ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Chair for Mathematical Information Science

Verner Vlačić
Sternwartstrasse 7
CH-8092 Zürich

Mathematics of Information
Hilbert spaces and linear operators

These notes are based on [1, Chap. 6, Chap. 8], [2, Chap. 1], [3, Chap. 4], [4, App. A] and [5].

1 Vector spaces

Let us consider a field $\mathbb{F}$, which can be, for example, the field of real numbers $\mathbb{R}$ or the field of complex numbers $\mathbb{C}$. A *vector space* over $\mathbb{F}$ is a set whose elements are called *vectors* and in which two operations, addition and multiplication by any of the elements of the field $\mathbb{F}$ (referred to as *scalars*), are defined with some algebraic properties. More precisely, we have

Definition 1 (Vector space)**.** A set $\mathcal{X}$ together with two operations $(+, \cdot)$ is a *vector space* over $\mathbb{F}$ if the following properties are satisfied:

(i) the first operation, called vector addition: $\mathcal{X} \times \mathcal{X} \to \mathcal{X}$ denoted by $+$ satisfies

- $(x + y) + z = x + (y + z)$ for all $x, y, z \in \mathcal{X}$ (associativity of addition)
- $x + y = y + x$ for all $x, y \in \mathcal{X}$ (commutativity of addition)

(ii) there exists an element $0 \in \mathcal{X}$, called the *zero vector*, such that $x + 0 = x$ for all $x \in \mathcal{X}$

(iii) for every $x \in \mathcal{X}$, there exists an element in $\mathcal{X}$, denoted by $-x$, such that $x + (-x) = 0$.

(iv) the second operation, called scalar multiplication: $\mathbb{F} \times \mathcal{X} \to \mathcal{X}$ denoted by $\cdot$ satisfies

- $1 \cdot x = x$ for all $x \in \mathcal{X}$
- $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$ for all $\alpha, \beta \in \mathbb{F}$ and $x \in \mathcal{X}$ (associativity for scalar multiplication)
- $(\alpha+\beta)\cdot x = \alpha\cdot x+\beta\cdot x$ for all $\alpha, \beta \in \mathbb{F}$ and $x \in \mathcal{X}$ (distributivity of scalar multiplication with respect to field addition)
- $\alpha\cdot(x+y) = \alpha\cdot x+\alpha\cdot y$ for all $\alpha \in \mathbb{F}$ and $x, y \in \mathcal{X}$ (distributivity of scalar multiplication with respect to vector addition).

We refer to $\mathcal{X}$ as a real vector space when $\mathbb{F} = \mathbb{R}$ and as a complex vector space when $\mathbb{F} = \mathbb{C}$.

Examples.

1. $\mathbb{C}$ is both a real and a complex vector space.
2. The set $\mathbb{F}^N \triangleq \{(x_1, x_2, \ldots, x_N) \colon x_k \in \mathbb{F}\}$ of all N-tuples forms a vector space over $\mathbb{F}$.
3. The set $\mathbb{F}[x]$ of all polynomials with coefficients in $\mathbb{F}$ is a vector space over $\mathbb{F}$.
4. The space $\mathbb{F}^{\mathbb{Z}}$ of all sequences of $\mathbb{F}$ is a vector space over $\mathbb{F}$.

5. The set $\mathbb{F}^{M \times N}$ of all matrices of size $M \times N$ with entries in $\mathbb{F}$ forms a vector space over $\mathbb{F}$ under the laws of matrix addition and scalar multiplication.
6. If $\mathcal{X}$ is an arbitrary set and $\mathcal{Y}$ an arbitrary vector space over $\mathbb{F}$, the set $\mathcal{F}(\mathcal{X}, \mathcal{Y})$ of all functions $\mathcal{X} \to \mathcal{Y}$ is a vector space over $\mathbb{F}$ under pointwise addition and multiplication.

A subspace of a vector space $\mathcal{X}$ over $\mathbb{F}$ is a subset of $\mathcal{X}$ which is itself a vector space over $\mathbb{F}$ (with respect to the same operations). One can easily verify that a subset $\mathcal{Y}$ of $\mathcal{X}$ is a subspace using the following proposition.

Proposition 1.1. Let $\mathcal{X}$ be a vector space over $\mathbb{F}$ and $\mathcal{Y}$ a nonempty subset of $\mathcal{X}$. Then $\mathcal{Y}$ is a subspace of $\mathcal{X}$ if it contains 0 and if it is stable under linear combinations, that is,

$$\alpha \cdot y + \beta \cdot z \in \mathcal{Y}$$

for all $\alpha, \beta \in \mathbb{F}$ and $y, z \in \mathcal{Y}$.

Examples. 1. $\mathbb{R}$ and $i\mathbb{R}$ are subspaces of the real vector space $\mathbb{C}$.

2. The set of N-tuples $(x_1, x_2, \ldots, x_{N-1}, 0)$ with $x_k \in \mathbb{R}$ is a subspace of $\mathbb{R}^N$.
3. The set $\ell^p(\mathbb{Z})$, $p \in [1, \infty]$, of all complex-valued sequence $\{u_k\}_{k \in \mathbb{Z}}$ which satisfy

$$\begin{array}{ll} \sum_{k=-\infty}^{+\infty} |u_k|^p < \infty, & \text{if } p \in [1, \infty), \\ \sup_{k \in \mathbb{Z}} |u_k| < \infty, & \text{if } p = \infty, \end{array}$$

   is a subspace of $\mathbb{C}^{\mathbb{Z}}$.
4. If $\mathcal{X}$ is an arbitrary set, $\mathcal{Y}$ a vector space over $\mathbb{F}$, and $\mathcal{Z}$ a subspace of $\mathcal{Y}$, then the set $\mathcal{F}(\mathcal{X}, \mathcal{Z})$ of functions $\mathcal{X} \to \mathcal{Z}$ is a subspace of $\mathcal{F}(\mathcal{X}, \mathcal{Y})$.
5. The space $C^n[a, b]$ of all complex-valued functions with continuous derivatives of order $0 \leqslant k \leqslant n$ on the closed, bounded interval $[a, b]$ of the real line is a subspace of $\mathcal{F}([a, b], \mathbb{C})$.
6. If $(\mathcal{S}, \Sigma, \mu)$ is a measure space, we let $L^p(\mathcal{S}, \Sigma, \mu)$ denote the space of all measurable functions mapping $\mathcal{S}$ to $\mathbb{C}$ whose absolute value raised to the p-power is μ-integrable, that is,

$$L^p(\mathcal{S}, \Sigma, \mu) = \left\{ f \colon \mathcal{S} \to \mathbb{C} \text{ measurable} \colon \int_{\mathcal{S}} |f|^p \mathrm{d}\mu < \infty \right\},$$

   two elements of $L^p(\mathcal{S}, \Sigma, \mu)$ being considered as equivalent if they differ only on a set whose measure is zero. $L^p(\mathcal{S}, \Sigma, \mu)$ is a subspace of $\mathcal{F}(\mathcal{S}, \mathbb{C})$. For simplicity, we often use the notation $L^p(\mathcal{S})$ when $\mathcal{S}$ is a subset of $\mathbb{R}^N$, Σ is the Borel σ-algebra over $\mathcal{S}$, and μ the Lebesgue measure on $\mathcal{S}$. In this case, two functions are equivalent when they are equal except possibly on a set of Lebesgue measure zero (e.g., a finite or countable[1] set of points of $\mathbb{R}^N$).

Theorem 1 (Intersection of vector spaces)**.** The intersection of any collection of subspaces of a vector space $\mathcal{X}$ over $\mathbb{F}$ is again a subspace of $\mathcal{X}$.

Definition 2. If $\mathcal{S}$ and $\mathcal{T}$ are linear subspaces of a vector space $\mathcal{X}$ with $\mathcal{S} \cap \mathcal{T} = \{0\}$, then we define the *direct sum* $\mathcal{S} \oplus \mathcal{T}$ by

$$\mathcal{S} \oplus \mathcal{T} = \{x + y \mid x \in \mathcal{S}, y \in \mathcal{T}\}.$$

Note that the union of two subspaces is in general *not* a subspace. Take for example $\mathcal{X} = \mathbb{R}^2$. The lines $\mathbf{x}\mathbb{R}$ and $\mathbf{y}\mathbb{R}$, with $\mathbf{x} = (1, 0)$ and $\mathbf{y} = (0, 1)$, are both subspaces of $\mathbb{R}^2$, but $\mathbf{x}\mathbb{R} \cup \mathbf{y}\mathbb{R}$ is a not subspace, given that $\mathbf{x} + \mathbf{y} = (1, 1) \notin \mathbf{x}\mathbb{R} \cup \mathbf{y}\mathbb{R}$.

Definition 3 (Subspace spanned by $\mathcal{S}$)**.** Let $\mathcal{S}$ be a (possibly infinite) subset of a vector space $\mathcal{X}$ over $\mathbb{F}$. The subspace spanned by $\mathcal{S}$ is the intersection of all subspaces containing $\mathcal{S}$. It is the smallest subspace containing $\mathcal{S}$. It is denoted by span($\mathcal{S}$) and may be written the set of all finite combinations of elements of $\mathcal{S}$, that is,

$$\operatorname{span}(\mathcal{S}) \triangleq \left\{ \sum_{\ell=1}^{k} \lambda_\ell x_\ell \colon k \in \mathbb{N}, x_\ell \in \mathcal{S}, \lambda_\ell \in \mathbb{F} \right\}.$$

Examples.

1. If $\mathcal{S}$ is empty, then the subspace spanned by $\mathcal{S}$ is $\{0\}$.
2. In the real vector space $\mathbb{C}$, we have the following:
   - the subspace spanned by $\{1\}$ is $\mathbb{R}$,
   - the subspace spanned by $\{i\}$ is $i\mathbb{R}$,
   - the subspace spanned by $\{1, i\}$ is $\mathbb{C}$.
3. In $\mathbb{F}[x]$, the subspace spanned by $\{1, x, \ldots, x^N\}$ is the space $\mathbb{F}_N[x]$ of all polynomials whose degree is less than N.

Definition 4 (Linear independence)**.** Let $\mathcal{X}$ be a vector space over $\mathbb{F}$. A finite set $\{x_1, x_2, \ldots, x_N\}$ of vectors of $\mathcal{X}$ is said to be linearly independent if for every $\lambda_1, \lambda_2, \ldots, \lambda_N \in \mathbb{F}$, the equality

$$\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_N x_N = 0$$

implies that $\lambda_1 = \lambda_2 = \ldots = \lambda_N = 0$. An infinite set $\mathcal{S}$ of vectors of $\mathcal{X}$ is linearly independent[2] if every *finite* subset of $\mathcal{S}$ is linearly independent.

Examples.

1. Any set of vectors containing the zero vector is linearly dependent.
2. The real vector space $\mathbb{C}$, the set $\{1, i\}$ is linearly independent.
3. The basic monomials $\{1, x, x^2, \ldots, x^N\}$ form a linearly independent set of $\mathbb{F}[x]$.
4. The set of trigonometric functions $\{1, \cos(x), \sin(x), \cos^2(x), \cos(x)\sin(x), \sin^2(x)\}$ is linearly dependent.

[1] Note, however, that there exist uncountable sets (e.g., the Cantor set) of Lebesgue measure zero.

Definition 5 (Dimension)**.** A vector space $\mathcal{X}$ is N-dimensional if there exists N linearly independent vectors in $\mathcal{X}$ and any $N+1$ vectors in $\mathcal{X}$ are linearly dependent.

Definition 6 (Finite-dimensional space)**.** A vector space $\mathcal{X}$ is finite-dimensional if $\mathcal{X}$ is N-dimensional for some integer N. Otherwise, $\mathcal{X}$ is infinite dimensional.

> **Examples.**
> 1. The spaces $\mathbb{F}^N$ and $\mathbb{F}_N[x]$ are finite-dimensional (their dimension is N).
> 2. The space $\mathcal{F}(\mathcal{X},\mathcal{X})$, $\mathbb{F}[x]$, $C^n[a,b]$, $\ell^p(\mathbb{Z})$ are infinite-dimensional.

Definition 7 (Basis)**.** A basis[3] of a vector space $\mathcal{X}$ over $\mathbb{F}$ is a set of linearly independent and spanning vectors of $\mathcal{X}$.

> **Examples.**
> 1. The set $\{\mathbf{e_k}\}_{k=1}^3$, where $\mathbf{e}_1 = [1\ 0\ 0]^T$, $\mathbf{e}_2 = [0\ 1\ 0]^T$, and $\mathbf{e}_3 = [0\ 0\ 1]^T$, forms a basis for $\mathbb{R}^3$.
> 2. The set $\{1, x, x^2, \ldots, x^N\}$ forms a basis for $\mathbb{F}_N[x]$.
> 3. The set $\{1, x, x^2, x^3, \ldots\}$ forms a basis for $\mathbb{F}[x]$.

Theorem 2. Let $\mathcal{X}$ be a finite dimensional vector space. Any set of linearly independent vectors can be extended to a basis of $\mathcal{X}$.

2 Inner products and norms

From here, we will assume that $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.

Definition 8 (Norm)**.** Let $\mathcal{X}$ be a vector space over $\mathbb{F}$. A norm on $\mathcal{X}$ is a function which maps $\mathcal{X}$ to $\mathbb{R}$ and satisfies the following properties:

(i) for all $x \in \mathcal{X}$, we have $\|x\| \geqslant 0$ and $\|x\| = 0$ implies $x = 0$ (positivity)

(ii) for all $x \in \mathcal{X}$ and for all $\alpha \in \mathbb{F}$, we have $\|\alpha x\| = |\alpha|\,\|x\|$ (homogeneity)

(iii) for all $x, y \in \mathcal{X}$, we have $\|x+y\| \leqslant \|x\| + \|y\|$ (triangle inequality).

Definition 9. A vector space $\mathcal{X}$ over $\mathbb{F}$ is a *normed vector space* (or a *pre-Banach space*) if it is equipped with a norm $\|\cdot\|$.

[3] This type of basis is sometimes referred to as an *algebraic basis* or a *Hamel basis* to highlight the difference with Schauder bases in infinite-dimensional Banach spaces, and orthonormal bases in infinite-dimensional Hilbert space. In finite dimensions, all these concepts coincide (cf. notes on bases in infinite-dimensional spaces for more details).

Examples. 1. We can provide the space $C^0[a,b]$ with the norm

$$\|f\| = \int_a^b |f(t)| \mathrm{d}t.$$

2. We can define a norm on $\mathbb{F}^N$ as follows:

$$\|\mathbf{x}\|_2 = \left(\sum_{k=1}^{N} |x_k|^2 \right)^{1/2}.$$

3. If we write $u = \{u_k\}_{k\in\mathbb{Z}}$, then the following defines a norm on $\ell^p(\mathbb{Z})$, $p \in [1, \infty]$,

$$\begin{aligned} \|u\|_p &= \left(\textstyle\sum_{k=-\infty}^{+\infty} |u_k|^p\right)^{1/p}, \quad && p \in [1,\infty), \\ \|u\|_\infty &= \sup_{k\in\mathbb{Z}} |u_k|, && p = \infty. \end{aligned}$$

4. We can define the following norm on $L^p(\mathcal{S}, \Sigma, \mu)$, $p \in [1, \infty)$:

$$\|f\|_p = \left(\int_{\mathcal{S}} |f|^p \mathrm{d}\mu \right)^{1/p}.$$

Definition 10 (Inner product)**.** Let $\mathcal{X}$ be a vector space over $\mathbb{F}$. An *inner product* (or *scalar product*) over $\mathbb{F}$ is a function $\mathcal{X} \times \mathcal{X} \to \mathbb{F}$ such that, for all $x, y \in \mathcal{X}$ and $\alpha \in \mathbb{F}$, the following holds

(i) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ (symmetry)

(ii) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ (sesquilinearity if $\mathbb{F} = \mathbb{C}$, bilinearity if $\mathbb{F} = \mathbb{R}$)

(iii) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ (sesquilinearity if $\mathbb{F} = \mathbb{C}$, bilinearity if $\mathbb{F} = \mathbb{R}$)

(iv) $\langle x, x \rangle \in \mathbb{R}$ and $\langle x, x \rangle \geqslant 0$ (positivity)

(v) $\langle x, x \rangle = 0$ implies that $x = 0$ (positivity).

The inner product on a vector space $\mathcal{X}$ induces a norm on $\mathcal{X}$. But note that not all norms come from an inner product (see Problem 3 of Homework 1).

Theorem 3. Let $\mathcal{X}$ be a vector space. If $\langle \cdot, \cdot \rangle$ is an inner product on $\mathcal{X}$, then, $\|x\| = \sqrt{\langle x, x \rangle}$ is a norm on $\mathcal{X}$ fulfilling the parallelogram law:

$$2\|x\|^2 + 2\|y\|^2 = \|x + y\|^2 + \|x - y\|^2, \quad x, y \in \mathcal{X}.$$

On the other hand, every norm $\|x\|$ on $\mathcal{X}$ fulfilling the parallelogram law is the induced norm of exactly one inner product, which is defined as

$$\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2), \quad \mathbb{F} = \mathbb{R} \tag{1}$$

$$\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2), \quad \mathbb{F} = \mathbb{C}. \tag{2}$$

Formulas (1) and (2) are called polarization formulas.

Definition 11 (Inner product space). A vector space $\mathcal{X}$ over $\mathbb{F}$ is an *inner product space* or a *pre-Hilbert space* if it is equipped with an inner product $\langle \cdot, \cdot \rangle$.

Theorem 4 (Cauchy-Schwarz inequality). Let $\mathcal{X}$ be an inner product space. Then, we have

$$|\langle x, y \rangle| \leqslant \|x\| \, \|y\| \,, \quad x, y \in \mathcal{X}$$

with equality if and only if x and y are linearly dependent.

In what follows, we will always assume that for an inner product space, $\|\cdot\|$ is the norm induced by the corresponding inner product.

Examples. 1. The space $\mathbb{C}^N$ can be equipped with the inner product

$$\langle \mathbf{x}, \mathbf{y} \rangle \triangleq \sum_{k=1}^{N} x_k \overline{y_k},$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_N)$ are in $\mathbb{C}^N$. The norm induced is the one given in the previous set of examples.

2. The space $\ell^2(\mathbb{Z})$ of square summable sequences is an inner product space. If we write $u = \{u_k\}_{k \in \mathbb{Z}}$ and $v = \{v_k\}_{k \in \mathbb{Z}}$, the inner product on $\ell^2(\mathbb{Z})$ and the norm induced are

$$\langle u, v \rangle \triangleq \sum_{k \in \mathbb{Z}} u_k \overline{v_k} \quad \text{and} \quad \|u\|_2 \triangleq \left(\sum_{k=-\infty}^{+\infty} |u_k|^2 \right)^{1/2}.$$

3. When $(\mathcal{S}, \Sigma, \mu)$ is a measure space, the space $L^2(\mathcal{S}, \Sigma, \mu)$ is an inner product space if it is equipped with the inner product:

$$\langle f, g \rangle \triangleq \int_{\mathcal{S}} f \overline{g} \, \mathrm{d}\mu$$

with the induced norm

$$\|f\|_2 \triangleq \left(\int_{\mathcal{S}} |f|^2 \mathrm{d}\mu \right)^{1/2}.$$

4. The space $C[0,1]$ of all continuous functions on $[0,1]$ can be equipped with the inner product

$$\langle f, g \rangle = \int_0^1 f(x) \overline{g(x)} \mathrm{d}x.$$

Definition 12 (Convergence of a sequence). Let $\mathcal{X}$ be a normed space equipped with the norm $\|\cdot\|$. A sequence $\{x_n\}_{n \in \mathbb{N}}$ of elements of $\mathcal{X}$ *converges* to x if for every $\varepsilon > 0$ there exists an integer $N(\varepsilon)$ such that $\|x_n - x\| \leqslant \varepsilon$ whenever $n \geqslant N(\varepsilon)$.

Theorem 5 (Norm is continuous). Let $\mathcal{X}$ be a normed vector space. Then, $\|\cdot\|$ is continuous, i.e., if a sequence $\{x_n\}_{n \in \mathbb{N}}$ of elements of $\mathcal{X}$ converges to $x \in \mathcal{X}$, then $\|x_n\|$ converges to $\|x\|$.

Theorem 6 (Inner product is continuous)**.** Let $\mathcal{X}$ be an inner product space. Then, $\langle \cdot, \cdot \rangle$ is continuous, i.e., if two sequences $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ of elements of $\mathcal{X}$ converge to $x \in \mathcal{X}$ and $y \in \mathcal{X}$, respectively, then $\langle x_n, y_n \rangle$ converges to $\langle x, y \rangle$.

3 Banach and Hilbert spaces

Definition 13 (Cauchy sequence)**.** Let $\mathcal{X}$ be a normed space equipped with the norm $\|\cdot\|$. A sequence $\{x_n\}_{n\in\mathbb{N}}$ of elements of $\mathcal{X}$ is called a *Cauchy sequence* if for every $\varepsilon > 0$ there exists an integer $N(\varepsilon)$ such that $\|x_n - x_m\| \leqslant \varepsilon$ whenever $n, m \geqslant N(\varepsilon)$.

Theorem 7. Every convergent sequence is Cauchy.

Definition 14 (Complete space)**.** A normed space $\mathcal{X}$ is complete if every Cauchy sequence in $\mathcal{X}$ converges in $\mathcal{X}$.

Definition 15 (Banach space)**.** A Banach space is a complete normed space.

Definition 16 (Hilbert space)**.** A Hilbert space is a complete inner product space.

Definition 17. A subset $\mathcal{S} \subseteq \mathcal{X}$ of a normed space $\mathcal{X}$ is called a closed set if it contains all limit points, i.e., if $\{x_n\}_{n\in\mathbb{N}}$ is a sequence of points $x_n \in \mathcal{S}$ with $\lim_{n\to\infty} x_n = x \in \mathcal{X}$, then $x \in \mathcal{S}$. We write $\overline{\mathcal{S}}$ for the smallest closed subset of $\mathcal{X}$ containing $\mathcal{S}$ in the following sense: if $\mathcal{A}$ is any other closed subset of $\mathcal{X}$ such that $\mathcal{A} \supset \mathcal{S}$, then $\mathcal{A} \supset \overline{\mathcal{S}}$. We say $\overline{\mathcal{S}}$ is the closure of $\mathcal{S}$ in $\mathcal{X}$.

Examples.

1. Consider $\mathbb{R}^2$ with norm $\|\cdot\|_2$. The open disk $\mathcal{D} = \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\|_2 < 1\}$ is not a closed set (consider the sequence $\{(1 - 1/n, 0)\}_{n\in\mathbb{N}}$). The closure of $\mathcal{D}$ is the closed disk $\overline{\mathcal{D}} = \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\|_2 \leqslant 1\}$.

2. Let $\mathcal{X} = L^2[0,1]$ be the space of square-integrable functions $[0,1] \to \mathbb{C}$ with the inner product
$$\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}\mathrm{d}x, \tag{3}$$
and let $\mathcal{S} = C[0,1] \subset \mathcal{X}$ be the set of all continuous functions $[0,1] \to \mathbb{C}$. Then $\overline{\mathcal{S}} = \mathcal{X}$. To see this, fix a function $f \in \mathcal{X}$ and let $\hat{f}_n = \langle f, e^{2\pi i n \cdot} \rangle$ denote its n-th Fourier coefficient. Then we have $\|f - f_N\| \to 0$, where $f_N = \sum_{n=-N}^{N} \hat{f}_n e^{2\pi i n \cdot}$ is the N-th partial Fourier sum. But, since f_N is a linear combination of continuous functions $x \to e^{2\pi i n x}$, f_N is itself continuous, for each $N \in \mathbb{N}$. Therefore $f \in \overline{\mathcal{S}}$. Since f was arbitrary, we obtain $\mathcal{X} \subset \overline{\mathcal{S}}$. But $\overline{\mathcal{S}} \subset \mathcal{X}$ in general, so $\overline{\mathcal{S}} = \mathcal{X}$, as desired.

3. To say that a set $\mathcal{S}$ is closed only makes sense relative to the space $\mathcal{X}$ that $\mathcal{S}$ is a subset of. To see this, let $\mathcal{X}_1 = C[0,1]$ be the space of all continuous functions from $[0,1]$ to $\mathbb{C}$ with the inner product (3), and let $\mathcal{X}_2 = L^2[0,1]$ be the space of all square-integrable functions $f : [0,1] \to \mathbb{C}$ with the same inner product. Let $\mathcal{S} = \{f \in C[0,1] : f(x) = 0 \text{ for } x \in [0, \frac{1}{2}]\}$ and note that $\mathcal{S}$ is a subset of both $\mathcal{X}_1$ and $\mathcal{X}_2$. However, $\mathcal{S}$ is closed in $\mathcal{X}_1$, but it is not closed in $\mathcal{X}_2$. Indeed, suppose

$\{f_n\}_{n\in\mathbb{N}}$ is a sequence in $\mathcal{S}$ that converges to some $f \in \mathcal{X}_1 = C[0,1]$. Then

$$\int_0^{\frac{1}{2}} |f(x)|^2 \mathrm{d}x = \int_0^{\frac{1}{2}} |f(x)-f_n(x)|^2 \mathrm{d}x \leqslant \int_0^1 |f(x)-f_n(x)|^2 \mathrm{d}x = \|f-f_n\|^2 \to 0,$$

as $n \to \infty$, so $f(x) = 0$, for all $x \in [0, \frac{1}{2}]$. Since we have assumed $f \in C[0,1]$, we have $f \in \mathcal{S}$ and so $\mathcal{S}$ is closed in $\mathcal{X}_1$. To see that $\mathcal{S}$ is *not* closed in $\mathcal{X}_2$, consider the following sequence of functions:

$$f_n(x) := \begin{cases} 0 & x \in [0, \frac{1}{2}], \\ n\left(x - \frac{1}{2}\right) & x \in (\frac{1}{2}, \frac{1}{2} + \frac{1}{n}], \\ 1 & x \in (\frac{1}{2} + \frac{1}{n}, 1]. \end{cases}$$

Note that $f_n \in \mathcal{S}$ for all $n \in \mathbb{N}$. Moreover we have

$$\int_0^1 |f_n(x) - \mathbb{1}_{[\frac{1}{2},1]}(x)|^2 \mathrm{d}x \leqslant \int_{\frac{1}{2}}^{\frac{1}{2}+\frac{1}{n}} 1 \, \mathrm{d}x = \frac{1}{n} \to 0,$$

as $n \to \infty$, so $f_n \to \mathbb{1}_{[\frac{1}{2},1]}$ in $\mathcal{X}_2$. But $\mathbb{1}_{[\frac{1}{2},1]} \notin \mathcal{S}$, so $\mathcal{S}$ is not closed in $\mathcal{X}_2$.

4. Since the concept of closedness depends on the ambient space, so does the concept of closure. In the above example, the closure of $\mathcal{S}$ in $\mathcal{X}_1$ is the set $\mathcal{S}$ itself, whereas the closure of $\mathcal{S}$ in $\mathcal{X}_2$ is the set $\overline{S} = \{f \in L^2[0,1] : f(x) = 0 \text{ for } x \in [0, \frac{1}{2}]\} \supsetneq S$.

Definition 18. Let $\mathcal{A} \subset \mathcal{B}$ be subsets of a normed space $\mathcal{X}$. We say that $\mathcal{A}$ is dense in $\mathcal{B}$ if $\overline{\mathcal{A}} = \mathcal{B}$. Equivalently, $\mathcal{A}$ is dense in $\mathcal{B}$ if, for every $x \in \mathcal{B}$ and every $\epsilon > 0$, there exists a $y \in \mathcal{A}$ such that $\|x - y\| < \epsilon$.

Theorem 8 (Completion of normed space)**.** Let $\mathcal{X}$ be a normed space with norm $\|\cdot\|$. Then, there exists a Banach space $\widehat{\mathcal{X}}$ with norm $\widehat{\|\cdot\|}$ such that the following properties hold.

- $\mathcal{X}$ is a subspace of $\widehat{\mathcal{X}}$;
- On $\mathcal{X}$, $\widehat{\|\cdot\|} = \|\cdot\|$;
- $\widehat{\mathcal{X}}$ is the closure of $\mathcal{X}$.

Theorem 9 (Completion of inner product spaces)**.** Let $\mathcal{X}$ be an inner product space with inner product $\langle\cdot,\cdot\rangle$. Then, there exists a Hilbert space $\widehat{\mathcal{X}}$ with inner product $\widehat{\langle\cdot,\cdot\rangle}$ such that the following properties hold.

- $\mathcal{X}$ is a subspace of $\widehat{\mathcal{X}}$;
- On $\mathcal{X}$, $\widehat{\langle\cdot,\cdot\rangle} = \langle\cdot,\cdot\rangle$;
- $\widehat{\mathcal{X}}$ is the closure of $\mathcal{X}$.

Furthermore, $\widehat{\mathcal{X}}$ is unique up to linear isometries.

Examples. 1. The set of rational numbers $\mathbb{Q}$ is an inner product space with inner product equal to pq for $p, q \in \mathbb{Q}$. This space is not complete.

2. The set of real numbers $\mathbb{R}$ is a Hilbert space space with inner product equal to rs for $r, s \in \mathbb{R}$.

3. $\mathbb{C}^N$ is a Hilbert space.

4. $\ell^2(\mathbb{Z})$ is a Hilbert space.

5. When $(\mathcal{S}, \Sigma, \mu)$ is a measure space, $L^2(\mathcal{S}, \Sigma, \mu)$ is a Hilbert space.

6. The space $C[0,1]$ equipped with the inner product

$$\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}\mathrm{d}x.$$

is not a Hilbert space, because it is not complete. The completion of $C[0,1]$ is $L^2([0,1])$.

4 Orthogonality in Hilbert spaces

Definition 19. Two vectors x and y in an inner product space $\mathcal{X}$ are called orthogonal if $\langle x, y \rangle = 0$. In this case we write $x \perp y$. Two sets $\mathcal{U}, \mathcal{V} \subseteq \mathcal{X}$ are called orthogonal if $\langle x, y \rangle = 0$ for all $x \in \mathcal{U}$ and $y \in \mathcal{V}$. We write $\mathcal{U} \perp \mathcal{V}$ if the sets $\mathcal{X}, \mathcal{Y}$ are orthogonal.

Definition 20. Let $\mathcal{S}$ be a nonempty subset of the inner product space $\mathcal{X}$. We define the orthogonal complement $\mathcal{S}^\perp$ of $\mathcal{S}$ as $\mathcal{S}^\perp = \{x \in \mathcal{X} \colon x \perp \mathcal{S}\}$.

Lemma 1. Let $\mathcal{S}$ be a nonempty subset of the inner product space $\mathcal{X}$. Then $\mathcal{S}^\perp$ is a closed linear subspace of $\mathcal{X}$.

Proof. Follows from the linearity and continuity of the inner product. □

The following theorem expresses one of the fundamental geometrical properties of Hilbert spaces. While the result may appear obvious (see Figure 1), the proof is not trivial.

Theorem 10 (Projection on closed subspaces)**.** Let $\mathcal{S}$ be a closed subspace of a Hilbert space $\mathcal{X}$. Then, the following properties hold.

(a) For each $x \in \mathcal{X}$ there is a unique closest point $y \in \mathcal{S}$ such that

$$\|x - y\| = \min_{z \in \mathcal{S}} \|x - z\|.$$

(b) The point $y \in \mathcal{S}$ closest to $x \in \mathcal{X}$ is the unique element of $\mathcal{S}$ such that the "error" $(x - y) \in \mathcal{S}^\perp$.

Corollary 4.1. Let $\mathcal{S}$ be a closed subspace of a Hilbert space $\mathcal{X}$. Then $\mathcal{X} = \mathcal{S} \oplus \mathcal{S}^\perp$ and $\mathcal{S}^{\perp\perp} = \mathcal{S}$.

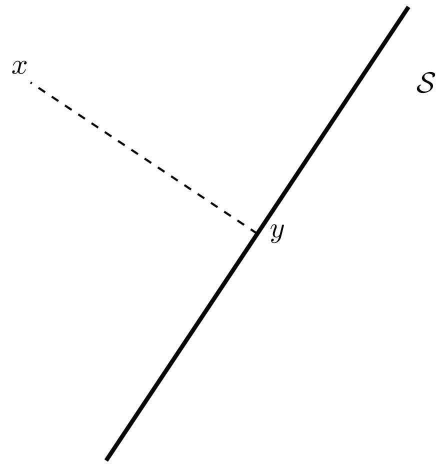

Figure 1: $y \in \mathcal{S}$ is the point in the closed subspace $\mathcal{S}$ closest to x and the unique point such that the "error" $(x - y) \in \mathcal{S}^\perp$.

Lemma 2. Let $\mathcal{S}$ be a subset of a Hilbert space $\mathcal{X}$. Then, the following properties hold.

(a) $\mathcal{S}^\perp = \overline{\text{span}(\mathcal{S})}^\perp$;

(b) $\mathcal{S}^{\perp\perp} = \overline{\text{span}(\mathcal{S})}$;

(c) $\mathcal{S}^\perp = \{0\}$ if and only if $\overline{\text{span}(\mathcal{S})} = \mathcal{X}$.

Proof.

- Proof of (a): $\overline{\text{span}(\mathcal{S})}^\perp \subseteq \mathcal{S}^\perp$ follows from $\mathcal{S} \subseteq \overline{\text{span}(\mathcal{S})}$. It remains to show that $\mathcal{S}^\perp \subseteq \overline{\text{span}(\mathcal{S})}^\perp$, which follows from the linearity and continuity of the inner product.
- Proof of (b): Follows from (a) and Corollary (4.1).

Proof of (c): Follows from (a) and Corollary (4.1). □

Corollary 4.1 and Lemma 2 tell us that subspaces of Hilbert spaces behave analogously to subspaces of finite-dimensional vector spaces familiar to us. In these results the completeness of $\mathcal{X}$ is crucial, as illustrated by the following example:

> **Example.**
> Let $\mathcal{X} = C[0,1]$ equipped with the inner product
>
> $$\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}\mathrm{d}x.$$
>
> As we have already seen, $\mathcal{X}$ is an inner product space, but not a Hilbert space. Let $\mathcal{S} = \{f \in C[0,1] : f(x) = 0 \text{ for } x \in [0, \frac{1}{2}]\}$ and note that this is a closed subspace of $\mathcal{X}$. Its orthogonal complement is $\mathcal{S}^\perp = \{f \in C[0,1] : f(x) = 0 \text{ for } x \in [\frac{1}{2}, 1]\}$, but $\mathcal{X} \neq \mathcal{S} \oplus \mathcal{S}^\perp$.

5 Orthonormal bases in Hilbert spaces

Definition 21 (Orthogonal/orthonormal set)**.** A subset $\mathcal{S}$ of nonzero vectors of a Hilbert space $\mathcal{X}$ is orthogonal if any two distinct elements in $\mathcal{S}$ are orthogonal. An orthogonal set $\mathcal{S}$ is called orthonormal if $\|x\| = 1$ for all $x \in \mathcal{S}$.

Examples. 1. The set $\{t \mapsto e^{2i\pi nt}\}_{n\in\mathbb{Z}}$ of complex exponentials is orthonormal in $L^2[0,1]$.

2. The set $\{\mathbf{e}_n\}_{n=0}^{N-1}$, with $\mathbf{e}_n[k] = e^{2i\pi kn/N}/\sqrt{N}$ for $0 \leqslant k \leqslant N-1$, is an orthonormal set in $\mathbb{C}^N$.

3. Considering the space

$$L^2_w(\mathbb{R}) = \left\{ f\colon \mathbb{R} \to \mathbb{R}\colon \int_{-\infty}^{+\infty} |f(x)|^2 w(x)\mathrm{d}x < \infty \right\}$$

equipped with the inner product

$$\langle f, g\rangle = \int_{-\infty}^{+\infty} f(x)g(x)w(x)\mathrm{d}x$$

with $w(x) = e^{-x^2/2}$. We define the polynomial H_n, $n \in \mathbb{N}$ such that

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-x^2/2} = (-1)^n H_n(x) e^{-x^2/2}.$$

The polynomials H_n, $n \in \mathbb{N}$, are the *Hermite polynomials* and they an orthogonal set in $L^2_w(\mathbb{R})$.

4. A function that is a sum of finitely many periodic functions is said to be quasiperiodic. If the ratios of the periods of the terms in the sum are rational, then the sum is itself periodic, but if at least one of the ratios is irrational, then the sum is not periodic. For example,

$$f(t) = e^{it} + e^{i\pi t}$$

is quasiperiodic but not periodic. Let $\mathcal{X}$ be the space of quasiperiodic functions $f\colon \mathbb{R} \to \mathbb{C}$ of the form

$$f(t) = \sum_{k=1}^{n} a_k e^{i\omega_k t}, \quad n \in \mathbb{N}, a_k \in \mathbb{C}, \omega_k \in \mathbb{R}.$$

Then $\mathcal{X}$ is a vector space. Furthermore, $\mathcal{X}$ is an inner product space with inner product

$$\langle f, g\rangle = \lim_{T\to\infty} \frac{1}{T} \int_{-T}^{T} f(t)\overline{g(t)}. \tag{4}$$

The set of functions

$$\{e^{i\omega t} \mid \omega \in \mathbb{R}\} \tag{5}$$

is orthonormal in $\mathcal{X}$. Note that $\mathcal{X}$ is an inner product space but not complete. The completion of $\mathcal{X}$ is the space of L^2-almost periodic functions and consists of equivalence classes of functions of the form

$$f(t) = \sum_{k=1}^{\infty} a_k e^{i\omega_k t}, \quad a_k \in \mathbb{C}, \omega_k \in \mathbb{R},$$

where $\sum_{k=1}^{\infty} |a_k|^2 < \infty$. The set in (5) is an uncountable orthogonal subset of this Hilbert space.

We next introduce the concept of unconditional convergence, which allows us to define sums of uncountable number of terms.

Definition 22 (Unconditional convergence)**.** Let $\{x_\alpha \in \mathcal{X} \mid \alpha \in \mathcal{I}\}$ be an indexed subset of a normed space $\mathcal{X}$, where $\mathcal{I}$ may be countable or not. For each finite subset $\mathcal{J}$ of $\mathcal{I}$, we define the partial sum

$$S_{\mathcal{J}} = \sum_{\alpha \in \mathcal{J}} x_\alpha.$$

The *unordered sum* $\sum_{\alpha \in \mathcal{I}} x_\alpha$ converges to $x \in \mathcal{X}$, written as

$$x = \sum_{\alpha \in \mathcal{I}} x_\alpha,$$

if for every $\varepsilon > 0$ there exists a *finite* subset $\mathcal{J}_\varepsilon \subseteq \mathcal{I}$ such that $\|S_{\mathcal{J}} - x\| < \varepsilon$ for all *finite* subsets $\mathcal{J} \subseteq \mathcal{I}$ containing $\mathcal{J}_\varepsilon$.

Note that unconditional convergence is independent of permutations of the index set $\mathcal{I}$.

Example.
Let $\{x_\alpha \in \mathcal{X} \mid \alpha \in \mathcal{I}\}$ be an indexed set of non-negative real numbers $x_\alpha \geqslant 0$ and set

$$M = \sup\Big\{ \sum_{\alpha \in \mathcal{J}} x_\alpha \mid \mathcal{J} \subseteq \mathcal{I}, \mathcal{J} \text{ finite}\Big\}.$$

- If $M = \infty$, then $\sum_{\alpha \in \mathcal{J}} x_\alpha$ is arbitrarily large for sufficiently large $\mathcal{J}$. Therefore, $\sum_{\alpha \in \mathcal{I}} x_\alpha$ does not converge unconditionally.
- Suppose that $0 \leqslant M < \infty$. Then, for each $\varepsilon > 0$ there exists a finite set $\mathcal{J}_\varepsilon \subseteq \mathcal{J}$ such that

  $$M - \varepsilon < \sum_{\alpha \in \mathcal{J}_\varepsilon} x_\alpha \leqslant M.$$

  It follows that for each finite set $\mathcal{J} \subseteq \mathcal{I}$ containing J_ε we have

  $$M - \varepsilon < \sum_{\alpha \in \mathcal{J}} x_\alpha \leqslant M, \tag{6}$$

  showing that the unordered sum converges to M.

Definition 23. An unordered sum $\sum_{\alpha\in\mathcal{I}} x_\alpha$ is Cauchy if for every $\varepsilon > 0$ there exists a finite subset $\mathcal{J}_\varepsilon \subseteq \mathcal{I}$ such that $\|S_\mathcal{K}\| < \varepsilon$ for all finite sets $\mathcal{K} \subseteq \mathcal{I} \setminus \mathcal{J}_\varepsilon$.

Proposition 5.1. An unordered sum in a Banach space converges if and only if it is Cauchy. If an unordered sum in a Banach space converges then $\{\alpha \in \mathcal{I} \mid x_\alpha \neq 0\}$ is at most countable.

Lemma 3. Let $\mathcal{S} = \{x_\alpha \in \mathcal{X} \mid \alpha \in \mathcal{I}\}$ be an indexed orthogonal subset of a Hilbert space $\mathcal{X}$. Then, the sum $\sum_{\alpha\in\mathcal{I}} x_\alpha$ converges unconditionally to some $x \in \mathcal{X}$ if and only if $\sum_{\alpha\in\mathcal{I}} \|x_\alpha\|^2 < \infty$. In that case, we have

$$\Big\|\sum_{\alpha\in\mathcal{I}} x_\alpha\Big\|^2 = \sum_{\alpha\in\mathcal{I}} \|x_\alpha\|^2. \tag{7}$$

Proof. For every finite set $\mathcal{K} \subseteq \mathcal{I}$ we have

$$\Big\|\sum_{\alpha\in\mathcal{K}} x_\alpha\Big\|^2 = \sum_{\alpha,\beta\in\mathcal{K}} \langle x_\alpha, x_\beta\rangle = \sum_{\alpha\in\mathcal{K}} \|x_\alpha\|^2.$$

Therefore, $\sum_{\alpha\in\mathcal{I}} x_\alpha$ is Cauchy if and only if $\sum_{\alpha\in\mathcal{I}} \|x_\alpha\|^2$ is Cauchy. Thus, one of the sums converges unconditionally if and only if the other does. Finally, (7) follows from the continuity of the norm and the fact that the sum is the limit of a sequence of finite partial sums. □

Proposition 5.2 (Bessel's inequality)**.** Let $\mathcal{X}$ be a Hilbert space, $\mathcal{S} = \{x_\alpha \in \mathcal{X} \mid \alpha \in \mathcal{I}\}$ an orthonormal subset of $\mathcal{X}$, and $x \in \mathcal{X}$. Then

(a) $\sum_{\alpha\in\mathcal{I}} |\langle x, x_\alpha\rangle|^2 \leqslant \|x\|^2$;

(b) $x_\mathcal{S} \triangleq \sum_{\alpha\in\mathcal{I}} \langle x, x_\alpha\rangle\, x_\alpha$ converges unconditionally;

(c) $x - x_\mathcal{S} \in \mathcal{S}^\perp$.

Proof. Let $\mathcal{J} \subseteq \mathcal{I}$ be a finite set. It follows immediately from the orthogonality that

$$\sum_{\alpha\in\mathcal{J}} |\langle x, x_\alpha\rangle|^2 = \|x\|^2 - \Big\|x - \sum_{\alpha\in\mathcal{J}} \langle x, x_\alpha\rangle\, x_\alpha\Big\|^2 \leqslant \|x\|^2. \tag{8}$$

Therefore, (see Example 5) (a) holds. Furthermore, (a) implies (b) because of Lemma 3. Finally, the statement (c) means that $\langle x - x_\mathcal{S}, x_\alpha\rangle = 0$ for all $\alpha \in \mathcal{I}$, which is true by the continuity of the inner product and the fact that $x - \sum_{\beta\in\mathcal{K}} \langle x, x_\beta\rangle x_\beta \perp x_\alpha$ for all finite subsets $\mathcal{K}$ of $\mathcal{I}$ and all $\alpha \in \mathcal{A}$. □

Definition 24 (Closed linear span)**.** Given a subset $\mathcal{S}$ of a Hilbert space $\mathcal{X}$, we define the *closed* linear span $[\mathcal{S}]$ of $\mathcal{S}$ by

$$[\mathcal{S}] = \Big\{\sum_{x\in\mathcal{S}} c_x x \mid c_x \in \mathbb{F}, \sum_{x\in\mathcal{S}} c_x x \text{ converges unconditionally}\Big\}.$$

Lemma 3 implies that the closed linear span $[\mathcal{S}]$ of an orthonormal subset $\mathcal{S} = \{x_\alpha \in \mathcal{X} \mid \alpha \in \mathcal{I}\}$ of a Hilbert space $\mathcal{X}$ can be written as

$$[\mathcal{S}] = \Big\{\sum_{\alpha\in\mathcal{I}} c_\alpha x_\alpha \mid c_\alpha \in \mathbb{F}, \sum_{\alpha\in\mathcal{I}} |c_\alpha|^2 < \infty\Big\}.$$

Theorem 11. If $\mathcal{S} = \{x_\alpha \in \mathcal{X} \mid \alpha \in \mathcal{I}\}$ is an orthonormal subset of a Hilbert space $\mathcal{X}$, then the following conditions are equivalent:

(a) $\langle x, x_\alpha \rangle = 0$ for all $\alpha \in \mathcal{I}$ implies that $x = 0$;

(b) $x = \sum_{\alpha\in\mathcal{I}} \langle x, x_\alpha \rangle x_\alpha$ for all $x \in \mathcal{X}$;

(c) $\|x\|^2 = \sum_{\alpha\in\mathcal{I}} |\langle x, x_\alpha \rangle|^2$ for all $x \in \mathcal{X}$;

(d) $[\mathcal{S}] = \mathcal{X}$;

(e) $\mathcal{S}$ is a maximal orthonormal set, i.e. it has the property that if $\mathcal{S}' \supset \mathcal{S}$ is another orthonormal set, then necessarily $S' = S$.

Proof.

- (a) implies (b): (a) states that $\mathcal{S}^\perp = \{0\}$. Therefore, Proposition 5.2 implies (b);
- (b) implies (c): Follows from Lemma 3;
- (c) implies (d): (c) implies that $\mathcal{S}^\perp = \{0\}$. Therefore, we also have $[\mathcal{S}]^\perp = \{0\}$;
- (d) implies (e): (d) implies that $x = \sum_{\alpha\in\mathcal{I}} c_\alpha x_\alpha$ for all $x \in \mathcal{X}$. Therefore, if $x \perp x_\alpha$ for all $\alpha \in \mathcal{I}$, then $x = 0$.
- (e) implies (a): Suppose for a contradiction that there exists an $x \in \mathcal{X}$ such that $\langle x, x_\alpha \rangle = 0$ for all $\alpha \in \mathcal{I}$, but $x \neq 0$. Then $\mathcal{S} \cup \{x/\|x\|\}$ is an orthonormal set and is a strict superset of $\mathcal{S}$.

□

Definition 25. An orthonormal subset $\mathcal{S} = \{x_\alpha \in \mathcal{X} \mid \alpha \in \mathcal{I}\}$ of a Hilbert space $\mathcal{X}$ is *complete* if it satisfies any of the equivalent conditions (a)–(e) in Theorem 11. A complete orthonormal subset of $\mathcal{X}$ is called *orthonormal basis* of $\mathcal{X}$.

Condition (a) is the easiest to verify. Condition (b) is used most often. Condition (c) is called *Parseval's identity.* Condition (d) simply expresses completeness. Condition (e) can be used to prove that every Hilbert space has an orthonormal basis using Zorn's lemma.

If a Hilbert space $\mathcal{X}$ has an orthonormal basis $\mathcal{S} = \{x_\alpha \in \mathcal{X} \mid \alpha \in \mathcal{I}\}$, then it is isomorphic to the sequence space $l^2(\mathcal{I})$. In what follows, we are mostly interested in Hilbert spaces that have a countable orthonormal bases, which are called *separable.*

Examples.

1. The set $\{t \mapsto e^{2i\pi nt}\}_{n\in\mathbb{Z}}$ of complex exponentials forms an orthonormal basis for $L^2[0,1]$.
2. The set $\{H_n\}_{n\in\mathbb{N}}$ of Hermite polynomials forms an orthogonal basis for $L_w(\mathbb{R})$. It can be transformed into an orthonormal basis by noting that $\|H_n\|^2 = \sqrt{2\pi}n!$ for all $n \in \mathbb{N}$.

6 Bounded linear operators on a Hilbert space

Definition 26. Let $\mathcal{X}$, $\mathcal{Y}$ be vector spaces over $\mathbb{F}$. A linear operator (or linear transformation, or linear map) from $\mathcal{X}$ to $\mathcal{Y}$ is a function $\mathbb{T}\colon \mathcal{X} \to \mathcal{Y}$ that satisfies

$$\mathbb{T}(\alpha x + \beta y) = \alpha \mathbb{T}x + \beta \mathbb{T}y$$

for all $x, y \in \mathcal{X}$ and $\alpha, \beta \in \mathbb{F}$. The set $\mathcal{R}(\mathbb{T}) = \{\mathbb{T}x \colon x \in \mathcal{X}\}$ is called the *range* of $\mathbb{T}$. The dimension of $\mathcal{R}(\mathbb{T})$ is called the *rank*. The set $\mathcal{N}(\mathbb{T}) = \{x \in \mathcal{X} \colon \mathbb{T}x = 0\}$ is called the *null space* or *kernel* of $\mathbb{T}$.

Example.
If $\mathcal{X}$ is N-dimensional with basis $\{x_n\}_{n=1}^N$ and $\mathcal{Y}$ is M-dimensional with basis $\{y_m\}_{m=1}^M$, then $\mathbb{T}$ is completely determined by its matrix representation $\mathbf{U} = \{U_{n,m}\}_{\substack{1 \leqslant n \leqslant N \\ 1 \leqslant m \leqslant M}}$ with respect to these two bases:

$$\mathbb{T}x_n = \sum_{m=1}^{M} U_{n,m} y_m, \quad n = 1, 2, \ldots, N.$$

Therefore, if $x = \sum_{n=1}^N \alpha_n x_n$ and $y = \mathbb{T}x$, we have

$$y = \sum_{m=1}^{M} \beta_m y_m \qquad \text{with} \quad \beta_m = \sum_{n=1}^{N} U_{m,n} \alpha_n.$$

That is, if $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N) \in \mathbb{F}^N$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_M) \in \mathbb{F}^M$ are the coefficient vectors of x and y with respect to the bases $\{x_n\}_{n=1}^N$ and $\{y_m\}_{m=1}^M$, respectively, then we have the matrix-vector relation: $\boldsymbol{\beta} = \mathbf{U}\boldsymbol{\alpha}$.

Proposition 6.1. $\mathcal{N}(\mathbb{T})$ is a subspace of $\mathcal{X}$ and $\mathcal{R}(\mathbb{T})$ is a subspace of $\mathcal{Y}$. Furthermore, if $\mathbb{T}$ is continuous then $\mathcal{N}(\mathbb{T})$ is closed.

Definition 27 (Bounded linear operators)**.** A linear operator $\mathbb{T} \colon \mathcal{X} \to \mathcal{Y}$ is bounded if there exists a constant $K > 0$ such that

$$\|\mathbb{T}x\|_{\mathcal{Y}} \leqslant K \|x\|_{\mathcal{X}}. \tag{9}$$

The set of bounded linear operators $\mathbb{T} \colon \mathcal{X} \to \mathcal{Y}$ is denoted by $\mathcal{B}(\mathcal{X}, \mathcal{Y})$.

Theorem 12. Let $\mathcal{X}, \mathcal{Y}$ be normed spaces. Then, $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is a normed space with norm

$$\|\mathbb{T}\| = \sup_{\|x\|_{\mathcal{X}}=1} (\|\mathbb{T}x\|_{\mathcal{Y}}).$$

This is called the *operator norm* of an operator between two normed vector spaces. Furthermore, if $\mathcal{Y}$ is a Banach space, then $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is a Banach space.

Proposition 6.2. Let $\mathcal{X}, \mathcal{Y}$ be normed spaces and $\mathbb{T} \colon \mathcal{X} \to \mathcal{Y}$ a linear operator. The following properties are equivalent.

(a) $\mathbb{T}$ is continuous;

(b) $\|\mathbb{T}\| < \infty$;

(c) $\mathbb{T}$ is bounded.

Proof.

(a) implies (b): Suppose that (b) does not hold. Then, there exists a sequence $\{x_n\}_{n=1}^{\infty}$ of elements $x_n \in \mathcal{X}$ with $\|x_n\|_{\mathcal{X}} = 1$ such that $\|\mathbb{T}x_n\|_{\mathcal{Y}} > n$. Set $y_n = x_n/n$. Then, $y_n \to 0$, but $\|\mathbb{T}y_n\|_{\mathcal{Y}} \geqslant 1$. Therefore, $\mathbb{T}$ is not continuous as $\lim_{n\to\infty} \mathbb{T}y_n \neq \mathbb{T}0 = 0$;

(b) implies (c): Obvious for $x = 0$. For $x \neq 0$, we have

$$\|\mathbb{T}x\|_{\mathcal{Y}} = \|x\|_{\mathcal{X}} \left\| \mathbb{T}\frac{x}{\|x\|_{\mathcal{X}}} \right\|_{\mathcal{Y}} \leqslant \|\mathbb{T}\| \, \|x\|_{\mathcal{X}} \, ;$$

(c) implies (a): (c) implies that $\mathbb{T}$ is uniformly continuous. □

Lemma 4. Let $\mathcal{X}, \mathcal{Y}$ be normed spaces and $\mathbb{T} \colon \mathcal{X} \to \mathcal{Y}$ a linear operator. Then,

$$\|\mathbb{T}x\|_{\mathcal{Y}} \leqslant \|\mathbb{T}\| \, \|x\|_{\mathcal{X}} \, , \quad x \in \mathcal{X}.$$

Proof. Obvious for $x = 0$. For $x \neq 0$, we have

$$\|\mathbb{T}x\|_{\mathcal{Y}} = \|x\|_{\mathcal{X}} \left\| \mathbb{T}\frac{x}{\|x\|_{\mathcal{X}}} \right\|_{\mathcal{Y}} \leqslant \|\mathbb{T}\| \, \|x\|_{\mathcal{X}} \, .$$

□

Lemma 5. If $\mathbb{T}_1$ and $\mathbb{T}_2$ are linear operators between normed spaces for which the range of $\mathbb{T}_2$ is contained in the domain of $\mathbb{T}_1$, we can consider the composed operator $\mathbb{T}_1\mathbb{T}_2$. If $\mathbb{T}_1$ and $\mathbb{T}_2$ are bounded, then also $\mathbb{T}_1\mathbb{T}_2$ is bounded and

$$\|\mathbb{T}_1\mathbb{T}_2\| \leqslant \|\mathbb{T}_1\| \, \|\mathbb{T}_2\| \, .$$

Proof.

$$\begin{aligned} \|\mathbb{T}_1\mathbb{T}_2\| &= \sup_{\|x\|=1} \left(\|\mathbb{T}_1(\mathbb{T}_2 x)\|\right) \\ &\leqslant \sup_{\|x\|=1} \left(\|\mathbb{T}_1\| \, \|(\mathbb{T}_2 x)\|\right) \\ &= \|\mathbb{T}_1\| \, \|\mathbb{T}_2\| \, . \end{aligned}$$

□

Definition 28. Let $\mathcal{X}, \mathcal{Y}$ be a normed space with norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{Y}}$, and let $\mathbb{T} : \mathcal{X} \to \mathcal{Y}$ be a linear operator. We say that $\mathbb{T}$ is a linear isometry if $\|\mathbb{T}x\|_{\mathcal{Y}} = \|x\|_{\mathcal{X}}$, for all $x \in \mathcal{X}$.

6.1 (Orthogonal) projections

Definition 29. A *projection* on a vector space $\mathcal{X}$ is a linear operator $\mathbb{P} \colon \mathcal{X} \to \mathcal{X}$ such that $\mathbb{P}^2 = \mathbb{P}$.

Theorem 13. Let $\mathcal{X}$ be a vector space.

(a) If $\mathbb{P} : \mathcal{X} \to \mathcal{X}$ is a projection, then $\mathcal{X} = \mathcal{R}(\mathbb{P}) \oplus \mathcal{N}(\mathbb{P})$;

(b) If $\mathcal{X} = \mathcal{S} \oplus \mathcal{T}$, then there exists a unique projection $\mathbb{P} \colon \mathcal{X} \to \mathcal{X}$ such that $\mathcal{S} = \mathcal{R}(\mathbb{P})$ and $\mathcal{T} = \mathcal{N}(\mathbb{P})$, called the projection onto $\mathcal{S}$ along $\mathcal{T}$.

Proof. We first show that $x \in \mathcal{R}(\mathbb{P})$ if and only if $x = \mathbb{P}x$. Clearly, $\mathbb{P}x \in \mathcal{R}(\mathbb{P})$. Suppose that $x \in \mathcal{R}(\mathbb{P})$. Then $x = \mathbb{P}z$ for some $z \in \mathcal{X}$ and, therefore, $\mathbb{P}x = \mathbb{P}\mathbb{P}z = \mathbb{P}z = x$.

Now we prove (a). Let $x \in \mathcal{N}(\mathbb{P}) \cap \mathcal{R}(\mathbb{P})$. Then $x = \mathbb{P}x$ and $\mathbb{P}x = 0$, which implies that $x = 0$. Therefore, $\mathcal{N}(\mathbb{P}) \cap \mathcal{R}(\mathbb{P}) = \{0\}$. Furthermore, we can decompose each $x \in \mathcal{X}$ as $x = \mathbb{P}x + (\mathbb{I} - \mathbb{P})x$ with $\mathbb{P}x \in \mathcal{R}(\mathbb{P})$ and $(\mathbb{I} - \mathbb{P})x \in \mathcal{N}(\mathbb{P})$. Therefore, $\mathcal{X} = \mathcal{R}(\mathbb{P}) \oplus \mathcal{N}(\mathbb{P})$.

Now we show that (b) holds. Each $x \in \mathcal{X}$ has a unique decomposition $x = s + t$ with $s \in \mathcal{S}$ and $t \in \mathcal{T}$, and $P(x) = s$ defines the required projection. □

Example.
Let

$$\mathcal{S} = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in \mathbb{R} \right\} \subseteq \mathbb{R}^2$$
$$\mathcal{T} = \left\{ \begin{pmatrix} x \\ x \end{pmatrix} \mid x \in \mathbb{R} \right\} \subseteq \mathbb{R}^2.$$

Then, $\mathcal{S}$ and $\mathcal{T}$ are linear subspaces of $\mathbb{R}^2$ with $\mathcal{S} \cap \mathcal{T} = \{0\}$. Furthermore, we can decompose each vector in $\mathbb{R}^2$ as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \underbrace{\begin{pmatrix} x - y \\ 0 \end{pmatrix}}_{\in \mathcal{S}} + \underbrace{\begin{pmatrix} y \\ y \end{pmatrix}}_{\in \mathcal{T}}.$$

Therefore, $\mathbb{R}^2 = \mathcal{S} \oplus \mathcal{T}$. Define the operator

$$\begin{aligned} \mathbb{P} \colon \mathbb{R}^2 &\to \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} &\mapsto \begin{pmatrix} x - y \\ 0 \end{pmatrix}. \end{aligned}$$

This operator is linear and $\mathbb{P}^2 = \mathbb{P}$. Therefore, $\mathbb{P}$ is a projection.

Definition 30. An *orthogonal projection* on a Hilbert space $\mathcal{X}$ is a linear operator $\mathbb{P} \colon \mathcal{X} \to \mathcal{X}$ such that $\mathbb{P}^2 = \mathbb{P}$ and

$$\langle \mathbb{P}x, y \rangle = \langle x, \mathbb{P}y \rangle, \quad x, y \in \mathcal{X}.$$

Theorem 14. Let $\mathcal{X}$ be a Hilbert space.

(a) If $\mathbb{P} \colon \mathcal{X} \to \mathcal{X}$ is an orthogonal projection, then $\mathcal{R}(\mathbb{P})$ is closed and $\mathcal{R}(\mathbb{P})^\perp = \mathcal{N}(\mathbb{P})$. Therefore, $\mathcal{X} = \mathcal{R}(\mathbb{P}) \oplus \mathcal{N}(\mathbb{P})$ is an orthogonal direct sum;

(b) If $\mathcal{S}$ is a closed subspace, then there exists a unique orthogonal projection $\mathbb{P} \colon \mathcal{X} \to \mathcal{X}$ such that $\mathcal{S} = \mathcal{R}(\mathbb{P})$ and $\mathcal{S}^\perp = \mathcal{N}(\mathbb{P})$.

Proof. We first prove (a). Let $\mathbb{P} \colon \mathcal{X} \to \mathcal{X}$ be an orthogonal projection. According to Property (a) in Theorem 13, $\mathcal{X} = \mathcal{R}(\mathbb{P}) \oplus \mathcal{N}(\mathbb{P})$. If $x = \mathbb{P}z \in \mathcal{R}(\mathbb{P})$ and $y \in \mathcal{N}(\mathbb{P})$, we have

$$\langle x, y \rangle = \langle \mathbb{P}z, y \rangle = \langle z, \mathbb{P}y \rangle = 0.$$

Therefore, $\mathcal{R}(\mathbb{P}) = \mathcal{N}(\mathbb{P})^\perp$ and, in particular, a closed subspace.

Now we show that (b) holds. Let $\mathcal{S}$ be a closed subspace of $\mathcal{X}$. Then, Corollary 4.1 implies that $\mathcal{X} = \mathcal{S} \oplus \mathcal{S}^\perp$. We define the projection $\mathbb{P}$ by

$$\mathbb{P}x = s, \quad \text{where } x = s + t \text{ with } s \in \mathcal{S}, t \in \mathcal{S}^\perp.$$

Then, $\mathcal{R}(\mathbb{P}) = \mathcal{S}$ and $\mathcal{N}(\mathbb{P}) = \mathcal{S}^\perp$. Now let $x_1, x_2 \in \mathcal{X}$ and $x_i = s_i + t_i$ be the unique decomposition of x_i with $s_i \in \mathcal{S}$ and $t_i \in \mathcal{S}^\perp$, $i = 1, 2$. Then, we have

$$\langle x_1, \mathbb{P}x_2 \rangle = \langle x_1, \mathbb{P}(s_2 + t_2) \rangle = \langle x_1, s_2 \rangle = \langle s_1 + t_1, s_2 \rangle = \langle s_1, s_2 \rangle = \langle s_1, s_2 + t_2 \rangle = \langle \mathbb{P}x_1, s_2 + t_2 \rangle .$$

Therefore, $\mathbb{P}$ is an orthogonal projection. The uniqueness follows from Theorem 13(b). □

Example.
Let

$$\mathcal{S} = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in \mathbb{R} \right\} \subseteq \mathbb{R}^2.$$

Then, $\mathcal{S}$ is a (closed) linear subspaces of $\mathbb{R}^2$. Furthermore, we have

$$\mathcal{S}^\perp = \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} \mid y \in \mathbb{R} \right\}.$$

Define the operator

$$\begin{aligned} \mathbb{P} \colon \mathbb{R}^2 &\to \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} &\mapsto \begin{pmatrix} x \\ 0 \end{pmatrix}. \end{aligned}$$

This operator is linear, $\mathbb{P}^2 = \mathbb{P}$, and

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}^{\mathrm{T}} \mathbb{P} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = x_1 x_2 = \left(\mathbb{P} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \right)^{\mathrm{T}} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \in \mathbb{R}^2.$$

Therefore, $\mathbb{P}$ is an orthogonal projection with $\mathcal{S} = \mathcal{R}(\mathbb{P})$ and $\mathcal{S}^\perp = \mathcal{N}(\mathbb{P})$.

Lemma 6. Let $\mathbb{P}$ be a non-zero orthogonal projection. Then $\|\mathbb{P}\| = 1$.

Proof. Recall that

$$\|\mathbb{P}\| = \sup_{\|x\|=1} (\|\mathbb{P}x\|).$$

If $x \neq 0$, the Cauchy-Schwarz inequality implies that

$$\|\mathbb{P}x\| = \frac{\langle \mathbb{P}x, \mathbb{P}x \rangle}{\|\mathbb{P}x\|} = \frac{\langle x, \mathbb{P}x \rangle}{\|\mathbb{P}x\|} \leqslant \|x\| . \tag{10}$$

Therefore, $\|\mathbb{P}\| \leqslant 1$. For $\mathbb{P}x \neq 0$, we have $\|\mathbb{P}(\mathbb{P}x)\| = \|\mathbb{P}x\|$, which implies that $\|\mathbb{P}\| \geqslant 1$. □

Proposition 6.3. Let $\mathcal{S} = \{x_\alpha : \alpha \in \mathcal{I}\}$ be an orthonormal system (but not necessarily an ONB) in a Hilbert space $\mathcal{X}$. Let $\mathcal{V} = [S]$ be the closed linear span of $\mathcal{S}$. Then the orthogonal projection $\mathbb{P}_\mathcal{V}$ is given explicitly by

$$\mathbb{P}_\mathcal{V}\, x = \sum_{\alpha \in \mathcal{I}} \langle x, x_\alpha \rangle x_\alpha.$$

Proof. The set $\mathcal{S}$ can be extended to an ONB $\mathcal{B}$ of $\mathcal{X}$ (if $\mathcal{X}$ is separable, then this can be done by applying the Gram-Schmidt orthogonalization procedure, or in the general case, by using Zorn's Lemma). Thus we can enumerate $\mathcal{B} = \{x_\alpha : \alpha \in \mathcal{J}\}$ by an index set $\mathcal{J}$ containing $\mathcal{I}$. Now, given an arbitrary $x \in \mathcal{X}$, define $s, t \in \mathcal{X}$ according to

$$s = \sum_{\alpha \in \mathcal{I}} \langle x, x_\alpha \rangle x_\alpha, \qquad t = \sum_{\alpha \in \mathcal{J} \setminus \mathcal{I}} \langle x, x_\alpha \rangle x_\alpha.$$

Then, by Theorem 11, we have

$$x = \sum_{\alpha \in \mathcal{J}} \langle x, x_\alpha \rangle x_\alpha = \sum_{\alpha \in \mathcal{I}} \langle x, x_\alpha \rangle x_\alpha + \sum_{\alpha \in \mathcal{J} \setminus \mathcal{I}} \langle x, x_\alpha \rangle x_\alpha = s + t.$$

Moreover,

$$\begin{aligned} \langle s, t \rangle &= \left\langle \sum_{\alpha \in \mathcal{I}} \langle x, x_\alpha \rangle x_\alpha, \sum_{\beta \in \mathcal{J} \setminus \mathcal{I}} \langle x, x_\beta \rangle x_\beta \right\rangle \\ &= \sum_{\substack{\alpha \in \mathcal{I} \\ \beta \in \mathcal{J} \setminus \mathcal{I}}} \langle x, x_\alpha \rangle \langle x, x_\beta \rangle \underbrace{\langle x_\alpha, x_\beta \rangle}_{=0, \text{ as } \alpha \neq \beta} = 0. \end{aligned}$$

Therefore, by Theorem 14, we have $\mathbb{P}_\mathcal{V}\, x = s$, as desired. ☐

7 The dual of a Hilbert space

Definition 31. Let $\mathcal{X}$ be a Hilbert space. An element of $\mathcal{B}(\mathcal{X}, \mathbb{F})$ is called a bounded linear functional.

> **Example.**
> Let $\mathcal{X}$ be a Hilbert space. For every $y \in \mathcal{X}$, the mapping
>
> $$\mathcal{X} \to \mathbb{F} \tag{11}$$
> $$x \mapsto \langle x, y \rangle \tag{12}$$
>
> is a bounded linear functional, which we denote by f_y.

Theorem 15 (Riesz representation theorem)**.** Let $f \colon \mathcal{X} \to \mathbb{F}$ be a bounded linear functional on a Hilbert space $\mathcal{X}$. Then there exists a unique $y_f \in \mathcal{X}$ such that $f(x) = \langle x, y_f \rangle$ for all $x \in \mathcal{X}$.

Therefore, the mapping

$$\begin{aligned} \theta_\mathcal{X} \colon \mathcal{X} &\to \mathcal{B}(\mathcal{X}, \mathbb{F}) \\ y &\mapsto f_y = \langle \cdot, y \rangle \end{aligned} \tag{13}$$

is a bijection. In the case of $\mathbb{F} = \mathbb{C}$, this mapping is antilinear (as can be seen from $\theta_{\mathcal{X}}(\alpha y) = \langle \cdot, \alpha y \rangle = \overline{\alpha} \langle \cdot, y \rangle = \theta_{\mathcal{X}}(y)$). Furthermore, the mapping θ is an isometry:

$$\|f_y\| = \sup_{\|x\|=1} |f_y(x)| = \sup_{\|x\|=1} |\langle x, y \rangle| = \|y\|, \quad y \neq 0.$$

What is more, $\mathcal{B}(\mathcal{X}, \mathbb{F})$ is a Hilbert space with inner product

$$\langle f_x, f_y \rangle = \langle y, x \rangle .$$

Theorem 16. Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ a bounded linear operator. Then, there exists a unique operator $\mathbb{T}^* \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$ such that

$$\langle \mathbb{T}x, y \rangle = \langle x, \mathbb{T}^* y \rangle , \quad x \in \mathcal{X}, y \in \mathcal{Y} \tag{14}$$

with $\|\mathbb{T}^*\| = \|\mathbb{T}\|$. Furthermore, we have $\mathbb{T}^{**} = \mathbb{T}$.

Proof. We first prove the existence. Let $y \in \mathcal{Y}$ be arbitrary but fixet. The mapping $x \mapsto \langle \mathbb{T}x, y \rangle$ is a bounded linear functional on $\mathcal{X}$. Therefore, according to Riesz's representation theorem, there exists a unique $z_y \in \mathcal{X}$ such that

$$\langle \mathbb{T}x, y \rangle = \langle x, z_y \rangle , \quad x \in \mathcal{X}.$$

We set $\mathbb{T}^* y = z_y$.

We now show that $\mathbb{T}^*$ is unique. Suppose that

$$\langle \mathbb{T}x, y \rangle = \langle x, \mathbb{T}_1^* y \rangle = \langle x, \mathbb{T}_2^* y \rangle , \quad x \in \mathcal{X}, y \in \mathcal{Y}$$

Then, we have

$$\langle x, (\mathbb{T}_1^* - \mathbb{T}_2^*) y \rangle = 0, \quad x \in \mathcal{X}, y \in \mathcal{Y}.$$

Therefore, $(\mathbb{T}_1^* - \mathbb{T}_2^*) y \in \mathcal{X}^\perp = \{0\}$ for all $y \in \mathcal{Y}$, which implies that $\mathbb{T}_1^* - \mathbb{T}_2^* = 0$ and, in turn, that $\mathbb{T}_1^* = \mathbb{T}_2^*$

The linearity of $\mathbb{T}^*$ follows from the uniqueness and the linearity of the inner product. Furthermore, we have

$$\begin{aligned}
\|\mathbb{T}^*\| &= \sup_{\|y\|=1} \|\mathbb{T}^* y\| \\
&= \sup_{\|y\|=1} \|z_y\| \\
&= \sup_{\|y\|=1} \|\langle \mathbb{T}\cdot, y \rangle\| \qquad (15) \\
&= \sup_{\|y\|=1} \sup_{\|x\|=1} |\langle \mathbb{T}x, y \rangle| \\
&\leqslant \sup_{\|y\|=1} \sup_{\|x\|=1} \|\mathbb{T}x\| \, \|y\| \qquad (16) \\
&= \|\mathbb{T}\| , \qquad (17)
\end{aligned}$$

where (15) follows from the fact that the mapping θ in (13) is an isometry and in (16) we applied the Cauchy-Schwarz inequality. $\mathbb{T}^{**} = \mathbb{T}$ follows immediately from (14). Furthermore, we have

$$\|\mathbb{T}\| = \|\mathbb{T}^{**}\| \leqslant \|\mathbb{T}^*\| \leqslant \|\mathbb{T}\| ,$$

which shows that $\|\mathbb{T}^*\| = \|\mathbb{T}\|$. ☐

Definition 32. Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. The operator $\mathbb{T}^* \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$ in Theorem 16 is called the adjoint of $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$.

Definition 33. Let $\mathcal{X}$ be a Hilbert space and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{X})$. Then, $\mathbb{T}$ is

1. self-adjoint if $\mathbb{T} = \mathbb{T}^*$;
2. unitary if $\mathbb{T}$ is invertible and $\mathbb{T}^{-1} = \mathbb{T}^*$;
3. normal if $\mathbb{T}\mathbb{T}^* = \mathbb{T}^*\mathbb{T}$.

Examples. 1. Every orthogonal projection $\mathbb{P} \colon \mathcal{X} \to \mathcal{X}$ is a self-adjoint operator.

2. If $\mathbb{T} \colon \mathcal{X} \to \mathcal{X}$ is a unitary operator then it is a linear isometry. Indeed,

$$\|\mathbb{T}x\|^2 = \langle \mathbb{T}x, \mathbb{T}x \rangle = \langle \mathbb{T}^*\mathbb{T}x, x \rangle = \langle \operatorname{Id} x, x \rangle = \|x\|^2, \text{ for all } x \in \mathcal{X}.$$

The converse is not true, that is, a linear isometry may not be unitary. To see this, consider the Hilbert space $\mathcal{X} = \{\boldsymbol{x} = (x_1, x_2, \dots) \in \mathbb{R}^{\mathbb{N}} : \sum_{n \in \mathbb{N}} |x_n|^2 < \infty\}$ with inner product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{n \in \mathbb{N}} x_n y_n$, and let $\mathbb{T}(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots)$ be the right-shift operator. Then $\mathbb{T}$ is a linear isometry, but it is not invertible, so it cannot be unitary.

Lemma 7. Let $\mathcal{X}$ be a Hilbert space and $\mathbb{T}$ self-adjoint. Then, we have

$$\|\mathbb{T}\| = \sup \{ |\langle \mathbb{T}x, x \rangle| : \|x\| = 1 \}.$$

Proof. Let

$$\alpha = \sup \{ |\langle \mathbb{T}x, x \rangle| : \|x\| = 1 \}.$$

From $|\langle \mathbb{T}x, x \rangle| \leqslant \|\mathbb{T}x\| \, \|x\| \leqslant \|\mathbb{T}\| \, \|x\|^2$ it follows that $\alpha \leqslant \|\mathbb{T}\|$. It remains to show that $\|\mathbb{T}\| \leqslant \alpha$. It follows from the Cauchy Schwarz inequality that

$$|\langle \mathbb{T}x, y \rangle| \leqslant \|\mathbb{T}x\| \, \|y\|, \quad x, y \in \mathcal{X}$$

with equality if and only if $\mathbb{T}x$ and y are colinear. Therefore, we can rewrite

$$\|\mathbb{T}x\| = \sup \{ |\langle \mathbb{T}x, y \rangle| : \|y\| = 1 \}, \tag{18}$$

which implies that

$$\begin{aligned} \|\mathbb{T}\| &= \sup_{\|x\|=1} (\|\mathbb{T}x\|) \\ &= \sup \{ |\langle \mathbb{T}x, y \rangle| : \|x\| = 1, \|y\| = 1 \} \\ &= \sup \{ |\langle \mathbb{T}x, y \rangle| : \|x\| = 1, \|y\| = 1, \langle \mathbb{T}x, y \rangle \in \mathbb{R} \}. \end{aligned} \tag{19}$$

It follows from the polarization formula ((2) for $\mathbb{F} = \mathbb{C}$ and (1) for $\mathbb{F} = \mathbb{R}$) in Theorem 3 for $\langle \mathbb{T}x, y\rangle \in \mathbb{R}$ that

$$\begin{aligned}
|\langle \mathbb{T}x, y\rangle| &= \frac{1}{4}\left| \|\mathbb{T}x + y\|^2 - \|\mathbb{T}x - y\|^2 \right| \\
&= \frac{1}{4}\left|\langle \mathbb{T}x + y, \mathbb{T}x + y\rangle - \langle \mathbb{T}x - y, \mathbb{T}x - y\rangle\right| \\
&= \frac{1}{4}\left|\langle \mathbb{T}(x + y), x + y\rangle - \langle \mathbb{T}(x - y), x - y\rangle\right| \\
&\leqslant \frac{1}{4}\left|\langle \mathbb{T}(x + y), x + y\rangle\right| + \frac{1}{4}\left|\langle \mathbb{T}(x - y), x - y\rangle\right| \\
&\leqslant \frac{\alpha}{4}(\|x + y\|^2 + \|x - y\|^2) \\
&= \frac{\alpha}{2}(\|x\|^2 + \|y\|^2) \quad , \quad x, y \in \mathcal{X} \text{ with } \langle \mathbb{T}x, y\rangle \in \mathbb{R}.
\end{aligned}$$

Using this result in (19) gives $\|\mathbb{T}\| \leqslant \alpha$. □

Definition 34. Let $\mathcal{X}$ be a Hilbert space. A self-adjoint operator $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{X})$ is

1. positive semidefinite if $\langle \mathbb{T}x, x\rangle \geqslant 0$ for all $x \in \mathcal{X}$. In this case, we write $\mathbb{T} \geqslant 0$.
2. positive definite if it is positive semidefinite and $\langle \mathbb{T}x, x\rangle = 0$ implies that $x = 0$. In this case, we write $\mathbb{T} > 0$.

Definition 35. Let $\mathcal{X}$ be a Hilbert space and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{X})$, $\mathbb{T} \geqslant 0$. We call a self-adjoint operator $\mathbb{A} \in \mathcal{B}(\mathcal{X}, \mathcal{X})$ a square root of $\mathbb{T}$ if $\mathbb{T} = \mathbb{A}^2$.

Proposition 7.1. Let $\mathcal{X}$ be a Hilbert space and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{X})$, $\mathbb{T} > 0$. Then, $\mathbb{T}$ has a unique positive definite square root.

Definition 36. An operator $\mathbb{T} \colon \mathcal{X} \to \mathcal{Y}$ is invertible if it is surjective and injective.

A Proof of Theorem 10

Proof of Theorem 10. We first prove the existence of y in (a). Set

$$d = \inf\{\|x - z\| \mid z \in \mathcal{S}\}. \tag{20}$$

From the definition of d, there exists a sequence $\{y_n\}_{n\in\mathbb{N}}$ of elements $y_n \in \mathcal{S}$ such that

$$\lim_{n\to\infty} \|x - y_n\| = d. \tag{21}$$

Suppose that this sequence is Cauchy. Since $\mathcal{X}$ is complete, there is a $y \in \mathcal{X}$ such that $\lim_{n\to\infty} y_n = y$, and since $\mathcal{S}$ is closed, we have $y \in \mathcal{S}$. Finally, it follows from the continuity of $\|\cdot\|$ that $\|x - y\| = \lim_{n\to\infty} \|x - y_n\| = d$.

It remains to show that $\{y_n\}_{n\in\mathbb{N}}$ is Cauchy. The parallelogram law applied to $(x - y_m)$ and $(x - y_n)$ implies that

$$\begin{aligned}
2\|x - y_n\|^2 + 2\|x - y_m\|^2 &= \|2x - y_m - y_n\|^2 + \|y_m - y_n\|^2 \\
&= 4\|x - (y_m + y_n)/2\|^2 + \|y_m - y_n\|^2 \\
&\geqslant 4d^2 + \|y_m - y_n\|^2,
\end{aligned} \tag{22}$$

where we used (20) and the fact that $(y_m + y_n)/2 \in \mathcal{S}$ in the last step. Furthermore, (21) implies that for each $\varepsilon > 0$ there exists an $N(\varepsilon)$ such that $\|x - y_n\| \leqslant d + \varepsilon$. Combining this with (22), we find that

$$\begin{aligned}\|y_m - y_n\|^2 &\leqslant 4(d+\varepsilon)^2 - 4d^2 \\ &= 4\varepsilon(2d+\varepsilon).\end{aligned}$$

Therefore, $\{y_n\}_{n\in\mathbb{N}}$ is Cauchy.

Now let $x \in \mathcal{X}$ and $y \in \mathcal{S}$ be the unique closest point from (a). Then, for each $\lambda \in \mathbb{F}$ and $z \in \mathcal{S}$, we have

$$\begin{aligned}\|x - y\|^2 &\leqslant \|x - y + \lambda z\|^2 \\ &= \|x-y\|^2 + |\lambda|^2\|z\|^2 + \langle x - y, \lambda z\rangle + \langle \lambda z, x - y\rangle \\ &= \|x-y\|^2 + |\lambda|^2\|z\|^2 + 2\Re(\bar{\lambda}\langle x - y, z\rangle).\end{aligned}$$

It follows that

$$-2\Re(\bar{\lambda}\langle x - y, z\rangle) \leqslant |\lambda|^2\|z\|^2, \quad \lambda \in \mathbb{C}, z \in \mathcal{S}.$$

We can write $\langle x - y, z\rangle = |\langle x - y, z\rangle| e^{i\varphi_z}$ (with $\varphi_z \in \{0, \pi\}$ in the case of $\mathbb{F} = \mathbb{R}$). Therefore, for $\lambda = -\varepsilon e^{i\varphi_z}$, we get

$$2|\langle x - y, z\rangle| \leqslant \varepsilon\|z\|^2, \quad \varepsilon > 0, z \in \mathcal{S}.$$

Taking the limit $\varepsilon \to 0$ gives

$$|\langle x - y, z\rangle| = 0, \quad z \in \mathcal{S}.$$

Now we prove the uniqueness in (a). Suppose y_1 and y_2 are elements of $\mathcal{S}$ such that $\|x - y_1\| = \|x - y_2\| = d$. Then since $x - y_1 \in \mathcal{S}^\perp$ and $y_1 - y_2 \in \mathcal{S}$, we have

$$\begin{aligned}d^2 = \|x - y_2\|^2 &= \|x - y_1\|^2 + \|y_1 - y_2\|^2 \\ &= d^2 + \|y_1 - y_2\|^2,\end{aligned}$$

so $y_1 = y_2$.

Finally, we prove the uniqueness in (b). Let y_1 and y_2 be points in $\mathcal{S}$ such that $x - y_1 \in \mathcal{S}^\perp$, $x - y_2 \in \mathcal{S}^\perp$. Since $x - y_1 \perp y_1 - z$ for all $z \in \mathcal{S}$, we have

$$\begin{aligned}d^2 = \inf_{z\in\mathcal{S}} \|x - z\|^2 &= \inf_{z\in\mathcal{S}} \left(\|x - y_1\|^2 + \|y_1 - z\|^2\right) \\ &= \|x - y_1\|^2 + \inf_{z\in\mathcal{S}} \|y_1 - z\|^2 \\ &= \|x - y_1\|^2.\end{aligned}$$

Therefore $\|x - y_1\| = d$, and by the same argument $\|x - y_2\| = d$. By the proof of uniqueness in (a) it follows that $y_1 = y_2$.

□

B Proof of Theorem 15

Proof. If $\mathcal{N}(f) = \mathcal{X}$ we have $y_f = 0$. Suppose that $\mathcal{N}(f) \neq \mathcal{X}$. Since f is continuous, $\mathcal{N}(f)$ is closed. Therefore, according to Corollary 4.1, we can decompose $\mathcal{X} = \mathcal{N}(f) \oplus \mathcal{N}(f)^\perp$. Since $\mathcal{N}(f) \neq \mathcal{X}$, there exists a non-zero $z \in \mathcal{N}(f)^\perp$. We can rewrite any $x \in \mathcal{X}$ as

$$x = \underbrace{\left(x - \frac{f(x)}{f(z)} z\right)}_{\in \mathcal{N}(f)} + \frac{f(x)}{f(z)} z$$

Therefore for any $x \in \mathcal{X}$ there exist $\alpha \in \mathbb{F}$ and $n \in \mathcal{N}(f)$ such that

$$x = \alpha z + n, \quad \alpha.$$

Taking the inner product with z implies that

$$\alpha = \frac{\langle x, z\rangle}{\|z\|^2}.$$

Evaluating f on $x = \alpha z + n$ gives

$$f(x) = \alpha f(z) = \frac{f(z)\langle x, z\rangle}{\|z\|^2}.$$

Therefore, we can set $z_f = (\overline{f(z)}/\|z\|^2) z$.

To prove uniqueness, suppose that

$$f(x) = \langle x, y_1\rangle = \langle x, y_2\rangle, \quad x \in \mathcal{X}.$$

In particular, we have

$$f(y_1 - y_2) = \langle y_1 - y_2, y_1\rangle = \langle y_1 - y_2, y_2\rangle,$$

which implies that $\|y_1 - y_2\| = 0$. Therefore, $y_1 = y_2$. □

C Spectrum of self-adjoint operators

Definition 37. Let $\mathcal{X}$ be a Hilbert space and $\mathbb{T}\colon \mathcal{X} \to \mathcal{X}$. The set

$$\sigma(\mathbb{T}) = \{\lambda \in \mathbb{F} \mid (\mathbb{T} - \lambda\mathbb{I}) \text{ is not invertible}\}$$

is called the spectrum of $\mathbb{T}$.

We can split the spectrum into $\sigma = \sigma_1 \cup \sigma_2$, where

$$\sigma_1(\mathbb{T}) = \{\lambda \in \mathbb{F} \mid (\mathbb{T} - \lambda\mathbb{I}) \text{ is not injective})\}$$
$$\sigma_2(\mathbb{T}) = \{\lambda \in \mathbb{F} \mid (\mathbb{T} - \lambda\mathbb{I}) \text{ is not surjective})\}.$$

Lemma 8. Let $\mathcal{X}$ be a Hilbert space and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{X})$ self-adjoint. Then, we have

(a) For all $x \in \mathcal{X}$ we have $\langle \mathbb{T}x, x\rangle \in \mathbb{R}$ and $\sigma_1(\mathbb{T}) \subseteq \mathbb{R}$;

(b) if $\lambda \notin \sigma_1(\mathbb{T})$, then $\overline{\mathcal{R}(\mathbb{T} - \lambda\mathbb{I})} = \mathcal{X}$;

(c) if there exists an $A > 0$ such that $\|(\mathbb{T} - \lambda\mathbb{I})x\| \geqslant A\|x\|$ for all $x \in \mathcal{X}$, then $\lambda \notin \sigma(\mathbb{T})$.

Proof. • Proof of (a): Let $x \in \mathcal{X}$. Then

$$\langle \mathbb{T}x, x\rangle = \langle x, \mathbb{T}x\rangle = \overline{\langle \mathbb{T}x, x\rangle} \in \mathbb{R}.$$

Now let $\lambda_1 \in \sigma_1(\mathbb{T})$. Then there exists an $x \in \mathcal{X}$, $x \neq 0$, such that $\mathbb{T}x = \lambda x$. Then $\lambda\|x\|^2 = \langle \mathbb{T}x, x\rangle \in \mathbb{R}$, so $\lambda \in \mathbb{R}$.

- Proof of (b): We show that $\mathcal{R}(\mathbb{T} - \lambda\mathbb{I})^\perp = \{0\}$. Let $x \in \mathcal{R}(\mathbb{T} - \lambda\mathbb{I})^\perp$. Then, we have

$$\begin{aligned} 0 &= \langle (\mathbb{T} - \lambda\mathbb{I})y, x\rangle \\ &= \left\langle y, (\mathbb{T} - \overline{\lambda}\mathbb{I})x\right\rangle, \quad y \in \mathcal{X}. \end{aligned}$$

Therefore, $(\mathbb{T} - \overline{\lambda}\mathbb{I})x = 0$. If $x \neq 0$, this implies that $\overline{\lambda} \in \sigma_1(\mathbb{T}) \subseteq \mathbb{R}$ and, in turn, that $\lambda \in \sigma_1(\mathbb{T})$, which is a contradiction. Therefore, $x = 0$.

- Proof of (c): it follows immediately that $\lambda \notin \sigma_1(\mathbb{T})$. Therefore, it remains to show that $\lambda \notin \sigma_2(\mathbb{T})$. According to (b), it remains to show that $\mathcal{R}(\mathbb{T} - \lambda\mathbb{I})$ is closed. Let $\{y_n = (\mathbb{T} - \lambda\mathbb{I})x_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{R}(\mathbb{T} - \lambda\mathbb{I})$ such that $y = \lim_{n\to\infty} y_n$. Then, we have

$$\begin{aligned} \|y_n - y_m\| &= \|(\mathbb{T} - \lambda\mathbb{I})(x_n - x_m)\| \\ &\geqslant A\,\|(x_n - x_m)\|\,. \end{aligned}$$

Therefore, $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in a Hilbert space and, therefore, converges to x. By the continuity of $\mathbb{T} - \lambda\mathbb{I}$, we have $y = (\mathbb{T} - \lambda\mathbb{I})x \in \mathcal{R}(\mathbb{T} - \lambda\mathbb{I})$.

□

Lemma 9. Let $\mathcal{X}$ be a Hilbert space and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{X})$ self-adjoint. Then $\sigma(\mathbb{T}) \subseteq \mathbb{R}$.

Proof. Let $\lambda = \alpha + i\beta \in \mathbb{C}$ with $\beta > 0$. We have to show that $\lambda \notin \sigma(\mathbb{T})$. We have

$$\langle (\mathbb{T} - \lambda\mathbb{I})x, x\rangle = \underbrace{\langle \mathbb{T}x, x\rangle}_{\in\mathbb{R}} - \lambda \underbrace{\|x\|^2}_{\in\mathbb{R}}, \quad x \in \mathcal{X}.$$

This implies that

$$\begin{aligned} 2i\,\mathrm{Im}(\langle (\mathbb{T} - \lambda\mathbb{I})x, x\rangle) &= (\overline{\lambda} - \lambda)\,\|x\|^2\,, \quad x \in \mathcal{X} \\ &= -2i\beta\,\|x\|^2\,, \quad x \in \mathcal{X} \end{aligned}$$

and, in turn, that

$$\begin{aligned} \|x\|\,\|(\mathbb{T} - \lambda\mathbb{I})x\| &\geqslant |\langle (\mathbb{T} - \lambda\mathbb{I})x, x\rangle| \\ &\geqslant |\beta|\,\|x\|^2\,, \quad x \in \mathcal{X}, \end{aligned}$$

where we used the Cauchy-Schwarz inequality in the first step. Therefore, Lemma 8 implies that $\lambda \notin \sigma(\mathbb{T})$. □

Theorem 17. Let $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{X})$ self-adjoint and set

$$m = \inf_{\|x\|=1} \langle \mathbb{T}x, x \rangle \qquad \text{and} \qquad M = \sup_{\|x\|=1} \langle \mathbb{T}x, x \rangle .$$

Then $\sigma(\mathbb{T}) \subseteq [m, M]$.

Proof. First note that

$$m \|x\|^2 \leqslant \underbrace{\|x\|^2 \left\langle \mathbb{T}\frac{x}{\|x\|}, \frac{x}{\|x\|} \right\rangle}_{\langle \mathbb{T}x, x\rangle} \leqslant M \|x\|^2 , \quad x \in \mathcal{X}.$$

Suppose that $\lambda = m - c$ with $c > 0$. Then, the Cauchy-Schwarz inequality implies that

$$\begin{aligned} \|x\| \, \|(\mathbb{T} - \lambda\mathbb{I})x\| &\geqslant |\langle (\mathbb{T} - \lambda\mathbb{I})x, x\rangle| \\ &\geqslant \langle (\mathbb{T} - \lambda\mathbb{I})x, x\rangle \\ &\geqslant \|x\|^2 (m - \lambda) \\ &= c \|x\|^2 , \quad x \in \mathcal{X}. \end{aligned}$$

Therefore, Lemma 8 implies that $\lambda \notin \sigma(\mathbb{T})$.

Now suppose that $\lambda = M + c$ with $c > 0$. Then, the Cauchy-Schwarz inequality implies that

$$\begin{aligned} \|x\| \, \|(\mathbb{T} - \lambda\mathbb{I})x\| &\geqslant |\langle (\mathbb{T} - \lambda\mathbb{I})x, x\rangle| \\ &\geqslant - \langle (\mathbb{T} - \lambda\mathbb{I})x, x\rangle \\ &\geqslant \|x\|^2 (\lambda - M) \\ &= c \|x\|^2 , \quad x \in \mathcal{X}. \end{aligned}$$

Again, Lemma 8 implies that $\lambda \notin \sigma(\mathbb{T})$. □

D Pseudo-inverse of a bounded linear operator

Let $\mathcal{X}$ and $\mathcal{Y}$ be Hilbert spaces. It is often desirable to find some kind of "inverse" for an operator $\mathbb{T} \colon \mathcal{X} \to \mathcal{Y}$ that is not invertible in the strict sense.

Theorem 18 (Inverse mapping theorem)**.** Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces. If $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is invertible, then its inverse $\mathbb{T}^{-1} \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$. Consequently, $\mathbb{T}(\mathcal{S})$ is closed if and only if $\mathcal{S}$ is closed.

Theorem 19. Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Then, we have

$$\overline{\mathcal{R}(\mathbb{T})} = \mathcal{N}(\mathbb{T}^*)^{\perp} \tag{23}$$

$$\mathcal{N}(\mathbb{T}) = \mathcal{R}(\mathbb{T}^*)^{\perp}. \tag{24}$$

Proof. We first show (23). If $y \in \mathcal{R}(\mathbb{T})$, there exists a $z \in \mathcal{X}$ such that $y = \mathbb{T}z$. Then, we have

$$\langle y, \tilde{y}\rangle = \langle \mathbb{T}z, \tilde{y}\rangle = \langle z, \mathbb{T}^*\tilde{y}\rangle = 0, \quad \tilde{y} \in \mathcal{N}(\mathbb{T}^*),$$

which implies that $\mathcal{R}(\mathbb{T}) \subseteq \mathcal{N}(\mathbb{T}^*)^{\perp}$. Since $\mathcal{N}(\mathbb{T}^*)^{\perp}$ is closed, it follows that $\overline{\mathcal{R}(\mathbb{T})} \subseteq \mathcal{N}(\mathbb{T}^*)^{\perp}$. On the other hand, if $y \in \mathcal{R}(\mathbb{T})^{\perp}$, we have

$$0 = \langle \mathbb{T}x, y\rangle = \langle x, \mathbb{T}^*y\rangle , \quad x \in \mathcal{X},$$

which implies that $\mathbb{T}^* y = 0$. This means that $\mathcal{R}(\mathbb{T})^\perp \subseteq \mathcal{N}(\mathbb{T}^*)$. By taking the complement of this relation we get $\mathcal{N}(\mathbb{T}^*)^\perp \subseteq \mathcal{R}(\mathbb{T})^{\perp\perp} = \overline{\mathcal{R}(\mathbb{T})}$. To prove (24), we first apply (23) to $\mathbb{T}^*$, use $\mathbb{T}^{**} = \mathbb{T}$, and apply complements. $\square$

An equivalent formulation of this theorem is that if $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ for two Hilbert spaces $\mathcal{X}, \mathcal{Y}$, then $\mathcal{Y}$ can be written as the orthogonal direct sum

$$\mathcal{Y} = \overline{\mathcal{R}(\mathbb{T})} \oplus \mathcal{N}(\mathbb{T}^*).$$

The equation $y = \mathbb{T}x$ has a solution $x \in \mathcal{X}$ if and only if $y \in \mathcal{R}(\mathbb{T})$. Therefore, if $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ has closed range, the equation $y = \mathbb{T}x$ has a solution $x \in \mathcal{X}$ if and only if $y \perp \mathcal{N}(\mathbb{T}^*)$.

Theorem 20. Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Then, $\mathcal{R}(\mathbb{T})$ is closed if and only if $\mathcal{R}(\mathbb{T}^*)$ is closed.

Proof. We show that $\mathcal{R}(\mathbb{T}^*)$ closed implies that $\mathcal{R}(\mathbb{T})$ is closed. We have the following situation:

$$\mathcal{X} \xrightarrow{\mathbb{T}} (\overline{\mathcal{R}(\mathbb{T})} \oplus \mathcal{R}(\mathbb{T})^\perp) \xrightarrow{\mathbb{T}^*} \mathcal{R}(\mathbb{T}^*) \subseteq \mathcal{Y}.$$

Using (23), if follows that the mapping

$$\begin{aligned} \mathbb{U} \colon \overline{\mathcal{R}(\mathbb{T})} &\to \mathcal{R}(\mathbb{T}^*) \\ y &\mapsto T^*(y) \end{aligned}$$

is a bounded invertible linear operator. Since $\mathcal{R}(\mathbb{R}^*)$ is closed in $\mathcal{X}$ which is complete, $\mathcal{R}(\mathbb{T}^*)$ is itself complete and hence a Hilbert space in its own right. Therefore, the adjoint

$$\mathbb{U}^* \colon \mathcal{R}(\mathbb{T}^*) \to \overline{\mathcal{R}(\mathbb{T})}$$

exists and is a bounded invertible linear operator (with inverse $\mathbb{U}^{*-1} = \mathbb{U}^{-1*}$). We now show that $\overline{\mathcal{R}(\mathbb{T})} \subseteq \mathcal{R}(\mathbb{T})$, which implies that $\mathcal{R}$ is closed. Let $y \in \overline{\mathcal{R}(\mathbb{T})}$ and denote by $\mathbb{P}$ the orthogonal projection onto $\overline{\mathcal{R}(\mathbb{T})}$. Then, $y = \mathbb{U}^*(x)$ for some $x \in \mathcal{R}(\mathbb{T}^*)$, and we have

$$\begin{aligned} \langle y, z \rangle &= \langle y, \mathbb{P}z \rangle \\ &= \langle \mathbb{U}^* x, \mathbb{P}z \rangle \\ &= \langle x, \mathbb{U}\mathbb{P}z \rangle \\ &= \langle x, \mathbb{T}^*\mathbb{P}z \rangle \\ &= \langle x, \mathbb{T}^* z \rangle \\ &= \langle \mathbb{T}x, z \rangle, \quad z \in \mathcal{Y}. \end{aligned}$$

Therefore, $y = \mathbb{T}x \in \mathcal{R}(\mathbb{T})$. $\square$

The following lemma gives a condition that ensures the existence of a right-inverse.

Lemma 10. Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ with closed range $\mathcal{R}(\mathbb{T})$. Then there exists a unique operator $\mathbb{T}^\dagger \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ satisfying

$$\mathcal{N}(\mathbb{T}^\dagger) = \mathcal{R}(\mathbb{T})^\perp \tag{25}$$

$$\mathcal{R}(\mathbb{T}^\dagger) = \mathcal{N}(\mathbb{T})^\perp \tag{26}$$

$$\mathbb{T}\mathbb{T}^\dagger y = y, \quad y \in \mathcal{R}(\mathbb{T}). \tag{27}$$

Proof. We have the orthogonal direct sums

$$\mathcal{X} = \mathcal{N}(\mathbb{T}) \oplus \mathcal{N}(\mathbb{T})^{\perp}$$
$$\mathcal{Y} = \mathcal{R}(\mathbb{T}) \oplus \mathcal{R}(\mathbb{T})^{\perp}.$$

To prove the existence of $\mathbb{T}^{\dagger}$, we consider the linear operator

$$\widetilde{\mathbb{T}} = \mathbb{T}|_{\mathcal{N}(\mathbb{T})^{\perp}} : \mathcal{N}(\mathbb{T})^{\perp} \to \mathcal{R}(\mathbb{T}).$$

It is clear that $\widetilde{\mathbb{T}}$ is linear and bounded. $\widetilde{\mathbb{T}}$ is injective because if $\widetilde{\mathbb{T}}x = 0$, it follows that $x \in \mathcal{N}(\mathbb{T})^{\perp} \cap \mathcal{N}(\mathbb{T}) = \{0\}$. To show that $\widetilde{\mathbb{T}}$ is surjective, let $y \in \mathcal{R}(\mathbb{T})$. Then, there exists an $x \in \mathcal{X}$ such that $y = \mathbb{T}(x)$. We can write $x = u + v$ with $u \in \mathcal{N}(\mathbb{T})$ and $v \in \mathcal{N}(\mathbb{T})^{\perp}$. Therefore, $y = \mathbb{T}(u+v) = \mathbb{T}(u) + \mathbb{T}(v) = \mathbb{T}(u) = \widetilde{\mathbb{T}}(u) \in \mathcal{R}(\widetilde{\mathbb{T}})$. It follows that $\widetilde{\mathbb{T}}^{-1}$ exists and $\widetilde{\mathbb{T}}^{-1} \in \mathcal{B}(\mathcal{R}(\mathbb{T}), \mathcal{N}(\mathbb{T})^{\perp})$ by Theorem 18. We define the operator $\mathbb{T}^{\dagger}$ as

$$\mathbb{T}^{\dagger}(y) = \widetilde{\mathbb{T}}^{-1}(y), \quad y \in \mathcal{R}(\mathbb{T}) \tag{28}$$
$$\mathbb{T}^{\dagger}(y) = 0, \quad y \in \mathcal{R}(\mathbb{T})^{\perp}. \tag{29}$$

It follows immediately that $\mathbb{T}^{\dagger}$ has the desired properties. To prove uniqueness, suppose that $\mathbb{T}_1^{\dagger}$ and $\mathbb{T}_2^{\dagger}$ fulfill Properties (25)–(27) and let $y \in \mathcal{Y}$. We can decompose $y = w + z$ with $w \in \mathcal{R}(\mathbb{T})$ and $z \in \mathcal{R}(\mathbb{T})^{\perp}$. Then, we have

$$\mathbb{T}_i^{\dagger}(y) = \mathbb{T}_i^{\dagger}(w) + \mathbb{T}_i^{\dagger}(z) = \mathbb{T}_i^{\dagger}(w) = \widetilde{\mathbb{T}}^{-1}(w), \quad i = 1, 2.$$

Therefore, $\mathbb{T}_1^{\dagger} = \mathbb{T}_2^{\dagger}$. □

Definition 38 (Pseudo-inverse)**.** Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ with closed range $\mathcal{R}(\mathbb{T})$. The unique operator $\mathbb{T}^{\dagger} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ satisfying Properties (25)–(27) is called the pseudo-inverse of $\mathbb{T}^{\dagger}$.

Proposition D.1 (Properties of $\mathbb{T}^{\dagger}$)**.** Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ with closed range $\mathcal{R}(\mathbb{T})$. Then, the following properties hold.

(a) The orthogonal projection of $\mathcal{Y}$ onto $\mathcal{R}(\mathbb{T})$ is given by $\mathbb{T}\mathbb{T}^{\dagger}$.

(b) The orthogonal projection of $\mathcal{X}$ onto $\mathcal{R}(\mathbb{T}^{\dagger})$ is given by $\mathbb{T}^{\dagger}\mathbb{T}$.

(c) $\mathbb{T}^{*}$ has closed range and $(\mathbb{T}^{*})^{\dagger} = (\mathbb{T}^{\dagger})^{*}$.

(d) On $\mathcal{R}(\mathbb{T})$, the operator $\mathbb{T}^{\dagger}$ is given explicitly by $\mathbb{T}^{\dagger} = \mathbb{T}^{*}(\mathbb{T}\mathbb{T}^{*})^{-1}$.

(e) If $\mathbb{T}$ is surjective, then given $y \in \mathcal{Y}$, the equation $\mathbb{T}x = y$ has a unique solution of minimal norm, namely $x = \mathbb{T}^{\dagger}y$.

Proof.

- (a) Follows from Theorem 14 and Properties (25) and (27);
- (b): Property (26) implies that $\mathbb{T}^{\dagger}\mathbb{T}x = 0$ for all $x \in \mathcal{R}(\mathbb{T}^{\dagger})^{\perp}$. Now let $x \in \mathcal{R}(\mathbb{T}^{\dagger})$. Then, $x = \mathbb{T}^{\dagger}y$ for some $y \in \mathcal{Y}$. Decomposing $y = u + v$ with $u \in \mathcal{R}(\mathbb{T})$ and $v \in \mathcal{R}(\mathbb{T})^{\perp}$, Property (25) implies that $x = \mathbb{T}^{\dagger}u$. Therefore, Property (26) implies that $\mathbb{T}^{\dagger}\mathbb{T}x = \mathbb{T}^{\dagger}\mathbb{T}\mathbb{T}^{\dagger}u = \mathbb{T}^{\dagger}u = x$. Theorem 14 now shows that $\mathbb{T}^{\dagger}\mathbb{T}$ is an orthogonal projection onto $\mathcal{R}(\mathbb{T}^{\dagger})$;

- (c): It follows from Theorem 20 that $\mathbb{T}^*$ has closed range. This implies that $\mathbb{T}^{*\dagger}$ is well defined and is the unique operator satisfying

$$\begin{aligned} &\mathcal{N}(\mathbb{T}^{*\dagger}) = \mathcal{R}(\mathbb{T}^*)^\perp \\ &\mathcal{R}(\mathbb{T}^{*\dagger}) = \mathcal{N}(\mathbb{T}^*)^\perp \\ &\mathbb{T}^*\mathbb{T}^{*\dagger}x = x, \quad x \in \mathcal{R}(\mathbb{T}^*). \end{aligned}$$

  Furthermore, it follows from (25)–(27) and (23) together with (24), that $\mathbb{T}^{\dagger *}$ satisfies the same properties. Therefore, $(\mathbb{T}^*)^\dagger = (\mathbb{T}^\dagger)^*$;

- It follows from $\mathcal{R}(\mathbb{T}) = \mathcal{N}(\mathbb{T}^*)^\perp$ and $\mathcal{N}(\mathbb{T}) = \mathcal{R}(\mathbb{T}^*)^\perp$ that the operator $\mathbb{U} : \mathcal{R}(\mathbb{T}) \to \mathcal{R}(\mathbb{T})$, $y \mapsto \mathbb{T}\mathbb{T}^*y$ is invertible and the operator

$$\mathbb{T}^*(\mathbb{T}\mathbb{T}^*)^{-1} \text{ on } \mathcal{R}(\mathbb{T}) \qquad \text{and} \qquad 0 \text{ on } \mathcal{R}(\mathbb{T})^\perp$$

  fulfills (25)–(27);

- $x = \mathbb{T}^\dagger y$ is a solution of $y = \mathbb{T}x$, i.e., $y = \mathbb{T}\mathbb{T}^\dagger y$. Suppose that $y = \mathbb{T}z$ for some $z \in \mathcal{X}$ and set $e = \mathbb{T}^\dagger y - z$. It follows that $e \in \mathcal{N}(\mathbb{T})$. Since $\mathbb{T}^\dagger y \in \mathcal{N}(\mathbb{T})^\perp$, we have $e \perp \mathbb{T}^\dagger y$, which implies that $\|z\|^2 = \left\|\mathbb{T}^\dagger y - e\right\|^2 = \left\|\mathbb{T}^\dagger y\right\|^2 + \|e\|^2$, which is minimal for $e = 0$.

□

E Application to frame theory

Let $\{f_k\}_{k=0}^\infty$ be a frame with frame bounds $0 < A \leqslant B < \infty$ for a Hilbert space $\mathcal{X}$. It has been seen in the lecture that the frame operator

$$\begin{aligned} \mathbb{S}\colon\ & \mathcal{X} \to \mathcal{X} \\ & x \mapsto \textstyle\sum_{k=0}^\infty \langle x, f_k\rangle f_k \end{aligned}$$

is bounded, invertible, self-adjoint, and positive, and that $\{\mathbb{S}^{-1}f_k\}_{k=0}^\infty$ is a frame with frame bounds $0 < B^{-1} \leqslant A^{-1} < \infty$. In fact, $\{\mathbb{T}f_k\}_{k=0}^\infty$ is a frame for a larger class of operators $\mathbb{T}$, as stated in the following theorem.

Theorem 21. Let $\mathcal{X}$ be a Hilbert spaces and $\mathbb{T} \in \mathcal{B}(\mathcal{X}, \mathcal{X})$ with closed range $\mathcal{R}(\mathbb{T})$ and $\{f_k\}_{k=0}^\infty$ be a frame with

$$A\|x\|^2 \leqslant \sum_{k=0}^\infty |\langle x, f_k\rangle|^2 \leqslant B\|x\|^2, \quad x \in \mathcal{X}.$$

Then, $\{\mathbb{T}f_k\}_{k=0}^\infty$ is a frame in $\overline{\operatorname{span}}\{\mathbb{T}f_k\}_{k=0}^\infty$ with

$$A\left\|\mathbb{T}^\dagger\right\|^{-2}\|y\|^2 \leqslant \sum_{k=0}^\infty |\langle y, \mathbb{T}f_k\rangle|^2 \leqslant B\|\mathbb{T}\|^2\|y\|^2, \quad y \in \overline{\operatorname{span}}\{\mathbb{T}f_k\}_{k=0}^\infty. \tag{30}$$

Proof. In order to prove Theorem 21, we make use of the following lemma, which shows that it is enough to check the frame condition on a dense set.

Lemma 11. Suppose that $\{f_k\}_{k=0}^\infty$ is a sequence of elements in $\mathcal{X}$ and there exist constants $0 < A \leqslant B < \infty$ such that

$$A\|x\|^2 \leqslant \sum_{k=0}^\infty |\langle x, f_k\rangle|^2 \leqslant B\|x\|^2 \tag{31}$$

for all x in a dense subset of $\mathcal{X}$. Then (31) holds for all $x \in \mathcal{X}$.

The upper frame bound in (30) follows from

$$\sum_{k=0}^{\infty} |\langle y, \mathbb{T}f_k\rangle|^2 = \sum_{k=0}^{\infty} |\langle \mathbb{T}^* y, f_k\rangle|^2 \leqslant B \|\mathbb{T}^* y\|^2 \leqslant B \|\mathbb{T}\|^2 \|y\|^2, \quad y \in \mathcal{X}.$$

Now we prove that the lower frame bound in (30) holdes for all $y \in \operatorname{span}\{\mathbb{T}f_k\}_{k=0}^{\infty}$. We know that the operator $\mathbb{T}\mathbb{T}^\dagger$ is the orthogonal projection onto $\mathcal{R}(\mathbb{T})$ and, in particular, it is self-adjoint. Let $y \in \operatorname{span}\{\mathbb{T}f_k\}_{k=0}^{\infty}$. Then, $y \in \mathcal{R}(\mathbb{T})$, and we obtain that

$$y = (\mathbb{T}\mathbb{T}^\dagger)y = (\mathbb{T}\mathbb{T}^\dagger)^* y = (\mathbb{T}^\dagger)^* \mathbb{T}^* y.$$

This gives

$$\begin{aligned} \|y\|^2 &\leqslant \|(\mathbb{T}^\dagger)^*\|^2 \|\mathbb{T}^* y\|^2 &(32)\\ &= \|(\mathbb{T}^\dagger)\|^2 \|\mathbb{T}^* y\|^2. &(33)\end{aligned}$$

Furthermore, we have

$$\|\mathbb{T}^* y\|^2 \leqslant \frac{1}{A} \sum_{k=0}^{\infty} |\langle \mathbb{T}^* y, f_k\rangle|^2 = \frac{1}{A} \sum_{k=0}^{\infty} |\langle y, \mathbb{T}f_k\rangle|^2. \tag{34}$$

Plugging (34) into (32) gives

$$\|y\|^2 \leqslant \frac{\left\|\mathbb{T}^\dagger\right\|^2}{A} \sum_{k=0}^{\infty} |\langle y, \mathbb{T}f_k\rangle|^2,$$

which shows that the lower frame condition is satisfied for all $y \in \operatorname{span}\{\mathbb{T}f_k\}_{k=0}^{\infty}$. Using Lemma 11, we can conclude that $\{\mathbb{T}f_k\}_{k=0}^{\infty}$ forms a frame for $\overline{\operatorname{span}}\{\mathbb{T}f_k\}_{k=0}^{\infty}$. □

In the special case where $\mathbb{T}$ is a surjective bounded operator, $\mathbb{T}$ has closed range. Furthermore, every $y \in \mathcal{Y}$ can be written as $y = \mathbb{T}x$, where $x \in \mathcal{X}$. Since $\{f_k\}_{k=0}^{\infty}$ is a frame, one has

$$x = \sum_{k=0}^{\infty} \left\langle x, \mathbb{S}^{-1} f_k \right\rangle f_k,$$

and it follows that

$$y = \sum_{k=0}^{\infty} \left\langle x, \mathbb{S}^{-1} f_k \right\rangle \mathbb{T}f_k,$$

which shows that $\mathcal{Y} = \overline{\operatorname{span}}\{\mathbb{T}f_k\}_{k=0}^{\infty}$. This result leads to the following corollary.

Corollary E.1. Assume that $\{f_k\}_{k=0}^{\infty}$ is a frame for $\mathcal{X}$ with frame bounds $0 < A \leqslant B < \infty$ and that $\mathbb{T}\colon \mathcal{X} \to \mathcal{Y}$ is a bounded surjective operator. Then $\{\mathbb{T}f_k\}_{k=0}^{\infty}$ is a frame for $\mathcal{Y}$ with frame bounds $A \left\|\mathbb{T}^\dagger\right\|^{-2}$, $B \|\mathbb{T}\|^2$.

As a consequence, if we have at our disposal a frame, we can construct many other frames by applying surjective operators to it.

References

[1] J. K. Hunter and B. Nachtergaele, *Applied Analysis.* Singapore: World Scientific Publishing, 2005.

[2] S. B. Damelin and W. Miller, *The Mathematics of Signal Processing*, ser. Cambridge texts in applied mathematics. Cambridge, UK: Cambridge University Press, 2012.

[3] E. M. Stein and R. Sharachi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, ser. Princeton lectures in analysis. Princeton, NJ, USA: Princeton University Press, 2005.

[4] O. Christensen, *An Introduction to Frames and Riesz Bases*, ser. Applied and Numerical Harmonic Analysis. Boston, MA, USA: Birkhäuser, 2003.

[5] C. Heil, "Operators on Hilbert spaces," Functional analysis lecture notes, February 2006.