Linear Analysis

Contents

Part I
Class Notes

1 Inner-Product Spaces

Definition 1.1. Let V be a vector space over $\mathbb{C}$. A map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ is called an **inner-product on** V if it satisfies the following properties:

1. Linearity in the first argument: $\langle x, z \rangle + \langle y, z \rangle = \langle x + y, z \rangle$ and $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$ for all $x, y, z \in V$ and $\lambda \in \mathbb{C}$.
2. Conjugate symmetric: $\langle x, y \rangle = \overline{\langle y, x \rangle}$ for all $x, y \in V$.
3. Positive definite: $\langle x, x \rangle > 0$ for all nonzero $x \in V$.

A vector space equipped with an inner-product is called an **inner-product space**. We often write $\mathcal{V}$ to denote an inner-product space.

Proposition 1.1. *Let $\mathcal{V}$ be an inner-product space. Then*

1. $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$ *for all* $x, y, z \in \mathcal{V}$*;*
2. $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle$ *for all* $x, y \in \mathcal{V}$ *and* $\lambda \in \mathbb{C}$
3. $\langle x, 0 \rangle = 0 = \langle 0, x \rangle$ *for all* $x \in \mathcal{V}$*;*
4. *Let* $x, y \in \mathcal{V}$*. If* $\langle x, z \rangle = \langle y, z \rangle$ *for all* $z \in \mathcal{V}$*, then* $x = y$*.*

Proof.

1. Let $x, y, z \in \mathcal{V}$. Then
$$\begin{aligned} \langle x, y + z \rangle &= \overline{\langle y + z, x \rangle} \\ &= \overline{\langle y, x \rangle + \langle z, x \rangle} \\ &= \overline{\langle y, x \rangle} + \overline{\langle z, x \rangle} \\ &= \langle x, y \rangle + \langle x, z \rangle. \end{aligned}$$
2. Let $x, y \in \mathcal{V}$ and $\lambda \in \mathbb{C}$. Then
$$\begin{aligned} \langle x, \lambda y \rangle &= \overline{\langle \lambda y, x \rangle} \\ &= \overline{\lambda \langle y, x \rangle} \\ &= \overline{\lambda} \overline{\langle y, x \rangle} \\ &= \overline{\lambda} \langle x, y \rangle. \end{aligned}$$
3. Let $x \in \mathcal{V}$. Then
$$\begin{aligned} \langle x, 0 \rangle &= \langle x, 0 + 0 \rangle \\ &= \langle x, 0 \rangle + \langle x, 0 \rangle \end{aligned}$$
implies $\langle x, 0 \rangle = 0$. A similar argument gives $\langle 0, x \rangle = 0$.
4. Assuming $\langle x, z \rangle = \langle y, z \rangle$ for all $z \in \mathcal{V}$, then we have $\langle x - y, z \rangle = 0$ for all $z \in \mathcal{V}$. In particular, setting $z = x - y$, we have $\langle x - y, x - y \rangle = 0$. Since the inner-product is positive definite, we must have $x - y = 0$, and hence $x = y$.

$\square$

1.1 Examples of Inner-Product Spaces

If we are given a complex vector space V, then we can **give V the structure of an inner-product space** by equipping V with an inner-product
$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}.$$
In the following examples, we give many familiar complex vector spaces the structure of an inner-product space.

1.1.1 Giving $\mathbb{C}^n$ the structure of an inner-product space

Proposition 1.2. *Let $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ be given by*

$$\langle x, y \rangle = \langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = \sum_{i=1}^n x_i \overline{y}_i.$$

for all $x, y \in \mathbb{C}^n$. Then the pair $(\mathbb{C}^n, \langle \cdot, \cdot \rangle)$ forms an inner-product space.

Proof. For linearity in the first argument follows from linearity, let $x, y, z \in \mathbb{C}^n$. Then

$$\begin{aligned} \langle x + y, z \rangle &= \sum_{i=1}^n (x_i + y_i) \overline{z}_i \\ &= \sum_{i=1}^n x_i \overline{z}_i + \sum_{i=1}^n y_i \overline{z}_i \\ &= \langle x, z \rangle + \langle y, z \rangle. \end{aligned}$$

For conjugate symmetry of $\langle \cdot, \cdot \rangle$, let $x, y \in \mathbb{C}^n$. Then

$$\begin{aligned} \langle x, y \rangle &= \sum_{i=1}^n x_i \overline{y}_i \\ &= \sum_{i=1}^n \overline{\overline{x_i \overline{y}_i}} \\ &= \sum_{i=1}^n \overline{y_i \overline{x}_i} \\ &= \overline{\langle y, x \rangle}. \end{aligned}$$

For positive-definiteness of $\langle \cdot, \cdot \rangle$, let $x \in \mathbb{C}^n$. Then

$$\begin{aligned} \langle x, x \rangle &= \sum_{i=1}^n x_i \overline{x}_i \\ &= \sum_{i=1}^n |x_i|^2. \end{aligned}$$

is a sum of its components absolute squared. This implies positive-definiteness. □

1.1.2 Giving $\mathrm{M}_{m \times n}(\mathbb{C})$ the structure of an inner-product space

Since $\mathrm{M}_{m \times n}(\mathbb{C}) \cong \mathbb{C}^{mn}$, we can give $\mathrm{M}_{m \times n}(\mathbb{C})$ the structure of an inner-product space by equipping it with the inner-product described in the previous example. In more detail, let E_{ij} be the standard matrix in $\mathrm{M}_{m \times n}(\mathbb{C})$ whose entry in the (i, j)-th component is 1 and whose entry everywhere else is 0, let e_k be the standard basis vector in $\mathbb{C}^{mn}$ whose entry in the k-th compenent is 1 and whose entry everywhere else is 0, and let $\varphi : \mathrm{M}_{m \times n}(\mathbb{C}) \to \mathbb{C}^{mn}$ be the isomorphism such that

$$\varphi(E_{ij}) = e_{n(i-1)+j}$$

for all $E_{ij} \in \mathrm{M}_{m \times n}(K)$ where $1 \leq i \leq m$ and $1 \leq j \leq n$. So under φ, we have

$$A := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \mapsto \begin{pmatrix} a_{11} \\ \vdots \\ a_{1n} \\ \vdots \\ a_{m1} \\ \vdots \\ a_{mn} \end{pmatrix} := a$$

Using this isomorphism, we give $\mathrm{M}_{m \times n}(\mathbb{C})$ the structure of an inner-product space by defining

$$\langle \cdot, \cdot \rangle : \mathrm{M}_{m \times n}(\mathbb{C}) \times \mathrm{M}_{m \times n}(\mathbb{C}) \to \mathbb{C}$$

by the formula

$$\langle A, B\rangle := \mathrm{Tr}(AB^*) = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}\overline{b}_{ij} = \langle a, b\rangle$$

where

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}, \quad \text{and} \quad B^* = \begin{pmatrix} \overline{b}_{11} & \cdots & \overline{b}_{m1} \\ \vdots & \ddots & \vdots \\ \overline{b}_{1n} & \cdots & \overline{b}_{nm} \end{pmatrix}.$$

1.1.3 Giving $\ell^2(\mathbb{N})$ the structure of an inner-product space

Lemma 1.1. *Let a and b be nonnegative real numbers. Then we have*

$$ab \leq \frac{1}{2}(a^2 + b^2) \tag{1}$$

with equality if and only if $a = b$.

Proposition 1.3. *Let $\ell^2(\mathbb{N})$ be the set of all sequence (x_n) in $\mathbb{C}$ such that*

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty$$

and let $\langle \cdot, \cdot\rangle \colon \ell^2(\mathbb{N}) \times \ell^2(\mathbb{N}) \to \mathbb{C}$ be given by

$$\langle (x_n), (y_n)\rangle = \sum_{n=1}^{\infty} x_n \overline{y}_n.$$

for all $(x_n), (y_n) \in \ell^2(\mathbb{N})$. Then the pair $(\ell^2(\mathbb{N}), \langle \cdot, \cdot\rangle)$ forms an inner-product space.

Proof. We first need to show that $\ell^2(\mathbb{N})$ is indeed a vector space. In fact, we will show that $\ell^2(\mathbb{N})$ is a subspace of $\mathbb{C}^N$, the set of all sequences in $\mathbb{C}$. Let $(x_n), (y_n) \in \ell^2(\mathbb{N})$ and $\lambda \in \mathbb{C}$. Then Lemma (6.1) implies

$$\begin{aligned} \sum_{n=1}^{\infty} |\lambda x_n + y_n|^2 &\leq \sum_{n=1}^{\infty} |\lambda x_n|^2 + \sum_{n=1}^{\infty} |y_n|^2 + \sum_{n=1}^{\infty} 2|\lambda x_n||y_n| \\ &\leq \lambda^2 \sum_{n=1}^{\infty} |x_n|^2 + \sum_{n=1}^{\infty} |y_n|^2 + \lambda^2 \sum_{n=1}^{\infty} |x_n|^2 + |y_n|^2 \\ &< \infty. \end{aligned}$$

Therefore $(\lambda x_n + y_n) \in \ell^2(\mathbb{N})$, which implies $\ell^2(\mathbb{N})$ is a subspace of $\mathbb{C}^{\mathbb{N}}$.

Next, let us show that the inner product converges, and hence is defined everywhere. Let $(x_n), (y_n) \in \ell^2(\mathbb{N})$. Then it follows from Lemma (6.1) that

$$\begin{aligned} \sum_{n=1}^{\infty} |x_n \overline{y}_n| &= \sum_{n=1}^{\infty} |x_n||y_n| \\ &\leq \sum_{n=1}^{\infty} \frac{|x_n|^2 + |y_n|^2}{2} \\ &= \frac{1}{2}\sum_{n=1}^{\infty} |x_n|^2 + \frac{1}{2}\sum_{n=1}^{\infty} |y_n|^2 \\ &< \infty. \end{aligned}$$

Therefore $\sum_{n=1}^{\infty} x_n \overline{y}_n$ is absolutely convergent, which implies it is convergent. (We can't use Cauchy-Schwarz here since we haven't yet shown that $\langle \cdot, \cdot\rangle$ is in fact an inner-product).

Finally, let us shows that $\langle \cdot, \cdot\rangle$ is an inner-product. Linearity in the first argument follows from distrubitivity of multiplication and linearity of taking infinite sums. For conjugate symmetry, let $(x_n), (y_n) \in \ell^2(\mathbb{N})$. Then

$$\begin{aligned}
\langle (x_n), (y_n) \rangle &= \sum_{n=1}^{\infty} x_n \overline{y}_n \\
&= \lim_{N \to \infty} \sum_{n=1}^{N} x_n \overline{y}_n \\
&= \overline{\overline{\lim_{N \to \infty} \sum_{n=1}^{N} x_n \overline{y}_n}} \\
&= \overline{\lim_{N \to \infty} \sum_{n=1}^{N} \overline{x_n \overline{y}_n}} \\
&= \overline{\lim_{N \to \infty} \sum_{n=1}^{N} y_n \overline{x}_n} \\
&= \overline{\sum_{n=1}^{\infty} y_n \overline{x}_n} \\
&= \overline{\langle (y_n), (x_n) \rangle},
\end{aligned}$$

where we were allowed to bring the conjugate inside the limit since the conjugate function is continuous on $\mathbb{C}$. For positive-definiteness, let $(x_n) \in \ell^2(\mathbb{N})$. Then

$$\begin{aligned}
\langle (x_n), (x_n) \rangle &= \sum_{n=1}^{\infty} x_n \overline{x}_n \\
&= \sum_{n=1}^{\infty} |x_n|^2 \\
&\geq 0.
\end{aligned}$$

If $\sum_{n=1}^{\infty} |x_n|^2 = 0$, then clearly we must have $x_n = 0$ for all n. □

1.1.4 Giving $C[a, b]$ the structure of an inner-product space

Proposition 1.4. *Let $C[a, b]$ be the space of all continuous functions defined on the closed interval $[a, b]$ and let $\langle \cdot, \cdot \rangle \colon C[a, b] \times C[a, b] \to \mathbb{C}$ be given by*

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

for all $f, g \in C[a, b]$. Then the pair $(C[a, b], \langle \cdot, \cdot \rangle)$ forms an inner-product space.

Proof. Exercise. □

1.2 Norm Induced by Inner-Product

Definition 1.2. The **norm** of $x \in \mathcal{V}$, denoted $\|x\|$, is defined by

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

Lemma 1.2. *(Pythagorean Theorem) Let x and y be nonzero vectors in $\mathcal{V}$ such that $\langle x, y \rangle = 0$ (we call such vectors **orthogonal** to one another). Then*

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

Proof. We have

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

□

1.2.1 Properties of Norm

Proposition 1.5. *If $x, y \in \mathcal{V}$ and $\lambda \in \mathbb{C}$, then*

1. *Positive-Definiteness:* $\|x\| \geq 0$ *with equality if and only if* $x = 0$;
2. *Absolutely Homogeneous:* $\|\lambda x\| = |\lambda| \|x\|$;
3. *Cauchy-Schwarz:* $|\langle x, y\rangle| \leq \|x\| \|y\|$ *with equality if and only if x and y are linearly dependent.*
4. *Subadditivity:* $\|x + y\| \leq \|x\| + \|y\|$

Proof.

1. This follows from positive-definiteness of $\langle \cdot, \cdot \rangle$.
2. We have

$$\begin{aligned}
\|\lambda x\| &= \sqrt{\langle \lambda x, \lambda x\rangle} \\
&= \sqrt{\lambda \overline{\lambda} \langle x, x\rangle} \\
&= \sqrt{|\lambda|^2 \langle x, x\rangle} \\
&= |\lambda| \sqrt{\langle x, x\rangle} \\
&= |\lambda| \|x\|.
\end{aligned}$$

3. We may assume that both x and y are nonzero, since it is trivial in this case. Let

$$z = x - \mathrm{pr}_y(x) = x - \frac{\langle x, y\rangle}{\langle y, y\rangle} y.$$

Then by linearity of the inner product in the first argument, one has

$$\begin{aligned}
\langle z, y\rangle &= \left\langle x - \frac{\langle x, y\rangle}{\langle y, y\rangle} y, y \right\rangle \\
&= \langle x, y\rangle - \frac{\langle x, y\rangle}{\langle y, y\rangle} \langle y, y\rangle \\
&= 0.
\end{aligned}$$

Therefore z is a vector orthogonal to the vector y. We can thus apply the Pythagorean theorem to

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

which gives

$$\begin{aligned}
\|x\|^2 &= \left| \frac{\langle x, y\rangle}{\langle y, y\rangle} \right|^2 \|y\|^2 + \|z\|^2 \\
&= \frac{|\langle x, y\rangle|}{(\|y\|^2)^2}^2 \|y\|^2 + \|z\|^2 \\
&= \frac{|\langle x, y\rangle|}{\|y\|^2}^2 + \|z\|^2 \\
&\geq \frac{|\langle x, y\rangle|^2}{\|y\|^2},
\end{aligned}$$

and after multiplication by $\|y\|^2$ and taking square root, we get the Cauchy-Schwarz inequality. Moreover, if the relation $\geq$ in the above expression is actually an equality, then $\|z\|^2 = 0$ and hence $z = 0$; the definition of z then establishes a relation of linear dependence between x and y. On the other hand, if x and y are linearly dependent, then there exists $\lambda \in \mathbb{C}$ such that $x = \lambda y$. Then

$$\begin{aligned}
|\langle x, y\rangle| &= |\langle \lambda y, y\rangle| \\
&= |\lambda| |\langle y, y\rangle| \\
&= |\lambda| \|y\|^2 \\
&= \|\lambda y\| \|y\| \\
&= \|x\| \|y\|.
\end{aligned}$$

4. By the Cauchy-Schwarz inequality, we have

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

now we take square roots on both sides to get the desired result.

□

Proposition 1.6. *(Parallelogram Identity) Let $x, y \in \mathcal{V}$. Then*

$$\|x-y\|^2 + \|x+y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{2}$$

Proof. We have

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

and

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

Adding these together gives us our desired result. □

Proposition 1.7. *(Polarization Identity) Let $x, y \in \mathcal{V}$. Then*

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

Proof. We have

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

and

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

and

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

and

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

Adding these together gives us our desired result. □

1.2.2 Normed Vector Spaces

Definition 1.3. Let V be a $\mathbb{C}$-vector space. A **norm** on V is a nonnegative-valued scalar function $\|\cdot\|: V \to [0,\infty)$ such that for all $\lambda \in \mathbb{C}$ and $x, y \in V$, we have

1. (Subadditivity) $\|x+y\| \leq \|x\| + \|x\|$,
2. (Absolutely Homogeneous) $\|\lambda x\| = |\lambda| \|x\|$,
3. (Positive-Definite) $\|x\| = 0$ if and only if $x = 0$.

We call the pair $(V, \|\cdot\|)$ a **normed vector space**.

Proposition (1.5) implies $\mathcal{V}$ is a normed vector space. This justifies our choice of the word "norm" in Definition (1.2). Proposition (1.5) also tells us that $\mathcal{V}$ satisfies an extra property which is not satisfied by other normed vector spaces, namely the Cauchy-Schwarz inequality. In fact, it turns out that inner-product spaces are just normed vector spaces which satisfy the parallelogram law.

1.2.3 Metric Induced By Norm

Definition 1.4. A **metric** on a set X is a function $d: X \times X \to \mathbb{R}$ which satisfies the following three properties:

1. (Identity of Indiscernibles) $d(x,y) = 0$ if and only if $x = y$ for all $x, y \in X$,
2. (Symmetric) $d(x,y) = d(y,x)$ for all $x, y \in X$,
3. (Triangle Inequality) $d(x,z) \leq d(x,y) + d(y,z)$ for all $x, y, z \in X$.

A set X together with a choice of a metric d is called a **metric space** and is denoted (X, d), or just denoted X if the metric is understood from context.

Remark. Given the three axioms above, we also have positive-definiteness: $d(x,y) \geq 0$ with equality if and only if $x = y$ for all $x, y \in X$. Indeed,

$$\begin{aligned} 0 &= d(x,x) \\ &\leq d(x,y) + d(y,x) \\ &= d(x,y) + d(x,y) \\ &= 2d(x,y). \end{aligned}$$

This implies $d(x,y) \geq 0$.

Proposition 1.8. *Let $(V, \|\cdot\|)$ be a normed vector space. Define $d: V \times V \to \mathbb{R}$ by $d(x,y) = \|x-y\|$ for all $(x,y) \in V \times V$. Then (V, d) is a metric space.*

Proof. Let us first check that d satisfies the identity of indiscernibles property. Since $\|\cdot\|$ is positive-definite, $d(x,y) = 0$ implies $\|x-y\| = 0$ which implies $x = y$. On the other hand, suppose $x = y$. Then since $\|\cdot\|$ is absolutely homogeneous, we have $\|0\| = |0|\|0\| = 0$, and so $d(x,y) = \|0\| = 0$.

Next we check that d is symmetric. For all $(x,y) \in V \times V$, we have

$$\begin{aligned} d(x,y) &= \|x-y\| \\ &= \|-(y-x)\| \\ &= |-1| \|y-x\| \\ &= \|y-x\| \\ &= d(y,x). \end{aligned}$$

Finally, triangle inequality for d follows from subadditivity of $\|\cdot\|$. Indeed, for all $x, y, z \in V$, we have

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

$\square$

The distance between points $x, y \in \mathcal{V}$ is measured using the metric induced by the norm:

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

Definition 1.5. A sequence (x_n) in $\mathcal{V}$ is said to converge to a point $x \in \mathcal{V}$ if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$n \geq N \text{ implies } \|x_n - x\| < \varepsilon.$$

In this case we write $x_n \to x$ as $n \to \infty$ or more simply $x_n \to x$. We also write

$$\lim_{n\to\infty} x_n = x.$$

Proposition 1.9. *Let (x_n) and (y_n) be two sequences in $\mathcal{V}$ and let (λ_n) be a sequence in $\mathbb{C}$. Then the following statements hold:*

1. *There exists $M \geq 0$ such that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$.*
2. *If $x_n \to x$ and $y_n \to y$, then $x_n + y_n \to x + y$.*
3. *If (λ_n) is a sequence in $\mathbb{C}$ and $\lambda_n \to \lambda$, then $\lambda_n x_n \to \lambda x$.*
4. *If $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$. In particular, $\|x_n\| \to \|x\|$.*

Proof.

1. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $\|x_n - x\| \leq 1$. Now set M to be

$$M = \max\{\|x_1\|, \dots, \|x_{N-1}\|, \|x\| + 1\}$$

2. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $\|x_n - x\| < \varepsilon/2$ and $\|y_n - y\| < \varepsilon/2$. Then $n \geq N$ implies

$$\begin{aligned} \|x_n + y_n - (x+y)\| &\leq \|x_n - x\| + \|y_n - y\| \\ &< \varepsilon/2 + \varepsilon/2 \\ &= \varepsilon. \end{aligned}$$

3. Since $x_n \to x$, there exists $M \geq 0$ such that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$. Choose such an M and let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $\|x_n - x\| < \varepsilon/2|\lambda|$ and $|\lambda_n - \lambda| < \varepsilon/2M$. Then $n \geq N$ implies

$$\begin{aligned} \|\lambda_n x_n - \lambda x\| &= \|\lambda_n x_n - \lambda x_n + \lambda x_n - \lambda x\| \\ &\leq \|\lambda_n x_n - \lambda x_n\| + \|\lambda x_n - \lambda x\| \\ &\leq \|(\lambda_n - \lambda) x_n\| + \|\lambda(x_n - x)\| \\ &= |\lambda_n - \lambda|\|x_n\| + |\lambda|\|x_n - x\| \\ &\leq |\lambda_n - \lambda| M + |\lambda|\|x_n - x\| \\ &< \varepsilon/2 + \varepsilon/2 \\ &= \varepsilon. \end{aligned}$$

4. Since $y_n \to y$, there exists $M \geq 0$ such that $\|y_n\| \leq M$ for all $n \in \mathbb{N}$. Choose such an M and let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $\|x_n - x\| < \varepsilon/2M$ and $\|y_n - y\| < \varepsilon/2\|x\|$. Then $n \geq N$ implies

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n\rangle - \langle x, y_n\rangle + \langle x, y_n\rangle - \langle x, y\rangle| \\ &\leq |\langle x_n, y_n\rangle - \langle x, y_n\rangle| + |\langle x, y_n\rangle - \langle x, y\rangle| \\ &= |\langle x_n - x, y_n\rangle| + |\langle x, y_n - y\rangle| \\ &\leq \|x_n - x\|\|y_n\| + \|x\|\|y_n - y\| \\ &\leq \|x_n - x\| M + \|x\|\|y_n - y\| \\ &< \varepsilon/2 + \varepsilon/2 \\ &= \varepsilon. \end{aligned}$$

   To see that $\|x_n\| \to \|x\|$, we just set $y_n = x_n$. Then

$$\begin{aligned} \|x_n\| &= \sqrt{\langle x_n, x_n\rangle} \\ &\to \sqrt{\langle x, x\rangle} \\ &= \|x\|, \end{aligned}$$

   where we were allowed to take limits inside the square root function since the square root function is continuous on $\mathbb{R}_{\geq 0}$.

$\square$

Definition 1.6. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on a vector space V. We say $\|\cdot\|_1$ is **stronger** than $\|\cdot\|_2$ (or $\|\cdot\|_2$ is **weaker** than $\|\cdot\|_1$) if there exists a constant $C > 0$ such that

$$\|x\|_2 \leq C\|x\|_1$$

for all $x \in V$.

Remark. Observe if a sequence (x_n) in V converges to x in the metric spaced induced by the $\|\cdot\|_1$ norm, then it also converges in the metric space induced by the $\|\cdot\|_2$ norm. Indeed, let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $\|x_n - x\|_1 < \varepsilon/C$. Then $n \geq N$ implies

$$\begin{aligned} \|x_n - x\|_2 &\leq C\|x_n - x\|_1 \\ &< \varepsilon. \end{aligned}$$

Thus a sequence (x_n) converging in the $\|\cdot\|_1$ norm is a *stronger* condition than the sequence (x_n) converging in the $\|\cdot\|_2$ norm. An alternative way of thinking about this is that the topology induced by the $\|\cdot\|_1$ norm is *finer* than the topology induced by the $\|\cdot\|_2$ norm. Indeed, if $B_r^2(0)$ is the open ball of radius r centered at 0 in the $\|\cdot\|_2$ norm and $B_{r/C}^1(0)$ is the open ball of radius r/C in the $\|\cdot\|_1$ norm, then we have

$$B_r^2(0) \subseteq B_{r/C}^1(0).$$

More generally, if $a \in V$, then

$$\begin{aligned} B_r^2(a) &= a + B_r^2(0) \\ &\subseteq a + B_{r/C}^1(0) \\ &= B_{r/C}^1(a). \end{aligned}$$

This implies the topology induced by $\|\cdot\|_1$ is finer than the topology induced by $\|\cdot\|_2$.

1.3 Closure

Definition 1.7. Let $A \subseteq \mathcal{V}$. A point $x \in \mathcal{V}$ is said to be a **closure point** of A if every open neighborhood of x meets A. This means that for any $\varepsilon > 0$, we have $B_\varepsilon(x) \cap A \neq \emptyset$, where

$$B_\varepsilon(x) := \{y \in \mathcal{V} \mid \|x - y\| < \varepsilon\}.$$

The **closure of** A is the set of all closure points of A and is denoted $\overline{A}$.

Remark. If every open neighborhood of a point $x \in \mathcal{V}$ meets A, then we can find a sequence (x_n) of points in A such that $x_n \to x$. Conversely, if (x_n) is a sequence of points in A such that $x_n \to x$, then every open neighborhood of x meets A. Thus, an equivalent condition for x to be a closure point of A is that there exists a sequence (x_n) of points in A such that $x_n \to x$. We will prove this in a moment.

Proposition 1.10. *Let $A, B \subseteq \mathcal{V}$. Then*

1. $A \subseteq \overline{A}$;
2. *If* $A \subseteq B$, *then* $\overline{A} \subseteq \overline{B}$;
3. $\overline{A \cup B} = \overline{A} \cup \overline{B}$;
4. $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$;
5. $\overline{\overline{A}} = \overline{A}$.

Proof.

1. Let $x \in A$. Then every open neighborhood of x meets A since, in particular, every open neighborhood meets $x \in A$. Therefore $x \in \overline{A}$.

2. Let $x \in \overline{A}$. Then every open neighborhood of x meets B since, in particular, each open neighborhood meets A. Therefore $x \in \overline{B}$.

3. First note that $\overline{A \cup B} \supseteq \overline{A} \cup \overline{B}$ follows from 2. For the reverse inclusion, let $x \in \overline{A \cup B}$ and assume for a contradiction that $x \notin \overline{A} \cup \overline{B}$. Then there exists a open neighborhood U of x such that $U \cap A = \emptyset$ and

an open neighborhood V of x such that $V \cap A = \emptyset$. Choose such open neighborhoods U and V and set $W = U \cap V$. Then W is a open neighborhood of x such that

$$W \cap (A \cup B) = (W \cap A) \cup (W \cap B) = \emptyset,$$

which contradicts the fact that $x \in \overline{A \cup B}$.

4. This follows from 2. To see why we do not necessarily get the reverse equality, consider $A = (0,1)$ and $B = (1,2)$ in $\mathbb{R}$. Then $\overline{A} \cap \overline{B} = \{1\}$ but $\overline{A \cap B} = \overline{\emptyset} = \emptyset$.

5. Exercise.

$\square$

Proposition 1.11. *$x \in \overline{A}$ if and only if there exists a sequence (x_n) of elements in A such that $x_n \to x$.*

Proof. Assume $x \in \overline{A}$. For each $n \in \mathbb{N}$, choose $x_n \in B_{1/n}(x) \cap A$ (this set is nonempty since $x \in \overline{A}$). Then one readily checks that $x_n \to x$.

Conversely, suppose that (x_n) in A such that $x_n \to x$ and assume for a contradiction that $x \notin \overline{A}$. Then there exists an open neighborhood $B_\varepsilon(x)$ of x such that $B_\varepsilon(x) \cap A = \emptyset$. Choose such an open neighborhood and choose $N \in \mathbb{N}$ such that $1/N < \varepsilon$. Then $n \geq N$ implies

$$x_n \in B_{1/N}(x) \cap A \subseteq B_\varepsilon(x) \cap A,$$

which is a contradiction. $\square$

Definition 1.8. A set $A \subseteq \mathcal{V}$ is said to be a **closed set** if $A = \overline{A}$.

In finite dimensional inner-product spaces every subspace is a closed set. However this is no longer true in infinite dimensional spaces.

Example 1.1. Let $\ell_0(\mathbb{N})$ be the subset of $\ell^2(\mathbb{N})$ consisting of all square summable sequences (x_n) with only finitely many nonzero terms. It is easy to see that $\ell_0(\mathbb{N})$ is in fact a subspace of $\ell^2(\mathbb{N})$. However $\ell_0(\mathbb{N})$ is not a *closed* subspace of $\ell^2(\mathbb{N})$. Indeed, consider the sequence of elements in $\ell_0(\mathbb{N})$ given by

$$\begin{aligned} x^1 &= (1, 0, 0, 0, \cdots) \\ x^2 &= (1, 1/2, 0, 0, \cdots) \\ x^3 &= (1, 1/2, 1/3, 0, \dots) \\ &\vdots \end{aligned}$$

Then the sequence (x^n) of sequences x^n converges to the sequence $(1/n) \notin \ell_0(\mathbb{N})$. Therefore $\ell_0(\mathbb{N})$ is not closed.

To see why we have $(x^n) \to (1/n)$, let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $\sum_{k=N}^\infty 1/k^2 < \varepsilon$ (there exists such an N since $\sum_{k=0}^\infty 1/k^2 < \infty$). Then for all $n \geq N$, we have

$$\|(x^n) - (1/n)\| \leq \sum_{k=N}^\infty \frac{1}{k^2} < \varepsilon.$$

Theorem 1.3. *Let $\mathcal{U}$ be a subspace of $\mathcal{V}$. Then $\overline{\mathcal{U}}$ is a subspace of $\mathcal{V}$.*

Proof. Let $x, y \in \overline{\mathcal{U}}$ and $\lambda \in \mathbb{C}$. Let (x_n) and (y_n) be two sequences of elements in $\mathcal{U}$ such that $x_n \to x$ and $y_n \to y$. Then $(\lambda x_n + y_n)$ is a sequence of elements in $\mathcal{U}$ such that $\lambda x_n + y_n \to \lambda x + y$. Therefore $\lambda x + y \in \overline{\mathcal{U}}$, which implies $\overline{\mathcal{U}}$ is a subspace of $\mathcal{V}$. $\square$

Proposition 1.12. *If $\mathcal{U}$ is a subspace of $\mathcal{V}$, then the closure $\overline{\mathcal{V}}$ is a closed subspace.*

Proof. Clearly $\overline{\mathcal{U}}$ is a closed set. Therefore it suffices to show that $\overline{\mathcal{U}}$ is a subspace of $\mathcal{V}$. Let $x, y \in \overline{\mathcal{U}}$ and let $\lambda \in \mathbb{C}$. Choose sequence (x_n) and (y_n) in $\mathcal{U}$ such that $x_n \to x$ and $y_n \to y$. Then since $\lambda x_n + y_n \in \mathcal{U}$ for each $n \in \mathbb{N}$ and

$$\lim_{n \to \infty} (\lambda x_n + y_n) = \lambda x + y,$$

we have $\lambda x + y \in \overline{\mathcal{U}}$. Therefore $\overline{\mathcal{U}}$ is a subspace of $\mathscr{V}$. $\square$

2 Hilbert Spaces

Let $\mathcal{V}$ be an inner-product space. A sequence $(x_n) \in \mathcal{V}$ is said to be a **Cauchy sequence** if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$n, m \geq N \text{ implies } \|x_n - x_m\| < \varepsilon.$$

Every convergent sequence in $\mathcal{V}$ is a Cauchy sequence. Indeed:

Proposition 2.1. *Let $\mathcal{V}$ be an inner-product space and let (x_n) be a convergent sequence in $\mathcal{V}$. Then (x_n) is a Cauchy sequence.*

Proof. Let $x \in \mathcal{V}$ be the limit of (x_n) and let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $\|x_n - x\| < \varepsilon/2$. Then $m, n \geq N$ implies

$$\begin{aligned}
\|x_n - x_m\| &= \|x_n - x + x - x_m\| \\
&\leq \|x_n - x\| + \|x - x_m\| \\
&< \varepsilon/2 + \varepsilon/2 \\
&= \varepsilon.
\end{aligned}$$

It follows that (x_n) is Cauchy. □

Though every convergent sequence is Cauchy, the converse need not hold. For instance, in $\mathbb{Q}$, we can construct a sequence of rational numbers which gets closer and closer to π. Namely, such a sequence starts out as

$$(3, 3.1, 3.14, 3.141, \ldots). \tag{3}$$

It's easy to see that such a sequence is in fact a Cauchy sequence. However, by construction, this sequence converges to π, which is not a rational number. On the other hand, since π is a real number, the Cauchy sequence (3) of real numbers converges to a real number. Is every Cauchy sequence in $\mathbb{R}$ convergent? It turns out that the answer to this is, yes! We give a special name to inner-product spaces which share this property with $\mathbb{R}$.

Definition 2.1. Let $\mathcal{V}$ be an inner-product space. We say $\mathcal{V}$ is a **Hilbert space** if every Cauchy sequence in $\mathcal{V}$ converges to a limit in $\mathcal{V}$.

The most basic examples with Hilbert spaces are $\mathbb{R}$, $\mathbb{R}^n$, and $\mathbb{C}^n$ (with their usual inner-product). We will show later on that $\ell^2(\mathbb{N})$ is also a Hilbert space. Nonexamples of Hilbert spaces include $\mathbb{Q}$ and $C[a, b]$. Hilbert spaces are usually denoted by $\mathcal{H}$ and $\mathcal{K}$.

2.1 Distances

Definition 2.2. Let $\mathcal{V}$ be an inner-product space, let $A, B \subseteq \mathcal{V}$, and let $x \in \mathcal{V}$. We define the **distance from** x **to** A, denoted $\mathrm{d}(x, A)$, by

$$\mathrm{d}(x, A) := \inf\{\|x - a\| \mid a \in A\}.$$

More generally, we defined the **distance from** A **to** B, denoted $\mathrm{d}(A, B)$, by

$$\mathrm{d}(A, B) := \inf\{\|a - b\| \mid a \in A \quad \text{and} \quad b \in B\}.$$

Proposition 2.2. *With the notation as in Definition (2.2), we have $\mathrm{d}(x, A) = 0$ if and only if $x \in \overline{A}$.*

Proof. Suppose $\mathrm{d}(x, A) = 0$. Then for each $n \in \mathbb{N}$ we can choose $a_n \in A$ such that

$$\mathrm{d}(x, A) \leq \|x - a_n\| < \mathrm{d}(x, A) + \frac{1}{n}.$$

In other words, since $\mathrm{d}(x, A) = 0$, we can find a sequence (a_n) in A such that

$$\|x - a_n\| < \frac{1}{n}$$

for all $n \in \mathbb{N}$. This implies $x \in \overline{A}$.

Conversely, suppose $x \in \overline{A}$. Then we can find a sequence (a_n) in A such that

$$\mathrm{d}(x, A) \leq \|x - a_n\| < \frac{1}{n}$$

for all $n \in \mathbb{N}$. Taking $n \to \infty$ gives us $\mathrm{d}(x, A) = 0$. □

Remark. A common technique that we do is we choose a sequence (a_n) of elements in A such that

$$\|x - a_n\| < d(x, A) + \frac{1}{n}$$

for all $n \in \mathbb{N}$. Such a sequence must exist since otherwise there would exist an $n \in \mathbb{N}$ such that

$$d(x, A) + \frac{1}{n} \leq \|x - a\|$$

for all $a \in A$, and this contradicts the fact that $d(x, A)$ is the infimum.

2.1.1 Absolute Homogeneity of Distances

Proposition 2.3. *Let $\mathcal{V}$ be an inner-product space, let $\mathcal{A}$ be a subspace of $\mathcal{V}$, let $x \in \mathcal{V}$, and let $\lambda \in \mathbb{C}$. Then*

$$d(\lambda x, \mathcal{A}) = |\lambda| d(x, \mathcal{A}).$$

Proof. Choose a sequence (y_n) of elements in $\mathcal{A}$ such that

$$\|x - y_n\| < d(x, \mathcal{A}) + \frac{1}{|\lambda n|}$$

for all $n \in \mathbb{N}$. Then since $\mathcal{A}$ is a subspace, we have

$$\begin{aligned} d(\lambda x, \mathcal{A}) &\leq \|\lambda x - \lambda y_n\| \\ &= |\lambda| \|x - y_n\| \\ &< |\lambda| \left(d(x, \mathcal{A}) + \frac{1}{|\lambda n|} \right) \\ &= |\lambda| d(x, \mathcal{A}) + \frac{1}{n} \end{aligned}$$

for all $n \in \mathbb{N}$. In particular, this implies $d(\lambda x, \mathcal{A}) \leq |\lambda| d(x, \mathcal{A})$.

Conversely, choose a sequence (z_n) of elements in $\mathcal{A}$ such that

$$\|\lambda x - z_n\| < d(\lambda x, \mathcal{A}) + \frac{1}{n}$$

Then since $\mathcal{A}$ is a subspace, we have

$$\begin{aligned} |\lambda| d(x, \mathcal{A}) &\leq |\lambda| \|x - z_n / |\lambda|\| \\ &= \|\lambda x - z_n\| \\ &< d(\lambda x, \mathcal{A}) + \frac{1}{n} \end{aligned}$$

for all $n \in \mathbb{N}$. In particular, this implies $|\lambda| d(x, \mathcal{A}) \leq d(\lambda x, \mathcal{A})$. □

2.1.2 Subadditivity of Distances

Proposition 2.4. *Let $\mathcal{V}$ be an inner-product space, let $\mathcal{A}$ be a subspace of $\mathcal{V}$, and let $x, y \in \mathcal{V}$. Then*

$$d(x + y, \mathcal{A}) \leq d(x, \mathcal{A}) + d(y, \mathcal{A}).$$

Proof. Choose a sequences (w_n) and (z_n) of elements in $\mathcal{A}$ such that

$$\|x - w_n\| < d(x, \mathcal{A}) + \frac{1}{2n} \quad \text{and} \quad \|y - z_n\| < d(y, \mathcal{A}) + \frac{1}{2n}$$

for all $n \in \mathbb{N}$. Then since $\mathcal{A}$ is a subspace, we have

$$\begin{aligned} d(x + y, \mathcal{A}) &\leq \|(x + y) - (w_n + z_n)\| \\ &\leq \|x - w_n\| + \|y - z_n\| \\ &< d(x, \mathcal{A}) + \frac{1}{2n} + d(y, \mathcal{A}) + \frac{1}{2n} \\ &= d(x, \mathcal{A}) + d(y, \mathcal{A}) + \frac{1}{n} \end{aligned}$$

for all $n \in \mathbb{N}$. In particular, this implies $d(x + y, \mathcal{A}) \leq d(x, \mathcal{A}) + d(y, \mathcal{A})$. □

2.2 Orthogonal Projection

Let K be a 2-dimensional subspace in $\mathbb{R}^3$. Such a subspace corresponds to a plane in $\mathbb{R}^3$ which passes through the origin. One of the main tools that we use in this setting is the concept of projecting onto K. For instance, if K corresponds to the plane $\{z = 0\}$ in $\mathbb{R}^3$, then the projection of the vector $(1,2,1)^\top$ onto K gives the vector $(1,2,0)^\top$. In fact, $(1,2,0)^\top$ is the *closest* vector to $(1,2,1)^\top$ which belongs to K. In other words, if $(a,b,0)^\top \in K$, then

$$\begin{aligned}\sqrt{(1-a)^2+(2-b)^2+1} &= \|(1,2,1)^\top - (a,b,0)^\top\| \\ &\geq \|(1,2,1)^\top - (1,2,0)^\top\| \\ &= 1.\end{aligned}$$

We wish to generalize these concepts from $\mathbb{R}^3$ to an arbitrary Hilbert space.

Theorem 2.1. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{K}$ be a closed subspace of $\mathcal{H}$. Suppose $x \in \mathcal{H}\backslash\mathcal{K}$. Then there exists a unique $a \in \mathcal{K}$ such that $\mathrm{d}(x,\mathcal{K}) = \|x-a\|$.*

Proof. Choose a sequence (a_n) of elements in $\mathcal{K}$ such that

$$\mathrm{d}(x,\mathcal{K}) \leq \|x-a_n\| < \mathrm{d}(x,\mathcal{K}) + 1/n. \tag{4}$$

for all $n \in \mathbb{N}$. We claim that (a_n) is a Cauchy sequence. The key to showing this is to use the parallelogram identity, which says

$$\|y-z\|^2 + \|y+z\|^2 = 2\|y\|^2 + 2\|z\|^2 \tag{5}$$

for all $y,z \in \mathcal{H}$. Setting $y = x - a_m$ and $z = x - a_n$ in (5) and rearranging terms, we have

$$\begin{aligned}\|a_m - a_n\|^2 &= 2\|x-a_m\|^2 + 2\|x-a_n\|^2 - \|(x-a_m)+(x-a_n)\|^2 \\ &< 2(\mathrm{d}(x,\mathcal{K})+1/m)^2 + 2(\mathrm{d}(x,\mathcal{K})+1/n)^2 - \|(2(x-(a_n-a_m)/2)\|^2 \\ &= 4\mathrm{d}(x,\mathcal{K})^2 + (4/m+4/n)\mathrm{d}(x,\mathcal{K}) + 2/m^2 + 2/n^2 - 4\|(x-(a_n-a_m)/2\|^2 \\ &\leq 4\mathrm{d}(x,\mathcal{K})^2 + (4/m+4/n)\mathrm{d}(x,\mathcal{K}) + 2/m^2 + 2/n^2 - 4\mathrm{d}(x,\mathcal{K})^2 \\ &= (4/n+4/m)\mathrm{d}(x,\mathcal{K}) + 2/n^2 + 2/m^2.\end{aligned}$$

Thus if $\varepsilon > 0$, then we choose $N \in \mathbb{N}$ such that $n, m \geq N$ implies

$$(4/n+4/m)\mathrm{d}(x,\mathcal{K}) + 2/n^2 + 2/m^2 < \varepsilon^2.$$

which implies $\|a_m - a_n\| < \varepsilon$. Therefore the sequence (a_n) is a Cauchy sequence, and since we are in a Hilbert space, (a_n) must be convergent, say $a_n \to a$, with $a \in \mathcal{K}$ since $\mathcal{K}$ is closed. Then taking the limit of (4) as $n \to \infty$ gives us $\mathrm{d}(x,\mathcal{K}) = \|x-a\|$.

To see uniqueness of a, let b be another point in $\mathcal{K}$ such that $\|x-b\| = \mathrm{d}(x,\mathcal{K})$. We again use the parallelogram identity. We have

$$\begin{aligned}\|b-a\|^2 &= \|(x-a)-(x-b)\|^2 \\ &= 2\|x-a\|^2 + 2\|x-b\|^2 - \|(x-a)+(x-b)\|^2 \\ &= 4\mathrm{d}(x,\mathcal{K})^2 - \|2x-(a+b)\|^2 \\ &= 4\mathrm{d}(x,\mathcal{K})^2 - 4\|x-(a+b)/2\|^2 \\ &\leq 0,\end{aligned}$$

which implies $a = b$. □

Theorem 2.2. *Let $\mathcal{H}$ be a Hilbert space, let $\mathcal{K}$ be a closed subspace of $\mathcal{H}$, and let $x \in \mathcal{H}$. Then $\mathrm{P}_{\mathcal{K}}x$ is the unique point in $\mathcal{K}$ such that*

$$\langle x - \mathrm{P}_{\mathcal{K}}x, y\rangle = 0$$

for all $y \in \mathcal{K}$.

Proof. Let $y \in \mathcal{K}$. Define a function $f\colon \mathbb{R} \to \mathbb{R}$ by

$$f(t) = \|x - \mathrm{P}_{\mathcal{K}}x + ty\|^2$$

for all $t \in \mathbb{R}$. Then $f(0) = \mathrm{d}(x, \mathcal{K})^2$ and $f(t) > \mathrm{d}(x, \mathcal{K})^2$ for all $t \neq 0$ (note we have a *strict* inequality here by uniqueness of $\mathrm{P}_{\mathcal{K}} x$). For each $t \in \mathbb{R}$, we have

$$\begin{aligned} f(t) &= \|x - \mathrm{P}_{\mathcal{K}} x + ty\|^2 \\ &= \langle x - \mathrm{P}_{\mathcal{K}} x + ty, x - \mathrm{P}_{\mathcal{K}} x + ty \rangle \\ &= \|x - \mathrm{P}_{\mathcal{K}} x\|^2 + 2\mathrm{Re}(\langle x - \mathrm{P}_{\mathcal{K}} x, ty \rangle) + \|ty\|^2 \\ &= \mathrm{d}(x, \mathcal{K})^2 + 2t\mathrm{Re}(\langle x - \mathrm{P}_{\mathcal{K}} x, y \rangle) + t^2 \|y\|^2. \end{aligned}$$

So the function f is just a quadratic function in t. In particular, it is differentiable at $t = 0$, and since it has a global minimum at $t = 0$, we have

$$\begin{aligned} 0 &= f'(t)\Big|_{t=0} \\ &= \Big(2\mathrm{Re}(\langle x - \mathrm{P}_{\mathcal{K}} x, y \rangle) + 2t\|y\|^2\Big)\Big|_{t=0} \\ &= 2\mathrm{Re}(\langle x - \mathrm{P}_{\mathcal{K}} x, y \rangle). \end{aligned}$$

Therefore $\mathrm{Re}\langle x - \mathrm{P}_{\mathcal{K}} x, y \rangle = 0$ for all $y \in \mathcal{K}$. Note that this also implies

$$\begin{aligned} \mathrm{Im}(\langle x - \mathrm{P}_{\mathcal{K}} x, y \rangle) &= -\mathrm{Re}(i\langle x - \mathrm{P}_{\mathcal{K}} x, y \rangle) \\ &= -\mathrm{Re}(\langle x - \mathrm{P}_{\mathcal{K}} x, -iy \rangle) \\ &= 0 \end{aligned}$$

for all $y \in \mathcal{K}$. Combining these together gives us $\langle x - \mathrm{P}_{\mathcal{K}} x, y \rangle = 0$ for all $y \in \mathcal{K}$.

This proves existence. To prove uniqueness, let $a \in \mathcal{K}$ such that $\langle x - a, y \rangle = 0$ for all $y \in \mathcal{K}$. Then

$$\begin{aligned} 0 &= \langle x - a, y \rangle \\ &= \langle x - \mathrm{P}_{\mathcal{K}} x + \mathrm{P}_{\mathcal{K}} x - a, y \rangle \\ &= \langle x - \mathrm{P}_{\mathcal{K}} x, y \rangle + \langle \mathrm{P}_{\mathcal{K}} x - a, y \rangle \\ &= \langle \mathrm{P}_{\mathcal{K}} x - a, y \rangle \end{aligned}$$

for all $y \in \mathcal{K}$. In particular, setting $y = \mathrm{P}_{\mathcal{K}} - a$ gives us $\|\mathrm{P}_{\mathcal{K}} x - a\|^2 = 0$, which implies $\mathrm{P}_{\mathcal{K}} x = a$. □

2.2.1 The Orthogonal Projection Map

The unique point a in Theorem (2.1) is called the **orthogonal projection** x onto the subspace $\mathcal{K}$ and is denoted by $\mathrm{P}_{\mathcal{K}} x$. More generally, we define the map $\mathrm{P}_{\mathcal{K}} \colon \mathcal{H} \to \mathcal{K}$ by $x \mapsto \mathrm{P}_{\mathcal{K}} x$ for all $x \in \mathcal{H}$. The map $\mathrm{P}_{\mathcal{K}}$ is called the **orthogonal projection** with respect to $\mathcal{K}$. Actually in linear analysis, the words "orthogonal projection" describe a certain class of linear maps:

Definition 2.3. Let $\mathcal{V}$ be an inner-product space and let $P \colon \mathcal{V} \to \mathcal{V}$ be a linear map. We say P is a **projection** if $P^2 = P$. We say P is an **orthogonal projection** if it is a projection and it satisfies

$$\langle Px, y \rangle = \langle x, Py \rangle$$

for all $x, y \in V$.

We need to make sure that our terminology is not contradictory. So we need to justify that $\mathrm{P}_{\mathcal{K}}$ really is an orthononal projection in the sense of Definition (2.3). This will be established in the next proposition.

Proposition 2.5. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{K}$ be a closed subspace of $\mathcal{H}$. Then*

1. $\mathrm{P}_{\mathcal{K}}$ *is $\mathbb{C}$-linear, that is,* $\mathrm{P}_{\mathcal{K}}(\lambda x + \mu y) = \lambda \mathrm{P}_{\mathcal{K}} x + \mu \mathrm{P}_{\mathcal{K}} y$ *for all $x, y \in \mathcal{H}$ and $\lambda, \mu \in \mathbb{C}$.*
2. $\|\mathrm{P}_{\mathcal{K}} x\| \leq \|x\|$ *for all $x \in \mathcal{H}$.*
3. *Let $x \in \mathcal{H}$. Then $\mathrm{P}_{\mathcal{K}} x = x$ if and only if $x \in \mathcal{K}$.*
4. $\mathrm{P}_{\mathcal{K}}$ *is an orthogonal projection, that is,* $\mathrm{P}_{\mathcal{K}}(\mathrm{P}_{\mathcal{K}} x) = \mathrm{P}_{\mathcal{K}} x$ *and* $\langle \mathrm{P}_{\mathcal{K}} x, y \rangle = \langle x, \mathrm{P}_{\mathcal{K}} y \rangle$ *for all $x, y \in \mathcal{H}$.*

Proof. 1. Let $x, y \in \mathcal{H}$ and $\lambda, \mu \in \mathbb{C}$. Then for all $z \in \mathcal{K}$, we have

$$\begin{aligned} \langle \lambda x + \mu y - \lambda \mathrm{P}_{\mathcal{K}} x - \mu \mathrm{P}_{\mathcal{K}} y, z \rangle &= \langle \lambda x - \lambda \mathrm{P}_{\mathcal{K}} x, z \rangle + \langle \mu y - \mu \mathrm{P}_{\mathcal{K}} y, z \rangle \\ &= \lambda \langle x - \mathrm{P}_{\mathcal{K}} x, z \rangle + \mu \langle y - \mathrm{P}_{\mathcal{K}} y, z \rangle \\ &= \lambda \cdot 0 + \mu \cdot 0 \\ &= 0. \end{aligned}$$

In particular, this implies

$$\lambda P_{\mathcal{K}}x + \mu P_{\mathcal{K}}y = P_{\mathcal{K}}(\lambda x + \mu y)$$

by Proposition (2.2).

2. Let $x \in \mathcal{H}$. By the Pythagorean theorem, we have

$$\begin{aligned} \|x\|^2 &= \|P_{\mathcal{K}}x\|^2 + \|P_{\mathcal{K}}x - x\|^2 \\ &\geq \|P_{\mathcal{K}}x\|^2, \end{aligned}$$

which implies $\|P_{\mathcal{K}}x\| \leq \|x\|$.

3. Suppose $P_{\mathcal{K}}x = x$. Then $x \in \mathcal{K}$ since $P_{\mathcal{K}}x \in \mathcal{K}$. Conversely, suppose $x \in \mathcal{K}$. Then

$$\begin{aligned} 0 &= \|x - x\| \\ &\geq d(x, \mathcal{K}) \\ &= \|x - P_{\mathcal{K}}x\| \\ &\geq 0. \end{aligned}$$

It follows that $\|x - P_{\mathcal{K}}x\| = 0$, which implies $x = P_{\mathcal{K}}x$.

4. We first show $P_{\mathcal{K}}$ is a projection. Let $x \in \mathcal{H}$. Then since $P_{\mathcal{K}}x \in \mathcal{K}$, we have $P_{\mathcal{K}}(P_{\mathcal{K}}x) = P_{\mathcal{K}}x$ by part 3. Thus $P_{\mathcal{K}}$ is a projection. Now we show it is an orthogonal projection. Let $x, y \in \mathcal{H}$. Then

$$\begin{aligned} 0 &= \langle x - P_{\mathcal{K}}x, P_{\mathcal{K}}y \rangle \\ &= \langle x, P_{\mathcal{K}}y \rangle - \langle P_{\mathcal{K}}x, P_{\mathcal{K}}y \rangle \\ &= \langle x, P_{\mathcal{K}}y \rangle - \langle P_{\mathcal{K}}x, (P_{\mathcal{K}}y - y) + y \rangle \\ &= \langle x, P_{\mathcal{K}}y \rangle - \langle P_{\mathcal{K}}x, P_{\mathcal{K}}y - y \rangle - \langle P_{\mathcal{K}}x, y \rangle \\ &= \langle x, P_{\mathcal{K}}y \rangle - \langle P_{\mathcal{K}}x, y \rangle, \end{aligned}$$

which implies $\langle x, P_{\mathcal{K}}y \rangle = \langle P_{\mathcal{K}}x, y \rangle$. Thus $P_{\mathcal{K}}$ is an orthogonal projection. □

2.2.2 Orthogonal Complements

Definition 2.4. Let $\mathcal{V}$ be an inner-product space, $x, y \in \mathcal{V}$, and let $A, B \subseteq \mathcal{V}$.

1. We say x is **orthonal to** y, denoted $x \perp y$, if $\langle x, y \rangle = 0$.
2. We say x is **orthogonal to** A, denoted $x \perp A$, if $\langle x, a \rangle = 0$ for all $a \in A$.
3. We say A is **orthogonal to** B, denoted $A \perp B$, if $\langle a, b \rangle = 0$ for all $a \in A$ and $b \in B$.
4. The **orthogonal complement of** A, denoted $A^{\perp}$, is defined to be

$$A^{\perp} := \{z \in \mathcal{V} \mid z \perp A\}.$$

Theorem 2.3. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{K} \subseteq \mathcal{L} \subseteq \mathcal{H}$. Then*

1. *we have $\mathcal{K}^{\perp} \supseteq \mathcal{L}^{\perp}$.*
2. *$\mathcal{K}^{\perp}$ is a closed subspace of $\mathcal{H}$.*
3. *If $\mathcal{K}$ is a closed subspace of $\mathcal{H}$, then every $x \in \mathcal{H}$ can be decomposed in a* unique *way as a sum of a vector in $\mathcal{K}$ and a vector in $\mathcal{K}^{\perp}$. In other words, we have $\mathcal{H} = \mathcal{K} \oplus \mathcal{K}^{\perp}$.*
4. *If $\mathcal{K}$ is a closed subspace of $\mathcal{H}$, then $(\mathcal{K}^{\perp})^{\perp} = \mathcal{K}$.*

Proof. 1. We have

$$\begin{aligned} x \in \mathcal{L}^{\perp} &\implies \langle x, y \rangle = 0 \text{ for all } y \in \mathcal{L} \\ &\implies \langle x, y \rangle = 0 \text{ for all } y \in \mathcal{K} \\ &\implies x \in \mathcal{K}^{\perp}. \end{aligned}$$

Thus $\mathcal{K}^{\perp} \supseteq \mathcal{L}^{\perp}$.

2. First we show that $\mathcal{K}^\perp$ is a subspace of $\mathcal{V}$. Let $x, z \in \mathcal{K}^\perp$ and $\lambda \in \mathbb{C}$. Then

$$\begin{aligned}\langle x + \lambda z, y\rangle &= \langle x, y\rangle + \lambda\langle z, y\rangle \\ &= 0\end{aligned}$$

for all $y \in \mathcal{K}$. This implies $\mathcal{K}^\perp$ is a subspace of $\mathcal{V}$. Now we will show that $\mathcal{K}^\perp$ is closed. Let (x_n) be a sequence of points in $\mathcal{K}^\perp$ such that $x_n \to x$ for some $x \in \mathcal{H}$. Then since $\langle x_n, y\rangle = 0$ for all $n \in \mathbb{N}$ and $y \in \mathcal{K}$, we have

$$\begin{aligned}\langle x, y\rangle &= \lim_{n\to\infty} \langle x_n, y\rangle \\ &= \lim_{n\to\infty} 0 \\ &= 0.\end{aligned}$$

for all $y \in \mathcal{K}$. Therefore $x \in \mathcal{K}^\perp$, which implies $\mathcal{K}^\perp$ is closed.

3. Let $x \in \mathcal{H}$. Then

$$x = \mathrm{P}_{\mathcal{K}}x + (x - \mathrm{P}_{\mathcal{K}}x),$$

where $\mathrm{P}_{\mathcal{K}}x \in \mathcal{K}$ and where $x - \mathrm{P}_{\mathcal{K}}x \in \mathcal{K}^\perp$ by Theorem (2.2). This establishes existence. For uniqueness, first note that $\mathcal{K} \cap \mathcal{K}^\perp = \{0\}$. Indeed, if $y \in \mathcal{K} \cap \mathcal{K}^\perp$, then we must have $\langle y, y\rangle = 0$, which implies $y = 0$. Now suppose that

$$x = y + z$$

is another decomposition of x where $y \in \mathcal{K}$ and $z \in \mathcal{K}^\perp$. Then we have

$$\mathrm{P}_{\mathcal{K}}x + (x - \mathrm{P}_{\mathcal{K}}x) = x = y + z,$$

which implies

$$\mathrm{P}_{\mathcal{K}}x - y = (x - \mathrm{P}_{\mathcal{K}}x) - z.$$

In particular, we see that

$$\mathrm{P}_{\mathcal{K}}x - y \in \mathcal{K} \cap \mathcal{K}^\perp = \{0\} \quad \text{and} \quad (x - \mathrm{P}_{\mathcal{K}}x) - z \in \mathcal{K} \cap \mathcal{K}^\perp = \{0\}.$$

So $\mathrm{P}_{\mathcal{K}}x = y$ and $(x - \mathrm{P}_{\mathcal{K}}x) = z$. This establishes uniqueness.

4. Let $x \in \mathcal{K}$. Then $\langle x, y\rangle = 0$ for all $y \in \mathcal{K}^\perp$. Thus $x \in (\mathcal{K}^\perp)^\perp$, and so $\mathcal{K} \subseteq (\mathcal{K}^\perp)^\perp$. Conversely, let $x \in (\mathcal{K}^\perp)^\perp$. Then $\langle x, y\rangle = 0$ for all $y \in \mathcal{K}^\perp$. In particular, we have

$$\begin{aligned}\|x - \mathrm{P}_{\mathcal{K}}x\|^2 &= \langle x - \mathrm{P}_{\mathcal{K}}x, x - \mathrm{P}_{\mathcal{K}}x\rangle \\ &= \langle x, x - \mathrm{P}_{\mathcal{K}}x\rangle - \langle \mathrm{P}_{\mathcal{K}}x, x - \mathrm{P}_{\mathcal{K}}x\rangle \\ &= 0 - 0 \\ &= 0,\end{aligned}$$

which implies $x = \mathrm{P}_{\mathcal{K}}x$. This implies $x \in \mathcal{K}$, and hence $(\mathcal{K}^\perp)^\perp \subseteq \mathcal{K}$. □

2.3 Separable Hilbert Spaces

Definition 2.5. Let $\mathcal{V}$ be an inner-product space and let $\mathcal{H}$ be a Hilbert space.

1. A sequence (e_n) of vectors in $\mathcal{V}$ is said to be an **orthonormal sequence** if

$$\langle e_m, e_n\rangle = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{else.}\end{cases}$$

2. A sequence (x_n) of vectors in $\mathcal{H}$ is said to be **complete** if

$$\overline{\operatorname{span}}(\{x_n \mid n \in \mathbb{N}\}) = \mathcal{H}.$$

3. A sequence (e_n) of vectors in $\mathcal{H}$ is said to be an **orthonormal basis** if it is both orthonormal and complete. If $\mathcal{H}$ contains an orthonormal basis, then we say $\mathcal{H}$ is **separable**.

To give motivation for what follows, let $\mathcal{H}$ be a separable Hilbert space and let (e_n) be an orthonormal basis for $\mathcal{H}$. We will show that every $x \in \mathcal{H}$ can be represented as

$$x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n. \tag{6}$$

where the $\langle x, e_n \rangle$ are uniquely determined (with respect to the orthonormal basis (e_n)). Moreover we will show that the sequence $(\langle x, e_n \rangle)$ of complex numbers is square summable, and so $(\langle x, e_n \rangle) \in \ell^2(\mathbb{N})$. We will arrive at this in the following steps:

2.3.1 Orthonormal Sequences

We first show that if $(a_n) \in \ell^2(\mathbb{N})$ and (e_n) is an orthonormal sequence of vectors in $\mathcal{H}$, then the infinite sum

$$\sum_{n=1}^{\infty} a_n e_n$$

converges, and hence (6) converges as long as $(\langle x, e_n \rangle) \in \ell^2(\mathbb{N})$.

Proposition 2.6. *Let $\mathcal{H}$ be a Hilbert space. Suppose $(a_n) \in \ell^2(\mathbb{N})$ and (e_n) is an orthonormal sequence of vectors in $\mathcal{H}$. Then the sequence of partial sums (s_N), where $s_N = \sum_{n=1}^{N} a_n e_n$, converges in $\mathcal{H}$ and the limit, which we denote by $\sum_{n=1}^{\infty} a_n e_n$, satisfies*

$$\left\| \sum_{n=1}^{\infty} a_n e_n \right\|^2 = \sum_{n=1}^{\infty} |a_n|^2. \tag{7}$$

Proof. We first show that the sequence of partial sums (s_N) converges in $\mathcal{H}$. To do this, we will show (s_N) is Cauchy. Let $\varepsilon > 0$. Since the sequence of partial sums (t_N) converges, where $t_N = \sum_{n=1}^{N} |a_n|^2$, there exists $N_0 \in \mathbb{N}$ such that $M, N \geq N_0$ (with $M \leq N$) implies $|t_N - t_M|^2 < \varepsilon$. Choose such an N_0. Then $M, N \geq N_0$ (with $M \leq N$) implies

$$\begin{aligned}
\|s_N - s_M\|^2 &= \left\| \sum_{n=M}^{N} a_n e_n \right\|^2 \\
&= \sum_{n=M}^{N} |a_n| \, \|e_n\|^2 \\
&= \sum_{n=M}^{N} |a_n|^2 \\
&= |t_N - t_M|^2 \\
&< \varepsilon,
\end{aligned}$$

where we used the Pythagorean Theorem to get from the first line to the second line. This implies (s_N) is a Cauchy sequence. To show (7), write

$$\begin{aligned}
\left\| \sum_{n=1}^{\infty} a_n e_n \right\|^2 &= \left\| \lim_{N\to\infty} \sum_{n=1}^{N} a_n e_n \right\|^2 \\
&= \lim_{N\to\infty} \left\| \sum_{n=1}^{N} a_n e_n \right\|^2 \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} |a_n|^2. \\
&= \sum_{n=1}^{\infty} |a_n|^2.
\end{aligned}$$

□

Proposition 2.7. *Suppose (e_n) is an orthonormal sequence in a Hilbert space $\mathcal{H}$. Then for every $x \in \mathcal{H}$ we have*

1. *(Bessel's Inequality) $\sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 \leq \|x\|^2$ (in particular $(\langle x, e_n \rangle) \in \ell^2(\mathbb{N})$);*
2. *$\sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$ exists.*

Proof. 1. For each $N \in \mathbb{N}$, we have

$$\begin{aligned}
0 &\leq \left\| x - \sum_{n=1}^{N} \langle x, e_n \rangle e_n \right\|^2 \\
&= \left\langle x - \sum_{n=1}^{N} \langle x, e_n \rangle e_n, x - \sum_{n=1}^{N} \langle x, e_n \rangle e_n \right\rangle \\
&= \|x\|^2 - 2\mathrm{Re} \left\langle x, \sum_{n=1}^{N} \langle x, e_n \rangle e_n \right\rangle + \left\| \sum_{n=1}^{N} \langle x, e_n \rangle e_n \right\|^2 \\
&= \|x\|^2 - 2\mathrm{Re} \sum_{n=1}^{N} \overline{\langle x, e_n \rangle} \langle x, e_n \rangle + \sum_{n=1}^{N} |\langle x, e_n \rangle|^2 \\
&= \|x\|^2 - 2 \sum_{n=1}^{N} |\langle x, e_n \rangle|^2 + \sum_{n=1}^{N} |\langle x, e_n \rangle|^2 \\
&= \|x\|^2 - \sum_{n=1}^{N} |\langle x, e_n \rangle|^2,
\end{aligned}$$

where we applied the Pythagorean Theorem to get to the fourth line from the third line. Since this holds for all $N \in \mathbb{N}$, this gives us Bessel's inequality

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

2. This follows from 1 and Proposition (2.6) with $(a_n) = (\langle x, x_n \rangle)$.

□

2.3.2 Complete Sequences

Here is a criterion to determine if a sequence in a Hilbert space is complete.

Proposition 2.8. *A sequence (x_n) in a Hilbert space $\mathcal{H}$ is complete if and only if the only $x \in \mathcal{H}$ with the property $\langle x, x_n \rangle = 0$ for all $n \in \mathbb{N}$ is $x = 0$.*

Proof. Let $E = \{x_n \mid n \in \mathbb{N}\}$. We first observe that

$$\begin{aligned}
\langle x, x_n \rangle = 0 \text{ for all } n \in \mathbb{N} &\iff x \in \mathrm{span}(E)^\perp \\
&\iff x \in \overline{\mathrm{span}}(E)^\perp.
\end{aligned}$$

Thus we are trying to show that $\mathcal{H} = \overline{\mathrm{span}}(E)$ if and only if $\overline{\mathrm{span}}(E)^\perp = 0$. If $\mathcal{H} = \overline{\mathrm{span}}(E)$, then

$$\begin{aligned}
0 &= \mathcal{H}^\perp \\
&= \overline{\mathrm{span}}(E).
\end{aligned}$$

Conversely, if $\overline{\mathrm{span}}(E)^\perp = 0$, then

$$\begin{aligned}
\mathcal{H} &= 0^\perp \\
&= (\overline{\mathrm{span}}(E)^\perp)^\perp \\
&= \overline{\mathrm{span}}(E),
\end{aligned}$$

where the last equality follows from the fact that $\overline{\mathrm{span}}(E)$ is a closed subspace. □

2.3.3 Unique Representations of Vectors in a Separable Hilbert Spaces

Theorem 2.4. *Let $\mathcal{H}$ be a separable Hilbert space and let (e_n) be an orthonormal basis for $\mathcal{H}$. Then for every $x \in \mathcal{H}$ we have*

$$x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n \quad \text{and} \quad \|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2.$$

In addition,

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

for all $x, y \in \mathcal{H}$.

Proof. Consider $y = x - \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n \in \mathcal{H}$. Then for all $k \in \mathbb{N}$, we have

$$\begin{aligned}
\langle y, e_k \rangle &= \left\langle x - \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n, e_k \right\rangle \\
&= \langle x, e_k \rangle - \left\langle \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n, e_k \right\rangle \\
&= \langle x, e_k \rangle - \left\langle \lim_{N \to \infty} \sum_{n=1}^{N} \langle x, e_n \rangle e_n, e_k \right\rangle \\
&= \langle x, e_k \rangle - \lim_{N \to \infty} \sum_{n=1}^{N} \langle x, e_n \rangle \langle e_n, e_k \rangle \\
&= \langle x, e_k \rangle - \langle x, e_k \rangle \\
&= 0.
\end{aligned}$$

Therefore $y = 0$ by Proposition (2.8).

Next, let $x, y \in \mathcal{H}$. Then

$$\begin{aligned}
\langle x, y \rangle &= \left\langle \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n, \sum_{k=1}^{\infty} \langle y, e_k \rangle e_k \right\rangle \\
&= \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \langle x, e_n \rangle \overline{\langle y, e_k \rangle} \langle e_n, e_k \rangle \\
&= \sum_{n=1}^{\infty} \langle x, e_n \rangle \overline{\langle y, e_n \rangle}.
\end{aligned}$$

To see that the $\langle x, e_n \rangle$ are uniquely determined. Suppose we have another representation of x, say

$$x = \sum_{n=1}^{\infty} \lambda_n e_n.$$

Then o

$$\begin{aligned}
0 &= \|0\| \\
&= \|x - \sum_{n=1}^{\infty} \lambda_n e_n\| \\
&= \| \sum_{n=1}^{\infty} \langle x, e_n \rangle x - \sum_{n=1}^{\infty} \lambda_n e_n \| \\
&=
\end{aligned}$$

□

2.3.4 Gram-Schmidt

We want to show that every inner-product space contains an orthonormal sequence.

Proposition 2.9. *Let $\{x_n \mid n \in \mathbb{N}\}$ be a linearly independent set of vectors in an inner-product space $\mathcal{V}$. Consider the so called Gram-Schmidt process: set $e_1 = \frac{1}{\|x_1\|} x_1$. Proceed inductively. If $e_1, e_2, \ldots, e_{n-1}$ are computed, compute e_n in two steps by*

$$f_n := x_n - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k, \text{ and then set } e_n := \frac{1}{\|f_n\|} f_n.$$

Then

1. *for every $N \in \mathbb{N}$ we have span$\{x_1, x_2, \ldots, x_N\} =$ span$\{e_1, e_2, \ldots, e_N\}$;*
2. *the set $\{e_n \mid n \in \mathbb{N}\}$ is an orthonormal set in $\mathcal{H}$;*

Remark. Note that $\|f_n\| \neq 0$ follows from linear independence of $\{x_n \mid n \in \mathbb{N}\}$.

Proof.

1. Let $N \in \mathbb{N}$. Then for each $1 \le n \le N$, we have

$$x_n = \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k + \|f_n\| e_n.$$

This implies $\operatorname{span}\{x_1, x_2, \ldots, x_N\} \subseteq \operatorname{span}\{e_1, e_2, \ldots, e_N\}$. We show the reverse inclusion by induction on n such that $1 \le n \le N$. The base case $n = 1$ being $\operatorname{span}\{x_1\} \supseteq \operatorname{span}\{e_1\}$, which holds since $e_1 = \frac{1}{\|x_1\|} x_1$. Now suppose for some n such that $1 \le n < N$ we have

$$\operatorname{span}\{x_1, x_2, \ldots, x_k\} \supseteq \operatorname{span}\{e_1, e_2, \ldots, e_k\} \tag{8}$$

for all $1 \le k \le n$. Then

$$e_{n+1} = \frac{1}{\|f_n\|} x_n - \sum_{k=1}^{n} \frac{1}{\|f_n\|} \langle x_n, e_k \rangle e_k \in \operatorname{span}\{x_1, x_2, \ldots, x_n\}.$$

where we used the induction step (8) on the e_k's ($1 \le k \le n$). Therefore

$$\operatorname{span}\{x_1, x_2, \ldots, x_k\} \supseteq \operatorname{span}\{e_1, e_2, \ldots, e_k\}$$

for all $1 \le k \le n+1$, and this proves our claim.

2. By construction, we have $\langle e_n, e_n \rangle = 1$ for all $n \in \mathbb{N}$. Thus, it remains to show that $\langle e_m, e_n \rangle = 0$ whenever $m \ne n$. We prove by induction on $n \ge 2$ that $\langle e_n, e_m \rangle = 0$ for all $m < n$. Proving this also give us $\langle e_m, e_n \rangle = 0$ for all $m < n$, since

$$\begin{aligned} \langle e_m, e_n \rangle &= \overline{\langle e_n, e_m \rangle} \\ &= \overline{0} \\ &= 0. \end{aligned}$$

The base case is

$$\begin{aligned} \langle e_2, e_1 \rangle &= \frac{1}{\|x_1\| \|f_2\|} \left\langle \left(x_2 - \frac{\langle x_2, x_1 \rangle}{\langle x_1, x_1 \rangle} x_1 \right), x_1 \right\rangle \\ &= \frac{1}{\|x_1\| \|f_2\|} \left(\langle x_2, x_1 \rangle - \langle x_2, x_1 \rangle \right) \\ &= 0 \end{aligned}$$

Now suppose that $n > 2$ and that $\langle e_n, e_m \rangle = 0$ for all $m < n$. Then

$$\begin{aligned} \langle e_{n+1}, e_m \rangle &= \frac{1}{\|f_{n+1}\|} \langle x_{n+1} - \sum_{k=1}^{n} \langle x_{n+1}, e_k \rangle e_k, e_m \rangle \\ &= \frac{1}{\|f_{n+1}\|} \left(\langle x_{n+1}, e_m \rangle - \sum_{k=1}^{n} \langle x_{n+1}, e_k \rangle \langle e_k, e_m \rangle \right) \\ &= \frac{1}{\|f_{n+1}\|} \left(\langle x_{n+1}, e_m \rangle - \langle x_{n+1}, e_m \rangle \langle e_m, e_m \rangle \right) \\ &= \frac{1}{\|f_{n+1}\|} \left(\langle x_{n+1}, e_m \rangle - \langle x_{n+1}, e_m \rangle \right) \\ &= 0, \end{aligned}$$

for all $m < n+1$, where we used the induction hypothesis to get from the second line to the third line. This proves the induction step, which finishes the proof of part 2 of the proposition.

3. By 2, we know that $\{e_n \mid n \in \mathbb{N}\}$ is an orthonormal set. Thus, it suffices to show that $\{e_n \mid n \in \mathbb{N}\}$ is complete. To do this, we use the criterion that the set $\{e_n \mid n \in \mathbb{N}\}$ is complete if and only if the only $x \in \mathcal{H}$ such that $\langle x, e_n \rangle = 0$ for all $n \in \mathbb{N}$ is $x = 0$.

Let $x \in \mathcal{H}$ and suppose $\langle x, e_n \rangle = 0$ for all $n \in \mathbb{N}$. Then

$$\begin{aligned} \langle x, x_n \rangle &= \left\langle x, \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k + \|f_n\| e_n \right\rangle \\ &= \sum_{k=1}^{n-1} \langle x_n, e_k \rangle \langle x, e_k \rangle + \|f_n\| \langle x, e_n \rangle \\ &= 0 \end{aligned}$$

for all $n \in \mathbb{N}$. Since $\{x_n \mid n \in \mathbb{N}\}$ is complete, this implies $x = 0$. Therefore $\{e_n \mid n \in \mathbb{N}\}$ is complete. □

3 Operators

In linear analysis, an **operator** is just another word for a linear map. We will stick to tradition and often refer to linear maps as operators.

3.1 Unitary Operators

Definition 3.1. An operator $U \colon \mathcal{V}_1 \to \mathcal{V}_2$ between inner-product spaces $\mathcal{V}_1$ and $\mathcal{V}_2$ is said to be **unitary** if it is an isomorphism as $\mathbb{C}$-vector spaces and it satisfies

$$\langle U(x), U(y) \rangle = \langle x, y \rangle$$

for all $x, y \in \mathcal{V}_1$. We say $\mathcal{V}_1$ and $\mathcal{V}_2$ are **unitarily equivalent** if there exists a unitary map $U \colon \mathcal{V}_1 \to \mathcal{V}_2$ which is unitary. In this case, the inverse map $U^{-1} \colon \mathcal{V}_2 \to \mathcal{V}_1$ is necessarily unitary (as one *should* check!).

Corollary. *Let $\mathcal{H}$ be an infinite dimensional separable Hilbert space. Then $\mathcal{H}$ is unitarily equivalent to $\ell^2(\mathbb{N})$.*

Proof. Let (e_n) be an orthonormal basis for $\mathcal{H}$. Define $U \colon \mathcal{H} \to \ell^2(\mathbb{N})$ by $U(x) = (\langle x, e_n \rangle)$ for all $x \in \mathcal{H}$. The proof that U is linear is very easy. That U is injective follows from completeness of (e_n) and Proposition (2.8). To show that U is surjective, let $(a_n) \in \ell^2(\mathbb{N})$. Since (e_n) is an orthonormal sequence, the series $x = \sum_{n=1}^{\infty} a_n e_n$ converges in $\mathcal{H}$ by Proposition (2.6). Then

$$\begin{aligned} U(x) &= (\langle x, e_n \rangle) \\ &= \left(\left\langle \sum_{m=1}^{\infty} a_m e_m, e_n \right\rangle \right) \\ &= \left(\sum_{m=1}^{\infty} a_n \langle e_m, e_n \rangle \right) \\ &= (a_n) \end{aligned}$$

implies U is surjective. Finally, we need to show that

$$\langle U(x), U(y) \rangle = \langle x, y \rangle$$

for all $x, y \in \mathcal{H}$. To see this, let $x, y \in \mathcal{H}$. Then

$$\begin{aligned} \langle U(x), U(y) \rangle &= \langle (\langle x, e_n \rangle), (\langle y, e_n \rangle) \rangle \\ &= \sum_{n=1}^{\infty} \langle x, e_n \rangle \overline{\langle y, e_n \rangle} \\ &= \langle x, y \rangle, \end{aligned}$$

by Theorem (2.4). □

Corollary. *$\ell^2(\mathbb{N})$ is a Hilbert space.*

Proof. Let $(a^k)_{k=1}^{\infty}$ be a Cauchy sequence in $\ell^2(\mathbb{N})$. Then for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $i, j \geq N$ implies $\|a^i - a^j\| < \varepsilon$. By the previous corollary, for each $k \in \mathbb{N}$ there exists $x_k \in \mathcal{H}$ such that $U(x_k) = a^k$. Then

$$\begin{aligned} \|a^i - a^j\| &= \|U(x_i) - U(x_j)\| \\ &= \|U(x_i - x_j)\| \\ &= \|x_i - x_j\| \end{aligned}$$

which implies $\|a^i - a^j\| = \|x_i - x_j\|$ for all $i, j \in \mathbb{N}$, and hence $i, j \geq N$ implies $\|x_i - x_j\| < \varepsilon$. So (x_n) is a Cauchy sequence in $\mathcal{H}$. Since $\mathcal{H}$ is a Hilbert space, the sequence (x_n) must converge, say $x_n \to x$. Then using the same argument that we just used it is easy to show $a^k \to U(x)$ as $k \to \infty$. So (a^k) is convergent, which implies $\ell^2(\mathbb{N})$ is a Hilbert space. □

Proposition 3.1. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{K}$ be a closed subspace. Suppose (e_n) is an orthonormal basis for $\mathcal{K}$. Then*

$$P_{\mathcal{K}}(x) = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$$

for all $x \in \mathcal{H}$.

Proof. We'll use the following fact that for any set $E \subseteq H$, we have $\overline{\operatorname{span}}(E)^{\perp} = E^{\perp}$. Recall that $\mathrm{P}_{\mathcal{K}} x$ is the unique vector such that

$$\langle x - \mathrm{P}_{\mathcal{K}} x, y \rangle = 0$$

for all $y \in \mathcal{K}$. So by the uniqueness, it suffices to show that $\sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$ has the same property. Then

$$\begin{aligned}
\left\langle x - \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n, e_k \right\rangle &= \langle x, e_k \rangle - \left\langle \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n, e_k \right\rangle \\
&= \langle x, e_k \rangle - \sum_{n=1}^{\infty} \langle x, e_n \rangle \langle e_n, e_k \rangle \\
&= \langle x, e_k \rangle - \langle x, e_k \rangle \\
&= 0.
\end{aligned}$$

for all $k \in \mathbb{N}$. In particular, this implies

$$\left\langle x - \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n, y \right\rangle = 0$$

for all $y \in \mathcal{K}$, which implies the proposition. □

3.2 Bounded Operators

Definition 3.2. Let $\mathcal{U}$ and $\mathcal{V}$ be normed linear spaces. A **bounded operator** $T \colon \mathcal{U} \to \mathcal{V}$ is a linear map such that

$$\sup \{ \|Tx\| \mid x \in \mathcal{U} \text{ and } \|x\| \leq 1 \} < \infty.$$

In this case we define

$$\|T\| := \sup \{ \|Tx\| \mid x \in \mathcal{U} \text{ and } \|x\| \leq 1 \}$$

and call it the **operator norm** of T.

3.2.1 Composition of Bounded Operators is Bounded

Proposition 3.2. *Let $T \colon \mathcal{H}_1 \to \mathcal{H}_2$ and $S \colon \mathcal{H} \to \mathcal{H}$ be two bounded operators. Then ST is bounded*

1. *TS is bounded and $\|TS\| \leq \|T\| \|S\|$;*
2. *$(TS)^* = S^* T^*$.*

Proof.

1. Let $x \in \mathcal{H}$ such that $\|x\| = 1$. Then

$$\begin{aligned}
\|TSx\| &\leq \|T\| \|Sx\| \\
&\leq \|T\| \|S\| \|x\| \\
&= \|T\| \|S\|.
\end{aligned}$$

Thus TS is bounded and $\|TS\| \leq \|T\| \|S\|$.

2. Let $y \in \mathcal{H}$. Then for all $x \in \mathcal{H}$, we have

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

In particular, this implies $(TS)^* y = S^* T^* y$ for all $y \in \mathcal{H}$, which implies $(TS)^* = S^* T^*$. □

Proposition 3.3. *Let $T \colon \mathcal{U} \to \mathcal{V}$ be a bounded operator. Then*

$$\|Tx\| \leq \|T\|\|x\| \tag{9}$$

for all $x \in \mathcal{U}$. *Moreover, let $M > 0$ and suppose there exists some $x_0 \in \mathcal{U}$ such that*

$$\|Tx_0\| \geq M\|x_0\|.$$

Then $M \leq \|T\|$.

Proof. Let $x \in \mathcal{U}$. If $x = 0$, then (9) is obvious, so assume $x \neq 0$. Then

$$\begin{aligned}\frac{\|Tx\|}{\|x\|} &= \left\| T\left(\frac{x}{\|x\|}\right) \right\| \\ &\leq \|T\|,\end{aligned}$$

and this implies (9).

For the latter statement, let $x_0 \in \mathcal{U}$ be such an element. Then

$$\begin{aligned}\|T\| &\geq \left\| T\left(\frac{x_0}{\|x_0\|}\right) \right\| \\ &= \frac{\|Tx_0\|}{\|x_0\|} \\ &\geq M.\end{aligned}$$

□

Proposition 3.4. *Let $\mathcal{H}$ be a Hilbert space and let $T \colon \mathcal{H} \to \mathcal{H}$ be a bounded operator. Then*

$$\|T\| = \sup\{\|Tx\| \mid \|x\| = 1\}.$$

Proof. First note that

$$\begin{aligned}\sup\{\|Tx\| \mid \|x\| = 1\} &\leq \sup\{\|Tx\| \mid \|x\| \leq 1\} \\ &= \|T\|.\end{aligned}$$

We prove the reverse inequality by contradiction. Assume that $\|T\| > \sup\{\|Tx\| \mid \|x\| = 1\}$. Choose $\varepsilon > 0$ such that

$$\|T\| - \varepsilon > \sup\{\|Tx\| \mid \|x\| = 1\} \tag{10}$$

Next, choose $x \in \mathcal{H}$ such that $\|x\| \leq 1$ and $\|Tx\| \geq \|T\| - \varepsilon$. Then since $\|x\| \leq 1$ and $\left\| \frac{x}{\|x\|} \right\| = 1$, we have

$$\begin{aligned}\|T\| &\geq \left\| T\left(\frac{x}{\|x\|}\right) \right\| \\ &= \frac{\|Tx\|}{\|x\|} \\ &\geq \|Tx\| \\ &> \|T\| - \varepsilon,\end{aligned}$$

and this contradicts (10). □

Exercise 1. Let $T \colon \mathcal{U} \to \mathcal{V}$ be a bounded operator and let $M \geq 0$. The following statements are equivalent.

1. $\|Tx\| \leq M\|x\|$ for all $x \in \mathcal{V}$ if and only if $\|T\| \leq M$.
2. $\|Tx\| > M\|x\|$ for some $x \in \mathcal{V}$ if and only if $\|T\| > M$.
3. $\|Tx\| \geq M\|x\|$ for some $x \in \mathcal{V}$ if and only if $\|T\| \geq M$.

3.2.2 Bounded Linear Operators and Normed Vector Spaces

We now want to justify our choice in terminology.

Definition 3.3. Let $\mathcal{V}$ and $\mathcal{W}$ be inner-product spaces. We define

$$\mathrm{Hom}(\mathcal{V},\mathcal{W}) := \{T\colon \mathcal{V}\to\mathcal{W} \mid T \text{ is a bounded linear operator}\}.$$

$\mathrm{Hom}(\mathcal{V},\mathcal{W})$ has the structure of a $\mathbb{C}$-vector space, where addition and scalar multiplication are defined by

$$(T+U)(x) = Tx + Ux \quad \text{and} \quad (\lambda T)(x) = T(\lambda x)$$

for all $T, U \in \mathrm{Hom}(\mathcal{V},\mathcal{W})$, $\lambda \in \mathbb{C}$, and $v \in \mathcal{V}$.

Proposition 3.5. *Let $\mathcal{V}$ and $\mathcal{W}$ be inner-product spaces. Then $(Hom(\mathcal{V},\mathcal{W}), \|\cdot\|)$ is a normed vector space, where $\|\cdot\|$ is the map which sends a bounded linear operator T to its norm $\|T\|$.*

Proof. An easy exercise in linear algebra shows that $\mathrm{Hom}(\mathcal{V},\mathcal{W})$ has the structure of a $\mathbb{C}$-vector space, where addition and scalar multiplication are defined by

$$(T+U)(x) = T(x) + U(x) \quad \text{and} \quad (\lambda T)(x) = T(\lambda x)$$

for all $T, U \in \mathrm{Hom}(\mathcal{V},\mathcal{W})$, $\lambda \in \mathbb{C}$, and $v \in \mathcal{V}$. The details of this are left as an exercise. We are more interested in the fact that $\mathrm{Hom}(\mathcal{V},\mathcal{W})$ is a *normed* vector space. We just need to check that $\|\cdot\|$ satisfies the conditions laid out in Definition (17.1).

We first check for subadditivity. Let $T, U \in \mathrm{Hom}(\mathcal{V},\mathcal{W})$. Then

$$\begin{aligned}\|(T+U)(x)\| &= \|Tx + Ux\| \\ &\leq \|Tx\| + \|Ux\| \\ &\leq \|T\|\|x\| + \|U\|\|x\| \\ &= (\|T\| + \|U\|)\|x\|\end{aligned}$$

for all $x \in \mathcal{V}$. In particular, this implies $\|T+U\| \leq \|T\| + \|U\|$. Thus we have subadditivity.

Next we check that $\|\cdot\|$ is absolutely homogeneous. Let $T \in \mathrm{Hom}(\mathcal{V},\mathcal{W})$ and let $\lambda \in \mathbb{C}$. Then

$$\begin{aligned}\|(\lambda T)(x)\| &= \|T(\lambda x)\| \\ &= \|\lambda Tx\| \\ &= |\lambda|\|Tx\|\end{aligned}$$

for all $x \in \mathcal{V}$. In particular, this implies $\|\lambda T\| = |\lambda|\|T\|$. Thus $\|\cdot\|$ is absolutely homogeneous.

Finally we check for positive-definitiness. Let $T \in \mathrm{Hom}(\mathcal{V},\mathcal{W})$. Clearly $\|T\|$ is greater than or equal to 0 since it is the supremum of terms which are greater than or equal to 0. Suppose $\|T\| = 0$. Then

$$\begin{aligned}\|Tx\| &\leq \|T\|\|x\| \\ &= 0 \cdot \|x\| \\ &= 0\end{aligned}$$

for all $x \in \mathcal{V}$. In particular, this implies $Tx = 0$ for all $x \in \mathcal{V}$ (by positive-definiteness of the norm for $\mathcal{W}$). Therefore $T = 0$ since they agree on all $x \in \mathcal{V}$. □

3.2.3 Bounded Linear Operators and Continuity

Theorem 3.1. *Let $T\colon \mathcal{U} \to \mathcal{V}$ be a linear operator. Then the following are equivalent.*

1. *T is uniformly continuous;*
2. *T is continuous at $0 \in \mathcal{U}$;*
3. *T is bounded.*

Proof. We may assume that $T \neq 0$ since the theorem is obvious in this case. That 1 implies 2 is clear. Let us show that 2 implies 3. Assume T is continous at 0. Choose $\delta > 0$ such that $\|x\| < \delta$ implies $\|Tx\| < 1$ (we can do this since T is continuous at 0). Then for any nonzero $y \in \mathcal{U}$ such that $\|y\| \leq 1$, we have

$$\begin{aligned}\|Ty\| &= \frac{2\delta}{2\delta}\|Ty\| \\ &= \frac{2}{\delta}\|T(\delta y/2)\| \\ &< \frac{2}{\delta}.\end{aligned}$$

It follows that $\|T\| < 2/\delta$ which implies T is bounded.

To finish the proof, we just need to show 3 implies 1. Assume T is bounded. Let $\varepsilon > 0$ and choose $\delta = \varepsilon/\|T\|$. Then $\|x - y\| < \delta$ implies

$$\begin{aligned} \|Tx - Ty\| &= \|T(x-y)\| \\ &\leq \|T\|\|x-y\| \\ &< \|T\|\frac{\varepsilon}{\|T\|} \\ &= \varepsilon. \end{aligned}$$

It follows that T is uniformly continuous. □

Proposition 3.6. *Let $T\colon \mathcal{U} \to \mathcal{V}$ be a bounded linear operator. Then Ker(T) is a closed linear subspace of $\mathcal{U}$.*

Proof. We first show that $\mathrm{Ker}(T)$ is a linear subspace of $\mathcal{U}$. Let $x, y \in \mathrm{Ker}(T)$ and let $\lambda \in \mathbb{C}$. Then

$$\begin{aligned} T(x + \lambda y) &= T(x) + \lambda T(y) \\ &= 0 + \lambda \cdot 0 \\ &= 0 \end{aligned}$$

implies $x + \lambda y \in \mathrm{Ker}(T)$. Thus, $\mathrm{Ker}(T)$ is a linear subspace of $\mathcal{U}$.

To see that $\mathrm{Ker}(T)$ is closed, let (x_n) be a sequence of elements in $\mathrm{Ker}(T)$ such that $x_n \to x$ where $x \in \mathcal{U}$. Since T is bounded, it is uniformly continuous (and in particular continuous at x). Therefore

$$\begin{aligned} 0 &= \lim_{n\to\infty} 0 \\ &= \lim_{n\to\infty} T(x_n) \\ &= T(\lim_{n\to\infty} x_n) \\ &= T(x), \end{aligned}$$

which implies $x \in \mathrm{Ker}(T)$. Thus $\mathrm{Ker}(T)$ is closed in $\mathcal{U}$. □

Remark. Note that $\mathrm{Im}(T)$ is not always closed.

3.3 Examples of Bounded Operators

3.3.1 Multiplication by Continuous Function is Bounded Operator

Proposition 3.7. *Let $k \in C[a,b]$. Then the operator $T\colon C[a,b] \to C[a,b]$ defined by*

$$Tf = kf$$

for all $f \in C[a,b]$ is bounded. It's norm will be explicitly computed in the proof below.

Proof. We first show it is linear. Let $f, g \in C[a,b]$ and let $\lambda, \mu \in \mathbb{C}$. Then we have

$$\begin{aligned} T(\lambda f + \mu g) &= k(\lambda f + \mu g) \\ &= \lambda k f + \mu k g \\ &= \lambda T(f) + \mu T(g). \end{aligned}$$

Thus, T is linear.

Next we show it is bounded. If $k = 0$, then $\|T\| = 0$, so assume $k \neq 0$. Since k is continuous on the compact interval $[a,b]$, there exists $c \in [a,b]$ such that $|k(x)| \leq |k(c)|$ for all $x \in [a,b]$. Choose such a $c \in [a,b]$ and let $f \in C[a,b]$ such that $\|f\| \leq 1$. Then

$$\begin{aligned} \|Tf\| &= \|kf\| \\ &= \sqrt{\int_a^b |k(x)|^2 |f(x)|^2 dx} \\ &\leq |k(c)| \sqrt{\int_a^b |f(x)|^2 dx} \\ &\leq |k(c)|. \end{aligned}$$

implies $\|T\| \leq |k(c)|$, and hence T is bounded.

To find the norm of T, let $\varepsilon > 0$ such that $\varepsilon < |k(c)|$. Without loss of generality, assume that $c < b$ (if $c = b$, then we swap the role of b with a in the argument which follows). Choose $c' \in (c, b)$ such that $|k(x)| \geq |k(c)| - \varepsilon$ for all $x \in (c, c')$ (such a c' must exist since k is continuous) and choose f to be a nonzero continuous function in $C[a, b]$ which vanishes outside the interval (c, c'). Then

$$|k(x)||f(x)| \geq (|k(c)| - \varepsilon)|f(x)|$$

for all $x \in (a, b)$. In particular, this implies

$$\begin{aligned}
\|Tf\| &= \|kf\| \\
&= \sqrt{\int_a^b |k(x)f(x)|^2 dx} \\
&\geq \sqrt{\int_a^b (|k(c)| - \varepsilon)|f(x)|^2 dx} \\
&= (|k(c)| - \varepsilon)\sqrt{\int_a^b |f(x)|^2 dx} \\
&= (|k(c)| - \varepsilon)\|f\|.
\end{aligned}$$

Therefore $\|T(f/\|f\|)\| \geq |k(c)| - \varepsilon$, and this implies

$$\|T\| \geq |k(c)| - \varepsilon \tag{11}$$

Since (30) holds for all $\varepsilon > 0$, we must have $\|T\| \geq |k(c)|$. Thus $\|T\| = |k(c)|$. □

3.3.2 Orthogonal Projections are Bounded

Example 3.1. Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{K}$ be a closed subspace of $\mathcal{H}$. Then the orthogonal projection map $P_{\mathcal{K}} \colon \mathcal{H} \to \mathcal{K}$ is a bounded operator. This is because $P_{\mathcal{K}}$ is a linear map and $\|P_{\mathcal{K}}x\| \leq \|x\|$ for all $x \in \mathcal{H}$. In particular, If $\mathcal{K} \neq 0$, then $\|P_{\mathcal{K}}\| \leq 1$. In fact, we have equality here: choose any nonzero $x \in \mathcal{K}$ such that $\|x\| = 1$. Then $\|P_{\mathcal{K}}x\| = 1$.

3.3.3 Unitary Operators are Bounded

Example 3.2. Let $U \colon \mathcal{H}_1 \to \mathcal{H}_2$ be a unitary operator. Then U is bounded. Indeed, this is because

$$\begin{aligned}
\|U\| &= \sup\{\|Ux\| \mid \|x\| \leq 1\} \\
&= \sup\{\|x\| \mid \|x\| \leq 1\} \\
&= 1.
\end{aligned}$$

3.3.4 Diagonal Operator is Bounded Operator

Example 3.3. Let (a_n) be a bounded sequence of complex numbers. Let $M = \sup(|a_n|)$. Define $T \colon \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ by

$$T((x_n)) = (a_n x_n)$$

for all $(x_n) \in \ell^2(\mathbb{N})$. Then T is a linear map and note that

$$\begin{aligned}
\|T((x_n))\| &= \|(a_n x_n)\| \\
&= \sqrt{\sum_{n=1}^{\infty} |a_n x_n|^2} \\
&\leq \sqrt{\sum_{n=1}^{\infty} M^2 |x_n|^2} \\
&= M\sqrt{\sum_{n=1}^{\infty} |x_n|^2} \\
&= M\|(x_n)\|.
\end{aligned}$$

Thus

$$\begin{aligned}
\|T\| &= \sup\{\|T((x_n))\| \mid \|(x_n)\| \leq 1\} \\
&\leq \sup\{M\|(x_n)\| \mid \|(x_n)\| \leq 1\} \\
&= M.
\end{aligned}$$

We claim that $\|T\| = M$. Assume for a contradiction that $||T|| < M$. Choose $k \in \mathbb{N}$ such that $|a_k| > \|T\|$. Let $e^k = (0, 0, \ldots, 0, 1, 0, \ldots,)$. Then $\|e^k\| = 1$ and $\|Te^k\| = |a_k| > \|T\|$. Contradiction.

3.3.5 Shift Operator is Bounded Operator

Let $S\colon \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be the unique linear map such that $S(e_n) = e_{n+1}$ for all $n \in \mathbb{N}$, where e_n is the standard orthonormal basis vector with 1 in its nth coordinate and 0 everywhere else. Then

$$\begin{aligned}
\|Sx\| &= \left\| S\left(\sum_{n=1}^{\infty} x_n e_n\right)\right\| \\
&= \left\| \sum_{n=1}^{\infty} x_n e_{n+1} \right\| \\
&= \sqrt{\sum_{n=1}^{\infty} |x_n|^2} \\
&= \|x\|
\end{aligned}$$

for all $x \in \ell^2(\mathbb{N})$ implies S is bounded and $\|S\| = 1$.

3.3.6 Operator on $(C[0,1], \langle \cdot, \cdot \rangle)$

Let $T\colon C[0,1] \to \mathbb{C}$ be defined by

$$T(f) = \int_0^1 f(x)dx$$

for all $f \in C[0,1]$. Then T is linear since the integral is linear. Furthermore

$$\begin{aligned}
\|Tf\| &= |Tf| \\
&= \left| \int_0^1 f(x)dx \right| \\
&\leq \sqrt{\int_0^1 |f(x)|^2} \sqrt{\int_0^1 |1|^2 dx} \\
&\leq \sqrt{\int_0^1 |f(x)|^2} \\
&= \|f\|,
\end{aligned}$$

where we applied Cauchy-Schwarz. Thus, T is bounded and $\|T\| \leq 1$. Moreover, $T(1) = 1$, and so $\|T(1)\| = 1$. In particular, this implies $\|T\| = 1$.

3.4 Riesz Representation Theorem

Theorem 3.2. *(Riesz representation theorem) Let $\mathcal{H}$ be a Hilbert space and let $\ell\colon \mathcal{H} \to \mathbb{C}$ be a bounded operator. Then there exists a unique vector $y \in \mathcal{H}$ such that $\ell = \ell_y$. In other words, we have*

$$\ell(x) = \langle x, y \rangle$$

for all $x \in \mathcal{H}$. Moreover, we have $\|\ell\| = \|y\|$.

Proof. If $\ell = 0$ then the theorem is clear, so assume $\ell \neq 0$. Denote $\mathcal{K} = \ker \ell$. Then $\mathcal{K}$ is a closed proper subspace of $\mathcal{H}$. Choose a nonzero vector z in $\mathcal{K}^\perp$. Note that $\ell(z) \neq 0$ since $\mathcal{K} \cap \mathcal{K}^\perp = \{0\}$. Now for any $x \in \mathcal{H}$, we can express it as

$$x = \left(x - \frac{\ell(x)}{\ell(z)} z\right) + \frac{\ell(x)}{\ell(z)} z \tag{12}$$

where $x - (\ell(x)/\ell(z))z \in \mathcal{K}$ and where $(\ell(x)/\ell(z))z \in \mathcal{K}^{\perp}$. Applying $\langle \cdot, z\rangle$ to both sides of (**??**) gives us

$$\begin{aligned}\langle x, z\rangle &= \left\langle x - \frac{\ell(x)}{\ell(z)}z, z\right\rangle + \left\langle \frac{\ell(x)}{\ell(z)}z, z\right\rangle \\ &= \left\langle \frac{\ell(x)}{\ell(z)}z, z\right\rangle \\ &= \frac{\ell(x)}{\ell(z)}\|z\|^2.\end{aligned}$$

In particular, we see that

$$\begin{aligned}\ell(x) &= \frac{\ell(z)}{\|z\|^2}\langle x, z\rangle \\ &= \left\langle x, \frac{\overline{\ell(z)}}{\|z\|^2}z\right\rangle.\end{aligned}$$

So setting $y = (\overline{\ell(z)}/\|z\|^2)z$, it follows that that $\ell = \ell_y$.

This proves the existence of $y \in \mathcal{H}$. To show uniqueness, suppose $y' \in \mathcal{H}$ such that $\ell_y = \ell = \ell_{y'}$. Then

$$\langle x, y\rangle = \ell(x) = \langle x, y'\rangle$$

for all $x \in \mathcal{H}$. However since $\langle \cdot, \cdot\rangle$ is positive-definite, this implies $y = y'$.

Finally, we need to check that $\|\ell\| = \|y\|$. Let $x \in \mathcal{H}$ such that $\|x\| \leq 1$. Then

$$\begin{aligned}|\ell(x)| &= |\langle x, y\rangle| \\ &\leq \|x\|\|y\| \\ &= \|y\|.\end{aligned}$$

Thus $\|\ell\| \leq \|y\|$. To see that equality is achieved, consider $x = y/\|y\|$. In this case, we have

$$\begin{aligned}\left|\ell\left(\frac{y}{\|y\|}\right)\right| &= \frac{1}{\|y\|}|\ell(y)| \\ &= \frac{1}{\|y\|}|\langle y, y\rangle| \\ &= \frac{1}{\|y\|}\|y\|^2 \\ &= \|y\|.\end{aligned}$$

It follows that $\|\ell\| = \|y\|$. □

3.4.1 Schur's test for boundedness

Let $T\colon \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be the operator given by

$$T(x)_i = \sum_{j=1}^{\infty} a_{ij}.$$

Then T is bounded if

$$\sup_i \left\{\sum_{j=1}^{\infty} |a_{ij}|\right\} \sup_j \left\{\sum_{i=1}^{\infty} |a_{ij}|\right\} < \infty$$

Set $\alpha := \sup_i \left\{\sum_{j=1}^{\infty} |a_{ij}|\right\}$ and $\beta := \sup_j \left\{\sum_{i=1}^{\infty} |a_{ij}|\right\}$. Then in this case, we have

$$\|T\| \leq \sqrt{\alpha\beta}.$$

Indeed, let $x \in \ell^2(\mathbb{N})$ such that $\|x\| \leq 1$. Then

$$\|Tx\| = \sup\left\{\langle Tx, y\rangle \mid \|y\| \leq 1\right\}.$$

Take $y \in \ell^2(\mathbb{N})$ such that $\|y\| \leq 1$. Then

$$
\begin{aligned}
|\langle Tx, y\rangle| &= |\langle \sum_{i=1}^{\infty} a_{ij} x_j), y_i\rangle| \\
&= \left| \sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} a_j x_j \right) \overline{y}_i \right| \\
&\leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{ij}||x_i||y_j| \\
&\leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{ij}| \frac{c|x_j|^2 + \frac{1}{c}|y_i|^2}{2} \\
&= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{c}{2}|a_{ij}||x_j|^2 + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{1}{2c}|a_{ij}||y_i|^2 \\
&= \sum_{j=1}^{\infty} \frac{c}{2}|x_j|^2 \sum_{i=1}^{\infty} |a_{ij}| + \sum_{i=1}^{\infty} \frac{1}{2c}|y_i|^2 \sum_{j=1}^{\infty} |a_{ij}| \\
&\leq \sum_{j=1}^{\infty} \frac{c}{2}|x_j|^2 \beta + \sum_{i=1}^{\infty} \frac{1}{2c}|y_i|^2 \alpha \\
&= \frac{c\beta}{2} \sum_{j=1}^{\infty} |x_j|^2 + \frac{\alpha}{2c} \sum_{i=1}^{\infty} |y_i|^2 \\
&= \frac{c\beta}{2}\|x\|^2 + \frac{\alpha}{2c}\|y\|^2 \\
&\leq \frac{c\beta}{2} + \frac{\alpha}{2c}
\end{aligned}
$$

where $c > 0$. Now choose $c = \sqrt{\frac{\alpha}{\beta}}$ (this makes the inequality minimal). Then we get

$$|\langle Tx, y\rangle| = \sqrt{\alpha\beta}$$

for all x such that $\|x\| \leq 1$. In particular, this implies $\|T\| \leq \sqrt{\alpha\beta}$.

3.5 Adjoint of an Operator

Theorem 3.3. *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces and let $T\colon \mathcal{H}_1 \to \mathcal{H}_2$ be a bounded operator. There exists a unique operator $T^*\colon \mathcal{H}_2 \to \mathcal{H}_1$ such that*

$$\langle Tx, y\rangle_{\mathcal{H}_2} = \langle x, T^* y\rangle_{\mathcal{H}_1}$$

for all $x \in \mathcal{H}_1$ and $y \in \mathcal{H}_2$.

Proof. We define $T^*\colon \mathcal{H}_2 \to \mathcal{H}_1$ as follows: Let $z \in \mathcal{H}_2$ and set $\ell = \ell_z \circ T$. Thus $\ell\colon \mathcal{H}_2 \to \mathbb{C}$ is defined by

$$\ell(x) = \langle Tx, z\rangle_{\mathcal{H}_2}$$

for all $x \in \mathcal{H}_1$. Since ℓ is a composition of bounded operators, it must also be a bounded operator with

$$\|\ell\| \leq \|T\|\|\ell_z\| = \|T\|\|z\|.$$

By the Riesz representation theorem, there exists a unique $y \in \mathcal{H}_1$ such that $\ell = \ell_y$. We set

$$T^* z = y.$$

Observe that T^* is well-defined because of the uniqueness part of the Riesz representation theorem!

Let us show that T^* is $\mathbb{C}$-linear. Let $\alpha, \alpha' \in \mathbb{C}$ and let $z, z' \in \mathcal{H}_2$. Then we have

$$
\begin{aligned}
\langle x, T^*(\alpha z + \alpha' z')\rangle &= \langle Tx, \alpha z + \alpha' z'\rangle \\
&= \overline{\alpha}\langle Tx, z\rangle + \overline{\alpha}'\langle Tx, z'\rangle \\
&= \overline{\alpha}\langle x, T^* z\rangle + \overline{\alpha}'\langle x, T^* z'\rangle \\
&= \langle x, \alpha T^* z + \alpha' T^* z'\rangle
\end{aligned}
$$

for all $x \in \mathcal{H}_1$. It follows that
$$T^*(\alpha z + \alpha' z') = \alpha T^* z + \alpha' T^* z',$$
which implies T^* is $\mathbb{C}$-linear.

Next let us show that T^* is bounded. For any $z \in \mathcal{H}_2$, we have

$$\begin{aligned} \|T^* z\|^2 &= \langle T^* z, T^* z \rangle \\ &= \langle TT^* z, z \rangle \\ &\leq \|T(T^* z)\| \|z\| \\ &\leq \|T\| \|T^* z\| \|z\|. \end{aligned}$$

It follows that $\|T^* z\| \leq \|T\| \|z\|$ which implies T^* is bounded with $\|T^*\| \leq \|T\|$. □

Definition 3.4. The operator T^* given in Theorem (3.3) is called the **adjoint** of T.

Proposition 3.8. *Let $T \colon \mathcal{H}_1 \to \mathcal{H}_2$ be a bounded operator between Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. Then $(T^*)^* = T$ and $\|T\| = \|T^*\|$.*

Proof. Let $x \in \mathcal{H}_1$. Then

$$\begin{aligned} \langle x, (T^*)^* y \rangle &= \langle T^* x, y \rangle \\ &= \overline{\langle y, T^* x \rangle}. \\ &= \overline{\langle Ty, x \rangle} \\ &= \langle x, Ty \rangle. \end{aligned}$$

Thus $(T^*)^* y = Ty$ for all $y \in \mathcal{H}_2$.

Now

$$\begin{aligned} \|T\| &= \|(T^*)^*\| \\ &\leq \|T^*\| \end{aligned}$$

implies $\|T\| \leq \|T^*\|$. Combining this with the previous theorem gives us $\|T^*\| \leq \|T\|$. □

Example 3.4. Let $S \colon \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be the shift operator, defined by

$$S((x_1, x_2, \dots)) = (0, x_1, x_2, \dots),$$

We compute S^*. We have

$$\langle S((x_n)), (y_n) \rangle = \langle ((x_n)), S^*(y_n) \rangle$$

if and ony if

$$\sum_{n=1}^{\infty} x_n y_{n+1} = \sum_{n=1}^{\infty} x_n \overline{(S^*((y_n))_n)}$$

Choose $(x_n) = e^k = (0, 0, \dots, 1, 0, 0, \dots)$. Then

$$\overline{y}_{k+1} = \overline{(S^*((y_n))_k}$$

which implies $S^*((y_n))_k = y_{k+1}$. So
$$S^*((y_1, y_2, \dots,)) = (y_2, y_3, \dots,).$$
We call S^* the **backwards shift operator**.

3.6 Self-Adjoint Operators

Definition 3.5. An operator $T \colon \mathcal{H} \to \mathcal{H}$ is said to be **self-adjoint** (or **symmetric, Hermitian**) if $T^* = T$, i.e.

$$\langle Tx, y \rangle = \langle x, Ty \rangle$$

for all $x, y \in \mathcal{H}$.

3.6.1 Examples and Non-examples of Self-Adjoint Operators

Example 3.5. Any orthogonal projection is self-adjoint.

Example 3.6. Let (a_n) be a sequence of complex numbers and let $T\colon \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be given by

$$T((x_n)) = (a_n x_n)$$

for all $(x_n) \in \ell^2(\mathbb{N})$. Then T is self-adjoint if and only if (a_n) is a sequence of real numbers. Indeed, if T is self-adjoint, then

$$\begin{aligned} a_n &= \langle a_n e^n, e^n \rangle \\ &= \langle T(e^n), e^n \rangle \\ &= \langle e^n, T(e^n) \rangle \\ &= \langle e^n, a_n e^n \rangle \\ &= \overline{a}_n \end{aligned}$$

for all $n \in \mathbb{N}$. Thus each a_n is real. Conversely, if each a_n is real, then

$$\begin{aligned} \langle T((x_n)), (y_n) \rangle &= \sum_{n=1}^{\infty} a_n x_n \overline{y}_n \\ &= \sum_{n=1}^{\infty} x_n \overline{a}_n \overline{y}_n \\ &= \langle (x_n), T((y_n)) \rangle. \end{aligned}$$

Example 3.7. The shift operator is not self-adjoint.

Example 3.8. Unitary operators are not self-adjoint.

Proposition 3.9. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be self-adjoint. Then*

$$\|T\| = \sup\{|\langle Tx, x \rangle| \mid \|x\| \leq 1\}.$$

Proof. First note that by Cauchy-Schwarz, we have

$$\begin{aligned} \sup\{|\langle Tx, x \rangle| \mid \|x\| \leq 1\} &\leq \sup\{\|Tx\| \mid \|x\| \leq 1\} \\ &= \|T\|. \end{aligned}$$

Conversely, let

$$M = \sup\{|\langle Tx, x \rangle| \mid \|x\| \leq 1\}.$$

Let $x, y \in \mathcal{H}$ with $\|x\|, \|y\| \leq 1$. Then

$$\begin{aligned} \langle T(x+y), x+y \rangle - \langle T(x-y), x-y \rangle &= \langle Tx, x \rangle + \langle Tx, y \rangle + \langle Ty, x \rangle + \langle Ty, y \rangle - \langle Tx, x \rangle + \langle Tx, y \rangle + \langle Ty, x \rangle - \langle Ty, y \rangle \\ &= \langle Tx, y \rangle + \langle Ty, x \rangle + \langle Tx, y \rangle + \langle Ty, x \rangle \\ &= 2\langle Tx, y \rangle + 2\overline{\langle x, Ty \rangle} \\ &= 2\langle Tx, y \rangle + 2\overline{\langle Tx, y \rangle} \\ &= 4\mathrm{Re}\langle Tx, y \rangle. \end{aligned}$$

Now observe that for any $z \in \mathcal{H}$ we have

$$\begin{aligned} |\langle Tz, z \rangle| &= |\langle \|z\| T(\frac{z}{\|z\|})|, \|z\| \frac{z}{\|z\|} \rangle \\ &= \|z\|^2 \langle T\left(\frac{z}{\|z\|}\right), \frac{z}{\|z\|} \rangle \\ &\leq \|z\|^2 M. \end{aligned}$$

Therefore

$$\begin{aligned} 4\mathrm{Re}\langle Tx, y \rangle &\leq |\langle T(x+y), x+y \rangle - \langle T(x-y), x-y \rangle| \\ &\leq M(\|x+y\|^2 + \|x-y\|^2) \\ &= 2M(\|x\|^2 + \|y\|^2) \\ &\leq 4M. \end{aligned}$$

Thus for any $x, y \in \mathcal{H}$ with $\|x\|, \|y\| \leq 1$, we proved

$$\mathrm{Re}\langle Tx, y\rangle \leq M. \tag{13}$$

Setting $y = Tx/\|Tx\|$, Then plugging in y into (13), we obtain

$$\begin{aligned}\|Tx\| &\leq= \frac{1}{\|Tx\|}\mathrm{Re}\langle\|Tx\|^2\rangle \\ &= \mathrm{Re}\langle Tx, \frac{Tx}{\|Tx\|}\rangle \\ &\leq M.\end{aligned}$$

Thus $\|T\| \leq M$. □

4 Compactness

In topology, one studies a class of spaces called **compact** spaces. Recall that a toplogical space X is said to be compact if it satisfies the following property: every open cover of X contains a finite subcover of X. In other words, X is compact if for all open covers $\{U_i\}_{i\in I}$ of X, where by an open cover we mean each U_i is an open subset of X and

$$X = \bigcup_{i\in I} U_i,$$

then there exists $U_{i_1}, \dots, U_{i_n} \in \{U_i\}_{i\in I}$ such that $\{U_{i_1}, \dots, U_{i_n}\}$ is an open cover of X, that is

$$X = U_{i_1} \cup \cdots \cup U_{i_n}.$$

There is an analagous notion of compactness called **sequential compactness**. Recall that a topological space X is said to be sequentially compact if every sequence of points in X has a convergent subsequence converging to a point in X. It turns out that for metric space, compactness and sequintial compactness are equivalent. Since we are always talking about sequences in linear analysis, it makes sense for us to study compact spaces as sequentially compact spaces. In particular, we make the following definitions:

Definition 4.1. Let $\mathcal{V}$ be an inner-product space and let $K \subseteq \mathcal{V}$.

1. We say K is **precompact** if every sequence in K has a convergent sequence.
2. We say K is **compact** if every sequence in K has a convergent sequence with a limit in K.

Precompactness can be thought of as "almost" compact. In fact, we have the following proposition.

Proposition 4.1. *Let $\mathcal{H}$ be an inner-product space and let $A \subseteq \mathcal{H}$. Then A is precompact if and only if $\overline{A}$ is compact.*

Proof. Suppose A is precompact. Let (a_n) be a sequence in $\overline{A}$. For each $n \in \mathbb{N}$ choose $b_n \in A$ such that

$$\|a_n - b_n\| < \frac{1}{n}.$$

Since A is precompact, there exists a convergent subsequence of (b_n), say $(b_{\pi(n)})$. We claim that the subsequence $(a_{\pi(n)})$ of (a_n) is Cauchy. Indeed, let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $\pi(n) \geq \pi(m) \geq N$ implies

$$\|b_{\pi(n)} - b_{\pi(m)}\| < \frac{\varepsilon}{3} \quad \text{and} \quad \frac{1}{\pi(m)} < \frac{\varepsilon}{3}.$$

Then $\pi(n) \geq \pi(m) \geq N$ implies

$$\begin{aligned}\|a_{\pi(n)} - a_{\pi(m)}\| &= \|a_{\pi(n)} - b_{\pi(n)} + b_{\pi(n)} - b_{\pi(m)} + b_{\pi(m)} - a_{\pi(m)}\| \\ &\leq \|a_{\pi(n)} - b_{\pi(n)}\| + \|b_{\pi(n)} - b_{\pi(m)}\| + \|b_{\pi(m)} - a_{\pi(m)}\| \\ &< \frac{1}{\pi(n)} + \frac{\varepsilon}{3} + \frac{1}{\pi(m)} \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon.\end{aligned}$$

This implies $(a_{\pi(n)})$ is Cauchy. Finally, since $(a_{\pi(n)})$ is Cauchy and since $\mathcal{H}$ is a Hilbert space, we must have $a_{\pi(n)} \to a$ for some $a \in \overline{A}$. This implies $\overline{A}$ is compact.

Conversely, suppose $\overline{A}$ is compact. Let (a_n) be a sequence in A. Then (a_n) is a sequence in $\overline{A}$. Since $\overline{A}$ is compact, the sequence (a_n) has a convergent subsequence. This implies A is precompact. □

4.0.1 Compactness = Sequential Compactness in Metric Spaces

As we mentioned before, sequential compactness and compactness are equivalent notions when it comes to metric spaces. It takes some work to show that sequential compactness implies compactness, so we will save that for later. On the other hand, we can prove that compactness implies sequential compactness relatively easily:

Proposition 4.2. *Let K be a compact subspace of $\mathcal{V}$. Then K is sequentially compact.*

Proof. Let (x_n) be a sequence in K. Assume for a contradiction that (x_n) does not have a convergent subsequence with a limit in K. We claim that for each $x \in K$, we can find an open neighborhood U_x of x such that only finitely many elements in the sequence (x_n) belongs to U_x. Indeed, assume for a contradiction that every open neighborhood of x contains infinitely many terms of (x_n). Then for each $n \in \mathbb{N}$, we can choose an $x_{\pi(n)}$ such that $x_{\pi(n)} \in B_{1/n}(x)$. Then the subsequence $(x_{\pi(n)})$ is a Cauchy sequence which converges to x, and this contradicts our first assumption.

Thus for each $x \in X$ we may choose an open neighborhood U_x of x such that only finitely many elements in the sequence (x_n) belongs to U_x. Since K is compact, the open cover $\{U_x\}_{x \in X}$ of K has a finite subcover, say $\{U_{x_1}, \dots, U_{x_k}\}$. But then each U_{x_i} contains only finitely many elements in the sequence (x_n), so the sequence only contains finitely many distinct elements. Such a sequence clearly has a convergent subsequence with a limit in K. This gives us our desired contradiction. $\square$

4.0.2 Failure of Heine-Borel Theorem in the Infinite-Dimensional Setting

As we've mentioned before, sequential compactness and compactness are equivalent notions when it comes to metric spaces. It turns out that there is another description of compactness when it comes to spaces like $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^n$, and $\mathbb{C}^n$. More generally, if $\mathcal{H}$ is a finite-dimensional Hilbert space and K is a subset of $\mathcal{H}$, then K is compact if and only if it is closed and bounded. This is essentially the content of the Heine-Borel Theorem. On the other hand, closed and bound subspaces of infinite-dimensional Hilbert spaces need not be compact subspaces. To see what goes wrong, consider the closed unit ball $\mathrm{B}_1[0]$ (centered at 0 and of radius 1) in $\ell^2(\mathbb{N})$. Clearly, $\mathrm{B}_1[0]$ is closed and bounded. However it is not compact. Indeed, the sequence $(\mathbf{e}_n)$ of standard coordinate vectors in $\mathrm{B}_1[0]$ cannot have a convergent subsequence. Indeed, the distance between any two standard coordinate vectors $\mathbf{e}_n$ and $\mathbf{e}_m$ is $\sqrt{2}$. Thus, any subsequence of $(\mathbf{e}_n)$ will fail to be Cauchy, and hence will not converge.

Even though closed and bound subspaces of infinite-dimensional Hilbert spaces need not be compact subspaces, the converse still holds. In other words, being closed and bounded is a necessary condition for a subspace to be compact, though it is not sufficient.

Proposition 4.3. *Let $\mathcal{H}$ be a Hilbert space and let $A \subseteq \mathcal{H}$ be a compact subspace. Then A is closed and bounded.*

Proof. In any Hausdorff space X (for example a metric space), a compact subspace $K \subseteq X$ is necessarily a closed subset of X. We leave the proof of this as an exercise for the reason. Let us prove that A is a bounded. Assume for a contradiction that A is not bounded. For each $a \in A$, let $U_a = \mathrm{B}_1(a)$ be the open ball of radius 1 centered at a. Then $\{U_a\}_{a \in A}$ forms an open cover of A. Since A is compact, there exists a finite subcover of $\{U_a\}$, say $\{U_{a_1}, \dots, U_{a_n}\}$. For each $1 \leq i, j \leq n$, set

$$L_{ij} = \mathrm{d}(U_{a_i}, U_{a_j}) = \sup\left\{ \|x_i - x_j\| \mid x_i \in U_{a_i} \text{ and } x_j \in U_{a_j} \right\}$$

Clearly L_{ij} is finite since $L_{ij} \leq \|a_i - a_j\| + 2$. Setting $L = \max_{1 \leq i,j \leq n}\{L_{ij}\}$, we see that for all $x, x' \in A$, we must have $\|x - x'\| \leq L$. Thus, A is bounded. $\square$

4.1 Weak Convergence

As we've seen, the sequence $(\mathbf{e}_n)$ in $\ell^2(\mathbb{N})$ has not convergent subsequences, so obviously the sequence $(\mathbf{e}_n)$ doesn't converge. On the other hand, there is a weaker notion of convergence in which the sequence $(\mathbf{e}_n)$ does converge. Let us discuss this weakened version of convergence now.

Definition 4.2. Let $\mathcal{V}$ be an inner-product space and let (x_n) be a sequence in $\mathcal{V}$. We say (x_n) **converges weakly** to an element $x \in \mathcal{V}$, which we denote by $x_n \xrightarrow{w} x$, if the sequence of complex numbers $(\langle x_n, y \rangle)$ converges (in the usual sense) to $\langle x, y \rangle \in \mathbb{C}$ for all $y \in \mathcal{V}$. In this case, we call x the **weak limit** of (x_n).

Remark. If $\langle x_n, y\rangle \to \langle x, y\rangle$, then $\langle y, x_n\rangle \to \langle y, x\rangle$. This is because

$$\begin{aligned}\lim_{n\to\infty}\langle y, x_n\rangle &= \lim_{n\to\infty}\overline{\langle x_n, y\rangle}\\ &= \overline{\lim_{n\to\infty}\langle x_n, y\rangle}\\ &= \overline{\langle x, y\rangle}\\ &= \langle y, x\rangle,\end{aligned}$$

where we are allowed to pull the conjugation outside of the limit operator since the conjugation function is continuous everywhere. A similar argument shows that if $\langle y, x_n\rangle \to \langle y, x\rangle$, then $\langle x_n, y\rangle \to \langle x, y\rangle$. Thus $\langle x_n, y\rangle \to \langle x, y\rangle$ if and only if $\langle y, x_n\rangle \to \langle y, x\rangle$.

Remark. Note that weak limits are unique. Indeed, this follows from positive-definiteness of the inner-product. If $\langle x_n, y\rangle \to \langle x, y\rangle$ and $\langle x_n, y\rangle \to \langle x', y\rangle$ for all $y \in \mathcal{V}$, then $\langle x, y\rangle = \langle x', y\rangle$ for all $y \in \mathcal{V}$, which implies $x = x'$ by positive-definiteness of the inner-product.

Example 4.1. Let $\mathcal{H}$ be an infinite-dimensional separable Hilbert space and let (e_n) be an orthonormal sequence in $\mathcal{H}$. We claim that $e_n \xrightarrow{w} 0$. Indeed, let $y \in \mathcal{H}$. Then since $\sum_{n=1}^{\infty}\langle y, e_n\rangle e_n$ is convergent, we must in particular have

$$\lim_{n\to\infty}\langle y, e_n\rangle = 0 = \langle y, 0\rangle.$$

Note that we need

Remark. Any orthonormal sequence is weakly convergent to 0.

We will prove that any bounded sequence in an infinite dimensional Hilbert space has a weakly convergent subsequence.

4.1.1 Weak Convergence Properties

Proposition 4.4. *Let $\mathcal{H}$ be an infinite dimensional Hilbert space. Then $x_n \to x$ implies $x_n \xrightarrow{w} x$. Conversely, if $x_n \xrightarrow{w} x$ and $\|x_n\| \to \|x\|$, then $x_n \to x$.*

Remark. In finite dimensional spaces we have $x_n \to x$ if and only if $x_n \xrightarrow{w} x$.

Proof. Suppose $x_n \to x$. Then for all $y \in \mathcal{H}$, we have

$$\begin{aligned}|\langle x_n, y\rangle - \langle x, y\rangle| &= |\langle x_n - x, y\rangle|\\ &\le \|x_n - x\|\|y\|\\ &\to 0\end{aligned}$$

as $n \to \infty$. It follows that $x_n \xrightarrow{w} x$.

Converesely, suppose $x_n \xrightarrow{w} x$ and $\|x_n\| \to \|x\|$. Since $x_n \xrightarrow{w} x$, we have in particular $\langle x_n, x\rangle \to \langle x, x\rangle$ and $\langle x, x_n\rangle \to \langle x, x\rangle$. Thus

$$\begin{aligned}\|x_n - x\|^2 &= \langle x_n, x_n\rangle - \langle x_n, x\rangle - \langle x, x_n\rangle + \langle x, x\rangle\\ &\to \|x\|^2 - \langle x, x\rangle - \langle x, x\rangle + \langle x, x\rangle\\ &= 0.\end{aligned}$$

It follows that $x_n \to x$. $\square$

Note that we really do need both $x_n \xrightarrow{w} x$ *and* $\|x_n\| \to \|x\|$ for the converse to be true. Indeed, let (x_n) be any orthonormal sequence in $\mathcal{H}$. Then for any $y \in \mathcal{H}$ we have the Bessel inequality

$$\sum_{n=1}^{\infty}|\langle y, x_n\rangle|^2 \le \|y\|^2.$$

In particular, the series on the left is convergent. Therefore $\langle x_n, y\rangle \to 0 = \langle 0, y\rangle$, and hence $x_n \xrightarrow{w} 0$. However, clearly (x_n) is not a convergent sequence since $\|x_n\| = 1$ for all $n \in \mathbb{N}$.

4.1.2 Weak convergence in Separable Hilbert Space

Lemma 4.1. *Let $\mathcal{H}$ be a separable Hilbert space, let (e_m) be an othonormal sequence in $\mathcal{H}$, let (x_n) be a bounded sequence in $\mathcal{H}$, and let $x \in \mathcal{H}$. Then $x_n \xrightarrow{w} x$ if and only if $\langle x_n, e_m \rangle \to \langle x, e_m \rangle$ for all $m \in \mathbb{N}$.*

Proof. Suppose $x_n \xrightarrow{w} x$. Then $\langle x_n, y \rangle \to \langle x, y \rangle$ for all $y \in \mathcal{H}$, so certainly $\langle x_n, e_m \rangle \to \langle x, e_m \rangle$ for all $m \in \mathbb{N}$. Conversely, suppose $\langle x_n, e_m \rangle \to \langle x, e_m \rangle$ for all $m \in \mathbb{N}$.

Step 1: We first show that $\langle x_n, y \rangle \to \langle x, y \rangle$ for all $y \in E$, where $E = \operatorname{span}\{e_m \mid m \in \mathbb{N}\}$, so let $y \in E$. Then

$$y = \lambda_1 e_{m_1} + \cdots + \lambda_r e_{m_r}$$

for some (unique) $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ and $m_1, \ldots, m_r \in \mathbb{N}$. Then

$$\begin{aligned}
\lim_{n\to\infty} \langle x_n, y \rangle &= \lim_{n\to\infty} \langle x_n, \lambda_1 e_{m_1} + \cdots + \lambda_r e_{m_r} \rangle \\
&= \overline{\lambda}_1 \lim_{n\to\infty} \langle x_n, e_{m_1} \rangle + \cdots + \overline{\lambda}_r \lim_{n\to\infty} \langle x_n, e_{m_r} \rangle \\
&= \overline{\lambda}_1 \langle x, e_{m_1} \rangle + \cdots + \overline{\lambda}_r \langle x, e_{m_r} \rangle \\
&= \langle x, \lambda_1 e_{m_1} + \cdots + \lambda_r e_{m_r} \rangle \\
&= \langle x, y \rangle.
\end{aligned}$$

Step 2: Now we will show that $\langle x_n, z \rangle \to \langle x, z \rangle$ for all $z \in \mathcal{H}$, so let $z \in \mathcal{H}$. Let $\varepsilon > 0$. Choose $M \geq 0$ such that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$ (we can do this by the assumption that (x_n) is bounded). Choose an element $y \in \operatorname{span}\{e_m \mid m \in \mathbb{N}\}$ such that $\|y - z\| < \frac{\varepsilon}{3\max(\|x\|, M)}$ (we can do this since $\mathcal{H}$ is separable). Finally, choose $N \in \mathbb{N}$ such that $n \geq N$ implies $|\langle x_n - x, y \rangle| < \varepsilon/3$ (we can do this by step 1). Then $n \geq N$ implies

$$\begin{aligned}
|\langle x_n, z \rangle - \langle x, z \rangle| &= |\langle x_n, z \rangle - \langle x_n, y \rangle + \langle x_n, y \rangle - \langle x, y \rangle + \langle x, y \rangle - \langle x, z \rangle| \\
&= |\langle x_n, z - y \rangle + \langle x_n - x, y \rangle + \langle x, y - z \rangle| \\
&\leq |\langle x_n, z - y \rangle| + |\langle x_n - x, y \rangle| + |\langle x, y - z \rangle| \\
&\leq \|x_n\| \|z - y\| + |\langle x_n - x, y \rangle| + \|x\| \|y - z\| \\
&< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \\
&= \varepsilon.
\end{aligned}$$

□

4.2 Baby Version of the Uniform Boundedness Principle and its Consequences

4.2.1 Baby Version of the Uniform Boundedness Principle

Lemma 4.2. *Let (x_n) be a sequence in $\mathcal{H}$. Assume there exists an $M > 0$ and a closed ball $\mathrm{B}_r[a] \subseteq \mathcal{H}$ such that*

$$|\langle x_n, y \rangle| \leq M$$

for all $n \in \mathbb{N}$ and for all $y \in \mathrm{B}_r[a]$. Then the sequence (x_n) is bounded.

Proof. The key is to translate everything from the closed ball $\mathrm{B}_r[a]$ to the closed ball $\mathrm{B}_1[0]$. Let $z \in \mathrm{B}_1[0]$. Then

$$\begin{aligned}
|\langle x_n, z \rangle| &= \frac{1}{r} |\langle x_n, rz \rangle| \\
&= \frac{1}{r} |\langle x_n, rz + a - a \rangle| \\
&= \frac{1}{r} |\langle x_n, rz + a \rangle - \langle x_n, a \rangle| \\
&\leq \frac{1}{r} |\langle x_n, rz + a \rangle| + \frac{1}{r} |\langle x_n, a \rangle \\
&\leq \frac{1}{r} M + \frac{1}{r} M \\
&= \frac{2M}{r}
\end{aligned}$$

for all $n \in \mathbb{N}$. In particular, fixing x_n and setting $z = x_n / \|x_n\|$, we have $\|x_n\| \leq 2M/r$. Since x_n was arbitrary, we have $\|x_n\| \leq 2M/r$ for all $n \in \mathbb{N}$. □

Theorem 4.3. *(Uniform Boundedness Principle) Let $\mathcal{H}$ be a Hilbert space and let (x_n) be a sequence in $\mathcal{H}$. Assume that for every $y \in \mathcal{H}$ there exists an $M_y > 0$ such that*

$$|\langle x_n, y\rangle| \leq M_y \tag{14}$$

for all $n \in \mathbb{N}$. Then the sequence (x_n) is bounded.

Proof. We claim that there exists $M > 0$ and a closed ball $\mathrm{B}_r[a] \subseteq \mathcal{H}$ such that

$$|\langle x_n, y\rangle| \leq M$$

for all $n \in \mathbb{N}$ and for all $y \in \mathrm{B}_r[a]$. This will imply (x_n) is bounded by Lemma (4.2).

For each $m \in \mathbb{N}$, we set

$$E_m = \{y \in \mathcal{H} \mid |\langle x_n, y\rangle| \leq m \text{ for all } n \in \mathbb{N}\}.$$

Observe that (E_m) is an ascending sequence of closed sets. Indeed, it is clearly ascending, to see that E_m is closed, view it as an infinite intersection of closed sets, namely

$$E_m = \bigcap_{n=1}^{\infty} \{y \in \mathcal{H} \mid |\langle x_n, y\rangle| \leq m\}.$$

Moreover, for any $y \in \mathcal{H}$, (14) implies $y \in \bigcup_{m=1}^{\infty} E_m$. Thus

$$\mathcal{H} = \bigcup_{m=1}^{\infty} E_m.$$

Now to prove the claim, it suffices to show that one of the E_m's contains an open ball of the form $\mathrm{B}_r(a)$ (the closure $\mathrm{B}_r[a]$ will then also belong to E_m since E_m is itself closed). Let $\mathrm{B}_{r_1}(a_1)$ be any open ball. If $\mathrm{B}_{r_1}(a_1) \subseteq E_1$, then we are done, so assume $\mathrm{B}_{r_1}(a_1) \cap E_1^c \neq \emptyset$. Choose $a_2 \in \mathcal{H}$ and $r_2 > 0$ such that $r_2 < r_1/2$ and $\mathrm{B}_{r_2}[a_2] \subseteq \mathrm{B}_{r_1}(a_1) \cap E_1^c$: we can find an open ball $\mathrm{B}_{r_2}(a_2)$ which is contained in $\mathrm{B}_{r_1}(a_1) \cap E_1^c$ since the intersection is a nonempty open set, and by replacing r_2 by a smaller positive number if necessary, we may assume that $\mathrm{B}_{r_2}[a_2]$ is contained in $\mathrm{B}_{r_2}[a_2] \cap E_1^c$). Continuing this process, we will either stop after finitely many iterations (and we'd be done!) or we can construct a sequence (a_n) in $\mathcal{H}$ and a sequence $(r_n) \in \mathbb{R}$ such that

$$r_{n+1} < r_n/2 < \cdots < r_1/2^n \quad \text{and} \quad \mathrm{B}_{r_{n+1}}[a_{n+1}] \subseteq \mathrm{B}_{r_n}(a_n) \cap E_n^c$$

for all $n \in \mathbb{N}$.

Assume for a contradiction that such a sequence has been constructed. We claim that (a_n) is a Cauchy sequence. Indeed, let $\varepsilon > 0$. Since $r_n \to 0$, there exists $N \in \mathbb{N}$ such that $n \geq N$ implies $2r_n < \varepsilon$. Then $n, m \geq N$ implies

$$\begin{aligned} \|a_n - a_m\| &\leq \|a_n - a_N\| + \|a_N - a_m\| \\ &< r_N + r_N \\ &= 2r_N \\ &< \varepsilon. \end{aligned}$$

since $a_n, a_m \in \mathrm{B}_{r_N}[a_N]$ for all $n, m \geq N$. Thus (a_n) is a Cauchy sequence and hence converges (since we are in a Hilbert space) say to $a \in \mathcal{H}$. Now observe that for any $m \in \mathbb{N}$, we have $a \in \mathrm{B}_{r_m}[a_m] \subseteq E_{m-1}^c$. In particular

$$\begin{aligned} a &\in \bigcap_{m=1}^{\infty} E_m^c \\ &= \left(\bigcup_{m=1}^{\infty} E_m\right)^c \\ &= \mathcal{H}^c \\ &= \emptyset, \end{aligned}$$

which is a contradiction. □

Corollary. *Let $\mathcal{H}$ be a Hilbert space, let (x_n) be a sequence of elements in $\mathcal{H}$, and let $x \in \mathcal{H}$. If $x_n \xrightarrow{w} x$, then the sequence (x_n) is bounded. Moreover, we have*

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|. \tag{15}$$

Proof. Let $y \in \mathcal{H}$. Since the sequence $(\langle x_n, y \rangle)$ of complex numbers is convergent, it must be bounded. Thus there exists $M_y > 0$ such that $|\langle x_n, y \rangle| \leq M_y$ for all $n \in \mathbb{N}$. It follows from Theorem (4.3) that the sequence (x_n) is bounded. For the last part, we have

$$\begin{aligned} \|x\|^2 &= |\langle x, x \rangle| \\ &= \lim_{n\to\infty} |\langle x_n, x \rangle| \\ &\leq \liminf_{n\to\infty} \|x_n\| \|x\| \\ &= \|x\| \liminf_{n\to\infty} \|x_n\|, \end{aligned}$$

which implies (15). □

4.2.2 Weak Convergence Plus Convergence Implies Convergence

Proposition 4.5. *Let $\mathcal{H}$ be a Hilbert space. If $x_n \xrightarrow{w} x$ and $y_n \to y$. Then $\langle x_n, y_n \rangle \to \langle x, y \rangle$.*

Proof. Choose $M \geq 0$ such that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$ (we can do this by the previous theorem). Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $\|y - y_n\| < \varepsilon/2M$ and $|\langle x_n, y \rangle - \langle x, y \rangle| < \varepsilon/2$. We have

$$\begin{aligned} |\langle x_n, y_n \rangle - \langle x, y \rangle| &\leq |\langle x_n, y_n \rangle - \langle x_n, y \rangle| + |\langle x_n, y \rangle - \langle x, y \rangle| \\ &= |\langle x_n, y - y_n \rangle| + |\langle x_n, y \rangle - \langle x, y \rangle| \\ &\leq \|x_n\| \|y - y_n\| + \frac{\varepsilon}{2} \\ &< M \frac{\varepsilon}{2M} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

□

4.2.3 Every Bounded Sequence in Hilbert Space has Weakly Convergent Subsequence

Lemma 4.4. *Let $\mathcal{H}$ be a separable Hilbert space. Then there exists a countable dense subset of $\mathcal{H}$.*

Proof. Choose an orthonormal basis (e_n) of $\mathcal{H}$. Let

$$\mathcal{Y} = \operatorname{span}_{\mathbb{Q}(i)}\{e_n \mid n \in \mathbb{N}\} = \{\lambda_1 e_{n_1} + \cdots + \lambda_k e_{n_k} \mid \lambda_1, \ldots, \lambda_k \in \mathbb{Q}(i)\}.$$

Then $\mathcal{Y}$ is countable since $\mathbb{Q}(i)$ is countable. Moreover, $\mathcal{Y}$ is dense in $\mathcal{H}$ since $\mathbb{Q}(i)$ is dense in $\mathbb{C}$ and since

$$\operatorname{span}_{\mathbb{C}}\{e_n \mid n \in \mathbb{N}\} = \{\lambda_1 e_{n_1} + \cdots + \lambda_k e_{n_k} \mid \lambda_1, \ldots, \lambda_k \in \mathbb{C}\}$$

is dense in $\mathcal{H}$. □

Remark. In general, a topological space X is said to be **separable** if it contains a countable dense subset. Every continuous function on a separable space whose image is a subset of a Hausdorff space is determined by its values on the countable dense subset.

Theorem 4.5. *Let $\mathcal{H}$ be an infinite dimensional separable Hilbert space. Then any bounded sequence in $\mathcal{H}$ has a weakly convergent subsequence.*

Remark. If $\mathcal{H}$ is finite-dimensional, then every bounded sequence has a convergent subsequence. This is a consequence of the Bolzano-Weierstrass Theorem. However this theorem is no longer true in infinite dimensions.

Proof. Let (x_n) be a bounded sequence in $\mathcal{H}$. Choose a countably dense susbset $\mathcal{Y}$ of $\mathcal{H}$ and order the elements in $\mathcal{Y}$ as a sequence, say $\mathcal{Y} = (y_m)$. Choose $M \geq 0$ such that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$. Consider the sequence $(\langle y_1, x_n \rangle)$ of complex numbers. Since

$$\begin{aligned} |\langle y_1, x_n \rangle| &\leq \|y_1\| \|x_n\| \\ &\leq \|y_1\| M \end{aligned}$$

for all $n \in \mathbb{N}$, we see that the sequence $(\langle y_1, x_n \rangle)$ is bounded. By the Bolzano-Weierstrass Theorem, it must have a convergent subsequence, say $(\langle y_1, x_{\pi_1(n)} \rangle)$[1] where $\langle y_1, x_{\pi_1(n)} \rangle \to \lambda_1$ for some (uniquely determined) $\lambda_1 \in \mathbb{C}$. Repeating the same argument for the bounded sequence $(\langle y_2, x_{\pi_1(n)} \rangle)$, we can find a convergent subsequence, say $(\langle y_2, x_{\pi_2(n)} \rangle)$ where $\langle y_2, x_{\pi_2(n)} \rangle \to \lambda_2$ for some (uniquely determined) $\lambda_2 \in \mathbb{C}$ *and* $\langle y_1, x_{\pi_2(n)} \rangle \to \lambda_1$ since

[1] Here we view π_1 as a strictly increasing function from $\mathbb{N}$ to $\mathbb{N}$ whose range consists of the indices in the subsequence.

$(\langle y_1, x_{\pi_2(n)} \rangle)$ is a subsequence of $(\langle y_1, x_{\pi_1(n)} \rangle)$ and hence must converge to the same limit. Repeating this process for each $m \in \mathbb{N}$, we can find a subsequence $(\langle y_m, x_{\pi_m(n)} \rangle)$ such that $\langle y_m, x_{\pi_m(n)} \rangle \to \lambda_m$ for some (uniquely determined) $\lambda_m \in \mathbb{C}$ *and* $\langle y_k, x_{\pi_m(n)} \rangle \to \lambda_k$ for all $1 \leq k < m$. Now we apply the diagonal trick. Define $\pi \colon \mathbb{N} \to \mathbb{N}$ by $\pi(n) = \pi_n(n)$ for all $n \in \mathbb{N}$ and consider the sequence $(x_{\pi(n)})$. Observe that the sequence $(x_{\pi(n)})$ is a subsequence of (x_n). Moreover, $(x_{\pi(n)})$ is essentially a subsequence of $(x_{\pi_m(n)})$ (minus the first m terms) for all $m \in \mathbb{N}$. Therefore $\langle y_m, x_{\pi(n)} \rangle \to \lambda_m$ for all $m \in \mathbb{N}$. In particular, for each $m \in \mathbb{N}$, the sequence $(\langle y_m, x_{\pi(n)} \rangle)$ is Cauchy.

Let $y \in \mathcal{H}$. We claim that the sequence $(\langle y, x_{\pi(n)} \rangle)$ of complex numbers is Cauchy. Indeed, let $\varepsilon > 0$. Choose $m_0 \in \mathbb{N}$ such that

$$\|y_{m_0} - y\| < \frac{\varepsilon}{3M}$$

(we can do this since $\mathcal{Y}$ is dense in $\mathcal{H}$). Choose $N \in \mathbb{N}$ such that $m, n \geq N$ implies

$$|\langle y_{m_0}, x_{\pi(n)} \rangle - \langle y_{m_0}, x_{\pi(m)} \rangle| < \frac{\varepsilon}{3}$$

(we can do this since the sequence $(\langle y_{m_0}, x_{\pi(n)} \rangle)$ is Cauchy). Then $m, n \geq N$ implies

$$\begin{aligned}
|\langle y, x_{\pi(n)} \rangle - \langle y, x_{\pi(m)} \rangle| &= |\langle y, x_{\pi(n)} \rangle - \langle y_{m_0}, x_{\pi(n)} \rangle + \langle y_{m_0}, x_{\pi(n)} \rangle - \langle y_{m_0}, x_{\pi(m)} \rangle + \langle y_{m_0}, x_{\pi(m)} \rangle - \langle y, x_{\pi(m)} \rangle| \\
&\leq |\langle y - y_{m_0}, x_{\pi(n)} \rangle| + |\langle y_{m_0}, x_{\pi(n)} \rangle - \langle y_{m_0}, x_{\pi(m)} \rangle| + |\langle y_{m_0} - y, x_{\pi(m)} \rangle| \\
&\leq \|y - y_{m_0}\| M + |\langle y_{m_0}, x_{\pi(n)} \rangle - \langle y_{m_0}, x_{\pi(m)} \rangle| + \|y_{m_0} - y\| M \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&= \varepsilon.
\end{aligned}$$

Thus the sequence $(\langle y, x_{\pi(n)} \rangle)$ is Cauchy.

Since the sequence $(\langle y, x_{\pi(n)} \rangle)$ is a Cauchy sequence of complex numbers for each $y \in \mathcal{H}$, we are justified in defining $\ell \colon \mathcal{H} \to \mathbb{C}$ by

$$\ell(y) = \lim_{n \to \infty} \langle y, x_{\pi(n)} \rangle$$

for all $y \in \mathcal{H}$. The map ℓ is linear since the limit operator is linear and since the inner-product is linear in the first argument. The map ℓ is also bounded since

$$\begin{aligned}
|\ell(y)| &= |\lim_{n \to \infty} \langle y, x_{\pi(n)} \rangle| \\
&= \lim_{n \to \infty} |\langle y, x_{\pi(n)} \rangle| \\
&\leq \lim_{n \to \infty} \|y\| M \\
&= \|y\| M
\end{aligned}$$

for all $y \in \mathcal{H}$. By the Riesz Representation Theorem, there is a unique $x \in \mathcal{H}$ such that

$$\begin{aligned}
\langle y, x \rangle &= \ell(y) \\
&= \lim_{n \to \infty} \langle y, x_{\pi(n)} \rangle
\end{aligned}$$

for all $y \in \mathcal{H}$. In other words, there exists a (necessarily unique) $x \in \mathcal{H}$ such that $x_{\pi(n)} \xrightarrow{w} x$. □

Corollary. *Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. Then every bounded sequence has a weakly convergent subsequence.*

Proof. Let (x_n) be a bounded sequence in $\mathcal{H}$. Define

$$\mathcal{H}_0 := \overline{\operatorname{span}}\{x_n \mid n \in \mathbb{N}\}.$$

Then (x_n) is a bounded sequence in the separable Hilbert space $\mathcal{H}_0$. Therefore by Theorem (4.5), there exists a bounded subsequence, say $(x_{\pi(n)})$, of the sequence (x_n) which weakly converges to some element, say x, in $\mathcal{H}_0$. Since $\mathcal{H}_0$ is a subspace of $\mathcal{H}$, it follows that $(x_{\pi(n)})$ is a subsequence of (x_n) which weakly converges to the element x, where $x \in \mathcal{H}_0 \subseteq \mathcal{H}$. □

4.3 Compact Operators

Definition 4.3. An operator $T\colon \mathcal{H} \to \mathcal{H}$ is said to be **compact** (also called **completely continuous**) if for any sequence (x_n) of elements in $\mathcal{H}$ such that $x_n \xrightarrow{w} x$, we have $Tx_n \to Tx$.

Remark. Thus compact operators improve convergence. Note that compact operators are continuous, and hence bounded. The set of all compact operators from $\mathcal{H}_1$ to $\mathcal{H}_2$ is denoted $\mathcal{C}(\mathcal{H}_1, \mathcal{H}_2)$. Note that $\mathcal{C}(\mathcal{H}, \mathcal{H})$ is a maximal ideal in $\mathcal{B}(\mathcal{H}, \mathcal{H})$.

Definition 4.4. A bounded operator T with finite-dimensional range is called a **finite rank operator**.

Proposition 4.6. *Every* ***finite rank operator*** *is compact.*

Proof. Let $T\colon \mathcal{H} \to \mathcal{H}$ be a finite rank operator. Let (x_n) be a sequence of elements in $\mathcal{H}$ such that $x_n \xrightarrow{w} x$. Then $Tx_n \xrightarrow{w} Tx$ since T is bounded. Let $e_1, \dots, e_k$ be an orthonormal basis for $\operatorname{im}(T)$. Then

$$\begin{aligned} Tx_n &= \langle Tx_n, e_1 \rangle e_1 + \cdots + \langle Tx_n, e_k \rangle e_k \\ &\to \langle Tx, e_1 \rangle e_1 + \cdots + \langle Tx, e_k \rangle e_k \\ &= Tx. \end{aligned}$$

□

4.4 Eigenvalues

Let T be a linear map from a complex vector space V to itself. Recall from linear algebra that $\lambda \in \mathbb{C}$ is said to be an **eigenvalue of** T if there exists a nonzero $v \in V$ such that $Tv = \lambda v$. In this case, we call v an **eigenvector of** T **corresponding to the eigenvalue** λ. We denote by Λ to be the set of all eigenvalues of T and we denote by E_λ to be the set of all eigenvectors of T corresponding to λ. Observe that $E_\lambda = \ker(\lambda - T)$, hence E_λ is in fact a subspace of V. We call this subspace the **eigenspace of** T **corresponding to** λ. When context is clear, we often refer to λ, v, and E_λ as "an eigenvalue", "an eigenvector", and "an eigenspace" respectively.

We'd like to have a good understanding of Λ. If V is finite dimensional, say $\dim V = n$, then the eigenvalues of T are completely characterized by the **characteristic polynomial of** T which is defined by the equation

$$\chi_T(X) = \det(X\mathrm{I}_n - [T]_\beta) \tag{16}$$

where $[T]_\beta$ is the matrix representation of T with respect to some basis β of V. Indeed, λ is an eigenvalue of T if and only if λ is a root of $\chi_T(X)$. Any polynomial in $\mathbb{C}[X]$ splits in $\mathbb{C}$ since $\mathbb{C}$is algebraically closed, and therefore $\chi_T(X)$ factors as

$$\chi_T(X) = \prod_{\lambda \in \Lambda} (X - \lambda)^{m_T(\lambda)} \tag{17}$$

where $m_T(\lambda)$ is called the **algebraic multiplicity of** λ. Therefore if we have a good understanding $\chi_T(X)$, then we will have a good understanding Λ as well.

If V is infinite dimensional, then the formulas (16) and (17) may no longer make sense due to convergence issues (for instance what does $\prod_{n=1}^\infty (X - n)$ mean?). Therefore we need to find alternative methods to better understand Λ. This is where linear analysis comes in. Throughout the rest of this subsection, let $\mathcal{H}$ be a separable Hilbert space.

4.4.1 Eigenspaces of Compact Operators

Proposition 4.7. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a compact operator. Then E_λ is finite dimensional for all $\lambda \in \Lambda \backslash \{0\}$.*

Proof. Let λ be a nonzero eigenvalue of T and let (x_n) be a bounded sequence in E_λ. Choose a subsequence $(x_{\pi(n)})$ of (x_n) such that $x_{\pi(n)} \xrightarrow{w} x$ for some $x \in \mathcal{H}$ (such a subsequence exists by the Uniform Boundedness Principle). Since T is compact, we have

$$\begin{aligned} \lambda x_{\pi(n)} &= Tx_{\pi(n)} \\ &\to Tx \\ &= \lambda x, \end{aligned}$$

and since $\lambda \neq 0$, this implies $x_{\pi(n)} \to x$. Thus $(x_{\pi(n)})$ is a convergent subsequence of (x_n). It follows from Proposition (12.7) (in HW7 Problem 6.a) that E_λ is finite dimensional. □

Remark. Note that the eigenspace E_0 corresponding to the eigenvalue 0 is precisely the kernel of T, and this may be *infinite* dimensional! For instance, if $T\colon \mathcal{H} \to \mathcal{H}$ is the zero map, then E_0 is all of $\mathcal{H}$.

4.4.2 Eigenspaces of Self-Adjoint Operators

Proposition 4.8. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a self-adjoint operator. Then Λ is bounded above by $\|T\|$.*

Proof. Let $\lambda \in \Lambda$ and choose an eigenvector x corresponding to the eigenvalue λ. By scaling if necessary, we may assume $\|x\| = 1$. Then

$$\begin{aligned} \|T\| &= \sup\{|\langle Ty, y\rangle| \mid \|y\| \leq 1\} \\ &\geq |\langle Tx, x\rangle| \\ &= |\langle \lambda x, x\rangle| \\ &= |\lambda|. \end{aligned}$$

Therefore $|\lambda| \leq \|T\|$ for all $\lambda \in \Lambda$. In other words, Λ is bounded above by $\|T\|$. □

Proposition 4.9. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a self-adjoint operator. Suppose λ and μ are two distinct eigenvalues of T. Then $E_\lambda \perp E_\mu$.*

Proof. Let $x \in E_\lambda$ and let $y \in E_\mu$. Then we have

$$\begin{aligned} (\lambda - \mu)\langle x, y\rangle &= \langle \lambda x, y\rangle - \langle x, \overline{\mu} y\rangle \\ &= \langle \lambda x, y\rangle - \langle x, \mu y\rangle \\ &= \langle Tx, y\rangle - \langle x, Ty\rangle \\ &= \langle Tx, y\rangle - \langle Tx, y\rangle \\ &= 0. \end{aligned}$$

Therefore $E_\lambda \perp E_\mu$. □

Proposition 4.10. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a self-adjoint operator. Then $\Lambda \subseteq \mathbb{R}$.*

Proof. Let λ be an eigenvalue of T. Choose an eigenvector of λ, say $x \in \mathcal{H}$. By scaling if necessary, we may assume $\|x\| = 1$. Then

$$\begin{aligned} \lambda &= \lambda\langle x, x\rangle \\ &= \langle \lambda x, x\rangle \\ &= \langle Tx, x\rangle \\ &= \langle x, Tx\rangle \\ &= \langle x, \lambda x\rangle \\ &= \overline{\lambda}\langle x, x\rangle \\ &= \overline{\lambda}, \end{aligned}$$

which implies λ is real. Therefore $\Lambda \subseteq \mathbb{R}$. □

4.4.3 Eigenspaces of Compact Self-Adjoint Operators

Proposition 4.11. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a compact self-adjoint operator. Assume that Λ is infinite. Then 0 is an accumulation point of Λ. Moreover, the set of all accumulation points of Λ is $\Lambda \cup \{0\}$. In other words, the closure of Λ (as a subset of $\mathbb{C}$) is $\overline{\Lambda} = \Lambda \cup \{0\}$.*

Remark. We give two remarks before we prove this proposition.

1. We are not saying that 0 is an eigenvalue of T. We are merely saying that 0 is an accumulation point of Λ and in fact the only accumulation point of Λ. Equivalently, we are saying that there exists a sequence of distinct eigenvalues (λ_n) such that $\lambda_n \to 0$ and that any convergent sequence of distinct eigenvalues must converge to 0.

2. The proposition tells us that Λ must be countable. This is because any uncountable subset of $\mathbb{C}$ must have infinitely many accumulation points.

Proof. Since Λ is bounded above by $\|T\|$, there exists a convergent sequence of distinct eigenvalues, say (λ_n). For each $n \in \mathbb{N}$, choose an eigenvector x_n of λ_n. By scaling if necessary, we may assume that $\|x_n\| = 1$ for all $n \in \mathbb{N}$. Since (λ_n) consists of distinct eigenvalues and since T is self-adjoint, we have $\langle x_m, x_n\rangle = 0$ whenever $m \neq n$. Thus (x_n) is an orthonormal sequence. In particular, this implies $x_n \xrightarrow{w} 0$. Since T is compact, we have $Tx_n \to 0$. Thus

$$\begin{aligned} \lambda_n x_n &= Tx_n \\ &\to 0, \end{aligned}$$

Taking norms gives us $|\lambda_n| \to 0$, which implies $\lambda_n \to 0$. □

4.4.4 Spectral Theorem

Since $\mathcal{H}$ is a separable Hilbert space, we know that there exists an orthonormal basis of $\mathcal{H}$. It turns out that we can get something even better:

Theorem 4.6. *(Spectral Theorem) Let $T\colon \mathcal{H} \to \mathcal{H}$ be a compact self-adjoint operator. Then there exists an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors of T.*

Remark. Note that the sum (19) may be finite. Also note that the sequence (λ_n) of eigenvalues corresponding to the sequence (e_n) of eigenvectors may have repeated terms.

Proof. If $T = 0$, then the theorem is clear, so assume $T \neq 0$.

Step 1: A nonzero eigenvalue of T exists: Denote by S to be the set $\{\langle Tx, x\rangle \mid \|x\| \leq 1\}$. Then S is nonempty and bounded above by $\|T\|$. Therefore the supremum of S exists. Denote by λ_0 to be this supremum. We claim that $\lambda_0 \in S$ and hence is in fact a maximum of S. Indeed, choose a sequence $(\langle Tx_n, x_n\rangle)$ in S such that $\langle Tx_n, x_n\rangle \to \lambda_0$. Since the sequence (x_n) is bounded, it has a weakly convergent subsequence, say $(x_{\pi(n)})$ where $x_{\pi(n)} \xrightarrow{w} x_0$ for some $x_0 \in \mathcal{H}$. Since T is compact, we will have $Tx_{\pi(n)} \to Tx_0$, and this implies $\langle Tx_{\pi(n)}, x_{\pi(n)}\rangle \to \langle Tx_0, x_0\rangle$ by Proposition (4.5). Since every subsequence of a convergence subsequence is convergent and moreover converges to the same limit, we have $\langle Tx_0, x_0\rangle = \lambda_0$. Finally, since

$$\begin{aligned}\|x_0\| &\leq \liminf_{n\to\infty} \|x_{\pi(n)}\| \\ &\leq 1,\end{aligned}$$

our claim is proved. In fact, we must have $\|x_0\| = 1$, since if $\|x_0\| < 1$, then

$$\begin{aligned}\left\langle T\left(\frac{x_0}{\|x_0\|}\right), \frac{x_0}{\|x_0\|}\right\rangle &= \frac{1}{\|x_0\|}\langle Tx_0, x_0\rangle \\ &> \lambda_0,\end{aligned}$$

which contradicts the fact that λ_0 is the supremum. Thus $\|x_0\| = 1$.

We next show that x_0 is an eigenvector of T with eigenvalue λ_0. Define a function $R_T\colon \mathcal{H}\backslash\{0\} \to \mathbb{R}$ (this is called the **Rayleigh quotient**) by

$$R_T(x) = \frac{\langle Tx, x\rangle}{\|x\|^2}$$

for all $x \in \mathcal{H}\backslash\{0\}$. Observe that $R_T(x) \leq R_T(x_0)$ for all $x \in \mathcal{H}\backslash\{0\}$. Define a function $f\colon \mathbb{R} \to \mathbb{R}$ by

$$f(t) = R_T(x_0 + tw)$$

where w is an arbitrary fixed vector in $\mathcal{H}$. Then f is maximized at $t = 0$. Moreover since

$$\begin{aligned}f(t) &= \frac{\langle T(x_0 + tw), x_0 + tw\rangle}{\|x_0 + tw\|^2} \\ &= \frac{\langle Tx_0, x_0\rangle + t(\langle Tx_0, w\rangle + \langle Tw, x_0\rangle) + t^2\langle Tw, w\rangle}{\|x_0\|^2 + 2t\mathrm{Re}\langle x_0, w\rangle + t^2\|w\|^2},\end{aligned}$$

we see that f is differentiabe at $t = 0$. It follows that

$$\begin{aligned}0 &= f'(0) \\ &= \frac{\langle Tx_0, w\rangle + \langle Tw, x_0\rangle}{\|x_0\|^2} - \langle Tx_0, x_0\rangle\frac{\langle x_0, w\rangle + \langle w, x_0\rangle}{\|x_0\|^4} \\ &= \langle Tx_0, w\rangle + \langle Tw, x_0\rangle - \lambda_0(\langle x_0, w\rangle + \langle w, x_0\rangle) \\ &= \langle Tx_0, w\rangle + \langle w, Tx_0\rangle - \lambda_0(\langle x_0, w\rangle + \langle w, x_0\rangle) \\ &= 2\mathrm{Re}\langle Tx_0, w\rangle - \lambda_0(2\mathrm{Re}\langle x_0, w\rangle) \\ &= 2\mathrm{Re}\langle Tx_0 - \lambda_0 x_0, w\rangle\end{aligned}$$

for all $w \in \mathcal{H}$. Plugging in $w = Tx_0 - \lambda_0 x_0$, we get $2\mathrm{Re}\|Tx_0 - \lambda_0 x_0\|^2 = 0$ which implies $\|Tx_0 - \lambda_0 x_0\| = 0$ which implies $Tx_0 = \lambda_0 x_0$. Thus λ_0 is an eigenvalue corresponding to the eigenvector x_0.

Now observe that $-\lambda_0$ is an eigenvalue of $-T$. Denote by S' to be the set $\{\langle -Tx, x\rangle \mid \|x\| \leq 1\}$. Then S' is nonempty and bounded above by $\|T\|$. Therefore the supremum of S' exists. Denote by λ_0' to be this supremum. Running through the same argument above, we find that λ_0' is an eigenvalue of $-T$ and hence $-\lambda_0'$ is an eigenvalue of T. If $\lambda_0 \geq \lambda_0'$, then it follows that

$$\begin{aligned}\lambda_0 &= \sup\{|\langle Tx, x\rangle| \mid \|x\| \leq 1\} \\ &= \|T\|.\end{aligned}$$

If $\lambda_0' \geq \lambda_0$, then we would get $\lambda_0' = \|T\|$. Thus either $\|T\|$ or $-\|T\|$ is an eigenvalue of T (and in fact the largest eigenvalue of T in absolute value).

Step 2: An orthonormal basis of T consisting of eigenvectors exists. Denote by $\Lambda_{\neq 0}$ to be the set of all nonzero eigenvalues of T. Then $\Lambda_{\neq 0} \neq \emptyset$ by step 1. For each $\lambda \in \Lambda_{\neq 0}$, let $\{e_{\lambda,i} \mid 1 \leq i \leq n_\lambda\}$ be an orthonormal basis of E_λ where $n_\lambda := \dim E_\lambda$ is finite since T is compact. The set

$$\bigcup_{\lambda \in \Lambda_{\neq 0}} \{e_{\lambda,i} \mid 1 \leq i \leq n_\lambda\} \tag{18}$$

is an orthonormal set consisting of eigenvectors. Indeed, we have $e_{\lambda,i} \perp e_{\mu,j}$ for all $\lambda, \mu \in \Lambda_{\neq 0}$ such that $\lambda \neq \mu$ and for all $1 \leq i \leq n_\lambda$ and $1 \leq j \leq n_\lambda$ since T is self-adjoint. We also have $e_{\lambda,i} \perp e_{\lambda,j}$ for all $\lambda \in \Lambda_{\neq 0}$ and for all $1 \leq i, j \leq n_\lambda$ such that $i \neq j$. By placing an order on (18) and relabeling indices based on that order, we obtain an orthonormal sequence (e_n) consisting of eigenvectors corresponding to nonzero eigenvalues of T. Let

$$\mathcal{K} = \overline{\operatorname{span}}(\{e_n \mid n \in \mathbb{N}\}).$$

Observe that $\mathcal{K}$ is T-invariant, meaning $Tx \in \mathcal{K}$ for all $x \in \mathcal{K}$. In fact, $\mathcal{K}^\perp$ is T-invariant too. Indeed, let $y \in \mathcal{K}^\perp$. Then

$$\begin{aligned}\langle x, Ty\rangle &= \langle Tx, y\rangle \\ &= 0\end{aligned}$$

for all $x \in \mathcal{K}$. Therefore $Ty \in \mathcal{K}^\perp$. Thus the restriction of T to $\mathcal{K}^\perp$ lands in $\mathcal{K}^\perp$, denote this restriction by $T|_{\mathcal{K}^\perp} \colon \mathcal{K}^\perp \to \mathcal{K}^\perp$. The operator $T|_{\mathcal{K}^\perp}$ is compact and self-adjoint since it inherits these properties from T. We claim that $T|_{\mathcal{K}^\perp}$ is the zero operator. Indeed, assume for a contradiction that $T|_{\mathcal{K}^\perp}$ is not the zero operator. Then from what we just proved, the operator $T|_{\mathcal{K}^\perp}$ must have a nonzero eigenvalue say λ. Choose an eigenvector $x \in \mathcal{K}^\perp$ of $T|_{\mathcal{K}^\perp}$ corresponding to the eigenvalue λ. Then $x \in \mathcal{K}^\perp$ is also an eigenvector of T. Therefore $x \in \mathcal{K}^\perp \cap \mathcal{K} = \{0\}$, which is a contradiction since eigenvectors are nonzero. Thus $T|_{\mathcal{K}^\perp}$ must be the zero operator. In particular this implies $\mathcal{K}^\perp \subseteq \ker(T)$. In fact, we already have the reverse inclusion. Indeed, let $x \in \ker(T)$ and let $n \in \mathbb{N}$. Then

$$\begin{aligned}\lambda_n \langle e_n, x\rangle &= \langle \lambda_n e_n, x\rangle \\ &= \langle Te_n, x\rangle \\ &= \langle e_n, Tx\rangle \\ &= \langle e_n, 0\rangle \\ &= 0,\end{aligned}$$

and since $\lambda_n \neq 0$, this implies $\langle e_n, x\rangle = 0$. Therefore $\langle e_n, x\rangle = 0$ for all $n \in \mathbb{N}$, and hence $x \in \mathcal{K}^\perp$. Finally, choose an orthonormal basis for $\ker T = \mathcal{K}^\perp$ and combine the orthonormal basis for $\mathcal{K}$ to get an orthonormal basis of $\mathcal{H} = \mathcal{K} \oplus \mathcal{K}^\perp$ consisting of eigenvectors of T. □

Corollary. *Let $T \colon \mathcal{H} \to \mathcal{H}$ be a compact self-adjoint operator and let (e_n) be an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors. Then*

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n\rangle e_n \tag{19}$$

for all $x \in \mathcal{H}$, where λ_n is the eigenvalue corresponding to the eigenvector e_n for all $n \in \mathbb{N}$.

Proof. Let $x \in \mathcal{H}$. Then

$$\begin{aligned}Tx &= T\left(\sum_{n=1}^{\infty} \langle x, e_n\rangle e_n\right) \\ &= \sum_{n=1}^{\infty} \langle x, e_n\rangle Te_n \\ &= \sum_{n=1}^{\infty} \langle x, e_n\rangle \lambda_n e_n \\ &= \sum_{n=1}^{\infty} \lambda_n \langle x, e_n\rangle e_n.\end{aligned}$$

$\square$

4.5 Functional Calculus

Let $T\colon \mathcal{H} \to \mathcal{H}$ be a positive compact self-adjoint operator. Thus $\langle Tx, x\rangle \geq 0$ for all $x \in \mathcal{H}$. This implies

$$\begin{aligned} 0 &\leq \langle Te_n, e_n\rangle \\ &= \langle \lambda_n e_n, e_n\rangle \\ &= \lambda_n \|e_n\|^2, \end{aligned}$$

which implies $\lambda_n \geq 0$ for all $n \in \mathbb{N}$. Define $S\colon \mathcal{H} \to \mathcal{H}$ by

$$Sx = \sum_{n=1}^{\infty} \sqrt{\lambda_n}\langle x, e_n\rangle e_n$$

for all $x \in \mathcal{H}$. Then

$$\begin{aligned} S^2x &= S(Sx) \\ &= \sum_{n=1}^{\infty} \sqrt{\lambda_n}\langle Sx, e_n\rangle e_n \\ &= \sum_{n=1}^{\infty} \sqrt{\lambda_n}\left\langle \sum_{m=1}^{\infty} \sqrt{\lambda_n}\langle x, e_m\rangle e_m, e_n \right\rangle e_n \\ &= \sum_{n=1}^{\infty} \sqrt{\lambda_n} \sum_{m=1}^{\infty} \sqrt{\lambda_n}\langle x, e_m\rangle \langle e_m, e_n\rangle e_n \\ &= \sum_{n=1}^{\infty} \sqrt{\lambda_n}\sqrt{\lambda_n}\langle x, e_n\rangle e_n \\ &= \sum_{n=1}^{\infty} \lambda_n\langle x, e_n\rangle e_n \\ &= Tx. \end{aligned}$$

More generally, let $f\colon D \subseteq \mathbb{R} \to \mathbb{R}$ be a bounded function. If $\{0\} \cup \Lambda \subseteq D$ where Λ is the set of all eigenvalues of T, then we can define $f(T)\colon \mathcal{H}' \to \mathcal{H}$ by

$$f(T)x = \sum_{n=1}^{\infty} f(\lambda_n)\langle x, e_n\rangle e_n.$$

Then $f(T) + g(T) = (f+g)(T)$ and $f(T) \circ g(T) = (f \circ g)(T)$ and moreover we will have

$$\|f(T)\| = \sup_{x \in D} |f(x)| \cdot \|T\|.$$

4.6 Singular-Value Decomposition

Recall that the polarization decomposition of a complex number

$$z = |z|e^{i\theta} \text{ where } |z|^2 = \bar{z}z.$$

We want to find an analogue of this for compact operators. Let $T\colon \mathcal{H} \to \mathcal{H}$ be a compact operator. Then T^*T is compact, self-adjoint, and positive. Then there exists a unique positive compact self-adjoint operator S such that $S^2 = T^*T$. [2] We denote this operator S by $|T|$. Thus

$$|T|x = \sum_{n=1}^{\infty} \sqrt{\lambda_n}\langle x, e_n\rangle e_n$$

for all $x \in \mathcal{H}$ where (λ_n) is a sequence consisting of all eigenvalues of T^*T and (e_n) a corresponding orthonormal basis of eigenvectors. Let $U\colon \mathcal{H} \to \mathcal{H}$ be defined by

$$Ux = \sum_{n=1}^{\infty} \frac{1}{\sqrt{\lambda_n}}\langle x, e_n\rangle Te_n,$$

[2] In fact, if T is not compact, then it turns out that we still obtain a unique positive compact self-adjoint operator S such that $S^2 = T^*T$, but this requires some measure theory which is outside the scope of this class.

where if $\lambda_n = 0$, then it is understood that $1/\sqrt{\lambda_n} = 0$. Then observe that

$$\begin{aligned}
U(|T|x) &= \sum_{n=1}^{\infty} \frac{1}{\sqrt{\lambda_n}} \langle |T|x, e_n \rangle T e_n \\
&= \sum_{n=1}^{\infty} \frac{1}{\sqrt{\lambda_n}} \left\langle \sum_{m=1}^{\infty} \sqrt{\lambda_m} \langle x, e_m \rangle e_m, e_n \right\rangle T e_n \\
&= \sum_{n=1}^{\infty} \frac{1}{\sqrt{\lambda_n}} \sum_{m=1}^{\infty} \sqrt{\lambda_m} \langle x, e_m \rangle \langle e_m, e_n \rangle T e_n \\
&= \sum_{n=1}^{\infty} \frac{1}{\sqrt{\lambda_n}} \sqrt{\lambda_n} \langle x, e_n \rangle T e_n \\
&= \sum_{n=1}^{\infty} \langle x, e_n \rangle T e_n \\
&= T \left(\sum_{n=1}^{\infty} \langle x, e_n \rangle e_n \right) \\
&= Tx
\end{aligned}$$

for all $x \in \mathcal{H}$. Therefore $U|T| = T$. We call this the **polar decomposition of** T. The numbers $\sqrt{\lambda_n}$ are called the **singular values of** T. Observe that

$$\begin{aligned}
Tx &= U|T|x \\
&= U \left(\sum_{n=1}^{\infty} \sqrt{\lambda_n} \langle x, e_n \rangle e_n \right) \\
&= \sum_{n=1}^{\infty} \sqrt{\lambda_n} \langle x, e_n \rangle U e_n
\end{aligned}$$

for all $x \in \mathcal{H}$.

5 Normed Linear Spaces

Definition 5.1. Let $\mathcal{X}$ be a vector space (over $\mathbb{C}$). A function $\|\cdot\| \colon \mathcal{X} \to \mathbb{R}$ satisfying

1. (Positive-Definite) $\|x\| \geq 0$ for all $x \in \mathcal{X}$ with equality if and only if $x = 0$.
2. (Absolutely Homogeneous) $\|\alpha x\| = |\alpha| \|x\|$ for all $x \in \mathcal{X}$ and $\alpha \in \mathbb{C}$.
3. (Subadditivity) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathcal{X}$

is called a **norm** on $\mathcal{X}$. We call the pair $(\mathcal{X}, \|\cdot\|)$ a **normed linear space**.

Example 5.1. Every inner-product space is a normed linear space.

Example 5.2. Let $1 \leq p < \infty$. Define $\|\cdot\|_p \colon \mathbb{C}^n \to \mathbb{R}$ by

$$\|(x_1, \dots, x_n)^\top\|_p = \sqrt[p]{|x_1|^p + \cdots + |x_n|^p}$$

for all $(x_1, \dots, x_n)^\top \in \mathbb{C}^n$. Then $\|\cdot\|_p$ is a norm on $\mathbb{C}^n$. More generally, we define

$$\ell^p(\mathbb{N}) := \left\{ (x_n) \in \mathbb{C}^\infty \mid \sum_{n=1}^{\infty} |x_n|^2 < \infty \right\}$$

and we define $\|\cdot\|_p \colon \ell^p(\mathbb{N}) \to \mathbb{R}$ by

$$\|(x_n)\|_p = \sqrt{\sum_{n=1}^{\infty} |x_n|^p}$$

Example 5.3. Define $\|\cdot\|_\infty \colon \mathbb{C}^n \to \mathbb{R}$ by

$$\|(x_1, \dots, x_n)^\top\|_\infty = \max\{|x_1|, \dots, |x_n|\}.$$

for all $(x_1,\dots,x_n)^\top \in \mathbb{C}^n$. Then $\|\cdot\|_\infty$ is a norm on $\mathbb{C}^n$. In fact we have $\|(x_1,\dots,x_n)^\top\|_p \to \|(x_1,\dots,x_n)^\top\|_\infty$ as $p\to\infty$ for all $(x_1,\dots,x_n)^\top \in \mathbb{C}^n$. More generally, we define

$$\ell^\infty(\mathbb{N}) := \{(x_n)\in\mathbb{C}^\infty \mid \sup\{|x_n| \mid n\in\mathbb{N}\}<\infty\}$$

and we define $\|\cdot\|_\infty \colon \ell^\infty(\mathbb{N})\to\mathbb{R}$ by

$$\|(x_n)\|_\infty = \sup\{|x_n| \mid n\in\mathbb{N}\}.$$

Example 5.4. Define $\|\cdot\|_{\sup}\colon C[a,b]\to\mathbb{R}$ by

$$\|f\|_{\sup} = \sup\{|f(x)| \mid x\in[a,b]\}.$$

Example 5.5. Let $\mathcal{H}$ be a Hilbert space and

$$\mathcal{L}(\mathcal{H}) = \{T\colon \mathcal{H}\to\mathcal{H} \mid T \text{ is bounded}\}.$$

We want to prove that $\|\cdot\|_p$ is a norm. We first prove it in the finite dimensional case. For the first property we have

$$\|x\|_p = \sqrt[p]{|x_1|^p+\cdots+|x_n|^p} \geq 0$$

with equality if and only if $x_1=\cdots=x_p=0$. For the second property, we have

$$\begin{aligned}\|\alpha x\|_p &= \sqrt[p]{|\alpha x_1|^p+\cdots+|\alpha x_n|^p}\\ &= |\alpha|\|x\|_p.\end{aligned}$$

For the triangle inequality, we need to show that

$$\sqrt[p]{|x_1+y_1|^p+\cdots+|x_n+y_n|^p} \leq \sqrt[p]{|x_1|^p+\cdots+|x_n|^p}+\sqrt[p]{|y_1|^p+\cdots+|y_n|^p}. \tag{20}$$

To prove this we will use the following analog of the Cauchy-Schwarz inequality called the **Hölder inequality**: for positive numbers $a_1,\dots,a_n$ and $b_1,\dots,b_n$, we have

$$\sum_{i=1}^n a_ib_i \leq \sqrt[p]{\sum_{i=1}^n a_i^p}\sqrt[q]{\sum_{i=1}^n b_i^q} \tag{21}$$

where $1\leq p,q<\infty$ such that $1=\frac{1}{p}+\frac{1}{q}$. To prove Hölder inequality we will use

$$ab \leq \frac{a^p}{p}+\frac{b^q}{q} \tag{22}$$

for all $a,b\geq 0$. If $\sum_{i=1}^n a_i^p=0$ or if $\sum_{i=1}^n b_i^p=0$ then the inequality (21) is trivial. So assume both are nonzero. In this case, the inequality (21) is equivalent to

$$\sum_{i=1}^n \frac{a_i}{A}\frac{b_i}{B}\leq 1. \tag{23}$$

where $A=\sqrt[p]{\sum_{i=1}^n a_i^p}$ and $B=\sqrt[q]{\sum_{i=1}^n b_i^q}$. Applying (22), we obtain

$$\begin{aligned}\sum_{i=1}^n \frac{a_i}{A}\frac{b_i}{B} &\leq \sum_{i=1}^n\left(\frac{\left(\frac{a_i}{A}\right)^p}{p}+\frac{\left(\frac{b_i}{B}\right)^q}{q}\right)\\ &= \frac{1}{A^pp}\sum_{i=1}^n a_i^p+\frac{1}{B^qq}\sum_{i=1}^n b_i^q\\ &= \frac{A^p}{A^pp}+\frac{B^q}{B^qq}\\ &= \frac{1}{p}+\frac{1}{q}\\ &= 1.\end{aligned}$$

This establishes (21). Next we prove (20). Let $x, y \in (\mathbb{C}^n, \|\cdot\|_p)$. Then

$$\begin{aligned}
\|x+y\|_p^p &= \left(\sum_{i=1}^n |x_i+y_i|^p\right) \\
&= \sum_{i=1}^n |x_i+y_i||x_i+y_i|^{p-1} \\
&\le \sum_{i=1}^n (|x_i|+|y_i|)|x_i+y_i|^{p-1} \\
&= \sum_{i=1}^n |x_i||x_i+y_i|^{p-1} + \sum_{i=1}^n |y_i||x_i+y_i|^{p-1} \\
&\le \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n |x_i+y_i|^{(p-1)q}\right)^{\frac{1}{q}} + \left(\sum_{i=1}^n |y_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n |x_i+y_i|^{(p-1)q}\right)^{\frac{1}{q}} \\
&= \|x\|_p \left(\sum_{i=1}^n |x_i+y_i|^p\right)^{\frac{1}{q}} + \|y\|_p \left(\sum_{i=1}^n |x_i+y_i|^p\right)^{\frac{1}{q}} \\
&= (\|x\|_p + \|y\|_p) \left(\sum_{i=1}^n |x_i+y_i|^p\right)^{\frac{1}{q}} \\
&= (\|x\|_p + \|y\|_p)\, \|x+y\|_p^{\frac{p}{q}} \\
&= (\|x\|_p + \|y\|_p)\, \|x+y\|_p^{p-1},
\end{aligned}$$

which implies $\|x+y\|_p \le \|x\|_p + \|y\|_p$.

Now consider $\ell^p(\mathbb{N})$. We want to show that the ℓ^p norm

$$\|x\|_p := \left(\sum_{i=1}^\infty |x_i|^p\right)^{\frac{1}{p}}$$

is a norm on $\ell^p(\mathbb{N})$. Properties (1) and (2) are easy to prove. We know for each $N \in \mathbb{N}$, we have

$$\begin{aligned}
\left(\sum_{i=1}^N |x_i+y_i|^p\right)^{\frac{1}{p}} &\le \left(\sum_{i=1}^N |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^N |y_i|^p\right)^{\frac{1}{p}} \\
&\le \left(\sum_{i=1}^\infty |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^\infty |y_i|^p\right)^{\frac{1}{p}} \\
&= \|x\|_p + \|y\|_p.
\end{aligned}$$

Since the left-hand side is monotone increasing and bounded in N, taking the limit $N \to \infty$ gives us $\|x+y\|_p \le \|x\|_p + \|y\|_p$.

5.1 Topology Induced by Norm

Just as in the case of inner-product spaces, we define convergence in normed linear spaces by $x_n \to x$ if $\|x_n - x\| \to 0$. We also define closure, open/closed balls, open/closed sets in exactly the same way. Also the notion of Cauchy sequence is defined in exactly the same way.

Proposition 5.1. *Let $\mathcal{X}$ be a normed linear space. Then*

1. *Every convergent sequence is Cauchy;*
2. *Every Cauchy (and hence every convergent sequence) is bounded;*
3. *If $x_n \to x$ and $y_n \to y$ then $\alpha x_n + \beta y_n \to \alpha x + \beta y$;*
4. *If $\alpha_n \to \alpha$ and $x_n \to x$, then $\alpha_n x_n \to \alpha x$.*

Proof. The proof is the same as in the inner-product case. □

Note that the converse of (1) in the previous proposition is wrong in general.

Definition 5.2. Let $\mathcal{X}$ be a normed linear space. We say $\mathcal{X}$ is a **Banach space** if every Cauchy sequence in $\mathcal{X}$ is convergent.

5.2 Bounded Operators on Normed Linear Spaces

We define the notion of a bounded operator on normed linear spaces in the same way as before.

Definition 5.3. Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces and let $T\colon \mathcal{X} \to \mathcal{Y}$ be a map. We say T is a **bounded linear operator** if T is linear and if

$$\sup\{\|Tx\|_{\mathcal{Y}} \mid \|x\|_{\mathcal{X}} \leq 1\} < \infty.$$

In this case, we define the **operator norm** of T to be

$$\|T\| := \sup\{\|Tx\|_{\mathcal{Y}} \mid \|x\|_{\mathcal{X}} \leq 1\}.$$

A bounded linear operator $\ell\colon \mathcal{X} \to \mathbb{C}$ is called a **bounded linear functional**.

Proposition 5.2. *Let $\mathcal{X}$ and $\mathcal{Y}$ be normed linear spaces and let $T\colon \mathcal{X} \to \mathcal{Y}$ be a linear map. Then T is bounded if and only if T is continuous at 0 if and only if T is uniformly continuous.*

Proposition 5.3. *If T is bounded then $\|Tx\| \leq \|T\|\|x\|$ for all $x \in \mathcal{X}$.*

5.3 Dual Spaces

Definition 5.4. Given a normed linear space X, the **dual space** of X, denoted by X^*, is the vector space which consists of all bounded linear functionals on X equipped with the operator norm. In other words,

$$X^* = \{\ell\colon X \to \mathbb{C} \mid \ell \text{ is bounded and linear.}\}$$

5.3.1 Hahn-Banach Theorem and its Consequences

Theorem 5.1. *(Hahn-Banach Theorem) Let X be a normed linear space and let Y be a subspace of X. Then every bounded linear functional $\psi\colon Y \to \mathbb{C}$ can be extended to X with the same norm. More precisely, there exists a bounded linear functional $\widetilde{\psi}\colon X \to \mathbb{C}$ such that $\widetilde{\psi}|_Y = \psi$ and moreover $\|\widetilde{\psi}\| = \|\psi\|$.*

Remark. Note that $\|\widetilde{\psi}\| \leq \|\psi\|$ is the nontrivial direction.

Proposition 5.4. *For any nonzero vector $x_0 \in X$ there exists $\ell \in X^*$ with $\|\ell\| = 1$ such that $\ell(x_0) = \|x_0\|$.*

Proof. Let $Y = \text{span}(\{x_0\})$. Define $\psi\colon Y \to \mathbb{C}$ by

$$\psi(\lambda x_0) = \lambda\|x_0\|$$

for all $\lambda x_0 \in \text{span}(\{x_0\})$. It is easy to check that ψ is linear and bounded with $\psi(x_0) = \|x_0\|$ and $\|\psi\| = 1$. By the Hahn-Banach Theorem, we can extend ψ to a bounded linear functional $\ell\colon X \to \mathbb{C}$ such that $\|\ell\| = \|\psi\| = 1$ and such that $\ell(x_0) = \psi(x_0) = \|x_0\|$. This completes the proof. □

Proposition 5.5. *For any $x \in X$, we have*

$$\|x\| = \max\{|\ell(x)| \mid \ell \in X^* \textit{ and } \|\ell\| \leq 1\}.$$

Proof. First note that for any $\ell \in X^*$ such that $\|\ell\| \leq 1$ we always have

$$\begin{aligned} |\ell(x)| &\leq \|\ell\|\|x\| \\ &\leq \|x\|. \end{aligned}$$

For the reverse inequality, choose an $\ell \in X^*$ such that $\|\ell\| = 1$ and such that $\ell(x) = \|x\|$ (such a choice exists by Proposition (5.4)). This gives us the reverse direction. □

Proposition 5.6. *X^* is always a Banach space.*

Proof. It suffices to show X^* is complete. Let (ℓ_n) be a Cauchy sequence in X^*. Let $x \in X$ and let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $m, n \geq N$ implies

$$\|\ell_n - \ell_m\| < \frac{\varepsilon}{\|x\|}.$$

Then $n, m \geq N$ implies

$$\begin{aligned} |\ell_n(x) - \ell_m(x)| &= |(\ell_n - \ell_m)(x)| \\ &\leq \|\ell_n - \ell_m\|\|x\| \\ &< \frac{\varepsilon}{\|x\|}\|x\| \\ &= \varepsilon. \end{aligned}$$

In particular, $(\ell_n(x))$ is a Cauchy sequence of complex numbers for all $x \in X$. Since $\mathbb{C}$ is complete, we may define $\ell\colon X \to \mathbb{C}$ by

$$\ell(x) = \lim_{n\to\infty} \ell_n(x)$$

for all $x \in X$.

We claim that $\ell \in X^*$ and $\ell_n \to \ell$ as $n \to \infty$. Linearity of ℓ follows from linearity of ℓ_n and linearity of the limit operator. To see that ℓ is bounded, we consider the inequality

$$|\ell_n(x) - \ell_m(x)| < \varepsilon \|x\|.$$

Then for $\|x\| \leq 1$ and setting $m \to \infty$, we find that $n \geq N$ implies

$$|\ell_n(x) - \ell(x)| \leq \varepsilon.$$

Therefore $\ell_n - \ell$ is bounded. In particular, $\ell = (\ell_n - \ell) + \ell_n$ is bounded since X^* is a vector space. This also implies $\|\ell_n\| \to \|\ell\|$. □

Definition 5.5. Two normed linear spaces $(X_1, \|\cdot\|_1)$ and $(X_2, \|\cdot\|_2)$ are said to be **isometricaly isomorphic** if there exists a bijective linear map $T\colon X_1 \to X_2$ such that $\|Tx\|_2 = \|x\|_1$ for all $x \in X_1$. In this case, we write $X \cong Y$. If T is not necessarily surjective, then we say $(X_1, \|\cdot\|_1)$ is **isometrically embedded** in $(X_2, \|\cdot\|_2)$ and we call such T an **(isometrical) embedding**.

Proposition 5.7. *Suppose $1 \leq p < \infty$ and q is the conjugate of p (i.e. $1/p + 1/q = 1$). The dual space of $\ell^p(\mathbb{N})$ is isometrically isomorphic to $\ell^q(\mathbb{N})$.*

Definition 5.6. Let $\mathcal{X}$ be a normed linear space and let A be a subset of $\mathcal{X}$. The **annihilator of** A **in** $\mathcal{X}$ is defined to be the set of all bounded linear functionals $\ell\colon \mathcal{X} \to \mathbb{C}$ that annihilate A, i.e. $\ell(a) = 0$ for all $a \in A$. We denote the annihilator of A by $A^\perp$.

Proposition 5.8. *For any $A \subseteq \mathcal{X}$ we have $A^\perp$ is a closed subspace of $\mathcal{X}^*$.*

Just like in Hilbert spaces, we define

$$d(x, A) = \inf\{\|x - a\| \mid a \in A\}$$

to be the distance from x to A.

Theorem 5.2. *Let $\mathcal{X}$ be a normed linear space and let $\mathcal{Y}$ be a subspace of $\mathcal{X}$. Then*

$$d(x, \mathcal{Y}) = \max\{|\ell(x)| \mid \ell \in \mathcal{Y}^\perp \text{ and } \|\ell\| \leq 1\}.$$

Proof. Let $x \in \mathcal{X}$ and let $\ell \in \mathcal{Y}^\perp$ with $\|\ell\| \leq 1$. Then

$$\begin{aligned} |\ell(x)| &= |\ell(x - y)| \\ &\leq \|\ell\| \|x - y\| \\ &\leq \|x - y\| \end{aligned}$$

for all $y \in \mathcal{Y}$. Therefore

$$\begin{aligned} d(x, \mathcal{Y}) &\geq \inf_{y \in \mathcal{Y}} \|x - y\| \\ &\geq \inf_{y \in \mathcal{Y}} |\ell(x)| \\ &= |\ell(x)|. \end{aligned}$$

This implies

$$\sup\{|\ell(x)| \mid \ell \in \mathcal{Y}^\perp \text{ and } \|\ell\| \leq 1\} \leq d(x, \mathcal{Y}).$$

Now we want to show that the supremum is actually a maximum. If $x \in \mathcal{Y}$ then it is obvious, so assume $x \notin \mathcal{Y}$. Define $\mathcal{Z} = \operatorname{span}\{x, \mathcal{Y}\}$ and define $\psi\colon \mathcal{Z} \to \mathbb{C}$ by

$$\psi(\lambda x + y) = \lambda d(x, \mathcal{Y})$$

for all $\lambda x + y \in \mathcal{Z}$. Note that any element in $\mathcal{Z}$ can be uniquely expressed as $\lambda x + y$ for some $\lambda \in \mathbb{C}$ and $y \in \mathcal{Y}$ and hence ψ is well-defined. It is easy to show that ψ is a linear functional. Also $\psi(y) = 0$ for all $y \in \mathcal{Y}$ and

$\psi(x) = d(x, \mathcal{Y})$. Let $z \in \mathcal{Z}$. If $z \in \mathcal{Y}$, then $|\psi(z)| \leq \|z\|$, otherwise let $z = \lambda x + y$ with $\lambda \neq 0$. Then

$$\begin{aligned}
|\psi(z)| &= |\lambda| d(x, \mathcal{Y}) \\
&= \frac{|\lambda| \|z\|}{\|z\|} d(x, \mathcal{Y}) \\
&= \frac{|\lambda| \|z\|}{\|\lambda x + y\|} d(x, \mathcal{Y}) \\
&= \frac{|\lambda| \|z\|}{\|\lambda(x + \lambda^{-1} y)\|} d(x, \mathcal{Y}) \\
&= \frac{\|z\|}{\|(x + \lambda^{-1} y)\|} d(x, \mathcal{Y}) \\
&\leq \|z\|
\end{aligned}$$

So for any $z \in \mathcal{Z}$, we have $|\psi(z)| \leq \|z\|$. By the Hahn-Banach Theorem there exists $\ell \in \mathcal{X}^*$ such that $\ell(z) = \psi(z)$ for all $z \in \mathcal{Z}$ and $\|\ell\| = \|\psi\| \leq 1$. Then $\ell(y) = \psi(y) = 0$ for all $y \in \mathcal{Y}$. In otherwords, $\ell \in \mathcal{Y}^\perp$. Also $\ell(x) = \psi(x) = d(x, \mathcal{Y})$. Therefore the supremum is a maximum. □

5.4 Reflexivity

Any $x \in \mathcal{X}$ defines a natural linear functional $x^{**} \colon \mathcal{X}^* \to \mathbb{C}$ by

$$x^{**}(\ell) = \ell(x)$$

for all $\ell \in \mathcal{X}^*$. Indeed, for linearity let $\alpha_1, \alpha_2 \in \mathbb{C}$ and $\ell_1, \ell_2 \in \mathcal{X}^*$. Then

$$\begin{aligned}
x^{**}(\alpha_1 \ell_1 + \alpha_2 \ell_2) &= (\alpha_1 \ell_1 + \alpha_2 \ell_2)(x) \\
&= \alpha_1 \ell_1(x) + \alpha_2 \ell_2(x) \\
&= \alpha_1 x^{**}(\ell) + \alpha_2 x^{**}(\ell).
\end{aligned}$$

Moreover x^{**} is bounded above by $\|x\|$ since

$$\begin{aligned}
|x^{**}(\ell)| &= |\ell(x)| \\
&\leq \|\ell\| \|x\|
\end{aligned}$$

for all $\ell \in \mathcal{X}^*$. In fact $\|x^{**}\| = \|x\|$.

Proposition 5.9. *Define* $\Phi \colon \mathcal{X} \to \mathcal{X}^{**}$ *by*

$$\Phi(x) = x^{**}$$

for all $x \in \mathcal{X}$. *Then* Φ *is an isometrically embedding.*

Proof. For linearity, let $\alpha, \beta \in \mathbb{C}$ and $x, y \in \mathcal{X}$ and let $\ell \in \mathcal{X}^*$. Then

$$\begin{aligned}
(\alpha x + \beta y)^{**}(\ell) &= \ell(\alpha x + \beta y) \\
&= \alpha \ell(x) + \beta \ell(y) \\
&= \alpha x^{**}(\ell) + \beta y^{**}(\ell) \\
&= (\alpha x^{**} + \beta y^{**})(\ell).
\end{aligned}$$

Therefore $(\alpha x + \beta y)^{**} = \alpha x^{**} + \beta y^{**}$, and so Φ is linear. Also we proved

$$\begin{aligned}
\|\Phi(x)\| &= \|x^{**}\| \\
&= \|x\|.
\end{aligned}$$

In other words Φ is an isometry. □

Definition 5.7. A normed linear space $\mathcal{X}$ is said to be **reflexive** if the natural embedding Φ is surjective, i.e. $\mathcal{X} \cong \mathcal{X}^{**}$.

5.4.1 Examples of Reflexive Banach Spaces

Example 5.6. Let $1 < p < \infty$. Then $\ell^p(\mathbb{N})$ is reflexive. On the other hand, $\ell^1(\mathbb{N})$, $\ell^\infty(\mathbb{N})$, and $(C[a,b], \|\cdot\|_{\sup})$ are not reflexive.

Lemma 5.3. *Let $\mathcal{Y}$ be a closed subspace of a normed linear space $\mathcal{X}$. Then there exists $z \in \mathcal{X}$ with $\|z\| = 1$ such that $d(z, \mathcal{Y}) \geq 1/2$.*

Proof. Let $x \notin \mathcal{Y}$. Since $\mathcal{Y}$ is closed we must have $d(x, \mathcal{Y}) > 0$. Choose $y_0 \in \mathcal{Y}$ such that

$$\|x - y_0\| < d(x, \mathcal{Y}) + d(x, \mathcal{Y}).$$

Let $z' = x - y_0$. Then

$$\begin{aligned}\|z' - y\| &= \|x - y_0 - y\| \\ &= \|x - (y + y_0)\| \\ &\geq d(x, \mathcal{Y})\end{aligned}$$

for all $y \in \mathcal{Y}$. Now set $z = z'/\|z'\|$. Then $\|z\| = 1$ and

$$\begin{aligned}\|z - y\| &= \left\| \frac{z'}{\|z'\|} - y \right\| \\ &\frac{\|z' - \|z'\| \cdot y\|}{\|z'\|} \\ &> \frac{d(x, \mathcal{Y})}{\|x - y_0\|} \\ &\geq \frac{d(x, \mathcal{Y})}{2d(x, \mathcal{Y})} \\ &= \frac{1}{2}.\end{aligned}$$

□

Theorem 5.4. *The closed (unit) ball in any infinite-dimensional normed linear space is not compact.*

Proof. Let y_1 with $\|y_1\| = 1$ be arbitrary. Set $\mathcal{Y}_1 = \operatorname{span}\{y_1\}$. Then $\mathcal{Y}_1$ is a closed proper subspace of $\mathcal{X}$. Therefore by the lemma, there exists y_2 with $\|y_2\| = 1$ such that $d(y_2, \mathcal{Y}_1) \geq 1/2$. Choose such y_2. Then clearly

$$\|y_2 - y_1\| \geq \frac{1}{2}.$$

Next set $\mathcal{Y}_2 = \operatorname{span}\{y_1, y_2\}$. Again, $\mathcal{Y}_2$ is a proper closed subspace of $\mathcal{X}$. Choose y_3 with $\|y_3\| = 1$ such that $d(y_3, \mathcal{Y}_2) \geq 1/2$. In particular, we have

$$\|y_3 - y_2\| \geq \frac{1}{2} \quad \text{and} \quad \|y_3 - y_1\| \geq \frac{1}{2}.$$

Continuing in this manner, we construct a sequence (y_n) such that

$$\|y_n - y_m\| \geq \frac{1}{2}$$

for all $n \geq m \geq 1$. Clearly this sequence doesn't have a convergent subsequence. □

Recall that in a Hilbert space $\mathcal{H}$, a sequence (x_n) converges weakly to x if

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

for all $y \in \mathcal{H}$. We do not have inner-products in a normed linear space, however there is still an analogue of weak convergence in normed linear spaces:

Definition 5.8. A sequence (x_n) in a normed linear space $\mathcal{X}$ is said to be **weakly convergent** if there exists $x \in \mathcal{X}$ such that

$$\ell(x_n) \to \ell(x)$$

for all $\ell \in \mathcal{X}^*$. If such x exists it is unique in this case we write $x_n \xrightarrow{w} x$.

Theorem 5.5. *Every weakly convergent sequence in a normed linear space is bounded. Moreover, if $x_n \xrightarrow{w} x$, then*

$$\|x\| \leq \liminf_{n \to \infty} \|x_n\|.$$

Proof. The proof relies on the uniform boundedness principle just as in the Hilbert space case. □

Theorem 5.6. *Every bounded sequence in a (separable) reflexive Banach space has a weakly convergent subsequence.*

Remark. The separable assumption can be dropped but the reflexive assumption cannot be dropped.

Proof. Diagonal argument. We need to use reflexivity in the last step in place of the Riesz Representation Theorem. □

Proposition 5.10. *Let $\mathcal{Y}$ be a closed subspace of a normed linear space $\mathcal{X}$. If (x_n) is a sequence in $\mathcal{Y}$ and $x_n \xrightarrow{w} x$, then $x \in \mathcal{Y}$.*

Proof. Similar to the Hilbert space case. We need to use

$$d(x, \mathcal{Y}) = \sup\{|\ell(x)| \mid \ell \in \mathcal{Y}^{\perp} \text{ and } \|\ell\| \leq 1\}.$$

□

Theorem 5.7. *Let $\mathcal{X}$ be a reflexive Banach space and let $\mathcal{Y}$ be a closed subspace of $\mathcal{X}$. Then for any $x \in \mathcal{X}$ there exists $y_0 \in \mathcal{Y}$ such that*

$$\|x - y_0\| = d(x, \mathcal{Y}).$$

Remark. Note that y_0 is not necessarily unique.

Proof. For each $n \in \mathbb{N}$, there exists $y_n \in \mathcal{Y}$ such that

$$d(x, \mathcal{Y}) \leq \|x - y_n\| < d(x, \mathcal{Y}) + \frac{1}{n}.$$

Then

$$\begin{aligned}\|y_n\| &\leq \|y_n - x\| + \|x\| \\ &< d(x, \mathcal{Y}) + \frac{1}{n} + \|x\| \\ &\leq d(x, \mathcal{Y}) + 1 + \|x\|\end{aligned}$$

for all $n \in \mathbb{N}$. Therefore (y_n) is a bounded sequence and hence contains a weakly convergent subsequence, say $(y_{\pi(n)})$ with $y_{\pi(n)} \xrightarrow{w} y_0$. Since $(y_{\pi(n)}) \subseteq \mathcal{Y}$, the weak limit y_0 is also in $\mathcal{Y}$. Then $x - y_{\pi(n)} \xrightarrow{w} x - y_0$, and hence

$$\|x - y_0\| \leq \liminf_{n \to \infty} \|x - y_{\pi(n)}\|.$$

Therefore

$$\begin{aligned}d(x, \mathcal{Y}) &\leq \|x - y_0\| \\ &\leq \liminf_{n \to \infty} \|x - y_{\pi(n)}\| \\ &\leq \liminf_{n \to \infty} \left(d(x, \mathcal{Y}) + \frac{1}{\pi(n)} \right) \\ &= d(x, \mathcal{Y}),\end{aligned}$$

which implies $\|x - y_0\| = d(x, \mathcal{Y})$. □

Part II

Homework Problems and Solutions

6 Homework 1

6.1 Polarization Identity

Proposition 6.1. *(Polarization Identity) For $x, y \in \mathscr{V}$ we have*

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

Proof. We calculate

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

and

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

and

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

and

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

Adding these together gives us our desired result. □

6.2 Parallelogram Identity

Proposition 6.2. *(Parallelogram Identity) For $x, y \in \mathscr{V}$ we have*

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

Proof. We calculate

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

and

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

Adding these together gives us our desired result. □

The geometric interpretation of Proposition (6.2) in the case where $\mathscr{V} = \mathbb{R}^3$ can be seen below:

6.3 Pythagorean Theorem

Proposition 6.3. *(Pythagorean Theorem) Let x and y be nonzero vectors in $\mathscr{V}$ such that $\langle x, y\rangle = 0$ (we call such vectors **orthogonal** to one another). Then*

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

Proof. We have

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\
&= \langle x, x\rangle + \langle y, y\rangle \\
&= \|x\|^2 + \|y\|^2.
\end{aligned}$$

□

6.4 $x_n \to x$ and $y_n \to y$ implies $\langle x_n, y_n\rangle \to \langle x, y\rangle$

Proposition 6.4. *Let (x_n) and (y_n) be two sequences in $\mathscr{V}$. Then the following statements hold:*

1. *If $x_n \to x$ and $y_n \to y$, then $x_n + y_n \to x + y$.*
2. *If $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$. In particular, $\|x_n\| \to \|x\|$.*

Proof.

1. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $\|x_n - x\| < \varepsilon/2$ and $\|y_n - y\| < \varepsilon/2$. Then $n \geq N$ implies

$$\begin{aligned}
\|(x_n + y_n) - (x+y)\| &\leq \|x_n - x\| + \|y_n - y\| \\
&< \varepsilon/2 + \varepsilon/2 \\
&= \varepsilon.
\end{aligned}$$

2. Since $y_n \to y$, there exists $M \geq 0$ such that $\|y_n\| \leq M$ for all $n \in \mathbb{N}$. Choose such an M and let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $\|x_n - x\| < \varepsilon/2M$ and $\|y_n - y\| < \varepsilon/2\|x\|$. Then $n \geq N$ implies

$$\begin{aligned}
|\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n\rangle - \langle x, y_n\rangle + \langle x, y_n\rangle - \langle x, y\rangle| \\
&\leq |\langle x_n, y_n\rangle - \langle x, y_n\rangle| + |\langle x, y_n\rangle - \langle x, y\rangle| \\
&= |\langle x_n - x, y_n\rangle| + |\langle x, y_n - y\rangle| \\
&\leq \|x_n - x\|\|y_n\| + \|x\|\|y_n - y\| \\
&\leq \|x_n - x\|M + \|x\|\|y_n - y\| \\
&< \varepsilon/2 + \varepsilon/2 \\
&= \varepsilon.
\end{aligned}$$

To see that $\|x_n\| \to \|x\|$, we just set $y_n = x_n$. Then

$$\begin{aligned}
\|x_n\| &= \sqrt{\langle x_n, x_n\rangle} \\
&\to \sqrt{\langle x, x\rangle} \\
&= \|x\|,
\end{aligned}$$

where we were allowed to take limits inside the square root function since the square root function is continuous on $\mathbb{R}_{\geq 0}$.

□

6.5 Inner-product on $\mathrm{M}_{m\times n}(\mathbb{R})$

Proposition 6.5. *Let $\langle \cdot, \cdot\rangle \colon M_{m\times n}(\mathbb{R}) \times M_{m\times n}(\mathbb{R}) \to \mathbb{R}$ be given by*

$$\langle A, B\rangle = Tr(B^\top A),$$

for all $A, B \in M_n(\mathbb{C})$. Then the pair $(M_n(\mathbb{C}), \langle \cdot, \cdot\rangle)$ forms an inner-product space.

Proof. Linearity in the first argument follows from distributivity of matrix multiplication and from linearity of the trace function: Let $A, B, C \in \mathrm{M}_{m\times n}(\mathbb{R})$. Then

$$\begin{aligned}\langle A+B, C\rangle &= \mathrm{Tr}(C^\top(A+B))\\ &= \mathrm{Tr}(C^\top A + C^\top B)\\ &= \mathrm{Tr}(C^\top A) + \mathrm{Tr}(C^\top B)\\ &= \langle A, C\rangle + \langle B, C\rangle.\end{aligned}$$

Symmetry of $\langle\cdot,\cdot\rangle$ follows from the fact that $\mathrm{Tr}(A) = \mathrm{Tr}(A^\top)$ for all $A \in \mathrm{M}_{m\times n}(\mathbb{R})$: Let $A, B \in \mathrm{M}_{m\times n}(\mathbb{R})$. Then

$$\begin{aligned}\langle A, B\rangle &= \mathrm{Tr}(B^\top A)\\ &= \mathrm{Tr}((B^\top A)^\top)\\ &= \mathrm{Tr}(A^\top B)\\ &= \langle B, A\rangle.\end{aligned}$$

Finally, to see positive-definiteness of $\langle\cdot,\cdot\rangle$, let

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{m1} & \cdots & a_{mn}\end{pmatrix} \in \mathrm{M}_{m\times n}(\mathbb{R}).$$

Then

$$\begin{aligned}\langle A, A\rangle &= \mathrm{Tr}(A^\top A)\\ &= \mathrm{Tr}\begin{pmatrix} a_{11} & \cdots & a_{m1}\\ \vdots & \ddots & \vdots\\ a_{1n} & \cdots & a_{mn}\end{pmatrix}\begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{m1} & \cdots & a_{mn}\end{pmatrix}\\ &= \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}^2.\end{aligned}$$

is a sum of its entries squared. This implies positive-definiteness. □

6.6 Inner-product on $\mathbb{C}^n$

Proposition 6.6. *Let $\langle\cdot,\cdot\rangle \colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ be given by*

$$\langle x, y\rangle = \langle (x_1, \dots, x_n), (y_1, \dots, y_n)\rangle = \sum_{i=1}^{n} x_i \overline{y}_i.$$

for all $x, y \in \mathbb{C}^n$. Then the pair $(\mathbb{C}^n, \langle\cdot,\cdot\rangle)$ forms an inner-product space.

Proof. For linearity in the first argument follows from linearity, let $x, y, z \in \mathbb{C}^n$. Then

$$\begin{aligned}\langle x+y, z\rangle &= \sum_{i=1}^{n}(x_i + y_i)\overline{z}_i\\ &= \sum_{i=1}^{n} x_i\overline{z}_i + \sum_{i=1}^{n} y_i\overline{z}_i\\ &= \langle x, z\rangle + \langle y, z\rangle.\end{aligned}$$

For conjugate symmetry of $\langle\cdot,\cdot\rangle$, let $x, y \in \mathbb{C}^n$. Then

$$\begin{aligned}\langle x, y\rangle &= \sum_{i=1}^{n} x_i\overline{y}_i\\ &= \sum_{i=1}^{n} \overline{\overline{x_i\overline{y}_i}}\\ &= \sum_{i=1}^{n} \overline{y_i\overline{x}_i}\\ &= \overline{\langle y, x\rangle}.\end{aligned}$$

For positive-definiteness of $\langle \cdot, \cdot \rangle$, let $x \in \mathbb{C}^n$. Then

$$\begin{aligned}\langle x, x \rangle &= \sum_{i=1}^{n} x_i \overline{x}_i \\ &= \sum_{i=1}^{n} |x_i|^2.\end{aligned}$$

is a sum of its components absolute squared. This implies positive-definiteness. $\square$

6.7 Cauchy-Schwarz

This follows from an easy application of Cauchy-Schwarz, but here's another method (which turns out to be equivalent to Cauchy-Schwarz). We need the following two lemmas:

Lemma 6.1. *Let a and b be nonnegative real numbers. Then we have*

$$2ab \leq a^2 + b^2. \tag{24}$$

Proof. We have

$$\begin{aligned}0 &\leq (a-b)^2 \\ &= a^2 - 2ab + b^2.\end{aligned}$$

Therefore the inequality (24) follows from adding $2ab$. $\square$

Lemma 6.2. *Let $a_1, \dots, a_n$ and $b_1, \dots, b_n$ be nonnegative real numbers. Then*

$$\left(\sum_{i=1}^{n} a_i b_i \right)^2 \leq \left(\sum_{i=1}^{n} a_i^2 \right) \left(\sum_{i=1}^{n} b_i^2 \right).$$

Proof. We have

$$\begin{aligned}\left(\sum_{i=1}^{n} a_i b_i \right)^2 &= \sum_{i=1}^{n} a_i^2 b_i^2 + \sum_{1 \leq i < j \leq n} 2 a_i b_j a_j b_i \\ &\leq \sum_{i=1}^{n} a_i^2 b_i^2 + \sum_{1 \leq i < j \leq n} (a_i^2 b_j^2 + a_j^2 b_i^2) \\ &= \left(\sum_{i=1}^{n} a_i^2 \right) \left(\sum_{i=1}^{n} b_i^2 \right)\end{aligned}$$

where the inequality in the second line follows from Lemma (6.1) applied to $a_i b_j$ and $a_j b_i$. $\square$

Corollary. *Let $x, y \in \mathbb{C}^n$. Then*

$$\sum_{i=1}^{n} |x_i||y_i| \leq \sqrt{\sum_{i=1}^{n} |x_i|^2} \sqrt{\sum_{i=1}^{n} |y_i|^2}.$$

Proof. This follows from by taking squares on both sides and applying Lemma (6.2) since the $|x_i|$ and $|y_i|$ are nonnegative real numbers. $\square$

6.8 Inner-product on $\ell^2(\mathbb{N})$

Proposition 6.7. *Let $\ell^2(\mathbb{N})$ be the set of all sequence (x_n) in $\mathbb{C}$ such that*

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty$$

and let $\langle \cdot, \cdot \rangle \colon \ell^2(\mathbb{N}) \times \ell^2(\mathbb{N}) \to \mathbb{C}$ be given by

$$\langle (x_n), (y_n) \rangle = \sum_{n=1}^{\infty} x_n \overline{y}_n.$$

for all $(x_n), (y_n) \in \ell^2(\mathbb{N})$. Then the pair $(\ell^2(\mathbb{N}), \langle \cdot, \cdot \rangle)$ forms an inner-product space.

Proof. We first need to show that $\ell^2(\mathbb{N})$ is indeed a vector space. In fact, we will show that $\ell^2(\mathbb{N})$ is a subspace of $\mathbb{C}^N$, the set of all sequences in $\mathbb{C}$. Let $(x_n),(y_n) \in \ell^2(\mathbb{N})$ and $\lambda \in \mathbb{C}$. Then Lemma (6.1) implies

$$\begin{aligned}
\sum_{n=1}^{\infty} |\lambda x_n + y_n|^2 &\leq \sum_{n=1}^{\infty} |\lambda x_n|^2 + \sum_{n=1}^{\infty} |y_n|^2 + \sum_{n=1}^{\infty} 2|\lambda x_n||y_n| \\
&\leq \lambda^2 \sum_{n=1}^{\infty} |x_n|^2 + \sum_{n=1}^{\infty} |y_n|^2 + \lambda^2 \sum_{n=1}^{\infty} |x_n|^2 + |y_n|^2 \\
&< \infty.
\end{aligned}$$

Therefore $(\lambda x_n + y_n) \in \ell^2(\mathbb{N})$, which implies $\ell^2(\mathbb{N})$ is a subspace of $\mathbb{C}^{\mathbb{N}}$.

Next, let us show that the inner product converges, and hence is defined everywhere. Let $(x_n),(y_n) \in \ell^2(\mathbb{N})$. Then it follows from Lemma (6.1) that

$$\begin{aligned}
\sum_{n=1}^{\infty} |x_n \overline{y}_n| &= \sum_{n=1}^{\infty} |x_n||y_n| \\
&\leq \sum_{n=1}^{\infty} \frac{|x_n|^2 + |y_n|^2}{2} \\
&= \frac{1}{2}\sum_{n=1}^{\infty} |x_n|^2 + \frac{1}{2}\sum_{n=1}^{\infty} |y_n|^2 \\
&< \infty.
\end{aligned}$$

Therefore $\sum_{n=1}^{\infty} x_n \overline{y}_n$ is absolutely convergent, which implies it is convergent. (We can't use Cauchy-Schwarz here since we haven't yet shown that $\langle \cdot, \cdot \rangle$ is in fact an inner-product).

Finally, let us shows that $\langle \cdot, \cdot \rangle$ is an inner-product. Linearity in the first argument follows from distrubitivity of multiplication and linearity of taking infinite sums. For conjugate symmetry, let $(x_n),(y_n) \in \ell^2(\mathbb{N})$. Then

$$\begin{aligned}
\langle (x_n),(y_n) \rangle &= \sum_{n=1}^{\infty} x_n \overline{y}_n \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} x_n \overline{y}_n \\
&= \overline{\overline{\lim_{N\to\infty} \sum_{n=1}^{N} x_n \overline{y}_n}} \\
&= \overline{\lim_{N\to\infty} \sum_{n=1}^{N} \overline{x_n \overline{y}_n}} \\
&= \overline{\lim_{N\to\infty} \sum_{n=1}^{N} y_n \overline{x}_n} \\
&= \overline{\sum_{n=1}^{\infty} y_n \overline{x}_n} \\
&= \overline{\langle (y_n),(x_n) \rangle},
\end{aligned}$$

where we were allowed to bring the conjugate inside the limit since the conjugate function is continuous on $\mathbb{C}$. For positive-definiteness, let $(x_n) \in \ell^2(\mathbb{N})$. Then

$$\begin{aligned}
\langle (x_n),(x_n) \rangle &= \sum_{n=1}^{\infty} x_n \overline{x}_n \\
&= \sum_{n=1}^{\infty} |x_n|^2 \\
&\geq 0.
\end{aligned}$$

If $\sum_{n=1}^{\infty} |x_n|^2 = 0$, then clearly we must have $x_n = 0$ for all n. $\square$

6.9 Cauchy-Schwarz Application

Proposition 6.8. *Let $(x_n) \in \ell^2(\mathbb{N})$ such that $\sum_{n=1}^{\infty} |x_n|^2 = 1$. Then*

$$\sum_{n=1}^{\infty} \frac{|x_n|}{2^n} \leq \frac{1}{\sqrt{3}}. \tag{25}$$

where the inequality (25) becomes an equality if and only if $|x_n| = \sqrt{3} \cdot 2^{-n}$ for all n.

Proof. By Cauchy-Schwarz, we have

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{|x_n|}{2^n} &= |\langle (|x_n|), (2^{-n}) \rangle| \\
&\leq \|(|x_n|)\| \|(2^{-n})\| \\
&= \sqrt{\sum_{n=1}^{\infty} |x_n|^2} \sqrt{\sum_{n=1}^{\infty} 2^{-2n}} \\
&1 \cdot \sqrt{\sum_{n=0}^{\infty} \left(\frac{1}{4}\right)^n - 1} \\
&= \sqrt{\frac{1}{1 - 1/4} - 1} \\
&= \sqrt{\frac{4}{3} - 1} \\
&= \frac{1}{\sqrt{3}}.
\end{aligned}$$

where the inequality becomes an equality if and only if $(|x_n|)$ and (2^{-n}) are linearly dependent. This means that there is a $\lambda \in \mathbb{C}$ such that $|x_n| = \lambda 2^{-n}$ for all n. To find this λ, write

$$\begin{aligned}
1 &= \sum_{n=1}^{\infty} |x_n|^2 \\
&= \sum_{n=1}^{\infty} |\lambda 2^{-n}|^2 \\
&= |\lambda|^2 \sum_{n=1}^{\infty} \left(\frac{1}{4}\right)^n \\
&= \frac{|\lambda|^2}{3}.
\end{aligned}$$

Thus, any $\lambda \in \mathbb{C}$ such that $|\lambda| = \sqrt{3}$ works. (Actually, we must have $\lambda = \sqrt{3}$ since $\lambda = |x_n| 2^n$ is positive). □

6.10 Cauchy-Schwarz Application

Proposition 6.9. *Let $f \in C[0,1]$ such that $\int_0^1 |f(x)|^2 dx = 1$. Then*

$$\int_0^1 |f(x)| \sin(\pi x) dx \leq \frac{1}{\sqrt{2}},$$

where the inequality becomes an equality if and only if $|f(x)| = \sqrt{2} \sin(\pi x)$.

Proof. First note that

$$\begin{aligned}
\int_0^1 \sin^2(\pi x) dx &= \int_0^1 \cos^2(\pi x) dx \\
&= \int_0^1 (1 - \sin^2(\pi x)) dx
\end{aligned}$$

implies $\int_0^1 \sin^2(\pi x)dx = 1/2$, where in the first equality above we used integration by parts with $u = \sin(\pi x)$ and $dv = \sin(\pi x)dx$. Therefore, by Cauchy-Schwarz, we have

$$\begin{aligned}\int_0^1 |f(x)|\sin(\pi x)dx &\leq \sqrt{\int_0^1 |f(x)|^2 dx} \cdot \sqrt{\int_0^1 \sin^2(\pi x)dx} \\ &= 1 \cdot \frac{1}{\sqrt{2}} \\ &= \frac{1}{\sqrt{2}},\end{aligned}$$

where the inequality becomes an equality if and only if $|f(x)|$ and $\sin(\pi x)$ and linearly dependent. This means that there is a $\lambda \in \mathbb{C}$ such that $|f(x)| = \lambda \sin(\pi x)$ for all x. To find this λ, write

$$\begin{aligned}1 &= \int_0^1 |f(x)|^2 dx \\ &= \int_0^1 |\lambda \sin(\pi x)|^2 dx \\ &= |\lambda|^2 \int_0^1 \sin^2(\pi x)dx \\ &= \frac{|\lambda|^2}{2}.\end{aligned}$$

Thus, any $\lambda \in \mathbb{C}$ such that $|\lambda| = \sqrt{2}$ works. (Actually, we must have $\lambda = \sqrt{2}$ since $\lambda = |f(x)|/\sin(\pi x)$ is positive). $\square$

Remark. If we tried to apply Lemma (6.1) at each $x \in [0,1]$, we'd only get the weaker result:

$$\begin{aligned}\int_0^1 |f(x)|\sin(\pi x)dx &\leq \frac{1}{2}\left(\int_0^1 |f(x)|^2 dx + \int_0^1 \sin^2(\pi x)dx\right) \\ &= \frac{1}{2} + \frac{1}{4} \\ &= \frac{3}{4}.\end{aligned}$$

7 Homework 2

Throughout this homework, let $\mathscr{V}$ be an inner-product space over $\mathbb{C}$. If $x \in \mathscr{V}$ and $r > 0$, then we define

$$B_r(x) := \{y \in \mathscr{V} \mid \|y - x\| < r\}$$

to be the **open ball centered at** x **and of radius** r. We also define

$$B_r[x] := \{y \in \mathscr{V} \mid \|y - x\| \leq r\}$$

to be the **closed ball centered at** x **and of radius** r.

7.1 Translating Open Balls

Proposition 7.1. *Let $a \in \mathscr{V}$ and $r > 0$. Then*

$$B_r(a) = a + rB_1(0).$$

Proof. We prove this in two steps.

Step 1: We show $B_r(a) = a + B_r(0)$: Let $x \in B_r(a)$, so $\|x - a\| < r$. This implies $x - a \in B_r(0)$. Thus

$$\begin{aligned}x &= a + (x - a) \\ &\in a + B_r(0).\end{aligned}$$

Therefore $B_r(a) \subseteq a + B_r(0)$.

Conversely, let $a + y \in a + B_r(0)$ where $y \in B_r(0)$, so $\|y\| < r$. This implies $\|(a + y) - a\| < r$. In other words, $a + y \in B_r(a)$. Therefore $a + B_r(0) \subseteq B_r(a)$.

Step 2: We show $B_r(0) = rB_1(0)$: Let $x \in B_r(0)$, so $\|x\| < r$. Then since $r > 0$, we have

$$\begin{aligned} 1 &> (1/r)\|x\| \\ &= \|x/r\|. \end{aligned}$$

In other words, $x/r \in B_1(0)$. Thus

$$\begin{aligned} x &= r(x/r) \\ &\in rB_1(0). \end{aligned}$$

Therefore $B_r(0) \subseteq rB_1(0)$.

Conversely, let $ry \in rB_1(0)$ where $y \in B_1(0)$, so $\|y\| < 1$. Then since $r > 0$, we have

$$\begin{aligned} \|ry\| &= r\|y\| \\ &< 1. \end{aligned}$$

In other words, $ry \in B_1(0)$. Therefore $rB_1(0) \subseteq B_r(0)$. □

7.2 Closure of Open Balls

Lemma 7.1. *Let $x \in \mathscr{V}$ and for each $n \in \mathbb{N}$ let $x_n \in \mathscr{V}$ such that $\|x_n - x\| < 1/n$. Then $x_n \to x$.*

Proof. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $1/N < \varepsilon$. Then $n \geq N$ implies

$$\begin{aligned} \|x_n - x\| &< 1/n \\ &\leq 1/N \\ &< \varepsilon. \end{aligned}$$

□

Proposition 7.2. *Let $a \in \mathscr{V}$ and $r > 0$. Then*

$$\overline{B_r(a)} = B_r[a].$$

Proof. Let $x \in \overline{B_r(a)}$. Choose a sequence (x_n) of elements in $B_r(a)$ such that $x_n \to x$. Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that $\|x_N - x\| < \varepsilon$. Then

$$\begin{aligned} \|x - a\| &= \|x - x_N + x_N - a\| \\ &\leq \|x - x_N\| + \|x_N - a\| \\ &< \varepsilon + r. \end{aligned}$$

Thus $\|x - a\| < r + \varepsilon$ for all $\varepsilon > 0$. This implies $\|x - a\| \leq r$, or in other words, $x \in B_r[a]$. Thus $\overline{B_r(a)} \subseteq B_r[a]$.

Conversely, let $x \in B_r[a]$ and let $n \in \mathbb{N}$. We first observe that for each $t \in (0,1)$, we have

$$\begin{aligned} \|(x + t(a - x)) - a\| &= \|(1-t)x - (1-t)a\| \\ &= (1-t)\|x - a\| \\ &< r. \end{aligned}$$

Thus $x + t(a - x) \in B_r(a)$ for all $t \in (0,1)$. Now let $n \in \mathbb{N}$. Choose $t_n \in (0,1)$ such that $t_n < \|x - a\|/n$. Then

$$\begin{aligned} \|(x + t_n(a - x)) - x\| &= \|t_n(x - a)\| \\ &= t_n\|x - a\| \\ &< 1/n. \end{aligned}$$

Thus $(x + t_n(a - x))$ is a sequence of elements of elements in $B_r(a)$ such that $x + t_n(a - x) \to x$ (by Lemma (7.1)), hence $x \in \overline{B_r(a)}$. Thus $B_r[a] \subseteq \overline{B_r(a)}$. □

7.3 Closed Set Properties

Lemma 7.2. *Let $A \subseteq \mathscr{V}$.*

Proposition 7.3. *Let $A \subseteq \mathscr{V}$ and let $C_1, C_2 \subseteq \mathscr{V}$ such that C_1 and C_2 are closed. Then*

1. *$\overline{A}$ is a closed set.*
2. *$\overline{A}$ is the smallest closed set that contains A, i.e., for any closed set B such that $A \subseteq B$ we have $\overline{A} \subseteq B$. In particular, $\overline{A} = A$ if and only if A is closed.*
3. *The union of C_1 and C_2 is closed.*
4. *The intersection of C_1 and C_2 is closed.*
5. *An infinite union of closed sets may not be closed.*

Proof.

1. We will show that $\overline{A}$ is closed by showing that $\mathscr{V} \backslash \overline{A}$ is open. To show that $\mathscr{V} \backslash \overline{A}$ is open, it suffices to show that for each $x \in \mathscr{V} \backslash \overline{A}$ there exists an open neighborhood of x which is contained in $\mathscr{V} \backslash \overline{A}$. Assume (for a contradiction) that $\mathscr{V} \backslash \overline{A}$ is not open. Choose $x \in \mathscr{V} \backslash \overline{A}$ such that every open neighborhood of x meets $\overline{A}$. In particular, for each $n \in \mathbb{N}$, there exists $x_n \in B_{1/n}(x) \cap \overline{A}$. Choose such x_n for all $n \in \mathbb{N}$. Then by Lemma (7.1), we must have $x_n \to x$, and hence $x \in \overline{\overline{A}} = \overline{A}$. This is a contradiction.
2. Let B be any closed set which contains A. Suppose $x \in \overline{A}$. Choose a sequence (x_n) of elements in A such that $x_n \to x$. Assume (for a contradiction) that $x \in \mathscr{V} \backslash B$. Choose $\varepsilon > 0$ such that $B_\varepsilon(x) \cap B = \emptyset$ (we can do this since $\mathscr{V} \backslash B$ is open). But then the sequence (x_n) of elements in B cannot converge to x since $x_n \notin B_\varepsilon(x)$ for all $n \in \mathbb{N}$. This is a contradiction. For the last statement. If A is closed, then since $\overline{A}$ is the smallest closed set containing A, we must have $A = \overline{A}$. And if $A = \overline{A}$, then since $\overline{A}$ is the smallest closed set containing A, the set A itself must be closed.
3. Combining 2 with an identity we proved in class, we have
$$\begin{aligned} C_1 \cup C_2 &= \overline{C}_1 \cup \overline{C}_2 \\ &= \overline{C_1 \cup C_2}. \end{aligned}$$
Therefore $C_1 \cup C_2$ is closed.
4. Combining 2 with a couple identities that we proved in class, we have
$$\begin{aligned} \overline{C_1 \cap C_2} &\supseteq C_1 \cap C_2 \\ &= \overline{C}_1 \cap \overline{C}_2 \\ &\supseteq \overline{C_1 \cap C_2}. \end{aligned}$$
Therefore $C_1 \cap C_2$ is closed.
5. Consider $\mathscr{V} = \mathbb{R}$ and $C_n = [0, 1 - 1/n]$ for all $n \in \mathbb{N}$. Then $\bigcup_{n=1}^\infty C_n = [0, 1)$, which is not closed in $\mathbb{R}$.

□

7.4 Distance Between Point and Set Properties

Proposition 7.4. *Let $E \subseteq \mathscr{V}$ and let $x, y \in \mathscr{V}$. Then*

1. *$d(x, E) = 0$ if and only if $x \in \overline{E}$;*
2. *$|d(x, E) - d(y, E)| \leq \|x - y\|$.*

Proof.

1. First suppose that $d(x, E) = 0$. For each $n \in \mathbb{N}$, choose $x_n \in E$ such that $\|x_n - x\| < 1/n$ (if we couldn't find such an x_n, then 0 would not be the infinum). Now we apply Lemma (7.1) to find that (x_n) is a sequence of elements in E such that $x_n \to x$. Therefore $x \in \overline{E}$. Conversely, suppose that $x \in \overline{E}$. Choose a sequence (x_n) of elements in E such that $x_n \to x$. Then we have
$$0 \leq d(x, E) < \|x_n - x\|$$
for all $n \in \mathbb{N}$. This implies $d(x, E) = 0$.

2. Without loss of generality, we may assume that $d(x,E) \geq d(y,E)$. Thus we are trying to show that $d(x,E) \leq \|x-y\| + d(y,E)$. Choose $y_n \in E$ such that $\|y_n - y\| < d(y,E) + 1/n$ for all $n \in \mathbb{N}$. Then

$$\begin{aligned} d(x,E) &\leq \|x - y_n\| \\ &= \|x - y + y - y_n\| \\ &\leq \|x-y\| + \|y - y_n\| \\ &< \|x-y\| + d(y,E) + 1/n. \end{aligned}$$

Taking $n \to \infty$ gives us our desired result.

□

7.5 Distance is a Norm

Proposition 7.5. *Let $\mathscr{H}$ be a Hilbert space of $\mathbb{C}$, let $\mathscr{K}$ be a closed subspace of $\mathscr{H}$, let $x, y \in \mathscr{H}$, and let $\lambda \in \mathbb{C}$. Then*

1. $d(\lambda x, \mathscr{K}) = |\lambda| d(x, \mathscr{K})$;
2. $d(x+y, \mathscr{K}) \leq d(x, \mathscr{K}) + d(y, \mathscr{K})$.

Proof.

1. Choose a sequence (y_n) of elements in $\mathcal{A}$ such that

$$\|x - y_n\| < d(x, \mathcal{A}) - 1/n$$

for all $n \in \mathbb{N}$. Then

$$\begin{aligned} \|\lambda x - \lambda y_n\| &= |\lambda| \|x - y_n\| \\ &< |\lambda| d(x, \mathcal{A}) - |\lambda|/n \end{aligned}$$

$$|\lambda| d(x, \mathcal{A}) \leq |\lambda| \|x - y_n\| < d(x, \mathcal{A}) - 1/n$$

2. Let a be the unique element in $\mathscr{K}$ such that $d(x, \mathscr{K}) = \|x - a\|$. Then $\lambda a \in \mathscr{K}$, and so

$$\begin{aligned} |\lambda| d(x, \mathscr{K}) &= |\lambda| \|x - a\| \\ &= \|\lambda x - \lambda a\| \\ &\geq d(\lambda x, \mathscr{K}). \end{aligned}$$

Conversely, let b be the unique element in $\mathscr{K}$ such that $d(\lambda x, \mathscr{K}) = \|x - b\|$. Then $b/\lambda \in \mathscr{K}$, and so

$$\begin{aligned} d(\lambda x, \mathscr{K}) &= \|\lambda x - b\| \\ &= |\lambda| \|x - b/\lambda\| \\ &\geq |\lambda| d(x, \mathscr{K}). \end{aligned}$$

3. Let a be the unique element in $\mathscr{K}$ such that $d(x, \mathscr{K}) = \|x - a\|$ and let b be the unique element in $\mathscr{K}$ such that $d(y, \mathscr{K}) = \|y - b\|$. Then $a + b \in \mathscr{K}$, and so

$$\begin{aligned} d(x+y, \mathscr{K}) &\leq \|x + y - (a+b)\| \\ &= \|x - a\| + \|y - b\| \\ &= d(x, \mathscr{K}) + d(y, \mathscr{K}). \end{aligned}$$

□

8 Homework 3

8.1 Orthogonal Projection $P_{\mathcal{K}}$ Properties

Proposition 8.1. *Let $\mathcal{K}$ be a closed subspace of a Hilbert space $\mathcal{H}$ and let $x \in \mathcal{H}$. Then*

1. $P_{\mathcal{K}} x = x$ *if and only if* $x \in \mathcal{K}$.
2. $\|P_{\mathcal{K}} x\| = \|x\|$ *if and only if* $x \in \mathcal{K}$.
3. $\langle P_{\mathcal{K}} x, x \rangle = \|P_{\mathcal{K}} x\|^2$.

Proof.

1. If $\mathrm{P}_{\mathcal{K}}x = x$, then it is clear that $x \in \mathcal{K}$ since $\mathrm{P}_{\mathcal{K}}x \in \mathcal{K}$. For the reverse direction, suppose $x \in \mathcal{K}$. Then

$$\begin{aligned} 0 &= \|x - x\| \\ &\geq d(x, \mathcal{K}) \\ &= \|x - \mathrm{P}_{\mathcal{K}}x\| \\ &\geq 0 \end{aligned}$$

implies $\|x - x\| = d(x, \mathcal{K}) = \|x - \mathrm{P}_{\mathcal{K}}x\|$, and so by uniqueness of $\mathrm{P}_{\mathcal{K}}x$, we must have $x = \mathrm{P}_{\mathcal{K}}x$.

2. If $x \in \mathcal{K}$, then it is clear that $\|\mathrm{P}_{\mathcal{K}}x\| = \|x\|$ since $x = \mathrm{P}_{\mathcal{K}}x$ by 1. For the reverse direction, suppose $\|\mathrm{P}_{\mathcal{K}}x\| = \|x\|$. Since $\langle x - \mathrm{P}_{\mathcal{K}}x, \mathrm{P}_{\mathcal{K}}x\rangle = 0$, the Pythagorean Theorem[3] implies

$$\begin{aligned} \|x\|^2 &= \|x - \mathrm{P}_{\mathcal{K}}x + \mathrm{P}_{\mathcal{K}}x\|^2 \\ &= \|x - \mathrm{P}_{\mathcal{K}}x\|^2 + \|\mathrm{P}_{\mathcal{K}}x\|^2 \\ &= \|x - \mathrm{P}_{\mathcal{K}}x\|^2 + \|x\|^2. \end{aligned}$$

Thus $\|x - \mathrm{P}_{\mathcal{K}}x\|^2 = 0$, which implies $x = \mathrm{P}_{\mathcal{K}}x$ since the metric is positive definite.

3. We have

$$\begin{aligned} 0 &= \langle x - \mathrm{P}_{\mathcal{K}}x, \mathrm{P}_{\mathcal{K}}x\rangle \\ &= \langle x, \mathrm{P}_{\mathcal{K}}x\rangle - \langle \mathrm{P}_{\mathcal{K}}x, \mathrm{P}_{\mathcal{K}}x\rangle \\ &= \langle x, \mathrm{P}_{\mathcal{K}}x\rangle - \|\mathrm{P}_{\mathcal{K}}x\|^2, \end{aligned}$$

which implies $\langle x, \mathrm{P}_{\mathcal{K}}x\rangle = \|\mathrm{P}_{\mathcal{K}}x\|^2$. Since $\|\mathrm{P}_{\mathcal{K}}x\|^2$ is a real number, this implies $\langle \mathrm{P}_{\mathcal{K}}x, x\rangle = \|\mathrm{P}_{\mathcal{K}}x\|^2$.

□

8.2 $\mathcal{K}_1 \subseteq \mathcal{K}_2$ if and only if $\langle \mathrm{P}_{\mathcal{K}_1}x, x\rangle \leq \langle \mathrm{P}_{\mathcal{K}_2}x, x\rangle$ for all $x \in \mathcal{H}$

Proposition 8.2. *Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be closed subspaces of a Hilbert space $\mathcal{H}$. Then $\mathcal{K}_1 \subseteq \mathcal{K}_2$ if and only if $\langle P_{\mathcal{K}_1}x, x\rangle \leq \langle P_{\mathcal{K}_2}x, x\rangle$ for all $x \in \mathcal{H}$.*

Proof. By Proposition (8.1), we can replace the condition $\langle \mathrm{P}_{\mathcal{K}_1}x, x\rangle \leq \langle \mathrm{P}_{\mathcal{K}_2}x, x\rangle$ for all $x \in \mathcal{H}$ with $\|\mathrm{P}_{\mathcal{K}_1}x\|^2 \leq \|\mathrm{P}_{\mathcal{K}_2}x\|^2$ for all $x \in \mathcal{H}$. Suppose $\mathcal{K}_1 \subseteq \mathcal{K}_2$. Then

$$\begin{aligned} \|x - \mathrm{P}_{\mathcal{K}_2}x\| &= \mathrm{d}(x, \mathcal{K}_2) \\ &= \inf\{\|x - y\| \mid y \in \mathcal{K}_2\} \\ &\leq \inf\{\|x - y\| \mid y \in \mathcal{K}_1\} \\ &= \mathrm{d}(x, \mathcal{K}_1) \\ &= \|x - \mathrm{P}_{\mathcal{K}_1}x\|. \end{aligned}$$

Therefore by the Pythagorean Theorem, we have

$$\begin{aligned} \|\mathrm{P}_{\mathcal{K}_1}x\|^2 &= \|x\|^2 - \|x - \mathrm{P}_{\mathcal{K}_1}x\|^2 \\ &\leq \|x\|^2 - \|x - \mathrm{P}_{\mathcal{K}_2}x\|^2 \\ &= \|\mathrm{P}_{\mathcal{K}_2}x\|^2. \end{aligned}$$

Conversely, suppose $\|\mathrm{P}_{\mathcal{K}_1}x\|^2 \leq \|\mathrm{P}_{\mathcal{K}_2}x\|^2$ for all $x \in \mathcal{H}$. Equivalently, by the Pythagorean Theorem, we have

$$\begin{aligned} \|x - \mathrm{P}_{\mathcal{K}_1}x\|^2 &= \|x\|^2 - \|\mathrm{P}_{\mathcal{K}_1}x\|^2 \\ &\leq \|x\|^2 - \|\mathrm{P}_{\mathcal{K}_2}x\|^2 \\ &= \|x - \mathrm{P}_{\mathcal{K}_2}x\|^2 \end{aligned}$$

for all $x \in \mathcal{K}_1$. Now let $x \in \mathcal{K}_1$. Then $x = \mathrm{P}_{\mathcal{K}_1}x$ by Proposition (8.1). Thus

$$\begin{aligned} 0 &= \|x - x\|^2 \\ &= \|x - \mathrm{P}_{\mathcal{K}_1}x\|^2 \\ &\geq \|x - \mathrm{P}_{\mathcal{K}_2}x\|^2, \end{aligned}$$

which implies $x = \mathrm{P}_{\mathcal{K}_2}x$ since the metric is positive definite. Applying Proposition (8.1) again, we see that $x \in \mathcal{K}_2$, and hence $\mathcal{K}_1 \subseteq \mathcal{K}_2$. □

[3] Theorem (8.1) in the Appendix.

8.3 $\|P_{\mathcal{K}^\perp}x\| = d(x, \mathcal{K})$ for all $x \in \mathcal{H}$

Proposition 8.3. *Let $\mathcal{K}$ be a closed subspace of a Hilbert space $\mathcal{H}$. Then $\|P_{\mathcal{K}^\perp}x\| = d(x, \mathcal{K})$ for all $x \in \mathcal{H}$.*

Proof. From a theorem[4] we proved in class, we know that x can be uniquely decomposed as

$$x = P_{\mathcal{K}}x + (x - P_{\mathcal{K}}x), \tag{26}$$

for unique $P_{\mathcal{K}}x \in \mathcal{K}$ and unique $x - P_{\mathcal{K}}x \in \mathcal{K}^\perp$. Since $\mathcal{K}^\perp$ is another closed subspace[5] of $\mathcal{H}$, we can uniquely decompose x as

$$x = P_{\mathcal{K}^\perp}x + (x - P_{\mathcal{K}^\perp}x) \tag{27}$$

for unique $P_{\mathcal{K}^\perp}x \in \mathcal{K}^\perp$ and unique $x - P_{\mathcal{K}^\perp}x \in (\mathcal{K}^\perp)^\perp = \mathcal{K}$. It follows from uniqueness of (26) and (27) that

$$P_{\mathcal{K}^\perp}x = x - P_{\mathcal{K}}x \quad \text{and} \quad P_{\mathcal{K}}x = x - P_{\mathcal{K}^\perp}x$$

In particular, we have

$$\begin{aligned} d(x, \mathcal{K}) &= \|x - P_{\mathcal{K}}x\| \\ &= \|P_{\mathcal{K}^\perp}x\|. \end{aligned}$$

□

8.4 span(E) Properties

Proposition 8.4. *Let $\mathcal{V}$ be an inner-product space and let $E \subseteq V$. Define*

$$Span(E) := \left\{ \sum_{i=1}^{n} \lambda_i v_i \mid n \in \mathbb{N},\, \lambda_i \in \mathbb{C}, \text{ and } v_i \in E \text{ for } 1 \le i \le n \right\}$$

Then

1. *$Span(E)$ is a subspace of $\mathcal{V}$.*
2. *$Span(E)$ is the smallest subspace containing E.*

Proof.

1. Let $\lambda \in \mathbb{C}$ and let $v, w \in \text{Span}(E)$ where

   $$v = \sum_{i=1}^{m} \lambda_i v_i \text{ and } w = \sum_{j=1}^{n} \mu_j w_j$$

   where $\lambda_i, \mu_j \in \mathbb{C}$ and $v_i, w_j \in E$ for all $1 \le i \le m$ and $1 \le j \le n$. Then

   $$\begin{aligned} \lambda v + w &= \sum_{i=1}^{m} \lambda\lambda_i v_i + \sum_{j=1}^{n} \lambda\mu_j w_j \\ &= \sum_{i=1}^{m} \kappa_k u_k \in \text{Span}(E), \end{aligned}$$

   where $\kappa_k = \lambda\lambda_k$ and $u_k = v_k$ for $1 \le k \le m$ and $\kappa_k = \lambda\mu_{k-m}$ and $u_k = w_{k-m}$ for $m < k \le m + n$. Therefore Span(E) is a subspace of $\mathcal{V}$.

2. Let $\mathcal{U}$ be any subspace of $\mathcal{V}$ which contains E. Suppose that $v \in \text{Span}(E)$, where

   $$v = \sum_{i=1}^{n} \lambda_i v_i \tag{28}$$

   where $\lambda_i \in \mathbb{C}$ and $v_i \in E$ for all $1 \le i \le n$. As $\mathcal{U}$ is a subspace of $\mathcal{V}$, it must be closed under taking finite linear combinations of elements in $\mathcal{U}$. Since for each $1 \le i \le n$, we have $\lambda_i \in \mathbb{C}$ and $v_i \in E \subseteq \mathcal{U}$, it is clear that from (28) that $v \in \mathcal{U}$. Thus Span$(E) \subseteq \mathcal{U}$.

□

[4] Theorem (8.3) in the Appendix.
[5] This was also shown in class and is given in Theorem (8.3) in the Appendix

8.5 $\overline{\text{Span}}(E)$ Properties

Proposition 8.5. *Let $\mathcal{V}$ be an inner-product space and let $E \subseteq V$. Define the closed span of E, denoted $\overline{Span}(E)$, as the closure of $Span(E)$. Then*

1. *$\overline{Span}(E)$ is a closed subspace of $\mathcal{V}$.*

2. *$\overline{Span}(E)$ is the smallest closed subspace containing E.*

Proof.

1. By Proposition (8.4), we know that $\text{Span}(E)$ is a subspace. By a theorem[6] which we proved in class, the closure of a subspace is a closed subspace. Therefore $\overline{\text{Span}}(E)$ is a closed subspace of $\mathcal{V}$.

2. Let $\mathcal{U}$ be any closed subspace of $\mathcal{V}$ which contains E. By Proposition (8.4), we know that $\text{Span}(E) \subseteq \mathcal{U}$. Therefore

$$\begin{aligned}\overline{\text{Span}}(E) &= \overline{\text{Span}(E)} \\ &\subseteq \overline{\mathcal{U}} \\ &= \mathcal{U},\end{aligned}$$

where $\mathcal{U} = \overline{\mathcal{U}}$ since $\mathcal{U}$ is closed (this was proved in the second homework).

□

8.6 $(E^{\perp})^{\perp} = \overline{\text{Span}}(E)$

Proposition 8.6. *Let $\mathcal{H}$ be a Hilbert space and let $E \subseteq \mathcal{H}$. Then*

$$(E^{\perp})^{\perp} = \overline{Span}(E).$$

Proof. First note that $E \subseteq (E^{\perp})^{\perp}$. Indeed, if $x \in E$, then $\langle x, y \rangle = 0$ for all $y \in E^{\perp}$, hence $x \in (E^{\perp})^{\perp}$. Also, from a theorem[7] we proved in class, we know that $(E^{\perp})^{\perp}$ is a closed subspace. Thus $(E^{\perp})^{\perp}$ is a closed subspace which contains E, which implies $(E^{\perp})^{\perp} \supseteq \overline{\text{Span}}(E)$ by Proposition (8.5).

Converesely, since taking orthononal complements is inclusion-reversing[8], $E \subseteq \overline{\text{Span}}(E)$ implies $E^{\perp} \supseteq \overline{\text{Span}}(E)^{\perp}$ which implies $(E^{\perp})^{\perp} \subseteq (\overline{\text{Span}}(E)^{\perp})^{\perp} = \overline{\text{Span}}(E)$, where the last equality follows from a theorem[9] in class. □

Appendix

Theorem 8.1. *(Pythagorean Theorem) Let x and y be nonzero vectors in $\mathcal{V}$ such that $\langle x, y \rangle = 0$ (we call such vectors **orthogonal** to one another). Then*

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

Proof. We have

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

□

Theorem 8.2. *Let $\mathcal{U}$ be a subspace of $\mathcal{V}$. Then $\overline{\mathcal{U}}$ is a subspace of $\mathcal{V}$.*

Proof. Let $x, y \in \overline{\mathcal{U}}$ and $\lambda \in \mathbb{C}$. Let (x_n) and (y_n) be two sequences of elements in $\mathcal{U}$ such that $x_n \to x$ and $y_n \to y$. Then $(\lambda x_n + y_n)$ is a sequence of elements in $\mathcal{U}$ such that $\lambda x_n + y_n \to \lambda x + y$. Therefore $\lambda x + y \in \overline{\mathcal{U}}$, which implies $\overline{\mathcal{U}}$ is a subspace of $\mathcal{V}$. □

[6] Theorem (8.2) in the Appendix.
[7] Theorem (8.3) in the Appendix
[8] Theorem (8.3) in the Appendix
[9] Theorem (8.3) in the Appendix

Theorem 8.3. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{K} \subseteq \mathcal{L} \subseteq \mathcal{H}$. Then*

1. *we have $\mathcal{K}^\perp \supseteq \mathcal{L}^\perp$.*
2. *$\mathcal{K}^\perp$ is a closed subspace of $\mathcal{H}$.*
3. *If $\mathcal{K}$ is a closed subspace of $\mathcal{H}$, then every $x \in \mathcal{H}$ can be decomposed in a* unique *way as a sum of a vector in $\mathcal{K}$ and a vector in $\mathcal{K}^\perp$. In other words, we have $\mathcal{H} = \mathcal{K} \oplus \mathcal{K}^\perp$.*
4. *If $\mathcal{K}$ is a closed subspace of $\mathcal{H}$, then $(\mathcal{K}^\perp)^\perp = \mathcal{K}$.*

Proof.

1. We have
$$\begin{aligned} x \in \mathcal{L}^\perp &\implies \langle x, y\rangle = 0 \text{ for all } y \in \mathcal{L} \\ &\implies \langle x, y\rangle = 0 \text{ for all } y \in \mathcal{K} \\ &\implies x \in \mathcal{K}^\perp. \end{aligned}$$
Thus $\mathcal{K}^\perp \supseteq \mathcal{L}^\perp$.

2. First we show that $\mathcal{K}^\perp$ is a subspace of $\mathcal{V}$. Let $x, z \in \mathcal{K}^\perp$ and $\lambda \in \mathbb{C}$. Then
$$\begin{aligned} \langle x + \lambda z, y\rangle &= \langle x, y\rangle + \lambda\langle z, y\rangle \\ &= 0 \end{aligned}$$
for all $y \in \mathcal{K}$. This implies $\mathcal{K}^\perp$ is a subspace of $\mathcal{V}$. Now we will show that $\mathcal{K}^\perp$ is closed. Let (x_n) be a sequence of points in $\mathcal{K}^\perp$ such that $x_n \to x$ for some $x \in \mathcal{H}$. Then since $\langle x_n, y\rangle = 0$ for all $n \in \mathbb{N}$ and $y \in \mathcal{K}$, we have
$$\begin{aligned} \langle x, y\rangle &= \lim_{n\to\infty} \langle x_n, y\rangle \\ &= \lim_{n\to\infty} 0 \\ &= 0. \end{aligned}$$
for all $y \in \mathcal{K}$. Therefore $x \in \mathcal{K}^\perp$, which implies $\mathcal{K}^\perp$ is closed.

3. Let $x \in \mathcal{H}$. Then $x = \mathrm{P}_\mathcal{K} x + x - \mathrm{P}_\mathcal{K} x$ where $\mathrm{P}_\mathcal{K} x \in \mathcal{K}$ and $x - \mathrm{P}_\mathcal{K} x \in \mathcal{K}^\perp$. This establishes existence. For uniqueness, first note that $\mathcal{K} \cap \mathcal{K}^\perp = \{0\}$. Indeed, if $y \in \mathcal{K} \cap \mathcal{K}^\perp$, then we must have $\langle y, y\rangle = 0$, which implies $y = 0$. Now suppose that $x = y + z$ is another decomposition of x where $y \in \mathcal{K}$ and $z \in \mathcal{K}^\perp$. Then we have
$$(\mathrm{P}_\mathcal{K} x) + (x - \mathrm{P}_\mathcal{K} x) = x = y + z$$
implies $\mathrm{P}_\mathcal{K} x - y = (x - \mathrm{P}_\mathcal{K} x) - z$ which implies $\mathrm{P}_\mathcal{K} x - y \in \mathcal{K} \cap \mathcal{K}^\perp = \{0\}$ and $(x - \mathrm{P}_\mathcal{K} x) - z \in \mathcal{K} \cap \mathcal{K}^\perp = \{0\}$.

4. Let $x \in \mathcal{K}$. Then $\langle x, y\rangle = 0$ for all $y \in \mathcal{K}^\perp$. Thus $x \in (\mathcal{K}^\perp)^\perp$, and so $\mathcal{K} \subseteq (\mathcal{K}^\perp)^\perp$. Conversely, let $x \in (\mathcal{K}^\perp)^\perp$. Then $\langle x, y\rangle = 0$ for all $y \in \mathcal{K}^\perp$. In particular, we have
$$\begin{aligned} \|x - \mathrm{P}_\mathcal{K} x\|^2 &= \langle x - \mathrm{P}_\mathcal{K} x, x - \mathrm{P}_\mathcal{K} x\rangle \\ &= \langle x, x - \mathrm{P}_\mathcal{K} x\rangle - \langle \mathrm{P}_\mathcal{K} x, x - \mathrm{P}_\mathcal{K} x\rangle \\ &= 0 - 0 \\ &= 0, \end{aligned}$$
which implies $x = \mathrm{P}_\mathcal{K} x$. This implies $x \in \mathcal{K}$, and hence $(\mathcal{K}^\perp)^\perp \subseteq \mathcal{K}$.

□

9 Homework 4

9.1 Equivalent Definitions of Norm of Operator

Proposition 9.1. *Let $\mathcal{H}$ be a Hilbert space and let $T \colon \mathcal{H} \to \mathcal{H}$ be a bounded operator. Then*

1. *$\|T\| = \sup\{\|Tx\| \mid \|x\| = 1\}$;*
2. *$\|T\| = \sup\left\{\frac{\|Tx\|}{\|x\|} \mid x \in \mathcal{H} \backslash \{0\}\right\}$.*

Proof.

1. First note that

$$\begin{aligned}\sup\{\|Tx\| \mid \|x\| = 1\} &\leq \sup\{\|Tx\| \mid \|x\| \leq 1\} \\ &= \|T\|.\end{aligned}$$

We prove the reverse inequality by contradiction. Assume that $\|T\| > \sup\{\|Tx\| \mid \|x\| = 1\}$. Choose $\varepsilon > 0$ such that

$$\|T\| - \varepsilon > \sup\{\|Tx\| \mid \|x\| = 1\} \tag{29}$$

Next, choose $x \in \mathcal{H}$ such that $\|x\| \leq 1$ and $\|Tx\| \geq \|T\| - \varepsilon$. Then since $\|x\| \leq 1$ and $\left\|\frac{x}{\|x\|}\right\| = 1$, we have

$$\begin{aligned}\|T\| &\geq \left\|T\left(\frac{x}{\|x\|}\right)\right\| \\ &= \frac{\|Tx\|}{\|x\|} \\ &\geq \|Tx\| \\ &> \|T\| - \varepsilon,\end{aligned}$$

and this contradicts (29).

2. We have

$$\begin{aligned}\sup\left\{\frac{\|Tx\|}{\|x\|} \mid x \in \mathcal{H}\backslash\{0\}\right\} &= \sup\left\{\left\|T\left(\frac{x}{\|x\|}\right)\right\| \mid x \in \mathcal{H}\backslash\{0\}\right\} \\ &= \sup\{\|Ty\| \mid \|y\| = 1\} \\ &= \|T\|,\end{aligned}$$

where the last equality follows from 1. □

9.2 Multiplication by $k \in C[a,b]$ is Bounded Operator

Proposition 9.2. *Let $k \in C[a,b]$. Then the operator $T\colon C[a,b] \to C[a,b]$ defined by*

$$Tf = kf$$

for all $f \in C[a,b]$ is bounded. It's norm will be explicitly computed in the proof below.

Proof. We first show it is linear. Let $f, g \in C[a,b]$ and let $\lambda, \mu \in \mathbb{C}$. Then we have

$$\begin{aligned}T(\lambda f + \mu g) &= k(\lambda f + \mu g) \\ &= \lambda kf + \mu kg \\ &= \lambda T(f) + \mu T(g).\end{aligned}$$

Thus, T is linear.

Next we show it is bounded. If $k = 0$, then $\|T\| = 0$, so assume $k \neq 0$. Since k is continuous on the compact interval $[a,b]$, there exists $c \in [a,b]$ such that $|k(x)| \leq |k(c)|$ for all $x \in [a,b]$. Choose such a $c \in [a,b]$ and let $f \in C[a,b]$ such that $\|f\| \leq 1$. Then

$$\begin{aligned}\|Tf\| &= \|kf\| \\ &= \sqrt{\int_a^b |k(x)|^2 |f(x)|^2 dx} \\ &\leq |k(c)| \sqrt{\int_a^b |f(x)|^2 dx} \\ &\leq |k(c)|.\end{aligned}$$

implies $\|T\| \leq |k(c)|$, and hence T is bounded.

To find the norm of T, let $\varepsilon > 0$ such that $\varepsilon < |k(c)|$. Without loss of generality, assume that $c < b$ (if $c = b$, then we swap the role of b with a in the argument which follows). Choose $c' \in (c, b)$ such that $|k(x)| \geq |k(c)| - \varepsilon$

for all $x \in (c, c')$ (such a c' must exist since k is continuous) and choose f to be a nonzero continuous function in $C[a,b]$ which vanishes outside the interval (c, c'). Then

$$|k(x)||f(x)| \geq (|k(c)| - \varepsilon)|f(x)|$$

for all $x \in (a,b)$. In particular, this implies

$$\begin{aligned}
\|Tf\| &= \|kf\| \\
&= \sqrt{\int_a^b |k(x)f(x)|^2 dx} \\
&\geq \sqrt{\int_a^b (|k(c)| - \varepsilon)|f(x)|^2 dx} \\
&= (|k(c)| - \varepsilon)\sqrt{\int_a^b |f(x)|^2 dx} \\
&= (|k(c)| - \varepsilon)\|f\|.
\end{aligned}$$

Therefore $\|T(f/\|f\|)\| \geq |k(c)| - \varepsilon$, and this implies

$$\|T\| \geq |k(c)| - \varepsilon \tag{30}$$

Since (30) holds for all $\varepsilon > 0$, we must have $\|T\| \geq |k(c)|$. Thus $\|T\| = |k(c)|$. □

9.3 Gram_Schmidt Properties

Proposition 9.3. *Let $\{x_n \mid n \in \mathbb{N}\}$ be a linearly independent set of vectors in a Hilbert space $\mathcal{H}$. Consider the so called Gram-Schmidt process: set $e_1 = \frac{1}{\|x_1\|}x_1$. Proceed inductively. If $e_1, e_2, \ldots, e_{n-1}$ are computed, compute e_n in two steps by*

$$f_n := x_n - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k, \text{ and then set } e_n := \frac{1}{\|f_n\|} f_n.$$

Then

1. *for every $N \in \mathbb{N}$ we have span$\{x_1, x_2, \ldots, x_N\}$ = span$\{e_1, e_2, \ldots, e_N\}$;*
2. *the set $\{e_n \mid n \in \mathbb{N}\}$ is an orthonormal set in $\mathcal{H}$;*
3. *if $\overline{\text{span}}\{x_n \mid n \in \mathbb{N}\} = \mathcal{H}$, then $\{e_n \mid n \in \mathbb{N}\}$ is an orthonormal basis for $\mathcal{H}$.*

Proof.
1. Let $N \in \mathbb{N}$. Then for each $1 \leq n \leq N$, we have

$$x_n = \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k + \|f_n\| e_n.$$

This implies $\text{span}\{x_1, x_2, \ldots, x_N\} \subseteq \text{span}\{e_1, e_2, \ldots, e_N\}$. We show the reverse inclusion by induction on n such that $1 \leq n \leq N$. The base case $n = 1$ being $\text{span}\{x_1\} \supseteq \text{span}\{e_1\}$, which holds since $e_1 = \frac{1}{\|x_1\|}x_1$. Now suppose for some n such that $1 \leq n < N$ we have

$$\text{span}\{x_1, x_2, \ldots, x_k\} \supseteq \text{span}\{e_1, e_2, \ldots, e_k\} \tag{31}$$

for all $1 \leq k \leq n$. Then

$$e_{n+1} = \frac{1}{\|f_n\|} x_n - \sum_{k=1}^{n} \frac{1}{\|f_n\|} \langle x_n, e_k \rangle e_k \in \text{span}\{x_1, x_2, \ldots, x_n\}.$$

where we used the induction step (31) on the e_k's ($1 \leq k \leq n$). Therefore

$$\text{span}\{x_1, x_2, \ldots, x_k\} \supseteq \text{span}\{e_1, e_2, \ldots, e_k\}$$

for all $1 \leq k \leq n+1$, and this proves our claim.

2. By construction, we have $\langle e_n, e_n \rangle = 1$ for all $n \in \mathbb{N}$. Thus, it remains to show that $\langle e_m, e_n \rangle = 0$ whenever $m \neq n$. We prove by induction on $n \geq 2$ that $\langle e_n, e_m \rangle = 0$ for all $m < n$. Proving this also give us $\langle e_m, e_n \rangle = 0$ for all $m < n$, since

$$\begin{aligned} \langle e_m, e_n \rangle &= \overline{\langle e_n, e_m \rangle} \\ &= \overline{0} \\ &= 0. \end{aligned}$$

The base case is

$$\begin{aligned} \langle e_2, e_1 \rangle &= \frac{1}{\|x_1\| \|f_2\|} \left\langle \left(x_2 - \frac{\langle x_2, x_1 \rangle}{\langle x_1, x_1 \rangle} x_1 \right), x_1 \right\rangle \\ &= \frac{1}{\|x_1\| \|f_2\|} \left(\langle x_2, x_1 \rangle - \langle x_2, x_1 \rangle \right) \\ &= 0 \end{aligned}$$

Now suppose that $n > 2$ and that $\langle e_n, e_m \rangle = 0$ for all $m < n$. Then

$$\begin{aligned} \langle e_{n+1}, e_m \rangle &= \frac{1}{\|f_{n+1}\|} \langle x_{n+1} - \sum_{k=1}^{n} \langle x_{n+1}, e_k \rangle e_k, e_m \rangle \\ &= \frac{1}{\|f_{n+1}\|} \left(\langle x_{n+1}, e_m \rangle - \sum_{k=1}^{n} \langle x_{n+1}, e_k \rangle \langle e_k, e_m \rangle \right) \\ &= \frac{1}{\|f_{n+1}\|} \left(\langle x_{n+1}, e_m \rangle - \langle x_{n+1}, e_m \rangle \langle e_m, e_m \rangle \right) \\ &= \frac{1}{\|f_{n+1}\|} \left(\langle x_{n+1}, e_m \rangle - \langle x_{n+1}, e_m \rangle \right) \\ &= 0, \end{aligned}$$

for all $m < n + 1$, where we used the induction hypothesis to get from the second line to the third line. This proves the induction step, which finishes the proof of part 2 of the proposition.

3. By 2, we know that $\{e_n \mid n \in \mathbb{N}\}$ is an orthonormal set. Thus, it suffices to show that $\{e_n \mid n \in \mathbb{N}\}$ is complete. To do this, we use the criterion that the set $\{e_n \mid n \in \mathbb{N}\}$ is complete if and only if the only $x \in \mathcal{H}$ such that $\langle x, e_n \rangle = 0$ for all $n \in \mathbb{N}$ is $x = 0$.

Let $x \in \mathcal{H}$ and suppose $\langle x, e_n \rangle = 0$ for all $n \in \mathbb{N}$. Then

$$\begin{aligned} \langle x, x_n \rangle &= \left\langle x, \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k + \|f_n\| e_n \right\rangle \\ &= \sum_{k=1}^{n-1} \langle x_n, e_k \rangle \langle x, e_k \rangle + \|f_n\| \langle x, e_n \rangle \\ &= 0 \end{aligned}$$

for all $n \in \mathbb{N}$. Since $\{x_n \mid n \in \mathbb{N}\}$ is complete, this implies $x = 0$. Therefore $\{e_n \mid n \in \mathbb{N}\}$ is complete. □

9.4 Gram-Schmidt Example Worked out on Legendre Polynomials

Example 9.1. The first three Legendre polynomials are

$$P_1(x) = 1, \quad P_2(x) = x, \quad , P_3(x) = \frac{1}{2}(3x^2 - 1).$$

We apply Gram-Schmidt process to the polynomials $1, x, x^2$ in the space $C[-1, 1]$ to get scalar multiples of the Legendre polynomials above. First we set $f_1(x) = 1$ and then calculate

$$\begin{aligned} \|f_1(x)\| &= \sqrt{\int_{-1}^{1} dx} \\ &= \sqrt{2}. \end{aligned}$$

Thus we set $e_1(x) = 1/\sqrt{2}$. Next we calculate

$$\begin{aligned} f_1(x) &= x - \left\langle x, \frac{1}{\sqrt{2}} \right\rangle \frac{1}{\sqrt{2}} \\ &= x - \frac{1}{2}\int_{-1}^{1} x dx \\ &= x. \end{aligned}$$

Next we calculate

$$\begin{aligned} \|f_1(x)\| &= \sqrt{\int_{-1}^{1} x^2 dx} \\ &= \sqrt{\frac{2}{3}}. \end{aligned}$$

Thus we set $e_2(x) = \sqrt{3/2}x$. Next we calculate

$$\begin{aligned} f_2(x) &= x^2 - \left\langle x^2, \sqrt{\frac{3}{2}}x \right\rangle \sqrt{\frac{3}{2}}x - \left\langle x^2, \sqrt{\frac{1}{2}} \right\rangle \sqrt{\frac{1}{2}} \\ &= x^2 - \frac{3}{2}x\int_{-1}^{1} x^3 dx - \frac{1}{2}\int_{-1}^{1} x^2 dx \\ &= x^2 - \frac{1}{3}. \end{aligned}$$

Then we finally calculate

$$\begin{aligned} \|f_2(x)\| &= \sqrt{\int_{-1}^{1} \left(x^2 - \frac{1}{3}\right)^2 dx} \\ &= \sqrt{\int_{-1}^{1} \left(x^4 - \frac{2}{3}x^2 + \frac{1}{9}\right) dx} \\ &= \sqrt{\int_{-1}^{1} x^4 dx - \frac{2}{3}\int_{-1}^{1} x^2 dx + \frac{1}{9}\int_{-1}^{1} dx} \\ &= \sqrt{\frac{2}{5} - \frac{4}{9} + \frac{2}{9}} \\ &= \sqrt{\frac{8}{45}}. \end{aligned}$$

Thus we set $e_3(x) = \sqrt{45/8}(x^2 - 1/3)$. Now observe that

$$\begin{aligned} P_1(x) &= \sqrt{2}e_1(x) \\ P_2(x) &= \sqrt{\frac{2}{3}}e_2(x) \\ P_3(x) &= \sqrt{\frac{2}{5}}e_3(x) \end{aligned}$$

9.5 Minimizing Integral Example

For this problem, we needed to establish some basic results which we proved in the Appendix.

Proposition 9.4. *The expression*

$$\int_{-1}^{1} |x^3 - a - bx - cx^2|^2 \mathrm{d}x. \tag{32}$$

is minimized in $a, b, c \in \mathbb{C}$ if and only if $a = 0$, $b = 3/5$, and $c = 0$.

Proof. Let

$$\mathcal{H} = \{p(x) \in \mathbb{C}[x] \mid \deg(p(x)) \leq 3\} \quad \text{and} \quad \mathcal{K} = \{p(x) \in \mathbb{C}[x] \mid \deg(p(x)) \leq 2\}.$$

Then $\mathcal{H}$ and $\mathcal{K}$ are subspaces of $C[-1, 1]$, Proposition (9.7) implies they are inner-product spaces with the inner-product inherited from $C[-1, 1]$. Since $\mathcal{H}$ is finite dimensional, Proposition (9.8) implies $\mathcal{H}$ is a separable Hilbert space. Since $\mathcal{K}$ is a finite dimensional subspace of $\mathcal{H}$, Proposition (9.9) implies $\mathcal{K}$ is closed in $\mathcal{H}$. Let $\{e_1, e_2, e_3\}$ be the orthonormal basis computed in problem 4. A proposition proved in class implies

$$\begin{aligned} P_{\mathcal{K}}(x^3) &= \langle x^3, e_1\rangle e_1 + \langle x^3, e_2\rangle e_2 + \langle x^3, e_3\rangle e_3 \\ &= \frac{1}{2}\int_{-1}^{1} x^3 dx + \frac{3}{2}x\int_{-1}^{1} x^4 dx + \frac{45}{8}\left(x^2 - \frac{1}{3}\right)\int_{-1}^{1} x^3\left(x^2 - \frac{1}{3}\right) \mathrm{d}x \\ &= \frac{3}{5}x. \end{aligned}$$

where we used the fact that $x^3(x^2 - 1/3)$ is an odd function to get $\int_{-1}^{1} x^3(x^2 - 1/3)dx = 0$. Therefore

$$\begin{aligned} \int_{-1}^{1}\left|x^3 - \frac{3}{5}x\right|^2 \mathrm{d}x &= \|x^3 - P_{\mathcal{K}}(x^3)\|^2 \\ &= \inf\left\{\|x^3 - (a + bx + cx^2)\|^2 \mid a + bx + cx^2 \in \mathcal{K}\right\} \\ &= \inf\left\{\int_{-1}^{1} |x^3 - a - bx - cx^2|^2 dx \mid a, b, c \in \mathbb{C}\right\}. \end{aligned}$$

By uniqueness of $P_{\mathcal{K}}x^3$, (32) is minimized in $a, b, c \in \mathbb{C}$ if and only if $a = 0$, $b = 3/5$, and $c = 0$. $\square$

9.6 $\ell^2(\mathbb{N})$ is a Hilbert Space

Proposition 9.5. *$\ell^2(\mathbb{N})$ is a Hilbert space.*

Proof. Let $(a^n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\ell^2(\mathbb{N})$.

Step 1: We show that for each $k \in \mathbb{N}$, the sequence of kth coordinates $(a_k^n)_{n\in\mathbb{N}}$ is a Cauchy sequence of complex numbers, and hence must converge (as $\mathbb{C}$ is complete). Let $k \in \mathbb{N}$ and let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $m, n \geq N$ implies $\|a^n - a^m\| < \varepsilon^2$. Then $n, m \geq N$ implies

$$\begin{aligned} |a_k^n - a_k^m|^2 &\leq \sum_{i=1}^{\infty} |a_i^n - a_i^m|^2 \\ &= \|a^n - a^m\| \\ &< \varepsilon^2, \end{aligned}$$

which implies $|a_k^n - a_k^m| < \varepsilon$. Therefore $(a_k^n)_{n\in\mathbb{N}}$ is a Cauchy sequence of complex numbers. In particular, the sequence $(a_k^n)_{n\in\mathbb{N}}$ converges to some element, say $a_k^n \to a_k$.

Step 2: We show that the sequence $(a_k)_{k\in\mathbb{N}}$ defined in step 1 is square summable. Since (a^n) is a Cauchy sequence of elements in $\ell^2(\mathbb{N})$, there exists an $M > 0$ such that $\|a^n\| < M$ for all $n \in \mathbb{N}$ (see Lemma ((16.1) for a proof of this). Choose such an $M > 0$. Let $\varepsilon > 0$ and let $K \in \mathbb{N}$. Choose $N \in \mathbb{N}$ such that

$$|a_k|^2 < |a_k^N|^2 + \varepsilon/K$$

for all $1 \leq k \leq K$. Then

$$\begin{aligned} \sum_{k=1}^{K} |a_k|^2 &< \sum_{k=1}^{K} |a_k^N|^2 + \varepsilon \\ &\leq \|a^N\| + \varepsilon \\ &\leq M + \varepsilon. \end{aligned}$$

Taking the limit $K \to \infty$, we see that

$$\begin{aligned} \|a\| &= \sum_{k=1}^{\infty} |a_k|^2 \\ &\leq M + \varepsilon \\ &\leq 0. \end{aligned}$$

In particular, a is square summable.

Step 3: Let a be the sequence $(a_k)_{k\in\mathbb{N}}$ defined in step 1. We show that $a^n \to a$ in the ℓ^2 norm. Let $\varepsilon > 0$ and let $K \in \mathbb{N}$. Choose $N \in \mathbb{N}$ such that $m, n \geq N$ implies $\|a^n - a^m\|^2 < \varepsilon/2$. Then

$$\begin{aligned}\sum_{k=1}^{K} |a_k^n - a_k^m|^2 &\leq \sum_{k=1}^{\infty} |a_k^n - a_k^m|^2 \\ &= \|a^n - a^m\|^2 \\ &< \varepsilon/2\end{aligned}$$

for all $n, m \geq N$. Since $a_k^m \to a_k$ as $m \to \infty$ implies

$$\sum_{k=1}^{K} |a_k^n - a_k^m|^2 \to \sum_{k=1}^{K} |a_k^n - a_k|^2$$

as $m \to \infty$, we see that after taking the limit $m \to \infty$, we have

$$\sum_{k=1}^{K} |a_k^n - a_k|^2 \leq \varepsilon/2. \tag{33}$$

for all $n \geq N$. Taking the limit $K \to \infty$ in (33) gives us

$$\|a^n - a\|^2 < \varepsilon$$

for all $n \geq N$. It follows that $a^n \to a$. □

9.7 $C[a,b]$ is not a Hilbert Space

Proposition 9.6. *$C[a,b]$ is not a Hilbert space.*

Proof. For each $n \in \mathbb{N}$, define $f_n \in C[a,b]$ by

$$f_n(x) = \begin{cases} 0 & x \in [a, c - \frac{1}{n}] \\ nx + 1 - nc & x \in [c - \frac{1}{n}, c] \\ 1 & x \in [c, b], \end{cases}$$

where $c = \frac{a+b}{2}$. We will show that the sequence (f_n) is a Cauchy sequence which is not convergent.

Step 1: We first show that the sequence (f_n) is a Cauchy sequence. Let $\varepsilon > 0$ and let $m, n \in \mathbb{N}$ such that $n \geq m$. Then

$$\begin{aligned}\|f_n - f_m\|^2 &= \int_{c-\frac{1}{m}}^{c-\frac{1}{n}} |mx + 1 - mc|^2\, dx + \int_{c-\frac{1}{n}}^{c} |nx + 1 - nc - (mx + 1 - mc)|^2\, dx \\ &= \int_{c-\frac{1}{m}}^{c-\frac{1}{n}} |m(x-c) + 1|^2\, dx + (n-m)^2 \int_{c-\frac{1}{n}}^{c} |x - c|^2\, dx \\ &\leq \left(\frac{1}{m} - \frac{1}{n}\right) \left|1 - \frac{m}{n}\right|^2 + \frac{(n-m)^2}{n^3} \\ &\leq \frac{1}{m} - \frac{1}{n} + \frac{(n-m)^2}{n^3}.\end{aligned}$$

Choose $N \in \mathbb{N}$ such that $n \geq m \geq N$ implies

$$\frac{1}{m} - \frac{1}{n} + \frac{(n-m)^2}{n^3} < \varepsilon^2.$$

Then $n \geq m \geq N$ implies $\|f_n - f_m\| < \varepsilon$. Therefore (f_n) is a Cauchy sequence.

Step 2: We show that the sequence (f_n) is not convergent. Assume for a contradiction that $f_n \to f$ where $f \in C[a,b]$. Then

$$\begin{aligned}\|f_n - f\|^2 &= \int_{a}^{c-\frac{1}{n}} |f(x)|^2 dx + \int_{c-\frac{1}{n}}^{c} |f(x) - f_n(x)|^2 dx + \int_{c}^{b} |f(x) - 1|^2 dx \\ &\leq (c - a - \frac{1}{n}) \sup_{x\in[a,c-\frac{1}{n}]} |f(x)|^2 + \int_{c-\frac{1}{n}}^{c} |f(x) - f_n(x)|^2 dx + (b-c) \sup_{x\in[c,c-\frac{1}{n}]} |f(x) - 1|^2 dx.\end{aligned}$$

Since $\|f_n - f\| \to 0$ as $n \to \infty$, we see that (after taking the limit $n \to \infty$) we must have

$$f(x) = \begin{cases} 0 & \text{if } x \in [a, c] \\ 1 & \text{if } x \in [c, b] \end{cases}$$

but this is not a continuous function. Thus we obtain a contradiction. □

Appendix

Proposition 9.7. *Let $(V, \langle \cdot, \cdot \rangle)$ be an inner-product space and let W be a subspace of V. Then $(W, \langle \cdot, \cdot \rangle|_{W \times W} \rangle$ is an inner-product space, where $\langle \cdot, \cdot \rangle|_{W \times W} \colon W \times W \to \mathbb{C}$ is the restriction of $\langle \cdot, \cdot \rangle$ to $W \times W$.*

Proof. All of the required properties for $\langle \cdot, \cdot \rangle|_{W \times W}$ to be an inner-product are *inherited* by $\langle \cdot, \cdot \rangle$ since W is a subset of V. For instance, let $x, y, z \in V$ and let $\lambda \in \mathbb{C}$. Then

$$\begin{aligned} \langle x + \lambda y, z \rangle|_{W \times W} &= \langle x + \lambda y, z \rangle \\ &= \langle x, z \rangle + \lambda \langle y, z \rangle \\ &= \langle x, z \rangle|_{W \times W} + \lambda \langle y, z \rangle|_{W \times W} \end{aligned}$$

gives us linearity in the first argument. The other properties follow similarly. □

Remark. As long as context is clear, then we denote $\langle \cdot, \cdot \rangle|_{W \times W}$ simply by $\langle \cdot, \cdot \rangle$.

Proposition 9.8. *Let $(V, \langle \cdot, \cdot \rangle)$ be an n-dimensional inner-product space. Then $(V, \langle \cdot, \cdot \rangle)$ is unitarily equivalent to $(\mathbb{C}^n, \langle \cdot, \cdot \rangle_e)$, where $\langle \cdot, \cdot \rangle_e$ is the standard Euclidean inner-product on $\mathbb{C}^n$. In particular, $(V, \langle \cdot, \cdot \rangle)$ is a separable Hilbert space.*

Proof. Let $\{v_1, \ldots, v_n\}$ be a basis for V. By applying the Gram-Schmidt process to $\{v_1, \ldots, v_n\}$, we can get an orthonormal basis, say $\{u_1, \ldots, u_n\}$, of V. Let $\varphi \colon V \to \mathbb{C}^n$ be the unique linear isomorphism such that

$$\varphi(u_i) = e_i$$

where e_i is the standard ith coordinate vector in $\mathbb{C}^n$ for all $1 \leq i \leq n$. Then φ is a unitary equivalence. Indeed, it is an isomorphism since it restricts to a bijection on basis sets. Moreover we have

$$\langle u_i, u_j \rangle = \langle \varphi(u_i), \varphi(u_j) \rangle_{\mathrm{e}} = \langle e_i, e_j \rangle_{\mathrm{e}}$$

for all $1 \leq i, j \leq n$. This implies

$$\langle x, y \rangle = \langle \varphi(x), \varphi(y) \rangle_{\mathrm{e}}$$

for all $x, y \in V$. □

Proposition 9.9. *Let $\mathcal{V}$ be an inner-product space over $\mathbb{C}$ and let $\mathcal{W}$ be a finite dimensional subspace of $\mathcal{V}$. Then $\mathcal{W}$ is a closed.*

Proof. Let $\{w_1, \ldots, w_k\}$ be an orthonormal basis for $\mathcal{W}$ and let (x_n) be a sequence of vectors in $\mathcal{W}$ such that $x_n \to x$ where $x \in \mathcal{V}$. For each $n \in \mathbb{N}$, express x_n in terms of the basis $\{w_1, \ldots, w_k\}$ say as

$$x_n = \lambda_{1n} w_1 + \cdots + \lambda_{kn} w_k,$$

where $\lambda_{1n}, \ldots, \lambda_{kn} \in \mathbb{C}$. Since $x_n \to x$ as $n \to \infty$, the sequence (x_n) is a Cauchy sequence. This implies the sequence $(\lambda_{jn})_{n \in \mathbb{N}}$ of complex numbers is a Cauchy sequence, for each $1 \leq j \leq k$. Indeed, letting $\varepsilon > 0$, choose $N \in \mathbb{N}$ such that $n, m \geq N$ implies $\|x_n - x_m\| < \varepsilon$. Then $n, m \geq N$ implies

$$\begin{aligned} |\lambda_{jn} - \lambda_{jm}| &\leq |\lambda_{1n} - \lambda_{1m}| + \cdots + |\lambda_{kn} - \lambda_{km}| \\ &= \|(\lambda_{1n} - \lambda_{1m}) w_1 + \cdots + (\lambda_{kn} - \lambda_{km}) w_k\| \\ &= \|x_n - x_m\| \\ &< \varepsilon \end{aligned}$$

for each $1 \leq j \leq k$. Now since $\mathbb{C}$ is complete, we must have $\lambda_{jn} \to \lambda_j$ as $n \to \infty$ for some $\lambda_j \in \mathbb{C}$ for all $1 \leq j \leq n$. In particular, we have

$$\begin{aligned} x &= \lim_{n \to \infty} x_n \\ &= \lim_{n \to \infty} (\lambda_{1n} w_1 + \cdots + \lambda_{kn} w_k) \\ &= \lim_{n \to \infty} (\lambda_{1n} w_1) + \cdots + \lim_{n \to \infty} (\lambda_{kn} w_k) \\ &= \lambda_1 w_1 + \cdots + \lambda_k w_k, \end{aligned}$$

and this implies $x \in \mathcal{W}$, which implies $\mathcal{W}$ is closed. □

Lemma 9.1. *Let (x_n) be a Cauchy sequence in $\mathcal{V}$. Then (x_n) is bounded.*

Proof. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $m, n \geq N$ implies $\|x_n - x_m\| < \varepsilon$. Thus, fixing $m \in \mathbb{N}$, we see that $n \geq N$ implies

$$\|x_n\| < \|x_m\| + \varepsilon.$$

Now we let

$$M = \max\{\|x_1\|, \|x_2\|, \dots, \|x_m\| + \varepsilon\}.$$

Then M is a bound for (x_n). □

Proposition 9.10. *Let (x_n) and (y_n) be Cauchy sequences of vectors in $\mathcal{V}$. Then $(\langle x_n, y_n \rangle)$ is a Cauchy sequence of complex numbers.*

Proof. Let $\varepsilon > 0$. Choose M_x and M_y such that $\|x_n\| < M_x$ and $\|y_n\| < M_y$ for all $n \in \mathbb{N}$. We can do this by Lemma (16.1). Next, choose $N \in \mathbb{N}$ such that $n, m \geq N$ implies $\|x_n - x_m\| < \frac{\varepsilon}{2M_y}$ and $\|y_n - y_m\| < \frac{\varepsilon}{2M_x}$. Then $n, m \geq M$ implies

$$\begin{aligned}
|\langle x_n, y_n \rangle - \langle x_m, y_m \rangle| &= |\langle x_n, y_n \rangle - \langle x_m, y_n \rangle + \langle x_m, y_n \rangle - \langle x_m, y_m \rangle| \\
&= |\langle x_n - x_m, y_n \rangle + \langle x_m, y_n - y_m \rangle| \\
&\leq |\langle x_n - x_m, y_n \rangle| + |\langle x_m, y_n - y_m \rangle| \\
&\leq \|x_n - x_m\| \|y_n\| + \|x_m\| \|y_n - y_m\| \\
&\leq \|x_n - x_m\| M_y + M_x \|y_n - y_m\| \\
&< \varepsilon.
\end{aligned}$$

This implies $(\langle x_n, y_n \rangle)$ is a Cauchy sequence of complex numbers in $\mathbb{C}$. The latter statement in the proposition follows from the fact that $\mathbb{C}$ is complete. □

Homework 2, Problem 5

Proposition 9.11. *Let $\mathcal{V}$ be an inner-product space, let $\mathcal{A}$ be a subspace of $\mathcal{V}$, let $x \in \mathcal{V}$, and let $\lambda \in \mathbb{C}$. Then*

$$d(\lambda x, \mathcal{A}) = |\lambda| d(x, \mathcal{A}).$$

Proof. Choose a sequence (y_n) of elements in $\mathcal{A}$ such that

$$\|x - y_n\| < d(x, \mathcal{A}) + \frac{1}{|\lambda n|}$$

for all $n \in \mathbb{N}$. Then since $\mathcal{A}$ is a subspace, we have

$$\begin{aligned}
d(\lambda x, \mathcal{A}) &\leq \|\lambda x - \lambda y_n\| \\
&= |\lambda| \|x - y_n\| \\
&< |\lambda| \left(d(x, \mathcal{A}) + \frac{1}{|\lambda n|} \right) \\
&= |\lambda| d(x, \mathcal{A}) + \frac{1}{n}
\end{aligned}$$

for all $n \in \mathbb{N}$. In particular, this implies $d(\lambda x, \mathcal{A}) \leq |\lambda| d(x, \mathcal{A})$.

Conversely, choose a sequence (z_n) of elements in $\mathcal{A}$ such that

$$\|\lambda x - z_n\| < d(\lambda x, \mathcal{A}) + \frac{1}{n}$$

Then since $\mathcal{A}$ is a subspace, we have

$$\begin{aligned}
|\lambda| d(x, \mathcal{A}) &\leq |\lambda| \|x - z_n / |\lambda|\| \\
&= \|\lambda x - z_n\| \\
&< d(\lambda x, \mathcal{A}) + \frac{1}{n}
\end{aligned}$$

for all $n \in \mathbb{N}$. In particular, this implies $|\lambda| d(x, \mathcal{A}) \leq d(\lambda x, \mathcal{A})$. □

Proposition 9.12. *Let $\mathcal{V}$ be an inner-product space, let $\mathcal{A}$ be a subspace of $\mathcal{V}$, and let $x, y \in \mathcal{V}$. Then*

$$d(x+y, \mathcal{A}) \leq d(x, \mathcal{A}) + d(y, \mathcal{A}).$$

Proof. Choose a sequences (w_n) and (z_n) of elements in $\mathcal{A}$ such that

$$\|x - w_n\| < d(x, \mathcal{A}) + \frac{1}{2n} \quad \text{and} \quad \|y - z_n\| < d(y, \mathcal{A}) + \frac{1}{2n}$$

for all $n \in \mathbb{N}$. Then since $\mathcal{A}$ is a subspace, we have

$$\begin{aligned} d(x+y, \mathcal{A}) &\leq \|(x+y) - (w_n + z_n)\| \\ &\leq \|x - w_n\| + \|y - z_n\| \\ &< d(x, \mathcal{A}) + \frac{1}{2n} + d(y, \mathcal{A}) + \frac{1}{2n} \\ &= d(x, \mathcal{A}) + d(y, \mathcal{A}) + \frac{1}{n} \end{aligned}$$

for all $n \in \mathbb{N}$. In particular, this implies $d(x+y, \mathcal{A}) \leq d(x, \mathcal{A}) + d(y, \mathcal{A})$. □

10 Homework 5

Throughout this homework, let $\mathcal{H}$ be a Hilbert space.

10.1 Map Sending Bounded Operator to its Adjoint Operator is Conjugate-Linear

Proposition 10.1. *Let $T \colon \mathcal{H} \to \mathcal{H}$ and $S \colon \mathcal{H} \to \mathcal{H}$ be two bounded operators. Then*

$$(\alpha T + \beta S)^* = \overline{\alpha} T^* + \overline{\beta} S^* \tag{34}$$

for all $\alpha, \beta, \in \mathbb{C}$.

Proof. Let $\alpha, \beta, \in \mathbb{C}$ and let $y \in \mathcal{H}$. Then for all $x \in \mathcal{H}$, we have

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

In particular, this implies $(\alpha T + \beta S)^* y = (\overline{\alpha} T^* + \overline{\beta} S^*) y$ for all $y \in \mathcal{H}$ (by positive-definiteness of the inner-product) which implies (34). □

10.2 Composite of Bounded Operators is Bounded

Proposition 10.2. *Let $T \colon \mathcal{H} \to \mathcal{H}$ and $S \colon \mathcal{H} \to \mathcal{H}$ be two bounded operators. Then*

1. *TS is bounded and $\|TS\| \leq \|T\| \|S\|$;*
2. *$(TS)^* = S^* T^*$.*

Proof.

1. Let $x \in \mathcal{H}$ such that $\|x\| = 1$. Then

$$\begin{aligned} \|TSx\| &\leq \|T\| \|Sx\| \\ &\leq \|T\| \|S\| \|x\| \\ &= \|T\| \|S\|. \end{aligned}$$

Thus TS is bounded and $\|TS\| \leq \|T\| \|S\|$.

2. Let $y \in \mathcal{H}$. Then for all $x \in \mathcal{H}$, we have

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

In particular, this implies $(TS)^* y = S^* T^* y$ for all $y \in \mathcal{H}$, which implies $(TS)^* = S^* T^*$. □

10.3 The Adjoint Operator of $T_{u,v}$ is $T_{v,u}$

Proposition 10.3. *Let $u, v \in \mathcal{H}$ be fixed vectors.*

1. *The operator $T_{u,v} \colon \mathcal{H} \to \mathcal{H}$ defined by*
$$T_{u,v}x = \langle x, u\rangle v$$
for all $x \in \mathcal{H}$ is bounded. Moreover, we have $\|T_{u,v}\| = \|u\|\|v\|$.

2. *The adjoint of $T_{u,v}$ is given by $T_{v,u}$, that is,*
$$(T_{u,v})^* y = \langle y, v\rangle u$$
for all $y \in \mathcal{H}$.

Proof.

1. Let $x \in \mathcal{H}$. Then

$$\begin{aligned}\|T_{u,v}x\| &= \|\langle x, u\rangle v\| \\ &= |\langle x, u\rangle|\|v\| \\ &\leq \|x\|\|u\|\|v\|,\end{aligned}$$

where we used Cauchy-Schwarz to get from the second to the third line. This implies $\|T_{u,v}\| \leq \|u\|\|v\|$. We have equality at the Cauchy-Schwarz step if and only if $x = \lambda u$ for some $\lambda \in \mathbb{C}$. In particular, setting $x = u/\|u\|$ gives us $\|T_{u,v}\| = \|u\|\|v\|$.

2. Let $y \in \mathcal{H}$. Then

$$\begin{aligned}\langle x, (T_{u,v})^* y\rangle &= \langle T_{u,v}x, y\rangle \\ &= \langle \langle x, u\rangle v, y\rangle \\ &= \langle x, u\rangle\langle v, y\rangle \\ &= \langle x, \overline{\langle v, y\rangle} u\rangle \\ &= \langle x, \langle y, v\rangle u\rangle\end{aligned}$$

for all $x \in \mathcal{H}$. This implies $(T_{u,v})^* y = \langle y, v\rangle u$ for all $y \in \mathcal{H}$. □

10.4 Computing Adjoint of Operator from $\ell^2(\mathbb{N})$ to Itself

Corollary. *Let $T \colon \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be operator defined by*

$$T(x)_n = \sum_{m=1}^{\infty} \frac{x_m}{2^n 3^m},$$

for all $x = (x_m) \in \ell^2(\mathbb{N})$, where $T(x)_n$ denotes the n-th coordinate of $T(x) \in \ell^2(\mathbb{N})$. Then T is bounded with

$$\|T\| = \sqrt{\frac{1}{24}}.$$

The adjoint of T is given by

$$T^*(y)_n = \sum_{m=1}^{\infty} \frac{y_m}{2^m 3^n},$$

for all $y \in \ell^2(\mathbb{N})$.

Proof. Set $u = (1/3^m)$ and $v = (1/2^n)$. Then

$$\begin{aligned}T(x)_n &= \sum_{m=1}^{\infty} \frac{x_m}{2^n 3^m} \\ &= \langle x, u\rangle \frac{1}{2^n} \\ &= \langle x, u\rangle v_n\end{aligned}$$

for all $x \in \mathcal{H}$. Thus $Tx = \langle x, u \rangle v$ for all $x \in \mathcal{H}$. Therefore we can apply Proposition (10.3) and obtain

$$\begin{aligned}
\|T\| &= \|u\| \|v\| \\
&= \sqrt{\sum_{n=1}^{\infty} 9^{-n}} \sqrt{\sum_{n=1}^{\infty} 4^{-n}} \\
&= \sqrt{\left(\frac{1}{1 - \frac{1}{9}} - 1 \right) \left(\frac{1}{1 - \frac{1}{4}} - 1 \right)} \\
&= \sqrt{\frac{1}{24}}.
\end{aligned}$$

The adjoint of T is given by

$$\begin{aligned}
T^*(y)_n &= \langle y, v \rangle u_n \\
&= \sum_{m=1}^{\infty} \frac{y_m}{2^m 3^n}
\end{aligned}$$

for all $y \in \mathcal{H}$. □

10.5 Properties of T^*T

Proposition 10.4. *Let $T \colon \mathcal{H} \to \mathcal{H}$ be a bounded operator. Then*

1. $\|T^*T\| = \|T\|^2$;
2. $Ker(T^*T) = Ker(T)$.

Proof.

1. First note that Proposition (10.2) implies $\|T^*T\| \leq \|T^*\| \|T\| = \|T\|^2$. For the reverse inequality, let $x \in \mathcal{H}$ such that $\|x\| = 1$. Then

$$\begin{aligned}
\|Tx\|^2 &= \langle Tx, Tx \rangle \\
&= \langle x, T^*Tx \rangle \\
&\leq \|x\| \|T^*Tx\| \\
&= \|T^*Tx\|,
\end{aligned}$$

where we used Cauchy-Schwarz to get from the second line to the third line. In particular, this implies

$$\begin{aligned}
\|T\|^2 &= \sup\{\|Tx\|^2 \mid \|x\| \leq 1\} \\
&\leq \sup\{\|T^*Tx\| \mid \|x\| \leq 1\} \\
&= \|T^*T\|,
\end{aligned}$$

where the first line is justifed in the Appendix.

2. Let $x \in \mathrm{Ker}(T)$. Then

$$\begin{aligned}
T^*Tx &= T^*(Tx) \\
&= T^*(0) \\
&= 0
\end{aligned}$$

implies $x \in \mathrm{Ker}(T^*T)$. Thus $\mathrm{Ker}(T) \subseteq \mathrm{Ker}(T^*T)$.
For the reverse inclusion, let $x \in \mathrm{Ker}(T^*T)$. Then

$$\begin{aligned}
\langle Tx, Tx \rangle &= \langle x, T^*Tx \rangle \\
&= \langle x, 0 \rangle \\
&= 0
\end{aligned}$$

implies $Tx = 0$ (by positive-definiteness of inner-product) which implies $x \in \mathrm{Ker}(T)$. Therefore $\mathrm{Ker}(T) \supseteq \mathrm{Ker}(T^*T)$. □

10.6 $\ker(T^*) = (\operatorname{im} T)^\perp$ and $(\ker T)^\perp = \overline{\operatorname{im}(T^*)}$

Proposition 10.5. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a bounded operator. Then*

1. $\ker(T^*) = (\operatorname{im} T)^\perp$;
2. $(\ker T)^\perp = \overline{\operatorname{im}(T^*)}$.

Proof.

1. Let $x \in \ker(T^*)$. Then

$$\begin{aligned}\langle Ty, x\rangle &= \langle y, T^*x\rangle \\ &= \langle y, 0\rangle \\ &= 0\end{aligned}$$

for all $Ty \in \operatorname{im} T$. This implies $x \in (\operatorname{im} T)^\perp$ and so $\ker(T^*) \subseteq (\operatorname{im} T)^\perp$. For the reverse inclusion, let $x \in (\operatorname{im} T)^\perp$. Then

$$\begin{aligned}0 &= \langle x, TT^*x\rangle \\ &= \langle T^*x, T^*x\rangle\end{aligned}$$

implies $T^*x = 0$ (by positive-definiteness of inner-product) which implies $x \in \ker(T^*)$ and so $\ker(T^*) \supseteq (\operatorname{im} T)^\perp$.

2. Let $T^*y \in \operatorname{im}(T^*)$. Then for all $x \in \ker T$, we have

$$\begin{aligned}\langle x, T^*y\rangle &= \langle Tx, y\rangle \\ &= \langle 0, y\rangle \\ &= 0.\end{aligned}$$

In particular, this implies $\operatorname{im}(T^*) \subseteq (\ker T)^\perp$ which implies $\overline{\operatorname{im}(T^*)} \subseteq (\ker T)^\perp$ (as $(\ker T)^\perp$ is a closed subspace which contains $\operatorname{im}(T^*)$). For the reverse inclusion, we have

$$\begin{aligned}(\ker T)^\perp &= \ker((T^*)^*)^\perp \\ &= (\operatorname{im}(T^*)^\perp)^\perp \\ &\subseteq ((\overline{\operatorname{im}(T^*)})^\perp)^\perp \\ &= \overline{\operatorname{im}(T^*)},\end{aligned}$$

where we used part 1 of this proposition to get from the first line to the second line. □

10.7 T is an Isometry if and only if $T^*T = 1_{\mathcal{H}}$

Definition 10.1. An **isometry** between normed vector spaces $\mathcal{V}_1$ and $\mathcal{V}_2$ is an operator $T\colon \mathcal{V}_1 \to \mathcal{V}_2$ such that

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

for all $x, y \in \mathcal{V}$.

Proposition 10.6. *Let $\mathcal{V}_1$ and $\mathcal{V}_2$ be inner-product spaces and let $T\colon \mathcal{V}_1 \to \mathcal{V}_2$ be an operator. Then T is an isometry (where $\mathcal{V}_1$ and $\mathcal{V}_2$ are viewed as the induced normed vector spaces with respect to their inner-products) if and only if*

$$\langle x, y\rangle = \langle Tx, Ty\rangle \tag{35}$$

for all $x, y \in \mathcal{V}_1$.

Proof. Suppose (35) holds for all $x, y \in \mathcal{V}_1$. Then

$$\begin{aligned}\|Tx - Ty\| &= \sqrt{\langle Tx - Ty, Tx - Ty\rangle} \\ &= \sqrt{\langle Tx, Tx\rangle - \langle Tx, Ty\rangle - \langle Ty, Tx\rangle + \langle Ty, Ty\rangle} \\ &= \sqrt{\langle x, x\rangle - \langle x, y\rangle - \langle y, x\rangle + \langle y, y\rangle} \\ &= \sqrt{\langle x - y, x - y\rangle} \\ &= \|x - y\|.\end{aligned}$$

for all $x, y \in \mathcal{V}_1$. Thus T is an isometry.

Conversely, suppose T is an isometry and let $x, y \in \mathcal{V}_1$. Then

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

implies $\mathrm{Re}(\langle x, y\rangle) = \mathrm{Re}(\langle Tx, Ty\rangle)$ for all $x, y \in \mathcal{V}_1$. Note that this also implies

$$\begin{aligned}
\mathrm{Im}(\langle x, y\rangle) &= -\mathrm{Re}(i\langle x, y\rangle) \\
&= -\mathrm{Re}(\langle ix, y\rangle) \\
&= -\mathrm{Re}(\langle T(ix), Ty\rangle) \\
&= -\mathrm{Re}(i\langle Tx, Ty\rangle) \\
&= \mathrm{Im}(\langle Tx, Ty\rangle)
\end{aligned}$$

for all $x, y \in \mathcal{V}_1$. Thus we have (35) for all $x, y \in \mathcal{V}_1$. □

Proposition 10.7. *Let $T \colon \mathcal{H} \to \mathcal{H}$ be a bounded operator. Then*

1. *T is an isometry if and only if $T^*T = 1_{\mathcal{H}}$.*
2. *There exists isometries T such that $TT^* \neq 1_{\mathcal{H}}$.*

Proof.

1. Suppose T is an isometry. Then for all $y \in \mathcal{H}$, we have

$$\begin{aligned}
\langle x, 1_{\mathcal{H}} y\rangle &= \langle x, y\rangle \\
&= \langle Tx, Ty\rangle \\
&= \langle x, T^*Ty\rangle
\end{aligned}$$

for all $x \in \mathcal{H}$. In particular, this implies $T^*Ty = 1_{\mathcal{H}} y$ for all $y \in \mathcal{H}$, which implies $T^*T = 1_{\mathcal{H}}$.

Conversely, suppose $T^*T = 1_{\mathcal{H}}$. Then

$$\begin{aligned}
\langle Tx, Ty\rangle &= \langle x, T^*Ty\rangle \\
&= \langle x, 1_{\mathcal{H}} y\rangle \\
&= \langle x, y\rangle
\end{aligned}$$

for all $x, y \in \mathcal{H}$. This implies T is an isometry.

2. Consider the shift operator $S \colon \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$, given by

$$S(x_n) = (x_{n-1})$$

for all $(x_n) \in \ell^2(\mathbb{N})$, where $x_0 = 0$. In class, it was shown that

$$S^*(x_n) = (x_{n+1})$$

for all $(x_n) \in \ell^2(\mathbb{N})$. Thus, whenever $x_1 \neq 0$, we have

$$\begin{aligned}
SS^*(x_n) &= SS^*(x_1, x_2, \dots) \\
&= S(x_2, x_3, \dots) \\
&= (0, x_2, x_3, \dots) \\
&\neq (x_n).
\end{aligned}$$

On the other hand, S is an isometry. Indeed, let $(x_n), (y_n) \in \ell^2(\mathbb{N})$. Then

$$\begin{aligned}
\langle S(x_n), S(y_n)\rangle &= \langle (x_{n-1}), (y_{n-1})\rangle \\
&= \sum_{n=1}^{\infty} x_{n-1}\overline{y}_{n-1} \\
&= \sum_{m=0}^{\infty} x_m \overline{y}_m \\
&= x_0 y_0 + \sum_{m=1}^{\infty} x_m \overline{y}_m \\
&= \sum_{m=1}^{\infty} x_m \overline{y}_m \\
&= \langle (x_n), (y_n)\rangle.
\end{aligned}$$

□

Appendix

Proposition 10.8. *Let $T\colon \mathcal{U} \to \mathcal{V}$ be a bounded linear operator. Then*

$$\|T\|^2 = \sup\{\|Tx\|^2 \mid \|x\| \leq 1\}$$

Proof. For any $x \in \mathcal{U}$ such that $\|x\| \leq 1$, we have $\|Tx\|^2 \leq \|T\|^2$. Thus

$$\|T\|^2 \geq \sup\{\|Tx\|^2 \mid \|x\| \leq 1\}. \tag{36}$$

To show the reverse inequality, we assume (for a contradiction) that (36) is a strictly inequality. Choose $\delta > 0$ such that

$$\|T\|^2 - \delta > \sup\{\|Tx\|^2 \mid \|x\| \leq 1\}.$$

Now let $\varepsilon = \delta/2\|T\|$, and choose $x \in \mathcal{U}$ such that $\|x\| \leq 1$ and such that

$$\|T\| - \varepsilon < \|Tx\|.$$

Then

$$\begin{aligned}
\|Tx\|^2 &> (\|T\| - \varepsilon)^2 \\
&= \|T\|^2 - 2\varepsilon\|T\| + \varepsilon^2 \\
&\geq \|T\|^2 - 2\varepsilon\|T\| \\
&= \|T\|^2 - \delta
\end{aligned}$$

gives us a contradiction. □

11 Homework 6

Throughout this homework, let $\mathcal{H}$ be a Hilbert space.

11.1 Decomposition of T as $T = A + iB$ where A and B are Self-Adjoint

Proposition 11.1. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a bounded operator. There exists unique self-adjoint operators $A\colon \mathcal{H} \to \mathcal{H}$ and $B\colon \mathcal{H} \to \mathcal{H}$ such that $T = A + iB$.*

Proof. Define

$$A = \frac{1}{2}(T + T^*) \quad \text{and} \quad B = \frac{-i}{2}(T - T^*).$$

Then

$$\begin{aligned}
A + iB &= \frac{1}{2}(T + T^*) + \frac{1}{2}(T - T^*) \\
&= \left(\frac{1}{2} + \frac{1}{2}\right) T + \left(\frac{1}{2} - \frac{1}{2}\right) T^* \\
&= T
\end{aligned}$$

Furthermore, A and B are self-adjoint. Indeed,

$$\begin{aligned}
A^* &= \left(\frac{1}{2}(T+T^*)\right)^* \\
&= \frac{1}{2}(T^*+T^{**}) \\
&= \frac{1}{2}(T^*+T) \\
&= \frac{1}{2}(T+T^*) \\
&= A,
\end{aligned}$$

and similarly

$$\begin{aligned}
B^* &= \left(\frac{-i}{2}(T-T^*)\right)^* \\
&= \frac{i}{2}(T^*-T^{**}) \\
&= \frac{i}{2}(T^*-T) \\
&= \frac{-i}{2}(T-T^*) \\
&= B.
\end{aligned}$$

This establishes existence.

For uniqueness, suppose that $A' \colon \mathcal{H} \to \mathcal{H}$ and $B' \colon \mathcal{H} \to \mathcal{H}$ are two other self-adjoint operators such that $T = A' + iB'$. Then since

$$\begin{aligned}
T^* &= (A+iB)^* \\
&= A^* - iB^* \\
&= A - iB,
\end{aligned}$$

and since

$$\begin{aligned}
T^* &= (A'+iB')^* \\
&= A'^* - iB'^* \\
&= A' - iB',
\end{aligned}$$

we have

$$\begin{aligned}
A + iB &= A' + iB' \\
A - iB &= A' - iB'.
\end{aligned}$$

Adding these together gives us $2A = 2A'$, and hence $A = A'$. Similarly, subtracting these gives us $2iB = 2iB'$, and hence $B = B'$. □

11.2 S^*S and Orthogonal Projection are Positive Operators

Definition 11.1. A self-adjoint operator $T \colon \mathcal{H} \to \mathcal{H}$ is said to be **positive** if $\langle Tx, x\rangle \geq 0$ for all $x \in \mathcal{H}$. We say T is **strictly positive** if $\langle Tx, x\rangle > 0$ for all $x \in \mathcal{H}\backslash\{0\}$.

Remark. Equivalently, $T \colon \mathcal{H} \to \mathcal{H}$ is positive if and only if $\langle x, Tx\rangle \geq 0$ for all $x \in \mathcal{H}$. Indeed, if $\langle Tx, x\rangle \geq 0$ for all $x \in \mathcal{H}$, then $\langle Tx, x\rangle$ is real for all $x \in \mathcal{H}$, and so

$$\begin{aligned}
\langle x, Tx\rangle &= \overline{\langle Tx, x\rangle} \\
&= \langle Tx, x\rangle \\
&\geq 0.
\end{aligned}$$

Similarly, $\langle x, Tx\rangle \geq 0$ for all $x \in \mathcal{H}$ implies $\langle Tx, x\rangle \geq 0$ for all $x \in \mathcal{H}$.

Problem 2.a

Proposition 11.2. *Let $S\colon \mathcal{H} \to \mathcal{H}$ be a bounded operator. Then S^*S is positive.*

Proof. Let $x \in \mathcal{H}$. Then

$$\begin{aligned}\langle S^*Sx, x\rangle &= \langle Sx, Sx\rangle \\ &\geq 0\end{aligned}$$

by positive-definiteness of the inner-product. It follows that S^*S is positive. □

Remark. I think we do not need S to be bounded here, but we only defined the adjoint of a bounded operator in class.

Problem 2.b

Proposition 11.3. *Let $\mathcal{K}$ be a closed subspace of $\mathcal{H}$. Then the orthogonal projection $P_{\mathcal{K}}\colon \mathcal{H} \to \mathcal{H}$ is positive.*

Proof. Let $x \in \mathcal{H}$. Then

$$\begin{aligned}0 &= \langle x - P_{\mathcal{K}}x, P_{\mathcal{K}}x\rangle \\ &= \langle x, P_{\mathcal{K}}x\rangle - \langle P_{\mathcal{K}}x, P_{\mathcal{K}}x\rangle \\ &= \langle x, P_{\mathcal{K}}x\rangle - \|P_{\mathcal{K}}x\|^2.\end{aligned}$$

It follows that $\langle x, P_{\mathcal{K}}x\rangle = \|P_{\mathcal{K}}x\|^2 \geq 0$ which implies $P_{\mathcal{K}}$ is positive by Remark (11.2). □

11.3 Another Version of Polarization Identity and $\langle Tx, x\rangle = 0$ for all $x \in \mathcal{H}$ implies $T = 0$

Problem 3.a

Proposition 11.4. *(Another Version of Polarization Identity) Let $T\colon \mathcal{H} \to \mathcal{H}$ be any operator. Then*

$$4\langle Tx, y\rangle = \sum_{i=0}^{3} \langle T(x + i^k y), x + i^k y\rangle \tag{37}$$

Proof. We have

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

and

$$\begin{aligned}i\langle T(x+iy), x+iy\rangle &= i\langle Tx + iTy, x+iy\rangle \\ &= i\langle Tx, x\rangle + i\langle Tx, iy\rangle + i\langle iTy, x\rangle + i\langle iTy, iy\rangle \\ &= i\langle Tx, x\rangle + \langle Tx, y\rangle - \langle Ty, x\rangle + i\langle Ty, y\rangle\end{aligned}$$

and

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

and

$$\begin{aligned}-i\langle T(x-iy), x-iy\rangle &= -i\langle Tx - iTy, x-iy\rangle \\ &= -i\langle Tx, x\rangle - i\langle Tx, -iy\rangle - i\langle -iTy, x\rangle - i\langle -iTy, -iy\rangle \\ &= -i\langle Tx, x\rangle + \langle Tx, y\rangle - \langle Ty, x\rangle - i\langle Ty, y\rangle.\end{aligned}$$

Adding these together gives us our desired result. □

Problem 3.b

Proposition 11.5. *Let $T \colon \mathcal{H} \to \mathcal{H}$ be any operator such that $\langle Tx, x \rangle = 0$ for all $x \in \mathcal{H}$. Then $T = 0$.*

Proof. Let $x \in \mathcal{H}$. Then it follows from the polarization identity proved above that

$$\begin{aligned} 4\langle Tx, y \rangle &= \sum_{i=0}^{3} \langle T(x + i^k y), x + i^k y \rangle \\ &= \sum_{i=0}^{3} 0 \\ &= 0 \end{aligned}$$

for all $y \in \mathcal{H}$. It follows that $\langle Tx, y \rangle = 0$ for all $y \in \mathcal{H}$. This implies $Tx = 0$ by positive-definiteness of the inner-product. Since x was arbitrary, this implies $T = 0$. $\square$

11.4 Twisting Inner-Product by Strictly Positive Self-Adjoint Operator

Problem 4.a

Proposition 11.6. *Let $T \colon \mathcal{H} \to \mathcal{H}$ be a strictly positive self-adjoint operator. Define a map $\langle \cdot, \cdot \rangle_T \colon \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ by*

$$\langle x, y \rangle_T = \langle Tx, y \rangle$$

for all $x, y \in \mathcal{H}$. Then $\langle \cdot, \cdot \rangle_T$ is an inner-product.

Proof. We first check that $\langle \cdot, \cdot \rangle_T$ is linear in the first argument. Let $x, y, z \in \mathcal{H}$. Then

$$\begin{aligned} \langle x + z, y \rangle_T &= \langle T(x + z), y \rangle \\ &= \langle Tx + Tz, y \rangle \\ &= \langle Tx, y \rangle + \langle Tz, y \rangle \\ &= \langle x, y \rangle_T + \langle z, y \rangle_T. \end{aligned}$$

Next we check that $\langle \cdot, \cdot \rangle_T$ is conjugate-symmetric. Let $x, y \in \mathcal{H}$. Then since T is self-adjoint, we have

$$\begin{aligned} \langle x, y \rangle_T &= \langle Tx, y \rangle \\ &= \overline{\langle y, Tx \rangle} \\ &= \overline{\langle Ty, x \rangle} \\ &= \overline{\langle y, x \rangle}_T. \end{aligned}$$

Next we check that $\langle \cdot, \cdot \rangle_T$ is positive-definite. Let $x \in \mathcal{H}$. Then since T is strictly positive, we have

$$\begin{aligned} \langle x, x \rangle_T &= \langle Tx, x \rangle \\ &> 0, \end{aligned}$$

where $\langle x, x \rangle_T = 0$ if and only if $x = 0$. $\square$

Problem 4.b

Proposition 11.7. *Let $T \colon \mathcal{H} \to \mathcal{H}$ be a strictly positive self-adjoint operator. Then*

$$|\langle Tx, y \rangle|^2 \leq \langle Tx, x \rangle \langle Ty, y \rangle \tag{38}$$

for all $x, y \in \mathcal{H}$.

Proof. We have

$$\begin{aligned} |\langle Tx, y \rangle|^2 &= |\langle x, y \rangle_T|^2 \\ &=\leq \|x\|_T^2 \|y\|_T^2 \\ &= \langle x, x \rangle_T \langle y, y \rangle_T \\ &= \langle Tx, x \rangle \langle Ty, y \rangle, \end{aligned}$$

where we applied Cauchy-Schwarz for the $\langle \cdot, \cdot \rangle_T$ inner-product. $\square$

Problem 4.c

Proposition 11.8. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a strictly positive self-adjoint operator. Then*

$$\|Tx\|^2 \leq \|T\|\langle Tx, x\rangle \tag{39}$$

for all $x \in \mathcal{H}$.

Proof. Let $x \in \mathcal{H}$. Then

$$\begin{aligned}\|Tx\|^4 &= \langle Tx, Tx\rangle^2 \\ &\leq \langle Tx, x\rangle\langle T^2x, Tx\rangle \\ &\leq \langle Tx, x\rangle\|T^2x\|\|Tx\| \\ &\leq \langle Tx, x\rangle\|T\|\|Tx\|\|Tx\| \\ &= \langle Tx, x\rangle\|T\|\|Tx\|^2,\end{aligned}$$

where we used (38) to get from the first line to the second line. Now dividing both sides by $\|Tx\|^2$[10], we obtain $\|Tx\|^2 \leq \langle Tx, x\rangle\|T\|$. □

11.5 T is Self-Adjoint if and only if $\langle Tx, x\rangle \in \mathbb{R}$ for all $x \in \mathcal{H}$

Proposition 11.9. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a bounded operator. Then T is self-adjoint if and only if $\langle Tx, x\rangle \in \mathbb{R}$ for all $x \in \mathcal{H}$.*

Proof. Suppose that T is self-adjoint. Let $x \in \mathcal{H}$. Then

$$\begin{aligned}\langle Tx, x\rangle &= \langle x, Tx\rangle \\ &= \overline{\langle Tx, x\rangle}\end{aligned}$$

implies $\langle Tx, x\rangle \in \mathbb{R}$.

Conversely, suppose that $\langle Tx, x\rangle \in \mathbb{R}$ for all $x \in \mathcal{H}$. Then

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

for all $x \in \mathcal{H}$. Therefore by Proposition (11.5), we see that $T - T^* = 0$, i.e. $T = T^*$. □

11.6 If T is Self-Adjoint Operator, then $\|T^n\| = \|T\|^n$

Problem 6.a

Proposition 11.10. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a self-adjoint operator. Then*

$$\|T^n\|^2 \leq \|T^{n+1}\|\|T^{n-1}\| \tag{40}$$

for all $n \in \mathbb{N}$.

Proof. Let $n \in \mathbb{N}$ and let $x \in \mathcal{H}$ such that $\|x\| \leq 1$. Then

$$\begin{aligned}\|T^nx\|^2 &= \langle T^nx, T^nx\rangle \\ &= \langle T^{n+1}x, T^{n-1}x\rangle \\ &\leq \|T^{n+1}x\|\|T^{n-1}x\| \\ &\leq \|T^{n+1}\|\|x\|\|T^{n-1}\|\|x\| \\ &\leq \|T^{n+1}\|\|T^{n-1}\|,\end{aligned}$$

which implies (40). □

[10] If $Tx = 0$, then we clearly have (39), thus we assume $Tx \neq 0$.

Problem 6.b

Proposition 11.11. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a self-adjoint operator. Then*

$$\|T^n\| = \|T\|^n \tag{41}$$

for all $n \in \mathbb{N}$.

Proof. We prove (41) by induction on $n \geq 0$. The base case $n = 0$ and the case $n = 1$ are trivial. Assume that (41) holds for some $n \geq 1$. Then by (40), we have

$$\begin{aligned}
\|T^{n+1}\| &\geq \|T^{n-1}\|^{-1}\|T^n\|^2 \\
&= \|T\|^{1-n}\|T\|^{2n} \\
&= \|T\|^{n+1},
\end{aligned}$$

where we used the induction step to get from the first line to the second line.

For the reverse inequality, let $x \in \mathcal{H}$ such that $\|x\| \leq 1$. Then

$$\begin{aligned}
\|T^{n+1}x\| &\leq \|T^n x\|\|Tx\| \\
&\leq \|T^n\|\|x\|\|Tx\| \\
&\leq \|T^n\|\|Tx\| \\
&\leq \|T^n\|\|T\| \\
&= \|T\|^n\|T\| \\
&= \|T\|^{n+1},
\end{aligned}$$

where we used the induction step to get from the fourth line to the fifth line. It follows that $\|T^{n+1}\| \leq \|T\|^{n+1}$. □

12 Homework 7

Throughout this homework, let $\mathcal{H}$ be a Hilbert space.

12.1 Weak Convergence Properties

Problem 1.a

Proposition 12.1. *Let (x_n) and (y_n) be two sequences in $\mathcal{H}$ such that $x_n \xrightarrow{w} x$ and $y_n \xrightarrow{w} y$ and let $\alpha, \beta \in \mathbb{C}$. Then*

$$\alpha x_n + \beta y_n \xrightarrow{w} \alpha x + \beta y.$$

Proof. Let $z \in \mathcal{H}$. Then

$$\begin{aligned}
\lim_{n\to\infty}(\langle \alpha x_n + \beta y_n, z\rangle) &= \lim_{n\to\infty}(\alpha\langle x_n, z\rangle + \beta\langle y_n, z\rangle) \\
&= \alpha \lim_{n\to\infty}(\langle x_n, z\rangle) + \beta \lim_{n\to\infty}(\langle y_n, z\rangle) \\
&= \alpha\langle x, z\rangle + \beta\langle y, z\rangle \\
&= \langle \alpha x + \beta y, z\rangle.
\end{aligned}$$

Therefore $\alpha x_n + \beta y_n \xrightarrow{w} \alpha x + \beta y$. □

Problem 1.b

Proposition 12.2. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a bounded operator and let (x_n) be a sequence in $\mathcal{H}$ such that $x_n \xrightarrow{w} x$. Then*

$$Tx_n \xrightarrow{w} Tx.$$

Proof. Let $z \in \mathcal{H}$. Then

$$\begin{aligned}
\lim_{n\to\infty}(\langle Tx_n, z\rangle) &= \lim_{n\to\infty}(\langle x_n, T^*z\rangle) \\
&= \langle x, T^*z\rangle \\
&= \langle Tx, z\rangle.
\end{aligned}$$

Therefore $Tx_n \xrightarrow{w} Tx$. □

Remark. Note that we may not have $Tx_n \to Tx$. Indeed, suppose $\mathcal{H}$ is separable. Let (e_n) be an orthonormal sequence in $\mathcal{H}$ and let $T\colon \mathcal{H} \to \mathcal{H}$ be the identity map. Then $e_n \xrightarrow{w} 0$ but $Te_n = e_n \not\to 0$. In fact (Te_n) doesn't even converge.

12.2 Weak Convergence and Orthonormal Basis

Problem 2.a

Proposition 12.3. *Let $\mathcal{Y}$ be a dense subset of $\mathcal{H}$. Let (x_n) be a bounded sequence of elements in $\mathcal{H}$ and suppose $\langle x_n, y\rangle \to \langle x, y\rangle$ for all $y \in \mathcal{Y}$. Then $x_n \xrightarrow{w} x$.*

Proof. Let $z \in \mathcal{H}$. Choose $M \geq 0$ such that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$. Choose $y \in \mathcal{Y}$ such that

$$\|z - y\| < \frac{\varepsilon}{3\max\{\|x\|, M\}}$$

(we can do this since $\mathcal{Y}$ is dense in $\mathcal{H}$). Choose $N \in \mathbb{N}$ such that $n \geq N$ implies

$$|\langle x_n, y\rangle - \langle x, y\rangle| < \frac{\varepsilon}{3}.$$

Then $n \geq N$ implies

$$\begin{aligned}
|\langle x_n, z\rangle - \langle x, z\rangle| &= |\langle x_n, z\rangle - \langle x_n, y\rangle + \langle x_n, y\rangle - \langle x, y\rangle + \langle x, y\rangle - \langle x, z\rangle| \\
&\leq |\langle x_n, z - y\rangle| + |\langle x_n, y\rangle - \langle x, y\rangle| + |\langle x, y - z\rangle| \\
&\leq M\|z - y\| + |\langle x_n, y\rangle - \langle x, y\rangle| + \|x\|\|y - z\| \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&= \varepsilon.
\end{aligned}$$

Therefore $x_n \xrightarrow{w} x$. □

Problem 2.b

Lemma 12.1. *Let $\mathcal{H}$ be a separable Hilbert space, let (e_m) be an othonormal basis in $\mathcal{H}$, let (x_n) be a bounded sequence in $\mathcal{H}$, and let $x \in \mathcal{H}$. Then $x_n \xrightarrow{w} x$ if and only if $\langle x_n, e_m\rangle \to \langle x, e_m\rangle$ for all $m \in \mathbb{N}$.*

Proof. Suppose $x_n \xrightarrow{w} x$. Then $\langle x_n, y\rangle \to \langle x, y\rangle$ for all $y \in \mathcal{H}$, so certainly $\langle x_n, e_m\rangle \to \langle x, e_m\rangle$ for all $m \in \mathbb{N}$. Conversely, suppose $\langle x_n, e_m\rangle \to \langle x, e_m\rangle$ for all $m \in \mathbb{N}$. We first show that $\langle x_n, y\rangle \to \langle x, y\rangle$ for all $y \in \mathcal{Y}$, where $\mathcal{Y} = \operatorname{span}\{e_m \mid m \in \mathbb{N}\}$, so let $y \in \mathcal{Y}$. Then

$$y = \lambda_1 e_{m_1} + \cdots + \lambda_r e_{m_r}$$

for some (unique) $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ and $m_1, \dots, m_r \in \mathbb{N}$. Then

$$\begin{aligned}
\lim_{n\to\infty} \langle x_n, y\rangle &= \lim_{n\to\infty} \langle x_n, \lambda_1 e_{m_1} + \cdots + \lambda_r e_{m_r}\rangle \\
&= \overline{\lambda}_1 \lim_{n\to\infty} \langle x_n, e_{m_1}\rangle + \cdots + \overline{\lambda}_r \lim_{n\to\infty} \langle x_n, e_{m_r}\rangle \\
&= \overline{\lambda}_1 \langle x, e_{m_1}\rangle + \cdots + \overline{\lambda}_r \langle x, e_{m_r}\rangle \\
&= \langle x, \lambda_1 e_{m_1} + \cdots + \lambda_r e_{m_r}\rangle \\
&= \langle x, y\rangle.
\end{aligned}$$

Therefore $\langle x_n, y\rangle \to \langle x, y\rangle$ for all $y \in \mathcal{Y}$. Now since $\overline{\mathcal{Y}} = \mathcal{H}$, we may use Proposition (12.3) to conclude that $\langle x_n, z\rangle \to \langle x, z\rangle$ for all $z \in \mathcal{H}$. In other words, we have $x_n \xrightarrow{w} x$. □

Corollary. *Let (x^n) be a sequence in $\ell^2(\mathbb{N})$ that converges coordinate-wise to $x = (x_m) \in \ell^2(\mathbb{N})$. Then $x^n \xrightarrow{w} x$.*

Proof. Saying (x^n) converges coordinate-wise to $x = (x_m)$ is equivalent to saying

$$\begin{aligned}
\lim_{n\to\infty} (\langle x^n, e^m\rangle) &= x_m \\
&= \langle x, e^m\rangle
\end{aligned}$$

for all $m \in \mathbb{N}$. Thus Lemma (12.1) implies $x^n \xrightarrow{w} x$. □

12.3 If $\mathcal{K}$ is a Closed Subspace of $\mathcal{H}$ and (x_n) is a Sequence in $\mathcal{K}$ such that $x_n \xrightarrow{w} x$, then $x \in \mathcal{K}$

Proposition 12.4. *Let $\mathcal{K}$ be a closed subspace of $\mathcal{H}$. If (x_n) is a sequence of elements in $\mathcal{K}$ and $x_n \xrightarrow{w} x$, then $x \in \mathcal{K}$.*

Proof. Let $y \in \mathcal{K}^\perp$. Then

$$\begin{aligned} 0 &= \lim_{n\to\infty} (0) \\ &= \lim_{n\to\infty} (\langle x_n, y\rangle) \\ &= \langle x, y\rangle. \end{aligned}$$

This implies $x \in (\mathcal{K}^\perp)^\perp = \mathcal{K}$. □

Remark. The same proof shows that if A is a subset of $\mathcal{H}$ and (x_n) is a sequence of elements in A and $x_n \xrightarrow{w} x$, then $x \in \overline{\text{span}}(A)$.

12.4 T Bounded and S Compact Implies TS and ST Compact

Proposition 12.5. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a bounded operator and let $S\colon \mathcal{H} \to \mathcal{H}$ be a compact operator. Then $TS\colon \mathcal{H} \to \mathcal{H}$ and $ST\colon \mathcal{H} \to \mathcal{H}$ are both compact operators.*

Proof. We first show ST is compact. Let (x_n) be a sequence in $\mathcal{H}$ such that $x_n \xrightarrow{w} x$. Since T is a bounded operator, we have $Tx_n \xrightarrow{w} Tx$ by Proposition (12.2). Since S is compact, we have $S(Tx_n) \to S(Tx)$. Thus, ST is compact. Now we show TS is compact. Let (x_n) be a sequence in $\mathcal{H}$ such that $x_n \xrightarrow{w} x$. Since S is compact, we have $Sx_n \to Sx$. Since T is bounded (and in particular continuous) we have $T(Sx_n) \to T(Sx)$. Thus, TS is compact. □

12.5 Equivalent Definition of Compact Operator

Lemma 12.2. *Let (x_n) be a sequence in $\mathcal{H}$ such that $x_n \xrightarrow{w} x$ and let $(x_{\pi(n)})$ be a subsequence of (x_n). Then $x_{\pi(n)} \xrightarrow{w} x$.*

Remark. Here we view π as a strictly increasing function from $\mathbb{N}$ to $\mathbb{N}$ whose range consists of the indices in the subsequence.

Proof. Let $y \in \mathcal{H}$ and let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies

$$|\langle x_n, y\rangle - \langle x, y\rangle| < \varepsilon.$$

Then $\pi(n) \geq N$ implies

$$|\langle x_{\pi(n)}, y\rangle - \langle x, y\rangle| < \varepsilon$$

□

Proposition 12.6. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a bounded operator with the property that for any bounded sequence (x_n) in $\mathcal{H}$, the sequence (Tx_n) has a convergent subsequence. Then T is compact.*

Proof. Let (x_n) be a sequence in $\mathcal{H}$ such that $x_n \xrightarrow{w} x$. Assume for a contradiction that $Tx_n \not\to Tx$. Choose $\varepsilon > 0$ and choose a subsequence $(Tx_{\pi(n)})$ of (Tx_n) such that $\|Tx_{\pi(n)} - Tx\| > \varepsilon$ for all $n \in \mathbb{N}$. By the (baby version) of the uniform boundedness principle, the sequence (x_n) is bounded, and hence the subsequence $(x_{\pi(n)})$ must be bounded too. Thus the sequence $(Tx_{\pi(n)})$ has a convergent subsequence (by the hypothesis on T). Choose a convergent subsequence of $(Tx_{\pi(n)})$, say $(Tx_{\rho(n)})$. Since $(x_{\rho(n)})$ is a subsequence of (x_n), we must have $x_{\rho(n)} \xrightarrow{w} x$ by Lemma (12.2), and since T is a bounded operator, we must have $Tx_{\rho(n)} \xrightarrow{w} Tx$ by Proposition (12.2). Since $(Tx_{\rho(n)})$ is a convergent sequence and since $Tx_{\rho(n)} \xrightarrow{w} Tx$, we must in fact have $Tx_{\rho(n)} \to Tx$. But since $(Tx_{\rho(n)})$ is a subsequence of $(Tx_{\pi(n)})$, we have $\|Tx_{\rho(n)} - Tx\| > \varepsilon$ for all $n \in \mathbb{N}$, and so $Tx_{\rho(n)} \not\to Tx$. This is a contradiction. □

12.6 Finite-Dimensional Spaces, Bounded Sequences, and Compact Operators

Let $\mathcal{K}$ be a Hilbert space.

Problem 6.a

Proposition 12.7. *If every bounded sequence in $\mathcal{K}$ has a convergent subsequence, then $\mathcal{K}$ must be finite-dimensional.*

Proof. We prove the contrapositive statement: if $\mathcal{K}$ is infinite-dimensional, then there exists a bounded sequence which has no convergent subsequence. Assume $\mathcal{K}$ is infinite-dimensional. Let (e_n) be an orthonormal sequence in $\mathcal{K}$. Then the sequence (e_n) cannot have a convergent subsequence. Indeed, the distance squared between any two elements e_n, e_m (with $m \neq n$) in the sequence is

$$\begin{aligned}\|e_n - e_m\|^2 &= \|e_n\|^2 + \|e_m\|^2 \\ &= 1 + 1 \\ &= 2,\end{aligned}$$

by the Pythagorean Theorem. Thus any subsequence of (e_n) will fail to be Cauchy (and in particular will fail to converge). □

Problem 6.b

Proposition 12.8. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be compact. Then $\ker(1_{\mathcal{H}} - T)$ is finite-dimensional.*

Proof. Let (x_n) be a bounded sequence in $\ker(1_{\mathcal{H}} - T)$. Choose a subsequence $(x_{\pi(n)})$ of (x_n) such that $x_{\pi(n)} \xrightarrow{w} x$ for some $x \in \mathcal{H}$ (such a subsequence exists by a theorem proved in class). Since T is compact, we have

$$\begin{aligned}x_{\pi(n)} &= Tx_{\pi(n)} \\ &\to Tx \\ &= x.\end{aligned}$$

Thus $(x_{\pi(n)})$ is a convergent subsequence of (x_n). It follows from Proposition (12.7) that $\ker(1_{\mathcal{H}} - T)$ is finite-dimensional. □

Problem 6.c

Lemma 12.3. *Let $S\colon \mathcal{H} \to \mathcal{H}$ be a bounded operator. Assume $\mathrm{im}(S)$ is closed. Then $\mathrm{im}(S^*)$ is closed.*

we have $\mathcal{H} = \ker(S^*) \oplus \mathrm{im}(S)$

Proof. It suffices to show that $\ker(S)^\perp \subseteq \mathrm{im}(S^*)$. Let $x \in \ker(S)^\perp$, so $\langle y, x\rangle = 0$ for all $y \in \ker(S)$. We claim that $x = S^*z$ for some $S^*z \in \mathrm{im}(S^*)$. Indeed we have

$$\begin{aligned}\langle y, x\rangle &= \\ &= \langle Sy, z\rangle \\ &= \langle y, S^*z\rangle\end{aligned}$$

for all $y \in \mathcal{H}$.
Let $y \in \overline{\mathrm{im}}(S^*)$. Observe that $\langle y, x\rangle = 0$ for all $x \in \ker(S)$. Chooose a sequence (S^*x_n) in $\mathrm{im}(S^*)$ such that such that $S^*x_n \to y$. Since (S^*x_n) is convergent, the sequence (Sx_n) is convergent too. Indeed, let $\varepsilon > 0$.
For all $Sx \in \mathrm{im}(S)$, we have

$$\begin{aligned}\langle S^*x_n, Sx\rangle &\to \langle y, Sx\rangle \\ &= \langle S^*y, x\rangle.\end{aligned}$$

Similarly, for all $x \in \ker(S^*)$, we have

$$\begin{aligned}\langle S^*x_n, Sx\rangle &\to \langle y, x\rangle \\ &= \langle S^*y, x\rangle.\end{aligned}$$

□

Proposition 12.9. *Let $S\colon \mathcal{H} \to \mathcal{H}$ be a bounded operator whose image $\mathrm{im}(S)$ is closed and infinite-dimensional. Then S cannot be compact.*

Proof. Let (y_n) be a bounded sequence in $\operatorname{im}(S)$ and let $\mathcal{K} := \ker(S)$. Choose $N > 0$ such that $\|y_n\| \leq N$ for all $n \in \mathbb{N}$. Choose a sequence (x_n) in $\mathcal{H}$ such that $Sx_n = y_n$ for all $n \in \mathbb{N}$. By replacing x_n with $x_n - P_{\mathcal{K}} x_n$ if necessary, we may assume that $\langle x_n, z \rangle = 0$ for all $n \in \mathbb{N}$ and for all $z \in \mathcal{K}$. We claim that for each $z \in \mathcal{H}$, the sequence $(|\langle x_n, z\rangle|)$ is bounded. Indeed, since $\mathcal{H} = \ker(S) \oplus \overline{\operatorname{im}}(S^*)$, it suffices to show that for each $z \in \overline{\operatorname{im}}(S^*)$, the sequence $(|\langle x_n, z\rangle|)$. Let $z \in \overline{\operatorname{im}}(S^*)$ and choose a sequence $(S^* z_m)$ in $\operatorname{im}(S^*)$ such that $S^* z_m \to z$. Then

$$\begin{aligned} |\langle x_n, S^* z_m\rangle| &= |\langle Sx_n, z_m\rangle| \\ &= |\langle y_n, z_m\rangle| \\ &\leq \|y_n\| \|z_m\| \\ &\leq N \|z_m\|. \end{aligned}$$

and choose a sequence $(S^* z_m)$ in $\operatorname{im}(S^*)$ such that $S^* z_m \to z$. Then

$$\begin{aligned} |\langle x_n, S^* x\rangle| &= |\langle Sx_n, x\rangle| \\ &= |\langle y_n, x\rangle| \\ &\leq \|y_n\| \|x\| \\ &\leq N \|x\|. \end{aligned}$$

Thus for each $z \in \mathcal{H}$, the sequence $(|\langle x_n, z\rangle|)$ is bounded. It follows from the Uniform Boundedness Principle (stated in Theorem (4.3) in the Appendix) that the sequence (x_n) is bounded. Since S is compact and (x_n) is bounded, the sequence $(Sx_n) = (y_n)$ must have a convergent subsequence. This contradicts the fact that $\mathcal{K}$ is infinite-dimensional. □

Proof using Open Mapping Theorem:

Proof. Assume (for a contradiction) that S is compact. Denote $\mathcal{K} := \operatorname{im}(S)$ and let (y_n) be a bounded sequence in $\mathcal{K}$. Let $M > 0$ and choose $N > 0$ such that if $y \in \mathcal{K}$ and $\|y\| < N$, then there exists an $x \in \mathcal{H}$ such that $Sx = y$ and $\|x\| < M$ (this follows from the Open Mapping Theorem). By scaling the sequence (y_n) if necessary, we may assume that $\|y_n\| < N$ for all $n \in \mathbb{N}$. Thus there exists $x_n \in \mathcal{H}$ such that $Sx_n = y$ and $\|x_n\| < M$ for all $n \in \mathbb{N}$. Thus (x_n) is a bounded sequence in $\mathcal{H}$. Since S is compact and (x_n) is bounded, the sequence $(Sx_n) = (y_n)$ must have a convergent subsequence. This contradicts the fact that $\mathcal{K}$ is infinite-dimensional. □

13 Homework 8

Throughout this homework, let $\mathcal{H}$ be a separable Hilbert space. If $x \in \mathcal{H}$ and $r > 0$, then we write

$$B_r(x) := \{y \in \mathcal{H} \mid \|y - x\| < r\}$$

for the open ball centered at x and of radius r. We also write

$$B_r[x] := \{y \in \mathcal{H} \mid \|y - x\| \leq r\}$$

for the closed ball centered at x and of radius r.

13.1 Equivalent Definition of Compact Operator

Proposition 13.1. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a bounded linear operator. Then T is compact if and only if $\overline{T(B_1[0])}$ is a compact space.*

Proof. Suppose T is compact. To show that $\overline{T(B_1[0])}$ is compact, it suffices to show that $T(B_1[0])$ is precompact, by Proposition (13.9) (stated and proved in the Appendix). Let (Tx_n) be a sequence in $T(B_1[0])$. Then (x_n) is a bounded sequence in $B_1[0]$. Since T is compact, it follows that (Tx_n) has a convergent subsequence (by homework 7 probem 5). It follows that $T(B_1[0])$ is precompact.

Conversely, suppose $\overline{T(B_1[0])}$ is compact. Then $T(B_1[0])$ is precompact by Proposition (13.9). Let (x_n) be a bounded sequence in $\mathcal{H}$. Choose $M > 0$ such that $\|x_n\| < M$ for all $n \in \mathbb{N}$. Then $(T(x_n/M))$ is a sequence in the precompact space $T(B_1[0])$, and hence must have a convergent subsequence, say $(T(x_{\pi(n)}/M))$. This impies $(T(x_{\pi(n)}))$ is a convergent subsequence $(T(x_n))$. Thus, T is compact (again by homework 7 probem 5). □

13.2 Sequence of Compact Operators (T_n) Which Converge in Operator Norm to T Implies T is Compact

Proposition 13.2. *Let $(T_n \colon \mathcal{H} \to \mathcal{H})$ be a sequence of compact operators that converges in the operator norm to an operator $T \colon \mathcal{H} \to \mathcal{H}$. Then T is compact.*

Proof. Let (x_k) be a weakly convegent sequence. We claim that (Tx_k) is Cauchy. Indeed, let $\varepsilon > 0$. Since (x_k) is weakly convergent, it must be bounded. Choose $M > 0$ such that $\|x_k\| \leq M$ for all $k \in \mathbb{N}$. Choose $N \in \mathbb{N}$ such that $\|T - T_N\| < \varepsilon/3M$. Since the sequence $(T_N x_k)_{k \in \mathbb{N}}$ is Cauchy, there exists $K \in \mathbb{N}$ such that $j, k \geq K$ implies $\|T_N x_k - T_N x_j\| < \varepsilon/3$. Choose such a $K \in \mathbb{N}$. Then $j, k \geq K$ implies

$$
\begin{aligned}
\|Tx_k - Tx_j\| &= \|Tx_k - T_N x_k + T_N x_k - T_N x_j + T_N x_j - Tx_j\| \\
&\leq \|Tx_k - T_N x_k\| + \|T_N x_k - T_N x_j\| + \|T_N x_j - Tx_j\| \\
&\leq \|T - T_N\| \|x_k\| + \|T_N x_k - T_N x_j\| + \|T_N - T\| \|x_j\| \\
&< \frac{\varepsilon}{3M} \cdot M + \frac{\varepsilon}{3} + \frac{\varepsilon}{3M} \cdot M \\
&= \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&= \varepsilon.
\end{aligned}
$$

Thus (Tx_k) is a Cauchy sequence. It follows that T is compact. □

13.3 Hilbert-Schmidt Operator

Proposition 13.3. *Let $T \colon \mathcal{H} \to \mathcal{H}$ be a bounded operator and let (e_n) and (f_m) be any two orthonormal bases for $\mathcal{H}$. Then*

$$
\sum_{n=1}^{\infty} \|Te_n\|^2 = \sum_{m=1}^{\infty} \|T^* f_m\|^2.
$$

Proof. Since $\mathcal{H}$ is a separable Hilbert space, we have

$$
\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 \quad \text{and} \quad \|x\|^2 = \sum_{m=1}^{\infty} |\langle x, f_m \rangle|^2
$$

for every $x \in \mathcal{H}$. Therefore

$$
\begin{aligned}
\sum_{n=1}^{\infty} \|Te_n\|^2 &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |\langle Te_n, f_m \rangle|^2 \\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\langle Te_n, f_m \rangle|^2 \\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\langle e_n, T^* f_m \rangle|^2 \\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\langle T^* f_m, e_n \rangle|^2 \\
&= \sum_{m=1}^{\infty} \|T^* f_m\|^2,
\end{aligned}
$$

where we are justified in changing the order of the infinite sums by Lemma (13.1) (stated and proved in the Appendix). By swapping the roles of T with T^* in the proof above, we see that the quantity $\sum_{n=1}^{\infty} \|Te_n\|^2$ doesn't depend on the choice of the orthonormal basis (e_n). □

13.4 Hilbert-Schmidt Operator

Definition 13.1. An operator $T \colon \mathcal{H} \to \mathcal{H}$ is said to be a **Hilbert-Schmidt** operator if if

$$
\sum_{n=1}^{\infty} \|Te_n\|^2 < \infty
$$

for some or equivalently any orthonormal basis (e_n) of $\mathcal{H}$. In this case, the Hilbert-Schmidt norm of T is defined by

$$
\|T\|_{\mathrm{HS}} := \sqrt{\sum_{n=1}^{\infty} \|Te_n\|^2}.
$$

Problem 4.a

Proposition 13.4. *Let (e_n) be an orthonormal basis of $\mathcal{H}$. For each $k \in \mathbb{N}$ define a projection operator $P_k \colon \mathcal{H} \to \mathcal{H}$ onto span$\{e_1, e_2, \ldots, e_k\}$ by*

$$P_k(x) = \sum_{n=1}^{k} \langle x, e_n \rangle e_n$$

for all $x \in \mathcal{H}$. If $T \colon \mathcal{H} \to \mathcal{H}$ is a Hilbert-Schmidt operator, then $\|T - P_k T\| \to 0$ as $k \to \infty$.

Proof. Let $\varepsilon > 0$ and let $x \in B_1[0]$. Since the sum $\sum_{n=1}^{\infty} \|T^* e_n\|^2$ converges, there exists $K \in \mathbb{N}$ such that

$$\sum_{n=K}^{\infty} \|T^* e_n\|^2 < \varepsilon.$$

Choose such $K \in \mathbb{N}$. Then $k \geq K$ implies

$$\begin{aligned}
\|Tx - \mathrm{P}_k Tx\|^2 &= \left\| \sum_{n=1}^{\infty} \langle Tx, e_n \rangle e_n - \sum_{n=1}^{k} \langle Tx, e_n \rangle e_n \right\|^2 \\
&= \left\| \sum_{n=k+1}^{\infty} \langle Tx, e_n \rangle e_n \right\|^2 \\
&= \left\| \sum_{n=k+1}^{\infty} \langle x, T^* e_n \rangle e_n \right\|^2 \\
&= \sum_{n=k+1}^{\infty} |\langle x, T^* e_n \rangle|^2 \\
&\leq \sum_{n=k+1}^{\infty} \|T^* e_n\|^2 \\
&\leq \sum_{n=K}^{\infty} \|T^* e_n\|^2 \\
&< \varepsilon.
\end{aligned}$$

This implies $\|T - \mathrm{P}_k T\| \to 0$ as $k \to \infty$ by Remark (13.6) (stated in the Appendix). □

Problem 4.b

Proposition 13.5. *Every Hilbert-Schmidt operator is compact.*

Proof. Let $T \colon \mathcal{H} \to \mathcal{H}$ be a Hilbert-Schmidt operator. To show that T is compact, it suffices to show that $\mathrm{P}_k T$ is compact for all $k \in \mathbb{N}$ since Proposition (13.4) implies $\|\mathrm{P}_k T - T\| \to 0$ as $k \to \infty$ and Proposition (13.2) would then imply T is compact.

Let $k \in \mathbb{N}$ and let (x_n) be a weakly convergent sequence in $\mathcal{H}$, say $x_n \xrightarrow{w} x$. We claim that $\mathrm{P}_k x_n \to \mathrm{P}_k x$ as $n \to \infty$. Indeed, let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies

$$|\langle x_n, e_m \rangle - \langle x, e_m \rangle| < \frac{\varepsilon}{k}$$

for all $m = 1, \ldots, k$. Then $n \geq N$ implies

$$\begin{aligned}
\|\mathrm{P}_k x_n - \mathrm{P}_k x\| &= \left\| \sum_{m=1}^{k} \langle x_n, e_m \rangle e_m - \sum_{m=1}^{k} \langle x, e_m \rangle e_m \right\| \\
&= \left\| \sum_{m=1}^{k} (\langle x_n, e_m \rangle - \langle x, e_m \rangle) e_m \right\| \\
&\leq \sum_{m=1}^{k} |\langle x_n, e_m \rangle - \langle x, e_m \rangle| \\
&< \sum_{m=1}^{k} \frac{\varepsilon}{k} \\
&= \varepsilon.
\end{aligned}$$

□

Problem 4.c

Proposition 13.6. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a Hilbert-Schmidt operator. Then $\|T\| \leq \|T\|_{HS}$.*

Proof. Let $x \in B_1[0]$. Then

$$\begin{aligned}
\|Tx\|^2 &= \sum_{n=1}^{\infty} |\langle Tx, e_n \rangle|^2 \\
&= \sum_{n=1}^{\infty} |\langle x, T^* e_n \rangle|^2 \\
&\leq \sum_{n=1}^{\infty} \|T^* e_n\|^2 \\
&= \|T\|_{\mathrm{HS}}^2.
\end{aligned}$$

In particular this implies

$$\begin{aligned}
\|T\|^2 &= \sup\{\|Tx\|^2 \mid x \in B_1[0]\} \\
&\leq \|T\|_{\mathrm{HS}}^2,
\end{aligned}$$

where the first line is justified in the Appendix. Thus $\|T\| = \|T\|_{\mathrm{HS}}$. □

13.5 If T is Compact Self-Adjoint and $T^m = 0$, then $T = 0$

Proposition 13.7. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a compact self-adjoint operator. Suppose $T^m = 0$ for some $m \in \mathbb{N}$. Then we must have $T = 0$.*

Proof. If $T^m = 0$ for some $m \in \mathbb{N}$, then 0 is the only eigenvalue for T. Indeed, suppose λ is an eigenvalue of T. Choose an eigenvector of λ, say x. Then

$$\begin{aligned}
0 &= T^m x \\
&= \lambda^m x,
\end{aligned}$$

which implies $\lambda^m = 0$, and hence $\lambda = 0$. Now choose an orthonormal basis (e_n) consisting of eigenvectors of T (the existence of such basis is guaranteed by the spectral theorem for compact self-adjoint operators). Then for all $x \in \mathcal{H}$, we have

$$\begin{aligned}
Tx &= \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n \\
&= \sum_{n=1}^{\infty} 0 \cdot \langle x, e_n \rangle e_n \\
&= 0.
\end{aligned}$$

□

13.6 Every Compact Self-Adjoint Operator is Limit of Operators with Finite-Dimensional Range

Proposition 13.8. *Let $\mathcal{H}$ be a separable Hilbert space and let $T\colon \mathcal{H} \to \mathcal{H}$ be a compact self-adjoint operator. Then there exists a sequence T_m of operators with finite dimensional range such that $\|T - T_m\| \to 0$ and $m \to \infty$.*

Proof. Choose an orthonormal basis (e_n) consisting of eigenvectors of T and let (λ_n) be the corresponding sequence of eigenvalues. By reindexing if necessary, we may assume that $|\lambda_n| \geq |\lambda_{n+1}|$ for all $n \in \mathbb{N}$. For each $m \in \mathbb{N}$, we define $T_m\colon \mathcal{H} \to \mathcal{H}$ by

$$T_m x = \sum_{n=1}^{m} \lambda_n \langle x, e_n \rangle e_n$$

for all $x \in \mathcal{H}$. Observe that $\operatorname{im}(T_m) = \operatorname{span}(\{e_1, \dots, e_m\})$ is finite dimensional. We claim that $\|T - T_m\| \to 0$ and $m \to \infty$. Indeed, let $\varepsilon > 0$ and let Λ denote the set of all eigenvalues of T. If Λ is finite, then the claim is clear by the spectral theorem for compact self-adjoint operators, so assume Λ is infinite. Then 0 must be an accumulation point of Λ. In particular, $|\lambda_m| \to 0$ as $m \to \infty$. Choose $M \in \mathbb{N}$ such that $m \geq M$ implies $|\lambda_m| < \varepsilon$. Then $m \geq M$ implies

$$\begin{aligned}
\|Tx - T_m x\|^2 &= \left\| \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n - \sum_{n=1}^{m} \lambda_n \langle x, e_n \rangle e_n \right\|^2 \\
&= \left\| \sum_{n=m+1}^{\infty} \lambda_n \langle x, e_n \rangle e_n \right\|^2 \\
&= \sum_{n=m+1}^{\infty} |\lambda_n \langle x, e_n \rangle|^2 \\
&\leq |\lambda_M|^2 \sum_{n=m+1}^{\infty} |\langle x, e_n \rangle|^2 \\
&\leq |\lambda_M|^2 \|x\|^2 \\
&< \varepsilon^2.
\end{aligned}$$

for all $x \in B_1[0]$. This implies $\|T - T_m\| \to 0$ and $m \to \infty$. □

Appendix

Problem 1

Definition 13.2. A subspace $A \subseteq \mathcal{H}$ is said to be **precompact** if every sequence in A has a convergent subsequence.

Proposition 13.9. *Let A be a subspace of $\mathcal{H}$. Then A is precompact if and only if $\overline{A}$ is compact.*

Proof. Suppose A is precompact. Let (a_n) be a sequence in $\overline{A}$. For each $n \in \mathbb{N}$ choose $b_n \in A$ such that

$$\|a_n - b_n\| < \frac{1}{n}.$$

Choose a convergent subsequence of (b_n), say $(b_{\pi(n)})$ (we can do this since A is precompact). We claim that the sequence $(a_{\pi(n)})$ is Cauchy, and hence convergent subsequence of (a_n). Indeed, let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $\pi(n) \geq \pi(m) \geq N$ implies

$$\|b_{\pi(n)} - b_{\pi(m)}\| < \frac{\varepsilon}{3} \quad \text{and} \quad \frac{1}{\pi(m)} < \frac{\varepsilon}{3}.$$

Then $\pi(n) \geq \pi(m) \geq N$ implies

$$\begin{aligned}
\|a_{\pi(n)} - a_{\pi(m)}\| &= \|a_{\pi(n)} - b_{\pi(n)} + b_{\pi(n)} - b_{\pi(m)} + b_{\pi(m)} - a_{\pi(m)}\| \\
&\leq \|a_{\pi(n)} - b_{\pi(n)}\| + \|b_{\pi(n)} - b_{\pi(m)}\| + \|b_{\pi(m)} - a_{\pi(m)}\| \\
&< \frac{1}{\pi(n)} + \frac{\varepsilon}{3} + \frac{1}{\pi(m)} \\
&\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&= \varepsilon.
\end{aligned}$$

Finally, since $(a_{\pi(n)})$ is Cauchy and since $\mathcal{H}$ is a Hilbert space, we must have $a_{\pi(n)} \to a$ for some $a \in \overline{A}$. Therefore $\overline{A}$ is compact.

Conversely, suppose $\overline{A}$ is compact. Let (a_n) be a sequence in A. Then (a_n) is a sequence in $\overline{A}$. Since $\overline{A}$ is compact, the sequence (a_n) has a convergent subsequence. Therefore A is precompact. □

Convergence in Operator Norm

Remark. Let $\mathcal{V}$ be an inner-product space and let $(T_n : \mathcal{V} \to \mathcal{V})$ be a sequence of bounded linear operators. If we want to show $\|T_n - T\| \to 0$ as $n \to \infty$, then it suffices to show that for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\|T_n x - Tx\| < \varepsilon$$

for all $x \in B_1[0]$. Indeed, assuming this is true, choose $M \in \mathbb{N}$ such that $n \geq M$ implies

$$\|T_n x - Tx\| < \varepsilon/2$$

for all $x \in B_1[0]$. Then $n \geq M$ implies

$$\begin{aligned} \|T_n - T\| &= \sup\{\|T_n x - Tx\| \mid x \in B_1[0]\} \\ &\leq \varepsilon/2 \\ &< \varepsilon. \end{aligned}$$

Problem 3

Lemma 13.1. *Let f be a nonnegative function defined on $\mathbb{N} \times \mathbb{N}$. Then*

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} f(m,n) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} f(m,n).$$

Proof. Let $M \in \mathbb{N}$. Then

$$\begin{aligned} \sum_{m=1}^{M} \sum_{n=1}^{\infty} f(m,n) &= \sum_{m=1}^{M} \lim_{N\to\infty} \sum_{n=1}^{N} f(m,n) \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \sum_{m=1}^{M} f(m,n) \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{M} f(m,n) \\ &\leq \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} f(m,n). \end{aligned}$$

Taking the limit as $M \to \infty$ gives us

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} f(m,n) \leq \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} f(m,n).$$

A similar argument gives us

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} f(m,n) \geq \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} f(m,n).$$

□

Problem 4.c

Proposition 13.10. *Let $T : \mathcal{U} \to \mathcal{V}$ be a bounded linear operator. Then*

$$\|T\|^2 = \sup\{\|Tx\|^2 \mid \|x\| \leq 1\}$$

Proof. For any $x \in \mathcal{U}$ such that $\|x\| \leq 1$, we have $\|Tx\|^2 \leq \|T\|^2$. Thus

$$\|T\|^2 \geq \sup\{\|Tx\|^2 \mid \|x\| \leq 1\}. \tag{42}$$

To show the reverse inequality, we assume (for a contradiction) that (42) is a strict inequality. Choose $\delta > 0$ such that

$$\|T\|^2 - \delta > \sup\{\|Tx\|^2 \mid \|x\| \leq 1\}.$$

Now let $\varepsilon = \delta/2\|T\|$, and choose $x \in \mathcal{U}$ such that $\|x\| \leq 1$ and such that

$$\|T\| - \varepsilon < \|Tx\|.$$

Then

$$\begin{aligned}\|Tx\|^2 &> (\|T\| - \varepsilon)^2 \\ &= \|T\|^2 - 2\varepsilon\|T\| + \varepsilon^2 \\ &\geq \|T\|^2 - 2\varepsilon\|T\| \\ &= \|T\|^2 - \delta\end{aligned}$$

gives us a contradiction. □

14 Homework 9

Throughout this homework, let $\mathcal{H}$ be a separable Hilbert space.

14.1 Singular Values and Eigenvalues of Compact Positive Self-Adjoint Operator Coincide

Proposition 14.1. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a compact positive self-adjoint operator. Then $T = |T|$, and consequently the eigenvalues of T coincide with the singular values of T.*

Proof. Choose an orthonormal eigenbasis (e_n) of T with $Te_n = \lambda_n e_n$ for all $n \in \mathbb{N}$ (this exists since T is compact and self-adjoint). Then (e_n) is an orthonormal basis consisting of eigenvectors of $T^2 = T^*T$ with $T^2 e_n = \lambda_n^2 e_n$ for all $n \in \mathbb{N}$. Then since $\lambda_n \geq 0$ for all $n \in \mathbb{N}$ (since T is positive and self-adjoint), we have

$$\begin{aligned}|T|x &= \sum_{n=1}^{\infty} s_n \langle x, e_n\rangle e_n \\ &= \sum_{n=1}^{\infty} \sqrt{\lambda_n^2}\langle x, e_n\rangle e_n \\ &= \sum_{n=1}^{\infty} \lambda_n \langle x, e_n\rangle e_n \\ &= Tx\end{aligned}$$

for all $x \in \mathcal{H}$. It follows that $T = |T|$, and consequently $s_n = \lambda_n$ for all $n \in \mathbb{N}$. □

14.2 Compact Operator that is not Hilbert-Schmidt

Proposition 14.2. *Let (e_n) be an orthonormal basis for $\mathcal{H}$. Define $T\colon \mathcal{H} \to \mathcal{H}$ by*

$$T(x) = \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}\langle x, e_n\rangle e_n.$$

for all $x \in \mathcal{H}$. Then $T\colon \mathcal{H} \to \mathcal{H}$ is compact but not Hilbert-Schmidt.

Remark. For this problem, I decided to prove this in an arbitrary separable Hilbert space than just $\ell^2(\mathbb{N})$.

Proof. We first show T is compact. For each $k \in \mathbb{N}$, define $T_k\colon \mathcal{H} \to \mathcal{H}$ by

$$T_k(x) = \sum_{n=1}^{k} \frac{1}{\sqrt{n}}\langle x, e_n\rangle e_n$$

for all $x \in \mathcal{H}$. First note that for each $k \in \mathbb{N}$, the operator T_k is bounded and has finite rank, and hence must be compact. Moreover, we have $\|T - T_k\| \to 0$ as $k \to \infty$. Indeed, let $\varepsilon > 0$ and let $x \in B_1[0]$ (so $\sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2 \leq 1$).

Choose $K \in \mathbb{N}$ such that $1/K < \varepsilon$. Then $k \geq K$ implies

$$\begin{aligned}
\|Tx - T_k x\|^2 &= \left\| \sum_{n=k+1}^{\infty} \frac{1}{\sqrt{n}} \langle x, e_n \rangle e_n \right\|^2 \\
&= \sum_{n=k+1}^{\infty} \left| \frac{\langle x, e_n \rangle}{\sqrt{n}} \right|^2 \\
&= \sum_{n=k+1}^{\infty} \frac{|\langle x, e_n \rangle|^2}{n} \\
&\leq \frac{1}{K} \sum_{n=k+1}^{\infty} |\langle x, e_n \rangle|^2 \\
&\leq \frac{1}{K} \\
&< \varepsilon.
\end{aligned}$$

This implies $\|T - T_k\| \to 0$ as $k \to \infty$. Thus (T_k) is a sequence of compact operators such that $\|T - T_k\| \to 0$ as $k \to \infty$. Therefore T is compact.

To see that T is not Hilbert-Schmidt, observe that

$$\begin{aligned}
\sum_{n=1}^{\infty} \|Te_n\|^2 &= \sum_{n=1}^{\infty} \|\frac{1}{\sqrt{n}} e_n\|^2 \\
&= \sum_{n=1}^{\infty} \frac{1}{n}
\end{aligned}$$

is the harmonic series which does not converge. □

14.3 Eigenvalue for Self-Adjoint Operator is less than or equal to the Operator Norm

Proposition 14.3. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a self-adjoint operator and let λ be an eigenvalue of T. Then $|\lambda| \leq \|T\|$.*

Proof. Choose an eigenvector x corresponding to the eigenvalue λ. By scaling if necessary, we may assume $\|x\| = 1$. Then

$$\begin{aligned}
\|T\| &= \sup\{|\langle Ty, y \rangle| \mid \|y\| \leq 1\} \\
&\geq |\langle Tx, x \rangle| \\
&= |\langle \lambda x, x \rangle| \\
&= |\lambda|.
\end{aligned}$$

□

If T is Compact, then $\||T|\| = \|T\|$

Lemma 14.1. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a compact operator. Then $\||T|\| = \|T\|$.*

Proof. Combining problem 5 on HW5 and problem 6.b on HW6, we have

$$\begin{aligned}
\||T|\|^2 &= \||T|^2\| \\
&= \|T^*T\| \\
&= \|T\|^2.
\end{aligned}$$

It follows that $\||T|\| = \|T\|$ since the norm of an operator is nonnegative. □

Proposition 14.4. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a compact operator and let s be a singular value of T. Then we have $0 \leq s \leq \|T\|$.*

Proof. Clearly we have $s \geq 0$ by definition. Combining Lemma (14.1) and Proposition (14.3) gives us

$$\begin{aligned}
|s| &\leq \||T|\| \\
&= \|T\|.
\end{aligned}$$

□

14.4 The Square of Hilbert-Schmidt Norm of Compact Operator equals sum of Singular Values Squared

Proposition 14.5. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a compact operator. Let (s_n) be the sequence of singular values of T. Then $\|T\|_{HS} = \sqrt{\sum_{n=1}^{\infty} s_n^2}$.*

Proof. Let (x_n) be an orthonormal basis for T^*T. Then

$$
\begin{aligned}
\|T\|_{\mathrm{HS}} &= \sqrt{\sum_{n=1}^{\infty} \|Tx_n\|^2} \\
&= \sqrt{\sum_{n=1}^{\infty} \langle Tx_n, Tx_n\rangle} \\
&= \sqrt{\sum_{n=1}^{\infty} \langle T^*Tx_n, x_n\rangle} \\
&= \sqrt{\sum_{n=1}^{\infty} \langle s_n^2 x_n, x_n\rangle} \\
&= \sqrt{\sum_{n=1}^{\infty} s_n^2}.
\end{aligned}
$$

□

14.5 Compact Self-Adjoint Operator which is not Zero

Proposition 14.6. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a compact self-adjoint operator. Then $T^2 + T + 1$ cannot be the zero operator.*

Proof. Choose an orthonormal eigenbasis (e_n) of T with $Te_n = \lambda_n e_n$ for all $n \in \mathbb{N}$. Assume for a contradiction that $T^2 + T + 1 = 0$. Then

$$
\begin{aligned}
0 &= (T^2 + T + 1)e_n \\
&= \sum_{n=1}^{\infty} (\lambda_n^2 + \lambda_n + 1)\langle e_n, e_n\rangle e_n \\
&= (\lambda_n^2 + \lambda_n + 1)e_n,
\end{aligned}
$$

which implies $\lambda_n^2 + \lambda_n + 1 = 0$ for all $n \in \mathbb{N}$. Therefore $\lambda_n = \pm e^{2\pi i/3}$ for all $n \in \mathbb{N}$, but this contradicts the fact that the λ_n must be real. □

14.6 Compact Operator is Limit of Operators with Finite-Dimensional Range

Proposition 14.7. *Let $T\colon \mathcal{H} \to \mathcal{H}$ be a compact operator. Then there exists a sequence $T_n\colon \mathcal{H} \to \mathcal{H}$ of operators with finite dimensional range such that $\|T_n - T\| \to 0$ as $n \to \infty$.*

Proof. Let $T = U|T|$ be the polar decomposition of T. Choose a sequence (S_n) of bounded operators with finite dimensional range such that $\|S_n - |T|\| \to 0$ as $n \to \infty$ (such a sequence exists by problem 6 HW8). Then for each $n \in \mathbb{N}$, the operator $T_n := US_n$ has finite dimensional range since S_n has finite dimensional range. Moreover we have $\|T - T_n\| \to 0$ as $n \to \infty$. Indeed, let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $\||T| - S_n\| < \frac{\varepsilon}{\|U\|}$. Then $n \geq N$ implies

$$
\begin{aligned}
\|T - T_n\| &= \|U|T| - US_n\| \\
&= \|U(|T| - S_n)\| \\
&\leq \|U\|\||T| - S_n\|| \\
&< \|U\|\frac{\varepsilon}{\|U\|} \\
&= \varepsilon.
\end{aligned}
$$

□

15 Homework 10

15.1 $(C[a,b], \|\cdot\|_\infty)$ is a Banach Space

Proposition 15.1. *Let $\|\cdot\|_\infty \colon C[a,b] \times C[a,b] \to \mathbb{R}$ be given by*

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in [a,b]\} \tag{43}$$

for all $f \in C[a,b]$. Then $\|\cdot\|_\infty$ is a norm. Moreover, the pair $(C[a,b], \|\cdot\|_\infty)$ forms a Banach space.

Proof. Let us first show $\|\cdot\|_\infty$ is a norm. First note that the set $\{|f(x)| \mid x \in [a,b]\}$ is non-empty and bounded above (since f is continuous on a compact interval and hence attains a maximum). Therefore the supremum (43) exists.

For positive-definiteness, let $f \in C[a,b]$. Then

$$\begin{aligned} \|f\|_\infty &= \sup\{|f(x)| \mid x \in [a,b]\} \\ &\geq \sup\{0 \mid x \in [a,b]\} \\ &= 0. \end{aligned}$$

We have equality if and only if $|f(x)| = 0$ for all $x \in [a,b]$, and since $|\cdot|$ is positive-definite, this is equivalent to f being the zero function.

For absolute-homogeneity, let $f \in C[a,b]$ and $\alpha \in \mathbb{C}$. Then

$$\begin{aligned} \|\alpha f\|_\infty &= \sup\{|\alpha f(x)| \mid x \in [a,b]\} \\ &= \sup\{|\alpha||f(x)| \mid x \in [a,b]\} \\ &= |\alpha| \sup\{|f(x)| \mid x \in [a,b]\} \\ &= |\alpha|\|f\|_\infty, \end{aligned}$$

where the equality at the third line is justified by Proposition (15.10) (stated and proved in the Appendix).

For subadditivity, let $f, g \in C[a,b]$. Then

$$\begin{aligned} \|f+g\|_\infty &= \sup\{|f(x)+g(x)| \mid x \in [a,b]\} \\ &\leq \sup\{|f(x)| + |g(x)| \mid x \in [a,b]\} \\ &= \sup\{|f(x)| \mid x \in [a,b]\} + \sup\{|g(x)| \mid x \in [a,b]\} \\ &= \|f\|_\infty + \|g\|_\infty, \end{aligned}$$

where the equality at the third line is justified by Proposition (15.11) (stated and proved in the Appendix).

Finally, to show that $(C[a,b], \|\cdot\|_\infty)$ forms a Banach space, we need to show that every Cauchy sequence in $(C[a,b], \|\cdot\|_\infty)$ is convegent. Throughout the rest of the proof, we drop the notation $(C[a,b], \|\cdot\|_\infty)$ and simply write $C[a,b]$ instead. Let (f_n) be a Cauchy sequence in $C[a,b]$. We first make the observation that for each $x \in [a,b]$, the sequence $(f_n(x))$ forms a Cauchy sequence of complex numbers. Indeed, let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $m, n \geq N$ implies $\|f_n - f_m\|_\infty < \varepsilon$. In other words, $m, n \geq N$ implies

$$\sup\{|f_n(x) - f_m(x)| \mid x \in [a,b]\} < \varepsilon.$$

In particular $m, n \geq N$ implies

$$|f_n(x) - f_m(x)| < \varepsilon \tag{44}$$

for all $x \in [a,b]$. This proves our claim.

Since $\mathbb{C}$ is complete, we are justified in defining $f \colon [a,b] \to \mathbb{C}$ by

$$f(x) := \lim_{n \to \infty} f_n(x)$$

for all $x \in [a,b]$. By taking $m \to \infty$ in (44), we see that (f_n) converges *uniformly* to f. In particular, this implies f is continuous (by the usual $\varepsilon/3$ trick). Thus $f \in C[a,b]$. Finally, we note that convergence in $\|\cdot\|_\infty$ is equivalent to uniform convergence. Thus the Cauchy sequence (f_n) converges in the $\|\cdot\|_\infty$ norm to f. $\square$

15.2 Normed Linear Space + Parallelogram Law = Inner-Product Space

Proposition 15.2. *Let $(V, \|\cdot\|)$ be a normed linear space over $\mathbb{C}$ which satisfies the parallelogram law. Then the map $\langle\cdot,\cdot\rangle\colon V\times V\to\mathbb{C}$ defined by*

$$\langle x,y\rangle = \frac{1}{4}\left(\|x+y\|^2 + i\|x+iy\|^2 - \|x-y\|^2 - i\|x-iy\|^2\right) \tag{45}$$

for all $x,y\in V$ is an inner-product. Moreover, the norm induced by this inner-product is precisely $\|\cdot\|$. In other words, we have

$$\langle x,x\rangle = \|x\|^2$$

for all $x\in V$.

Proof. The most difficult part of this proof is showing that (45) is linear in the first argument. Before we do this, let us show that (55) is positive-definite and conjugate-symmetric.

For positive-definiteness, let $x\in V$. Then

$$\begin{aligned}\langle x,x\rangle &= \frac{1}{4}\left(\|2x\|^2 + i\|x+ix\|^2 - i\|x-ix\|^2\right)\\ &= \frac{1}{4}\left(\|2x\|^2 + i((|1+i|^2 - |1-i|^2)\|x\|^2)\right)\\ &= \|x\|^2\\ &\geq 0,\end{aligned}$$

with equality if and only if $x=0$. Note that this also gives us $\langle x,x\rangle = \|x\|^2$ for all $x\in V$.

For conjugate-symmetry, let $x,y\in V$. Then

$$\begin{aligned}\overline{\langle y,x\rangle} &= \frac{1}{4}\overline{(\|y+x\|^2 + i\|y+ix\|^2 - \|y-x\|^2 - i\|y-ix\|^2)}\\ &= \frac{1}{4}\left(\|y+x\|^2 - i\|y+ix\|^2 - \|y-x\|^2 + i\|y-ix\|^2\right)\\ &= \frac{1}{4}\left(\|x+y\|^2 - i\|i(x-iy)\|^2 - \|x-y\|^2 + i\|i(x+iy)\|^2\right)\\ &= \frac{1}{4}\left(\|x+y\|^2 - i\|x-iy\|^2 - \|x-y\|^2 + i\|x+iy\|^2\right)\\ &= \frac{1}{4}\left(\|x+y\|^2 + i\|x+iy\|^2 - \|x-y\|^2 - i\|x-iy\|^2\right)\\ &= \langle x,y\rangle\end{aligned}$$

Now we come to the difficult part, namely showing that (55) is linear in the first argument. We do this in several steps:

Step 1: We show that (55) is additive in the first argument (i.e. $\langle x+z,y\rangle = \langle x,y\rangle + \langle z,y\rangle$ for all $x,y,z\in V$). Let $x,y,z\in V$. First note that by the parallelogram law, we have

$$\begin{aligned}\|x+z+y\|^2 - \|x+z-y\|^2 &= 2\|x+y\|^2 + 2\|z\|^2 - \|x+y-z\|^2 - 2\|x-y\|^2 - 2\|z\|^2 + \|x-y-z\|^2\\ &= 2\|x+y\|^2 - 2\|x-y\|^2 - \|z-y-x\|^2 + \|z+y-x\|^2\\ &= 2\|x+y\|^2 - 2\|x-y\|^2 - 2\|z-y\|^2 - 2\|x\|^2 + \|z-y+x\|^2 + 2\|z+y\|^2 + 2\|x\|^2 - \|z+y+x\|^2\\ &= 2\|x+y\|^2 - 2\|x-y\|^2 + 2\|z+y\|^2 - 2\|z-y\|^2 + \|x+z-y\|^2 - \|x+z+y\|^2.\end{aligned}$$

Adding $\|x+z-y\|^2 - \|x+z+y\|^2$ to both sides gives us

$$2(\|x+z+y\|^2 - \|x+z-y\|^2) = 2(\|x+y\|^2 - \|x-y\|^2 + \|z+y\|^2 - \|z-y\|^2),$$

and after cancelling 2 from both sides, we obtain

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

Therefore

$$
\begin{aligned}
\langle x+z,y\rangle &= \frac{1}{4}\left(\|x+z+y\|^2+i\|x+z+iy\|^2-\|x+z-y\|^2-i\|x+z-iy\|^2\right) \\
&= \frac{1}{4}\left(\|x+z+y\|^2-\|x+z-y\|^2+i(\|x+z+iy\|^2-\|x+z-iy\|^2)\right) \\
&= \frac{1}{4}\left(\|x+y\|^2-\|x-y\|^2+\|z+y\|^2-\|z-y\|^2+i(\|x+iy\|^2-\|x-iy\|^2+\|z+iy\|^2-\|z-iy\|^2)\right) \\
&= \frac{1}{4}\left(\|x+y\|^2-\|x-y\|^2+\|z+y\|^2-\|z-y\|^2+i(\|x+iy\|^2-\|x-iy\|^2+\|z+iy\|^2-\|z-iy\|^2)\right) \\
&= \frac{1}{4}\left(\|x+y\|^2+i\|x+iy\|^2-\|x-y\|^2-i\|x-iy\|^2+\|z+y\|^2+i\|z+iy\|^2-\|z-y\|^2-i\|z-iy\|^2\right) \\
&= \langle x,y\rangle+\langle z,y\rangle.
\end{aligned}
$$

Thus we have additivity in the first argument.

Step 2: We show that (55) respects $\mathbb{Z}$-scaling in the first argument (i.e. $m\langle x,y\rangle = \langle mx,y\rangle$ for all integers $m\in\mathbb{Z}$ and for all $x,y\in V$). It suffices to show that (55) respects $\mathbb{N}$-scaling in the first argument since additivity implies

$$
\begin{aligned}
0 &= \langle 0,y\rangle \\
&= \langle x-x,y\rangle \\
&= \langle x,y\rangle+\langle -x,y\rangle,
\end{aligned}
$$

which implies $\langle -x,y\rangle = -\langle x,y\rangle$ for all $x,y\in V$. We prove (55) respects $\mathbb{N}$-scaling in the first argument using induction on $m\geq 2$. The base case $m=2$ follows from Step 1. Now assume that for some $m\geq 2$ and for all $x,y\in V$, we have $\langle mx,y\rangle = m\langle x,y\rangle$. Then we have

$$
\begin{aligned}
\langle (m+1)x,y\rangle &= \langle mx+x,y\rangle \\
&= \langle mx,y\rangle+\langle x,y\rangle \\
&= m\langle x,y\rangle+\langle x,y\rangle \\
&= (m+1)\langle x,y\rangle,
\end{aligned}
$$

where we applied the induction step at the third line.

Step 3: We show that (55) respects $\mathbb{Q}$-scaling in the first argument. Let $\frac{m}{n}\in\mathbb{Q}$ and let $x,y\in V$. Then since (55) is additive in the first argument and since V is a $\mathbb{C}$-vector space, we have

$$
\begin{aligned}
\frac{m}{n}\langle x,y\rangle &= \frac{m}{n}\left\langle \frac{n}{n}x,y\right\rangle \\
&= \frac{mn}{n}\left\langle \frac{1}{n}x,y\right\rangle \\
&= m\left\langle \frac{1}{n}x,y\right\rangle \\
&= \left\langle \frac{m}{n}x,y\right\rangle.
\end{aligned}
$$

Therefore (55) respects $\mathbb{Q}$-scaling in the first argument.

Step 4: We show that (55) respects $\mathbb{R}$-scaling in the first argument. First note that for each $y\in V$, the map $\langle\cdot,y\rangle\colon V\to\mathbb{C}$ is continuous since the norm is continuous. Let $x,y\in V$ and let $r\in\mathbb{R}$. Choose a sequence (r_n) of rational numbers such that $r_n\to r$ (we can do this since $\mathbb{Q}$ is dense in $\mathbb{R}$). Then we have

$$
\begin{aligned}
\langle rx,y\rangle &= \lim_{n\to\infty}\langle r_nx,y\rangle \\
&= \lim_{n\to\infty} r_n\langle x,y\rangle \\
&= r\langle x,y\rangle.
\end{aligned}
$$

Therefore (55) respects $\mathbb{R}$-scaling in the first component.

Step 5: We show that (55) respects $\mathbb{C}$-scaling in the first component. We first show that $\langle ix,y\rangle = i\langle x,y\rangle$ for all $x,y\in V$.

Let $x, y \in V$. Then we have

$$\begin{aligned}
\langle ix, y\rangle &= \frac{1}{4}\left(\|ix+y\|^2 + i\|ix+iy\|^2 - \|ix-y\|^2 - i\|ix-iy\|^2\right) \\
&= \frac{1}{4}\left(\|x-iy\|^2 + i\|x+y\|^2 - \|x+iy\|^2 - i\|x-y\|^2\right) \\
&= \frac{1}{4}\left(i\|x+y\|^2 - \|x+iy\|^2 - i\|x-y\|^2 + \|x-iy\|^2\right) \\
&= \frac{i}{4}\left(\|x+y\|^2 + i\|x+iy\|^2 - \|x-y\|^2 - i\|x-iy\|^2\right) \\
&= i\langle x, y\rangle.
\end{aligned}$$

Now let $\lambda = r + is \in \mathbb{C}$. Then we have

$$\begin{aligned}
\langle \lambda x, y\rangle &= \langle (r+is)x, y\rangle \\
&= \langle rx + isx, y\rangle \\
&= \langle rx, y\rangle + \langle isx, y\rangle \\
&= r\langle x, y\rangle + s\langle ix, y\rangle \\
&= r\langle x, y\rangle + is\langle x, y\rangle \\
&= (r+is)\langle x, y\rangle \\
&= \lambda\langle x, y\rangle
\end{aligned}$$

for all $x, y \in V$. Therefore (55) respects $\mathbb{C}$-scaling in the first component. □

15.3 Example of Bounded Operator in $(C[0,1], \|\cdot\|_\infty)$

Proposition 15.3. *Consider $C[0,1]$ equipped with the supremum norm. Let $T\colon C[0,1] \to C[0,1]$ be the linear operator defined by*

$$(Tf)(x) = \int_0^x f(y)\mathrm{d}y$$

for all $x \in [0,1]$. Then T is bounded with $\|T\| = 1$.

Proof. Let $f \in C[0,1]$ such that $\|f\|_\infty \leq 1$. Then

$$\begin{aligned}
\|Tf\|_\infty &= \sup\{|(Tf)(x)| \mid x \in [0,1]\} \\
&= \sup\left\{\left|\int_0^x f(y)\mathrm{d}y\right| \mid x \in [0,1]\right\} \\
&\leq \sup\left\{\int_0^x |f(y)|\mathrm{d}y \mid x \in [0,1]\right\} \\
&\leq \sup\left\{\int_0^x \mathrm{d}y \mid x \in [0,1]\right\} \\
&= \sup\{x \mid x \in [0,1]\} \\
&= 1.
\end{aligned}$$

Thus $\|T\| \leq 1$. To see that $\|T\| = 1$, let $f\colon [0,1] \to \mathbb{C}$ be the constant function $f = 1$. Then $\|f\|_\infty = 1$ and

$$\begin{aligned}
\|Tf\|_\infty &= \sup\{|(Tf)(x)| \mid x \in [0,1]\} \\
&= \sup\left\{\left|\int_0^x \mathrm{d}y\right| \mid x \in [0,1]\right\} \\
&= \sup\{|x| \mid x \in [0,1]\} \\
&= \sup\{x \mid x \in [0,1]\} \\
&= 1.
\end{aligned}$$

□

15.4 Example of Linear Functional in $(C[a,b], \|\cdot\|_\infty)$

Proposition 15.4. *Consider $C[a,b]$ equipped with the supremum norm. Define a linear functional $\ell\colon C[a,b] \to \mathbb{R}$ by*

$$\ell(f) := f(a) - f(b).$$

for all $f \in C[a,b]$. Then ℓ is bounded. Moreover the set

$$\{f \in C[a,b] \mid f(a) = f(b)\}$$

is a closed subspace of $C[a,b]$.

Proof. Let $f \in C[a,b]$ such that $\|f\|_\infty \le 1$. Then

$$\begin{aligned} |\ell(f)| &= |f(a) - f(b)| \\ &\le |f(a)| + |f(b)| \\ &\le 1 + 1 \\ &= 2. \end{aligned}$$

Thus $\|\ell\| \le 2$. To see that $\|\ell\| = 2$, let $f\colon [a,b] \to \mathbb{C}$ be given by

$$f(x) = \frac{2}{b-a}(x-a) - 1$$

for all $x \in [a,b]$. So the graph of f is just the line segment from $(a,-1)$ to $(b,1)$. In particular, $\|f\|_\infty = 1$ and

$$\begin{aligned} |\ell(f)| &= |f(a) - f(b)| \\ &= |-1-1| \\ &= 2. \end{aligned}$$

The last part of the proposition follows from

$$\ker\ell = \{f \in C[a,b] \mid f(a) = f(b)\},$$

and $\ker\ell$ is a closed subspace since ℓ is a bounded linear operator. □

15.5 Example of Linear Functional in $(C[a,b], \|\cdot\|_\infty)$ and Closed Subspace

Lemma 15.1. *Consider $C[a,b]$ equipped with the supremum norm. Let $[c,d] \subseteq [a,b]$ and define $\ell_{c,d}\colon C[a,b] \to \mathbb{C}$ by*

$$\ell_{c,d}(f) = \int_c^d f(t)\mathrm{d}t$$

for all $f \in C[a,b]$. Then $\ell_{c,d}$ is a bounded linear functional with $\|\ell_{c,d}\| = d - c$.

Proof. Linearity of $\ell_{c,d}$ follows from linearity of integration. So it suffices to check that $\ell_{c,d}$ is bounded. Let $f \in C[a,b]$ such that $\|f\|_\infty \le 1$. Then

$$\begin{aligned} |\ell_{c,d}(f)| &= \left|\int_c^d f(t)\mathrm{d}t\right| \\ &\le \int_c^d |f(t)|\mathrm{d}t \\ &\le \int_c^d \mathrm{d}t \\ &= d - c. \end{aligned}$$

Thus $\|\ell\| \le d - c$. To see that $\|\ell\| = d - c$, let $f\colon [a,b] \to \mathbb{C}$ be the constant function $f = 1$. Then $\|f\|_\infty = 1$ and

$$\begin{aligned} |\ell_{c,d}(f)| &= \left|\int_c^d f(t)\mathrm{d}t\right| \\ &= \left|\int_c^d \mathrm{d}t\right| \\ &= |d - c| \\ &= d - c. \end{aligned}$$

□

Proposition 15.5. *Consider $C[-1,1]$ equipped with the supremum norm. Let $\mathcal{Y}$ be the subset of $C[-1,1]$ consisting of all functions $g \in C[-1,1]$ such that*

$$\int_{-1}^{0} g(x)\mathrm{d}x = \int_{0}^{1} g(x)\mathrm{d}x = 0.$$

Then $\mathcal{Y}$ is a closed subspace.

Proof. Note that $\mathcal{Y} = \ker \ell_{-1,0} \cap \ker \ell_{0,1}$ is an intersection of two closed subspaces (since $\ell_{-1,0}$ and $\ell_{0,1}$ are bounded linear functionals by Lemma (15.1)). Thus $\mathcal{Y}$ is a closed subspace. $\square$

Proposition 15.6. *With the notation as in Proposition (15.5) above, let $h \in C[-1,1]$ be given by*

$$h(x) = 2x$$

for all $x \in [-1,1]$. Then there does not exist a $g \in \mathcal{Y}$ such that $\|g - h\|_\infty = d(h, \mathcal{Y})$.

Proof.

Step 1: We will first show that $d(h, \mathcal{Y}) = 1$. To prove $d(h, \mathcal{Y}) \geq 1$, assume for a contradiction that $d(h, \mathcal{Y}) < 1$. Choose $\varepsilon > 0$ and $g \in \mathcal{Y}$ such that

$$\|g - h\|_\infty < 1 - \varepsilon.$$

Write g in terms of its real and imaginary parts, say $g = u + iv$. Then

$$\begin{aligned} 0 &= \int_{-1}^{0} g(x)\mathrm{d}x \\ &= \int_{-1}^{0} u(x)\mathrm{d}x + i\int_{-1}^{0} v(x)\mathrm{d}x \end{aligned}$$

implies $\int_{-1}^{0} u(x)\mathrm{d}x = 0$ and $\int_{-1}^{0} v(x)\mathrm{d}x = 0$. Similarly,

$$\begin{aligned} 0 &= \int_{0}^{1} g(x)\mathrm{d}x \\ &= \int_{0}^{1} u(x)\mathrm{d}x + i\int_{0}^{1} v(x)\mathrm{d}x \end{aligned}$$

implies $\int_{0}^{1} u(x)\mathrm{d}x = 0$ and $\int_{0}^{1} v(x)\mathrm{d}x = 0$. Moreover, we have

$$\begin{aligned} 1 - \varepsilon &> \|g - h\|_\infty \\ &= \sup_{x \in [-1,1]} \sqrt{(u(x) - h(x))^2 + v(x)^2} \\ &\geq \sup_{x \in [-1,1]} \sqrt{(u(x) - h(x))^2} \\ &= \|u - h\|_\infty. \end{aligned}$$

Therefore $u \in \mathcal{Y}$, $\|u - h\|_\infty < 1 - \varepsilon$, and u is a real-valued function. Since $\|u - h\|_\infty < 1 - \varepsilon$, $h(x) = 2x$ for all $x \in [-1,0]$, and both u and h are real-valued functions, we have

$$u(x) \leq 2x + 1 - \varepsilon$$

for all $x \in [-1,0]$. This implies

$$\begin{aligned} 0 &= \int_{-1}^{0} u(x)\mathrm{d}x \\ &\leq \int_{-1}^{0} (2x + 1 - \varepsilon)\mathrm{d}x \\ &= (x^2 + x - \varepsilon x)|_{-1}^{0} \\ &= \varepsilon \\ &> 0, \end{aligned}$$

which gives us our desired contradiction. Therefore $d(h, \mathcal{Y}) \geq 1$.

Now we will show that $d(h, \mathcal{Y}) \leq 1$. Let $t \in (0,1]$ and define $g_t \colon [-1,0] \to \mathbb{R}$ by the formula

$$g_t(x) = \begin{cases} 2x + 1 + t & \text{if } -1 \leq x \leq \frac{-2t}{1+t} \\ -\frac{(t-1)^2}{2t}x & \text{if } \frac{-2t}{1+t} \leq x \leq 0. \end{cases}$$

Extend g_t to all of $[-1,1]$ by the formula

$$g_t(x) = g_t(-x)$$

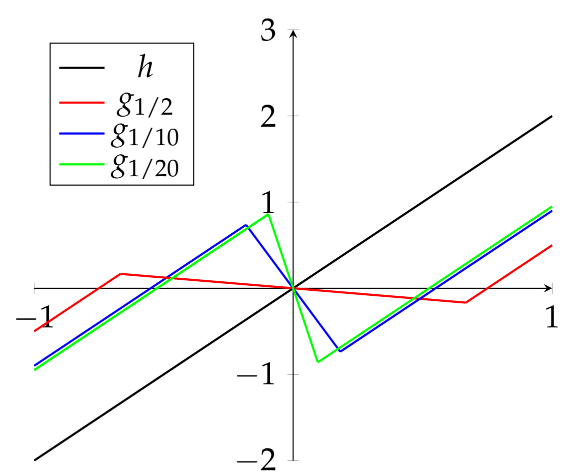
for all $x \in [0,1]$. So g_t is an odd function. Moreover g_t is continous since each segment of g_t is linear and since they agree on their boundaries:

$$\begin{aligned}
2\left(\frac{-2t}{1+t}\right)+1+t &= \frac{-4t}{1+t}+\frac{(1+t)^2}{1+t} \\
&= \frac{t^2-2t+1}{1+t} \\
&= \frac{(t-1)^2}{1+t} \\
&= -\frac{(t-1)^2}{2t}\left(\frac{-2t}{1+t}\right)
\end{aligned}$$

and

$$\begin{aligned}
-\frac{(t-1)^2}{2t}\cdot 0 &= 0 \\
&= \frac{(t-1)^2}{2t}\cdot 0.
\end{aligned}$$

The image below gives the graphs for h, $g_{1/2}$, and $g_{1/10}$:



Now observe that

$$\begin{aligned}
\int_{-1}^{0} g_t(x)\mathrm{d}x &= \int_{-1}^{\frac{-2t}{1+t}}(2x+1+t)\mathrm{d}x+\int_{\frac{-2t}{1+t}}^{0}-\frac{(t-1)^2}{2t}x\mathrm{d}x \\
&= (x^2+x+tx)|_{-1}^{\frac{-2t}{1+t}}+\left(\frac{-(t-1)^2}{4t}x^2\right)|_{\frac{-2t}{1+t}}^{0} \\
&= \left(\frac{2t}{1+t}\right)^2+\left(\frac{-2t}{1+t}\right)+t\left(\frac{-2t}{1+t}\right)-(1-1-t)+\frac{(t-1)^2}{4t}\left(\frac{2t}{1+t}\right)^2 \\
&= \left(\frac{(t-1)^2}{4t}+1\right)\left(\frac{2t}{1+t}\right)^2+(1+t)\left(\frac{-2t}{1+t}\right)+t \\
&= \frac{(t+1)^2}{4t}\frac{4t^2}{(1+t)^2}-t \\
&= t-t \\
&= 0.
\end{aligned}$$

Therefore $g_t \in \mathcal{Y}$ for all $t \in (0,1]$. Moreover, by construction we have

$$\|g_t - h\|_\infty = 1+t$$

for all $t \in (0,1]$. This implies $d(h,\mathcal{Y}) \leq 1$.

Step 2: We claim that there does not exist a $g \in \mathcal{Y}$ such that $\|g-h\|_\infty = 1$. Indeed, assume for a contradiction there does exist a $g \in \mathcal{Y}$ such that $\|g-h\|_\infty = 1$. Choose such a $g \in \mathcal{Y}$. We may assume that g is real-valued: if g is not real-valued, then we pass to its real-valued part u and as argued above we obtain $u \in \mathcal{Y}$ and

$$\begin{aligned} 1 &= \|g-h\|_\infty \\ &= \|u-h\|_\infty \\ &\geq 1. \end{aligned}$$

Since g is real-valued and $\|g-h\|_\infty = 1$, we have

$$2x-1 \leq g(x) \leq 2x+1$$

for all $x \in [-1,1]$. Since g is continous, we cannot have

$$g(x) = \begin{cases} 2x+1 & \text{for all } x \in (-1,0) \\ 2x-1 & \text{for all } x \in (0,1). \end{cases}$$

Assume $g(x) \neq 2x-1$ on the interval $(0,1)$. Choose $c \in (0,1)$ such that $g(c) \neq 2c-1$. Since g is continuous and since $g(c) > 2c-1$, there exists $\varepsilon > 0$ and $\delta > 0$ such that

$$g(x) > 2x-1+\varepsilon$$

for all $x \in (c-\delta, c+\delta)$. Choose such ε and δ so that $(c-\delta, c+\delta) \subset (0,1)$. Then

$$\begin{aligned} 0 &= \int_0^1 g(x)\mathrm{d}x \\ &= \int_0^1 g(x)\mathrm{d}x + \int_{c-\delta}^{c+\delta} g(x)\mathrm{d}x + \int_{c+\delta}^1 g(x)\mathrm{d}x \\ &> \int_0^{c-\delta} (2x-1)\mathrm{d}x + \int_{c-\delta}^{c+\delta} (2x-1+\varepsilon)\mathrm{d}x + \int_{c+\delta}^1 (2x-1)\mathrm{d}x \\ &= \int_0^1 (2x-1)\mathrm{d}x + \int_{c-\delta}^{c+\delta} \varepsilon \mathrm{d}x \\ &= (x^2-x)|_0^1 + \varepsilon x|_{c-\delta}^{c+\delta} \\ &= 2\varepsilon\delta \\ &> 0 \end{aligned}$$

gives us a contradiction.

Thus $g(x) \neq 2x+1$ on the interval $(-1,0)$. Choose $c \in (-1,0)$ such that $g(c) \neq 2c+1$. Then by a similar argument as above, we have

$$\begin{aligned} 0 &= \int_{-1}^0 g(x)\mathrm{d}x \\ &< \int_{-1}^0 (2x+1)\mathrm{d}x - \int_{c-\delta}^{c+\delta} \varepsilon \mathrm{d}x \\ &= -2\varepsilon\delta \\ &< 0, \end{aligned}$$

which also gives us a contradiction. Therefore there does not exist a $g \in \mathcal{Y}$ such that $\|g-h\|_\infty = 1$. □

15.6 Closed Subspaces of $\mathcal{X}^*$ and $\mathcal{X}$

Definition 15.1. Let $\mathcal{X}$ be a normed linear space. For a set $A \subseteq \mathcal{X}$ we define $A^\perp$ to be the subset of $\mathcal{X}^*$ consisting of all $\ell \in \mathcal{X}^*$ such that $\ell(a) = 0$ for all $a \in A$. Similarly, for a set $M \subseteq \mathcal{X}^*$ we define $M_\perp$ to be the subset of $\mathcal{X}$ consisting of all vectors $x \in \mathcal{X}$ such that $\ell(x) = 0$ for all $\ell \in M$.

Proposition 15.7. *Let $\mathcal{X}$ be a normed linear space, let A be a subset of $\mathcal{X}$, and let M be a subset of $\mathcal{X}^*$. Then $A^\perp$ and $M_\perp$ are closed subspaces of $\mathcal{X}^*$ and $\mathcal{X}$ respectively.*

Proof. Let $x \in \mathcal{X}$. Define $\widehat{x} \colon \mathcal{X}^* \to \mathbb{C}$ by

$$\widehat{x}(\ell) = \ell(x)$$

for all $\ell \in \mathcal{X}^*$. We claim that $\widehat{x}$ is a bounded linear functional. To see that $\widehat{x}$ is linear, let $\ell, \ell' \in \mathcal{X}^*$ and let $\lambda, \lambda' \in \mathbb{C}$. Then

$$\begin{aligned}
\widehat{x}(\lambda\ell + \lambda'\ell') &= (\lambda\ell + \lambda'\ell')(x) \\
&= \lambda\ell(x) + \lambda'\ell'(x) \\
&= \lambda\widehat{x}(\ell) + \lambda'\widehat{x}(\ell').
\end{aligned}$$

To see that $\widehat{x}$ is bounded, let $\ell \in \mathcal{X}^*$. Then

$$\begin{aligned}
|\widehat{x}(\ell)| &= |\ell(x)| \\
&\leq \|x\|\|\ell\|.
\end{aligned}$$

Therefore $\widehat{x}$ is a bounded linear functional. In particular $\ker\widehat{x}$ is a closed subspace. Thus

$$A^\perp = \bigcap_{a \in A} \ker\widehat{a} \quad \text{and} \quad M_\perp = \bigcap_{\ell \in M} \ker\ell$$

are closed subspaces since an arbitrary intersection of closed subspaces is a closed subspace. □

Proposition 15.8. *Let $\mathcal{X}$ be a normed linear space, let A be a subset of $\mathcal{X}$, and let M be a subset of $\mathcal{X}^*$. Then $\overline{\text{span}}(A) \subseteq (A^\perp)_\perp$ and $\overline{\text{span}}(M) \subseteq (M_\perp)^\perp$.*

Proof. Proposition (15.7) implies $(A^\perp)_\perp$ and $(M_\perp)^\perp$ are closed subspace. Thus, it suffices to show

$$\text{span}(A) \subseteq (A^\perp)_\perp \quad \text{and} \quad \text{span}(M) \subseteq (M_\perp)^\perp.$$

First we show the former. Let $\lambda_1 a_1 + \cdots + \lambda_n a_n \in \text{span}(A)$ and let $\ell \in A^\perp$. Then since $\ell(a) = 0$ for all $a \in A$, we have

$$\begin{aligned}
\ell(\lambda_1 a_1 + \cdots + \lambda_n a_n) &= \lambda_1\ell(a_1) + \cdots + \lambda_n\ell(a_n) \\
&= \lambda_1 \cdot 0 + \cdots + \lambda_n \cdot 0 \\
&= 0.
\end{aligned}$$

Since ℓ was arbitrary, this implies $\lambda_1 a_1 + \cdots + \lambda_n a_n \in (A^\perp)_\perp$, and hence $\text{span}(A) \subseteq (A^\perp)_\perp$.

Now we show the latter. Let $\lambda_1\ell_1 + \cdots + \lambda_n\ell_n \in \text{span}(M)$ and let $x \in M_\perp$. Then since $\ell(x) = 0$ for all $\ell \in M$, we have

$$\begin{aligned}
(\lambda_1\ell_1 + \cdots + \lambda_n\ell_n)(x) &= \lambda_1\ell_1(x) + \cdots + \lambda_n\ell_n(x) \\
&= \lambda_1 \cdot 0 + \cdots + \lambda_n \cdot 0 \\
&= 0.
\end{aligned}$$

Since x was arbitrary, this implies $\lambda_1\ell_1 + \cdots + \lambda_n\ell_n \in (M_\perp)^\perp$, and hence $\text{span}(M) \subseteq (M_\perp)^\perp$. □

15.7 $(\ell^1)^*$ is isometrically isomorphic to ℓ^∞.

Proposition 15.9. *$(\ell^1)^*$ is isometrically isomorphic to ℓ^∞.*

Proof. For each $n \in \mathbb{N}$, let e^n denote the sequence with entry 1 in the nth component and entry 0 everywhere else. Define $\Phi\colon (\ell^1)^* \to \ell^\infty$ by

$$\Phi(\psi) = (\psi(e^n))$$

for all $\psi \in (\ell^1)^*$. Note that for any $\psi \in (\ell^1)^*$, we have $|(\psi(e^n))| \leq \|\psi\|$, and therefore $(\psi(e^n)) \in \ell^\infty$. We claim that $\|\psi\| = \|\Phi(\psi)\|_\infty$. Indeed,

$$\begin{aligned}
\|\Phi(\psi)\|_\infty &= \sup\{|\psi(e^n)| \mid n \in \mathbb{N}\} \\
&\leq \sup\left\{\left|\psi\left(\sum_{n=1}^\infty a_n e^n\right)\right| \mid \sum_{n=1}^\infty |a_n| \leq 1\right\} \\
&= \|\psi\|.
\end{aligned}$$

To prove the reverse inequality assume for a contradiction that $\|\psi\| > \|\Phi(\psi)\|_\infty$. Choose $\varepsilon > 0$ and $\sum_{n=1}^\infty a_n e^n \in \ell^1$ such that $\sum_{n=1}^\infty |a_n| \leq 1$ and

$$\left|\psi\left(\sum_{n=1}^\infty a_n e^n\right)\right| > \|\Phi(\psi)\|_\infty + \varepsilon. \tag{46}$$

Choose $N \in \mathbb{N}$ such that $\sum_{n=N}^{\infty} |a_n| < \varepsilon / \|\psi\|$ (we can find such an N since $\sum_{n=1}^{\infty} |a_n| < \infty$). Then

$$\begin{aligned}
\left|\psi\left(\sum_{n=1}^{\infty} a_n e^n\right)\right| &= \left|\psi\left(\sum_{n=1}^{N} a_n e^n + \sum_{n=N+1}^{\infty} a_n e^n\right)\right| \\
&= \left|\psi\left(\sum_{n=1}^{N} a_n e^n\right) + \psi\left(\sum_{n=N+1}^{\infty} a_n e^n\right)\right| \\
&= \left|\sum_{n=1}^{N} a_n \psi(e^n) + \psi\left(\sum_{n=N+1}^{\infty} a_n e^n\right)\right| \\
&\leq \left|\sum_{n=1}^{N} a_n \psi(e^n)\right| + \left|\psi\left(\sum_{n=N+1}^{\infty} a_n e^n\right)\right| \\
&\leq \sum_{n=1}^{N} |a_n||\psi(e^n)| + \|\psi\| \sum_{n=N+1}^{\infty} |a_n| \\
&< \|\Phi(\psi)\|_\infty \sum_{n=1}^{N} |a_n| + \|\psi\| \cdot \frac{\varepsilon}{\|\psi\|} \\
&\leq \|\Phi(\psi)\|_\infty + \varepsilon.
\end{aligned}$$

This contradicts (**??**).

Next we show Φ is linear. Let $\varphi, \psi \in (\ell^1)^*$ and let $\lambda, \mu \in \mathbb{C}$. Then

$$\begin{aligned}
\Phi(\lambda\varphi + \mu\psi) &= ((\lambda\varphi + \mu\psi)(e^n)) \\
&= \lambda(\varphi(e^n)) + \mu(\psi(e^n)) \\
&= \lambda\Phi(\varphi) + \mu\Phi(\psi).
\end{aligned}$$

Therefore Φ is an isometric embedding.

Now show that Φ is surjective, and hence an isometric isomorphism. Let $(a_n) \in \ell^\infty$, let $M = \sup\{|a_n|\}$, and let $E = \operatorname{span}\{e^n \mid n \in \mathbb{N}\}$. Define $\varphi \colon E \to \mathbb{C}$ to be the unique linear map such that

$$\varphi(e^n) = a_n$$

for all $n \in \mathbb{N}$. Let $x = x_{n_1} e^{n_1} + \cdots + x_{n_k} e^{n_k} \in E$ such that $|x_{n_1}| + \cdots + |x_{n_k}| \leq 1$. Then

$$\begin{aligned}
|\varphi(x_{n_1} e^{n_1} + \cdots + x_{n_k} e^{n_k})| &= |x_{n_1}\varphi(e^{n_1}) + \cdots + x_{n_k}\varphi(e^{n_k})| \\
&= |x_{n_1} a_{n_1} + \cdots + x_{n_k} a_{n_k}| \\
&\leq |x_{n_1}||a_{n_1}| + \cdots + |x_{n_k}||a_{n_k}| \\
&\leq |x_{n_1}|M + \cdots + |x_{n_k}|M \\
&= (|x_{n_1}| + \cdots + |x_{n_k}|)M \\
&\leq M
\end{aligned}$$

It follows that φ is bounded. By the Hahn-Banach Theorem, there exists a bounded linear functional $\widetilde{\varphi}$ defined on all of ℓ^1 such that $\widetilde{\varphi}|_E = \varphi$ and $\|\widetilde{\varphi}\| = \|\varphi\|$. Choose such a $\widetilde{\varphi} \in (\ell^1)^*$. Then clearly $\Phi(\widetilde{\varphi}) = (a_n)$. Therefore Φ is surjective, and hence an isometric isomorphism. □

Appendix

Problem 1

Proposition 15.10. *Let A be a non-emtpy set of real numbers which is bounded above and let λ be any non-negative real number. Then*

$$\sup(\lambda A) = \lambda \sup(A). \tag{47}$$

Proof. If $\lambda = 0$, then (47) is obvious, so assume $\lambda > 0$. Let α denote $\sup(A)$. Choose any element in λA, say λa where $a \in A$. Then since $a \leq \alpha$ and λ is non-negative, we have $\lambda a \leq \lambda\alpha$. This implies

$$\sup(\lambda A) \leq \lambda \sup(A).$$

For the reverse direction, observe that

$$\begin{aligned}
\sup(A) &= \sup(\lambda^{-1}\lambda A) \\
&\leq \lambda^{-1} \sup(\lambda A),
\end{aligned}$$

and this implies

$$\sup(\lambda A) \geq \lambda \sup(A).$$

$\square$

Proposition 15.11. *Let A and B be non-empty sets of non-negative real numbers both of which are bounded above. Then*

$$\sup(A+B) = \sup(A) + \sup(B). \tag{48}$$

Proof. Let α denote $\sup(A)$, let β denote $\sup(B)$, and let $a+b$ be an arbitrary element in $A+B$. Then $a \leq \alpha$ and $b \leq \beta$ implies $a+b \leq \alpha+\beta$. Therefore

$$\sup(A+B) \leq \sup(A) + \sup(B). \tag{49}$$

To show the reverse inequality, we assume (for a contradiction) that the inequality (49) is strict

$$\sup(A+B) < \sup(A) + \sup(B).$$

Choose $\varepsilon > 0$ such that

$$\sup(A+B) < \sup(A) + \sup(B) - \varepsilon. \tag{50}$$

Choose $a \in A$ and $b \in B$ such that $a > \alpha - \varepsilon/2$ and $b > \beta - \varepsilon/2$. Then

$$\begin{aligned} a+b &> \alpha - \frac{\varepsilon}{2} + \beta - \frac{\varepsilon}{2} \\ &= \alpha + \beta - \varepsilon. \end{aligned}$$

But this contradicts (50). Therefore

$$\sup(A+B) \geq \sup(A) + \sup(B).$$

$\square$

Part III
Appendix

16 Completion

Let $\mathcal{V}$ be an inner-product space. In this section, we describe a procedure called **completion** which constructs a Hibert space $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$ and an injective linear map $\iota\colon \mathcal{V} \to \mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$ such that ι respects then inner-product structure on both $\mathcal{V}$ and $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$ (namely we will show that ι is an isometry) and such that $\iota(\mathcal{V})$ is dense in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$.

16.1 Constructing Completions

Let $\mathscr{C}_{\mathcal{V}}$ denote the set of all Cauchy sequences in $\mathcal{V}$. We can give $\mathscr{C}_{\mathcal{V}}$ the structure of a $\mathbb{C}$-vector space as follows: let $(x_n), (y_n) \in \mathscr{C}_{\mathcal{V}}$ and let $\lambda, \mu \in \mathbb{C}$. Then we define

$$a(x_n) + b(y_n) := (ax_n + by_n). \tag{51}$$

Scalar multiplication and addition as in (51) are easily seen to give $\mathscr{C}_{\mathcal{V}}$ the structure of a $\mathbb{C}$-vector space.

16.1.1 Pseudo Inner-Product

A natural contender for an inner-product on $\mathscr{C}_{\mathcal{V}}$ is the map $\langle \cdot, \cdot \rangle \colon \mathscr{C}_{\mathcal{V}} \times \mathscr{C}_{\mathcal{V}} \to \mathbb{C}$ defined by

$$\langle (x_n), (y_n) \rangle := \lim_{n\to\infty} \langle x_n, y_n \rangle \tag{52}$$

for all $(x_n), (y_n) \in \mathscr{C}_{\mathcal{V}}$. In fact, (52) will not be an inner-product, but rather a *pseudo* inner-product. Before we explain this however, let us first show that the righthand side of (52) converges in $\mathbb{C}$

Lemma 16.1. *Let (x_n) be a Cauchy sequence in $\mathcal{V}$. Then (x_n) is bounded.*

Proof. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n, m \geq N$ implies $\|x_n - x_m\| < \varepsilon$. Thus, fixing $m \in \mathbb{N}$, we see that $n \geq N$ implies

$$\|x_n\| < \|x_m\| + \varepsilon.$$

Now we let

$$M = \max\{\|x_1\|, \|x_2\|, \dots, \|x_m\| + \varepsilon\}.$$

Then M is a bound for (x_n). $\square$

Proposition 16.1. *Let (x_n) and (y_n) be Cauchy sequences of vectors in $\mathcal{V}$. Then $(\langle x_n, y_n\rangle)$ is a Cauchy sequence of complex numbers in $\mathbb{C}$. In particular, (52) converges in $\mathbb{C}$.*

Proof. Let $\varepsilon > 0$. Choose M_x and M_y such that $\|x_n\| < M_x$ and $\|y_n\| < M_y$ for all $n \in \mathbb{N}$. We can do this by Lemma (16.1). Next, choose $N \in \mathbb{N}$ such that $n, m \geq N$ implies $\|x_n - x_m\| < \frac{\varepsilon}{2M_y}$ and $\|y_n - y_m\| < \frac{\varepsilon}{2M_x}$. Then $n, m \geq M$ implies

$$\begin{aligned}
|\langle x_n, y_n\rangle - \langle x_m, y_m\rangle| &= |\langle x_n, y_n\rangle - \langle x_m, y_n\rangle + \langle x_m, y_n\rangle - \langle x_m, y_m\rangle| \\
&= |\langle x_n - x_m, y_n\rangle + \langle x_m, y_n - y_m\rangle| \\
&\leq |\langle x_n - x_m, y_n\rangle| + |\langle x_m, y_n - y_m\rangle| \\
&\leq \|x_n - x_m\|\|y_n\| + \|x_m\|\|y_n - y_m\| \\
&\leq \|x_n - x_m\|M_y + M_x\|y_n - y_m\| \\
&< \varepsilon.
\end{aligned}$$

This implies $(\langle x_n, y_n\rangle)$ is a Cauchy sequence of complex numbers in $\mathbb{C}$. The latter statement in the proposition follows from the fact that $\mathbb{C}$ is complete. $\square$

16.1.2 Quotienting Out To get an Inner-Product

As mentioned above, (52) is not an inner-product. It is what's called a pseudo inner-product:

Definition 16.1. Let V be a vector space over $\mathbb{C}$. A map $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}$ is called a **pseudo inner-product on** V if it satisfies the following properties:

1. Linearity in the first argument: $\langle x, z\rangle + \langle y, z\rangle = \langle x + y, z\rangle$ and $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$ for all $x, y, z \in V$ and $\lambda \in \mathbb{C}$.
2. Conjugate symmetric: $\langle x, y\rangle = \overline{\langle y, x\rangle}$ for all $x, y \in V$.
3. Pseudo positive definiteness: $\langle x, x\rangle \geq 0$ for all nonzero $x \in V$.

A vector space equipped with a pseudo inner-product is called a **pseudo inner-product space**.

To see why (52) is a pseudo inner-product, note that linearity in the first argument and conjugate symmetry are clear. What makes (52) a pseudo inner-product and not an inner-product is that we have pseudo positive definiteness:

$$\begin{aligned}
\langle (x_n), (x_n)\rangle &= \lim_{n\to\infty} \langle x_n, x_n\rangle \\
&= \lim_{n\to\infty} \|x_n\|^2 \\
&\geq 0.
\end{aligned}$$

In particular, we may have $\langle (x_n), (x_n)\rangle = 0$ with $(x_n) \neq 0$. To remedy this situation, we define

$$\mathscr{C}_{\mathcal{V}}^0 := \{(x_n) \in \mathscr{C}_{\mathcal{V}} \mid x_n \to 0\}.$$

Then $\mathscr{C}_{\mathcal{V}}^0$ is a subspace of $\mathscr{C}_{\mathcal{V}}$ (if $\lambda \in \mathbb{C}$ and $(x_n), (y_n) \in \mathscr{C}_{\mathcal{V}}^0$, then $(\lambda x_n + y_n) \to 0$ and hence $(\lambda x_n + y_n) \in \mathscr{C}_{\mathcal{V}}^0$). Therefore we obtain a quotient space $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$. Now we claim that the pseudo inner-product (52) induces a genuine inner-product, which we denote again by $\langle \cdot, \cdot \rangle$, on $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$, defined by

$$\langle \overline{(x_n)}, \overline{(y_n)} \rangle := \lim_{n\to\infty} \langle x_n, y_n\rangle. \tag{53}$$

for all $\overline{(x_n)}$[11] and $\overline{(y_n)}$ in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$. We need to be sure that (53) is well-defined. Let (x_n') and (y_n') be different

[11] When we write $\overline{(x_n)}$ for a coset in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$, then it is implicitly understood that (x_n) is an element $\mathscr{C}_{\mathcal{V}}$ which represents the coset $\overline{(x_n)}$ in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$.

representatives of the cosets $\overline{(x_n)}$ and $\overline{(y_n)}$ respectively (so $x_n - x'_n \to 0$ and $y_n - y'_n \to 0$). Then

$$\begin{aligned}\langle \overline{(x'_n)}, \overline{(y'_n)} \rangle &= \lim_{n\to\infty} \langle x'_n, y'_n \rangle \\ &= \lim_{n\to\infty} \langle x'_n, y'_n \rangle + \lim_{n\to\infty} \langle x_n - x'_n, y'_n \rangle + \lim_{n\to\infty} \langle x_n, y_n - y'_n \rangle \\ &= \lim_{n\to\infty} (\langle x'_n, y'_n \rangle + \langle x_n - x'_n, y'_n \rangle + \langle x_n, y_n - y'_n \rangle) \\ &= \lim_{n\to\infty} (\langle x_n, y'_n \rangle + \langle x_n, y_n - y'_n \rangle) \\ &= \lim_{n\to\infty} \langle x_n, y_n \rangle \\ &= \langle \overline{(x_n)}, \overline{(y_n)} \rangle.\end{aligned}$$

Thus (53) is well-defined (meaning it is independent of the choice of representatives of cosets).

Now linearity in the first argument of (53) and conjugate symmetry of (53) are clear. This time however, we have positive definiteness: if $\overline{(x_n)} \in \mathscr{C}_{\mathcal{V}}/\mathscr{C}^0_{\mathcal{V}}$ such that $\langle \overline{(x_n)}, \overline{(x_n)} \rangle = 0$, then

$$\begin{aligned}\lim_{n\to\infty} \langle x_n, x_n \rangle &= \langle \overline{(x_n)}, \overline{(x_n)} \rangle \\ &= 0\end{aligned}$$

implies $x_n \to 0$, which implies $\overline{(x_n)} = 0$ in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}^0_{\mathcal{V}}$.

16.1.3 The map $\iota\colon \mathcal{V} \to \mathscr{C}_{\mathcal{V}}/\mathscr{C}^0_{\mathcal{V}}$

Let $\iota\colon \mathcal{V} \to \mathscr{C}_{\mathcal{V}}/\mathscr{C}^0_{\mathcal{V}}$ be defined by

$$\iota(x) = \overline{(x)}_{n\in\mathbb{N}}$$

for all $x \in \mathcal{V}$, where $(x)_{n\in\mathbb{N}}$ is a constant sequence in $\mathscr{C}_{\mathcal{V}}$.

Definition 16.2. An **isometry** between inner-product spaces $\mathcal{V}_1$ and $\mathcal{V}_2$ is an operator $T\colon \mathcal{V}_1 \to \mathcal{V}_2$ such that

$$\langle Tx, Ty \rangle = \langle x, y \rangle$$

for all $x, y \in \mathcal{V}_1$.

Remark. Note that an isometry is automatically injective. Indeed, let $x \in \operatorname{Ker}(T)$. Then

$$\begin{aligned}\langle x, y \rangle &= \langle Tx, Ty \rangle \\ &= \langle 0, Ty \rangle \\ &= 0\end{aligned}$$

for all $y \in \mathcal{V}_1$. It follows that $x = 0$.

Proposition 16.2. *The map $\iota\colon \mathcal{V} \to \mathscr{C}_{\mathcal{V}}/\mathscr{C}^0_{\mathcal{V}}$ is an isometry.*

Proof. Linearity of ι is clear. Let $x, y \in \mathcal{V}$. Then

$$\begin{aligned}\langle \langle \iota(x), \iota(y) \rangle &:= \lim_{n\to\infty} \langle x, y \rangle \\ &= \langle x, y \rangle.\end{aligned}$$

Thus ι is an isometry. □

Proposition 16.3. *The image of $\mathcal{V}$ under ι is dense in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}^0_{\mathcal{V}}$. In other words, the closure of $\iota(\mathcal{V})$ in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}^0_{\mathcal{V}}$ is all of $\mathscr{C}_{\mathcal{V}}/\mathscr{C}^0_{\mathcal{V}}$.*

Proof. Let $\overline{(x_n)}$ be a coset in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}^0_{\mathcal{V}}$. To show that the closure of $\iota(\mathcal{V})$ is is all of $\mathscr{C}_{\mathcal{V}}/\mathscr{C}^0_{\mathcal{V}}$, we construct a sequence of cosets in $\iota(\mathcal{V})$ which converges to $\overline{(x_n)}$. Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that $n, m \geq N$ implies

$$\|x_n - x_m\| < \varepsilon/2.$$

Then $n, m \geq N$ implies

$$\begin{aligned}\|\iota(x_m) - \overline{(x_n)}\| &= \lim_{n\to\infty} \|x_m - x_n\| \\ &< \lim_{n\to\infty} \varepsilon \\ &= \varepsilon.\end{aligned}$$

Thus, $(\iota(x_m))$ is a sequence of cosets in $\iota(\mathcal{V})$ which converges to $\overline{(x_n)}$. □

16.1.4 $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$ is a Hilbert Space

Proposition 16.4. *$\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$ is a Hilbert space.*

Proof. Let $(\overline{x}^n)$ be a Cauchy sequence of cosets in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$ where

$$\overline{x}^n = \overline{(x_k^n)}_{k\in\mathbb{N}}$$

for each $n \in \mathbb{N}$. Throughout the remainder of this proof, let $\varepsilon > 0$.

Since each $x^n = (x_k^n)_{k\in\mathbb{N}}$ is a Cauchy sequence of elements in $\mathcal{V}$, there exists a $\pi(n) \in \mathbb{N}$ such that $k, l \geq \pi(n)$ implies

$$\|x_k^n - x_l^n\| < \frac{\varepsilon}{3}.$$

For each $n \in \mathbb{N}$, choose such $\pi(n) \in \mathbb{N}$ in such a way so $\pi(n) \geq \pi(m)$ whenever $n \geq m$.

Step 1: We show that the sequence $(x^n_{\pi(n)})$ of elements in $\mathcal{V}$ is a Cauchy sequence. Since $(\overline{x}^n)$ is a Cauchy sequence of cosets in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$, there exists an $N \in \mathbb{N}$ such that $n, m \geq N$ implies

$$\|(\overline{x}^n) - (\overline{x}^m)\| = \lim_{k\to\infty} \|x_k^n - x_k^m\| < \frac{\varepsilon}{4}. \tag{54}$$

Choose such an $N \in \mathbb{N}$. It follows from (54) that for each $n \geq m \geq N$, there exists $\pi(n,m) \geq \pi(n)$ such that

$$\|x_k^n - x_k^m\| < \frac{\varepsilon}{3}$$

for all $k \geq \pi(n,m)$. Choose such $\pi(n,m)$ for each $n \geq m \geq N$. Then if $n \geq m \geq N$, we have

$$\begin{aligned}
\|x^n_{\pi(n)} - x^m_{\pi(m)}\| &= \|x^n_{\pi(n)} - x^n_{\pi(n,m)} + x^n_{\pi(n,m)} - x^m_{\pi(n,m)} + x^m_{\pi(n,m)} - x^m_{\pi(m)}\| \\
&\leq \|x^n_{\pi(n)} - x^n_{\pi(n,m)}\| + \|x^n_{\pi(n,m)} - x^m_{\pi(n,m)}\| + \|x^m_{\pi(n,m)} - x^m_{\pi(m)}\| \\
&< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \\
&= \varepsilon.
\end{aligned}$$

Therefore $(x^n_{\pi(n)})$ is a Cauchy sequence of elements in $\mathcal{V}$ and hence represents a coset $\overline{(x^n_{\pi(n)})}$ in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$.

Step 2: Let $x = (x^k_{\pi(k)})$[12]. We want to show that the sequence $(\overline{x}^n)$ of cosets in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$ converges to the coset $\overline{x}$ in $\mathscr{C}_{\mathcal{V}}/\mathscr{C}_{\mathcal{V}}^0$. In particular, we need to find an $N \in \mathbb{N}$ such that $n \geq N$ implies

$$\|\overline{x}^n - \overline{x}\| = \lim_{k\to\infty} \|x_k^n - x^k_{\pi(k)}\| < \varepsilon$$

or in other words, $n \geq N$ implies

$$\|x_k^n - x^k_{\pi(k)}\| < \varepsilon$$

for all k sufficiently large.

Since x is a Cauchy sequence of elements in $\mathcal{V}$, there exists an $M \in \mathbb{N}$ such that $n, k \geq M$ implies

$$\|x^n_{\pi(n)} - x^k_{\pi(k)}\| < 2\varepsilon/3.$$

Choose such an $M \in \mathbb{N}$. Then $n \geq M$ implies

$$\begin{aligned}
\|x_k^n - x^k_{\pi(k)}\| &\leq \|x_k^n - x^n_{\pi(n)}\| + \|x^n_{\pi(n)} - x^k_{\pi(k)}\| \\
&< \frac{\varepsilon}{3} + \frac{2\varepsilon}{3} \\
&= \varepsilon.
\end{aligned}$$

for all $k \geq \max\{M, \pi(n)\}$. □

[12] Note the change in index from n to k.

17 Normed Vector Spaces

Definition 17.1. Let V be a $\mathbb{C}$-vector space. A **norm** on V is a nonnegative-valued scalar function $\|\cdot\|: V \to [0,\infty)$ such that for all $\lambda \in \mathbb{C}$ and $u, v \in V$, we have

1. (Subadditivity) $\|u+v\| \leq \|u\| + \|v\|$,
2. (Absolutely Homogeneous) $\|\lambda v\| = |\lambda|\|v\|$,
3. (Positive-Definite) $\|v\| = 0$ if and only if $v = 0$.

We call the pair $(V, \|\cdot\|)$ a **normed vector space**.

17.1 Bounded Linear Operators and Normed Vector Spaces

Definition 17.2. Let $\mathcal{V}$ and $\mathcal{W}$ be inner-product spaces. We define

$$\mathcal{B}(\mathcal{V},\mathcal{W}) := \{T: \mathcal{V} \to \mathcal{W} \mid T \text{ is a bounded linear operator}\}.$$

$\mathcal{B}(\mathcal{V},\mathcal{W})$ has the structure of a $\mathbb{C}$-vector space, where addition and scalar multiplication are defined by

$$(T+U)(x) = Tx + Ux \quad \text{and} \quad (\lambda T)(x) = T(\lambda x)$$

for all $T, U \in \mathcal{B}(\mathcal{V},\mathcal{W})$, $\lambda \in \mathbb{C}$, and $v \in \mathcal{V}$.

Proposition 17.1. *Let $\mathcal{V}$ and $\mathcal{W}$ be inner-product spaces. Then $(\mathcal{B}(\mathcal{V},\mathcal{W}), \|\cdot\|)$ is a normed vector space, where $\|\cdot\|$ is the map which sends a bounded linear operator T to its norm $\|T\|$.*

Proof. An easy exercise in linear algebra shows that $\mathcal{B}(\mathcal{V},\mathcal{W})$ has the structure of a $\mathbb{C}$-vector space, where addition and scalar multiplication are defined by

$$(T+U)(x) = T(x) + U(x) \quad \text{and} \quad (\lambda T)(x) = T(\lambda x)$$

for all $T, U \in \mathcal{B}(\mathcal{V},\mathcal{W})$, $\lambda \in \mathbb{C}$, and $v \in \mathcal{V}$. The details of this are left as an exercise. We are more interested in the fact that $\mathcal{B}(\mathcal{V},\mathcal{W})$ is a *normed* vector space. We just need to check that $\|\cdot\|$ satisfies the conditions laid out in Definition (17.1).

We first check for subadditivity. Let $T, U \in \mathcal{B}(\mathcal{V},\mathcal{W})$. Then

$$\begin{aligned} \|(T+U)(x)\| &= \|Tx + Ux\| \\ &\leq \|Tx\| + \|Ux\| \\ &\leq \|T\|\|x\| + \|U\|\|x\| \\ &= (\|T\| + \|U\|)\|x\| \end{aligned}$$

for all $x \in \mathcal{V}$. In particular, this implies $\|T+U\| \leq \|T\| + \|U\|$. Thus we have subadditivity.

Next we check that $\|\cdot\|$ is absolutely homogeneous. Let $T \in \mathcal{B}(\mathcal{V},\mathcal{W})$ and let $\lambda \in \mathbb{C}$. Then

$$\begin{aligned} \|(\lambda T)(x)\| &= \|T(\lambda x)\| \\ &= \|\lambda Tx\| \\ &= |\lambda|\|Tx\| \end{aligned}$$

for all $x \in \mathcal{V}$. In particular, this implies $\|\lambda T\| = |\lambda|\|T\|$. Thus $\|\cdot\|$ is absolutely homogeneous.

Finally we check for positive-definitiness. Let $T \in \mathcal{B}(\mathcal{V},\mathcal{W})$. Clearly $\|T\|$ is greater than or equal to 0 since it is the supremum of terms which are greater than or equal to 0. Suppose $\|T\| = 0$. Then

$$\begin{aligned} \|Tx\| &\leq \|T\|\|x\| \\ &= 0 \cdot \|x\| \\ &= 0 \end{aligned}$$

for all $x \in \mathcal{V}$. In particular, this implies $Tx = 0$ for all $x \in \mathcal{V}$ (by positive-definiteness of the norm for $\mathcal{W}$). Therefore $T = 0$ since they agree on all $x \in \mathcal{V}$. $\square$

17.2 Normed Vector Spaces Which Satisfy Parallelogram Law are Inner-Product Spaces

Proposition 17.2. *Let $(V, \|\cdot\|)$ be a normed vector space over $\mathbb{C}$ which satisfies the parallelogram law (2). Then the map $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}$ defined by*

$$\langle x, y \rangle = \frac{1}{4}\left(\|x+y\|^2 + i\|x+iy\|^2 - \|x-y\|^2 - i\|x-iy\|^2\right) \tag{55}$$

for all $x, y \in V$ is an inner-product. Moreover, the norm induced by this inner-product is precisely $\|\cdot\|$. In other words, we have

$$\langle x, x \rangle = \|x\|^2$$

for all $x \in V$.

Proof. The most difficult part of this proof is showing that (55) is linear in the first argument. Before we do this, let us show that (55) is positive-definite and conjugate-symmetric.

For positive-definiteness, let $x \in V$. Then

$$\begin{aligned}
\langle x, x \rangle &= \frac{1}{4}\left(\|2x\|^2 + i\|x+ix\|^2 - i\|x-ix\|^2\right) \\
&= \frac{1}{4}\left(\|2x\|^2 + i((|1+i|^2 - |1-i|^2)\|x\|^2)\right) \\
&= \|x\|^2 \\
&\geq 0,
\end{aligned}$$

with equality if and only if $x = 0$. Note that this also gives us $\langle x, x \rangle = \|x\|^2$ for all $x \in V$.

For conjugate-symmetry, let $x, y \in V$. Then

$$\begin{aligned}
\overline{\langle y, x \rangle} &= \overline{\frac{1}{4}(\|y+x\|^2 + i\|y+ix\|^2 - \|y-x\|^2 - i\|y-ix\|^2)} \\
&= \frac{1}{4}\left(\|y+x\|^2 - i\|y+ix\|^2 - \|y-x\|^2 + i\|y-ix\|^2\right) \\
&= \frac{1}{4}\left(\|x+y\|^2 - i\|i(x-iy)\|^2 - \|x-y\|^2 + i\|i(x+iy)\|^2\right) \\
&= \frac{1}{4}\left(\|x+y\|^2 - i\|x-iy\|^2 - \|x-y\|^2 + i\|x+iy\|^2\right) \\
&= \frac{1}{4}\left(\|x+y\|^2 + i\|x+iy\|^2 - \|x-y\|^2 - i\|x-iy\|^2\right) \\
&= \langle x, y \rangle
\end{aligned}$$

Now we come to the difficult part, namely showing that (55) is linear in the first argument. We do this in several steps:

Step 1: We show that (55) is additive in the first argument (i.e. $\langle x+z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$ for all $x, y, z \in V$). Let $x, y, z \in V$. First note that by the parallelogram law (2), we have

$$\begin{aligned}
\|x+z+y\|^2 - \|x+z-y\|^2 &= 2\|x+y\|^2 + 2\|z\|^2 - \|x+y-z\|^2 - 2\|x-y\|^2 - 2\|z\|^2 + \|x-y-z\|^2 \\
&= 2\|x+y\|^2 - 2\|x-y\|^2 - \|z-y-x\|^2 + \|z+y-x\|^2 \\
&= 2\|x+y\|^2 - 2\|x-y\|^2 - 2\|z-y\|^2 - 2\|x\|^2 + \|z-y+x\|^2 + 2\|z+y\|^2 + 2\|x\|^2 - \|z+y+x\|^2 \\
&= 2\|x+y\|^2 - 2\|x-y\|^2 + 2\|z+y\|^2 - 2\|z-y\|^2 + \|x+z-y\|^2 - \|x+z+y\|^2.
\end{aligned}$$

Adding $\|x+z-y\|^2 - \|x+z+y\|^2$ to both sides gives us

$$2(\|x+z+y\|^2 - \|x+z-y\|^2) = 2(\|x+y\|^2 - \|x-y\|^2 + \|z+y\|^2 - \|z-y\|^2),$$

and after cancelling 2 from both sides, we obtain

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

Therefore

$$\begin{aligned}
\langle x+z,y\rangle &= \frac{1}{4}\left(\|x+z+y\|^2+i\|x+z+iy\|^2-\|x+z-y\|^2-i\|x+z-iy\|^2\right)\\
&= \frac{1}{4}\left(\|x+z+y\|^2-\|x+z-y\|^2+i(\|x+z+iy\|^2-\|x+z-iy\|^2)\right)\\
&= \frac{1}{4}\left(\|x+y\|^2-\|x-y\|^2+\|z+y\|^2-\|z-y\|^2+i(\|x+iy\|^2-\|x-iy\|^2+\|z+iy\|^2-\|z-iy\|^2)\right)\\
&= \frac{1}{4}\left(\|x+y\|^2-\|x-y\|^2+\|z+y\|^2-\|z-y\|^2+i(\|x+iy\|^2-\|x-iy\|^2+\|z+iy\|^2-\|z-iy\|^2)\right)\\
&= \frac{1}{4}\left(\|x+y\|^2+i\|x+iy\|^2-\|x-y\|^2-i\|x-iy\|^2+\|z+y\|^2+i\|z+iy\|^2-\|z-y\|^2-i\|z-iy\|^2\right)\\
&= \langle x,y\rangle+\langle z,y\rangle.
\end{aligned}$$

Thus we have additivity in the first argument.

Step 2: We show that (55) respects $\mathbb{Q}$-scaling in the first argument (i.e. $\frac{m}{n}\langle x,y\rangle = \langle \frac{m}{n}x,y\rangle$ for all rational numbers $\frac{m}{n}\in\mathbb{Q}$ and for all $x,y\in V$. Let $\frac{m}{n}\in\mathbb{Q}$ and let $x,y\in V$. Then since (55) is additive in the first argument and since V is a $\mathbb{C}$-vector space, we have

$$\begin{aligned}
\frac{m}{n}\langle x,y\rangle &= \frac{m}{n}\left\langle \frac{n}{n}x,y\right\rangle\\
&= \frac{mn}{n}\left\langle \frac{1}{n}x,y\right\rangle\\
&= m\left\langle \frac{1}{n}x,y\right\rangle\\
&= \left\langle \frac{m}{n}x,y\right\rangle.
\end{aligned}$$

Therefore (55) respects $\mathbb{Q}$-scaling in the first argument.

Step 3: We show that (55) respects $\mathbb{R}$-scaling in the first argument. First note that $y\in V$, the map $\langle\cdot,y\rangle\colon V\to\mathbb{C}$ is continuous. Let $x,y\in V$ and let $r\in\mathbb{R}$. Choose a sequence (r_n) of rational numbers such that $r_n\to r$ (we can do this since $\mathbb{Q}$ is dense in $\mathbb{R}$). Then we have

$$\begin{aligned}
\langle rx,y\rangle &= \lim_{n\to\infty}\langle r_n x,y\rangle\\
&= \lim_{n\to\infty} r_n\langle x,y\rangle\\
&= r\langle x,y\rangle.
\end{aligned}$$

Therefore (55) respects $\mathbb{R}$-scaling in the first component.

Step 4: We show that (55) respects $\mathbb{C}$-scaling in the first component. We first show that $\langle ix,y\rangle = i\langle x,y\rangle$ for all $x,y\in V$.

Let $x,y\in V$. Then

$$\begin{aligned}
\langle ix,y\rangle &= \frac{1}{4}\left(\|ix+y\|^2+i\|ix+iy\|^2-\|ix-y\|^2-i\|ix-iy\|^2\right)\\
&= \frac{1}{4}\left(\|x-iy\|^2+i\|x+y\|^2-\|x+iy\|^2-i\|x-y\|^2\right)\\
&= \frac{1}{4}\left(i\|x+y\|^2-\|x+iy\|^2-i\|x-y\|^2+\|x-iy\|^2\right)\\
&= \frac{i}{4}\left(\|x+y\|^2+i\|x+iy\|^2-\|x-y\|^2-i\|x-iy\|^2\right)\\
&= i\langle x,y\rangle.
\end{aligned}$$

Now let $\lambda = r+is\in\mathbb{C}$. Then we have

$$\begin{aligned}
\langle \lambda x,y\rangle &= \langle (r+is)x,y\rangle\\
&= \langle rx+isx,y\rangle\\
&= \langle rx,y\rangle+\langle isx,y\rangle\\
&= r\langle x,y\rangle+s\langle ix,y\rangle\\
&= r\langle x,y\rangle+is\langle x,y\rangle\\
&= (r+is)\langle x,y\rangle\\
&= \lambda\langle x,y\rangle
\end{aligned}$$

for all $x,y\in V$. Therefore (55) respects $\mathbb{C}$-scaling in the first component. □

17.3 Distances and Pseudo Normed Vector Spaces

Let $\mathcal{V}$ be an inner-product space and let $\mathcal{A}$ be a subspace of $\mathcal{V}$. We define

$$d(x, \mathcal{A}) = \inf\{\|x - a\| \mid a \in A\} \tag{56}$$

for all $x \in \mathcal{V}$. The map $d(-, \mathcal{A})$ is a good candidate for a norm on $\mathcal{V}$. It turns out however that $d(-, \mathcal{A})$ is just a **pseudo norm**.

Definition 17.3. Let V be a vector space over $\mathbb{C}$. A map $p\colon V \to [0, \infty)$ is called a **pseudo norm on** V if it satisfies the following properties:

1. (Subadditivity) $p(u + v) \leq p(u) + p(v)$ for all $u, v \in V$,
2. (Absolutely Homogeneous) $p(\lambda v) = |\lambda| p(v)$ for all $v \in V$ and $\lambda \in \mathbb{C}$.

A vector space equipped with a pseudo inner-product is called a **pseudo normed vector space**.

Remark. Thus a norm is just a pseudo norm with the positive-definiteness property.

17.3.1 Absolute Homogeneity of Distances

Proposition 17.3. *Let $\mathcal{V}$ be an inner-product space, let $\mathcal{A}$ be a subspace of $\mathcal{V}$, let $x \in \mathcal{V}$, and let $\lambda \in \mathbb{C}$. Then*

$$d(\lambda x, \mathcal{A}) = |\lambda| d(x, \mathcal{A}).$$

Proof. Choose a sequence (y_n) of elements in $\mathcal{A}$ such that

$$\|x - y_n\| < d(x, \mathcal{A}) + \frac{1}{|\lambda n|}$$

for all $n \in \mathbb{N}$. Then since $\mathcal{A}$ is a subspace, we have

$$\begin{aligned} d(\lambda x, \mathcal{A}) &\leq \|\lambda x - \lambda y_n\| \\ &= |\lambda| \|x - y_n\| \\ &< |\lambda| \left(d(x, \mathcal{A}) + \frac{1}{|\lambda n|} \right) \\ &= |\lambda| d(x, \mathcal{A}) + \frac{1}{n} \end{aligned}$$

for all $n \in \mathbb{N}$. In particular, this implies $d(\lambda x, \mathcal{A}) \leq |\lambda| d(x, \mathcal{A})$.

Conversely, choose a sequence (z_n) of elements in $\mathcal{A}$ such that

$$\|\lambda x - z_n\| < d(\lambda x, \mathcal{A}) + \frac{1}{n}$$

Then since $\mathcal{A}$ is a subspace, we have

$$\begin{aligned} |\lambda| d(x, \mathcal{A}) &\leq |\lambda| \|x - z_n / |\lambda|\| \\ &= \|\lambda x - z_n\| \\ &< d(\lambda x, \mathcal{A}) + \frac{1}{n} \end{aligned}$$

for all $n \in \mathbb{N}$. In particular, this implies $|\lambda| d(x, \mathcal{A}) \leq d(\lambda x, \mathcal{A})$. □

17.3.2 Subadditivity of Distances

Proposition 17.4. *Let $\mathcal{V}$ be an inner-product space, let $\mathcal{A}$ be a subspace of $\mathcal{V}$, and let $x, y \in \mathcal{V}$. Then*

$$d(x + y, \mathcal{A}) \leq d(x, \mathcal{A}) + d(y, \mathcal{A}).$$

Proof. Choose a sequences (w_n) and (z_n) of elements in $\mathcal{A}$ such that

$$\|x - w_n\| < d(x, \mathcal{A}) + \frac{1}{2n} \quad \text{and} \quad \|y - z_n\| < d(y, \mathcal{A}) + \frac{1}{2n}$$

for all $n \in \mathbb{N}$. Then since $\mathcal{A}$ is a subspace, we have

$$\begin{aligned} d(x+y, \mathcal{A}) &\leq \|(x+y) - (w_n + z_n)\| \\ &\leq \|x - w_n\| + \|y - z_n\| \\ &< d(x, \mathcal{A}) + \frac{1}{2n} + d(y, \mathcal{A}) + \frac{1}{2n} \\ &= d(x, \mathcal{A}) + d(y, \mathcal{A}) + \frac{1}{n} \end{aligned}$$

for all $n \in \mathbb{N}$. In particular, this implies $d(x+y, \mathcal{A}) \leq d(x, \mathcal{A}) + d(y, \mathcal{A})$. □

17.3.3 Quotienting out to get a Norm

To see why (56) is just a pseudo norm and not a norm, note that $d(x, \mathcal{A}) = 0$ if and only if $x \in \overline{\mathcal{A}}$. To remedy this situation, we quotient out by $\overline{\mathcal{A}}$. First we need a lemma.

Lemma 17.1. *Let $x \in \mathcal{V}$. Then $d(x, \mathcal{A}) = d(x, \overline{\mathcal{A}})$.*

Proof. We have $d(x, \mathcal{A}) \geq d(x, \overline{\mathcal{A}})$ since $\mathcal{A} \subseteq \overline{\mathcal{A}}$. For the reverse inequality, we assume (for a contradiction) that $d(x, \mathcal{A}) > d(x, \overline{\mathcal{A}})$. For the reverse inequality, let $\varepsilon > 0$. Choose $a \in \overline{\mathcal{A}}$ such that

$$\|x - a\| < d(x, \overline{\mathcal{A}}) + \varepsilon/2.$$

Choose $b \in \mathcal{A}$ such that $\|a - b\| < \varepsilon/2$. Then

$$\begin{aligned} d(x, \mathcal{A}) &\leq \|x - b\| \\ &\leq \|x - a\| + \|a - b\| \\ &< d(x, \overline{\mathcal{A}}) + \varepsilon/2 + \varepsilon/2 \\ &= d(x, \overline{\mathcal{A}}) + \varepsilon. \end{aligned}$$

Since this holds for all $\varepsilon > 0$, we have $d(x, \mathcal{A}) \leq d(x, \overline{\mathcal{A}})$. □

Proposition 17.5. *The pseudo norm $d(-, \mathcal{A})$ on $\mathcal{V}$ induces a well-defined norm $\|\cdot\|$ on $\mathcal{V}/\overline{\mathcal{A}}$, defined by*

$$\|\overline{x}\| = d(x, \mathcal{A}) \tag{57}$$

for all $\overline{x} \in \mathcal{V}/\overline{\mathcal{A}}$.

Proof. We first check that (57) is well-defined. Let $\overline{x} \in \mathcal{V} \backslash \overline{\mathcal{A}}$[13]. Choose another representative of the coset $\overline{x}$, say $x + a$ where $a \in \overline{\mathcal{A}}$. Then

$$\begin{aligned} \|\overline{x+a}\| &= d(x+a, \mathcal{A}) \\ &= d(x+a, \overline{\mathcal{A}}) \\ &= \inf\{\|x + a - b\| \mid b \in \overline{\mathcal{A}}\} \\ &= \inf\{\|x - c\| \mid c \in \overline{\mathcal{A}}\} \\ &= d(x, \mathcal{A}) \\ &= \|\overline{x}\|. \end{aligned}$$

Thus $\|\cdot\|$ is well-defined.

Now $\|\cdot\|$ is a pseudo norm on $\mathcal{V} \backslash \overline{\mathcal{A}}$ since it inherits these properties from $d(-, \mathcal{A})$ on $\mathcal{V}$. Indeed, for subadditivity, we have

$$\begin{aligned} \|\overline{x} + \overline{y}\| &= d(x+y, \mathcal{A}) \\ &\leq d(x, \mathcal{A}) + d(y, \mathcal{A}) \\ &= \|\overline{x}\| + \|\overline{y}\| \end{aligned}$$

for all $\overline{x}, \overline{y} \in \mathcal{V} \backslash \overline{\mathcal{A}}$, and for absolute homogeneity, we have

$$\begin{aligned} \|\lambda \overline{x}\| &= d(\lambda x, \mathcal{A}) \\ &= |\lambda| d(x, \mathcal{A}) \\ &= |\lambda| \|\overline{x}\| \end{aligned}$$

[13] Do not confuse the overline over $\mathcal{A}$ with the overline over x. One denotes the closure of $\mathcal{A}$ and the other denotes a coset in $\mathcal{V}/\overline{\mathcal{A}}$ with a given represenative $x \in \mathcal{V}$.

for all $\overline{x} \in \mathcal{V} \backslash \overline{\mathcal{A}}$ and $\lambda \in \mathbb{C}$.

Moreover $\|\cdot\|$ is a norm on $\mathcal{V} \backslash \overline{\mathcal{A}}$ since we also have positive-definiteness. Indeed, let $\overline{x} \in \mathcal{V}/\overline{\mathcal{A}}$. Then

$$\begin{aligned} \overline{x} = 0 &\iff x \in \overline{\mathcal{A}} \\ &\iff d(x, \mathcal{A}) = 0 \\ &\iff \|\overline{x}\| = 0. \end{aligned}$$

□

18 Duality

Let K be a field, let V be a K-vector space with basis $\beta = \{\beta_1, \dots, \beta_m\}$, and let W be a K-vector space with basis $\gamma = \{\gamma_1, \dots, \gamma_n\}$. Recall from linear algebra that we define the **(algebraic) dual** of V to be the K-vector space

$$V^\star := \{\varphi\colon V \to K \mid \varphi \text{ is linear}\}.$$

where addition and scalar multiplication are defined by

$$(\varphi + \psi)(v) = \varphi(v) + \psi(v) \quad \text{and} \quad (\lambda\varphi)(v) = \varphi(\lambda v)$$

for all $\varphi, \psi \in V^\star$, $\lambda \in \mathbb{C}$, and $v \in V$. The **(algebraic) dual** of β is defined to be the basis of $V^\star$ given by $\beta^\star := \{\beta_1^\star, \dots, \beta_m^\star\}$, where each $\beta_i^\star$ is uniquely determined by

$$\beta_i^\star(\beta_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

If $T\colon V \to W$ is a linear map, we define its **(algebraic) dual** to be the linear map $T^\star\colon W^\star \to V^\star$ defined by

$$T^\star(\varphi) = \varphi \circ T$$

for all $\varphi \in W^\star$. One learns in linear algebra that the transpose of the matrix representation of T with respect to the bases β and γ is equal to the matrix representation of $T^\star$ with respect to the bases $\gamma^\star$ and $\beta^\star$. In terms of notation, this is written as

$$([T]_\beta^\gamma)^\top = [T^\star]_{\gamma^\star}^{\beta^\star}.$$

We want to describe an analog of this situation for inner-product spaces over $\mathbb{C}$.

Definition 18.1. Let $\mathcal{V}$ be an inner-product space over $\mathbb{C}$. We define its **(continuous) dual space**[a] to be

$$\begin{aligned} \mathcal{V}^* &:= \{\ell\colon \mathcal{V} \to \mathbb{C} \mid \ell \text{ is linear and continuous}\}. \\ &= \{\ell\colon \mathcal{V} \to \mathbb{C} \mid \ell \text{ is a bounded operator}\}. \end{aligned}$$

[a] When speaking about the dual space of an inner-product space, we will always mean its continuous dual.

Remark. Thus $\mathcal{V}^*$ captures both topological and linear aspects of $\mathcal{V}$.

18.1 Riesz Representation Theorem Revisited

Theorem 18.1. *(Riesz Representation Theorem) Let $\mathcal{H}$ be a Hilbert space. Then there exists an isometric antiisomorphism $\Phi\colon \mathcal{H} \to \mathcal{H}^*$.*

Proof. Define $\Phi\colon \mathcal{H} \to \mathcal{H}^*$ by

$$\Phi(x) = \langle \cdot, x \rangle$$

for all $x \in \mathcal{H}$. We first show Φ is antilinear. Let $x, y \in \mathcal{H}$ and let $\lambda, \mu \in \mathbb{C}$. Then

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

for all $z \in \mathcal{H}$. Therefore Φ is antilinear.

Note Φ is an injective antilinear map since the inner-product is positive-definite. Also, the Riesz representation theorem implies Φ is surjective. Finally Example (**??**) implies Φ is an isometry. Therefore Φ is an isometric antiisomorphism. □

19 Limit Infimum

Let (a_n) be a sequence of positive real numbers. We define the **limit infimum** of (a_n), denoted $\liminf(a_n)$, to be the limit

$$\liminf(a_n) := \lim_{N\to\infty}(\inf\{a_n \mid n \geq N\}). \tag{58}$$

Since the sequence $(\inf\{a_n \mid n \geq N\})_{N\in\mathbb{N}}$ is a monotone increasing sequence in N, the limit (58) always exists or equals $-\infty$.

Proposition 19.1. *Let (a_n) be a sequence of positive real-valued numbers.*

1. *If $\liminf(a_n) = A$, then for each $\varepsilon > 0$ and $N \in \mathbb{N}$, there exists some $n \geq N$ such that $a_n > A + \varepsilon$. In other words, for all $\varepsilon > 0$, the sequence (a_n) is frequently strictly less than $A + \varepsilon$.*
2. *If $\liminf(a_n) = A$, then for each $\varepsilon > 0$, there exists $N \in N$ such that $a_n > A - \varepsilon$ for all $n \geq N$. In other words, for all $\varepsilon > 0$, the sequence (a_n) is eventually strictly greater than $A - \varepsilon$.*
3. *Conversely, if $A \geq 0$ satisfies 1 and 2, then liminf$(a_n) = A$.*

Proof.

1. We prove this by contradiction. To obtain a contradiction, assume that there there exists $\varepsilon > 0$ and $N \in \mathbb{N}$ such that there does not exist any $n \geq N$ with $a_n < A + \varepsilon$. Choose such $\varepsilon > 0$ and $N \in \mathbb{N}$. Then $a_n \geq A + \varepsilon$ for all $n \geq N$. This implies $\inf\{a_n \mid n \geq N\} > A + \varepsilon$. Since $\inf\{a_n \mid n \geq N\}$ is a monotone increasing function of N, this implies $\liminf(a_n) \geq A + \varepsilon$, which is a contradiction.

2. We prove this by contradiction. To obtain a contradiction, assume that there exists $\varepsilon > 0$ such that there does not exist an $N \in \mathbb{N}$ with $a_n > A - \varepsilon$ for all $n \geq N$. Choose such $\varepsilon > 0$ and let $N \in \mathbb{N}$. Then there exists $n \geq N$ such that $a_n \leq A - \varepsilon$. In particular, this implies

$$\inf\{a_n \mid n \geq N\} \leq A - \varepsilon$$

Since N is arbitrary, this further implies

$$\begin{aligned}\liminf(a_n) &= \lim_{N\to\infty}(\inf\{a_n \mid n \geq N\})\\ &\leq A - \varepsilon,\end{aligned}$$

which contradicts the fact that $\liminf(a_n) = A$.

3. Let $A \geq 0$ satisfy 1 and 2 and let $A' = \liminf(a_n)$. We prove by contradiction that $A = A'$. Assume for a contradiction that $A < A'$. Let $\varepsilon = A' - A$. Since A' satisfies 2, the sequence (a_n) is eventually greater than $A' - \varepsilon/2$. On the other hand, since A satisfies 1, the sequence (a_n) is frequently less than $A + \varepsilon/2 = A' - \varepsilon/2$. This is a contradiciton. An analagous argument gives a contradiction when we assume $A > A'$. Therefore $A = A'$. □

Lemma 19.1. *Let (a_n) and (b_n) be two sequences of positive real numbers such that $\liminf(a_n) = A$ and $\liminf(b_n) = B$. Then*

1. $\liminf(a_n b_n) = AB$
2. $\liminf(a_n + b_n) = A + B$

Proof.

1. Let $\varepsilon > 0$. Since the sequence (a_n) is eventually greater than $A - \varepsilon/(A + B)$ and since the sequence (b_n) is eventually greater than $B - \varepsilon/(A + B)$, the sequence $(a_n b_n)$ is eventually greater than

$$\begin{aligned}\left(A - \frac{\varepsilon}{A+B}\right)\left(B - \frac{\varepsilon}{A+B}\right) &= AB - \frac{\varepsilon(A+B)}{(A+B)} + \frac{\varepsilon^2}{(A+B)^2}\\ &\geq AB - \varepsilon.\end{aligned}$$

An analagous argument shows that $(a_n b_n)$ is frequently less than $AB + \varepsilon$.

2. This is proved in the same way as 1. □