FUNCTIONAL ANALYSIS

G.H.

Contents

1. Introduction 1
2. Normed and Banach spaces 1
  2.1. Beginnings, metrics spaces and examples 1
  2.2. Convergence and linear operators 3
  2.3. Separable spaces 12
  2.4. Hahn-Banach theorems and applications 12
3. Fundamental results in functional analysis 15
  3.1. Baire Category Theorem 15
  3.2. Modes of convergence 19
4. Hilbert spaces 22
  4.1. Inner product spaces 22
  4.2. Bounded linear operators on a Hilbert space 29
  4.3. Spectral theory of operators 31
  4.4. Arbitrary measures, Radon measures, and more spectral theory 48
5. Addendum 51
  5.1. Addendum and errata 51

1. Introduction

These notes were taken in University of Delaware's MATH806 (Functional Analysis) course, taught by Dr. Ivan Todorov in Fall 2021. I typed them based on hand-written notes taken during class each week- the hope was that a typed version would provide a better record in the future and be much more useful. Dr. Todorov's lecture notes were self-contained, though we took material from:

- *Linear Analysis, Béla Bollobás*
- *General Topology, John Kelley*

These notes are a work in progress; all mistakes are mine and mine alone (either through mistyping or a misunderstanding of the material). If you have any error corrections, tips, or general comments, please reach out to me at: ghoefer@udel.edu.

2. Normed and Banach spaces

2.1. **Beginnings, metrics spaces and examples.** As a bit of background introduction, we may think of functional analysis as the study of functions defined on other functions. We'll take a brief look at some specific applications.

Applications:

(i) PDEs, ODEs, and integral equations: these are (in general) specific operators, and the purpose is to use functional analysis to solve equations of operators.

Date: September 2021.

Definition 2.1. *An integral equation is an equation of the form*

$$\int k(x,y)f(x)dx = g(y),$$

where g is known, while f is not.

The purpose for integral equations as above is to solve for f when g is known.

General context:

Functions $\rightsquigarrow$ vectors

We then apply operators to vectors- as a specific example, we can apply the operator

$$T = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

to the vector (function) h, obtaining the function Th. Here the specific operator T is known as the *Laplacian operator*.

For functional analysis, we begin with linear transformations T defined on infinite dimensional spaces. We can roughly state:

Functional analysis $\approx$ Infinite dimensional linear algebra $+$ analysis.

Recall the following definition.

Definition 2.2. *A metric space (X, d) is a set X equipped with function $d : X \times X \to \mathbb{R}_+$ such that for all $x, y, z \in X$:*

- *(i)* $d(x,z) \leq d(x,y) + d(y,z)$;
- *(ii)* $d(x,y) = d(y,x)$;
- *(iii)* $d(x,y) = 0$ *if and only if* $x = y$.

Example 2.3.

- *(i) The space*

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

  with metric

  $$d(f,g) = \sup_{x \in [a,b]} |f(x) - g(x)|.$$

- *(ii) The space $\mathbb{C}^n$, with elements $(x_i)_{i=1}^n$ and $x_i \in \mathbb{C}$. There are many possible metrics which we can apply to $\mathbb{C}^n$- here we consider the Euclidean metric*

  $$d((x_i),(y_i)) = \left(\sum_{i=1}^n |x_i - y_i|^2\right)^{1/2}.$$

The previous two examples share a common feature- namely, both underlying spaces $\mathbb{C}^n$ and $C_{[a,b]}$ are vector spaces. The metrics in these cases are derived through *norms.*

Definition 2.4. *A normed space is a pair $(X, \|\cdot\|)$, where X is a vector space and $\|\cdot\| : X \to \mathbb{R}_+$ such that*

- *(i)* $\|x + y\| \leq \|x\| + \|y\|$;
- *(ii)* $\|\lambda x\| = |\lambda|\|x\|$;
- *(iii)* $\|x\| = 0$ *only if* $x = 0$,

for all $x, y \in X$ and $\lambda \in \mathbb{F}$. We call $\|\cdot\|$ a norm.

Example 2.5.

- *(i) As seen previously, the spaces $C_{[a,b]}$ and $\mathbb{C}^n$ are both normed metric spaces; the latter's norm is induced by the Euclidean metric.*

(ii) $(\mathbb{C}^n, \|\cdot\|_1)$ *with*

$$\|(x_i)_{i=1}^n\|_1 = \sum_{i=1}^n |x_i|.$$

(iii) $(\mathbb{C}^n, \|\cdot\|_p)$ *with*

$$\|(x_i)_{i=1}^n\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}.$$

(iv) $(\mathbb{C}^n, \|\cdot\|_\infty)$ *with*

$$\|(x_i)_{i=1}^n\|_\infty = \max_{1\le i\le n} |x_i|.$$

(v) ℓ^p *spaces, where* $1 \le p < +\infty$. *These are the spaces defined by*

$$\ell^p = \{(x_i)_{i=1}^\infty : \ x_i \in \mathbb{C}, \ \sum_{i=1}^\infty |x_i|^p < \infty\}.$$

The norm on the space is given by

$$\|(x_i)_{i=1}^\infty\|_p = \left(\sum_{i=1}^\infty |x_i|^p\right)^{1/p}.$$

We can also define the space when $p = \infty$. *We have*

$$\ell^\infty = \{(x_i)_{i=1}^\infty : \ x_i \in \mathbb{C}, \ \sup_{i\in\mathbb{N}} |x_i| < \infty\}.$$

The norm here is given by

$$\|(x_i)_{i=1}^\infty\|_\infty = \sup_{i\in\mathbb{N}} |x_i|.$$

One final special example is the space

$$c_0 = \{(x_i)_{i=1}^\infty : \ \textit{convergent sequences with } \lim_{i\to\infty} x_i = 0.\}.$$

The norm on c_0 *is the same as the* ℓ^∞ *norm, and we have* $c_0 \subseteq \ell^\infty$.

(vi) $L^p([a,b])$ *and* $L^\infty([a,b])$ *(the* L^p *spaces) where* $1 \le p < \infty$ *and the spaces are formed with regards to Lebesgue measure. As we've seen before, the norm here is*

$$\|f\|_p = \left(\int_a^b |f|^p dx\right)^{1/p},$$

for $f : \mathbb{R} \to \mathbb{R}$ *measurable and* $|f|$ *Riemann integrable.*

2.2. **Convergence and linear operators.** In the following, let $(X, \|\cdot\|)$ be a fixed normed space.

Definition 2.6. *A sequence* $(x_n)_{n\in\mathbb{N}} \subseteq X$ *is called Cauchy if for every* $\epsilon > 0$ *there exists an* $N \in \mathbb{N}$ *such that for all* $n, m \ge N$ *then* $\|x_n - x_m\| < \epsilon$. *Similarly, we say* $(x_n)_{n\in\mathbb{N}}$ *is convergent to* $x \in X$ *if*

$$\lim_{n\to\infty} \|x_n - x\| = 0.$$

Note: Every convergent sequence is Cauchy (as we expect).

Definition 2.7. *A Banach space is a normed space* $(X, \|\cdot\|)$ *in which every Cauchy sequence is convergent- i.e.,* X *is complete with regards to* $\|\cdot\|$.

Note: All previous examples of normed spaces we have discussed are Banach spaces.

Example 2.8. *Take $C_{[0,1]}$ with $\|\cdot\|$ given by*

$$\|f\| = \int_0^1 |f| dx.$$

This is an example of a normed space which is not complete- hence, not a Banach space. Note that while the space $C_{[0,1]}$ endowed with norm $\|\cdot\|_{\sup}$ is complete, the same underlying space when endowed with norm $\|\cdot\|_1$ is no longer complete.

As a historical aside, we note that much of the modern foundations of functional analysis began in the 1930's- specifically, it was developed by Banach and von Neumann (although by no means were those two the only to develop the modern foundations).

Definition 2.9 (New spaces from old ones)**.** *Let X, Y be normed spaces. We create the new space*

$$X \oplus_\infty Y = \{(x, y) : \ x \in X, y \in Y\}$$

and define norm $\|\cdot\|_\infty$ by

$$\|(x, y)\|_\infty = \max\{\|x\|, \|y\|\}.$$

Alternatively, we can define the space

$$X \oplus_1 Y = \{(x, y) : \ x \in X, y \in Y\}$$

and norm $\|\cdot\|_1$ given by

$$\|(x, y)\|_1 = \|x\| + \|y\|.$$

Note that the underlying vector space structure is the same in both cases.

Already, some questions begin to arise based on the construction of new spaces from others.

Question: Can a general Banach space be split into a space of the form $X \oplus_1 Y$?

Answer: In general, not often. The full answer is much more complicated.

For a bit more history, we briefly mention some important recent results. One specific problem that remained unanswered for a long time up until the 90's was known as the homogeneous space problem. Timothy Gowers gave an answer to the problem in 1998 and won the Fields medal as a result. His work also provided several counterexamples to many famous unsolved conjectures in functional analysis at the time. Boris Tsirelson introduced the concept of combinatorial norms in the 1970's, beginning the connection between functional analysis and combinatorics which would be further developed by Gowers a few decades later. Finally, Argyros and Hayden showed in 2013 that there exist Banach spaces X such that

$$\mathcal{B}(X) = \mathcal{K}(X) + \mathbb{C}I,$$

where $\mathcal{K}(X)$ denotes the space of compact operators on X.

Definition 2.10. *Let X, Y be normed spaces. For linear operators $T : X \to Y$, we say:*

(i) T is continuous at $x_0 \in X$ if $x_n \to x_0$ implies $Tx_n \to Tx_0$ as $n \to \infty$. We note that this means T is continuous at x_0 in the metric space determined by $\|\cdot\|$, i.e.

$$\|x_n - x_0\| \to 0 \Rightarrow \|Tx_n - Tx_0\| \to 0$$

as $n \to \infty$.

(ii) T is continuous if T is continuous at every $x \in X$.

(iii) T is bounded if there exists a $c > 0$ such that

$$\|Tx\| \leq c\|x\|$$

for all $x \in X$.

Remark: We cannot have some $M > 0$ such that $\|Tx\| \leq M$ for all $x \in X$ if T is a non-zero linear operator. Suppose there did exist such an M- if we take some $\lambda \in \mathbb{C}$, we see

$$\|T(\lambda x)\| = |\lambda|\|Tx\| \leq |\lambda|M.$$

This implies $\|Tx\| = 0$ for $x \neq 0$, as the inequality holds for arbitrary $\lambda \in \mathbb{C}$. Hence $T = 0$.

Notation: We use the following:

- $\mathcal{B}(X,Y)$ as the space of all bounded linear operators from X into Y.
- $\mathcal{B}(X) = \mathcal{B}(X,X)$.
- $X^* = \mathcal{B}(X,\mathbb{C})$ for the dual space of X. Recall that linear and bounded maps $f : X \to \mathbb{C}$ are called bounded functionals.
- For $x_0 \in X$, $r > 0$ we set
$$B_r(x_0) = \{x \in X : \ \|x - x_0\| \leq r\},$$
$$D_r(x_0) = \{x \in X : \ \|x - x_0\| < r\}.$$
Here $B_r(x_0)$ stands for the closed ball, and $D_r(x_0)$ the open ball.
- For $A, B \subseteq X$ and $x_0 \in X$
$$\lambda A = \{\lambda x : \ x \in A\} \subseteq X,$$
$$A + B = \{x + y : \ x \in A, y \in B\},$$
$$A + x_0 = A + \{x_0\} = \{x + x_0 : x \in A\}.$$

Recall:

(i) Open sets: $U \subseteq X$ is open if $U \neq \emptyset$ and for any $x_0 \in U$, there exists $r > 0$ such that $D_r(x_0) \subseteq U$.
(ii) Closed sets: set $F \subseteq X$ is closed if F^C is open. This is also equivalent to saying F is closed if for any $(x_n)_{n\in\mathbb{N}} \subseteq F$ such that $x_n \to x$, then $x \in F$ as well.
(iii) Map $T : X \to Y$ is continuous if $T^{-1}(U)$ is open in X for every open set $U \subseteq Y$.

Theorem 2.11. *For linear operator $T : X \to Y$ (where X, Y are normed spaces) the following are equivalent:*

(i) T is continuous at 0;
(ii) T is continuous at x_0 for some $x_0 \in X$;
(iii) T is continuous;
(iv) T is bounded.

Proof.
That $(i) \Rightarrow (ii)$ is trivial.

$(ii) \Rightarrow (i)$ Suppose we take $x_n \to 0$ in X, and fix $x_0 \in X$. As addition is a continuous operation, we also know that $x_n + x_0 \to x_0$ as $n \to \infty$. Additionally, we note that if $\lambda_n \to \lambda$ (with $\lambda_n, \lambda \in \mathbb{C}$) and $x_n \to x$, then $\lambda_n x_n \to \lambda x$ (i.e. scalar multiplication is continuous). Applying T to both sides, we find
$$T(x_n + x_0) \to Tx_0$$
$$= Tx_n + Tx_0 \to Tx_0$$
as T is linear. Therefore, $Tx_n \to 0$.

$(iii) \Rightarrow (i)$ is trivial.

$(i) \Rightarrow (iii)$ Fix any x_0, and apply the previous arguments.

$(iv) \Rightarrow (i)$ Assume T is bounded with constant $C > 0$- so $\|Tx\| \leq C\|x\|$ for all $x \in X$. If we take $x_n \to 0$ as $n \to \infty$, as $\|Tx_n\| \leq C\|x_n\|$ and $\|x_n\| \to 0$, by the Sandwich Theorem (for limits) $\|Tx_n\| \leq C\|x_n\| \to 0$. Hence, $\|Tx_n\| \to 0$, and so T is continuous at 0.

$(i) \Rightarrow (iv)$ Take $D_1(0) \subseteq Y$- this is an open set (clearly). By definition, $T^{-1}(D_1(0))$ is open as T is continuous at 0. We note that $0 \in T^{-1}(D_1(0))$, as $T(0) = 0$. Then as $T^{-1}(D_1(0))$ is open in X, there exists some $r > 0$ such that $D_r(0) \subseteq T^{-1}(D_1(0))$.

This means for all $x \in D_r(0)$, $Tx \in D_1(0)$. Hence if $\|x\| < r$, then $\|Tx\| \leq 1$. If we take $\|x\| \leq 1$, then $\|x/2\| \leq 1$ as well. This implies for any $r > 0$, $\|rx/2\| < r$, and so $\|T(\frac{rx}{2})\| < 1$. This implies $\|Tx\| \leq 2/r$. If we now take $C = 2/r$, we've shown that $\|x\| \leq 1$ implies $\|Tx\| \leq C$. Taking any $x \neq 0$, clearly $\frac{x}{\|x\|}$ has norm 1- so $\|T(\frac{x}{\|x\|})\| \leq C$. Then

$$\|T(\frac{x}{\|x\|})\| \leq C \Rightarrow \frac{1}{\|x\|}\|Tx\| \leq C$$
$$\Rightarrow \|Tx\| \leq C\|x\|.$$

This shows T is bounded for all $x \in X$. □

Definition 2.12. *If $T \in \mathcal{B}(X,Y)$ we let*

$$\|T\| = \inf\{C > 0 : \|Tx\| \leq C\|x\|, x \in X\}.$$

This is known as the norm of operator T.

Note: We have the equivalent formulation

$$\|T\| = \sup\{\|Tx\| : \|x\| \leq 1, x \in X\}$$

by the proof of the previous theorem. In fact,

$$\|T\| = \sup\{\|Tx\| : \|x\| = 1, x \in X\}.$$

To see this, if we suppose $\|x\| < 1$, then there exists $r > 1$ such that $\|rx\| = r\|x\| < 1$. Letting $r \to 1$, we have $\|rx\| \to \|x\|$. Since our values of r can allow us to "jump up" above whatever our supposed supremum for the norm of T when only considering $\|x\| < 1$, this implies we cannot achieve a supremum of $\|Tx\|$ when $\|x\| < 1$.

Theorem 2.13. *Let X, Y be normed spaces.*

(i) $\mathcal{B}(X,Y)$, equipped with the operator norm $\|\cdot\|_{op}$, is a normed space.
(ii) If Y is a Banach space, then $(\mathcal{B}(X,Y), \|\cdot\|)$ is a Banach space.

Proof. To show (i), we'll focus on the Triangle Inequality. Let $T, S \in \mathcal{B}(X,Y)$. By definition

$$\|S + T\| = \sup\{\|Sx + Tx\| : \|x\| \leq 1\}.$$

We have

$$\sup\{\|Sx + Tx\| : \|x\| \leq 1\} \leq \sup_{\|x\| \leq 1}\{\|Sx\| + \|Tx\|\} \leq \sup_{\|x\| \leq 1}\|Sx\| + \sup_{\|y\| \leq 1}\|Ty\|$$
$$= \|S\| + \|T\|.$$

The other properties are easy to show- this takes care of (i).

To show (ii), assume Y is complete. Suppose $(T_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{B}(X,Y)$- i.e. $\|T_n - T_m\| \to 0$ as $n, m \to \infty$. Fix $x \in X$, and consider $(T_nx)_{n\in\mathbb{N}}$- we claim this is a Cauchy sequence in Y. As x is fixed, we have

$$\|T_nx - T_mx\| = \|(T_n - T_m)x\| \leq \|T_n - T_m\|\|x\| \to 0,$$

as $n, m \to \infty$. Hence, $(T_nx)_{n\in\mathbb{N}}$ is indeed Cauchy. As Y is complete, $(T_nx)_{n\in\mathbb{N}}$ converges in Y- i.e. there exists some $y \in Y$ such that $T_nx \to y$ as $n \to \infty$. Set $y = Tx$. We now have a map which sends

$$x \mapsto Tx,$$

(so a map between X and Y).

To check additivity of this map, for $x_1, x_2 \in X$ we apply

$$T(x_1 + x_2) = \lim_{n\to\infty} T_n(x_1 + x_2) = \lim_{n\to\infty}(T_nx_1 + T_nx_2) = Tx_1 + Tx_2.$$

Similarly, scalar multiplication is preserved (as we wish). This shows T is a linear map.

To show T is bounded, we note

$$\begin{aligned}\|Tx\| = \|Tx - T_m x + T_m x\| &\leq \|Tx - T_m x\| + \|T_m x\| \\ &= \lim_{n\to\infty} \|T_n x - T_m x\| + \|T_m x\|.\end{aligned}$$

Choose $m \in \mathbb{N}$ such that if $n \geq m$, $\|T_n - T_m\| < 1$. Then

$$\|T_n x - T_m x\| = \|(T_n - T_m)x\| \leq \|T_n - T_m\|\|x\| \leq \|x\|.$$

For this choice of m,

$$\|T_x\| \leq \|x\| + \|T_m x\| \leq \|x\| + \|T_m\|\|x\| = (1 + \|T_m\|)\|x\|$$

for all $x \in X$. This shows T is bounded.

Finally, we show $T_n \to T$ in norm. For $x \in X$,

$$(T - T_n)x = \lim_{m\to\infty} (T_m - T_n)x.$$

Hence,

$$\lim_{n\to\infty} \|(T - T_n)x\| = \lim_{n,m\to\infty} \|T_m x - T_n x\| = 0$$

when looking at Cauchy sequence $(T_n)_{n\in\mathbb{N}}$. Hence, $\|T_n - T\| \to 0$ as $n \to \infty$. This completes the proof. □

Corollary 2.14. *For any normed space X, X^* is Banach.*

Definition 2.15 (The dual operator)**.** *For normed spaces X, Y and $T \in \mathcal{B}(X, Y)$ define $T^* : Y^* \to X^*$ by*

$$T^*(g)(x) = g(Tx),$$

for $g \in Y^*$ and for all $x \in X$. We say T^* is the dual operator of T.

Note:

(i) T^* is linear, as

$$T^*(g_1 + g_2)(x) = (g_1 + g_2)(Tx) = g_1(Tx) + g_2(Tx) = T^*(g_1)(x) + T^*(g_2)(x).$$

(ii) T^* is bounded: let $\|g\| \leq 1$ for $g \in Y^*$. We see

$$\begin{aligned}\|T^*(g)\| &= \sup\{|T^*(g)(x)| : \ \|x\| \leq 1\} \\ &= \sup\{|g(Tx)| : \ \|x\| \leq 1\} \\ &\leq \sup\{\|Tx\| : \ \|x\| \leq 1\} = \|T\|.\end{aligned}$$

This implies $\|T^*\| \leq \|T\|$. (In fact, $\|T^*\| = \|T\|$, which we will show later).

(iii) If we have $S : X \to Y$ and $T : Y \to Z$, then $T \circ S : X \to Z$ is (obviously) well-defined. We also have
   (i) $\|TS\| \leq \|T\|\|S\|$ (directly from the definition),
   (ii) $(T \circ S)^* = S^* \circ T^*$.

Proposition 2.16. *Every vector space has a Hamel basis.*

Proof. We use Zorn's Lemma. Let

$$\mathcal{B} = \{(v_i)_{i\in I} : \ \text{linearly independent set }\}.$$

We can order $\mathcal{B}$ by inclusion of sets- i.e., for $A, B \in \mathcal{B}$ we say $A \leq B$ if $A \subseteq B$. By Zorn's Lemma, we have a maximal element $H \in \mathcal{B}$. This element H is our Hamel basis for the space. □

Projections:

Definition 2.17. *For a normed space X, we say $P : X \to X$ is a projection if P is a linear operator and $P^2 = P$.*

Note: For operator $T : X \to X$, $T^k = \underbrace{T \cdot T \cdot \dots \cdot T}_{k \text{ times}}$.

Let $Y = \text{rng}P = \{Px : x \in X\}$, where P is a projection. If $y \in Y$, then $Py = y$. To see this, we note

$$Py = P(Px) = P^2x = Px = y,$$

for some $x \in X$.

Example 2.18.

(i) *Take $Y_n \subseteq \ell^p$ where*

$$Y_n = \{(\lambda_i)_{i=1}^{\infty} : \lambda_i = 0 \text{ if } i > n\},$$

and define $P_n : \ell^p \to \ell^p$ by

$$P((\lambda_1, \dots, \lambda_n, \lambda_{n+1}, \dots)) = (\lambda_1, \dots, \lambda_n, 0, 0, \dots).$$

We claim P_n is bounded; to see this, take any $x = (\lambda_i)_{i \in \mathbb{N}} \in \ell^p$. We have

$$\|P_n x\|_p = \|(\lambda_1, \dots, \lambda_n, 0, \dots)\|_p = \left(\sum_{i=1}^{n} |\lambda_i|^p \right)^{1/p}$$
$$\leq \left(\sum_{i=1}^{\infty} |\lambda_i|^p \right)^{1/p} = \|x\|_p.$$

Hence, $\|P_n\| \leq 1$. We also have $\|P_n\| \geq 1$- taking $e_1 = (1, 0, 0, \dots)$ we have $\|e_1\|_p = 1$ (clearly) and $P_n e_1 = e_1$ for all $n \in \mathbb{N}$. Hence $\|P_n e_1\|_p = \|e_1\|_p$, so $\|P_n\|_p \geq 1$. This shows $\|P_n\| = 1$.

It is also clear that $P_n^2 = P_n$, and so P_n is a projection with $\text{rng}P_n = Y_n$. This same projection also holds on c_0.

Note: *If $(x_n)_{n \in \mathbb{N}}$ is a "basis", one can define projections as above where*

$$P_n : \sum_{i=1}^{\infty} \lambda_i x_i \mapsto \sum_{i=1}^{n} \lambda_i x_i.$$

(ii) *We'll look at the shift operator. Let $S : \ell^p \to \ell^p$ be given by*

$$S((\lambda_1, \lambda_2, \dots)) = (0, \lambda_1, \lambda_2, \dots).$$

Clearly, $\|Sx\| = \|x\|$ for all $x \in \ell^p$. We first note that S is not surjective: as an example, $e_1 \notin \text{rng}S$. We have

$$\text{rng}S = \{(\lambda_i)_{i \in \mathbb{N}} \in \ell^p : \lambda_1 = 0\}.$$

The backward shift $B : \ell^p \to \ell^p$, on the other hand, is given by

$$B((\lambda_1, \lambda_2, \lambda_3, \dots)) = (\lambda_2, \lambda_3, \dots).$$

We have $\|B\| = 1$, and B is surjective, but not injective. Finally, we note we may write

$$\ell^p = \text{rng}S \oplus \text{rng}P_1.$$

(iii) *We'll look at operators on the space* $C[0,1]$*; recall that*

$$C[0,1] = \{f : [0,1] \to \mathbb{C} \text{ continuous}\},$$

and $C[0,1]$ *is endowed with the norm*

$$\|f\|_\infty = \max_{0\le t\le 1} |f(t)|.$$

Fix $t \in [0,1]$*, and let* $F_t : C[0,1] \to \mathbb{C}$ *be given by*

$$F_t(f) = f(t).$$

This is a bounded functional, as

$$\|F_t(f)\| = |F_t(f)| = |f(t)| \le \|f\|_\infty.$$

This implies $\|F_t\| \le 1$*. In fact, we may say* $\|F_t\| = 1$*, as we may evaluate at any* $f : [0,1] \to \mathbb{C}$ *continuous where* $f(t) = 1$ *and* $f(t') < 1$ *for all other* $t' \in [0,1]$*.*

The functions F_t *are called evaluation functionals. We note that we cannot define evaluation functionals on* $L^\infty(0,1)$*, as we deal here with equivalence classes of functions- in this space, the expression* $f(t)$ *makes no sense. If we take* $\sum_{i=1}^{k} \lambda_i F_{t_i}$ *where* λ_i *are scalars, this also defines a linear functional on* $C[0,1]$*.*

Fact: *Every bounded linear functional on* $C[0,1]$ *is a limit of functionals of the form* $\sum_{i=1}^{k} \lambda_i F_{t_i}$*. In fact, every bounded linear functional* $F : C[0,1] \to \mathbb{C}$ *has the form*

$$F(t) = \int f d\mu,$$

for some measure μ*. Here* μ *is the complex Borel measure on* $[0,1]$*.*

For all $t \in [0,1]$*, we can define function* δ_t *where*

$$\delta_t = \begin{cases} 1, & t \in E, \\ 0, & t \notin E. \end{cases}$$

Then δ_t *gives rise to* F_t *by the above.*

(iv) *For* $1 \le p < \infty$*, what is* $(\ell^p)^*$*? Take* q *with* $1 < q \le \infty$ *such that* q *is conjugate to* p*. We can show (**in the homework!**) that*

$$\ell^q \cong (\ell^p)^*,$$

i.e. the two spaces are isometrically isomorphic.

(v) *For* $(c_0)^*$*, we follow the same idea as in (iv); we can show that the dual space of* c_0 *is* ℓ^1*. We note that* $(\ell^\infty)^* \neq \ell^1$*, but must use* c_0 *instead. This is because* ℓ^∞ *is not separable, while* ℓ^1 *is.*

The quotient normed space:

Let X be a normed space, and let $Y \subseteq X$ be a closed subspace. Define a binary relation $\sim$ on X, where for $x, x' \in X$ we say $x \sim x'$ if $x - x' \in Y$. This is an equivalence relation on X (which is easy to check), and we write

$$[x] = \{x' \in X : \ x' \sim x\}.$$

If we quotient by the equivalence relation $\sim$ on X, we end up with a set X/Y whose elements are the equivalence classes $[x]$, for $x \in X$. Addition and scalar multiplication are defined in the obvious way (and are well defined), which turns X/Y into a complex vector space. Furthermore, for $x \in X$ define

$$\|[x]\| = \inf\{\|x + y\| : \ y \in Y\} = \inf_{z\sim x} \|z\| = \inf_{z\in[x]} \|z\|.$$

We claim X/Y equipped with $\|[\cdot]\|$ is a normed vector space.

To show this, we need to check the three conditions of a norm. We begin by taking any two $x, y \in X$. We see

$$\begin{aligned}\|[x]+[y]\| &= \|[x+y]\| = \inf\{\|x+y+z\| : z \in Y\} = \inf\{\|x+y+z_1+z_2\| : z_1, z_2 \in Y\} \\ &\leq \inf\{\|x+z_1\|+\|y+z_2\| : z_1, z_2 \in Y\} = \inf_{z_1 \in Y}\|x+z_1\| + \inf_{z_2 \in Y}\|y+z_2\| = \|[x]\|+\|[y]\|.\end{aligned}$$

This shows $\|[\cdot]\|$ satisfies the triangle inequality. For scalar multiplication, it is quite easy to check (it follows almost immediately by definition). Therefore, we'll finish by checking definiteness of $\|[\cdot]\|$. If we suppose $\|[x]\| = 0$. By definition,

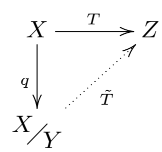
$$0 = \|[x]\| = \inf\{\|x+z\| : z \in Y\}.$$

This tells us we can find a sequence of elements $(z_n)_{n\in\mathbb{N}} \subseteq Y$ such that $\|x+z_n\| \to 0$, which implies $x + z_n \to 0$. As Y is closed with $z_n \to -x$, this implies $-x \in Y$, and hence $x \in Y$. Thus, $[x] = [0] = Y$. This shows $\|[\cdot]\|$ is a norm on X/Y.

Definition 2.19. *For a normed space X and a closed subspace $Y \subseteq X$, we define the quotient normed space X/Y as the vector space formed on X when quotient-ing by relation $\sim$ and endowed with the norm $\|[\cdot]\|$ defined above.*

The natural quotient map $q : X \to X/Y$ is given by $q(x) = [x]$, for $x \in X$. We also have that q is a linear bounded operator with $\|q\| \leq 1$.

Theorem 2.20. *Let X be a normed space, and $Y \subseteq X$ a closed subspace. Denote $q : X \to X/Y$ as the standard quotient map. For every normed space Z, and every bounded operator $T : X \to Z$ with $Y \subseteq \ker T$ there exists a unique operator $\tilde{T} : X/Y \to Z$ such that $T = \tilde{T} \circ q$, i.e. the following diagram commutes:*



Moreover, $\|\tilde{T}\| = \|T\|$.

Proof. (Sketch) Define operator $\tilde{T}$ by letting $\tilde{T}([x]) = Tx$. This map is well-defined, as $Y \subseteq \ker T$. We also have $\tilde{T}(q(x)) = T(x)$ for all $x \in X$, which implies $\tilde{T} \circ q = T$. To show $\|\tilde{T}\| = \|T\|$, **exercise!** □

Note: If $T = S \circ q$ for some operator S, then $Y \subseteq \ker T$. Why is this so? If $y \in Y$, then $q(y) = 0$. Since S is linear, $S(0) = 0$. Hence, $Ty = S(q(y)) = S(0) = 0$. This implies inclusion.

To actually use quotient spaces, we'll look at the distance of a point from a subspace.

Definition 2.21. *If $Y \subseteq X$ is a closed subspace, for $x \in X$ then*

$$\mathit{dist}(x, Y) = \|q(x)\|_{X/Y}.$$

Note: In a Hilbert space, this is always achieved.

Example 2.22. *Take ℓ^p, and $Y = \{(\lambda_1, \lambda_2, \ldots, \lambda_n, 0, 0, \ldots) : \lambda_i \in \mathbb{C}\}$. The subspace $Y \subseteq \ell^p$ is closed, and if we take $x = (x_n)_{n\in\mathbb{N}}$ then the norm $\|q(x)\|$ in ℓ^p/Y is*

$$\|q(x)\| = \left(\sum_{i>n} |x_i|^p\right)^{1/p}.$$

The closest point to x in Y is clearly $(x_1, x_2, \ldots, x_n, 0, 0, \ldots)$.

Example 2.23. *In* $C[-1,1]$*, consider* $\mathcal{P}_n = \{p : \text{ polynomials with degree at most } n-1\}$.

Question: *what is the distance of* x^n *to* $\mathcal{P}_n$*?*

We take

$$\inf_{a_0,\cdots,a_{n-1}} \|x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0\|_\infty$$

to try and find the answer.

Answer: *Chebyshev polynomials* $T_n(x) = \frac{1}{2^{n-1}} \cos(n \arccos(x))$ *are the closet points, and the distance is* $\frac{1}{2^{n-1}}$.

2.3. **Separable spaces.**

Definition 2.24. *A normed space* X *is called separable if there exists a countable set* S *dense in* X.

Remark: A space X is separable if and only if there exists a countable set S whose linear combinations are dense in X.

$$\text{dense} \iff \text{linear span is dense}$$

Example 2.25.

(i) *The space* ℓ^p *is separable for* $1 \leq p < \infty$*. Take the standard basis* $\{e_i\}_{i\in\mathbb{N}}$*- this is a countable set (clearly). We also have*

$$\text{span}\{e_1, e_2, \cdots\} = c_{00} := \{(x_n)_{n\in\mathbb{N}} : \text{ there exists } N > 0 \text{ so that } x_n = 0 \text{ if } n > N\}.$$

(ii) c_0 *is separable.*
(iii) $C[0,1]$ *is separable.*
(iv) ℓ^∞ *is not separable. Take the power set of* $\mathbb{N}$*, and* χ_S *as the indicator function for a set* $S \in \mathcal{P}(\mathbb{N})$*. Then* $\{\chi_S : S \subseteq \mathbb{N}\}$ *is an uncountable family lying in* ℓ^∞*. As the distance between any two elements in this family is* 1*, this implies inseparability of* ℓ^∞.

2.4. **Hahn-Banach theorems and applications.**

Definition 2.26. *If* V *is a vector space over* $\mathbb{R}$*, a subset* $C \subseteq V$ *is called convex if for* $x, y \in C$ *and* $t, s \in [0,1]$ *such that* $t + s = 1$*, then* $sx + ty \in C$.

Definition 2.27. *A hyperplane in a vector space* V *over* $\mathbb{R}$ *is a set of the following form:* $H = x_0 + Z$*, where* Z *is a subspace of* V *of co-dimension* 1.

Notes:

(i) If $Z \subseteq V$ is a subspace, we define the co-dimension of Z as $\text{codim}(Z) = \dim {}^{V}\!/_{Z}$.
(ii) Suppose that $f : V \to \mathbb{R}$ is a linear functional with $f \neq 0$. Then $\ker f$ is a subspace of co-dimension 1 in V.

Proof. (Loosely speaking, if a space has co-dimension 1 we only have one possible way to “move”) As $f \neq 0$, $H = \ker f$ is distinct from V. Let $x_0 \in V \setminus H$, and let $x \in V$. Then

$$f\left(x - \frac{f(x)}{f(x_0)}x_0\right) = f(x) - \frac{f(x)}{f(x_0)}f(x_0) = 0.$$

Hence, $x - \frac{f(x)}{f(x_0)}x_0 \in H$, and so $x = \frac{f(x)}{f(x_0)}x_0 + z$, where $z \in H$. This also implies that for any $x \in V$, there exists $\lambda \in \mathbb{R}$ such that $x = \lambda x_0 + z$ for some $z \in H$. Hence, $\ker f$ has co-dimension 1 in V. $\square$

(iii) Conversely, if $Z \subseteq V$ is a subspace with co-dimension 1, then there exists a linear functional $f : V \to \mathbb{R}$ such that $Z = \ker f$.

Proof. Let $q : V \to {}^{V}\!/\!_{Z}$ be the canonical quotient map. Then ${}^{V}\!/\!_{Z} \cong \mathbb{R}$ (as both are 1-dimensional vector spaces over $\mathbb{R}$). Let $\phi : {}^{V}\!/\!_{Z} \to \mathbb{R}$ be an isomorphism. Then $f : V \to \mathbb{R}$ where $f = \phi \circ q$ is a linear map with kernel Z. □

Question: If X is a normed space over $\mathbb{R}$, which hyperplanes are closed (in the norm topology)?

Answer: Precisely the ones of the form $x_0 + Z$, where $Z = \ker f$ for a continuous (equivalently, bounded) linear functional f.

We'll prove the statement above as follows:

Proof. If f is bounded, then $\ker f$ is closed. Suppose $(x_n)_{n\in\mathbb{N}} \subseteq \ker f$, and $x_n \to x$ as $n \to \infty$. Then $f(x_n) = 0$ for all $n \in \mathbb{N}$. As boundedness of f is equivalent to continuity, we have $f(x) = \lim_{n\to\infty} f(x_n) = 0$. Hence, $x \in \ker f$.

Conversely, if $Z \subseteq X$ is a closed subspace of co-dimension 1, then the quotient normed space ${}^{X}\!/\!_{Z}$ is well-defined, with ${}^{X}\!/\!_{Z} \cong \mathbb{R}$. For an isomorphism $\phi : {}^{X}\!/\!_{Z} \to \mathbb{R}$, as the quotient map $q : X \to {}^{X}\!/\!_{Z}$ is continuous, $f = \phi \circ q$ is a continuous linear functional with $\ker f = Z$. □

We'll now move to discussing the Hahn-Banach Theorem, and some of its applications. The Hahn-Banach theorem "splits" into two potential forms of interpretation, which are equivalent: an interpretation as an extension theorem, and interpretation as a separation theorem (which is its geometric form).

Definition 2.28. *A semi-norm on a normed space over $\mathbb{R}$ is a function $p : X \to \mathbb{R}_+$ such that*

(i) $p(x + y) \leq p(x) + p(y)$;
(ii) $p(\lambda x) = \lambda p(x)$, for all $\lambda \in \mathbb{R}_+$.

Lemma 2.29. *Let X be a real vector space, $Y \subseteq X$ a subspace with co-dimension 1, $p : X \to \mathbb{R}_+$ a semi-norm, and $f : Y \to \mathbb{R}$ a linear functional such that $f(y) \leq p(y)$ for all $y \in Y$. Then there exists $g : X \to \mathbb{R}$ where $g|_Y = f$, and $g(x) \leq p(x)$ for all $x \in X$.*

Proof. Since Y has co-dimension 1, there exists a vector $z \in X \setminus Y$- so for all $x \in X$, there exists $\lambda \in \mathbb{R}$ and $y \in Y$ such that $x = y + \lambda z$. We wish to determine $g(z)$; assuming (for a short while) that g already exists, let $c = g(z)$. We need to determine what conditions hold on c. We know that as g must be linear, $g(y + tz) = g(y) + g(tz) = f(y) + tc$. We want $g(y + tz) \leq p(y + tz)$- this means $f(y) + tc \leq p(y + tz)$. Consider the following two cases:

(i) If $t > 0$: then
$$c \leq \frac{1}{t}[p(y + tz) - f(y)] = p(y/t + z) - f(y/t).$$

(ii) If $t < 0$: then $t = -s$, for some $s > 0$. We have
$$f(y) - sc \leq p(y - sz) \Rightarrow \frac{1}{s}[f(y) - p(y - sz)] \leq c$$
$$\Rightarrow f(y/s) - p(y/s - z) \leq c.$$

These two cases imply that the constraints on c are the following:
$$f(y') - p(y' - z) \leq c \leq p(y + z) - f(y),$$

for all $y, y' \in Y$. So to be able to find c, it is sufficient and necessary that the following inequalities are fulfilled:

$$f(y') - p(y' - z) \leq p(y + z) - f(y), \qquad \text{for all } y, y' \in Y.$$

Equivalently,

$$f(y) + f(y') = f(y + y') \leq p(y + z) + p(y' - z),$$

for $y, y' \in Y$. But this is entirely possible, for

$$f(y + y') \leq p(y + y') = p(y + y' + z - z) \leq p(y + z) + p(y' - z),$$

as p is a semi-norm. Thus, such a c is possible- choosing any c which satisfies this inequality, we uniquely determine our g which completes the proof. □

Theorem 2.30. *Let X be a real vector space, $Y \subseteq X$ a subspace, $p : X \to \mathbb{R}_+$ a semi-norm, and $f : Y \to \mathbb{R}$ a linear functional such that $f(y) \leq p(y)$ for all $y \in Y$. Then there exists $g : X \to \mathbb{R}$ such that $g|_Y = f$, and $g(x) \leq p(x)$ for all $x \in X$.*

Proof. Let $\mathcal{E} = \{(Z, h) : \ Z \text{ subspace with } Y \subseteq Z, \ h : Z \to \mathbb{R}, h|_Y = f, \text{ and } h(z) \leq p(z) \ z \in Z\}$. we note that $\mathcal{E} \neq \emptyset$, as $(Y, f) \in \mathcal{E}$. Equip $\mathcal{E}$ with the partial order given by:

$$(Z_1, h_1) \leq (Z_2, h_2) \text{ if } Z_1 \subseteq Z_2 \text{ and } h_2|_{Z_1} = h_1.$$

Suppose $((Z_\alpha, h_\alpha))_{\alpha \in A}$ is a chain of elements in $\mathcal{E}$. Let $Z = \cup_{\alpha \in A} Z_\alpha$; this is a subspace as a total ordering of the chain. Let $h : Z \to \mathbb{R}$ be defined by $h(z) = h_\alpha(z)$ where $\alpha \in A$ is any index with $z \in Z_\alpha$. Then (Z, h) is an upper bound of our chain. By Zorn's Lemma, there then exists a maximal element in $\mathcal{E}$- say (W, g). We claim $W = X$- if not, i.e. if $x_0 X \setminus W$ then consider the space $W + \mathbb{R}x_0$. Clearly $W \subseteq W + \mathbb{R}x_0$ has co-dimension 1, and so by the previous lemma we have a further extension of g. But as (W, g) is a maximal element, this contradicts our exact choice. Hence, $W = X$, which completes the proof. □

Theorem 2.31 (Hahn-Banach)**.** *Let X be a normed real space, $Y \subseteq X$ a closed subspace, and $f : Y \to \mathbb{R}$ a bounded linear functional. Then there exists a linear functional $g : X \to \mathbb{R}$ which is bounded, $g|_Y = f$ and $\|g\| = \|f\|$.*

Proof. Define a semi-norm $p : X \to \mathbb{R}_+$ by $p(x) = \|f\|\|x\|$, for $x \in X$. It is easy to check p is indeed a semi-norm. We also see that $f(x) \leq \|f\|\|x\|$ by properties of the norm on X. Then by the previous theorem, there exists a $g : X \to \mathbb{R}$, $g|_Y = f$ and $g(x) \leq \|f\|\|x\|$ for all $x \in X$. Then $|g(x)| \leq \|f\|\|x\|$ as well- this holds, as we can take either x or $-x$ while still bound above by $\|x\|$. As $\|g\|$ cannot decrease, and $\|g\| \leq \|f\|$ we necessarily have $\|g\| = \|f\|$. □

Theorem 2.32 (Complex Hahn-Banach)**.** *Take the same conditions and statements as above, but for X a complex normed space.*

Proof. (Sketch) We'll use the following "trick": every vector space V over $\mathbb{C}$ is also a vector space over $\mathbb{R}$. Denote this space as $V_\mathbb{R}$. If $f : V \to \mathbb{C}$ is a linear functional, then $\Re f : V \to \mathbb{R}$ where $\Re f(v) = \Re(f(v))$ is a linear vector functional on $V_\mathbb{R}$. Conversely, a real functional $g : V_\mathbb{R} \to \mathbb{R}$ gives rise to the complex functional $f : V \to \mathbb{C}$ given by $f(v) = g(v) - ig(iv)$. All that remains to finish the proof is showing that the norm is preserved. □

Theorem 2.33. *Let X be a normed space, and $x_0 \in X$. Then $\|x_0\| = \sup\{|f(x_0)| : \ f \in X^*, \|f\| \leq 1\}$.*

Note: The right hand side above is at most $\|x_0\|$, as

$$|f(x_0)| \leq \|f\|\|x_0\| \leq \|x_0\|,$$

if $\|f\| \leq 1$. Here, the Hahn-Banach theorem will allow us to translate problems in the norm down to the level of problems with the dual space.

Proof. Assume $x_0 \neq 0$. Consider the space $\mathbb{C}x_0$, and define a linear functional $f_0 : \mathbb{C}x_0 \to \mathbb{C}$ by $f_0(\lambda x_0) = \lambda \cdot \|x_0\|$. Note that $\|f_0\| = 1$. Extending f_0 to X by the Hahn-Banach theorem gives $f : X \to \mathbb{C}$ with $\|f\| \leq 1$, and $f(x_0) = \|x_0\|$. □

Theorem 2.34. *Let $T \in \mathcal{B}(X, Y)$ for normed spaces X and Y. Then $\|T^*\| = \|T\|$.*

Note: We have already observed that $\|T^*\| \leq \|T\|$.

Proof. Let $\epsilon > 0$, and $x_0 \in X$ with $\|x_0\| = 1$ such that $\|Tx_0\| > \|T\| - \epsilon$. Let $g : Y \to \mathbb{C}$ be a linear functional such that $\|g\| = 1$ and $g(Tx_0) = \|Tx_0\|$ (this is possible through the application of the previous theorem). Then

$$\|T^*\| \geq \|T^*(g)\| \geq |T^*(g)(x_0)| = |g(Tx_0)| = \|Tx_0\| > \|T\| - \epsilon.$$

As $\epsilon > 0$ was arbitrary, letting $\epsilon \to 0$ we have $\|T^*\| \geq \|T\|$, and hence $\|T^*\| = \|T\|$. □

Recall: Take map $\iota : X \to X^{**}$ where $\iota(x)(f) = f(x)$ for $f \in X^*$ and $x \in X$. If $f : X \to \mathbb{C}$, fixing $x \in X$ we vary across f with the mapping $f \mapsto f(x)$. Then $x \in X$ can be viewed as a functional on X^*, i.e. $x \in X^{**}$. We claim $\iota : X \to X^{**}$ is isometric and linear. It is not always surjective.

Example 2.35. *Take $X = c_0$. Then $X^{**} = \ell^\infty$, with $c_0 \subset \ell^\infty$.*

Geometric form of the Hann-Banach Theorem
Let A, B be convex sets in a normed space X, with $A \cap B = \emptyset$.

(i) If A is open, then there exists a linear functional $f : X \to \mathbb{C}$ bounded such that $\Re f(a) < \Re f(b)$ for every $a \in A$, and $b \in B$.
(ii) If A is compact and B is closed, there exists $f : X \to \mathbb{C}$ which is bounded and linear, and γ_1, γ_2 such that $\Re f(a) \leq \gamma_1 < \gamma_2 \leq \Re f(b)$ for all $a \in A, b \in B$.

Proof. (Sketch/idea) For part (i), take $A - B = \{a - b : \ a \in A, b \in B\}$ and consider space $W = A - B + (y_0 - x_0)$ for elements $x_0 \in A, y_0 \in B$. This is an open set which now contains 0. As W is open, there exists an $r > 0$ such that $D_r(0) \subseteq W$. Define $p : X \to \mathbb{R}_+$ where $p(z) = \inf\{t > 0 : \ (1/t)z \in W\}$. For $f : \mathbb{C}(y_0 - x_0) \to \mathbb{R}$ where $f(\lambda(y_0 - x_0)) \approx \lambda$, if we check to make sure f is dominated by p it extends to a functional which is exactly the one given in the statement above. □

3. Fundamental results in functional analysis

3.1. Baire Category Theorem.

This section is devoted to taking a look at three major theorems in the study of functional analysis:

- The Open Mapping Theorem
- The Uniform Boundedness Principle/Banach-Steinhaus Theorem
- The Closed Graph Theorem

Definition 3.1. *A subset F of a metric space X is called nowhere dense if $\overline{F}$ has empty interior (i.e., $\overline{F}^C$ is dense).*

Definition 3.2. *A set $A \subseteq X$ (where X is a metric space) is called of first category if $A = \bigcup_{n=1}^{\infty} F_n$, for nowhere dense subsets F_n, $n \in \mathbb{N}$. These are also called meagre sets. A set $A \subseteq X$ is called of second category if it is not of first category.*

Theorem 3.3 (Baire Category Theorem)**.** *Let X be a complete metric space and $(U_n)_{n=1}^{\infty}$ is a sequence of dense open sets. Then $\bigcap\limits_{n=1}^{\infty} U_n$ is a dense set. Equivalently, a complete metric space is of second category.*

Proof. Let $x_0 \in X, r > 0$ and consider the open ball $D_r(x_0)$ at x_0. We wish to show that $\Omega \cap D_r(x_0)$ is non-empty, where $\Omega = \bigcap\limits_{n=1}^{\infty} \Omega_n$. Since Ω_1 is dense and open in X, there exists a point $x_1 \in \Omega_1 \cap D_r(x_0)$, and $r_1 > 0$ such that $B_{r_1}(x_1) \subseteq \Omega_1 \cap D_r(x_0)$. Now, $D_{r_1}(x_1)$ intersects Ω_2, as Ω_2 is dense in X. Hence, there exists $x_2 \in \Omega_2 \cap D_{r_1}(x_1)$ and $r_2 > 0$ such that $B_{r_2}(x_2) \subseteq \Omega_2 \cap D_{r_1}(x_1)$. Continuing inductively, we obtain a sequence $\{x_i\}_{i=1}^{\infty} \subseteq X$ which are contained in balls of progressively smaller radii. By completeness of X, we then know (as $B_{r_{n+1}}(x_{n+1}) \subseteq B_{r_n}(x_n)$) there exists a $y \in X$ such that $y = \lim\limits_{n\to\infty} x_n$. Therefore, $y \in \Omega$. As $B_n \subseteq D_r(x_0)$ for all $n \in \mathbb{N}$, this also shows $y \in D_r(x_0)$. Hence, Ω is dense in X. □

Theorem 3.4. *A complete metric space is of second category.*

Proof. Assume towards contradiction that X is of first category. So $X = \bigcup\limits_{n=1}^{\infty} F_n$, where F_n is nowhere dense for all $n \in \mathbb{N}$. We note that $X = \bigcup\limits_{n=1}^{\infty} \overline{F_n}$. Now, $\overline{F}_n^C$ is dense and open- so by the previous theorem, $\bigcap\limits_{n=1}^{\infty} \overline{F}_n^C$ is dense in X. However,

$$\bigcap_{n=1}^{\infty} \overline{F}_n^C = \left(\bigcup_{n=1}^{\infty} \overline{F}_n\right)^C = X^C = \emptyset.$$

This is clearly not dense in X, and hence we have reached a contradiction. Therefore, we conclude that X must be of second category. □

Lemma 3.5. *Let X be a complete metric space, and $\mathcal{F}$ a family of continuous functions which are real-valued. Assume that for some set $U \subseteq X$ of second category,*

$$\sup\{|f(u)| : f \in \mathcal{F}\} < \infty, \quad u \in U.$$

Then there exists a ball $B_r(x_0)$ with $r > 0$ such that

$$\sup\{|f(u)| : f \in \mathcal{F},\ u \in B_r(x_0)\} < \infty.$$

Proof. Let $F_n = \{u \in U : |f(u)| \leq n, \text{ for all } f \in \mathcal{F}\}$. This is a closed set, as each f is a continuous function. By the assumption, we have $U = \bigcup\limits_{n=1}^{\infty} F_n$. As U is of second category, there exists some $n \in \mathbb{N}$ such that F_n has non-empty interior. Therefore, there exists $r > 0$ and $x_0 \in U$ such that $B_r(x_0) \subseteq F_n$. Then

$$\sup\{|f(u)| : f \in \mathcal{F},\ u \in B_r(x_0)\} \leq n,$$

which completes the proof. □

Theorem 3.6 (Banach-Steinhas/Uniform Boundedness Principle)**.** *Let X, Y be normed spaces, suppose $U \subseteq X$ is of second category, $\mathcal{F} \subseteq \mathcal{B}(X, Y)$ is a family of operators, and*

$$\sup\{\|Tu\| : T \in \mathcal{F}\} < \infty,$$

for all $u \in U$. Then

$$\sup\{\|T\| : T \in \mathcal{F}\} < \infty.$$

Proof. Let $f_T : X \to \mathbb{R}$, where $T \in \mathcal{F}$ and the map is given by $f_T(u) = \|Tu\|$ for all $u \in X$. By the previous lemma, for family of functions $\{f_T : \ T \in \mathcal{F}\}$ there exists a positive $r > 0$ and $x_0 \in X$ such that

$$\sup\{\|Tu\| : \ T \in \mathcal{F},\ u \in B_r(x_0)\} < \infty.$$

Say $\|Tu\| \leq C$ for all $T \in \mathcal{F}$, and $u \in B_r(x_0)$. We see

$$Tx = \frac{1}{2r} T(x_0 + rx - (x_0 - rx)),$$

where $x_0 - rx, x_0 + rx \in B_r(x_0)$ for $\|x\| \leq 1$. This implies

$$\|Tx\| \leq \frac{1}{2r}(\|T(x_0 + rx)\| + \|T(x_0 - rx)\|) \leq \frac{1}{2r} 2C = \frac{C}{r}.$$

As x is an arbitrary vector with $\|x\| \leq 1$, this implies $\|T\| \leq \frac{C}{r}$ for all $T \in \mathcal{F}$. □

Corollary 3.7. *Let X be a Banach space, $\mathcal{F} \subseteq \mathcal{B}(X, Y)$, and* $\sup\{\|Tx\| : \ T \in \mathcal{F}\} < \infty$ *for all $x \in X$. Then* $\sup\{\|T\| : \ T \in \mathcal{F}\} < \infty$.

Proof. Direct from choosing X as your subset of second category, and applying the previous theorem. □

Notation: Let $D_X = \{x \in X : \ \|x\| < 1\}$ be the open unit ball in X.

Lemma 3.8. *Let X be a Banach space, Y a normed space, and $T \in \mathcal{B}(X, Y)$. If $D_Y \subseteq \overline{T(D_X)}$, then $D_Y \subseteq T(D_X)$.*

Proof. Let $A = D_Y \cap T(D_X)$. We first claim that $D_Y \subseteq \overline{A}$. If we let $y \in D_Y$, we know that $y \in \overline{T(D_X)}$; so for any $\epsilon > 0$, there exists $y' \in T(D_X) \cap D_Y$ which is ϵ-close to y. Through this approximation, we have $y \in T(D_X)$, and hence $D_Y \subseteq \overline{A}$. We wish to show that $D_Y \subseteq T(D_X)$. Take $z \in D_Y$- so $\|z\| < 1$. Let $\delta > 0$ such that $\|z\| < 1 - \delta < 1$. Letting $y = \frac{z}{1-\delta}$, we have $\|y\| = \frac{\|z\|}{1-\delta} < 1$. If we first choose $y_0 = 0$, as $D_Y \subseteq \overline{A}$ there exists $y_1 \in A$ such that $\|y_1 - y\| < \delta$. This implies $y - y_1 \in \delta D_Y$. But as $\delta D_Y \subseteq D_Y$, there exists $y_2 \in A$ such that $y_2 - y_1 \in \delta A$ and $\|y_2 - y_1\| < \delta^2$. Continue this process inductively, finding $y_n \in A$ such that $y_n - y_{n-1} \in \delta^{n-1} A$ and $\|y_n - y_{n-1}\| < \delta^n$. This constructs a sequence $\{y_n\}_{n=0}^{\infty}$ where $y_n - y_{n-1} = Tx_n$, with $x_n \in \delta^{n-1} D_X$ (as $y_n \in A$ for $n \in \mathbb{N}$). Consider the convergent series $x_0 = \sum\limits_{n=1}^{\infty} x_n$. We see $\|x_0\| < \frac{1}{1-\delta}$, and $Tx_0 = \sum y_n - y_{n-1} = y$. Therefore, $T((1-\delta)x_0) = (1-\delta)y = z$, with $(1-\delta)x_0 \in D_X$. This shows every $z \in D_Y$ has the form $z = Tx_0$ for some $x_0 \in D_X$. Hence, $D_Y \subseteq T(D_X)$. □

Theorem 3.9 (Open Mapping Theorem)**.** *If X, Y are Banach spaces and $T : X \to Y$ is a bounded surjective operator then T is an open mapping- i.e., $T(U)$ is open in Y for every open $U \subseteq X$.*

Proof. Let $G = D_X$. Now, $X = \bigcup\limits_{n=1}^{\infty} nG$; if we apply T, we find $T(X) = \bigcup\limits_{n=1}^{\infty} nT(G)$. Let $F = \overline{T(G)}$. As T is surjective, $T(X) = Y$; this implies $Y = \bigcup\limits_{n=1}^{\infty} nF$. By the Baire Category Theorem there exists an $n_0 \in \mathbb{N}$ such that $D_r(y_0) \subseteq n_0 F$ for some $r > 0$ and $y_0 \in Y$. But this implies $D_\delta(y_1) \subseteq F$ for scaled $\delta > 0$ and $y_1 \in Y$. We note that F is symmetric (i.e if $y \in F$, then $-y \in F$ as well) and contains 0. As we have an open ball lying inside F, we may also conclude $D_\delta(-y_1) \subseteq F$. By convexity of F, we then have translation $D_\delta(0) \subseteq F$. By the previous lemma, we may then conclude $D_\delta(0) \subseteq T(G)$. This tells us 0 is an interior point of $T(G)$. Then, through the same process of translations of $D_\delta(0)$ if we let $Tx_0 \in T(G)$ where $\|x_0\| < 1$ by considering $\epsilon > 0$ such that $x_0 + \epsilon G \subseteq G$, we have $Tx_0 + \epsilon T(G) \subseteq T(G)$. Hence, $Tx_0 + \epsilon \cdot \delta G \subseteq T(G)$, which implies $T(G)$ is open. From this, it is clear $T(D_r(0))$ is open in Y for any $r > 0$. Again, by translations we may claim $T(D_r(x_0)) = T(x_0 + D_r(0)) = Tx_0 + T(D_r(0))$ is open in Y as well. This shows T is an open mapping, which completes the proof. □

Theorem 3.10 (Inverse Mapping Theorem)**.** *Let X, Y be Banach spaces, and $T \in \mathcal{B}(X, Y)$ be bijective. Then T^{-1} is bounded.*

Proof. Applying the Open Mapping Theorem for T, we know that T is an open and bijective mapping. Hence, T^{-1} is continuous, and therefore bounded. □

Theorem 3.11 (Closed Graph Theorem)**.** *Let X, Y be Banach spaces and $T : X \to Y$ be a linear map. The following are equivalent:*

(i) T is bounded;
(ii) $GrT = \{(x, Tx) : x \in X\}$ is closed in the Banach space $X \oplus_1 Y$ (i.e. the graph of T is closed).

Proof.
(i) ⇒ *(ii)* Suppose T is bounded, and let $\{(x_n, y_n)\}_{n\in\mathbb{N}} \subseteq \text{Gr}T$ with $(x_n, y_n) \to (x, y)$ as $n \to \infty$. By definition, $y_n = Tx_n$ for each $n \in \mathbb{N}$. Convergence in $\text{Gr}T$ tells us $x_n \to x$ and $Tx_n \to y$ as $n \to \infty$. As T is bounded, it is continuous and hence $Tx_n \to Tx$ in X; then by the uniqueness of limits in a Banach space, we have $Tx = y$. Hence, $\text{Gr}T$ is closed.

(ii) ⇒ *(i)* Suppose $\text{Gr}T \subseteq X \oplus_1 Y$ is closed. Consider the map $\Gamma : \text{Gr}T \to X$ given by $\Gamma((x, Tx)) = x$ for all $x \in X$. We claim Γ is injective- if $\Gamma((x, Tx)) = 0$, then $x = 0$ by definition of Γ. As $Tx = 0$ when $x = 0$, we have $\ker \Gamma = \{(0, 0)\}$. Hence, Γ is injective. It is also clear that Γ is a surjective map. Finally, as $\text{Gr}T$ is closed, it is complete as a space in itself. Then Γ satisfies the assumptions of the Inverse Mapping Theorem. This means Γ^{-1} is bounded- so there exists a $C > 0$ such that $\|\Gamma^{-1}(x)\| \le C\|x\|$ for all x. But then

$$\begin{aligned}\|\Gamma^{-1}(x)\| &= \|(x, Tx)\| \le C\|x\| \\ &= \|x\| + \|Tx\| \le C\|x\| \\ &\Rightarrow \|Tx\| \le (C-1)\|x\|,\end{aligned}$$

for all $x \in X$. Hence, T is a bounded map. □

Note: The meaning of $\text{Gr}T$ being closed is as follows: if $(x_n, Tx_n) \to (x, y)$ as $n \to \infty$, then $y = Tx$ and so $Tx_n \to Tx$. On the other hand, the meaning of T being continuous is if $x_n \to x$, then $Tx_n \to Tx$. There is an important subtlety here- the closure of $\text{Gr}T$ *guarantees* the existence in the graph of a limit for a sequence; it is not necessarily a given that a sequence in the graph must converge.

Example 3.12 (Schur multipliers)**.**

Definition 3.13. *If $A, B \in M_n$, the Schur product (also known as the Hadamard product) $A * B$ is the entry-wise product of the two matrices.*

We'll look at the infinite dimensional space ℓ^2 (although the general argument we make can be made for any ℓ^p space with $1 \le p < \infty$).

Take $T \in \mathcal{B}(\ell^2)$ (i.e. $T : \ell^2 \to \ell^2$ is bounded and linear) and let T be such that

$$\begin{aligned}T(1, 0, 0, \ldots) &= (a_{11}, a_{21}, a_{31}, \ldots), \\ T(0, 1, 0, \ldots) &= (a_{12}, a_{22}, a_{32}, \ldots),\end{aligned}$$

and onward. We write $T = [a_{ij}]_{i,j}$ where $i, j \in \mathbb{N}$ as an infinite matrix. As in finite dimensional spaces, if we apply T to the n^{th} basis vector e_n we get $Te_n = n^{th}$ column of T.

Just as in finite dimensions, the infinite-dimensional matrices correspond exactly with maps $T \in \mathcal{B}(\ell^2)$; we are allowed to say this as the canonical basis vectors span a dense set in ℓ^2- so we can approximate sequences in ℓ^2 by finite linear combinations of basis vectors.

Note: *Not all $\mathbb{N} \times \mathbb{N}$ matrices arise from bounded operators. As an example, take*

$$J = \begin{bmatrix} 1 & 1 & \cdots \\ 1 & \ddots & \\ \vdots & & \end{bmatrix}.$$

The space $\mathcal{B}(\ell^2)$ is closed under Schur multiplication- to see why, if we take $A, B \in \mathcal{B}(\ell^2)$ we at least know $A \otimes B \in \mathcal{B}(\ell^2 \otimes \ell^2)$. As $A * B = P(A \otimes B)P$ for a projection P (i.e. $A * B$ is a compression of $A \otimes B$) we have $A * B \in \mathcal{B}(\ell^2)$. We say $T = (t_{ij})$ is a Schur multiplier if $A \in \mathcal{B}(\ell^2)$ implies $T * B \in \mathcal{B}(\ell^2)$.

Theorem 3.14. *If T is a Schur multiplier on $\mathcal{B}(\ell^2)$, then the corresponding map*

$$A \mapsto T * A$$

from $\mathcal{B}(\ell^2)$ into $\mathcal{B}(\ell^2)$ is bounded.

Proof. Apply the Closed Graph Theorem: if we take sequence of matrices $(a_{ij}^n) \to 0$, and $(t_{ij}a_{ij}^n) \to (b_{ij})$ as $n \to \infty$ we necessarily have that $(t_{ij}a_{ij}^n) \to 0$ as well. Then $(b_{ij}) = 0$, and hence $b_{ij} = 0$ for all $i, j \in \mathbb{N}$. This shows the graph of the map above is closed, and hence the map itself must be bounded. □

3.2. **Modes of convergence.** We'll begin by discussing the motivation for a theorem we will soon be proving. Suppose X is an infinite dimensional normed space- then $B_X := \{x \in X : \|x\| \leq 1\}$ is not compact. This means we cannot pick a convergent subsequence given an arbitrary sequence in B_X.

Recall that if K is compact and $(x_n)_{n=1}^\infty \subseteq K$, there exists a subsequence $(x_{n_k})_{n_k}$ and $y \in K$ such that $x_{n_k} \to y$ as $k \to \infty$. Given that compactness is a nice property to have, how can we get around this? The fact that B_X is not compact for an infinite dimensional space is not ideal, as this means things do not necessarily converge nicely. This all plays a role in motivating our discussion of Alaoglu's Theorem; our aim is to show that the unit ball of X^* is compact, but for a weaker mode of convergence. To introduce this different mode of convergence, we replace "open-ness" in terms of distance on spaces by "open sets".

Definition 3.15. *Let X be a non-empty set. Suppose τ is a family of subsets of X which we call open, and which satisfy:*

- *(i) If $U_\alpha \in \tau$ for all $\alpha \in I$, then $\cup_{\alpha \in I} U_\alpha \in \tau$.*
- *(ii) If $U_1, U_2 \in \tau$, then $U_1 \cap U_2 \in \tau$.*
- *(iii) $\emptyset, X \in \tau$.*

Such a $\tau \subseteq \mathcal{P}(X)$ is called a topology on X, and the elements of τ are a replacement of the open sets in a metric space.

Definition 3.16. *Let $(x_n)_{n=1}^\infty \subseteq X$. We say the sequence converges to $x \in X$ if for all $U \in \tau$ with $x \in U$, there exists an $n_0 \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq n_0$.*

Next, we'll replace the index set $\mathbb{N}$ with a directed set $\mathbb{A}$, where $\mathbb{A}$ has a partial order $\leq$ which necessarily satisfies:

- (i) $a \leq a$ for all $a \in \mathbb{A}$;
- (ii) $a \leq b$ and $b \leq c$ implies $a \leq c$;
- (iii) $a \leq b$ and $b \leq a$ implies $a = b$.

Definition 3.17. *$\mathbb{A}$ is called a directed set if for all $a, b \in \mathbb{A}$, there exists a $c \in \mathbb{A}$ such that $a \leq c$ and $b \leq c$.*

Definition 3.18. *A net in X (where X is a topological space) is a mapping $f : \mathbb{A} \to X$ where $\mathbb{A}$ is directed. Slightly less formally, we consider the mapping as an indexing of elements in X by $\mathbb{A}$- i.e., we have the net $(x_\alpha)_{\alpha\in\mathbb{A}}$, with $x_\alpha \in X$.*

We may now talk about general convergence.

Definition 3.19. *We say that a net $(x_\alpha)_{\alpha\in\mathbb{A}}$ converges to an element $x \in X$ if for every $U \in \tau$ with $x \in U$, there exists $\alpha_0 \in \mathbb{A}$ such that $x_\alpha \in U$ for every $\alpha \geq \alpha_0$.*

Example 3.20 (Partial sums in Riemann integration)**.** *Suppose we wish to find the Riemann integral of a function f. We recall that the integral is defined using limit sup's and inf's on partitions of our interval, and we use these to index the sequence of partial sums. The issue we note is that these partitions are not guaranteed to be comparable- for any partition P and Q, we don't necessarily have $P \subseteq Q$ or $Q \subseteq P$ depending on our choice of points in the partition. However, we can define a partial order based on inclusion- as $P \subseteq P \cup Q$, and $Q \subseteq P \cup Q$ for any two partitions, we can use this to turn the set of partitions into a directed set. This allows us to define integration by convergence of nets of partial sums.*

Definition 3.21. *A subset K of a topological space X is compact if for every net $(x_\alpha)_{\alpha\in\mathbb{A}} \subseteq K$ there exists a subnet $(x_\beta)_{\beta\in\mathbb{B}}$ that converges to a point $x \in K$. (Here $\mathbb{B} \subseteq \mathbb{A}$).*

For more information on nets and convergence, see John Kelly's book "General Topology".

Example 3.22. *Consider ℓ^p with $1 \leq p < \infty$, and the canonical basis $(e_i)_{i=1}^\infty$ for ℓ^p. It is clear $\|e_i\|_p = 1$, with $(e_i)_{i=1}^\infty$ <u>not</u> convergent in norm. However, if we look at the number of 0's before the first 1 occurs in vector e_i, we have $i - 1$ 0's for basis element e_i. Letting $i \to \infty$, this suggests an "almost convergence" of $e_i \to 0$. This forms some of the basic intuition for weak convergence.*

Definition 3.23. *Let X be a normed space. The weak topology on X is defined by: $(x_\alpha)_{\alpha\in\mathbb{A}}$ converges weakly to x if $f(x_\alpha) \to_{\alpha\in\mathbb{A}} f(x)$, for all $f \in X^*$.*

Note: Look at $(e_n)_{n=1}^\infty \in \ell^2$. Then $\|e_n\| = 1$, and hence $e_n \not\to 0$ in norm. However, $e_n \to_w 0$ in the weak topology (and we write $x_\alpha \to_w x$ if x_α is weakly convergent to x). Indeed- to see this, let $f \in (\ell^2)^*$. There exists a sequence $(y_n)_{n\in\mathbb{N}} \in \ell^2$ such that $f(x) = \sum_{n=1}^{\infty} x_n y_n$, for every $x = (x_n)_{n=1}^\infty \in \ell^2$ (this holds, as 2 is its own conjugate index). We have $f(e_n) = y_n$, and as $(y_n)_{n=1}^\infty \in \ell^2$ we know $y_n \to 0$ as $n \to \infty$. So $f(e_n) \to 0 = f(0)$ for $n \to \infty$. Hence, $e_n \to_w 0$.

Definition 3.24. *Let X be a normed space with dual X^*. The weak* topology on X^* is defined by the convergence: $(f_\alpha)_{\alpha\in\mathbb{A}}$ converges to $f \in X^*$ if $f_\alpha(x) \to_{\alpha\in\mathbb{A}} f(x)$ for every $x \in X$.*

Notes:

(i) X^* has the following two convergence modes:
- w^* : $f_\alpha \to_{w^*} f$ if $f_\alpha(x) \to_{\alpha\in\mathbb{A}} f(x)$ for all $x \in X$.
- w : $f_\alpha \to_w f$ if $F(f_\alpha) \to_{\alpha\in\mathbb{A}} F(f)$ for every $F \in X^{**}$.

(ii) We also see that if $f_\alpha \to_w f$, then $f_\alpha \to_{w*} f$, as we can take F to be evaluation on a chosen point $x \in X$.

(iii) Weak convergence and weak* convergence are equivalent for reflexive spaces.

Example 3.25. *Recall that* $(L^1(0,1))^* = L^\infty(0,1)$ *(here we use the Lebesgue measure). We'll look at* $\chi_{[0,1/n]}$*. It is clear that* $\|\chi_{[0,1/n]}\|_\infty = 1$ *for all* $n \in \mathbb{N}$*. We also note* $\chi_{[0,1/n]} \to_{w*} 0$ *as* $n \to \infty$*, even though* $\chi_{[0,1/n]}$ *clearly does not converge to* 0 *in norm.*

Definition 3.26. *Given two topological spaces* X_1, X_2 *the product* $X_1 \times X_2$ *is defined by:*

$$(x_\alpha^{(1)}, x_\alpha^{(2)}) \to_{\alpha\in\mathbb{A}} (x^{(1)}, x^{(2)})$$

if $x_\alpha^{(1)} \to_{\alpha\in\mathbb{A}} x^{(1)}$ *and* $x_\alpha^{(2)} \to_{\alpha\in\mathbb{A}} x^{(2)}$*.*

In particular: by definition $(x_n, y_n) \to (x, y)$ as $n \to \infty$ if $x_n \to x$ and $y_n \to y$ as $n \to \infty$.

Note: Originally, we are given two convergence structures: one on X_1, and the other on X_2. From these two, I define convergence on $X_1 \times X_2$.

Just as we are able to form a product space given only two topological spaces, suppose we are given some family of topological spaces $\{X_i\}_{i\in I}$, where I is an index set. We can define a product space in exactly the same way.

Definition 3.27. *For* $\{X_i\}_{i\in I}$ *as above, the product space* $X = \prod_{i\in I} X_i$ *consists of families* $(x_i)_{i\in I}$ *with* $x_i \in X_i$ *for every* $i \in I$*. The convergence in* X *is pointwise/coordinate-wise/entry-wise.*

Theorem 3.28 (Tychonoff)**.** *The product of any family of compact topological spaces is a compact topological space.*

Note: The result is straightforward in the case the family consists of finitely many topological spaces (as we can prove the result for two, and then apply induction).

Example 3.29 (The Cantor set)**.** *We look at the Cantor set* $C = \prod_{i\in\mathbb{N}} \{0,1\} = \{0,1\}^{\mathbb{N}}$*. The elements of* $\{0,1\}^{\mathbb{N}}$ *are* $(0,1)$*-strings with* $(x_i)_{i\in\mathbb{N}}$ *and* $x_i \in \{0,1\}$*. If* $x_i = 0$ *this indicates* x_i *is in the left* $1/3^{rd}$ *interval, and if* $x_i = 1$ *this indicates* x_i *is in the right* $1/3^{rd}$ *interval.*

Quote of the day: "There's a lot of space in infinite dimensional spaces." - Dr. Todorov

Notation: Let $B_{X^*} = \{f \in X^* : \|f\| \le 1\}$.

Theorem 3.30 (Banach-Alaoglu)**.** *The closed unit ball of* X^* *is compact in the weak** *topology for any normed space* X*.*

What does this mean? Specifically, it means that if $f_n \in X^*$ for $n \in \mathbb{N}$ where $\|f_n\| \le 1$ there exists a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ and $f \in X^*$ such that $f_{n_k}(x) \to f(x)$ as $k \to \infty$ for all $x \in X$. The same will be true for arbitrary nets.

Proof. Every functional $f \in X^*$ is completely determined by the family of scalars: $(f(x))_{x\in B_X}$. Consider the map $\gamma : B_{X^*} \to \prod_{x\in B_X} \mathbb{D}_X$ (here $\mathbb{D}_X$ is the unit complex disk) where

$$\gamma(f) = (f(x))_{x\in B_X}.$$

As $\|f\| \le 1$, we see $|f(x)| \le \|f\|\|x\| \le 1$. This function is well-defined, and injective. Let $\Gamma = \gamma(B_{X^*})$. By Tychonoff's theorem, we know that $\prod_{x\in B_X} \mathbb{D}_X$ is compact (as each $\mathbb{D}_X$ is compact), with $\Gamma \subseteq \prod_{x\in B_X} \mathbb{D}_X$. We claim Γ is closed (and hence compact). To see why showing this suffices to finish the proof: as $f_\alpha \to_{w^*} f$ if and only if $f_\alpha(x) \to f(x)$ for all $x \in X$, this is equivalent to saying $f_\alpha \to_{w^*} f$ if and only if $f_\alpha(x) \to f(x)$ for all

$x \in B_X$. This holds if and only if $\gamma(f_\alpha) \to \gamma(f)$ in the product topology, by the injectivity of γ. Therefore, the compactness of Γ would imply (through the chain of equalities just mentioned) that B_{X^*} is compact. Therefore, to show Γ is closed: exercise! (It should be straightforward from the definition of the convergence in $\prod_{x \in B_X} \mathbb{D}_X$). □

Example 3.31. *Recall that* $(\ell^\infty)^* \neq \ell^1$. *We will prove this via the Banach-Alaoglu theorem.*

Proof. For every $n \in \mathbb{N}$, let $\mu_n : \ell^\infty \to \mathbb{C}$ where $\mu_n((x_i)_{i\in\mathbb{N}}) = \frac{x_1+\cdots+x_n}{n}$. We note that μ_n satisfies the conditions for Banach-Alaoglu's theorem- so there exists a $\mu : \ell^\infty \to \mathbb{C}$ where $\mu_{n_k} \to_{w*} \mu$ as $k \to \infty$. Then $\mu \in (\ell^\infty)^*$, but $\mu \notin \ell^1$. If $\mu((x_i)_{i\in\mathbb{N}}) = \sum x_i y_i$, then $\mu(e_n) = y_n = \lim_{k\to\infty} \mu_{n_k}(e_n) = 0$. Then $y_n = 0$ for all $n \in \mathbb{N}$. However, $\mu \neq 0$. As we have reached a contradiction, we conclude $(\ell^\infty)^* \neq \ell^1$. □

4. Hilbert spaces

4.1. Inner product spaces.

Definition 4.1. *An inner product space is a complex vector space $\mathcal{H}$ equipped with a map $\mathcal{H} \times \mathcal{H} \to \mathbb{C}$ denoted $(\cdot,\cdot)$ satisfying:*

(i) $(x + \lambda y, z) = (x, z) + \lambda(y, z)$ *for all* $x, y, z \in \mathcal{H}$ *and* $\lambda \in \mathbb{C}$;

(ii) $(x, y) = \overline{(y, x)}$;

(iii) $(x, x) \geq 0$ *and* $(x, x) = 0$ *if and only if* $x = 0$ *for all* $x \in \mathcal{H}$.

The map $(\cdot,\cdot)$ is called an inner product.

The norm of an inner-product

Proposition 4.2 (Cauchy-Schwarz). *If $\mathcal{H}$ is an inner product space and $x, y \in \mathcal{H}$ then*

$$|(x, y)| \leq \sqrt{(x, x)}\sqrt{(y, y)}.$$

Notation: We set $\|x\| = \sqrt{(x, x)}$.

Proof. By properties of the inner product, for $y \neq 0$ we know

$$(x - \frac{(x,y)}{\|y\|^2}y, x - \frac{(x,y)}{\|y\|^2}y) \geq 0.$$

Then

$$\begin{aligned}(x - \frac{(x,y)}{\|y\|^2}y, x - \frac{(x,y)}{\|y\|^2}y) &= (x, x) - \frac{(x,y)}{\|y\|^2}(y, x) - \frac{(y,x)}{\|y\|^2}(x, y) + \frac{(x,y)(y,x)}{\|y\|^4}(y, y) \\ &= \|x\|^2 - \frac{|(x,y)|^2}{\|y\|^2} - \frac{|(x,y)|^2}{\|y\|^2} + \frac{|(x,y)|^2}{\|y\|^2} \\ &= \|x\|^2 - \frac{|(x,y)|^2}{\|y\|^2} \geq 0.\end{aligned}$$

But then $|(x, y)|^2 \leq \|x\|^2\|y\|^2$. Taking square roots, this completes the proof. □

Note: Equality in the Cauchy-Schwarz inequality holds if and only if the vectors are proportional/collinear. To see this, we first note that if $|(x, y)| = \|x\|\|y\|$, then $(x - \frac{(x,y)}{\|y\|^2}y, x - \frac{(x,y)}{\|y\|^2}y) = 0$. This means $x - \frac{(x,y)}{\|y\|^2}y = 0$, and hence $x = \frac{(x,y)}{\|y\|^2}y$. Conversely, if $x = \lambda y$ (where $\lambda \in \mathbb{C}$) then $|(x, y)| = |(\lambda y, y)| = |\lambda|\|y\|^2$ while $\|x\|\|y\| = |\lambda|\|y\|^2$.

Proposition 4.3. *The function* $\|\cdot\| : \mathcal{H} \to \mathbb{R}^+$ *where* $x \mapsto \|x\| = (x,x)^{1/2}$ *is a norm.*

Proof. We focus on subadditivity, as the others are immediate. Let $x, y \in \mathcal{H}$. We see

$$\|x+y\|^2 = (x+y, x+y) = \|x\|^2 + (x,y) + (y,x) + \|y\|^2$$
$$= \|x\|^2 + 2\mathrm{Re}(x,y) + \|y\|^2 \le \|x\|^2 + 2|(x,y)| + \|y\|^2 \le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2$$
$$= (\|x\| + \|y\|)^2,$$

by Cauchy-Schwarz. Taking square roots, we have $\|x+y\| \le \|x\| + \|y\|$. □

Definition 4.4. *An inner product space* $\mathcal{H}$ *is called a Hilbert space if* $\mathcal{H}$ *is complete (with respect to the norm induced by the inner product).*

Definition 4.5. *Two vectors* $x, y \in \mathcal{H}$ *are orthogonal if* $(x,y) = 0$. *We write* $x \perp y$ *in this case.*

Theorem 4.6 (Pythagoreas)**.** *If* $\mathcal{H}$ *is an inner product space and* $x_1, \ldots, x_n$ *are pairwise orthogonal vectors then*

$$\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2.$$

Proof. By induction, it suffices to show the above when we look only at x_1 and x_2. We have

$$\|x_1 + x_2\|^2 = (x_1 + x_2, x_1 + x_2) = \|x_1\|^2 + (x_1, x_2) + (x_2, x_1) + \|x_2\|^2$$
$$= \|x_1\|^2 + \|x_2\|^2.$$

□

Proposition 4.7 (Parallelogram identity)**.** *Let* $\mathcal{H}$ *be an inner product space, and let* $x, y \in \mathcal{H}$. *Then*

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

<u>Fact:</u> A normed space is an inner product space if and only if its norm satisfies the parallelogram identity.

For more reference on inner product spaces, see Paul Halmos' "Hilbert Space Problem Book".

Proposition 4.8 (Polarization identity)**.** *If* $\mathcal{H}$ *is an inner product space and* $T : \mathcal{H} \to \mathcal{H}$ *is a linear operator (not necessarily bounded) then*

$$(Tx, y) = \frac{1}{4}[(T(x+y),(x+y)) - (T(x-y),(x-y))) + i(T(x+iy),(x+iy)) - i(T(x-iy),(x-iy))].$$

Corollary 4.9. *A linear operator* $T : \mathcal{H} \to \mathcal{H}$ *is equal to* 0 *if and only if* $(Tx, x) = 0$ *for all* $x \in \mathcal{H}$.

Proof. Suppose $(Tx, x) = 0$ for all $x \in \mathcal{H}$. By the Polarization Identity, this implies $(Tx, y) = 0$ for all $y \in \mathcal{H}$. Choosing $y = Tx$, then $(Tx, Tx) = 0$- hence $Tx = 0$, and so $T = 0$. □

Theorem 4.10. *Let $\mathcal{H}$ be an inner product space. The metric space completion of $\mathcal{H}$ is a Hilbert space.*

Proof. (Sketch) Recall that the completion of $\mathcal{H}$ has elements

$$(x_1, x_2, x_3, \ldots)$$

where the sequences are all Cauchy. The inner product extends to the space of these sequences by letting

$$((x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}) = \lim_{n\to\infty}(x_n, y_n).$$

The previous limit exists, as $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}$ are Cauchy sequences, and we are defining the inner product for the completion ourselves. □

Example 4.11.

(i) Euclidean space $\mathbb{C}^n$ (or ℓ^2_n) where $\mathbb{C}^n$ is the underlying vector space and

$$((x_i)_{i=1}^n, (y_i)_{i=1}^n) = \sum_{i=1}^n x_i\overline{y_i}.$$

(ii) $\ell^2 = \{(x_i)_{i=1}^\infty : \ x_i \in \mathbb{C}, \text{ and } \sum_{i=1}^\infty |x_i|^2 < \infty\}$. The inner product here is

$$((x_i), (y_i)) = \sum_{i=1}^\infty x_i\overline{y_i}.$$

This pairing makes sense, as $(\ell^2)^* = \ell^2$.

(iii) $\mathcal{L}^2[a,b]$, where the underlying space is the space of all functions $f : [a,b] \to \mathbb{C}$ which are measurable and $\int_a^b |f|^2 d\lambda < \infty$. The inner product here is

$$(f, g) = \int_a^b f(x)\overline{g(x)}d\lambda.$$

Note that this example and the previous few are all Hilbert spaces.

(iv) This is an example of an inner product space which is not a Hilbert space. Take $C[a,b]$ as our underlying vector space, and define

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

(here we use Riemann integration, not the Lebesgue integral). This space is not complete, but its completion can be identified with $\mathcal{L}^2[a,b]$.

Theorem 4.12. *Let $\mathcal{H}$ be a Hilbert space and $E \subseteq \mathcal{H}$ a closed subspace. Let $E^\perp = \{y \in \mathcal{H} : \ y \perp x, \text{ for all } x \in E\}$. For every $x \in \mathcal{H}$, there exists a unique $y \in E$ such that*

$$dist(x, E) = \inf_{y'\in E} \|x - y'\| = \|x - y\|.$$

Moreover, $x - y \in E^\perp$.

Proof. Assume $\mathcal{H}$ and E are as above, and let $x \in \mathcal{H}$. We have

$$d := \text{dist}(x, E) = \inf_{y'\in E} \|x - y'\| \geq 0$$

by definition of norm $\|\cdot\|$. Let $y_n \in E$, for $n \in \mathbb{N}$ such that $\|x - y_n\| \to d$ as $n \to \infty$. We wish to show that $(y_n)_{n\in\mathbb{N}}$ is Cauchy. The parallelogram identity implies

$$\|y_n - y_m\|^2 = \|(x - y_m) - (x - y_n)\|^2 = 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - \|2x - y_n - y_m\|^2.$$

We note that $\|2x - y_n - y_m\|^2 = 4\|x - \frac{y_n+y_m}{2}\|^2 \geq 4d^2$. Furthermore, as $m, n \to \infty$ both $2\|x - y_n\|^2 \to 2d^2$ and $2\|x - y_m\|^2 \to 2d^2$. Putting everything together, we have

$$2\|x - y_n\|^2 + 2\|x - y_m\|^2 - \|2x - y_n - y_m\|^2 \to 2d^2 + 2d^2 - (4d^2 + \epsilon) \leq 0.$$

As the individual terms in the previous limit are equal to $\|y_n - y_m\|^2$, and $\|y_n - y_m\|^2 \geq 0$ for $n, m \in \mathbb{N}$, this implies $\|y_n - y_m\|^2 \to \epsilon' \geq 0$. Then $\|y_n - y_m\|^2 \to 0$, and hence $(y_n)_{n\in\mathbb{N}}$ must be a Cauchy sequence. Since $\mathcal{H}$ is a Hilbert space, there exists $y \in \mathcal{H}$ such that $y_n \to y$ as $n \to \infty$. Furthermore, as E is a closed subspace with $(y_n)_{n\in\mathbb{N}}$, necessarily we have $y \in E$. But then by continuity of the norm,

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

For the second claim, let $z = x - y$, and consider the vectors $z - \epsilon y'$ where $\epsilon > 0$ and $y' \in E$ are temporarily fixed. We have

$$\|z - \epsilon y'\|^2 = \|z\|^2 - 2\epsilon \mathrm{Re}(z, y') + \epsilon^2\|y'\|^2.$$

Note that when $\epsilon > 0$ is small enough, the sign of the previous sum is determined by $-2\epsilon\mathrm{Re}(z, y')$. We also note that $\|z - \epsilon y'\|^2 \geq d^2$, by definition of z. If we replace y' by $y'' = (z, y')y'$ (which we are allowed to do without loss of generality) then $\mathrm{Re}(z, y'') = |(z, y')|^2 \geq 0$. This would mean (for $\epsilon \to 0$)

$$\|z\|^2 - 2\epsilon|(z, y')|^2 + \epsilon^2\|y''\|^2 = d^2 - 2\epsilon|(z, y')|^2 + \epsilon^2\|y''\|^2 < d^2.$$

However, as $\|z - \epsilon y''\|^2 \geq d^2$, and

$$\|z - \epsilon y''\|^2 = \|z\|^2 - 2\epsilon|(z, y')|^2 + \epsilon^2\|y''\|^2,$$

we have both $\|z - \epsilon y''\|^2 \geq d^2$ and $\|z - \epsilon y''\|^2 < d^2$. As this is only possible when $(z, y') = 0$, we must have $(x - y, y') = 0$. Then as $y' \in E$ was arbitrary, this means $x - y \in E^\perp$. □

Remark: As a consequence of the previous theorem, given $x \in \mathcal{H}$ and $E \subseteq \mathcal{H}$ a closed subspace then $x = y + z$ where $y \in E, z \in E^\perp$ and the decomposition is unique- i.e., $\mathcal{H} = E \oplus E^\perp$.

To see uniqueness, suppose $x = y + z = y' + z'$ where $y, y' \in E$ and $z, z' \in E^\perp$. Then $y - y' = z' - z$, and so $y - y', z' - z \in E \cap E^\perp = \{0\}$. This forces $y = y'$, and $z = z'$. The rest of the consequence follows directly from the statement of the theorem.

Definition 4.13. *Let $\mathcal{H}$ be a Hilbert space, and $E \subseteq \mathcal{H}$ a closed subspace. The projection $P_E : \mathcal{H} \to \mathcal{H}$ is defined by letting $P_E(x) = y$, where $x = y + z$ is the unique decomposition in $E \oplus E^\perp$ for $x \in \mathcal{H}$.*

Proposition 4.14. *For Hilbert space $\mathcal{H}$ and $E \subseteq \mathcal{H}$ a closed subspace, then:*

- *(i) P_E is bounded; in fact, $\|P_E\| = 1$ if $E \neq \{0\}$.*
- *(ii) P_E is idempotent: $P_E^2 = P_E$.*
- *(iii) $\mathrm{rng} P_E = E$.*
- *(iv) $\ker P_E = E^\perp$.*
- *(v) $d(x, E) = \|x - P_E(x)\|$.*
- *(vi) $I - P_E$ is the projection onto $E^\perp$.*

Proof.
(i) If $E = \{0\}$, then $P_E = 0$- so $\|P_E\| = 0$. Else, suppose $E \neq \{0\}$. Fix any $x \in \mathcal{H}$, where $x = y + z$. Then

$$\|P_E(x)\|^2 = \|y\|^2 \leq \|y\|^2 + \|z\|^2 = \|y + z\|^2 = \|x\|^2,$$

by the Pythagorean Theorem. As $x \in \mathcal{H}$ was arbitrary, this shows $\|P_E\| \leq 1$. If we choose any $y \in E$, then $\|P_E(y)\| = \|y\|$, showing $\|P_E\| = 1$ exactly.

(ii) The proof is fairly trivial- just apply P_E and P_E^2 to any vector in $\mathcal{H}$.

(iii) We look at how P_E is defined: for $x = y + z$ with $y \in E, z \in E^\perp$ then $P_E(x) = y$. Clearly, $\text{rng} P_E \subseteq E$ as $y \in E$. On the other hand, if we take $y \in E$ then $y = y + 0$- so $y = P_E(y)$. This implies $E \subseteq \text{rng} P_E$, and hence the two are equal.

(iv) Let $z \in E^\perp$. Then $z = 0 + z$ where $0 \in E$. By definition of P_E, $P_E(z) = 0$- so $z \in \ker P_E$. If we instead take $z \in \ker P_E$, $P_E(z) = 0$ by definition. So $z = 0 + z$, with $0 \in E$. This implies $z \in E^\perp$, which shows $E^\perp = \ker P_E$.

(v) We've shown this previously.

(vi) We know that for each $x \in \mathcal{H}$, $x = y + (x - y)$ where $x - y \in E^\perp$ and $y \in E$. By definition, $y = P_E(x)$- so

$$x - y = x - P_E(x) = I(x) - P_E(x) = (I - P_E)(x).$$

This implies $I - P_E = P_{E^\perp}$. □

Theorem 4.15 (Riesz Representation Theorem)**.** *Let $\mathcal{H}$ be a Hilbert space, and $f \in \mathcal{H}^*$. Then there exists a unique $y \in \mathcal{H}$ such that*

$$f(x) = (x, y), \qquad \text{for all } x \in \mathcal{H}.$$

Note: If $y \in \mathcal{H}$ is fixed, the map $f_y : \mathcal{H} \to \mathbb{C}$ which sends $x \mapsto (x, y)$ is bounded and linear.

Indeed, linearity is straightforward (from linearity of the inner product). Also,

$$|f_y(x)| = |(x, y)| \leq \|x\|\|y\|$$

by the Cauchy-Schwarz inequality- so $\|f_y\| \leq \|y\|$. In fact, $\|f_y\| = \|y\|$ (just estimate using $x = y$).

Proof. Fix $f : \mathcal{H} \to \mathbb{C}$ bounded and linear. If $f = 0$, just take $y = 0$ and we are done. So suppose $f \neq 0$; then $E = \ker f$ is a closed subspace in $\mathcal{H}$ of codimension 1. We have $\mathcal{H} = E \oplus E^\perp$, and $\dim(E^\perp) = 1$. Let $z \in E^\perp$, with $\|z\| = 1$, and set $y = \overline{f(z)} z$. For any $x \in \mathcal{H}$, we may write $x = \psi + \lambda z$, where $\psi \in E$ and $\lambda \in \mathbb{C}$. From this, we see

$$f(x) = f(\psi + \lambda z) = f(\psi) + \lambda f(z) = \lambda f(z),$$

as $\psi \in \ker f$. On the other hand,

$$\begin{aligned} (x, y) = (\psi + \lambda z, \overline{f(z)} z) &= (\psi, \overline{f(z)} z) + \lambda (z, \overline{f(z)} z) \\ &= f(z)(\psi, z) + \lambda f(z)(z, z) = 0 + \lambda f(z) \\ &= \lambda f(z). \end{aligned}$$

As $x \in \mathcal{H}$ was arbitrary, this shows $f(x) = (x, y)$ for all $x \in \mathcal{H}$. The uniqueness of y is trivial. □

Note: Consider the map $\mathcal{H} \to \mathcal{H}^*$ given by $y \mapsto f_y$. We know the map is

(i) Isometric, as $\|y\| = \|f_y\|$ for all $y \in \mathcal{H}$.
(ii) Surjective, by the Riesz Representation Theorem.
(iii) Conjugate linear, where $f_{\lambda y_1 + y_2} = \overline{\lambda} f_{y_1} + f_{y_2}$.

Note: $\mathcal{H}^*$ becomes a Hilbert space under the inner product $(f_y, f_z) = (y, z)$.

Definition 4.16. *Let $\mathcal{H}$ be a Hilbert space. A family $\mathcal{E} \subseteq \mathcal{H}$ is called*

(i) Orthogonal, if $x, y \in \mathcal{E}$ implies $(x, y) = 0$.
(ii) Orthonormal, if $\mathcal{E}$ is orthogonal and $\|x\| = 1$ for all $x \in \mathcal{E}$.
(iii) Total, if $\mathcal{E}$ has a dense linear span

$$[\mathcal{E}] := span\{\mathcal{E}\}$$

(i.e. $\overline{[\mathcal{E}]} = \mathcal{H}$).
(iv) Complete, if it is a maximal orthogonal set.
(v) A basis, if it is a complete orthonormal set.

Note: Complete families and bases exist for every non-trivial Hilbert space $\mathcal{H}$.

Proof. We will show the existence of a basis (as the case for complete families is similar).

Consider all orthonormal families, and call their class $\mathcal{B}$. Order this class by set inclusion- this defines a partial order, where every ascending chain has an upper bound. Then by Zorn's Lemma, there exists a maximal element- this should be our basis for the space $\mathcal{H}$. □

We will show that a basis allows us to decompose any vector $x \in \mathcal{H}$- i.e., we will be able to write

$$x = \sum \lambda_i e_i, \quad \lambda_i \in \mathbb{C},\ \{e_i\} \text{ a basis}$$

where the previous sum is a convergent series.

Proposition 4.17. *Let $\mathcal{H}$ be a Hilbert space, and $\mathcal{E} \subseteq \mathcal{H}$ an orthogonal set. The following are equivalent:*

(i) $\mathcal{E}$ is total.
(ii) $\mathcal{E}$ is complete.

Proof.
$(i) \Rightarrow (ii)$ Assume that $\mathcal{E}$ is total- i.e., $\overline{[\mathcal{E}]} = \mathcal{H}$. This implies $\overline{[\mathcal{E}]}^{\perp} = \{0\}$. In particular, $\mathcal{E}^{\perp} = \{0\}$. If we suppose $u \in \mathcal{E}^{\perp}$, then $u \in [\mathcal{E}]^{\perp}$. As the inner product is continuous, we can pass to the closure to find $u \in \overline{[\mathcal{E}]}^{\perp}$. Then $u = 0$, as $\overline{[\mathcal{E}]}^{\perp} = \{0\}$. This means $\mathcal{E}$ is maximal (as a subset in $\mathcal{H}$), and therefore complete.

$(ii) \Rightarrow (i)$ Now suppose $\mathcal{E}$ is complete; this means $\mathcal{E}^{\perp} = \{0\}$ by definition. Then $\mathcal{E}^{\perp\perp} = \mathcal{H}$; however, $\mathcal{E}^{\perp\perp} = \overline{[\mathcal{E}]}$. Indeed, $\mathcal{H} = \overline{[\mathcal{E}]} \oplus \overline{[\mathcal{E}]}^{\perp}$ and $\mathcal{H} = \mathcal{E}^{\perp\perp} \oplus \mathcal{E}^{\perp}$. Then

$$\mathcal{H} = \mathcal{E}^{\perp\perp} \oplus \mathcal{E}^{\perp} = \mathcal{E}^{\perp\perp} \oplus \overline{[\mathcal{E}]}^{\perp},$$

which implies $\overline{[\mathcal{E}]} = \mathcal{E}^{\perp\perp}$. This shows $\overline{[\mathcal{E}]} = \mathcal{H}$, and so $\mathcal{E}$ is total. □

In a standard linear algebra course, an important tool for finite-dimensional inner product spaces is known as the "Gram-Schmidt" process. This process allows us to take a finite set of linearly independent vectors, and orthonormalize it. The Gram-Schmidt process can still be applied in infinite dimensional Hilbert spaces: if $\mathcal{E} = \{x_1, x_2, \ldots, \} \subseteq \mathcal{H}$ is a countable family of vectors, we can still create a linearly independent orthonormal set. As an example, first set $e_1 = \frac{x_1}{\|x_1\|}$. Let $M = [e_1]$. and $e_2 = \frac{x_2 - P_{M_1}(x_2)}{\|x_2 - P_{M_1}(x_2)\|}$. Continue inductively, with $M_k = [e_1, \ldots, e_k]$ and $e_{k+1} = \frac{x_{k+1} - P_{M_k}(x_{k+1})}{\|x_{k+1} - P_{M_k}(x_{k+1})\|}$. The family $\{e_1, e_2, \ldots, \}$ is orthonormal, and has the same closed span as $\mathcal{E}$.

Example 4.18.

(i) *In ℓ^2, consider the "standard basis" $\{e_i : i \in \mathbb{N}\}$ with*

$$e_1 = (1, 0, 0, \ldots),$$
$$e_2 = (0, 1, 0, \ldots),$$
$$\ldots$$

To see completeness for this family, consider $x \in \ell^2$ where $x \perp e_n$ for all $n \in \mathbb{N}$. Write $x = (x_1, x_2, x_3, \ldots)$ where $x_n \in \mathbb{C}$ for each $n \in \mathbb{N}$. Then $x_n = (x, e_n) = 0$, which forces $x = 0$.

(ii) *Let $\mathcal{H} = L^2(\mathbb{T})$ (here $\mathbb{T} = [0, 2\pi)$ equipped with the Lebesgue measure). The inner product here is*

$$(f, g) = \frac{1}{2\pi} \int_0^{2\pi} f(t)\overline{g(t)}dt.$$

For a basis of the space: for each $n \in \mathbb{Z}$, let $\xi_n : \mathbb{T} \to \mathbb{C}$ be such that

$$\xi_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}.$$

Then $\{\xi_n\}_{n\in\mathbb{Z}}$ is a basis. The Fourier coefficients of a function $\xi \in L^2(\mathbb{T})$ are:

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} \xi(t)e^{-int}dt = (\xi, \xi_n),$$

for $n \in \mathbb{Z}$.

We can associate to each function a sequence of complex numbers which encodes all necessary information by letting

$$\xi \mapsto (c_n)_{n\in\mathbb{Z}}.$$

One can then reconstruct the function ξ from $(c_n)_{n\in\mathbb{Z}}$.

Theorem 4.19. *Let $\mathcal{H}$ be a Hilbert space, and $e_1, e_2, \ldots$ be orthonormal vectors. Let $M = \overline{[e_1, e_2, \ldots]}$. Then*

(i) *For each $x \in \mathcal{H}$, $P_M(x) = \sum_{n=1}^{\infty} (x, e_n)e_n$ as a convergent series in $\mathcal{H}$.*

(ii) *For every $x \in \mathcal{H}$, $\sum_{n=1}^{\infty} |(x, e_n)|^2 \leq \|x\|^2$ (Bessel's Inequality)*

(iii) *If $M = \mathcal{H}$ and $x, y \in \mathcal{H}$, then*

$$(x, y) = \sum_{n=1}^{\infty} (x, e_n)\overline{(y, e_n)}.$$

This implies $\sum_{n=1}^{\infty} |(x, e_n)|^2 = \|x\|^2$.

Proof.

(i) The formula $P_M(x) = \sum_{n=1}^{N} (x, e_n)e_n$ is valid for any finite set $\{e_1, \ldots, e_N\}$ (by an exercise in a previous homework assignment). For any arbitrary countable family, let $x_n = \sum_{i=1}^{n} (x, e_i)e_i$, $x_n \in \mathcal{H}$ and $n \in \mathbb{N}$. We note $(x, e_i) = (x_n, e_i)$ for $i = 1, \ldots, n$ as

$$(x_n, e_i) = \left(\sum_{j=1}^{n} (x, e_j)e_j, e_i \right) = \sum_{j=1}^{n} (x, e_j)(e_j, e_i) = (x, e_i),$$

as our vectors are orthonormal. Then $(x - x_n, e_i) = 0$ for $i = 1, \dots, n$, with $x - x_n \perp M_n$. As $x = x_n + (x - x_n)$ where $x_n \in M_n$ and $x - x_n \in M_n^\perp$, then $\|x_n\| \le \|x\|$ for all $n \in \mathbb{N}$. We see

$$\|x_n\|^2 = \left(\sum (x, e_i) e_i, \sum (x, e_j) e_j \right) = \sum_{i,j=1}^{n} (x, e_i)\overline{(x, e_j)}(e_j, e_i)$$
$$= \sum_{i=1}^{n} (x, e_i)\overline{(x, e_i)} = \sum_{i=1}^{n} |(x, e_i)|^2 \le \|x\|^2.$$

As the above holds for all $n \in \mathbb{N}$, we have $\sum\limits_{n=1}^{\infty} (x, e_n) e_n$ is convergent.

(ii) Follows immediately by part (i). Consider (for $n < m$) the term

$$\|x_n - x_m\|^2 = \sum_{i=n+1}^{m} |(x, e_i)|^2.$$

We have $\|x_n - x_m\|^2 \to 0$ as $n, m \to \infty$ by convergence of the series- this means $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, and therefore has some limit y. Note that $y \in M$; it is also essentially immediate that as $x_n \to y$ in $\mathcal{H}$, then $y = \sum\limits_{i=1}^{\infty} (x, e_i) e_i$ as the limit of a sequence of partial sums.

(iii) Let $x, y \in \mathcal{H}$ and assume $M = \mathcal{H}$. Then

$$(x, y) = \lim_{n \to \infty} (x_n, y_n) = \lim_{n \to \infty} \left(\sum_{i=1}^{n} (x, e_i) e_i, \sum_{j=1}^{n} (y, e_j) e_j \right)$$
$$= \lim_{n \to \infty} \sum_{i=1}^{n} (x, e_i)\overline{(y, e_i)} = \sum_{i=1}^{\infty} (x, e_i)\overline{(y, e_i)}.$$

□

4.2. **Bounded linear operators on a Hilbert space. Recall:** We use $\mathcal{H}$ to denote a Hilbert space, and $\mathcal{B}(\mathcal{H})$ to denote the space of all bounded linear operators on $\mathcal{H}$. We note that $\mathcal{B}(\mathcal{H})$ is a Banach space, as $\mathcal{H}$ is a Banach space.

Note: If $\mathcal{H}$ is finite-dimensional (say $\mathcal{H} = \mathbb{C}^n$) then $\mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathbb{C}^n) \equiv M_n$. Here we have a natural inner product: $(S, T) = \mathrm{Tr}(T^* S)$. If $\mathcal{H}$ is infinite-dimensional, one can define a "natural" inner product only on a subspace of $\mathcal{B}(\mathcal{H})$: the space of Hilbert-Schmidt operators (denoted $S_2(\mathcal{H})$).

We have both a linear structure on $\mathcal{B}(\mathcal{H})$, and operator multiplication: if we take $S, T \in \mathcal{B}(\mathcal{H})$ then ST is modeled by composition. We note that $ST \ne TS$ in most cases. Our norm on the space is, of course, the operator norm. This turns $\mathcal{B}(\mathcal{H})$ into a Banach algebra.

Adjoint operation (involution):
For every $T \in \mathcal{B}(\mathcal{H})$, there exists $T^* \in \mathcal{B}(\mathcal{H})$ such that

$$(Tx, y) = (x, T^* y), \quad x, y \in \mathcal{H}.$$

The operator T^* is unique for this property, and is called the adjoint of T.

Proposition 4.20. *The adjoint satisfies the following properties:*

(i) $(T + S)^* = T^* + S^*$
(ii) $(\lambda T)^* = \overline{\lambda} T^*$
(iii) $(ST)^* = T^* S^*$

(iv) $(T^*)^* = T$
(v) $\|T^*\| = \|T\|$
(vi) $\|T^*T\| = \|T\|^2$,

for all $S, T \in \mathcal{B}(\mathcal{H})$ and $\lambda \in \mathbb{C}$. The last identity is called the C^*-property of the norm. These properties together imply $\mathcal{B}(\mathcal{H})$ is a C^*-algebra.

Proof.
(i)-(v) The proofs for these are straightforward, by definition of T^*.

(vi) For $x \in \mathcal{H}$, $(T^*Tx, x) = (Tx, Tx) = \|Tx\|^2$. If we take the supremum over all x with $\|x\| \leq 1$, we find

$$\|T\|^2 \leq \|T^*T\|\|x\| \leq \|T^*T\| \leq \|T^*\|\|T\| = \|T\|\|T\| = \|T\|^2$$

by Cauchy-Schwarz. Therefore, $\|T\|^2 = \|T^*T\|$. □

Definition 4.21. *Let $T \in \mathcal{B}(\mathcal{H})$. We say T is:*

(i) normal, if $T^*T = TT^*$.
(ii) self-adjoint, if $T = T^*$.
(iii) positive, if $(Tx, x) \geq 0$ for all $x \in \mathcal{H}$.
(iv) an isometry, if $\|Tx\| = \|x\|$ for all $x \in \mathcal{H}$. Equivalently, $T^*T = I$. To see why these are equivalent, we note

$$\|Tx\| = \|x\| \iff \|Tx\|^2 = \|x\|^2 \iff (Tx, Tx) = (x, x) \text{ for all } x \in \mathcal{H}$$
$$\iff (T^*Tx, x) = (x, x) \text{ for all } x \in \mathcal{H},$$

which implies $T^*T = I$ by polarization.
(v) unitary, if $T^*T = TT^* = I$.
(vi) compact, if $\overline{T(B_1(0)}$ is a compact subset of $\mathcal{H}$. Note that if $\text{rng}T$ is finite dimensional, then T is compact.

Remark: An operator $T \in \mathcal{B}(\mathcal{H})$ is normal if and only if $\|Tx\| = \|T^*x\|$ for all $x \in \mathcal{H}$.

Proof. We have

$$\|Tx\| = \|T^*x\| \iff \|Tx\|^2 = \|T^*x\|^2 \iff (Tx, Tx) = (T^*x, T^*x)$$
$$\iff (T^*Tx, x) = (TT^*x, x) \text{ for all } x \in \mathcal{H},$$

which by polarization means $TT^* = T^*T$. □

Remark: For operator $T \in \mathcal{B}(\mathcal{H})$, then $(\text{rng}T)^\perp = \ker T^*$ and $\overline{\text{rng}T} = (\ker T^*)^\perp$.

Proof. We have

$$y \in (\text{rng}T)^\perp \iff (z, y) = 0 \text{ for all } z \in \text{rng}T$$
$$\iff (Tx, y) = 0 \text{ for all } x \in \mathcal{H} \iff (x, T^*y) = 0 \text{ for all } x \in \mathcal{H}$$
$$\iff T^*y = 0 \iff y \in \ker T^*.$$

This shows $(\text{rng}T)^\perp = \ker T^*$. The second part of the statement above follows directly from the first. □

Example 4.22.

(i) The shift operator: take $\mathcal{H} = \ell^2$, and $S : \ell^2 \to \ell^2$ where

$$S(\lambda_1, \lambda_2, \ldots) = (0, \lambda_1, \lambda_2, \ldots).$$

It is clear that S is an isometry.
Question: *Is every isometry a shift?*
Answer: *More or less- see Wold's decomposition.*

(ii) *Diagonal operators: take $\mathcal{H} = \ell^2$. For every $a \in \ell^\infty$, we define $D_a : \ell^2 \to \ell^2$ such that*

$$D_a((\lambda_i)_{i=1}^\infty) = (a_i\lambda_i)_{i=1}^\infty.$$

We have the following properties for diagonal operators:
- $D_a \in \mathcal{B}(\ell^2)$*, and* $\|Da\| = \|a\|_\infty$.
- $D_a + D_b = D_{a+b}$ *(similarly for scalar product* λD_a*).*
- $D_a D_b = D_{ab} = D_{ba}$.
- $D_a^* = D_{\overline{a}}$*, where* $\overline{a} = (\overline{a_i})_{i=1}^\infty$*. From this, it is clear that any* D_a *is normal, as* ℓ^∞ *is commutative.*
- D_a *is compact if and only if* $a \in c_0$.

(iii) *Let $[0,1]$ be equipped with the Lebesgue measure λ. Take $\mathcal{H} = L^2(0,1)$, with inner product*

$$(f,g) = \int_0^1 f(x)\overline{g(x)}dx.$$

For every $\phi \in L^\infty(0,1)$, we define an operator on $\mathcal{H}$:

$$T_\phi(f) = \phi f, \quad f \in L^2(0,1).$$

Here $(\phi f)(x) = \phi(x)f(x)$, for $x \in [0,1]$.
The operator T_ϕ is bounded on $L^2(0,1)$:

$$\|T_\phi f\|^2 = \|\phi f\|^2 = \int_0^1 |\phi(x)f(x)|^2 dx \le C^2 \int_0^1 |f(x)|^2 dx$$
$$= C^2\|f\|,$$

as ϕ is bounded with $\|\phi\|_\infty \le C$. Hence, $\|T_\phi\| \le C^2$. In fact, $\|T_\phi\| = \|\phi\|_\infty$ (which we will not show). We also see
- $T_\phi + T_\psi = T_{\phi+\psi}$.
- $T_\phi T_\psi = T_{\phi\psi} = T_{\psi\phi}$.
- $T_\phi^* = T_{\overline{\phi}}$*- so* T_ϕ *is normal.*

We can think of T_ϕ as the "diagonal form".
Note: *$\phi \mapsto T_\phi$ as a mapping from $L^\infty(0,1) \to \mathcal{B}(\mathcal{H})$ is a bounded homomorphism.*

4.3. **Spectral theory of operators.**

Definition 4.23. *An operator $T \in \mathcal{B}(\mathcal{H})$ is invertible if there exists an $S \in \mathcal{B}(\mathcal{H})$ such that $TS = ST = I$.*

Proposition 4.24. *Assume $T \in \mathcal{B}(\mathcal{H})$, with $\|T\| < 1$. Then $I - T$ is an invertible operator.*

Proof. Consider the series $I + T + T^2 + \cdots = \sum_{k=0}^\infty T^k$ (with operator terms). It is absolutely convergent: as $\|T\| < 1$,

$$\sum_{k=0}^\infty \|T^k\| \le \sum_{k=0}^\infty \|T\|^k < \infty.$$

This means the series $\sum_{k=0}^{\infty} T^k$ is convergent as well, as $\mathcal{B}(\mathcal{H})$ is a Banach space. Let $S = \sum_{k=0}^{\infty} T^k \in \mathcal{B}(\mathcal{H})$. Now,

$$S(I-T) = \lim_{n\to\infty}(\sum_{k=0}^{n} T^k)(I-T) = \lim_{n\to\infty}(I + T + T^2 + \cdots + T^n - T - T^2 - \cdots - T^{n+1})$$
$$= \lim_{n\to\infty} I - T^{n+1}.$$

However, as $\|T\| < 1$, then $\lim_{n\to\infty} T^{n+1} = 0$. Thus, $S(I-T) = I$. A similar argument is made to show $(I-T)S = I$, which means $I - T$ is invertible. □

Note: By the proposition, if T is "close" to I then T is invertible. Here we mean "close" as in close in norm, from the correct side.

Proposition 4.25. *Let $T \in \mathcal{B}(\mathcal{H})$ be invertible. If $S \in \mathcal{B}(\mathcal{H})$ such that $\|S - T\| < \|T^{-1}\|^{-1}$, then S is invertible.*

Proof. Consider the operator $I - T^{-1}(T - S)$. As

$$\|T^{-1}(T-S)\| \le \|T^{-1}\|\|S - T\| < 1,$$

the operator $I - T^{-1}(T-S)$ is invertible with inverse $\sum_{k=0}^{\infty}[T^{-1}(T-S)]^k$, by the previous proposition. Additionally, we note

$$(I - T^{-1}(T-S))^{-1} = (I - T^{-1}T - T^{-1}S)^{-1} = (I - I - T^{-1}S)^{-1}$$
$$= (-T^{-1}S)^{-1}.$$

This means $T^{-1}S$ is invertible, which therefore implies S is invertible. □

Note: Since $(T^{-1}S)^{-1} = \sum_{k=0}^{\infty}[T^{-1}(T-S)]^k = S^{-1}T$, then

$$S^{-1} = \sum_{k=0}^{\infty}[T^{-1}(T-S)]^k T^{-1}.$$

This also allows us to find estimates for norms on operators like $S^{-1} - T^{-1}$ (for example), as

$$S^{-1} - T^{-1} = \sum_{k=1}^{\infty}(T^{-1}(T-S))^k T^{-1}.$$

Hence,

$$\|S^{-1} - T^{-1}\| \le \|T^{-1}\| \sum_{k=1}^{\infty} \|T^{-1}(T-S)\|^k \le \frac{\|T^{-1}\|^2\|T-S\|}{1 - \|T^{-1}(T-S)\|}.$$

Remarks: From the previous propositions, it follows that

(i) The set of all invertible elements is open.
(ii) The operation $T \mapsto T^{-1}$ is continuous (on the set of all invertible operators).

Definition 4.26. *Let $T \in \mathcal{B}(\mathcal{H})$. The spectrum $\sigma(T)$ is the set of all complex scalars $\lambda \in \mathbb{C}$ such that $\lambda I - T$ is not invertible. The complement $\mathbb{C} \setminus \sigma(T)$ is called the resolvent set of T.*

Note: When $\mathcal{H}$ is finite dimensional, $\sigma(T)$ is just the set of eigenvalues for the operator T.

Theorem 4.27. *For any $T \in \mathcal{B}(\mathcal{H})$, the set $\sigma(T)$ is a non-empty compact subset of $\mathbb{C}$.*

First, some initial observations. Suppose $\lambda \notin \sigma(T)$; then $\lambda I - T$ is invertible. Write $\lambda I - T = \lambda(I - (T/\lambda))$, if $\lambda \neq 0$. So $(\lambda I - T)^{-1} = \frac{1}{\lambda}(I - (T/\lambda))^{-1}$. If we also assume $\|T/\lambda\| < 1$, then $|\lambda| > \|T\|$. Then the proposition above implies that invertible $\lambda I - T$ has inverse

$$(\lambda I - T)^{-1} = \frac{1}{\lambda}\sum_{k=0}^{\infty}\left(\frac{T}{\lambda}\right)^k = \sum_{k=0}^{\infty}\frac{T^k}{\lambda^{k+1}}.$$

Proof. If $|\lambda| > \|T\|$, then $\|T/\lambda\| < 1$ and hence $\lambda I - T$ is invertible (see the previous observation). From this, we may infer that $\sigma(T)$ is bounded, as $\mathbb{C}\setminus\sigma(T) \subseteq \{\lambda \in \mathbb{C} : |\lambda| > \|T\|\}$ (specifically, $\sigma(T) \subseteq \{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}$). We also note that $\mathbb{C}\setminus\sigma(T)$ is open: suppose that $\lambda I - T$ is invertible. If $\mu \in \mathbb{C}$ such that $|\mu - \lambda| < \|(\lambda I - T)^{-1}\|^{-1}$, then $\mu I - T$ is invertible (by a previous proposition) with $\mu \notin \sigma(T)$. This means

$$N_{\|(\lambda I - T)^{-1}\|^{-1}}(\lambda) \subseteq \mathbb{C}\setminus\sigma(T),$$

and hence $\mathbb{C}\setminus\sigma(T)$ is open. This means $\sigma(T)$ is closed, and therefore compact (as a closed and bounded subset of $\mathbb{C}$).

Now, assume towards contradiction that $\sigma(T) = \emptyset$. Then $\mathbb{C}\setminus\sigma(T) = \mathbb{C}$. The function defined by $\lambda \mapsto (\lambda I - T)^{-1}$ is therefore defined on all of $\mathbb{C}$; we note this is the same function as

$$\lambda \mapsto \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{T^k}{\lambda^k}.$$

Take $\phi \in \mathcal{B}(\mathcal{H})^*$, and consider the function defined by

$$\lambda \mapsto \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{\phi(T^k)}{\lambda^k}.$$

This is an analytic function on $\mathbb{C}$- hence, it is entire. Furthermore, as $\lambda \to \infty$ the function tends to 0, and therefore must be bounded. By Liouville's Theorem, it must be constant-thus, it is zero everywhere. This means $\phi((\lambda I - T)^{-1}) = 0$ for all $\phi \in \mathcal{B}(\mathcal{H})^*$. However, by the Hahn-Banach Theorem, this implies $(\lambda I - T)^{-1} = 0$, which is impossible (as the zero operator is clearly not invertible). As we have reached a contradiction, we conclude that $\sigma(T) \neq \emptyset$, which completes the proof. ∎

Definition 4.28. *An operator $T \in \mathcal{B}(\mathcal{H})$ is called compact if the set $\{Tx : \|x\| \leq 1\}$ is relatively compact, i.e. if $\overline{T(B_1(0))}$ is compact.*

Definition 4.29. *For $T \in \mathcal{B}(\mathcal{H})$, the rank of T is defined by*

$$\mathrm{rank}(T) = \dim T\mathcal{H} = \dim(\mathrm{rng}(T)).$$

If $\mathrm{rank}(T) < \infty$, *then T is called a finite rank operator.*

Definition 4.30. *For $\xi, \eta \in \mathcal{H}$ we define $\xi\eta^*$ to be the operator in $\mathcal{B}(\mathcal{H})$ given by*

$$\xi\eta^*(x) = (x, \eta)\xi.$$

Note: The range of $\xi\eta^*$ is just $\mathbb{C}\xi$. Hence, $\xi\eta^*$ is finite rank with rank 1.

Conversely, suppose that $T \in \mathcal{B}(\mathcal{H})$ has rank 1- i.e., $\mathrm{rng}(T) = \mathbb{C}\xi$ for some $\xi \neq 0$ with $\|\xi\| = 1$. Define $f : \mathcal{H} \to \mathbb{C}$ by $f(x) = (Tx, \xi)$ for all $x \in \mathcal{H}$. As T is bounded, so is f; furthermore, f is clearly linear by properties of the inner product (and as T is linear). Therefore, $f \in \mathcal{H}^*$. By the Riesz-Representation Theorem, there exists a $\eta \in \mathcal{H}$ such that $f(x) = (x, \eta)$ for each $x \in \mathcal{H}$. Thus, $Tx = (x, \eta)\xi$ for all $x \in \mathcal{H}$. This shows $T = \xi\eta^*$.

Starting with $\xi, \eta \in \mathcal{H}$, consider $\xi' = \theta\xi$, and $\eta' = \theta\eta$ where $\theta \in \mathbb{C}$ with $|\theta| = 1$. Then $\xi\eta^* = \xi'\eta'^*$, as

$$\xi'\eta'^* = \theta\xi\overline{\theta}\eta^* = \theta\overline{\theta}\xi\eta^* = |\theta|^2\xi\eta^* = \xi\eta^*.$$

This shows our decomposition of T as a rank-one operator is not unique.

Remark: If $\text{rank}(T) < \infty$, then T is a compact operator. This follows, as

$$\{Tx : \|x\| \le 1\} \subseteq \{y \in \text{rng}(T) : \|y\| \le \|T\|\}.$$

As the latter set is bounded and lies inside a finite dimensional (and hence closed) space, we know that $\overline{T(B_1(0))}$ is closed and bounded- therefore, compact.

Notation: We let $\mathcal{K}(\mathcal{H})$ be the set of all compact operators on $\mathcal{H}$.

Definition 4.31. *We say a set $\{y_1, \ldots, y_n\}$ such that for every $x \in K$ there exists an i such that $d(x, y_i) \le \epsilon$ is a finite ϵ-net for K. This means*

$$K \subseteq \bigcup_{i=1}^{n} B_\epsilon(y_i),$$

where the latter are closed balls.

Note:

(i) This implies if for some $\epsilon > 0$, K does not have a finite ϵ-net then there exists a sequence $x_1, x_2, \ldots$ such that $d(x_i, x_j) \ge \epsilon$ for all $i \ne j$.

(ii) A subset K of a metric space (X, d) is pre-compact (i.e. $\overline{K}$ is compact) if for every $\epsilon > 0$, K has a finite ϵ-net.

Up until now, most likely the only experience the reader has had with operators was with operators on finite-dimensional Hilbert spaces. As our space is finite-dimensional, all operators are compact- so there was no meaningful distinction between compactness and non-compactness for an operator. The following provide a few examples of non-compact operators once we turn our attention to infinite-dimensional Hilbert spaces.

Example 4.32.

(i) If $\dim \mathcal{H} = \infty$, take $T = I$. Then as $T(B_1(0)) = B_1(0)$, and $B_1(0)$ is not compact (as $\mathcal{H}$ is not finite-dimensional) then T is a bounded but non-compact operator.

(ii) Diagonal operators: take $\mathcal{H} = \ell^2$. For every $a \in \ell^\infty$, we define $D_a : \ell^2 \to \ell^2$ by

$$D_a(\xi) = (a_i\xi_i)_{i\in\mathbb{N}}.$$

Then D_a is a compact operator if and only if $a \in c_0$. Hence, D_a is bounded but non-compact if $a \in \ell^\infty \setminus c_0$.

As an analogy, when $\mathcal{H}$ is infinite-dimensional and separable we can consider $\mathcal{B}(\mathcal{H})$ as a non-commutative version of ℓ^∞, and $\mathcal{K}(\mathcal{H})$ as a non-commutative version of c_0.

Theorem 4.33. *The space $\mathcal{K}(\mathcal{H})$ is a closed ideal in $\mathcal{B}(\mathcal{H})$.*

Proof. First, let $T \in \mathcal{K}(\mathcal{H})$ and take an arbitrary $\lambda \in \mathbb{C}$; it is trivial that $\lambda T \in \mathcal{K}(\mathcal{H})$. So suppose we have $S \in \mathcal{K}(\mathcal{H})$ as well, and look at $S + T$. We see

$$\begin{aligned}\{(S+T)x : \|x\| \le 1\} &= \{Sx + Tx : \|x\| \le 1\} \\ &\subseteq \{Sx : \|x\| \le 1\} + \{Ty : \|y\| \le 1\}.\end{aligned}$$

We note that the sum of two pre-compact sets is pre-compact: we can show this using the two finite ϵ-nets. As $\{(S+T)x : \|x\| \le 1\}$ is contained in a pre-compact set, it is also pre-compact. Therefore, $S + T \in \mathcal{K}(\mathcal{H})$, which shows $\mathcal{K}(\mathcal{H})$ is a linear subspace.

Now, take $T \in \mathcal{K}(\mathcal{H})$ and suppose $B \in \mathcal{B}(\mathcal{H})$. We know as $T(B_1(0))$ is pre-compact, then $B(T(B_1(0)))$ is pre-compact as well as B is a bounded (hence continuous) operator. This implies $BT \in \mathcal{K}(\mathcal{H})$. Considering TB, we note that

$$\{TBx : \ \|x\| \leq 1\} \subseteq \{Ty : \ \|y\| \leq \|B\|\}.$$

As the set on the right is pre-compact, so is the set $\{TBx : \ \|x\| \leq 1\}$. Hence, $TB \in \mathcal{K}(\mathcal{H})$, which shows $\mathcal{K}(\mathcal{H})$ is an ideal in $\mathcal{B}(\mathcal{H})$.

Finally, we show the closure of $\mathcal{K}(\mathcal{H})$. Suppose $(T_n)_{n=1}^{\infty} \subseteq \mathcal{K}(\mathcal{H})$, and $T_n \to T$ as $n \to \infty$ in norm. Fix $\epsilon > 0$, and let $N \in \mathbb{N}$ such that $\|T_N - T\| < \epsilon$. Let $\{y_1, \ldots, y_n\}$ be an ϵ-net for $T_N(B_1(0))$. Then $\{y_1, \ldots, y_n\}$ is a 2ϵ-net for $T(B_1(0))$: to see why, we note that for any $x \in \mathcal{H}$ where $\|x\| \leq 1$, there exists y_j such that $\|T_N x - y_j\| \leq \epsilon$. Then

$$\begin{aligned}\|Tx - y_j\| &\leq \|Tx - T_N x\| + \|T_N x - y_j\| \leq \|T - T_N\|\|x\| + \epsilon \\ &\leq \epsilon + \epsilon = 2\epsilon.\end{aligned}$$

This shows $T \in \mathcal{K}(\mathcal{H})$, and so $\mathcal{K}(\mathcal{H})$ is closed. □

Definition 4.34. *Pointwise convergence of operators is called strong convergence.*

Note:

(i) If $\|T_n - T\| \to 0$ as $n \to \infty$, then $T_n \to T$ strongly (i.e. uniform convergence implies strong convergence).

(ii) Unless $\mathcal{H}$ is a finite-dimensional Hilbert space, compact operators are not invertible: if $K \in \mathcal{K}(\mathcal{H})$ is invertible, then $KK^{-1} = I \in \mathcal{K}(\mathcal{H})$, which is a contradiction when $\mathcal{H}$ is infinite-dimensional.

Proposition 4.35. *If $K \in \mathcal{K}(\mathcal{H})$ is compact and $(T_n)_{n=1}^{\infty} \subseteq \mathcal{B}(\mathcal{H})$ such that $T_n x - Tx \to 0$ as $n \to \infty$ for all $x \in \mathcal{H}$, then $\|T_n K - TK\| \to 0$ as $n \to \infty$.*

Proof. Suppose not; so there exists a $\delta > 0$ such that $\|T_{n_k} K - TK\| > \delta$, for all $k \in \mathbb{N}$. Let $x_{n_k} \in \mathcal{H}$ with $\|x_{n_k}\| = 1$ such that $\|T_{n_k} K x_{n_k} - TK x_{n_k}\| > \delta$, for all $k \in \mathbb{N}$. Consider the sequence $(Kx_{n_k})_{k \in \mathbb{N}}$. As K is compact, there exists a further subsequence $Kx_{n_{k_j}} \to y$. We see

$$\begin{aligned}\|T_{n_{k_j}} K x_{n_{k_j}} - Ty\| &\leq \|T_{n_{k_j}} K x_{n_{k_j}} - T_{n_{k_j}} y\| + \|T_{n_{k_j}} y - Ty\| \\ &\leq \|T_{n_{k_j}}\|\|Kx_{n_{k_j}} - y\| + \|T_{n_{k_j}} y - Ty\|,\end{aligned}$$

where the latter two converge to 0 by assumption and by application of the Uniform Boundedness Principle (for the $T_{n_{k_j}}$'s). Hence, $\|T_{n_{k_j}} K x_{n_{k_j}} - TK x_{n_{k_j}}\| \to 0$. However, this contradicts $\|T_{n_k} K - TK\| > \delta$ for all $k \in \mathbb{N}$. Thus, we must have uniform convergence. □

Proposition 4.36. *Let $T = T^* \in \mathcal{B}(\mathcal{H})$, λ, μ be eigenvalues with $\mu \neq \lambda$, and x, y be their corresponding eigenvectors. Then $x \perp y$.*

Proof. We know that $Tx = \lambda x$, and $Ty = \mu y$. We then see

$$\mu(x, y) = \overline{\mu}(x, y) = (x, \mu y) = (x, Ty) = (Tx, y) = (\lambda x, y) = \lambda(x, y).$$

This means $(\mu - \lambda)(x, y) = 0$, which implies $(x, y) = 0$ as $\mu \neq \lambda$. □

Note: For $T \in \mathcal{B}(\mathcal{H})$,

$$\|T\| = \sup\{|(Tx, y)| : \ \|x\| \leq 1, \|y\| \leq 1\}$$

as

$$\begin{aligned}\|T\| &= \sup\{\|Tx\| : \ \|x\| \leq 1\} = \sup\{|f(Tx)| : \ \|x\| \leq 1, f \in \mathcal{H}^*, \|f\| \leq 1\} \\ &= \sup\{|(Tx, y)| : \ \|x\| \leq 1, \|y\| \leq 1\}\end{aligned}$$

by the Hahn-Banach Theorem and the Riesz Representation Theorem.

Proposition 4.37. *Let $T = T^* \in \mathcal{B}(\mathcal{H})$. Then*

$$\|T\| = \sup\{|(Tx,x)| : \|x\| \le 1\}.$$

Proof. Write $m = \sup\{|(Tx,x)| : \|x\| \le 1\}$. We note that $m \ge 0$ (with $m > 0$ if $T \ne 0$). We want to show that $\|T\| = m$. By the previous note, we already have $m \le \|T\|$.

For $x, y \in \mathcal{H}$, we note that

$$(T(x+y), x+y) = (Tx,x) + (Ty,y) - 2\Re(Tx,y),$$
$$(T(x-y), x-y) = (Tx,x) + (Ty,y) - 2\Re(Tx,y).$$

Subtracting, this shows

$$4\Re(Tx,y) = (T(x+y),x+y) - (T(x-y),x-y) \le m\|x+y\|^2 + m\|x-y\|^2 = m(\|x+y\|^2 + \|x-y\|^2)$$
$$\le m(2\|x\|^2 + 2\|y\|^2).$$

Thus, $2\Re(Tx,y) \le m(\|x\|^2+\|y\|^2)$. Let $\theta \in \mathbb{C}$, with $\|\theta\| = 1$ such that $\theta(Tx,y) = |(Tx,y)|$. Replacing x with θx, we find

$$2|(Tx,y)| \le m(\|x\|^2 + \|y\|^2).$$

If we then set $y = \frac{\|x\|(Tx)}{\|Tx\|}$, we have

$$2|(Tx, \frac{\|x\|}{\|Tx\|}Tx)| \le m\left(\|x\|^2 + \left\|\frac{\|x\|}{\|Tx\|}Tx\right\|^2\right)$$
$$\Rightarrow 2\frac{\|x\|}{\|Tx\|}\|Tx\|^2 \le m(\|x\|^2 + \|y\|^2) \Rightarrow 2\|x\|\|Tx\| \le 2m\|x\|^2$$
$$\Rightarrow \|Tx\| \le m\|x\|,$$

for all $x \in \mathcal{H}$. However, by definition this just means $\|T\| \le m$. Thus, $\|T\| = m$. □

Note: If $A \in M_n$ and $A \ge 0$, then $\|A\|$ is just the largest eigenvalue of A. Thinking in similar terms, the (following) proposition then is an infinite-dimensional formulation of this fact. We first provide the following lemma.

Lemma 4.38. *Let $T = T^* \in \mathcal{B}(\mathcal{H})$, $\lambda = \inf\{(Tx,x) : \|x\| = 1\}$ and $\mu = \sup\{(Tx,x) : \|x\| = 1\}$. If λ is attained at a vector x_1, then x_1 is an eigenvector and λ is an eigenvalue for T. A similar statement holds for μ.*

Proof. Suppose λ is attained at $x_1 \in \mathcal{H}$- so $\|x_1\| = 1$ and $(Tx_1, x_1) = \lambda$. The definition of λ implies $(Ty,y) \ge \lambda(y,y)$ for all $y \in \mathcal{H}$. Consider $y = x_1 + \alpha z$, where $z \in \mathcal{H}$, with $\|z\| = 1$ and $\alpha \in \mathbb{C}$. Then

$$(T(x_1+\alpha z), (x_1+\alpha z)) \ge \lambda(x_1+\alpha z, x_1+\alpha z)$$
$$\Rightarrow (Tx_1,x_1) + 2\Re\overline{\alpha}(Tx_1,z) + |\alpha|^2(Tz,z) \ge \lambda(\|x_1\|^2 + 2\Re\overline{\alpha}(x_1,z) + |\alpha|^2\|z\|^2).$$

As $(Tx_1,x_1) = \lambda$ and $\|x_1\| = 1$, the latter inequality reduces to

$$2\Re\overline{\alpha}(Tx_1,z) + |\alpha|^2(Tz,z) \ge 2\lambda\Re\overline{\alpha}(x_1,z) + \lambda|\alpha|^2\|z\|^2.$$

Now, let $\alpha = r\overline{(z,(T-\lambda I)x_1)}$ where $r \in \mathbb{R}$. Then

$$2r\Re(z,(T-\lambda I)x_1)(Tx_1,z) + r^2|(z,(T-\lambda I)x_1)|^2(Tz,z)$$
$$\ge 2\lambda r\Re(z,(T-\lambda I)x_1)(x_1,z) + \lambda r^2|(z,(T-\lambda I)x_1)|^2\|z\|^2$$
$$= 2r\Re(z,(T-\lambda I)x_1)((T-\lambda I)x_1,z)$$
$$+r^2|(z,(T-\lambda I)x_1)|^2(Tz,z) - \lambda r^2|(z,(T-\lambda I)x_1)|^2\|z\|^2 \ge 0.$$

This implies

$$2r|(z,(T-\lambda I)x_1)|^2 + r^2|(z,(T-\lambda I)x_1)|^2(Tz,z) - \lambda r^2|(z,(T-\lambda I)x_1)|^2\|z\|^2 \ge 0.$$

Set $s = |(z, (T - \lambda I)x_1)|^2$. The inequality above reduces to

$$2rs + r^2 s(Tz, z) - \lambda r^2 s\|z\|^2 \geq 0,$$

for all $z \in \mathcal{H}, r \in \mathbb{R}$. When $\|z\| = 1$, we get

$$s(2r + r^2(Tz, z) - \lambda r^2) \geq 0.$$

If $s = 0$ for all such z, then $(T - \lambda I)x_1 = 0$ and we are done. So suppose not- that is, suppose there exists some z such that $s \neq 0$. We then have

$$r(2 + r((Tz, z) - \lambda)) = r(2 + rt) \geq 0.$$

If $r > 0$, the statement above holds. However, if $r < 0$, we necessarily have $r < -2/t$. If we specifically choose r which is negative but between $-2/t$ and 0, we derive a contradiction (as the statement should hold for *all* $r \in \mathbb{R}$). Thus, $s = 0$, and so x_1 is an eigenvector. □

Proposition 4.39. *Let $T \in \mathcal{B}(\mathcal{H})$, with $T = T^*$ and T compact. Then either $\|T\|$ or $-\|T\|$ is an eigenvalue.*

Proof. Recall that $(Tx, x) \in \mathbb{R}$ for all $x \in \mathcal{H}$. Furthermore, recall the quantities

$$\lambda = \sup\{(Tx, x) : \ x \in \mathcal{H}, \ \|x\| = 1\},$$
$$\mu = \inf\{(Tx, x) : \ x \in \mathcal{H}, \ \|x\| = 1\}$$

and that by a previous proposition we have

$$\|T\| = \sup\{|(Tx, x)| : \ \|x\| = 1\}.$$

Then $\|T\| = \lambda$ or $\|T\| = -\mu$. Without loss of generality, assume that $\|T\| = \lambda$ and let $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{H}$ with $\|x_n\| = 1$ for all $n \in \mathbb{N}$ such that $(Tx_n, x_n) \to \lambda$ as $n \to \infty$. We observe:

$$\|Tx_n - \lambda x_n\|^2 = (Tx_n, Tx_n) - \lambda(Tx_n, x_n) - \lambda(x_n, Tx_n) + \lambda^2(x_n, x_n)$$
$$= \|Tx_n\|^2 - 2\lambda(Tx_n, x_n) + \lambda^2 \leq \|T\|^2 - 2\lambda(Tx_n, x_n) + \lambda^2 = 2\lambda^2 - 2\lambda(Tx_n, x_n) \geq 0,$$

and $2\lambda^2 - 2\lambda(Tx_n, x_n) \to 0$ as $n \to \infty$. Therefore, as $\|Tx_n - \lambda x_n\|^2 \leq 2\lambda^2 - 2\lambda(Tx_n, x_n)$ for each $n \in \mathbb{N}$, we see $\|Tx_n - \lambda x_n\| \to 0$ as $n \to \infty$. As T is compact, there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that $Tx_{n_k} \to y$ as $k \to \infty$. Furthermore, by convergence in norm of $(Tx_n)_{n\in\mathbb{N}}$ we have $\lambda x_{n_k} \to y$ as $k \to \infty$. As the case when $T = 0$ is trivial, we may assume $\|T\| > 0$, and hence $\lambda > 0$. Then $x_{n_k} \to (1/\lambda)y$ as $k \to \infty$. Applying T to the convergence relation, we have $Tx_{n_k} \to (1/\lambda)Ty$. However, $Tx_{n_k} \to y$ as $k \to \infty$ as well. By uniqueness of limits, this implies $y = (1/\lambda)Ty$, and so $Ty = \lambda y$. This shows $y \neq 0$ is an eigenvector with eigenvalue λ. □

Definition 4.40.

(i) A closed subspace K of a Hilbert space $\mathcal{H}$ is called an invariant subspace for an operator T if $TK \subseteq K$.
(ii) K is called a reducing subspace for T if $K, K^\perp$ are both invariant under T.

Example 4.41. *Consider $S : \ell^2 \to \ell^2$ given by*

$$S((x_n)_{n=1}^{\infty}) = (0, x_1, x_2, x_3, \ldots).$$

For every $k \in \mathbb{N}$, the subspace $K_k = \{(0, 0, \ldots, 0, x_{k+1}, \ldots)\}$ is invariant for S.

Beurling's Theorem: *Characterizes the invariant subspaces of S (using complex analysis). The theorem can be used to show that $\ell^2 \cong L^2(\mathbb{T})$, where $\mathbb{T}$ is the torus.*

Question: Does every bounded operator on a Hilbert space possess a closed proper invariant subspace?

Answer: This is not known- it is still an open question.

Note: The question was resolved for Banach spaces by C. Read in the mid-to-late 1980's.

Note: (On reducing spaces for an operator T) Let K be reducing; we know that $\mathcal{H} = K \oplus K^\perp$. Any operator on $\mathcal{H}$ can be written in the form

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}, \quad \text{with regards to } \mathcal{H} = K \oplus K^\perp,\ h = x + y,\ x \in K, y \in K^\perp.$$

The matrix for T acts on h as

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} T_{11}x + T_{12}y \\ T_{21}x + T_{22}y \end{bmatrix}$$
$$\Rightarrow T \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} T_{11}x \\ T_{21}x \end{bmatrix}, \quad T \begin{bmatrix} 0 \\ y \end{bmatrix} = \begin{bmatrix} T_{12}y \\ T_{22}y \end{bmatrix}.$$

When K is reducing for T, the operator matrix of T is diagonal. Thus,

$$T = \begin{bmatrix} T_{11} & 0 \\ 0 & T_{22} \end{bmatrix}, \quad T \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} T_{11}x \\ 0 \end{bmatrix} \in K, \quad T \begin{bmatrix} 0 \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ T_{22}y \end{bmatrix} \in K^\perp.$$

We can even write out the individual components for a clearer picture:

$$T_{11} : K \to K, \quad T_{12} : K^\perp \to K,$$
$$T_{21} : K \to K^\perp, \quad T_{22} : K^\perp \to K^\perp.$$

Notes:

(i) If x is an eigenvector for T, then $\mathbb{C}x = \text{span}\{x\}$ is invariant for T.

Proof. Let λ be the eigenvalue. Then for all $c \in \mathbb{C}$,

$$T(cx) = cTx = c\lambda x = (c\lambda)x.$$

□

(ii) A subspace K is invariant for T if and only if $K^\perp$ is invariant for T^*.

Proof. Suppose $TK \subseteq K$. Take $y \in K^\perp$, $x \in K$. Then $(T^*y, x) = (y, Tx) = 0$ as $Tx \in K, y \in K^\perp$. This implies $T^*y \perp K$- so $T^*y \in K^\perp$. As $y \in K^\perp$ was arbitrary, we have $K^\perp$ is invariant for T^*. Reversing the steps completes the proof. □

(iii) If $T = T^* \in \mathcal{B}(\mathcal{H})$, and x is an eigenvector for T then $\mathbb{C}x$ is a reducing subspace for T.

Proof. We have that $\mathbb{C}x$ is invariant for T by (ii). Then $(\mathbb{C}x)^\perp$ is invariant for T^*, and thus T. □

In fact, if $T = T^*$ and $K \subseteq \mathcal{H}$ is invariant for T, then K is reducing for T.

Theorem 4.42 (Spectral Theorem for Compact Self-Adjoint Operators)**.** *Let T be a compact self-adjoint operator on a Hilbert space $\mathcal{H}$. Then there exist scalars $\lambda_1, \lambda_2, \dots \in \mathbb{R}$ (possibly finitely many) such that*

$$|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \cdots$$

and there exist $e_n \in \mathcal{H}$ with $\|e_n\| = 1$ such that

(i) $Tx = \sum_{n=1}^{\infty} \lambda_n (x, e_n) e_n$ for all $x \in \mathcal{H}$. (Intuitively, we give T a diagonal form- think finite-dimensional case.)

(ii) We have $\lambda_n \to 0$ as $n \to \infty$ provided the sequence is infinite, and $e_n \perp e_m$ if $n \neq m$.

Proof. By the lemma, either $\|T\|$ or $-\|T\|$ is an eigenvalue- say λ_1. Let e_1 be a unit eigenvector for λ_1, and let $M_1 := \mathbb{C}e_1$. We know that M_1 is an invariant subspace for T, and hence is reducing as T is self-adjoint. Consider $T_1 := T|_{M_1^\perp}$. As $M_1^\perp$ is invariant, this restriction is well-defined. Continuing inductively, we construct $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$ as follows: if $T_1, \ldots, T_{n-1}$ and $\lambda_1, \ldots, \lambda_{n-1}$ are constructed, we let $M_n := \text{span}\{e_1, \ldots, e_n\}$ with $T_n := T|_{M_n^\perp}$. Then let λ_n be the eigenvalue ensured by the same lemma, and e_n to be an eigenvector for λ_n of norm 1. We claim that if the sequence of λ_n's does not terminate, it must converge to 0.

Suppose not: let $\delta > 0$ and choose a subsequence $(\lambda_{n_k})_{k\in\mathbb{N}}$ such that $|\lambda_{n_k}| \geq \delta$ for all $k \in \mathbb{N}$. Then $Te_{n_k} = \lambda_{n_k} e_{n_k}$. So,

$$\|Te_{n_k} - Te_{n_m}\|^2 = \|\lambda_{n_k} e_{n_k} - \lambda_{n_m} e_{n_m}\|^2 = |\lambda_{n_k}|^2 + |\lambda_{n_m}|^2 \geq 2\delta^2$$

for all $k, m \in \mathbb{N}$. However, this contradicts the compactness of T. Therefore, if the sequence is infinite, it must converge to 0. This shows that there are at most a countably infinite number of λ_n's.

We now break the problem into two cases to finish:

<u>Case I:</u> If the sequence terminates, then $T_n = 0$ for some $n \in \mathbb{N}$. Then for $x \in \mathcal{H}$, $x - \sum_{i=1}^{n} (x, e_i)e_i \in M_n^\perp$. Applying T_n, we have

$$T_n\left(x - \sum_{i=1}^{n}(x, e_i)e_i\right) = 0$$

$$\Rightarrow Tx - \sum_{i=1}^{n}(x, e_i)\lambda_i e_i = 0 \Rightarrow Tx = \sum_{i=1}^{n} \lambda_i (x, e_i) e_i.$$

Thus, the formula for T has finitely many terms.

<u>Case II:</u> If $\lim_{n\to\infty} \lambda_n = 0$, we must have $T_n \neq 0$ for all $n \in \mathbb{N}$. For all $x \in \mathcal{H}$, if we pick any $n \in \mathbb{N}$ we again have the decomposition in $M_n^\perp$ as noted above. Then

$$\left\|Tx - \sum_{i=1}^{n} \lambda_i (x, e_i) e_i\right\| \leq \|T_n\|\|x\| = |\lambda_n|\|x\|.$$

As $\lambda_n \to 0$, then $\|T_n\|\|x\| \to 0$ as well. Thus, $\|Tx - \sum_{i=1}^{n} \lambda_i(x, e_i)e_i\| \to 0$ as $n \to \infty$. From this, we have

$$Tx = \sum_{i=1}^{\infty} \lambda_i (x, e_i) e_i$$

for all $x \in \mathcal{H}$. □

Start with compact and self-adjoint operator T, and take eigenvectors $\{e_n\}$ guaranteed by the Spectral Theorem. Extend the set to an orthonormal basis of $\mathcal{H}$- say B. With respect to B, every $S \in \mathcal{B}(\mathcal{H})$ has a certain matrix representation

$$S = \begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & & \ddots \end{bmatrix}, \quad a_{ij} = (Se_j, e_i).$$

The matrix of the compact operator T will be

$$T = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \end{bmatrix}.$$

Note that $\lambda_n \to 0$ as $n \to \infty$ if there are infinitely many λ_i's.

Note: It is possible to define an extension of the determinant to infinite-dimensional matrices for certain subspaces of $\mathcal{B}(\mathcal{H})$. The Fuglede-Kadison determinant is defined for certain von Neumann algebras.

Lemma 4.43.

(i) If P, Q are projections then $PQ = 0 \iff \text{rng}P \perp \text{rng}Q$.
(ii) Let $PQ = 0$, where P, Q are projections; furthermore, let $S, T \in \mathcal{B}(\mathcal{H})$ such that $S = PSP, T = QTQ$ (i.e. $\text{supp}(S) \subseteq P\mathcal{H}, \text{supp}(T) \subseteq Q\mathcal{H}$). Then $\|S+T\| = \max\{\|S\|, \|T\|\}$.

Note: Suppose $S = PSP$. Then with respect to the decomposition $\mathcal{H} = P\mathcal{H} \oplus P\mathcal{H}^\perp$, the matrix of S has the form

$$\begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad S_1 : P\mathcal{H} \to P\mathcal{H}.$$

This also implies the matrix of $S+T$ has the form

$$\begin{bmatrix} S_1 & 0 \\ 0 & T_1 \end{bmatrix}, \quad S_1 : P\mathcal{H} \to P\mathcal{H},\ T_1 : Q\mathcal{H} \to Q\mathcal{H}.$$

The meaning of (ii) in the above then becomes clearer to see: $\|S+T\| = \max\{\|S_1\|, \|T_1\|\}$, which specifically holds for the scalar diagonal matrices.

Proof. (i) We have

$$PQ = 0 \iff (PQ\xi, \eta) = 0 \text{ for all } \xi, \eta \in \mathcal{H} \iff (Q\xi, P^*\eta) = 0 \text{ for all } \xi, \eta \in \mathcal{H}$$
$$\iff (Q\xi, P\eta) = 0 \text{ for all } \xi, \eta \in \mathcal{H} \iff \text{rng}P \perp \text{rng}Q.$$

(ii) Assume $PQ = 0, S = PSP$, and $T = QTQ$. Take any $\xi \in \mathcal{H}$, where $\xi = x + y + z$ for $x \in \text{rng}P, y \in \text{rng}Q$, and $z \in (\text{rng}P + \text{rng}Q)^\perp$. We note that we potentially could have $(\text{rng}P + \text{rng}Q)^\perp = \{0\}$. For our choice of ξ, we have

$$(S+T)\xi = PSP\xi + QTQ\xi = PSx + QTy.$$

Then

$$\|(S+T)\xi\|^2 = \|PSx + QTy\|^2 = \|PSx\|^2 + \|QTy\|^2 \leq \|S\|^2\|x\|^2 + \|T\|^2\|y\|^2$$
$$\leq \max\{\|S\|^2, \|T\|^2\}(\|x\|^2 + \|y\|^2) \leq \max\{\|S\|^2, \|T\|^2\}(\|x\|^2 + \|y\|^2 + \|z\|^2) = \max\{\|S\|^2, \|T\|^2\}\|\xi\|^2.$$

This implies $\|(S+T)\xi\| \leq \max\{\|S\|, \|T\|\}\|\xi\|$ for all $\xi \in \mathcal{H}$. Therefore, $\|S+T\| \leq \max\{\|S\|, \|T\|\}$.

Now, consider $\xi = x$, for $x \in \text{rng}P$ with $\|x\| = 1$. Then

$$\|S+T\| \geq \|(S+T)x\| = \|PSPx + QTQx\| = \|PSPx\| = \|Sx\|.$$

Taking the supremum over all $x \in \text{rng}P$ with $\|x\| = 1$, we obtain $\|S\| \leq \|S+T\|$. Through symmetry, we get $\|T\| \leq \|S+T\|$ as well. Thus, $\|S+T\| = \max\{\|S\|, \|T\|\}$. □

Note: It is possible to have bounded operators P with $P^2 = P$ while $P^* \neq P$. This means one can have a direct sum decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{N}$ where $\mathcal{M} \cap \mathcal{N} = \{0\}$, but $\mathcal{M} \not\perp \mathcal{N}$. This can happen even in finite dimensional spaces- for example, let $\mathcal{H} = \mathbb{C}^2, \mathcal{M} = \mathbb{C}e_1$, and $\mathcal{N} = \mathbb{C}(e_1 + e_2)$.

Theorem 4.44 (Spectral Theorem- Strong Form)**.** *Let $T \in \mathcal{B}(\mathcal{H})$, $T^* = T$, and T be compact. Then there exist projections $P_1, P_2, \ldots$ (possibly finitely many) such that $P_iP_j = 0$ for $i \neq j$, and scalars $\mu_1, \mu_2, \ldots$ such that $\mu_i \neq \mu_j$ if $i \neq j$, $|\mu_1| \geq |\mu_2| \geq \cdots$, $\mu_n \to 0$ as $n \to \infty$ and*

$$T = \sum_{n=1}^{\infty} \mu_n P_n$$

as a convergent series in the norm of $\mathcal{B}(\mathcal{H})$.

Proof. By the first form of the Spectral Theorem, we have $\{e_n\}_{n\in\mathbb{N}}$ and $\lambda_1, \lambda_2, \ldots$ such that

$$Tx = \sum_{n=1}^{\infty} \lambda_n (x, e_n) e_n$$

where $\|e_n\| = 1$ and $Te_n = \lambda_n e_n$ for all $n \in \mathbb{N}$. Furthermore, we have $\lambda_n \to 0$ as $n \to \infty$ when we have infinitely many.

For every element of the set $\{\lambda_i : i \in \mathbb{N}\}$- say μ_n- let P_n be the projection onto $\text{span}\{e_i : \lambda_i = \mu_n\}$. As T is compact, the set $\{e_i : \lambda_i = \mu_n\}$ is finite for all μ_n. To see why, consider $Te_i = \mu_n e_i$ and through arguing by contradiction make a similar argument as shown in Homework 4.

We claim the series $\sum\limits_{n=1}^{\infty} \mu_n P_n$ is convergent in the norm of $\mathcal{B}(\mathcal{H})$. The difference of partial sums of the series are of the form

$$\mu_n P_n + \mu_{n+1} P_{n+1} + \cdots + \mu_m P_m$$

where $n < m$. By induction and application of the previous lemma, we have

$$\|\mu_n P_n + \cdots + \mu_m P_m\| = \max\{\|\mu_n P_n\|, \ldots, \|\mu_m P_m\|\} = \max\{|\mu_n|, \ldots, |\mu_m|\}$$

as P_n is a projection for each $n \in \mathbb{N}$. As $\mu_n \to 0$ necessarily (as $\lambda_i \to 0$ when $n \to \infty$), this implies $\max\{|\mu_n|, \ldots, |\mu_m|\} \to 0$. Therefore, the sequence of differences of partial sums is Cauchy in $\mathcal{B}(\mathcal{H})$, and thus the series is convergent in $\mathcal{B}(\mathcal{H})$.

Let $S = \sum\limits_{n=1}^{\infty} \mu_n P_n$; we also claim that $S = T$. For any $x \in \mathcal{H}$, we have

$$Sx = \sum_{n=1}^{\infty} \mu_n P_n x = \sum_{n=1}^{\infty} \mu_n \left(\sum_{\lambda_i = \mu_n} (x, e_i) e_i \right) = \sum_{i=1}^{\infty} \lambda_i (x, e_i) e_i = Tx.$$

As $x \in \mathcal{H}$ was arbitrary, this shows $S = T$. □

Note: If $P_i P_j = 0$ for $i \neq j$, $\text{rank} P_i < \infty$ for all i, and $\mu_i \to 0$ as $i \to \infty$ for scalars $(\mu_i)_{i\in\mathbb{N}}$ then the operator $\sum\limits_{i=1}^{\infty} \mu_i P_i$ is compact. Indeed, the partial sums all have finite rank, as they are of the form $\sum\limits_{i=1}^{n} \mu_i P_i$. They are necessarily compact, and as $\mathcal{K}(\mathcal{H})$ is a closed ideal so must be their limit in $\mathcal{B}(\mathcal{H})$.

Recall: An operator $T \in \mathcal{B}(\mathcal{H})$ is called positive if $(T\xi, \xi) \geq 0$ for all $\xi \in \mathcal{H}$.

Theorem 4.45. *The following are equivalent for an operator $T \in \mathcal{B}(\mathcal{H})$:*

(i) T is positive.
(ii) There exists $A \in \mathcal{B}(\mathcal{H})$ such that $T = A^*A$.
(iii) There exists a positive operator $B \in \mathcal{B}(\mathcal{H})$ such that $T = B^2$.

Notes:

(i) The operator B in (iii) is unique, and is called the square root.
(ii) Some directions are easy: for $(iii) \Rightarrow (ii)$, take $A = B$. Then $A^* = B$, with B self-adjoint. For $(ii) \Rightarrow (i)$, we have

$$(Tx, x) = (A^*Ax, x) = (Ax, Ax) \geq 0$$

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

Proof. We shall prove this at a later date. □

Question: Suppose $T = T^*$ is compact. How is the positivity of T reflected in its spectral decomposition?

Answer: Write T in the form given by the Spectral Theorem:

$$T = \sum_{n=1}^{\infty} \lambda_n P_n.$$

Then

$$T \geq 0 \iff \lambda_n \geq 0 \text{ for all } n \in \mathbb{N}.$$

Proof. Suppose $\lambda_1 < 0$. Let $x \in \text{rng}P_1$, with $\|x\| = 1$; then

$$(Tx, x) = \left(\sum_{n=1}^{\infty} \lambda_n P_n x, x \right) = \sum_{n=1}^{\infty} \lambda_n (P_n x, x) = \lambda_1 (P_1 x, x) = \lambda(x, x) = \lambda_1 < 0.$$

However, this contradicts the positivity of T- so $\lambda_n \geq 0$ for all $n \in \mathbb{N}$.

Conversely, suppose $\lambda_n \geq 0$ for all $n \in \mathbb{N}$. For $x \in \mathcal{H}$, we see

$$(Tx, x) = \left(\sum_{n=1}^{\infty} \lambda_n P_n x, x \right) = \sum_{n=1}^{\infty} \lambda_n (P_n x, x) = \sum_{n=1}^{\infty} \lambda_n (P_n x, P_n x) = \sum_{n=1}^{\infty} \lambda_n \|P_n x\|^2 \geq 0,$$

as $P_n^* = P_n = P_n^2$. □

Square root of a positive compact operator:

Suppose that $T = T^*$ is compact, and $T \geq 0$. Write $T = \sum_{n=1}^{\infty} \lambda_n P_n$, where $\lambda_n \geq 0$. Let $S := \sum_{n=1}^{\infty} \sqrt{\lambda_n} P_n$; we note that as $\lambda_n \to 0$ when $n \to \infty$, then $\sqrt{\lambda_n} \to 0$ as well. Thus, S is indeed well-defined. Moreover,

$$\begin{aligned} S^2 = \left(\sum_{n=1}^{\infty} \sqrt{\lambda_n} P_n \right) \left(\sum_{m=1}^{\infty} \sqrt{\lambda_m} P_m \right) &= \sum_{n=1}^{\infty} \sqrt{\lambda_n} P_n \left(\sum_{m=1}^{\infty} \sqrt{\lambda_m} P_m \right) \\ &= \sum_{n=1}^{\infty} \sqrt{\lambda_n}^2 P_n^2 = \sum_{n=1}^{\infty} \lambda_n P_n = T. \end{aligned}$$

Also, S is a positive operator, as $\sqrt{\lambda_n} \geq 0$ for all $n \in \mathbb{N}$. This shows S is the square root of T, as desired.

Theorem 4.46 (Fredholm Alternative)**.** *Let T be a compact operator on a Hilbert space $\mathcal{H}$. Let $\lambda \neq 0$, with $\lambda \in \mathbb{C}$. Then precisely one of the following holds:*

(i) There exists a non-zero $x \in \mathcal{H}$ such that $Tx = \lambda x$- i.e., x is an eigenvector of T with eigenvalue λ.

(ii) $T - \lambda I$ is invertible.

Proof. Let $\lambda \neq 0$ with $\lambda \in \mathbb{C}$, and assume that λ is not an eigenvalue of T. We will show that $T - \lambda I$ is invertible.

Step 1: We claim that $T - \lambda I$ is "bounded below", in that there exists a $c > 0$ such that $\|(T - \lambda I)x\| \geq c\|x\|$ for all $x \in \mathcal{H}$. To see this, we first note that through scaling without loss of generality we may assume $\|x\| = 1$. If we suppose that $(T - \lambda I)$ is not bounded below, this means there exists a subsequence $(x_n)_{n \in \mathbb{N}}$ with $\|x_n\| = 1$ such that $\|(T - \lambda I)x_n\| \to 0$ as $n \to \infty$. We have $\|\lambda x_n\| = |\lambda|$ for all $n \in \mathbb{N}$, and hence the vectors λx_n are bounded away from 0. As T is compact, the sequence $(Tx_n)_{n \in \mathbb{N}}$ must have a convergent sequence- say $(Tx_{n_k})_{k \in \mathbb{N}}$, where $Tx_{n_k} \to y$ for some $y \in \mathcal{H}$ as $k \to \infty$. Since $\|Tx_n - \lambda x_n\| \to 0$, this means our $Tx_{n_k} \to \lambda x_{n_k}$- so $y \neq 0$, as $\lambda x_{n_k} \not\to 0$.

As $Tx_{n_k} \to y$, then (by the boundedness and hence continuity of $(T - \lambda I)$) we have $(T - \lambda I)Tx_{n_k} \to (T - \lambda I)y$. We also have

$$(T - \lambda I)Tx_{n_k} = T(T - \lambda I)x_{n_k} = T(Tx_{n_k} - \lambda x_{n_k}) \to 0$$

as $k \to \infty$. Hence,

$$(T - \lambda I)y = 0 \Rightarrow Ty - \lambda y = 0 \Rightarrow Ty = \lambda y.$$

As $y \neq 0$, this means y is an eigenvector with eigenvalue λ; however, this directly contradicts our choice of λ. This shows $(T - \lambda I)$ must be bounded below.

Step 2: We claim $\text{rng}(T - \lambda I)$ is closed. More generally: every operator S which is bounded below has a closed range.

To show the above, suppose that there exist $x_n \in \mathcal{H}$ for $n \in \mathbb{N}$ such that $Sx_n \to y$ as $n \to \infty$. We wish to show that $y \in \text{rng}S$. As $(Sx_n)_{n=1}^{\infty}$ is a convergent sequence and S is bounded below, we see

$$c\|x_n - x_m\| \leq \|S(x_n - x_m)\| = \|Sx_n - Sx_m\| \to 0$$

as $n, m \to \infty$. As $c > 0$ was constant, this implies the sequence $(x_n)_{n=1}^{\infty}$ is Cauchy. As we are in a Hilbert space, this sequence must converge to some $x \in \mathcal{H}$; as S is bounded, it is continuous, and hence

$$Sx_n \to Sx, \quad n \to \infty.$$

However, as $Sx_n \to y$ as well, this shows $y = Sx$; thus, $y \in \text{rng}S$. This shows the range of any bounded below operator must be closed. As we specifically have shown that $T - \lambda I$ is bounded below (by Step 1), this shows $\text{rng}(T - \lambda I)$ is closed.

Step 3: We claim $(T - \lambda I)$ is injective. As $(T - \lambda I)$ is bounded below, if $x \in \ker(T - \lambda I)$ then

$$\|(T - \lambda I)x\| = 0 \geq c\|x\|.$$

This immediately implies $x = 0$, as $c > 0$ and $\|x\| \geq 0$ for any $x \in \mathcal{H}$. Hence, $T - \lambda I$ has trivial kernel.

Step 4: Finally, we claim that $T - \lambda I$ is surjective. Suppose not: let $\mathcal{H}_1 = \text{rng}(T - \lambda I)$. By Step 2, $\mathcal{H}_1$ is a closed proper subspace of $\mathcal{H}$. Then, let $\mathcal{H}_2 = (T - \lambda I)\mathcal{H}_1$; we should have $\mathcal{H}_2 \subset \mathcal{H}_1$ strictly (if not, we could apply the inverse of $(T - \lambda I)|_{\mathcal{H}_1}$, which is guaranteed by the Open Mapping Theorem and show that $\mathcal{H}_1 = \mathcal{H}$, which is a contradiction). Continue inductively, forming a nested sequence of spaces

$$\mathcal{H}_1 \supset \mathcal{H}_2 \supset \mathcal{H}_3 \supset \cdots$$

Let $y_n \in \mathcal{H}_n$ such that $d(y_n, \mathcal{H}_{n+1}) \geq 1/2$, with $\|y_n\| = 1$. Consider the sequence $(Ty_n)_{n=1}^{\infty}$; as T is compact, it must have a convergent subsequence. However, for $n < m$ we have

$$Ty_n - Ty_m = (T - \lambda I)y_n - (T - \lambda I)y_m + \lambda y_n - \lambda y_m,$$

which is in $\lambda y_n + \mathcal{H}_{n+1}$. As $d(y_n, \mathcal{H}_{n+1}) \geq 1/2$, we have $\|Ty_n - Ty_m\| \geq |\lambda|/2 > 0$ always. Therefore, T could not be compact- a contradiction. This shows $(T - \lambda I)$ is surjective, and hence invertible.

□

Notes:

(i) The above is equivalent to saying that the non-zero elements of the spectrum $\sigma(T)$ are necessarily eigenvalues.

(ii) The only assumption we make on T in the above is that it is compact- it does not have to be self-adjoint.

Multiplication Operators
Recall that if $[a,b] \subseteq \mathbb{R}$ and $\phi : [a,b] \to \mathbb{R}$ is bounded and measurable, then the multiplication operator M_ϕ is the operator acting on the Hilbert space $L^2[a,b]$ defined by

$$M_\phi f(x) = \phi(x) f(x), \quad x \in \mathcal{H},\ f \in L^2[a,b].$$

Question: Given a continuous function $f : \mathbb{R} \to \mathbb{C}$, can I define/talk about the operator $f(M_\phi)$?

Example 4.47. *If f is a polynomial where*

$$f(s) = a_n s^n + \cdots + a_1 s + a_0,$$

then

$$f(A) = a_n A^n + \cdots + a_1 A + a_0 I.$$

Answer: Yes: $f(M_\phi) = M_{f\circ\phi}$, where we assume ϕ is real-valued.

Properties: For the functions f, g:
(i) $(fg)(M_\phi) = f(M_\phi) g(M_\phi)$.
(ii) $(f+g)(M_\phi) = f(M_\phi) + g(M_\phi)$.
(iii) $\overline{f}(M_\phi) = f(M_\phi)^*$.

By the properties above, we have a map

$$C(\mathbb{R}) \to \mathcal{B}(\mathcal{H}),$$
$$f \mapsto f(M_\phi).$$

In general, we have a homomorphism $f \mapsto f(A)$ between algebras.

One may ask- why are we considering this at all? While we may state the results above for multiplication operators, this by no means exhausts all possible operators in $\mathcal{B}(\mathcal{H})$. What we ultimately wish to develop is a similar result, but for any self-adjoint operator. While up until now we have spent a lot of time considering compact self-adjoint operators, compactness is ultimately not necessary; however, self-adjointness is absolutely necessary. Specifically, we wish to develop what is known as the *functional calculus* for self-adjoint operators.

Lemma 4.48. *Let $T = T^*$ for $T \in \mathcal{B}(\mathcal{H})$ and $c > 0$ such that $\|Tx\| \geq c\|x\|$, for all $x \in \mathcal{H}$. Then T is invertible.*

Note: If T is invertible, there automatically exists a constant $c > 0$ such that $\|Tx\| \geq c\|x\|$ for all $x \in \mathcal{H}$. Indeed, let $S = T^{-1}$; then $x = STx$, with

$$\|x\| = \|STx\| \leq \|S\|\|Tx\| \Rightarrow c = 1/\|S\|.$$

Proof. Let $T = T^*$, and assume we have boundedness from below (as stated above). We first claim that T is injective: if $Tx = 0$, then as $\|Tx\| \geq c\|x\|$, $0 \geq c\|x\|$ with $c > 0$. This forces $\|x\| = 0$, and hence $x = 0$. Thus, $\ker T = \{0\}$, and so T must be injective.

We next claim that T is surjective. Suppose $z \perp \text{rng}T$. Then $(z, Tx) = 0$ for all $x \in \mathcal{H}$. As T is self-adjoint, this means $(Tz, x) = 0$ for all $x \in \mathcal{H}$- hence, $Tz = 0$. Then as T is injective, we have $z = 0$. This implies $\overline{\text{rng}T} = \mathcal{H}$. Then, using the same reasoning as in Claim 2 of our proof of the Fredholm Alternative, we have that $\text{rng}T$ is closed- hence, T is surjective.

Our next claim is that T is invertible. Indeed, as T is bijective (as we have just shown) and $\mathcal{H}$ is a Banach space, the Open Mapping Theorem directly tells us that T must be invertible. This completes the proof. $\square$

Lemma 4.49. *If $T = T^*$, then $\sigma(T) \subseteq \mathbb{R}$.*

Proof. We will first show that $i \notin \sigma(T)$. We see

$$\begin{aligned}\|Tx - ix\|^2 = (Tx - ix, Tx - ix) &= \|Tx\|^2 - i(x, Tx) + i(Tx, x) + \|x\|^2 \\ &= \|Tx\|^2 - i(Tx, x) + i(Tx, x) + \|x\|^2 \\ &= \|Tx\|^2 + \|x\|^2 \geq \|x\|^2.\end{aligned}$$

Thus, $\|(T - iI)x\| \geq \|x\|$ for all $x \in \mathcal{H}$; by the previous lemma, this tells us $T - iI$ is invertible. Thus, $i \notin \sigma(T)$. Note that a similar statement can be made for $T + iI$ as well- hence, $T + \beta iI$ for every $\beta \in \mathbb{R} \setminus \{0\}$.

Now, take any $z \notin \mathbb{R}$ with $z = \alpha + \beta i$ where $\alpha, \beta \in \mathbb{R}$. We see $T - zI = (T - \alpha I) - i\beta I$, where $T - \alpha I$ is self-adjoint. Applying the previous part of our proof to $T - \alpha I$ alone, we know that it must be an invertible operator- thus, $(T - \alpha I) - \beta iI$ is invertible as well. This means $z \notin \sigma(T)$ for any $z \in \mathbb{C}$- thus, $\sigma(T) \subseteq \mathbb{R}$. ◻

Definition 4.50. *For $T \in \mathcal{B}(\mathcal{H})$, its spectral radius is the number*

$$r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}.$$

Note: For any $T \in \mathcal{B}(\mathcal{H})$, $r(T) \leq \|T\|$. Indeed, if $|\lambda| > \|T\|$, then $T - \lambda I$ is invertible- hence, $\lambda \notin \sigma(T)$.

Theorem 4.51. *If $T = T^*$ for $T \in \mathcal{B}(\mathcal{H})$, then $\|T\| = r(T)$.*

Proof. By our note above, we automatically know that $r(T) \leq \|T\|$. We wish to show that $\|T\| \leq r(T)$.

Recall that $\alpha = \inf\{(Tx, x) : x \in \mathcal{H}, \|x\| = 1\}$, $\beta = \sup\{(Tx, x) : x \in \mathcal{H}, \|x\| = 1\}$. Assume without loss of generality that $\|T\| = \beta$ (we know this is possible by a previous proposition). We have $\sigma(T) \subseteq [\alpha, \beta]$; we claim $\beta \in \sigma(T)$.

To see why, let $x_n \in \mathcal{H}$ with $\|x_n\| = 1$ for each $n \in \mathbb{N}$ such that $(Tx_n, x_n) \to \beta$ as $n \to \infty$. We find

$$\begin{aligned}\|Tx_n - \beta x_n\|^2 = (Tx_n - \beta x_n, Tx_n - \beta x_n) &= \|Tx_n\|^2 - \beta(x_n, Tx_n) - \beta(Tx_n, x_n) + \beta^2\|x_n\|^2 \\ &= \|Tx_n\|^2 - 2\beta(Tx_n, x_n) + \beta^2 \\ &\leq \|T\|^2\|x_n\|^2 - 2\beta(Tx_n, x_n) + \beta^2 \\ &= 2\beta^2\|x_n\|^2 - 2\beta(Tx_n, x_n) \\ &= 2\beta^2 - 2\beta(Tx_n, x_n)\end{aligned}$$

As $(Tx_n, x_n) \to \beta$, the quantity $2\beta^2 - 2\beta(Tx_n, x_n)$ converges to 0 as $n \to \infty$. Thus, $\|Tx_n - \beta x_n\|^2 \to 0$ as $n \to \infty$, and so $(T - \beta I)x_n \to 0$. This implies $(T - \beta I)$ is not invertible: if it were, with inverse S, then $S(T - \beta I)x_n \to 0 \Rightarrow x_n \to 0$ as $n \to \infty$, which is impossible as $\|x_n\| = 1$ for all $n \in \mathbb{N}$. Therefore, $\beta \in \sigma(T)$. This means $\|T\| \leq r(T)$, and hence

$$r(T) = \|T\|.$$

◻

Theorem 4.52 (Spectral Mapping Theorem)**.** *Let p be a polynomial, and $T = T^*$ for $T \in \mathcal{B}(\mathcal{H})$. Then*

$$\sigma(p(T)) = p(\sigma(T)).$$

Proof. Factor the polynomial

$$p(z) - \lambda = \lambda_0(z - \lambda_1) \cdots (z - \lambda_n),$$

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

$$p(T) - \lambda I = \lambda_0(T - \lambda_1 I) \cdots (T - \lambda_n I).$$

Note that the terms in the factorization of $p(T) - \lambda I$ above are mutually commuting. By commutativity of the terms, if $T - \lambda_i I$ is invertible for all i where each has inverse S_i, then $(T - \lambda_1 I) \cdots (T - \lambda_n I)$ is invertible as well; specifically, it has inverse $S_1 \cdots S_n$. This implies that if $\lambda_i \notin \sigma(T)$ for each $i = 1, \ldots, n$, then $p(T) - \lambda I$ is invertible; so,

$$p(\sigma(T))^C \subseteq \sigma(p(T))^C \Rightarrow \sigma(p(T)) \subseteq p(\sigma(T)).$$

Conversely, suppose that $\lambda_i \in \sigma(T)$ for some i. Then $T - \lambda_i I$ is not invertible, and so is either not injective or not surjective (this ultimately is a result of the Open Mapping Theorem, as a bijective bounded map on $\mathcal{H}$ necessarily has a bounded inverse).

Case I: Suppose $T - \lambda_i I$ is not injective. As the factors $T - \lambda_i I$ are all mutually commuting, we may conclude that product $(T - \lambda_1 I) \cdots (T - \lambda_i I) \cdots (T - \lambda_n I)$ is not injective as well (just commute $(T - \lambda_i I)$ to the rightmost location in the product). Therefore, $p(T) - \lambda I$ is not invertible.

Case II: Suppose $T - \lambda_i I$ is not surjective. Again, by commuting our factors so that $T - \lambda_i I$ is at the leftmost position in the product, we know that $(T - \lambda_1 I) \cdots (T - \lambda_n I)$ is not surjective either. Thus, the operator is not invertible.

As this handles both cases, we have shown that

$$p(\sigma(T)) \subseteq \sigma(p(T)),$$

which completes the proof. □

Theorem 4.53. *If p is a polynomial and $T = T^*$, then*

$$\|p(T)\| = \max\{|p(\lambda)| : \lambda \in \sigma(T)\}.$$

Note: This generalizes the fact that $\|T\| = r(T)$ when T is self-adjoint. Indeed: consider polynomial $p(z) = z$.

Proof. Consider $q = p\overline{p}$, where $\overline{p}(z) = \overline{p(z)}$ is the complex conjugate of the polynomial $p(z)$. For any $z \in \mathbb{C}$, we have that $q(z) = |p(z)|^2 \geq 0$. As $q(z)$ is always positive, this implies that the coefficients of q must all be real. In addition, as T is a self-adjoint operator, when paired with the fact that the coefficients of q are real this implies $q(T)$ is a self-adjoint operator as well.

By applying the C^*-property, using the self-adjointness of $q(T)$, and the Spectral Mapping Theorem we see

$$\begin{aligned} \|p(T)\|^2 &= \|p(T)^* p(T)\| = \|\overline{p}p(T)\| = \|q(T)\| \\ &= r(q(T)) = \max\{|\mu| : \mu \in \sigma(q(T))\} \\ &= \max\{|\mu| : \mu \in q(\sigma(T))\} = \max\{|p(\lambda)|^2 : \lambda \in \sigma(T)\}. \end{aligned}$$

This completes the proof. □

Theorem 4.54 (Continuous Functional Calculus)**.** *Let $T \in \mathcal{B}(\mathcal{H})$ with $T = T^*$. Then there exists a map*

$$F : C(\sigma(T)) \to \mathcal{B}(\mathcal{H})$$

such that

(i) F is a homomorphism, i.e.:

$$F(f+g) = F(f) + F(g),$$
$$F(fg) = F(f)F(g).$$

Note that this automatically implies linearity of F.

(ii) $F(p) = p(T)$.

(iii) $\|F(f)\| = \|f\|_\infty = \sup\{|f(z)| : z \in \sigma(T)\}$.

Note: By part (iii), we have an isometric embedding

$$C(\sigma(T)) \hookrightarrow \mathcal{B}(\mathcal{H})$$

under F.

Proof. Let $\mathcal{P}(X)$ denote the normed space of complex polynomials over X, and consider the map

$$F : \mathcal{P}(\sigma(T)) \to \mathcal{B}(\mathcal{H}),$$
$$F(p) = p(T).$$

By the previous theorem, we know F is a well-defined isometry. Considering $\mathcal{P}(\sigma(T)) \subseteq C(\sigma(T))$ as a normed subspace of a Banach space, we can then extend F to an isometry (denoted in the same way)

$$F : \overline{\mathcal{P}(\sigma(T))} \to \mathcal{B}(\mathcal{H}).$$

We claim that $\overline{\mathcal{P}(\sigma(T))} = C(\sigma(T))$; that this holds is immediate from the Stone-Weierstrass theorem. Therefore, $F : C(\sigma(T)) \to \mathcal{B}(\mathcal{H})$ is an isometry, and so (iii) holds.

Using the fact that F is a homomorphism on $\mathcal{P}(\sigma(T))$, by extending to the closure this property is preserved (think of approximations). Thus, F as defined above is a homomorphism, and so (i) holds. □

Corollary 4.55. *Let $T \in \mathcal{B}(\mathcal{H})$, where T is positive. Then given $n \in \mathbb{N}$, there exists a positive operator $T^{1/n} \in \mathcal{B}(\mathcal{H})$ such that $(T^{1/n})^n = T$.*

Proof. Apply the Continuous Functional Calculus to the function

$$f(x) = \sqrt[n]{x}, \quad x \geq 0.$$

Then $f^n = \mathrm{id}$, and so $f(T)^n = T$. Under F, we get our positive operator $F(f) = T^{1/n}$. □

In particular, for any positive operator $T \geq 0$ we have a square root $T^{1/2}$ where

$$T^{1/2} \geq 0, \quad (T^{1/2})^2 = T.$$

It is often denoted by $\sqrt{T}$ (at the risk of abusing notation).

Another application of the Continuous Functional Calculus is to define the modulus of an operator. Let $T \in \mathcal{B}(\mathcal{H})$ be arbitrary. Then $T^*T \geq 0$, and so a square root for it exists. Define

$$|T| := (T^*T)^{1/2}.$$

Note the analogous formulation for the modulus of a complex number z, where $|z| = (\overline{z}z)^{1/2}$.

Furthermore, any $z \in \mathbb{C}$ has a polar decomposition $z = \omega|z|$, where $\omega \in \mathbb{C}$ with $|\omega| = 1$.

Unsurprisingly, this notion also extends to operators on a Hilbert space. We pause to introduce the following definition.

Definition 4.56. *Let $V : \mathcal{H}_1 \to \mathcal{H}_2$ be a bounded linear map. We say V is a partial isometry if there exist subspaces $\mathcal{V}_1 \subseteq \mathcal{H}_1$ and $\mathcal{V}_2 \subseteq \mathcal{H}_2$ such that for $h \in \mathcal{H}_1$ if we write $h = v_1 + v_1^\perp$ where $v_1 \in \mathcal{V}_1, v_1^\perp \in \mathcal{V}_1^\perp$ then $Vh = Vv_1 \in \mathcal{V}_2$ and*

$$\|Vh\| = \|Vv_1\| = \|v_1\|,$$
$$V(\mathcal{V}_1) = \mathcal{V}_2.$$

We call $\mathcal{V}_1$ the initial space of V, and $\mathcal{V}_2$ the final space.

Polar decomposition of operators:
Let $T \in \mathcal{B}(\mathcal{H})$. Then there exists a partial isometry V such that $T = V|T|$.

4.4. **Arbitrary measures, Radon measures, and more spectral theory.**

Definition 4.57. *Let X be a compact metric space. The Borel σ-algebra of X is the σ-algebra generated by the open sets of X.*

Recall: A σ-algebra is a collection $A \subseteq 2^X$ such that

(i) $\emptyset, X \in A$;
(ii) If $\alpha \in A$, so is α^C;
(iii) For $\alpha_i \in A$ with $i \in \mathbb{N}$, then $\cup_{i=1}^\infty \alpha_i \in A$.

Definition 4.58. *Let A be a σ-algebra. A measure μ is a function $\mu : A \to \mathbb{R}_+$ such that*

(i) $\mu(\emptyset) = 0$.
(ii) If $\alpha_1, \alpha_2, \ldots \in A$ where $\alpha_i \cap \alpha_j = \emptyset$ for $i \neq j$, then

$$\mu\Big(\bigcup_{i=1}^\infty \alpha_i\Big) = \sum_{i=1}^\infty \mu(\alpha_i).$$

To build up integration, we start (as expected) from characteristic functions. For $\alpha \in A$, let χ_α be the characteristic function for α. We define

$$\int \chi_\alpha d\mu = \mu(\alpha).$$

These then give rise to linear combinations (called simple functions) where

$$f = \sum_{i=1}^k \lambda_i \chi_{\alpha_i}, \quad \alpha_1, \ldots, \alpha_k \in A \text{ are mutually disjoint.}$$

Integration for a simple function is defined by

$$\int f d\mu = \sum_{i=1}^k \lambda_i \mu(\alpha_i).$$

This sum is indeed well-defined (which can be shown using the standard argument encountered during MATH602). Once we have established integration for simple functions, we can move onto positive functions: let $f \geq 0$, and f be measurable. The integral of f with respect to μ is defined by

$$\int f d\mu = \sup\Big\{ \int g d\mu : \ g \text{ simple}, \ g \leq f \Big\}.$$

From here, we decompose any measurable function into the sum of its two positive parts and extend integration in this way.

Example 4.59. *The following is an example of a type of measure known as a "point mass measure". Let $s \in X$ for some space X. Define the measure δ_s by*

$$\delta_s(\alpha) = \begin{cases} 1, & \text{if } s \in \alpha, \\ 0, & \text{if } s \notin \alpha. \end{cases}$$

For any integrable function f, we then have

$$\int f d\delta_s = f(s), \quad s \in X.$$

We pause our discussion of measures here to discuss a distinct (but related) topic.

Let X be a compact metric space, and recall that $C(X)$ is the space of all continuous functions $f : X \to \mathbb{C}$. We have seen that this is a Banach space under the $\|\cdot\|_\infty$ norm. The space has a natural cone of positive elements- namely, $C(X)^+$, the set of all non-negative functions in $C(X)$.

Definition 4.60. *A positive functional $\phi : C(X) \to \mathbb{C}$ is an element $\phi \in C(X)^*$ such that if $f \geq 0, \phi(f) \geq 0$.*

Example 4.61. *Let μ be a measure on the Borel σ-algebra $B(X)$ for some compact space X (we call such a measure a Borel measure). Then the functional $\phi : C(X) \to \mathbb{C}$ given by*

$$\phi(f) = \int f d\mu$$

is positive (provided $\mu(X) < \infty$).

Our discussion of measures and the introduction of positive linear functionals has a greater purpose: to develop the most general form of the Spectral Theorem for self-adjoint operators. Before we reach this point, we first include the following theorem (which is a quite stunning result in the opinion of the author).

Theorem 4.62 (Riesz-Markov-Kakutani)**.** *Let X be a compact metric space. Let $\phi : C(X) \to \mathbb{C}$ be a positive functional. Then there exists a (finite) Borel measure μ on X such that*

$$\phi(f) = \int f d\mu, \quad f \in C(X).$$

Moreover, the measure μ is unique as a Borel measure.

We omit a proof of the result above, as it requires a bit more theory than we have time to develop.

Recall: For a given operator $T \in \mathcal{B}(\mathcal{H})$ with $T = T^*$, we constructed a functional calculus

$$C(\sigma(T)) \to \mathcal{B}(\mathcal{H}),$$
$$f \mapsto f(T)$$

where the mapping was an algebraic homomorphism and an isometry; that is, we have

$$\|f(T)\| = \|f\|_\infty = \sup\{|f(s)| : s \in \sigma(T)\}.$$

The Riesz Representation Theorem will then be applied for the space $X = \sigma(T)$, where T is a fixed self-adjoint operator.

Definition 4.63. *A vector $\xi \in \mathcal{H}$ is called cyclic for T if the span of*

$$\{\xi, T\xi, T^2\xi, \dots\}$$

is dense in $\mathcal{H}$. The set above is called the orbit of T and ξ.

In the discussion that follows, we assume T is self-adjoint and has a cyclic vector $v \in \mathcal{H}$. When we drop this assumption, we will make it explicit. Again, assuming $X = \sigma(T)$ we consider the map

$$\phi : C(X) \to \mathbb{C},$$
$$f \mapsto (f(T)v, v), \quad f \in C(X).$$

We claim that ϕ is a positive functional on $C(X)$: if $f \geq 0$, then by the functional calculus $f(T) \geq 0$ as well- hence $(f(T)v, v) \geq 0$.

By the Riesz Representation Theorem, there exists a Borel measure on X such that

$$\phi(f) = \int f d\mu, \quad f \in C(X)$$

where

$$(f(T)v, v) = \int f d\mu.$$

Considering $\overline{f}f$ instead of f alone, we get

$$\int |f|^2 d\mu = (f(T)^* f(T)v, v) = \|f(T)v\|^2.$$

Now, consider the mapping $U : C(X) \to \mathcal{H}$ defined via

$$U(f) = f(T)v.$$

By the identity in norm above, we have

$$\|Uf\|^2 = \|f(T)v\|^2 = \int |f|^2 d\mu = \|f\|_2^2.$$

This implies $\|Uf\| = \|f\|_2$ for all $f \in C(X)$, and so U is an isometry from $(C(X), \|\cdot\|_2) \to \mathcal{H}$. Extending to the closure of $C(X)$ in $L^2(X, \mu)$, we obtain an isometry (which we also denote in the same way)

$$U : L^2(X, \mu) \to \mathcal{H}.$$

Since v is a cyclic vector, U is onto. Thus, we may talk about the inverse mapping $U^{-1} : \mathcal{H} \to L^2(X, \mu)$, which is also a surjective isometry.

Writing this out via diagram, we have

$$\begin{array}{ccc} \mathcal{H} & \xrightarrow{T} & \mathcal{H} \\ {\scriptstyle U^{-1}} \downarrow & & \uparrow {\scriptstyle U} \\ L^2(X, \mu) & \xrightarrow{M_x} & L^2(X, \mu) \end{array}$$

where $M_x : L^2(X, \mu) \to L^2(X, \mu)$ via

$$f \mapsto xf,$$

for $x \in X$. In other words, we have

$$UTU^{-1} = M_x,$$

so our self-adjoint operator T is unitarily equivalent to a multiplication operator (when T has a cyclic vector).

Now, relax the assumption of cyclicity- i.e., suppose T is self-adjoint, but not necessarily possessing a cyclic vector. The idea here to work around this issue is to decompose $\mathcal{H}$ as a sum $\mathcal{H} = \oplus_{i \in I} \mathcal{H}_i$ where $T|_{\mathcal{H}_i}$ has a cyclic vector for each $i \in I$.

Pick up any $v \in \mathcal{H}$, with $v \neq 0$. Consider

$$\mathcal{H}_1 := \overline{\text{span}\{T^n v\}_{n=1}^{\infty}}.$$

If $\mathcal{H}_1 = \mathcal{H}$, then v is cyclic and we are done. Else, the subspace $\mathcal{H}_1$ is invariant for T. As T is self-adjoint, this means $\mathcal{H}_1$ is reducing for T, and so we may write T in the form

$$T = \begin{bmatrix} T_1 & 0 \\ 0 & T' \end{bmatrix}.$$

We can then apply our argument once more, but for operator T'. Continue inductively- and then use Zorn's Lemma to ensure we have the desired composition of $\mathcal{H}$. We summarize the previous discussion with the following theorem (in its most general form).

Theorem 4.64 (Spectral Theorem I)**.** *If $T = T^*$, with $T \in \mathcal{B}(\mathcal{H})$ then there exists a compact space X, a Borel measure μ on X and an element $g \in L^\infty(X, \mu)$ such that up to unitary equivalence we have*

$$T = M_g,$$

hwere $M_g : L^2(X, \mu) \to L^2(X, \mu)$ is given by

$$M_g(f) = gf, \quad f \in L^2(X, \mu).$$

Theorem 4.65 (Spectral Theorem II)**.** *Let $T = T^*$ for $T \in \mathcal{B}(\mathcal{H})$. Then there is a spectral measure E on $\mathbb{R}$ such that*

$$T = \int_{\sigma(T)} t dE(t).$$

To help interpret the previous theorem, we introduce the following definition.

Definition 4.66. *A spectral measure E is a mapping $E : B(X) \to P(\mathcal{H})$ which is countably additive. (Here $P(\mathcal{H})$ denotes the space of projections in $\mathcal{H}$).*

Thinking about the integral in Spectral Theorem II, we have that T is a limit of operators of the form

$$\sum_{i=1}^{k} t_i E[t_i, t_{i+1}] = \sum_{i=1}^{k} t_i E_i,$$

where $E_i \in \mathcal{B}(\mathcal{H})$ and $\sum_{i=1}^{k} t_i E_i \to T$ in <u>norm</u>. This is similar (at least in spirit) to the compact operators, as limits of finite rank operators.

5. Addendum

5.1. **Addendum and errata.**

The following section includes a few additional examples or proofs for material left out which the author found to be helpful to include as a reference.

Proposition 5.1. *For a Hilbert space $\mathcal{H}$, if $V \subseteq \mathcal{H}$ is finite dimensional then V is closed.*

Proof. Suppose $V \neq 0$, and let $\dim V = m$. Choose an orthonormal basis $e_1, \ldots, e_m$ for V, and let $v \in \overline{V}$ with $(v_n)_{n\in\mathbb{N}} \subset V$ such that $v_n \to v$ as $n \to \infty$. By the Cauchy-Schwarz inequality, for $1 \leq k \leq m$

$$|(e_k, v) - (e_k, v_n)| = |(e_k, v - v_n)| \leq \|e_k\| \|v - v_n\| = \|v - v_n\| \to 0$$

as $n \to \infty$. Hence, $\lim_{n\to\infty} (e_k, v_n) = (e_k, v)$ for $1 \leq k \leq m$. From this, we see

$$v = \lim_{n\to\infty} v_n = \lim_{n\to\infty} \sum_{k=1}^{m} (e_k, v_n) e_k = \sum_{k=1}^{m} (e_k, v) e_k.$$

This shows $v \in V$. As $v \in \overline{V}$ was arbitrary, this shows V is closed. □

We include the following theorem as a reference only- we omit the proof.

Theorem 5.2. *If $T \in \mathcal{B}(X)$ for a complex Banach space X, then $r(T) = \lim_{n\to\infty} \sqrt[n]{\|T^n\|}$.*

Example 5.3. *The following is an example when strong convergence of compact operators in $\mathcal{B}(\mathcal{H})$ does not imply that the limit is compact. For each $n \in \mathbb{N}$, define $T_n : \ell^2 \to \ell^2$ via*

$$T_n((x_i)_{i\in\mathbb{N}}) = (x_1, \dots, x_n, 0, 0, \dots).$$

It is easy to show that T_n is linear and bounded for each $n \in \mathbb{N}$ (indeed- it mostly falls out directly by definition). As the rank of T_n is finite for each n, they are all compact operators. It is also easy to show that $T_n x \to x$ as $n \to \infty$ for $x \in \ell^2$. This means $T_n \to I$ in the strong operator topology. However, I cannot be compact, as the space ℓ^2 is infinite dimensional. Hence, strong operator convergence of compact operators does not imply that the limit is compact.

Proposition 5.4. *If X is an infinite-dimensional normed space, there is no compact surjective operator $T \in \mathcal{B}(X)$.*

Proof. Assume towards contradiction that T exists; as T is surjective and bounded, by the Open Mapping Theorem it must be an open map. Therefore, there exists a $c \in \mathbb{R}$ with $c > 0$ such that $cB_X \subseteq T(B_X)$. As T is a compact operator, then $\overline{T(B_X)}$ is a compact set. However, this would imply cB_X is compact, and hence B_X must be as well; as X is infinite-dimensional, this is impossible. Therefore, as we have reached a contradiction we conclude that such a T cannot exist. □