, Qu\rangle \\ &= \|Pu\|^2 + \|Qu\|^2; \end{aligned}$$

in particular $\|Pu\|, \|Qu\| \leq \|u\|$. Clearly, $Q = P_{Z^{\perp}}$. □

8.2 Direct sum. Orthonormality

In a $\mathbb{K}$-vector space H, recall that the *span* of a non-empty subset $S \subset H$ is

$$\mathrm{span}(S) := \left\{ \sum_{j=1}^{k} \lambda_j u_j \in H : \ k \in \mathbb{Z}^+, \ \{u_j\}_{j=1}^k \subset S, \ \{\lambda_j\}_{j=1}^k \subset \mathbb{K} \right\},$$

i.e. $\mathrm{span}(S) \subset H$ is the smallest vector subspace in H containing S.

Definition. Let H be Hilbert space. We write *direct sum*

$$H = \bigoplus_{\alpha \in J} H_\alpha$$

if $\mathscr{F} = \{H_\alpha : \alpha \in J\}$ is a family of pairwise orthogonal closed vector subspaces H_α with $\mathrm{span}(\cup \mathscr{F})$ dense in H. If $\mathscr{F} = \{Z, Z^{\perp}\}$, we write

$$H = Z \oplus Z^{\perp},$$

which is the case in the Orthogonal Projection Theorem.

Exercise. Let V be a closed vector subspace of a Hilbert space H. Show that $V = (V^{\perp})^{\perp}$.

Exercise. Let $u_1, \ldots, u_n \in H$ be mutually orthogonal, i.e. assume that $\langle u_j, u_k \rangle = 0$ for all $j \neq k$. Prove the **Pythagorean equality**

$$\left\| \sum_{k=1}^{n} u_k \right\|^2 = \sum_{k=1}^{n} \|u_k\|^2. \tag{65}$$

Definition. Collection $(e_\alpha)_{\alpha \in J}$ is *orthonormal* if $\|e_\alpha\| = 1$ and $\langle e_\alpha, e_\beta \rangle = 0$ for all $\alpha, \beta \in J$, $\alpha \neq \beta$.

Exercise. Let $(e_\alpha)_{\alpha \in J}$ be an orthonormal collection in H and let $u \in H$. Use Pythagoras' (65) to show **Bessel's inequality**

$$\sum_{\alpha \in J} |\langle u, e_\alpha \rangle|^2 \leq \|u\|^2. \tag{66}$$

Deduce that the set of α with $\langle u, e_\alpha \rangle \neq 0$ is at most countable (we shall use this fact in proving the Orthonormal Lemma soon.)

8.3 Orthonormal bases

An orthonormal basis is a maximal orthonormal set in a Hilbert space: often this allows decomposing vectors neatly into "simpler pieces". It turns out that every Hilbert space has an orthonormal basis. We start preparing for this result:

Orthonormal Lemma. *Let $(e_\alpha)_{\alpha\in J}$ orthonormal in Hilbert space H. Then conditions* (i), (ii), (iii) *are equivalent:*

(i) $\langle u, e_\alpha\rangle \neq 0$ *for at most countably many $\alpha \in J$, and series*

$$u = \sum_{\alpha\in J} \langle u, e_\alpha\rangle e_\alpha \tag{67}$$

converges in norm for any $u \in H$, regardless arranging terms.

(ii) *If $\langle u, e_\alpha\rangle = 0$ for all $\alpha \in J$, then $u = 0$.*

(iii) *Parseval–Plancherel formula holds:*

$$\|u\|^2 = \sum_{\alpha\in J} |\langle u, e_\alpha\rangle|^2 \tag{68}$$

for all $u \in H$.

Proof. Fix $u \in H$. Let $f_1, f_2, f_3, \cdots$ enumerate those e_α for which $\langle u, e_\alpha\rangle \neq 0$ (this family is countable by the previous Bessel exercise!).

(i) $\Rightarrow$ (iii): By (i), the left-hand-side of

$$\begin{aligned}\left\| u - \sum_{k=1}^{n} \langle u, f_k\rangle f_k \right\|^2 &= \|u\|^2 + \sum_{k=1}^{n} \|\langle u, f_k\rangle f_k\|^2 - 2\sum_{k=1}^{n} \mathrm{Re}\langle u, \langle u, f_k\rangle f_k\rangle \\ &= \|u\|^2 - \sum_{k=1}^{n} |\langle u, f_k\rangle|^2\end{aligned}$$

tends to 0 as $n \to \infty$. Hence this implies (iii).

(iii) $\Rightarrow$ (ii) is trivial.

(ii) $\Rightarrow$ (i): We can define $v := u - \sum_{k=1}^{\infty} \langle u, f_k\rangle f_k$, because

$$\left\| \sum_{k=k_1}^{k_2} \langle u, f_k\rangle f_k \right\|^2 \overset{(65)}{=} \sum_{k=k_1}^{k_2} |\langle u, f_k\rangle|^2 \overset{(66)}{\longrightarrow} 0 \quad \text{as} \quad k_1, k_2 \to \infty.$$

Here $\langle v, e_\alpha\rangle = 0$ for all $\alpha \in J$, because

$$\langle v, e_\alpha\rangle = \langle u - \sum_{k=1}^{\infty} \langle u, f_k\rangle f_k, e_\alpha\rangle = \langle u, e_\alpha\rangle - \sum_{k=1}^{\infty} \langle u_k, f_k\rangle\langle f_k, e_\alpha\rangle = \langle u, e_\alpha\rangle - \langle u, e_\alpha\rangle.$$

Thus $v = 0$ by (ii). □

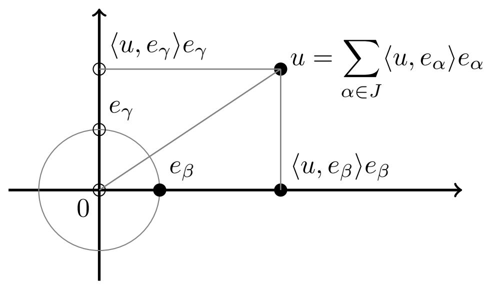


Figure 14: "Fourier" series $u = \sum_{\alpha \in J} \langle u, e_\alpha \rangle e_\alpha$ for orthonormal basis $(e_\alpha)_{\alpha \in J}$.

Definition. An orthonormal collection $(e_\alpha)_{\alpha \in J}$ satisfying conditions of the Orthonormal Lemma is an *orthonormal basis* of H. Alternatively, an orthonormal basis of H is a maximal orthonormal collection in H.

Example. For $H = \ell^2(J)$, orthonormal basis $(\delta_\alpha)_{\alpha \in J}$ contains "Kronecker delta vectors" $\delta_\alpha \in \ell^2(J)$, where

$$\delta_\alpha(\beta) = \begin{cases} 1 & \text{if } \beta = \alpha, \\ 0 & \text{if } \beta \neq \alpha. \end{cases}$$

Example. For $H = L^2(\mathbb{R}/\mathbb{Z})$, orthonormal basis $(e_k)_{k \in \mathbb{Z}}$ contains vectors $e_k \in H$, where

$$e_k(x) := \mathrm{e}^{\mathrm{i}2\pi x \cdot k}.$$

For $u \in H$, number

$$\widehat{u}(k) := \langle u, e_k \rangle = \int_0^1 u(x)\, \mathrm{e}^{-\mathrm{i}2\pi x \cdot k}\, \mathrm{d}x$$

is the kth Fourier coefficient of u, which has the Fourier series

$$u = \sum_{k \in \mathbb{Z}} \widehat{u}(k)\, e_k.$$

Example. *Haar wavelet* $\psi : \mathbb{R} \to \mathbb{R}$ is defined by

$$\psi(x) := \begin{cases} +1 & \text{if } 0 < x < 1/2, \\ -1 & \text{if } 1/2 < x < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{69}$$

The *Haar orthonormal basis* $(\psi_{n,m})_{n,m \in \mathbb{Z}}$ of $H = L^2(\mathbb{R})$ consists of functions $\psi_{m,n} : \mathbb{R} \to \mathbb{R}$, where

$$\psi_{n,m}(x) := 2^{n/2}\, \psi(2^n x - m). \tag{70}$$

Example. *Hermite functions* $\psi_n : \mathbb{R} \to \mathbb{R}$ provide another orthonormal basis $(\psi_n)_{n=0}^{\infty}$ for $H = L^2(\mathbb{R})$. Here ψ_n is defined by

$$\psi_n(x) := (-1)^n \, 2^{n/2} \, (n!)^{1/2} \, \pi^{-1/4} \, \mathrm{e}^{x^2/2} \, \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2}. \tag{71}$$

They are eigenfunctions to the Schrödinger equation of a harmonic oscillator:

$$\psi_n''(x) + (2n + 1 - x^2)\psi_n(x) = 0. \tag{72}$$

Definition. Metric space is called *separable* if it has a countable dense subset.

Remark. Many of the Hilbert spaces in applications are separable, so that numerical computations are feasible. For example, $\ell^2(\mathbb{Z})$, $L^2(\mathbb{R}/\mathbb{Z})$ and $L^2(\mathbb{R})$ are separable. If J is uncountable then $\ell^2(J)$ is not separable.

Orthonormal Basis Theorem. *Every Hilbert space H has an orthonormal basis. An orthonormal basis is countable if and only if H is separable (and then any other orthonormal basis is countable).*

Exercise. Prove the Orthonormal Basis Theorem. (Hint: In the first part, use property (ii) of the Orthonormal Lemma, order orthonormal collections by inclusion, applying Zorn's Lemma. In the second part, use the following Gram–Schmidt process.)

Remark. The *Gram–Schmidt process* gives an orthonormal sequence $(e_k)_{k=1}^{\infty}$ from linearly independent $(u_k)_{k=1}^{\infty}$ as follows: Let $e_1 := u_1/\|u_1\|$. Then inductively let $e_k := v_k/\|v_k\|$, where

$$v_k = u_k - \sum_{j=1}^{k-1} \langle u_k, e_j \rangle \, e_j.$$

Moreover, $\mathrm{span}\{u_j\}_{j=1}^{k} = \mathrm{span}\{e_j\}_{j=1}^{k}$ for all $k \in \mathbb{Z}^+$. But beware: this process is numerically unstable.

Example. We start with vectors $u_1 = (3, 4)$ and $u_2 = (10, -5)$ in Euclidean plane $\mathbb{R}^2$. Then $e_1 = u_1/\|u_1\| = (3/5, 4/5)$, next

$$v_2 = u_2 - \langle u_2, e_1 \rangle e_1 = (44/5, -33/5),$$

and $e_2 = v_2/\|v_2\| = (4/5, -3/5)$.

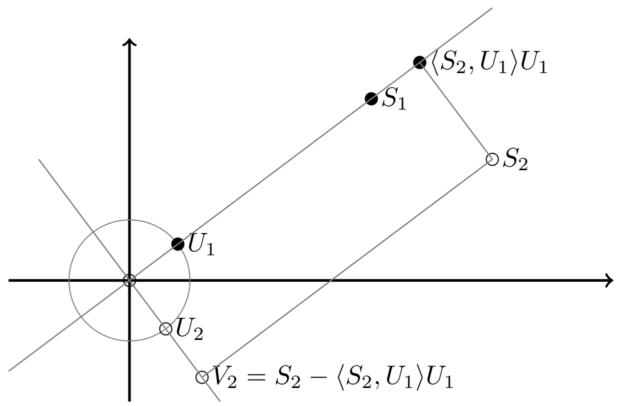


Figure 15: Gram–Schmidt: from invertible $S = [S_1 \cdots S_n] \in \mathbb{C}^{n\times n}$ to unitary $U = [U_1 \cdots U_n] \in \mathbb{C}^{n\times n}$ such that $\mathrm{span}\{S_1, \cdots, S_k\} = \mathrm{span}\{U_1, \cdots, U_k\}$.

Example. Next, let us find a presentation of an orthogonal projection $P = [P_1 \cdots P_n] \in \mathbb{K}^{n\times n}$. Here P projects orthogonally onto $Z = P(\mathbb{K}^n) = \mathrm{span}\{P_j\}_{j=1}^n$. Here $Pz = z$ for all $z \in Z$, as $P^2 = P$. Let $I \in \mathbb{K}^{n\times n}$ be the identity matrix, and let $R = [R_1 \cdots R_{2n}] := [P\ I] \in \mathbb{K}^{n\times 2n}$. Let $Z_j := \mathrm{span}\{R_1, \cdots, R_j\}$. There are smallest $j_\ell \in \{1, \cdots, 2n\}$ with

$$\dim(Z_{j_\ell}) = \ell \in \{1, \cdots, n\}.$$

For $S_\ell := R_{j_\ell}$, vectors $S_1, \cdots, S_n \in \mathbb{K}^n$ are linearly independent, from which the Gram-Schmidt process finds orthonormal $U_1, \cdots, U_n \in \mathbb{K}^n$, with $Z = \mathrm{span}\{U_j\}_{j=1}^k$ for $k = \dim(Z)$, so $P(U_\ell) = U_\ell$ for $\ell \le k$. And $P(U_\ell) = 0$ for $\ell > k$, because

$$\langle P(U_\ell), y\rangle = \langle U_\ell, P^* y\rangle \overset{P^*=P}{=} \langle U_\ell, Py\rangle \overset{Py\in Z}{=} \langle U_\ell, \sum_{j=1}^k \lambda_j U_j\rangle \overset{\ell\neq j}{=} 0$$

for all $y \in \mathbb{K}^n$. Thus, the orthogonal projection $P : \mathbb{K}^n \to \mathbb{K}^n$ has presentation

$$Px = \sum_{j=1}^k \langle x, U_j\rangle U_j \tag{73}$$

for all $x \in \mathbb{K}^n$.

Remark. Formally we can identify any two Hilbert spaces which have orthonormal bases of the same cardinality. However, e.g. separable spaces $L^2(M)$ and $L^2(N)$ for non-diffeomorphic manifolds M, N should not be just trivially identified: extra structures might be otherwise interesting, like when $\mathbb{R}^m = M \not\cong N = \mathbb{R}^n/\mathbb{Z}^n$.

8.4 Matrix

Let $(e_\alpha)_{\alpha\in J}$ be an orthonormal basis for Hilbert space H. Bounded operator $A \in \mathscr{B}(H)$ can be recovered also from its *matrix elements*

$$A_{\alpha\beta} := \langle Ae_\beta, e_\alpha\rangle \in \mathbb{K}$$

Why? Let $u_\alpha := \langle u, e_\alpha\rangle$. Thus $e_{\alpha\alpha} = \langle e_\alpha, e_\alpha\rangle = 1$ and $e_{\beta\alpha} = \langle e_\beta, e_\alpha\rangle = 0$ whenever $\beta \neq \alpha$. Just as in the finite-dimensional matrix case,

$$(Au)_\alpha = \sum_{\beta\in J} A_{\alpha\beta}\, u_\beta, \tag{74}$$

So, any bounded operator can be represented by a matrix. However, when defining operators on H by matrix formula (74), we have to be careful: then A is bounded if and only if

$$\sum_{\alpha\in J}\left|\sum_{\beta\in J} A_{\alpha\beta}\, u_\beta\right|^2 \leq C\,\|u\|^2 \tag{75}$$

for every $u \in H$, for a constant $C < \infty$. And beware: this matrix representation is often a bit too non-redundant for real-life applications.

Exercise. Prove this boundedness assertion (75).

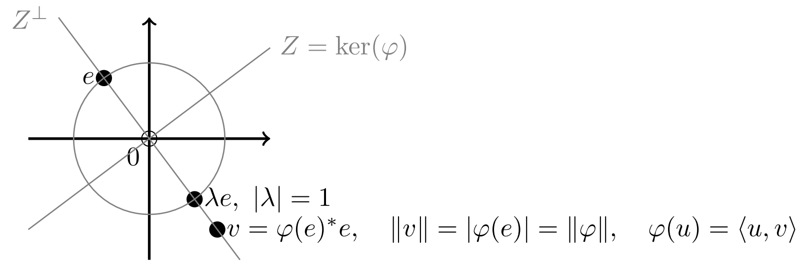


Figure 16: Depicting the proof of the Fréchet–Riesz Representation Theorem.

9 Hilbert duality

Linear mappings $\varphi : H \to \mathbb{K}$ on a normed $\mathbb{K}$-vector space H are called *linear functionals*. The dual H' of H consists of those linear functionals $\varphi : H \to \mathbb{K}$ which are *bounded* in the sense that

$$|\varphi(u)| \leq C\,\|u\|$$

for a constant $C < \infty$, for all $u \in H$. Then the norm of φ is

$$\|\varphi\| := \sup\{|\varphi(u)| : \ u \in H,\ \|u\| \leq 1\}.$$

Exercise. Assume the Basis Lemma, i.e. existence of the algebraic vector space basis, equivalent to the Axiom of Choice. Show that there exist unbounded linear functionals $\varphi : H \to \mathbb{K}$ when Hilbert space H is infinite-dimensional.

9.1 Bounded linear functionals in Hilbert spaces

In a Hilbert space H, the mapping $\varphi_v = (u \mapsto \langle u, v\rangle)$ belongs to H'. Clearly,

$$|\varphi_v(u)| \leq \|v\|\,\|u\|$$

by the Cauchy-Schwarz inequality. Hence $\|\varphi_v\| = \|v\|$, as $\varphi_v(v) = \langle v, v\rangle = \|v\|^2$. Actually, there are no other bounded linear functionals on Hilbert spaces:

Fréchet–Riesz Representation Theorem. *Let $\varphi \in H'$, where H is a Hilbert space. Then $\varphi = (u \mapsto \langle u, v\rangle)$ for unique $v \in H$.*

Proof. Here $Z := \ker(\varphi) \subset H$ is a closed vector subspace, as $\varphi \in H'$. By the Orthogonal Projection Theorem, $H = Z^\perp \oplus Z$. Case $\varphi = 0$ would be trivial; so let $e \in Z^\perp$ such that $\|e\| = 1$. Since $\varphi(e) = \langle e, \varphi(e)^* e\rangle$, could it be that $\varphi(u) = \langle u, v\rangle$ for $v := \varphi(e)^* e \in H$? This is indeed the case, as

$$\langle u, \varphi(e)^* e\rangle - \varphi(u) = \langle \varphi(e) u, e\rangle - \langle \varphi(u) e, e\rangle \overset{z:=\phi(e)u-\phi(u)e}{=} \langle z, e\rangle \overset{z\perp e}{=} 0,$$

where $z \in Z$, because $\varphi(z) = \varphi(e)\varphi(u) - \varphi(u)\varphi(e) = 0$. Finally, if $\langle u, v\rangle = \langle u, w\rangle$ for all $u \in H$ then $v = w$, because

$$0 = \langle u, v\rangle - \langle u, w\rangle = \langle u, v - w\rangle \stackrel{u=v-w}{=} \|v - w\|^2.$$

This completes the proof. □

Remark. When $\varphi \not\equiv 0$ in the proof of the Fréchet–Riesz Representation Theorem, then $Z^\perp = \ker(\varphi)^\perp$ is a 1-dimensional subspace of H. Think for instance of $\varphi \in H'$ for $H = \ell^2(\mathbb{Z}^+)$ such that $\varphi(u) := u(1)$. Then $\varphi(u) = \langle u, v\rangle$ for the Kronecker delta vector $v = \delta_1$, and

$$\begin{aligned} Z &= \{u \in H : \ u(1) = 0\}, \\ Z^\perp &= \{\lambda v \in H : \ \lambda \in \mathbb{K}\} \quad \cong \quad \mathbb{K}. \end{aligned}$$

Remark. Let H be a Hilbert space. For $v \in H$, define $\varphi_v \in H'$ by $\varphi_v(u) := \langle u, v\rangle$. Then we can endow H' the structure of a Hilbert space from H via the bijective mapping $(v \mapsto \varphi_v) : H \to H'$. Thereby H and H' can be seen as isomorphic Hilbert spaces (however, notice that here $\varphi_{\lambda v}(u) = \lambda^* \varphi_v(u)$ for $u, v \in H$ and $\lambda \in \mathbb{K}$).

Exercise. Let $\Phi \in H'' = (H')'$. Show that there exists unique $u \in H$ such that $\Phi(\varphi_v) = \langle v, u\rangle$ for all $v \in H$, and that $\|\Phi\| = \|u\|$. In this sense, we have the Hilbert space isomorphism $H'' \cong H$: Hilbert spaces are examples of reflexive Banach spaces.

Exercise. Let $\varphi : H \to \mathbb{K}$ be a linear functional, and let $(e_\alpha)_{\alpha \in J}$ be an orthonormal basis of Hilbert space H. Show that if $\varphi \in H'$ then $(\varphi(e_\alpha))_{\alpha \in J} \in \ell^2(J)$. Show that if H is infinite-dimensional then there exists an unbounded linear functional $\psi : H \to \mathbb{K}$ such that $\psi(e_\alpha) = 0$ for all $\alpha \in J$. (Hint: Basis Lemma, i.e. the existence of an algebraic vector space basis.)

Natural topologies of H. Let τ_n be the *norm topology* of a Hilbert space H. That is, $U \subset H$ belongs to τ_n if for all $u \in U$ there exists $\varepsilon > 0$ such that

$$\mathbb{B}(u, \varepsilon) := \{v \in H : \ \|v - u\| < \varepsilon\} \subset U.$$

The *weak topology* τ_w of H is the smallest topology for which all bounded linear functionals $\varphi \in H'$ are continuous.

Remark. By definition, $\tau_w \subset \tau_n$. Actually, $\tau_w = \tau_n$ only if H is finite-dimensional. By the Fréchet–Riesz Representation Theorem, notice that the convergence of a sequence $(u_k)_{k=1}^\infty$ to $u \in H$ in the weak topology means

$$\langle u_k - u, v\rangle \to 0$$

as $k \to \infty$, for all $v \in H$. Naturally, the norm convergence $\|u_k - u\| \to 0$ implies the weak convergence, but often not the other way round. For instance, in $H = \ell^2(\mathbb{Z}^+)$, take vectors $u_k \in H$ such that $u_k(k) = 1$ and $u_k(j) = 0$ otherwise: then the sequence $(u_k)_{k=1}^\infty$ of "travelling bumps" does not converge in norm, but converges still weakly to $0 \in H$. Nevertheless:

Corollary. *Suppose $(u_k)_{k=1}^\infty$ converges to $u \in H$ in the weak topology. If $\lim_{k\to\infty} \|u_k\| = \|u\|$ then we have the norm convergence $\lim_{k\to\infty} \|u_k - u\| = 0$.*

Proof. Simply calculate

$$\begin{aligned} \|u_k - u\|^2 &= \|u_k\|^2 + \|u\|^2 - \langle u_k, u\rangle - \langle u_k, u\rangle^* \\ &\overset{k\to\infty}{\longrightarrow} \|u\|^2 + \|u\|^2 - \langle u, u\rangle - \langle u, u\rangle^* = 0. \end{aligned}$$ □

9.2 Weak formulation of linear operators

Let H be a Hilbert space. Linear operator $A : H \to H$ can be recovered from data $(\langle Au, v\rangle)_{u,v\in H}$. (Well, the matrix elements $A_{\alpha\beta} := \langle Ae_\beta, e_\alpha\rangle \in \mathbb{C}$ would be enough, where $(e_\alpha)_{\alpha\in J}$ is an orthonormal basis.) Actually, the situation is even better for the scalar field $\mathbb{K} = \mathbb{C}$:

Weak Formulation Theorem. *Let $A, B : H \to H$ be linear in a $\mathbb{C}$-Hilbert space H. Then $A = B$ if for all $u \in H$*

$$\langle Au, u\rangle = \langle Bu, u\rangle.$$

Proof. Let $C = A - B$, $u, v \in H$, $\lambda \in \mathbb{C}$. Now

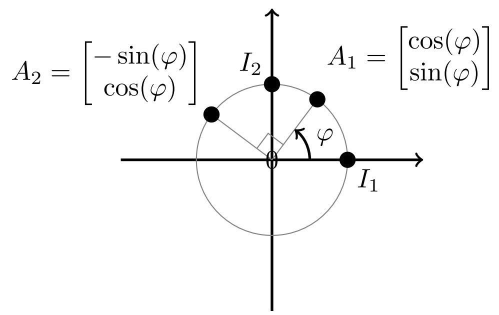
$$\begin{aligned} 0 &\overset{0=\lambda\langle Cw,w\rangle}{=} \lambda\langle C(u + \lambda v), u + \lambda v\rangle \\ &= |\lambda|^2\langle Cu, v\rangle + \lambda^2\langle Cv, u\rangle. \end{aligned}$$

Plug in $\lambda \in \{1, \mathrm{i}\}$ to get

$$\begin{cases} 0 = \langle Cu, v\rangle + \langle Cv, u\rangle, \\ 0 = \langle Cu, v\rangle - \langle Cv, u\rangle. \end{cases}$$

Clearly, $\langle Cu, v\rangle = 0$ for all $u, v \in H$. Thus $A - B = C = 0$. □

Remark. This weak formulation statement does not hold if $\mathbb{K} = \mathbb{R}$: If we only had $\langle Au, u\rangle = \langle Bu, u\rangle$ for all $u \in H$ in a $\mathbb{R}$-Hilbert space H, we still could have $A \neq B$. Rotations of the real plane $H = \mathbb{R}^2$ give an easy counter-example. For instance, $A = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix}$ rotates $\mathbb{R}^2$ by angle $\varphi \in \mathbb{R}$ around the origin. Then $\langle Au, u\rangle = \cos(\varphi)\,\|u\|^2$, which does not identify A when $|\cos(\varphi)| < 1$.



$$A = [A_1 \ A_2] = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix}.$$

$$\text{Identity matrix } I = [I_1 \ I_2] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Figure 17: Rotation $A \in \mathbb{R}^{2\times 2}$ by angle φ around the origin.

Exercise. Let H be a Hilbert space and $A \in \mathscr{B}(H)$. Show that

$$\|A\| = \sup_{u,v:\|u\|,\|v\|\leq 1} |\langle Au, v\rangle| \,.$$

Example. For nice-enough functions $v : \mathbb{R} \to \mathbb{C}$, let

$$Av(x) := \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}2\pi(x-y)\cdot\eta}\, b(x,y,\eta)\, v(y)\, \mathrm{d}y\, \mathrm{d}\eta,$$

where the *amplitude* $b : \mathbb{R}\times\mathbb{R}\times\mathbb{R} \to \mathbb{C}$ is obtained from a nice-enough *symbol* $a : \mathbb{R}\times\mathbb{R} \to \mathbb{C}$ by

$$b(x,y,\eta) := \frac{1}{y-x} \int_x^y a(t,\eta)\, \mathrm{d}t.$$

This is equivalent to the following weak formulation:

$$\langle u, Av\rangle_{L^2(\mathbb{R})} = \langle Q(u,v), a\rangle_{L^2(\mathbb{R}\times\mathbb{R})},$$

where *Born–Jordan transform* $Q(u,v) : \mathbb{R}\times\mathbb{R} \to \mathbb{C}$ is defined by

$$\begin{aligned} Q(u,v)(x,\eta) &= \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}2\pi y\cdot\eta}\, R(u,v)(x,y)\, \mathrm{d}y, \\ R(u,v)(x,y) &= \frac{1}{y} \int_{x-y/2}^{x+y/2} u(t+y/2)\, v(t-y/2)^*\, \mathrm{d}t. \end{aligned}$$

9.3 Adjoint operator

The Hilbert space adjoint $A^* \in \mathscr{B}(H)$ is a sort of a natural mirror image of operator $A \in \mathscr{B}(H)$. (Notice, though, that A^* may differ from the Banach space adjoint $A' \in \mathscr{B}(H')$.)

Definition. Let H, G be Hilbert spaces. Notice that $(u \mapsto \langle Au, w\rangle) \in H'$ if $A \in \mathscr{B}(H,G)$ and $w \in G$. By Fréchet–Riesz Representation Theorem there is unique $A^*w \in H$ such that for all $u \in H$

$$\langle Au, w\rangle = \langle u, A^*w\rangle. \tag{76}$$

This defines *adjoint* (or more specifically *Hilbert adjoint*) mapping $A^* : G \to H$. Operator $A \in \mathscr{B}(H)$ is called *self-adjoint* if $A^* = A$.

Exercise. Let $\lambda \in \mathbb{K}$, $A, B \in \mathscr{B}(H,G)$. Show that $A^* \in \mathscr{B}(G,H)$ with

$$\|A^*\| = \|A\|, \quad (A^*)^* = A, \quad (\lambda A)^* = \lambda^* A^*, \quad (A+B)^* = A^* + B^*.$$

Exercise. Let $A \in \mathscr{B}(H,G)$ and $B \in \mathscr{B}(G,F)$. Show that $(BA)^* = A^*B^*$.

Exercise. Let $A \in \mathscr{B}(H,G)$. Prove that

$$\ker(A)^\perp \text{ is the closure of } \mathrm{ran}(A^*). \tag{77}$$

Example. If $B \in \mathscr{B}(H,G)$ then $B^*B \in \mathscr{B}(H)$ is self-adjoint, as

$$(B^*B)^* = B^*(B^*)^* = B^*B.$$

Informal example. If $A : L^2(N) \to L^2(M)$ with $K_A \in L^2(M \times N)$,

$$\begin{aligned} Av(x) &= \int_N K_A(x,y)\, v(y)\, \mathrm{d}y, \text{ then} \\ A^*u(y) &= \int_M K_A(x,y)^* u(x)\, \mathrm{d}x. \end{aligned}$$

Self-adjointness would mean $K_A(y,x)^* = K_A(x,y)$ for almost all $x, y \in M = N$.

Exercise. Suppose $A : H \to H$ is linear and $\langle Au, v\rangle = \langle u, Av\rangle$ for all $u, v \in H$. Show that $A = A^* \in \mathscr{B}(H)$. (Hint: apply the Closed Graph Theorem.)

Proposition. *Operator $P \in \mathscr{B}(H)$ is an orthogonal projection if and only if $P = P^* = P^2$.*

Proof. Let $P = P^* = P^2 \in \mathscr{B}(H)$ and $Q = I - P$. Then $\ker(P)$ and $\ker(Q)$ are closed vector subspaces of H, because $P, Q \in \mathscr{B}(H)$. For any $u \in H$,

$$\langle Pu, Qu\rangle \overset{Q=I-P}{=} \langle Pu, u - Pu\rangle \overset{P^*=P}{=} \langle u, Pu - P^2u\rangle \overset{P=P^2}{=} \langle u, 0\rangle = 0.$$

Thus $\|u\|^2 = \|Pu + Qu\|^2 = \|Pu\|^2 + \|Qu\|^2$, so that P is the orthogonal projection onto the closed vector subspace $P(H) = \ker(Q)$.

Now let $P_Z = P_Z^2 \in \mathscr{B}(H)$ be the orthogonal projection onto a closed vector subspace $Z \subset H$. Denoting $Q_Z = I - P_Z$, we see that $P_Z^* = P_Z$, because

$$\langle P_Z u, v\rangle = \langle P_Z u, P_Z v + Q_Z v\rangle = \langle P_Z u, P_Z v\rangle = \langle P_Z u + Q_Z u, P_Z v\rangle = \langle u, P_Z v\rangle.$$

This completes the proof. □

Exercise. Let H be a Hilbert space and $P = P^2 \in \mathscr{B}(H)$, where $\|P\| = 1$. Show that P is an orthogonal projection.

Remark. In general for $A \in \mathscr{B}(H)$, we have

$$\|A\| = \sup_{u,v \in H:\ \|u\|,\|v\| \le 1} |\langle Au, v\rangle|. \tag{78}$$

For self-adjoint operators, the norm has an interesting property that will be useful later when studying spectral properties:

Norm Symmetry Lemma. *Let $A^* = A \in \mathscr{B}(H)$. Then*

$$\|A\| = \sup_{u \in H:\ \|u\| \le 1} |\langle Au, u\rangle| . \tag{79}$$

Proof. First, $|\langle Au, u\rangle| \le \|Au\|\,\|u\| \le \|A\|\,\|u\|^2$ by Cauchy–Schwarz. Let

$$r := \sup \{|\langle Au, u\rangle| :\ u \in H,\ \|u\| \le 1\} .$$

Assume $Au \ne 0$ for $\|u\| = 1$ (case $A = 0$ would be trivial anyway), and let $v := Au/\|Au\|$. Then

$$\begin{aligned}
\|Au\| &= \langle Au, v\rangle/2 + \langle v, Au\rangle/2 \\
&\overset{A^*=A}{=} \langle Au, v\rangle/2 + \langle Av, u\rangle/2 \\
&= \langle A(u+v), u+v\rangle/4 - \langle A(u-v), u-v\rangle/4 \\
&\le |\langle A(u+v), u+v\rangle|/4 + |\langle A(u-v), u-v\rangle|/4 \\
&\le r\left(\|u+v\|^2 + \|u-v\|^2\right)/4 \\
&\overset{\text{Parallelogram Id.}}{=} r\left(\|u\|^2 + \|v\|^2\right)/2 \quad = \quad r.
\end{aligned}$$

All in all, we have now $r \le \|A\| \le r$. □

Now it is easy to prove the following consequence:

Corollary. *If $B \in \mathscr{B}(H)$ then $\|B^*B\| = \|B\|^2$.* □

Exercise. Use the Spectral Radius Formula (53) to show that $\rho(B^*B) = \|B\|^2$ for all $B \in \mathscr{B}(H)$.

9.4 Almost orthogonality (Cotlar–Stein Lemma)

Next we present a useful tool for checking that some operators on Hilbert spaces are bounded:

Cotlar–Stein Lemma (Almost orthogonality). *Let H, G be Hilbert spaces. Let bounded linear operators $A_\alpha : H \to G$ and constants $C_\alpha < \infty$ satisfy*

$$C = \sum_{\alpha \in \mathbb{Z}^n} C_\alpha < \infty, \quad \|A_\alpha^* A_\beta\|_{H\to H} \le C_{\alpha-\beta}^2, \quad \|A_\alpha A_\beta^*\|_{G\to G} \le C_{\alpha-\beta}^2.$$

Then $A = \sum_{\alpha\in\mathbb{Z}^n} A_\alpha$ converges in the strong operator topology, $\|A\|_{H\to G} \le C$.

Proof. First,

$$\|A\|_{H\to G}^2 = \sup_{\|u\|_H \le 1} \langle Au, Au\rangle_G = \sup_{\|u\|_H\le 1} \langle A^*Au, u\rangle_H \le \|A^*A\|_{H\to H}.$$

So if $m = 2^k$ then

$$\|A\|_{H\to G}^{2m} \le \|A^*A\|_{H\to H}^m = \|(A^*A)^m\|_{H\to H} = \left\| \sum_{\alpha_1,\ldots,\alpha_{2m}} A_{\alpha_1}^* A_{\alpha_2} \cdots A_{\alpha_{2m-1}}^* A_{\alpha_{2m}} \right\|_{H\to H}. \tag{80}$$

Grouping the terms as $(A_{\alpha_1}^* A_{\alpha_2})(A_{\alpha_3}^* A_{\alpha_4}) \cdots (A_{\alpha_{2m-1}}^* A_{\alpha_{2m}})$, we have

$$\|A_{\alpha_1}^* A_{\alpha_2} \cdots A_{\alpha_{2m-1}}^* A_{\alpha_{2m}}\|_{H\to H} \le C_{\alpha_1-\alpha_2}^2\, C_{\alpha_3-\alpha_4}^2 \cdots C_{\alpha_{2m-1}-\alpha_{2m}}^2. \tag{81}$$

Grouping the terms as $A_{\alpha_1}^* (A_{\alpha_2} A_{\alpha_3}^*) \cdots (A_{\alpha_{2m-2}} A_{\alpha_{2m-1}}^*) A_{\alpha_{2m}}$, we have

$$\|A_{\alpha_1}^* A_{\alpha_2} \cdots A_{\alpha_{2m-1}}^* A_{\alpha_{2m}}\|_{H\to H} \le C^2\, C_{\alpha_2-\alpha_3}^2\, C_{\alpha_4-\alpha_5}^2 \cdots C_{\alpha_{2m-2}-\alpha_{2m-1}}^2. \tag{82}$$

From (80) taking the geometric mean of (81) and (82), we obtain

$$\|A\|_{H\to G}^{2m} \le \sum_{\alpha_1,\ldots,\alpha_{2m}} C\, C_{\alpha_1-\alpha_2}\, C_{\alpha_2-\alpha_3} \cdots C_{\alpha_{2m-1}-\alpha_{2m}}.$$

This leads to

$$\|A\|_{H\to G}^{2m} \le C^{2m} \sum_{\alpha_{2m}} 1.$$

If there are only $N < \infty$ non-zero operators A_α, we have

$$\|A\|_{H\to G} \le C\, N^{1/(2m)} \overset{m\to\infty}{\longrightarrow} C.$$

This estimate is uniform in $N \in \mathbb{Z}^+$. □

9.5 Hermitian forms

Above we studied linear functionals $\varphi : H \to \mathbb{K}$. These are needed when treating Hermitian forms $B : H \times H \to \mathbb{K}$.

Definition. *Hermitian form* on $\mathbb{K}$-Hilbert space H is mapping

$$B : H \times H \to \mathbb{K}$$

where $(u \mapsto B(u,v)) : H \to \mathbb{K}$ is linear and $B(v,u) = B(u,v)^*$. Hermitian form B is *bounded* if there is a constant $C < \infty$ such that for all $u, v \in H$

$$|B(u,v)| \leq C\,\|u\|\,\|v\|. \tag{83}$$

Hermitian form B is *coercive* if there is a constant $c > 0$ such that for all $u \in H$

$$B(u,u) \geq c\,\|u\|^2. \tag{84}$$

Example. If $A^* = A \in \mathscr{B}(H)$ then $B(u,v) := \langle Au, v\rangle$ defines a bounded Hermitian form, where $C = \|A\|$. Such B is coercive when A is invertible and positive (defined later). Especially, $B(u,v) := \langle u, v\rangle$ defines a bounded coercive Hermitian form, where $C = 1 = c$.

Exercise. Let $B : H \times H \to \mathbb{K}$ be a Hermitian form such that

$$\sup_{v \in H : \|v\| \leq 1} |B(u,v)| < \infty$$

for each $u \in H$. Show that B is bounded. (Hint: Banach–Steinhaus Theorem.)

Remark. The following Lax–Milgram Theorem can be used to find weak solutions v to partial differential equations, encoded by a Hermitian form B:

Lax–Milgram Theorem. *Let $B : H \times H \to \mathbb{K}$ be a bounded coercive Hermitian form. For each $f \in H'$ there is unique $u \in H$ such that for all $v \in H$*

$$B(u,v) = f(v)^*.$$

Moreover, $\|u\| \leq \|f\|/c$, where $B(u,u) \geq c\|u\|^2$ for all $u \in H$.

Proof: Exercise! (Hint: Fréchet–Riesz Representation Theorem, Hermitian form by $B(u,v) = \langle u, Av\rangle$...)

Informal example. Let $\Omega \subset \mathbb{R}^d$ be a bounded open domain with nice-enough boundary $\partial\Omega$. The Poisson problem (in mechanics and electrostatics) is to find $u : \Omega \to \mathbb{R}$ with

$$\begin{cases} -\Delta u(x) = f(x) & \text{when } x \in \Omega, \\ u(x) = 0 & \text{when } x \in \partial\Omega, \end{cases} \tag{85}$$

for given $f : \Omega \to \mathbb{R}$, where $\Delta = \nabla\cdot\nabla$ is the Laplacian. Relating this to the Lax–Milgram Theorem, here $H = H_0^1(\Omega) \subset L^2(\Omega)$ is the Sobolev space obtained by completing test function space $C_c^\infty(\Omega) \subset L^2(\Omega)$ of compactly supported smooth functions with respect to the inner product given by

$$\langle u, v\rangle := \langle \nabla u, \nabla v\rangle_{L^2(\Omega)} = \int_\Omega \nabla u(x) \cdot \overline{\nabla v(x)}\, \mathrm{d}x =: B(u, v).$$

Integrating by parts, we obtain

$$\int_\Omega \nabla u(x) \cdot \overline{\nabla v(x)}\, \mathrm{d}x \overset{(85)}{=} \int_\Omega f(x)\, \overline{v(x)}\, \mathrm{d}x$$

which is the weak formulation $B(u, v) = f(v)^*$ of the Poisson problem. In this example, $B : H \times H \to \mathbb{C}$ is bounded and coercive, with $C = 1 = c$. According to the Lax–Milgram Theorem, for every $f \in H' \supset L^2(\Omega)$, the Poisson problem has a unique solution $u \in H$ such that $\|u\|_H \leq \|f\|_{H'}$. For a nice-enough domain Ω there is the *Poincaré inequality*

$$\|u\|_{L^2(\Omega)} \leq C_\Omega \|\nabla u\|_{L^2(\Omega)} = C_\Omega \|u\|.$$

10 Operator families in Hilbert spaces

We have seen already orthogonal projections, which are building blocks of more complicated linear operators, and more generally self-adjoint operators. Next we shall deal with other important subfamilies of bounded linear operators on Hilbert spaces, explaining their inherent importance.

10.1 Compact operators in Hilbert space

Here we do not require the reader to know about compact operators in Banach spaces. Let H, G be $\mathbb{K}$-Hilbert spaces. Operator $A \in \mathscr{B}(H,G)$ is *compact* if $(Au_k)_{k=1}^{\infty}$ has a converging subsequence $(Au_{k_j})_{j=1}^{\infty}$ (where $k_j \in \mathbb{Z}^+$, $k_j < k_{j+1}$) whenever $(u_k)_{k=1}^{\infty}$ is a bounded sequence in H, i.e. when there is a constant $C < \infty$ such that $\|u_k\| \leq C$ for all $k \in \mathbb{Z}^+$. The family of compact linear operators $A : H \to G$ is denoted by $\mathscr{K}(H,G) \subset \mathscr{B}(H,G)$. We mostly study $\mathscr{K}(H) := \mathscr{K}(H,H)$.

Example. If H is finite-dimensional then $\mathscr{K}(H) = \mathscr{B}(H)$.

Example. If $A \in \mathscr{K}(H)$ and $B \in \mathscr{B}(H)$ then $AB, BA \in \mathscr{K}(H)$. Why? Take a bounded sequence $(u_k)_{k=1}^{\infty}$ in H. Then $\|Bu_k\| \leq \|B\|\,\|u_k\|$, so $(Bu_k)_{k=1}^{\infty}$ is another bounded sequence, and thus $(ABu_k)_{k=1}^{\infty}$ has a convergent subsequence by the compactness of A. Similarly, $(Au_k)_{k=1}^{\infty}$ has a subsequence $(Au_{k_j})_{j=1}^{\infty}$ converging to some $v \in H$, so that $(BAu_{k_j})_{j=1}^{\infty}$ converges to $Bv \in H$ by the continuity of B.

Example. Naturally, an orthonormal sequence $(u_k)_{k=1}^{\infty}$ in H cannot have a converging subsequence. Thus the identity operator $I \in \mathscr{B}(H)$ in an infinite-dimensional Hilbert space H cannot be compact. Consequently, if $A \in \mathscr{K}(H)$ has inverse $A^{-1} \in \mathscr{B}(H)$ then H must be finite-dimensional by the previous example!

Remark. Actually, due to the Open Mapping Theorem, if $A \in \mathscr{K}(H)$ is bijective then $A^{-1} : H \to H$ must be bounded, so that H must be finite-dimensional! In infinite-dimensional Hilbert (or Banach) spaces there are no compact linear bijections.

Exercise. Let $(e_k)_{k=1}^{\infty}$ be an orthonormal sequence in H, and let $(\lambda_k)_{k=1}^{\infty}$ be a bounded sequence in $\mathbb{K}$. Show that

$$Au := \sum_{k=1}^{\infty} \lambda_k \langle u, e_k \rangle e_k \tag{86}$$

defines a bounded operator $A \in \mathscr{B}(H)$. Moreover, show that $A \in \mathscr{K}(H)$ here if and only if $\lim_{k\to\infty} \lambda_k = 0$. Hint: Prove that $\|A - A_N\| \to 0$ as $N \to \infty$, where

$A_N \in \mathscr{K}(H)$,

$$A_N u := \sum_{k=1}^{N} \lambda_k \langle u, e_k \rangle e_k.$$

You can find help from Chapter 5 on compact operators in Banach spaces.

Remark. Bounded operator A in (86) in the previous exercise is an example of a *normal operator*, for which $A^*A = AA^*$. Here the self-adjointness $A^* = A$ would mean that $\lambda_k \in \mathbb{R}$ for all $k \in \mathbb{Z}^+$. When later diagonalizing compact self-adjoint operators, we learn that this phenomenon is actually quite general.

10.2 Normal operators

Loosely speaking, normal operators behave like scalars when finding polynomials of them and their adjoints:

Definition. Operator $A \in \mathscr{B}(H)$ is *normal* if $A^*A = AA^*$.

Exercise. Let $A \in \mathscr{B}(H)$ in a complex Hilbert space H. Show that A is normal if and only if

$$\|A^*u\| = \|Au\| \tag{87}$$

for all $u \in H$. What might go wrong in a real Hilbert space?

Example. Self-adjoint operators $A = A^*$ are clearly normal.

Example. Let $H = \mathbb{C}$. Then $\mathscr{B}(H) \cong \mathbb{C}^{1\times 1} \cong \mathbb{C}$. All matrices $[A] \in \mathbb{C}^{1\times 1}$ are normal. Here $[A]^* = [A]$ if and only if $A \in \mathbb{R}$.

Example. Diagonal matrix $[D] \in \mathbb{C}^{n\times n}$ is always normal.

Example. Let $A = \begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix} \in \mathbb{C}^{2\times 2}$. Then $A^* = \begin{bmatrix} 0 & c^* \\ b^* & 0 \end{bmatrix}$. Thus here A is selfadjoint if and only if $c = b^*$. Moreover,

$$\begin{aligned} A^*A &= \begin{bmatrix} 0 & c^* \\ b^* & 0 \end{bmatrix} \begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix} &= \begin{bmatrix} |c|^2 & 0 \\ 0 & |b|^2 \end{bmatrix}, \\ AA^* &= \begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix} \begin{bmatrix} 0 & c^* \\ b^* & 0 \end{bmatrix} &= \begin{bmatrix} |b|^2 & 0 \\ 0 & |c|^2 \end{bmatrix}. \end{aligned}$$

Hence here A is normal if and only if $|b| = |c|$.

$$\Lambda = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} & \cdots & \Lambda_{1n} \\ 0 & \Lambda_{22} & \Lambda_{23} & \cdots & \Lambda_{2n} \\ 0 & 0 & \Lambda_{33} & \cdots & \Lambda_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \Lambda_{nn} \end{bmatrix} \overset{\Lambda^*\Lambda=\Lambda\Lambda^*}{\Longrightarrow} \Lambda = \begin{bmatrix} \Lambda_{11} & 0 & 0 & \cdots & 0 \\ 0 & \Lambda_{22} & 0 & \cdots & 0 \\ 0 & 0 & \Lambda_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \Lambda_{nn} \end{bmatrix}.$$

Figure 18: Any normal (upper or lower) triangular matrix is diagonal! And what if the matrix would be symmetric and triangular?

Normal triangular matrix is diagonal!

Let $\Lambda \in \mathbb{C}^{n\times n}$. Then

$$\begin{aligned} (\Lambda^*\Lambda)_{kk} &= \sum_{\ell=1}^{n} \overline{\Lambda_{\ell k}}\,\Lambda_{\ell k} &&= \sum_{\ell=1}^{n} |\Lambda_{\ell k}|^2, \\ (\Lambda\Lambda^*)_{kk} &= \sum_{\ell=1}^{n} \Lambda_{k\ell}\overline{\Lambda_{k\ell}} &&= \sum_{\ell=1}^{n} |\Lambda_{k\ell}|^2. \end{aligned}$$

Let Λ be normal ($\Lambda^*\Lambda = \Lambda\Lambda^*$) and upper triangular ($\Lambda_{ij} = 0$ if $i > j$). Then

$$|\Lambda_{nn}|^2 = \sum_{\ell=1}^{n} |\Lambda_{n\ell}|^2 = (\Lambda\Lambda^*)_{nn} = (\Lambda^*\Lambda)_{nn} = \sum_{\ell=1}^{n} |\Lambda_{\ell n}|^2 = |\Lambda_{nn}|^2 + \sum_{\ell=1}^{n-1} |\Lambda_{\ell n}|^2,$$

which shows that $\Lambda_{\ell n} = 0$ for $\ell \in \{1, \cdots, n-1\}$. Thus

$$\Lambda = \begin{bmatrix} \widetilde{\Lambda} & O \\ O^* & \Lambda_{nn} \end{bmatrix},$$

where $\widetilde{\Lambda} \in \mathbb{C}^{(n-1)\times(n-1)}$ is a normal upper triangular matrix and $O \in \mathbb{R}^{(n-1)\times 1}$ is the zero vector. By reducing dimensions $n > n-1 > \cdots > 1$, we get:

Theorem. *Normal (upper or lower) triangular matrices are diagonal.*
(A trivial result: symmetric triangular matrices are diagonal and real.)

10.3 Unitary operators

Unitary operators are invertible linear mappings that preserve inner products (i.e. distances) between the Hilbert spaces. Informally, unitary operators behave much like unimodular complex numbers.

Definition. Operator $U \in \mathscr{B}(H, G)$ is *unitary* if $U^* = U^{-1}$. That is,

$$U^*U = I_H \quad \text{and} \quad UU^* = I_G, \tag{88}$$

where $I_G : G \to G$ and $I_H : H \to H$ are the identity operators.
(Especially, unitary operators are normal.)

Remark. Here $U^* = U^{-1}$ is equivalent to that $U \in \mathscr{B}(H, G)$ is surjective with

$$\langle Uv, Uw \rangle_G = \langle v, w \rangle_H \tag{89}$$

for all $v, w \in H$; alternatively, $U \in \mathscr{B}(H, G)$ is surjective satisfying

$$\|Uv\|_G = \|v\|_H \tag{90}$$

for all $v \in H$. Notice that (90) just implies

$$\langle U^*Uv, v \rangle = \langle Uv, Uv \rangle = \|Uv\|^2 = \|v\|^2 = \langle v, v \rangle,$$

giving $U^*U = I$ by the Weak Formulation Theorem in a complex Hilbert space; however, (90) does not mean that U would be automatically surjective, as we shall soon see in an example in a sequence space.

Example. Let $H = \mathbb{C}$. Then $\mathscr{B}(H) \cong \mathbb{C}^{1\times 1} \cong \mathbb{C}$. Matrix $[U] \in \mathbb{C}^{1\times 1}$ would be unitary if and only if $|U| = 1$ for $U \in \mathbb{C}$. That is, $U = \mathrm{e}^{\mathrm{i}t}$ for some $t \in \mathbb{R}$.

Example. Diagonal matrix $[D] \in \mathbb{C}^{n\times n}$ is unitary if and only if all its diagonal elements D_{kk} have absolute value $|D_{kk}| = 1$.

Example. Matrix $[U] \in \mathbb{K}^{n\times n}$ is unitary if and only if its column vectors (respectively row vectors) form an orthonormal basis for $\mathbb{K}^n$.

Example. Let $H = \ell^2(\mathbb{Z})$. Define $L, R : H \to H$ by

$$Lv(k) := v(k+1)$$

and

$$Rv(k) := v(k-1)$$

for all $k \in \mathbb{Z}$. Then $L, R \in \mathscr{B}(H)$ are unitary, and actually $L^* = R$ and $R^* = L$. (Here L stands for $Left$ and R for $Right$, for some reason...).

Example. For $v \in H := \ell^2(\mathbb{Z}^+)$, let $v_k := v(k)$. Define $L, R : H \to H$ by

$$Lv = L(v_1, v_2, v_3, \cdots) := (v_2, v_3, v_4, \cdots)$$

and

$$Rv = R(v_1, v_2, v_3, \cdots) := (0, v_1, v_2, \cdots).$$

Then $L, R \in \mathscr{B}(H)$ are not unitary: $L^* = R$, $R^* = L$, but $LR = I \neq RL$, as

$$RLv = (0, v_2, v_3, \cdots).$$

Here $\langle Ru, Rv \rangle = \langle u, v \rangle$ for all $u, v \in H$, but R is not surjective.

Example. For $u \in \mathscr{S} := \mathscr{S}(\mathbb{R}^d)$, the Fourier transform $\widehat{u} \in \mathscr{S}$ is defined by

$$\widehat{u}(\eta) = \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}2\pi y\cdot\eta}\, u(y)\, \mathrm{d}y.$$

For all $u, v \in \mathscr{S} \subset L^2 := L^2(\mathbb{R}^d)$, here

$$\langle \widehat{u}, \widehat{v} \rangle_{L^2} = \langle u, v \rangle_{L^2},$$

so that the bijective linear mapping $(u \mapsto \widehat{u}) : \mathscr{S} \to \mathscr{S}$ extends to the unitary Fourier transform $\mathscr{F} = (u \mapsto \widehat{u}) : L^2 \to L^2$.

Exercise. Let H be a Hilbert space. The *exponential* of $A \in \mathscr{B}(H)$ is

$$\exp(A) := \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

Show that $A^* = A$ if and only if $\exp(\mathrm{i}tA)$ is unitary for all $t \in \mathbb{R}$.

10.4 Positive operators

Positive operators on complex Hilbert space turn out to be self-adjoint operators which in certain ways resemble positive real numbers in their behavior.

Definition. Let H be a complex Hilbert space. $P \in \mathscr{B}(H)$ is *positive*, denoted by $P \geq 0$ (or $0 \leq P$), if

$$\langle Pu, u \rangle \geq 0$$

for all $u \in H$.

Example. Let $H = \mathbb{C}$. Then $\mathscr{B}(H) \cong \mathbb{C}^{1\times 1} \cong \mathbb{C}$. Matrix $[P] \in \mathbb{C}^{1\times 1}$ would be positive if and only if $0 \leq P \in \mathbb{R}$.

Example. Diagonal matrix $[D] \in \mathbb{C}^{n\times n}$ is positive if and only if all its diagonal elements D_{kk} are non-negative.

Remark. If $P \in \mathscr{B}(H)$ is positive then for all $u \in H$ we have

$$0 \leq \langle Pu, u \rangle = \langle u, Pu \rangle^* \overset{\text{real}}{=} \langle u, Pu \rangle = \langle P^* u, u \rangle,$$

so that $\langle Pu, u \rangle = \langle P^* u, u \rangle$. As H is a complex Hilbert space, this means self-adjointness $P = P^*$ by the Weak Formulation Theorem.

Example. If $A \in \mathscr{B}(H)$ then $A^* A \geq 0$, as

$$\langle A^* A u, u \rangle = \langle Au, Au \rangle = \|Au\|^2 \geq 0.$$

Especially, orthogonal projections P are positive, as $P = P^* P$; alternatively,

$$\langle Pu, u \rangle = \langle Pu, Pu \rangle + \langle Pu, (I - P)u \rangle = \|Pu\|^2 \geq 0.$$

Example. If $P \geq 0$ then $P^k \geq 0$ for all $k \in \mathbb{N}$: If here k is even, then clearly

$$\langle P^k u, u \rangle = \langle P^{k/2} P^{k/2} u, u \rangle = \langle P^{k/2} u, P^{k/2} u \rangle = \|P^{k/2} u\|^2 \geq 0,$$

and

$$\langle P^{k+1} u, u \rangle = \langle P^{k/2} P P^{k/2} u, u \rangle = \langle P(P^{k/2} u), P^{k/2} u \rangle \overset{P \geq 0}{\geq} 0.$$

Example. Let $P \geq 0$ and $\|P\| \leq 1$. Then also $I - P \geq 0$:

$$\langle (I - P)u, u \rangle = \|u\|^2 - \langle Pu, u \rangle \in [0, \|u\|^2]$$

by $P \geq 0$ and by the Cauchy–Schwarz inequality. Combining this with the previous Example, we notice that here $(I - P)^k \geq 0$ for all $k \in \mathbb{N}$. Also, we see that $\|I - P\| \leq 1$ by the Norm Symmetry Lemma.

Lemma (Existence of Positive Square Root). *Let $0 \leq P \in \mathscr{B}(H)$. Then there exists $R \geq 0$ such that $R^2 = P$ (here $R = P^{1/2}$ is called the* positive square root *of P).*

Proof. By scaling $P \mapsto \lambda P$, we may assume that P itself satisfies $\|P\| \leq 1$. For $f(x) = (1 - x)^{1/2}$ the Taylor–Maclaurin series

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \, x^k$$

converges when $|x| \leq 1$. Here $f^{(0)}(0) = f(0) = 1$, and otherwise $f^{(k)}(0) < 0$. As $f(1 - x) = x^{1/2}$, we define

$$R = f(I - P) := \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \, (I - P)^k.$$

Clearly as $f(1 - x)^2 = x$, also $R^2 = P$, and by the previous Example,

$$\begin{aligned} \langle Ru, u \rangle &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \, \langle (I - P)^k u, u \rangle \\ &\overset{(I-P)^k \geq 0,\ \|I-P\| \leq 1,\ f^{(k+1)}(0) < 0}{\geq} \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \, \|u\|^2 \\ &= f(1) \, \|u\|^2 \overset{f(1)=0}{=} 0, \end{aligned}$$

so that $R \geq 0$. ∎

Remark. From the proof of the previous Lemma, we see that

$$R = P^{1/2} = \|P\|^{1/2} f(I - P/\|P\|)$$

is a limit of polynomials in variable P. Notice that here

$$\|Pu\|^2 \leq \left\|P^{1/2}\right\|^2 \left\|P^{1/2}u\right\|^2 = \|P^{1/2}P^{1/2}\| \langle P^{1/2}u, P^{1/2}u\rangle = \langle Pu, \|P\|u\rangle.$$

The positive square root is unique by the following result:

Proposition (Uniqueness of Positive Square Root). *Suppose $Q, R \geq 0$ such that $Q^2 = P = R^2$. Then $Q = R$.*

Proof. Notice that $PR = RP$ and $PQ = QP$. Above, $R = P^{1/2} \geq 0$ was obtained as the limit of polynomials in P. Thus also $QR = RQ$. Let $u \in H$ and $v := Qu - Ru$. Then

$$\begin{aligned}
\|Q^{1/2}v\|^2 + \|R^{1/2}v\|^2 &= \langle Qv, v\rangle + \langle Rv, v\rangle \\
&= \langle (Q+R)v, v\rangle \\
&= \langle (Q+R)(Q-R)u, v\rangle \\
&= \langle (Q^2 - R^2 - QR + RQ)u, v\rangle \\
&\overset{QR=RQ}{=} \langle (Q^2 - R^2)u, v\rangle \overset{Q^2=R^2}{=} 0.
\end{aligned}$$

Thus $Q^{1/2}v = 0 = R^{1/2}v$, so that $Qv = 0 = Rv$. Finally,

$$\|Qu - Ru\|^2 = \langle (Q-R)^2u, u\rangle = \langle (Q-R)v, u\rangle = 0,$$

implying $Q = R$. □

Definition. For $A \in \mathscr{B}(H)$, let the *absolute value* be $|A| := (A^*A)^{1/2} \in \mathscr{B}(H)$. Notice that $|A^*| = |A|$ if and only if A is normal.

Polar decomposition in $\mathscr{B}(H)$. Next we generalize the usual polar decomposition $\lambda = \mathrm{e}^{\mathrm{i}\varphi}|\lambda| \in \mathbb{C}$:

Theorem. *Let $A \in \mathscr{B}(H)$. Then*

$$A = E|A|, \tag{91}$$

where $E \in \mathscr{B}(H)$ is a partial isometry: this means $\|Ew\| = \|w\|$ whenever $w \in \ker(E)^\perp = \ker(A)^\perp$.

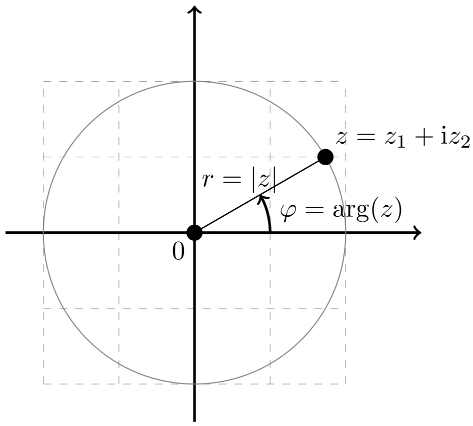


Figure 19: Polar coordinates (r, φ) for point $z = z_1 + \mathrm{i}z_2 \in \mathbb{C}$.

Proof. We have $M := \ker(A) = \ker(|A|)$, since

$$\||A|v\|^2 = \langle |A|v, |A|v\rangle = \langle |A|^2 v, v\rangle = \langle A^*Av, v\rangle = \langle Av, Av\rangle = \|Av\|^2.$$

Moreover, $M \perp \mathrm{ran}(|A|)$, because

$$\mathrm{ran}(|A|)^\perp \overset{(77)}{=} \ker(|A|^*) \overset{|A|^*=|A|}{=} \ker(|A|) = M.$$

Thus $\mathrm{ran}(|A|)$ is a dense subspace of $M^\perp$. Define $E : M + \mathrm{ran}(|A|) \to H$ by

$$E(z + |A|v) := Av,$$

where $(z, v) \in M \times H$, so that $\|Ew\| = \|w\|$ for all $w \in \mathrm{ran}(|A|)$. Finally, operator E extends by continuity to $H = M \oplus M^\perp$. $\square$

Example. Let $R \in \mathscr{B}(H)$ be the right-shift operator in Hilbert space $H = \ell^2(\mathbb{Z}^+)$. Then $R^* = L$ is the left-shift operator,

$$\begin{aligned}
&R(u_1, u_2, u_3, u_4, u_5, \cdots) = (0, u_1, u_2, u_3, u_4, \cdots),\\
&R^*(v_1, v_2, v_3, v_4, v_5, \cdots) = (v_2, v_3, v_4, v_5, v_6, \cdots),\\
&R^*R = LR = I.
\end{aligned}$$

Thereby we have the polar decomposition $R = E|R|$, where $|R| = (R^*R)^{1/2} = I$ with the partial isometry $E = R$. And since

$$L^*L(v_1, v_2, v_3, v_4, v_5, \cdots) = (0, v_2, v_3, v_4, v_5, v_6, \cdots),$$

we have the polar decomposition $L = D|L|$, where $|L| = (L^*L)^{1/2} = L^*L$ with the partial isometry $D = L$.

11 Spectral properties in Hilbert spaces

Recall that the *spectrum* of linear $A : V \to V$ is

$$\sigma(A) := \{\lambda \in \mathbb{C} : \ \lambda I - A \text{ is not bijective}\}.$$

If $Au = \lambda u$ where $\lambda \in \mathbb{C}$ and $0 \neq u \in V$ then $\lambda \in \sigma(A)$ is called an *eigenvalue* corresponding to the *eigenvector* u. The subset of the eigenvalues is called the *point spectrum* of A.

Proposition. *Possible eigenvalues of positive $A \in \mathscr{B}(H)$ are non-negative.*

Proof. Let $Au = \lambda u$, where A is positive, $0 \neq u \in H$ and $\lambda \in \mathbb{C}$. Then

$$0 \leq \langle Au, u\rangle = \langle \lambda u, u\rangle = \lambda \langle u, u\rangle = \lambda \|u\|^2,$$

so that $0 \leq \lambda$. □

Proposition. *Possible eigenvalues of self-adjoint $A \in \mathscr{B}(H)$ are real.*

Proof. Let $\lambda \in \mathbb{C}$ be an eigenvalue of self-adjoint $A = A^* \in \mathscr{B}(H)$ with eigenvector $u \in H$. Then

$$\begin{aligned} \lambda\langle u, u\rangle &= \langle \lambda u, u\rangle \overset{Au=\lambda u}{=} \langle Au, u\rangle = \langle u, A^* u\rangle \overset{A^*=A}{=} \langle u, Au\rangle \\ &\overset{Au=\lambda u}{=} \langle u, \lambda u\rangle = \overline{\lambda}\langle u, u\rangle. \end{aligned}$$

Thus $\lambda = \overline{\lambda}$ (i.e. $\lambda \in \mathbb{R}$), because $\langle u, u\rangle = \|u\|^2 \neq 0$. □

11.1 Extreme eigenvalues of compact self-adjoint operator

Recall that operator $A \in \mathscr{B}(H)$ is *compact* if $(Au_k)_{k=1}^\infty$ has a converging subsequence whenever $(u_k)_{k=1}^\infty$ is a bounded sequence in H, and then we wrote $A \in \mathscr{K}(H)$. Notice that $0 \in \sigma(A)$ for $A \in \mathscr{K}(H)$ in infinite-dimensional H, because there are no compact linear bijections in infinite-dimensional Hilbert (or Banach) spaces, due to the Open Mapping Theorem. The following result will be the initial step when diagonalizing compact self-adjoint operators:

Compact Self-Adjoint Eigenvalue Lemma. *Let $A^* = A \in \mathscr{K}(H)$ in Hilbert space H. Then A has an eigenvalue $\lambda \in \{+\|A\|, -\|A\|\}$.*

Proof. The non-trivial case $\|A\| > 0$ is enough for us. By the Norm Symmetry Lemma, the set

$$\{\langle Au, u\rangle \in \mathbb{R} : \ u \in H, \ \|u\| \leq 1\}$$

has accumulation point $\lambda \in \{\pm\|A\|\}$. For all $k \in \mathbb{Z}^+$, choose $u_k \in H$ so that $\|u_k\| \leq 1$ and

$$\lim_{k\to\infty} \langle Au_k, u_k\rangle = \lambda.$$

Since A is compact, sequence $(Au_k)_{k=1}^{\infty}$ has a converging subsequence; we may simply assume that $v := \lim_{k\to\infty} Au_k \in H$ exists. Now

$$\begin{array}{rcl} 0 & \leq & \|Au_k - \lambda u_k\|^2 \\ & = & \|Au_k\|^2 + \lambda^2\|u_k\|^2 - 2\lambda\langle Au_k, u_k\rangle \\ & \leq & \|A\|^2 + \lambda^2 - 2\lambda\langle Au_k, u_k\rangle \\ & \xrightarrow[k\to\infty]{} & \lambda^2 + \lambda^2 - 2\lambda^2 \quad = \quad 0, \end{array}$$

implying that

$$\lim_{k\to\infty} \lambda u_k = \lim_{k\to\infty} Au_k = v.$$

Moreover, $\|v\| > 0$ as $\|A\| > 0$. By continuity,

$$Av = A(\lim_{k\to\infty} \lambda u_k) = \lambda \lim_{k\to\infty} Au_k = \lambda v.$$

This completes the proof. □

11.2 Diagonalization of finite-dimensional matrices

Before going to diagonalization in infinite-dimensional Hilbert spaces, we review the finite-dimensional case:

Recall that a matrix $A \in \mathbb{C}^{n\times n}$ can be *diagonalized* if there is invertible $S \in \mathbb{C}^{n\times n}$ for which $\Lambda = S^{-1}AS \in \mathbb{C}^{n\times n}$ is diagonal, i.e. $\Lambda_{jk} = 0$ whenever $j \neq k$: then $A = S\Lambda S^{-1}$ and

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \in \mathbb{C}^{n\times n},$$

where $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{C}$ are the eigenvalues of A. The kth column of S is the eigenvector of A corresponding to the eigenvalue λ_k:

$$(AS)_{jk} = (S\Lambda S^{-1}S)_{jk} = (S\Lambda)_{jk} = \sum_{\ell=1}^{n} S_{j\ell}\,\Lambda_{\ell k} = \lambda_k\, S_{jk}.$$

Notice that a matrix $A \in \mathbb{C}^{n\times n}$ can be diagonalized if and only if the geometric and algebraic multiplicities of each eigenvalue coincide: $m_g(\lambda) = m_a(\lambda)$ for all the eigenvalues λ of A.

Example. Let $\lambda \in \mathbb{C}$ be the only eigenvalue of diagonalizable $A \in \mathbb{C}^{n\times n}$. Thus $A = S(\lambda I)S^{-1} = \lambda SS^{-1} = \lambda I$.

Example. $M \in \mathbb{C}^{n\times n}$ is *triangular* if it is *lower triangular* ($M_{jk} = 0$ whenever $j < k$) or *upper triangular* ($M_{jk} = 0$ whenever $j > k$). If triangular A has zero diagonal, then $\lambda = 0$ is its only eigenvalue. By the previous example, such A can be diagonalized only if $A = 0$.

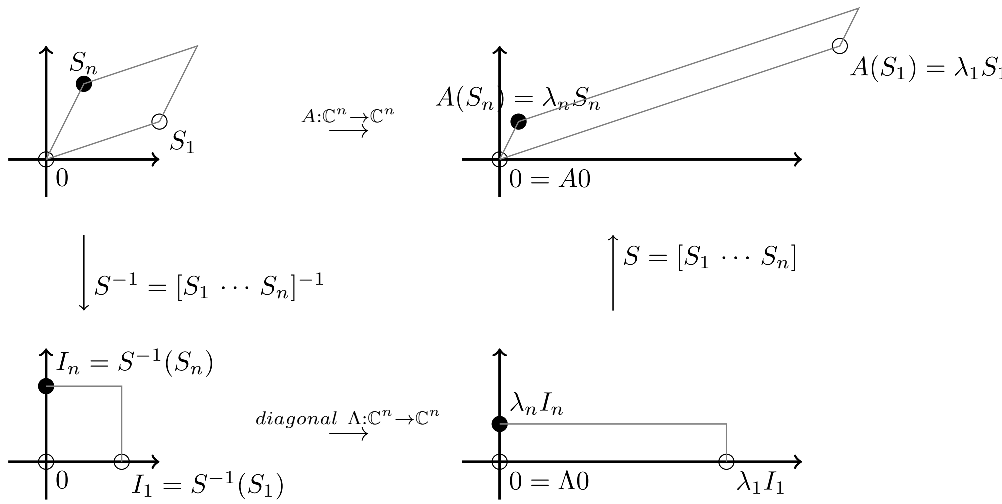


Figure 20: Diagonalization $\Lambda = S^{-1}AS$. In other words, $A = S\Lambda S^{-1}$.

Example. $AS = S\Lambda$ when

$$A = \begin{bmatrix} a & b & c \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad S = \begin{bmatrix} 1 & -b & -c \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix}, \quad \Lambda = \begin{bmatrix} a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus if $a \neq 0$ we have $A = S\Lambda S^{-1}$ (if $a = 0$ then S is not invertible).

Remark. Soon we recall that a matrix is *normal* if and only if it has a unitary diagonalization. In the example above, A is a normal matrix if and only if $b = 0 = c$: thus some non-normal matrices can be diagonalized (yet not unitarily diagonalized).

Example. Let $[A] = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$, where $b \neq 0$. Now we have the *characteristic polynomial*

$$p_A(z) = \det[A - zI] = \det \begin{bmatrix} 1-z & b \\ 0 & 1-z \end{bmatrix} = (1-z)^2.$$

Then the equation $(A - \lambda I)(x) = 0$ has only solution $x = (t, 0)$ for constants $t \in \mathbb{C}$. This $[A]$ is an example of a non-diagonalizable matrix.

Functions of square matrices. An analytic function $f : \mathbb{C} \to \mathbb{C}$ can be presented as the power series

$$f(z) = \sum_{k=0}^{\infty} c_k \, z^k, \tag{92}$$

where $c_k = f^{(k)}(0)/k!$. For instance, functions exp, cos, sin are analytic. Define $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ by

$$f(A) = \sum_{k=0}^{\infty} c_k\, A^k. \tag{93}$$

Here $f(A)_{ij} \neq f(A_{ij})$ often. But with diagonalization $A = S\Lambda S^{-1}$,

$$f(A) = \sum_{k=0}^{\infty} c_k\, (S\Lambda S^{-1})^k = \ldots = S\left(\sum_{k=0}^{\infty} c_k\, \Lambda^k\right) S^{-1} = S\, f(\Lambda)\, S^{-1},$$

where $f(\Lambda) \in \mathbb{C}^{n\times n}$ is diagonal with $f(\Lambda)_{jj} = f(\Lambda_{jj}) \in \mathbb{C}$. Nice!

Example. Let $A = S\Lambda S^{-1}$, where $\Lambda = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$. Then

$$\begin{aligned} \exp(A) &= S\, \exp(\Lambda)\, S^{-1}, \\ |A|^{1/2} &= S|\Lambda|^{1/2}\, S^{-1}, \end{aligned}$$

where

$$\exp(\Lambda) = \begin{bmatrix} \mathrm{e}^a & 0 & 0 \\ 0 & \mathrm{e}^b & 0 \\ 0 & 0 & \mathrm{e}^c \end{bmatrix}, \quad |\Lambda|^{1/2} = \begin{bmatrix} \sqrt{|a|} & 0 & 0 \\ 0 & \sqrt{|b|} & 0 \\ 0 & 0 & \sqrt{|c|} \end{bmatrix}.$$

So, what would be $|A|$ then?

Application to differential equations: Let $A \in \mathbb{C}^{n\times n}$ and let the unknown functions $u_1, \cdots, u_n : \mathbb{R} \to \mathbb{C}$ satisfy

$$u'(t) = A\, u(t),$$

where naturally $(u'(t))_j = (u_j)'(t) = \frac{\mathrm{d}}{\mathrm{d}t} u_j(t)$. Then

$$u(t) = \exp(tA)\, u(0).$$

If here $A = S\Lambda S^{-1}$ then $\exp(tA) = S \exp(t\Lambda) S^{-1}$, which is easy to find.

Unitary triangulation of square matrices. For a square matrix $A_n \in \mathbb{C}^{n\times n}$, we next find a unitary matrix $U_n \in \mathbb{C}^{n\times n}$ and an upper triangular matrix $\Lambda_n \in \mathbb{C}^{n\times n}$ such that

$$A_n = U_n\, \Lambda_n\, U_n^*. \tag{94}$$

It is enough to reduce case n to case $n-1$ (here case $n = 1$ is trivial). Take an eigenvalue $\lambda \in \mathbb{C}$ with a normalized eigenvector $v \in \mathbb{C}^{n\times 1}$:

$$A_n v = \lambda v, \quad \|v\| = 1.$$

By the Gram–Schmidt process, find a unitary matrix $V \in \mathbb{C}^{n\times n}$ with first column v. So

$$A_n = V \begin{bmatrix} \lambda & w \\ 0 & A_{n-1} \end{bmatrix} V^* \quad \text{for some } w \in \mathbb{C}^{1\times(n-1)}, \quad \text{where}$$

$A_{n-1} = U_{n-1}\Lambda_{n-1}U_{n-1}^* \in \mathbb{C}^{(n-1)\times(n-1)}$ by case $n-1$ of (94). Let

$$U_n := V \begin{bmatrix} 1 & 0 \\ 0 & U_{n-1} \end{bmatrix}, \quad \Lambda_n = \begin{bmatrix} \lambda & wU_{n-1} \\ 0 & \Lambda_{n-1} \end{bmatrix}.$$

Then $A_n = U_n\Lambda_nU_n^*$, where U_n is unitary and Λ_n upper triangular.

Unitary diagonalization of normal matrices. Thus, for any $A \in \mathbb{C}^{n\times n}$ there is a unitary triangulation $A = U\Lambda U^*$, where $U \in \mathbb{C}^{n\times n}$ is unitary and $\Lambda \in \mathbb{C}^{n\times n}$ is upper triangular. It is easy to see that here Λ is normal if and only if A is normal. As normal triangular matrices are diagonal (see page 94), we get:

Normal matrices can be diagonalized by unitary matrices!

More precisely:

Theorem. *Conditions (1) and (2) are equivalent:*
(1) $A \in \mathbb{C}^{n\times n}$ *is normal (that is,* $A^*A = AA^*$*).*
(2) $A = U\Lambda U^*$ *for unitary* $U \in \mathbb{C}^{n\times n}$ *and diagonal* $\Lambda \in \mathbb{C}^{n\times n}$.

Remark: In this result on the unitary diagonalization $A = U\Lambda U^*$, it is easy to see that $A^* = A$ if and only if $\Lambda^* = \Lambda$ (Why?). So, a normal matrix is symmetric if and only if its eigenvalues are real. However, there are non-normal diagonalizable matrices with real eigenvalues: see the example on page 102.

Example. Let $A \in \mathbb{C}^{n\times n}$ be normal, that is $A^*A = AA^*$. We saw that this is equivalent to the existence of unitary diagonalization $A = U\Lambda U^*$. For normal $A \in \mathbb{C}^{n\times n}$ and **all** its eigenvalues $\lambda \in \mathbb{C}$, it is then easy to prove:

- $A^* = A^{-1}$ (unitary A) iff $|\lambda| = 1$.
- $A^* = A$ (symmetric A) iff $\lambda \in \mathbb{R}$.
- $\langle Au, u\rangle \geq 0$ for all $u \in \mathbb{C}^n$ (positive A) iff $\lambda \geq 0$.
- $A^* = A = A^2$ (orthogonal projection A) iff $\lambda \in \{0, 1\}$.

In particular, orthogonal projections are always positive, and positive operators are always symmetric. The only unitary positive operator in $\mathbb{C}^{n\times n}$ is the identity I. The only unitary orthogonal projection in $\mathbb{C}^{n\times n}$ is the identity I.

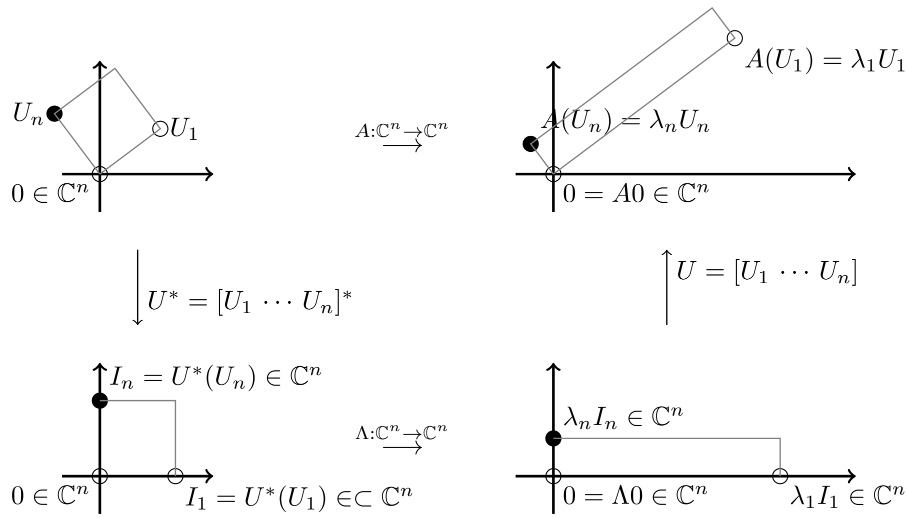


Figure 21: Idea of unitary diagonalization $\Lambda = U^* A U$ of normal $A \in \mathbb{C}^{n\times n}$. Equivalently, this means $A = U\Lambda U^*$. Unitary operations U^* and U preserve distances and angles.

Example. Let us find a unitary diagonalization for the rotation matrix

$$A = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

It is easy to find the characteristic polynomial

$$\det[A - zI] = z^2 - 2\cos(\varphi) + 1 = (z - \lambda_1)(z - \lambda_2),$$

where $\lambda_k = \cos(\varphi) \pm \mathrm{i}\sin(\varphi) \overset{\text{Euler}}{=} \mathrm{e}^{\pm \mathrm{i}\varphi}$. Then $A = U\Lambda U^*$, where e.g.

$$\Lambda = \begin{bmatrix} \mathrm{e}^{\mathrm{i}\varphi} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\varphi} \end{bmatrix}, \qquad U = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathrm{i} & \mathrm{i} \\ 1 & -1 \end{bmatrix}.$$

11.3 Diagonalization of compact self-adjoint operators

Now we treat diagonalization in infinite-dimensional Hilbert spaces.

Diagonalization Theorem. *Let H be infinite-dimensional Hilbert space and $A^* = A \in \mathscr{K}(H)$. Then A has eigenvalues $\lambda_k \in \mathbb{R}$ and orthonormal eigenvectors $u_k \in H$ with $|\lambda_{k+1}| \le |\lambda_k|$, $\lim_{k\to\infty} \lambda_k = 0$, and for all $u \in H$*

$$Au = \sum_{k=1}^{\infty} \lambda_k \langle u, u_k \rangle u_k. \tag{95}$$

Especially, $\sigma(A) = \{0\} \cup \{\lambda_k\}_{k=1}^{\infty}$. Moreover, $\|(\lambda I - A)^{-1}\| = 1/\mathrm{dist}(\{\lambda\}, \sigma(A))$ whenever $\lambda \notin \sigma(A)$.

Remark. Formula (95) is the *spectral decomposition* of $A = A^* \in \mathscr{K}(H)$. This is an analogue of the unitary diagonalization

$$A = U\Lambda U^*$$

of symmetric matrices $A^* = A \in \mathbb{C}^{n\times n}$, where $\Lambda \in \mathbb{R}^{n\times n}$ is the diagonal matrix of eigenvalues $\lambda_k \in \mathbb{R}$, and the columns of unitary matrix $U \in \mathbb{C}^{n\times n}$ are the corresponding orthonormal eigenvectors $u_k \in \mathbb{C}^{n\times 1}$: $Au_k = \lambda_k u_k$.

Proof of the Diagonalization Theorem. We shall inductively obtain

$$\begin{aligned} u &= \sum_{k=1}^{n} \langle u, u_k\rangle\, u_k + v_n, & v_n \in H_n := \left(\{u_k\}_{k=1}^n\right)^{\perp}, \\ Au &= \sum_{k=1}^{n} \lambda_k \,\langle u, u_k\rangle\, u_k + A(v_n), & A(v_n) \in H_n. \end{aligned}$$

How? The Compact Self-Adjoint Eigenvalue Lemma gives $\lambda_1 \in \mathbb{R}$ and $u_1 \in H_0 := H$ such that for $A_0 := A$

$$\|u_1\| = 1, \quad A_0 u_1 = \lambda_1 u_1, \quad \|A_0\| = |\lambda_1|.$$

Closed vector subspace $H_n := \left(\{u_k\}_{k=1}^n\right)^{\perp}$ is proper for each $n \geq 1$, as H is infinite-dimensional. Notice that $A(v_n) \in H_n$, because if $k \leq n$ then

$$\langle u_k, A(v_n)\rangle \overset{A=A^*}{=} \langle A(u_k), v_n\rangle = \langle \lambda_k\, u_k, v_n\rangle = \lambda_k \,\langle u_k, v_n\rangle = 0.$$

Operator $A_n := A|_{H_n} \in \mathscr{B}(H_n)$ is compact and self-adjoint, so that by the Compact Self-Adjoint Eigenvalue Lemma we can take $\lambda_{n+1} \in \mathbb{R}$ and $u_{n+1} \in H_n$ such that

$$\|u_{n+1}\| = 1, \quad A(u_n) = \lambda_{n+1}\, u_{n+1}, \quad \|A_n\| = |\lambda_{n+1}|.$$

Clearly, sequence $(u_n)_{n=1}^\infty \subset H$ is orthonormal, and $|\lambda_{n+1}| \leq |\lambda_n|$. By compactness, $(Au_n)_{n=1}^\infty$ has a converging subsequence, so that

$$\|Au_k - Au_n\| = \|\lambda_k u_k - \lambda_n u_n\| \overset{k\neq n}{=} \sqrt{\lambda_k^2 + \lambda_n^2} \geq |\lambda_n| \geq |\lambda_{n+1}|$$

implies $\lim_{n\to\infty} |\lambda_n| = 0$. Therefore also

$$\lim_{n\to\infty} Au_n = \lim_{n\to\infty} \lambda_n u_n = 0.$$

We now obtain (95) by

$$\|Av_n\| = \|A_n v_n\| \leq \|A_n\|\,\|v_n\| = |\lambda_n|\,\|v_n\| \leq |\lambda_n|\,\|u\| \xrightarrow[n\to\infty]{} 0.$$

Next, $0 \in \sigma(A)$, because the spectrum is closed and $\sigma(A) \ni \lambda_n \to 0$ as $n \to \infty$. (Alternatively: there are no compact linear bijections in infinite-dimensional

Banach spaces, so $0 \in \sigma(A)$ for $A \in \mathscr{K}(H)$ here.) Thus, $\{0\} \cup \{\lambda_k\}_{k=1}^{\infty} \subset \sigma(A)$. Finally, suppose $0 \neq \lambda \neq \lambda_k$ for all $k \in \mathbb{Z}^+$. We have to show that $\lambda \notin \sigma(A)$:

$$(\lambda I - A)u = \lambda\, Pu + \sum_{k=1}^{\infty} (\lambda - \lambda_k)\, \langle u, u_k \rangle\, u_k,$$

where $Pu := u - \sum_{k=1}^{\infty} \langle u, u_k \rangle u_k \in \ker(A)$. It is easy to see that

$$(\lambda I - A)^{-1} v = \lambda^{-1} Pv + \sum_{k=1}^{\infty} (\lambda - \lambda_k)^{-1}\, \langle v, u_k \rangle\, u_k.$$

Moreover, here

$$\begin{aligned} \|(\lambda I - A)^{-1}\| &= \sup \left\{ |\lambda|^{-1}, |\lambda - \lambda_k|^{-1} :\ k \in \mathbb{Z}^+ \right\} \\ &= 1/\mathrm{dist}(\{\lambda\}, \sigma(A)). \end{aligned}$$ □

Functions of operators. Now linear compact $A^* = A : H \to H$ was spectrally decomposed as

$$Au = \sum_{k=1}^{\infty} \lambda_k\, \langle u, u_k \rangle\, u_k.$$

For nice-enough $f : \mathbb{C} \to \mathbb{C}$, we can define operator $f(A)$ by

$$f(A)u := \sum_{k=1}^{\infty} f(\lambda_k)\, \langle u, u_k \rangle\, u_k.$$

Example. From the spectral decomposition (95) we see that $0 \leq A \in \mathscr{K}(H)$ means $\lambda_k \geq 0$ for all $k \in \mathbb{Z}^+$. Let $f(\lambda) = \sqrt{\lambda}$ for $\lambda \geq 0$. Then we obtain the positive square root $f(A) = A^{1/2} \in \mathscr{K}(H)$.

Example. Define $g : \mathbb{C} \to \mathbb{C}$ by $g(\lambda) := |\lambda|^p$. This gives operator $g(A) = |A|^p$. So, if A is as in (95) then

$$|A|^p u = \sum_{k=1}^{\infty} |\lambda_k|^p\, \langle u, u_k \rangle\, u_k.$$

Hilbert–Schmidt Spectral Theorem. *Let $A^* = A \in \mathscr{K}(H)$ in Hilbert space H. Then $\sigma(A)$ is at most countable; $\ker(\lambda I - A)$ is finite-dimensional if $0 \neq \lambda \in \sigma(A)$. Also, $\sigma(A) \setminus \{0\}$ is discrete, and*

$$H = \bigoplus_{\lambda \in \sigma(A)} \ker(\lambda I - A).$$

Exercise. Prove the Hilbert–Schmidt Spectral Theorem by using the diagonalization of compact self-adjoint operators.

Exercise. Let $H = \ell^2(\mathbb{Z}^+)$. Define linear operator $B : H \to H$ by

$$(Bu)_k = u_{k+1}/k.$$

In other words, for $u = (u_k)_{k=1}^{\infty} \in H$ we have

$$Bu = B(u_1, u_2, u_3, \cdots) = (u_2/1, u_3/2, u_4/3, \cdots). \tag{96}$$

(a) Find the adjoint operator B^*.
(b) Diagonalize $A = B^*B$. In other words, write the spectral decomposition for the compact self-adjoint operator $A = B^*B$.

Finite-dimensional approximation of compact operators. Operator $A \in \mathscr{B}(H)$ has respective *real and imaginary parts* $\mathrm{Re}(A), \mathrm{Im}(A) \in \mathscr{B}(H)$, which are the self-adjoint operators defined by

$$\begin{aligned}\mathrm{Re}(A) &:= (A + A^*)/2,\\ \mathrm{Im}(A) &:= (A - A^*)/(2\mathrm{i}).\end{aligned}$$

That is, $A = \mathrm{Re}(A) + \mathrm{i}\,\mathrm{Im}(A)$. Thus $A \in \mathscr{B}(H)$ is compact if and only if $\|A_k - A\| \to 0$ as $k \to \infty$ for some operators $A_k \in \mathscr{B}(H)$ with $\dim(A_k(H)) < \infty$. There is no analogous statement for arbitrary Banach spaces, by Enflo [6].

Informal example. Hilbert–Schmidt operator $A : L^2(M) \to L^2(M)$ is of the form

$$Av(x) = \int_M K_A(x, y)\, v(y)\, \mathrm{d}y,$$

where $K_A \in L^2(M \times M)$. Then A is compact, and

$$\|A\|_{L^2(M) \to L^2(M)} \leq \|K_A\|_{L^2(M \times M)}.$$

Here $\langle u, Av\rangle = \langle A^*u, v\rangle$ gives

$$A^*u(y) = \int_M K_A(x, y)^* u(x)\, \mathrm{d}x,$$

so $K_A(x, y)^* = K_{A^*}(y, x)$, and

$$\begin{aligned}K_{\mathrm{Re}(A)} &= (K_A(x, y) + K_A(y, x)^*)/2,\\ K_{\mathrm{Im}(A)} &= (K_A(x, y) - K_A(y, x)^*)/(2\mathrm{i}).\end{aligned}$$

Exercise. Let $(e_k)_{k=1}^{\infty}$ be an orthonormal basis of a separable Hilbert space H. Let $w = (w_k)_{k=1}^{\infty} \in \ell^{\infty} := \ell^{\infty}(\mathbb{Z}^+)$. Show that

$$Au := \sum_{k=1}^{\infty} w_k \langle u, e_k \rangle e_k$$

defines a normal operator $A = (u \mapsto Au) : H \to H$.
For which $w \in \ell^{\infty}$ is A self-adjoint?
For which $w \in \ell^{\infty}$ is A unitary?
For which $w \in \ell^{\infty}$ is A positive?
For which $w \in \ell^{\infty}$ is A compact?

12 Singular value decomposition (SVD)

Next we study singular value decomposition (SVD) and its consequences. Roughly speaking, the SVD presents a compact linear operator A as a composition $A = U\Sigma V^*$, where U, V are unitary operators (i.e. "rotations of the spaces") and Σ a positive diagonal operator. The SVD will be obtained from the diagonalization of compact self-adjoint operators.

12.1 Finite-dimensional case of SVD

Definition. A *singular value decomposition* (SVD) of a matrix $A \in \mathbb{C}^{m\times n}$ is a matrix triple (U, Σ, V) such that

$$A = U\Sigma V^*,$$

where $U = [U_1 \cdots U_m] \in \mathbb{C}^{m\times m}$ and $V = [V_1 \cdots V_n] \in \mathbb{C}^{n\times n}$ are unitary, and $\Sigma \in \mathbb{R}^{m\times n}$ is the diagonal matrix of *singular values* $\Sigma_{jj} = \sigma_j \geq 0$ of A: that is $\Sigma_{jk} = 0$ when $j \neq k$. We also demand that $\Sigma_{jj} = \sigma_j \geq \sigma_{j+1}$ for all j.

Remark: If $A = U\Sigma V^*$ as above, then $A(V_j) = \sigma_j\, U_j$, and

$$\begin{aligned} A^*A &= V\,(\Sigma^*\Sigma)\,V^*, \\ AA^* &= U\,(\Sigma\Sigma^*)\,U^*, \end{aligned}$$

where $\Sigma^*\,\Sigma \in \mathbb{R}^{n\times n}$ and $\Sigma\,\Sigma^* \in \mathbb{R}^{m\times m}$ are positive diagonal matrices, where

$${\sigma_j}^2 = (\Sigma^*\,\Sigma)_{jj} = (\Sigma\,\Sigma^*)_{jj}\,.$$

This suggests that an SVD could be found by the unitary diagonalization!

How to find the SVD? For $A \in \mathbb{C}^{m\times n}$, matrix $A^*A \in \mathbb{C}^{n\times n}$ is normal (A^*A is even positive), so the unitary diagonalization gives us

$$A^*A = V\Lambda V^*,$$

where the diagonal matrix $\Lambda \in \mathbb{C}^{n\times n}$ has the eigenvalues $\lambda_k := \Lambda_{kk}$ of A^*A, with the unitary matrix $V = [V_1 \cdots V_n] \in \mathbb{C}^{n\times n}$ having the corresponding eigenvectors $V_k \in \mathbb{C}^{n\times 1}$. Now

$$\begin{aligned} \langle A(V_j), A(V_k)\rangle &= \langle A^*A(V_j), V_k\rangle \\ &= \langle \lambda_j V_j, V_k\rangle \\ &= \lambda_j \langle V_j, V_k\rangle \\ &\overset{V^*V=I}{=} \begin{cases} \lambda_j & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases} \end{aligned}$$

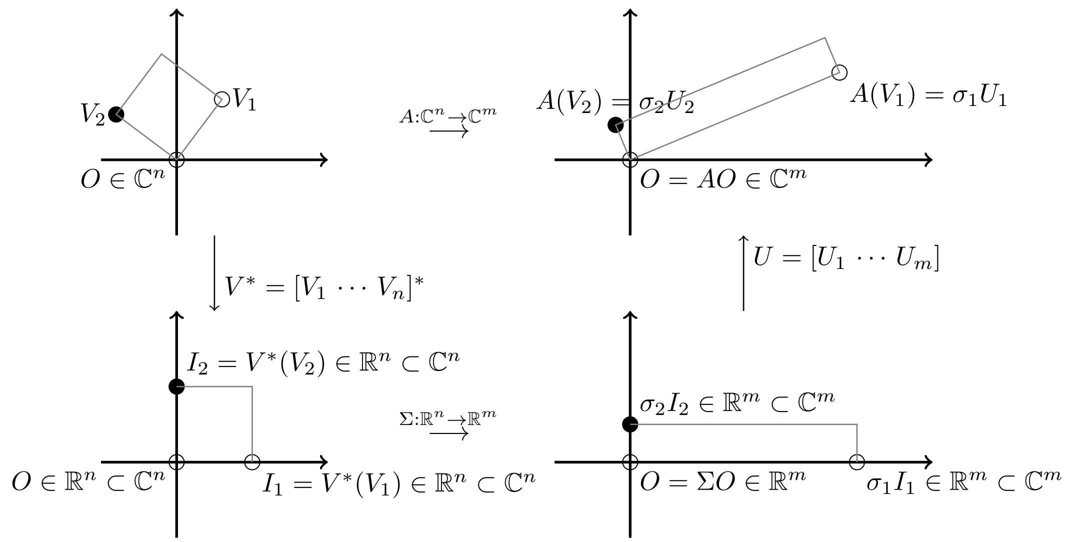


Figure 22: Idea of SVD, or Singular Value Decomposition $A = U\Sigma V^*$: here matrices V^* and U are "rotations" of the vector spaces, and matrix Σ is a "scaling/projection/embedding".

So let $\sigma_j := \|A(V_j)\| = \sqrt{\lambda_j}$ (conventionally here we have already arranged the order so that $\lambda_j \geq \lambda_{j+1}$ when $1 \leq j < n$). Then find a unitary matrix $U = [U_1 \cdots U_m] \in \mathbb{C}^{m\times m}$ for which

$$A(V_j) = \sigma_j U_j.$$

More precisely: If $\sigma_j > 0$ then $U_j = A(V_j)/\sigma_j$. If $m > j > n$ or if $\sigma_j = 0$ then we have more freedom of choosing U_j. Finally, define $\Sigma \in \mathbb{R}^{m\times n}$ such that $\Sigma_{jk} = 0$ whenever $j \neq k$, and $\Sigma_{jj} := \sigma_j$ whenever $j \leq \min\{m, n\}$, i.e.

$$\Sigma \overset{m<n}{=} \begin{bmatrix} \sigma_1 & 0 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & \cdots & 0 \\ 0 & 0 & \sigma_m & \cdots & 0 \end{bmatrix}, \quad \Sigma \overset{m=n}{=} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n \end{bmatrix}, \quad \Sigma \overset{m>n}{=} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus (U, Σ, V) is an SVD for $A \in \mathbb{C}^{m\times n}$, because clearly

$$\begin{aligned} AV &= U\Sigma, \\ A &= U\Sigma V^*. \end{aligned}$$

Moreover, since here

$$A^* = V\Sigma^* U^*,$$

we notice that (V, Σ^*, U) is an SVD for $A^* \in \mathbb{C}^{n\times m}$.

Example. Let $A = \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{C}^{2\times 1}$. Then

$$A^*A = \begin{bmatrix} \overline{a} & \overline{b} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} |a|^2 + |b|^2 \end{bmatrix} \in \mathbb{C}^{1\times 1},$$

and clearly $A^*A = V\Lambda V^*$ for

$$V = \begin{bmatrix} V_1 \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \lambda_1 \end{bmatrix} = \begin{bmatrix} |a|^2 + |b|^2 \end{bmatrix}.$$

Now

$$\Sigma = \begin{bmatrix} \sigma_1 \\ 0 \end{bmatrix} \in \mathbb{R}^{2\times 1},$$

with the singular value $\sigma_1 = \sqrt{\lambda_1} = \sqrt{|a|^2 + |b|^2}$. As $\sigma_1 U_1 = A(V_1) = \begin{bmatrix} a \\ b \end{bmatrix}$, we have $U_1 = \begin{bmatrix} a/\sigma_1 \\ b/\sigma_1 \end{bmatrix}$. We can choose e.g. $U_2 = \begin{bmatrix} -\overline{b}/\sigma_1 \\ \overline{a}/\sigma_1 \end{bmatrix}$. Thus $A = U\Sigma V^*$ reads now

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a/\sigma_1 & -\overline{b}/\sigma_1 \\ b/\sigma_1 & \overline{a}/\sigma_1 \end{bmatrix} \begin{bmatrix} \sigma_1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix}.$$

Another example. In the example above, $A^* = V\Sigma^* U^*$ means

$$\begin{bmatrix} \overline{a} & \overline{b} \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \end{bmatrix} \begin{bmatrix} \overline{a}/\sigma_1 & \overline{b}/\sigma_1 \\ -b/\sigma_1 & a/\sigma_1 \end{bmatrix}.$$

The SVD can be found for any matrix $A \in \mathbb{C}^{m\times n}$: here $A = U\Sigma V^*$. $\Sigma \in \mathbb{R}^{m\times n}$ is unique, but there is some freedom in choosing U, V. $\lambda_j = \sigma_j^2$ are the common eigenvalues of A^*A and AA^*. $U_j \in \mathbb{C}^{m\times 1}$ are eigenvectors of symmetric $AA^* \in \mathbb{C}^{m\times m}$, and

$$A^* = V\Sigma^* U^*$$

(of course, this is the SVD for A^*). In case of $\sigma_j = 0$, it does not matter how vectors $U_j \in \mathbb{C}^{m\times 1}$ and $V_j \in \mathbb{C}^{n\times 1}$ are chosen.

Remark. Above, $A = U\Sigma V^* \in \mathbb{C}^{m\times n}$ and $A^* = V\Sigma^* U^* \in \mathbb{C}^{n\times m}$. Is there some essential difference in finding these SVDs? Well, we first diagonalize either $A^*A \in \mathbb{C}^{n\times n}$ or $AA^* \in \mathbb{C}^{m\times m}$ ... which of the dimensions m, n is smaller... ?

Remark! Define $\tilde{U}, \tilde{\Sigma}, \tilde{V}$ by putting 0 to the columns $k+1, k+2, k+3, \cdots$ of the corresponding SVD-matrices U, Σ, V. Then $\tilde{A} = \tilde{U}\tilde{\Sigma}\tilde{V}^*$ is the best kth rank approximation to $A \in \mathbb{C}^{m\times n}$.

Example. Grey-scale $m \times n$-pixel image $A \in \mathbb{R}^{m\times n}$: $A_{jk} \in [0,1]$ is intensity at (j,k): 0 is black, 0.5 middle grey, 1 white. Storing A takes mn numbers A_{jk}, but storing $\tilde{A}$ takes only

$$m\,k + k + n\,k = (m+n+1)\,k$$

numbers, where often $k \ll m$, $k \ll n$. In this fashion, the SVD can also be used to remove noise from photographs: noise is mostly loaded to relatively small singular values.

Pseudo-inverse. For $A \in \mathbb{C}^{m\times n}$, equation $A(x) = b$ may have no solutions at all. However, $A(x) = b$ has the "best SVD-solution"

$$\tilde{x} := A^+(b),$$

where $A^+ = V\Sigma^+U^* \in \mathbb{C}^{n\times m}$ is the *pseudo-inverse* of $A = U\Sigma V^*$; here $\Sigma^+ \in \mathbb{R}^{n\times m}$ is diagonal matrix, where the non-zero diagonal elements are $1/\sigma_j$ for singular values $\sigma_j > 0$. This *least squares solution* $\tilde{x} \in \mathbb{C}^n$ is best solution in sense that

$$\|A(\tilde{x}) - b\| \leq \|A(x) - b\|$$

for all $x \in \mathbb{C}^n$.

Example. From

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}$$

we see that $U\Sigma V^* = A$ has the pseudo-inverse

$$A^+ = V\Sigma^+U^* = \begin{bmatrix} 0 & 0 \\ 1/3 & 0 \end{bmatrix},$$

where

$$V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \; \Sigma^+ = \begin{bmatrix} 1/3 & 0 \\ 0 & 0 \end{bmatrix}, \; U^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then

$$A^+A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad A\,A^+ = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Notice that A is not surjective:

$$Ax = \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_2 \\ 0 \end{bmatrix}.$$

For $b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \in \mathbb{C}^{2\times 1}$, the "least squares solution" to $Ax = b$ is

$$\widetilde{x} = A^+(b) = \begin{bmatrix} 0 & 0 \\ 1/3 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0 \\ b_1/3 \end{bmatrix}.$$

Polar decomposition. The *polar decomposition* of $z \in \mathbb{C}$ is $z = e^{i\theta}|z|$, where $\theta = \arg(z) \in \mathbb{R}$ is the argument and $|z| \geq 0$ is the absolute value of z. A *polar decomposition* of $A \in \mathbb{C}^{n\times n}$ is

$$A = E|A|,$$

where matrices $E, |A| \in \mathbb{C}^{n\times n}$ are obtained from the SVD

$$A = U\Sigma V^* = (UV^*)(V\Sigma V^*):$$

$$E := UV^*, \quad |A| := V\Sigma V^*. \tag{97}$$

So $E \in \mathbb{C}^{n\times n}$ is unitary, and $|A| = (A^*A)^{1/2} \in \mathbb{C}^{n\times n}$ is positive:

$$\langle |A|u, u\rangle \geq 0$$

for all $u \in \mathbb{C}^n$.

Example. From

$$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} = \begin{bmatrix}0 & 3\\ 0 & 0\end{bmatrix}$$

we see that $U\Sigma V^* = A := \begin{bmatrix}0 & 3\\ 0 & 0\end{bmatrix}$ has the polar decomposition

$$A = E|U| = (UV^*)(V\Sigma V^*),$$

where

$$E = UV^* = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix},$$

$$|A| = V\Sigma V^* = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & 3\end{bmatrix}.$$

Notice that

$$|A| = (A^*A)^{1/2} = \left(\begin{bmatrix}0 & 0\\ 3 & 0\end{bmatrix}\begin{bmatrix}0 & 3\\ 0 & 0\end{bmatrix}\right)^{1/2} = \begin{bmatrix}0 & 0\\ 0 & 9\end{bmatrix}^{1/2}.$$

Matrix norm and singular values. The norm of a matrix $A \in \mathbb{C}^{m\times n}$ is

$$\|A\| = \max_{u\in\mathbb{C}^n : \|u\|\leq 1} \|Au\|.$$

Here $\|A\| = \sigma_1$, the largest singular value of A: this follows from

$$\|Au\|^2 = \langle Au, Au\rangle = \langle A^*Au, u\rangle = \langle u, AA^*u\rangle,$$

because by the unitary diagonalization it is clear that

$$\|A^*A\| = {\sigma_1}^2 = \|AA^*\|.$$

Some other applications: SVD is used in the Google search algorithm [19], and SVD finds regular features in statistics (in tables of numbers; this is the PCA, Principal Component Analysis).

12.2 SVD for compact operators in Hilbert spaces

Now we present an infinite-dimensional version of the SVD:

SVD Theorem. *Let $A \in \mathscr{K}(H, G)$, where G, H are infinite-dimensional Hilbert spaces. Then there are* singular values $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots \geq 0$ *and orthonormal collections $(u_k)_{k=1}^{\infty}$ in G and $(v_k)_{k=1}^{\infty}$ in H so that for all $v \in H$*

$$Av = \sum_{k=1}^{\infty} \sigma_k \langle v, v_k \rangle u_k. \tag{98}$$

Remark. This is an infinite-dimensional analogue of SVD

$$A = U \Sigma V^*$$

of matrices $A \in \mathbb{C}^{m\times n}$, where $\Sigma \in \mathbb{R}^{m\times n}$ is the diagonal matrix of the singular values $\sigma_k \geq 0$, and the columns of the unitary matrices $U \in \mathbb{C}^{m\times m}, V \in \mathbb{C}^{n\times n}$ are the singular vectors $u_k \in \mathbb{C}^{m\times 1}, v_k \in \mathbb{C}^{n\times 1}$ with $Av_k = \sigma_k u_k$.

Proof. Diagonalizing compact self-adjoint operator $A^*A \in \mathscr{K}(H)$, we have

$$A^*Av = \sum_{k=1}^{\infty} \lambda_k \langle v, v_k \rangle v_k, \tag{99}$$

where the orthonormal eigenvectors $v_k \in H$ correspond to the eigenvalues

$$\lambda_k = \langle A^*Av_k, v_k \rangle = \|Av_k\|^2 \geq 0.$$

Thus we may define $\sigma_k := \sqrt{\lambda_k}$. For $\sigma_k > 0$ define $u_k := Av_k/\sigma_k \in G$. Let δ_{jk} be the Kronecker delta: $\delta_{kk} = 1$ and $\delta_{jk} = 0$ if $j \neq k$. If $\sigma_j \sigma_k > 0$ then

$$\langle u_j, u_k \rangle = \langle Av_j/\sigma_j, Av_k/\sigma_k \rangle = \langle A^*Av_j, v_k \rangle/(\sigma_j\sigma_k) = \lambda_j\, \delta_{jk}/(\sigma_j\sigma_k) = \delta_{jk}.$$

When $\sigma_k = 0$, choosing $u_k \in G$ is rather flexible: for such vectors u_k, take any sequence of orthonormal vectors in $\operatorname{ran}(A)^{\perp} \subset G$. Any $v \in H$ is of the form

$$v = \sum_{k=1}^{\infty} \langle v, v_k \rangle v_k + w,$$

where $w \in (\{v_k\}_{k=1}^{\infty})^{\perp}$. Here $Aw = 0$, because

$$\|Aw\|^2 = \langle Aw, Aw \rangle = \langle A^*Aw, w \rangle \overset{(99)}{=} \langle 0, w \rangle = 0.$$

Now it is easy to verify that (98) holds. $\square$

We already had polar decomposition for any bounded operators in (91). However, let us still consider the case of compact operators, starting from the SVD:

Corollary (Polar decomposition for compact operators). *Let $H = G$ above in* (98)*. Then*

$$A = E|A|,$$

where $|A| := (A^*A)^{1/2}$ satisfies

$$|A|v = \sum_{k=1}^{\infty} \sigma_k \langle v, v_k \rangle v_k, \tag{100}$$

and $E \in \mathscr{B}(H)$ is a partial isometry: this means

$$\|Ew\| = \|w\|$$

whenever $w \in \ker(E)^\perp = \ker(A)^\perp$.

Proof. We obtain (100) from the proof of the SVD, as

$$A^*Av = \sum_{k=1}^{\infty} \sigma_k^2 \langle v, v_k \rangle v_k.$$

We have $M := \ker(A) = \ker(|A|)$, since

$$\||A|v\|^2 = \langle |A|v, |A|v \rangle \overset{|A|^*=|A|}{=} \langle |A|^2 v, v \rangle = \langle A^*Av, v \rangle = \langle Av, Av \rangle = \|Av\|^2.$$

Moreover, $M \perp \operatorname{ran}(|A|)$, because

$$\operatorname{ran}(|A|)^\perp \overset{(77)}{=} \ker(|A|^*) \overset{|A|^*=|A|}{=} \ker(|A|) = M.$$

Thus $\operatorname{ran}(|A|)$ is a dense subspace of $M^\perp$. Define $E : M + \operatorname{ran}(|A|) \to H$ by

$$E(z + |A|v) := Av,$$

where $(z, v) \in M \times H$, so that $\|Ew\| = \|w\|$ for all $w \in \operatorname{ran}(|A|)$. Finally, operator E extends by continuity to $H = M \oplus M^\perp$. $\square$

Remark: In the proof above, $Ev_k = Av_k/\sigma_k$ if $\sigma_k > 0$.

Exercise. Let $H = \ell^2(\mathbb{Z}^+)$, and let $A \in \mathscr{K}(H)$ such that

$$(Au)_k := \frac{u_{k+1}}{k+1}.$$

Find polar decompositions of A and A^*.

12.3 Schatten classes $\mathscr{B}_p \subset \mathscr{K}(H)$ (operator kindred of ℓ^p)

In the sequel, let H be an infinite-dimensional separable Hilbert space.

Definition. Let σ_k be the singular values of $A \in \mathscr{K}(H)$. Then A is in *Schatten class* $\mathscr{B}_p$ (for $1 \leq p < \infty$) if $\|A\|_p < \infty$, where

$$\|A\|_p := \left(\sum_{k=1}^{\infty} \sigma_k^p \right)^{1/p}. \tag{101}$$

We also define $\mathscr{B}_\infty := \mathscr{K}(H)$, setting

$$\|A\|_\infty := \|A\| = \sigma_1 \geq \sigma_k. \tag{102}$$

We have the inclusions $\mathscr{B}_p \subset \mathscr{B}_q$ if $p < q$, especially

$$\mathscr{B}_1 \subset \mathscr{B}_2 \subset \mathscr{B}_\infty := \mathscr{K}(H) \subset \mathscr{B}(H).$$

In a sense, the operator spaces $\mathscr{B}_p$ behave much like the sequence spaces ℓ^p.

Trace. Independent of orthonormal basis $(e_k)_{k=1}^\infty$ of the separable Hilbert space H, the *trace* functional $\mathrm{tr} : \mathscr{B}_1 \to \mathbb{C}$ is defined by

$$\mathrm{tr}(A) := \sum_{k=1}^{\infty} \langle A e_k, e_k \rangle \in \mathbb{C}. \tag{103}$$

Here $\|A\|_p^p = \mathrm{tr}(|A|^p)$. *Hilbert–Schmidt class* $\mathscr{B}_2$ has inner product

$$\langle A, B \rangle := \mathrm{tr}(B^* A). \tag{104}$$

Schatten duality. If $1/p + 1/q = 1$ for $p, q \in [1, \infty]$ then we have the Hölder-type inequality

$$\|AB\|_1 \leq \|A\|_p \|B\|_q \tag{105}$$

for all $A \in \mathscr{B}_p$ and $B \in \mathscr{B}_q$, yielding $\varphi_A \in (\mathscr{B}_q)'$, where

$$\varphi_A := (B \mapsto \mathrm{tr}(AB)) : \mathscr{B}_q \to \mathbb{C}, \tag{106}$$

the mapping $(A \mapsto \varphi_A)$ giving an isomorphism $\mathscr{B}_p \cong (\mathscr{B}_q)'$, when $1 < q$. Similar duality by the trace functionals holds for the trace class $\mathscr{B}_1$, but then

$$(\mathscr{B}_1)' \cong \mathscr{B}(H) \neq \mathscr{B}_\infty = \mathscr{K}(H).$$

This is also related to quantum mechanics: separable H describes a quantum system (quantum states are normalized vectors $u \in H$), $A^* = A$ represents a physical observable (e.g. position, momentum, angular momentum, spin, energy), its eigenvectors represent pure states of the physical system, with eigenvalues corresponding to the physical observable quantities, and the density matrix $\rho \in \mathscr{B}_1$ gives the state probabilities (where $\rho \geq 0$ and $\mathrm{tr}(\rho) = 1$).

12.4 Further generalizations

Previous decompositions represented compact operators as series, based on diagonalization of compact self-adjoint operators. Analogous statements hold for bounded operators on Hilbert space, but the spectral series have to be replaced by spectral integrals (operator-valued integrals with respect to spectral measures). Also unbounded self-adjoint operators can be spectrally decomposed, but there are technicalities here (think e.g. of Laplacian $A = \Delta$ on $H = L^2(M)$, which is defined only in a dense subset of smooth-enough functions $u : M \to \mathbb{C}$). And above, when we mentioned "self-adjoint" we could have said "normal" with the analogous more general results; here $\sigma(A) \subset \mathbb{R}$ when $A^* = A$, but $\sigma(B) \subset \mathbb{C}$ when $B^*B = BB^*$.

Informal example. Let M be a closed Riemannian manifold (compact, without boundary). Then $H = L^2(M)$ is a separable infinite-dimensional Hilbert space having an orthonormal basis of eigenvectors $e_k \in C^\infty(M) \subset H$ of the Laplace–Beltrami operator $\Delta : C^\infty(M) \to C^\infty(M)$, where $\Delta e_k = \lambda_k\, e_k$ such that $0 \geq \lambda_k$. Think that e_k describes a "standing wave" on M, where $\sqrt{-\lambda_k}$ is the corresponding frequency for the vibration of M. Here Δ is not bounded on H, but $(I - \Delta)^{-1}$ is compact and self-adjoint there. The Sobolev space $H^{2j}(M)$ is the completion of $C^\infty(M)$ with respect to the norm given by

$$\|u\|_{H^{2j}} := \left\|(I - \Delta)^j u\right\|_{L^2}.$$

As a concrete example of "Fourier eigenfunction expansions" on manifolds, think of $M = \mathbb{R}/\mathbb{Z}$, where $(e_k)_{k\in\mathbb{Z}}$ is an orthonormal basis for $L^2(M)$ when $e_k(x) = \mathrm{e}^{\mathrm{i}2\pi x\cdot k}$, and here $\Delta e_k = e_k'' = -(2\pi k)^2 e_k$, so that $\sqrt{-\lambda_k} = 2\pi|k| = \sqrt{-\lambda_{-k}}$. Here

$$\begin{aligned}
\langle u, v\rangle &= \int_0^1 u(x)\, v(x)^*\, \mathrm{d}x \\
\langle u, e_k\rangle &= \int_0^1 u(x)\, \mathrm{e}^{-\mathrm{i}2\pi x\cdot k}\, \mathrm{d}x \quad = \quad \widehat{u}(k), \\
u &= \sum_{k\in\mathbb{Z}} \widehat{u}(k)\, e_k, \\
\Delta u &= \sum_{k\in\mathbb{Z}} -(2\pi k)^2\, \widehat{u}(k)\, e_k, \\
(I - \Delta)^{-1} v &= \sum_{k\in\mathbb{Z}} \frac{1}{1 + (2\pi k)^2}\, \widehat{v}(k)\, e_k.
\end{aligned}$$

13 What to study next?

Due to the lack of time, we had to skip many important features in functional analysis. Let us list some of things worth studying more:

- Applications to PDEs, to harmonic and Fourier analysis.
- Unbounded operators, function spaces and distribution theory.
- Further spectral properties in Banach and Hilbert spaces.
- Operator algebras (Banach, C^*, von Neumann algebras).
- Weak and weak* topologies.
- Reflexive spaces.
- Locally convex spaces.
- Topological vector spaces, in general.
- Multilinear functional analysis.
- Non-linear functional analysis.
- Et cetera...

14 Some selected general references

For functional analysis, see [11, 12, 13, 14, 16, 22].
For measure theory, see [7, 8, 15, 16].
For distribution theory, see [1, 4, 9, 14, 16, 17, 20].

References

[1] P. Antosik, J. Mikusiński and R. Sikorski: *Theory of distributions. The sequential approach.* Warszawa. PWN—Polish Scientific Publishers, 1973.

[2] S. Banach and H. Steinhaus: *Sur le principe de la condensation de singularités.* Fundamenta Mathematicae **9** (1927), 50–61.

[3] R. Baire: *Sur les fonctions de variables réelles.* Ann. di Mat. **3** (1899), 1–123.

[4] J. Barros-Neto: *An introduction to the theory of distributions.* Marcel Dekker, Inc., 1973.

[5] D. Crespin: *Hahn–Banach Theorem implies Riesz Theorem.* Portugaliae Mathematica **51** (1994), 217–218.

[6] P. Enflo: *A counterexample to the approximation problem in Banach spaces.* Acta Mathematica **130** (1973), 309–317.

[7] G. B. Folland: *Real Analysis. Modern techniques and their applications.* Second edition. A Wiley-Interscience Publication. John Wiley and Sons, Inc., 1999.

[8] R. Gariepy and W. Ziemer: *Modern real analysis.*

[9] I. M. Gelfand and G. E. Shilov: *Generalized functions.* Vols. 1–3, Academic Press, 1968.

[10] L. Grafakos: *An elementary proof of the square summability of the discrete Hilbert transform.* The American Mathematical Monthly **101** (1994), 456–458.

[11] V. Hutson and J.S. Pym: *Applications of functional analysis and operator theory.* Academic Press, 1980.

[12] J. L. Kelley, I. Namioka et al.: *Linear topological spaces.* D. Van Nostrand Company, Inc. Princeton, New Jersey, 1963.

[13] E. Kreyszig: *Introductory functional analysis with applications.* John Wiley & Sons 1989.

[14] W. Rudin: *Functional analysis.* Tata McGraw-Hill, 1974.

[15] W. Rudin: *Real and complex analysis.* McGraw-Hill 1987.

[16] M. Ruzhansky and V. Turunen: *Pseudo-differential operators and symmetries.* Birkhäuser 2010.

[17] L. Schwartz: *Théorie des distributions, I, II.* 2nd ed., Hermann, Paris, 1957.

[18] A. Sokal: *A really simple elementary proof of the uniform boundedness theorem.* Amer. Math. Monthly **118** (2011), 450–452.

[19] G. Strang: *Introduction to linear algebra.* Fourth edition. Wellesley–Cambridge Press. Wellesley, 2009.

[20] F. Treves: *Topological vector spaces, distributions and kernels.* Academic Press. New York, 1967.

[21] David C. Ullrich: *A simple elementary proof of Hilbert's inequality.* The American Mathematical Monthly **120** (2), 2013, 161–164.

[22] K. Yosida: *Functional analysis.* Reprint of the sixth (1980) edition. Springer-Verlag, Berlin, 1995.

[23] P. P. Zabreiko: *A theorem for semiadditive functionals.* Functional Analysis and Its Applications **3** (1), 1969, 70–72.

15 Appendix A

Appendix A contains some nice-to-know things about the course in functional analysis. The material in this appendix is not part of the course requirements!

15.1 Topological and metric spaces

Let us quickly recall some topological and metric concepts.

Definition. A collection τ of subsets of a set X is a *topology* if

$$\begin{aligned} &\emptyset \in \tau,\ X \in \tau, \\ &\{U_1, U_2\} \subset \tau \Rightarrow U_1 \cap U_2 \in \tau, \\ &\mathcal{U} \subset \tau \Rightarrow \bigcup \mathcal{U} \in \tau. \end{aligned}$$

A set $U \in \tau$ is called *open*, and then the set $X \setminus U$ is called *closed*.

The *closure* of a subset $S \subset X$ is the smallest closed set $\overline{S} \supset S$; here $S \subset X$ is *dense* if $\overline{S} = X$. The *interior* of $S \subset X$ is $\mathrm{int}(S) := X \setminus \overline{X \setminus S}$ (i.e. the largest open set within S). The *boundary* of $S \subset X$ is $\partial S := \overline{S} \setminus \mathrm{int}(S) = \overline{S} \cap \overline{X \setminus S}$.

(X, τ) is called a *topological space* (abbreviation: X is a topological space, if τ is implicitely known). The *product topology* on $X \times Y$ has those open sets $U \times V$ where $U \subset X, V \subset Y$ are open.

15.1.1 Hausdorff spaces. Compact spaces. Continuity

Definition. Topological space (X, τ) is a *Hausdorff space* if for all $p, q \in X$ for which $p \neq q$ there exist $U_1, U_2 \in \tau$ such that $p \in U_1, q \in U_2$ and $U_1 \cap U_2 = \emptyset$.

Definition. A subset $K \subset X$ is *compact* if its open covers always have a finite subcover: that is, if $K \subset \bigcup \mathscr{U}$ for a family $\mathscr{U} \subset \tau$ of open sets, then there is a finite subset $\mathscr{V} \subset \mathscr{U}$ such that $K \subset \bigcup \mathscr{V}$.

Exercise. Assuming that $K \subset X$ is compact, prove:
(1) If $C \subset X$ is closed, then $C \cap K$ is compact.
(2) If X is Hausdorff, then $K \subset X$ is closed.

Definition. A mapping $f : X \to Y$ is *continuous* for topological spaces (X, τ_X) and (Y, τ_Y) if the preimages of open sets are open: that is, $f^{-1}(W) \in \tau_X$ for all $W \in \tau_Y$.

Exercise. Let $f : X \to Y$ be continuous and $K \subset X$ be compact. Show that $f(K) \subset Y$ is compact.

15.1.2 "Automatic continuity". Closed graph property

Once you learn the major results about Banach spaces, the following two exercises about topological properties shall look rather familiar in spirit:

Exercise. Let $f : X \to Y$ be a continuous bijection, X compact and Y Hausdorff. Show that also $f^{-1} : Y \to X$ is continuous (i.e. f is a homeomorphism, essentially identifying the topological spaces).

Remark. Compact Hausdorff topology is an interesting borderline case, which is an outcome of the previous exercise: If τ_1, τ_2 are topologies on a set X such that $\tau_1 \subset \tau_2$ where τ_2 is compact and τ_1 is Hausdorff, then $\tau_1 = \tau_2$.

Exercise. Let X, Y be compact Hausdorff spaces. Show that $f : X \to Y$ is continuous if and only if its graph $\Gamma(f) := \{(x, f(x)) \in X \times Y \mid x \in X\}$ is closed in $X \times Y$.

15.1.3 Metric topology

Definition. A *metric* on a set X is a function $d : X \times X \to \mathbb{R}$ such that $d(u, u) = 0$, $u \neq v \Rightarrow d(u, v) = d(v, u) > 0$, and $d(u, w) \leq d(u, v) + d(v, w)$ for all $u, v, w \in X$. Then (X, d) (or X for short when d is implicitely known) is called a *metric space*.

The *distance* between sets $S_1, S_2 \subset X$ is

$$\operatorname{dist}(S_1, S_2) := \inf\{d(u, v) : \ u \in S_1, \ v \in S_2\}.$$

The smallest topology on X that contains the *open balls*

$$\mathbb{B}(u, r) := \{v \in X : \ d(u, v) < r\}$$

for all $u \in X$ and $r > 0$ is called the *metric topology* on X.

Exercise. In metric topology, show that the closure of $S \subset X$ is

$$\overline{S} = \{u \in X : \ \operatorname{dist}(\{u\}, S) = 0\}.$$

15.1.4 Sequences in metric spaces

Definition. A sequence $(u_k)_{k=1}^{\infty}$ in X *converges* to $p \in X$ if $d(u_k, p) \to 0$ as $k \to \infty$: then we denote $u_k \to p$.

Definition. A sequence $(u_k)_{k=1}^{\infty}$ in X is a *Cauchy sequence* if for all $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{Z}^+$ such that $d(u_j, u_k) < \varepsilon$ whenever $j, k \geq n_\varepsilon$. Metric space is called *complete* if its every Cauchy sequence converges.

Remark. In a metric space X, compactness of $K \subset X$ is equivalent to the following *sequential compactness*: if $u_j \in K$ for all $j \in \mathbb{Z}^+$ then there exists a subsequence $(u_{j_k})_{k=1}^{\infty}$ converging to a point $p \in K$.

15.1.5 Continuity in metric spaces

Remark. Mapping $f : X \to Y$ between metric spaces (X, d_X) and (Y, d_Y) is continuous if and only if it is *sequentially continuous*, i.e. $f(u_k) \to f(u)$ whenever $u_k \to u$. Equivalently, continuity means

$$\forall u, v \in X \ \forall \varepsilon > 0 \ \exists \delta > 0 : \ d_X(u, v) < \delta \ \Rightarrow \ d_Y(f(u), f(v)) < \varepsilon.$$

Idea of continuity: Continuity of $f : X \to Y$ roughly means that $f(u) \approx f(v)$ if $u \approx v$.

Exercise. Let $C(M)$ be the set of continuous functions $u : M \to \mathbb{C}$ on a compact space M. For $u \in C(M)$, let

$$\|u\| := \sup\{|u(x)| : \ x \in M\}.$$

Show that $C(M)$ is a complete metric space with the metric $(u, v) \mapsto \|u - v\|$.

15.1.6 Heine–Borel Theorem

The Euclidean metric on $\mathbb{R}^n$ is defined by $d(u, v) := \|u - v\|$, where

$$\|u\| := \left(\sum_{k=1}^{n} u_k^2 \right)^{1/2}$$

is the Euclidean norm of $u = (u_1, \cdots, u_n) \in \mathbb{R}^n$. A set $K \subset \mathbb{R}^n$ is *bounded* if there exists a constant $c < \infty$ such that $\|u\| \le c$ for all $u \in \mathbb{R}^n$.

Theorem (Heine–Borel). *A set $K \subset \mathbb{R}^n$ is compact in the Euclidean topology if and only if it is closed and bounded.*

Remark. In Heine–Borel Theorem, it was essential that the dimension n was finite. On the other hand, Riesz' Compactness Theorem shows that the closed balls are never compact in infinite-dimensional normed spaces.

15.1.7 Banach Fixed Point Theorem

Banach contraction has a unique fixed point:

Theorem. *Let $X \neq \emptyset$ be a complete metric space. Let $f : X \to X$ such that $d(f(u), f(v)) \le \lambda\, d(u, v)$ for all $u, v \in X$, where $\lambda < 1$. Then there exists unique $p \in X$ such that $f(p) = p$.*

Proof. Take any $u = u_0 \in X$. Let $u_{k+1} := f(u_k)$. Then $(u_k)_{k=1}^{\infty}$ is a Cauchy sequence, because

$$\begin{aligned} d(u_k, u_{k+m}) &\leq \sum_{j=0}^{m-1} d(u_{k+j}, u_{k+j+1}) \\ &\leq \sum_{j=0}^{m-1} \lambda^{k+j} d(u_0, u_1) \\ &\leq \frac{\lambda^k}{1-\lambda} d(u_0, u_1) \\ &\overset{k\to\infty}{\longrightarrow} 0. \end{aligned}$$

By completeness, $u_k \to p_u \in X$. By continuity, we have $f(p_u) = f(\lim_k u_k) = \lim_k f(u_k) = \lim_k u_{k+1} = p_u$. If $u, v \in X$ then

$$d(p_u, p_v) = d(f(p_u), f(p_v)) \leq \lambda\, d(p_u, p_v).$$

Here necessarily $p_u = p_v$, as $\lambda < 1$. □

Example. Let us apply the Fixed Point Theorem to differential equations: Let us study initial value problem

$$\begin{cases} y'(t) &= b(t, y(t)), \\ y(0) &= y_0, \end{cases}$$

where b nice-enough when $t \approx 0$. Then

$$y(t) = y_0 + \int_0^t b(s, y(s))\, \mathrm{d}s.$$

In the Picard–Lindelöf iteration, find approximate solutions u_k to the initial value problem as follows: let $u_0(t) \equiv y_0$, and let

$$u_{k+1}(t) := y_0 + \int_0^t b(s, u_k(s))\, \mathrm{d}s.$$

If $|b(s, r_1) - b(s, r_2)| \leq c\, |r_1 - r_2|$ whenever $|s| \leq |t|$ then

$$\left| \int_0^t [b(s, u(s)) - b(s, v(s))]\ \mathrm{d}s \right| \leq |t|\, c \sup_{|s|<|t|} |u(s) - v(s)|$$

so we may apply the Fixed Point Theorem for small-enough $|t|$.

16 Index of Banach space results (valid in Hilbert spaces, too)

Quotient norm construction
Banach space of continuous functions on compact space
Banach spaces L^p, ℓ^p
Linear boundedness is continuity
Banach spaces of linear operators
Banach completion
Baire's Theorem
Uniform Boundedness Principle (Banach–Steinhaus Theorem)
Strong convergence (defining bounded operators)
Open Mapping Theorem (Banach–Schauder Theorem)
Closed Graph Theorem
Hahn–Banach Theorem(s), functional separation
Banach–Alaoglu Theorem
Closure Lemma (for compact operators)
Almost Orthogonality Lemma [F. Riesz]
Riesz' Compactness Theorem
Fredholm Index Theorem
Characterisation of Banach algebras
Continuity of inversion
Gelfand's Spectrum Theorem
Gelfand–Mazur Theorem
Spectral Radius Formula

17 Index of Hilbert space results