FAQI NOTES

VERN I. PAULSEN

ABSTRACT.

1. Topics to be covered

- Review of metric spaces
  - Definition of metric and examples
  - Convergence of sequences
  - $\epsilon - \delta$ definition of Continuity
  - Open, closed and compact sets
  - Sequential characterizations
  - Connected and pathwise connected sets
  - Equivalence and Uniform equivalence of metrics
  - Cauchy sequences and completeness
  - $C(\mathbb{R})$ as a metric space
  - Baire's theorem
- General topological spaces and nets
  - Open and closed sets, continuity
  - Nets and directed sets
  - Characterizations of closed sets, compact sets, and continuity in terms of nets
- Normed spaces
  - ℓ^p spaces, Holder and Minkowski inequalities, Riesz-Fischer
  - The space $C(X)$
  - Quotients and conditions for completeness, the 2/3's theorem
  - Finite dimensional normed spaces, equivalence of norms
  - Convexity, Absolute convexity, the bipolar theorem
  - Consequences of Baire's theorem:
    - Principle of Uniform boundedness, Resonance principle
    - Open mapping, closed graph and bounded inverse theorems
  - Hahn-Banach theorem
  - Krein-Milman theorem
  - Dual spaces and adjoints
  - The double dual
  - Weak topologies, weak convergences
- Hilbert spaces

Date: January 9, 2018.

Cauchy-Schwarz inequality
Polarization identity. Parallelogram Law
Jordan-vonNeumann theorem
Orthonormal bases and Parseval identities
Direct sums
Bilinear maps and Tensor products of Banach and Hilbert spaces
Infinite tensor products and quantum spin chains

- Operators on Hilbert spaces
  Special families of operators: adjoints, projections, Hermitian, unitaries, partial isometries, polar decomposition
  Density matrices and trace class operators
  B(H) as dual of trace class
- Spectral Theory
  Spectrum and resolvent
  Spectrum versus point spectrum; approximate points spectrum
  Neumann series
  $\sigma(T)$ is non-empty and compact
  Unbounded operators and self-adjointness problems
- The Riesz functional calculus
  The functional calculus for normal operators
- Compact operators
  Singular values and Schmidt's theorem
  Schmidt decomposition of tensors
  The Schatten classes
- Unbounded operators
  Hellinger-Toeplitz theorem
  Closable and non-closable operators
  Self-adjointness problems
  Stone's theorem
- Von Neumann algebras
  Measurement operators and observables
  The Double Commutant theorem
  Theory of types
  Connes' embedding conjecture and the Tsirelson problems
- C*-algebras
  Key examples
  GNS theorem
- Completely Positive and Completely Bounded Maps
  Matrix norm and matrix order
  Stinespring's theorem and Choi-Kraus
  Arveson's Hahn-Banach and Radon-Nikodym theorems
  Wittstock's extension and decomposition theorems

2. Metric Spaces

- Definition of metric and examples
- Convergence of sequences
- $\epsilon - \delta$ definition of Continuity
- Open, closed and compact sets
- Sequential characterizations
- Connected and pathwise connected sets
- Equivalence and Uniform equivalence of metrics
- Cauchy sequences and completeness
- $C(\mathbb{R})$ as a metric space
- Baire's theorem

Definition 2.1. Given a set X a **metric** on X is a function, $d : X \times X \to \mathbb{R}$ that satisfies:

(1) $d(x, y)ge0$,
(2) $d(x, y) = 0 \iff x = y$,
(3) $d(x, y) = d(y, x)$,
(4) $d(x, z) \leq d(x, y) + d(y, z)$.

The pair (X, d) is called a **metric space**. If in place of 2) we only have that $d(x, x) = 0$, then d is called a **semimetric**.

Some examples:

- $(\mathbb{R}, d)$ where $d(x, y) = |x - y|$.
- $(\mathbb{R}^n, d_p)$ where $d_p(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p\right)^{1/p}$, for any $1 \leq p < +\infty$. When $p = 2$ we call this the **Euclidean distance**.
- $(\mathbb{R}^n, d_\infty)$ where $d_\infty(x, y) = max\{|x_i - y_i| : 1 \leq i \leq n\}$.
- X any non-empty set, define $d(x, y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$. This is called the **discrete metric**.
- Given a metric space (X, d), and a subset $W \subseteq X$, (W, d) is also a metric space. This is called a **metric subspace of X**.

Definition 2.2. Let $(X, d), (Y, \rho)$ be metric spaces, let $f : X \to Y$ be a function, and let $x_0 \in X$. We say that f is **continuous at** x_0 provided that for every $\epsilon > 0$, there exists $\delta > 0$ such that $d(x, x_0) < \delta \implies \rho(f(x), f(x_0) < \epsilon$. We call f **continuous** if it is continuous at every $x_0 \in X$.

Consider the subspace $(-1, +1) \subseteq \mathbb{R}$ of $(\mathbb{R}, d)$. We have that $f : \mathbb{R} \to (-1, +1)$ given by $f(x) = \frac{x}{1+|x|}$ is one-to-one, onto and continuous. The inverse function $g(t) = \frac{t}{1-|t|}$ is also continuous.

Definition 2.3. Let (X, d) be a metric space. A subset $\mathcal{O}$ is **open** provided that whenever $x \in \mathcal{O}$ then there exists $\delta > 0$ such that the set

$$B(x_0; \delta) := \{x | d(x_0, x) < \delta\} \subseteq \mathcal{O}.$$

The set $B(x_0;\delta)$ is called the **ball centered at x_0 of radius δ**. A set is called **closed** if its complement is open.

These concepts also characterize continuity.

Proposition 2.4. *Let $(X,d),(Y,\rho)$ be metric spaces. Let $f:X\to Y$. TFAE:*

(1) *f is continuous,*
(2) *for every open set $U\subseteq Y$, the set $f^{-1}(U):=\{x: f(x)\in U\}$ is open,*
(3) *for every closed set $C\subseteq Y$, the set $f^{-1}(C)$ is closed.*

Convergence of sequences also gives good characterizations.

Definition 2.5. Let (X,d) be a metric space. A sequence $\{x_n\}_{n\in\mathbb{N}}$ **converges to x_0** provided that for every $\epsilon>0$ there is N such that $n>N \implies d(x_n,x_0)<\epsilon$. We write $\lim_n x_n = x_0$ or $x_n\to x_0$.

Proposition 2.6. *Let $(X,d),(Y,\rho)$ be metric spaces and let $f:X\to Y$. Then*

(1) *f is continuous at $x_0\in X$ iff for every sequence $\{x_n\}$ such that $\lim_n x_n = x_0$ we have that $\lim_n f(x_n) = f(x_0)$.*
(2) *$C\subseteq X$ is closed iff for every convergent sequence $\{x_n\}\subseteq C$ we have that $\lim_n x_n \in C$.*

2.1. Equivalent Metrics.

Definition 2.7. Let X be a set and let d and ρ be metrics on X. We say that d and ρ are **equivalent** provided that a set is open in the d metric iff it is open in the ρ metric. We say that d and ρ are **uniformly equivalent** provided that there are constants, $A,B>0$ such that $d(x,y)\le A\rho(x,y)$ and $\rho(x,y)\le Bd(x,y)$ for all $x,y\in X$.

It is easy to see that if two metrics are uniformly equivalent then they are equivalent. Also, two metrics are equivalent iff the function $id:X\to X$ is continuous from (X,d) to(X,ρ) and from (X,ρ) to (X,d).

A map $f:X\to Y$ between metric spaces such that f is one-to-one, onto with both f and f^{-1} continuous is called an **homeomorphism**. Thus, two metrics are equivalent iff the identity map is an isomorphism.

For $x=(x_1,...,x_n), y=(y_1,..,y_n)\in\mathbb{R}^n$ we have that

- $d_\infty(x,y)\le d_1(x,y)$ and $d_1(x,y)\le nd_{infty}(x,y)$,
- $d_{infty}(x,y)\le d_2(x,y)$ and $d_2(x,y)\le\sqrt{n}d_{infty}(x,y)$,
- $d_1(x,y)\le\sqrt{n}d_2(x,y)$(use Cauchy-Schwarz) and $d_2(x,y)\le d_1(x,y)$,

so that these are all uniformly equivalent metrics.

Problem 2.8. *Let (X,d) be a metric space. Define $\rho(x,y)=\frac{d(x,y}{1+d(x,y)}$. Prove that ρ is a metric on X and that d and ρ are equivalent metrics.*

2.2. **Cauchy Sequences and Completeness.**

Definition 2.9. Let (X,d) be a metric space. A sequence $\{x_n\}$ is **Cauchy** provided that for every $\epsilon > 0$ there exists N so that when $n, m > N$ then $d(x_n, x_m) < \epsilon$. A metric space is called **complete** provided that every Cauchy sequence converges to a point in X.

Definition 2.10. Let (Y, ρ) be a metric space. A subset $X \subseteq Y$ is called **dense** provided that for every $y \in Y$ and every $\epsilon > 0$ there is a point $x \in X$ with $\rho(y, x) < \epsilon$.

The canonical example is that the rational numbers $\mathbb{Q}$ together with $d(x, y) = |x - y|$ is a metric space that is not complete. They way that we "construct" $\mathbb{R}$ from $\mathbb{Q}$ by "adding" limits of Cauchy sequences, generalizes to any metric space.

Theorem 2.11. *Let (X, d) be a metric space. Then there is a metric space $(\hat{X}, \hat{d})$ so that*

(1) *$X \subseteq \hat{X}$,*
(2) *$\hat{d}(x, y) = d(x, y)$ for every pair $x, y \in X$,*
(3) *X is a dense subset of $\hat{X}$.*

Moreover, if $(\tilde{X}, \tilde{d})$ is another metric space satisfying !0,2),3), then there is an homeomorphism $h : \hat{X} \to \tilde{X}$ such that $h(x) = x$ for every $x \in X$ and $\tilde{d}(h(\hat{x}), h(\hat{y})) = \hat{d}(\hat{x}, \hat{y})$ for every pair $\hat{x}, \hat{y} \in \hat{X}$.

Definition 2.12. The (unique) metric space $(\hat{X}, \hat{d})$ given in the above theorem is called the **completion** of (X, d).

The following problem shows that the property of being complete is NOT invariant under equivalence of metric.

Problem 2.13. *On $\mathbb{R}$ define $d(x, y) = |x - y|$, $\rho(x, y) = \frac{d(x,y)}{1+d(x,y)}$ and $\gamma(x, y) = |\frac{x}{1+|x|} - \frac{y}{1+|y|}|$. Prove that*

(1) *γ is a metric on $\mathbb{R}$,*
(2) *d, ρ and γ are all equivalent metrics,*
(3) *$(\mathbb{R}, d)$ and $(\mathbb{R}, \rho)$ are complete metric spaces,*
(4) *$(\mathbb{R}, \gamma)$ is not complete.*

2.3. **Compact Sets.**

Definition 2.14. Let (X, d) be a metric space. Then a subset $K \subseteq X$ is called **compact** provided that whenever $\{U_a\}_{a \in A}$ is a collection of open sets such that $K \subseteq \cup_{a \in A} U_a$ then there is a finite subset $F \subseteq A$ such that $K \subseteq \cup_{a \in F} U_a$.

The following gives a nice characterization of this property.

Theorem 2.15. *Let (X, d) be a metric space. T.F.A.E.*

(1) *K is compact*

(2) *whenever* $\{x_n\}_{n\in\mathbb{N}} \subseteq K$ *is a sequence, then it has a subsequence* $\{x_{n_k}\}$ *that converges to a point in* K *(this is often called the* **Bolzano-Weierstrass property***),*
(3) (K,d) *is complete and for each epsilon* > 0 *there is a finite subset* $\{x_1,...,x_n\} \subseteq K$ *such that for each* $x \in K$ *there is an* i *with* $d(x,x_i) < \epsilon$ *(Such a subset is called an* ϵ**-net***.)*

One corollary of this is:

Corollary 2.16 (Heine-Borel)**.** *A subset* $K \subseteq \mathbb{R}^n$ *is compact iff it is closed and bounded.*

2.4. **Uniform Convergence.**

Proposition 2.17. *Let* (K,d) *be a non-empty compact metric space and let* $f : K \to \mathbb{R}$ *be continuous. Then there are points* $x_M, x_m \in K$ *such that* $f(X_M) = \sup\{f(x) : x \in K\}$, $f(x_m) = \inf\{f(x) : x \in K\}$*. In particular, both the sup and inf are finite.*

Proof. Choose a sequence $\{x_n\}$ so that $\lim_n f(x_n) = \sup\{f(x) : x \in K\}$. This then has a convergent subsequence with $\lim_k x_{n_k} = x_M \in K$. By continuity, $f(x_M) = \lim_k f(x_{n_k}) = \sup\{f(x) : x \in K\}$.

The rest of the proof is similar. □

One motivation for studying metrics is that they can often be built to capture various kinds of convergence. We illustrate this with one example.

Definition 2.18. Let (X,d) be a metric space and let $K \subseteq X$. We say that a sequence of functions $\{f_n\} \subset C(K)$ **converges uniformly to f on K** provided that

$$\lim_n \sup\{|f(x) - f_n(x)| : x \in K\} = 0.$$

We say that $\{f_n\}$ **converges uniformly on compact subsets** to f provided that $\{f_n\}$ converges uniformly to f on K for every compact subset K.

We now construct a metric that for $(\mathbb{R}, d)$ captures uniform convergence on compact subsets. Let $C(\mathbb{R})$ denote the continuous real-valued functions on $\mathbb{R}$ and let $f, g \in C(\mathbb{R})$. For each n, set $d_n(f,g) = \sup\{|f(t) - g(t)| : -n \leq t \leq +n\}$ and let $\rho_n(f,g) = \frac{d_n(f,g)}{1+d_n(f,g)}$. We define

$$\rho(f,g) = \sum_{n=1}^{\infty} \frac{\rho_n(f,g)}{2^n}.$$

Note that since $0 \leq \rho_n(f,g) \leq 1$ this series converges.

Lemma 2.19. *Let* $a, b, c \geq 0$*. Then* $a \leq b \iff \frac{a}{1+a} \leq \frac{b}{1+b}$*. If* $a \leq b + c$ *then* $\frac{a}{1+a} \leq \frac{b}{1+b} + \frac{c}{1+c}$*.*

Theorem 2.20. *Let* $f_n, f \in C(\mathbb{R})$*. Then* $\lim_n \rho(f_n, f) = 0$ *if and only if* $\{f_n\}$ *converges uniformly to* f *on compact subsets. Moreover,* $(C(\mathbb{R}), \rho)$ *is a complete metric space.*

Proof. First we show that ρ is a metric. It is clear that $\rho(f,g) = 0 \iff f = g$ and that $\rho(f,g) = \rho(g,f)$. To see the triangle inequality, let $f, g, h \in C(\mathbb{R})$. It is clear that $d_n(f,g) \leq d_n(f,h) + d_n(h,g)$ so by the lemma, $\rho_n(f,g) \leq \rho_n(f,h) + \rho_n(h,g)$. Now it follows that $\rho(f,g) \leq \rho(f,h) + \rho(h,g)$.

Suppose that $\lim_n \rho(f_n, f) = 0$. Given a compact subset K and $epsilon > 0$, pick an m so that $K \subseteq [-m, +m]$. Choose an N so that $n > N$ implies that $\rho(f_n, f) \leq 2^{-m}\frac{\epsilon}{1+\epsilon}$. Then for $n > N$ we have that

$$2^{-m}\frac{d_m(f_n,f)}{1+d_m(f_n,f)} \leq \rho(f_n,f) < 2^{-m}\frac{\epsilon}{1+\epsilon} \implies d_m(f_n,f) < \epsilon.$$

But $\sup\{|f_n(t) - f(t)| : t \in K\} \leq d_m(f_n, f) < \epsilon$. Hence, $\{f_n\}$ converges uniformly to f on K.

Conversely, assume that $\{f_n\}$ converges uniformly to f on every K and that $\epsilon > 0$ is given. Pick M so that $\sum_{m=M+1}^{\infty} 2^{-m} < \epsilon/2$. Pick δ so that $\frac{\delta}{1+\delta} = \epsilon/2$, and finally pick an N so that for $n > N$, we have that $d_M(f_n, f) < \delta$.

Then for $n > N$ we have that

$$\rho(f_n,f) \leq \sum_{m=1}^{M} 2^{-m}\rho_m(f_n,f) + \sum_{m=M+1}^{\infty} 2^{-m} \leq$$
$$\sum_{m=1}^{M} 2^{-m}\rho_M(f_n,f) + \epsilon/2 <$$
$$\frac{d_M(f_n,f)}{1+d_M(f_n,f)} + \epsilon/2 < \frac{\delta}{1+\delta} + \epsilon/2 = \epsilon.$$

Thus, uniform convergence on compact sets implies convergence in the metric.

Finally, if a sequence $\{f_n\}$ is Cauchy in the metric ρ, then it is pointwise Cauchy and so there is a function, $f(x) = \lim_n f_n(x)$. Also it is easy to show that for each M and $\epsilon > 0$, there must be an N so that $n, m > N$ implies that $d_M(f_n, f_m) < \epsilon$. From this it follows that, f_n converges uniformly to f on $[-M, +M]$ and so is continuous on $[-M, +M]$. Thus, f is continuous. More of the same shows that $\rho(f_n, f) \to 0$, and so the space is complete. ☐

2.5. **Baire's Theorem.** Some of the deepest results in functional analysis are consequences of Baire's theorem.

Definition 2.21. Let (X, d) be a metric space. A subset $DsubseteqX$ is **dense** provided that for every $x \in X$ and every $\epsilon > 0$ there is $y \in D$ with $d(x, y) < \epsilon$.

Theorem 2.22 (Baire)**.** *Let (X, d) be a complete metric space and let $\{U_n\}_{n\in\mathbb{N}}$ be a countable collection of open, dense sets. Then $\cap_{n\in\mathbb{N}} U_n$ is dense in X.*

Proof. Given $\epsilon > 0$ and $x \in X$, since U_1 is dense, we can find $x_1 \in U_1$ so that $d(x, x_1) < \epsilon/2$. Because U_1 is open we can pick $\delta_1 < \epsilon/2$ so that

$$B(x_1; \delta_1)^- \subseteq U_1 \cap B(x; \epsilon).$$

Applying the same reasoning to U_2 we may pick $x_2 \in U_2$ so that $d(x_1, x_2) < \delta_1/2$ and a $\delta_2 < \delta_1/2$ so that $B(x_2, \delta_2)^- \subseteq U_2 \cap B(x_1; \delta_1)$.

Continuing inductively, we define x_n, δ_n such that $\delta_{n+1} < \delta_n/2 \leq \epsilon/2^n$ and $B(x_{n+1}; \delta_{n+1})^- \in U_{n+1} \cap B(x_n; \delta_n)$.

The fact that $d(x_n, x_{n+1}) < \epsilon/2^n$ is enought to show that $\{x_n\}$ is Cauchy. So there is a point $x_0 = \lim_n x_n$. The containments imply that,

$$x_n, x_{n+1}, \ldots \text{ are all in } B(x_n, \delta_n)^-,$$

hence, $x_0 \in B(x_n; \delta_n)^- \subseteq U_n$. Thus, $x_0 \in \cap_{n \in \mathbb{N}} U_n$. Also, since each $B(x_n, \delta_n)^- \subseteq B(x, \epsilon)$ we have that $d(x, x_0) < \epsilon$. □

Definition 2.23. A set E is called **nowhere dense** provided that the complement of E^- is dense, i.e., $X \backslash E^-$ is dense.

Note that U_n is dense and open if and only if $E_n = X \backslash U_n$ is closed and nowhere dense.

A much weaker statement than saying that $\cap U_n$ is dense, is to say that it is non-empty. Taking complements, this is just the statement that $X \neq X \backslash \cap U_n = \cup (X \backslash U_n)$. Thus, the following result appears weaker than Baire's theorem, but it has a name that makes it sound better!

Theorem 2.24 (Baire's Category Theorem)**.** *A complete metric space cannot be written as a countable union of nowhere dense sets.*

The name comes from the following. A set that can be written as a countable union of nowhere dense sets is called a set of **first category**. A set that cannot is called a set of **second category**. In this language, the above theorem says that a complete metric space is of second category. Hence, the name. Another name for sets of first category is to call them **meagre sets**.

An example of this theorem, is to take $(\mathbb{Q}, d)$ since $\mathbb{Q}$ is countalbe we may write it as $\mathbb{Q} = \{r_n : n \in \mathbb{N}\}$. Now set $E_n = \{r_n\}$ then this is a countable collection of nowhere dense sets and $\mathbb{Q} = \cup E_n$. Thus, by the theorem, $(\mathbb{Q}, d)$ is not complete, something that we knew already. What about $(\mathbb{N}, d)$ and $E_n = \{n\}$?

3. Topological Spaces

Not all notions of convergence can be defined by a metric, this lead to a generalization of metric spaces. Recall that in a metric space, arbitrary unions of open sets are open and finite intersections of open sets are open. This motivates the following definition.

Definition 3.1. Let X be a set $\mathcal{T}$ be a collection of subsets of X. We call $\mathcal{T}$ a **topology on X** and call $(X, \mathcal{T})$ **a topological space**, provided that:

(1) $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$,
(2) whenever $U_a \in \mathcal{T}$, $\forall a \in A$, then $\left(\cup_{a\in A} U_a\right) \in \mathcal{T}$,
(3) whenever $U_i \in \mathcal{T}$, $1 \le i \le n$, then $\left(\cap_{i=1}^{n} U_i\right) \in \mathcal{T}$.

We call a subset of X **open** if it is in $\mathcal{T}$ and we call a subset of X **closed** if its complement is in $\mathcal{T}$. Given a point $x \in X$, we let $\mathcal{N}_x = \{U \in \mathcal{T} : x \in U\}$ and call this collection the **neighborhoods of x**.

In order to define continuity, we first need to see how to characterize continuity in metric spaces using only open sets. This leads to the following.

Definition 3.2. Given two topological spaces $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ and a function $f : X \to Y$ we say that:

- f is **continuous at** x_0 provided that for every $V \in \mathcal{S}$ with $f(x_0) \in V$ there is $U \in \mathcal{T}$ with $x_0 \in U$ such that $f(U) \subseteq V$, i.e., every neighborhood of $f(x_0)$ contains $f(U)$ for some neighborhood of x_0;
- f is **continuous** provided that for every $V \in \mathcal{S}$ we have that $f^{-1}(V) := \{x \in X : f(x) \in V\} \in \mathcal{T}$.

It is not hard to see that when X and Y are metric spaces these reduce to the usual definitions of continuity.

Problem 3.3. *Prove that $f : X \to Y$ is continuous if and only if it is continuous at every point in X.*

3.1. **Nets and Directed Sets.** Many results about topological spaces have the same proofs as in the metric space case if one replaces sequences with nets.

Definition 3.4. A pair $(\Lambda, \le)$ consisting of a set Λ together with a relation $\le$ on Λ is called a **directed set** if the relation satisfies:

(1) $x \le y$ and $y \le x$ implies that $x = y$(**symmetric**),
(2) $x \le y$ and $y \le z$ implies that $x \le z$(**transitive**),
(3) for every $x_1, x_2 \in \Lambda$ there is $x_3 \in \Lambda$ such that $x_1 \le x_3$ and $x_2 \le x_3$.

Below are some examples of directed sets.

Let Λ be either $\mathbb{N}, \mathbb{Z}$, or $\mathbb{R}$ with the usual $\le$.

Let S be a set and let $\mathcal{F}$ be the collection of all finite subsets of S. Given $F_1, F_2 \in \mathcal{F}$ define $F_1 \le F_2 \iff F_1 \subseteq F_2$.

Given a topological space $(X, \mathcal{T})$ and a point x, let $\Lambda = \mathcal{N}_x$–the set of all open neighborhoods of x and define $U_1 \le U_2 \iff U_2 \subseteq U_1$. The intuition of this order is that "farther out" means smaller set and so "closer" to x.

The most complicated example is the one that you use in calculus!

Given an interval $[a, b]$ recall that a **partition** is a set $P = \{x_0 = a < x_1 \cdots < x_n = b\}$. These are used to subdivide the interval. An **augmentation of P** is a collection, $P^* = \{x_1' < \cdots < x_n'\}$ where $x_{i-1} \le x_i' \le x_i$. The pair $\mathcal{P} = (P, P^*)$ is called an **augmented partition**. We define $\mathcal{P}_1 = (P_1, P_1^*) \le \mathcal{P}_2 = P_2, P_2^*)$ provided that $P_1 \subseteq P_2$ and $P_1^* \subseteq P_2^*$. It i easy to see that $\leq$ satisfies the 1st two properties needed to be a directed set. The 3rd is trickier.

Given any two augmented partitions $\mathcal{P}_1 = (P_1, P_1^*)$ and $\mathcal{P}_2 = (P_2, P_2^*)$, let $P_1 \cup P_1^* \cup P_2 \cup P_2^* = \{a = x_0 < ... < x_m = b\}$ add one extra point, c with $x_{m-1} < c < b$ and let $P_3 = \{x_0 < ... < x_{m-1} < c < x_m = b\}$ Now let $P_3^* = \{x_0 < ... < x_{m-1} < x_m\}$, then this is an augmentation of P_3 and $P_1 \cup P_2 \subseteq P_3$ and $P_1^* \cup P_2^* \subseteq P_3$ so that $\mathcal{P}_1 \leq \mathcal{P}_3$ and $\mathcal{P}_2 \leq \mathcal{P}_3$.

Definition 3.5. Let $(X, \mathcal{T})$ be a topological space. Then a **net in X** is a directed set $(\Lambda, \leq)$ together with a function $f : \Lambda \to X$. As with sequences we prefer to set $x_\lambda = f(\lambda)$ and write the net as $\{x_\lambda\}_{\lambda \in \Lambda}$. We say that the net $\{x_\lambda\}_{\lambda \in \Lambda}$ **converges to x** and write $\lim_\lambda x_\lambda = x$ provided that for each $U \in \mathcal{N}_x$ there is $\lambda_0 \in \Lambda$ such that when $\lambda_0 \leq \lambda$ then $x_\lambda \in U$.

Here is the careful statement of the Riemann sum theorem.

Let $f : [a, b] \to \mathbb{R}$ be continuous. For each augmented partition $\mathcal{P} = (\{a = x_0 < \cdots < x_n = b\}.\{x_1' < \cdots < x_n'\})$, the **Riemann sum** is given by

$$S_{\mathcal{P}} = \sum_{i=1}^{n} f(x_i')(x_i - x_{i-1}).$$

The set of all Riemann sums forms a net of real numbers.

Theorem 3.6. *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then the net of Riemann sums converges, and we call this limit the* **Riemann integral of f***.*

3.2. **Unordered vs Ordered Sums.** Given a set A and real numbers $x_a, a \in A$ we wish to define $\sum_{a \in A} r_a$. To do this consider the directed set $(\mathcal{F}, \leq)$ of all finite subsets of A. Given $F \in \mathcal{F}$ we set

$$s_F = \sum_{a \in F} r_a$$

and we call this number the **partial sum over F**. The collection $\{s_F\}_{F \in F}$ is a net of real numbers, called the **net of partial sums**.

Definition 3.7. We say that $\sum_{a \in A} r_a$ **converges to s** provided that s is the limit of the net of partial sums.

Thus, $s = \sum_{a \in A} r_a$ iff for each $\epsilon > 0$ there is a finite set $F_0 \subseteq A$ such that for every finite set F with $F_0 \subseteq F$ we have that

$$|s - s_F| < \epsilon.$$

Given $r_n, n \in \mathbb{N}$ it is interesting to compare the convergence of the unordered series, $\sum_{n \in \mathbb{N}} r_n$ with the convergence of the ordered series, $\sum_{n=1}^{+\infty} r_n$. For the latter case we only consider partial sums of the form $\sum_{n=1}^{K} r_n$, a very small collection of all finite subsets of $\mathbb{N}$!

Since we only need this smaller collection of partial sums to approach a value, it is easy to see that whenever $\sum_{n \in \mathbb{N}} r_n$ converges, then $\sum_{n=1}^{+\infty} r_n$ will converge.

The converse is not true. In fact, $\sum_{n=1}^{+\infty} \frac{(-1)^n}{n}$ converges, but $\sum_{n\in\mathbb{N}} \frac{(-1)^n}{n}$ does not converge.

Problem 3.8. *Prove that $\sum_{n\in\mathbb{N}} r_n$ converges iff $\sum_{n=1}^{+\infty} |r_n|$ converges, i.e. iff the series* **converges absolutely***.*

Here are just some of the reasons that nets are convenient.

Proposition 3.9. *Let $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ be topological spaces and let $f : X \to Y$. Then*

(1) *f is continuous at x_0 iff for every net $\{x_\lambda\}_{\lambda\in\Lambda}$ such that $\lim_\lambda x_\lambda = x_0$ we have that $\lim_\lambda f(x_\lambda) = f(x_0)$.*
(2) *f is continuous iff whenever a net $\{x_\lambda\}_{\lambda\in\Lambda}$ in X converges to a point x, then $\lim_\lambda f(x_\lambda) = f(x)$.*
(3) *A set $C \subseteq X$ is closed iff whenever $\{x_\lambda\}_{\lambda\in\Lambda}$ is a net in C that converges to a point $x \in X$, then $x \in C$.*
(4) *A set $K \subseteq X$ is compact iff every net in K has a subnet that converges to a point in K.*

Note that these results are the counterparts of many results for metric spaces. In general they are not true if you only use sequences instead of nets.

We have not yet defined subnets.

Definition 3.10. Let Λ and D be two directed sets. A function $g : D \to \Lambda$ is called **final** provided that $a \le b \implies g(a) \le g(b)$ and given any $\lambda_0 \in \Lambda$ there is $d_0 \in D$ such that $d_0 \le d \implies \lambda_0 \le g(d)$. Given a net $\{x_\lambda\}_{\lambda\in\Lambda}$ a **subnet** is any net of the form $\{x_{g(d)}\}_{d\in D}$ for some directed set D and final function $g : D \to \Lambda$.

Just as with subsequences, we will often write a subnet as $\{x_{\lambda_d}\}_{d\in D}$ where really, $\lambda_d = g(d)$.

CAREFUL: When $\Lambda = D = \mathbb{N}$ then every subsequence is a subnet but a subnet need not be a subsequence. Explain why.

3.3. **The key separation axiom.** A topological space $(X, \mathcal{T})$ is called **Hausdorff** provided that for any $x \ne y$ there are open set U, V such that $x \in U$, $y \in V$ and $U \cap V = \emptyset$.

This axiom guarantees that there are lots of continuous functions.

Theorem 3.11 (Urysohn's Lemma)**.** *Let $(K, \mathcal{T})$ be a compact Hausdorff space and let A, B be closed subsets with $A \cap B = \emptyset$. Then there is a cntinuous function $f : K \to [0, 1]$ such that $f(a) = 0$, $\forall a \in A$ and $f(b) = 1$, $\forall b \in B$.*

Theorem 3.12 (Tietze's ExtensionTheorem)**.** *Let $(K, \mathcal{T})$ be a compact, Hausdorff space and let $A \subseteq K$ be a closed subset and let $f : A \to [0, 1]$ be continuous. Then there is a cntinuous function $F : K \to [0, 1]$ such that $F(a) = f(a)$, $\forall a \in A$.*

4. Banach Spaces

All vector spaces will be over the field $\mathbb{R}$ or $\mathbb{C}$. When I want to let the field be either one I will write $\mathbb{F}$.

Definition 4.1. Let X be a vector space over $\mathbb{F}$. Then a **norm** on X is a function, $\|\cdot\| : X \to \mathbb{R}^+$ satisfying:

(1) $\|x\| = 0 \iff x = 0$,
(2) $\|\lambda x\| = |\lambda| \cdot \|x\|$, $\forall \lambda \in \mathbb{F}$, $\forall x \in X$,
(3) (Triangle Inequality) $\|x + y\| \le \|x\| + \|y\|$, $\forall x, y \in X$.

We call the pair $(X, \|\cdot\|)$ a **normed linear space** or **n.l.s.**, for short.

Proposition 4.2. *Let $(X, \|\cdot\|)$ be a n.l.s., then $d(x, y) = \|x - y\|$ is a metric on X(called the* **induced metric***) that satisfies $d(x + z, y + z) = d(x, y)$***translation invariance** *and $d(\lambda x, \lambda y) = |\lambda| d(x, y)$***scaled***. Conversely, if X is a vector space and d is a metric on X that is scaled and translation invariant, then $\|x\| = d(0, x)$ is a norm on X.*

Definition 4.3. A n.l.s. $(X, \|\cdot\|)$ is a **Banach space** if and only if it is complete in the induced metric, $d(x, y) = \|x - y\|$.

It is not hard to see that the induced metric on a n.l.s. satsfies: $\lambda_n \to \lambda \implies \lambda_n x \to \lambda x$ and $x_n \to x, y_n \to y \implies x_n + y_n \to x + y$. If $(X$ is a vector space and ρ is a complete metric on X satsfying these two properties, then (X, ρ) is called a **Frechet space**. The vector space $C(\mathbb{R})$ with the metric ρ corresponding to uniform convergence on compact subsets is an example of a Frechet space that is not a Banach space.

Example 4.4. For $1 \le p < +\infty$ and $n \in \mathbb{N}$, and $x = (x_1, ..., x_n) \in \mathbb{R}^n$, set

$$\|(x_1, ..., x_n)\|_p = \big(\sum_{j=1}^{n} |x_j|^p\big)^{1/p},$$

then this is a norm called the **p-norm** and we write ℓ_n^p to denote $\mathbb{R}^n$ endowed with this norm. It is not so easy to see that this satisfies the triangle inequality, this depends on some results below. For $p = +\infty$ we set

$$\|(x_1, ..., x_n)\| = \max\{|x_1|, ..., |x_n|\}.$$

Example 4.5. More generally, for $1 \le p < +\infty$, we let

$$\ell^p = \{(x_1, x_2, ...) : \sum_{j=1}^{+\infty} |x_j|^p < +\infty\},$$

and for $x \in \ell^p$ we set $\|x\|_p = \big(\sum_{j=1}^{+\infty} |x_j|^p\big)^{1/p}$.

For $p = +\infty$, we set

$$\ell^\infty = \{(x_1, x_2, ...) : \sup_j |x_j| < +\infty\},$$

and define $\|x\|_\infty = \sup_j |x_j|$.

Then for $1 \le p \le +\infty$, $(\ell^p, \|\cdot\|_p)$ are all Banach spaces. This fact relies on a number of theorems that we will state below.

Lemma 4.6 (Young's Inequality)**.** *Let $1 < p, q < +\infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then for any $a, b \ge 0$ we have $ab \le \frac{a^p}{p} + \frac{b^q}{q}$.*

Problem 4.7. *Prove this inequality.*

Proposition 4.8 (Holder's Inequality)**.** *Let $1 < p, q < +\infty$, and let $x = (x_1, ...) \in \ell^p$, $y = (y_1, ...) \in \ell^q$, then $(x_1 y_1, ...) \in \ell^1$ and*

$$\sum_n |x_n y_n| \le \|x\|_p \|y\|_q.$$

Problem 4.9. *Prove Holder's inequality.*

Theorem 4.10 (Minkowski's Inequality)**.** *Let $1 \le p \le +\infty$, and let $x, y \in \ell^p$, then $x + y \in \ell^p$ and $\|x + y\|_p \le \|x + y\|_p$.*

Proof. We only do the case that $1 < p < +\infty$. First note that

$$|x_i + y_i| \le \begin{cases} 2|x_i| & \text{when } |x_i| \ge |y_i| \\ 2|y_i| & \text{when } |x_i| \le |y_i| \end{cases}.$$

Hence, $|x_i + y_i|^p \le \max\{2^p|x_i|^p,\ 2^p|y_i|^p\}$, from which it follows that $\sum_i |x_i + y_i|^p < +\infty$.

Next notice that

$$|x_i + y_i|^p \le |x_i + y_i|^{p-1}(|x_i| + |y_i|).$$

Also, $p + q = pq$ so that $(p-1)q = p$. From this it follows that the sequence $|x_i + y_i|^{p-1}$ is q-summable. By Holder's inequality,

$$\begin{aligned} \sum_i |x_i + y_i|^p &\le \big(\sum_i (|x_i + y_i|^{p-1})^q\big)^{1/q} \big((\sum_i |x_i|^p)^{1/p} + (\sum_i |y_i|^p)^{1/p}\big) \\ &= \big(\sum_i |x_i + y_i|^p\big)^{1/q} \big((\sum_i |x_i|^p)^{1/p} + (\sum_i |y_i|^p)^{1/p}\big). \end{aligned}$$

Cancelling the common term from each side and using $\frac{1}{q} = 1 - \frac{1}{p}$ yields the result.

□

Theorem 4.11 (Riesz-Fischer)**.** *The spaces $(\ell^p, \|\cdot\|_p)$ are Banach spaces.*

Example 4.12. Let $(K, \mathcal{T})$ be a compact Hausdorff space and let $C(K)$ denote the space of continuous real or complex valued functions on K. This is a vector space. If we set $\|f\|_\infty = \sup\{|f(x)|; x \in K\}$, then this defines a norm and $(C(K), \|\cdot\|_\infty)$ is a Banach space. Completeness follows from the fact that convergence in this norm is uniform convergence and the fact that uniformly convergent sequences of continuous function converge to a continuous function.

4.1. **Quotient Spaces.** A concept that plays a big role in functional analysis are quotient spaces. Recall from algebra, that if we have a vector space X and a subspace Y then there is a **quotinet space**, denoted X/Y consisting of *cosets*,

$$X/Y := \{x + Y : x \in Y\},$$

and defining, $(x_1 + Y) + (x_2 + Y) = (x_1 + x_2) + Y$, $\lambda(x + Y) = (\lambda x) + Y$ gives well-defined operations that makes this into a vector space. One often used fact is that if $T : X \to Z$ is a linear map and $Y = ker(T)$ then there is a well-defined linear map $\hat{T} : X/Y \to Z$ given by $\hat{T}(x + Y) = T(x)$.

To form quotients of normed spaces, we want one more condition. Let $(X, \|\cdot\|)$ be a normed space and let Y be a closed subspace. On the quotient space X/Y, we define

$$\|x + Y\| := \inf\{\|x + y\| : y \in Y\} = \inf\{\|x - y\| : y \in Y\}.$$

The second formula shows that $\|x + Y\|$ is the distance from x to the set Y. This is the reason that we want Y closed, otherwise there would be a point in the closure that is not in Y and any such point would be distance 0 from Y. Thus, such a vector would be one for which $x + Y \neq 0 + Y$ but $\|x + Y\| = 0$. This quantity is called the **quotient norm**.

Proposition 4.13. *Let $(X, \|\cdot\|)$ be a normed space and let Y be a closed subspace, then the quotient norm is a norm on X/Y.*

The vector space together with the quotient norm is called the **quotient space** or if we want to be very clear the **normed quotient space**.

4.2. **Conditions for Completeness.** Here are two results that are often useful for proving that a Banach space is complete.

Theorem 4.14. *Let $(X, \|\cdot\|)$ be a normed space. It is a Banach space if and only if whenever $\{x_n\} \subseteq X$ is a sequence such that $\sum_{n=1}^{\infty} \|x_n\| < +\infty$ then there is an element $x \in X$ such that $\sum_{n=1}^{+\infty} x_n = x$.(This is sometimes stated as* **absolutely convergent sequences are convergent***.)*

Next is one of many *2 out of 3* theorems.

Theorem 4.15 (The 2/3's Theorem). *Let $(X, \|\cdot\|)$ be a normed space, let $Y \subseteq X$ be a closed subspace and let X/Y be the normed quotient space. If any 2 of these normed spaces is a Banach space, then the 3rd space is also a Banach space.*

Proof. We only prove the case, that if Y and X/Y are Banach spaces, then so is X. To this end let $\{x_n\}$ be a Cauchy sequence in X and let $[x_n] = x_n + Y$ be the image of this sequence in X/Y. Since $\|[x_n] - [x_m]\| \leq \|x_n - x_m\|$, this is a Cauchy sequence in X/Y, and hence, converges to some $[x] \in X/Y$. Since $\|[x - x_n]\| \to 0$ we can pick $y_n \in Y$ such that $\|x - x_n - y_n\| \to 0$. Note that

$$\|y_n - y_m\| = \|(x_n + y_n - x) - (x_m + y_m - x) + (x_m - x_n)\|$$
$$\leq \|x_n + y_n - x\| + \|x_m + y_m - x\| + \|x_n - x_m\|,$$

since the first two terms tend to 0 and the 3rd is Cauchy, it is easy to see that $\{y_n\}$ is a Cauchy sequence in Y and hence $y_n \to y$.

This implies that

$$x_n = (x_n + y_n - x) + x - y_n \to x - y,$$

and so X is complete. □

To see how this last result can be used we will show that every finite dimensional normed space is automatically complete(although this is not the shortest way to prove this fact).

Theorem 4.16. *Every finite dimensional normed space is complete.*

Proof. We do the case that the field is $\mathbb{R}$, the case for $\mathbb{C}$ is identical.

This is easy to see when the dimension is one. To see pick any $x \in X$ with $\|x\| = 1$ Then every vector is of the form λx and $\|\lambda_1 x - \lambda_2 x\| = |\lambda_1 - \lambda_2|$. Thus, a sequence $\{\lambda_n x\}$ is Cauchy in X iff $\{\lambda_n\}$ is Cauchy in $\mathbb{R}$.

Now, inductively assume that it is true for spaces of dimension n and let X be $(n+1)$-dimensional take a n-dimensional subspace Y. It will be complete by the inductive hypothesis and hence closed. Also the quotient X/Y is one-dimensional so it is complete. Hence by the 2/3 theorem X is complete. □

4.3. **Norms and Unit Balls.** Let $(X, \|\cdot\|)$ be a normed space, we call $\{x : \|x\| < 1\}$ the **open unit ball** and $\{x : \|x\| \leq 1\}$ the **closed unit ball**. It is not hard to see that these sets are in fact open and closed, respectively and that the closed unit ball is the closure of the open unit ball. We wish to characterize norms in terms of properties of these balls.

Definition 4.17. A subset C of a vector space V is called:

- **convex** if whenever $x, y \in C$ and $0 \leq t \leq 1$ then $tx + (1-t)y \in C$, i.e., provided that the line segment from x to y is in C,
- **absolutely convex** if whenever $x_1, ..., x_N \in C$ and $|\lambda_1| + \cdots + |\lambda_n| = 1$, then $\lambda_1 x_1 + \cdots + \lambda_n x_n \in C$,
- **absorbing** if whenever $y \in V$ then there is a $t > 0$ such that $ty \in C$.

A **ray** is any set of the form $\{ty : t > 0\}$ for some $y \in V$.

Proposition 4.18. *Let $(X, \|\cdot\|)$ be a normed space and let $\mathcal{B}$ be the open(or closed) unit ball, then $\mathcal{B}$ is absolutely convex, absorbing and contains no rays. Moreover,*

$$\|x\| = \inf\{t > 0 : t^{-1}x \in \mathcal{B}\} = \big(\sup\{r > 0 : rx \in \mathcal{B}\}\big)^{-1}.$$

Conversely, if $C \subseteq X$ is any absolutely convex, absorbing set that contains no rays and we define

$$\|x\| = \inf\{t > 0 : t^{-1}x \in C\},$$

then $\|\cdot\|$ is a norm on X and

$$\{x : \|x\| < 1\} \subseteq C \subseteq \{x : \|x\| \leq 1\}.$$

This latter fact gives us a way to generate many norms, by just defining sets C that satisfy the above conditions. Also notice that $\|x\|_1 \leq \|x\|_2$ iff and only if the unit balls satisfy $\mathcal{B}_2 \subseteq \mathcal{B}_1$. Since the p-norms on $\mathbb{R}^n$ satisfy $p \leq p' \implies \|x\|_p \geq \|x\|_{p'}$ we see that $\mathcal{B}_p \subseteq \mathcal{B}_{p'}$.

Suppose that $\mathcal{B}_1$ and $\mathcal{B}_2$ are the unit balls of two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X. It is easy to see that $\mathcal{B}_1 \cap \mathcal{B}_2$ satisfies the conditions needed to be a unit ball. What is the norm?

Consider $K = \{(x, y) : xy \geq 0 \text{ and } x^2 + y^2 \leq 1\}$. Find the smallest set containing K that is absolutely convex, absorbing find the norm that this set defines.

4.4. **Polars and the Absolutely Convex Hull.** Given any set K the intersection of all absolutely convex sets containing K is the smallest absolutely convex set containing K, this set is called the **absolutely convex hull of K**. Its closure is called the **closed absolutely convex hull of K**.

Given any set $K \subseteq \mathbb{R}^n$ its **polar** is the set

$$K^0 := \{y : |x \cdot y| \leq 1\}.$$

The set $K^{00} := (K^0)^0$ is called the **bipolar**

Theorem 4.19 (The Bipolar Theorem)**.** *Let $K \subseteq \mathbb{R}^n$. Then K^0 is a closed, absolutely convex set. The bipolar K^{00} is the closed absolutely convex hull of K.*

If $\mathcal{B}_i \subseteq \mathbb{R}^n$ are closed unit balls for norms $\|\cdot\|_i$, then $\mathcal{B} = \big(\mathcal{B}_1 \cup \mathcal{B}_2\big)^{00}$ is the closed unit ball for a norm called the **decomposition norm** and it satisfies

$$\|x\| = \inf\{\|y\|_1 + \|z\|_2 : x = y + z\}.$$

4.5. **Reverse Triangle Inequality.**

Proposition 4.20. *Let $(X, \|\cdot\|)$ be a normed space. Then*

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

4.6. **Bounded and Continuous.**

Proposition 4.21. *Let $(X, \|\cdot\|_1)$ and $(Y, \|\cdot\|_2)$ be normed spaces and let $T : X \to Y$ be a linear map. Then the following are equivalent:*

(1) *T is continuous,*
(2) *T is continuous at 0,*
(3) *there is a constant M such that $\|Tx\|_2 \leq M\|x\|_1$, $\forall x \in X$*

Definition 4.22. Let X and Y be normed spaces and let $T : X \to Y$ be a linear map. Then T is **bounded** if there exists a constant M such that $\|TX\| \leq M\|x\|, \forall x \in X$. We set

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

and this is called the **operator norm** of T.

Some other formulas for this value, when $X \neq 0$, include

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

4.7. **Equivalence of Norms.** Suppose that $\|\cdot\|_i, i = 1, 2$ are norms on X and that $\rho_i, i = 1, 2$ are the metrics that they induce. Recall that we say that the metrics are equivalent iff the maps $id : (X, \rho_1) \to (X, \rho_2)$ and $id : (X, \rho_2) \to (X, \rho_1)$ are both continuous. But since these are normed spaces this is the same as requiring that both maps be bounded, which is the same as both metrics being uniformly equivalent. Thus, the metrics that come from norms are equivalent iff they are uniformly equivalent.

This leads to the following definition.

Definition 4.23. Let $(X, \|\cdot\|_i), i = 1, 2$ be two norms on X. Then we say that these norms are **equivalent** provided taht there exists constants, $A, B > 0$ such that

$$A\|x\|_1 \leq \|x\|_2 \leq B\|x\|_1, \forall x \in X.$$

Note that these last inequalities are the same as requiring that there are constants $C, D > 0$ such that $\|x\|_2 \leq C\|x\|_1$ and $\|x\|_1 \leq D\|x\|_2$. If we think geometrically then this is the same as requiring that $A \cdot \mathcal{B}_1 \subseteq \mathcal{B}_2 \subseteq B \cdot \mathcal{B}_1$, where $\mathcal{B}_i$ denotes the unit balls in each norm.

Finally, note that "equivalence of norms" really is an equivalence relation on the set of all norms on a space X, that is, it is a symmetric($\|\cdot\|_1$ equivalent to $\|\cdot\|_2$ iff $\|\cdot\|_2$ is equivalent to $\|\cdot\|_1$) and transitive($\|\cdot\|_1$ equivalent to $\|\cdot\|_2$ and $\|\cdot\|_2$ equivalent to $\|\cdot\|_3$ implies that $\|\cdot\|_1$ is equivalent to $\|\cdot\|_3$) relation.

Theorem 4.24. *All norms on $\mathbb{R}^n$ are equivalent.*

Proof. It will be enough to show that an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^n$ is equivalent to the Euclidean norm $\|\cdot\|_2$. Let $e_i = |i>$ denote the standard basis for $\mathbb{R}^n$, so that $x = (x_1, ..., x_n) = x_1e_1 + \cdots + x_ne_n$. Hence,

$$\|x\| \leq |x_1|\|e_1\| + \cdots + |x_n|\|e_n\| \leq \big(\sum_i |x_i|^2\big)^{1/2}\big(\sum_i \|e_i\|^2\big)^{1/2} \leq C\|x\|_2,$$

where $C = \big(\sum_i \|e_i\|^2\big)^{1/2}$.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = \|x\|$. By the reverse triangle inequality we have that,

$$|f(x) - f(y)| \leq \|x - y\| \leq C\|x - y\|_2,$$

which shows that f is continuous from $\mathbb{R}^n$ to $\mathbb{R}$ when $\mathbb{R}^n$ is given the usual Euclidean topology.

Now $K = \{x : \|x\|_2 = 1\}$ is a compact set and hence,

$$A = \inf\{f(x) : x \in K\} = f(x_0) = \|x_0\| \neq 0.$$

Now for any non-zero $x \in \mathbb{R}^n$, $\frac{x}{\|x\|_2} \in K$ and hence,

$$A \leq \|\frac{x}{\|x\|_2}\| = \frac{\|x\|}{\|x\|_2},$$

so that $A\|x\|_2 \leq \|x\|$ and we are done. □

5. Consequences of Baire's Theorem

Many results are true for Banach spaces that are not true for general normed spaces. This is because Banach spaces are complete metric spaces and so we have Baire's theorem as a tool. In this section we look at many facts about Banach spaces that rely on Baire's theorem. The version of Baire's theorem that we shall use most often is the version that says that in a complete metric space, a countable union of nowhere dense sets is still nowhere dense.

5.1. **Principle of Uniform Boundedness.** We begin with an example to show why the following result is perhaps surprising. Let $f_n : [0, 1] \to \mathbb{R}$ be defined by

$$f_n(t) = \begin{cases} n^2 t, & 0 \leq t \leq 1/n \\ 2n - n^2 t, & 1/n \leq t \leq 2/n \\ 0, & 2/n \leq t \leq 1 \end{cases},$$

so that each f_n is continuous.

We have that $\|f_n\|_\infty = n$ so that each function is bounded. Also for each x we have that for n large enough, $f_n(x) = 0$. Hence, we also have that for each x there is a constant M_x so that $|f_n(x)| \leq M_x$, $\forall n$. However, clearly $\sup\{\|f_n\|_\infty\} = +\infty$.

Note that in this case the domain is a complete metric space.

Theorem 5.1 (Principle of Uniform Boundedness)**.** *Let X be a Banach space, Y a normed space, and let $\mathcal{F}$ be a set of linear mappings from X to Y. If for each $A \in \mathcal{F}$ there is a constant C_A such that $\|Ax\| \leq C_A\|x\|$(i.e., each A is bounded) and for each $x \in X$ there is a constant M_x so that $\|Ax\| \leq M_x$(i.e., the collection $\mathcal{F}$ is bounded at each point), then*

$$\sup\{\|A\| : A \in \mathcal{F}\}$$

is finite.

Proof. Let $E_n = \{x : \|Ax\| \leq n, \forall A \in \mathcal{F}\}$. Since each A is continuous, it is not hard to see that each E_n is closed. By the hypothesis of pointwise boundedness, $X = \cup_n E_n$. Hence, one of these sets is not nowhere dense, say for n_0. Since E_{n_0} is already closed, there must exist x_0 and a constant $r > 0$ so that $\{y : \|y - x_0\| < r\} \subseteq E_{n_0}$.

Given any $0 \neq z \in X$, we have that $\|\frac{rz}{2\|z\|}\| < r$ so that $x_0 + \frac{rz}{2\|z\|} \in E_{n_0}$. Now given any $A \in \mathcal{F}$ we have that

$$\|A(\frac{rz}{2\|z\|})\| = \|A(x_0 + \frac{rz}{2\|z\|}) - A(x_0)\| \leq n_0 + n_0.$$

Hence, for any $0 \neq z$,

$$\|Az\| \leq \frac{4n_0\|z\|}{r} \text{ and } \|A\| \leq \frac{2n_0}{r}, \forall A \in \mathcal{F}.$$

□

The following is just the contrapositive statement of the above theorem, but it is often useful.

Theorem 5.2 (Resonance Principle)**.** *Let X be a Banach space and let Y be a normed space. If $\mathcal{F}$ is a set of bounded linear maps from X to Y with $\sup\{\|A\|; A \in \mathcal{F}\} = +\infty$, then there is a vector $x \in X$ such that $\sup\{\|Ax\| : A \in \mathcal{F}\} = +\infty$.*

The vector given by the above result is often called a **resonant vector** for the set $\mathcal{F}$.

Here is an example to show what can go wrong if the domain X is not a Banach space.

Definition 5.3. Let $C_{00} = span\{e_n : n \in \mathbb{N}\} \subseteq \ell^\infty$

Thus, $x = (x_1, x_2, ...) \in C_{00}$ iff there is an N_x so that $x_n = 0$ for all $n > N_x$. The number N_x varies from vector to vector.

Example 5.4. Define $A_n : C_{00} \to C_{00}$ by $A_n(x) = (x_1, 2x_2, 3x_3, ..., nx_n, 0, 0, ...)$. Clearly, each A_n is a bounded linear map with $\|A_n\| = n$. Also for each $x \in C_{00}$ we have an N_x as above and so $\sup_n \|A_n x\| \leq N_x < +\infty$.

Problem 5.5. *Let X and Y be normed spaces with X Banach, and let $A_n : X \to Y$ be a sequence of bounded linear maps such that for each $x \in X$ the sequence $A_n(x)$ converges in nomr to an element of Y. Prove that if we set $A(x) = \lim_n A_n(x)$ then A defines a bounded limear map. (The sequence $\{A_n\}$ is said to converge* **strongly** *to A and we write $A_n \xrightarrow{S} A$.*

5.2. **Open Mapping, Closed Graph, and More!** The following results are also all consequences of Baire's theorem applied to Banach spaces. In fact, given any one as the starting point, it is possible to deduce the others as consequences. These require that the domain and range both be Banach spaces.

Lemma 5.6 (No Name)**.** *Let X and Y be Banach spaces, and let $A : X \to Y$ be a bounded linear map that is onto. Then there exists $\delta > 0$ such that*

$$\mathcal{B}_Y(0;\delta) := \{y \in Y : \|y\| < \delta\} \subseteq A\big(\mathcal{B}_X(0;1)\big),$$

where $\mathcal{B}_X(0;1) = \{x \in X : \|x\| < 1\}$.

Theorem 5.7 (Bounded Solutions). *Let X and Y be Banach spaces and let $A : X \to Y$ be a bounded linear map. If for each $y \in Y$ there is $x \in X$ solving $Ax = y$, then there is a constant $C > 0$ such that for each $y \in Y$ there is a solution to $Ax = y$ with $\|x\| \leq C\|y\|$.*

Theorem 5.8 (Bounded Inverse Theorem). *Let X and Y be Banach spaces and let $A : X \to Y$ be a bounded linear map. If A is one-to-one and onto then the inverse map $A^{-1} : Y \to X$ is bounded.*

Theorem 5.9 (Equivalence of Norms). *Let X be a space with two norms $\|\cdot\|_1$ and $\|\cdot\|_2$, such that X is a Banach space in both norms. If there is a constant such that $\|x||_1 \leq C\|x\|_2$, $\forall x \in X$, then the norms are equivalent.*

As an application of the equivalence of norms theorem, note that once we proved that

Theorem 5.10 (Open Mapping Theorem). *Let X and Y be Banach spaces, and let $A : X \to Y$ be a bounded linear map. If A is onto, then for every open set $U \subseteq X$ the set $A(U)$ is open in Y.*

If X and Y are vector spaces, then their Cartesian product is also a vector space with operations, $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ and $\lambda(x, y) = (\lambda x, \lambda y)$. This space is sometimes denoted $X \oplus Y$. There are many equivalent norms that can be put on $X \oplus Y$, we will use $\|(x, y)\|_1 := \|x\| + \|y\|$. It is fairly easy to see that a subset $C \subseteq X \oplus Y$ is closed in the $\|\cdot\|_1$ norm iff $\|x_n - x\| \to 0$ and $\|y_n - y\| \to 0$ implies that $(x, y) \in C$.

Given a linear map, $A : X \to Y$ its **graph** is the subset $G_A := \{(x, Ax) : x \in X\}$.

Theorem 5.11 (Closed Graph Theorem). *Let X and Y be Banach spaces, and let $A : X \to Y$ be a linear map. If $\|x_n - x\| \to 0$ and $\|Ax_n - y\| \to 0$ implies that $Ax = y$(this is equivalent to the statement that G_A is a closed subset), then A is bounded.*

In general, a linear map $A : X \to Y$ is called **closed** provided that G_A is closed. When X and Y are Banach spaces, then closed implies bounded by the above. But if X is not Banach then the map need not be bounded. Here is an eample.

Example 5.12. Let $C^1([0, 1]) := \{f : [0, 1] \to \mathbb{R} : f'$ exists and is continuous $\}$. Set $\|f\| = \|f$
$_\infty$ and let $Y = C([0, 1])$, also with $\|\cdot\|_\infty$. Let $D : C^1([0, 1]) \to C([0, 1])$ by $D(f) = f'$. By Rudin, Theorem 7.1.7, if $\|f_n - f\|_\infty \to 0$ and $\|f_n' - g\|_\infty \to 0$, then f is differentiable and $g = f'$. This result is equivalent to saying that the graph of D is closed. But D is not bounded since $\|D(t^n)\|_\infty = n\|t^n\|_\infty$. What goes wrong is that $C^1([0, 1])$ is not complete in $\|\cdot\|_\infty$. This last fact can be seen directly, or it can be deduced as an application of the closed graph theorem.

Problem 5.13. *Let X be a Banach space and let Y and Z be two closed subspaces of X such that $Y \cap Z = (0)$ and $Y + Z := \{y + z : y \in Y, z \in Z\} = X$. Thus, each $x \in X$ has a unique representation as $x = y + z$. Define $\|x\|_1 = \|y\| + \|z\|$. Prove that $(X, \|\cdot\|_1)$ is a Banach space and that $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent norms on X.*

6. Hahn-Banach Theory

The Hahn-Banach theorem is usually stated for norm extensions, but it is really more general with the proofs being no harder, so we state it in full generality.

Definition 6.1. Let X be a vector space. Then a function $p : X \to \mathbb{R}$ is **sublinear** if:

(1) $p(x + y) \le p(x) + p(y)$, $\forall x, y \in X$,
(2) for all $t > 0$, $p(tx) = tp(x)$.

Note that, unlike a norm, it is not required that $p(-x) = p(x)$.

Lemma 6.2 (One-Step Extension)**.** *Let X be a real vector space, $Y \subseteq X$ a subspace, $p : X \to \mathbb{R}$ a sublinear functional and $f : Y \to \mathbb{R}$ a linear functional satisfying $f(y) \le p(y)$, $\forall y \in Y$. If $x \in X \backslash Y$ and $Z = span\{Y, x\}$, then there exists a linear functional $g : Z \to \mathbb{R}$ such that $g(y) = f(y), \forall y \in Y$ (i.e., g is an* **extension** *of f) and $g(z) \le p(z)$, $\forall z \in Z$.*

Proof. Since every vector in Z is of the form $z = rx + y$ for some $r \in \mathbb{R}$ every possible extension of f is of the form $g_\alpha(rx + y) = r\alpha + f(y)$. So we need to show that for some choice of α we will have that $r\alpha + f(y) \le p(rx + y)$, $\forall r, y$.

First note that if $y_1, y_2 \in Y$, then

$$f(y_1 + y_2) \le p(y_1 + y_2) \le p(y_1 + x) + p(y_2 - x).$$

Hence,

$$f(y_2) - p(y_2 - x) \le p(y_1 + x) - f(y_1), \ \forall y_1, y_2 \in Y.$$

Thus,

$$\alpha_0 := \sup\{f(y) - p(y - x) : y \in Y\} \le \inf\{p(y + x) - f(y) : y \in Y\} =: \alpha_1.$$

The proof is completed by showing that for any α with $\alpha_0 \le \alpha \le \alpha_1$ we will have that g_α is an extension with $g_\alpha(rx + y) \le p(rx + y)$. □

Theorem 6.3 (Hahn-Banach Extension Theorem for Sublinear Functions)**.** *Let X be a real vector space, $p : X \to \mathbb{R}$ a sublinear functional, $Y \subseteq X$ a subspace and $f : Y \to \mathbb{R}$ a real linear functional satisfying $f(y) \le p(y)$, $\forall y \in Y$. Then there exist a linear functional $g : X \to \mathbb{R}$ that extends f and satisfies, $g(x) \le p(x) < \forall x \in X$.*

The idea of the proof is to use the one-step extension lemma and then apply Zorn's lemma to show that an extension exists whose domain is maximal among all extensions and argue that necessarily this maximal extension has domain equal to all of X. Zorn's lemma is equivalent to the Axiom of

Choice, so this is one theorem that requires this additional axiom, which almost every functional analyst is comfortable with assuming.

Next we look at complex versions. The follwoing result is often useful for passing between real linear maps and complex linear maps.

Proposition 6.4. *Let X be a complex vector space.*

(1) *If $f : X \to \mathbb{R}$ is a real linear functional, then setting*
$$g(x) = f(x) - if(ix)$$
defines a complex linear functional.

(2) *If $f : X \to \mathbb{C}$ is a complex linear functional, then $Ref(x) = Re(f(x))$, is a real linear functional and*
$$f(x) = Ref(x) - iRef(ix).$$

Theorem 6.5. *Let X be a complex vector space, $p : X \to \mathbb{R}$ be a function that satisfies, $p(x + y) \le p(x) + p(y)$ and $p(\lambda x) = |\lambda|p(x), \forall \lambda \in \mathbb{C}$. If $Y \subseteq X$ is a complex subspace and $f : Y \to \mathbb{C}$ is a complex liner functional satisfying $|f(y)| \le p(y), \forall y \in Y$, then there exists a comple linear functional $g : X \to \mathbb{C}$ that extends f and satisfies $|g(x)| \le p(x), \forall x \in X$.*

Proof. Consider the real linear functional Ref, observe that it satisfies $Ref(y) \le p(y)$. Apply the previous theorem to extend it to a real linear functional $h : X \to \mathbb{R}$ satisfying $h(x) \le p(x)$.

Now set $g(x) = h(x) - ih(ix)$ and check that g works. □

Given a bounded linear functional $f : Z \to \mathbb{F}$ we define $\|f\| = \sup\{|f(z)| : \|z\| \le 1\}$.

Theorem 6.6 (Hahn-Banach Extension Theorem)**.** *Let X be a normed linear space over $\mathbb{F}$, let $Y \subseteq X$ be a subspace, and let $f : Y \to \mathbb{F}$ be a bounded linear functional. Then there exists an extension $g : X \to \mathbb{F}$ that is also a bounded linear functional and satisfies, $\|g\| = \|f\|$.*

Proof. Apply the above result with $p(x) = \|f\| \cdot \|x\|$. □

Corollary 6.7. *Let X be a normed space and let $x_0 \in X$. Then there exists a bounded linear functional $g : X \to \mathbb{F}$ with $\|g\| = 1$ such that $g(x_0) = \|x_0\|$.*

Proof. Let Y be the one dimensional space spanned by x_0 and let $f(\lambda x_0) = \lambda\|x_0\|$. Check that $\|f\| = 1$ and apply the last result. □

6.1. **Krein-Milman Theory.** This result is a consequence of the Hahn-Banach theorem. But its proof uses the version for sublinear functions.

Definition 6.8. Given a real normed space X, a bounded linear functional $f : X \to \mathbb{R}$ and $\alpha \in \mathbb{R}$ we set
$$H_{f,\alpha} := \{y \in X : f(y) \le \alpha\},$$
and we call $H_{f,\alpha}$ a **closed 1/2-space**.

Theorem 6.9 (Krein-Milman)**.** *Let X be a real normed space and let $K \subseteq X$ be a convex subset.*

(1) *If* $\inf\{\|x - y\| : y \in K\} \neq 0$*, then there exists a bounded linear functional* $f : X \to \mathbb{R}$ *and* $\alpha \in \mathbb{R}$ *such that* $f(y) \leq \alpha, \forall y \in K$*, while* $\alpha < f(x_0)$*.*
(2) *Every closed convex set is an intersection of 1/2-spaces.*

7. Dual Spaces

Let X be a normed linear space. We let X^* denote the set of bounded linear functionals on X. It is easy to see that this is a vector space and that $\|f\| = \sup\{|f(x)| : \|x\| \leq 1\}$ is a norm on this space. The normed space $(X^*, \|\cdot\|)$ is called the **dual space of X**.

Other notations used for the dual space are $X', X^d$and sometimes $X^\dagger$.

Proposition 7.1. *Let $(X, \|\cdot\|)$ be a normed space, then $(X^*, \|\cdot\|)$ is a Banach space.*

Here are some examples of dual spaces.

Example 7.2 (The dual of ℓ^1)**.** Given $y \in \ell^\infty$ define $f_Y : \ell^1 \to \mathbb{R}$ by $f_y(x) = \sum_n x_n y_y$, i.e., the dot product! It is not hard to see that f_y is bounded and that $\|f_y\| = \|y\|_\infty$. This gives a map from ℓ^∞ into $(\ell^1)^*$.

It is easy to see that this is a linear map, i.e., that $f_{y_1+y_2} = f_{y_1} + f_{y_2}$ and $f_{\lambda y} = \lambda f_y$. Also we have seen that it satisfies $\|f_y\| = \|y\|_\infty$, i.e., that it is isometric and hence must be one-to-one. Finally, we show that it is onto. To see this start with $f : \ell^1 \to \mathbb{R}$ and let $y_n = f(e_n)$ where e_n is the vector that is 1 in the n-th entry and 0 elsewhere. We have that $|y_n| \leq \|f\|\|e_n\|_1 = \|f\|$, so that $y = (y_1, ...) \in \ell^\infty$. Now check that $f_y(x) = f(x)$ for every x so that $f_y = f$ and we have onto.

The above results are often summarized as saying that $(\ell^1)^* = \ell^\infty$, when what they really mean is that there is the above linear, onto isometry between these two spaces.

Example 7.3 (The dual of ℓ^p, $1 < p < +\infty$)**.** Let $\frac{1}{p} + \frac{1}{q} = 1$ and for each $y \in \ell^q$ define

$$f_y : \ell^p \to \mathbb{R} \text{ by setting } f_y(x) = \sum_{n=1}^{+\infty} y_n x_n.$$

By Minkowski's inequality f_y is a bounded linear functional and $\|f_y\| = \|y\|_q$. Moreover, given any $f : \ell^p \to \mathbb{R}$ if we set $y_n = f(e_n)$ where e_n is the vector with 1 in the n-th coordinate and 0's elsewhere, then it follows that $y = (y_n) \in \ell^q$. This proves that the map $L : \ell^q \to (\ell_p)^*$ given by $L(y) = f_y$ is onto. It is clear that it is linear and by the above it is isometric. Thus, briefly $(\ell^p)^* = \ell^q$.

However, this fails when $p = +\infty$, as we shall show.

Example 7.4 (The space C_0)**.** We set $C_0 = \{x = (x_1, x_2, ...) \in \ell^\infty : \lim_n x_n = 0\}$. It is easy to show that this is a closed subspace. In fact C_0 is the closed linear span of $e_n : n \in \mathbb{N}$. Given a bounded linear functional $f : C_0 \to \mathbb{R}$ if we set $y_n = f(e_n)$, then the same calculations as above show that $y \in \ell^1$ and that $f = f_y$. This leads to $C_0^* = \ell^1$.

Example 7.5 ($(\ell^\infty)^* \neq \ell^1$)**.** Now let $f : \ell^\infty \to \mathbb{R}$ be a bounded linear functional. As before let $y_n = f(e_n)$ and show that $y \in \ell^1$. The functional $f - f_y$ has the property that $(f - f_y)(e_n) = 0$ and hence, $(f - f_y)(C_0) = 0$. But this does not guarantee that $f = f_y$.

In fact, since C_0 is a closed subspace and $C_0 \neq \ell^\infty$ we may form the quotient, ℓ^∞ / C_0 and by the Hahn-Banach theorem there will be plenty of linear functionals of norm 1 on this space. Take any such $g : \ell^\infty / C_0 \to \mathbb{R}$ and since $e_n \in C_0$ we have that $g(e_n + C_0) = 0$. Thus, if we define $f : \ell^\infty \to \mathbb{R}$ by $f(x) = g(x + C_0)$, then f is a norm one bounded linear functional such that $f(e_n) = 0$ for all n.

Thus, not every bounded linear functional on ℓ^∞ is of the form $f_y, y \in \ell^1$. What is true is that every $f \in \ell^\infty$ decomposes uniquely as $f = f_y + h$ where $y \in \ell^1$ and $h(x) = g(x + C_0)$ for some bounded linear functional $g : \ell^\infty / C_0 \to \mathbb{R}$.

7.1. **Banach Generalized Limits.** It is often useful to be able to take limits even when one doesn't exist! This is one reason that we use $\liminf$ and $\limsup$ but these two operations are not linear. For example if we let x_n be 1 for odd n and 0 for even n and let y_n be one for even n and 0 for odd n, then $\limsup_n x_n = \limsup_n y_n = 1 = \limsup(x_n + y_n)$. Banach generalized limits are one way that we can assign "limits" in a linear fashion.

Proposition 7.6 (Banach Generalized Limits)**.** *There exists a bounded linear functional, $G : \ell^\infty \to \mathbb{R}$ such that:*

- $\|G\| = 1$,
- *when $x = (x_1, x_2, ...)$ and $lim_n x_n$ exists, then $G(x) = \lim_n x_n$,*
- $\forall x \in \ell^\infty$, $\liminf_n x_n \leq G(X) \leq \limsup_n x_n$,
- *given K if we let $y_K = (x_{K+1}, x_{K+2}, ...)$, then $G(y_K) = G(x)$, i.e., the generalized limit doesn't care where we start the sequence.*

The existence of such functional follows from the Hahn-Banach theorem and a Banach generalized limit has the property that $G(C_0) = 0$ so these are examples of functionals that do not come from ℓ^1 sequences.

Example 7.7. For those of you familiar with measure theory. Given any $(X, \mathcal{B}, \mu)$ measure space and $1 \leq p < +\infty$, the set $L^p(X, \mathcal{B}, \mu)$ of equivalence classes of p-integrable functions(*essentially bounded functions* in the case $p = +\infty$) is a Banach space in the p-norm.

The dual spaces behave like the ℓ^p spaces with one important exception. In fact, $\ell^p = L^p(\mathbb{N}, 2^{\mathbb{N}}, \mu)$ where μ is counting measure.

Theorem 7.8 (Riesz Representation Theorem)**.** *Let* $1 < p \leq +\infty$ *and* $\frac{1}{p} + \frac{1}{q} = 1$ *and let* $(X, \mathcal{B}, \mu)$ *be a measure space. Then every bounded linear functional* $L : L^p(X, \mathcal{B}, \mu) \to \mathbb{R}$ *is of the form*

$$L(f) = \int_X fg d\mu, \, \exists g \in L^q(X, \mathcal{B}, \mu),$$

and conversely, every g *defines a bounded linear functional via integration. Moreover,* $\|L\| = \|g\|_q$*, briefly,* $L^p(X, \mathcal{B}, \mu)^* = L^q(X, \mathcal{B}, \mu)$*. If, in addition,* μ *is* σ*-finite, then* $L^1(X, \mathcal{B}, \mu)^* = L^\infty(X, \mathcal{B}, \mu)$*.*

There is one more theorem that has the same name.

Theorem 7.9 (Riesz Representation Theorem)**.** *Let* $(K, \mathcal{T})$ *be a compact, Hausdorff space, let* $C(K)$ *denote the space of continuous* $\mathbb{R}$*-valued functions on* K*. If* $L : C(K) \to \mathbb{R}$ *is a bounded linear functional, then there exists a bounded, Borel measure* μ *such that* $L(f) = \int_K f d\mu$ *and conversely, every bounded Borel measure defines a bounded linear functional. Moreover,* $\|L\| = |\mu|(X)$*.*

7.2. **The Double Dual and the Canonical Embedding.** Given a normed linear space, X, we let $X^{**} = (X^*)^*$, i.e., the bounded linear functionals on the space of bounded linear functionals. Given $x \in X$ we define a bounded linear functional $\hat{x} : X^* \to \mathbb{R}$ by setting $\hat{x}(f) = f(x)$, i.e., reversing the roles of dependent and independent variables. Thus, $\hat{x} \in X^{**}$.

The map $J : X \to X^{**}$ defined by $J(x) = \hat{x}$ is easily seen to be linear, one-to-one and an isometry. This map is called the *canonical injection* or the *canonical embedding* of X into X^{**}.

Definition 7.10. A normed space X is called **reflexive** provided that $J(X) = X^{**}$, which is often written simply as $X = X^{**}$.

Note that if X is reflexive, then so is X^*.

The spaces ℓ^p and $L^p(X, \mathcal{B}, \mu)$ are reflexive for $1 < p < +\infty$. The spaces C_0, ℓ^1, ℓ^∞ are not reflexive.

This concept is very important when one discusses the **weak topologies**.

7.3. **The Weak Topology.** Give a normed space X, we say that a subset $U \subseteq X$ is **weakly open** provided that whenever $x_0 \in U$ then there exists a finite set $f_1, ..., f_n \in X^*$ and $\epsilon_1, ..., \epsilon_n > 0$ such that

$$\mathcal{B}_{f,\epsilon} := \{x \in X : |f_i(x - x_0)| < \epsilon, \, \forall 1 \leq i \leq n\} \subseteq U.$$

The set $\mathcal{T}_w$ of all weakly open subsets of X is a topology on X called the **weak topology.**

We say that a net converges **weakly to** x_0 if it converges to x_0 in the weak topology.

Theorem 7.11. *A net* $\{x_\lambda\}_{\lambda \in D} \subseteq X$ *converges in the weak topology to* $x_0 \in X$ *if and only if*

$$\lim_\lambda f(x_\lambda) = f(x_0), \, \forall f \in X^*.$$

We write $x_\lambda \xrightarrow{w} x_0$ to indicate that a net converges weakly to x_0.

7.4. **The Weak* Topology.** Given a normed space X, we define a subset $U \subseteq X^*$ to be **weak* open** provided that for every $f_0 \in U$ there exists a finite set $x_1, ..., x_n \in X$ and $\epsilon_1, ..., \epsilon_n > 0$ such that

$$\{f \in X^* : |f(x_i) - f_0(x_i)| < \epsilon_i, \forall 1 \le i \le n\} \subseteq U.$$

The set $\mathcal{T}_{w*}$ of all weak* open subsets of X^* is a topology on X^* called the **weak* topology**.

A net of functionals converges **weak* to** f_0 provided that it converges in the weak* topology to f_0.

Theorem 7.12. *A net $\{f_\lambda\}_{\lambda \in D}$ converges in the weak* topology to f_0 if and only if*

$$\lim_\lambda f_\lambda(x) = f_0(x), \forall x \in X.$$

We write $f_\lambda \xrightarrow{w*} f_0$ to indicate that a net converges in the weak* topology to f_0.

The following is the most important property of the weak* topology.

Theorem 7.13 (Banach-Alaoglu)**.** *Let X be a normed space and let $K = \{f \in X^* : \|f\| \le 1\}$. Then K is compact in the weak* topology.*

Consequently, any time that we have a bounded net in X^*, we may select a weak* convergent subnet.

Note that when X is reflexive, then regarding $X = (X^*)^*$ we have a weak* topology on X, but this is also the same as the weak topology on X. So weak convergence and weak* convergence mean the same thing.

In particular, in $\ell^p, 1 < p < +\infty$, we see that every bounded net will have a weakly convergent subnet.

7.5. **Completion of Normed Linear Spaces.** We discussed earlier that every metric space (X, ρ) has a canonical complete metric space $(\hat{X}, \hat{\rho})$ that it is isometrically contained in as a dense subset. So if $(X, \|\cdot\|)$ is a normed space then it is contained in a complete metric space, $\hat{X}$. But there is nothing in the earlier theorem that says that $\hat{X}$ can be taken to be a vector space and that the metric $\hat{\rho}$ can be taken to come from a norm. Both of these tings are true and can be shown with some care.

However, the following gives us an easier way to see all of this. Given a normed space $(X, \|\cdot\|)$, we have that $X^{**} = (X^*)^*$ and so it is a Banach space. The map $J : X \to X^{**}$ is a linear, isometric embedding. So if we just take the closure of $J(X)$ inside the Banach space X^{**} we will get a closed subspace and that subspace will be a Banach space. Because the completion of a metric space is unique, this shows that the completion of X actually has the structure of a vector space.

8. Hilbert Spaces

Definition 8.1. Given a vector space V over $\mathbb{F}$ a map, $< | >: V \times V \to \mathbb{F}$ is called an **inner product** provided that it for all vectors $v_1, v_2, w_1, w_2 \in V$ and $\lambda \in \mathbb{F}$, it satisfies:

(1) $\langle v_1 + v_2, w_1 \rangle = \langle v_1, w_1 \rangle + \langle v_2, w_1 \rangle$,
(2) $\langle v_1, w_1 + w_2 \rangle = \langle v_1, w_1 \rangle + \langle v_1, w_2 \rangle$,
(3) $\langle \lambda v_1, w_1 \rangle = \overline{\lambda} \langle v_1, w_1 \rangle$,
(4) $\langle v_1, \lambda w_1 \rangle = \lambda \langle v_1, w_1 \rangle$,
(5) $\langle v_1, w_1 \rangle = \overline{\langle w_1, v_1 \rangle}$,
(6) $\langle v_1, v_1 \rangle \geq 0$,
(7) $\langle v_1, v_1 \rangle = 0 \iff v_1 = 0$.

The pair $(V, \langle | \rangle)$ is called an inner product space. If $\langle | \rangle$ only satisifes 1–6, then it is called a **semi-inner product**.

Some examples of inner products:

Example 8.2. The space $\mathbb{F}^n$ with $\langle (x_1, ..., x_n), (y_1, ..., y_n) \rangle = \sum_i \overline{x_i} y_i$

Example 8.3. The vector space $C([a, b])$ of continuous $\mathbb{F}$-valued functions on $[a, b]$ with

$$\langle f, g \rangle = \int_a^b \overline{f(t)} g(t) dt.$$

Example 8.4. The vector space ℓ^2 with $\langle x, y \rangle = \sum_{n=1}^{\infty} \overline{x_n} y_n$.

Proposition 8.5 (Cauchy-Schwarz Inequality)**.** *Let V be an inner product space then*

$$|\langle v, w \rangle| \leq \sqrt{\langle v, v \rangle} \sqrt{\langle w, w \rangle}.$$

Corollary 8.6. *Let V be an inner product space, then $\|x\| = \sqrt{\langle x, x \rangle}$ defines a norm on V.*

This is called the **norm induced by the inner product**.

Definition 8.7. An inner product space is called a **Hilbert space** if it is complete in the norm induced by the inner product.

Examples 8.2 and 8.4 are Hilbert spaces. To see that 8.3 is not, consider the case of [0,1]. It is enough to consider the sequence

$$f_n(t) = \begin{cases} 0, & 0 \leq t \leq 1/2, \\ n(t - 1/2), & 1/2 \leq t \leq 1/2 + 1/n, \\ 1, & 1/2 + 1/n \leq t \leq 1 \end{cases},$$

which can be shown to be Cauchy but not converge to any continuous function.

Problem 8.8. *Let V be a semi-inner product space. Prove that $\mathcal{N} = \{v : \langle v, v \rangle = 0\}$ is a subspace, that setting $\langle v + \mathcal{N}, w + \mathcal{N} \rangle = \langle v, w \rangle$ is well-defined and gives an inner product on the quotient $V/\mathcal{N}$. (Hint: First show that the Cauchy-Schwarz inequality is still true.)*

Proposition 8.9. *Let V be an inner product space and let $\|\cdot\|$ be the norm induced by the inner product. Then:*

Polarization Identity $\langle x, y\rangle = 1/4\sum_{k=0}^{3} i^k\|x + i^k y\|^2$,
Parallelogram Law $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$.

Theorem 8.10 (Jordan-von Neumann)**.** *Let $(X, \|\cdot\|)$ be a Banach space whose norm satisfies the parallelogram law. Then defining $\langle|\rangle$ via the polarization identity, defines an inner product on X and the norm on X is the norm induced by the inner product. Thus, any Banach space satisfying the parallelogram law is a Hilbert space.*

Corollary 8.11. *Let $(V, \langle|\rangle)$ be an inner product space, then there exists a Hilbert space $\mathcal{H}$ and a one-to-one inner product preserving map $J : V \to \mathcal{H}$ such that $J(V)$ is a dens subspace of $\mathcal{H}$.*

The space $\mathcal{H}$ is called the **Hilbert space completion** of V. For example, the Hilbert space completion of 8.3 is the space $L^2([a, b], \mathcal{M}, \lambda)$ of equivalence classes of measurable square integrable functions with respect to Lebesgue measure.

9. Bases in Hilbert Space

Definition 9.1. Given an inner product space a set of vectors S is called **orthonormal** if $v \in S$ implies $\|v\| = 1$ and whenever $v, w \in S$ and $v \neq w$ then $\langle v, w\rangle = 0$. We will often write $v \perp w \iff \langle v, w\rangle = 0$.

Proposition 9.2 (Pythagoras)**.** *Let V be an inner product space and let $\{x_n : 1 \leq n \leq N\}$ be an otrhonormal set. Then*

$$\|x\|^2 = \sum_{n=1}^{N} |\langle x_n, x\rangle|^2 + \|x - \sum_{n=1}^{N}\langle x_n, x\rangle x_n\|^2.$$

Proposition 9.3 (Bessel's Inequality)**.** *Let $\{x_a : a \in A\}$ be an orthonormal set. Then $\sum_{a\in A} |\langle x_a, x\rangle|^2$ converges and is less than $\|x\|^2$.*

Definition 9.4. We say that an orthonormal set $\{e_a : a \in A\}$ in a Hilbert space $\mathcal{H}$ is **maximal** if there does not exist an orthonormal set that contains it as a proper subset. A maximal orthonormal set is called an **orthonormal basis(o.n.b.)**.

Maximal orthonormal sets exist by Zorn's Lemma(which is equivalent to the Axiom of Choice). The following explains the reason for calling these bases.

Theorem 9.5 (Parseval's Identities)**.** *Let $\mathcal{H}$ be a Hilbert space and let $\{e_a : a \in A\}$ be an orthonormal basis. Then for any $x, y \in \mathcal{H}$,*

(1) $\|x\|^2 = \sum_{a\in A} |\langle e_a, x\rangle|^2$,
(2) $x = \sum_{a\in A}\langle e_a, x\rangle e_a$,
(3) $\langle x, y\rangle = \sum_{a\in A}\langle x, e_a\rangle\langle e_a, y\rangle$.

Problem 9.6. *Prove the 3rd Parseval identity.*

Definition 9.7. Two sets S and T are said to have the same **cardinality** if there is a one-to-one onto function between them and we write **card(S) = card(T)**. When there exists a one-to-one function from S into T we write that **card(S) $\leq$ card(T)**.

It is a fact that $card(S) = card(T) \iff Card(S) \leq card(T)$ and $card(T) \leq card(S)$. For each natural number $\{0, 1, ...\}$ we have a cardinal number. We set $\aleph_0 := card(\mathbb{N})$ and we set $\aleph_1 := card(\mathbb{R})$. The **continuum hypothesis** asks if there can be a set with $card(\mathbb{N}) < card(S) < card(\mathbb{R})$. This statement is known to be independent of the ZFC= Zermelo-Frankl + Axiom of Choice. So one can add as another axiom the statement that this is true or that it is false to one's axiom system.

We call a set A **countably infinite** if $card(A) = card(\mathbb{N})$. We call a set **countable** if it is either finite or countably infinite.

Cardinal arithmetic is defined as follows: Given disjoint sets A and B we set $card(A) + card(B) := card(A \cup B)$ and we set $card(A) \cdot card(B) := card(A \times B)$, where $A \times B = \{(a, b) : a \in A, b \in B\}$ is the Cartesian product.

Theorem 9.8 (Basis Theorem)**.** *Let $\mathcal{H}$ be a Hilbert space and let $\{e_a : a \in A\}$ and $\{f_b : b \in B\}$ each be an orthonormal basis for $\mathcal{H}$. Then $card(A) = card(B)$.*

Definition 9.9. We call the cardinality of any orthonormal basis for $\mathcal{H}$ the **Hilbert space dimension of** $\mathcal{H}$. We denote it by $dim_{HS}(\mathcal{H})$.

Example 9.10. For every cardinal number, there is a Hilbert space of that dimension. Here is a way to see this fact. Given any non-empty set A set

$$\ell^2_A := \{f : A \to \mathbb{F} : \sum_{a \in A} |f(a)|^2 < +\infty\}.$$

On this vector space we define

$$\langle f, g \rangle = \sum_{a \in A} \overline{f(a)} g(a).$$

The fact that this converges and defines an inner product goes much like the proof for ℓ^2. Now define $e_a \in \ell^2_A$ by

$$e_a(t) = \begin{cases} 1, & t = a \\ 0, & t \neq a \end{cases}.$$

Then the set $\{e_a : a \in A\}$ can be shown to be an o.n.b. for ℓ^2_A. Hence, $dim_{HS}(\ell^2_A) = card(A)$.

Proposition 9.11 (Gram-Schmidt)**.** *Let $\{x_n : 1 \leq n \leq N\}$ be a linearly independent set in a Hilbert space. Then there exists an orthonormal set $\{e_n : 1 \leq n \leq N\}$ such that*

$$span\{x_1, ..., x_k\} = span\{e_1, ..., e_k\}, \ \forall k, 1 \leq k \leq N.$$

Proposition 9.12 (Gram-Schmidt). *Let $\{x_n : n \in \mathbb{N}\}$ be a linearly independent set in a Hilbert space. Then there exists an orthonormal set $\{e_n : n \in \mathbb{N}\}$ such that*

$$span\{x_1, ..., x_k\} = span\{e_1, ..., e_k\}, \ \forall k \in \mathbb{N}.$$

One produces this set by the **Gram-Schmidt orthogonalization process**. In the finite case. For the general case one uses the finite case together with the Principle of Recursive Definition.

Recall that a metric space is called **separable** if it contains a countable dense subset.

Theorem 9.13. *Let $\mathcal{H}$ be a Hilbert space. Then $\mathcal{H}$ is separable(as a metric space) if and only if $dim_{HS}(\mathcal{H}) \leq \aleph_0$, i.e., any o.n.b. for $\mathcal{H}$ is countable.*

9.1. **Direct Sums.** Given two vector spaces V and W we can make a new vector space by taking the Cartesian product

$$\{(v, w) : v \in V, w \in W\},$$

and defining $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$ and $\lambda(v, w) = (\lambda v, \lambda w)$. Even though it uses the Cartesian *product*, this space is generally denoted $V \oplus W$. One reason is that in the finite dimensional case if $\{v_1, ..., v_n\}$ is a basis for V and $\{w_1, ..., w_m\}$ is a basis for W, then $\{(v_1, 0), ..., (v_n, 0)\} \cup \{(0, w_1), ..., (0, w_m)\}$ is a basis for $V \oplus W$ so that $dim(V \oplus W) = dim(V) + dim(W)$.

Given Hilbert spaces $(\mathcal{H}_i, \langle , \rangle_i), i = 1, 2$ on their Cartesian product we define

$$\langle (h_1, h_2), (h_1', h_2') \langle = \langle h_1, h_1' \rangle_1 + \langle h_2, h_2' \rangle_2,$$

and it is easy to check that this is an inner product and that the Cartesian product is complete in this inner product. To see this one checks that $(h_{1,n}, h_{2,n})$ is Cauchy if and only if $h_{1,n}$ is Cauchy in $\mathcal{H}_1$ and $h_{2,n}$ is Cauchy in $\mathcal{H}_2$. Hence, there are vectors such that $h_1 = \lim_n h_{1,n}$ and $h_2 = \lim h_{2,n}$. Now one checks that $\|(h_1, h_2) - (h_{1,n}, h_{2,n})\| \to 0$.

This Hilbert space is denoted by $\mathcal{H}_1 \oplus \mathcal{H}_2$. Note that in this space, $(h_1, 0) \perp (0, h_2)$ and that $\|(h_1, 0)\| = \|h_1\|_1$ while $\|(0, h_2)\| = \|h_2\|_2$ so that these subspaces can be identified with the original HIlbert spaces and they are arranged as orthogonal subspaces of $\mathcal{H}_1 \oplus \mathcal{H}_2$.

It is not hard to verify that if $\{e_a : a \in A\}$ is an o.n.b. for $\mathcal{H}_1$ and $\{f_b : b \in B\}$ is an o.n.b. for $\mathcal{H}_2$, then

$$\{(e_a, 0) : a \in A\} \cup \{(0, f_b) : b \in B\},$$

is an o.n.b. for $\mathcal{H}_1 \oplus \mathcal{H}_2$. Thus, $dim_{HS}(\mathcal{H}_1 \oplus \mathcal{H}_2) = dim_{HS}(\mathcal{H}_1) + dim_{HS}(\mathcal{H}_2)$, where the addition is either ordinary in the finite dimensional case or cardinal arithmetic in general.

Note that when we form the direct sum, we can identify $\mathcal{H}_1$ and $\mathcal{H}_2$ canonically as subspaces of the direct sum in the following sense. Namely, if we look at $\widetilde{\mathcal{H}}_1 = \{(h_1, 0) : h_1 \in \mathcal{H}_1\}$ and $\widetilde{\mathcal{H}}_2 = \{(0, h_2) : h_2 \in \mathcal{H}_2\}$, then we have that, as vector spaces these are isomorphic to the original spaces and for any $h_1, h_1' \in \mathcal{H}_1$,

$$\langle h_1 | h_1' \rangle_{\mathcal{H}_1} = \langle (h_1, 0) | (h_1', 0) \rangle_{\mathcal{H}_1 \oplus \mathcal{H}_2},$$

and similarly for any pair of vectors in $\mathcal{H}_2$. We also have that

$$\widetilde{\mathcal{H}_1}^{\perp} = \widetilde{\mathcal{H}_2}.$$

We now generalize this construction to arbitrary collections. Given a family of Hilbert spaces $(\mathcal{H}_a, \langle , \rangle_a)$, $a \in A$. We define a new Hilbert space which is a subset of the Cartesian product via

$$\oplus_{a \in A} \mathcal{H}_a = \{(h_a)_{a \in A} : \sum_{a \in A} \|h_a\|_a^2 < +\infty\},$$

and inner product

$$\langle (h_a), (k_a) \rangle = \sum_{a \in A} \langle h_a, k_a \rangle_a.$$

Proofs similar to the case of ℓ^2_A show that this is indeed a Hilbert space.

9.2. **Bilinear Maps and Tensor Products.** Given vector spaces V, W, Z a map $B : V \times W \to Z$ is called **bilinear** provided that it is linear in each variable, i.e.,

- $B(v_1 + v_2, w) = B(v_1, w) + B(v_2, w)$,
- $B(v, w_1 + w_2) = B(v, w_1) + B(v, w_2)$
- $B(\lambda v, w) = \lambda B(v, w) = B(v, \lambda w)$.

Tensor products were created to linearize bilinear maps. This concept takes some development but, in summary, given two vector spaces their tensor product,

$$V \otimes W = span\{v \otimes w : v \in V, w \in W\},$$

where $v \otimes w$ is called an **elementary tensor** and the rules for adding these mimic the bilinear rules:

- $v_1 \otimes w + v_2 \otimes w = (v_1 + v_2) \otimes w$,
- $v \otimes w_1 + v \otimes w_2 = v \otimes (w_1 + w_2)$,
- $\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w)$.

If V and W have bases $\{v_a : a \in A\}$ and $\{w_b : b \in B\}$, respectively, then $\{v_a \otimes w_b : a \in A, b \in B\}$ is a basis for $V \otimes W$. Thus, the vector space dimension satisfies

$$dim(V \otimes W) = dim(V) \cdot dim(W),$$

where this is just the product of integers in the finite dimensional case and is true in the sense of cardinal arithmetic in the general case.

Proposition 9.14. *There is a one-to-one correspondence between bilinear maps, $B : V \times W \to Z$ and linear maps $T : V \otimes W \to Z$ given by $T(v \otimes w) = B(v, w)$.*

When V and W are both normed spaces, we would like the tensor product to be a normed space too. There are several "natural" norms to put on the tensor product and on bilinear maps.

Definition 9.15. Let X, Y and Z be normed spaces and let $B : X \times Y$ be a bilinear map. Then we say that B is **bounded** provided that

$$\|B\| = \sup\{||B(x, y)|| : \|x\| \leq 1, \|y\| \leq 1\} < +\infty,$$

and we call this quantity the **norm of B**. For $u \in X \otimes Y$ the **projective norm** of u is given by

$$\|u\|_{\vee} = \inf\{\sum \|x_i\| \cdot \|y_i\| : u = \sum_i x_i \otimes y_i\},$$

where the infimum is taken over all ways to express u as a finite sum of elementary tensors. The completion of $X \otimes Y$ in this norm is denoted $X \otimes^{\vee} Y$ and is called the **projective tensor product of X and Y**.

Here is the relationship these two concepts.

Proposition 9.16. *Let X, Y and Z be Banach spaces. Then $B : X \times Y \to Z$ is bounded if and only if the linear map $T_B : X \otimes Y \to Z$ is bounded when $X \otimes Y$ is given the projective norm. In this case $\|B\| = \|T_B\|$ and T_B can be extended to a bounded linear map of the same norm from $X \otimes_{\vee} Y$ into Z.*

The extension of T_B to the completion is, generally, still denoted by T_B.

There is one other tensor norm that is very important. Note that if $f : X \to \mathbb{F}$ and $g : Y \to \mathbb{F}$ are linear, then there is a linear map $f \otimes g : X \otimes Y \to \mathbb{F}$ given by $f \otimes g(x \otimes y) = f(x) \cdot g(y)$. To see this just check that $B(x, y) := f(x)g(y)$ is bilinear.

Definition 9.17. Let X and Y be normed spaces. Given $u \in X \otimes Y$, the **injective norm of u** is given by the formula

$$\|u\|_{\wedge} := \sup\{|f \otimes g(u)| : f \in X^*, g \in Y^*, \|f\| \leq 1, \|g\| \leq 1\}.$$

This is indeed a norm on $X \otimes Y$ and the completion in this norm is denoted by $X \otimes_{\wedge} Y$ and is called the **injective tensor product of X and Y**.

Grothendieck was the first to systematically study and classify tensor norms. He defined a **reasonableness condition** for tensor norms and proved that if X and Y are normed spaces, then any reasonable tensor norm on $X \otimes Y$ satisfies, $\|u\|_{\wedge} \leq \|u\| \leq \|u\|_{\vee}$.

.

9.3. **Tensor Products of Hilbert Spaces.** All of the above discussion is preliminary to defining a norm on the tensor product of Hilbert spaces that yields a new Hilbert space.

Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. We wish to define an inner product on $\mathcal{H} \otimes \mathcal{K}$. Given $u = \sum_i h_i \otimes k_i$ and $v = \sum_j v_j \otimes w_j$ two finite sums representing vectors in $\mathcal{H} \otimes \mathcal{K}$, we set

$$\langle u, v\rangle = \sum_{i,j} \langle h_i | v_j\rangle_{\mathcal{H}} \langle k_i | w_j\rangle_{\mathcal{K}}.$$

It is not hard to see that if this formula is well-defined, i.e., does not depend on the particular way that we write u and v as a sum of elementary tensors, then it is sesquilinear.

Proposition 9.18. *The above formula is well-defined and gives an inner product on $\mathcal{H} \otimes \mathcal{K}$.*

Proof. We omit the details of showing that it is well-defined. As noted above, given that it is well-defined it is elementary that it is sesquilinear. So we only prove that it satisfies $u \neq 0 \implies \langle u, u\rangle > 0$.

To this end let $u = \sum_{i=1}^{n} h_i \otimes k_i$. Let $\mathcal{F} = span\{k_i : 1 \leq i \leq n\}$ and let $\{e_j : 1 \leq j \leq m\}$ be an o.n.b. for $\mathcal{F}$ so that $m \leq n$. Write each $k_i = \sum_j a_{i,j} e_j$. Then we have that

$$u = \sum_i \sum_j h_i \otimes (a_{i,j} e_j) = \sum_j \sum_i (a_{i,j} h_i) \otimes e_j = \sum_j v_j \otimes e_j,$$

where $v_j = \sum_i a_{i,j} h_i$. Now since $u \neq 0$ we must have that some $v_j \neq 0$. Using the fact that the inner product is independent of the way that we express u as a sum of elementary tensors, we have that

$$\langle u | u\rangle = \sum_{i,j} \langle v_i | v_j\rangle_{\mathcal{H}} \langle e_i | e_j\rangle_{\mathcal{K}} = \sum_i \|v_i\|^2 \neq 0,$$

and we are done. □

Definition 9.19. Given Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ as above, we let $\mathcal{H} \overline{\otimes} \mathcal{K}$ denote the completion of $\mathcal{H} \otimes \mathcal{K}$ in the above inner product and call it the **Hilbert space tensor product**.

Many authors use $\mathcal{H} \otimes \mathcal{K}$ to denote both the vector space tensor product AND its completion, or they sometimes invent a non-standard notation to denote the vector space tensor product, like $\mathcal{H} \odot \mathcal{K}$, reserving $\mathcal{H} \otimes \mathcal{K}$ for the completion.

Proposition 9.20. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces, let $\{e_a : a \in A\}$ be an o.n.b. for $\mathcal{H}$ and $\{f_b : b \in B\}$ be an o.n.b. for $\mathcal{K}$. Then*

$$\{e_a \otimes f_b : a \in A, \, b \in B\}.$$

is an o.n.b. for $\mathcal{H} \overline{\otimes} \mathcal{K}$. Hence,

$$dim_{HS}(\mathcal{H} \overline{\otimes} \mathcal{K}) = dim_{HS}(\mathcal{H}) \cdot dim_{HS}(\mathcal{K}).$$

In a similar way one can form a Hilbert space tensor product of any finite collection of Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_n$, denoted $\mathcal{H}_1 \overline{\otimes} \cdots \overline{\otimes} \mathcal{H}_n$. The inner product on any two elementary tensors is given by

$$\langle v_1 \otimes \cdots \otimes v_n | w_1 \otimes \cdots \otimes w_n\rangle = \langle v_1 | w_1\rangle_{\mathcal{H}_1} \cdots \langle v_n | w_n\rangle_{\mathcal{H}_n},$$

and then is extended linearly to finite sums of elementary tensor products.

Unlike for the direct sum construction, there is no natural way that we can regard $\mathcal{H}$ and $\mathcal{K}$ as subspaces of $\mathcal{H}\overline{\otimes}\mathcal{K}$. In particular, $\{h \otimes 0 : h \in \mathcal{H}\} = (0 \otimes 0)$. One way that we can regard $\mathcal{H}$ as a subspace of the tensor product is to pick a unit vector $k \in \mathcal{K}$. Then we have that $\widetilde{\mathcal{H}} := \{h \otimes k : h \in \mathcal{H}\}$ is isomorphic to $\mathcal{H}$ as a vector space, since $(h_1 + h_2) \otimes k = h_1 \otimes k + h_2 \otimes k$ and $(\lambda h) \otimes k = \lambda(h \otimes k)$ and has the property that for any $h_1, h_2 \in \mathcal{H}$,

$$\langle h_1|h_2\rangle_{\mathcal{H}} = \langle h_1 \otimes k|h_2 \otimes k\rangle_{\mathcal{H}\otimes\mathcal{K}}.$$

Similarly, fixing a unit vector in $h \in \mathcal{H}$ allows us to include $\mathcal{K}$ into the tensor product as $\widetilde{\mathcal{K}} := \{h \otimes k : k \in \mathcal{K}\}$.

However, unlike the direct sum these two subspaces are not orthogonal, in fact, their intersection is the one-dimensional subspace spanned by $h \otimes k$.

9.4. **Infinite Tensor Products.** Many physical models call for tensor products of infinitely many Hilbert spaces. This arises for example where one has an infinite lattice of electrons. Each electron has a state space that is a Hilbert space $\mathcal{H}_a$ and the models say that the state space of the whole ensemble should be the tensor product over all electrons.

The first problem that one encounters is that if one has two elementary tensor products of infinitely many vectors, $v_1 \otimes v_2 \otimes \cdots$ and $w_1 \otimes w_2 \otimes \cdots$, then one would like to set

$$\langle v_1 \otimes v_2 \otimes \cdots |w_1 \otimes w_2 \otimes \cdots\rangle = \prod_{i=1}^{\infty}\langle v_i|w_i\rangle,$$

but such an infinite product need not converge. Second, if we had a set of unit vectors $v_1, \ldots$ and we sat $u = v_1 \otimes (v_2/2) \otimes (v_3/3) \otimes \cdots$, then we would have that

$$\|u\|^2 = \langle u|u\rangle = \prod_{i=1}^{\infty}\frac{1}{i^2} = 0.$$

However, in physics one usually has a lot more. Each electron has a distinguished unit vector in its state space called the *ground state* and it turns out that given Hilbert spaces $\mathcal{H}_a$ and distinguished unit vectors $\psi_a \in \mathcal{H}_a$, then this is enough data to form an infinite tensor product.

The key is, remember that when we have a unit vector then we can "include" $\mathcal{H}$ into $\mathcal{H} \otimes \mathcal{K}$ by sending $h \to h \otimes k$, where k is some fixed unit vector.

We form the infinite tensor product of a collection $(\mathcal{H}_a, \psi_a)$, $a \in A$, where $\psi_a \in \mathcal{H}_a$ is a unit vector as described below.

For each finite subset $F \subseteq A$ there is a well-defined tensor product Hilbert space of these finitely many spaces, which we denote by $\mathcal{H}_F := \otimes_{a\in F}\mathcal{H}_a$(we do not need to complete at this time).

Now given two finite sets $F_1 \subseteq F_2 \subseteq A$, say $F_1 = \{a_1, \ldots, a_n\}$ and $F_2 = \{a_1, \ldots, a_n, \ldots a_m\}$ with $m > n$, we can regard $\mathcal{H}_{F_1}$ as a subspace of $\mathcal{H}_{F_2}$ by identfying $u \in \mathcal{H}_{F_1}$ with $u \otimes \psi_{a_{n+1}} \otimes \cdots \otimes \psi_{a_m} \in \mathcal{H}_{F_2}$. Formally, this means that we have a linear map, $V_{F_1,F_2} : \mathcal{H}_{F_1} \to \mathcal{H}_{F_2}$.

Now take the union over all finite subsets, $\cup_{F \subseteq A} \mathcal{H}_F$, and define a relation on this union by declaring $u_1 \in \mathcal{H}_{F_1}$ **equivalent** to $u_2 \in \mathcal{H}_{F_2}$ provided that there exists a finite set F_3 with $F_1 \subseteq F_3$ and $F_2 \subseteq F_3$ such that $V_{F_1,F_3}(u_1) = V_{F_2,F_3}(u_2)$(we will write $u_1 \sim u_2$ for this relation). It is not hard to show that this is an equivalence relation and we will write $[u]$ for the equivalence class of a vector.

It is not hard to show that $\mathcal{K} = \big(\cup_{F \subseteq A} \mathcal{H}_F \big) / \sim = \{[u] : \exists F, u \in \mathcal{H}_F\}$ is an inner product space with operations defined as follows. Given $\lambda \in \mathbb{F}$ and $[u]$ set $\lambda[u] = [\lambda u]$. Given $[u_1], [u_2]$ with $u_1 \in \mathcal{H}_{F_1}$ and $u_2 \in \mathcal{H}_{F_2}$ pick F_3 with $F_1, F_2 \subseteq F_3$ and set

$$[u_1] + [u_2] = [V_{F_1,F_3}(u_1) + V_{F_2,F_3}(u_2)].$$

Finally, the inner product is defined by

$$\langle [u_1] | [u_2] \rangle_{\mathcal{K}} = \langle V_{F_1,F_3}(u_1) | V_{F_2,F_3}(u_2) \rangle_{\mathcal{H}_{F_3}}.$$

Showing that all of these operations are well-defined, that is, only depend on the equivalence class, not on the particular choices, and that the last formula is an inner product is tedious, but does all work. The completion of the space $\mathcal{K}$ is denoted

$$\overline{\otimes}_{a \in A} (\mathcal{H}_a, \psi_a).$$

In the end this tensor product does not depend on the choice of vectors, in the following sense. If we also choose unit vectors $\phi_a \in \mathcal{H}_a$, then for each Hilbert space there is a unitary map $U_a : \mathcal{H}_a \to \mathcal{H}_a$ such that $U_a(\psi_a) = \phi_a$(we will discuss unitaries in a later section) and these define unitaries $U_F : \mathcal{H}_F \to \mathcal{H}_F$ via $U_F(h_{a_1} \otimes \cdots \otimes h_{a_n}) = U_{a_1}(h_{a_1}) \otimes \cdots \otimes U_{a_n}(h_{a_n})$. Together these define a unitary U on $\mathcal{K}$ by for $u \in \mathcal{H}_F$ setting $U([u]) = [U_F(u)]$. Then this unitary satisfies $U([\psi_{a_1} \otimes \cdots \psi_{a_n}]) = [\phi_{a_1} \otimes \cdots \otimes \phi_{a_n}]$, and so induces a unitary on the completions, $U : \overline{\otimes}_{a \in A} (\mathcal{H}_a, \psi_a) \to \overline{\otimes}_{a \in A} (\mathcal{H}_a, \phi_a)$.

10. Operators on Hilbert Space

Definition 10.1. Let $\mathcal{H}$ be a Hilbert space. We write $h \perp k \iff \langle h\, k \rangle = 0$. Let $S \subseteq \mathcal{H}$ be a subset, then we set

$$S^{\perp} = \{h : \langle h | k \rangle = 0, \forall k \in S\}.$$

It is easy to see that $S^{\perp}$ is a closed subspace of $\mathcal{H}$.

Proposition 10.2. *Let $\mathcal{M} \subseteq \mathcal{H}$ be a closed subspace. Then every vector $x \in \mathcal{H}$ has a decomposition as $x = h + k$ with $h \in \mathcal{M}$ and $k \in \mathcal{M}^{\perp}$ and this decomposition is unique.*

Proof. First uniqueness. Suppose that $x = h_1 + k_1 = h_2 + k_2$ with $h_1, h_2 \in \mathcal{M}$ and $k_1, k_2 \in \mathcal{M}^{\perp}$. Subtracting we have $h_1 - h_2 = k_2 - k_1 := v$ and since these are both subspaces, we have that $v \in \mathcal{M} \cap \mathcal{M}^{\perp}$. But this implies that $\langle v | v \rangle = 0$ and so $v = 0$ which implies the uniqueness.

Now we show existence. Since $\mathcal{M}$ is itself a HIlbert space, it has an o.n.b. Let $\{e_a; a \in A\}$ be an o.n.b. for $\mathcal{M}$ and set

$$P(x) = \sum_{a \in A} \langle e_a | x \rangle e_a.$$

By Bessel's inequality $\sum_{a \in A} |\langle e_a | x \rangle|^2$ converges and this is enough to show that the sum for $P(x)$ converges and that $\|P(x)\| \le \|x\|$. Hence, the formula defines a bounded linear operator.

Next note that $\langle x - P(x) | e_a \rangle = 0$ for every $a \in A$. Hence, $x - P(x) \in \mathcal{M}^\perp$. So that $x = P(x) + (x - P(x))$ is the decomposition. □

Corollary 10.3. *Let $\{e_a : a \in A\}$ and $\{f_b : b \in B\}$ be two o.n.b.'s for $\mathcal{M}$. Then for every $x \in \mathcal{H}$, we have that $\sum_{a \in A} \langle e_a | x \rangle e_a = \sum_{b \in B} \langle f_b | x \rangle f_b$.*

Proof. This follows from the uniqueness of the decomposition. □

Definition 10.4. We call the map P given above the **orthogonal projection of $\mathcal{H}$ onto $\mathcal{M}$**.

Corollary 10.5. *Let $\mathcal{H}$ be a HIlbert space and $\mathcal{M} \subseteq \mathcal{H}$ a subspace. If $\{e_a : a \in A\}$ is an o.n.b. for $\mathcal{M}$ then there exists an orhtonormal set $\{f_b : b \in B\}$ such that $\{e_a : a \in A\} \cup \{f_b : b \in B\}$ is an o.n.b. for $\mathcal{H}$.*

Proof. Let $\{f_b : b \in B\}$ be an o.n.b. for $\mathcal{M}^\perp$. □

The above result is often summarized by saying that every o.n.b. for $\mathcal{M}$ can be **extended** to an o.n.b. for $\mathcal{H}$.

10.1. **The dual of a Hilbert Space.** Let $\mathcal{H}$ be a Hilbert space and let $h \in \mathcal{H}$ then the map $f_h : \mathcal{H} \to \mathbb{F}$ defined by

$$f_h(k) = \langle h \; k \rangle,$$

is linear and $|f_h(k)| \le \|h\| \|k\|$ so it is a bounded linear functional with $\|f_h\| \le \|h\|$. Since

$$\|h\|^2 = f_h(h) \le \|f_h\| \|h,$$

we see that $\|f_h\| = \|h\|$. We now show that every bounded linear functional is of this form.

Theorem 10.6. *Let $\mathcal{H}$ be a HIlbert space and let $f : \mathcal{H} \to \mathbb{F}$ be a bounded linear functional. Then there exists a vector $h \in \mathcal{H}$ such that $f = f_h$. Consequently, the map*

$$L : \mathcal{H} \to \mathcal{H}^*, \; L(h) = f_h,$$

is a one-to-one isometric map onto the dual space that is linear in the $\mathbb{R}$ case and conjugate linear in the $\mathbb{C}$ case.

Proof. The case of the 0 functional is trivia, so assume tht $f \neq 0$. Let $\mathcal{K} = ker(f)$. Since f is continuous this is a closed subspace of $\mathcal{H}$ and is not all of $\mathcal{H}$. Hence, $\mathcal{K}^\perp \neq (0)$. Pick $k_0 \in \mathcal{K}^\perp$ with $f(k_0) = 1$. Then for every $h \in \mathcal{H}$, we have that $f(h - f(h)k_0) = 0$ so that $h - f(h)k_0 \in \mathcal{K}$.

Hence, this vector is orthogonal to k_0 and

$$0 = \langle k_0 | h - f(h) k_0 \rangle = \langle k_0 | h \rangle - f(h) \|k_0\|^2.$$

Thus, $f(h) = \langle \frac{k_0}{\|k_0\|^2} | h \rangle$ and so $f = f_{k_1}$ with $k_1 = \frac{k_0}{\|k_0\|^2}$.

Finally, note that $L(h_1 + h_2) = f_{h_1+h_2} = f_{h_1} + f_{h_2} = L(h_1) + L(h_2)$, while $L(\lambda h) = f_{\lambda h} = \overline{\lambda} f_h = \overline{\lambda} L(h)$. □

10.2. **The Hilbert Space Adjoint.** Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $T : \mathcal{H} \to \mathcal{K}$ be a bounded linear map. Then for each $k \in \mathcal{K}$ the map from $\mathcal{H}$ to $\mathbb{F}$ given by

$$h \to g_k(h) = \langle k | T(h) \rangle_{\mathcal{K}},$$

is easily seen to be linear and it is bounded since,

$$|\langle k | T(h) \rangle| \leq \|k\| \|T(h)\| \leq (\|k\| \|T\|) \|h\|.$$

Thus, it is a bounded linear functional and hence there is a unique vector, denoted $T^*(k) \in \mathcal{H}$ such that

$$\langle T^*(k) | h \rangle_{\mathcal{H}} = \langle k | T(h) \rangle_{\mathcal{K}}.$$

This last equation is known as the **adjoint equation**.

It is easy to check that the map $T^* : \mathcal{K} \to \mathcal{H}$ is linear. Also, $\|T^*(k)\| = \|g_k\| \leq (\|k\| \|T\|)$ so that T^* is bounded with $\|T^*\| \leq \|T\|$. Equality follows since

$$\|T^*\| = \sup\{|\langle T^*(k) | h \rangle| : \|k\| \leq 1, \|h\| \leq 1\} = \sup\{|\langle k | T(h) \rangle| : \|k\| \leq 1, \|h\| \leq 1\} = \|T\|.$$

Here are some basic properties:

- $(T_1 + T_2)^* = T_1^* + T_2^*$,
- $(\lambda T)^* = \overline{\lambda} T^*$,
- $(T_1 T_2)^* = T_2^* T_1^*$,
- $(T^*)^* = T$,
- if $T = (t_{i,j}) : \mathbb{C}^m \to \mathbb{C}^n$ is a $n \times m$ matrix and we regard $\mathbb{C}^k$ as a Hilbert space in the usual way, then $T^* = (b_{i,j})$ where $b_{i,j} = \overline{t}_{j,i}$, i.e., T^* is the conjugate transpose.

If $T : \mathcal{H} \to \mathcal{K}$, we let

$$\mathcal{R}(T) = \{T(x) : x \in \mathcal{H}\} \subseteq \mathcal{K}$$

denote the **range of** T, and let

$$\mathcal{N}(T) = \{x \in \mathcal{H} : T(x) = 0\} \subseteq \mathcal{H},$$

denote the **kernel of T** also called the **nullspace of T**.

Proposition 10.7. *$\mathcal{R}(T)^\perp = \mathcal{N}(T^*)$ and $\mathcal{R}(T)^- = \mathcal{N}(T^*)^\perp$.*

10.3. **The Banach Space Adjoint.** Operators on normed spaces also have adjoints but there is a bit of a difference. Given X, Y normed spaces and $T : X \to Y$ a bounded linear map, for each $f \in Y^*$–the dual space, we have that $f \circ T : X \to \mathbb{F}$ is a bounded linear functional, i.e., $f \circ T \in X^*$.

We define $T^* : X^* \to Y^*$ by $T^*(f) = T \circ f$.

We have that:

- $(T_1 + T_2)^* = T_1^* + T_2^*$,
- $(\lambda T)^* = \lambda T^*$,
- $(T_1 T_2)^* = T_2^* T_1^*$,
- if $T = (t_{i,j}) : \mathbb{F}^m \to \mathbb{F}^n$ is a matrix, both spaces have a norm, and we identify the dual spaces as $\mathbb{F}^m$ and $\mathbb{F}^n$ with the inner product pairing and a dual norm, then $T^* = (b_{i,j})$ where $b_{i,j} = t_{j,i}$, i.e., T^* is the transpose.

Finally, if $T : X \to Y$, then $T^{**} : X^{**} \to Y^{**}$ and generally $X^{**} \neq X$ and $Y^{**} \neq Y$ so $T^{**} \neq T$. However, we do have that $T^{**}(\hat{x}) = \widehat{T(x)}$ where $\hat{x} \in X^{**}$ is the functional defined by $\hat{x}(f) = f(x)$.

11. Some Important Classes of Operators

In this section we introduce some of the important types of operators that we will encounter later. One common theme is to find spatial characterizations of operators that are given by an algebraic characterization.

Definition 11.1. A bounded linear operator $P : \mathcal{H} \to \mathcal{H}$ is called an **idempotent** if $P^2 = P$.

Proposition 11.2. *If P is an idempotent then $\mathcal{R}(P)$ is a closed subspace and every $h \in \mathcal{H}$ decomposes uniquely as $h = h_1 + h_2$ with $h_1 \in \mathcal{R}(P)$ and $Ph_2 = 0$.*

Proof. If $Ph_n \in \mathcal{R}(P)$ and $P(h_n) \to k$, then $P^2(h_n) \to P(k)$, but $P^2(h_n) = P(h_n)$ so $k = P(k)$, so k is also in the range of P.

Now $h = Ph + (I - P)h$ with $h_1 = Ph \in \mathcal{R}(P)$ and $P(I - P)h = (P - P^2)h = 0$ so $h_2 = (I - P)h \in \mathcal{N}(P)$.

To see uniqueness, if $h = h_1 + h_2 = k_1 + k_2$ is two ways to decompose h as a sum, then $h_1 - k_1 = k_2 - h_2$ with $h_1 - k_1 \in \mathcal{R}(P)$ and $k_2 - h_2 \in \mathcal{N}(P)$. This implies that

$$h_1 - k_1 = P(h_1 - k_1) = P(k_2 - h_2) = 0,$$

so that $h_1 = k_1$ and similarly, $h_2 = k_2$. □

In general, we need not have that $\mathcal{R}(P)$ and $\mathcal{N}(P)$ are orthogonal when P is idempotent.

A good example is to consider $P = \begin{pmatrix} 1 & t \\ 0 & 0 \end{pmatrix}$ then $\mathcal{R}(P) = span\{e_1\}$ while $\mathcal{N}(P) = span\{-te_1 + e_2\}$ so that these two subspaces are orthogonal iff $t = 0$.

Proposition 11.3. *If $P \in B(\mathcal{H})$ and $P = P^2 = P^*$ then P is the orthogonal projection onto its range.*

Proof. By the above every vector decomposes uniquely as $h = h_1 + h_2$ with $h_1 \in \mathcal{R}(P)$ and $h_2 \in \mathcal{N}(P)$. But

$$\langle h_1 | h_2 \rangle = \langle Ph_1 | h_2 \rangle = \langle h_1 | P^* h_2 \rangle = \langle h_1 | Ph_2 \rangle = 0,$$

since $h_2 \in \mathcal{N}(P)$. Hence, $h_1 \perp h_2$. □

By a **projection** we will always mean an operator P such that $P = P^2 = P^*$, i.e., an orthogonal projection.

Definition 11.4. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. A map $V : \mathcal{H} \to \mathcal{K}$ is called an **isometry** iff $\|Vh\| = \|h\|, \forall h \in \mathcal{H}$. A map $U : \mathcal{H} \to \mathcal{K}$ is called a **unitary** if it is an onto isometry.

Note that an isometry is always one-to-one since $h \in \mathcal{N}(V) \iff \|Vh\| = 0 \iff \|h\| = 0 \iff h = 0$. Also the range of an isometry is easily seen to be closed.

The **unilateral shift** $S : \ell^2 \to \ell^2$ given by $S((x_1, x_2, \ldots)) = (0, x_1, x_2, \ldots)$ is easily seen to be an isometry and is not onto.

Proposition 11.5. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $V : \mathcal{H} \to \mathcal{K}$ be linear. The following are equivalent:*

(1) *V is an isometry,*
(2) *$\forall h_1, h_2 \in \mathcal{H}$, $\langle Vh_1 | Vh_2 \rangle_{\mathcal{K}} = \langle h_1 | h_2 \rangle_{\mathcal{H}}$,(this property is called* **inner product preserving***)*
(3) *$V^*V = I_{\mathcal{H}}$.*

Problem 11.6. *Prove the above proposition.*

Proposition 11.7. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $U : \mathcal{H} \to \mathcal{K}$. The following are equivalent:*

(1) *U is a unitary,*
(2) *U is an invertible siometry,*
(3) *U is invertible and $U^{-1} = U^*$,*
(4) *U and U^* are both isometries.*

Definition 11.8. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. A map $W : \mathcal{H} \to \mathcal{K}$ is called a **partial isometry** provided that $\|Wh\| = \|h\|$, $\forall h \in \mathcal{N}(W)^\perp$. We shall call $\mathcal{N}(W)^\perp$ the **initial space of W** and $\mathcal{R}(W)$ the **final space of W**.

Proposition 11.9. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $W \in B(\mathcal{H}, \mathcal{K})$.The following are equivalent:*

(1) *W is a partial isometry,*
(2) *W^*W is a projection,*
(3) *WW^* is a projection,*
(4) *W^* is a partial isometry.*

Moreover, we have that W^*W is the projection onto the initial space of W and WW^* is the projection onto the final space of W.

Definition 11.10. Let $\mathcal{H}$ be a Hilbert space. A bounded linear map $H : \mathcal{H} \to \mathcal{H}$ is called **self-adjoint** or **Hermitian** if $H = H^*$.

Proposition 11.11 (Cartesian Decomposition)**.** *Let $\mathcal{H}$ be a Hilbert space and let $T \in B(\mathcal{H})$. Then there are unique self-adjoint operators, H, K such that $T = H + iK$.*

Proof. Let

$$H = \frac{T + T^*}{2} \text{ and } K = \frac{T - T^*}{2i},$$

then $H = H^*, K = K^*$ and $T = H + iK$. If we also had $T = H_1 + iK_1$, then $H_1 = \frac{T+T^*}{2} = H$ and $K_1 = \frac{T-T^*}{2i}$ and so uniqueness follows. □

The notation $Re(T) = \frac{T+T^*}{2}$ and $Im(T) = \frac{T-T^*}{2i}$ is often used.

Definition 11.12. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. A bounded linear map $T : \mathcal{H} \to \mathcal{K}$ is called **finite rank** provided that $dim(\mathcal{R}(T))$ is finite.

Proposition 11.13. *T is finite rank if and only if for some n there are vectors $h_1, ..., h_n \in \mathcal{H}$ and $k_1, ..., k_n \in \mathcal{K}$ such that*

$$T(h) = \sum_{i=1}^{n} \langle h_i | h \rangle k_i.$$

Proof. If T has this fomr then $\mathcal{R}(T) \subseteq span\{k_1, ..., k_n\}$ so that T is finite rank.

Conversely, assume that T is finite rank and choose an orthonormal basis $\{k_1, ..., k_n\}$ for $\mathcal{R}(T)$. By Parseval,

$$Th = \sum_{i=1}^{n} \langle k_i | Th \rangle k_i = \sum_{i=1}^{n} \langle T^* k_i | h \rangle k_i,$$

so we have the desired form with $h_i = T^* k_i$. □

In physics the above operator is often denoted

$$T = \sum_{i=1}^{n} |k_i\rangle\langle h_i|.$$

Definition 11.14. Let $\mathcal{H}$ and $\mathcal{K}$ be HiIbert spaces. A bounded linear map $K : \mathcal{H} \to \mathcal{K}$ is called **compact** if it is in the closure of the finite rank operators. We let $\mathbb{K}(\mathcal{H}, \mathcal{K})$ denote the set of compact operators.

It is not hard to see that $\mathbb{K}(\mathcal{H}, \mathcal{K})$ is a closed linear subspace of $B(\mathcal{H}, \mathcal{K})$. We will return to the compact operators later.

12. Spectral Theory

For a finite square matrix, its eigenvalues, which are characterized as the roots of the **characteristic polynomial**, $p_A(t) = det(tI - A)$, play a special role in understanding the matrix. This polynomial need not factor over $\mathbb{R}$ which is why we generally assume that our spaces are complex when studying eigenvalues.

The spectrum is the set that plays a similar role in infinite dimensions. For similar reasons we will from now on assume that our Hilbert space is over $\mathbb{C}$.

Definition 12.1. Let $T \in B(\mathcal{H})$. Then the **spectrum of T** is the set

$$\sigma(T) = \{z \in \mathbb{C} : (zI - T) \text{ is not invertible}\},$$

the complement of this set is called the **resolvent of T** and is denoted $\rho(T)$.

Theorem 12.2 (Neumann Series). *Let $\mathcal{H}$ be a Hilbert space and let $A, B, X, T \in B(\mathcal{H})$.*

1. *If $\|X\| < 1$, then $I - X$ is invertible and the series $\sum_{n=0}^{\infty} X^n$ converges in norm to $(I - X)^{-1}$.*
2. *If A is invertible and $\|A - B\| < \|A^{-1}\|^{-1}$, then B is invertible.*
3. *If $|z| > \|T\|$, then $(zI - T)$ is invertible.*

Theorem 12.3. *Let $T \in B(\mathcal{H})$, then $\sigma(T)$ is a non-empty compact set.*

Proof. By (3), $\sigma(T) \subseteq \{z : |z| \le \|T\|\}$, so $\sigma(T)$ is bounded. By (2), the set of invertible operators is open, so $\rho(T)$ is open and so $\sigma(T)$ is closed. Thus, $\sigma(T)$ is compact.

The deepest part is that $\sigma(T)$ is non-empty. This uses some non-trivial complex analysis. Suppose that $\sigma(T)$ is empty. Fix $h, k \in \mathcal{H}$ and define $f_{h,k} : \mathbb{C} \to \mathbb{C}$, by

$$f_{h,k}(z) = \langle h | (zI - T)^{-1} k \rangle.$$

One shows that this function is analytic on all of $\mathbb{C}$ and uses the Neumann series to show that as $|z| \to +\infty$ we have $|f_{h,k}(z)| \to 0$. This in turn implies that $f_{h,k}$ is bounded. A result in complex analysis says that any bounded entire function is constant. Since it tends to 0 it must be the 0 function. Setting $z = 0$ we see that $\langle h \; T^{-1} k \rangle = 0$ for every pair of vectors. Choosing $h = T^{-1} k$ yields that $\|T^{-1} k\| = 0$ so that $T^{-1} = 0$, a contradiction.

Hence, $\sigma(T)$ must be non-empty. □

Definition 12.4. We let $\sigma_p(T)$ denote the set of eigenvalues of T, i.e., the $\lambda \in \mathbb{C}$ such that there eixsts $h \neq 0$ with $Th = \lambda h$.

If $\lambda \in \sigma_p(T)$ then $\mathcal{N}(\lambda I - T) \neq (0)$ so is not invertible. Thus, $\sigma_p(T) \subseteq \sigma(T)$. In finite dimensions we have that $\sigma(T) = \sigma_p(T)$. Also in finite dimensions, if λ is an eigenvalue of T then $\overline{\lambda}$ is an eigenvalue of T^*. The following example shows that both of these are far from the case in infinite dimensions.

First a little result.

Proposition 12.5. $\sigma(T) = \{\overline{\lambda} : \lambda \in \sigma(T)\}$.

Proof. If $(\lambda I - T)^{-1} = R$ then check that R^* is the inverse of $(\overline{\lambda} I - T^*)$. This means that if $\lambda \in \rho(T)$, then $\overline{\lambda} \in \rho(T^*)$ and the result follows. □

Example 12.6. Let $S : \ell^2 \to \ell^2$ denote the unilateral shift. Since this is an isometry, $Sx = 0 \implies x = 0$ so $0 \notin \sigma_p(S)$. Now let $\lambda \neq 0$ and suppose that $Sx = \lambda x$ with $x = (x_1, x_2, ...)$. We have that $(\lambda x_1, \lambda x_2, ...) = (0, x_1, x_2, ...)$ so that $\lambda x_1 = 0 \implies x_1 = 0$ but then, $\lambda x_2 = x_1 \implies x_2 = 0$ and, inductively, $x_n = 0$ for all n. Hence $\lambda \notin \sigma_p(S)$.

Recall $S^*((x_1, x_2, ...)) = (x_2, x_3, ...)$. So $S^*x = \mu x$ if and only if $\mu x_n = x_{n+1}$. Now if $x_1 = 0$ then all entries would be 0, so we can assume that $x_1 \neq 0$ and that after scaling $x_1 = 1$. Then shows that then $x_n = \mu^{n-1}$. So that the eigenvector would need to have the form

$$(1, \mu, \mu^2, ...).$$

Now it is not hard to see that this vector is in ℓ^2 iff $|\mu| < 1$. Thus,

$$\sigma_p(S^*) = \{\mu : |\mu| < 1\} \subseteq \sigma(S^*).$$

Since $\|S^*$
$+ 1$ we have that $\sigma(S^*) \subseteq \{\mu ; |\mu| \leq 1\}$, so using the fact that it is a closed set, we have that

$$\sigma(S^*) = \{\mu : |\mu| \leq 1\},$$

and by the last result,

$$\sigma(S) = \{\lambda : |\lambda| \leq 1\}.$$

Example 12.7. Let $\mathcal{H} = L^2([0,1], \mathcal{M}, m)$ be the Hilbert space of equivalence classes of square-integrable measurable functions with respect to Lebesgue measure. Recall that two functions are equivalent iff they are equal almost everywhere(a.e.). We shall write $[f]$ for the equivalence class of a function. We define M_t to be the operator of multiplication by the variable t, so that $[f] = [tf(t)]$. It is easy to check that this is a bounded linear operator and is self-adjoint.

In finite dimensions we are often use that every self-adjoint matrix an orthonormal basis of eigenvectors with real eigenvalues. We show that M_t has no eigenvalues. To see this suppose that $M_t[f] = \lambda[f]$ then we would have that $tf(t) = \lambda f(t)$ a.e. so that $(t - \lambda)f(t) = 0$ a.e. But this implies that $f(t) = 0$ a.e. and so $[f] = 0$. Thus, $\sigma_p(M_t) = \emptyset$. To find the spectrum we will need another new idea.

Definition 12.8. An operator $T \in B(\mathcal{H}, \mathcal{K})$ is **bounded below** if there exist $C > 0$ such that $\|Th\| \geq C\|h\|$ for all $h \in \mathcal{H}$.

Proposition 12.9. *If $T \in B(\mathcal{H})$ is invertible then T is bounded below. In fact, $\|Th\| \geq \|T^{-1}\|^{-1}\|h\|$.*

Proof. This follows from, $\|h\| = \|T^{-1}(Th)\| \leq \|T^{-1}\|\|Th\|$. □

Definition 12.10. Let $T \in B(\mathcal{H})$ we say that $\lambda \in \mathbb{C}$ is an **approximate eigenvalue** of T provided that there exists a sequence $\|h_n\| = 1$ such that $\lim_n \|(Th_n - \lambda h_n\| = 0$. We let $\sigma_{ap}(T)$ denote the set of approximate eigenvalues of T.

Note that if $(T - \lambda I)h_n \to 0$ then $(T - \lambda I)$ is not bounded below and so not invertible. Thus,

$$\sigma_{ap}(T) \subseteq \sigma(T).$$

We now show that $[0, 1] \subseteq \sigma_{ap}(M_t)$. To this end let $0 < \lambda < 1$, and for n sufficiently large, set

$$f_n(t) = \begin{cases} 0, & 0 \le t < \lambda - \frac{1}{n} \\ \sqrt{n/2}, & \lambda - \frac{1}{n} \le t \le \lambda + \frac{1}{n} \\ 0, & \lambda + \frac{1}{n} < t \le 1 \end{cases}.$$

It is easily checked that $f_n \in L^2$ and that $\|f_n\| = 1$. Also,

$$\|(M_t f_n - \lambda f_n\|^2 = \int_{\lambda - \frac{1}{n}}^{\lambda + \frac{1}{n}} (t - \lambda)^2 \frac{n}{2} dt = \frac{1}{3n^2} \to 0.$$

Thus, $\lambda \in \sigma_{ap}(M_t)$. The cases of $\lambda = 0$ and $\lambda = 1$ are similar.

Finally, if $\lambda \notin [0, 1]$, then the function $(t-\lambda)^{-1}$ is bounded on $[0, 1]$. Hence the operator $M_{(t-\lambda)^{-1}}$ of multiplication by this function defines a bounded operator and it is easily seen that it is the inverse of the operator $M_t - \lambda I$. Hence, $\lambda \notin \sigma(M_t)$ and we have that

$$\sigma(M_t) = [0, 1].$$

13. Spectral Mapping Theorems and Functional Calculi

In the study of eigenvalues of matrices a special role is played by being able to take polynomials of the matrix. Similarly, if $T \in B(\mathcal{H})$, and $p(z) = a_0 + a_1 z + \cdots + a_n z^n$ then we set $p(T) := a_0 I + a_1 T + \cdots + a_n T^n$.

Note that if we let $\mathcal{P}$ denote the algebra of polynomials, then the mapping $\pi_T : \mathcal{P} \to B(\mathcal{H})$ defined by $\pi_T(p) = p(T)$ satisfies:

- $\pi_T(p + q) = \pi_T(p) = \pi_T(q)$,
- $\pi_T(\lambda p) = \lambda \pi_T(p)$,
- $\pi_T(pq) = \pi_T(p)\pi_T(q)$.

That is it is linear and multiplicative. Such a map between two algebras is called a **homomorphism**. This mapping is sometimes called the **polynomial functional calculus**.

We now wish to see how spectrum behaves under such maps.

Lemma 13.1. *Let $A, B, T \in B(\mathcal{H})$ with $AT = I$ and $TB = I$, then $A = B$ and hence T is invertible.*

Theorem 13.2 (Spectral Mapping Theorem for Polynomials)**.** *Let $T \in B(\mathcal{H})$ and let p be a polynomial. Then*

$$\sigma(p(T)) = \{p(\lambda) : \lambda \in \sigma(T)\}.$$

Proof. For a polynomial p we always assume that $a_n \neq 0$. Given $\lambda \in \sigma(T)$ set $q(z) = p(z) - p(\lambda)$. Since $q(\lambda) = 0$ this means that $(z - \lambda)$ divides q and we may write $q(z) = (z - \lambda)q_1(z)$ for some polynomial $q_1(z)$. Hence, $p(T) - p(\lambda)I = (T - \lambda I)q_1(T) = q_1(T)(T - \lambda I)$. Suppose that $p(T) - p(\lambda)I$ is invertible. Then if B is the inverse then,

$$I = \big(Bq_1(T)\big)(T - \lambda I) = (T - \lambda I)\big(q_1(T)B\big),$$

and so $T - \lambda I$ is invertible by the lemma. Hence, $p(T) - p(\lambda)I$ is not invertible and we have that $\{p(\lambda) : \lambda \in \sigma(T)\} \subseteq \sigma(p(T))$.

On the other hand, if $\mu \in \sigma(p(T))$,(we assume that $a_n \neq 0$) factor

$$p(z) - \mu = a_n(z - \mu_1)\cdots(z - \mu_n).$$

If $(T - \mu_i I)$ is invertible for all i, then $p(T) - \mu$ is a product of invertible operators and so is invertible. Thus, there exists an i so that $\mu_i \in \sigma(T)$. But $p(\mu_i) - \mu = 0$ so that $\mu = p(\lambda)$ for some $\lambda \in \sigma(T)$. □

Definition 13.3. Let $T \in B(\mathcal{H})$, then the **spectral radius** of T is the number

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

Theorem 13.4 (Spectral Radius Theorem)**.** *Let $T \in B(\mathcal{H})$. Then*

- $r(T) \leq \|T^n\|^{1/n}$, $\forall n$,
- $\lim_n \|T^n\|^{1/n}$ *converges,*
- $r(T) = \lim_n \|T^n\|^{1/n}$.

Proof. We sketch the key ideas. First we know that $r(X) \leq \|X\|$ for any X. But by the spectral mapping theorem, $r(T^n) = r(T)^n$, so $r(T) = r(T^n)^{1/n} \leq \|T^n\|^{1/n}$.

From this it follows that $r(T) \leq \liminf_n \|T^n\|^{1/n}$. The proof is completed by showing that $\limsup_n \|T^n\|^{1/n} \leq r(T)$. This is done by setting $w = z^{-1}$ and observing that $zI - T$ is invertible iff $I - wT$ is invertible.

But $(I - wT)^{-1} = \sum_n (wT)^n$ for w small enough(hence z large enough). But we know that this power series is analytic for $|z| > r(T)$ and hence for $|w| < r(T)^{-1}$and w cannot be any larger. This guarantees that the radius of this power series is $R = r(T)^{-1}$. But even for power series with into Banach spaces we have that, $R^{-1} = \limsup \|T^n\|^{1/n}$. We have that $r(T) = \limsup_n \|T^n\|^{1/n}$. □

14. The Riesz Functional Calculus

In this section we will assume that the reader is familiar with complex analysis.

Let $K \subseteq \mathbb{C}$ be a compact set. We let $Hol(K)$ denote the set of functions that are analytic on some open set containing K. Sums and products of such functions are again in this set. Recall from complex analysis that given a simple closed curve or set of closed curves, $\Gamma = \{\Gamma_1, ..., \Gamma_N\}$ then each point that is not on the curves has a **winding number**. For a family of curves the winding number of a point is the sum of its winding numbers. Let $\gamma_j : [a_j, b_j] \to \mathbb{C}$ be a smooth parametrization of each curve Γ_j. If $g(z)$ is continuous on Γ then we set

$$\int_\Gamma g(z)dz = \sum_{j=1}^n \int_{a_j}^{b_j} g(\gamma_j(t))\gamma_j'(t)dt.$$

Note that since everything is continuous these integrals are defined as limits of Riemann sums.

Cauchy's integral formula says that if $g \in Hol(K)$ and Γ has the property that each point in K has winding number 1, then

$$g(w) = \frac{1}{2\pi i}\int_\Gamma g(z)(z-w)^{-1}dz.$$

The **Riesz functional calculus** is defined as follows. Given $T \in B(\mathcal{H})$, pick a system of curves Γ so that each point in $\sigma(T)$ has winding number 1. We define

$$f(T) := \frac{1}{2\pi i}\int_\Gamma f(z)(zI-T)^{-1}dz = \sum_{j=1}^n \frac{1}{2\pi i}\int_{a_j}^{b_j} f(\gamma(t))(\gamma(t)I-T)^{-1}\gamma_j'(t)dt.$$

Again since everything is continuous in the norm topology, these integrals can be defined as limits of Riemann sums, alternatively, for each pair of vectors h, k we can set

$$\langle h|f(T)k\rangle := \sum_{j=1}^n \frac{1}{2\pi i}\int_{a_j}^{b_j} f(\gamma(t))\langle h|(\gamma(t)I-T)^{-1}k\rangle\gamma_j'(t)dt,$$

and check that this defines a bounded operator.

The Riesz functional calculus refers to a group of results that explain how these integral formulas behave. Here is a brief summary of the key results.

Theorem 14.1 (Riesz Functional Calculus)**.** *Let $T \in B(\mathcal{H})$, let $f \in Hol(\sigma(T))$ and let Γ and $\widetilde{\Gamma}$ be two systems of curves that have winding number one about $\sigma(T)$.*

- $\int_\Gamma f(z)(zI-T)^{-1}dz = \int_{\widetilde{\Gamma}} f(z)(zI-T)^{-1}dz$*, so that the definition of $f(T)$ is independent of the particular curves used.*
- *if $f(z) = \sum_{n=0}^\infty a_n z^n$ is a power series with radius of convergence $R, r(T) < R$, then*

$$f(T) = \sum_{n=0}^\infty a_n T^n.$$

- *the map $\pi_T : Hol(\sigma(T)) \to B(\mathcal{H})$ given by $\pi_T(f) = f(T)$ is a homomorphism.*
- $\sigma(f(T)) = \{f(\lambda) : \lambda \in \sigma(T)\}$*.*

14.1. **The Spectrum of a Hermitian.** It is easy to see that any eigenvalue of a Hermitian matrix must be real, but Hermitian operators need not have any eigenvalues. We illustrate how to use the functional calculus to prove that the spectrum of a Hermitian operator is contained in the reals.

Proposition 14.2. *If T is invertible, then $\sigma(T^{-1}) = \{\lambda^{-1}; \lambda \in \sigma(T)\}$.*

Proof. First, since T^{-1} is invertible, $0 \notin \sigma(T)$. Now let $\mu \neq 0$, then

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

Since the first and third terms are invertible, the left hand side is invertible iff $(T - \mu^{-1} I)$ is invertible and the result follows. □

Proposition 14.3. *Let $U \in B(\mathcal{H})$ be unitary. Then $\sigma(U) \subseteq \{\lambda : |\lambda| = 1\}$.*

Proof. If $\lambda \in \sigma(U)$ then $|\lambda| \leq \|U\| = 1$. Since $U^{-1} = U^*$, $\|U^{-1}\| \leq !$. Hence, $\lambda \in \sigma(U) \implies \lambda^{-1} \in \sigma(U^{-1}) \implies |\lambda^{-1}| \leq \|U^{-1}\| = 1$. Thus, $|\lambda| \geq 1$ and the result follows. □

We will often use $\mathbb{T} := \{\lambda : |\lambda| = 1\}$ to denote the unit circle in the complex plane.

Proposition 14.4. *Let $H \in B(\mathcal{H})$, $H = H^*$. Then $\sigma(H) \subseteq \mathbb{R}$.*

Proof. We have that $e^{iz} = \sum_{n=0}^{\infty} \frac{i^n z^n}{n!}$ has infinite radius of convergence. Set $U = e^{iH}$, then $U^* = \sum_{n=0}^{\infty} \frac{(-i)^n H^{*n}}{n!} = e^{-iH}$.

Hence, $UU^* = U^*U = I$ so U is unitary. Hence,

$$\{e^{iz} : z \in \sigma(H)\} = \sigma(U) \subseteq \mathbb{T}.$$

But it is easily checked that $e^{iz} \in \mathbb{T}$ iff $z \in \mathbb{R}$. □

15. Normal Operators

Definition 15.1. An operator $N \in B(\mathcal{H})$ is **normal** provided that $NN^* = N^*N$.

We see that Hermitian and unitary operators are normal.

Proposition 15.2. *Let $N \in B(\mathcal{H})$ and let $N = H + iK$ be its Cartesian decomposition. Then N is normal iff $HK = KH$.*

Given a polynomial, $p(z, \overline{z})) = \sum_{i,j} a_{i,j} z^i \overline{z}^j$ in z and $\overline{z}$ we can evaluate it for any normal operator, by setting $p(N, N^*) = \sum_{i,j} a_{i,j} N^i N^{*j}$. If we let $\mathcal{P}(z, \overline{z})$ denote the set of such polynomials, then it is again an algebra and we see that when N is normal, the map $\pi_N : \mathcal{P}(z, \overline{z}) \to B(\mathcal{H})$ is a homomorphism. In fact, if it is a homomorphism, then necessarily N is normal, since in the polynomials, $z\overline{z} = \overline{z}z$.

Proposition 15.3 (Gelfand-Naimark)**.** *Let N be normal and let $p(z, \overline{z})$ be a polynomial. Then*

$$\|p(N, N^*)\| = \{|p(\lambda, \overline{\lambda})| : \lambda \in \sigma(N)\}.$$

The right hand side of the above formula is the supremum norm $\|\cdot\|_\infty$ of the polynomial over $\sigma(N)$. By the Stone-Weierstrass theorem, the polynomials in z and $\overline{z}$ are dense in the continuous functions on $\sigma(N)$, Thus, given any continuous function $f \in C(\sigma(N))$ there is a Cauchy sequence of polynomials $\{p_n(z, \overline{z})\}$ that converge to it in the supremum norm. By the above result, $\|p_n(N, N^*) - p_m(N, N^*)\| = \|p_n - p_m\|_\infty$ and so this sequence of operators is also Cauchy in norm and hence converges to an operator that we shall denote by $f(N)$. It is not hard to check that the operator $f(N)$ is independent of the particular Cauchy sequence chosen.

This gives us a **continuous functional calculus for normal operators**.

Theorem 15.4 (Gelfand-Naimark). *Let $N \in B(\mathcal{H})$ be normal and let $\pi_N : C(\sigma(N)) \to B(\mathcal{H})$ be defined by $\pi_N(f) = f(N)$. Then*

- *π_N is a homomorphism,*
- *$\pi_N(\overline{f}) = f(N)^*$(maps satisfying these two properties are called* ***-homomorphisms***),*
- *$\|f(N)\| = \sup\{|f(\lambda)| : \lambda \in \sigma(N)\}$,*
- *$\sigma(f(N)) = \{f(\lambda) : \lambda \in \sigma(N)\}$.*

15.1. **Positive Operators.**

Definition 15.5. An operator $P \in B(\mathcal{H})$ is called **positive**, denoted $P \geq 0$ provided that $\langle h|Ph\rangle \geq 0$ for every $h \in \mathcal{H}$.

CAUTION: In the theory of matrices these are often called **positive semidefinite** but the same notation $P \geq 0$ is used and positive is reserved for matrices such that $\langle h|Ph\rangle > 0$ for all $h \neq 0$, for which they use the notation $P > 0$. So the terminology is different for the two fields.

By one of our earlier results, $P \geq 0 \implies P = P^*$, since $\langle h|Ph\rangle \in \mathbb{R}$. So positive operators are self-adjoint.

Note that if $T \in B(\mathcal{H}, \mathcal{K})$ then $T^*T \geq 0$.

Problem 15.6. *Let $H = H^*$. Prove that $\sigma_{ap}(H) = \sigma(H)$. (HINT: For $\lambda \in \mathbb{R}$, $\mathcal{N}(H - \lambda I) = \mathcal{R}(H - \lambda I)^\perp$, so that if λ is not an eigenvalue, then $\mathcal{R}(H - \lambda I)$ is dense. Show that if $H - \lambda I$ is bounded below and has dense range, then it is invertible.)*

Problem 15.7. *Prove that the following are equivalent:*

1. *$P \geq 0$,*
2. *$P = P^*$ and $\sigma(P) \subseteq [0, +\infty)$,*
3. *there exists $B \in B(\mathcal{H})$ such that $P = B^*B$.*

By the above results, since P is normal and $\sigma(P) \subseteq [0, +\infty)$ the function $\sqrt{t}$ is continuous on $\sigma(P)$ and so by the continuous functional calculus, there is an operator $\sqrt{P}$ defined.

Definition 15.8. Let $T \in B(\mathcal{H}, \mathcal{K})$ we define the **absolute value of T** to be the operator,

$$|T| := \sqrt{T^*T}.$$

CAUTION: Unlike numbers, $|T|$ is not always equal to $|T^*|$. In fact, if $T \in B(\mathcal{H}, \mathcal{K})$, then $|T| \in B(\mathcal{H})$ while $|T^*| \in B(\mathcal{K})$.

Here is another application of the continuous functional calculus. Let $H = H^*$. Define

$$f^+(t) = \begin{cases} 0, & t \leq 0 \\ t, & t > 0 \end{cases} \text{ and } f^-(t) = \begin{cases} -t, & t \leq 0 \\ 0, & t > 0 \end{cases}.$$

These are both continuous functions and they satisfy

- $f^+(t) + f^-(t) = |t| = \sqrt{t^2}$,
- $f^+(t) - f^-(t) = t$,
- $f^+(t)f^-(t) = 0$.

By the functional calculus, we may set $H^+ = f^+(H)$ and $H^- = f^-(H)$. Then we have that

- $H^+ + H^- = |H| = \sqrt{H^2}$,
- $H^+ - H^- = H$,
- $H^+H^- = H^-H^+ = 0$.

Note that if H was just a matrix in diagonal form, then H^+ would be the matrix gotten by leaving the positive eigenvalues alone and setting the negative eigenvalues equal to 0, with H^- similarly defined. These two matrices would satisfy the above properties. Thus, the functional calculus allows us to abstractly carry out such operations even when we have no eigenvalues.

15.2. **Polar Decomposition.** Every complex number z can be written as $z = e^(it)|z|$ for some unique "rotation". The following is the analogue for operators.

Theorem 15.9 (Polar Decomposition)**.** *Let $T \in B(\mathcal{H}, \mathcal{K})$. Then there exists a partial isometry W such that $T = W|T|$. Moreover, W is unique if we require that*

- *W^*W is the projection onto $\mathcal{N}(T)^\perp$,*
- *WW^* is the projection onto $\mathcal{R}(T)^-$.*

Because $\mathcal{R}(|T|)^- = \mathcal{N}(|T|)^\perp = \mathcal{N}(T)^\perp$ we have that $W^*T = W^*W|T| = |T|$.

16. Compact Operators

Theorem 16.1. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces, let $\mathbb{B}_1 = \{h \in \mathcal{H} : \|h\| \leq 1\}$ and let $T \in B(\mathcal{H}, \mathcal{K})$. The following are equivalent:*

- *T is a norm limit of finite rank operators,*
- *$T(\mathbb{B}_1)^-$ is compact,*
- *$T(\mathbb{B}_1)$ is compact.*

Such operators are called **compact** and the set of all compact operators in $B(\mathcal{H}, \mathcal{K})$ is denoted $\mathbb{K}(\mathcal{H}, \mathcal{K})$. When $\mathcal{H} = \mathcal{K}$ we write more simply, $\mathbb{K}(\mathcal{H})$.

Theorem 16.2. *$\mathbb{K}(\mathcal{H},\mathcal{K})$ is a norm closed subspace of $B(\mathcal{H},\mathcal{K})$. If $A \in B(\mathcal{H})$ and $B \in B(\mathcal{K})$ and $K \in \mathbb{K}(\mathcal{H},\mathcal{K})$, then $BKA \in \mathbb{K}(\mathcal{H},\mathcal{K})$.*

Theorem 16.3. *Let $H \in \mathbb{K}(\mathcal{H})$ with $H = H^*$. Then*

(1) *every $0 \neq \lambda \in \sigma(H)$ is an eigenvalue,*
(2) *H has at most countably many non-zero eigenvalues, and each eigenvalue has a finite dimensional eigenspace,*
(3) *there exists a countable orthonormal set of eigenvectors $\{\phi_n : n \in A\}$ with corresponding eigenvalues $\{\lambda_n : n \in A\}$ such that*

$$H = \sum_{n \in A} \lambda_n |\phi_n\rangle\langle\phi_n|,$$

Moreover, if $H = \sum_j \beta_j |\gamma_j\rangle\langle\gamma_j|$ is another such representation with orthonormal vectors, then $\{\beta_j\} = \{\lambda_n\}$ including multplicities.

Proof. First if $0 \neq \lambda \in \sigma(H)$, then we know that there exists a sequence of vectors $\{h_n\}$ such that $Hh_n - \lambda h_n\| \to 0$. Since H is compact, there is a subsequence $\{h_{n_k}\}$ such that $\{Hh_{n_k}\}$ converges. Set $v = \lim_k Hh_{n_k} = \lim_k \lambda h_{n_k}$, since $\|Hh_{n_k} - \lambda h_{n_k}\| \to 0$. But this implies that $\lim_k h_{n_k} = \lambda^{-1} v := w$. Since each h_{n_k} is a unit vector, $\|w\| =!$ and in particular $w \neq 0$. Now $Hw = \lim_k Hh_{n_k} = v = \lambda w$, so we have that λ is an eigenvalue.

Now suppose that for some $0 \neq \lambda \in \sigma(H)$ the eigenspace was infinite dimensional. Then by choosing an orthonormal basis for the eigenspace, we could find countably many orthonormal vectors $\{h_n\}$ with $Hh_n = \lambda h_n$ Since for $n \neq m$ we have that $\|\lambda h_n - \lambda h_m\| = |\lambda|\sqrt{2}$, we see that no subsequence could converge, contradicting (2). Thus, each eigenspace is at most finite dimensional.

We prove that the set of non-zero eigenvalues is countable by contradiction. Now suppose that there was an uncountable number of non-zero points in $\sigma(H)$. For each n, consider $S_n = \{\lambda \in \sigma(H) : |\lambda| \geq 1/n\}$. If this set was finite for all n then

$$\sigma(H)\backslash\{0\} = \cup_n S_n,$$

would be a countable union of finite sets and hence countable. Thus, the assumption implies that for some n_0 the set S_{n_0} is infinite. This allows us to choose a countably infinite subset $\{\lambda_k : k \in \mathbb{N}\} \subset S_{n_0}$ of distinct points. For each k there is unit eigenvector, v_k with $Hv_k = \lambda_k v_k$. But then for $k \neq j$,

$$\lambda_k \langle v_k | v_j\rangle = \langle Hv_k | v_j\rangle = \langle v_k | Hv_j\rangle = \langle v_k | v_j\rangle \lambda_j.$$

Since $\lambda_k \neq \lambda_j$ this equation is possible iff $\langle v_k | v_j\rangle = 0$. Thus, the eigenvectors are orthogonal. This implies that $\|Hv_k - Hv_j\| = \|\lambda_k v_k - \lambda_j v_j\| \geq \frac{\sqrt{2}}{n_0}$ which means that no subsequence of these vectors could converge. This contradicts that H is compact.

The third statement follows by choosing an orthonormal basis for each finite dimensional eigenspace and taking the union of all these vectors to for $\{\phi_n\}$.

Finally, the last statement follows from the fact that both sets must be the eigenvalues with the appropriate multiplicities. □

Definition 16.4. If $K \in \mathbb{K}(\mathcal{H}, \mathcal{K})$. Then $|K|$ is a positive, compact operator and hence by the above theorem it has countably many distinct non-zero eigenvalues, every non-zero eigenvalue is positive and each non-zero eigenvalue has at most a finite dimensional eigenspace. We let $\lambda_1 \geq \lambda_2 \geq \dots$, denote these eigenvalues with the convention that each eigenvalue is repeated a finite number of times corresponding to the dimension of its eigenspace. The number

$$s_j(K) := \lambda_j,$$

is called the **j-th singular value of K**.

Theorem 16.5 (Schmidt)**.** *Let $K \in \mathbb{K}(\mathcal{H}, \mathcal{K})$. Then there exist countable collections of orthonormal vectors $\{\phi_n\} \subseteq \mathcal{H}$ and $\{\psi_n\} \subseteq \mathcal{K}$, such that*

$$K(h) = \sum_n s_n(K)\langle \phi_n | h \rangle \psi_n,$$

i.e., $K = \sum_n s_n(K)|\psi_n\rangle\langle\phi_n|$. Moreover, if $K = \sum_j \beta_j \langle|\gamma_j\rangle\langle\delta_j|$ is another representation with $\beta_j > 0$ and $\{\gamma_j\}$ and $\{\delta_j\}$ orthonormal families, then $\{\beta_j\} = \{s_n(K)\}$, including multiplicities.

Proof. Use the polar decomposition to write $K = W|K|$ with W a partial isometry. Since $|K|$ is compact and positive with eigenvalues $s_j(K)$, we can write $|K| = \sum_j s_j(K)|\phi_j\rangle\langle\phi_j|$, with $\{\phi_j\}$ a countable orthonormal set. Since the vectors $\phi_j \in \mathcal{R}(|K|)$ and W acts as an isometry on that space the vectors $\psi_j = W\phi_j$ are also orthonormal, by the inner product preserving property of isometries.

For any h we have that $K(h) = W\big(\sum_j s_j(K)\langle\phi_j|h\rangle\phi_j\big) = \sum_j s_j(K)\langle\phi_j|h\rangle\psi_j$ and so $K = \sum_j s_j(K)|\phi_j\rangle\langle\psi_j|$.

Finally for the last statement, given such a representation, we have that $K^* = \sum_j \beta_j|\delta_j\rangle\langle\gamma_j|$. Hence, $K^*K = \sum_j \beta_j^2|\gamma_j\rangle\langle\gamma_j|$. Thus, $|K| = \sum_j \beta_j|\gamma_j\rangle\langle\gamma_j|$. From the earlier result we conclude that, $\{\beta_j\} = \{s_n(K)\}$, including multiplicities. □

16.1. **The Schatten Classes.** Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. For $1 < p < +\infty$ the **Schatten p-class** is defined to be

$$\mathcal{C}_p(\mathcal{H}, \mathcal{K}) = \{K \in \mathbb{K}(\mathcal{H}, \mathcal{K}) : \sum_j s_j(K)^p < +\infty\}.$$

Note also that since $\{s_j(K)\}$ are the eigenvalues of $|K|$ including multiplicities, that if we use a basis for $\mathcal{H}$ that includes the eigenvectors for $|K|$, then we see that

$$Tr(|K|^p) = \sum_j s_j(K)^p,$$

so that $\mathcal{C}_p$ is exactly the set of operators for which $|K|^p$ is in the trace-class.

For $K \in \mathcal{C}_p$ we set $\|K\|_p = \left(\sum_j s_j(K)^p\right)^{1/p}$ and this is called the **p-norm** of the operator.

The space $\mathcal{C}_1$ is called the **trace class operators** and the space $\mathcal{C}_2$ is called the **Hilbert-Schmidt** operators.

Theorem 16.6. *We have that:*

1. $\mathcal{C}_p(\mathcal{H}, \mathcal{K})$ *is a Banach space in the p-norm,*
2. *for* $A \in B(\mathcal{H}), B \in B(\mathcal{K})$ *and* $K \in \mathcal{C}_p(\mathcal{H}, \mathcal{K})$ *we have that* $BKA \in \mathcal{C}_p(\mathcal{H}, \mathcal{K})$ *and* $\|BKA\|_p \leq \|B\|\|K\|_p\|A\|$.

Theorem 16.7. *Let* $K \in \mathcal{C}_1(\mathcal{H})$ *and let* $\{e_a : a \in A\}$ *be any orthonormal basis for* $\mathcal{H}$. *Then* $\sum_a \langle e_a | K e_a \rangle$ *converges and its value is independent of the particular orthonormal basis.*

Definition 16.8. Let $K \in \mathcal{C}_1(\mathcal{H})$, then we set

$$Tr(K) = \sum_a \langle e_a | K e_a \rangle,$$

where $\{e_a : a \in A\}$ is any orthonormal basis.

Theorem 16.9. *Let* $\mathcal{H}$ *be a Hilbert space.*

1. *for* $A \in B(\mathcal{H})$ *and* $K \in \mathcal{C}_1(\mathcal{H})$, $Tr(AK) = Tr(KA)$.
2. *For each* $A \in B(\mathcal{H})$ *define*
$$f_A : \mathcal{C}_1(\mathcal{H}) \to \mathbb{C},$$
by $f_A(K) = Tr(AK)$, *then* f_A *is a bounded linear functional and* $\|f_A\| = \|A\|$.
3. *If* $f : \mathcal{C}_1(\mathcal{H}) \to \mathbb{C}$ *is any bounded, linear functional, then there exists a unique* $A \in B(\mathcal{H})$ *such that* $f = f_A$. *Consequently, the map* $A \to f_A$ *is an isometric isomorphism of* $B(\mathcal{H})$ *onto the Banach space dual* $\mathcal{C}_1(\mathcal{H})^*$.
4. *If* $K \in \mathcal{C}_1(\mathcal{H})$ *then the map*
$$f_K : \mathbb{K}(\mathcal{H}) \to \mathbb{C},$$
defined by $f_K(T) = Tr(KT)$ *is a bounded linear functional and* $\|f_K\| = \|K\|_1$.
5. *If* $f : \mathbb{K}(\mathcal{H}) \to \mathbb{C}$ *is any bounded linear functional, then there exists a unique* $K \in \mathcal{C}_1(\mathcal{H})$ *such that* $f = f_K$. *Consequently, the map* $K \to f_K$ *is an isometric isomorphism of* $\mathcal{C}_1(\mathcal{H})$ *onto the Banach space dual* $\mathbb{K}(\mathcal{H})^*$.

Theorem 16.10. *Let* $1 < p < +\infty$ *and let* $\frac{1}{p} + \frac{1}{q} = 1$. *Then*

1. *for* $T \in \mathcal{C}_q(\mathcal{H})$ *and* $R \in \mathcal{C}_p(\mathcal{H})$, $TR \in \mathcal{C}_1(\mathcal{H})$, $Tr(RT) = Tr(TR)$ *and* $|Tr(TR)| \leq \|R\|_p\|T\|_q$.
2. *the map* $f_T : \mathcal{C}_p(\mathcal{H}) \to \mathbb{C}$ *given by* $f_T(R) = Tr(TR)$ *is a bounded linear functional with* $\|f_T\| = \|T\|_q$.
3. *if* $f : \mathcal{C}_p(\mathcal{H}) \to \mathbb{C}$ *is any bounded linear functional, then there exists a unique* $T \in \mathcal{C}_q(\mathcal{H})$ *such that* $f = f_T$.

Consequently, the map $T \to f_T$ is an isometric isomorphism of $\mathcal{C}_q(\mathcal{H})$ onto the Banach space dual $\mathcal{C}_p(\mathcal{H})^*$.

16.2. **Hilbert-Schmidt and Tensor Products.** Given orthonormal bases $\{e_a : a \in A\}$ for $\mathcal{H}$ and $\{f_b : b \in B\}$ for $\mathcal{K}$ we have a matrix representation for any $T \in B(\mathcal{H}, \mathcal{K})$, namely,

$$T_{mat} = (t_{b,a}), \text{ with } t_{b,a} = \langle f_b | T e_a \rangle.$$

Proposition 16.11. *$T \in \mathcal{C}_2(\mathcal{H}, \mathcal{K})$ if and only if $\sum_{a \in A, b \in B} |t_{b,a}|^2 < +\infty$. Moreover,*

$$\|T\|_2^2 = \sum_{a \in A, b \in B} |t_{b,a}|^2.$$

Proof. $T \in \mathcal{C}_2$ if and only if $\|T\|_2^2 = Tr(T^*T) < +\infty$. But

$$Tr(T^*T) = \sum_{a \in A} \langle e_a | T^*T e_a \rangle = \sum_{a \in A} \|Te_a\|^2 = \sum_{a \in A} \sum_{b \in B} |\langle f_b, Te_a \rangle|^2 = \sum_{a \in A, b \in B} |t_{b,a}|^2,$$

and the result follows. □

Let $\phi \in \mathcal{H} \overline{\otimes} \mathcal{K}$ and let $\{e_a : a \in A\}$ and $\{f_b : b \in B\}$ be orthonormal bases for $\mathcal{H}$ and $\mathcal{K}$, respectively. Expanding ϕ with respect to this basis we have that $\phi = \sum_{a,b} t_{b,a} e_a \otimes f_b$ with $\|\phi\|^2 = \sum_{a,b} |t_{b,a}|^2$. Hence, $(t_{b,a})$ is the matrix of a Hilbert-Schmidt operator T from $\mathcal{H}$ to $\mathcal{K}$, with $\|T\|_2 = \|\phi\|$. Thus, if we fix bases, then we see that there is an isometric linear map, $\Gamma : \mathcal{H} \overline{\otimes} \mathcal{K} \to \mathcal{C}_2(\mathcal{H}, \mathcal{K})$, that is onto.

In the case of Hilbert spaces over $\mathbb{R}$ it is possible to define this map in a basis free manner. If $\phi = \sum_i h_i \otimes k_i$ is a finite sum, then define an operator T_ϕ by

$$T_\phi(h) = \sum_i \langle h_i | h \rangle k_i,$$

i.e., $T_\phi = \sum_i |k_i\rangle\langle h_i|$. It is not hard to check that T_ϕ is Hilbert-Schmidt and that the map, $\Delta : \mathcal{H} \otimes \mathcal{K} \to \mathcal{C}_2(\mathcal{H}, \mathcal{K})$ given by $\Delta(\phi) = T_\phi$ is linear and isometric. Moreover, if we pick bases for $\mathcal{H}$ and $\mathcal{K}$ as above, then the matrix of the linear map T_ϕ is $\Gamma(\phi)$. Thus, in the real case the map Γ is the continuous extension of the map Δ after picking bases.

Unfortunately, in the complex case, we can no longer define the map Δ map, since the identification of the vector h_i with the functional $|h_i\rangle$ is conjugate linear, and so we are forced to pick bases and deal with the map Γ, which in this case really is a basis dependent map.

Here is one application of these ideas.

Proposition 16.12 (Schmidt Decomposition of Vectors)**.** *Let $\phi \in \mathcal{H} \overline{\otimes} \mathcal{K}$. Then there are countable orthonormal families of vectors $\{u_n\} \subseteq \mathcal{H}$ and $\{v_n\} \subseteq \mathcal{K}$ and constants $\beta_n > 0$ such that*

$$\phi = \sum_n \beta_n u_n \otimes v_n.$$

Moreover, if $\phi = \sum_m \alpha_m w_m \otimes y_m$ *is another such representation, then* $\{\alpha_m\} = \{\beta_n\}$*, including multiplicities. Moreover,* $\|\phi\|^2 = \sum_n \beta_n^2$.

Proof. We first do the case of real Hilbert spaces. Consider the Hilbert-Schmidt operator $T_\phi : \mathcal{H} \to \mathcal{K}$. By the above results we may write

$$T_\phi = \sum_n s_n(T_\phi)|v_n\rangle\langle u_n|,$$

where $\{u_n\}$ and $\{v_n\}$ are orthonormal. Consequently, $\phi = \sum_n s_n(T_\phi) u_n \otimes v_n$, as desired.

If we have a representation of $\phi = \sum_m \alpha_m w_m \otimes y_m$, then $T_\phi = \sum_m \alpha_m |y_m\rangle\langle w_m|$ and by our earlier results this forces, $\{\alpha_m\} = \{s_n(T_\phi)\}$ along with multiplicities.

For the complex case, we fix a basis to define the matrix and Hilbert-Schmidt operator $T = (t_{b,a})$. This yields the representation of the vector with coefficients $s_n(T)$. Now one shows that if one chooses a different basis, then the operator defined by the new matrix, say R is unitarily to T and so $s_n(R) = s_n(T)$ for all n. Finally, given any representation, one extends the sets $\{w_m\}$ and $\{y_m\}$ to orthonormal bases, and observes that then $\{\alpha_m\}$ is then the singular values of the matrix obtained from ϕ for this basis. ☐

17. Unbounded Operators

Many naturally defined maps in mathematical physics, such as position and momentum operators, are often unbounded. So are many natural differential operators. Unbounded operators and self-adjointness questions play an important role in these areas. We provide a brief glimpse into this theory.

Definition 17.1. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. A map $B : \mathcal{K} \times \mathcal{H} \to \mathbb{C}$ such that

- $B(k, h_1 + \lambda h_2) = B(k, h_1) + \lambda B(k, h_2)$,
- $B(k_1 + \lambda k_2, h) = B(k_1, h) + \overline{\lambda} B(k_2, h)$

is called a **sesquilinear form**. A sesquilinear form is called **bounded** if there is a constant $C \geq 0$ such that

$$|B(k, h)| \leq C\|k\|\|h\|, \forall h \in \mathcal{H}, k \in \mathcal{K}.$$

In this case, the least constant C satisfying this inequality is called the **norm** of the sesquilinear form and is denoted $\|B\|$.

Proposition 17.2. *If* $T \in B(\mathcal{H}, \mathcal{K})$*, then* $B(k, h) = \langle k|Th\rangle$ *is a bounded sesquilinear form with* $\|B\| = \|T\|$*. Conversely, if* B *is a bounded sesquilinear form, then there exists a unique* $T \in B(\mathcal{H}, \mathcal{K})$ *such that* $B(k, h) = \langle k|Th\rangle$*.*

Problem 17.3. *Prove the above proposition.*

Theorem 17.4 (Hellinger-Toeplitz)**.** *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $T : \mathcal{H} \to \mathcal{K}$ be a linear map. If there exists a linear map $R : \mathcal{K} \to \mathcal{H}$ such that*

$$\langle Rk|h\rangle = \langle k|Th\rangle, \forall h \in \mathcal{H},\ k \in \mathcal{K},$$

then T is bounded.

Problem 17.5. *Prove the Hellinger-Toeplitz theorem.*

By the above theorem, if one wants to study notions of adjoints in the unbounded setting, then necessarily the domain of the map will not be all of the Hilbert space. We will think of a linear map as a pair consisting of the domain of the map and the map itself. To this end let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces, let $\mathcal{D}(T) \subseteq \mathcal{H}$ be a vector subspace, not necessarily closed, and let $T : \mathcal{D}(T) \to \mathcal{K}$ be a linear map. By the **graph of T** we mean the set

$$\Gamma(T) = \{(h, Th) \in \mathcal{H} \oplus \mathcal{K} : h \in \mathcal{D}(T)\}.$$

It is not hard to see that $\Gamma(T)$ is a subspace.

We say that $T_1 : \mathcal{D}(T_1) \to \mathcal{K}$ **extends T** provided that $\mathcal{D}(T) \subseteq \mathcal{D}(T_1) \subseteq \mathcal{H}$ and for every $h \in \mathcal{D}(T)$ we have that $T_1 h = Th$. In this case we write $T \subseteq T_1$**.**

It is not hard to see that $T \subseteq T_1 \iff \Gamma(T) \subseteq \Gamma(T_1)$. The following problem clarifies some of these connections.

Problem 17.6. *Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces and let $\Gamma \subseteq \mathcal{H} \oplus \mathcal{K}$ be a vector subspace. Prove that Γ is the graph of a linear map if and only if $(0, k) \in \Gamma \implies k = 0$.*

Such a subspace is called a **graph.**

Definition 17.7. We say that $T : \mathcal{D}(T) \to \mathcal{K}$ is **closed** provided that $\Gamma(T)$ is a closed subspace of $\mathcal{H} \oplus \mathcal{K}$. We say that T is **closalbe** if there exists T_1 with $T \subseteq T_1$ such that T_1 is closed. If T is closable, the smallest closed extension of T is called the **closure of T** and is denoted $\bar{T}$.

Proposition 17.8. *T is closable if and only if the closure of $\Gamma(T)$ is a graph. In this case, $\Gamma(T)^- = \Gamma(\bar{T})$.*

Thus, by the last problem we see that T is closable if and only if $(0, k) \in \Gamma(T)^-$ implies that $k = 0$.

For the following examples, we assume some measure theory. It is useful to note that if f, g are continuous on $[0, 1]$ and $f = g$ a.e., then $f = g$. Thus, we can think of $C([0, 1])$ as a subspace of $L^2([0, 1])$, where this latter space denotes the Hilbert space of equivalence classes of square-integrable functions with respect to Lebesgue measure. We let $\mathcal{C}^1([0, 1]) \subseteq C([0, 1])$ denote the vector space of continuously differentiable functions.

Example 17.9. Let $\mathcal{H} = \mathcal{K} = L^2([0, 1])$, let $\mathcal{D}(T_1) = \mathcal{C}^1([0, 1])$, let $\mathcal{D}(T_2) = \{f \in \mathcal{C}^1([0, 1]) : f(0) = f(1) = 0\}$ and define $T_1(f) = T_2(f) = f'$. Note that $f_\lambda(t) = e^{\lambda t} \in \mathcal{D}(T_1)$ but is never in $\mathcal{D}(T_2)$ and $T_1(f_\lambda) = \lambda f_\lambda$. Thus, every complex number is an eigenvalue of T_1. From calculus we know that this is the unique solution to $f' = \lambda f$, so that T_2 has no eigenvalues. Hence, we have that

$$\sigma_p(T_1) = \mathbb{C}, \ \ \sigma_p(T_2) = \emptyset.$$

Are these operators closed or closable? For closed, we need to decide that if $\{f_n\} \subseteq \mathcal{D}(T_i)$ such that $\|f_n - f\|_2 \to 0$ and $\|f_n' - g\|_2 \to 0$, then do we necessarily have that $f \in \mathcal{D}(T_i)$ with $g = f'$? The answer to this is no. For example, the functions,

$$f(t) = |t - 1/2|, \ g(t) = \begin{cases} 1, & 0 \le t < 1/2 \\ -1, & 1/2 < t \le 1 \end{cases}$$

can be approximated as above by just "smoothing f near 1/2" and f is in neither domain. For closable, we need to rule out $(0, g) \in \Gamma(T_i)^-$, i.e., if $\|f_n\|_2 \to 0$ and $\|f_n' - g\|_2 \to 0$ does this force $g = 0$ a.e. ? Trying one sees that this is hard to do with no tools. The next result gives a very useful tool.

Definition 17.10. We say that $T : \mathcal{D}(T) \to \mathcal{K}$ is **densely defined** if $\mathcal{D}(T) \subseteq \mathcal{H}$ is a dense subspace. When T is densely defined, we set

$$\mathcal{D}(T^*) = \{k \in \mathcal{K} : \exists h_1 \in \mathcal{H}, \langle k|Th\rangle = \langle h_1|h\rangle, \forall h \in \mathcal{D}(T)\}.$$

It follows readily, using the fact that $\mathcal{D}(T)$ is dense, that if such a vector h_1 exists, then it is unique and we define $T^* : \mathcal{D}(T^*) \to \mathcal{H}$ by $T^*(k) = h_1$. This operator(along with this domain) is called the **adjoint of T**.

Theorem 17.11. (1) *T^* is closed.*
(2) *T is closable if and only if $\mathcal{D}(T^*)$ is a dense subset of $\mathcal{K}$. In this case $\overline{T} = T^{**}$.*
(3) *If T is closable, then $(\overline{T})^* = T^*$.*

Returning to the examples, we see that $g \in \mathcal{D}(T_i^*)$ if and only if there is a $h_1 \in L^2$ such that

$$\langle g|T_i(f)\rangle = \int_0^1 g(t)f'(t)dt = \int_0^1 h_1(t)f(t)dt = \langle h_1|f\rangle, \forall f \in \mathcal{D}(T_i).$$

This suggests integration by parts:

$$\int_0^1 gf'dt = -\int_0^1 g'(t)f(t)dt + (f(1)g(1) - f(0)g(0)).$$

Thus, if $g \in \mathcal{C}^1([0,1])$ and $g(1) = g(0) = 0$, then

$$\langle g|T_i(f)\rangle = \langle -g'|f\rangle,$$

so that in both cases every such g is in $\mathcal{D}(T_i)$. Because such g's are dense in L^2, we have that both operators are closable.

To understand the domains of either operator one needs more measure theory.

Definition 17.12. A function $f : [a, b] \to \mathbb{R}$ is **absolutely continuous(AC)** provided that for every $\epsilon > 0$ there exists $\delta > 0$ such that whenever we have finitely many $[a_i, b_i]$ that are non-intersecting subintervals with $\sum_i |b_i - a_i| < \delta$, then $\sum_i |f(b_i) - f(a_i)| < \epsilon$.

Not that if we only use one subinterval, then this is precisely uniform continuity.

Theorem 17.13. *Let $f : [a, b] \to \mathbb{R}$. Then f is AC if and only if*

1. *$f'(t)$ exists for almost all t,*
2. *f' is integrable,*
3. *for any $a \le c < d \le b$, $\int_c^d f'(t)dt = f(d) - f(c)$.*

For our two examples we have that

$$\mathcal{D}(T_1^*) = \{g : \text{g is AC, g(0) = g(1) =0}\},\ \mathcal{D}(T_2^*) = \{g :\ \text{g is AC }\},$$

and in both cases $T_1^*(g) = T_2^*(g) = -g'$.

We now turn to the definition of self-adjoint in the unbounded case.

Definition 17.14. A densely defined operator on $\mathcal{H}$ is **symmetric**(some times called **Hermitian**) if $T \subseteq T^*$. i.e., $\mathcal{D}(T) \subseteq \mathcal{D}(T^*)$ and $T^*h = Th, \forall h \in \mathcal{D}(T)$. A densely defined operator is called **self-adjoint** if $T = T^*$.

If T is self-adjoint, then $T = T^* = T^** = \bar{T}$.

Just as with bounded self-adjoint operators there is a nice functional calculus for unbounded self-adjoint operators.

Example 17.15. For the operator of differentiation we saw that a -1 was involved in the adjoint. For this reason we consider $\mathcal{H} = \mathcal{K} = L^2([0, 1])$, let $\mathcal{D}(S_1) = \mathcal{C}^1([0, 1])$, let $\mathcal{D}(S_2) = \{f \in \mathcal{C}^1([0, 1]) : f(0) = f(1) = 0\}$ and define $S_1(f) = S_2(f) = if'$. Then we have that $S_1^*(g) = S_2^*(g) = -ig'$ when g is in $\mathcal{C}^1$ but a careful look at the boundary values shows that $S_2 \subseteq S_2^*$ while $\mathcal{D}(S_1^*)$ is not a subset of $\mathcal{D}(S_1)$. Thus, S_2 is symmetric while S_1 is not.

A natural question is what kinds of boundary values do give self-adjoint extensions?

Example 17.16. Let $\mathcal{H} = L^2([0, 1])$, fix $|\alpha| = 1$ and let $\mathcal{D}(R_\alpha) = \{f \in AC([0, 1]) : f(0) = \alpha f(1)\}$ and set $R_\alpha(f) = if'$. Now it is easily checked that $R_\alpha^* = R_\alpha$.

Notice that R_α is a one-dimensional extension of S_2.

Since self-adjoint operators must be closed a natural question is if we start with a symmetric operator, when will its closure be self-adjoint? The following results address this issue.

Definition 17.17. A symmetric operator is **essentially self-adjoint** if its closure is self-adjoint.

Theorem 17.18. *Let T be symmetric. The following are equivalent:*

(1) *T is self-adjoint,*
(2) *$T = \bar{T}$ and $\mathcal{N}(T^* \pm iI) = (0)$,*
(3) *$\mathcal{R}(T \pm iI) = \mathcal{H}$.*

Theorem 17.19. *Let T be symmetric. The following are equivalent:*
(1) *T is essentially self-adjoint,*
(2) *$\mathcal{N}(T^* \pm iI) = (0)$,*
(3) *$\mathcal{R}(T \pm iI) = \mathcal{H}$.*

Example 17.20. We apply these results to S_1 and S_2. We have that $S_1^*(g) = S_2^*(g) = -ig'$ for g in their domains and $\mathcal{D}(S_1^*) = \{g \in AC([0,1]) : g(0) = g(1) = 0\}$ while $\mathcal{D}(S_2^*) = AC([0,1])$. The unique solutions to $ig' = \pm ig$ are $g_\pm(x) = e^{\pm x}$. Both these functions belong to $\mathcal{D}(S_2^*)$ and so $\mathcal{N}(S_2^* \pm iI) \neq (0)$. Hence, S_2 is not essentially self-adjoint. However, neither of these functions belong to $\mathcal{D}(S_1^*)$. Tempting to think S_1 is essentially self-adjoint, but remember we already showed that it is not even symmetric! So some care must be taken not to fall into a trap!

Here is one reason that these concepts are important. If we imagine that our Hilbert space represents states of some system and that system evolves over time, then we would expect that there is a family of unitaries, $U(t)$ so that if we are in state h_0 at time 0, then at time t we are in state $U(t)h_0$,

Definition 17.21. A family of operators $\{U(t) : t \in \mathbb{R}\}$ on a Hilbert space $\mathcal{H}$ is called a **strongly continuous one-parameter unitary group** provided that:
(1) $U(t)$ is a unitary operator for all $t \in \mathbb{R}$,
(2) for each $h \in \mathcal{H}$ the function $t \to U(t)h$ is continuous,
(3) $\forall t, s \in \mathbb{R}$, $(t)U(s) = U(t+s)$.

Theorem 17.22 (Stone's Theorem)**.** *Let A be a self-adjoint operator on $\mathcal{H}$ and use the functional calculus to define $U(t) = e^{iAt}$. Then $U(t)$ is a strongly continuous one-parameter unitary group and for $h \in \mathcal{D}(A)$, $\|\frac{U(t)h-h}{t} - iAh\| \to 0$ as $t \to 0$. Conversely, let $U(t)$ be a strongly continuous one-parameter unitary group, set*

$$\mathcal{D}(A) = \{h \in \mathcal{H} : \lim_{t\to 0} \frac{U(t)h - h}{t} \textit{exists}\},$$

and set

$$Ah = -i \lim_{t \to 0} \frac{U(t)h - h}{t}.$$

Then A is self-adjoint and $U(t) = e^{iAt}$.

One instructive example is to consider $\mathcal{H} = L^2(\mathbb{R})$ and set $[U(t)f](x) = f(x+t)$. Then $f \in \mathcal{D}(A)$ precisely, when the difference quotients, $\frac{f(x+t)-f(x)}{t}$ converge in norm to a square-integrable function. That function must be equal to f' almost everywhere. Again measure theory shows that $\mathcal{D}(A) = \{f \in AC(\mathbb{R}) : f' \in L^2(\mathbb{R})\}$. Thus, $Af = -if'$.

18. Von Neumann Algebras

These were originally called "Algebras of Operators" by von Neumann but their name has since been changed in his honor. This is a difficult theory, but his motivation for studying these was quantum mechanics, so we start with that motivation.

In quantum mechanics the unit vectors in a Hilbert space represent the states of a system. Suppose that the system is in state $\psi \in \mathcal{H}$ and we want to perform a measurement that has K possible outcomes. Then the theory says that there will exist K measurement operators, $M_1, ...M_K \in B(\mathcal{H})$ such that:

- the probability of getting outcome k is $p_k = \|M_k\psi\|^2$,
- if outcome k is observed then the state of the system changes to $\frac{M_k\psi}{\|M_k\psi\|}$.

The fact that $1 = \sum_{k=1}^K p_k$ implies that $I = \sum_{k=1}^K M_k^* M_k$. Von Neumann argued that in certain settings the underlying state space could have a family of unitaries that acted upon it $\{U_a : a \in A\}$ such that if the system was in state $U_a\psi$ and we observed outcome k, then the state of the system should change to $\frac{U_a M_k \psi}{\|M_k U_a \psi\|}$. This implies that $U_a M_k = M_k U_a, \forall a \in A$.

This lead him to study sets of operators that commute with a set of unitary operators. He denoted such sets by $\mathcal{M}$ since this is what he thought sets of measurement operators should look like. For this reason von Neumann algebras are still generally denoted by the letter $\mathcal{M}$.

Definition 18.1. Given a set $\mathcal{S} \subseteq B(\mathcal{H})$ we call the set

$$\mathcal{S}' = \{T \in B(\mathcal{H}) : TS = ST, \forall S \in \mathcal{S}\},$$

the **commutant** of $\mathcal{S}$.

Note that

$$\mathcal{S}' = \mathcal{S}''',$$

and

$$\mathcal{S}'' = \mathcal{S}''''.$$

Note that if U is a unitary and $UT = TU$, then $TU^* = U^*(UT)U^* = U^*(TU)U^* = U^*T$, so T commutes with the adjoint too. Also T always commutes with I.

We briefly recall weak, strong and weak* convergence. A net of operators $\{T_\lambda\}_{\lambda \in D}$ converges to T in

- the weak-topology if $\langle k|T_\lambda h\rangle \to \langle k|Th\rangle$ for all $h, k \in \mathcal{H}$,
- the strong-topology if $\|T_\lambda h - Th\| \to 0$ for all $h \in \mathcal{H}$,
- the weak*-topology if $Tr(T_\lambda K) \to Tr(TK)$ for all $K \in \mathcal{C}_1(\mathcal{H})$.

We use $\mathcal{S}^{-w}$, $\mathcal{S}^{-s}$ and $\mathcal{S}^{-wk*}$, to denote the sets of operators that are limits of nets of operators from $\mathcal{S}$ in the weak, strong, and weak* sense.

Recall that a set $\mathcal{A}$ is called an algebra if it is a vector space and $X, Y \in \mathcal{A} \implies XY \in \mathcal{A}$.

Proposition 18.2. *Let $\mathcal{S} \subseteq B(\mathcal{H})$ be a set such that $I \in \mathcal{S}$ and $X \in \mathcal{S} \implies X^* \in \mathcal{S}$. Then $\mathcal{S}'$ is an algebra of operators that is closed in the weak, strong and weak* topologies.*

Theorem 18.3 (von Neumanns bicommutant Theorem)**.** *Let $\mathcal{A} \subseteq B(\mathcal{H})$ be an algebra of operators such that $I \in \mathcal{A}$ and $X \in \mathcal{A} \implies X^* \in \mathcal{A}$. Then $\mathcal{A}'' = \mathcal{A}^{-w} = \mathcal{A}^{-s} = \mathcal{A}^{-wk*}$.*

Thus, not only are operators that can be realized as limits in these three senses all equal, but something defined purely algebraically, $\mathcal{A}''$ is equal to something defined as a topological closure.

Corollary 18.4. *Let $\mathcal{M} \subseteq B(\mathcal{H})$ be an algebra of operators such that $I \in \mathcal{M}$ and $X \in \mathcal{M} \implies X^* \in \mathcal{M}$. The following are equivalent:*

- $\mathcal{M} = \mathcal{M}''$,
- $\mathcal{M} = \mathcal{M}^{-s}$,
- $\mathcal{M} = \mathcal{M}^{-w}$,
- $\mathcal{M} = \mathcal{M}^{-wk*}$.

Corollary 18.5. *Let $\mathcal{S} \subseteq B(\mathcal{H})$ satisfy $I \in \mathcal{S}$ and $X \in \mathcal{S} \implies X^* \in \mathcal{S}$. Then $\mathcal{S}''$ is the smallest weak* closed algebra containing $\mathcal{S}$.*

Definition 18.6. Any $\mathcal{M} \subseteq B(\mathcal{H})$ satisfying $I \in \mathcal{M}$, $X \in \mathcal{M} \implies X^* \in \mathcal{M}$ and $\mathcal{M} = \mathcal{M}''$ is called a **von Neumann algebra**.

Murray and von Neumann set about to classify all such algebras. This program goes on to this day and has an influence on quantum mechanics, which we will try to explain. First, we discuss the classification program.

Definition 18.7. A von Neumann algebra $\mathcal{M} \subseteq B(\mathcal{H})$ is called a **factor** if $\mathcal{M} \cap \mathcal{M}' = \mathbb{C} \cdot I$.

Just like all integers can be decomposed into products of primes, Von Neumann proved that every von Neumann algebra could be built from factors by something called he called **direct integration theory**. Hence, to understand all von Neumann algebras we only need to understand all factors.

The next step that Murray and von Neumann made was to break factors down into three types.

Definition 18.8. Let $\mathcal{M}$ be a von Neumann algebra and let $E, F \in \mathcal{M}$ be projections. We write $E \leq F$ if $\mathcal{R}(E) \subseteq \mathcal{R}(F)$ and $E < F$ when $E \leq F$ and $E \neq F$. We say that $F \neq 0$ is **minimal** if $E < F \implies E = 0$. We say that E and F are **Murray-von Neumann equivalent** and write $E \sim F$ if there exists a partial isometry $W \in \mathcal{M}$ such that $E = W^*W, F = WW^*$. We say that F is **finite** if there is no E such that $E \sim F$ and $E < F$.

Some examples are useful.

Example 18.9. Let $\mathcal{M} = B(\mathbb{C}^m)$. Then F is minimal if and only if F is a rank one projection, $E \sim F$ if and only if E and F are projections onto subspaces of the same dimension. Hence, every projection is finite.

Example 18.10. Let $\mathcal{M} = B(\mathcal{H})$, where $\mathcal{H}$ is infinite dimensional. Again a projection is minimal if and only if it is rank one. Since a partial isometry with $E = W^*W, F = WW^*$ is an isometry from $\mathcal{R}(E)$ to $\mathcal{R}(F)$ and isometries preserve inner products, they take an onb for one space to an onb to the other space. Hence, $E \sim F \iff dim_{HS}(\mathcal{R}(E)) = dim_{HS}(\mathcal{R}(F))$. On the other hand as soon as a set is infinite we can throw away one element and the subset has the same cardinality. Hence, as soon as $\mathcal{R}(F)$ is infinite dimensional, we can take an onb, throw away one element and that will be an onb for a subspace $\mathcal{R}(E) \subseteq \mathcal{R}(F)$ of the same dimension. We can now take a partial isometry that sends the onb for $\mathcal{R}(F)$ to the onb for $\mathcal{R}(E)$. This shows that $E \sim F$ with $E < F$. Hence, F is NOT finite in the M-vN sense.

Hence, the only finite projections are the projections onto finite dimensional subspaces.

Definition 18.11. A von Neumann factor is called **Type I** if it has a minimal projection. It is **Type II** if it has no minimal projections, but has a finite projection. It is called **Type** II_1 is it is Type II and the identity is a finite projection. If it is Type II but not Type Type II_1, then it is called **Type** II_∞. It is called **Type III** if it is not Type I or Type II.

Theorem 18.12 (Murray-von Neumann)**.** *A von Neumann algebra is Type I if and only if it is *-isomorphic to $B(\mathcal{H})$ for some $\mathcal{H}$.*

Definition 18.13. Let $\mathcal{M}$ be a von Neumann algebra. A map $\tau : \mathcal{M} \to \mathbb{C}$ is called a **state** is $\tau(I) = 1$ and $p \geq 0 \implies \tau(p) \geq 0$. A map is called a **tracial state** if it is a state and satisfies $\tau(XY) = \tau(YX)$. It is **faithful** if $\tau(X^*X) = 0 \implies X = 0$.

For $B(\mathbb{C}^n)$, there is only one faithful tracial state, namely $\tau(X) = \frac{1}{n}Tr(X)$. Note that in this setting the possible traces of projections are the numbers $\{\frac{k}{n} : 0 \leq k \leq n\}$, which represent the "fractional" dimension of the corresponding subspace.

Theorem 18.14 (Murray-von Neumann)**.** *Let $\mathcal{M}$ be a Type II_1 factor. Then:*

- *there exists a faithful tracial state, $\tau : \mathcal{M} \to \mathbb{C}$ that is also continuous in the weak*-topology,*
- *for every $0 \leq t \leq 1$ there exists a projection $P \in \mathcal{M}$ with $\tau(P) = t$,*
- *if $P, Q \in \mathcal{M}$ are projections, then $P \sim Q \iff \tau(P) = \tau(Q)$.*

This lead von Neumann to refer to "continuous geometries" since, unlike finite dimensions, in a Type II_1 setting there are subspaces of every (fractional) dimension $t, 0 \leq t \leq 1$.

Institute for Quantum Computing and Department of Pure Mathematics, University of Waterloo, Waterloo, ON, Canada N2L 3G1
E-mail address: vpaulsen@uwaterloo.ca