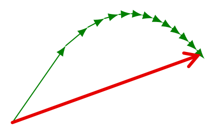

FUNCTIONAL ANALYSIS:

Hilbert Spaces

Luke Collins*

maths.com.mt/notes

Version 0.8† (13th January, 2019)

The infinite! No other question has ever moved so profoundly the spirit of man.

— David Hilbert

Contents

*with additional reference to the following texts: *Functional Analysis* by Prof Joseph Muscat (ISBN 978-3-319-06727-8), and *Introductory Functional Analysis* by Prof Erwin Kreyszig (ISBN 978-0-471-50459-7).

†If you find any mathematical, grammatical or typographical errors whilst reading these notes, please let the author know via email: luke.collins@um.edu.mt.

1 Introduction

We assume basic facts about real and complex numbers, vector spaces and metric spaces. We use $\mathbb{F}$ to denote one of the fields $\mathbb{R}$ and $\mathbb{C}$. For any complex number z, we denote the conjugate by $\bar{z}$, and its real and imaginary parts by $\Re z$ and $\Im z$, so that $z = \Re z + i\Im z$. For any vector v, we denote the transpose by v^T.

1.1 Inner Product Spaces

Recall that in geometry, the angle $\theta \in [0, \pi]$ between two vectors u and v is defined by the equation $u \cdot v = \|u\|\|v\| \cos\theta$, where the operation $\cdot$ maps two vectors to a number ($\cdot : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$). Here we present a general notion of such a product which we refer to as an *inner product*. A vector space together with an inner product is called an *inner product space*.

Definition 1.1 (Inner product space)**.** Let X be a vector space over the field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. An *inner product* on X is a function

$$\langle \cdot\,, \cdot \rangle \colon X \times X \to \mathbb{F}$$

such that for all $x, y, z \in X$ and $\lambda \in \mathbb{F}$, we have:

(i) $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$, and $\langle x, \lambda y\rangle = \lambda\langle x, y\rangle$, (Linearity in the second coordinate)

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

(iii) $\langle x, x\rangle \geqslant 0$, and $\langle x, x\rangle = 0 \iff x = 0$. (Positivity)

The pair $(X, \langle\cdot, \cdot\rangle)$ is said to be an *inner product space* or a *pre-Hilbert space*.

We will almost exclusively be using the brackets $\langle \cdot\,, \cdot \rangle$ to denote inner products in all contexts. Therefore we relax the pair notation $(X, \langle \cdot\,, \cdot \rangle)$, simply referring to the corresponding space as "*the inner product space X*".

Examples 1.2. (i) Let $X = \mathbb{R}^n$, and define $\langle \cdot\,, \cdot \rangle$ by $\langle x, y\rangle \stackrel{\text{def}}{=} x^\mathsf{T} y$. We show that this defines an inner product. Indeed, we have

$$\langle x, y + z\rangle = x^\mathsf{T}(y + z) = x^\mathsf{T} y + x^\mathsf{T} z = \langle x, y\rangle + \langle x, z\rangle$$

and

$$\langle x, \lambda y\rangle = x^\mathsf{T}(\lambda y) = \lambda(x^\mathsf{T} y) = \lambda\langle x, y\rangle,$$

which proves linearity. Conjugate symmetry is also easy to verify:

$$\begin{aligned}\langle x, y\rangle &= x^{\mathsf{T}} y \\ &= \sum_{k=1}^{n} x_k y_k \quad \text{(by definition of matrix multiplication)} \\ &= y^{\mathsf{T}} x = \langle y, x\rangle = \overline{\langle y, x\rangle},\end{aligned}$$

since $\langle y, x\rangle$ is real. Finally for positivity we have

$$\langle x, x\rangle = \sum_{k=1}^{n} {x_k}^2 \geqslant 0,$$

and clearly $\sum_{k=1}^{n} {x_k}^2 = 0 \iff x_k = 0$ for $k = 1, \dots n \iff x = 0$. Thus $\langle\,\cdot\,,\,\cdot\,\rangle$ defines an inner product on $X = \mathbb{R}^n$, and $(\mathbb{R}^n, \langle\,\cdot\,,\,\cdot\,\rangle)$ defines an inner product space.

(ii) Similarly for $X = \mathbb{C}^n$, the function $\langle w, z\rangle \stackrel{\text{def}}{=} \overline{w}^{\mathsf{T}} z$ defines an inner product (exercise: verify).

(iii) We may also consider vectors with infinitely many components (i.e., sequences). The set $\ell^2(\mathbb{F}) = \{(x_i)_{i\in\mathbb{N}} : x_i \in \mathbb{F} \text{ and } \sum_{i=1}^{\infty} |x_i|^2 < \infty\}$ of *square summable* sequences, together with the product

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle \stackrel{\text{def}}{=} \sum_{i=1}^{\infty} \bar{x}_i y_i$$

where $\boldsymbol{x} = (x_i)$ and $\boldsymbol{y} = (y_i)$, define an inner product space. We prove this later (proposition 1.13).

(iv) The set

$$C[a, b] = \{f \colon [a, b] \to \mathbb{F} : f \text{ is continuous}\}$$

of continuous functions defined on the interval $[a, b]$ together with the product $\langle f, g\rangle \stackrel{\text{def}}{=} \int_a^b \overline{f} g = \int_a^b \overline{f(t)} g(t)\, dt$ define an inner product space.

Indeed, for linearity we have

$$\begin{aligned}\langle f, g + h\rangle &= \int_a^b \overline{f}(g + h) = \int_a^b (\overline{f} g + \overline{f} h) \\ &= \int_a^b \overline{f} g + \int_a^b \overline{f} h = \langle f, g\rangle + \langle f, h\rangle\end{aligned}$$

and

$$\langle f, \lambda g\rangle = \int_a^b \bar{f}(\lambda g) = \int_a^b \lambda(\bar{f}g) = \lambda \int_a^b \bar{f}g = \lambda\langle f, g\rangle,$$

as required. For conjugate symmetry, we have

$$\begin{aligned}
\langle f, g\rangle = \int_a^b \bar{f}g &= \int_a^b (\Re f - i\Im f)(\Re g + i\Im g) \\
&= \int_a^b (\Re f \Re g + \Im f \Im g) + i\int_a^b (\Re f \Im g - \Im f \Re g) \\
&= \int_a^b (\Re f \Re g + \Im f \Im g) - i\int_a^b (\Im f \Re g - \Re f \Im g) \\
&= \overline{\int_a^b (\Re f \Re g + \Im f \Im g) + i\int_a^b (\Im f \Re g - \Re f \Im g)} \\
&= \overline{\int_a^b (\Re f + i\Im f)(\Re g - i\Im g)} \\
&= \overline{\int_a^b \bar{g}f} = \overline{\langle g, f\rangle},
\end{aligned}$$

as required. Finally for positivity, observe that

$$\langle f, f\rangle = \int_a^b \bar{f}f = \int_a^b |f| \geqslant 0,$$

and since f is continuous, $\int_a^b |f| = 0 \Leftrightarrow f = 0$ on $[a, b]$. ■

Proposition 1.3. *Let X be an inner product space. Then we have the following.*

(i) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$, *and* $\langle \lambda x, y\rangle = \bar{\lambda}\langle x, y\rangle$, (Antilinearity in the first coordinate)

(ii) If $x \in X$ is such that $\langle x, y\rangle = 0$ for any $y \in X$, then $x = 0$.

Proof. For (i), we have

$$\langle x + y, z\rangle = \overline{\langle z, x + y\rangle} = \overline{\langle z, x\rangle + \langle z, y\rangle} = \langle x, z\rangle + \langle y, z\rangle,$$

and

$$\langle \lambda x, y\rangle = \overline{\langle y, \lambda x\rangle} = \overline{\lambda\langle y, x\rangle} = \bar{\lambda}\langle x, y\rangle.$$

For (ii), observe that taking $y = x$, we have $\langle x, x\rangle = 0$, so that $x = 0$ by positivity of the inner product. □

1.2 Normed Spaces

Here we generalise the geometric notion of the length $\|v\|$ of a vector v for arbitrary vector spaces.

Definition 1.4 (Normed space)**.** Let X be a vector space over the field $\mathbb{F}$. A *norm* on X is a function

$$[\![\cdot]\!] \colon X \to \mathbb{R}$$

such that for all $x, y \in X$ and $\lambda \in \mathbb{F}$, we have:

(i) $[\![x]\!] \geqslant 0$ and $[\![x]\!] = 0 \iff x = 0$, (Positivity)

(ii) $[\![\lambda x]\!] = |\lambda| [\![x]\!]$, (Absolute scalability)

(iii) $[\![x + y]\!] \leqslant [\![x]\!] + [\![y]\!]$. (Triangle inequality)

The pair $(X, [\![\cdot]\!])$ is said to be a *normed space*.

Any inner product space is a normed space. Indeed, we have the following definition.

Definition 1.5 (Induced norm)**.** Let X be an inner product space. Them the *induced norm* on X is the function $\| \cdot \| \colon X \to \mathbb{R}$ defined by

$$\|x\| \overset{\text{def}}{=} \sqrt{\langle x, x\rangle}$$

for all $x \in X$.

Unless stated otherwise, the vertical bars $\| \cdot \|$ always denote to the induced norm of an inner product space.

Let us prove that the induced norm does indeed define a norm. Positivity follows immediately from the positivity of the inner product:

$$\|x\| \geqslant 0 \iff \sqrt{\langle x, x\rangle} \geqslant 0 \iff \langle x, x\rangle \geqslant 0,$$
$$\|x\| = 0 \iff \sqrt{\langle x, x\rangle} = 0 \iff \langle x, x\rangle = 0 \iff x = 0.$$

Absolute scalability also follows by proposition 1.3:

$$\|\lambda x\|^2 = \langle \lambda x, \lambda x\rangle = \lambda \langle \lambda x, x\rangle = \lambda\bar{\lambda}\langle x, x\rangle = |\lambda|^2 \|x\|^2.$$

Thus all there remains to show is the triangle inequality. But first we need the following results.

Proposition 1.6. *Let X be an inner product space. Then*

$$\|x+y\|^2 = \|x\|^2 + 2\Re\langle x, y\rangle + \|y\|^2.$$

Proof. We have

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \langle x+y, x\rangle + \langle x+y, y\rangle \\ &= \langle x, x\rangle + \langle y, x\rangle + \langle x, y\rangle + \langle y, y\rangle \qquad \text{(by proposition 1.3)} \\ &= \|x\|^2 + \overline{\langle x, y\rangle} + \langle x, y\rangle + \|y\|^2 \\ &= \|x\|^2 + 2\Re\langle x, y\rangle + \|y\|^2 \end{aligned}$$

since for any complex number z, we have $\bar{z} + z = 2\Re z$. □

Corollary 1.7 (Pythagoras' Theorem)**.** *Let X be an inner product space. If $z = x + y$ and $\langle x, y\rangle = 0$, then*

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

Proof. Follows immediately from proposition 1.6 since $\Re\langle x, y\rangle = 0$. □

Lemma 1.8. *Let X be an inner product space. Then for every $x, y \in X$ where $y \neq 0$, there exists $z \in X$ and $\lambda \in \mathbb{F}$ such that*

$$x = \lambda y + z$$

and $\langle y, z\rangle = 0$ (see figure 1).

Proof. Define $\lambda = \langle y, x\rangle/\|y\|^2$ and $z = x - \lambda y$. Clearly by definition of z we have $x = \lambda y + z$, and

$$\langle y, z\rangle = \langle y, x - \lambda y\rangle = \langle y, x\rangle - \lambda\langle y, y\rangle = \langle y, x\rangle - \lambda\|y\|^2 = 0,$$

as required. □

Theorem 1.9. *Let X be an inner product space. Then for any $x, y \in X$, we have*

(i) $|\langle x, y\rangle| \leqslant \|x\|\|y\|$ *(Cauchy–Schwarz inequality)*

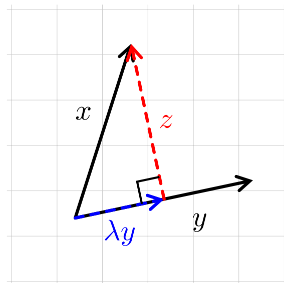

FIGURE 1: Illustration of lemma 1.8 in $\mathbb{R}^2$. For any given vector y, this result decomposes a given vector x in two components: one component along the vector y, the other perpendicular to y (Recall that in $\mathbb{R}^2$, $\langle y, z\rangle = 0$ means that the vectors y and z are perpendicular.)

(ii) $\|x + y\| \leqslant \|x\| + \|y\|$ *(Triangle inequality)*

Proof. Suppose $x, y \neq 0$ (if either of $x, y = 0$, the results are both trivial). By lemma 1.8, there exist λ, z such that $x = \lambda y + z$ with $\langle y, z\rangle = 0$. Also, $\langle \lambda y, z\rangle = \bar{\lambda}\langle y, z\rangle = 0$, so by Pythagoras' theorem, we have

$$\|x\|^2 = \|z\|^2 + \|\lambda y\|^2,$$

and in particular, $\|x\|^2 \geqslant \|\lambda y\|^2 = |\lambda|^2\|y\|^2$. Now in the proof of lemma 1.8, we had $\lambda = \langle y, x\rangle/\|y\|^2$, therefore

$$\|x\|^2 \geqslant |\lambda|^2\|y\|^2 \implies \|x\|^2 \geqslant \frac{|\langle y, x\rangle|^2}{\|y\|^4}\|y\|^2 \implies |\langle x, y\rangle| \leqslant \|x\|^2\|y\|^2,$$

which proves (i). Now for (ii), by proposition 1.6 we have

$$\begin{aligned}
\|x + y\|^2 &= \|x\|^2 + 2\Re\langle x, y\rangle + \|y\|^2 \\
&\leqslant \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 && (\Re z \leqslant |z| \text{ for all } z \in \mathbb{C}) \\
&\leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 && (\text{Cauchy–Schwarz}) \\
&= (\|x\| + \|y\|)^2,
\end{aligned}$$

as required. □

In obtaining the triangle inequality for the induced norm, we have completed the proof that every inner product space is also a normed space.

1.3 Hilbert Spaces

Recall that a metric space consists of a set X and a function $d\colon X \times X \to \mathbb{R}$ called the *metric* such that for all $x, y, z \in X$, we have

(i) $d(x, y) = 0 \iff x = y$, (Identity of indiscernibles)

(ii) $d(x, y) = d(y, x)$, (Symmetry)

(iii) $d(x, z) \leqslant d(x, y) + d(y, z)$. (Triangle inequality)

The goal of a metric space is to give a notion of distance. It shouldn't be hard to see that once we have a notion of length (a norm), a notion of distance is easily obtained. Indeed, if X is a normed space with norm $\|\cdot\|$, then the function

$$d(x, y) = \|x - y\|$$

defines a metric on X, which we call the *induced metric*.

Clearly we have identity of indiscernibles:

$$d(x, y) = 0 \iff \|x - y\| = 0 \iff x - y = 0 \iff x = y,$$

symmetry:

$$d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |-1|\|y - x\| = \|y - x\| = d(y, x),$$

and the triangle inequality:

$$\begin{aligned} d(x, z) = \|x - z\| &= \|x - y + y - z\| \\ &\leqslant \|x - y\| + \|y - z\| = d(x, y) + d(y, z), \end{aligned}$$

which concludes the proof. Therefore we have that every inner product space is a normed space, and every normed space is a metric space.

Now recall that a metric space (X, d) is said to be *complete* if every Cauchy sequence converges, that is, if a sequence $(x_n)_{n\in\mathbb{N}}$ satisfies $d(x_n, x_m) \to 0$, then there exists some $x \in X$ such that $x_n \to x$.

We are now ready to give the definition of a Hilbert space.

> **Definition 1.10** (Hilbert space)**.** An inner product space X is said to be a *Hilbert space* if it is complete with respect to the induced metric.

Let us prove that the set $\ell^2(\mathbb{C})$ introduced in examples 1.2(iii) is a Hilbert space. Before we do so, we need the following two propositions.

Proposition 1.11. *Let X be an inner product space over $\mathbb{C}$. Then for any $x, y \in X$, we have*

(i) $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$ (PARALLELOGRAM LAW)

(ii) $\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + i\|x-iy\|^2 - i\|x+iy\|^2)$ (POLARISATION IDENTITY)

Proof. By proposition 1.6,

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

which proves (i). Expanding in a similar way, we get that

$$\|x+y\|^2 - \|x-y\|^2 = 4\Re\langle x, y\rangle,$$

and that

$$\begin{aligned}i\|x-iy\|^2 - i\|x+iy\|^2 &= 2i\Re\langle x, -iy\rangle - 2i\Re\langle x, iy\rangle \\ &= 4i\Im\langle x, y\rangle,\end{aligned}$$

since $\Re(iz) = -\Im z$ for any complex number z. Adding and dividing by 4 yields (ii). □

Proposition 1.12 (Continuity of the inner product)**.** *Let X be an inner product space, and let (x_n), (y_n) be two sequences in X. Then if $x_n \to x$ and $y_n \to y$,*

$$\langle x_n, y_n\rangle \to \langle x, y\rangle.$$

Proof. Let $\epsilon > 0$. Since x_n converges, $\|x_n\|$ is bounded by some $M \in \mathbb{R}$. Also, by definition, there exists N_1 such that $\|x_n - x\| < \epsilon/2M$ for $n \geqslant N_1$, and similarly there exists N_2 such that $\|y_n - y\| < \epsilon/2\|y\|$ for $n \geqslant N_2$. Now

$$\begin{aligned}|\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n\rangle - \langle x_n, y\rangle + \langle x_n, y\rangle - \langle x, y\rangle| \\ &\leqslant |\langle x_n, y_n\rangle - \langle x_n, y\rangle| + |\langle x_n, y\rangle - \langle x, y\rangle| \\ &= |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\leqslant \|x_n\|\|y_n - y\| + \|y\|\|x_n - x\| \qquad \text{(Cauchy–Schwarz)} \\ &< M \cdot \epsilon/2M + \|y\| \cdot \epsilon/2\|y\| = \epsilon\end{aligned}$$

for $n \geqslant \max\{N_1, N_2\}$, as required. □

Proposition 1.13. *The set $\ell^2(\mathbb{C})$ defined by*

$$\ell^2(\mathbb{C}) \stackrel{\text{def}}{=} \left\{(x_i)_{i\in\mathbb{N}} : x_i \in \mathbb{C} \textit{ and } \sum_{i=1}^{\infty} |x_i|^2 < \infty\right\},$$

together with the inner product $\langle \boldsymbol{x}, \boldsymbol{y}\rangle \stackrel{\text{def}}{=} \sum_{i=1}^{\infty} \bar{x}_i y_i$, forms a Hilbert space.

Proof. First of all, $\ell^2(\mathbb{C})$ is a vector space under the operations

$$\boldsymbol{x} + \boldsymbol{y} = (x_i) + (y_i) \stackrel{\text{def}}{=} (x_i + y_i) \qquad \text{and} \qquad \lambda\boldsymbol{x} = \lambda(x_i) \stackrel{\text{def}}{=} (\lambda x_i).$$

Indeed, if $\boldsymbol{x}, \boldsymbol{y} \in \ell^2(\mathbb{C})$, then by the parallelogram law,

$$\sum_{i=1}^{\infty} |x_i + y_i|^2 \leqslant \sum_{i=1}^{\infty} \left(|x_i + y_i|^2 + |x_i - y_i|^2\right) = 2\sum_{i=1}^{\infty} |x_i|^2 + 2\sum_{i=1}^{\infty} |y_i|^2 < \infty,$$

and also $\sum_{i=1}^{\infty} |\lambda x_i|^2 = |\lambda|^2 \sum_{i=1}^{\infty} |x_i|^2 < \infty$; so it follows that $\ell^2(\mathbb{C})$ is closed under the operations. The remaining vector space properties follow from the field properties of $\mathbb{C}$ componentwise.

Now we check that $\langle\,\cdot\,,\,\cdot\,\rangle$ defines an inner product. That $\langle \boldsymbol{x}, \boldsymbol{y}\rangle$ is defined follows by the Cauchy–Schwarz inequality:

$$|\langle \boldsymbol{x}, \boldsymbol{y}\rangle| \leqslant \|\boldsymbol{x}\|\|\boldsymbol{y}\| = \sqrt{\sum_{i=1}^{\infty} |x_i|^2}\sqrt{\sum_{i=1}^{\infty} |y_i|^2} < \infty.$$

Linearity follows from the algebra of infinite series:

$$\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z}\rangle = \sum_{i=1}^{\infty} \bar{x}_i(y_i + z_i) = \sum_{i=1}^{\infty} \bar{x}_i y_i + \sum_{i=1}^{\infty} \bar{x}_i z_i = \langle \boldsymbol{x}, \boldsymbol{y}\rangle + \langle \boldsymbol{x}, \boldsymbol{z}\rangle,$$

$$\langle \boldsymbol{x}, \lambda\boldsymbol{y}\rangle = \sum_{i=1}^{\infty} \bar{x}_i(\lambda y_i) = \lambda\sum_{i=1}^{\infty} \bar{x}_i y_i = \lambda\langle \boldsymbol{x}, \boldsymbol{y}\rangle,$$

as does conjugate symmetry:

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{i=1}^{\infty} \bar{x}_i y_i = \sum_{i=1}^{\infty} \overline{x_i \bar{y}_i} = \overline{\sum_{i=1}^{\infty} x_i \bar{y}_i} = \overline{\langle \boldsymbol{y}, \boldsymbol{x}\rangle},$$

and for positivity, observe that $\langle \boldsymbol{x}, \boldsymbol{x}\rangle = \sum_{i=1}^{\infty} \bar{x}_i x_i = \sum_{i=1}^{\infty} |x_i|^2 \geqslant 0$, and that $\sum_{i=1}^{\infty} |x_i|^2 = 0 \Leftrightarrow x_i = 0$ for all $i \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$.

$$\begin{array}{rcccccc}
\boldsymbol{x}^1 & = & (x_1^1, & x_2^1, & x_3^1, & \dots) \\
\boldsymbol{x}^2 & = & (x_1^2, & x_2^2, & x_3^2, & \dots) \\
\text{claim}\big\downarrow & & \text{proved}\big\downarrow & \text{proved}\big\downarrow & \text{proved}\big\downarrow & \dots \\
\boldsymbol{x} & = & (x_1, & x_2, & x_3, & \dots)
\end{array}$$

FIGURE 2: Illustration of the claim that $\boldsymbol{x}^n \to \boldsymbol{x}$. We know that each component x_i^n converges to some x_i, we want to show that as a whole, $\boldsymbol{x}^n \to \boldsymbol{x}$.

Finally we show that ℓ^2 is complete. Indeed, suppose $(\boldsymbol{x}^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in ℓ^2. Its terms are themselves sequences of complex numbers:

$$\begin{aligned}
\boldsymbol{x}^1 &= (x_1^1, \quad x_2^1, \quad x_3^1, \quad \dots) \\
\boldsymbol{x}^2 &= (x_1^2, \quad x_2^2, \quad x_3^2, \quad \dots) \\
&\vdots
\end{aligned}$$

Fix $i \in \mathbb{N}$. We show that $(x_i^n)_{n\in\mathbb{N}}$ converges. Indeed, fix $\epsilon > 0$. Since $(\boldsymbol{x}^n)$ is Cauchy, there is $N \in \mathbb{N}$ such that $\|\boldsymbol{x}^n - \boldsymbol{x}^m\| < \epsilon$ for all $n, m \geqslant N$. But

$$|x_i^n - x_i^m|^2 \leqslant \sum_{j=1}^{\infty} |x_j^n - x_j^m|^2 = \|\boldsymbol{x}^n - \boldsymbol{x}^m\|^2 < \epsilon^2,$$

so it follows that $|x_i^n - x_i^m| < \epsilon$ for all $n, m \geqslant N$, i.e., that (x_i^n) is Cauchy. Hence, since $\mathbb{C}$ is complete, for each $i \in \mathbb{N}$, $\boldsymbol{x}_i^n \to x_i$ for some $x_i \in \mathbb{C}$.

Now we claim that $\boldsymbol{x}^n \to \boldsymbol{x} = (x_i)_{i\in\mathbb{N}}$, see figure 2. Indeed, fix $\epsilon > 0$. Just as before, there is $N \in \mathbb{N}$ such that $\|\boldsymbol{x}^n - \boldsymbol{x}^m\| < \sqrt{\epsilon}/2$ for all $n, m \geqslant N$. Thus for any k,

$$\sum_{i=1}^{k} |x_i^n - x_i^m|^2 \leqslant \sum_{i=1}^{\infty} |x_i^n - x_i^m|^2 = \|\boldsymbol{x}^n - \boldsymbol{x}^m\|^2 < \epsilon/4$$

for $n, m \geqslant N$. Let $f(a) = \sum_{i=1}^{k} |x_i^n - a|^2$. Clearly f is continuous, and since $x_i^m \to x_i$, we get $f(x_i^m) \to f(x_i) = \sum_{i=1}^{k} |x_i^n - x_i|^2$. By the squeeze theorem, since $f(x_i^m) < \epsilon$ for all m, it follows that the limit $f(x_i) \leqslant \epsilon/4$, i.e.,

$$\sum_{i=1}^{k} |x_i^n - x_i|^2 \leqslant \epsilon/4 < \epsilon/2.$$

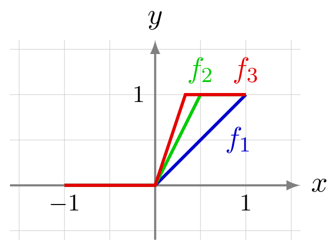

FIGURE 3: Illustration of the first three functions f_1, f_2, f_3 of the sequence $\boldsymbol{f}$ in remark 1.14.

By similar reasoning, interpreting the sum as a sequence in k, it follows that the limit $\sum_{i=1}^{\infty} |x_i^n - x_i|^2 \leqslant {}^{\epsilon}/_2 < \epsilon$, i.e., that $\|\boldsymbol{x}^n - \boldsymbol{x}\| < \epsilon$ for $n \geqslant N$, i.e., that $\boldsymbol{x}^n \to \boldsymbol{x}$.

Finally, pick an arbitrary $n \geqslant N$ (where N is the same as earlier). Since we had that $\sum_{i=1}^{\infty} |x_i^n - x_i|^2 < \epsilon$, it follows that $\boldsymbol{x}^n - \boldsymbol{x} \in \ell^2(\mathbb{C})$. Consequently $\boldsymbol{x}^n - (\boldsymbol{x}^n - \boldsymbol{x}) = \boldsymbol{x} \in \ell^2(\mathbb{C})$, which proves completeness. □

Remark 1.14. In examples 1.2(iv), we considered the set $C[a, b]$ of continuous functions on $[a, b]$, and showed that this is an inner product space with inner product $\langle f, g\rangle = \int_a^b fg$. Unlike ℓ^2 however, this is not a Hilbert space.

Indeed, consider $C[-1, 1]$ and the sequence $\boldsymbol{f} = (f_n)$ of functions defined by

$$f_n(t) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if } -1 \leqslant t \leqslant 0 \\ nt & \text{if } 0 \leqslant t \leqslant {}^1/_n \\ 1 & \text{otherwise.} \end{cases}$$

We show that $\boldsymbol{f}$ is Cauchy. Indeed, fix $\epsilon > 0$, and let $m, n \in \mathbb{N}$, supposing w.l.o.g. that $n > m$. Then

$$\|f_n - f_m\|^2 = \int_{-1}^{1} |f_n(t) - f_m(t)|^2 \, dt = \frac{n^2 - m^2}{3mn^2} < \frac{1}{3m} < \epsilon$$

for $m, n \geqslant N$, where N is large enough so that $3N\epsilon > 1$. This proves $\boldsymbol{f}$ is Cauchy.

Now we show that $\boldsymbol{f}$ does not converge in $C[-1, 1]$. Fix $\epsilon > 0$, and suppose that $\boldsymbol{f} \to f \in C[-1, 1]$, i.e., there exists $N \in \mathbb{N}$ such that for $n \geqslant N$,

$$\|f_n - f\|^2 = \int_{-1}^{1} |f_n(t) - f(t)|^2 \, dt < \epsilon.$$

In particular, since $f_n(t) = 0$ for $t \in [-1, 0]$ for all n,

$$0 \leqslant \int_{-1}^{0} |f(t)|^2 \, dt = \int_{-1}^{0} |f_n(t) - f(t)|^2 \, dt \leqslant \int_{-1}^{1} |f_n(t) - f(t)|^2 \, dt < \epsilon.$$

It follows by the squeeze theorem that $\int_{-1}^{0} |f(t)|^2 \, dt = 0$, and since f is continuous, it follows from integration theory that $f(t) = 0$ on $[-1, 0]$.

Now let $\delta \in (0, 1]$. Then for $n > \max\{N, 1/\delta\}$, we have

$$0 \leqslant \int_{\delta}^{1} |1 - f(t)|^2 \, dt = \int_{\delta}^{1} |f_n(t) - f(t)|^2 \, dt \leqslant \int_{-1}^{1} |f_n(t) - f(t)|^2 \, dt < \epsilon,$$

which again by the squeeze theorem yields $\int_{\delta}^{1} |1 - f(t)|^2 \, dt = 0$, and by the continuity of f, we have $1 - f(t) = 0$ for $t \in (\delta, 1]$. But δ is arbitrary, so it follows that $f(t) = 1$ for all $t \in (0, 1]$. We therefore have that f is the step function

$$f(t) = \begin{cases} 0 & \text{if } -1 \leqslant t \leqslant 0 \\ 1 & \text{if } \quad 0 < t \leqslant 1, \end{cases}$$

which contradicts that f is continuous. ■

We therefore see that the space $C[a, b]$ is not complete, but it does have a completion. Let $\mathbb{C}^{[a,b]}$ be the set of functions from $[a, b]$ to $\mathbb{C}$. Then the completion of $C[a, b]$ is isomorphic to the Lebesgue space

$$L^2[a, b] \stackrel{\text{def}}{=} \{f \in \mathbb{C}^{[a,b]} : f \text{ is Lebesgue measurable and } \textstyle\int_a^b |f(t)|^2 \, dt < \infty\}/\sim,$$

where the equivalence relation $\sim$ is defined by

$$f \sim g \iff \int_a^b |f(t) - g(t)|^2 \, dt = 0,$$

with inner product $\langle f, g \rangle = \int_a^b \overline{f} g$.[1]

Theorem 1.15. *Let $(X, \langle \cdot\,, \cdot \rangle_X)$ be an inner product space. Then there exists a Hilbert space $(H, \langle \cdot\,, \cdot \rangle_H)$ and a linear map $T \colon X \to H$ such that:*

(i) T is one-to-one,

(ii) for all $x, y \in X$, $\langle x, y \rangle_X = \langle Tx, Ty \rangle_H$, and

(iii) the image $T(X) = \{Tx : x \in X\}$ is dense in H (i.e., for all $\epsilon > 0$ and $x \in H$, $B(x; \epsilon) \cap T(x) \neq \emptyset$.)

[1] Note that the equivalence relation is important, because otherwise we have things like $f \in L^2[0, 1]$ where $f(x) \stackrel{\text{def}}{=} 2x - 1$, having $\|f\| = 0$ but $f \neq 0$.

2 Orthogonality

Before we go to the main topic of this section, we review some terminology and concepts from vector spaces and metric spaces. Let X be an inner product space over the field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Then X is a metric space with a metric induced from the inner product.

- A subset $A \subseteq X$ which is also a vector space with respect to the inherited addition and scalar multiplication operations is said to be a *linear subspace* of X.
- A (*finite*) *linear combination* is a sum of the form $\alpha_1 x_1 + \cdots + \alpha_n x_n$, also written $\sum_{i=1}^{n} \alpha_i x_i$, where $x_i \in X$ and $\alpha_i \in \mathbb{F}$ for $1 \leqslant i \leqslant n$.
- The (*linear*) *span* of a subset $A \subseteq X$, denoted by $\operatorname{span}(A)$, is the set of all finite linear combinations of points in A, i.e.,

  $$\operatorname{span}(A) \stackrel{\text{def}}{=} \{\alpha_1 x_1 + \cdots + \alpha_n x_n : \alpha_i \in \mathbb{F}, x_i \in A\}.$$

  By convention, $\operatorname{span}(\emptyset) \stackrel{\text{def}}{=} \{0\}$.

  For any $A \subseteq X$, $\operatorname{span}(A)$ forms a linear subspace of X; moreover, it is the smallest linear subspace of X which contains A. We say that A *spans* or *generates* $\operatorname{span}(A)$.
- The subset $A \subseteq X$ is linearly independent when no point $x \in A$ is a finite linear combination of the others, i.e., if $x \notin \operatorname{span}(A \setminus \{x\})$. This is equivalent to the following: for any finite subset $\{x_1, \ldots, x_n\} \subseteq A$,

  $$\alpha_1 x_1 + \cdots + \alpha_n x_n = 0 \implies \alpha_i = 0 \text{ for all } i = 1, \ldots, n.$$

- A point $x \in X$ is called *adherent* to A if there is a sequence (x_n) in A such that $x_n \to x$.
- A subset $A \subseteq X$ is *closed* if it contains all its adherent points, i.e., A is closed if whenever (x_n) is a sequence in A such that $x_n \to x$, then $x \in A$.
- If $A \subseteq X$, the closure of A in X, denoted by $\bar{A}$, is defined by

  $$\bar{A} \stackrel{\text{def}}{=} \{x \in X : x \text{ is adherent to } A\}.$$

  $\bar{A}$ is the smallest closed set containing A.
- If $\bar{A} = X$, then A is *dense* in X. Equivalently, A is dense in X if for all $\epsilon > 0$ and $x \in X$, $A \cap B(x; \epsilon) \neq \emptyset$.

Proposition 2.1. *Let X be an inner product space. Then*

(i) If A is a linear subspace of X, so is $\bar{A}$.

(ii) For any subset $A \subseteq X$, $\overline{\operatorname{span}(A)}$ is the smallest closed subspace of X that contains A.

Proof. For (i), it suffices to prove closure, since all the remaining vector space properties are hereditary. Indeed, let $x, y \in \bar{A}$, and $\alpha \in \mathbb{F}$. Then there are sequences (x_n), (y_n) in A such that $x_n \to x$ and $y_n \to y$. Since A is a linear subspace, $x_n + \alpha y_n \in A$ for all n, and it is not hard to see that $x_n + \alpha y_n \to x + \alpha y$, from which it follows that $x + \alpha y \in \bar{A}$.

For (ii), we clearly have that $\overline{\operatorname{span}(A)}$ is a closed linear subspace which contains A by (i). Now suppose Y is a closed linear subspace containing A. Since Y contains A, we have that $\operatorname{span}(A) \subseteq Y$. It follows that $\overline{\operatorname{span}(A)} \subseteq \overline{Y} = Y$, since Y is closed. □

2.1 Definition and Basic Properties

Recall that in geometry, two vectors $u, v \in \mathbb{F}^n$ are said to be perpendicular or orthogonal if $u \cdot v = 0$.

Definition 2.2 (Orthogonal space)**.** Let X be an inner product space. If $\langle x, y \rangle = 0$, we say that x is *orthogonal* or *perpendicular* to y.

For $A \subseteq X$, we define the *orthogonal space* of A, denoted by $A^\perp$, by

$$A^\perp \stackrel{\text{def}}{=} \{x \in X : \forall\, y \in A, \langle x, y \rangle = 0\}.$$

Example 2.3. Let $X \subseteq C[0,1]$ be the set of quadratic polynomial functions on $[0,1]$, i.e.,

$$X = \{\alpha x^2 + \beta x + \gamma : \alpha, \beta, \gamma \in \mathbb{C}\},$$

and let $A = \operatorname{span}\{ix^2\}$. What is $A^\perp$? If $p \in A^\perp$ and $p(x) = \alpha x^2 + \beta x + \gamma$, then

$$\langle ix^2, \alpha x^2 + \beta x + \gamma \rangle = \int_0^1 \overline{ix^2}(\alpha x^2 + \beta x + \gamma)\,dx = -i(\tfrac{\alpha}{5} + \tfrac{\beta}{4} + \tfrac{\gamma}{3}) = 0,$$

so $p(x) = \alpha(x^2 - {}^3\!/\!_5) + \beta(x - {}^3\!/\!_4)$ for $\alpha, \beta \in \mathbb{C}$, i.e.,

$$A^\perp = \operatorname{span}\{x^2 - {}^3\!/\!_5, x - {}^3\!/\!_4\}.$$

Proposition 2.4 (Basic Properties)**.** *Let X be an inner product space, and let $A, B \subseteq X$. Then*

(i) $A \cap A^\perp \subseteq \{0\}$,

(ii) $A^\perp$ *is a closed linear subspace of* X,

(iii) $X^\perp = \{0\}$,

(iv) $\{x\}^\perp = X$ *if and only if* $x = 0$,

(v) $A \subseteq A^{\perp\perp}$,

(vi) *if* $A \subseteq B$, *then* $B^\perp \subseteq A^\perp$,

(vii) $A^\perp = (\operatorname{span} A)^\perp = (\overline{\operatorname{span} A})^\perp$.

Proof. For (i), let $x \in A \cap A^\perp$. Then $x \in A^\perp$, so $\langle x, a\rangle = 0$ for all $a \in A$. But $x \in A$ also, so we may take $a = x$ to get $\langle x, x\rangle = 0$, i.e., $x = 0$.

For (ii), first let $a \in A$. Now let $x, y \in A^\perp$ and $\alpha \in \mathbb{F}$. Then $\langle x + \alpha y, a\rangle = \langle x, a\rangle + \bar{\alpha}\langle y, a\rangle = 0$, so $x + \alpha y \in A^\perp$ and therefore $A^\perp$ is a linear subspace. Now for closure, let (x_n) be a sequence in $A^\perp$, and suppose $x_n \to x$. Then by proposition 1.12, $\langle x, a\rangle = \langle \lim x_n, \lim a\rangle = \lim\langle x_n, a\rangle = \lim 0 = 0$, so $x \in A^\perp$, therefore $A^\perp$ is closed.

(iii) follows immediately from proposition 1.3(ii).

Now for (iv), observe that $\{x\}^\perp = X \Leftrightarrow \langle x, y\rangle = 0$ for all $y \in X \Leftrightarrow x = 0$ again by proposition 1.3(ii).

Next for (v), let $x \in A$. Then for all $y \in A^\perp$, $\langle x, y\rangle = 0$, i.e., $x \in A^{\perp\perp}$.

For (vi), suppose $A \subseteq B$, and let $x \in B^\perp$. Then $\langle x, y\rangle = 0$ for all $y \in B$, and in particular, for all $y \in A$. It follows that $x \in A^\perp$.

Finally for (vii), observe that $A \subseteq \operatorname{span}(A) \subseteq \overline{\operatorname{span}(A)}$, so then by (vi), we get $(\overline{\operatorname{span} A})^\perp \subseteq (\operatorname{span} A)^\perp \subseteq A^\perp$. We show that $A^\perp \subseteq (\operatorname{span} A)^\perp \subseteq (\overline{\operatorname{span} A})^\perp$. Let $x \in A^\perp$ and $y \in \operatorname{span} A$. Then $y = \sum_{i=1}^n \alpha_i a_i$ for some subset $\{a_1, \ldots, a_n\} \subseteq A$, and thus $\langle x, y\rangle = \langle x, \sum_{i=1}^n \alpha_i a_i\rangle = \sum_{i=1}^n \alpha_i\langle x, a_i\rangle = 0$, hence $x \in (\operatorname{span} A)^\perp$. Now take $x \in (\operatorname{span} A)^\perp$, and let $y \in \overline{\operatorname{span}(A)}$. By definition, there is a sequence (y_n) in $\operatorname{span} A$ such that $y_n \to y$. By proposition 1.12, $0 = \lim 0 = \lim\langle x, y_n\rangle = \langle x, \lim y_n\rangle = \langle x, y\rangle$, so that $x \in (\overline{\operatorname{span} A})^\perp$, as required. □

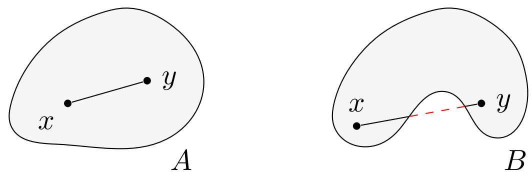

FIGURE 4: Illustration of two sets in $\mathbb{R}^2$. The set A is convex, the set B is not.

Proposition 2.5. *If A is a set of pairwise orthogonal non-zero points, i.e., $\langle x, y\rangle = 0$ for all $x, y \in A$ with $x \neq y$, then A is linearly independent.*

Proof. Let $\{x_1, \dots, x_n\} \subseteq A$, and suppose $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$. Then for all $1 \leqslant i \leqslant n$, $0 = \langle x_i, 0\rangle = \langle x_i, \alpha_1 x_1 + \cdots + \alpha_n x_n\rangle = \sum_{j=1}^{n} \alpha_j \langle x_i, x_j\rangle = \alpha_i \|x_i\|$, i.e., $\alpha_i = 0$, as required. $\square$

2.2 Closed Linear Subspaces and Least Distance

Definition 2.6 (Convexity)**.** Let X be an inner product space. A subset A of X is said to be *convex* if for any two points $x, y \in A$, and $t \in [0, 1]$, we have $tx + (1-t)y \in A$.

Figure 4 provides an illustration of convexity in $\mathbb{R}^2$.

Proposition 2.7. *Let X be an inner product space, let $a \in X$, and $r > 0$. Then the open ball $B(a; r) = \{x \in X : \|a - x\| < r\}$ is convex.*

Proof. Let $x, y \in B(a; r)$, and $t \in [0, 1]$. Then

$$\begin{aligned} \|a - tx - (1-t)y\| &= \|t(a-x) + (1-t)(a-y)\| \\ &\leqslant t\|a - x\| + (1-t)\|a - y\| \qquad \text{(triangle inequality)} \\ &< tr + (1-t)r = r, \end{aligned}$$

so $tx + (1-t)y \in B(a; r)$. $\square$

Proposition 2.8. *Let X be an inner product space, and let A be a subspace of X. Then A is convex.*

Proof. Follows immediately by closure: $x, y \in A \Rightarrow tx + (1-t)y \in A$. $\square$

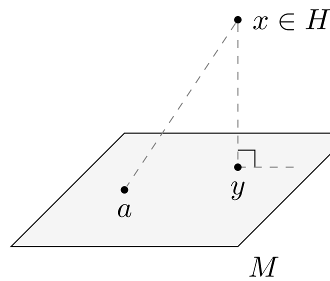

FIGURE 5: Illustration of the statement of theorem 2.9 in $\mathbb{R}^3$, with a plane as a linear subspace.

Theorem 2.9. *Let M be a non-empty, closed and convex subset of a Hilbert space H. Given $x \in H$, there exists a unique $y \in M$ which is the closest point to x, i.e., for all $a \in M$ different from y, $\|x - y\| < \|x - a\|$.*

Proof. Let $\delta = \inf\{\|x - a\| : a \in M\}$ be the least distance from M to x. Now there exists a sequence (y_n) in M such that $\|x - y_n\| \to \delta$.[2] Moreover, by the parallelogram law and the convexity of M, (y_n) is Cauchy:

$$\begin{aligned}\|y_n - y_m\|^2 &= 2(\|y_n - x\|^2 + \|y_m - x\|^2) - \|(y_n + y_m) - 2x\|^2 \\ &= 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\left\|\tfrac{y_n + y_m}{2} - x\right\|^2 \\ &\leqslant 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\delta^2 \to 0,\end{aligned}$$

as $n, m \to \infty$. Since H is a Hilbert space, $y_n \to y \in H$, since M is closed, $y \in M$, and by proposition 1.12,

$$\begin{aligned}\|x - y\|^2 = \|x - \lim y_n\|^2 = \langle x - \lim y_n, x - \lim y_n\rangle &= \lim\langle x - y_n, x - y_n\rangle \\ &= \lim \|x - y_n\|^2 = \delta^2,\end{aligned}$$

so y attains the minimum distance δ from x. Finally we show y is unique. Indeed, suppose $z \in M$ attains this minimum also, i.e., $\|z - x\| = \delta$. Then applying the parallelogram law similarly to the above, we have

$$\begin{aligned}\|z - y\|^2 &= 2\|z - x\|^2 + 2\|y - x\|^2 - 4\left\|\tfrac{y+z}{2} - x\right\|^2 \\ &\leqslant 4\delta^2 - 4\delta^2 = 0,\end{aligned}$$

so $\|z - y\|^2 = 0$, i.e., $y = z$. □

[2] By definition of infimum, for each n, there is $y_n \in M$ such that $\|x - y_n\| \leqslant \delta + 1/n$, otherwise we contradict its maximality. Using the axiom of choice, we can construct the sequence by choosing y_n from $Y_n = \{\eta \in M : \|x - \eta\| \leqslant \delta + 1/n\}$, and it is then straightforward to verify that $\|x - y_n\| \to \delta$.

Notice that the proof does not require all of H to be complete; it suffices if we are in an inner product and only M is complete:

Theorem 2.10. *Let M be a non-empty, convex and complete subset of an inner product space X. Given $x \in X$, there exists a unique $y \in M$ which is the closest point to x, i.e., for all $a \in M$ different from y, $\|x-y\| < \|x-a\|$.*

The proof is identical. Now consider the special case where M is a linear subspace of a Hilbert space H. We have the following.

Theorem 2.11. *Let M be a closed linear subspace of a Hilbert space H. Given $x \in H$ and $y \in M$, then*

$$y \text{ is the closest point in } M \text{ to } x \iff x - y \in M^{\perp}.$$

Proof. For the only if part, suppose y is the closest point in M to x, and let $a \in M$. By lemma 1.8, we may write $x-y = \lambda a+z$ where $\langle a, z\rangle = 0$. Thus by Pythagoras' theorem, $\|x-y\|^2 = \|\lambda a\|^2 + \|z\|^2 \geqslant \|z\|^2$, but $z = x-(y+\lambda a)$, so $\|x-(y+\lambda a)\|^2 \leqslant \|x-y\|^2$, i.e., $y+\lambda a$ is closer to x than y is, which is a contradiction, unless we take $\lambda = 0$. It follows that $z = x-y$, and therefore $\langle a, x-y\rangle = 0$. By the arbitrariness of a, it follows that $x-y \in M^{\perp}$.

For the converse, suppose $\langle a, x-y\rangle = 0$ for all $a \in M$. Since M is a subspace, then $y - a \in M$, so $\langle y-a, x-y\rangle = 0$. Thus by Pythagoras,

$$\|x-a\|^2 = \|x-y+y-a\|^2 = \|x-y\|^2 + \|y-a\|^2 \geqslant \|x-y\|^2,$$

which concludes the proof. □

Just as before, this result can be generalised to inner product spaces if we insist that M is complete, rather than closed:

Theorem 2.12. *Let M be a complete linear subspace of an inner product space X. Given $x \in X$ and $y \in M$, then y is the closest point in M to x if and only if $x-y \in M^{\perp}$.*

Remark 2.13. In order to avoid the duplication of further results, we draw the reader's attention to the following fact. In general, any result which makes use of the completeness of H and the closure of M can be adapted for inner product spaces if M itself is complete; so long as the proof of that result utilises the completeness of H only for sequences in M. ■

Recall. A vector space X is said to be the *direct sum* of two vector spaces Y and Z, written $X = Y \oplus Z$, if each $x \in X$ can be expressed as $x = y+z$, $y \in Y$, $z \in Z$ in a unique way.

Corollary 2.14. *If M is a closed subspace of a Hilbert space H, then*

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

Proof. Write $x = y + (x - y)$ where y is in M, and $x - y \in M^{\perp}$. It follows that y is the closest point in M to x by theorem 2.11, so this expression is unique. □

Corollary 2.15. *Let H be a Hilbert space, let M be a closed linear subspace of H, and let $A \subseteq H$ be any subset of H. Then*

(i) $M = M^{\perp\perp}$,

(ii) $\overline{\operatorname{span} A} = A^{\perp\perp}$,

(iii) $\overline{\operatorname{span} A} = H$ *if and only if* $A^{\perp} = \{0\}$.

Proof. For (i), note that we always have $M \subseteq M^{\perp\perp}$ by proposition 2.4(v), thus we simply show the converse inclusion. Indeed, let $x \in M^{\perp\perp} \subseteq H = M \oplus M^{\perp}$. Then $x = a + b$, where $a \in M, b \in M^{\perp}$, so

$$0 = \langle x, b\rangle = \langle a + b, b\rangle = \langle a, b\rangle + \langle b, b\rangle = \|b\|^2,$$

and thus $b = 0$, so $x = a \in M$.

Now (ii) follows by combining (i) and proposition 2.4(vii).

Finally for (iii), combine (ii) and proposition 2.4(iii). □

Example 2.16 (Least squares approximation). Let $H = L^2[0,1]$, let $f \in H$ be the exponential function $f(x) = e^x$, and let $M = \operatorname{span}\{1, x, x^2\}$. Then M is closed (exercise: verify). What is the closest point q in M to f?

Let $q(x) = ax^2 + bx + c$ for some $a, b, c \in \mathbb{C}$. Then $f - q \in M^{\perp}$, so we have the equations

$$\begin{cases} \langle e^x - q, 1\rangle = 0 \\ \langle e^x - q, x\rangle = 0 \\ \langle e^x - q, x^2\rangle = 0 \end{cases} \iff \begin{cases} \langle e^x, 1\rangle = \langle q, 1\rangle \\ \langle e^x, x\rangle = \langle q, x\rangle \\ \langle e^x, x^2\rangle = \langle q, x^2\rangle \end{cases}$$

$$\iff \begin{cases} \int_0^1 \overline{e^x}\, 1\, dx = \int_0^1 (\overline{ax^2 + bx + c}) 1\, dx \\ \int_0^1 \overline{e^x} x\, dx = \int_0^1 (\overline{ax^2 + bx + c}) x\, dx \\ \int_0^1 \overline{e^x} x^2\, dx = \int_0^1 (\overline{ax^2 + bx + c}) x^2\, dx \end{cases}$$

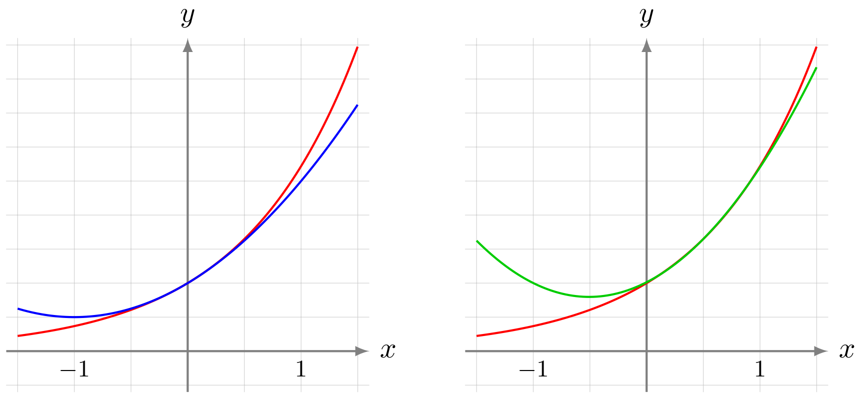

FIGURE 6: Comparison of the Maclaurin series expansion and the least squares approximation. The curve in red in both plots is $f(x) = e^x$, the curve in blue is the Maclaurin series expansion $1 + x + \frac{1}{2}x^2$, and the curve in green is the least squares approximation $1.013 + 0.851x + 0.839x^2$.

$$\iff \begin{cases} e - 1 = a/3 + b/2 + c \\ \quad\; 1 = a/4 + b/3 + c/2 \\ e - 2 = a/5 + b/4 + c/3, \end{cases}$$

which, when solved for a, b, c, yield $a = 30(7e - 19)$, $b = 12(49 - 18e)$ and $c = 3(13e - 35)$. Thus the closest point is the function

$$q(x) = 30(7e - 19)x^2 + 12(49 - 18e)x + 3(13e - 35),$$

approximately $1.013 + 0.851x + 0.839x^2$. Notice that these coefficients are close to, but not equal to, the coefficients of the Maclaurin expansion of e^x. The difference is that while the Maclaurin series expansion is accurate at 0 and progressively worsens as we get farther away, the approximation we obtained balances out the 'root-mean-square error' throughout the region $[0, 1]$. Refer to the illustration in figure 6. ■

Example 2.17. Recall the set

$$\ell^2(\mathbb{R}) = \{(x_i)_{i\in\mathbb{N}} : x_i \in \mathbb{R} \text{ and } \textstyle\sum_{i=1}^{\infty} |x_i|^2 < \infty\}$$

with inner product $\langle \boldsymbol{x}, \boldsymbol{y}\rangle \overset{\text{def}}{=} \sum_{i=1}^{\infty} x_i y_i$, and consider the subspace M defined by $M \overset{\text{def}}{=} \{(x_i) \in \ell^2 : x_1 = x_2\}$ and the sequence $\boldsymbol{y} = (1/n)_{n\in\mathbb{N}} \in \ell^2$. What is the closest point to $\boldsymbol{y}$ in M?

We first prove that M is closed. Let $\boldsymbol{x}^n = (x_i^n)_{i \in \mathbb{N}}$ be a sequence in M, and suppose it converges to some $\boldsymbol{x} \in \ell^2$:

$$\begin{aligned} \boldsymbol{x}^1 &= (x_1^1, \quad x_2^1, \quad x_3^1, \quad \dots) \\ \boldsymbol{x}^2 &= (x_1^2, \quad x_2^2, \quad x_3^2, \quad \dots) \\ &\vdots \\ \boldsymbol{x} \;\; &= (x_1, \quad x_2, \quad x_3, \quad \dots) \end{aligned}$$

By proposition 1.13, we have $x_1^n \to x_1$ and $x_2^n \to x_2$, and since $x_1^i = x_2^i$ for all i, it follows that $x_1 = x_2$, so $\boldsymbol{x} \in M$, proving closure.

Now write

$$\boldsymbol{y} = (1/n) = \boldsymbol{a} + \boldsymbol{b},$$

where $\boldsymbol{a} \in M$ is the closest point in M to $\boldsymbol{y}$, and $\boldsymbol{b} \in M^\perp$ by corollary 2.14. Since $\boldsymbol{a} \in M$, then $\boldsymbol{a} = (a_1, a_1, a_3, \dots)$, and since $\boldsymbol{b} \in M^\perp$, then $\langle \boldsymbol{b}, \boldsymbol{c} \rangle = 0$ for all $\boldsymbol{c} \in M$. In particular, take $\boldsymbol{c}_1 = (1, 1, 0, 0, 0, \dots) \in M$. Then

$$0 = \langle \boldsymbol{b}, \boldsymbol{c}_1 \rangle = \langle (b_1, b_2, b_3, \dots), (1, 1, 0, \dots) \rangle = b_1 + b_2,$$

which implies that $b_2 = -b_1$. Next, take $\boldsymbol{c}_2 = (0, 0, b_3, b_4, \dots) \in M$. Then

$$0 = \langle \boldsymbol{b}, \boldsymbol{c}_2 \rangle = \langle (b_1, b_2, b_3, \dots), (0, 0, b_3, \dots) \rangle = \sum_{i=3}^{\infty} {b_i}^2 = \|\boldsymbol{c}_2\|^2,$$

from which it follows that $b_i = 0$ for all $i \geqslant 3$. We therefore have that

$$M^\perp = \{(b_1, -b_1, 0, 0, \dots) : b_1 \in \mathbb{R}\}.$$

Thus writing $\boldsymbol{y}$ as $\boldsymbol{a} + \boldsymbol{b}$, i.e.,

$$(1, 1/2, 1/3, \dots) = (a_1, a_1, a_2, a_3 \dots) + (b_1, -b_1, 0, 0, \dots),$$

it follows that $a_1 = 3/4 = a_2$, $b_1 = 1/4$, and $a_n = 1/n$ for $n \geqslant 3$, i.e.,

$$(1, 1/2, 1/3, \dots) = (3/4, 3/4, 1/3, \dots) + (1/4, -1/4, 0, 0, \dots).$$

Therefore, the point closest to $\boldsymbol{y}$ in M is $\boldsymbol{a} = (3/4, 3/4, 1/3, \dots)$. ■

2.3 Orthogonal Projections

Recall. Let $f\colon X \to Y$ be a map between the vector spaces X and Y. The *kernel* of f, denoted by $\ker f$, is the subset

$$\ker f \stackrel{\text{def}}{=} \{x \in X : f(x) = 0\}$$

of X, and the *image* of f, denoted by $\operatorname{im} f$, is the subset

$$\operatorname{im} f \stackrel{\text{def}}{=} \{f(x) : x \in X\}$$

of Y. For any linear map f, $\ker f$ and $\operatorname{im} f$ are linear subspaces of X and Y respectively.

Definition 2.18 (Orthogonal projection)**.** Let X be an inner product space. A *projection* in X is a linear map $P\colon X \to X$ such that $P^2 = P$, i.e., such that $(P \circ P)(x) = P(x)$ for all $x \in X$.

A projection P is an *orthogonal projection* if $\ker P \perp \operatorname{im} P$.

Suppose M is a closed subspace of a Hilbert space H, and define the map $P_M\colon H \to H$ by $P_M(x) = y$, where y is the closest point in M to x. Then we have the following.

Proposition 2.19. *Let M be a closed linear subspace of a Hilbert space H. Then P_M is an orthogonal projection.*

Proof. First, we show P_M is linear. Let $x, y \in H$, and $\alpha \in \mathbb{F}$. Then $x - P_M(x) \in M^\perp$ by theorem 2.11, and similarly $y - P_M(y) \in M^\perp$. Since $M^\perp$ is a linear subspace, then the sum

$$(x - P_M(x)) + \alpha(y - P_M(y)) = (x + \alpha y) - (P_M(x) + \alpha P_M(y)) \in M^\perp,$$

so by theorem 2.11, $P_M(x + \alpha y) = P_M(x) + \alpha P_M(y)$, proving that P_M is linear.

Note that if $x \in M$, then $P_M(x) = x$. Consequently, $P_M^2(x) = P_M(x)$, so P_M is a projection.

Finally, note that $\operatorname{im} P_M = M$ (since for $y \in M$, $P_M(y) = y$), and that

$$x \in \ker P_M \iff P_M(x) = 0 \iff x = x - P_M(x) \in M^\perp,$$

so that $\ker P_M = M^\perp$, and therefore $\ker P_M \perp \operatorname{im} P_M$. $\square$

Another property about the map P_M is the following.

Proposition 2.20. *Let M be a closed linear subspace of a Hilbert space H. Then P_M is (Lipschitz) continuous, and for all $x, y \in H$,*

$$\langle P_M(x), y\rangle = \langle x, P_M(y)\rangle.$$

Proof. Let $x \in H$. Then $x = P_M(x) + (x - P_M(x)) \in M \oplus M^\perp$, and so by Pythagoras, $\|x\|^2 = \|P_M(x)\|^2 + \|x - P_M(x)\|^2$. In particular, it follows that $\|P_M(x)\| \leqslant \|x\|$, and therefore $\|P_M(x) - P_M(y)\| = \|P_M(x-y)\| \leqslant \|x-y\|$, so that P_M is Lipschitz.

Now observe that

$$\begin{aligned}\langle P_M(x), y\rangle &= \langle P_M(x), P_M(y) + (y - P_M(y))\rangle \\ &= \langle P_M(x), P_M(y)\rangle + \langle \underbrace{P_M(x)}_{\in M}, \underbrace{y - P_M(y)}_{\in M^\perp}\rangle = \langle P_M(x), P_M(y)\rangle,\end{aligned}$$

and similarly $\langle x, P_M(y)\rangle = \langle P_M(x), P_M(y)\rangle$. $\square$

2.4 Orthonormal Sets

Definition 2.21 (Orthonormal set)**.** Let X be an inner product space, and let $S \subseteq X$. Then S is said to be *orthonormal* if for any $x, y \in S$, we have:

(i) $\langle x, y\rangle = 0$ for $x \neq y$, (Orthogonality)

(ii) $\langle x, x\rangle = \|x\|^2 = 1$. (Normality)

If we have (i) only, we say then S is said to be *orthogonal*.

Examples 2.22. (i) In ℓ^2, the sequence $(\boldsymbol{e}_n)$ where $\boldsymbol{e}_n = (\delta_{in} : i = 1, 2, \dots)$ and $\delta_{in} = 1$ if $i = n$, and 0 otherwise, forms an orthonormal set:

$$\begin{aligned}\boldsymbol{e}_1 &= (1, 0, 0, 0, \dots)\\ \boldsymbol{e}_2 &= (0, 1, 0, 0, \dots)\\ \boldsymbol{e}_3 &= (0, 0, 1, 0, \dots)\\ &\vdots\end{aligned}$$

(ii) Let $X = C[-\pi, \pi]$, and define $x_n(t) = \cos(nt)$ and $y_n(t) = \sin(nt)$. Then the set $\{x_0, y_1, x_1, y_2, x_2 \dots\}$ is orthogonal. Indeed, it is straightforward to check that for integers $n, m \geqslant 1$,

$$\|x_0\|^2 = \int_{-\pi}^{\pi} dt = 2\pi, \qquad \langle x_n, y_m \rangle^2 = \int_{-\pi}^{\pi} \cos(nt)\sin(mt)\, dt = 0,$$

$$\int_{-\pi}^{\pi} \cos(nt)\cos(mt)\, dt = \int_{-\pi}^{\pi} \sin(nt)\sin(mt)\, dt = \begin{cases} \pi & \text{if } n = m \\ 0 & \text{otherwise.} \end{cases}$$

Hence $\{\frac{1}{\sqrt{2\pi}}, \frac{\cos t}{\sqrt{\pi}}, \frac{\sin t}{\sqrt{\pi}}, \frac{\cos 2t}{\sqrt{\pi}}, \frac{\sin 2t}{\sqrt{\pi}}, \dots\}$ is an orthonormal set. ■

Proposition 2.23 (Gram–Schmidt orthogonalisation)**.** *Let X be an inner product space, and let $S = \{u_1, u_2, \dots\} \subseteq X$ be a countable (possibly infinite) subset. Define*

$$v_1 \stackrel{\text{def}}{=} u_1, \qquad\qquad e_1 \stackrel{\text{def}}{=} v_1/\|v_1\|,$$
$$v_n \stackrel{\text{def}}{=} u_n - \sum_{i=1}^{n-1} \langle e_i, u_n \rangle e_i, \qquad e_n \stackrel{\text{def}}{=} v_n/\|v_n\|,$$

and for any n, if $v_n = 0$, discard u_n and v_n, and relabel u_{n+1}, v_{n+1} to u_n, v_n respectively. Then the set $B = \{e_1, e_2, \dots\}$ is an orthonormal set with $\operatorname{span}(B) = \operatorname{span}(S)$.

Proof. It is clear that $\|e_n\| = 1$ for all n by construction. We first show that $\langle e_n, e_m \rangle = 0$ for $n \neq m$ (w.l.o.g. $n > m$) by induction on n. When $n = 1$, this is vacuously true. Next,

$$\begin{aligned}
\langle e_n, e_m \rangle &= \tfrac{1}{\|v_n\|} \langle u_n - \sum_{i=1}^{n-1} \langle e_i, u_n \rangle e_i, e_m \rangle \\
&= \tfrac{1}{\|v_n\|} \langle u_n, e_m \rangle - \tfrac{1}{\|v_n\|} \sum_{i=1}^{n-1} \overline{\langle e_i, u_n \rangle} \langle e_i, e_m \rangle \\
&= \tfrac{1}{\|v_n\|} \langle u_n, e_m \rangle - \tfrac{1}{\|v_n\|} \overline{\langle e_m, u_n \rangle} \langle e_m, e_m \rangle \qquad \text{(induction hypothesis)} \\
&= \tfrac{1}{\|v_n\|} \langle u_n, e_m \rangle - \tfrac{1}{\|v_n\|} \langle u_n, e_m \rangle = 0,
\end{aligned}$$

Thus B is an orthonormal set. Now we show that $\operatorname{span}(B) = \operatorname{span}(S)$. Indeed, it is clear that $\operatorname{span}(B) \subseteq \operatorname{span}(S)$, since each e_n is a combination of u_i's. For the converse, let $x \in \operatorname{span}(S)$. Then there is a finite subset $I \subseteq \mathbb{N}$ such that $x = \sum_{i \in I} \alpha_i u_i$ and $\alpha_i \in \mathbb{F}$ for all $i \in I$. We proceed by induction on $i' \stackrel{\text{def}}{=} \max I$. Indeed, if $i' = 1$, then $x = \alpha_1 u_1 = \alpha_1 \|v_1\|\, e_1 \in \operatorname{span}(B)$.

Next, if $i' = n \geqslant 2$, then

$$x = \alpha_n u_n + \sum_{\substack{i \in I \\ i \neq n}} \alpha_i u_i = \alpha_n \Bigg(\|v_n\| e_n + \sum_{i=1}^{n-1} \langle e_i, u_n \rangle e_i \Bigg) + y,$$

where $y = \sum_{\substack{i \in I \\ i \neq n}} \alpha_i u_i \in \operatorname{span}(B)$ by the induction hypothesis. It follows that $x \in \operatorname{span}(B)$, which completes the proof. $\square$

Example 2.24. We orthogonalise the set $\{1, x, x^2\}$ in $L^2[0,1]$.

Indeed, we have $u_1 = v_1 = 1$, and $\|v_1\| = \int_0^1 dx = 1$, so $e_1 = v_1/\|v_1\| = 1$.

Next, we have $u_2 = x$, so $v_2 = x - \langle 1, x \rangle 1$, where $\langle x, 1 \rangle = \int_0^1 x\,dx = \frac{1}{2}$, thus $v_2 = x - \frac{1}{2} \cdot 1 = x - \frac{1}{2}$. Also, $\|x - \frac{1}{2}\|^2 = \int_0^1 (x - \frac{1}{2})^2\,dx = \frac{1}{12}$, and therefore $e_2 = v_2/\|v_2\| = 2\sqrt{3}(x - \frac{1}{2})$.

Finally we have $u_3 = x^2$, so

$$v_3 = x^2 - \langle 1, x^2 \rangle 1 - \langle 2\sqrt{3}(x - \tfrac{1}{2}), x^2 \rangle\, 2\sqrt{3}(x - \tfrac{1}{2}),$$

$\langle 1, x^2 \rangle = \int_0^1 x^2\,dx = \frac{1}{3}$, and $\langle 2\sqrt{3}(x - \frac{1}{2}), x^2 \rangle = \int_0^1 2\sqrt{3}(x - \frac{1}{2}) x^2\,dx = \frac{1}{2\sqrt{3}}$, and hence

$$u_3 = x^2 - \tfrac{1}{3} \cdot 1 - \tfrac{1}{2\sqrt{3}} \cdot 2\sqrt{3}(x - \tfrac{1}{2}) = \tfrac{1}{6} - x + x^2.$$

Also, $\|u_3\|^2 = \int_0^1 (\frac{1}{6} - x + x^2)^2\,dx = \frac{1}{180}$, so $e_3 = u_3/\|u_3\| = 6\sqrt{5}(\frac{1}{6} - x + x^2)$.

Therefore the set

$$B = \{1, 2\sqrt{3}(x - \tfrac{1}{2}), 6\sqrt{5}(\tfrac{1}{6} - x + x^2)\}$$

is an orthonormal set with the same span as $\{1, x, x^2\}$. ■

Theorem 2.25. *Let X be an inner product space, and let $\{e_1, \ldots, e_n\} \subseteq X$ be a finite orthonormal set. Then for any $x \in X$, we have*

$$\Big\|x - \sum_{i=1}^n \langle e_i, x \rangle e_i\Big\|^2 = \|x\|^2 - \sum_{i=1}^n |\langle e_i, x \rangle|^2.$$

Proof. Let $\lambda_i \in \mathbb{C}$ for $i = 1, \ldots, n$. Then

$$\Big\|x - \sum_{i=1}^n \lambda_i e_i\Big\|^2 = \Big\langle x - \sum_{i=1}^n \lambda_i e_i, x - \sum_{i=1}^n \lambda_i e_i \Big\rangle$$

$$= \|x\|^2 - \sum_{i=1}^{n} \bar{\lambda}_i \langle e_i, x\rangle - \sum_{i=1}^{n} \lambda_i \langle x, e_i\rangle + \sum_{i=1}^{n} \bar{\lambda}_i \lambda_i.$$

If $\lambda_i = \langle e_i, x\rangle$, then we have

$$\begin{aligned}\Big\|x - \sum_{i=1}^{n} \lambda_i e_i\Big\|^2 &= \|x\|^2 - \sum_{i=1}^{n} \bar{\lambda}_i \lambda_i - \sum_{i=1}^{n} \lambda_i \bar{\lambda}_i + \sum_{i=1}^{n} \bar{\lambda}_i \lambda_i \\ &= \|x\|^2 - \sum_{i=1}^{n} |\lambda_i|^2,\end{aligned}$$

as required. □

Corollary 2.26 (Bessel's inequality)**.** *Let X be an inner product space, and let $\{e_1, e_2, \dots\}$ be a countably infinite orthonormal set. Then for all $x \in X$,*

$$\sum_{i=1}^{\infty} |\langle e_i, x\rangle|^2 \leqslant \|x\|^2.$$

Proof. For any $n \in \mathbb{N}$, we have

$$0 \leqslant \Big\|x - \sum_{i=1}^{n} \langle e_i, x\rangle e_i\Big\|^2 = \|x\|^2 - \sum_{i=1}^{n} |\langle e_i, x\rangle|^2 \implies \sum_{i=1}^{n} |\langle e_i, x\rangle|^2 \leqslant \|x\|^2$$

by theorem 2.25. Letting $n \to \infty$ gives the result. □

Remark 2.27. In particular, $\sum_{i=1}^{\infty} |\langle e_i, x\rangle|^2$ converges. Consequently,

$$\langle e_n, x\rangle \to 0$$

as $n \to \infty$. ■

Example 2.28 (Application of Bessel's inequality)*.* We show that

$$\lim_{n\to\infty} \int_0^{\pi} \ln x \sin(nx)\, dx = 0.$$

This is an interesting application because the integral $\int_0^{\pi} \ln x \sin(nx)\, dx$ has no closed-form expression in n.

Note that $\{\sin(nx) : n = 1, 2, \dots\}$ is an orthogonal set in $L^2[0, \pi]$.[3] In particular, $\|\sin(nx)\| = \sqrt{\pi/2}$ for all n, so $\{\sqrt{2/\pi}\sin(n\pi) : n = 1, 2, \dots\}$ is an orthonormal set.

[3] Verification is similar to examples 2.22(ii).

Now we also have $\ln x \in L^2[0,\pi]$, and

$$\begin{aligned}\|\ln x\|^2 = \int_0^\pi (\ln x)^2\,dx = \lim_{a\to 0}\int_a^\pi (\ln x)^2\,dx \\ &= \pi + \pi(\ln \pi - 1)^2 - \lim_{a\to 0}\left(a + a(\ln a - 1)^2\right) \\ &= \pi + \pi(\ln\pi - 1)^2\end{aligned}$$

by L'Hôpital's rule. Thus by Bessel's inequality,

$$\sum_{i=1}^\infty \sqrt{2/\pi}\langle \ln x, \sin(nx)\rangle^2 \leqslant \|\ln x\|^2,$$

and in particular, convergence of this series implies that the inner product $\langle \ln x, \sin(nx)\rangle^2 = \int_0^\pi \ln x \sin(nx)\,dx$ goes to 0 as $n \to \infty$. ∎

Now we investigate when equality in Bessel's inequality is attained.

Definition 2.29 (Complete Orthonormal Set)**.** Let H be a Hilbert space. A *complete orthonormal set* $S \subseteq H$ is an orthonormal set such that $S^\perp = \{0\}$.

If H has a countable complete orthonormal set S, then H is said to be *separable*, and S is called an *orthonormal basis* or a *Hilbert basis*.

Theorem 2.30. *Let H be a Hilbert space and let $S = \{e_1, e_2, \dots\} \subseteq H$ be a countably infinite orthonormal sequence. Then the following are equivalent.*

(i) S is complete (i.e., $S^\perp = \{0\}$),

(ii) For all $x \in H$, $x = \sum_{n=1}^\infty \langle e_n, x\rangle e_n$,

(iii) $\operatorname{span}(S)$ is dense in H.

Proof. (i) ⇒ (ii). Let $x \in H$, and let $y = x - \sum_{n=1}^\infty \langle e_n, x\rangle e_n$. We show that $y = 0$, i.e., that $y \in S^\perp$. Fix $m \in \mathbb{N}$. Then

$$\begin{aligned}\langle e_m, y\rangle &= \langle e_m, x\rangle - \langle e_m, \textstyle\sum_{n=1}^\infty \langle e_n, x\rangle e_n\rangle \\ &= \langle e_m, x\rangle - \lim_{k\to\infty}\langle e_m, \textstyle\sum_{n=1}^k \langle e_n, x\rangle e_n\rangle && \text{(proposition 1.12)} \\ &= \langle e_m, x\rangle - \lim_{k\to\infty}\sum_{n=1}^k \langle e_n, x\rangle\langle e_m, e_n\rangle \\ &= \langle e_m, x\rangle - \langle e_m, x\rangle\langle e_m, e_m\rangle && (\langle e_m, e_n\rangle = 0 \text{ unless } m = n)\end{aligned}$$

$$= \langle e_m, x\rangle - \langle e_m, x\rangle = 0,$$

as required.

(ii) ⇒ (iii). Let $x \in H$. Then $x = \sum_{n=1}^{\infty} \langle e_n, x\rangle e_n$. We show that $x \in \overline{\operatorname{span} S}$. Let $x_k = \sum_{n=1}^{k} \langle e_n, x\rangle e_n \in \operatorname{span}(S)$ for all k. Clearly $x_k \to x$, as required.

(i) ⇔ (iii). This is the assertion of corollary 2.15(iii). □

Example 2.31. The set $S = \{\boldsymbol{e}_n : n = 1, 2, \dots\}$ in ℓ^2 is orthonormal (see examples 2.22(i)). It is also complete. Indeed, if $\boldsymbol{y} \in \ell^2$ and $\langle \boldsymbol{y}, \boldsymbol{e}_n\rangle = 0$ for all n, it follows that $\sum_{i=1}^{\infty} \bar{y}_i \delta_{in} = \bar{y}_n = 0$ for all n, i.e., $\boldsymbol{y} = (0, 0, \dots)$. Hence $S^\perp = \{0\}$, and S is therefore an orthonormal basis for ℓ^2. ■

Theorem 2.32 (Parserval's Identity)**.** *Let H be a Hilbert space, and let $\{e_1, e_2, \dots\}$ be an orthonormal basis. Then for all $x \in H$,*

$$\sum_{i=1}^{\infty} |\langle e_i, x\rangle|^2 = \|x\|^2.$$

Proof. Let $x, y \in H$. By theorem 2.30(ii), $x = \sum_{n=1}^{\infty} \langle e_n, x\rangle e_n$, and similarly $y = \sum_{n=1}^{\infty} \langle e_n, y\rangle e_n$. Thus

$$\begin{aligned}
\langle x, y\rangle &= \Big\langle \lim_{k\to\infty} \sum_{n=1}^{k} \langle e_n, x\rangle e_n, \lim_{k\to\infty} \sum_{n=1}^{k} \langle e_n, y\rangle e_n \Big\rangle \\
&= \lim_{k\to\infty} \Big\langle \sum_{n=1}^{k} \langle e_n, x\rangle e_n, \sum_{n=1}^{k} \langle e_n, y\rangle e_n \Big\rangle \\
&= \lim_{k\to\infty} \sum_{n=1}^{k} \overline{\langle e_n, x\rangle} \langle e_n, y\rangle \langle e_n, e_n\rangle \qquad \text{(by orthonormality[4])} \\
&= \sum_{n=1}^{\infty} \overline{\langle e_n, x\rangle} \langle e_n, y\rangle,
\end{aligned}$$

and if we put $x = y$, then

$$\|x\|^2 = \sum_{n=1}^{\infty} \overline{\langle e_n, x\rangle} \langle e_n, x\rangle = \sum_{n=1}^{\infty} |\langle e_n, x\rangle|^2,$$

as required. □

[4] Here we use the general fact that $\langle \sum_{i=1}^{n} \alpha_i u_i, \sum_{i=1}^{m} \beta_i v_i\rangle = \sum_{i=1}^{n} \sum_{j=1}^{m} \bar{\alpha}_i \beta_j \langle u_i, v_j\rangle$ which follows from linearity, and is straightforward to verify.

Note. Recall that $L^2[a,b]$ denotes the space of square (Lebesgue) integrable functions on $[a,b]$. A set $A \subseteq \mathbb{R}$ has *Lebesgue measure zero* (written $\mu(A) = 0$) if for all $\epsilon > 0$, there exists a set of intervals $\{I_n : n \in \mathbb{N}\}$ such that $A \subseteq \bigcup_{n=0}^{\infty} I_n$ and $\sum_{n=0}^{\infty} \ell(I_n) < \epsilon$.[5] Examples of sets with measure zero include any countable subset of $\mathbb{R}$, including $\mathbb{Q}$. The Cantor set $\mathscr{C}$ also as Lebesgue measure zero.

$L^2[a,b]$ is a Hilbert space with inner product $\langle f, g\rangle = \int_a^b \overline{f} g$. It is easy to verify that it is an inner product space; however we do not give a proof that $L^2[a,b]$ is complete. ■

Theorem 2.33. *The set of functions*

$$\left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots\right\}$$

is an orthonormal basis for $L^2[-\pi,\pi]$.

We do not give a proof of theorem 2.33. We also have the the following.

Theorem 2.34. *The set of functions*

$$\left\{\frac{e^{inx}}{\sqrt{2\pi}} : n \in \mathbb{Z}\right\}$$

is an orthonormal basis for $L^2[-\pi,\pi]$. *More generally, the set*

$$\left\{\frac{e^{2\pi i n x/(b-a)}}{\sqrt{b-a}} : n \in \mathbb{Z}\right\}$$

is an orthonormal basis for $L^2[a,b]$.

Proof. We prove the statement for $L^2[-\pi,\pi]$, the corresponding statement for $L^2[a,b]$ is proved similarly. That $B = \{\frac{e^{inx}}{\sqrt{2\pi}} : n \in \mathbb{Z}\}$ is an orthonormal set is easily verified, and is left as an exercise. We prove that B is complete.

Recall that $e^{inx} = \cos nx + i \sin nx$. Let $f \in L^2[a,b]$. Decomposing f into $\Re f + i\Im f$, by theorem 2.33 we have

$$(\Re f)(x) = \frac{\alpha_0}{\sqrt{2\pi}} + \frac{1}{\sqrt{\pi}} \sum_{n=1}^{\infty} (\alpha_n \cos nx + \beta_n \sin nx),$$

[5] Here $\ell(I)$ denotes the *length* of the interval I. Given an interval $I = [a,b]$ (or (a,b), $[a,b)$, $(a,b]$), its length is defined by $\ell(I) \stackrel{\text{def}}{=} b - a$.

where $\alpha_0 = \langle \frac{1}{\sqrt{2\pi}}, \Re f\rangle$, $\alpha_n = \langle \frac{\cos nx}{\sqrt{\pi}}, \Re f\rangle$ and $\beta_n = \langle \frac{\sin nx}{\sqrt{\pi}}, \Re f\rangle$ for $n \geqslant 1$, and similarly

$$(\Im f)(x) = \frac{\alpha_0'}{\sqrt{2\pi}} + \frac{1}{\sqrt{\pi}} \sum_{n=1}^{\infty} (\alpha_n' \cos nx + \beta_n' \sin nx)$$

where α_0', α_n' and β_n' are defined similarly. Substituting $\cos nx = \frac{e^{inx}+e^{-inx}}{2}$ and $\sin nx = \frac{e^{inx}-e^{-inx}}{2i}$ into $f = \Re f + i \Im f$, we obtain f as a single series

$$f(x) = \sum_{n\in\mathbb{Z}} c_n e^{inx}.$$

Hence

$$\begin{aligned}
\langle e^{imx}, f\rangle &= \langle e^{imx}, \textstyle\sum_{n\in\mathbb{Z}} c_n e^{inx}\rangle \\
&= \langle e^{imx}, \lim_{k\to\infty} \textstyle\sum_{n=-k}^{k} c_n e^{inx}\rangle \\
&= \lim_{k\to\infty} \langle e^{imx}, \textstyle\sum_{n=-k}^{k} c_n e^{inx}\rangle \\
&= \lim_{k\to\infty} \sum_{n=-k}^{k} c_n \langle e^{imx}, e^{inx}\rangle \\
&= c_m \langle e^{imx}, e^{imx}\rangle = 2\pi c_m,
\end{aligned}$$

since $\|e^{imx}\|^2 = 2\pi$, and therefore we get $c_m = \frac{1}{2\pi}\langle e^{imx}, f\rangle = \frac{1}{\sqrt{2\pi}}\langle \frac{e^{imx}}{\sqrt{2\pi}}, f\rangle$, so that

$$f(x) = \sum_{n\in\mathbb{Z}} \left\langle \frac{e^{inx}}{\sqrt{2\pi}}, f \right\rangle \frac{e^{inx}}{\sqrt{2\pi}},$$

and hence by theorem 2.30, we get that B is complete. □

Remark 2.35 (Fourier Series). If $f \in L^2[a,b]$, then the corresponding series representation

$$f(x) = \sum_{n\in\mathbb{Z}} \left\langle \frac{e^{2\pi inx/T}}{\sqrt{T}}, f \right\rangle \frac{e^{2\pi inx/T}}{\sqrt{T}}$$

where $T = b - a$, is called the *Fourier expansion* of f. ■

Examples 2.36. (i) Let $f \in L^2[0,1]$ be defined by $f(x) = |x - 1/2|$. We determine the Fourier expansion of f. Indeed, when $n \neq 0$,

$$\begin{aligned}
\langle e^{2\pi inx}, f\rangle &= \int_0^1 \overline{e^{2\pi inx}} |x - 1/2| \, dx \\
&= \int_0^{1/2} e^{-2\pi inx}(1/2 - x)\, dx + \int_{1/2}^{1} e^{-2\pi inx}(x - 1/2)\, dx
\end{aligned}$$

$$= \frac{1}{4n^2\pi^2}(1 + (-1)^{n+1} - \pi i n) + \frac{1}{4n^2\pi^2}(1 + (-1)^{n+1} + \pi i n)$$

$$= \frac{1 + (-1)^{n+1}}{2n^2\pi^2} = \begin{cases} 1/n^2\pi^2 & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even,} \end{cases}$$

and when $n = 0$, $\langle e^{2\pi i n x}, f\rangle = \int_0^1 |x - 1/2|\, dx = 1/4$. Therefore,

$$f(x) = \frac{1}{4} + \sum_{\substack{n\in\mathbb{Z} \\ n \text{ odd}}} \frac{1}{n^2\pi^2} e^{2\pi i n x}.$$

(ii) Consider the function $f \in L^2[-\pi, \pi]$ defined by

$$f(x) = \begin{cases} -1 & \text{if } -\pi \leqslant x \leqslant 0 \\ 1 & \text{if } \quad 0 < x \leqslant \pi. \end{cases}$$

We determine the Fourier expansion of f and then use Parserval's identity to prove that $\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}$.

Indeed, $f(x) = \sum_{n\in\mathbb{Z}} \langle \frac{e^{inx}}{\sqrt{2\pi}}, f, \rangle \frac{e^{inx}}{\sqrt{2\pi}} = \frac{1}{2\pi} \sum_{n\in\mathbb{Z}} \langle e^{inx}, f\rangle e^{inx}$, so when $n \neq 0$, we have

$$\begin{aligned} \langle e^{inx}, f\rangle &= -\int_{-\pi}^{0} e^{-inx}\, dx + \int_0^{\pi} e^{-inx}\, dx \\ &= \frac{1}{in}(1 - (-1)^n) - \frac{1}{in}((-1)^n - 1) \\ &= \frac{2}{in}(1 - (-1)^n) = \begin{cases} 4/in & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even,} \end{cases} \end{aligned}$$

and when $n = 0$, $\langle e^{inx}, f\rangle = \int_{-\pi}^{\pi} f = 0$. Therefore

$$f(x) = \sum_{\substack{n\in\mathbb{Z} \\ n \text{ odd}}} \frac{4}{\sqrt{2\pi} in} \frac{e^{inx}}{\sqrt{2\pi}}.$$

Thus by Parserval's identity,

$$\sum_{\substack{n\in\mathbb{Z} \\ n \text{ odd}}} \left| \frac{4}{\sqrt{2\pi} in} \right|^2 = \|f\|^2 = \int_{-\pi}^{\pi} |f|^2\, dx = 2\pi$$

$$\implies \sum_{\substack{n\in\mathbb{Z} \\ n \text{ odd}}} \frac{8}{\pi n^2} = 2\sum_{n=0}^{\infty} \frac{8}{\pi(2n+1)^2} = 2\pi \implies \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8},$$

as required. ■

Exercise 2.37. Let $f \in L^2[-\pi,\pi]$ be defined by

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

(i) Determine the Fourier expansion of $|\sin x| \in L^2[-\pi,\pi]$.

(ii) Using Parserval's identity, show that $\sum_{n=0}^{\infty} \frac{1}{(1-4n^2)}^2 = \frac{\pi^2}{16}$.

Finally, we revisit the problem of minimum distance.

Proposition 2.38. *If M is a closed linear subspace of a separable Hilbert space H and $\{e_1, e_2, \dots\}$ is an orthonormal basis for M, then*

$$y = \sum_{n=1}^{\infty} \langle e_n, x\rangle e_n$$

is the closest point in M to x.

Proof. Fix $x \in H$, and let $y = \sum_{n=1}^{\infty} \langle e_n, x\rangle e_n$. Then for all $m \in \mathbb{N}$,

$$\begin{aligned} \langle e_m, x-y\rangle &= \langle e_m, x\rangle - \langle e_m, \textstyle\sum_{n=1}^{\infty} \langle e_n, x\rangle e_n\rangle \\ &= \langle e_m, x\rangle - \sum_{n=1}^{\infty} \langle e_n, x\rangle \langle e_m, e_n\rangle \\ &= \langle e_m, x\rangle - \langle e_m, x\rangle \qquad (\langle e_m, e_n\rangle = 0 \text{ unless } m = n) \\ &= 0. \end{aligned}$$

Thus given $u \in M$, we may write $u = \sum_{i=1}^{\infty} \langle e_i, u_i\rangle e_i$, and therefore

$$\langle x-y, u\rangle = \sum_{n=1}^{\infty} \langle e_i, u_i\rangle \langle x-y, e_i\rangle = 0,$$

so that $x - y \in (\overline{\operatorname{span}\{e_1, e_2, \dots\}})^{\perp} = M^{\perp}$, and the result then follows by theorem 2.11. ☐

In the following example, we revisit the problem of example 2.16, making use of the proposition above.

Example 2.39. Consider the closed subspace $M = \operatorname{span}\{1, x, x^2\}$ in $L^2[0,1]$, and the function $f = e^x \in L^2[0,1]$. We determine the closest point in M to f. Indeed, from example 2.24, we have the orthonormal basis

$$B = \{e_1 = 1, \quad e_2 = 2\sqrt{3}(x - \tfrac{1}{2}), \quad e_3 = 6\sqrt{5}(\tfrac{1}{6} - x + x^2)\}$$

for M. The closest point in M to f is therefore

$$y = \langle e_1, e^x \rangle e_1 + \langle e_2, e^x \rangle e_2 + \langle e_3, e^x \rangle e_3.$$

Evaluating the necessary integrals, we get

$$\begin{aligned} y &= (e-1)\, e_1 - \sqrt{3}(e-3)\, e_2 + \sqrt{5}(7e-19)\, e_3 \\ &= 30(7e-19)x^2 + 12(49-18e) + 3(13e-35), \end{aligned}$$

which concurs with the findings of example 2.16. ■

3 Linear Operators on Hilbert Spaces

We use the term *operator* synonymously with the term map.

Definition 3.1 (Linear map)**.** Let X and Y be a vector spaces over a field $\mathbb{F}$. A map $T\colon X \to Y$ is said to be *linear* if for all $x, y \in X$ and $\alpha \in \mathbb{F}$,

$$T(x + \alpha y) = T(x) + \alpha\, T(y).$$

Notation. It is conventional to omit brackets when denoting the image of a point x under a linear operator T, writing Tx in lieu of $T(x)$. The composition notation $S \circ T$ of two linear maps S and T is also relaxed to the juxtaposition ST, and we write $S^2 = SS$, $S^3 = S(SS)$, etc.

Examples 3.2. (i) If X is an inner product space and $y \in X$ is fixed, then the map $T_y\colon X \to \mathbb{F}$ defined by $T_y x \overset{\text{def}}{=} \langle y, x \rangle$ is a linear map. Indeed,

$$T_y(x + \alpha z) = \langle y, x + \alpha z \rangle = \langle y, x \rangle + \alpha \langle y, z \rangle = T_y x + \alpha\, T_y z.$$

(ii) If $X = \{p \in C[a,b] : p \text{ is polynomial}\}$, the map $D\colon X \to X$ defined by $Dp \overset{\text{def}}{=} p'$ (i.e., the derivative of p) is a linear map.

(iii) If $X = \ell^2$, then the map $T\colon \ell^2 \to \ell^2$ defined by

$$T(x_1, x_2, x_3, \dots) \overset{\text{def}}{=} (x_1, {}^{x_2}\!/_2, {}^{x_3}\!/_3, \dots)$$

is linear. ■

Definition 3.3 (Bounded operator and operator norm)**.** Let X and

Y be normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ respectively, and let $T\colon X \to Y$ be an operator. We say that T is *bounded* if there exists $c \geqslant 0$ for all $x \in X$ such that

$$\|Tx\|_Y \leqslant c\,\|x\|_X.$$

We also define the *norm of* T, denoted by $\|T\|$, by

$$\|T\| \overset{\text{def}}{=} \sup\left\{\tfrac{\|Tx\|_Y}{\|x\|_X} : x \in X \text{ and } x \neq 0\right\}.$$

Clearly, $\|T\|$ exists if and only if T is bounded.

Proposition 3.4. *Let $T\colon X \to Y$ be a bounded linear operator between normed spaces. Then*

$$\begin{aligned}\|T\| &= \sup\{\|Tx\| : x \in X \text{ and } \|x\| = 1\}\\ &= \sup\{\|Tx\| : x \in X \text{ and } \|x\| \leqslant 1\}.\end{aligned}$$

Proof. $S_1 \overset{\text{def}}{=} \{\tfrac{\|Tx\|}{\|x\|} : x \in X \text{ and } x \neq 0\}$, $S_2 \overset{\text{def}}{=} \{\|Tx\| : x \in X \text{ and } \|x\| = 1\}$, and $S_3 \overset{\text{def}}{=} \{\|Tx\| : x \in X \text{ and } \|x\| \leqslant 1\}$.

Since $S_2 \subseteq S_3$, we have $\sup S_2 \leqslant \sup S_3$. Now let $x \in X$, $x \neq 0$. Then

$$\frac{\|Tx\|}{\|x\|} = \left\|\tfrac{1}{\|x\|}Tx\right\| = \left\|T(\tfrac{1}{\|x\|}\,x)\right\| = \|Ty\|$$

where $\|y\| = 1$, by scalability of the norm and linearity of T. It follows that $S_1 \subseteq S_2$, so $\sup S_1 \leqslant \sup S_2$.

Finally, if $x \in X$, $x \neq 0$ and $\|x\| \leqslant 1$, then $\|Tx\| \leqslant \|Tx\|/\|x\|$, so $S_3 \subseteq S_1$, and $\sup S_3 \leqslant \sup S_1$. $\square$

Proposition 3.5. *Let $T\colon X \to Y$ be a linear operator. Then the following are equivalent:*

(i) T is bounded,

(ii) T is (Lipschitz) continuous,

(iii) If (x_n) is a sequence in X such that $x_n \to x$, then $Tx_n \to Tx$.

Proof. (i) $\Rightarrow$ (ii). Since T is bounded, there exists a constant $c \geqslant 0$ such that for all $x \in X$, $\|Tx\| \leqslant c\|x\|$. In particular, $\|Tx - Ty\| = \|T(x-y)\| \leqslant c\|x-y\|$, so T is Lipschitz with constant c.

(ii) $\Rightarrow$ (iii). Suppose T is continuous, and let (x_n) be a sequence in X converging to x. Fix $\epsilon > 0$. Then there exists $\delta > 0$ such that if $\|x - x_n\| < \delta$, then $\|Tx - Tx_n\| < \epsilon$. But since $x_n \to x$, then there exists $N \in \mathbb{N}$ such that $\|x - x_n\| < \delta$ for all $n \geqslant N$. Consequently, $\|Tx - Tx_n\| < \epsilon$ for $n \geqslant N$.

(iii) $\Rightarrow$ (i). Suppose T is not bounded. Then for each $n \in \mathbb{N}$, there exists $x_n \in \mathbb{N}$ such that $\|Tx_n\| > n^2\|x_n\|$. Define $z_n = \frac{x_n}{n\|x_n\|}$ for all $n \in \mathbb{N}$. Clearly $z_n \to 0$, but

$$\|Tz_n\| = \frac{\|Tx_n\|}{n\|x_n\|} > \frac{n^2\|x_n\|}{n\|x_n\|} = n$$

for all n. In particular, $Tz_n \nrightarrow T0 = 0$, which contradicts (iii). □

Proposition 3.6. *Let $T\colon X \to \mathbb{F}$ be a linear operator. Then T is bounded if and only if $\ker T$ is closed.*

Proof. Suppose T is bounded, and let (x_n) be a sequence in $\ker T$ such that $x_n \to x$. Then $Tx_n \to Tx$ by proposition 3.5. But $Tx_n = 0$ for all n, so $Tx_n \to 0$. Consequently, $Tx = 0$, so $x \in \ker T$, proving closure.

Conversely, suppose T is not bounded. Then for all $n \in \mathbb{N}$, we can find $x_n \in X$ such that $|Tx_n| > n\|x_n\|$. In particular, $T \neq 0$, so there exists $e \in X$ such that $Te = y \neq 0$. Thus if $e' = e/y$, we have $Te' = 1$. Now define $y_n = e' - \frac{x_n}{Tx_n}$ for all $n \in \mathbb{N}$. Then

$$Ty_n = T(e' - \tfrac{x_n}{Tx_n}) = Te' - \tfrac{1}{Tx_n}Tx_n = 1 - 1 = 0,$$

so $y_n \in \ker T$ for all $n \in \mathbb{N}$. Also, $y_n \to e'$ since $\|\frac{x_n}{Tx_n}\| = \frac{\|x_n\|}{|Tx_n|} < \frac{1}{n} \to 0$. But $Te' = 1 \neq 0$, so $e' \notin \ker T$, and therefore $\ker T$ is not closed. □

3.1 The Dual Space

Definition 3.7 (Linear functional)**.** Let X be a normed space over a field $\mathbb{F}$. Then a *linear functional* is a linear map $T\colon X \to \mathbb{F}$ from X to the field $\mathbb{F}$.

The *dual space* of X, denoted by X^*, is the set of all linear functionals on X which are bounded, i.e.,

$$X^* \stackrel{\text{def}}{=} \{T \in \mathbb{F}^X : T \text{ is a bounded linear functional}\}.$$

Suppose X is an inner product space. In examples 3.2, we saw that for fixed $y \in X$, the map $\langle y, \cdot \rangle \colon X \to \mathbb{F}$ defined by $\langle y, \cdot \rangle(x) = \langle y, x \rangle$, is linear. This is an example of a linear functional.

Notation. We denote the map $\langle y, \cdot \rangle$ by y^*.

Proposition 3.8. *Let X be an inner product space, and let $y \in X$. Then y^* is bounded, and $\|y^*\| = \|y\|$.*

Proof. Let $x \in X$ such that $\|x\| = 1$. Then by Cauchy–Schwarz,

$$\|y^*(x)\| = |\langle y, x \rangle| \leqslant \|y\|\|x\| = \|y\|.$$

Also,

$$\|y\| = \frac{\langle y, y \rangle}{\|y\|} = \frac{y^* y}{\|y\|} = y^*(\tfrac{y}{\|y\|}) \leqslant \sup_{\|x\|=1} |y^* x| = \|y^*\|,$$

which completes the proof. □

Theorem 3.9 (Riesz representation theorem)**.** *Let X be an inner product space. If $\phi \colon X \to \mathbb{C}$ is a bounded linear functional, then there exists a unique $y \in X$ such that $\phi = y^*$.*

Proof. Notice that for any $x, z \in X$, we have

$$(\phi z)x - (\phi x)z \in \ker \phi.$$

Indeed, $\phi((\phi z)x - (\phi x)z) = (\phi z)(\phi x) - (\phi x)(\phi z) = 0$. If $\phi = 0$, then $\phi = 0^*$, so suppose $\phi \neq 0$, and pick $z \in (\ker \phi)^\perp$ such that $z \neq 0$. Then for any $x \in X$,

$$0 = \langle z, (\phi z)x - (\phi x)z \rangle = (\phi x)\langle z, z \rangle - (\phi z)\langle z, x \rangle,$$

and therefore

$$\phi x = \frac{\phi z}{\|z\|^2}\langle z, x \rangle = \langle y, x \rangle$$

for all $x \in X$, where $y = \frac{\overline{\phi z}}{\|z\|^2} z$.

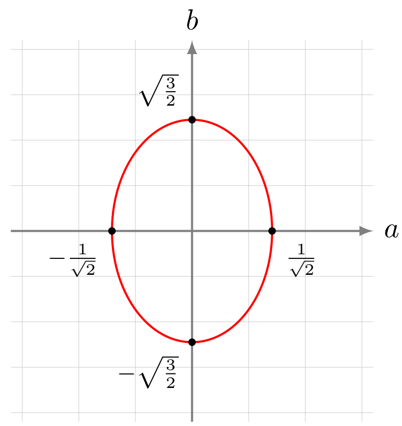

FIGURE 7: Plot of the ellipse $\frac{a^2}{1/2} + \frac{b^2}{3/2} = 1$. The function $|b|$ is maximised at $(0, \pm\sqrt{\frac{3}{2}})$.

Now we prove that this y is unique. Indeed, suppose y' exhibits the same property, i.e., $\phi x = \langle y', x\rangle$ for all $x \in X$. Then

$$\langle y - y', x\rangle = \langle y, x\rangle - \langle y', x\rangle = \phi x - \phi x = 0,$$

so $y - y' = 0$, i.e., $y = y'$. □

Example 3.10. Let $X = \{a + bx : a, b \in \mathbb{R}\}$ with $\langle p, q\rangle = \int_{-1}^{1} pq$, and define $\phi\colon X \to \mathbb{R}$ such that $\phi(a + bx) = b$.

Note that ϕ is bounded, since $\ker\phi = \operatorname{span}\{1\}$, which is closed. Thus by Riesz's theorem, there exists $y \in X$ such that $\phi = y^*$. Let $p \in (\ker\phi)^\perp = \{a + bx : a = 0\}$, say, $p(x) = x$. Then

$$y = \frac{\phi p}{\|p\|^2}\, p = \frac{x}{\|x\|^2} = \tfrac{3}{2}x,$$

and $\|\phi\| = \|y^*\| = \|y\| = \|\tfrac{3}{2}x\| = \sqrt{\tfrac{3}{2}}$.

An alternative way to calculate $\|\phi\|$ is the following. We have

$$\|\phi\| = \sup_{\|p\|=1} \|\phi p\| = \sup_{\|a+bx\|=1} |b|.$$

Now $\|a + bx\| = 1 \iff \int_{-1}^{1}(a + bx)^2\,dx = 1 \iff \frac{a^2}{1/2} + \frac{b^2}{3/2} = 1$. This equation corresponds to the ellipse illustrated in figure 7. We therefore have

$$\|\phi\| = \sup_{\|a+bx\|=1} |b| = \sup\{|b| : \tfrac{a^2}{1/2} + \tfrac{b^2}{3/2} = 1\} = \sqrt{\tfrac{3}{2}},$$

which concurs with our previous evaluation. ■

3.2 The Adjoint Operator

Recall that the transpose of a matrix satisfies $(A^\mathsf{T} y)^\mathsf{T} = y^\mathsf{T} A$. In terms of inner products, this can be expressed as $\langle A^\mathsf{T} y, x\rangle = \langle y, Ax\rangle$. In this form it can be generalised to any Hilbert space.

Proposition 3.11. *Let X and Y be Hilbert spaces, and let $T\colon X \to Y$ be a bounded linear operator. Then there exists a unique map $T^*\colon Y \to X$ such that for all $x \in X$ and $y \in Y$,*

$$\langle T^* y, x\rangle = \langle y, Tx\rangle.$$

Proof. For $y \in Y$, define the map $\langle y, T\,\cdot\,\rangle\colon X \to \mathbb{C}$, by $\langle y, T\,\cdot\,\rangle x \overset{\text{def}}{=} \langle y, Tx\rangle$ for all $x \in X$. It is clear that $\langle y, T\,\cdot\rangle$ is a bounded linear functional. Indeed, by Cauchy–Schwarz, for all $x \in X$, we have

$$\|\langle y, T\,\cdot\,\rangle x\| = |\langle y, Tx\rangle| \leqslant \|y\|\|Tx\| \leqslant \|y\|\|T\|\|x\| = c\|x\|$$

with $c = \|y\|\|T\|$. Thus by Riesz's theorem, there exists $z_y \in X$ such that $\langle y, T\,\cdot\,\rangle = z_y^*$. Therefore, for all $x \in X$,

$$\langle y, Tx\rangle = \langle y, T\,\cdot\,\rangle x = z_y^* x = \langle z_y, x\rangle.$$

Define $T^*\colon Y \to X$ by $T^* y = z_y$. Then this map has the required property. Now we show that it is unique. Indeed, if there is w_y which exhibits the same property, then for all $x \in X$,

$$\langle z_y - w_y, x\rangle = \langle z_y, x\rangle - \langle w_y, x\rangle = \langle y, Tx\rangle - \langle y, Tx\rangle = 0,$$

so that $z_y - w_y = 0$, i.e., $w_y = z_y$. □

Definition 3.12 (Adjoint operator)**.** Let X and Y be Hilbert spaces, and let $T\colon X \to Y$ be a bounded linear operator. Then the *(Hilbert)* *adjoint* of T is the unique operator $T^*\colon Y \to X$ such that

$$\langle T^* y, x\rangle_X = \langle y, Tx\rangle_Y$$

for all $x \in X$, $y \in Y$.

Lemma 3.13. *Let X and Y be Hilbert spaces, and let $T\colon X \to Y$ be a bounded linear operator. Then*

$$\|T\| = \sup\{|\langle Tx, y\rangle| : x \in X,\ y \in Y, \text{ and } \|x\| = \|y\| = 1\}.$$

Proof. Let $\|x\| = \|y\| = 1$. Then by Cauchy–Schwarz,

$$|\langle Tx, y\rangle| \leqslant \|Tx\|\|y\| \leqslant \|T\|\|x\|\|y\| = \|T\|$$

for all x, y, so $M \stackrel{\text{def}}{=} \sup_{\|x\|=\|y\|=1} |\langle Tx, y\rangle| \leqslant \|T\|$. Now suppose $T \neq 0$, and let $x \in X$ such that $Tx \neq 0$. Set $x_0 = \frac{x}{\|x\|}$ and $y_0 = \frac{Tx}{\|Tx\|}$. Then $\|x_0\| = \|y_0\| = 1$, and

$$\frac{\|Tx\|}{\|x\|} = \frac{\|Tx\|^2}{\|x\|\|Tx\|} = |\langle Tx_0, y_0\rangle| \leqslant M.$$

Since x was arbitrary, then $\|T\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} \leqslant M$. □

Theorem 3.14 (Basic Properties of T^*)**.** *Let X and Y be Hilbert spaces over $\mathbb{F}$, let S and T be bounded linear operators, and let $\alpha \in \mathbb{F}$. Then*

(i) T^* *is a bounded linear operator,*

(ii) $\|T^*\| = \|T\|$,

(iii) $(T^*)^* = T$,

(iv) $(S+T)^* = S^* + T^*$,

(v) $(\alpha T)^* = \bar{\alpha} T^*$,

(vi) $(ST)^* = T^* S^*$,

(vii) $\|T\|^2 = \|T^* T\|$.

Proof. (i). Let $x \in X$, $y, w \in Y$ and $\alpha \in \mathbb{F}$. Then

$$\langle T^*(y + \alpha w), x\rangle = \langle y + \alpha w, Tx\rangle$$

$$
\begin{aligned}
&= \langle y, Tx \rangle + \bar{\alpha}\langle w, Tx \rangle \\
&= \langle T^* y, x \rangle + \bar{\alpha}\langle T^* w, x \rangle \\
&= \langle T^* y + \alpha T^* w, x \rangle,
\end{aligned}
$$

therefore $\langle T^*(y + \alpha w) - (T^* y + \alpha T^* w), x \rangle = 0$ for all $x \in X$, and it follows that $T^*(y+\alpha w) = T^* y + \alpha T^* w$, therefore T^* is linear. Now for boundedness, observe that by lemma 3.13,

$$
\|T^*\| = \sup_{\|x\|=\|y\|=1} |\langle T^* y, x \rangle| = \sup_{\|x\|=\|y\|=1} |\langle y, Tx \rangle| = \|T\|,
$$

which also proves (ii).

(iii). Let $x \in X$ and $y \in Y$. Then

$$
\langle y, Tx \rangle = \langle T^* y, x \rangle = \overline{\langle x, T^* y \rangle} = \overline{\langle (T^*)^* x, y \rangle} = \langle y, (T^*)^* x \rangle.
$$

It follows that $\langle y, Tx - (T^*)^* x \rangle = 0$ for all $y \in Y$, so that $Tx = (T^*)^* x$ for all $x \in X$, i.e., $T = (T^*)^*$.

(iv). We start by showing that $S + T$, i.e., $(S+T)x = Sx + Tx$ for $x \in X$, is a bounded linear operator. Indeed, for $x, z \in X$ and $\alpha \in \mathbb{F}$,

$$
\begin{aligned}
(S+T)(x + \alpha z) &= S(x + \alpha z) + T(x + \alpha z) \\
&= Sx + \alpha Sz + Tx + \alpha Tz \\
&= (Sx + Tx) + \alpha(Sz + Tz) \\
&= (S+T)x + \alpha(S+T)z,
\end{aligned}
$$

so $S + T$ is linear. Now for all x with $\|x\| = 1$,

$$
\|Sx + Tx\| \leqslant \|Sx\| + \|Tx\| \leqslant \|S\| + \|T\|.
$$

It follows that $\|S + T\| \leqslant \|S\| + \|T\|$, so $S + T$ is bounded. Hence $(S+T)^*$ exists, and by (i) is also bounded. For all $x \in X$ and $y \in Y$,

$$
\begin{aligned}
\langle (S+T)^* y, x \rangle &= \langle y, (S+T)x \rangle \\
&= \langle y, Sx + Tx \rangle \\
&= \langle y, Sx \rangle + \langle y, Tx \rangle \\
&= \langle S^* y, x \rangle + \langle T^* y, x \rangle \\
&= \langle S^* y + T^* y, x \rangle \\
&= \langle (S^* + T^*) y, x \rangle,
\end{aligned}
$$

thus it follows that $(S+T)^* = S^* + T^*$.

(v) is proved similarly to (iv).

(vi). Composition of bounded linear maps (hence continuous) is linear and also continuous (hence bounded). Thus $(ST)^*$ exists, and for all $x \in X$ and $y \in Y$,

$$\langle (ST)^* y, x\rangle = \langle y, (ST)x\rangle = \langle y, S(Tx)\rangle = \langle S^* y, Tx\rangle = \langle (T^* S^*)y, x\rangle,$$

thus it follows that $(ST)^* = T^* S^*$.

(vii). We have

$$\begin{aligned}\|T^*T\| = \sup_{\|x\|=\|y\|=1} |\langle y, T^*Tx\rangle| &= \sup_{\|x\|=\|y\|=1} |\langle Ty, Tx\rangle| \\ &= \sup_{\|x\|=\|y\|=1} \|Tx\|\|Ty\| = \|T\|^2,\end{aligned}$$

where the penultimate equality follows by Cauchy–Schwarz with $y = x$. □

Examples 3.15. (i) Let $A\colon \mathbb{C}^m \to \mathbb{C}^n$, where $\langle w, z\rangle = \overline{w}^{\mathsf{T}} z$. Then A can be represented as a matrix, and

$$\langle w, Az\rangle = \overline{w}^{\mathsf{T}} Az = (\overline{\overline{A}^{\mathsf{T}} w})^{\mathsf{T}} z = \langle \overline{A}^{\mathsf{T}} w, z\rangle$$

so $A^* = \overline{A}^{\mathsf{T}}$.

(ii) Let $S\colon \ell^2 \to \ell^2$ be defined by

$$S(x_1, x_2, x_3, \dots) \stackrel{\text{def}}{=} S(0, x_1, x_2, \dots).$$

It is not hard to see that S is linear. Also, notice that for all $\boldsymbol{x} \in \ell^2$, $\|S\boldsymbol{x}\| = \|\boldsymbol{x}\|$. Hence $\|S\| = \sup_{\boldsymbol{x}\neq 0} \frac{\|S\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = 1$, so $S^*\colon \ell^2 \to \ell^2$ exists. Now we must have

$$\langle S^* \boldsymbol{y}, \boldsymbol{x}\rangle = \langle \boldsymbol{y}, S\boldsymbol{x}\rangle,$$

i.e., $\langle S^*(y_1, y_2, \dots), (x_1, x_2, \dots)\rangle = \langle (y_1, y_2, \dots), (0, x_1, x_2, \dots)\rangle$, or if we let $S^*\boldsymbol{y} = (z_1, z_2, \dots)$,

$$\begin{aligned}&\langle (z_1, z_2, \dots), (x_1, x_2, \dots)\rangle = \langle (y_1, y_2, \dots), (0, x_1, x_2, \dots)\rangle \\ \iff\ & \bar{z}_1 x_1 + \bar{z}_2 x_2 + \cdots = \bar{y}_2 x_1 + \bar{y}_2 x_2 + \cdots .\end{aligned}$$

If $z_i = y_{i+1}$ for all $i \in \mathbb{N}$, the equation holds, i.e.,

$$S^*(y_1, y_2, y_3, \dots) \stackrel{\text{def}}{=} (y_2, y_3, y_4, \dots)$$

exhibits the desired behaviour. But S^* is unique by proposition 3.11.

(iii) Let $T\colon L^2[0,1] \to \mathbb{C}^2$ be defined by

$$Tf \overset{\text{def}}{=} \left(\int_0^1 tf(t)\,dt, \int_0^1 t^2 f(t)\,dt \right).$$

That T is linear follows from linearity of integration. Also, for all f, by the Cauchy–Schwarz inequality in $L^2[0,1]$, we have

$$\begin{aligned}
\|Tf\|^2 &= \langle (I_1, I_2), (I_1, I_2) \rangle_{\mathbb{C}^2} \\
&= I_1^2 + I_2^2 \\
&= \langle \bar{t}, f \rangle_{L^2}^2 + \langle \overline{t^2}, f \rangle_{L^2}^2 \\
&\leqslant \|\bar{t}\|^2 \|f\|^2 + \|\overline{t^2}\|^2 \|f\|^2 \\
&= \left(\int_0^1 t^2\,dx \right) \left(\int_0^1 f^2 \right) + \left(\int_0^1 t^4\,dt \right) \left(\int_0^1 f^2 \right) = \frac{8}{15} \|f\|^2,
\end{aligned}$$

where I_1 and I_2 denote the two integrals in Tf. It follows that $\|T\| = \sup_{f \neq 0} \frac{\|Tf\|}{\|f\|} \leqslant 2\sqrt{\frac{2}{15}}$, so T is bounded and T^* exists. Now we must have

$$\int_0^1 \overline{T^*(z,w)} f(t)\,dt = \langle T^*(z,w), f \rangle = \langle (z,w), Tf \rangle = \int_0^1 (\bar{z}t + \bar{w}t^2) f(t)\,dt.$$

In particular, $T^*(z,w) \overset{\text{def}}{=} zt + wt^2$ satisfies this equation. Thus by uniqueness, it must be the adjoint. ∎

3.3 Optimisation in Hilbert Spaces

TBA

4 Self-Adjoint, Normal and Unitary Operators

In this section, we are mainly concerned with linear maps from and to the same Hilbert space H.

4.1 Self-Adjoint Operators

Definition 4.1 (Self-adjoint)**.** Let H be a Hilbert space, and let $T\colon H \to H$ be a bounded linear map. Then T is said to be *self-adjoint* if $T^* = T$.

Remarks 4.2. Let $T\colon H \to H$ and $S\colon H \to H$ be two bounded linear maps, and let $\lambda \in \mathbb{R}$.

(i) If S and T are self-adjoint, then $S + \lambda T$ and is self-adjoint:

$$(S + \lambda T)^* = S^* + \bar{\lambda} T^* = S + \lambda T.$$

(ii) For any T, $T + T^*$ and TT^* are self-adjoint:

$$(T + T^*)^* = T^* + T^{**} = T^* + T = T + T^*$$
$$(TT^*)^* = (T^*)^* T^* = TT^*.$$

(iii) If S and T are self-adjoint, then ST is self-adjoint iff $ST = TS$:

$$(ST)^* = T^* S^* = TS.$$

Proposition 4.3. *If $T\colon H \to H$ is a bounded linear map, then T has unique representation*

$$T = A + iB$$

where $A, B\colon H \to H$ are self-adjoint.

Proof. Define $A = \frac{T+T^*}{2}$ and $B = \frac{T-T^*}{2i}$. It is straightforward to verify that A and B are self-adjoint, and that $T = A + iB$.

For uniqueness, suppose that we may write $T = A + iB$ where A and B are self-adjoint. Then $T^* = A^* + \bar{i}B^* = A - iB$. Consequently, $T + T^* = 2A$, and $T - T^* = 2iB$. It follows that A and B are as defined above. □

Recall that if $\langle Tx, y\rangle = 0$ for all $x, y \in H$, it follows that $Tx = 0$ for all $x \in H$, i.e., $T = 0$. Here we prove a stronger version for complex Hilbert spaces.

Lemma 4.4. *Let H be a Hilbert space over $\mathbb{C}$, and let $T\colon H \to H$ be a bounded linear map. Then*

$$\langle Tx, x\rangle = 0 \text{ for all } x \in H \iff T = 0.$$

Proof. If $T = 0$, then $\langle Tx, x\rangle = \langle 0, x\rangle = 0$ for all $x \in H$.

Conversely, let $x, y \in H$. By the hypothesis, for all $\alpha \in \mathbb{C}$ we have

$$\begin{aligned} 0 = \langle T(x+\alpha y), x+\alpha y\rangle &= \langle Tx, x\rangle + \langle Tx, \alpha y\rangle + \langle T(\alpha y), x\rangle + \langle T, (\alpha y)\rangle \\ &= \langle Tx, \alpha y\rangle + \langle T(\alpha y), x\rangle \\ &= \alpha\langle Tx, y\rangle + \bar{\alpha}\langle Ty, x\rangle \end{aligned}$$

In particular, when $\alpha = 1$, we get $\langle Tx, y\rangle + \langle Ty, x\rangle = 0$, which implies that $i\langle Tx, y\rangle + i\langle Ty, x\rangle = 0$. On the other hand, when $\alpha = i$, we get $i\langle Tx, y\rangle - i\langle Ty, x\rangle = 0$. Adding gives $2i\langle Tx, y\rangle = 0$, i.e., $\langle Tx, y\rangle = 0$.

By the arbitrariness of x and y, it follows that $Tx = 0$ for all $x \in H$, i.e., that $T = 0$. □

Example 4.5. If H is a Hilbert space over $\mathbb{R}$, then lemma 4.4 is false. Indeed, in $\mathbb{R}^2$, the 90° rotation map $R\colon \mathbb{R}^2 \to \mathbb{R}^2$, given by the matrix

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

satisfies $\langle Rx, x\rangle = 0$ for all $x \in \mathbb{R}^2$.

Notice that R is bounded, since $\|Rx\| = \|x\|$ for all x. Thus R^* exists. In particular,

$$\begin{aligned} \langle y, Rx\rangle &= \langle R^* y, x\rangle \\ \implies \langle (y_1, y_2), R(x_1, x_2)\rangle &= \langle R^*(y_1, y_2), (x_1, x_2)\rangle \\ \implies \langle (y_1, y_2), (-x_2, x_1)\rangle &= \langle (z_1, z_2), (x_1, x_2)\rangle \\ \implies x_1 y_2 - x_2 y_1 &= x_1 z_1 + x_2 z_2, \end{aligned}$$

so $z_1 = y_2$ and $z_2 = -y_1$ is a solution, and therefore

$$R^*(y_1, y_2) = (y_2, -y_1) \neq (-y_2, y_1) = R(y_1, y_2),$$

so R is not self-adjoint. ■

Indeed, we have the following result.

Lemma 4.6. *Let H be a Hilbert space over $\mathbb{R}$, and let $T\colon H \to H$ be a self-adjoint bounded linear map. Then*

$$\langle Tx, y\rangle = 0 \text{ for all } x, y \in H \iff T = 0.$$

Proof. If $T = 0$, then $\langle Tx, y\rangle = \langle 0, y\rangle = 0$ for all $x, y \in H$.

Conversely, let $x, y \in H$. By the polarisation identity over $\mathbb{R}$,

$$\begin{aligned}
4\langle Tx, y\rangle &= \|Tx + y\|^2 - \|Tx - y\|^2 \\
&= \langle Tx + y, Tx + y\rangle - \langle Tx - y, Tx - y\rangle \\
&= \langle Tx, y\rangle + \langle y, Tx\rangle + \langle Tx, y\rangle + \langle y, Tx\rangle \\
&= \langle Tx, y\rangle + \langle Ty, x\rangle + \langle Tx, y\rangle + \langle Ty, x\rangle \qquad \text{(self-adjoint)} \\
&= \langle Tx, y\rangle + \langle Ty, x\rangle + \langle Tx, x\rangle + \langle Ty, y\rangle + \langle Tx, y\rangle + \langle Ty, x\rangle \\
&\qquad - \langle Tx, x\rangle - \langle Ty, y\rangle \\
&= \langle T(x + y), x + y\rangle - \langle T(x - y), x - y\rangle = 0
\end{aligned}$$

by the hypothesis. By the arbitrariness of x and y, it follows that $Tx = 0$ for all $x \in X$, i.e., that $T = 0$. $\square$

Proposition 4.7. *Let H be a Hilbert space over $\mathbb{C}$, and let $T\colon H \to H$ be a bounded linear map. Then T is self-adjoint if and only if $\langle Tx, x\rangle$ is real for all $x \in H$.*

Proof. Let $x \in H$. If T is self-adjoint, then

$$\overline{\langle Tx, x\rangle} = \langle x, Tx\rangle = \langle T^*x, x\rangle = \langle Tx, x\rangle,$$

so $\langle Tx, x\rangle$ is real. Conversely, suppose $\langle Tx, x\rangle$ is real. Then

$$\begin{aligned}
\langle (T - T^*)x, x\rangle &= \langle Tx, x\rangle - \langle T^*x, x\rangle \\
&= \langle Tx, x\rangle - \langle x, Tx\rangle \\
&= \overline{\langle Tx, x\rangle} - \langle x, Tx\rangle \qquad \text{(real)} \\
&= \langle x, Tx\rangle - \langle x, Tx\rangle = 0,
\end{aligned}$$

and since x is arbitrary, the result follows by lemma 4.4. $\square$

4.2 Normal Operators

Definition 4.8 (Normal operator)**.** Let H be a Hilbert space, and let $T\colon H \to H$ be a bounded linear map. Then T is said to be *normal* if

$$TT^* = T^*T.$$

Notice that any self-adjoint map is normal.

Proposition 4.9. *Let $T\colon H \to H$ be a bounded linear operator on a Hilbert space H. Write $T = A+iB$ where A and B are self-adjoint (proposition 4.3). Then T is normal if and only if $AB = BA$.*

Proof.

$$\begin{aligned}
(A+iB)(A+iB)^* &= (A+iB)(A-iB) \\
&= A^2 - iAB + iBA + B^2 \\
&= A^2 - iBA + iAB + B^2 \qquad \text{(iff } AB = BA\text{)} \\
&= (A-iB)(A+iB) \\
&= (A+iB)^*(A+iB).
\end{aligned}$$

□

Proposition 4.10. *Let H be a Hilbert space, and let $T\colon H \to H$ be a bounded linear operator. Then the following are equivalent.*

(i) T *is normal.*

(ii) $\|Tx\| = \|T^*x\|$ *for all* $x \in H$*.*

Notice that (ii) is stronger than $\|T\| = \|T^*\|$, which is true for any bounded linear operator (theorem 3.14).

Proof. (i) ⇒ (ii). Let $x \in H$, and let T be normal. Then

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle = \langle x, TT^*x\rangle = \langle T^*x, T^*x\rangle = \|T^*x\|.$$

(ii) ⇒ (i). Let $x \in H$. Then

$$\langle (TT^* - T^*T)x, x\rangle = \langle T^*x, T^*x\rangle - \langle Tx, Tx\rangle = \|T^*x\|^2 - \|Tx\|^2 = 0.$$

Thus if H is a complex Hilbert space, the result follows by lemma 4.4. On the other hand, if H is a real Hilbert space, the result follows from lemma 4.6 and the fact that $TT^* - T^*T$ is self-adjoint. □

Example 4.11. Let $H = \ell^2$ and $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots\}$ be an orthonormal basis. If $(\lambda_n) \in \ell^2$ is a sequence of complex numbers such that $\sup_n |\lambda_n| = M < \infty$, then we can define the linear map $T\colon H \to H$ such that $T\boldsymbol{e}_n = \lambda_n \boldsymbol{e}_n$ for all $n \in \mathbb{N}$. (If $\boldsymbol{x} = \sum_{i=1}^{\infty} \alpha_i \boldsymbol{e}_i$, then $T\boldsymbol{x} = \sum_{i=1}^{\infty} \alpha_i T\boldsymbol{e}_i$.)

It is straightforward to verify that $T^*\boldsymbol{e}_n = \overline{\lambda}_n \boldsymbol{e}_n$. Thus

$$\|T^*\boldsymbol{e}_n\| = \|\overline{\lambda}_n \boldsymbol{e}_n\| = |\lambda_n|\|\boldsymbol{e}_n\| = \|\lambda_n \boldsymbol{e}_n\| = \|T\boldsymbol{e}_n\|,$$

and it follows that $\|T^*\boldsymbol{x}\| = \|T\boldsymbol{x}\|$ for all $\boldsymbol{x} \in H$. Therefore T is normal. Moreover, T is self-adjoint if and only if $\lambda_n = \overline{\lambda}_n$ for all n, i.e., iff $\lambda_n \in \mathbb{R}$ for all $n \in \mathbb{N}$. ■

Note. A linear operator T is said to be *bounded below* if there exists $c > 0$ such that for all $x \in H$, $\|Tx\| \geqslant c\|x\|$.

Proposition 4.12. *Let H be a Hilbert space, and $T\colon H \to H$ a normal bounded linear operator. Then*

(i) $\ker T = \ker T^*$,

(ii) $\operatorname{im} T$ *is dense in H if and only if T is injective,*

(iii) *if T is bounded below, then* $\ker T^* = 0$.

Proof. For (i),

$$x \in \ker T \Leftrightarrow Tx = 0 \Leftrightarrow \|Tx\| = 0 = \|T^*x\| \Leftrightarrow T^*x = 0 \Leftrightarrow x \in \ker T^*.$$

For (ii),

$$\operatorname{im} T \text{ dense in } H \Leftrightarrow \overline{\operatorname{im} T} = H \Leftrightarrow \{0\} = (\overline{\operatorname{im} T})^{\perp} = \ker T^* = \ker T.$$

Finally for (iii), observe that if T is bounded below, then

$$T^*x = 0 \Leftrightarrow \|T^*x\| = \|Tx\| = 0 \Leftrightarrow c\|x\| = 0 \Leftrightarrow x = 0,$$

where c is the lower bound of T. □

Theorem 4.13 (Bounded inverse theorem)**.** *Let X, Y be a complete normed spaces, and let $T\colon X \to Y$ be a bounded invertible linear map. Then the inverse $T^{-1}\colon Y \to X$ is also bounded.*

We do not give a proof of this fact.

Theorem 4.14. *Let $T\colon H \to H$ be a normal bounded linear operator. Then T is invertible if and only if T is bounded below.*

Proof. Suppose T is invertible. Then there exists $T^{-1}\colon H \to H$ which is also bounded by the bounded inverse theorem, i.e., there exists $c > 0$ such that $\|T^{-1}y\| \leqslant c\|y\|$ for all $y \in H$.

Hence for $x \in H$, $\|x\| = \|T^{-1}Tx\| \leqslant c\|Tx\|$, as required.

Conversely, suppose T is bounded below. Then by proposition 4.12, $\ker T = \ker T^* = \{0\}$, so that T is injective. We also have that $\operatorname{im} T$ is dense in H. We show that $\operatorname{im} T$ is closed, so that $\operatorname{im} T = \overline{\operatorname{im} T} = H$.

Indeed, let y_n be a Cauchy sequence in $\operatorname{im} T$. Then it converges to $y \in H$. Being a sequence in $\operatorname{im} T$, then for each $n \in \mathbb{N}$, there exists $x_n \in H$ such that $y_n = Tx_n$. Thus $\|y_n - y_m\| = \|T(x_n - x_m)\| \geqslant c\|x_n - x_m\|$, which implies (x_n) is also Cauchy, so $x_n \to x \in H$. Thus since T is bounded, it follows that $Tx_n \to Tx$ (proposition 3.5). Therefore $y = Tx \in \operatorname{im} T$. □

Example 4.15. [Multiplication operation] Fix a continuous function $\kappa \in C[a, b]$, and define $T_\kappa\colon L^2[a, b] \to L^2[a, b]$ by $T_\kappa f \stackrel{\text{def}}{=} \kappa f$, where $(\kappa f)(t) = \kappa(t)f(t)$ for all $t \in [a, b]$.

Then by Cauchy–Schwarz,

$$\|T_\kappa f\|^2 = \int_a^b |\kappa(t)|^2|f(t)|^2\,dt \leqslant \int_a^b M^2|f(t)|^2\,dt = M^2\|f\|^2,$$

where M is an upper-bound for κ (continuous functions on closed intervals are bounded), so T_κ is bounded. Now let $f, g \in L^2[a, b]$. Then

$$\langle T_\kappa f, g\rangle = \int_a^b \overline{\kappa f} g = \int_a^b \overline{f}(\overline{\kappa} g) = \langle f, T_{\overline{\kappa}} g\rangle,$$

so $T_\kappa^* = T_{\overline{\kappa}}$, in particular, T is not self-adjoint. But for all $f \in L^2[a, b]$, $T_\kappa T_\kappa^* f = T_\kappa T_{\overline{\kappa}} f = \kappa\overline{\kappa} f = \overline{\kappa}\kappa f = T_{\overline{\kappa}} T_\kappa f = T_\kappa^* T_\kappa f$, so T_κ is normal. ■

4.3 Isometric and Unitary Operators

For a Hilbert space H, let $B(H)$ be the set

$$B(H) \stackrel{\text{def}}{=} \{T \in H^H : T \text{ is a bounded linear operator}\}.$$

Then we have the following definition.

Definition 4.16 (Isometry)**.** Let H be a Hilbert space, and let $T \in B(H)$. If for all $x, y \in H$,

$$\|Tx - Ty\| = \|x - y\|,$$

then T is said to be an *isometry*.

An isometry is a bounded linear map which "preserves distances". Indeed, the distance $\|x - y\|$ between x and y is equal to the distance $\|Tx - Ty\|$ between the images Tx and Ty.

Because of linearity, the isometry condition can be formulated equivalently as $\|Tx\| = \|x\|$ for all $x \in H$.

Proposition 4.17. *Let H be a Hilbert space, and let $T \in B(H)$. Then the following are equivalent.*

(i) *T is an isometry,*

(ii) *$T^*T = I$,*

(iii) *$\langle Tx, Ty \rangle = \langle x, y \rangle$ for all $x, y \in H$.*

Proof. (i) $\Rightarrow$ (ii). Let $x, y \in X$. Then

$$\begin{aligned}
&4\langle (T^*T - I)x, y \rangle \\
&= 4\langle T^*Tx, y \rangle - 4\langle x, y \rangle \\
&= 4\langle Tx, Ty \rangle - 4\langle x, y \rangle \\
&= \|Tx + Ty\|^2 - \|Tx - Ty\|^2 + i\|Tx - iTy\|^2 - i\|Tx + iTy\|^2 \\
&\qquad - \|x + y\|^2 + \|x - y\|^2 - i\|x - iy\|^2 + i\|x + iy\|^2 \\
&= \|x + y\|^2 - \|x - y\|^2 + i\|x - iy\|^2 - i\|x + iy\|^2 \qquad (T \text{ isometry}) \\
&\qquad - \|x + y\|^2 + \|x - y\|^2 - i\|x - iy\|^2 + i\|x + iy\|^2 \\
&= 0
\end{aligned}$$

by the polarisation identity. Therefore $(T^*T - I)x = 0$ for all $x \in H$, and the result follows.

(ii) $\Rightarrow$ (iii). Let $x, y \in H$. Then

$$\langle Tx, Ty \rangle = \langle T^*Tx, y \rangle = \langle Ix, y \rangle = \langle x, y \rangle,$$

as required.

(iii) $\Rightarrow$ (i). Let $x, y \in H$. Then

$$\begin{aligned} \|Tx - Ty\|^2 = \|T(x-y)\|^2 &= \langle T(x-y), T(x-y)\rangle \\ &= \langle x-y, x-y\rangle = \|x-y\|^2, \end{aligned}$$

as required. $\square$

Proposition 4.18. *Let H be a Hilbert space, and let $T \in B(H)$ be an isometry. Then*

(i) $\operatorname{im} T$ *is closed,*

(ii) $\ker T = \{0\}$.

Proof. (i). Let y_n be a Cauchy sequence in $\operatorname{im} T$. Then $y_n \to y \in H$, and there exists $x_n \in H$ such that $y_n = Tx_n$ for each $n \in \mathbb{N}$. Since $\|Tx\| = \|x\|$, then $x_n \to x$ also. Hence by proposition 3.5, it follows that $Tx_n \to Tx$, so $y = Tx \in \operatorname{im} T$. Therefore $\operatorname{im} T$ is closed.

(ii). $Tx = 0 \iff \|Tx\| = 0 = \|x\| \iff x = 0$, so $\ker T = \{0\}$. $\square$

Definition 4.19 (Unitary operator)**.** Let H be a Hilbert space, and let $T \in B(H)$. If

$$T^*T = TT^* = I,$$

then T is said to be *unitary.*

Notice that by proposition 4.10(ii), it follows that an operator T is unitary if and only if T^* is unitary.

Theorem 4.20. *Let H be a Hilbert space, and let $T \in B(H)$. Then the following are equivalent.*

(i) T is unitary,

(ii) T and T^* are isometries,

(iii) T is an isometry and T^* is injective,

(iv) T is a surjective isometry,

(v) T is bijective and $T^* = T^{-1}$.

Proof. (i) $\Rightarrow$ (ii). If T is unitary, then $T^*T = I$ and $TT^* = (T^*)^*T^* = I$, so T and T^* are isometries by proposition 4.17.

(ii) $\Rightarrow$ (iii). Since T is an isometry, $\ker T = \{0\}$ by proposition 4.18, so T is injective.

(iii) $\Rightarrow$ (iv). Since T^* is injective, $\ker T^* = \{0\}$; and since T is an isometry, $\operatorname{im} T = \overline{\operatorname{im} T}$. Thus $H = (\ker T^*)^\perp = \overline{\operatorname{im} T} = \operatorname{im} T$, so T is surjective.

(iv) $\Rightarrow$ (v). Since T is a surjective isometry, it is bijective ($\ker T = \{0\}$), therefore it is invertible, i.e., there exists T^{-1} such that $T^{-1}T = I = T^*T$. Therefore
$$T^* = T^*(TT^{-1}) = (T^*T)T^{-1} = IT^{-1} = T^{-1},$$
as required.

(v) $\Rightarrow$ (i). Since $T^* = T^{-1}$, then $T^*T = TT^* = I$. $\square$

Examples 4.21. (i) Let T_κ be the multiplication operator in $L^2[a,b]$ (see example 4.15). Then $T_\kappa^* T_\kappa f(t) = |\kappa(t)|^2 f(t)$

5 Spectral Theory

5.1 Compact Operators

TBA

Bibiliography

[1] J. Muscat, (2014). *Functional Analysis*, Springer-Verlag, 1st Edition, chapters 9, 10 and 14 (pp. 139–219, 307–343).

[2] E. Kreyszig, (1989). *Introductory Functional Analysis with Applications*, John Wiley & Sons, 1st Edition.

[3] W. Rudin, (1987). *Real and Complex Analysis*, McGraw–Hill, 3rd Edition.

[4] W. Rudin, (1991). *Functional Analysis*, McGraw–Hill, 2nd Edition.

[5] P. R. Halmos, (1974). *Measure Theory*, Graduate Texts in Mathematics, Springer-Verlag, 1st Edition, chapters 4 and 5 (pp. 73–112).

Index