GOALS PREREQUISITE NOTES

These notes represent a crash course in functional analysis, focusing on ideas and concepts that will be referenced during the GOALS mini-courses. While it is likely you have been exposed to some of this material in a graduate analysis course, for the purposes of GOALS there is no expectation that you have seen all of it before. Moreover, you are **not** expected to master all of the concepts presented here. Instead, we ask only that you familiarize yourself with these notes enough to not be caught off guard when these concepts arise in the mini-courses. We encourage you to post questions in the Piazza forums, or email the organizers directly.

Many of the results below are presented without proofs. In most cases, proving them yourself would be an excellent exercise. Detailed proofs can also be found in most standard textbooks on functional analysis. The presentation below is based on this MSU Functional Analysis Couse, which followed John B. Conway's *A Course in Functional Analysis*, Second Edition (Springer, 1990).

Throughout the notes, all vector spaces are taken to be over $\mathbb{C}$.

Contents

1. Hilbert Spaces

Definition 1.1. For a vector space V, an **inner product** is a map $\langle \cdot , \cdot \rangle : V \times V \to \mathbb{C}$ satisfying:

1. $\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z}\rangle = a\langle \mathbf{x}, \mathbf{z}\rangle + b\langle \mathbf{y}, \mathbf{z}\rangle$ for all $a, b \in \mathbb{C}$ and $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ (*linearity*);
2. $\langle \mathbf{x}, \mathbf{y}\rangle = \overline{\langle \mathbf{y}, \mathbf{x}\rangle}$ for all $\mathbf{x}, \mathbf{y} \in V$ (*conjugate symmetry*);
3. $\langle \mathbf{x}, \mathbf{x}\rangle \geq 0$ for all $\mathbf{x} \in V$ (*positive semi-definiteness*);
4. $\langle \mathbf{x}, \mathbf{x}\rangle = 0$ if and only if $\mathbf{x} = 0$. (*non-degeneracy*).

A **degenerate inner product** is a map satisfying (1)-(3), but not (4). That is, there exists a non-zero $\mathbf{x} \in V$ with $\langle \mathbf{x}, \mathbf{x} \rangle = 0$.

An **inner product space** is a pair $(V, \langle \cdot\,, \cdot \rangle)$ consisting of a vector space and a (non-degenerate) inner product on V.

Example 1.2.

(1) $\mathbb{C}^n$ has an inner product defined by
$$\langle (a_1, \ldots, a_n), (b_1, \ldots, b_n) \rangle := \sum_{j=1}^{n} a_j \overline{b_j}$$

(2) $M_{m\times n}(\mathbb{C})$ has an inner product defined by
$$\langle A, B \rangle := \mathrm{Tr}(AB^*)$$
where B^* is the conjugate transpose (*adjoint*) of B.

(3) Let (X, μ) be a measure space with positive measure μ. Then $L^2(X, \mu)$ has the inner product
$$\langle f, g \rangle := \int_X f(x)\overline{g(x)}\ d\mu(x).$$
In particular, if X is an countable set equipped with the counting measure μ, then $L^2(X, \mu)$ can be identified with
$$\ell^2(X) := \{(a_x)_{x\in X} \in \mathbb{C}^X : \sum_{x\in X} |a_x|^2 < \infty\},$$
which has the inner product
$$\langle (a_x)_{x\in X}, (b_x)_{x\in X} \rangle = \sum_{x\in X} a_x \overline{b_x}.$$

■

Exercise 1.3. Let $(V, \langle \cdot\,, \cdot \rangle)$ be an inner product space. Show that it satisfies *conjugate linearity*:
$$\langle \mathbf{x}, a\mathbf{y} + b\mathbf{z} \rangle = \bar{a} \langle \mathbf{x}, \mathbf{y} \rangle + \bar{b} \langle \mathbf{x}, \mathbf{z} \rangle \qquad a, b \in \mathbb{C},\ \mathbf{x}, \mathbf{y}, \mathbf{z} \in V,$$

Given an inner product space $(V, \langle \cdot\,, \cdot \rangle)$, consider the map $\|\cdot\|\colon V \to [0, \infty)$ defined by
$$\|\mathbf{x}\| := (\langle \mathbf{x}, \mathbf{x} \rangle)^{1/2} \qquad \mathbf{x} \in V.$$
This defines a *norm* on the vector space V (see Definition 2.1).

Exercise 1.4. Let $(V, \langle \cdot\,, \cdot \rangle)$ be an inner product space. Show that $\|a\mathbf{x}\| = |a|\|\mathbf{x}\|$ for all $a \in \mathbb{C}$ and $\mathbf{x} \in V$.

Note that the non-degeneracy of the inner product implies $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0$. Moreover, the previous exercise implies $\|\mathbf{x} - \mathbf{y}\| = \|\mathbf{y} - \mathbf{x}\|$ for all $\mathbf{x}, \mathbf{y} \in V$. Thus, up to establishing the triangle inequality, we have shown that $d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$ defines a metric on V. This will be a corollary of the next result.

Theorem 1.5 (Cauchy–Schwarz Inequality). *Let $(V, \langle \cdot\,, \cdot \rangle)$ be an inner product space. For any $\mathbf{x}, \mathbf{y} \in V$ we have*
$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|.$$
Moreover, if the above is an equality then one of $\mathbf{x}$ or $\mathbf{y}$ is in the linear span of the other.

Corollary 1.6 (Triangle Inequality). *$(V, \langle \cdot\,, \cdot \rangle)$ be an inner product space. Then for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ we have*
$$\|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{z}\| + \|\mathbf{z} - \mathbf{y}\|.$$
Consequently, $(\mathbf{x}, \mathbf{y}) \mapsto \|\mathbf{x} - \mathbf{y}\|$ defines a metric on V.

Exercise 1.7. Use the Cauchy–Schwarz inequality to prove the Triangle Inequality.

Recall that a metric space (X, d) is *complete* when all Cauchy sequences converge.

Definition 1.8. A **Hilbert space** is an inner product space which is complete with respect to the metric induced by its norm.

All of the examples in Example 1.2 are complete and hence Hilbert spaces. For a non-example, consider $(\mathbb{Q}+i\mathbb{Q})^n$ equipped with the inner product it inherits as a subspace from $\mathbb{C}^n$. It is not complete and hence not a Hilbert space, but its completion is $\mathbb{C}^n$.

A more sophisticated non-example comes from $C(0,1)$, the space of continuous functions on the interval $(0,1)$, which we equip with the inner product:

$$\langle f, g\rangle := \int_0^1 f(t)\overline{g(t)}\, dt.$$

Observe that this is simply the inner product it inherits as a subspace of $L^2(0,1)$. Since $C(0,1)$ is a proper dense subspace of $L^2(0,1)$, it follows that $C(0,1)$ is not a Hilbert space with this inner product, but its completion is $L^2(0,1)$.

Exercise 1.9. Let $(V, \langle\,\cdot\,,\,\cdot\,\rangle)$ be an inner product space. Show that if $(\mathbf{x}_n)_{n\in\mathbb{N}}, (\mathbf{y}_n)_{n\in\mathbb{N}} \subset V$ are Cauchy sequences then

$$\lim_{n\to\infty} \langle \mathbf{x}_n, \mathbf{y}_n\rangle$$

exists. Use this to define an inner product on the completion of V (as a metric space) making it into a Hilbert space.

Exercise 1.10. Let $\mathcal{H}$ be a Hilbert space and define $\bar{\mathcal{H}} := \{\bar{\mathbf{x}} \colon \mathbf{x} \in \mathcal{H}\}$; that is, $\bar{\mathcal{H}}$ as equivalent to $\mathcal{H}$ as a set but with elements decorated with formal '$-$' notation. Show that if we equip $\bar{\mathcal{H}}$ with the vector space operations

$$\bar{\mathbf{x}} + \bar{\mathbf{y}} = \overline{\mathbf{x}+\mathbf{y}} \qquad \mathbf{x}, \mathbf{y} \in \mathcal{H}$$
$$a\bar{\mathbf{x}} = \overline{\bar{a}\mathbf{x}} \qquad a \in \mathbb{C}$$

and inner product

$$\langle \bar{\mathbf{x}}, \bar{\mathbf{y}}\rangle := \langle \mathbf{y}, \mathbf{x}\rangle\,,$$

then $\bar{\mathcal{H}}$ is a Hilbert space. $\bar{\mathcal{H}}$ is called the **conjugate** Hilbert space to $\mathcal{H}$.

1.1 Orthogonality and Convexity

Definition 1.11. Let $(V, \langle\,\cdot\,,\,\cdot\,\rangle)$ be an inner product space. We say $\mathbf{x}, \mathbf{y} \in V$ are **orthogonal** and write $\mathbf{x} \perp \mathbf{y}$ if $\langle \mathbf{x}, \mathbf{y}\rangle = 0$. We say two subsets $X, Y \subset V$ are **orthogonal** and write $X \perp Y$ if $\mathbf{x} \perp \mathbf{y}$ for all $\mathbf{x} \in X$ and $\mathbf{y} \in Y$.

For $X \subset V$, the **orthogonal complement** of X is the set

$$X^\perp := \{\mathbf{y} \in V \colon \mathbf{y} \perp \mathbf{x}\ \forall \mathbf{x} \in X\}.$$

Observe that by the linearity (or conjugate linearity) of the inner product, $X^\perp$ is always a subspace of V (even if X isn't). Moreover, the Cauchy–Schwarz inequality implies it is always a closed subspace. It is also easy to check that $V^\perp = \{0\}$ (that is, the only vector orthogonal to every other vector is zero) and $\{0\}^\perp = V$.

Exercise 1.12. For $\mathbf{x}, \mathbf{y} \in V$ orthogonal, establish the *Pythagorean theorem*

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

One very important aspect of Hilbert spaces is how they interact with *convex* subsets.

Definition 1.13. Let V be a vector space. We say a subset $K \subset V$ is **convex** if for every $\mathbf{x}, \mathbf{y} \in K$ one has $t\mathbf{x} + (1-t)\mathbf{y} \subset K$ for all $0 \leq t \leq 1$.

Theorem 1.14. *Let $\mathcal{H}$ be a Hilbert space and let $K \subset \mathcal{H}$ be closed, convex, and non-empty. Then for any $\mathbf{x} \in \mathcal{H}$ there is a unique $\mathbf{y}_0 \in K$ satisfying*

$$\|\mathbf{x} - \mathbf{y}_0\| = \inf_{y\in K} \|\mathbf{x} - \mathbf{y}\| \qquad (= \mathit{dist}(\mathbf{x}, K)).$$

When K in the above theorem is a subspace, we can achieve a stronger conclusion:

Theorem 1.15. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{K} \subset \mathcal{H}$ be a closed non-empty subspace. For $\mathbf{x} \in \mathcal{H}$, $\mathbf{y}_0 \in \mathcal{K}$ is the unique vector satisfying*

$$\|\mathbf{x} - \mathbf{y}_0\| = \inf_{\mathbf{y} \in \mathcal{K}} \|\mathbf{x} - \mathbf{y}\|$$

if and only if $\mathbf{x} - \mathbf{y}_0 \in \mathcal{K}^\perp$.

Given a closed subspace $\mathcal{K} \subset \mathcal{H}$, define $P_\mathcal{K} \colon \mathcal{H} \to \mathcal{H}$ by letting $P_\mathcal{K}\mathbf{x}$ be the unique $y_0 \in \mathcal{K}$ which minimizes the distance from $\mathbf{x}$ to $\mathcal{K}$. Then $P_\mathcal{K}$ is linear (**Exercise:** check this) and the previous theorem implies $\mathbf{x} - P_\mathcal{K}\mathbf{x} \in \mathcal{K}^\perp$. Consequently, every $\mathbf{x} \in \mathcal{H}$ can be written uniquely as a sum of vectors in $\mathcal{K}$ and $\mathcal{K}^\perp$:

$$\mathbf{x} = P_\mathcal{K}\mathbf{x} + (\mathbf{x} - P_\mathcal{K}\mathbf{x}).$$

Definition 1.16. The linear operator $P_\mathcal{K}$ defined above is called the (**orthogonal**) **projection** onto $\mathcal{K}$.

Exercise 1.17. For a closed subspace $\mathcal{K} \subset \mathcal{H}$, show

(a) $\mathbf{x} \in \mathcal{K}$ if and only if $\mathbf{x} = P_\mathcal{K}\mathbf{x}$.
(b) $\mathbf{x} \in \mathcal{K}^\perp$ if and only if $P_\mathcal{K}\mathbf{x} = 0$.
(c) $\|P_\mathcal{K}\mathbf{x}\| \leq \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathcal{H}$.

Exercise 1.18. Let $\mathcal{H}$ be a Hilbert space. For any $X \subset \mathcal{H}$, show that $(X^\perp)^\perp$ equals $\overline{span}X$, the closure of the span of X.

1.2 Orthonormal Bases and Dimension

Definition 1.19. Let $\mathcal{H}$ be a Hilbert space. A subset $\mathcal{F} \subset \mathcal{H}$ is called **orthonormal** if for all $\mathbf{x}, \mathbf{y} \in \mathcal{F}$ we have

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0 \qquad \text{and} \qquad \|\mathbf{x}\| = 1$$

We say $\mathcal{E} \subset \mathcal{H}$ is an **orthonormal basis** for $\mathcal{H}$ if it is an orthonormal set satisfying $\overline{\text{span}}\,\mathcal{E} = \mathcal{H}$.

Just as every vector space admits a basis (i.e. a linearly independent spanning set) every Hilbert space admits an orthonormal basis. In fact, one can characterize an orthonormal basis as a *maximal* orthonormal set, and consequently the existence of an orthonormal basis can be shown using Zorn's lemma.

Example 1.20.

(1) In $\mathbb{C}^n$,

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

is an orthonormal basis.

(2) Let $E_{i,j} \in M_{m \times n}(\mathbb{C})$ with 1 in the (i,j) entry and zeros elsewhere. Then $\{E_{i,j} : 1 \leq i \leq m,\ 1 \leq j \leq n\}$ is an orthonormal basis.

(3) In $L^2([0,1], m)$, define $f_n(t) := e^{2\pi i n t}$ for each $n \in \mathbb{Z}$. Then $\{f_n : n \in \mathbb{Z}\}$ is orthonormal (**Exercise:** check this). Moreover, Fourier analysis tells us that it is actually an orthonormal basis. ■

Exercise 1.21. Let $\mathcal{E} \subset \mathcal{H}$ be an orthonormal set in a Hilbert space. Show that $\mathcal{E}$ is an orthonormal basis for $\mathcal{H}$ if and only if $\mathcal{E}^\perp = \{0\}$.

Theorem 1.22. *Let $\mathcal{H}$ be a Hilbert space with orthonormal basis $\mathcal{E}$. Then for all $\mathbf{x} \in \mathcal{H}$ one has*

$$\mathbf{x} = \sum_{\mathbf{e} \in \mathcal{E}} \langle \mathbf{x}, \mathbf{e} \rangle \, \mathbf{e} \qquad \textit{and} \qquad \|\mathbf{x}\|^2 = \sum_{\mathbf{e} \in \mathcal{E}} |\langle \mathbf{x}, \mathbf{e} \rangle|^2.$$

In particular, $\langle \mathbf{x}, \mathbf{e} \rangle = 0$ for all but countably many $\mathbf{e} \in \mathcal{E}$.

A vector space V does not have a unique basis, but any two bases have the same size. The same holds true for Hilbert spaces:

Theorem 1.23. *If $\mathcal{H}$ is a Hilbert space then any two orthonormal bases for $\mathcal{H}$ have the same cardinality.*

Definition 1.24. For a Hilbert space $\mathcal{H}$, the **dimension** of $\mathcal{H}$, denoted $\mathbf{dim}\,\mathcal{H}$, is the cardinality of any orthonormal basis for $\mathcal{H}$.

Moreover, just as any two vector spaces of the same dimension are isomorphic, we have the following.

Theorem 1.25. *If $\mathcal{H}, \mathcal{K}$ are Hilbert spaces and $dim(\mathcal{H}) = dim(\mathcal{K})$, then $\mathcal{H}, \mathcal{K}$ are isomorphic in the sense that there exist a bijection $U \in B(\mathcal{H}, \mathcal{K})$ satisfying*

$$\langle U\mathbf{x}, U\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$$

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

Exercise 1.26. Show that $U \in B(\mathcal{H}\,\mathcal{K})$ satisfies

$$\langle U\mathbf{x}, U\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$$

for all $\mathbf{x}, \mathbf{y} \in \mathcal{H}$ if and only if $\|U\mathbf{x}\| = \|\mathbf{x}\|$. Thus, in this case U is automatically injective.

Recall that a metric space is said to be *separable* if it admits a countable dense set.

Theorem 1.27. *A Hilbert space $\mathcal{H}$ is separable if and only if* $\dim(\mathcal{H})$ *is countable.*

1.3 Direct Sums and Tensor Products

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. The **direct sum** of $\mathcal{H}_1$ and $\mathcal{H}_2$ is the vector space

$$\mathcal{H}_1 \oplus H_2 := \{(\mathbf{x}, \mathbf{y}) : \mathbf{x} \in \mathcal{H}_1, \mathbf{y} \in \mathcal{H}_2\}$$

with operations

$$\alpha(\mathbf{x}_1, \mathbf{y}_1) + (\mathbf{x}_2, \mathbf{y}_2) = (\alpha\mathbf{x}_1 + \mathbf{x}_2, \alpha\mathbf{y}_1 + \mathbf{y}_2).$$

This is naturally a Hilbert space with the inner product

$$\langle(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2)\rangle := \langle \mathbf{x}_1, \mathbf{x}_2\rangle + \langle \mathbf{y}_1, \mathbf{y}_2\rangle$$

Note that

$$\|(\mathbf{x}, \mathbf{y})\| = \left(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\right)^{1/2}.$$

More generally, if $\{\mathcal{H}_i\}_{i\in I}$ is a family of Hilbert spaces indexed by a countable set I, we define the direct sum as the vector space

$$\bigoplus_{i\in I} \mathcal{H}_i := \left\{ (\mathbf{x}_i)_{i\in I} : \mathbf{x}_i \in \mathcal{H}_i, \sum_{i\in I} \|\mathbf{x}_i\|^2 < \infty \right\}.$$

with operations

$$\alpha(\mathbf{x}_i)_{i\in I} + (\mathbf{y}_i)_{i\in I} = (\alpha\mathbf{x}_i + \mathbf{y}_i)_{i\in I}$$

and inner product

$$\langle(\mathbf{x}_i)_{i\in I}, (\mathbf{y}_i)_{i\in I}\rangle = \sum_{i\in I} \langle \mathbf{x}_i, \mathbf{y}_i\rangle.$$

Exercise 1.28. Verify that $\bigoplus_{i\in I} \mathcal{H}_i$ is complete, and therefore a Hilbert space.

When $\mathcal{H}_i = \mathcal{H}$ for all $i \in I$, we write

$$\ell^2(I, \mathcal{H}) := \bigoplus_{i\in I} \mathcal{H}_i.$$

Note that for $\mathcal{H} = \mathbb{C}$, one has $\ell^2(I, \mathbb{C}) = \ell^2(I)$.

If $\mathcal{H}_1$ and $\mathcal{H}_2$ are a pair of Hilbert spaces, the **tensor product** of $\mathcal{H}_1$ and $\mathcal{H}_2$, denoted by $\mathcal{H}_1 \otimes \mathcal{H}_2$, is defined as follows. Consider first the algebraic tensor product

$$\mathcal{H}_1 \odot \mathcal{H}_2 := \left\{ \sum_{j=1}^{n} \mathbf{x}_j \otimes \mathbf{y}_j : n \in \mathbb{N},\ \mathbf{x}_j \in \mathcal{H}_1, \mathbf{y}_j \in \mathcal{H}_2 \right\}$$

with operations

$$\begin{aligned}
\mathbf{x}_1 \otimes \mathbf{y} + \mathbf{x}_2 \otimes \mathbf{y} &= (\mathbf{x}_1 + \mathbf{x}_2) \otimes \mathbf{y} \\
\mathbf{x} \otimes \mathbf{y}_1 + \mathbf{x} \otimes \mathbf{y}_2 &= \mathbf{x} \otimes (\mathbf{y}_1 + \mathbf{y}_2) \\
(a\mathbf{x}) \otimes \mathbf{y} = \mathbf{x} \otimes (a\mathbf{y}) &= a(\mathbf{x} \otimes \mathbf{y})
\end{aligned}$$

and inner product

$$\langle \mathbf{x}_1 \otimes \mathbf{y}_1, \mathbf{x}_2 \otimes \mathbf{y}_2 \rangle := \langle \mathbf{x}_1, \mathbf{x}_2 \rangle \langle \mathbf{y}_1, \mathbf{y}_2 \rangle.$$

$\mathcal{H}_1 \otimes \mathcal{H}_2$ is then defined as the completion of $\mathcal{H}_1 \odot \mathcal{H}_2$ with respect to metric induced by this inner product.

Exercise 1.29. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces with orthonormal bases $(e_i)_{i \in I}$ and $(f_j)_{j \in J}$. Compute orthonormal bases for $\mathcal{H}_1 \oplus \mathcal{H}_2$ and $\mathcal{H}_1 \otimes \mathcal{H}_2$. Use this to show that

$$\mathcal{H}_1 \otimes \mathcal{H}_2 \cong \ell^2(J, \mathcal{H}_1) \cong \ell^2(I, \mathcal{H}_2).$$

Exercise 1.30. Show that $\ell^2(I) \otimes \mathcal{H} \cong \ell^2(I, \mathcal{H})$.

1.4 Bounded Linear Operators

Proposition 1.31. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. For a linear operator $T \colon \mathcal{H} \to \mathcal{K}$, the following are equivalent:*

(i) *T is uniformly continuous.*
(ii) *T is continuous at zero.*
(iii) $\displaystyle\sup_{\mathbf{x} \in H \setminus \{0\}} \|T\mathbf{x}\| / \|\mathbf{x}\| < \infty$

Definition 1.32. We say a linear operator $T \colon \mathcal{H} \to \mathcal{K}$ is **bounded** if it satisfies any (hence all) of the conditions in Proposition 1.31. The **operator norm** of T is the quantity

$$\|T\| := \sup_{\mathbf{x} \in \mathcal{H} \setminus \{0\}} \frac{\|T\mathbf{x}\|}{\|\mathbf{x}\|}.$$

The collection of all bounded linear operators from $\mathcal{H}$ to $\mathcal{K}$ is denoted $B(\mathcal{H}, \mathcal{K})$, and we write $B(\mathcal{H}) := B(\mathcal{H}, \mathcal{H})$.

In the case that $\mathcal{K} = \mathbb{C}$, then we call $T \in B(\mathcal{H}, \mathbb{C})$ a **bounded linear functional**.

Note that for $T \in B(\mathcal{H}, \mathcal{K})$ and $\mathbf{x} \in \mathcal{H}$ we have $\|T\mathbf{x}\| \leq \|T\| \|\mathbf{x}\|$.

Exercise 1.33. Let $T \colon \mathcal{H} \to \mathcal{K}$ be a bounded linear operator. Show that

$$\|T\| = \sup_{\|\mathbf{x}\| \leq 1} \|T\mathbf{x}\| = \sup_{\|\mathbf{x}\| = 1} \|T\mathbf{x}\| = \inf\{c > 0 \colon \|T\mathbf{x}\| \leq c\|\mathbf{x}\| \ \forall \mathbf{x} \in \mathcal{H}\}.$$

Example 1.34. Let (X, μ) be a positive measure space. For $\phi \in L^\infty(X, \mu)$ define $M_\phi \in B(L^2(X, \mu))$ by $M_\phi f = \phi f$. It is easy to see that $\|M_\phi\| \leq \|\phi\|_\infty$, and if μ is σ-finite then one can prove that $\|M_\phi\| = \|\phi\|_\infty$. ■

On a Hilbert space $\mathcal{H}$, the linear functionals $\mathcal{H} \ni \mathbf{x} \mapsto \langle \mathbf{x}, \mathbf{y} \rangle$ are bounded by the Cauchy–Schwarz inequality. It turns out all bounded linear functionals are of this form.

Theorem 1.35. *If $T \colon \mathcal{H} \to \mathbb{C}$ is a bounded linear functional, then there exists a unique $\mathbf{y} \in \mathcal{H}$ so that $T\mathbf{x} = \langle \mathbf{x}, \mathbf{y} \rangle$ for all $\mathbf{x} \in \mathcal{H}$. Moreover, $\|T\| = \|\mathbf{y}\|$.*

Using this one can establish the existence of the *adjoint* of a bounded operator.

Theorem 1.36. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. For each $T \in B(\mathcal{H}, \mathcal{K})$ there exists a unique $S \in B(\mathcal{K}, \mathcal{H})$ satisfying*

$$\langle T\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, S\mathbf{y} \rangle \qquad \forall \mathbf{x} \in \mathcal{H}, \ \mathbf{y} \in \mathcal{K}$$

Proof. Fix $\mathbf{y} \in \mathcal{K}$ and use the Cauchy–Schwarz inequality to observe that

$$|\langle T\mathbf{x}, \mathbf{y}\rangle| \leq \|T\mathbf{x}\|\|\mathbf{y}\| \leq \|T\|\|\mathbf{x}\|\|\mathbf{y}\|.$$

Thus $\mathbf{x} \mapsto \langle T\mathbf{x}, \mathbf{y}\rangle$ is a bounded linear functional on $\mathcal{H}$. The previous theorem implies there exists $\mathbf{z} \in \mathcal{H}$ so that $\langle T\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{z}\rangle$. Define $S\colon \mathcal{K} \to \mathcal{H}$ by $S\mathbf{y} := \mathbf{z}$ so that

$$\langle T\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, S\mathbf{y}\rangle .$$

Using the uniqueness of the vector $\mathbf{z}$ and the linearity of the inner products, one can show S is linear (**Exercise:** check this). Observe that:

$$\|S\mathbf{y}\|^2 = \langle S\mathbf{y}, S\mathbf{y}\rangle = \langle TS\mathbf{y}, \mathbf{y}\rangle \leq \|TS\mathbf{y}\|\|\mathbf{y}\| \leq \|T\|\|S\mathbf{y}\|\|\mathbf{y}\|.$$

Consequently $\|S\mathbf{y}\| \leq \|T\|\|\mathbf{y}\|$, and so S is bounded.

It remains to check that S is unique. Suppose $R \in B(\mathcal{K}, \mathcal{H})$ also satisfies $\langle T\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, R\mathbf{y}\rangle$ for all $\mathbf{x} \in \mathcal{H}$ and $\mathbf{y} \in \mathcal{K}$. This implies

$$\langle \mathbf{x}, S\mathbf{y} - R\mathbf{y}\rangle = \langle \mathbf{x}, S\mathbf{y}\rangle - \langle \mathbf{x}, R\mathbf{y}\rangle = \langle T\mathbf{x}, \mathbf{y}\rangle - \langle T\mathbf{x}, \mathbf{y}\rangle = 0$$

for all $\mathbf{x} \in \mathcal{H}$ and $\mathbf{y} \in \mathcal{K}$. So in particular it holds for $\mathbf{x} = S\mathbf{y} - R\mathbf{y}$, which yields $\|S\mathbf{y} - R\mathbf{y}\|^2 = 0$. That is, $S\mathbf{y} - R\mathbf{y} = 0$ or $S\mathbf{y} = R\mathbf{y}$. □

Definition 1.37. Given an bounded linear operator $T \in B(\mathcal{H}, \mathcal{K})$, the unique $S \in B(\mathcal{K}, \mathcal{H})$ satisfying $\langle T\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, S\mathbf{y}\rangle$ for all $\mathbf{x} \in \mathcal{H}$ and $\mathbf{y} \in \mathcal{K}$ is called the **adjoint** of T and is denoted $T^* := S$.

Example 1.38. For $T \in M_{m\times n}(\mathbb{C})$, its adjoint when thought of as a element of $B(\mathbb{C}^n, \mathbb{C}^m)$ is the conjugate transpose of T. Indeed, let $[T]_{i,j}$ denote the (i, j)th entry of T. Then for any $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{C}^n$ and $\mathbf{y} = (y_1, \ldots, y_m) \in \mathbb{C}^m$ we have $[T\mathbf{x}]_i = \sum_{j=1}^n [T]_{i,j} x_j$, and therefore

$$\langle T\mathbf{x}, \mathbf{y}\rangle = \sum_{i=1}^m \sum_{j=1}^n [T]_{i,j} x_j \bar{y}_i.$$

In order for this to equal

$$\langle \mathbf{x}, T^*\mathbf{y}\rangle = \sum_{j=1}^n \sum_{i=1}^m x_j \overline{[T^*]_{j,i} y_i},$$

we must have $[T^*]_{j,i} = \overline{[T]_{i,j}}$. ■

Definition 1.39. We call an operator $T \in B(\mathcal{H})$:

- **self-adjoint** if $T = T^*$;
- **normal** if $T^*T = TT^*$;
- **invertible** if there exists $S \in B(\mathcal{H})$ with $ST = TS = I$.
- a **unitary** if $T^*T = TT^* = 1$;
- a **projection** if $T = T^* = T^2$;
- an **isometry** if $T^*T = 1$
- a **partial isometry** if $T = TT^*T$.

Note that self-adjoint operators, unitaries, and projections are all normal. A unitary is precisely a normal isometry; equivalently, a unitary is an invertible isometry. Isometries, unitaries, and projections are all partial isometries. (**Exercise:** convince yourself of these statements.)

Example 1.40. In $M_2(\mathbb{C})$

$$\begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$

is a unitary for all $t \in \mathbb{R}$,

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

are projections, and

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

are partial isometries. Also, $A \in M_2(\mathbb{C})$ is invertible if and only if $\det(A) \neq 0$, which is equivalent to $\ker(A) = \{0\}$. (However, this latter statement is **not** true for bounded operators on infinite dimensional Hilbert spaces.) ■

Note that an operator S is only invertible if it admits both a right and a left inverse. The existence of a left inverse for S does not imply the existence for a right inverse of S, and vice versa, as the following exercise shows.

Exercise 1.41. Consider $S \in B(\ell^2(\mathbb{N}))$ defined by

$$S(x_1, x_2, \ldots) := (0, x_1, x_2, \ldots).$$

Show that S is an isometry but **not** a unitary.

Proposition 1.42. *Let $T \in B(\mathcal{H})$.*

(1) If T is a projection, then $T(1-T) = 0$ and $T\mathcal{H} \perp (1-T)\mathcal{H}$.
(2) T is an isometry if and only if $\|T\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathcal{H}$.
(3) T is a partial isometry if and only if T^*T and TT^* are projections.

2. Banach Spaces

Definition 2.1. For a vector space V, a **norm** is a map $\|\cdot\|: V \to [0,\infty)$ satisfying

1. $\|\mathbf{x}+\mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$ for all $\mathbf{x},\mathbf{y} \in V$;
2. $\|a\mathbf{x}\| = |a|\|\mathbf{x}\|$ for all $a \in \mathbb{C}$ and $\mathbf{x} \in V$;
3. $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0$.

A **normed space** is a pair $(V, \|\cdot\|)$ consisting of a vector space and a norm.

Observe that a norm defines a metric $d(\mathbf{x},\mathbf{y}) := \|\mathbf{x}-\mathbf{y}\|$ on V.

Definition 2.2. A **Banach space** is a normed space which is complete with respect to the metric induced by the norm.

Example 2.3.

(1) Every inner product space is a normed space and every Hilbert space is a Banach space. However, a normed space (Banach space) is an inner product space (Hilbert space) iff the norm satisfies the *parallelogram law*:
$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$
In this case, the inner product associated to the norm is unique.

(2) Let (X,μ) be a positive measure space. For all $1 \le p \le \infty$, $L^p(X,\mu)$ is a Banach space with norm
$$\|f\|_p = \left(\int_X |f(x)|^p \, d\mu(x)\right)^{1/p}.$$
In particular, $\ell^p(\mathbb{N})$ is a Banach space with norm
$$\|(a_n)_{n\in\mathbb{N}}\|_p = \left(\sum_{n\in\mathbb{N}} |a_n|^p\right)^{1/p}.$$

(3) Let X be a locally compact Hausdorff space. Let $C_b(X)$ denote the collection of bounded continuous functions $f: X \to \mathbb{C}$. This is a vector space with pointwise operations and is a Banach space under the norm
$$\|f\| := \sup_{x\in X} |f(x)|.$$
Let $C_0(X)$ denote the subset of $C_b(X)$ which consists of functions *vanishing at infinity*: that is, $f \in C_0(X)$ iff for all $\epsilon > 0$ the set $\{x \in X : |f(x)| \ge \epsilon\}$ is compact. Then $C_0(X)$ is also a Banach space with the norm it inherits from $C_b(X)$. If X is compact, then $C_b(X) = C_0(X) = C(X)$, the collection of all continuous functions on X.
In particular, for $X = \mathbb{N}$ equipped with the discrete topology, we have $C_b(\mathbb{N}) = \ell^\infty(\mathbb{N})$ and $C_0(\mathbb{N}) = c_0(\mathbb{N})$ (the collection of sequences converging to zero). ■

Remark 2.4. For another perspective on functions vanishing at infinity, recall that any locally compact Hausdorff space X has a one-point compactification, often denoted $X \cup \{\infty\}$, whose topology is generated by all open sets of X along with all sets of the form $(X \setminus C) \cup \{\infty\}$ where $C \subset X$ is compact. With this perspective, we can view $C_0(X)$ as the subspace of $C(X \cup \{\infty\})$ consisting of continuous functions that are 0 at the point ∞. For example, when $X = (0,1]$, its one point compactification is $[0,1]$, and we view $C_0((0,1])$ as the functions f in $C([0,1])$ such that $f(0) = 0$.

The analogue of Proposition 1.31 holds for linear operators between two normed spaces V and W. The collection of all such bounded linear operators is denoted $B(V,W)$, and we write $B(V) := B(V,V)$. Note that the operator norm on $B(V,W)$ makes it into a normed space.

Exercise 2.5. Let V be a normed space and $\mathcal{X}$ a Banach space. Show that $B(V,\mathcal{X})$ is a Banach space under the operator norm. (In particular, $B(V,\mathbb{C})$ is always a Banach space.)

Example 2.6.

(1) Let (X,μ) be a positive measure space. For $\phi \in L^\infty(X,\mu)$ and $1 \le p \le \infty$, define $M_\phi \in B(L^p(X,\mu))$ by $M_\phi f = \phi f$. Then $\|M_\phi\| \le \|\phi\|_\infty$ and if μ is σ-finite then this is an equality.

(2) Let X be a locally compact Hausdorff space. For any $x_0 \in X$, $C_b(X) \ni f \mapsto f(x_0)$ defines a bounded linear functional of norm 1. ■

Definition 2.7. Let V and W be normed spaces. An **isomorphism** between V and W is a bijection $T \in B(V, W)$ such that $T^{-1} \in B(W, V)$. An **isometric isomorphism** between V and W is a bijection $T \in B(V, W)$ such that $\|T\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in V$.

If T is an isometric isomorphism between normed spaces then T^{-1} is automatically bounded with $\|T\| = \|T^{-1}\| = 1$.

Bounded linear functionals on a Banach space ($B(\mathcal{X}, \mathbb{C})$) are not as simply characterized as those on a Hilbert space. This will be explored further in the next section.

The following four theorems are often first encountered in a graduate analysis course, where they are stated in terms of L^p spaces. However, they are really theorems about general Banach spaces:

Theorem 2.8 (Open Mapping Theorem)**.** *If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces and $T \in B(\mathcal{X}, \mathcal{Y})$ is surjective, then $T(U) \subset \mathcal{Y}$ is open for any open $U \subset \mathcal{X}$.*

Theorem 2.9 (Inverse Mapping Theorem)**.** *If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces and $T \in B(\mathcal{X}, \mathcal{Y})$ is bijective, then $T^{-1} \in B(\mathcal{Y}, \mathcal{X})$.*

Theorem 2.10 (Closed Graph Theorem)**.** *If $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces and $T \colon \mathcal{X} \to \mathcal{Y})$ is a linear transformation whose graph*

$$\mathit{graph}(T) := \{(\mathbf{x}, T\mathbf{x}) \colon \mathbf{x} \in \mathcal{X}\} \subset \mathcal{X} \oplus \mathcal{Y}$$

is closed, then $T \in B(\mathcal{X}, \mathcal{Y})$.

Theorem 2.11 (Principle of Uniform Boundedness)**.** *Let $\mathcal{X}$ be a Banach space and V a normed space. For any collection of bounded linear operators $\mathscr{C} \subset B(\mathcal{X}, V)$, if*

$$\sup_{T \in \mathscr{C}} \|T\mathbf{x}\| < \infty \qquad \forall \mathbf{x} \in \mathcal{X},$$

then

$$\sup_{T \in \mathscr{C}} \|T\| < \infty.$$

2.1 Dual Spaces

Definition 2.12. For a normed space $(V, \|\cdot\|)$, its **dual space** is the set $V^* := B(V, \mathbb{C})$ of bounded linear functionals on V.

Recall that $V^* = B(V, \mathbb{C})$ is always a Banach space, even if V is not.

Example 2.13.

(1) Let $\mathcal{H}$ be a Hilbert space and let $\bar{\mathcal{H}}$ its conjugate Hilbert space (see Exercise 1.10). Theorem 1.35 implies that

$$\begin{aligned} \bar{\mathcal{H}} f &\to \mathcal{H}^* \\ \bar{\mathbf{x}} &\mapsto \langle \,\cdot\,, \mathbf{x} \rangle \end{aligned}$$

is an isometric isomorphism.

(2) Let (X, μ) be a σ-finite measure space. For $1 \le p < \infty$, let $1 < q \le \infty$ be such that $\frac{1}{p} + \frac{1}{q} = 1$ (when $p = 1$ we take $q = \infty$). Given $g \in L^q(X, \mu)$ define $F_g \in L^p(X, \mu)^*$ by

$$F_g(f) := \int_X fg \; d\mu \qquad f \in L^p(X, \mu).$$

Then real analysis tells us that

$$\begin{aligned} L^q(X, \mu) &\to L^p(X, \mu)^* \\ g &\mapsto F_g \end{aligned}$$

is an isometric isomorphism.

In particular, $\ell^q(\mathbb{N})$ is isometrically isomorphic to $\ell^p(\mathbb{N})^*$. ■

Another important example is the dual of $C_0(X)$ for a locally compact Hausdorff space X. However, in order to describe it we must first recall some facts about complex measures.

Definition 2.14. Let Ω be a σ-algebra over a set X. For a complex valued measure μ on (X, Ω), define for $E \in \Omega$

$$|\mu|(E) := \sup \left\{ \sum_{n=1}^{\infty} |\mu(E_n)| \colon \{E_n\}_{n=1}^{\infty} \text{ is a } \Omega\text{-measurable partition of } E \right\}.$$

Then $|\mu|$ is a positive measure on (X, Ω) called the **absolute value** of μ. The **total variation** of μ is the quantity $\|\mu\| := |\mu|(X)$.

If μ is a positive measure, then $|\mu| = \mu$. If μ is real valued, then by the Jordan decomposition theorem $\mu = \mu_+ - \mu_-$ for positive measures $\mu_\pm$ and consequently $|\mu| = \mu_+ + \mu_-$.

Definition 2.15. Let X be a locally compact Hausdorff space and let Ω be the Borel σ-algebra on X. A positive measure μ on (X, Ω) is called a **regular Borel** measure if

(1) $\mu(K) < \infty$ for all compact $K \subset X$; and
(2) For all $E \in \Omega$

$$\mu(E) = \sup\{\mu(K) \colon K \subset E \text{ compact}\} = \inf\{\mu(U) \colon U \supset E \text{ open}\}.$$

We say a complex valued measure ν is a **regular Borel** measure if $|\nu|$ is a regular Borel measure. The set of all complex valued regular Borel measures on X is denoted $M(X)$.

For a locally compact Hausdorff space X, $M(X)$ is a vector space with operations defined by

$$(a\mu + \nu)(E) := a\mu(E) + \nu(E).$$

It is also a normed space with the norm of μ given by its total variation $\|\mu\|$. In fact, the following example implies it is even a Banach space:

Theorem 2.16 (Riesz Representation Theorem)**.** *Let X be a locally compact Hausdorff space. For $\mu \in M(X)$, define $F_\mu \in C_0(X)^*$ by*

$$F_\mu(f) = \int_X f \ d\mu.$$

Then

$$\begin{aligned} M(X) &\to C_0(X)^* \\ \mu &\mapsto F_\mu \end{aligned}$$

is an isometric isomorphism.

Example 2.17. As a partcular example of the previous theorem, consider $\mathbb{N}$ equipped with the discrete topology so that $C_0(\mathbb{N}) = c_0(\mathbb{N})$. Note that every $\mu \in M(\mathbb{N})$ satisfies

$$\sum_{n \in \mathbb{N}} |\mu(\{n\})| = \|\mu\| < \infty.$$

Thus $c_0(\mathbb{N})^* \cong M(\mathbb{N})$ is isometrically isomorphic to $\ell^1(\mathbb{N})$ via the map $\mu \mapsto (\mu(\{n\}))_{n \in \mathbb{N}}$. ■

2.2 Weak and Weak* Topologies

The relationship between a Banach space and its dual space induces new topologies on each of them.

Definition 2.18. Let $\mathcal{X}$ be a Banach space. The **weak topology** on $\mathcal{X}$, denoted $\sigma(\mathcal{X}, \mathcal{X}^*)$, is the topology generated by sets of the form

$$\{\mathbf{x} \in \mathcal{X} \colon |\mathbf{x}_0^*(\mathbf{x} - \mathbf{x}_0)| < \epsilon\} \qquad \mathbf{x}_0^* \in \mathcal{X}^*, \ \mathbf{x}_0 \in \mathcal{X}, \ \epsilon > 0.$$

The **weak* topology** on $\mathcal{X}^*$, denoted $\sigma(\mathcal{X}^*, \mathcal{X})$, is the topology generated by sets of the form

$$\{\mathbf{x}^* \in \mathcal{X}^* \colon |(\mathbf{x}^* - \mathbf{x}_0^*)(\mathbf{x}_0)| < \epsilon\} \qquad \mathbf{x}_0^* \in \mathcal{X}^*, \ \mathbf{x}_0 \in \mathcal{X}, \ \epsilon > 0.$$

From the perspective of analysis, these topologies are best understood in terms of convergence. However, they are (generally) not *metrizable* and consequently convergence must be understood through nets rather than sequences (see the Appendix on nets).

A net $(\mathbf{x}_i)_{i\in I} \subset \mathcal{X}$ converges to some $\mathbf{x}_0 \in \mathcal{X}$ in the weak topology if and only if for all $\mathbf{x}^* \in \mathcal{X}^*$ one has

$$\lim_{i\to\infty} \mathbf{x}^*(\mathbf{x}_i) = \mathbf{x}^*(\mathbf{x}_0).$$

In this case we say the net converges **weakly** to $\mathbf{x}_0$.

A net $(\mathbf{x}_i^*)_{i\in I} \subset \mathcal{X}^*$ converges to some $\mathbf{x}_0^* \in \mathcal{X}^*$ in the weak* topology if and only if for all $\mathbf{x} \in \mathcal{X}$ one has

$$\lim_{i\to\infty} \mathbf{x}_i^*(\mathbf{x}) = \mathbf{x}_0^*(\mathbf{x}).$$

In this case we say that the net converges **weak*** to $\mathbf{x}_0^*$.

Note that if $(\mathbf{x}_n)_{n\in\mathbb{N}} \subset \mathcal{X}$ converges in norm, then it also converges weakly since each $\mathbf{x}^* \in \mathcal{X}^*$ is norm-continuous. Consequently any subset of $\mathcal{X}$ that is closed in the weak topology is also closed in the norm topology. (In topologlical terms: the weak topology is coarser than the norm topology.) Likewise, any subset of $\mathcal{X}^*$ that is closed in the weak* topology is also closed in the norm topology. The following examples show that the converse of these statements is not true.

Example 2.19.

(1) Let $e_n \in \ell^2(\mathbb{N})$ be the element which has a 1 in the nth position and zeros elsewhere. Then $(e_n)_{n\in\mathbb{N}}$ converges weakly to zero, but $\|e_n\|_2 = 1$ for all $n \in \mathbb{N}$, so $(e_n)_{n\in\mathbb{N}}$ does not converge to zero in norm.

(2) Let $\mathscr{F}$ be the collection of finite subsets of $[0,1]$, directed by inclusion. For each $F \in \mathscr{F}$, define $\mu_F \in M([0,1])$ by

$$\mu_F = \frac{1}{|F|}\sum_{x\in F} \delta_x,$$

where $\delta_x \in M([0,1])$ is the Dirac probability measure at x. If we identify $M([0,1])$ with the dual of $C([0,1])$, then the net $(\mu_F)_{F\in\mathscr{F}} \subset M([0,1])$ converges weak* to the Riemann integral, but not in the total variation norm. ■

Despite these subleties, the weak and weak* topologies can be easier to work with than the norm topologies. For example, the closed unit ball $(\mathcal{X}^*)_1 := \{\mathbf{x}^* \in \mathcal{X}^* \colon \|\mathbf{x}^*\| \leq 1\}$ is compact in the norm topology if and only if $\mathcal{X}^*$ is finite-dimensional (**Exercise:** prove this). However, it is always compact in the weak* topology:

Theorem 2.20 (Banach–Alaoglu Theorem)**.** *The closed unit ball in the dual of a normed space is weak* compact.*

2.3 Ultrafilters If we view $\ell^\infty(\mathbb{N})$ as bounded functions on $\mathbb{N}$, then elements $\omega \in \ell^\infty(\mathbb{N})^*$ can be viewed as *finitely* additive $\mathbb{C}$-valued measures on $\mathbb{N}$. Indeed, for $S \subset \mathbb{N}$ let 1_S denote the indicator function on the subset S (i.e. the sequence $(a_n)_{n\in\mathbb{N}}$ with $a_n = 1$ if $n \in S$ and $a_n = 0$ otherwise). Then linear combinations of such functions are the simple functions on $\mathbb{N}$, and these are dense in $\ell^\infty(\mathbb{N})$. Consequently ω, as a bounded linear functional, is completely determined by its values on indicator functions. Thus we write $\omega(S) := \omega(1_S)$. (**Exercise:** verify that ω is finitely additive.)

Now, given $n_0 \in \mathbb{N}$, define $\delta_{n_0} \in \ell^\infty(\mathbb{N})^*$ by

$$\delta_{n_0}((a_n)_{n\in\mathbb{N}}) = a_{n_0}$$

Let $\beta\mathbb{N}$ denote the weak* closure of $\{\delta_n \colon n \in \mathbb{N}\} \subset \ell^\infty(\mathbb{N})^*$. (This is the *Stone–Cěch compactification* of $\mathbb{N}$.) Viewing $\ell^\infty(\mathbb{N})^*$ as finitely additive $\mathbb{C}$-valued measures on $\mathbb{N}$ as above, elements in $\beta\mathbb{N}$ are finitely additive $\{0,1\}$-valued measures. Indeed, if $\omega \in \beta\mathbb{N}$ then it is the weak* limit of some net $(\delta_{n_i})_{i\in I}$. Since $\delta_{n_i}(S) \in \{0,1\}$ for each $i \in I$, we have $\omega(S) \in \{0,1\}$.

Definition 2.21. The elements $\omega \in \beta\mathbb{N}$ are called **ultrafilters** on $\mathbb{N}$. We say an ultrafilter is **principal** if $\omega \in \mathbb{N}$, and otherwise say it is **non-principal** (or **free**).

It turns out there is an easy way to distinguish the principal and non-principal ultrafilters:

Exercise 2.22. For $\omega \in \beta\mathbb{N}$, we have $\omega \in \mathbb{N}$ if and only if there exists a finite subset $F \subset \mathbb{N}$ with $\omega(F) = 1$.

Thus a non-principal ultrafilter, viewed as a measure, vanishes on all finite subsets of $\mathbb{N}$. Morally speaking, this means the non-principal ultrafilters ignore any finitary behavior of bounded sequences. Exploring this further yields another important feature of non-principal ultrafilters (when viewed as elements of $\ell^\infty(\mathbb{N})^*$):

Proposition 2.23. *For $(a_n)_{n\in\mathbb{N}} \subset \ell^\infty(\mathbb{N})$, $a_0 \in \mathbb{C}$ is a cluster point of the sequence if and only if there exists $\omega \in \beta\mathbb{N} \setminus \mathbb{N}$ with*

$$\omega((a_n)_{n\in\mathbb{N}}) = a_0.$$

In particular, if $(a_n)_{n\in\mathbb{N}} \subset \ell^\infty(\mathbb{N})$ is a convergent sequence then

$$\omega((a_n)_{n\in\mathbb{N}}) = \lim_{n\to\infty} a_n$$

for all $\omega \in \beta\mathbb{N} \setminus \mathbb{N}$.

Compare the above to the behavior of a principal ultrafilters. In light of the preceding proposition, we adopt the following notation

$$\lim_{n\to\omega} a_n := \omega((a_n)_{n\in\mathbb{N}})$$

for any $\omega \in \beta\mathbb{N} \setminus \mathbb{N}$ and any $(a_n)_{n\in\mathbb{N}} \subset \ell^\infty(\mathbb{N})$.

3. Banach Algebras

Definition 3.1. An **algebra** is a vector space $\mathcal{A}$ that admits a multiplication operation

$$\begin{aligned} \mathcal{A} \times \mathcal{A} &\to \mathcal{A} \\ (\mathbf{x}, \mathbf{y}) &\mapsto \mathbf{xy} \end{aligned}$$

satisfying

1. $(\mathbf{xy})\mathbf{z} = \mathbf{x}(\mathbf{yz})$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{A}$ (*associativity*);
2. $\mathbf{x}(\mathbf{y} + \mathbf{z}) = \mathbf{xy} + \mathbf{xz}$ and $(\mathbf{x} + \mathbf{y})\mathbf{z} = \mathbf{xz} + \mathbf{yz}$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{A}$ (*distributivity*);
3. $a(\mathbf{xy}) = (a\mathbf{x})\mathbf{y} = \mathbf{x}(a\mathbf{y})$ for all $a \in \mathbb{C}$ and $\mathbf{x}, \mathbf{y} \in \mathcal{A}$.

We say $\mathcal{A}$ is **unital** if it admits an element $\mathbf{e} \in \mathcal{A}$ such that $\mathbf{ex} = \mathbf{xe} = \mathbf{x}$ for all $\mathbf{x} \in \mathcal{A}$. We call $\mathbf{e}$ the **identity** (or **unit**) of $\mathcal{A}$. We say $\mathcal{A}$ is **abelian** if $\mathbf{xy} = \mathbf{yx}$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{A}$.

Exercise 3.2. Show that the identity in a unital algebra is unique.

Definition 3.3. A **Banach algebra** is an algebra $\mathcal{A}$ equipped with a norm that makes $\mathcal{A}$ into a Banach space and satisfies

$$\|\mathbf{xy}\| \leq \|\mathbf{x}\|\|\mathbf{y}\| \qquad \forall \mathbf{x}, \mathbf{y} \in \mathcal{A}.$$

If $\mathcal{A}$ is unital, it is assumed that $\|\mathbf{e}\| = 1$.

Observe that for a unital Banach algebra, $\|a\mathbf{e}\| = |a|$ and so $\mathbb{C} \ni a \mapsto a\mathbf{e}$ is an isometry. Consequently, we typically write $1 := \mathbf{e}$ and $a := a\mathbf{e}$ for all $a \in \mathbb{C}$.

Example 3.4.

(1) For a locally compact Hausdorff space X, $C_b(X)$ is a Banach algebra with multiplication operation given by point-wise multiplication

$$(fg)(x) := f(x)g(x)$$

and the same norm which makes it a Banach space. Moreover, it is unital with identity given by the constant function $x \mapsto 1$.

The same multiplication and norm make $C_0(X)$ into a Banach algebra. However, $C_0(X)$ is unital if and only if X is compact.

(2) For (X, Ω, μ) a σ-finite measure space, $L^\infty(X, \mu)$ is a unital Banach algebra with point-wise mutliplication and identity given by the constant function $1(x) = 1$ for μ-a.e. $x \in X$.

(3) For $\mathcal{H}$ a Hilbert space, $B(\mathcal{H})$ is a unital Banach algebra with multiplication operation given by composition, identity given by the identity map $\mathbf{x} \mapsto \mathbf{x}$, and norm given by the operator norm.

In particular, $M_n(\mathbb{C})$ is a Banach algebra under matrix multiplication and the operator norm. ■

Exercise 3.5. For $f, g \in L^1(\mathbb{R})$ define their convolution $f * g \in L^1(\mathbb{R})$ by

$$(f * g)(t) := \int_{\mathbb{R}} f(t - s)g(s)\, ds.$$

Show that with convolution as the multiplication operation, $L^1(\mathbb{R})$ is a Banach algebra that is abelian but not unital.

Banach algebras (and their subalgebras, homomorphisms, and related structures) are a cornerstone of the field of operator algebras. *Ideals* in Banach algebras will be particularly important for us.

Definition 3.6. For a Banach algebra $\mathcal{A}$, an **ideal** is a closed subspace $\mathcal{I} \subset \mathcal{A}$ which satisfies $\mathbf{xy}, \mathbf{yx} \in \mathcal{I}$ for all $\mathbf{x} \in \mathcal{A}$ and all $\mathbf{y} \in \mathcal{I}$.

Example 3.7. Let $\phi\colon \mathcal{A} \to \mathbb{C}$ be a continuous homomorphism. Then $\ker \phi$ is an ideal of $\mathcal{A}$. Indeed, it is closed subspace since ϕ is, in particular, a continuous linear functional. Also, if $\mathbf{x} \in \mathcal{A}$ and $\mathbf{y} \in \ker \phi$ then we have

$$\begin{aligned} \phi(\mathbf{xy}) &= \phi(\mathbf{x})\phi(\mathbf{y}) = 0 \\ \phi(\mathbf{yx}) &= \phi(\mathbf{y})\phi(\mathbf{x}) = 0. \end{aligned}$$

Thus $\mathbf{xy}, \mathbf{yx} \in \ker \phi$.

If $\mathcal{A}$ is unital, then for all $\mathbf{x} \in \mathcal{A}$, $\mathbf{x} - \phi(\mathbf{x})1 \in \ker(\phi)$. Since we can write $\mathbf{x} = (\mathbf{x} - \phi(\mathbf{x})1) + \phi(\mathbf{x})1$, it follows that $\mathcal{A}/\ker(\phi) \cong \mathbb{C}$. ■

Any ideal $\mathcal{I} \subset \mathcal{A}$ in a unital Banach algebra that satisfies $\mathcal{A}/\mathcal{I} \cong \mathbb{C}$ is of this form:

Theorem 3.8. *If $\mathcal{I} \subset \mathcal{A}$ is an ideal in a unital Banach algebra satisfying $\mathcal{A}/\mathcal{I} \cong \mathbb{C}$, then there exists a continuous homomorphism $\phi\colon \mathcal{A} \to \mathbb{C}$ with $\ker \phi = \mathcal{I}$.*

Remark 3.9. Our use of the word "ideal" here is more specific than what is typically seen in algebra (where one can consider *non-closed* ideals, *left* ideals, or *right* ideals). The above theorem is our motivation for adopting this convention throughout GOALS (and likewise is the motivation for the convention appearing in most of the operator algebras literature).

Sometimes, as in the case of $B(\mathcal{H})$ or $C_0(X)$, there is also a natural involution on a given Banach algebra.

Definition 3.10. A **Banach $*$-algebra** is a Banach algebra $\mathcal{A}$ with involution $*$ (called the **adjoint**) which is an anti-isomorphism, i.e. for all $\mathbf{x}, \mathbf{y} \in \mathcal{A}$ and $\lambda \in \mathbb{C}$,

- $(\mathbf{x} + \mathbf{y})^* = \mathbf{x}^* + \mathbf{y}^*$
- $(a\mathbf{x})^* = \bar{a}\mathbf{x}^*$
- $(\mathbf{x}^*)^* = \mathbf{x}$
- $(\mathbf{x}\mathbf{y})^* = \mathbf{y}^*\mathbf{x}^*$

When we are working in a Banach $*$-algebra, we will also assume ideals are $*$-closed.

3.1 The Spectrum

Definition 3.11. A element $\mathbf{x} \in \mathcal{A}$ of a unital algebra is said to be **invertible** if there exists $\mathbf{y} \in \mathcal{A}$ with $\mathbf{x}\mathbf{y} = \mathbf{y}\mathbf{x} = 1$. In this case, we write $\mathbf{x}^{-1} := \mathbf{y}$.

Example 3.12.

(1) Let X be a compact Hausdorff space. Then $f \in C(X)$ is invertible if and only if $f(x) \neq 0$ for all $x \in X$. In this case, its inverse is given by the function $g(x) = 1/f(x)$.

(2) Let (X, Ω, μ) be a σ-finite measure space. For $f\colon X \to \mathbb{C}$, the *essential range* of f is the set

$$\text{ess.im}(f) := \{a \in \mathbb{C}\colon \mu(\{x \in X\colon |f(x) - a| \leq \epsilon\}) > 0 \text{ for all } \epsilon > 0\}$$

Then $f \in L^\infty(X, \mu)$ is invertible if and only if $0 \notin \text{ess.im}(f)$.

(3) Let $\mathcal{H}$ be a Hilbert space. The Inverse Mapping Theorem implies $T \in B(\mathcal{H})$ is invertible if and only if T is bijective. ■

Theorem 3.13. *Let $\mathcal{A}$ be a unital Banach algebra. Then*

$$G := \{\mathbf{x} \in \mathcal{A}\colon \mathbf{x} \textit{ is invertible}\}$$

is open and the map $G \ni \mathbf{x} \mapsto \mathbf{x}^{-1}$ is continuous.

Definition 3.14. Let $\mathcal{A}$ be a unital Banach algebra. The **spectrum** of $\mathbf{x} \in \mathcal{A}$ is the set

$$\sigma(\mathbf{x}) := \{a \in \mathbb{C}\colon \mathbf{x} - a \text{ is not invertible}\}.$$

The **resolvent** of $\mathbf{x}$ is the set $\rho(\mathbf{x}) := \mathbb{C} \setminus \sigma(\mathbf{x})$.

Example 3.15.

(1) For $A \in M_n(\mathbb{C})$, $\sigma(A)$ is given by the eigenvalues of A. Indeed, $A - a$ is not invertible if and only if $\det(A - a) = 0$, which is equivalent to a being an eigenvalue of A.

(2) Let X be a compact Hausdorff space. For $f \in C(X)$, $\sigma(f) = f(X)$ since $f - a$ is invertible if and only if $0 \neq (f - a)(x) := f(x) - a$ for all $x \in X$.

(3) Let (X, Ω, μ) be a σ-finite measure space. For $f \in L^\infty(X, \mu)$, $\sigma(f) = \text{ess.im}(f)$. ■

Theorem 3.16. *Let $\mathcal{A}$ be a unital Banach algebra. For all $\mathbf{x} \in \mathcal{A}$, $\sigma(\mathbf{x})$ is a non-empty compact subset of $\{a \in \mathbb{C}\colon |a| \leq \|\mathbf{x}\|\}$.*

We will just prove that it is a subset of $\{a \in \mathbb{C}\colon |a| \leq \|\mathbf{x}\|\}$.

Proof. If $|\lambda| > \|a\|$, then the series

$$\lambda^{-1} \sum_{n \geq 0} \lambda^{-n} a^n$$

converges in norm. Notice that for each $N > 0$,

$$(\lambda 1 - a)\lambda^{-1} \sum_{n=0}^{N} \lambda^{-n} a^n = 1 - \frac{a^{N+1}}{\lambda^{N+1}} \to 1.$$

A similar computation with multiplication on the left shows that the limit of the series is $(\lambda 1 - a)^{-1}$. □

Corollary 3.17. *For any element a in a unital Banach algebra, if $\|a\| < 1$, then $1 - a$ is invertible with inverse $\sum_{n \geq 0} a^n$.*

Definition 3.18. Given $\mathbf{x} \in \mathcal{A}$ in a unital Banach algebra, the **spectral radius** of $\mathbf{x}$ is the quantity:

$$r(\mathbf{x}) := \sup_{a \in \sigma(\mathbf{x})} |a|.$$

Theorem 3.16 implies that we always have $r(\mathbf{x}) \leq \|\mathbf{x}\|$. However, the reverse inequality need not hold as the following example demonstrates.

Example 3.19. Consider

$$\mathbf{x} := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(\mathbb{C}).$$

From linear algebra we know the only eigenvalue of $\mathbf{x}$ is zero, so $\sigma(\mathbf{x}) = \{0\}$ and hence $r(\mathbf{x}) = 0$. On the other hand, $\|\mathbf{x}\| = 1$ (**Exercise:** verify this). ■

Observe that in the previous example we have $\mathbf{x}^2 = 0$ and so we do have the equality $r(\mathbf{x}) = \|\mathbf{x}^2\|^{1/2}$. It turns out one can always prove a modified equality:

Theorem 3.20 (Spectral Radius Theorem)**.** *Let $\mathcal{A}$ be a unital Banach algebra. For all $\mathbf{x} \in \mathcal{A}$*

$$r(\mathbf{x}) = \lim_{n \to \infty} \|\mathbf{x}^n\|^{1/n}.$$

4. The Hahn–Banach and Krein–Milman Theorems

4.1 The Hahn–Banach Theorem and Corollaries

Definition 4.1. For a vector space V, a **seminorm** is a map $p\colon V \to [0,\infty)$ satisfying

1. $p(\mathbf{x}+\mathbf{y}) \leq p(\mathbf{x}) + p(\mathbf{y})$ for all $\mathbf{x},\mathbf{y} \in V$;
2. $p(a\mathbf{x}) = |a|p(\mathbf{x})$ for all $a \in \mathbb{C}$ and $\mathbf{x} \in V$.

Every norm is a seminorm, but a seminorm p is only a norm if $p(\mathbf{x}) = 0$ implies $\mathbf{x} = 0$. (Note that $p(0) = 0$ always by the second part of the above definition.)

Example 4.2.

(1) Let X be a locally compact Hausdorff space and let $K \subset X$ be compact. Define
$$p_K(f) := \sup_{x\in K} |f(x)| \qquad f \in C(X).$$
Then p_K is a seminorm on $C(X)$, and $p_K(f) = 0$ for any f that vanishes on K.

(2) Let $\mathcal{X}$ be a Banach space. For fixed $\mathbf{x}_0 \in \mathcal{X}$ and $\mathbf{x}_0^* \in \mathcal{X}^*$, the maps
$$\mathbf{x} \mapsto |\mathbf{x}_0^*(\mathbf{x})| \qquad \text{and} \qquad \mathbf{x}^* \mapsto |\mathbf{x}^*(\mathbf{x}_0)|$$
are seminorms are $\mathcal{X}$ and $\mathcal{X}^*$, respectively. ■

Strictly speaking, the following is not *the* Hahn–Banach Theorem, but rather a corollary of it. However, this is what researchers typically mean when they invoke "The Hahn–Banach Theorem" and we will continue this tradition in the GOALS mini-courses.

Theorem 4.3 (Hahn–Banach Theorem)**.** *Let V be a vector space, $W \subset V$ a subspace, and p a seminorm on V. If $f\colon W \to \mathbb{C}$ is a linear functional satisfying*
$$|f(\mathbf{x})| \leq p(\mathbf{x}) \qquad \forall \mathbf{x} \in W,$$
then there exists a linear functional $F\colon V \to \mathbb{C}$ satisfying $F|_W = f$ and $|F(\mathbf{x})| \leq p(\mathbf{x})$ for all $\mathbf{x} \in V$.

We note that the power of the above theorem is *not* in the existence of an extension F (of which there are many), but in the fact that one can find an extension which is still "continuous" with respect to the seminorm p.

We mention a few useful corollaries of the Hahn–Banach theorem

Corollary 4.4. *If V is a normed space and $W \subset V$ is a closed subspace, then for any bounded linear functional $f\colon W \to \mathbb{C}$ there exists a bounded lienar functional $F\colon V \to \mathbb{C}$ satisfying $F|_W = f$ and $\|F\| = \|f\|$.*

Corollary 4.5. *If V is a normed space, then for any $\mathbf{x} \in V$,*
$$\|\mathbf{x}\| = \sup\{|\mathbf{x}^*(\mathbf{x})| \colon \mathbf{x}^* \in V^*, \|\mathbf{x}^*\| = 1\}.$$
In particular, there exists $\mathbf{x}_0^* \in V^*$ with $\|\mathbf{x}_0^*\| = 1$ and satisfying $\mathbf{x}_0^*(\mathbf{x}) = \|\mathbf{x}\|$.

Corollary 4.6. *If V is a normed space, then for any subspace $W \subset V$*
$$\overline{W} = \bigcap\{\ker(\mathbf{x}^*) \colon \mathbf{x}^* \in V^* \text{ with } W \subset \ker(\mathbf{x}^*)\}.$$
In particular, W is dense if and only if $W \subset \ker(\mathbf{x}^*)$ for some $\mathbf{x}^* \in V^*$ implies $\mathbf{x}^* = 0$.

By adopting a geometric perspective, one can obtain the following result as another corollary of the Hahn–Banach theorem:

Theorem 4.7 (Hahn–Banach Separation Theorem)**.** *Let V be a normed space. If $X, Y \subset V$ are disjoint closed convex subsets with Y compact, then there exists a continuous linear functional $f\colon V \to \mathbb{C}$, $\alpha \in \mathbb{R}$, and $\epsilon > 0$ such that*
$$\mathrm{Re} f(\mathbf{x}) < \alpha < \alpha + \epsilon < \mathrm{Re} f(\mathbf{y}) \qquad \forall \mathbf{x} \in X,\ \mathbf{y} \in Y.$$

The above theorem may same innocuous at first glance, but it is in fact quite useful for proving (by way of contradiction) equalities of various closures. For example:

Proposition 4.8. *Let $\mathcal{X}$ be a Banach space. If $C \subset \mathcal{X}$ is convex, then its norm closure is equal to its weak closure.*

Proof. Let C_1 and C_2 denote the norm and weak closures, respectively, of C. Recall that this means C_1 (resp. C_2) equals the intersection of all norm (resp. weak) closed sets containing C. Also recall that being weak closed implies being norm closed, so C_2 is norm closed and hence $C_1 \subset C_2$.

Now suppose, towards a contradiction, that there exists $\mathbf{x}_0 \in C_2 \setminus C_1$. Then, in the norm topology, C_1 and $\{\mathbf{x}_0\}$ are disjoint closed convex closed sets (**Exercise:** show that C_1 is indeed convex) and $\{\mathbf{x}_0\}$ is compact. Thus the Hahn–Banach Separation Theorem implies that there exists $\mathbf{x}^* \in \mathcal{X}^*$, $\alpha \in \mathbb{R}$, and $\epsilon > 0$ such that

$$\mathrm{Re}[\mathbf{x}^*(\mathbf{x})] < \alpha < \alpha + \epsilon < \mathrm{Re}[\mathbf{x}^*(\mathbf{x}_0)] \qquad \forall \mathbf{x} \in C_1.$$

In particular, $\mathrm{Re}[\mathbf{x}^*(\mathbf{x})] < \alpha$ for all $\mathbf{x} \in C$. Consider the set

$$D := \{\mathbf{x} \in \mathcal{X} \colon \mathrm{Re}[\mathbf{x}^*(\mathbf{x})] \leq \alpha\}.$$

It is convex and weakly closed (**Exercise:** check this) and contains C. Hence $C_2 \subset D$, but this contradicts $\mathbf{x}_0 \notin D$. Thus we must have $C_2 \setminus C_1 = \emptyset$ and so $C_1 = C_2$. □

4.2 The Krein–Milman Theorem

Definition 4.9. Let V be a vector space and $K \subset V$ a convex subset. A point $\mathbf{x}_0 \in K$ is called an **extreme point** of K if $\mathbf{x}_0 \neq t\mathbf{x} + (1-t)\mathbf{y}$ for any $0 < t < 1$ and $\mathbf{x}, \mathbf{y} \in K$. The set of all extreme points of K is dentoted $\mathrm{ext}(K)$.

Example 4.10.

(1) If $P \subset \mathbb{R}^2$ is a closed convex polygon, then $\mathrm{ext}(P)$ consists of its vertices.
(2) For $\mathbb{D}$ the open unit disc of $\mathbb{C}$, $\mathrm{ext}(\mathbb{D}) = \emptyset$ while $\mathrm{ext}(\overline{\mathbb{D}}) = \{z \in \mathbb{C} \colon |z| = 1\}$.
(3) For $K = \{f \in L^1([0,1], m) \colon \|f\|_1 \leq 1\}$, $\mathrm{ext}(K) = \emptyset$. Indeed, let $t_0 \in (0,1)$ be such that
$$\int_0^{t_0} |f(t)| \, dt = \frac{1}{2}.$$
Define $g := 2f1_{[0,t_0]}$ and $h := 2f1_{[t_0,1]}$. Then $g, h \in K$ and $f = \frac{1}{2}g + \frac{1}{2}h \in (g,h)$, so f is not an extreme point. ■

As the examples above demonstrate, it is possible for a convex set to have no extreme points. However, if the convex set is also compact, then this is not possible thanks to the following theorem.

Theorem 4.11 (Krein–Milman Theorem)**.** *Let V be a locally convex set. For a nonempty compact convex subset $K \subset V$, $ext(K) \neq \emptyset$ and K is equal to the intersection of all closed convex subsets containing $ext(K)$ (i.e. the* closed convex hull *of $ext(K)$).*

Recall that by the Banach–Alaoglu theorem, the closed unit ball of the dual of a normed space is weak* compact. Hence we obtain the following corollary to the Krein–Milman theorem:

Corollary 4.12. *For a normed space V, the closed unit ball of V^* is equal to the closed convex hull of its extreme points.*

5. Appendix: Nets

Roughly speaking, nets are a generalization of sequences wherein the indexing set $\mathbb{N}$ is replaced by a *directed set*. As the name suggests, these sets have a notion of direction much like $\mathbb{N}$ does ($1 \to 2 \to 3 \cdots$), however they may be uncountable and may have multiple "paths to infinity." The elements that are indexed by a directed set live in a topological space so that one can consider the notion of convergence of a net. Nets are essential for general topology in the sense that they can characterize closedness, compactness, and continuity in the same way that sequences do in metric spaces.

5.1 Directed Sets

Definition 5.1. A **directed set** I is a set equipped with a binary relation $\leq$ that satisfies:

1. $i \leq i$ for all $i \in I$ (*reflexive*);
2. if $i \leq j$ and $j \leq k$, then $i \leq k$ (*transitive*);
3. for any $i, j \in I$ there exists $k \in I$ with $i, j \leq k$ (*upper bound property*).

Typically reflexivity and transitivity are obvious, whereas the upper bound property may need to be justified.

Example 5.2. $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ are all directed sets with the usual ordering. In fact, any subset of $\mathbb{R}$ (even finite ones) are directed sets with the order they inherit from $\mathbb{R}$. ∎

Example 5.3. Let X be a set, and let $\mathscr{F}$ denote the collection of all finite subsets of X. For $A, B \in \mathscr{F}$, write $A \leq B$ if $A \subset B$. This makes $\mathscr{F}$ into a directed set. Note that $A \cup B$ serves as an upper bound for both A and B. ∎

Example 5.4. Let X be a topological space, and fix $x_0 \in X$. Let $\mathscr{N}(x_0)$ denote the collection of open neighborhoods of x_0. For $A, B \in \mathscr{N}(x_0)$, write $A \leq B$ if $A \supset B$. This makes $\mathscr{N}(x_0)$ into a directed set where $A \cap B$ is an upper bound for A and B. ∎

Example 5.5. Let X be a topological space. Then $\{(\epsilon, K) \colon \epsilon > 0,\ K \subset X \text{ compact}\}$ is a directed set where

$$(\epsilon, K) \leq (\epsilon', K')$$

if and only if $\epsilon \geq \epsilon'$ and $K \subset K'$. (**Exercise:** determine a common upper bound for (ϵ, K) and (ϵ', K').) ∎

Exercise 5.6. Let $\{a_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$. Show that the following are equivalent:

(a) The net $\left(\sum_{n\in F} a_n\right)_{F\in\mathscr{F}}$ converges, where $\mathscr{F}$ is the collection of finite subsets $F \subset \mathbb{N}$.
(b) For any permutation $\pi \colon \mathbb{N} \to \mathbb{N}$, $\sum_{n=1}^{\infty} a_{\pi(n)}$ converges.
(c) $\sum_{n=1}^{\infty} |a_n|$ converges.

5.2 Nets

Definition 5.7. Let X be a topological space. A **net** in X is a map $x \colon I \to X$ where I is a directed set.

A net $x \colon I \to X$ is usually denoted $(x(i))_{i\in I}$ or $(x_i)_{i\in I}$ where $x_i := x(i)$. This is supposed to remind you of sequence notation. As with sequences in a metric space, there is a notion of convergence:

Definition 5.8. A net $(x_i)_{i\in I}$ **converges** to $x \in X$ if for every open subset $U \subset X$ containing x there is $i_0 \in I$ so that $x_i \in U$ whenever $i \geq i_0$. In this case we call x the **limit** of the net and write

$$x = \lim_i x_i.$$

When $I = \mathbb{N}$, this is simply the usual notion of convergence for a sequence. When $I = \mathbb{R}$ this is also capturing familiar behavior:

Example 5.9. Let $f \colon \mathbb{R} \to \mathbb{R}$ be a function. Recall that we say f has a limit at ∞ if there exists $L \in \mathbb{R}$ so that for all $\epsilon > 0$ there exists $t_0 \in \mathbb{R}$ so that

$$|f(t) - L| < \epsilon \qquad \forall t \geq t_0.$$

But this is precisely saying that the net $(f(t))_{t\in\mathbb{R}}$ converges to L. ∎

Example 5.10. Let X be a topological space, let $x_0 \in X$ and let $\mathscr{N}(x_0)$ be as in Example 5.4. For each $U \in \mathscr{N}(x_0)$ pick any point in U and label it x_U. Then $(x_U)_{U\in\mathscr{N}(x_0)}$ is a net which converges to x_0. Indeed, let $U \subset X$ be an open set containing x_0. Then $U \in \mathscr{N}(x_0)$ and for any $U' \in \mathscr{N}(x_0)$ with $U' \geq U$, we have $x_{U'} \in U' \subset U$. ■

Example 5.11. Let X be a topological space and let $f\colon X \to \mathbb{C}$ be a function. For each pair (ϵ, K) as in Example 5.5, let $f_{(\epsilon,K)}$ be any function $g\colon X \to \mathbb{C}$ satisfying $|f(x) - g(x)| < \epsilon$ for all $x \in K$. Then the net $(f_{(\epsilon,K)})$ converges to f in the topology of uniform convergence on compact subsets. Indeed, fix $K \subset X$ compact and let $\epsilon > 0$. Then for any $(\epsilon', K') \geq (\epsilon, K)$ we have $|f(x) - f_{(\epsilon',K')}(x)| < \epsilon' \leq \epsilon$ for all $x \in K'$; in particular, for all $x \in K$. ■

Proposition 5.12. *Let X be a topological space. Then $V \subset X$ is closed if and only if for every convergent net $(x_i)_{i\in I} \subset V$ one has $\lim_i x_i \in V$.*

Proof. ($\Rightarrow$): Let $(x_i)_{i\in I} \subset V$ be a convergent net. Suppose, towards a contradiction, that $x := \lim_i x_i$ is not contained in V. Then $x \in V^c$ which is an open set. Consequently, by definition of the convergence of a net, there exists $i_0 \in I$ such that $x_i \in V^c$ for all $i \geq i_0$. But this contradicts $x_i \in V$ for all $i \in I$.
($\Leftarrow$): To show that V is closed, we will show that V^c is open. Suppose, towards a contradiction, that there exists $x \in V^c$ such that for all open subsets U containing x one has $U \cap V \neq \emptyset$. Let $\mathscr{N}(x)$ be as in Example 5.4. For each $U \in \mathscr{N}(x)$, let $x_U \in U \cap V$. Then $(x_U)_{U\in\mathscr{N}(x)} \subset V$ and it converges to x by Example 5.10. By assumption we must have $x \in V$, but this contradicts $x \in V^c$. Thus for any $x \in V^c$ there is an open set containing x which does not intersect V; that is, V^c is open. □

We say a subset $S \subset X$ in a topological space is *sequentially closed* if whenever $(x_n)_{n\in\mathbb{N}} \subset S$ is a convergent sequence one has $\lim_n x_n \in S$. Since sequences are particular kinds of nets, the above proposition implies that closed sets are sequentially closed. In a metric space, the two notions are equivalent. However, for general topological spaces sequentially closed does not imply closed, as the following example illustrates.

Example 5.13. [1] Consider $\mathbb{R}^{\mathbb{R}}$ with the product topology, which we think of as arbitrary functions $f\colon \mathbb{R} \to \mathbb{R}$. Recall that under the product topology, an open subset of $\mathbb{R}^{\mathbb{R}}$ is a union of subsets of the form

$$\prod_{t\in\mathbb{R}} U_t,$$

where $U_t \subset \mathbb{R}$ is open for all $t \in \mathbb{R}$ and $U_t \neq \mathbb{R}$ for only finitely many $t \in \mathbb{R}$. Consequently, a net $(f_i)_{i\in I} \subset \mathbb{R}^{\mathbb{R}}$ converges to $f \in \mathbb{R}^{\mathbb{R}}$ if and only if the functions $(f_i)_{i\in I}$ converge pointwise to f on $\mathbb{R}$. Let B be the subset of Borel functions. Then B is sequentially closed because we know from measure theory that the pointwise limit of a sequence of Borel functions is Borel. B is also also dense. Indeed, let $f \in \mathbb{R}^{\mathbb{R}}$. Let $\mathscr{F}$ be the collection of finite subsets of $\mathbb{R}$, ordered by inclusion. Then for each $F \in \mathscr{F}$ we can find a polynomial p_F such that $p_F(t) = f(t)$ for each $t \in F$. The net $(p_F)_{F\in\mathscr{F}}$ converges pointwise to f and consists of Borel functions. Therefore the closure of B is all of $\mathbb{R}^{\mathbb{R}}$. On the other hand, we know there are non-Borel functions, so B cannot be closed. ■

Proposition 5.14. *Let X and Y be topological spaces. Then $f\colon X \to Y$ is continuous if and only if for every convergent net $(x_i)_{i\in I} \subset X$ one has that $(f(x_i))_{i\in I} \subset Y$ is a convergent net with $\lim_i f(x_i) = f(\lim_i x_i)$.*

Proof. ($\Rightarrow$): Suppose f is continuous and $(x_i)_{i\in I} \subset X$ converges to some $x \in X$. Let $U \subset Y$ be an open subset containing $f(x)$. Then $f^{-1}(U) \subset X$ is an open subset containing x. Consequently there exists $i_0 \in I$ such that for all $i \geq i_0$ we have $x_i \in f^{-1}(U)$. Thus for all $i \geq i_0$ we have $f(x_i) \in U$. So $(f(x_i)_{i\in I})$ converges to $f(x)$.
($\Leftarrow$): Let $U \subset Y$ be an open subset. We must show $f^{-1}(U)$ is open. If not, then there is an $x \in f^{-1}(U)$ such that $N \cap f^{-1}(U)^c \neq \emptyset$ for all $N \in \mathscr{N}(x)$. We can then define a net by letting $x_N \in N \cap f^{-1}(U)^c$ for each $N \in \mathscr{N}(x)$. Then the net $(x_N)_{N\in\mathscr{N}(x)}$ converges to x by Example 5.10. By construction, $f(x_N) \in U^c$ for all $N \in \mathscr{N}(x)$. By assumption, $(f(x_N))_{N\in\mathscr{N}(x)}$ converges to $f(x)$, and since U^c is closed the previous proposition implies $f(x) \in U^c$. But this contradicts $x \in f^{-1}(U)$. Thus $f^{-1}(U)$ must be open and therefore f is continuous. □

[1]Thanks to Ben Hayes for supplying this example.

Let (X, d) be a metric space. We say a net $(x_i)_{i\in I} \subset X$ is *Cauchy* if for all $\epsilon > 0$ there exists $i_0 \in I$ so that whenever $i, j \geq i_0$ we have $d(x_i, x_j) < \epsilon$. We conclude this section by examining Cauchy nets in metric spaces. In particular, we will show that Cauchy nets in a complete metric space converge. The idea is to extract a Cauchy sequence from the Cauchy net, so as to use the completeness.

Proposition 5.15. *Let (X, d) be a complete metric space and let $(x_i)_{i\in I}$ be a Cauchy net. Then $(x_i)_{\in I}$ converges.*

Proof. Let $i(1) \in I$ be such that $d(x_i, x_j) < 1$. Let $i(2) \in I$ be such that $i(2) \geq i(1)$ and $d(x_i, x_j) < \frac{1}{2}$ for all $i, j \geq i(2)$. We inductively find $i(n) \in I$ for each $n \in \mathbb{N}$ such that $i(n) \geq i(n-1)$ and $d(x_i, x_j) < \frac{1}{n}$ for all $i, j \geq i(n)$. We claim that the sequence $(x_{i(n)})_{n\in\mathbb{N}}$ is Cauchy. Indeed, let $\epsilon > 0$. If $N \in \mathbb{N}$ satisfies $\frac{1}{N} < \epsilon$, then for $n, m \geq N$ we have $d(x_{i(n)}, x_{i(m)}) < \frac{1}{N} < \epsilon$. Since (X, d) is complete, $(x_{i(n)})_{n\in\mathbb{N}}$ converges to some $x \in X$. We claim the original net also converges to this x. Indeed, let $\epsilon > 0$ and choose $N \in \mathbb{N}$ such that for all $n \geq N$ we have $d(x_{i(n)}, x) < \frac{\epsilon}{2}$. By choosing a larger N if necessary, we may assume $\frac{1}{N} \leq \frac{\epsilon}{2}$. Then for any $i \geq i(N)$ we have

$$d(x_i, x) \leq d(x_i, x_{i(N)}) + d(x_{i(N)}, x) < \frac{1}{N} + \frac{\epsilon}{2} \leq \epsilon.$$

Hence the $(x_i)_{i\in I}$ converges to x. □

Remark 5.16. When (X, d) is a metric space, any Cauchy sequence $(x_n)_{n\in\mathbb{N}} \subset X$ is bounded. Indeed, let $N \in \mathbb{N}$ be such that $d(x_n, x_m) \leq 1$ for all $n, m \geq N$. Then setting $R := \max\{d(x_1, x_N), \ldots, d(x_{N-1}, x_N), 1\}$, we have $(x_n)_{n\in\mathbb{N}} \subset B(x_N, R)$. This same argument does **not** work for nets. We can still find $i_0 \in I$ such that $d(x_i, x_j) \leq 1$ for all $i, j \geq i_0$, but then there are not necessarily finitely many $i \leq i_0$. For example, the net $(e^{-t})_{t\in\mathbb{R}}$ converges in $\mathbb{R}$ to zero but is not bounded.

5.3 Subnets

Subnets are the analogue of subsequences, though they are a bit more subtle.

Definition 5.17. Let $(x_i)_{i\in I}$ be a net in a topological space. Then $(y_j)_{j\in J}$ is a **subnet** of $(x_i)_{i\in I}$ if there exists a map $\sigma\colon J \to I$ such that

(i) $x_{\sigma(j)} = y_j$ for all $j \in J$;
(ii) if $j_1 \leq j_2$ then $\sigma(j_1) \leq \sigma(j_2)$; (*monotone*)
(iii) for any $i \in I$ there exists $j \in J$ such that $\sigma(j) \geq i$. (*final*)

Example 5.18. For a sequence $(x_n)_{n\in\mathbb{N}}$, any subsequence $(x_{n_k})_{k\in\mathbb{N}}$ is a subnet where $\sigma(k) = n_k$. However, because we only require the map σ to be monotone (rather than strictly monotone) there are subnets of the sequence which are **not** subsequences. For example, $(x_1, x_1, x_1, x_2, x_3, \ldots)$ is a valid subnet, even though it is not a valid subsequence. ■

Proposition 5.19. *Let X be a topological space. If a net $(x_i)_{i\in I} \subset X$ converges, then every subnet converges to the same limit.*

Proof. Let $x := \lim_i x_i$. Let $(y_j)_{j\in J}$ be a subnet with monotone final map $\sigma\colon J \to I$. Let $U \subset X$ be an open subset containing x. Then there exists $i_0 \in I$ such that $x_i \in U$ for all $i \geq i_0$. By finality there exists $j_0 \in J$ such that $\sigma(j_0) \geq i_0$. Thus, by monotonicity, for all $j \geq j_0$ we have $\sigma(j) \geq \sigma(j_0) \geq i_0$ and hence $y_j = x_{\sigma(j)} \in U$. That is, $(y_j)_{j\in J}$ converges to x. □

Finally, we conclude this note by characterizing compactness in terms of convergent subnets. This is the analogue of the fact that in a metric space a set is compact if and only if every sequence in it has a convergent subsequence (which is sometimes called being sequentially compact).

Proposition 5.20. *Let X be a topological space. Then $K \subset X$ is compact if and only if every net $(x_i)_{i\in I} \subset K$ has a convergent subnet.*

Proof. ($\Rightarrow$): Let K be compact. We recall that it has the finite intersection property: if $\{C_i\}_{i\in I}$ is a collection of closed subsets of K satisfying $\bigcap_{i\in F} C_i \neq \emptyset$ for any finite subset $F \subset I$, then $\bigcap_{i\in I} C_i \neq \emptyset$. Indeed, otherwise $\{C_i^c\}_{i\in I}$ is an open cover for K with no finite subcover.

Now, let $(x_i)_{i\in I} \subset K$ be a net. Define $C_i := \overline{\{x_j : j \geq i\}}$. Then for $F \subset I$ finite, we can find j such that $j \geq i$ for each $i \in F$ and so

$$x_j \in \bigcap_{i\in F} C_i \neq \emptyset.$$

By the finite intersection property we therefore have $\bigcap_{i\in I} C_i \neq \emptyset$. Let y be an element of this set. Then for every $i \in I$, $y \in C_i$ which means for every neighborhood U of y, $U \cap \{x_j : j \geq i\} \neq \emptyset$. That is, for every $i \in I$ and every neighborhood U, there exists $j \geq i$ such that $x_j \in U$. Set $y_{(U,j)} := x_j$. Then $(y_{(U,j)})$ is a net (where $(U, j) \leq (U', j')$ means $U \supset U'$ and $j \leq j'$), which converges to y. Defining $\sigma(U, j) := j$ yields a monotone final map and so $(y_{(U,j)})$ is a (convergent) subnet of $(x_i)_{i\in I}$.

($\Leftarrow$): Towards a contradiction, let $\{U_i : i \in I\}$ be an open cover of K with no finite subcover. Let $\mathscr{F}$ be the collection of finite subsets of I, which we make into a directed set by ordering by inclusion. For each $F \in \mathscr{F}$ let x_F be any point in $K \setminus \bigcup_{i\in F} U_i$ (which exists by virtue of there being no finite subcover). Then $(x_F)_{F\in\mathscr{F}}$ is a net and consequently has a convergent subnet $(x_{\sigma(j)})_{j\in J}$, say with limit x. Then $x \in U_i$ for some $i \in I$ and consequently there is $j_0 \in J$ such that $x_{\sigma(j)} \in U_i$ for all $j \geq j_0$. Let $j_1 \in J$ be such that $\sigma(j_1) \geq \{i\} \in \mathscr{F}$. Then there exists $j \geq j_1$ and $j \geq j_0$. For this j we have $x_{\sigma(j)} \in U_i$ but $\sigma(j) \geq \sigma(j_1) \geq \{i\}$ implies $x_{\sigma(j)} \notin U_i$, a contradiction. Thus every open cover of K has a finite subcover and K is therefore compact. $\square$

6. Appendix: Inductive Limits

An inductive limit is a way to start with many "smaller" objects that are assembled in a specific way to construct a larger object of the same type. These objects may be groups, rings, vector spaces, or algebras depending on the context. How these small pieces come together to build the limiting object is determined by a family of group homomorphisms, ring homomorphisms, linear maps, or algebra homomorphisms between these smaller objects.

As a motivating example, we consider a sequence of subgroups $(G_n)_{n\in\mathbb{N}}$ such that $G_n \leq G_{n+1}$ for every $n \in \mathbb{N}$, i.e. we have an ascending sequence of subgroups. It is left as an exercise to verify that $G = \bigcup_{n\in\mathbb{N}} G_n$ is also a group. Here, there are families of homomorphisms $\phi_{n,m} : G_n \to G_m$ whenever $n \geq m$ given by the inclusion of subgroups $G_n \subseteq G_m$ with the property that if $n \leq k \leq m$ then the composition of the intermediary maps is $\phi_{n,m} = \phi_{k,m} \circ \phi_{n,k}$. Moreover, for every $n \in \mathbb{N}$ there is a map $\phi_n : G_n \to G$. These observations form the prototype of essential properties of the inductive limit of the sequence of G_n's.

Definition 6.1. Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of objects of the same type (groups, rings, etc) along with suitable maps $\phi_{n,m} : A_n \to A_m$ whenever $n \leq m$ so that

(1) $\phi_{n,n} = \mathrm{id}_{A_n}$, and
(2) for every $n \leq k \leq m$, $\phi_{n,m} = \phi_{k,m} \circ \phi_{n,k}$.

Then the collection $(A_n, \phi_{n,m})$ is called a **directed system**. An object B along with a family of maps $\phi_n : A_n \to B$ is **compatible** if for every $n \leq m$ we have that $\phi_n = \phi_m \circ \phi_{n,m}$.

Let A along with maps $\phi_n : A_n \to A$ be compatible. A is called an **inductive limit** of the directed system $(A_n, \phi_{n,m})$, denoted $A = \varinjlim A_n$, if it satisfies the following universal property: if whenever B with maps ψ_n is compatible we have that there exists a unique map $\rho : A \to B$ so that $\psi_n = \rho \circ \phi_n$.

Exercise: check that the definition above ensures that if an inductive limit exists, it is in fact unique up to isomorphism in the appropriate sense.

Remark 6.2. Sometimes inductive limits are also called direct limits, but they should not be confused with "co-limits" or inverse limits, which are more often seen in topology, where the arrows go the other way.

Inductive limits may be generalized by allowing the family of objects $(A_i)_{i\in\mathcal{I}}$ to be indexed by an arbitrary directed set $\mathcal{I}$.

Example 6.3. Let $A_n = M_{2^n}(\mathbb{C})$ be the algebra of $2^n \times 2^n$ matrices with maps $\phi_{n,n+1} : M_{2^n}(\mathbb{C}) \to M_{2^{n+1}}(\mathbb{C})$ defined by

$$x \mapsto \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}.$$

Letting $\phi_{n,m} := \phi_{m,m-1} \circ \cdots \circ \phi_{n,n+1}$ whenever $m > n$, we see that by construction this forms a directed system. Since these are inclusions, one can identify the inductive limit with $\bigcup_{n\in\mathbb{N}} A_n$. ■

Example 6.4. Let $\mathcal{H}$ be a separable infinite dimensional Hilbert space with orthonormal basis $(\mathbf{e}_n)_{n\in\mathbb{N}}$. Let P_n be the projection onto the subspace spanned by $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$, and let $A_n = P_n B(\mathcal{H}) P_n$ with $\phi_{n,n+1} : A_n \to A_{n+1}$ the canonical inclusion map. Again, the direct limit of this system can be identified with $\bigcup_{n\in\mathbb{N}} A_n$. ■

Suppose that (A_n) is a directed system of algebraic objects (e.g. rings, algebras, vector spaces). The direct limit can be constructed by considering

$$A = \bigsqcup A_n / \sim, \tag{6.1}$$

the disjoint union of the A_n's modulo the following equivalence relation: $x_n \in A_n$ and $x_m \in A_m$ are equivalent if and only if there exist $k \geq n, m$ so that $\phi_{n,k}(x_n) = \phi_{m,k}(x_m)$. Loosely, this means that two elements are equivalent under $\sim$ if they eventually become equal in some A_k. Notice that this was automatic in the prototype example of nested groups.

While this construction is well-behaved for algebraic objects (if each A_n is a group, then the object A defined in Equation (6.1) is clearly a group), this may not be the case when working with *analytic* objects such as Hilbert spaces or Banach algebras. In these situations, one must take extra care to show that an inductive limit of the appropriate type actually exists.

7. Introduction to Operator Theory

Given a Hilbert space $\mathcal{H}$, we will explore several ideals of operators in $B(\mathcal{H})$: *finite-rank*, *compact*, *trace class*, and *Hilbert–Schmidt*.

7.1 Projections, Partial Isometries, and Positive Semi-Definite Operators We will begin by taking a closer look at certain types of operators that will be important throughout the mini-courses: *projections*, *partial isometries*, and *positive semi-definite operators*. We encountered the first two types in the prerequisite notes, while the last type is defined below.

Let $\mathcal{H}$ be a Hilbert space and $\mathcal{K} \subset \mathcal{H}$ a closed subspace. Recall that we defined the projection onto $\mathcal{K}$ as the operator $P_{\mathcal{K}} \in B(\mathcal{H})$ such that, for all $\xi \in \mathcal{H}$, we have $P_{\mathcal{K}}\xi \in \mathcal{K}$ and

$$\|\xi - P_{\mathcal{K}}\xi\| = \inf_{\eta \in \mathcal{K}} \|\xi - \eta\|.$$

We also saw that $P_{\mathcal{K}}$ is equivalently defined to be the operator such that $P_{\mathcal{K}}\xi \in \mathcal{K}$ and $\xi - P_{\mathcal{K}}\xi \in \mathcal{K}^{\perp}$ for all $\xi \in \mathcal{H}$. In particular, if $\xi \in \mathcal{K}$ then $P_{\mathcal{K}}\xi = \xi$ and if $\xi \in \mathcal{K}^{\perp}$ then $P_{\mathcal{K}}\xi = 0$ (see Exercise 1.17).

From an operator algebras perspective, there is a much more natural way to think about projections, which is given in the following theorem.

Theorem 7.1. *For $p \in B(\mathcal{H})$, p is a projection if and only if $p = p^* = p^2$. In this case, $p\mathcal{H}$ is a closed subspace and $p = P_{p\mathcal{H}}$.*

Proof. First suppose $p = P_{\mathcal{K}}$ for some closed subspace $\mathcal{K} \subset \mathcal{H}$. Since $P_{\mathcal{K}}\xi \in \mathcal{K}$ for all $\xi \in \mathcal{H}$ and $\mathcal{P}_{\mathcal{K}}\eta = \eta$ for all $\eta \in \mathcal{K}$, we have $P_{\mathcal{K}}^2\xi = P_{\mathcal{K}}(P_{\mathcal{K}}\xi) = P_{\mathcal{K}}\xi$. Thus $P_{\mathcal{K}}^2 = P_{\mathcal{K}}$. To see that $P_{\mathcal{K}} = P_{\mathcal{K}}^*$, let $\xi, \eta \in \mathcal{H}$ and use the fact that $\mathcal{K}$ is closed (and so $\mathcal{H} = \mathcal{K} \oplus \mathcal{K}^{\perp}$) to compute

$$\langle P_{\mathcal{K}}\xi, \eta\rangle = \langle P_{\mathcal{K}}\xi, P_{\mathcal{K}}\eta + (1 - P_{\mathcal{K}})\eta\rangle = \langle P_{\mathcal{K}}\xi, P_{\mathcal{K}}\eta\rangle = \langle P_{\mathcal{K}}\xi + (1 - P_{\mathcal{K}})\xi, P_{\mathcal{K}}\eta\rangle = \langle \xi, P_{\mathcal{K}}\eta\rangle .$$

So $P_{\mathcal{K}}^* = \mathcal{P}_K$.

Conversely, suppose $p \in B(\mathcal{H})$ satisfies $p = p^* = p^2$. Let $\mathcal{K} := p\mathcal{H}$. Then $\mathcal{K}$ is a subspace and moreover it is closed: suppose $(p\xi_n)_{n\in\mathbb{N}} \subset \mathcal{K}$ converges to some $\eta \in \mathcal{H}$. Then since p is continuous we have

$$p\eta = p \lim_{n\to\infty} p\xi_n = \lim_{n\to\infty} p^2\xi_n = \lim_{n\to\infty} p\xi_n = \eta.$$

Thus $\eta = p\eta \in \mathcal{K}$, and so $\mathcal{K}$ is closed. Additionally, for $\xi \in \mathcal{H}$ and $p\eta \in \mathcal{K}$ we have

$$\langle \xi - p\xi, p\eta\rangle = \langle \xi, p\eta\rangle - \langle p\xi, p\eta\rangle = \langle \xi, p\eta\rangle - \langle \xi, p^*p\eta\rangle = \langle \xi, p\eta\rangle - \left\langle \xi, p^2\eta\right\rangle = \langle \xi, p\eta\rangle - \langle \xi, p\eta\rangle = 0.$$

Thus $\xi - p\xi \in \mathcal{K}^{\perp}$, and by definition of $\mathcal{K}$ we have $p\xi \in \mathcal{K}$ for all $\xi \in \mathcal{H}$. Consequently, $p = P_{\mathcal{K}}$. □

The previous theorem allows us to redefine a projection as an operator $p \in B(\mathcal{H})$ satisfying $p = p^* = p^2$. We will typically take this more algebraic perspective in the mini-courses. Note that if $p \in B(\mathcal{H})$ is a projection, then one can use this definition to quickly verify that $1-p$ is also a projection (see Exercise 7.34). In particular, $1 - p$ is the projection onto $(p\mathcal{H})^{\perp}$.

Definition 7.2. We say two projections $p, q \in B(\mathcal{H})$ are **orthogonal** to one another if $pq = 0$. We say a family of projections $\{p_i\}_{i\in I} \subset B(\mathcal{H})$ is (**pairwise**) **orthogonal** if $p_ip_j = 0$ for all $i, j \in I$ with $i \neq j$.

For a projection $p \in B(\mathcal{H})$, observe that $p = p^2$ is equivalent to $p(1-p) = 0$. Thus p and $1-p$ are always orthogonal projections. Also $p\mathcal{H}$ and $(1 - p)\mathcal{H}$ are always orthogonal subspaces (see Exercise 7.34). This is not a coincidence: two projection $p, q \in B(\mathcal{H})$ are orthogonal if and only if $p\mathcal{H} \perp q\mathcal{H}$ (see Exercise 7.36, which explains the terminology.

We will now move onto a related class of operators: partial isometries. Recall that we defined a partial isometry as an operator $v \in B(\mathcal{H})$ that satisfies $v = vv^*v$. We will soon see why this terminology makes sense, but we first make a few observations. By taking the adjoint of each side of $v = vv^*v$, we see that $v^* = v^*vv^*$, and so (since $(x^*)^* = x$ for all $x \in B(\mathcal{H})$) v^* is also a partial isometry. Moreover, vv^* and v^*v are projections: they are both self-adjoint and

$$(vv^*)^2 = vv^*vv^* = (vv^*v)v^* = vv^*,$$

and similarly $(v^*v)^2 = v^*v$. Thus v^*v and vv^* are projections by Theorem 7.1. In fact, this gives an alternate characterization of partial isometries (see Exercise 7.37).

Definition 7.3. For a partial isometry $v \in B(\mathcal{H})$, the **source projection** of v is v^*v and the **range projection** of v is vv^*.

Note that the source projection of v is the range projection of v^*, and vice versa.

Theorem 7.4. *Let $v \in B(\mathcal{H})$ be a partial isometry and consider the subspaces $\mathcal{S} := v^*v\mathcal{H}$ and $\mathcal{R} := vv^*\mathcal{H}$. Then $v|_{\mathcal{S}}$ and $v^*|_{\mathcal{R}}$ are isometries and $v|_{\mathcal{S}^\perp} \equiv 0 \equiv v^*|_{\mathcal{R}^\perp}$. Moreover,*

$$v|_{\mathcal{S}} \colon \mathcal{S} \to \mathcal{R}$$

is an isometric isomorphism with inverse $v^*|_{\mathcal{R}}$.

Proof. By definition of $\mathcal{S}$, we have $v^*v\xi = \xi$ for any $\xi \in \mathcal{S}$. Consequently,

$$\|v\xi\|^2 = \langle v\xi, v\xi\rangle = \langle v^*v\xi, \xi\rangle = \langle \xi, \xi\rangle = \|\xi\|^2.$$

Hence $v|_{\mathcal{S}}$ is an isometry. The proof of $v^*|_{\mathcal{R}}$ is similar. Also, using $v = vv^*v$, we see that

$$v(1 - v^*v) = v - vv^*v = v - v = 0.$$

Consequently, v is identically zero on $(1 - v^*v)\mathcal{H} = \mathcal{S}^\perp$. Similarly, $v^*|_{\mathcal{R}^\perp} \equiv 0$.

To see that $v|_{\mathcal{S}}$ is valued in $\mathcal{R}$, note that $v\xi = (vv^*v)\xi = (vv^*)v\xi \in \mathcal{R}$. Finally, for $\xi \in \mathcal{S}$ we have

$$v^*(v\xi) = (v^*v)\xi = \xi.$$

Since $\mathcal{R} \subseteq v\mathcal{H}$, this shows that $v^*|_{\mathcal{R}} = v|_{\mathcal{S}}^{-1}$. □

Thus if $v \in B(\mathcal{H})$ is a partial isometry, then its restriction to $v^*v\mathcal{H}$ is an isometry while its restriction to $(1 - v^*v)\mathcal{H} = (v^*v\mathcal{H})^\perp$ is identically zero. It turns out that this is actually equivalent to v being a partial isometry (see Exercise 7.39).

Definition 7.5. We say $x \in B(\mathcal{H})$ is **positive semi-definite** if $\langle x\xi, \xi\rangle \geq 0$ for all $\xi \in \mathcal{H}$.

If $x = y^*y$ for some $y \in B(\mathcal{H})$, then x is positive semi-definite: for $\xi \in \mathcal{H}$ we have

$$\langle x\xi, \xi\rangle = \langle y^*y\xi, \xi\rangle = \langle y\xi, y\xi\rangle = \|y\xi\|^2 \geq 0.$$

In particular, if $p \in B(\mathcal{H})$ is a projection, then $p = p^2 = p^*p$ and so is positive semi-definite. We will eventually see in the C*-algebras mini-course that $x = y^*y$ for some $y \in B(\mathcal{H})$ is the *only* way for x to be positive semi-definite. Moreover, we can require that y also be positive-semidefinite, in which it is unique and we denote it $x^{1/2} := y$. For now, we can check this when $\mathcal{H}$ is finite dimensional and $B(\mathcal{H}) = M_n(\mathbb{C})$:

Example 7.6. Suppose $A \in M_n(\mathbb{C})$ is positive semi-definite. If λ is an eigenvalue of A with eigenvector $\mathbf{x}$, then

$$\lambda\|\mathbf{x}\|^2 = \langle \lambda\mathbf{x}, \mathbf{x}\rangle = \langle A\xi, \xi\rangle \geq 0.$$

Thus $\lambda \geq 0$. Now, A is self-adjoint by Exercise 7.42, and consequently it is diagonalizable: $A = UDU^*$ for U a unitary matrix and

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A counting multiplicities. Since $\lambda_i \geq 0$ for $i = 1, \ldots, n$ by the argument above, we can define

$$D^{1/2} := \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix}.$$

Note that $D^{1/2}D^{1/2} = D$ and $(D^{1/2})^* = D^{1/2}$. If we set $B := UD^{1/2}U^*$ then we have

$$B^*B = (UD^{1/2}U^*)^*(UD^{1/2}U^*) = (UD^{1/2}U^*)(UD^{1/2}U^*) = UD^{1/2}D^{1/2}U^* = UDU^* = A.$$

Thus $A = B^*B$. ■

7.2 Finite-Rank Operators All of the classes of operators we will consider in this section contain this first class:

Definition 7.7. We say $x \in B(\mathcal{H})$ is a **finite-rank** operator if $\dim(x\mathcal{H}) < \infty$. In particular, if $\dim(x\mathcal{H}) = n$ for $n \in \mathcal{H}$ we say x is a **rank** n operator. The collection of finite-rank operators on $\mathcal{H}$ is denoted $FR(\mathcal{H})$.

Example 7.8. For $\xi, \eta \in \mathcal{H}$, define $\xi \otimes \bar{\eta} \in B(\mathcal{H})$ by

$$(\xi \otimes \bar{\eta})(\zeta) = \langle \zeta, \eta \rangle \xi \qquad \zeta \in \mathcal{H}.$$

Then $\xi \otimes \bar{\eta}\mathcal{H} = \operatorname{span}\{\xi\}$, so that $\xi \otimes \bar{\eta}$ is a rank 1 operator. Observe that for $\zeta_1, \zeta_2 \in \mathcal{H}$ we have

$$\langle (\xi \otimes \bar{\eta})(\zeta_1), \zeta_2 \rangle = \langle \langle \zeta_1, \eta \rangle \xi, \zeta_2 \rangle = \langle \zeta_1, \eta \rangle \langle \xi, \zeta_2 \rangle = \langle \zeta_1, \langle \zeta_2, \xi \rangle \eta \rangle = \left\langle \zeta_1, (\eta \otimes \bar{\xi})(\zeta_2) \right\rangle.$$

Thus $(\xi \otimes \bar{\eta})^* = \eta \otimes \bar{\xi}$. In particular, $\xi \otimes \bar{\xi}$ is self-adjoint.

If $\|\xi\| = 1$, then in fact $\xi \otimes \bar{\xi}$ is a *projection*: $(\xi \otimes \bar{\xi})^2 = (\xi \otimes \bar{\xi})^* = \xi \otimes \bar{\xi}$. To see that $(\xi \otimes \bar{\xi})^2 = \xi \otimes \bar{\xi}$, we compute:

$$(\xi \otimes \bar{\xi})(\xi \otimes \bar{\xi})(\zeta) = (\xi \otimes \bar{\xi})(\langle \zeta, \xi \rangle \xi) = \langle \zeta, \xi \rangle \langle \xi, \xi \rangle \xi,$$

which equals $(\xi \otimes \bar{\xi})\zeta$ since $\|\xi\| = \langle \xi, \xi \rangle^{1/2} = 1$.

More generally, it turns out that the product of any two of these operators is also a rank 1 operator of the form $\xi \otimes \bar{\eta}$: for $\xi_1, \xi_2, \eta_1, \eta_2, \zeta \in \mathcal{H}$ we have

$$(\xi_1 \otimes \overline{\eta_1})(\xi_2 \otimes \overline{\eta_2})(\zeta) = (\xi_1 \otimes \overline{\eta_1})\left(\langle \zeta, \eta_2 \rangle \xi_2\right) = \langle \langle \zeta, \eta_2 \rangle \xi_2, \eta_1 \rangle \xi_1 = \langle \zeta, \eta_2 \rangle \langle \xi_2, \eta_1 \rangle \xi_1 = \langle \zeta, \langle \eta_1, \xi_2 \rangle \eta_2 \rangle \xi_1.$$

In other words,

$$(\xi_1 \otimes \overline{\eta_1})(\xi_2 \otimes \overline{\eta_2}) = (\langle \xi_2, \eta_1 \rangle \xi_1) \otimes \overline{\eta_2} = \xi_1 \otimes \overline{\langle \eta_1, \xi_2 \rangle \eta_2}.$$

To remember this formula, it can be helpful to think of the special case where $\eta_1 = \xi_2$ has norm 1. In that case, we have

$$(\xi_1 \otimes \overline{\eta_1})(\eta_1 \otimes \overline{\eta_2}) = \xi_1 \otimes \overline{\eta_2}.$$

Finite rank operators behave a lot like matrices. For example, we have a "Rank-Nullity" type result.

Lemma 7.9. *For $x \in FR(\mathcal{H})$, the rank of x is $\dim(\ker(x)^\perp)$.*

Proof. First note that since $\mathcal{H} = \ker(x) \oplus \ker(x)^\perp$, we have $x\mathcal{H} = x\ker(x)^\perp$, and so the rank of x is $\dim(x\ker(x)^\perp) \leq \dim(\ker(x)^\perp)$.

On the other hand, let $\mathcal{E}$ be an orthonormal basis for $\ker(x)^\perp$. Since x is injective on $\ker(x)^\perp$, $|\mathcal{E}| = |x\mathcal{E}|$. So, we just need to show that the set $\{x\xi \colon \xi \in \mathcal{E}\}$ is linearly independent. If $\alpha_j \in \mathbb{C}$ and $\xi_j \in \mathcal{E}$ are such that

$$\sum_{j=1}^k \alpha_j x \xi_j = 0,$$

then $\sum_{j=1}^k \alpha_j \xi_j \in \ker(x) \cap \ker(x)^\perp$, and so we must have $\alpha_1 = \cdots = \alpha_k = 0$ since $\xi_1, \ldots, \xi_k$ are orthonormal. That means $|\{x\xi \colon \xi \in \mathcal{E}\}| \leq n$. $\square$

It turns out the operators in the previous example give *all* rank 1 operators. In fact, we have

Theorem 7.10. *If $x \in FR(\mathcal{H})$ is rank n for $n \in \mathbb{N}$, then there exist $\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n \in \mathcal{H}$ so that*

$$x = \sum_{j=1}^n \xi_j \otimes \overline{\eta_j}.$$

In particular, if x is the projection onto a finite dimensional subspace $\mathcal{K} \subset \mathcal{H}$, then

$$x = \sum_{j=1}^n \xi_j \otimes \overline{\xi_j}$$

for any orthonormal basis $\xi_1, \ldots, \xi_n$ for $\mathcal{K}$.

Proof. Let $\{\eta_1, \ldots, \eta_n\}$ be an orthonormal basis for $\ker(x)^\perp$, and set $\xi_j := x\eta_j$ for each $j = 1, \ldots, n$. For $\zeta \in \mathcal{H}$, we can write

$$\zeta = \sum_{j=1}^{n} \alpha_j \eta_j + \zeta_0$$

for some $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ and $\zeta_0 \in \ker(x)$. We then have

$$\left(\sum_{j=1}^{n} \xi_j \otimes \bar{\eta}_j\right)(\zeta) = \sum_{j=1}^{n} \langle \zeta, \eta_j \rangle \xi_j = \sum_{j=1}^{n} \left\langle \sum_{k=1}^{n} \alpha_k \xi_k + \zeta_0, \eta_j \right\rangle \xi_j x\zeta = \sum_{j=1}^{n} \alpha_j \langle \eta_j, \eta_j \rangle \xi_j \sum_{j=1}^{n} \alpha_j x\eta_j = x\zeta$$

Thus $x = \sum_{j=1}^{n} \xi_j \otimes \bar{\eta}_j$.

If x is a projection onto a subspace $\mathcal{K} \subset \mathcal{H}$ with orthonormal basis $\xi_1, \ldots, \xi_n$, then in the above argument we have $\eta_j = x\xi_j = \xi_j$. Thus x has the claimed form. □

Corollary 7.11. *$FR(\mathcal{H})$ is a (not necessarily norm-closed) two-sided $*$-ideal in $B(\mathcal{H})$.*

Proof. Let $x \in B(\mathcal{H})$. By the previous theorem, it suffices to show $x(\xi \otimes \bar{\eta}), (\xi \otimes \bar{\eta})x \in FR(\mathcal{H})$ for $\xi, \eta \in \mathcal{H}$. For $\zeta \in \mathcal{H}$ we have

$$x(\xi \otimes \bar{\eta})(\zeta) = x \langle \zeta, \eta \rangle \xi = \langle \zeta, \eta \rangle x\xi = ((x\xi) \otimes \bar{\eta})(\zeta).$$

So $x(\xi \otimes \bar{\eta}) = (x\xi) \otimes \bar{\eta} \in FR(\mathcal{H})$. Likewise one can show $(\xi \otimes \bar{\eta})x = \xi \otimes \overline{x^*\eta} \in FR(\mathcal{H})$. □

Corollary 7.12. *Any nonzero two-sided ideal in $B(\mathcal{H})$ contains $FR(\mathcal{H})$ – even if the ideal is not $*$-closed or norm-closed.*

Proof. Let $0 \neq J \triangleleft B(\mathcal{H})$. By Theorem 7.10, it suffices to show that J contains all rank one projections. (Pause and convince yourself of this.) Let $x \in J$ be a nonzero operator, $\xi \in \mathcal{H}$ such that $x\xi \neq 0$, and $p \in B(\mathcal{H})$ a rank one projection. Then by Theorem 7.10, $p = \eta \otimes \bar{\eta}$ for some $\eta \in \mathcal{H}$. Let $y \in B(\mathcal{H})$ such that $yx\xi = \eta$. (What would such an operator look like?) Then for any $\zeta \in \mathcal{H}$,

$$yx(\xi \otimes \bar{\xi})x^*y^*\zeta = yx\langle x^*y^*\zeta, \xi\rangle\xi = \langle x^*y^*\zeta, \xi\rangle yx\xi = \langle \zeta, yx\xi\rangle yx\xi = \langle \zeta, \eta\rangle\eta = p\zeta.$$

So, $p = yx(\xi \otimes \bar{\xi})x^*y^* \in J$. Since p was arbitrary, it follows that J contains all rank one projections. □

Exercise 7.13. Show that for any $*$-closed ideal $J \triangleleft B(\mathcal{H})$, if J contains one rank one projection, then J contains all of $FR(\mathcal{H})$.

If $\mathcal{H}$ is finite dimensional, say with $\dim(\mathcal{H}) = n$, then $FR(\mathcal{H}) = B(\mathcal{H})$ since every operator is at most rank n. However, if $\mathcal{H}$ is infinite dimensional then $FR(\mathcal{H})$ is not closed. Indeed, let $\{\xi_n\}_{n\in\mathbb{N}} \subset \mathcal{H}$ be an orthonormal set and let $(\alpha_n)_{n\in\mathbb{N}} \in c_0(\mathbb{N})$. Recall from Example 7.8 that $\xi_n \otimes \overline{\xi_n}$ is a projection and so has $\|\xi_n \otimes \bar{\xi}_n\| = 1$. From Exercise 7.51, it follows that

$$x := \sum_{n=1}^{\infty} \alpha_n \xi_n \otimes \overline{\xi_n}$$

is the norm limit of the partial sums (which are finite-rank operators). Since $x\xi_j = \alpha\xi_j$, we see that $\dim(x\mathcal{H}) = \infty$ and so x is not finite-rank. However, it is an example of the next class of operators.

7.3 Compact Operators As we specified in the Prerequisite Notes, we will primarily be concerned with norm-closed ideals. The *compact operators* are the norm closure of $FR(\mathcal{H})$.

Definition 7.14. We say $x \in B(\mathcal{H})$ is compact if $x = \lim_i x_i$ is the norm limit of a sequence of finite-rank operators. The collection of compact operators on $\mathcal{H}$ is denoted $K(\mathcal{H})$.

Of course, the distinction between $FR(\mathcal{H})$ and $K(\mathcal{H})$ is nonexistent when $\mathcal{H}$ is finite dimensional.

Exercise 7.15. How would $K(\mathcal{H})$ change if we took it to be the set of limits of *nets* of finite-rank operators?

Exercise 7.16. Prove that $K(\mathcal{H})$ is a (norm-closed) ideal in $B(\mathcal{H})$.

It follows from Corollary 7.12 that every norm-closed 2-sided ideal in $B(\mathcal{H})$ contains $K(\mathcal{H})$. In particular, Exercise 7.13 tells us that if p is any rank-one projection, then the norm-closed ideal generated by p is precisely $K(\mathcal{H})$.

Exercise 7.17. Fix a rank one operator $\xi\otimes\overline{\zeta}$ in $FR(\mathcal{H})$. Show that the $*$-closed, norm-closed ideal of $B(\mathcal{H})$ generated by $\xi\otimes\overline{\zeta}$ is $K(\mathcal{H})$.

Moreover, we shall see later in the C*-algebra course that if $J \triangleleft B(\mathcal{H})$ and $I \triangleleft J$, then $I \triangleleft B(\mathcal{H})$. It follows then that $K(\mathcal{H})$ is simple.

In fact, it turns out that when $\mathcal{H}$ has a countable basis, then $K(\mathcal{H})$ is the unique closed 2-sided ideal in $B(\mathcal{H})$. That means that the quotient $B(\mathcal{H})/K(\mathcal{H})$ is also simple. This quotient is called the *Calkin algebra* and has played a fundamental role in the theory of extensions of C*-algebras.

As seen in Example 3.15, the spectrum of an operator is a generalization of the notion of eigenvalues for matrices. The following definition gives another, more explicit generalization.

Definition 7.18. For $x \in B(\mathcal{H})$, the **point spectrum** of x is the set

$$\sigma_p(x) := \{\alpha \in \mathbb{C} \colon \exists \xi \in \mathcal{H}\setminus\{0\} \text{ satisfying } x\xi = \alpha\xi\}.$$

For any $\alpha \in \sigma_p(x)$, $x-\alpha$ has a non-trivial kernel and therefore is not invertible. Hence we always have the inclusion $\sigma_p(x) \subset \sigma(x)$. The reverse inclusion holds when $\mathcal{H}$ is finite dimensional (by Example 3.15.(1)), but does not hold in general.

Example 7.19. Let $S \in B(\ell^2(\mathbb{N}))$ be the shift operator:

$$S(a_1, a_2, \dots) = (0, a_1, a_2, \dots)$$

From Exercise 1.41 we know S is an isometry but not a unitary. That is, S is injective but not surjective. Consequently, $0 \in \sigma(S)\setminus\sigma_p(S)$. ■

When $\mathcal{H}$ is finite-dimensional, we can diagonalize any normal element $x \in B(\mathcal{H})$. In other words, there exists a unitary $u \in B(\mathcal{H})$ such that u^*xu is a diagonal operator whose entries are the eigenvalues of x. In infinite dimensions, this is no longer necessarily true. However, we do have the following result. Because it often appears in a graduate real analysis course, we will omit the proof. But one can be found in [Conway, Section 7.4].

Theorem 7.20 (Spectral Theorem for Compact Normal Operators)**.** *Let $x \in K(\mathcal{H})$ be normal. Then $\sigma_p(x)$ is countable, has no non-zero cluster points, and for each $\lambda \in \sigma_p(x)$ the subspace* $\ker(x-\lambda)$ *is finite-dimensional. Let $p_\lambda \in FR(\mathcal{H})$ be the projection onto* $\ker(x-\lambda)$*. Then $\{p_\lambda \colon \lambda \in \sigma_p(x)\}$ are pairwise orthogonal, sum in the SOT to* 1*, and*

$$x = \sum_{\lambda\in\sigma_p(x)} \lambda p_\lambda$$

where the series converges in the norm topology.

Observe that in the above theorem, $p_\lambda \in FR(\mathcal{H})$ for each $\lambda \in \sigma_p(x)$ implies that for any finite $F \subset \sigma_p(x)$, the partial sum $\sum_{\lambda\in F}\lambda p_\lambda$ lies in $FR(\mathcal{H})$. This gives an explicit description of a compact normal operator x as the norm limit of finite-rank operators.

7.4 Trace Class Operators Let $\mathcal{E} \subset \mathcal{H}$ be an orthonormal basis. For $x \in B(\mathcal{H})$ positive semi-definite, we have $\langle x\xi, \xi\rangle \geq 0$ for all $x \in \mathcal{H}$. Consequently, we can define

$$\operatorname{Tr}(x) := \sum_{\xi\in\mathcal{E}} \langle x\xi, \xi\rangle \in [0, \infty]$$

It turns out the above formula does not depend on the particular orthonormal basis (see Exercise 7.54).

Example 7.21. Let $\mathcal{K} \subset \mathcal{H}$ be a finite dimensional subspace and let $x \in B(\mathcal{H})$ be the projection onto $\mathcal{K}$. We claim $\operatorname{Tr}(x) = \dim(\mathcal{K})$. Indeed, let $\{\xi_1, \dots, \xi_n\} \subset \mathcal{K}$ be an orthonormal basis for $\mathcal{K}$ so that

$$x = \sum_{j=1}^n \xi_j \otimes \overline{\xi_j}$$

by Theorem 7.10. Extend $\{\xi_1, \dots, \xi_n\}$ to an orthonormal basis $\mathcal{E}$ for $\mathcal{H}$. We have

$$\operatorname{Tr}(x) = \sum_{\xi\in\mathcal{E}} \left\langle \sum_{j=1}^n (\xi_j\otimes\overline{\xi_j})(\xi), \xi \right\rangle = \sum_{j=1}^n \|\xi_j\|^2 = n = \dim(\mathcal{K}),$$

as claimed. ∎

Given $x \in B(\mathcal{H})$, recall that x^*x is positive semi-definite and so we can consider $|x| := \sqrt{x^*x}$, which we call the *absolute value* of x.

Definition 7.22. We say $x \in B(\mathcal{H})$ is **trace class** if

$$\|x\|_1 := \mathrm{Tr}(|x|) < \infty.$$

The collection of trace class operators is denoted $L^1(B(\mathcal{H}))$.

The notation $L^1(B(\mathcal{H}))$ is motivated by a concept in von Neumann algebras called the *predual*, which will be explored later in the von Neumann algebras mini-course. For now, let the following example serve as justification for the notation.

Example 7.23. Let $\mathcal{H}$ be a separable Hilbert space and let $\{\xi_n \in \mathcal{H} \colon n \in \mathbb{N}\}$ be an orthonormal basis. For any $(\alpha_n)_{n\in\mathbb{N}} \in \ell^1(\mathbb{N})$,

$$x := \sum_{n=1}^{\infty} \alpha_n \xi_n \otimes \overline{\xi_n} \in L^1(B(\mathcal{H})).$$

Indeed, first note that the series defines a bounded operator by Exercise 7.51 (using the fact that $\ell^1(\mathbb{N}) \subset c_0(\mathbb{N})$). In order to compute $\|x\|_1$, we must first determine $|x|$. By Example 7.8 we have

$$x^*x = \sum_{m,n=1}^{\infty} \overline{\alpha_m}\alpha_n (\xi_m \otimes \overline{\xi_m})(\xi_n \otimes \overline{\xi_n}) = \sum_{m,n=1}^{\infty} \overline{\alpha_m}\alpha_n (\langle \xi_n, \xi_m\rangle \xi_m) \otimes \overline{\xi_n} = \sum_{n=1}^{\infty} |\alpha_n|^2 \xi_n \otimes \overline{\xi_n}.$$

By a similar computation, the quantity on the right is the square of the positive semi-definite operator $\sum |\alpha_n| \xi_n \otimes \overline{\xi_n}$. So by the uniqueness of $|x|$, we obtain

$$|x| = \sum_{n=1}^{\infty} |\alpha_n| \xi_n \otimes \overline{\xi_n}.$$

We then compute

$$\begin{aligned} \|x\|_1 = \mathrm{Tr}(|x|) = \sum_{n=1}^{\infty} \langle |x|\xi_n, \xi_n\rangle &= \sum_{n=1}^{\infty} \left\langle \sum_{m=1}^{\infty} |\alpha_m| (\xi_m \otimes \overline{\xi_m})(\xi_n), \xi_n \right\rangle \\ &= \sum_{n=1}^{\infty} \langle |\alpha|_n \xi_n, \xi_n\rangle = \sum_{n=1}^{\infty} |\alpha_n| = \|(\alpha_n)_{n\in\mathbb{N}}\|_1 < \infty, \end{aligned}$$

and so $x \in L^1(B(\mathcal{H}))$ as claimed. We also note that by a similar computation

$$\sum_{n=1}^{\infty} \langle x\xi_n, \xi_n\rangle = \sum_{n=1}^{\infty} \alpha_n,$$

which converges. ∎

The series obtained at the end of the above example should be viewed as $\mathrm{Tr}(x)$. It turns out that for any $x \in L^1(B(\mathcal{H}))$ and any orthonormal basis $\mathcal{E} \subset \mathcal{H}$ one has

$$\sum_{\xi\in\mathcal{E}} |\langle x\xi, \xi\rangle| \leq \sum_{\xi\in\mathcal{E}} \langle |x|\xi, \xi\rangle = \|x\|_1.$$

We will delay proving this until later on in the mini-courses, but for now note that it implies the series $\sum_{\xi\in\mathcal{E}} \langle x\xi, \xi\rangle$ is absolutely convergent. Thus we make the following definition:

Definition 7.24. For $x \in L^1(B(\mathcal{H}))$, the **trace** of x is the quantity

$$\mathrm{Tr}(x) := \sum_{\xi\in\mathcal{E}} \langle x\xi, \xi\rangle$$

where $\mathcal{E}$ is an orthonormal basis for $\mathcal{H}$.

Just as with positive semi-definite operators, the trace is independent of the choice of orthonormal basis for $\mathcal{H}$. As the name suggests, when $\mathcal{H}$ is finite dimensional this quantity agrees with the usual trace on $M_n(\mathbb{C})$ (see Exercise 7.55). Moreover, like the trace on $M_n(\mathbb{C})$ it is invariant under cyclic permutations: $\mathrm{Tr}(xy) = \mathrm{Tr}(yx)$. We collect this fact along with several others in the next theorem, whose proof is delayed until later in the mini-courses.

Theorem 7.25. *On $L^1(B(\mathcal{H}))$, $\|\cdot\|_1$ is a norm satisfying $\|x\| \leq \|x\|_1$. $L^1(B(\mathcal{H}))$ is a (not necessarily closed) $*$-ideal in $B(\mathcal{H})$ such that $\|x^*\|_1 = \|x\|_1$, $|Tr(x)| \leq \|x\|_1$,*

$$\|axb\|_1 \leq \|a\|\|b\|\|x\|_1 \qquad a, b \in B(\mathcal{H}),\ x \in L^1(B(\mathcal{H})),$$

and

$$Tr(ax) = Tr(xa) \qquad a \in B(\mathcal{H}),\ x \in L^1(B(\mathcal{H})).$$

Moreover, $L^1(B(\mathcal{H}))$ equipped with $\|\cdot\|_1$ is a Banach algebra for which $FR(\mathcal{H})$ is a dense subalgebra.

We leave the proof of the following corollary as an exercise (see Exercise 7.56).

Corollary 7.26. *For any Hilbert space $\mathcal{H}$, $L^1(B(\mathcal{H})) \subset K(\mathcal{H})$.*

7.5 Hilbert–Schmidt Operators

Definition 7.27. We say $x \in B(\mathcal{H})$ is a **Hilbert–Schmidt** operator if

$$\|x\|_2 := \mathrm{Tr}(x^*x)^{1/2} < \infty.$$

The collection of Hilbert–Schmidt operators is denoted $HS(\mathcal{H})$.

Another way to characterize when $x \in B(\mathcal{H})$ is Hilbert–Schmidt is if $x^*x \in L^1(B(\mathcal{H}))$. Consequently, sometimes the notation $L^2(B(\mathcal{H}))$ is used for $HS(\mathcal{H})$. Since $L^1(B(\mathcal{H}))$ is closed under multiplication, we have $L^1(B(\mathcal{H})) \subset HS(\mathcal{H})$. Indeed, by Theorem 7.25 we have

$$\mathrm{Tr}(x^*x) = \|x^*x\|_1 \leq \|x^*\|\|x\|_1 = \|x\|\|x\|_1 \leq \|x\|_1^2.$$

Thus

$$\|x\|_2 = \mathrm{Tr}(x^*x)^{1/2} \leq \|x\|_1 \qquad \forall x \in L^1(B(\mathcal{H})).$$

In particular, we have $FR(\mathcal{H}) \subset HS(\mathcal{H})$.

Example 7.28. Let $\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n \in \mathcal{H}$ and set

$$x := \sum_{j=1}^n \xi_j \otimes \overline{\eta_j}.$$

Let $\mathcal{E} \subset \mathcal{H}$ be an orthonormal basis. Then

$$\begin{aligned}\|x\|_2^2 = \mathrm{Tr}(x^*x) &= \sum_{\xi\in\mathcal{E}} \langle x^*x\xi, \xi\rangle = \sum_{\xi\in\mathcal{E}} \langle x\xi, x\xi\rangle = \sum_{\xi\in\mathcal{E}} \left\langle \sum_{j=1}^n \langle \xi, \eta_j\rangle \xi_j, \sum_{k=1}^n \langle \xi, \eta_k\rangle \xi_k \right\rangle \\ &= \sum_{i,j=1}^n \langle \xi_j, \xi_k\rangle \sum_{\xi\in\mathcal{E}} \langle \eta_k, \xi\rangle \langle \xi, \eta_j\rangle = \sum_{i,j=1}^n \langle \xi_j, \xi_k\rangle \left\langle \sum_{\xi\in\mathcal{E}} \langle \eta_k, \xi\rangle \xi, \eta_j \right\rangle = \sum_{i,j=1}^n \langle \xi_j, \xi_k\rangle \langle \eta_k, \eta_j\rangle.\end{aligned}$$

Observe that final sum is precisely the square of the norm of $\sum \xi_j \otimes \overline{\eta_j}$ when viewed as a vector in the Hilbert space tensor product $\mathcal{H} \otimes \overline{\mathcal{H}}$ (recall from Exercise 1.10 that $\overline{\mathcal{H}}$ is the conjugate Hilbert space to $\mathcal{H}$). ■

The observation made at the end of the above example is precisely why the notation $\xi \otimes \bar{\eta}$ was used in the first place. It also hints that $HS(\mathcal{H})$ has a Hilbert space structure itself.

Theorem 7.29. *On $HS(\mathcal{H})$, $\|\cdot\|_2$ is a norm satisfying $\|x\| \leq \|x\|_2$. $HS(\mathcal{H}))$ is a (not necessarily closed) $*$-ideal in $B(\mathcal{H})$ such that $\|x^*\|_2 = \|x\|_2$ and*

$$\|axb\|_2 \leq \|a\|\|b\|\|x\|_2 \qquad a, b \in B(\mathcal{H}),\ x \in HS(\mathcal{H}).$$

For $x, y \in HS(\mathcal{H})$, $xy, yx \in L^1(B(\mathcal{H}))$ with $Tr(xy) = Tr(yx)$. Consequently, $HS(\mathcal{H})$ is a Hilbert space with inner product

$$\langle x, y\rangle_2 := Tr(y^*x) \qquad x, y \in HS(\mathcal{H}),$$

and $FR(\mathcal{H})$ is a dense subspace. As a Hilbert space, $HS(\mathcal{H})$ is isomorphic to $\mathcal{H} \otimes \overline{\mathcal{H}}$.

We leave the proof of the following corollary as an exercise (see Exercise 7.56).

Corollary 7.30. *For any Hilbert space $\mathcal{H}$, $L^1(B(\mathcal{H})) \subset HS(\mathcal{H}) \subset K(\mathcal{H})$.*

Exercises

In what follows $\mathcal{H}$ is a Hilbert space.

Exercise 7.31. For $x \in B(\mathcal{H})$, show that $\|x\| = \|x^*\|$.

Exercise 7.32. For $x \in B(\mathcal{H})$, show that $\|x\| = \|x^*x\|^{1/2}$.

Exercise 7.33. Suppose $(x_n)_n \in B(\mathcal{H})$ is a sequence that converges in norm to $x \in B(\mathcal{H})$. Show that x_n^* also converges to x^*.

Exercise 7.34. If $p \in B(\mathcal{H})$ is a projection, show that $1-p$ is also a projection and $(1-p)\mathcal{H} = (p\mathcal{H})^\perp$.

Exercise 7.35. If $p, q \in B(\mathcal{H})$ are projections such that $q\mathcal{H} \subset p\mathcal{H}$, show that $p-q$ is a projection with $p\mathcal{H} = q\mathcal{H} \oplus (p-q)\mathcal{H}$.

Exercise 7.36. For projections $p, q \in B(\mathcal{H})$, show that $pq = 0$ if and only if $p\mathcal{H} \perp q\mathcal{H}$.

Exercise 7.37. Show that $v \in B(\mathcal{H})$ is a partial isometry if and only if v^*v is a projection. [**Hint:** expand $\|(v - vv^*v)\xi\|^2$ for $\xi \in \mathcal{H}$.]

Exercise 7.38. Let $\mathcal{H}$ be a Hilbert space and $\mathcal{H}_1, \mathcal{H}_2$ two subspaces of $\mathcal{H}$. Suppose there is a bijective isometry $\tilde{u} : \mathcal{H}_1 \to \mathcal{H}_2$. Show that for any $\xi, \eta \in \mathcal{H}_1$, we have $\langle u\xi, u\eta\rangle = \langle \xi, \eta\rangle$.

Hint: Use the polarization identity: for any $\delta, \zeta \in \mathcal{H}$,

$$4\langle \delta, \zeta\rangle = \sum_{k=0}^{3} \mathrm{i}^k \|\delta + \mathrm{i}^k \zeta\|^2.$$

Exercise 7.39.

Exercise 7.40. Let $v, w \in B(\mathcal{H})$ be partial isometries. Show that if $ww^*\mathcal{H} \subset v^*v\mathcal{H}$, then vw is a partial isometry.

Exercise 7.41. Let $p \in B(\mathcal{H})$ be a projection of rank n. Show that $pB(\mathcal{H})p \simeq B(p\mathcal{H}) \simeq M_n(\mathbb{C})$.

Exercise 7.42. Let $x \in B(\mathcal{H})$.

(1) Show that $x = x^*$ if and only if $\langle x\xi, \xi\rangle \in \mathbb{R}$ for all $\xi \in \mathcal{H}$.
(2) Show that if x is positive semi-definite, then x is self-adjoint.

Exercise 7.43. For $x \in B(\mathcal{H})$ with $x = x^*$, show that

$$\sup_{\|\xi\|=1} |\langle x\xi, \xi\rangle| = \|x\|.$$

[**Hint:** show $\operatorname{Re}\langle x\xi, \eta\rangle = \frac{1}{2}\langle x(\xi+\eta), \xi+\eta\rangle + \frac{1}{2}\langle x(\xi-\eta), \xi-\eta\rangle$ for all $\xi, \eta \in \mathcal{H}$.]

Exercise 7.44. For $x \in B(\mathcal{H})$, show that $\ker(x^*) = (x\mathcal{H})^\perp$.

Exercise 7.45. For $x \in B(\mathcal{H})$, if $\langle x\xi, \eta\rangle = 0$ for all $\xi, \eta \in \mathcal{H}$, then $x = 0$. What if $\langle x\xi, \xi\rangle = 0$ for all $\xi \in \mathcal{H}$?

Exercise 7.46.

Exercise 7.47. Let $x, y \in B(\mathcal{H})$ be self-adjoint. We say $x \geq y$ if $x - y$ is positive semi-definite. This gives us a partial order on the the collection of self-adjoint operators on $B(\mathcal{H})$. Show the following for $x \in B(\mathcal{H})$.

(a) $x, y \geq 0 \Rightarrow x + y \geq 0$.
(b) $x \geq y$, $z = z^* \Rightarrow x + z \geq y + z$.
(c) $x, y \geq 0 \nRightarrow xy \geq 0$.
(d) $x = x^* \Rightarrow x^2 \geq 0$.
(e) $x \geq y \geq 0 \nRightarrow x^2 \geq y^2$.

Exercise 7.48. For $x \in B(\mathcal{H})$, its *numerical range* is the set $W(x) = \{\langle x\xi, \xi\rangle \ : \ \|\xi\| = 1\}$ and its numerical radius is $w(x) = \sup\{|z| \ : \ z \in W(x)$. Show that $\frac{1}{2}\|x\| \leq w(x) \leq \|x\|$.

Exercise 7.49. For $\xi, \eta \in \mathcal{H}$, show that $\|\xi \otimes \bar{\eta}\| \leq \|\xi\|\|\eta\|$.

Exercise 7.50.

Exercise 7.51. Let $\{p_n : n \in \mathbb{N}\} \subset B(\mathcal{H})$ be a family of pairwise orthogonal projections.

(a) For $m < n$ and $\alpha_m, \alpha_{m+1}, \ldots, \alpha_n \in \mathbb{C}$, show that

$$\left\| \sum_{j=m}^{n} \alpha_j p_j \right\| = \max_{m \leq j \leq n} |\alpha_j|.$$

(b) For $(\alpha_n)_{n\in\mathbb{N}} \in c_0(\mathbb{N})$, show that

$$\left(\sum_{j=1}^{n} \alpha_j p_j \right)_{n\in\mathbb{N}}$$

is a Cauchy sequence (with respect to the metric induced by the operator norm).

Exercise 7.52. Let $\mathcal{E} \subset \mathcal{H}$ be an orthonormal basis. For $x \in B(\mathcal{H})$ show that

$$\sum_{\xi\in\mathcal{E}} \|x\xi\|^2 = \sum_{\xi\in\mathcal{E}} \|x^*\xi\|^2.$$

[**Hint:** use Theorem 1.22 and Fubini's theorem.]

Exercise 7.53. Suppose $x, y \in B(\mathcal{H})$ are such that $x = y^*y$. Show that x is positive semi-definite.

Exercise 7.54. Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{E}$ be an orthonormal basis.

(a) Show that if $u \in B(\mathcal{H})$ is a unitary operator, then $\{u\xi : \xi \in \mathcal{E}\}$ is an orthonormal basis.

(b) Show that for any other orthonormal basis $\mathcal{F}$, there exists a unitary operator $u \in B(\mathcal{H})$ such that $u\mathcal{E} = \mathcal{F}$.

(c) Assume for $x \in B(\mathcal{H})$ that $x = y^*y$ for some $y \in B(\mathcal{H})$. Show that for any $u \in B(\mathcal{H})$ unitary,

$$\sum_{\xi\in\mathcal{E}} \langle xu\xi, u\xi\rangle = \sum_{\xi\in\mathcal{E}} \langle x\xi, \xi\rangle\,.$$

[**Hint:** use $yy^* = yuu^*y^*$ and Exercise 7.52.]

Exercise 7.55. For $\mathcal{H} = \mathbb{C}^n$ and $A = (A_{i,j})_{i,j=1}^n \in M_n(\mathbb{C})$, show that the trace defined in Definition 7.24 is given by

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} A_{i,i}.$$

Exercise 7.56. For a Hilbert space $\mathcal{H}$, prove the inclusions

$$FR(\mathcal{H}) \subset L^1(B(\mathcal{H})) \subset HS(\mathcal{H}) \subset K(\mathcal{H}).$$

[**Hint:** approximate by finite-rank operators in the appropriate norm.]

Exercise 7.57. Fill in the following table. (Group exercise.) Note that some cells can have more than one correct answer.

	Algebraic Definition	Spatial Definition
Normal	$TT^* = T^*T$	
Self-Adjoint	$T = T^*$	(hint: E.7.42)
Projection	$T = T^2 = T^*$	T is an orthogonal projection onto a closed subspace of H
Invertible		
Unitary	$T^*T = I = TT^*$	
Isometry		$\|T\xi\| = \|\xi\|$ for all $\xi \in \mathcal{H}$
Co-Isometry	$TT^* = I$	
Partial Isometry	$T = TT^*T$	For some closed subspace $\mathcal{K} \subset \mathcal{H}$, $T\|_{\mathcal{K}}$ is an isometry and $T\|_{\mathcal{K}^\perp} \equiv 0$