Functional Analysis

Alex Rutar*
University of Waterloo

Fall 2019†

*arutar@uwaterloo.ca
†Last updated: December 17, 2019

Contents

I. Analysis in Metric Spaces

0.1 Theorem. (Baire Category I) *If (X,d) is a complete metric space and $\{U_n\}_{n=1}^{\infty}$ is a countable collection of dense, open subsets, then $\bigcap_{n=1}^{\infty} U_n$ is dense in X.*

Definition. Let (X,d) be a metric space. A subset $F \subset X$ is **nowhere dense** if $X \setminus F$ is dense in X; equivalently, $\overline{F}$ contains no non-trivlal open subsets. We say that a subset $M \subseteq X$ is **meagre** (1st category) if $M = \bigcup_{n=1}^{\infty} F_n$ and each F_n is nowehere dense; and a set is **non-meagre** (2nd category) otherwise.

0.2 Theorem. (Baire Category II) *Let (X,d) be a complete metric space. Then a non-empty open $U \subseteq X$ is non-meagre.*

Proof Suppose not, so $U = \bigcup_{n=1}^{\infty} F_n \subseteq \bigcup_{n=1}^{\infty} \overline{F}_n$, each F_n (hence $\overline{F_n}$) nowhere dense. Then each $V_n = X \setminus \overline{F_n}$ is open and dense, and hence by BCT I, $G = \bigcap_{n=1}^{\infty} V_n$ is dense in X, and hence $U \cap G \neq \emptyset$, violating assumption ■

II. Set Theory and Topology

1 The Axiom of Choice

Definition. Let S be a non-empty set. A **partial ordering** is a binary relation $\leq$ on S which satisfies for $s, t, u \in S$,

(i) *(reflexivity)* $s \leq s$

(ii) *(transitivity)* $s \leq t$, $t \leq u$ implies $s \leq u$

(iii) *(anti-symmetry)* $s \leq t$, $t \leq s$ implies $s = t$

We call the pair $(S, \leq)$ a **partially ordered set**. We say that $(S, \leq)$ is **totally ordered** if, given $s, t \in S$, at least one of $s \leq t$ or $t \leq s$ holds. We say that $(S, \leq)$ is **well-ordered** if given any $\emptyset \neq S_0 \subseteq S$, there is some $s_0 \in S_0$ such that $s_0 \leq s$ for $s \in S_0$. A **chain** in a poset $(S, \leq)$ is any $\emptyset \neq C \subseteq S$ such that $(S, \leq|_C)$ is totally ordered.

Example. (i) $X \neq \emptyset$, $(\mathcal{P}(X), \subseteq)$ is a poset

(ii) $(\mathbb{R}, \leq)$ is a totally ordered set

(iii) $(\mathbb{N}, \leq)$, $(\omega = \mathbb{N} \cup \{\infty\}, \leq)$, are well-ordered sets.

1.1 Theorem. *The following are equivalent:*

(i) (Axiom of Choice 1): For any $x \neq \emptyset$, there is a function $\gamma : \mathcal{P}(X) \setminus \{\emptyset\} \to X$ such that $\gamma(A) \in A$ for each $A \in \mathcal{P})X_\setminus\{\emptyset\}$.

(ii) (Axiom of Choice 2): Given any $\{A_\lambda\}_{\lambda \in \Lambda}$ where $A_\lambda \neq \emptyset$ for each λ,

$$\prod_{\lambda \in \Lambda} A_\lambda = \{(a_\lambda)_{\lambda \in \Lambda} : a_\lambda \in A_\lambda \text{ for each } \lambda\} \neq \emptyset.$$

(iii) (Zorn's Lemma): In a partially ordered set $(S, \leq)$, if each chain $C \subseteq S$ admits an upper bound in S, then $(S, \leq)$ admits a maximal element.

(iv) (Well-ordering principle): Any $S \neq \emptyset$ admits a well-ordering.

Proof [***TODO: Write up proof***] ∎

2 Cardinal Arithmetic

[***TODO: Write up***]

3 Topology

3.1 Basic Definitions and Examples

Let X denote a non-empty set, and $\mathcal{P}(X)$ denote the power set of X.

Definition. A **topology** on a set X is a set τ of subsets of X such that

(i) $\emptyset, X \in \tau$

(ii) If $U_\alpha \in \tau$ for all $\alpha \in A$, then $\bigcup_{\alpha \in A} U_\alpha \in \tau$.

(iii) If $n \in \mathbb{N}$ and $U_i \in \tau$ for each $1 \leq i \leq n$, then $\bigcap_{i=1}^{n} U_i \in \tau$.

The sets $U \in \tau$ are the **open sets** in X, and sets $X \setminus U$ for some open set U are the **closed sets** in X. The pair (X, τ) is called a **topological space**.

Example. (i) *Sorgenfry line:* Set $X = \mathbb{R}$, and consider

$$\sigma = \{ V \subseteq \mathbb{R} \mid \text{ for any } s \in V, \text{ there is } \delta = \delta(s) > 0 \text{ s.t. } [s, s+\delta) \subseteq V \}$$

It is a straightforward exercise to verify that $\tau_{|\cdot|} \subsetneq \sigma$. We say that σ is **finer** than $\tau_{|\cdot|}$.

(ii) *Relative or subset topology*: let (X, τ) be a topological space and $\emptyset \neq A \subseteq X$. Then we can define a topology $\tau|_A = \{U \cap A : U \in \tau\}$.

(iii) *Metric topology:* A metric space (X, d) is naturally a topological space, where the topology is given by

$$\tau_d = \{ U \subseteq X \mid \text{ for each } x_0 \in U, \text{ there is } \delta = \delta(x_0) \text{ s.t. } B_\delta(x_0) \subseteq U \}.$$

Given two metrics d, ρ on X, we say that $d \sim \rho$ are **equivalent** if and only if there are $c, C > 0$ such that

$$cd(x,y) \leq \rho(x,y) \leq Cd(x,y) \text{ for any } x, y \in X$$

Note that $d \sim \rho$ implies that $\tau_d = \tau_\rho$, but the reverse implication is not true: for example, the metrics on $\mathbb{R}$ given by $d(x,y)$ and $\rho(x,y) = \frac{|x-y|}{1+|x-y|}$ are not equivalent.

Let (X, d), (Y, ρ) be metric spaces. A map $f : X \to Y$ is an **isometry** if for any $x, y \in X$, $d(x,y) = \rho(f(x), f(y))$. By non-degeneracy, f is automatically injective. In particular, when (X, d) is complete, then $(f(X), \rho|_{f(X)})$ is a complete metric space.

Definition. A set $F \subseteq X$ is **closed** if $X \setminus F \in \tau$. If $S \subseteq X$, then the **closure** of S is $\overline{S} = \cap\{F \subseteq X : S \subseteq F, X \setminus F \in \tau\}$.

Note that $\overline{S} = \{x \in X : \text{for any } U \in \tau \text{ with } x \in U, U \cap S \neq \emptyset\}$.

Definition. Let (X, τ) and (Y, σ) be topological spaces, and $f : X \to Y$. We say that f is **$(\tau - \sigma -)$continuous at x_0** in X if for any $V \in \sigma$ such that $f(x_0) \in V$, then there exists $U \in \tau$ such that $x_0 \in U$ and $f(U) \subseteq V$. We say that f is **$(\tau - \sigma -)$continuous** if it is continuous at each x_0 in X.

An easy application of definitions yields the following:

3.1 Proposition. *Let (X, τ), (Y, σ) be topological spaces and $f : X \to Y$. Then f is continuous if and only if for any $U \in \sigma$, $f^{-1}(U) \in \tau$.*

3.2 Lemma. *If $x_0 \in X$ where (X, τ) is a topological space, then*

$$\mathcal{I}(x_0) = \{ f \in C_b(X) \mid f(x_0) = 0 \}$$

is closed, hence complete, subspace of $C_b(X)$.

Proof If $(f_n)_{n=1}^{\infty} \subseteq \mathcal{I}(x_0)$ and $f = \lim_{n\to\infty} f_n$ with respect to $\|\cdot\|_\infty$ in $C_b(X)$, then $f(x_0) = \lim_{n\to\infty} f_n(x_0) = 0$. Thus $f \in \mathcal{I}(x_0)$, and closed subsets of complete spaces are themselves complete. ■

An important class of topologies which are particularly tractable are the Hausdorff spaces:

Definition. A topological space (X,τ) is **Hausdorff** if given $x \neq y$ in X, there are $U_x, V_y \in \tau$ such that $x \in U_x$ and $y \in V_y$ and $U_x \cap U_y = \emptyset$.

Example. (i) A metric space is Hausdorff.

(ii) X a normed space, $w = \sigma(X, X^*)$ is Hausdorff (by Hahn-Banach and A2Q1).

(iii) If X is a normed space, then $w^* = \sigma(X^*, \hat{X})$ on X^* is Hausdorff.

(iv) $\{(X_\alpha, \tau_\alpha)\}_{\alpha\in A}$ family of topological spaces, $X = \prod_{\alpha\in A} X_\alpha$ with π the product topology. Then (X,π) is Hausdorff if and only if all (X_α, τ_α) are Hausdorff. (Straightfoward exercise).

3.2 Compactness

Let (X,τ) be a topological space.

Definition. A subset $K \subseteq X$ is called **compact** if for any collection $\{U_\alpha\}_{\alpha\in A} \subseteq \tau$ with $\bigcup_{\alpha\in A} U_\alpha \supseteq K$, there exists some finite $U_1, \ldots, U_n$ covering K. If X itself is τ−compact, we call (X,τ) a compact space.

A basic application of the definitions yields the following:

3.3 Proposition. *Let (X,τ) be a compact space.*

(i) If $K \subseteq X$ is closed, then K is compact.

(ii) If (Y,σ) is a topological space and $f : X \to Y$ is continuous, then $f(X)$ is compact.

Interestingly, if the τ is a Hausdorff topology, we also have a converse for (i), which is again left as an interesting exercise:

3.4 Proposition. *Let (X,τ) be a Hausdorff space, $K \subseteq X$ τ−compact. Then K is τ−closed.*

This yields the following nice result for compact subsets of Hausdorff spaces.

3.5 Proposition. *Let (X,τ) be a compact space.*

(i) If (Y,σ) is a Hausdorff space and $\phi : X \to Y$ is continuous and bijective, then $\phi^{-1} : Y \to X$ is continuous.

(ii) If $\tau' \subseteq \tau$ is a Hausdorff topology on X, so $\tau' = \tau$.

Proof (i) If $F \subset X$ is closed, it is compact, so that $(\phi^{-1})^{-1}\phi(F) \subset Y$ is compact, and hence closed.

(ii) $\mathrm{id} : X \to X$ is continuous, so if $U \in \tau'$, then $\mathrm{id}^{-1}(U) = U \in \tau$, so id is continuous. Hence by (1) id^{-1} is continuous so $\tau \subseteq \tau'$. ■

Here we give an alternative way to characterize compactness in a topological space:

Definition. A family $\mathcal{F} \subseteq \mathcal{P}(X)$ has the **finite intersection property** if for any $F_1, \ldots, F_n \in \mathcal{F}$, $\bigcap_{l=1}^n F_k \neq \emptyset$.

3.6 Proposition. *Let (X,τ) be a topological space. Then (X,τ) is compact if and only if any $\mathcal{F} \subseteq \mathcal{P}(X)$ with the finite intersection property has $\bigcap_{F\in\mathcal{F}} \overline{F} \neq \emptyset$.*

Proof Suppose X is compact and $\mathcal{F} \subset \mathcal{P}(X)$ has the finite intersection property but with $\bigcap_{F\in\mathcal{F}} \overline{F}$, then $\{X \setminus \overline{F}\}_{F\in\mathcal{F}}$ is an open cover of X with no finite subcover.

Conversely, if $\mathcal{O} \subseteq \tau$ is an open cover of X, then $\mathcal{F} = \{X \setminus U\}_{U\in\mathcal{O}}$ satisfies $\bigcap_{F\in\mathcal{F}} = \emptyset$, so there is $F_1, \ldots, F_n \in \mathcal{F}$ with $\bigcap_{k=1}^n F_k = \emptyset$. Then $\{X \setminus F_i\}_{i=1}^k$ is a finite subcover. ■

3.3 Bases and Initial Topologies

Definition. Let (X, τ) be a topological space. A **base** for τ is any family $\beta \subseteq \tau$ such that for any $U \in \tau$ and $x \in U$, there is $B \in \beta$ such that $x \in B \subseteq U$. A **subbase** for τ is any family $\alpha \subseteq \tau$ such that $\{\bigcap_{k=1}^{n} U_k : n \in \mathbb{N}, U_1, \ldots, U_n \in \alpha\}$ is a base for τ.

Note that if $\emptyset \neq X$ and $\beta \subseteq \mathcal{P}(X)$ for which $\bigcup_{B\in\beta} B = X$ and β is closed under finite intersections, then

$$\tau_\beta = \left\{ \bigcup_{i\in I} B_i : \{B_i\}_{i\in I} \subset \beta \right\}$$

defines a topology.

Definition. Let $X \neq \emptyset$. Suppose we are given

- a family $\{(X_\alpha, \tau_\alpha)\}_{\alpha\in A}$ of topological spaces, and
- for each $\alpha \in A$, a function $f_\alpha : X \to X_\alpha$

Then the **initial topology** on X given this data is denoted

$$\sigma = \sigma(X, (f_\alpha)_{\alpha\in A}) = \sigma(X, (f_\alpha, \tau_\alpha)_{\alpha\in A})$$

and is the topology with base

$$\beta = \left\{ \bigcap_{k=1}^{n} f_{\alpha_k}^{-1}(U_{\alpha_k}), n \in \mathbb{N}, \alpha_1, \ldots, \alpha_n \in A, \text{ each } U_{\alpha_k} \in \tau_{\alpha_k} \right\}$$

In particular, $\{f_\alpha^{-1}(U_\alpha) : U_\alpha \in \tau_\alpha, \alpha \in A\}$ is a subbase for σ.

Remark. The topology is chosen so that each $f_\alpha : X \to X_\alpha$ is $\sigma - \tau_\alpha$–continuous. Furthermore, if $\tau \subseteq \mathcal{P}(X)$ is any topology for which every f_α is $\sigma - \tau_\alpha$–continuous, then $\sigma \subseteq \tau$. We say that σ is the **coarsest** topology so that all the f_α are continuous.

A number of common topologies can be defined using this notation:

Example. (i) *Metric topology:* If (X, d) is a metric space, for each $x \in X$, let d_x be given by $d_x(x') = d(x, x')$. Then $\sigma(X, (d_x)_{x\in X}) = \tau_d$.

(ii) *Relative topology:* If (Y, τ)–topological space, $\emptyset \neq X \subseteq Y$, and $i : X \to Y$ is the inclusion map. Then $\tau|_X = \sigma(X, \{i\})$.

(iii) *Product topology:* Let $\{(X_\alpha, \tau_\alpha)\}_{\alpha\in A}$ be a family of topological spaces. Let $X = \prod_{\alpha\in A} X_\alpha$. Let for $\alpha \in A$, $p_\alpha : X \to X_\alpha$ denote the projection map onto the component α. Then the product topology $\pi = \sigma(X, \{p_\alpha\}_{\alpha\in A})$. Hence, $V \in \mathcal{P}(X)$, then $V \in \pi$ if and only if for any $x \in V$, there is $\alpha_1, \ldots, \alpha_n \in A$ and $U_{\alpha_k} \in \tau_{\alpha_k}$ such that $x_{\alpha_k} = p_{\alpha_k(x)} \in U_{\alpha_k}$ and $x \in \bigcap_{k=1}^{n} p_{\alpha_k}^{-1}(U_{\alpha_j}) \subseteq V$.
Note that if $X = \prod_{n=1}^{\infty} X_n$, each (X_n, τ_n) is a topological space, then the basic open sets look like $U_1 \times U_2 \times \cdots \times U_m \times X_{m+1} \times X_{m+2} \times \cdots$.

3.4 Ultrafilters and Tychonoff's Theorem

Definition. Let X be a non-empty set. An **ultrafilter** is a family $\mathcal{U} \subset \mathcal{P}(X)$ such that

- $\mathcal{U}$ has the finite intersection property
- If $A \in \mathcal{P}(X)$, then either $A \in \mathcal{U}$ or $X \setminus A \in \mathcal{U}$.

Example. (i) *Principal / trivial ultrafilter:* If $x_0 \in X$, let $U_{x_0} = \{U \subseteq X : x_0 \in U\}$.

3.7 Lemma. (Ultrafilter) *If $\mathcal{F} \subseteq \mathcal{P}(X)$ is any set with the finite intersection property, then there is an ultrafilter $\mathcal{U}$ with $\mathcal{F} \subset \mathcal{U}$.*

Proof Let $\Phi = \{\mathcal{G} \subseteq \mathcal{P}(X) : \mathcal{F} \subseteq \mathcal{G}, \mathcal{G} \text{ has f.i.p.}\}$. Then Φ is partially ordered by inclusion. If $\Gamma \subseteq \Phi$ is a chain, then $\mathcal{G}_\Phi = \bigcup_{\mathcal{G}\in\Gamma} \mathcal{G}$ contains $\mathcal{F}$ and has the finite intersection property. Hence Φ admits a maximal element $\mathcal{U}$. Let $A \in \mathcal{P}(X) \setminus \mathcal{U}$. Then $U \cup \{A\} \supsetneq \mathcal{U}$, so $\mathcal{U} \cup \{A\}$ fails the finite intersection property. Hence get $U_1, \ldots, U_n$ so $A \cap \bigcap_{k=1}^n U_k = \emptyset$. Now if $V_1, \ldots, V_m \in \mathcal{U}$, then $\bigcap_{j=1}^n V_j \cap \bigcap_{k=1}^n U_j \subseteq \bigcap_{k=1}^n U_k \subseteq X \setminus A$, so $(X \setminus A) \cap \bigcap_{j=1}^m V_j$. Thus $\mathcal{U} \cup \{X \setminus A\}$ has finite intersection property, so $X \setminus A \in \mathcal{U}$ by maximality. ■

3.8 Corollary. *If $U \subseteq \mathcal{P}(X)$ is an ultrafilter, then*

(i) If $A \in \mathcal{P}(X)$, $A \in \mathcal{U}$ if and only if $A \cap U \neq \emptyset$ for each $U \in \mathcal{U}$

(ii) If $A, B \in \mathcal{P}(X)$, then $A \cup B \in \mathcal{U}$ implies at least one of A or B is in $\mathcal{U}$

(iii) If $A \in \mathcal{U}$ and $A \subseteq V$ implies $V \in \mathcal{U}$

Proof The forward implication of (i) follows since $\mathcal{U}$ has finite intersection. Conversely, $X \setminus A \notin \mathcal{U}$, so $A \in \mathcal{U}$. (ii) and (iii) follow consequently. ■

3.9 Corollary. *If X is an infinite set, it admits a non-principle ultrafilter.*

Proof Let $\mathcal{F} = \{F \in \mathcal{P}(X) : X \setminus F \text{ is finite}\}$. Then $\mathcal{F}$ has the finite intersection property. Apply the lemma. ■

3.10 Proposition. *There are at least $\mathfrak{c}$ many ultrafilters in $\mathcal{P}(\mathbb{N})$.*

Proof We let $\mathcal{F} \subset \mathcal{P}(\mathbb{N})$ be a collection of infinite sets such that $E \neq F$ in $\mathcal{F}$ implies $|E \cap F| < \infty$, and $|\mathcal{F}| = \mathfrak{c}$. For each $F \in \mathcal{F}$, we let $\mathcal{F}_F = \mathcal{F}_0 \cup \{F\}$, which has the finite intersection property. Moreover, if $E \in \mathcal{F} \setminus \{F\}$, then $\mathcal{F}_F \cup \{E\}$ would fail f.i.p. Hence, for $F \in \mathcal{F}$, let $\mathcal{U}_F$ be any ultrafilter containing $\mathcal{F}_F$, giving $\mathfrak{c}$ many ultrafilters. ■

Remark. It can be shown (with a lot more work) that $\mathbb{N}$ admits $2^{\mathfrak{c}}$ ultrafilters.

Let $\mathcal{U} \subset \mathcal{P}(\mathbb{N})$ be a non-principal ultrafilter. Define $\delta_{\mathcal{U}} : \mathcal{P}(\mathbb{N}) \to \{0,1\} \subset \mathbb{R}$ by $\delta_{\mathcal{U}}(A) = 1$ if $A \in \mathcal{U}$, and 0 if $X \setminus A \in \mathcal{U}$. Since $\mathbb{N} \in \mathcal{U}$, we see that $\delta_{\mathcal{U}}(\emptyset) = 0$. If $\emptyset \neq A, B \in \mathcal{P}(\mathbb{N})$ with $A \cap B = \emptyset$, then if $A \cup B \in \mathcal{U}$, then exactly one of A or B is in $\mathcal{U}$. Thus $\delta_U(A \cup B) = \delta_U(A) + \delta_U(B)$. If $E_1, \ldots, E_n \subseteq \mathbb{N}$ with $E_j \cap E_k = \emptyset$ for $j \neq k$, then $\sum_{k=1}^n |\delta_{\mathcal{U}}(E_k)| \leq 1$ so $\|\delta_{\mathcal{U}}\|_{\text{var}} \leq 1$. Since $\delta_{\mathcal{U}}(\mathbb{N}) = 1$, we have $\|\delta_{\mathcal{U}}\|_{\text{var}} = 1$. Let $L_{\mathcal{U}} \in \ell_\infty^*$ be the linear functional associated to $\delta_{\mathcal{U}}$. We then have (with some verification possibly needed)

(i) $L_{\mathcal{U}}(1) = 1$, $\|L_{\mathcal{U}}\| = 1$

(ii) $L_{\mathcal{U}}|_{\mathbf{c}_0} = 0$, so if $x \in \ell_\infty^{\mathbb{R}}$, then $\liminf_{n\to\infty} x_n \leq L_{\mathcal{U}} \leq \limsup_{n\to\infty} x_n$

(iii) Exactly one of $2\mathbb{N}$ and $2\mathbb{N}-1$ is in $\mathcal{U}$, so $L(\chi_{2\mathbb{N}}) \neq L_{\mathcal{U}}(\chi_{2\mathbb{N}-1})$, so $L_{\mathcal{U}}$ is not translation invariant.

(iv) Let $S \in \mathcal{B}(\ell_\infty)$ be given by $Sx = \left(\frac{x_1+\cdots+x_n}{n}\right)_{n=1}^\infty$. Then $L_{\mathcal{U}} \circ S$ is a Banach limit.

Definition. If (X, τ) is a topological space, $\mathcal{U}$ an ultrafilter on X, we say that $x_0 \in X$ is a $(\tau-)$limit point for $\mathcal{U}$ if for each $U \in \tau$ with $x_0 \in U$, we have $U \in \mathcal{U}$.

3.11 Proposition. *Let (X, τ) be a topological space. Then (X, τ) is compact if and only if any ultrafilter on X admits a $\tau-$limit point.*

Proof Let us begin with an observation: if $x \in X$ and $\mathcal{U}$ is an ultrafilter on X, then

$$x \in \bigcap_{V \in \mathcal{U}} \overline{V} \Leftrightarrow \text{for any } U \in \tau \text{ with } x \in U, U \cap V \neq \emptyset \text{ for each } V \in \mathcal{U}$$
$$\Leftrightarrow x \text{ is a } \tau\text{–limit point of } \mathcal{U} \qquad \blacksquare$$

If (X,τ) is compact, then $\bigcap_{V\in\mathcal{U}} \overline{V} \neq \emptyset$. If $\mathcal{F} \subseteq \mathcal{P}(X)$ has the finite intersection property, then there exists an ultrafilter $\mathcal{U} \supseteq \mathcal{F}$, so $\bigcap_{F\in\mathcal{F}} \overline{F} \supseteq \bigcap_{V\in\mathcal{U}} \overline{V} \neq \emptyset$.

3.12 Theorem. (Tychonoff) *Let $\{(X_\alpha,\tau_\alpha)\}_{\alpha\in A}$ be a family of compact spaces, and $X = \prod_{\alpha\in A} X_\alpha$ with the product topology π. Then (X,π) is compact.*

Proof Let $\mathcal{U}$ be an ultrafilter on X; we will show that it admits a π–limit point. Fix $\alpha \in A$ and let $\mathcal{U}_\alpha = \{p_\alpha(V) : V \in \mathcal{U}\}$, where p_α is the coordinate projection onto α. If $\emptyset \neq S_\alpha \subseteq X_\alpha$, then $S_\alpha = p_\alpha^{-1}(p_\alpha^{-1}(S_\alpha))$, so $S_\alpha \in \mathcal{U}_\alpha$ if and only if $p^{-1}(S_\alpha) \in \mathcal{U}$, and since p^{-1} commutes with complementation, $\mathcal{U}_\alpha$ is an ultrafilter. The last proposition provides a τ_α–limit point x_α for $\mathcal{U}_\alpha$. Now let $x = (x_\alpha)_{\alpha\in A}$, where x_α is found as above. If $W \in \pi$ with $x \in W$, then there are $\alpha_1,\dots,\alpha_n$ in A, $U_{\alpha_i} \in \tau_{\alpha_i}$ with $x \in \bigcap_{k=1}^n p_{\alpha_k}^{-1}(U_{\alpha_k}) \subseteq W$. Since each x_{α_k} is a τ_{α_k}–limit point of $\mathcal{U}_{\alpha_k}$, we see that each $U_{\alpha_k} \in \mathcal{U}_{\alpha_k}$, so $p_{\alpha_k}^{-1}(U_{\alpha_k}) \in \mathcal{U}$. Thus we see that $W \in \mathcal{U}$, so x is a π–limit point of $\mathcal{U}$. $\blacksquare$

Remark. (i) Tychonoff's theorem implies the axiom of choice. Given $\{X_\alpha\}_{\alpha\in A}$ be a family of non-empty sets. Find y which is not a member of any X_α, and let $Y_\alpha = X_\alpha \cup \{y\}$ and $\tau_\alpha = \{\emptyset, \{y\}, X_\alpha, Y_\alpha\}$, and $(Y_\alpha.\tau_\alpha)$ is compact. The constant element y is an element of Y, so by Tychonoff, (Y,π) is compact. Given $\alpha_1,\dots,\alpha_n \in A$, then $\bigcup_{k=1}^n p_{\alpha_k}^{-1}(\{y\})$. Since $\prod_{k=1}^n X_{\alpha_k} \neq 0$, we see that $Y \subsetneq \bigcup_{k=1}^n p_{\alpha_k}^{-1}(\{y\})$. Hence by compactness, $Y \not\subseteq \bigcup_{\alpha\in A} p_\alpha^{-1}(\{y\})$. Hence $\prod_{x\in A} X_\alpha = Y \setminus \bigcup_{\alpha\in A} p_\alpha^{-1}(\{y\}) \neq 0$.

(ii) If we are given $(X_\alpha,\tau_\alpha)_{\alpha\in A}$ a family of topological spaces, $X = \prod_{\alpha\in A} X_\alpha$, we can define the **box topology**, i.e. the topology with base $\{\prod_{\alpha\in A} U_\alpha : U_\alpha \in \tau_\alpha \setminus \{\emptyset\} \text{ for each } \alpha\}$ Of course, $\pi \subseteq \tau$, and the inclusion is proper on infinite products.

4 Nets

Definition. A pair $(N,\leq)$ is a **preorder** on N if

- $v \leq v$ for $v \in N$
- $v_1 \leq v_2$ and $v_2 \leq v_3$ implies $v_1 \leq v_3$.

This pair is **cofinal** if for any $v_1, v_2 \in N$, there is $v_3 \in N$ so $v_1 \leq v_3$ and $v_2 \leq v_3$. Then $(N,\leq)$ is a **directed set** if $\leq$ is a cofinal preorder. Given a non-empty set X, a **net** is a function $x : N \to X$.

Definition. If $(x_0)_{v\in N}$ is a net in X, $A \subseteq X$, we say that $(x_0)_{n\in\mathbb{N}}$ is

- **eventually** in A if there is $v_A \in N$ so $x_v \in A$ whenever $v \geq v_A$
- **frequently** in A if for any $v \in N$, there is $v' \in N$ with $v' \geq g$ so $x_{v'} \in A$.

Definition. Now, let $(M,\leq)$ be anther directed set A map $\phi : M \to N$ is **eventually cofinal** if for any $v \in N$, there is $\mu_v \in N$ s $\phi(u) \geq v$ whenever $\mu \geq \mu_0$. Given a net $(x_v)_{v\in N}$ and an eventually cofinal $\phi : M \to N$, we call $(x_{\phi(\mu)})_{\mu\in M}$ a **subnet**.

Definition. We call $\phi : M \to N$ a **directed map** if
(i) $\mu \leq \mu'$ in M implies $\phi(\mu) \leq \phi(\mu')$ in N
(ii) For any $v \in N$, there is $\mu \in M$ s $v \leq \phi(\mu)$.

Directed maps are always cofinal. Different sources use directed maps over eventually cofinal maps.

Example. (i) $(\mathbb{N}, \leq)$ is directed, and subsequences are special types of subnets.
(ii) $(\mathbb{R}, \leq)$ is directed
(iii) (*Riemann sums*) Let $a < b$ in $\mathbb{R}$. We let

$$N = \{(P, P^*) : P = \{a = t_0 < t_1 < \cdots < t_n = b\}, P^* = \{t_1^*, \ldots, t_n^*\}, t_k^* \in [t_{k-1}, t_k]\}$$

and say $(P, P^*) \leq (Q, Q^*)$ if $P \subseteq Q$. One can verify that this is a net (the Riemann sum net).
(iv) (*Nets from filtering families*). We say that $\mathcal{F} \subset \mathcal{P}(X) \setminus \{\emptyset\}$ is a **filtering family** if for each $F_1, F_2 \in \mathcal{F}$, there is $F_3 \in \mathcal{F}$ such that $F_3 \subseteq F_1 \cap F_2$. For example, an ultrafilter is a filtering family. Let

$$N_{\mathcal{F}} = \{(x, F) : x \in F, F \in \mathcal{F}\}$$

equipped with the pre-order $(x, F) \leq (x', F')$ if and only if $F \supseteq F'$. Since $\mathcal{F}$ is a filtering family, $(N_{\mathcal{F}}, \leq)$ is directed. Let $x_{(x,F)} = x$, so $(x)_{(x,F)\in N_{\mathcal{F}}}$ is the net built from $\mathcal{F}$. Note that if $F \in \mathcal{F}$, then $(x)_{(x,F)\in\mathcal{F}}$ is eventually in F.
An **ultranet** $(x_v)_{v\in N} \subset X$ is a net for which any $A \in \mathcal{P}(X)$, $(x_v)_{v\in N}$ is either eventually in A or eventually in $X \setminus A$. If $\mathcal{F}$ is an ultrafilter, then $(x)_{(x,F)\in N_{\mathcal{F}}}$ is an **ultranet**.

4.1 Nets and Topology

Now, suppose (X, τ) is a topological space.

Definition. We say that $x_0 \in X$ is
- Some $x_0 \in X$ is a **limit point** if for any $U \in \tau$ with $x_0 \in U$, $(x_v)_{v\in N}$ is eventually in U. That is, there is ν_U such that $x_\nu \in U$ whenever $\nu \geq \nu_U$. We write $x_0 = \lim_{v\in N} x_v$, the τ–limit of $(x_v)_{v\in N}$. Note that this is an abuse of notation, since limit points need not be unique (when (X, τ) is not Hausdorff).
- Some $x_0 \in X$ is a **cluster point** of $(x_v)_{v\in N}$ if for any $U \in \tau$ with $x_0 \in U$, $(x_v)_{v\in N}$ is frequently in U.

4.1 Proposition. *If $(x_v)_{v\in N}$ is a net in (X, τ) and $x_0 \in X$, then x_0 is a cluster point for $(x_v)_{v\in N}$ if and only if x_0 is a τ–limit point of x_{v_μ} for some subnet $(x_{v_\mu})_{\mu\in M}$ of $(x_v)_{v\in N}$.*

PROOF ($\Longrightarrow$) Suppose x_0 is a cluster point for $(x_\nu)_{\nu\in N}$. Then for each $v \in N$ and $U \in \tau$ containing x_0, define

$$F_{\nu,U} = \{\nu' \in N : \nu' \geq \nu, x_{\nu'} \in U\}$$

which is non-empty since x_0 is a cluster point. Then set

$$\mathcal{F} = \{F_{\nu,U} : \nu \in N, U \in \tau, x_0 \in U\} \subset \mathcal{P}(N)$$

Let's see that $\mathcal{F}$ is filtering: suppose $F_{\nu,U}$ and $F_{\nu',U'}$ are in $\mathcal{F}$. Get $\mu \geq \nu$ and $\mu \geq \nu'$ by definition of a net and set $V = U \cap U'$, which is open and contains x_0. Then since x_0 is a cluster point, get some $\mu' \geq \mu$ such that $x_{\mu'} \in V$, so $F_{\mu',V} \subseteq F_{\nu,U} \cap F_{\nu',U'}$ We then let $M = N_{\mathcal{F}}$ be the net construction from the filtering family and set $v_{(\nu,F)} = V$.

Now set $N_{\mathcal{F}} = \{(\nu, F) : \nu \in F, F \in \mathcal{F}\}$ with the standard preorder and $\nu_{(\nu,F)} = \nu$. Then the map $(\nu, F) \mapsto \nu$ from $N_{\mathcal{F}} \to N$ is eventually cofinal: if $\nu_0 \in N$ is arbitrary, take any $F_0 = F_{\nu_0,U} \in \mathcal{F}$. Then $F_0 = \{\nu \in N : \nu \geq \nu_0, x_\nu \in U\}$, so if $F_{\mu,V} \in \mathcal{F}$ with $F_{\mu,V} \subsetneq$ We let $M = N_{\mathcal{F}}$ as in (iv) above, and $v_{v,\mathcal{F}} = v$. Check that $(x_v)_{(v,F)\in N_{\mathcal{F}}}$ is eventually in U for any $U \in \tau$ with $x_0 \in U$. [Check: $(v,F) \mapsto v : N_{\mathcal{F}} \to N$ is cofinal, but is not evidently directed]

($\Leftarrow$) If for some subnet $(x_{v_\mu})_{\mu\in M}$ is eventually in U for any $U \in \tau$ with $x_0 \in U$, then $(x_v)_{v\in N}$ is frequently in U for such U by definition of a subnet. ∎

4.2 Proposition. *If (Y, σ) is another topological space, then $f : X \to Y$ is continuous if and only if for any $x_0 \in X$ and net $(x_v)_{v\in N}$ with having x_0 as a limit, $f(x_0) = \lim_{v\in N} f(x_v)$.*

Proof If $V \in \sigma$ with $f(x_0) \in V$, then $f^{-1}(V) \in \tau$ with $x_0 \in f^{-1}(V)$. Since $(x_v)_{v\in N}$ is eventually in $f^{-1}(V)$, so $(f(x_v))_{v\in N}$ is eventually in V.

Conversely, let $\tau_{x_0} = \{U \in \tau : x_0 \in U\}$, which is filtering on X. Let $N_{\tau_{x_0}} = \{(x, U) : x \in U, U \in \tau_{x_0}\}$ be directed by $(x, U) \leq (x', U')$ if and only if $U \supseteq U'$ as in (iv) above. Then $x_0 = \lim_{(x,U)\in N_{\tau_{x_0}}} x$. Now, let $V \in \sigma$ with $f(x_0)$. The assumptions on f tell us there is $v - V \in N_{\tau_{x_0}}$ such that for $v \geq v_V$, we have $f(x_0) \in V$ We have $v_V = (x, U)$ for some $U \in \tau_{x_0}$ and $x \in U$, so for any $x' \in U$, $(x', U) \geq (x, U)$ and $f(x') = f(x_{x',U}) \in V$, so that $x_0 \in U = \bigcup_{x'\in U}\{x'\} \subseteq f^{-1}(V)$, so f is continuous at x_0. But $x_0 \in X$ was arbitrary. ∎

Remark. We get the following consequences of this result:

(i) Given topologies τ, τ' on X, $\tau' \subseteq \tau$ if and only if $\tau' - \lim_{v\in N} x_v = x_0$ whenever $\tau - \lim_{v\in N} x_v = x_0$ for any $x_0 \in X$.

(ii) (limits in product topology) $\{(x_\alpha, \tau_\alpha)\}_{\alpha\in A}$ be topological space and $X = \prod_{\alpha\in A} X_\alpha$ equipped with the product topology π. If $(x^{(v)})_{v\in N}$ is a net in X and $x^{(0)} \in X$, then $\pi - \lim_{v\in N} x^{(v)} = x^{(0)}$ if and only if for every $\alpha \in A$, $\tau_\alpha - \lim_{v\in N} x_\alpha^{(v)} = x_\alpha^{(0)}$. Recall that π is the coarsest topology making each μ_α continuous.

(iii) If X is a normed space and $(f_v)_{v\in N} \subset X^*$, $f_0 \in X^*$, then $w^* - \lim_{v\in N} f_v = f_0$ if and only if $\lim_{v\in N} f_v(x) = f_0(x)$ for each $x \in X$.

III. Basic Elements of Functional Analysis

5 Banach Spaces

5.1 Vector Spaces

Definition. Let $\Bbbk$ be a field. Then X is a **vector space** if X is a $\Bbbk$−module. If $Y \subset X$ is a $\Bbbk$−vector space, we say that Y is a **subspace** of X.

Given a subset $S \subset X$, we write

$$\operatorname{span} S = \{\alpha_1 v_1 + \cdots + \alpha_n v_n : \alpha \in \Bbbk, v_i \in S\}$$

to denote the smallest subspace of X containing S.

A basic property of vector spaces is that they have bases:

Definition. Let X be a vector space over $\Bbbk$ A subset $S \subseteq X$ is called

- **linearly independent** if for any distinct $x_1, \ldots, x_n \in S$, the equation $0 = \alpha_1 x_1 + \cdots + \alpha_n x_n = 0$ where $\alpha_i \in \Bbbk$ implies $\alpha_1 = \cdots = \alpha_n = 0$,
- **spanning** if $\operatorname{span} S = X$, and
- a **Hamel basis** if it is both linearly independent and spanning

As a basic consequence of Zorn's lemma, we have the following result:

> ***5.1 Lemma.*** *Suppose $S \subset X$ is linearly independent. Then there exists a Hamel basis M containing S.*

Proof Let $\mathcal{L} = \{S \subset L \subseteq X : L \text{ is linearly independent}\}$. Then $(\mathcal{L}, \subseteq)$ is a poset. If $\mathcal{C} \subseteq \mathcal{L}$ is a chain, it is easy to verify that $U = \bigcup_{L \in \mathcal{C}} L \in \mathcal{L}$ and is an upper bound for $\mathcal{C}$. Thus by Zorn's lemma, $\mathcal{L}$ has some maximal element L; by maximality, M is spanning for X. ■

> ***5.2 Proposition.*** *Let X be a vector space. Then X admits a Hamel basis M. Moreover, if M, M' are distinct Hamel bases, them $|M| = |M'|$.*

Proof Apply Lemma 5.1 to $\emptyset \subset X$ to get some Hamel basis $M = \{v_i : i \in I\}$. For the second part, let $\{w_j : j \in J\}$ be any spanning set: it suffices to show that $|M| \leq |J|$. If J is finite, this is a standard exercise in (finite dimensional) linear algebra.

Now suppose J is infinite and for any $j \in J$, write $w_j = \sum_{i \in A_j \subset I} \lambda_{i,j} v_i$ where $A_j \subset I$ is finite. Thus we have an injection $\phi : \bigcup_{j \in J} \{j\} \times A_j \to M$. But then since J is infinite, $|J| = \left|\bigcup_{j \in J} A_j\right| \leq |M|$ as required. ■

The previous proposition implies that the following notion is well-defined:

Definition. The **dimension** of a vector space X is the cardinality of any Hamel basis for X.

An important feature of any category is to consider the set of maps between objects.

Definition. Given vector spaces X and Y over $\Bbbk$, let $\mathcal{L}(X,Y)$ denote the set of all linear maps from X to Y. In particular, when $Y = \Bbbk$, we write $\mathcal{L}(X,\Bbbk) = X'$ and call this the **algebraic dual** of X'.

Note that $\mathcal{L}(X,Y)$ also has a natural vector space structure given by $(\alpha S + T)(u) = \alpha S(u) + T(u)$. We will discuss such spaces in more details later.

5.2 Seminorms and Norms

The remainder of the chapter will focus particularly on vector spaces $\mathbb{F}$, to denote either the field $\mathbb{R}$ or the field $\mathbb{C}$. Note that if X is a $\mathbb{C}-$vector space, then X is naturally an $\mathbb{R}-$vector space by restriction of scalars, which we denote by $X_{\mathcal{R}}$. Moreover, if X is an $\mathbb{R}-$vector space, we can apply extension by scalars to define a $\mathbb{C}-$vector space structure on X, called the **complexification** of X.

We now focus on the essential additional structure we introduce in the study of normed vector spaces of Banach spaces:

Definition. Let X be a vector space over $\mathbb{F}$. A **seminorm** is a functional $\|\cdot\| : X \to \mathbb{R}$ such that it is

- *(non-negative)* $\|x\| \geq 0$ for any $x \in X$
- *(subadditive)* $\left\|x+y\right\| \leq \|x\| + \left\|y\right\|$ for $x, y \in X$
- *($|\cdot|-$homogenous)* $\|\alpha x\| = |\alpha|\|x\|$ for $\alpha \in \mathbb{F}$, $x \in X$.

If in addition, $\|\cdot\|$ satisfies the added requirement

- *(non-degenerate)* $\|x\| = 0$ if and only if $x = 0$

we call $\|\cdot\|$ a **norm** for X. In this case, the pair $(X, \|\cdot\|)$ a **normed vector space**. We say that $(X, \|\cdot\|)$ is a **Banach space** provided that X is complete with respect to the metric $\rho(x,y) = \left\|x-y\right\|$ induced by the norm.

Example. Here are some standard examples of Banach spaces:

(i) $(\mathbb{F}, |\cdot|)$ is probably the simplest example of a Banach space.

(ii) *Finite-dimensional space:* denoted $(\mathbb{F}^d, \|\cdot\|_p)$ with points $x = (x_j)_{j=1}^n$ equipped with the $p-$norm
$$\|x\|_p = \begin{cases} \left(\sum_{i=1}^n \left|x_j\right|^p\right)^{1/p} & 1 \leq p < \infty \\ \max_{j=1,\dots,n} |x_j| & p = \infty \end{cases}$$
is a Banach space

(iii) If you have a background in basic measure theory, the space $L_{p,\mathbb{F}}(\Omega)$, where Ω is a compact domain. For a concrete example, take for example
$$L^p\mathbb{F}([0,1]) = \left\{ f : [0,1] \to \mathbb{F} \;\middle|\; f \text{ is Lebesgue measurable,} \left(\int_0^1 |f|^p\right)^{1/p} < \infty \right\} \Big/ \sim_{\text{a.e.}}$$
where $1 \leq p < \infty$. To enforce non-degeneracy, we must mod out by equivalence almost everywhere.

(iv) The space of essentially bounded functions, $L_\infty^{\mathbb{F}}[0,1]$, $\|f\|_\infty = \operatorname{ess\,sup}_{t\in[0,1]} |f(t)|$.

(v) *Function spaces:* let (X,d) be a metric space, and define

$$C_b(X,\mathbb{F}) = \{ f : X \to \mathbb{F} \mid f \text{ is continuous and bounded} \}, \qquad \|f\|_\infty = \sup_{x\in X} |f(x)|.$$

Here, we define a more involved example.

Example. Let (X,d) be a metric space. We define the space of *Lipschitz functions*

$$\mathrm{Lip}_{\mathbb{F}}(X,d) = \left\{ f : X \to \mathbb{F} \;\middle|\; f \text{ is bounded}, L(f) = \sup_{\substack{x,y\in X\\ x\neq y}} \frac{|f(x)-f(y)|}{d(x,y)} < \infty \right\}$$

Note that for any $f : X \to \mathbb{F}$, $f \in \mathrm{Lip}_{\mathbb{F}}(X,d)$ if and only if there is some $L \geq 0$ such that $|f(x)-f(y)| \leq Ld(x,y)$ for all x,y in X. One may verify that $L(f)$ is the infimum over all values of L for which this inequality holds over X.

It is an easy exercise to see that $\mathrm{Lip}_{\mathbb{F}}(X,d)$ is a vector space and that $L : \mathrm{Lip}_{\mathbb{F}}(X,d) \to \mathbb{R}$ is a seminorm. However, we do not have non-degeneracy - for example, if f is constant, then $L(f) = 0$. To define a norm on the space of Lipschitz functions, we essentially force non-degeneracy by construction: we define the *Lipschitz norm*

$$\|f\|_{\mathrm{Lip}} = \|f\|_\infty + L(f).$$

In this case, we do in fact have what we want:

5.3 Proposition. *$(\mathrm{Lip}_{\mathbb{F}}(X,d), \|\cdot\|_{\mathrm{Lip}})$ is a Banach space.*

Proof Let $(f_n)_{n=1}^\infty$ be a Cauchy sequence in $(\mathrm{Lip}_{\mathbb{F}}(X,d), \|\cdot\|_{\mathrm{Lip}})$. Since $\|\cdot\|_\infty \leq \|\cdot\|_{\mathrm{Lip}}$ on $\mathrm{Lip}_{\mathbb{F}}(X,d)$, this sequence is uniformly Cauchy and hence converges to some $f \in C_b(X,\mathbb{F})$ with respect to the uniform norm. Moreover, if $x,y \in X$, then

$$\begin{aligned} |f(x)-f(y)| &= \lim_{n\to\infty} |f_n(x)-f_n(y)| \leq \sup_{n\in\mathbb{N}} |f_n(x)-f_n(y)| \\ &\leq \sup_{n\in\mathbb{N}} L(f_n)d(x,y) \leq \sup_{n\in\mathbb{N}} \|f_n\|_{\mathrm{Lip}}\, d(x,y). \end{aligned}$$

Since Cauchy sequences are bounded in norm, we have that $|f(x)-f(y)| \leq Ld(x,y)$ where $L = \sup_{n\in\mathbb{N}} \|f_n\|_{\mathrm{Lip}} < \infty$, so in fact $f \in \mathrm{Lip}_{\mathbb{F}}(X,d)$. It is easy to verify that $\lim_{n\to\infty} \|f - f_n\|_{\mathrm{Lip}} = 0$. ■

5.3 Sequence Spaces

Since we do not assume the background of measure theory in this treatment, one of our main basic examples of Banach spaces will be the sequence spaces. Let $\mathbb{F}^{\mathbb{N}}$ denote the set of all sequences in $\mathbb{F}$, and define

$$\ell^1 = \left\{ x = (x_j)_{j=1}^\infty \in \mathbb{F}^{\mathbb{N}} \;\middle|\; \|x\|_1 = \sum_{j=1}^\infty |x_j| < \infty \right\}.$$

It is easy to see that $(\ell^1, \|\cdot\|_1)$ is a normed vector space. Another common sequence space is the space

$$\ell^\infty = \left\{ x = (x_j)_{j=1}^\infty \in \mathbb{F}^{\mathbb{N}} \;\middle|\; \|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n| < \infty \right\}.$$

More generally, for $1 < p < \infty$, we may define

$$\ell^p = \left\{ x \in \mathbb{F}^{\mathbb{N}} \;\middle|\; \|x\|_p = \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{1/p} < \infty \right\}.$$

As always, it is easy to verify that the ℓ^p–spaces, for $1 \le p < \infty$, are in fact normed vector spaces. The interesting work is in proving that they are Banach spaces.

Let $q = p/(p-1)$ so that $1/p + 1/q = 1$. Then q is called the **conjugate index** to p. We have a number of standard inequalities on ℓ^p–spaces, the proofs of which can be found in general in [***TODO: eventually link measure theory result***].

5.4 Proposition. (Inequalities in ℓ^p–spaces) *Throughout, let $1 < p, q < \infty$ be conjugate exponents.*

- ***Young's Inequality:*** *If $a, b \ge 0$ in $\mathbb{R}$, then $ab \le \frac{a^p}{p} + \frac{b^q}{q}$, with equality if and only if $a^p = b^q$.*
- ***Hölder's Inequality:*** *If $x \in \ell^p$ and $y \in \ell^q$, then $xy = (x_i y_i)_{i=1}^{\infty} \in \ell^1$, with*

  $$\sum_{i=1}^{\infty} \left| x_i y_i \right| \le \|x\|_p \left\| y \right\|_q .$$

  Note that equality holds if and only if the following two conditions hold:

  (i) $\operatorname{sgn}(x_i y_i) = \operatorname{sgn}(x_k y_k)$ *for all $j, k \in \mathbb{N}$ where $x_i y_i \ne 0 \ne x_k y_k$, and*
  (ii) $|x|^p = (|x_j|^p)_{j=1}^{\infty}$ *and $|y|^q$ are linearly dependent in ℓ^1.*
- ***Minkowski's Inequality:*** *If $x, y \in \ell^p$, then $\left\| x + y \right\|_p \le \|x\|_p + \left\| y \right\|_p$ with equality exactly when one of x or y is a non-negative scalar combination of the other.*

In particular, Minkowski's Inequality [***TODO: ref certain labels by name? and also link - would be nice***]

We will also encounter the sequence spaces

$$\begin{aligned} \mathbf{c} &= \{(x_n)_{n=1}^{\infty} : \lim_{n\to\infty} x_n \text{ exists}\} \\ \mathbf{c}_0 &= \{(x_n)_{n=1}^{\infty} : \lim_{n\to\infty} x_n = 0\} \end{aligned}$$

where they are both equipped with $\|\cdot\|_\infty$. They are closed subspaces of ℓ^∞.

5.4 Bounded Continuous Functions into a Normed Space

Let $(Y, \|\cdot\|)$ be a normed space and $\tau = \tau_{\|\cdot\|}$ the topology induced by $\|\cdot\|$. Let (X, τ) be any topological space. We define the space

$$C_b(X, Y) = \left\{ f : X \to Y \;\middle|\; f \text{ is bounded and } \tau - \tau_{\|\cdot\|} - \text{continuous} \right\}$$

With pointwise operations, we see that $C_b(X, Y)$ is a vector space. We also define for $f \in C_b(X, Y)$, $\|f\|_\infty = \sup\{\|f(x)\| : x \in X\}$, making $(C_b(X, Y), \|\cdot\|_\infty)$ a normed vector space.

5.5 Theorem. *If $(Y, \|\cdot\|)$ is a Banach space, then $(C_b(X, Y), \|\cdot\|_\infty)$ is a Banach space.*

Proof Let $(f_n)_{n=1}^\infty$ be a Cauchy sequence in $(C_b(X,Y),\|\cdot\|_\infty)$. Then for any $x \in X$, we have that $(f_n(x))_{n=1}^\infty$ is Cauchy in $(Y,\|\cdot\|)$ since $\|f_n(x)-f_m(x)\| \le \|f_n - f_m\|_\infty$, and hence admis a limit $f(x)$. This defines a pointwise limit $f : X \to Y$. Fix $x_0 \in X$: we must show that f is continuous at x_0. Given $\epsilon > 0$, set

- n_1 so that whenever $n,m \ge n_1$, $\|f_n - f_m\|_\infty < \epsilon/4$.
- n_2 so that whenever $n \ge n_2$, $\|f_n(x_0) - f(x_0)\| < \epsilon/4$.
- $N = \max\{n_1,n_2\}$.
- $U \in \tau$, $x_0 \in U$ such that $f_N(U) \subseteq B_{\epsilon/4}(f(x_0)) \subset Y$.

Then for $x \in U$, we let n_x be so $n_x \ge n_1$ and $n \ge n_x$, so that $\|f_n(x) - f(x)\| < \epsilon/4$. We then have

$$\begin{aligned}\|f(x)-f(x_0)\| &\le \left\|f(x)-f_{n_x}(x)\right\| + \left\|f_{n_x}(x)-f_N(x)\right\| + \|f_N(x)-f_N(x_0)\| + \|f_N(x_0)-f(x_0)\| \\ &< \frac{\epsilon}{4} + \left\|f_{n_x}-f_N\right\|_\infty + \frac{\epsilon}{4} + \frac{\epsilon}{4} < \epsilon\end{aligned}$$

in other words that $f(U) \subseteq B_\epsilon(f(x_0))$ so that f is continuous.

Now let us check that $\|f\|_\infty < \infty$. Since $|\|f_n\|_\infty - \|f_m\|_\infty| \le \|f_n - f_m\|_\infty$, $(\|f_n\|_\infty)_{n=1}^\infty \subseteq \mathbb{R}$ is Cauchy, hence bounded. If $x \in X$, then

$$\|f(x)\| = \lim_{n\to\infty}\|f_n(x)\| \le \sup_{n\in\mathbb{N}}\|f_n(x)\| \le \sup_{n\in\mathbb{N}}\|f_n\|_\infty < \infty$$

so $\|f\|_\infty = \sup_{x\in X}\|f(x)\| < \infty$.

Finally, to show that the limit indeed converges approriately, if ϵ, n_1, x_0, N are as above, we have for $n \ge n_1$

$$\|f_n(x_0)-f(x_0)\| \le \|f_n(x_0)-f_N(x_0)\| + \|f_N(x_0)-f(x_0)\| < \frac{\epsilon}{2}$$

so $\|f_n - f\|_\infty = \sup_{x_0\in X}\|f_n(x_0)-f(x_0)\| \le \epsilon/2 < \epsilon$. The convergence is uniform since n_1 is chosen uniformly in X. ■

5.6 Corollary. *$(C_b(X,\mathbb{F}),\|\cdot\|_\infty)$ is a Banach space.*

Example. (i) Let T be a non-empty set and let

$$\ell^\infty(T) = \left\{ x = (x_t)_{t\in T} \in \mathbb{F}^T \;\middle|\; \|x\|_\infty \right\} < \infty$$

With pointwise operations, $(\ell^\infty,\|\cdot\|_\infty)$ is a normed space. In fact, it is a Banach space, since

$$f \mapsto (f(t))_{t\in T} : C_b(T,\mathcal{P}(T)) \to \ell^\infty(T)$$

is a surjective linear isometry, and the result follows.

(ii) Let $c = \{\, x \in \ell^\infty \mid \lim_{n\to\infty} x_n \text{ exists}\,\}$. Then $(c,\|\cdot\|_\infty)$ is a Banach space. Consider the topological space given by $\omega = \mathbb{N}\cup\{\infty\}$, with topology

$$\tau_\omega = \mathcal{P}(\mathbb{N}) \cup \bigcup_{n\in\mathbb{N}}\{k \in \mathbb{N} : k \ge n\}$$

The map $f \mapsto (f(n))_{n=1}^\infty : C_b(\omega) \to c$ is a linear surjective isometry.

(iii) Recall that $\mathcal{I}(\infty)$ is a closed, and hence complete, subspace of c. We may define $c_0 = \left\{ x \in \mathbb{F}^{\mathbb{N}} \,\middle|\, \lim_{n\to\infty} x_n = 0 \right\} \subseteq c \subseteq \ell^\infty$. In this case, $f \mapsto (f(n))_{n=1}^{\infty} : \mathcal{I}(\infty) \to c_0$ is a (linear) surjective isometry.

(iv) Consider the Sorgenfry line $(\mathbb{R}, \sigma)$. One may verify that

$$C_b((\mathbb{R},\sigma),\mathbb{F}) = \left\{ f : \mathbb{R} \to \mathbb{F} \,\middle|\, f \text{ is bounded and } \lim_{t\to t_0^+} f(t) = f(t_0) \text{ for } t \in \mathbb{R} \right\}$$

5.5 Linear Functionals and Operators

Suppose $(X, \|\cdot\|)$ is a normed space. We denote

$$D(X) = \{x \in X : \|x\| < 1\},$$
$$S(X) = \{x \in X : \|x\| = 1\},$$
$$B(X) = \{x \in X : \|x\| \leq 1\}.$$

In general, if Z is any of the above sets, we may write $x_0 + rZ = \{x_0 + rx : x \in Z\}$ to denote an open (open ball, sphere, closed ball) about x_0 with radius r.

5.7 Proposition. *If X, Y are normed spaces and $S \in \mathcal{L}(X,Y)$. Then the following are equivalent:*
(i) S is continuous
(ii) S is continuous at some $x_0 \in X$
(iii) $\|S\| := \sup_{x\in D(X)} \|Sx\| < \infty$.
Moreover, in this case, we have

$$\|S\| = \inf\{L > 0 : \|Sx\| \leq L\|x\| \text{ for } x \in X\} = \sup_{x\in S(X)} \|Sx\| = \sup_{x\in B(X)} \|Sx\|$$

Proof $(i \Rightarrow ii)$ By definition.
$(ii \Rightarrow iii)$ Since $Sx_0 + D(Y)$ is a neighbourhood of Sx_0, get some $\delta > 0$ so that

$$S(x_0 + \delta D(X)) \subseteq Sx_0 + D(Y)$$

and thus $\delta S(D(X)) \subseteq D(Y)$, or in other words that $\|Sx\| \leq 1/\delta$ for $x \in D(X)$.
$(iii \Rightarrow i)$ If $x \in X$ and $\epsilon > 0$, then

$$\|Sx\| = (\|x\| + \epsilon)\left\| S\left(\frac{1}{\|x\| + \epsilon} \|x\| \right) \right\| \leq (\|x\| + \epsilon)\|S\|$$

so that $\|Sx\| \leq \|S\|\|x\|$. Thus if $x, x' \in X$, then $\|Sx - S'x\| \leq \|S\|\|x - x'\|$ so that S is Lipschitz, hence continuous.

To complete the proof, the content of (iii) implies (i) informs us that the Lipschitz constant $L(S) \leq \|S\|$. Furthermore, taking $\|x\| = 1$, the preceding proof gives us that $\sup_{x\in S(X)} \|Sx\| \leq L(S)$. Finally,

$$\|S\| = \sup_{x\in D(X)\setminus\{0\}} \|Sx\| = \sup_{x\in D(X)\setminus\{0\}} \|x\| \left\| S\left(\frac{1}{\|x\|} x \right) \right\| \leq \sup_{x\in S(X)} \|Sx\|$$

so that the three quantities are equal. The remaining equivalence is obvious. ∎

We now let $\mathcal{B}(X,Y) = \{ S \in \mathcal{L}(X,Y) \mid S \text{ is bounded} \}$. We will see that $\|\cdot\|$, above, defines a norm on $B(X,Y)$.

5.8 Theorem. *Let X, Y, Z be normed spaces. Then*
(i) $(\mathcal{B}(X,Y), \|\cdot\|)$ is a normed space under the operator norm.
(ii) If $S \in \mathcal{B}(X,Y)$ and $T \in \mathcal{B}(Y,Z)$, then $T \circ S \in \mathcal{B}(X,Z)$ with $\|T \circ S\| \leq \|T\| \cdot \|S\|$.
(iii) If Y is a Banach space, then so is $(\mathcal{B}(X,Y), \|\cdot\|)$

Proof We have:
(i) Define $\Gamma : \mathcal{B}(X,Y) \to C_b(B(X),Y)$ by $\Gamma(S) = S|_{B(X)}$. By definition, Γ is linear and $\|\Gamma(S)\|_\infty = \sup_{x \in B(X)} \|Sx\| = \|S\|$ so that $\|\cdot\|$ is a norm and $\Gamma : \mathcal{B}(X,Y) \to C_b(\mathcal{B}(X),Y)$ is an isometry.
(ii) Certainly $T \circ S \in \mathcal{L}(X,Z)$, so it suffices to show the bound on the norm. We have

$$\|T \circ S\| = \sup_{x \in X} \frac{\|T \circ S(x)\|}{\|x\|} = \sup_{\substack{x \in X \\ S(x) \neq 0}} \frac{\|T \circ S(x)\|}{\|S(x)\|} \cdot \frac{\|S(x)\|}{\|x\|} \leq \|T\| \cdot \|S\|$$

as required.
(iii) Now suppose that Y is a Banach space. Since Γ is an isometry, by Theorem 5.5, it suffices to show that $\Gamma(\mathcal{B}(X,Y))$ is closed in the complete space $C_b(B(X),Y)$. Let $(S_n)_{n=1}^\infty \subset \mathcal{B}(X,Y)$ be $\|\cdot\|$–Cauchy so there is some $f \in C_b(B(X),Y)$ such that $\lim_{n\to\infty} \|\Gamma(S_n) - f\|_\infty = 0$. Let $S : X \to Y$ be given by

$$Sx = \begin{cases} \|x\| f\left(\frac{x}{\|x\|}\right) & x \neq 0 \\ 0 & x = 0 \end{cases};$$

in other words, the linear extension of f to X; we will show $\lim_{n\to\infty} S_n = S$. To see that this extension is indeed linear, let $x, x' \in X$ and $\alpha \in \mathbb{F}$ be such that $x, x', x+\alpha x' \neq 0$, then

$$\begin{aligned}
S(x+\alpha x') &= \left\|x+\alpha x'\right\| f\left(\frac{1}{x+\alpha x'}(x+\alpha x')\right) \\
&= \left\|x+\alpha x'\right\| \lim_{n\to\infty} S_n\left(\frac{1}{x+\alpha x'}(x+\alpha x')\right) \\
&= \lim_{n\to\infty} (S_n x + \alpha S_n x') = \lim_{n\to\infty} \left[\|x\| S_n\left(\frac{1}{\|x\|}x\right) + \alpha \left\|x'\right\| S_n\left(\frac{1}{\|x\|}x'\right)\right] \\
&= \|x\| f\left(\frac{x}{\|x\|}\right) + \alpha \left\|x'\right\| f\left(\frac{x'}{\|x\|}\right) \\
&= Sx + \alpha Sx'.
\end{aligned}$$

The above computation is easier if any of x, x', $x+\alpha x'$ are 0. It is immediate that S is continuous (say, at any point on $S(X)$), so $S \in \mathcal{B}(X,Y)$. Finally, as $S|_{B(X)} = f = \lim_{n\to\infty} S_n|_{B(X)}$ (with respect to the uniform norm), we have

$$\|S - S_n\| = \sup_{x \in B(X)} \|(S-S_n)x\| = \|f - \Gamma(S_n)\|_\infty \xrightarrow{n\to\infty} 0$$

as required. ■

5.6 Dual Spaces

A natural way to study an object is to study functions defined on an object:

Definition. If X is a normed space, we let $X^* = \mathcal{B}(X,\mathbb{F})$ denote the **(continuous) dual**. When unspecified, we will always mean the dual space to denote the continuous dual space. When X is an infinite dimensional space, this distinction is important:

5.9 Proposition. *If X is an infinite dimensional normed space, then there exists $f \in X' \setminus X^*$.*

Proof Let $\{e_n\}_{n=1}^\infty \subset X$ be linearly independent; normalizing each element, we may assume that $\|e_n\| = 1$ for each $n \in \mathbb{N}$. Let B be any Hamel basis containing $\{e_n\}_{n=1}^\infty$, and define $f : X = \operatorname{span} B \to \mathbb{F}$ by $f(e_n) = n \cdot e_n$, and $f(e) = 0$ for $e \in B \setminus \{e_n\}_{n=1}^\infty$ and extend by linearity. Notice that $\|f\| \geq |f(e_n)| = n$ for any $n \in \mathbb{N}$, so f is unbounded. ■

In the following lemma, we discuss consequences of extension and restriction by scalars on dual spaces.

5.10 Lemma. *Let X be a $\mathbb{C}-$vector space.*

(i) If $f \in X'_{\mathbb{R}}$ into $\mathbb{R}$, then $f_{\mathbb{C}}$ given by $f_{\mathbb{C}}(x) = f(x) - if(ix)$ defines an element of X'.

(ii) If $g \in X'$, then $f = \operatorname{Re} g$ in $X'_{\mathbb{R}}$ satisfies $g = f_{\mathbb{C}}$.

(iii) If X is a normed $\mathbb{C}-$vector space, then for $f \in X'_{\mathbb{R}}$, $f \in X^*_{\mathbb{R}}$ if and only if $f_{\mathbb{C}} \in X^*$. Moreover, $\|f\| = \|f_{\mathbb{C}}\|$.

Proof (i) and (ii) are straightforward and left as exercises; let's see (iii). Let $x \in X$ be arbitrary and set $z = \operatorname{sgn} f_{\mathbb{C}}(x)$. Then

$$\begin{aligned}|f_{\mathbb{C}}(x)| &= \bar{z} f_{\mathbb{C}}(x) = f_{\mathbb{C}}(\bar{z}x) = \operatorname{Re} f_{\mathbb{C}}(\bar{z}x) = f(\bar{z}x) = |f(\bar{z}x)| \\ &\leq \|f\|\|\bar{z}x\| = \|f\|\|x\|\end{aligned}$$

so we see that $\|f_{\mathbb{C}}\| \leq \|f\|$. Conversely, $|f(x)| = |\operatorname{Re} f_{\mathbb{C}}(x)| \leq |f_{\mathbb{C}}(x)| \leq \|f_{\mathbb{C}}\|\|x\|$ so that $\|f\| \leq \|f_{\mathbb{C}}\|$. ■

Since $\mathbb{F}$ is complete, we have the following corollary of Theorem 5.8:

5.11 Corollary. *If X is a normed spaces, then X^* is always a Banach space.*

Here is an example computation of the dual space of a Banach space:

5.12 Proposition. *Let for $x \in \ell^1$, $f_x : \mathbf{c}_0 \to \mathbb{F}$ be given by $f_x(y) = \sum_{j=1}^\infty x_j y_j$. Then $x \mapsto f_x : \ell^1 \to \mathbf{c}_0^*$ is a surjective isometry.*

Proof If $x \in \ell^1$ and $y \in \mathbf{c}_0 \subseteq \ell^\infty$, then

$$\sum_{j=1}^\infty |x_j y_j| \leq \sum_{j=1}^\infty |x_j| \left\|y\right\|_\infty = \|x\|_1 \left\|y\right\|_\infty < \infty$$

so $f_x(y) = \sum_{j=1}^\infty x_j y_j$ is well-defined and $\|f_x\| \leq \|x\|_1$. It is obvious that f_x is linear. To show that the norms are equal, let $y^n = (\overline{\operatorname{sgn} x_1}, \dots, \overline{\operatorname{sgn} x_n}, 0, 0, \dots) \in \mathbf{c}_0$ where $\left\|y^n\right\| = 1$. Then

$$\|f_x\| \geq |f_x(y^n)| = \sum_{j=1}^n x_i \overline{\operatorname{sgn} x_i} = \sum_{j=1}^n |x_i| \tag{5.1}$$

so that $\|f_x\| \geq \|x\|_1$ and we have equality, so the map is a surjective linear isometry.

Now let $f \in \mathbf{c}_0^*$, $e_n = (0,\dots,0,1,0,0,\dots) \in \mathbf{c}_0$, and let $x_n = f(e_n)$. Suppose $y \in \mathbf{c}_0$ is arbitrary and set $y^n = (y_1,\dots,y_n,0,0,\dots)$, so that $\lim_{n\to\infty} y_n = y$ uniformly. Since f is continuous, we have

$$f(y) = \lim_{n\to\infty} f(y^n) = \lim_{n\to\infty} \sum_{j=1}^{n} y_j x_j = \sum_{j=1}^{\infty} x_j y_j = f_x(y)$$

and moreover taking the sequence $(y_n)_{n=1}^{\infty}$ from (5.1) we have that

$$\sum_{j=1}^{n} |x_i| = |f(y^n)| \leq \|f\| < \infty$$

so $x \in \ell^1$ and $f = f_x$, as required. ∎

5.13 Corollary. *$\ell^1 \cong \mathbf{c}^*$ isometrically isomorphically.*

Proof For $y \in \mathbf{c}$, let $L(y) = \lim_{n\to\infty} y_n$ and $y^n = (y_1,\dots,y_n,L(y),L(y),\dots) \in c$. As before, $\lim_{n\to\infty} y^n = y$ uniformly. Let $f \in \mathbf{c}^*$ be arbitrary so that $f|_{\mathbf{c}_0} = f_x$ for some $x \in \ell^1$.

Let $\mathbf{1} = (1,1,\dots)$, and $\mathbf{1}_n = (0,\dots,0,1,1,\dots)$. Note that $f(\mathbf{1}_n)_{n=1}^{\infty}$ is Cauchy in $\mathbb{F}$ for if $m < n$ is arbitrary, since $\mathbf{1}_m - \mathbf{1}_n \in \mathbf{c}_0$,

$$|f(\mathbf{1}_n) - f(\mathbf{1}_m)| = |f(\mathbf{1}_n - \mathbf{1}_m)| \leq \sum_{j=m+1}^{n} |x_j|$$

which can be made arbitrarily small for sufficiently large m.

Thus we may set $x_0 = \lim_{n\to\infty} f(1_n)$ and $\tilde{x} = (x_0, x_1, \dots) \in \ell^1$ where $x_j = f(e_j)$. Then

$$f(y) = \lim_{n\to\infty} f(y^n) = \sum_{j=1}^{\infty} x_j y_j + x_0 L(y)$$

so we may consider the map $x \mapsto f_x : \ell^1 \to \mathbf{c}$ where $x = (x_n)_{n=0}^{\infty}$ given by $f_x(y) = \sum_{j=1}^{\infty} x_j y_j + x_0 L(y)$. Similarly to before, we may show that $\|f_x\| = \|x\|_1$. ∎

Again appealing to Theorem 5.5, we have:

5.14 Corollary. *$(\ell_1, \|\cdot\|_1)$ is complete.*

6 Fundamental Results in Banach Space Theory

6.1 The Hahn-Banach Theorem

The Hahn-Banach theorem is a fundamental constructive result in Banach space theory which implies that linear functionals exist in abundance.

Definition. Let X be a $\mathbb{R}$−vector space. A **sublinear functional** is any $\rho : X \to \mathbb{R}$ that satisfies

- *(non-negative homogenity)* $\rho(tx) = t\rho(x)$ for $t \geq 0$, $x \in X$.

- *(subadditivity)* $\rho(x+y) \le \rho(x) + \rho(y)$ for $x, y \in X$.

6.1 Theorem. (Hahn-Banach) *Let X be a $\mathbb{R}-$vector space, $\rho : X \to \mathbb{R}$ a sublinear functional, $Y \subseteq X$ a subspace and $f \in Y'$ such that $f \le \rho|_Y$. Then there exists $F \in X'$ such that $F|_Y = f$ and $F \le \rho$ on X.*

Proof The technique for this proof is to first do it by extensions by a single point $x \in X \setminus Y$, then to appeal to Zorn's lemma to extend the result to X. To extend the definition of f along the subspace $\operatorname{span}\{x\} \cup X$, it suffices find $c \in \mathbb{R}$ such that

$$f(y) + \alpha c \le \rho(y + \alpha x) \tag{6.1}$$

for $y \in Y$ and $\alpha \in \mathbb{R}$. Assuming we have such a c, we may set $F : \operatorname{span} Y \cup \{x\} \to \mathbb{R}$ by $F(y + \alpha x) = f(y) + \alpha c$. Then it is clear that F is linear and satisfies $F \le \rho$ on $\operatorname{span} Y \cup \{x\}$.

To do this, let $y^+, y^- \in Y$ be arbitrary and observe that by subadditivity

$$f(y^+) + f(y^-) = f(y^+ + y^-) \le \rho(y^+ + y^-) \le \rho(y^+ + x) + \rho(y^- - x)$$

so that $f(y^-) - \rho(y^- - x) \le \rho(y^+ + x) - f(y^+)$. Since y^-, y^+ were arbitrary, there is some $c \in \mathbb{R}$ such that

$$\sup\{f(y) - \rho(y - x) : y \in Y\} \le c \le \inf\{\rho(y + x) - f(y) : y \in Y\}$$

Now if $t > 0$ is arbitrary, for $y \in Y$, since $c \le \rho\left(\frac{1}{t}y + x\right) - f\left(\frac{1}{t}y\right)$, by non-negative homogeneity $tx \le \rho(y + tx); f(y)$ so that $f(y) + tc \le \rho(y + tx)$. In the exact same way, if $s > 0$, for $y \in Y$, we have $f(y) - sc \le \rho(y - sx)$. It is clear that $f(y) + 0 \le \rho(y + 0x)$. Thus we have our desired inequality in (6.1)

We now use Zorn's lemma to extend this result to the whole space. Consider the set of ρ-extensions of f,

$$\mathcal{E} = \left\{(\mathcal{M}, \psi) : Y \subseteq \mathcal{M} \subseteq X, \mathcal{M} \text{ is a subspace}, \psi \in \mathcal{M}', \psi|_Y = f, \psi \le \rho|_{\mathcal{M}}\right\}$$

equipped with the partial order where $(\mathcal{M}, \psi) \le (\mathcal{N}, \phi)$ if and only if $\mathcal{M} \subseteq \mathcal{N}$ and $\phi|_{\mathcal{M}} = \psi$. Suppose $\mathcal{C} \subseteq \mathcal{E}$ is a chain with respect to $\le$, and let

- $\mathcal{U} = \bigcup_{(\mathcal{M},\varphi)} \mathcal{M}$, and
- $\phi : \mathcal{U} \to \mathbb{R}$ by $\phi(x) = \psi(x)$ whenever $x \in \mathcal{M}$

where it is easily verified that $\mathcal{U}$ is a subspace and ϕ is well-defined since $\mathcal{C}$ is a chain. Moreover, we that $\phi \in U'$, since if $x, y \in \mathcal{U}$, get $x \in \mathcal{M}$, $y \in \mathcal{N}$ for some $(\mathcal{M}, \psi) \le (\mathcal{N}, \psi') \in \mathcal{C}$, so $\phi(\alpha x + y) = \psi'(\alpha x + y) = \alpha\psi'(x) + \psi'(y) = \alpha\phi(x) + \phi(y)$. In the same way, we have $\psi \le \rho|_{\mathcal{U}}$. Thus $(\mathcal{U}, \phi) \in \mathcal{E}$ and is indeed an upper bound for $\mathcal{C}$.

Thus by Zorn's lemma, $\mathcal{E}$ some maximal element $(\mathcal{M}, F)$. In particular, we must have $\mathcal{M} = X$, for if not, get $x \in X \setminus \mathcal{M}$ and apply the single point extension result to $\operatorname{span} \mathcal{M} \cup \{x\}$ to get some strictly larger element $(\operatorname{span}\{x\} \cup \mathcal{M}, F')$, violating maximality. ■

To put Hahn-Banach into a more useful form and define it for any normed vector space over $\mathbb{R}$ or $\mathbb{C}$, we have the following result:

6.2 Theorem. (Continuous Hahn-Banach) *Let X be a normed space.*

(i) If $Y \subseteq X$ a subspace and $f \in Y^*$, then there exists $F \in X^*$ such that $F|_Y = f$ and $\|F\| = \|f\|$.

(ii) If $x \in X$, then there is $f \in X^*$ such that $f(x) = \|x\|$ and $\|f\| = 1$.

Proof We have:

(i) Define $\rho : X \to \mathbb{R}$ be given by $\rho(x) = \|f\| \cdot \|x\|$, so ρ is sublinear and $\operatorname{Re} f \leq \rho|_Y$. Apply Hahn-Banach and get $\tilde{F} \in X^*_{\mathbb{R}}$ such that $\tilde{F}|_Y = \operatorname{Re} f$ and $\tilde{F} \leq p$. By Lemma 5.10, $F = \tilde{F}_{\mathbb{C}}$ satisfies the required properties.

(ii) If $x = 0$, this is trivial, so suppose not. Let $f_0 : \operatorname{span}\{x\} \to \mathbb{F}$ be given by $f_0(\alpha x) = \alpha \|x\|$. If $x \neq 0$, we have immediately that $\|f_0\| = 1$; apply (i) above. ■

As a nice application of Hahn-Banach, we get the following embedding result. Let X be a normed space and X^{**} denote the bidual, and for $x \in X$, define $\hat{x} : X^* \to \mathbb{F}$ by $\hat{x}(f) = f(x)$.

6.3 Theorem. *The map $\phi : X \to X^{**}$ given by $\phi(x) = \hat{x}$ is a linear isometry.*

Proof Linearity is immediate. Notice that $|\hat{x}(f)| = |f(x)| \leq \|f\|\|x\|$ so $\|\hat{x}\| \leq \|x\|$. Moreover, Theorem 6.2 provides for each $x \in X$ some $f_x \in S(X^*)$ with $|f_x(x)| = \|x\|$, so that $\|\hat{x}\| = \|x\|$. ■

This method gives a nice way to define the completion of a normed space. Naturally, we may do this construction for any metric space, but it is nicer for Banach space. Since X^{**} is a dual space, it is complete, so that $\hat{X} = \{\hat{x} : x \in X\}$ has $\overline{\hat{X}} \subseteq X^{**}$ is complete. But then $\overline{\hat{X}}$ is Banach space containing a dense isometric copy of X, and since the completion of a metric space is unique (up to surjective isometry), this is the completion of X, $\overline{X}$.

6.2 Hyperplane Separation Theorem

An important application of Hahn-Banach is the Hyperplane Separation Theorem. If $A, B \subset X$ with $A \cap B = \emptyset$ along with some other suitable assumptions, we will find a $\mathbb{R}$−hyperplane between A and B.

Definition. In a vector space, a **hyperplane** is any set of the form $x_0 + \ker f$ with $x_0 \in X$ and $f \in X'$. Then a $\mathbb{R}$−**hyperplane** is any set of the form $x_0 + \ker \operatorname{Re} f$.

6.4 Proposition. *Let X be a normed space.*

(i) If $f \in X^* \setminus \{0\}$, then $\ker f$ is closed and nowhere dense.

(ii) if $f \in X' \setminus X^*$, then $\overline{\ker f} = X$.

Thus a hyperplane in X is either closed and nowhere dense, or it is dense.

Proof (i) To see (i), $\ker f = f^{-1}(\{0\})$ is a closed set since f is continuous. Since f is not the zero map, $\ker f$ is a proper closed subspace. In general, if $Y \subsetneq X$ is a proper closed subspace, then it is nowhere dense. To see this, since Y is closed, if Y is dense at y_0, get $\delta > 0$ such that $y_0 + \delta D(X) \subseteq Y$. Then $D(X) \subseteq \frac{1}{\delta}(Y - y_0) = Y$, so $X = \operatorname{span} D(X) \subseteq Y$.

(ii) Suppose that $\ker f$ is not dense in X. Then there would be $x_0 \in X$ and $\delta > 0$ such that $(x_0 + \delta D(X)) \cap \ker f = \emptyset$, so $0 \notin f(x_0) + \delta f(D(X))$ and

$$\frac{1}{\delta} f(x_0) \notin -f(D(X)) = f(D(X)). \tag{6.2}$$

Then $\|f\| \leq \frac{1}{\delta} f(x_0)$: to see this, suppose $\|f\| > \frac{1}{\delta} f(x_0)$ and get $x \in D(X)$ such that $|f(x)| > \frac{1}{\delta}|f(x_0)|$. But then $\left|\frac{1}{\delta} f(x_0)\right| < c$ where $c \in f(D(X))$...[***TODO: Fix??***] ■

Definition. Let $\emptyset \neq A \subseteq X$. We say that A is

- **convex** if for $a, b \in A$ and $0 < \lambda < 1$, $(1-\lambda)a + \lambda b \in A$.
- **absorbing** at $a_0 \in A$ if for any $x \in X$, there is $\epsilon = \epsilon(a_0, x) > 0$ such that $a_0 + tx \in A$ for $0 \leq t < \epsilon$.

For example, if X is a normed space, then any open set is absorbing around any of its points.

6.5 Lemma. (Minkowski Functional) *Let $A \subset X$ be a convex set containing 0 and absorbing at 0. Then the map $\rho : X \to \mathbb{R}$ given by $\rho(x) = \inf\{t > 0 : x \in tA\}$ is a sublinear functional such that*

(i) $\{x \in X : \rho(x) < 1\} \subseteq A \subseteq \{x \in X : \rho(x) \leq 1\}$ *and*

(ii) if X is normed and A is a neighbourhood of 0, *then there is $N > 0$ such that $\rho(x) \leq N\|x\|$ for $x \in X$.*

Proof Note that ρ is well-defined since A is absorbing at 0. We verify that ρ is a sublinear functional:

- *(non-negative homogeneity):* Clearly $\rho(0) = 0$. If $s > 0$ and $x \in X$, then
$$\rho(sx) = \inf\{t > 0 : sx \in tA\} = s \cdot \inf\left\{\frac{t}{s} > 0 : x \in \frac{t}{s}A\right\} = s\rho(x).$$
- *(subadditivity):* If $s, t > 0$ and $a, b \in A$, then by convexity
$$sa + tb = (s+t)\left(\frac{s}{s+t}a + \frac{t}{s+t}b\right) \in (s+t)A$$
so that $sA + tA \subseteq (s+t)A$. Clearly $(s+t)A \subseteq sA + tA$ so that $sA + tA = (s+t)A$.
Now for $x, y \in X$, we have
$$\begin{aligned}\rho(x) + \rho(y) &= \inf\{s > 0 : x \in sA\} + \inf\{t > 0 : y \in tA\}\\ &= \inf\{s + t : s > 0, t > 0, x \in sA, y \in tA\}\\ &\geq \inf\{s + t : s > 0, t > 0, x + y \in sA + tA = (s+t)A\}\\ &= \inf\{r > 0 : x + y \in rA\} = \rho(x+y).\end{aligned}$$

Now we see that it satisfies the required properties:

(i) If $\rho(x) < 1$, then there is $0 < t < 1$ so $\frac{1}{t}x \in A$ and since $0 \in A$ and A is convex $x = (1-t)\cdot 0 + t \cdot \frac{1}{t}x \in A$. The second inclusion is obvious.

(ii) Get $\delta > 0$ so $\delta D(X) \subseteq A$. Then for $x \in X$ and any $\epsilon > 0$,
$$x \in (\|x\| + \epsilon)D(X) \subseteq \frac{\|x\| + \epsilon}{\delta}A$$
so $\rho(x) \leq \frac{\|x\|+\epsilon}{\delta}$; since $\epsilon > 0$ was arbitrary, $\rho(x) \leq \frac{1}{\delta}\|x\|$ so the result follows with $N = 1/\delta$. ■

6.6 Theorem. (Hyperplane Separation) *Let X be an $\mathbb{F}$ −vector space, $A, B \subset X$ be convex with $A \cap B = \emptyset$ and A absorbing at some a_0. Then there are $f \in X'$ and $\alpha \in \mathbb{R}$ such that*
$$\operatorname{Re} f(a) \geq \alpha \geq \operatorname{Re} f(b)$$
for $a \in A$ and $b \in B$. Moreover, if X is normed, then

- *if A is a neighbourhood of a_0, $f \in X^*$*
- *if A is absorbing around each of its points,* $\operatorname{Re} f(a) > \alpha \geq \operatorname{Re} f(b)$.

Proof We first recenter at 0. It is easy to verify that
- $A - B$ is absorbing at any $a_0 - b$, $b \in B$
- $A - B$ is convex
- if X is normed and A a neighbourhood of a_0, then $A - B$ is a neighbourhood of each $a_0 - b$, $b \in B$; and
- if A is absorbing around any of its points, then $A - B$ is absorbing around any of its points.

Let $x_0 = a_0 - b_0$ for some $b_0 \in B$ and set $C = x_0 - (A - B)$; note that $0 = x_0 - x_0 \in C$. Then by the above points, C is absorbing at 0, convex, and if X is normed and A a neighbourhood of a_0, then C is a neighbourhood of 0; and if A is absorbing at any of its points, then C is absorbing at each of its points.

Let ρ be the Minkowski functional of C. Notice that since $A \cap B = \emptyset$, $0 \notin A - B$ so $x_0 \notin C$. Thus by (i) of Lemma 6.5, $\rho(x_0) > 1$.

Let us find f and α. Let $f_0 : \operatorname{span}_{\mathbb{R}}\{x_0\} \to \mathbb{R}$, by $f_0(sx) = s\rho(x_0)$. Then f_0 is linear and $f_0 \leq \rho|_{\operatorname{span}_{\mathbb{R}}\{x_0\}}$, so by Hahn-Banach, get $f \in X'_{\mathbb{R}}$ such that $f \leq \rho$ on X. If $a \in A$ and $b \in B$, then $x_0 - (a - b) \in C$ so that

$$f(x_0 - (a - b)) \leq \rho(x_0 - (a - b)) \leq 1$$

again by (i) of Lemma 6.5. Thus $f(x_0) + f(b) \leq 1 + f(a)$ and since $f(x_0) = \rho(x_0)$, we have $f(b) \leq f(a)$. Thus there exists some $\alpha \in \mathbb{R}$ such that

$$\sup\{f(b) : b \in B\} \leq \alpha \leq \inf\{f(a) : a \in A\}$$

If $\mathbb{F} = \mathbb{R}$, we are done; otherwise, replace f by $f_{\mathbb{C}}$ and the result follows.

For the remainder of the proof, we suppose X is a normed space, and A is a neighbourhood of a_0. Then (ii) of Lemma 6.5 provides $N > 0$ so that $\rho(x) \leq N\|x\|$. Then for $x \in X$, $f(x) \leq \rho(x) \leq N\|x\|$ and $-f(x) = \rho(-x) \leq N\|-x\| = N\|x\|$ so $|f(x)| \leq N\|x\|$ and $f \in X^*$.

If A is absorbing around any of its points, then $f(a) > \alpha$ for any $a \in A$. Indeed, suppose $f(a) = \alpha$. Then there would be $t > 0$ so $a + t(-x_0) \in A$. But then $\alpha \leq f(a - tx_0) = f(a) - tf(x_0) < \alpha$ since $f(x_0) = \rho(x_0) > 0$, a contradiction. ■

Definition. If $\emptyset \neq S \subset X$, then its **convex hull** is given by

$$\operatorname{conv}(S) = \left\{ \sum_{i=1}^{n} \lambda_j x_j : n \in \mathbb{N}, x_j \in S, \lambda_j \geq 0, \sum_{j=1}^{n} \lambda_j = 1 \right\}$$

One can verify that $\operatorname{conv}(S)$ is in fact convex, and is the smallest convex set containing S, in other words that

$$\operatorname{conv}(S) = \bigcap_{\substack{S \subseteq C \subseteq X \\ C \text{ convex}}} C$$

If X is normed, we let $\overline{\operatorname{conv}}(S)$ denote the **closed convex hull**, i.e. the closure of the convex hull.

Definition. A **half-space** of X is any set of the form $H = \{x \in X : \operatorname{Re} f(x) \leq \alpha\}$ for some $f \in X'$, $\alpha \in \mathbb{R}$.

If X is normed, then Proposition 6.4 shows that H is closed if and only if f is bounded. We now have the following consequence of Theorem 6.6:

6.7 Corollary. *If X is a normed vector space and $\emptyset \neq S \subset X$, then*

$$\overline{\operatorname{conv}}(S) = \bigcap_{\substack{S \subseteq H \subset X \\ H \text{ closed half space}}} H.$$

Proof The forward inclusion is trivial, so we assume $x_0 \notin \overline{\operatorname{conv}}(S)$. Then there is $\delta > 0$ such that $(x_0 + \delta D(X)) \cap \overline{\operatorname{conv}}(S) = \emptyset$. Since $x_0 + \delta D(X)$ is open and convex, Theorem 6.6 gives provides $f \in X^*$ and $\alpha \in \mathbb{R}$ so $\operatorname{Re} f(a) > \alpha \geq \operatorname{Re} f(b)$ for $a \in x_0 + \delta D(X)$ and $b \in \overline{\operatorname{conv}}(S)$. Then $S \subset H = \{y \in X : \operatorname{Re} f(x) \leq \alpha\}$ but $x_0 \notin H$. ∎

6.3 Banach-Steinhaus

We now prove a straightforward but important corollary of the Baire Category Theorem:

6.8 Theorem. (Banach-Steinhaus) *Let X, Y be normed spaces, $U \subseteq X$ be non-meagre, and $\mathcal{F} \subset \mathcal{B}(X, Y)$ be such that for each $x \in U$, $\sup\{\|Tx\| : T \in \mathcal{F}\} < \infty$. Then $\sup\{\|T\| : T \in \mathcal{F}\} < \infty$.*

Proof Let for each $n \in \mathbb{N}$

$$F_n = \bigcap_{T \in \mathcal{F}} T^{-1}(nB(Y)) = \{x \in X : \|Tx\| \leq n \text{ for all } T \in \mathcal{F}\}$$

so each F_n is closed and, by the pointwise boundedness assumption, $U \subseteq \bigcup_{n=1}^{\infty} F_n$. By assumption of non-meagreness of U, at least one F_{n_0} admis an interior point: there is $x_0 \in F_{n_0}$ and $\delta > 0$ such that $x_0 + \delta D(X) \subseteq F_{n_0}$. Then if $x \in D(X)$, we have

$$Tx = \frac{1}{\delta}\left[T\left(x_0 + \frac{\delta}{2}x\right) - T\left(x_0 - \frac{\delta}{2}x\right)\right]$$

so $\|Tx\| \leq \frac{2}{\delta} n_0$, in other words

$$\|T\| = \sup_{x \in D(x)} \|Tx\| \leq \frac{2n_0}{\delta} < \infty$$

where the bound is independent of T. ∎

6.4 Open Mapping and Closed Graph

6.9 Theorem. (Open Mapping) *Let X, Y be Banach spaces, and $T \in B(X, Y)$ surjective. Then T is an open map; i.e. $T(U)$ is open in Y whenver U is open in X.*

Remark. Given $x \in X$ and $\alpha \in \mathbb{F} \setminus \{0\}$, non-empty $A \subset X$, we have that $\overline{x + \alpha A} = x + \alpha \overline{A}$. Indeed, note that for $(a_k)_{k=1}^{\infty} \subset A$, we have

$$a_k \to a \in \overline{A} \text{ if and only if } x + \alpha a_k \to x + \alpha a \in x + \alpha \overline{A}$$

6.10 Lemma. *With the assumptions as above, we have that if $\overline{T(D(X)}} \supset rB(Y)$ for some $r > 0$, then $T(D(X)) \supseteq rD(Y)$.*

Proof Let $z \in rD(Y)$ and let $0 < \delta < 1$ be so $\|z\| < r(1-\delta) < r$. Set $y = z/(1-\delta)$ so $\|y\| < r/(1-\delta)$. It suffices to show that $y \in \frac{1}{1-\delta} T(D(X))$. To begin, let $A = T(D(X)) \cap rB(Y)$, so $\overline{A} = rB(Y)$. Indeed, if $y \in rB(Y) \subseteq \overline{T(D(X))}$, then there is $(y_k)_{k=1}^{\infty} \subset \overline{T(D(X))}$, so $y = \lim y_k$. But then there is $x_k \in D(X)$ so each $\|y_k - T(x_k)\| < 1/k$ so $y = \lim T(x_k)$ with each $x_k \in D(X)$.

Now we inductively build a sequence $(y_n)_{n=1}^{\infty}$ as follows.

- Since $y \in rD(Y) \subseteq \overline{A}$, there is $\underline{y_1 \in A} \cap (y + \delta r D(Y))$
- $y \in y_1 + \delta r(D(Y)) \subseteq \underline{y_1 + \delta \overline{A}} = y_1 + \delta A$, so there is $y_2 \in (y_1 + \delta A) \cap (y + \delta^2 r D(Y))$
- $y \in y_n + \delta^n r D(Y) \subseteq \overline{y_n + \delta^n A}$, so there is $y_{n+1} \in (y_n + \delta^n A) \cap (y + \delta^{n+1} r D(Y))$

By construction, $y_{n+1} - y_n \in \delta^n A$, so $\|y_{n+1} - y_n\| \le \delta^n r$ and there is $x_n \in \delta^n D(X)$ such that $y_{n+1} - y_n = T x_n$. Likewise, $y_1 \in A \subseteq T(D(X))$ so $y = T(x_0)$ for some $x_0 \in D(X)$. Notice that each $y_n \in y + \delta^n r \in D(Y)$, so $\|y_n - y\| \le \delta^n r \to 0$. Since X is complete, we let $x = \sum_{n=0}^{\infty} x_n$, and by construction

$$\|x\| \le \sum_{n=0}^{\infty} \|x_n\| < \sum_{n=0}^{\infty} \delta^n = \frac{1}{1-\delta}$$

Then by linearity and continuity of T, we have

$$Tx = \sum_{n=0}^{\infty} T x_n = y_1 + \sum_{n=1}^{\infty} (y_{n+1} - y_n) = y_N + \sum_{n=N}^{\infty} (y_{n+1} - y_n) \to y$$

so that indeed $T(x) = y$, as required. ■

Remark. So far, we've only used completeness of X and continuity and linearity of T.

We now proceed with the proof of the open mapping theorem.

Proof It suffices to see that $T(D(X))$ contains a neighbourhood of 0 in Y. Indeed, if $\emptyset \ne U \subseteq X$ is open, $x \in U$, then there is $\delta > 0$ such that $x + \delta D(X) \subseteq U$, so $U - x \supseteq \delta D(X)$. If $T(D(X)) \supseteq rD(Y)$, then $T(U - x) \supseteq \delta T(D(X)) \supseteq r\delta D(Y)$ so that $Tx + r\delta D(Y) \subseteq T(U)$. In other words, $T(U)$ is a neighbourhood of any of its points, and thus open.

Now write $X = \bigcup_{n=1}^{\infty} nD(X)$, and we assume that $T(X) = Y$. Hence $Y = \bigcup_{n=1}^{\infty} nT(D(X))$, so $Y = \bigcup_{n=1}^{\infty} n\overline{T(D(X))}$. But Y is complete, so by Baire category theorem, there is some n so that $n\overline{T(D(X))}$ has non-empty interior. Since $nT(D(X))$ is convex and symmetric, and hence $n\overline{T(D(X))}$ is convex and symmetric as well. Thus if $y \in D(Y)$, then $y_0 \pm \epsilon \in y_0 + \epsilon D(Y)$ so

$$\epsilon y = \frac{1}{2}[y_0 + \epsilon y - (y_0 - \epsilon y)] \in n\overline{T(D(X))}$$

and $\frac{\epsilon}{n} y \in \overline{T(D(X))}$, i.e. $\frac{\epsilon}{n} D(Y) \subseteq \overline{T(D(X))}$. Thus applying the main lemma, $\frac{\epsilon}{n} D(Y) \subseteq T(D(X))$. ■

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

Proof Direct application of the open mapping theorem. ■

Let X, Y be normed spaces. Then we define for $(x, y) \in X \oplus Y$, and we let $\|(x, y)\|_1 = \|x\| + \|y\|$. It is easy to check that $\|\cdot\|_1$ is a norm on $X \oplus Y$, and if X, Y are Banach, then so is $(X \oplus Y, \|\cdot\|_1)$. In this case, we write $X \oplus_1 Y$.

6.12 Theorem. (Closed Graph) *Let X, Y be Banach spaces and $T \in \mathcal{L}(X,Y)$. Then $T \in \mathcal{B}(X,Y)$ if and only if $\Gamma(T) = \{(x,Tx) : x \in X\}$ is closed in $X \oplus_1 Y$.*

Proof Let $T \in \mathcal{B}(X,Y)$ and suppose $(x,y) = \lim_{n\to\infty}(x_n, Tx_n)$ with $(x_n, Tx_n) \in \Gamma(T)$. Then

$$\begin{aligned}\left\|y - Tx\right\| &\leq \left\|y - Tx_n\right\| + \|Tx_n - Tx\| \leq \left\|y - Tx_n\right\| + \|T\|\|x_n - x\| \\ &\leq \max\|T\|, 1 \cdot \left\|(x,y) - (x_n, Tx_n)\right\|_1\end{aligned}$$

so that $y = Tx$ and $(x,y) = (x,Tx) \in \Gamma(T)$.

Conversely, if $\Gamma(T)$ is closed in $X \oplus_1 Y$, then $\Gamma(T)$ is a Banach space. Define $\pi : \Gamma(T) \to X$ by $\pi(x,Tx) = x$ (projection onto the first coordinate). Notice that π is linear, and

$$\|\pi(x,Tx)\| = \|x\| \leq \|(x,Tx)\|_1$$

so $\|\pi\| \leq 1$ and π is bounded. Moreover, π is a bijection where $\pi^{-1} : X \to \Gamma(T)$ is given by $\pi^{-1}(x) = (x,Tx)$, so by Theorem 6.11 π^{-1} is bounded. Hence for any $x \in X$,

$$\|Tx\| \leq \|(x,Tx)\|_1 = \left\|\pi^{-1}x\right\| \leq \|x\|\left\|\pi^{-1}\right\|$$

so that T is bounded. ■

6.13 Theorem. (Closed graph test) *Given normed spaces and $T \in \mathcal{L}(X,Y)$, we have that $\Gamma(T)$ is closed in $X \oplus_1 Y$ if and only if whenever $x_n \to 0$ for which we may assume that Tx_n converges in Y, say $y = \lim Tx_n$, then $y = 0$ too.*

Proof We have $(x_n, Tx_n) \to (x,z) \in \overline{\Gamma(T)}$ if and only if $(x_n - x, T(x_n - x)) \to (x,z) - (x,Tx) = (0, z - Tx)$. Set $y = z - Tx$. We have $(x,z) \in \Gamma(T)$ if and only if $z = Tx$ if and only if $y = 0$. ■

6.5 Testing hypothesis of OMT

(i) Let $1 \leq p < r < \infty$. We have that $\ell_p \subseteq \ell_r$, with $\|x\|_r \leq \|x\|_p$ for $x \in \ell_p$. First, suppose $x \in B(\ell_p)$, so for each k, $|x_k| \leq \|x\|_p \leq 1$ so $|x_k|^{r/p} \leq |x_k|$. Hence

$$\|x\|_r = \left(\sum_{k=1}^{\infty} |x_k|^r\right)^{1/r} \leq \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{1/r} = \|x\|_p^{p/r} \leq 1$$

so if $x \in \ell_p \setminus \{0\}$, then the result follows.
Let $S : (\ell_p, \|\cdot\|_p) \to (\ell_p, \|\cdot\|_r)$ be the identity map. Then $\|S\| \leq 1$, and furthermore S is bijective. If S were open, then by the proof of inverse mapping theorem, we would see that $\left\|S^{-1}\right\| < \infty$. Define $x^{(n)} \in \ell_p$ by

$$x_k^{(n)} = \begin{cases} \frac{1}{ck^{1/p}} & k \leq n \\ 0 & k > n \end{cases}, c = \sum_{k=1}^{\infty} \frac{1}{k^{r/p}}$$

We compute that $\left\|x^{(n)}\right\|_r < 1$ while $\left\|x^{(n)}\right\|_p = \frac{1}{c}\left(\sum_{k=1}^{n} \frac{1}{k}\right)^{1/p}$. In other words, $\left\|S^{-1}x^{(n)}\right\|_p$ goes to infinity, while $\left\|x^{(n)}\right\|_r < 1$, contradicting $\left\|S^{-1}\right\| < \infty$. The moral of this is that if the range space is not complete, then OMT may not hold.

(ii) Take $X = C_b(0,1)$, $X_0 = \{f \in X : f \text{ is diferentiable on } (0,1), f' \in C_b(0,1)\}$. We have $X_0 \subseteq X$, and we put the uniform norm $\|\cdot\|_\infty$ on both spaces. We let $D : X_0 \to X$, $Df = f'$. If $h_n(t) = t^n$, then $\|h_n\|_\infty = 1$ while $\|Dh_n\|_\infty = n$, so D is not bounded. Despite this, we have that $\Gamma(D) = \{(f,f') : f \in X_0\}$ is closed in $X_0 \oplus_1 X$. We apply the closed graph test: let $(f_n, f_n') \to (0,g)$ in $X_0 \oplus_1 X$. Notice that $\|f_n'\|_\infty < \infty$, so f_n is Libschitz on $(0,1)$, so f_n is uniformly continuous on $(0,1)$, so $f_n(0^+) = \lim_{t\to 0^+} f(t)$ exists. Thus by the fundamental theorem of calculus, $f_n(t) = f_n(0^+) + \int_0^t f_n'$ for $t \in (0,1)$. In particular,

- $f_n \to 0$ uniformly, so $f_n(0^+) \to 9$
- $f_n' \to g$ uniformly, so for each $t \in (0,1)$,

$$\int_0^t g = \lim_{n\to\infty} \int_0^t f_n' = \lim_{n\to\infty} [f_n(t) - f_n(0^+)] = 0$$

and again, by the FT of C, $g(t) = 0$. Thus $g = 0$, so $\Gamma(D)$ is closed. We say that $D : X_0 \to X$ is a **closed** operator. The moral here is that if the domain is not complete, then closedness of the graph does not imply boundedness of the operator.

Now, let $J : X \to X_0$ have $Jg(t) = \int_0^t g$ for $t \in (0,1)$. By the FT of C, $D \circ J(G) = g$, in other words that $D \circ J = I$. We have for $g \in X$,

$$\|Jg\|_\infty = \sup_{t\in(0,1)} |\int_0^t g| \le \sup_{t\in(0,1)} t\|g\|_\infty \le \|g\|_\infty$$

so $\|J\| \le 1$. Hence $J(D(X)) \subseteq D(X_0)$, and we apply D to see $D(X) \subseteq D(D(X_0))$, in other words, that D is open. As an exercise, show that $C_b(0,1) = X$ is not separable, while X_0 is separable.

7 Geometry and Topology of Banach Spaces

7.1 Bounded Complementation

Let $X \subsetneq Y$ be $\mathbb{F}$ –vector spaces. We can always find a subspace $Z \subset Y$ so $X + Z = Y$ and $X \cap Z = \{0\}$. Indeed, let B be a basis for X, and $B' = B \cup B'$ is a basis for Y, and take $Z = \operatorname{span} B'$.

7.1 Theorem. *Let Y be a Banach space and $X \subsetneq Y$ a closed subspace. Then X admits a closed complement Z if and only if there is some $P \in \mathcal{B}(Y)$ such that $P \circ P = P$ and $\operatorname{im} P = P(Y) = X$.*

Remark. We say that $X \subsetneq Y$ is **boundedly complemented** if either of the above conditions hold.

Proof ($\Leftarrow$) Let $Z = \ker P$, which is closed. If $y \in Y$, then $y = Py + (I - Py$ where $Py \in X$ and $P(I-P)y = 0$ so $(I-P)y \in \ker P$. If $z \in Z \cap X$, then $z = Py$ for some $y \in Y$ so $Pz = P^2 y = Py = z$, but $z \in \ker P$, so $z = Pz = 0$.

($\Rightarrow$) Let $S : X \oplus_1 Z \to Y$ be given by $S(x,z) = x+z$. Then S is surjective and if $(x,z) \in \ker S$, then $x + z = 0$ so $x = -z \in X \cap Z = \{0\}$, hence S is injective. Furthermore,

$$\|S(x+z)\| = \|x+z\| \le \|(x,z)\|_1$$

so $\|S\| \leq 1$. Hence S is a bounded bijection between Banach space and hence S^{-1} is bounded by the inverse mapping theorem. Let $P_1 : X \oplus_1 Z \to X$ be given by $P_1(x,z) = x$; and $J : X \to Y$ by $Jx = x$. Notice that $\|P_1\| = 1$ and $\|J\| = 1$. Define $P : Y \to Y$ by $Py = JP_1S^{-1}y$. Then

- $\operatorname{im} J = X$, and each of P_1, S^{-1} are surjective, so $\operatorname{im} P = X$
- If $y \in Y$, $\|Py\| = \|JP_1S^{-1}y\| \leq \|S^{-1}\|\|y\|$ so $\|P\| \leq \|S^{-1}\|$
- Clearly $P^2 = JP_1S^{-1}JP_1S^{-1} = P$ ■

7.2 Theorem. *c_0 is not boundedly complemented in ℓ_∞.*

Proof Let us assume otherwise; hence, there is $P = P^2 \in \mathcal{B}(\ell_\infty)$ such that $\operatorname{im} P = c_0$. Note that $c_0 = \ker(I-P)$. As in A2, we let $\mathcal{F} \subset \mathcal{P}(\mathbb{N})$ be a family of infinite subsets such that for $E \neq F$ in $\mathcal{F}$, $|E \cap F| < \infty$ and $|\mathcal{F}| = \mathfrak{c}$. For each $F \in \mathcal{F}$, we let $y_F = (I_P)\chi_F \neq 0$. If $\alpha_1, \dots, \alpha_n \in F$ are pairwise distinct, $F_1, \dots, F_m \in \mathcal{F}$, then

$$\sum_{i=1}^{n} \alpha_i \chi_{F_i} = \underbrace{\sum_{i=1}^{m} \alpha_i \chi_{F_i \setminus \bigcup_{j\in[m]\setminus\{i\}} F_j}}_{:=z} + \underbrace{\sum_{k=2}^{m} \sum_{1\leq i<\dots<i_k\leq m} (\alpha_{i_1} + \dots + \alpha_{i_k})\chi_{F_{i_1}\cap\dots\cap F_{i_k}}}_{\in c_0}$$

where $\|z\|_\infty = \max_{k=1,\dots,m} |\alpha_k|$. Hence

$$\left\|\sum_{i=1}^{m} \alpha_i y_{F_i}\right\| = \|(I-P)z\| \leq \|I-P\|\|z\| = \|I-P\| \max_{k=1,\dots,m} |\alpha_k| \tag{7.1}$$

Now, let for $n,k \in \mathbb{N}$, $\mathcal{F}_{n,k} = \{F \in \mathcal{F} : |\delta_k(y_F)| \geq \frac{1}{n}\}$m where $\delta_k(x_i)_{i=1}^\infty = x_k$, so $\delta_k \in \ell_\infty^*$ with $\|\delta_k\| \leq 1$. Let $F_1, \dots, F_m$ be pairwise disjoint in $\mathcal{F}_{n,k}$, and $\alpha_i = \overline{\operatorname{sgn} \delta_k(y_{F_i})}$. Then we have each $|\alpha_i| = 1$, so by (7.1), we find

$$\|I-P\| \geq \left\|\sum_{i=1}^{\infty} \alpha_i y_{F_i}\right\|_\infty \geq |\delta_k \sum_{i=1}^{n} \alpha_i y_{F_i}| = \sum_{i=1}^{m} |\delta_k(y_{F_i})| \geq \frac{m}{n}$$

so $m \leq n\|I-P\|$ and it folows that $\mathcal{F}_{n,k}$ is finite. Since each $y_F \neq 0$ for $F \in \mathcal{F}$, we see that $\mathcal{F} = \bigcup_{n=1}^\infty \bigcup_{k=1}^\infty$, which contradicts that $|\mathcal{F}| = \mathfrak{c}$. Hence such a P must not exist. ■

7.2 Characterization of Finite Dimensionality

7.3 Theorem. *If X is a finite dimensional vector space over $\mathbb{F}$, then any two norms are equivalent.*

Proof Let $\|\cdot\|$ be a norm on X. Fix a basis $(e_1, \dots, e_n)$ for X, and let $x = \sum_{k=1}^n x_k e_k$, $x_i \in \mathbb{F}$, $\|x_k\|_\infty = \max_{k=1,\dots,n} |x_k|$. This is easily checked to be a norm. Moreover, $B_\infty = \{x \in X : \|x\|_\infty \leq 1\}$ admits a homeomorphic identification

$$B_\infty = \begin{cases} [-1,1]^n & \mathbb{F} = \mathbb{R} \\ \overline{D}^n & \mathbb{F} = \mathbb{C} \end{cases}$$

and hence is compact. Thus $S_\infty = \{x \in X : \|x\|_\infty = 1\}$ is compact as well. Hence, for $x = \sum_{k=1}^\infty x_k e_k$, we have

$$\|x\| \leq \sum_{k=1}^n |x_k| \|e_k\| \leq \|x\|_\infty \underbrace{\sum_{k=1}^n \|e_k\|}_{:=M}$$

Now for $x, y \in X$, we have $|\|x\| - \|y\|| \leq \|x-y\| \leq M\|x-y\|_\infty$ so $\|\cdot\|$ is Lipschitz with respect to $\|\cdot\|_\infty$, and hence $\tau_{\|\cdot\|_\infty}$−continuous. Thus the extreme value theorem tells us that $m = \inf_{x \in S_\infty} \|x\| > 0$. Hence for $x \in X \setminus \{0\}$, $\|x\| = \|x\|_\infty \cdot \left\| \frac{1}{\|x\|_\infty} x \right\| \geq \|x\|_\infty m$. In general, $m\|x\|_\infty \leq \|x\| \leq M\|x\|_\infty$. We thus have that $\|\cdot\| \sim \|\cdot\|_\infty$, so any norms are equivalent. ∎

7.4 Corollary. *Let $(X, \|\cdot\|)$ be a finite dimensional normed space. Then*
- *(i) $K \subseteq X$ is compact if an only if K is closed and bounded.*
- *(ii) $(X, \|\cdot\|)$ is a Banach space*
- *(iii) For any normed space Y, we have $\mathcal{L}(X,Y) = \mathcal{B}(X,Y)$*
- *(iv) We have $X' = X^*$.*

Proof (i) The forward direction is immediate. If K is closed and bounded, is contained in some scaled copy of B_∞, which is compact.
- (ii) Cauchy sequences are bounded, and thus contained in some scaled copy of B_∞, which is compact.
- (iii) Let $T \in \mathcal{L}(X,Y)$, and let $\|x\|_0 = \|x\| + \|Tx\|$. Then the result follows by equivalence of norms.
- (iv) Immediate. ∎

7.5 Proposition. *A finite dimensional subspace of normed space is always closed and boundedly complemented.*

Proof Let $Y \subseteq X$ be so Y is finite dimensional and X a normed space. We can find a basis $(e_1, \ldots, e_n)$ for Y. We may assume that each $\|e_k\| = 1$. We define $f_1, \ldots, f_n \in Y' = Y^*$ by

$$f_k\left(\sum_{j=1}^n \alpha_j e_j\right) = \alpha_k$$

By Hahn-Banach, get $F_1, \ldots, F_n \in X^*$ such that $F_k|_Y = f_k$ and $\|F_k\| = \|f_k\|$. Define $P : X \to X$ by $Px = \sum_{k=1}^n F_k(x) e_k$. Notice that $\operatorname{im} P \subseteq Y$ and by choice of $F_k|_Y = f_k$, we have $P|_Y = I_Y$. Thus $P^2 = P$. Finally, for $x \in X$, $\|Px\| \leq \sum_{k=1}^n \|f_k\| \|x\|$ so $\|P\| \leq \sum \|f_k\| < \infty$, i.e. P is bounded. Closedness of Y thus follows from the last corollary. Alternatively, $Y = \ker(I - P)$. ∎

7.3 On Compactness of the Unit Ball

7.6 Lemma. *Let X be a normed space and $Y \subsetneq X$ a closed subspace. Then for any $0 < \epsilon < 1$ there is some $x_0 \in D(X)$ such that $\operatorname{dist}(x_0, Y) > 1 - \epsilon$.*

Proof Let $x \in X \setminus Y$ and $f : \operatorname{span}(\{x\} \cup Y) \to \mathbb{F}$ be given by $f(y + \alpha x) = \alpha$ where $y \in Y$ and $\alpha \in \mathbb{F}$. In particular, since $\ker f = Y$ is a closed proper subspace of $\operatorname{span}(\{x\} \cup Y)$, by Proposition 6.4, f is bounded. Thus apply Hahn-Banach to get $F \in X^*$ with $\|F\| = \|f\|$.

Now, let $x_0 \in D(X)$ be so that $|F(x_0)| > (1-\epsilon)\|F\|$. Since $Y \subseteq \ker F$, for $y \in Y$,

$$\|F\| \left\|x_0 - y\right\| \geq |F(x_0 - y)| = |F(x_0)| > (1-\epsilon)\|F\|$$

so $\left\|x_0 - y\right\| > 1-\epsilon$. But $y \in Y$ was arbitrary, so $\operatorname{dist}(x_0, Y) = \inf_{y\in Y} \left\|x_0 - y\right\| \geq 1-\epsilon$, as required. ■

7.7 Theorem. *Let X be a normed space. Then $B(X)$ is compact if and only if X is finite dimensional.*

Proof To prove the interesting direction, suppose X is not finite dimensional. Let $\epsilon \in (0,1)$ and let $x_1 \in B(X) \setminus \{0\}$. Inductively, for each $n \in \mathbb{N}$, given $\{x_1, \ldots, x_n\}$, by Lemma 7.6 get $x_{n+1} \in B(X)$ so that $\operatorname{dist}(x_{n+1}, \operatorname{span}\{x_1, \ldots, x_n\}) \geq 1-\epsilon$. But then by construction for $i \neq j$ $d(x_i, x_j) > 1-\epsilon$ so $B(X)$ is not totally bounded, and hence not compact. ■

7.4 Weak and Weak* Topologies

Let X be a vector space and $Z \subseteq X'$ a subspace. Then $\sigma(X,Z)$ is the coarsest topology allowing each $f \in Z$ to be continuous, $f : X \to \mathbb{F}$. The basic open sets are given as follows: let $x_0 \in X$, $\epsilon > 0$, and $D = D(\mathbb{F})$, and we consider for $f \in Z$

$$f^{-1}(f(x_0) + \epsilon D) = \underbrace{\{x \in X : |f(x) - f(x_0)| < \epsilon\}}_{\text{"affine hypertube"}} = \{x \in X : |\frac{1}{\epsilon}f(x) - \frac{1}{\epsilon}f(x_0)| < 1\}$$

so that

$$\left\{ \bigcap_{k=1}^{n} \{x \in X : |f_k(x) - f_k(x_0)| < 1\} : f_1, \ldots, f_n \in Z, n \in \mathbb{N} \right\}$$

is a base for $\sigma(X,Z)$.

In partiular, if X is a normed space, then the **weak topology** on X is $\omega = \sigma(X, X^*)$. Certainly $\omega \subseteq \tau_{\|\cdot\|}$. Similarly, the **weak*-topology** on X^* is $\omega^* = \sigma(X^*, \hat{X})$ (recall for $x \in X$, $\hat{x}(f) = f(x)$). Since $\hat{X} \subseteq X^{**}$, we have $\omega^* \subseteq \omega = \sigma(X^*, X^{**}) \subseteq \tau_{\|\cdot\|}$.

Here we have a key feature of the w^*–topology on a normed space, which follows as a consequence to Tychonoff's Theorem:

7.8 Theorem. (Alaoglu) *Let X be a normed space. Then $B(X^*)$ is $w^* = \sigma(X^*, \hat{X})$–compact*

Proof Let $\Gamma : X^* \to \mathbb{F}^X$ be given by $\Gamma(f) = (f(x))_{x\in X}$, so Γ is injective. Let $\pi = \sigma(\mathbb{F}^X, \{p_x\}_{x\in X})$ be the product topology. If $U_1, \ldots, U_n \subseteq \mathbb{F}$ are open and $x_1, \ldots, x_n \in X$, then

$$\Gamma\left(\bigcap_{k=1}^{n} \hat{x}_n^{-1}(U_k)\right) = \bigcap_{k=1}^{n} \Gamma\left(\hat{x}_n^{-1}(U_k)\right) = \bigcap_{k=1}^{n} \hat{x}_n^{-1}(U_k) \cap \Gamma(X^*)$$

so Γ is an open map onto its image in $\mathbb{F}^X$. Similarly, it is easy to show that Γ^{-1} is also an open map, so in fact Γ is a homeomorphism onto its image.

We now consider $\overline{\Gamma(B(X^*))} \subset \mathbb{F}^X$. Let $g \in \overline{\Gamma(B(X^*))}$ and let $D = D(\mathbb{F})$. Given $x, y \in X$ and $\alpha \in \mathbb{F}$, and then given $\epsilon > 0$, we find $f \in B(X^*)$ such that

$$\Gamma(f) \in p_x^{-1}\left(g(x) + \frac{\epsilon}{3}D\right) \cap p_y^{-1}\left(g(y) + \frac{\epsilon}{3(|\alpha|+1)}D\right) \cap p_{x+\alpha y}^{-1}\left(g(x+\alpha y) + \frac{\epsilon}{3}D\right)$$

We have that f is linear with $\Gamma(f)(x) = f(x)$, etc. so we have

$$|g(x) + \alpha g(y) - g(x+\alpha y)| \le |g(x) - f(x)| + |\alpha||g(y) - f(y)| + |g(x+\alpha y) - f(x+\alpha y)| < \epsilon$$

and since $\|f\| \le 1$, we have $|g(x)| \le |g(x) - f(x)| + |f(x)| < \epsilon/3 + \|x\|$. Then since $\epsilon > 0$ is arbitrary, get $g \in X'$ and $|g(x)| \ge \|x\|$, i.e. $g \in B(X^*)$. Hence we have that $g = \Gamma(g)$.

Thus $\Gamma(B(X^*)) \subseteq \prod_{x\in X} \|x\|\overline{D} \subseteq \mathbb{F}^X$ is a closed subset of a compact subset of $\mathbb{F}^X$. Thus $B(X^*)$ is the continuous image of a compact set and hence compact. ■

Remark. If X is a normed space, $w^* = \sigma(X^*, \hat{X})$, if $x \in X$, $\hat{x} \in X^{**}$, $\hat{x}(f) = f(x)$, $\hat{X} = \{\hat{x} : x \in X\}$. If A, B are non-empty sets, $A^B \cong \{f : B \to A\}$.

Remark. If $r > 0$, then we may replace $B(X^*)$ with $rB(X^*)$ in the proof above, with trivial modifications. Thus any ball is w^*–compact. Hence bounded w^*–closed sets in X^* are automatically w^*–compact.

7.9 Theorem. (Metrization) *If X is a separable normed space, then $B(X^*)$ is w^*–metrizable.*

Proof Let $\{x_n\}_{n=1}^\infty \subset B(X)$ be any set which is separating for X^*, i.e. if $f \in X^* \setminus \{0\}$, then $f(x_n) \neq 0$ for some n (for example, take any dense subset of $D(X) \setminus \{0\}$). Let ρ be given by

$$\rho(f,g) = \sum_{k=1}^\infty \frac{|(f-g)(x_k)|}{2^k} \le 2$$

It is easy to see that this is a metric.

Given $f_0 \in B(X^*)$, take $\epsilon > 0$ and let

- n be so $\sum_{k=n+1}^\infty \frac{2}{2^k} < \frac{\epsilon}{2}$, and
- $V = \bigcap_{k=1}^n \{f \in B(X^*) : |\hat{x}_k(f) - \hat{x}_k(f_0)| < \epsilon/2\} \in w^*|_{B(X^*)}$, $f_0 \in V$.

Then if $f \in V$,

$$\rho(f, f_0) = \sum_{k=1}^n \frac{|f(x_k) - f_0(x_k)|}{2^k} + \sum_{k=n+1}^\infty \frac{|f(x_k) - f_0(x_k)|}{2^k} < \epsilon$$

so that V is contained in a metric neighbourhood of f_0. Since f_0 is arbitrary, we have $\tau_\rho \subseteq w^*|_{B(X^*)}$, but since $B(X^*)$ is w^*–compact and τ_ρ is Hausdorff, these must be equal. ■

(i) Note that different separating families from $B(X)$ may produce different metrics, but always the same topology.
(ii) The definition of ρ above extends to all of $X^* \times X^*$. However, X^* with the weak* topology is not in metrizable if X is infinite dimensional.
(iii) $X^* = \bigcup_{=1}^\infty nB(X^*)$, so each $nB(X^*)$ is metrizable and compact, and thus w^*–separable. Thus if X is separable, then X^* is itself separable.

7.5 Weak and Weak* Topologies in Convexity

We begin with a useful technical lemma:

7.10 Lemma. *Let X be an $\mathbb{F}$ –vector space and $f_0, f_1, \dots, f_n \in X'$ with $\ker f_0 \supseteq \bigcap_{i=1}^n \ker f_i$. Then $f \in \operatorname{span}\{f_1, \dots, f_n\}$.*

Proof Define $T : X \to \mathbb{F}^n$ by $Tx = (f_1(x),\dots,f_n(x))$ so $\ker T = \bigcap_{i=1}^n \ker f_i$. Let $R = \operatorname{im} T \subseteq \mathbb{F}^n$ and $g_0 \in R'$ by $g_0(Tx) = f_0(x)$. Certainly g_0 is linear, and g_0 is well-defined since $\ker T \subseteq \ker f_0$. Let $g \in (\mathbb{F}^n)'$ such that $g|_R = g_0$; get $\alpha_i \in \mathbb{F}$ so that $g(y_1,\dots,y_n) = \sum_{j=1}^n \alpha_j y_j$. Hence for $x \in X$,

$$f_0(x) = g_0(Tx) = g(Tx) = g(f_1(x),\dots,f_n(x)) = \sum_{j=1}^n \alpha_j f_j(x)$$

so that $f_0 = \sum_{j=1}^n \alpha_j f_j \in \operatorname{span}\{f_1,\dots,f_n\}$. ■

We now show that we have an analogue of norm separation with respect to the w^*–topology:

7.11 Theorem. (w^*–Separation) *Let X be a normed space, $A,B \subset X^*$ each be non-empty and convex, with $A \cap B = \emptyset$ and B w^*–open. Then there is $x \in X$ and $\alpha \in \mathbb{R}$ such that*

$$\operatorname{Re} f(x) \le \alpha < \operatorname{Re} g(x)$$

for $f \in A$ and $g \in B$.

Proof Since $w^* \subseteq \tau_{\|\cdot\|}$, the norm separation theorem gives $F \in X^{**}$ and $\alpha \in \mathbb{R}$ such that $\operatorname{Re} F(f) \le \alpha < \operatorname{Re} F(g)$ for $f \in A$, $g \in B$. It suffices to show that $F = \hat{x}$ for some $x \in X$: assuming this, we have $\operatorname{Re} f(x) = \operatorname{Re} \hat{x}(f) \le \alpha < \operatorname{Re} \hat{x}(g) = \operatorname{Re} g(x)$, which is the desired result.

Fix some $f_0 \in B$, and by w^*–continuity get $x_1,\dots,x_n$ in X such that

$$f_0 \in U = \bigcap_{i=1}^n \hat{x}_i^{-1}(V_i) \subseteq B$$

where each $V_i \subset \mathbb{F}$ is an open neighbourhood of $f_0(x_i)$. Let $Y = \bigcap_{i=1}^n \ker \hat{x}_i \subseteq X^*$. Then for each $i = 1,\dots,n$,

$$\hat{x}_i(f_0 + Y) = \{f_0(x_i)\} \subset V_i,$$

so that $f_0 + Y \subseteq U \subseteq B$. Thus if $f \in Y$ is arbitrary, then $f_0 + f \in B$ so that $\operatorname{Re} F(f_0 + f) > \alpha$, or equivalently $\operatorname{Re} F(f) > \alpha - \operatorname{Re} F(f_0) =: c$ where c is a fixed quantity. But Y is a vector subspace, and hence closed under scalar multiplication; thus $\alpha \operatorname{Re} F(f) > c$ for any $\alpha \in \mathbb{F}$, forcing $F(f) = 0$. Thus $Y \subseteq \ker F$, and apply Lemma 7.10 so that $F \in \operatorname{span}\{\hat{x}_1,\dots,\hat{x}_n\} \subseteq \hat{X}$. Thus $F = \hat{x}$ for some $x \in X$. ■

A set of the form $H = \{f \in X^* : \operatorname{Re} f(x) \le \alpha\}$ for some fixed $x \in X$ and $\alpha \in \mathbb{R}$ is called a real w^*–halfspace. Note that any such space is w^*–closed and convex.

7.12 Theorem. (w^*–Closed Convex Hull) *If $S \subset X^*$ and $\mathcal{H}$ is the set of real w^*–halfspaces in X, then*

$$\overline{\operatorname{conv}}^{w^*} S = \bigcap_{\substack{H \in \mathcal{H} \\ H \supseteq S}} H.$$

Proof Firstly, $\bigcap_{H \in \mathcal{H}, H \supseteq S} H$ is w^*–closed and convex since it is the intersection of such spaces. Conversely, if $f \in X^* \setminus \overline{\operatorname{conv}}^{w^*} S$, then there is a basic w^*–open neighbourhood

$$B = \bigcap_{j=1}^n \hat{x}_j^{-1}(V_j) \subseteq X^* \setminus \overline{\operatorname{conv}}^{w^*} S$$

so that $B \cap \overline{\text{conv}}^{w^*} S = \emptyset$. Then since B is convex, apply Theorem 7.11 to get a half space not containing B. ∎

The case for weak-closed spaces is easier, and follows from our earlier results on the norm-closed convex hull.

7.13 Corollary. *If $C \subset X$ is convex, then C is norm closed if and only if C is w–closed.*

Proof If C is w–closed then it is automatically norm-closed.

Conversely, suppose we are given a normed space X and closed half space $H = \{x \in X : \operatorname{Re} f(x) \le \alpha\}$ for some f in X^*, $\alpha \in \mathbb{R}$. Then

$$H = (\operatorname{Re} f)^{-1}([\alpha, \infty)) = f^{-1}(\{z \in \mathbb{C} : \operatorname{Re} z \ge \alpha\})$$

is w–closed as well. Since C is convex and norm-closed, by Corollary 6.7, C is the intersection of such half spaces, which are w–closed, so that C is w–closed. ∎

Definition. Let X be a normed space. If $E \subseteq X$, the **polar** of E is given by

$$E^\circ = \bigcap_{x \in E} \{f \in X^* : \operatorname{Re} f(x) \le 1\}.$$

Conversely, if $F \subseteq X^*$, the **pre-polar** of F is given by

$$F_\circ = \bigcap_{f \in F} \{x \in X : \operatorname{Re} f(x) \le 1\}.$$

Note that E° is convex, w^*-closed, and contains 0. Similarly, F_0 is convex, w–closed, and contains 0.

7.14 Theorem. (Bi-polar) *(i) If $\emptyset \ne E \subseteq X$, then $(E^\circ)_\circ = \overline{\text{conv}}(E \cup \{0\})$.*
(ii) If $\emptyset \ne F \subseteq X^*$, then $(F_\circ)^\circ = \overline{\text{conv}}^{w^*}(F \cup \{0\})$.

Proof (i) Since $E \cup \{0\} \subseteq (E^\circ)_\circ$, by Corollary 7.13, $\overline{\text{conv}}(E \cup \{0\}) \subseteq (E^\circ)_\circ$. Conversely, if $x_0 \in X \setminus \overline{\text{conv}}(E \cup \{0\})$, then Corollary 6.7 provides some $f \in X^*$ and $\alpha \in \mathbb{R}$ so that $\operatorname{Re} f(x_0) > \alpha \ge \operatorname{Re} f(x)$ for $x \in E \cup \{0\}$. Notice that $\operatorname{Re} f(x_0) > \alpha \ge \operatorname{Re} f(0) = 0$, so we may set

$$\beta := \frac{1}{2}(\operatorname{Re} f(x_0) + \alpha)$$

with $\beta > 0$. Then $\operatorname{Re} f(x_0) > \beta \ge \operatorname{Re} f(x)$ for any $x \in E \cup \{0\}$. Let $g = \frac{1}{\beta} f$ so that for any $x \in E$, $\operatorname{Re} g(x) = \frac{1}{\beta} \operatorname{Re} f(x) \le 1$ and $g \in E^\circ$, but $\operatorname{Re} g(x_0) > 1$ so $x_0 \notin (E^\circ)_\circ$.

(ii) The proof is similar to above; use Theorem 7.12. ∎

Suppose $Y \subseteq X$ is a subspace. If $f \in Y^0$ such that $\operatorname{Re} f(y) \le 1$ for $y \in Y$, then $f(y) = 0$ for all $y \in Y$. In particular, we define $Y^a = \{f \in X^* : f|_Y = 0\}$ to be the **annhilator** of Y, where $Y^a = Y^0$. Likewise, if $Z \subseteq X^*$ is a subspace, then we define $Z_a = \{x \in X : f(x) = 0 \text{ for each } f \in Z\}$ to be the **pre-annhilator**, where $Z_a = Z_0$. Notice that Y^a and Z_a are subspaces.

Since subspaces are closed and contain 0, we immediately have the following:

7.15 Corollary. *(i) If $Y \subseteq X$ is a subspace, then $(X^a)_a = \overline{X}$.*

(ii) If $Z \subseteq X^*$ is a subspace, then $(Z_a)^a = \overline{Z}^{w^*}$.

In the context of Corollary 7.13, if $S \subset X^*$ is norm-closed, it may not be w^*–closed. [***TODO: example??***] However, this relationship does hold in the special case when $S = B(X^*)$; and in fact, as a consequence of the Bi-polar theorem, we have a bit more:

7.16 Lemma. *If X is a normed space, then $B(X)^\circ = B(X^*)$ and $B(X^*)_\circ = B(X)$.*

Proof If $f \in B(X)^\circ$, then $\operatorname{Re} f(x) \leq 1$ for any $x \in B(X)$. Thus for $x \in B(X)$,

$$|f(x)| = \overline{\operatorname{sgn} f(x)} f(x) = f(\overline{\operatorname{sgn} f(x)} x) \leq 1,$$

so $\|f\| \leq 1$ and $f \in B(X^*)$. Conversely, if $f \in B(X^*)$ and $x \in B(X)$, then $\operatorname{Re} f(x) \leq |f(x)| \leq 1$ so $f \in B(X)^\circ$.

For the second equality, by Theorem 7.14,

$$B(X^*)_\circ = (B(X)^\circ)_\circ = \overline{\operatorname{conv}} B(X) = B(X). \qquad \blacksquare$$

7.17 Theorem. (Goldstine) *If X is a normed space, then $B(\hat{X})$ is weak*-dense in $B(X^{**})$.*

Proof Note that, by definition, $B(X)^\circ = B(\hat{X})_\circ$. Thus by Theorem 7.14 and Lemma 7.16, we have

$$\overline{B(\hat{X})}^{w^*} = \overline{\operatorname{conv}}^{w^*} B(\hat{X}) = (B(\hat{X})_\circ)^\circ = (B(X)^\circ)^\circ = B(X^*)^\circ = B(X^{**})$$

as required. $\blacksquare$

Remark. As an exercise, one can show that the conclusion of the theorem is not true for the norm topology by considering $\mathbf{c}_0$ with bidual ℓ^∞.

Example. (i) Recall that $c_0^* \cong \ell_1$ and $\ell_1^* \cong \ell_\infty$, wherein $c_0 \subseteq \ell_\infty$. Thus by Goldstine, $\overline{B(c_0)}^{w^*} = B(\ell_\infty)$, so $w^* = \sigma(\ell_\infty, \ell_1)$. Since ℓ_1 is separable, we have that $(B(\ell_\infty), w^*)$ is metrizable. In fact, if $x \in \ell_\infty$, then if $x^{(n)} = (x_1, \ldots, x_n, 0, 0, \ldots) \in c_0$, we have $x = w^* - \lim_{n\to\infty} x^{(n)}$.

(ii) $\ell_\infty^* \cong \operatorname{FA}(\mathbb{N})$. But $B(\operatorname{FA}(\mathbb{N}), w^*)$ is not metrizable. Since $\ell_1^* \cong \ell_\infty$, there is a natural isometric embedding $\ell_1 \hookrightarrow \operatorname{FA}(\mathbb{N})$. Then $y^{(n)} = \frac{1}{n}(1, 1, \ldots) \in B(\ell_1)$, and w^*–cluster point of $(y^{(n)})_{n=1}^\infty \subset B(\operatorname{FA}(\mathbb{N}))$ is a Banach limit.

7.18 Corollary. *If $F \in X^{**}$, there always exists a net $(x_\nu)_{\nu\in N} \subset X$ such that*

$$F = w^* - \lim_{\nu \in N} \hat{x}_\nu \text{ and } \|x_\nu\| \leq \|F\|$$

Proof If $F \neq 0$, $\frac{1}{\|F\|} F \in B(X^{**}) = \overline{B(\hat{X})}^{w^*}$, and we may find $(y_\nu)_{\nu\in N} \subset B(X)$ such that $(\hat{y}_\nu)_{\nu\in N} \subset B(\hat{X})$ and $\frac{1}{\|F\|} F = w^* - \lim_{\nu\in N} \hat{y}_\nu$. Let $x_\nu = \|F\| y_\nu$. $\blacksquare$

Consider $\mathcal{F} = w^*_{\frac{1}{\|F\|}F} = \{U \in w^*|_{B(X^{**})} : F \in U\}$ is a filtering family. Each $U \in w^*_{\frac{1}{\|F\|}F}$ has $U \cap B(\hat{X}) \neq \emptyset$ by Goldstine. Let $N_{\mathcal{F}} = \{(x, U) : x \in B(X), \hat{x} \in U, U \in \mathcal{F}\}$. Then $(x_\nu)_{\nu\in N_{\mathcal{F}}} = (x)_{(x,U)\in N_{\mathcal{F}}}$ works.

Definition. A normed space X is **reflexive** if $\hat{X} = X^{**}$.

Notice that $X^{**} = (X^*)^*$ is complete, and the bijection ι given by $\iota(x) = \hat{x}$ is an isometry, so a reflexive space is always complete. Moreover,

$$U = \bigcap_{i=1}^{n} f_i^{-1}(V_i) \iff \iota(U) = \bigcap_{i=1}^{n} \hat{f_i}^{-1}(V_i)$$

where $\hat{f_i} \in \widehat{X^*} \subseteq X^{***}$ is given by $\hat{f_i}(\hat{x}) = f_i(x)$, so that ι is a $w - w^*|_{\hat{X}}$–homeomorphism.

7.19 Theorem. *Let X be a Banach space. The following are equivalent:*
(i) X is reflexive
(ii) $B(X)$ is w–compact
(iii) $w^* = w$ on X^*
(iv) X^* is reflexive.

Proof $(i \Rightarrow ii)$ By assumption, $\widehat{B(X)} = B(\hat{X}) = B(X^{**})$. Since $B(X^{**})$ is w^*–compact by Theorem 7.8, $\iota^{-1}(B(X^{**})) = B(X)$ is w–compact.

$(ii \Rightarrow i)$ If $B(X)$ is w–compact, by continuity of ι, $B(\hat{X})$ is w^*–compact (and therefore w^*–closed), so by Theorem 7.17, $B(\hat{X}) = B(X^{**})$ and $\hat{X} = X^{**}$ so X is reflexive.

$(i \Rightarrow iii)$ We have $w = \sigma(X^*, X^{**}) = \sigma(X^*, \hat{X}) = w^*$.

$(iii \Rightarrow iv)$ Since $B(X^*)$ is w^*–compact, it is w–compact and by the same proof as (ii) implies (i) applied to X^*, we have that X^* is reflexive.

$(iv \Rightarrow i)$ Since $B(\hat{X}) \subset X^{**}$ is norm-closed, by Corollary 7.13, $B(\hat{X})$ is w–closed, where $w = \sigma(X^{**}, X^{***}) = \sigma(X^{**}, \widehat{X^*}) = w^*$, so $B(\hat{X})$ is w^*–closed. Thus by Theorem 7.17, $B(X^{**}) = B(\hat{X})$, as required. ∎

7.20 Corollary. *(i) Any finite dimensional normed space is reflexive.*
(ii) Any closed subspace Y of a normed space X is reflexive.

Proof (i) A finite dimensional normed space is complete, and its closed ball is compact, and thus w–compact as $\tau_{\|\cdot\|} \supseteq w$.

(ii) By Hahn-Banach, $Y^* = X^*|_Y$, so $\sigma(Y, Y^*) = \sigma(Y, X^*|_Y) = \sigma(X, X^*)|_Y$. Now $B(Y) = B(X) \cap Y$ is norm-closed and convex, hence w–closed in $B(X)$. But $B(X)$ is w–compact, so $B(Y)$ is a w–closed subset of a w–compact space and thus w–compact. ∎

7.6 Extreme Points and the Krein-Milman Theorem

Definition. Let X be a vector space and $C \subset X$ convex. A **face** F of C is any non-empty subset such that if $x \in F$, $x = (1-t)y + tz$, $t \in (0,1)$, $y, z \in C$ implies that $y, z \in F$. A **extreme point** of C is a singleton face, i.e. $\operatorname{ext} C = \{x \in C : \{x\} \text{ is a face of } C\}$. Hence $x \in \operatorname{ext} C$ if for any $t \in (0,1)$ and $y, z \in C$, if $x = (1-t)y + tz$ then $x = y = z$.

Remark. (i) Faces of C are not necessarily convex.
(ii) A face F' of a convex face F of C is itself a face of C.
(iii) $\operatorname{ext} F \subseteq \operatorname{ext} C$.
(iv) If $f \in X'$ and $\operatorname{Re} f(C) = [a, b]$, then $(\operatorname{Re} f)^{-1}(\{b\})$ is itself a face of C.

7.21 Theorem. (Krein-Milman) *Let X be a normed space and $C \subset X^*$ convex and w^*–compact. Then $C = \overline{co}^{w^*} \operatorname{ext} C$.*

Proof We first verify that any w^*–closed face of C admits an extreme point. We let $\mathcal{F} = \{F : F \text{ is a convex } w^*\text{–closed face of } C\}$, which is partially ordered by reverse inclusion. If $\mathcal{C}$ is a chain in $\mathcal{F}$ with $F_1, \ldots, F_n \in \mathcal{C}$, we may assume $F_1 \supseteq \cdots \supseteq F_n$ so that $\mathcal{C}$ has the finite intersection property. Thus $\emptyset \neq F_0 = \bigcap_{F \in \mathcal{C}} F$. If $x \in F_0$, $t \in (0,1), y, z \in C$ and $x = (1-t)y + tz$, then $x \in F$ for any $F \in \mathcal{C}$ so $y, z \in F$ for any $f \in \mathcal{C}$. Thus $y, z \in \bigcap_{F \in \mathcal{C}} F = F_0$. Also F_0 is closed, so $F_0 \in \mathcal{F}$. Thus F_0 is an upper bound in $\mathcal{F}$ for $\mathcal{C}$, so by Zorn, get some maximal element M.

Let M be a minimal w^*–closed convex face of F. Then given $x \in X$, $\operatorname{Re}\hat{x} : X^* \to \mathbb{R}$ is w^*–continuous, and hence $\operatorname{Re}\hat{x}(M) = [a_x, b_x]$ since the only compact convex subsets of $\mathbb{R}$ are compact intervals. But then $F_x = (\operatorname{Re}\hat{x})^{-1}(\{b_x\}) \cap M$ is a w^*–closed convex face in M, so that $F_x = M$. If $f, g \in M$, then $\operatorname{Re} f(x) = \operatorname{Re}\hat{x}(f) = b_x = \operatorname{Re}\hat{x}(g) = \operatorname{Re} g(x)$, so $f = g$ and hence $M = \{f\}$ and $f \in \operatorname{ext} F$.

Now let $f_0 \in X^* \setminus \overline{\operatorname{conv}}^{w^*} \operatorname{ext} C$. Since C is w^*–compact and convex, $\operatorname{Re}\hat{x}(C) = [a_x, b_x]$, so $C_x = (\operatorname{Re}\hat{x})^{-1}(\{b_x\}) \cap C$ is a w^*–closed convex face of C. Hence by above, there is $f \in \operatorname{ext} C_x \subseteq \operatorname{ext} C$ with $\operatorname{Re}\hat{x}(f) = b_x$. But then $\operatorname{Re}\hat{x}(f_0) > \alpha \geq \operatorname{Re}\hat{x}(f) = b_x$, so $\operatorname{Re}\hat{x}(f_0) \notin [a_x, b_x] = \operatorname{Re}\hat{x}(C)$, so $f_0 \notin C$. Thus $C \subseteq \overline{\operatorname{conv}}^{w^*} \operatorname{ext} C$, where the converse inclusion is obvious. ■

7.22 Corollary. *(i) If $C \subset X$ is a w–compact convex set, then $C = \overline{\operatorname{conv}} \operatorname{ext} C$.*
(ii) If $C \subset X$ is a norm-compact convex set, then $C = \overline{\operatorname{conv}} \operatorname{ext} C$.

Proof (i) We have that $x \mapsto \hat{x} : X \to \hat{X} \subseteq X^{**}$ is continuous. Hence $\hat{C}$ is w^*–compact in X^{**}, so $x \mapsto \hat{x} : C \to \hat{C}$ is a homeomorphism. In $\hat{C}$, we have

$$\widehat{\overline{\operatorname{conv}}^{w} \operatorname{ext} C} = \overline{\operatorname{conv}}^{w^*} \operatorname{ext} \hat{C} = \hat{C}$$

so that $C = \overline{\operatorname{conv}}^{w} \operatorname{ext} C = \overline{co} \operatorname{ext} C$ by the closed convex hull theorem.

(ii) Since $w \subseteq \tau_{\|\cdot\|}$, any norm-compact is w–compact. ■

Remark. Let X be a normed space. Then $\operatorname{ext} B(X) \subseteq S(X)$.

7.23 Proposition. *Let $1 < p < \infty$. Then $\operatorname{ext} B(\ell_p) = S(\ell_p)$.*

Proof Let $x \in S(\ell_p)$, so $x = (1-t)y + tz$. Then

$$1 = \|x\|_p \leq (1-t)\left\|y\right\|_p + t\|z\|_p \leq 1$$

so that $\left\|y\right\|_p = \|z\|_p = 1$ and $\|x\|_p = (1-t)\left\|y\right\|_p + t\|z\|_p$. Thus by the equality case for Minkowski, there is $s \geq 0$ so $s(1-t)y = tz$. Taking norms, we have $y = z$. ■

7.24 Proposition. *We have $\operatorname{ext} B(c_0) = \emptyset$.*

Proof Let $x = (x_1, x_2, \ldots) \in B(C_0)$. Since $\lim x_n = 0$, get n_0 so $|x_{n_0}| \leq 1/2$. If $x_{n_0} \neq 0$, let $y = (x_1, \ldots, x_{n_0-1}, 2x_{n_0}, x_{n_0+1}, \ldots)$ and $z = (x_1, \ldots, x_{n_0-1}, 0, x_{n_0+1}, \ldots)$, and similarly for $x_{n_0} = 0$. Thus we have in each case that $y, z \in B(c_0)$ and $x = y/2 + z/2$. ■

7.25 Corollary. *There exists no normed space X for which $c_0 \cong X^*$.*

Proof If there were such X, then $B(c_0)$ would be w^*–compact, and hence Krein-Milman would imply $\operatorname{ext} B(c_0) \neq \emptyset$. ■

Definition. Let (X,τ) be a compact Hausdorff space, and let

$$P(X) = \{\mu \in B(C^{\mathbb{R}}(X,\tau)^*) : \mu(1) = 1\}$$

7.26 Theorem. $\operatorname{ext} P(X) = \{\hat{x} : x \in X\}$, *where* $\hat{x}(f) = f(x)$. *Furthermore,* $\overline{\operatorname{conv}}^{w^*} \operatorname{ext} P(X) = P(X)$.

Proof Write $C = C^{\mathbb{R}}(X,\tau)$. Note that $P(X) = B(C^*) \cap \hat{\mathbf{1}}^{-1}(\{1\})$ is w^*–compact and convex. Hence by Krein-Milman, we have that $\overline{\operatorname{conv}}^{w^*} \operatorname{ext} P(X) = P(X)$. It remains to describe $\operatorname{ext} P(X)$.

(I) Some inequalities. Fix $\mu \in P(X)$. If $0 \le f \le 1$ in C, then $0 \le \mathbf{1} - f \le 1$ so $\|f\|_\infty$, $\|\mathbf{1} - f\|_\infty \le 1$. Thus $|\mu(f)| \le 1$ and $|\mathbf{1} - \mu(f)| = |\mu(\mathbf{1} - f)| \le 1$. Thus $0 \le \mu(f) \le 1$. Then if $g \ne 0$ and $g \ge 0$ in C, then we have $\mu(g/\|g\|_\infty) \ge 0$, so $\mu(g) > 0$; if $g \le h$ in C, then $h - g \ge 0$ and $\mu(h) \ge \mu(g)$.

If $g \in C$, $g^+ = \max\{g, 0\}$, $g^- = \max\{-g, 0\} \in C$, and $g = g^+ - g^-$ while $|g| = g^+ + g^-$. Hence if $0 \le f \le 1$ in C and let $\mu_f(g) = \mu(fg)$ for $g \in C$, we have

$$\begin{aligned} |\mu_f(g)| = |\mu_f(g^+ - g^-)| &= |\mu(fg^+) + \mu(fg^-)| \le \mu(fg^+) + \mu(fg^-) = \mu(f(g)) \\ &\le \mu(f\|g\|_\infty) = \mu(f)\|g\|_\infty \end{aligned} \tag{7.2}$$

and, with $f = \mathbf{1}$, we have

$$|\mu(g)| \le \mu(|g|) \tag{7.3}$$

(II) Let $\delta \in \operatorname{ext} P(X)$. We first show for h, g in C that $\delta(hg) = \delta(h)\delta(g)$. To see this, since $\delta \ne 0$, we may find $0 \le f \le 1$ such that $0 < \delta(f) < 1$. Now let $\mu = \frac{1}{\delta(f)}\delta_f$ so, for $g \in C$, (7.2) provides

$$|\mu(g)| = \frac{1}{\delta(f)}|\delta_f(g)| \le \frac{1}{\delta(f)}\delta(f)\left\|y\right\|_\infty = \left\|y\right\|_\infty$$

so that $\mu \in B(C^*)$. We also know that $\mu(\mathbf{1}) = 1$. Hence $\mu \in P(X)$. Likewise, $\nu = \frac{1}{1-\delta(f)}\delta_{\mathbf{1}-f} \in P(X)$. We have that

$$\delta(f)\mu + (1 - \delta(f))\nu = \delta$$

so by assumption on δ, $\mu = \delta$. Thus $\frac{1}{\delta(f)}\delta(fg) = \mu(g) = \delta(g)$, so that $\delta(fg) = \delta(f)\delta(g)$. Note that $C = \operatorname{span}\{f \in C : 0 \le f \le 1\}$, so we get multiplicativity of δ.

Suppose now for each $x \in X$, there exists some $f_x \in \ker\delta$ so that $f_x(x) \ne 0$. Let $U_x = f_x^{-1}(\mathbb{R}\setminus\{0\})$, so $X = \bigcup_{x\in X}\{x\} = \bigcup_{x\in X} U_x$ so there are $x_1, \ldots, x_n$ in X so $X = \bigcup_{j=1}^n U_{x_j}$. Then $f = \sum_{j=1}^n f_{x_j}^2 > 0$ on X (by definition of each U_{x_j}), so $1/f \in C$. Then

$$1 = \delta(\mathbf{1}) = \delta\left(\frac{1}{f}\right)\delta(f) = \delta\left(\frac{1}{f}\right)\sum_{j=1}^n \delta(f_{x_j})^2 = 0$$

since each $f_{x_i} \in \ker\delta$, which is absurd. Hence there is $x \in X$ so $f(x) = 0$ whenever $f \in \ker\delta$, so $\ker\delta \supsetneq \ker\hat{x}$, so $\delta \in \mathbb{R}\hat{x}$ and since $\delta(\mathbf{1}) = 1 = \hat{x}(\mathbf{1})$, so $\delta = \hat{x}$.

(III) If $\hat{x} = (1-t)\mu + t\nu$ and $t \in (0,1)$, $\mu, \nu \in P(X)$, then by (7.3),

$$t|\nu(f)| \le t\nu(|f|) \le \hat{x}(|f|) = |f(x)|$$

so $\ker\nu \supseteq \ker\hat{x}$ and as above, $\nu = \hat{x}$. Then $\mu = \hat{x}$. ∎

Remark. For $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, it is similar to show that $\operatorname{ext} B(C^{\mathbb{F}}(X,\tau)^*) = \{z\hat{x} : z \in \mathbb{F}, |z| = 1, x \in X^*\}$.

Let $PA(\mathbb{N}) = \{\mu \in \mathrm{FA}(\mathbb{N}) : \|\mu\|_{\mathrm{var}} \leq 1, \mu(\mathbb{N}) = 1\}$ so, as above, $PA(\mathbb{N})$ is a $w^* = \sigma(\mathrm{FA}(\mathbb{N}), \ell_\infty)$–compact set.

7.27 Proposition. $\operatorname{ext} PA(\mathbb{N}) = \{\delta_{\mathcal{U}} : \mathcal{U}$ *is an ultrafilter on* $\mathbb{N}\}$

Proof If $\delta \in \operatorname{ext} PA(\mathbb{N})$, let $f_\delta \in \ell_\infty^*$ be as in $A1$. As above, we compute that $f_\delta(\chi_E \chi_F) = f_\delta(\chi_E) f_\delta(\chi_F)$, and we have $\chi_E \chi_F = \chi_{E\cap F}$ and hence $\delta(E \cap F) = \delta(E)\delta(F)$. Hence

$$\mathcal{U} = \{E \subseteq \mathbb{N} : \delta(E) \neq 0\} = \{E \subseteq \mathbb{N} : \delta(E) = 1\}$$

is an ultrafilter. The converse is easy. ∎

8 Euclidean and Hilbert Spaces

Definition. Let X be a vector space over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). A form $[\cdot,\cdot] : X \to \mathbb{F}$ is called **Hermitian** if for x, x', y in X, $\alpha \in \mathbb{F}$, we have
(i) $[x + \alpha x', y] = [x,y] + \alpha [x', y]$
(ii) $\overline{[y,x]} = [x,y]$,
positive if in addition
(iii) $[x,x] \geq 0$ for all $x \in X$
and **non-degenerate** if
(iii') $[x,y] = 0$ for all $y \in X$ implies $x = 0$.

Note that positive Hermitian forms naturally induce a seminorm by $\rho(x) = [x,x]^{1/2}$. In particular, if this form is also non-degenerate, then ρ is a norm. We summarize this result, along with some additional properties, in the following proposition:

8.1 Proposition. *Let $[\cdot,\cdot]$ be a positive Hermitian form and $\rho(x) = [x,x]^{1/2} \in [0,\infty)$. Then for $x, y \in X$ and $\alpha \in \mathbb{F}$, we have*
(i) $\rho(\alpha x) = |\alpha|\rho(x)$
(ii) $|[x,y]| \leq \rho(x)\rho(y)$
(iii) $\rho(x+y) \leq \rho(x) + \rho(y)$
(iv) $[\cdot,\cdot]$ *is non-degenerate if and only if* $[x,x] > 0$ *for* $x \in X \setminus \{0\}$.
Furthermore, in the non-degenerate case,
- *Equality in (ii) occurs if and only if x, y are linearly dependent*
- $[x,y] = \rho(x)\rho(y)$ *if and only if there is $s \geq 0$ such that $x = sy$ or $y = sx$ if and only if equality holds in (iii).*

Proof (i) $\rho(\alpha x) = (\alpha\overline{\alpha}[x,x])^{1,2} = |\alpha|\rho(x)$
(ii) If $\alpha \in F$, then

$$\begin{aligned} 0 \leq [x - \alpha y, x - \alpha y] &= [x,x] - \overline{\alpha}[x,y] - \overline{\overline{\alpha}[x,y]} + |\alpha|^2 [y,y] \\ &= \rho(x)^2 - 2\operatorname{Re}\overline{\alpha}[x,y] + |\alpha|\rho(y)^2 \end{aligned}$$

Set $\alpha = \operatorname{sgn}[x,y]$ so that $\overline{\alpha}[x,y] = |[x,y]|$ so

$$|[x,y]| \leq \frac{1}{2}\left(\rho(x)^2 + \rho(y)^2\right)$$

Then if $t > 0$, by (i),

$$|[x,y]| = \left|\left[tx, \frac{1}{t}y\right]\right| \le \frac{1}{2}(t^2 p(x)^2 + \frac{1}{t^2}p(y)^2).$$

If $\rho(x) = 0$, we let $t \to \infty$ so that $[x,y] = 0$; if $\rho(y) = 0$, we let $t \to 0^+$ and again that $[x,y] = 0$. If $[x,y] \ne 0$, set $t = \rho(y)/\rho(x)$ and we are done.

(iii) By direct computation, we have

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

(iv) By (ii), if $\rho(x)^2 = [x,x] = 0$, then $[x,y] = 0$ for all y. Hence $[\cdot,\cdot]$ is non-degenerate if and only if $[x,x] > 0$ for $x \in X \setminus \{0\}$.
[***TODO: Flesh this out***] If x, y are linearly dependant, then equality holds in (ii) by direct computation. If x, y are not linearly dependent, then the choice of $\alpha = \mathrm{sgn}\,[x,y]$ in (ii) gives strict inequality. The condition $[x,y] = p(x)p(y)$ requires non-negativity of $[x,y]$, showing one is a $R_{\ge 0}$ multiple of the other. This is equivalent to having equality in (iii). ■

Definition. A non-degenerate positive Hermitian form on a vector space $\mathcal{E}$ is called an **inner product**. The pair $(\mathcal{E}, (\cdot,\cdot))$ is called a Euclidean space. If, further, $\mathcal{E}$ is complete with respect to the induced norm $\|x\| = (x,x)^{1/2}$, then we call $(\mathcal{E}, (\cdot,\cdot))$ a **Hilbert space**.

Example. (i) (Euclidean Space) $(C[0,1], \langle\cdot,\cdot\rangle)$ given by $(f,g) = \int_0^1 f\overline{g}$

(ii) (Euclidean Space) Recall $\ell = \{x \in \mathbb{F}^{\mathbb{N}} : x_n = 0$ for all but finitely many $n\}$, and $(\ell, \langle\cdot,\cdot\rangle)$ with $\langle x,y\rangle = \sum_{j=1}^{\infty} x_j \overline{y}_j$

(iii) (Hilbert Space) $(L_2[0,1], (\cdot,\cdot))$, $(f,g) = \int_{[0,1]} f\overline{g}$.

(iv) (Hilbert Space) $(\ell_2, (\cdot,\cdot))$, $(x,y) = \sum_{j=1}^{\infty} x_j \overline{y}_j$ (convergence by Hölder's inequality)

(v) (Non-separable Hilbert Space) Let Γ be an uncountable set. If $a = (a_\gamma)_{\gamma\in\Gamma} \in [0,\infty)^\Gamma$, we let $\mathcal{F} = \{F \subset \Gamma : |F| < \infty\}$. We define $\sum_{\gamma\in\Gamma} a_\gamma = \sup_{F\in\mathcal{F}} \sum_{\gamma\in F} a_\gamma = \lim_{F\in\mathcal{F}} \sum_{\gamma\in F} a_\gamma$ where $\mathcal{F}$ is pre-ordered by inclusion. Suppose that $\sum_{\gamma\in\Gamma} a_\gamma < \infty$. Let $\Gamma_n = \{\gamma \in \Gamma : a_\gamma \ge 1/n\}$ and we have

$$\infty > \sum_{\gamma\in\Gamma} a_\gamma \ge \sup_{F\in\mathcal{F}} \sum_{\gamma\in F\cap\Gamma_n} a_\gamma \ge \sum_{F\in\mathcal{F}} \frac{F\cap\Gamma_n}{n}$$

so that $|\Gamma_n| < \infty$. Thus $\Gamma_a = \{\gamma \in \Gamma : a_\gamma > 0\} = \bigcup_{n=1}^{\infty} \Gamma_n$ is countable.

Now, define $\ell_2(\Gamma) = \{x = (x_\gamma) \in \mathbb{F}^\Gamma : \sum_{\gamma\in\Gamma} |x_\gamma|^2 < \infty\}$. If $x, y \in \ell_2(\Gamma)$, then we may let $\Gamma_{|x|^2} \cup \Gamma_{|y|^2} \subseteq \{\gamma_k\}_{k=1}^{\infty}$ so Hölder's inequality for ℓ_2 says that

$$\sum_{k=1}^{\infty} |x_{\gamma_k} \overline{y}_{\gamma_k}| \le \left(\sum_{k=1}^{\infty} |x_{\gamma_k}|^2\right)^{1/2} \left(\sum_{k=1}^{\infty} |y_{\gamma_k}|^2\right)^{1/2} < \infty.$$

Thus, $\sum_{k=1}^{\infty} x_{\gamma_k}\overline{y_{\gamma_k}}$ is absolutely converging. Write $(x,y)=\sum_{\gamma\in\Gamma} x_\gamma\overline{y_\gamma}=\sum_{k=1}^{\infty} x_{\gamma_k}\overline{y_{\gamma_k}}$. Now if $(x^{(n)})_{n=1}^{\infty}\subset\ell_2(\Gamma)$ is $\|\cdot\|_2$–Cauchy, then $\Gamma'=\bigcup_{n=1}^{\infty}\Gamma_{|x^{(n)}|^2}$ is countable. Then since $\ell_2(\Gamma')\cong\ell_2$ (up to counting Γ'), so the Cauchy sequence has a limit. Thus $\ell_2(\Gamma)$ is a Hilbert space. It is immediate that $(\ell_2(\Gamma),\|\cdot\|_2)$ is non-separable.

(vi) Let $w:\mathbb{N}\to(0,\infty)$. Let $\ell_2^w=\{x\in\mathbb{F}^{\mathbb{N}}:\sum_{k=1}^{\infty}|x_k|^2w(k)<\infty\}$. Notice that if $x,y\in\ell_2^w$, then $(x_kw(k)^{1/2})_{k=1}^{\infty}$, $(y_kw(k)^{1/2})_{k=1}^{\infty}\in\ell_2$, so it follows that

$$(x,y)_w=\sum_{k=1}^{\infty}x_k\overline{y_k}w(k)$$

defines an inner product, and $W:\ell_2^2\to\ell_2$ by $W(x_k)_{k=1}^{\infty}=(x_kw(k)^{1/2})_{k=1}^{\infty}$ is a surjective linear isometry, so ℓ_2^w is a hilbert space.

8.1 Various Identities

Let $[\cdot,\cdot]$ be a Hermitian form on X. Then we have the *polarization identitiy*: then over $\mathbb{R}$, $4[x,y]=[x+y,x+y]-[x-y,x-y]$, and over $\mathbb{C}$, $4[x,y]=\sum_{k=0}^{3}i^k[x+i^ky,x+i^ky]$.

Now suppose $(\mathcal{E},(\cdot,\cdot))$ is a Euclidean space. We say that $x,y\in\mathcal{E}$ are **orthogonal** if $(x,y)=0$ and write $x\perp y$. We call a subset $E\subset\mathcal{E}$ **orthogonal** if $x\neq y\in E$ implies $x\perp y$. We write $x\perp E$ if $x\perp y$ for each $y\in E$. We have

- *Pythagoreans' identity*: if $\{x_1,\dots,x_n\}\subset\mathcal{E}$ orthogonal, then $\left\|\sum_{j=1}^{n}x_j\right\|^2=\sum_{j=1}^{n}\left\|x_j\right\|^2$.
- *Parallelogram law*: $\left\|x+y\right\|^2+\left\|x-y\right\|^2=2\|x\|^2+2\left\|y\right\|^2$.

Note that if $\mathbb{F}=\mathbb{C}$, $(x,y)=\frac{1}{4}\sum_{k=0}^{3}i^k\left\|x+i^ky\right\|^2$ defines an inner product, for any norm satisfying the parallelogram law.

8.2 Proposition. *If $y\in\mathcal{E}$ with $(\mathcal{E},(\cdot,\cdot))$ a Euclidean space, then $f_y:\mathcal{E}\to\mathbb{F}$ by $f_y(x)=(x,y)$ is linear with $\left\|f_y\right\|=\left\|y\right\|$. Furthermore, $|f_y(x)|=\left\|y\right\|$ for $y\neq0$, $x\in B(\mathcal{E})$ if and only if $x=\frac{\zeta}{\|y\|}y$ where $|\zeta|=1$.*

Proof Linearity isfrom an assumption on $(\cdot,\cdot)$. Furthermore, Cauchy-Schwarz tells us that

$$|f_y(x)|=|(x,y)|\le\|x\|\left\|y\right\|\Rightarrow\left\|f_y\right\|\le\left\|y\right\|$$

so the equality case for Cauchy-Schwarz provides the last statement of the proposition, and supplies $\left\|f_y\right\|\ge\left\|y\right\|$. ■

Definition. In a Euclidean space $(\mathcal{E},(\cdot,\cdot))$, a set $E\subset\mathcal{E}$ is called **orthonormal** provided that for $e,e'\in E$,

$$(e,e')=\begin{cases}1 & :e=e'\\ 0 & :e\neq e'\end{cases}$$

8.3 Lemma. (Closest Approximation to Finite) *Let $\{e_1,\dots,e_n\}$ be orthonormal in a Euclidean space $(\mathcal{E},(\cdot,\cdot))$ and $\mathcal{M}=\operatorname{span}\{e_1,\dots,e_n\}$. Then for $x\in\mathcal{E}$ we have that*

(i) $P_{\mathcal{M}}x=\sum_{j=1}^{n}\left(x,e_j\right)e_j$ satisfies that $x-P_{\mathcal{M}}x\perp\mathcal{M}$ and hence $\|x\|^2=\|P_{\mathcal{M}}\|^2+\|x-P_{\mathcal{M}}x\|^2$

(ii) $d(x,\mathcal{M})=\left\|x-\sum_{j=1}^{n}(x,e_j)e_j\right\|^{1/2}$

Proof (i) If $1 \le k \le n$, we have

$$(x - P_{\mathcal{M}}x, e_k) = (x, e_k) - \sum_{j=1}^{n} \left(e, e_j\right)\left(e_j, e_k\right) = (x, e_k) - (x, e_k) = 0$$

and it follows that $x - P_{\mathcal{M}}x \perp \mathcal{M}$. Pythagoras' law provides the second formula.

(ii) Endow $\mathbb{F}^n$ with the usual inner product $\|\cdot\|_2$. By Cauchy-Schwarz, for $x \in \mathcal{E}$ and $\alpha \in \mathbb{F}^n$,

$$\left|\left(((x, e_j))_{j=1}^{n}, \alpha\right)\right| = \left|\sum_{j=1}^{n} (x, e_j)\overline{\alpha}_j\right| \le \left(\sum_{j=1}^{n} \left|\left(x, e_j\right)\right|^2\right)^{1/2} = \|P_{\mathcal{M}}x\| \, \|\alpha\|_2$$

so that

$$\begin{aligned}\left\|x - \sum_{j=1}^{n} \alpha_j e_j\right\|^2 &= \|x\|^2 - 2\,\mathrm{Re} \sum_{i=1}^{n} \left(x, e_j\right)\overline{\alpha}_j + \sum_{j=1}^{n} |\alpha_j|^2 \\ &\ge \|x\|^2 - 2\left|\left(\left(\left(x, e_j\right)\right)_{j=1}^{n}, \alpha\right)\right| + \|\alpha\|_2^2 \\ &\ge \|x\|^2 - 2\,\|P_{\mathcal{M}}x\|\,\|\alpha\|_2 + \|\alpha\|_2^2 \\ &= \|x - P_{\mathcal{M}}x\|^2 + (\|P_{\mathcal{M}}x\| - \|\alpha\|_2)\end{aligned}$$

We get equality above if $x \perp \mathcal{M}$ or otherwise there is some $s \ge 0$ so $\alpha_j = s\left(x, e_j\right)$ for $j = 1, \dots, n$. Hence, in this case,

$$\left\|x - \sum_{j=1}^{n} s\left(x, e_j\right) e_j\right\|^2 = \|x - P_{\mathcal{M}}x\|^2 + \|P_{\mathcal{M}}x\|^2 \,(1 - s)^2$$

which is minimized when $s = 1$. ■

Remark. (i) The proof above shows that $P_{\mathcal{M}}x$ is the unique elemet of $\mathcal{M}$ satisfying $\mathrm{dist}(x, \mathcal{M}) = \|x - P_{\mathcal{M}}x\|$.

(ii) It may be shown that $P_{\mathcal{M}} : \mathcal{E} \to \mathcal{E}$ is linear with $\mathrm{im}\, P_{\mathcal{M}} = \mathcal{M}$, $P_{\mathcal{M}}^2 = P_{\mathcal{M}}$, and $\|P_{\mathcal{M}}\| = 1$ (in other words, this map is actually a projection operator)

8.4 Theorem. (Orthonormal Basis) *Let $(\mathcal{E}, (\cdot,\cdot))$ be a Euclidean space, $E \subset \mathcal{E}$ an orthonormal set. Then the following are equivalent:*

(i) $\overline{\mathrm{span}} E = \mathcal{E}$

(ii) for $x \in \mathcal{E}$, $x = \sum_{e \in E} (x, e)\, e = \lim_{F \in \mathcal{F}} \sum_{e \in F} (x, e)\, e$, where $\mathcal{F} = \{F \subseteq E : |f| < \infty\}$, directed by inclusion (Bessel's identity)

(iii) $\|x\|^2 = \sum_{e \in E} |(x, e)|^2$

(iv) For $x, y \in \mathcal{E}$, $(x, y) = \sum_{e \in E} (x, e)\,(e, y)$ (Parseval's identity).

Proof $(i \Rightarrow ii)$ For $F \in \mathcal{F}$, let $\mathcal{E}_F = \mathrm{span}\, F$, so that $\mathcal{E}_F \subseteq \mathcal{E}_{F'}$ if $F \subseteq F'$ in $\mathcal{F}$ and $\mathrm{span}\, E = \bigcup_{F \in \mathcal{F}} \mathcal{E}_F$. Hence for $x \in \mathcal{E}$, we have

$$0 = \mathrm{dist}(x, \mathrm{span}\, E) = \mathrm{dist}\left(x, \bigcup_{F \in \mathcal{F}} \mathcal{E}_F\right) = \inf_{F \in \mathcal{F}} \mathrm{dist}(x, \mathcal{E}_F) = \lim_{F \in \mathcal{F}} \mathrm{dist}(x, \mathcal{E}_F)$$

Thus by the f.d. approximation lemma, we have

$$0 = \lim_{F\in\mathcal{F}} \operatorname{dist}(x, \mathcal{E}_F) = \lim_{F\in\mathcal{F}} \left\| x - \sum_{e\in F} (x,e)\, e \right\|$$

$(ii \Leftrightarrow iii)$ We have

$$\begin{aligned}
0 &= \lim_{F\in\mathcal{F}} \left\| x - \sum_{e\in F} (x,e)\, e \right\|^2 \\
&= \lim_{F\in\mathcal{F}} \left(\|x\|^2 - 2\operatorname{Re} \sum_{e\in F} \overline{(x,e)}\,(x,e) + \sum_{e\in F} \|(x,e)\|^2 \right) \\
&= \lim_{F\in\mathcal{F}} \left(\|x\|^2 - \sum_{e\in F} |(x,e)^2| \right) \\
&= \|x\|^2 - \sum_{e\in E} |(x,e)|^2
\end{aligned}$$

$(ii \Rightarrow iv)$ Recall that $f_y = (\cdot, y) \in \mathcal{E}^*$ so that

$$(x,y) = f_y\left(\lim_{F\in\mathcal{F}} \sum_{e\in F} (x,e)\, e \right) = \lim_{F\in\mathcal{F}} \sum_{e\in F} (x,e)\, f_y(e) = \sum_{e\in E} (x,e)\,(e,y)$$

$(iv \Rightarrow ii)$ Take $x = y$.
$(iii \Rightarrow i)$ Obvious; $x = \lim_{F\in\mathcal{F}} \sum_{e\in F} (x,e)\, e \in \overline{\operatorname{span} E}$, i.e. $\mathcal{E} \subseteq \overline{\operatorname{span} E} \subseteq \mathcal{E}$. ∎

Definition. Any set $E \subset \mathcal{E}$ satisfying the above conditions is called a **orthonormal basis** for $\mathcal{E}$.

8.5 Theorem. (Gram-Schmidt) *Let $(x_1, x_2, \ldots)$ be a linearly independent sequence in a euclidean space $(\mathcal{E}, (\cdot,\cdot))$. There exists an orthogonal sequence $(z_1, z_2, \ldots)$ which satisfies* $\operatorname{span}\{z_1, \ldots, z_n\} = \operatorname{span}\{x_1, \ldots, x_n\}$ *for $n = 1, 2, \ldots$ so that* $\operatorname{span}\{z_1, z_2, \ldots\} = \operatorname{span}\{x_1, x_2, \ldots\}$.

Proof Let $\mathcal{E}_n = \operatorname{span}\{x_1, \ldots, x_n\}$. We set

$$\begin{aligned}
z_1 &= x_1 & e_1 &= \frac{z_1}{\|z_1\|} \\
z_2 &= x_2 - P_{e_1} x_2 & e_2 &= \frac{z_2}{\|z_2\|} \\
&\vdots \\
z_{n+1} &= x_{n+1} - P_{\mathcal{E}_n} x_{n+1} & e_{n+1} &= \frac{z_{n+1}}{\|z_{n+1}\|}
\end{aligned}$$

where $P_{\mathcal{E}_n} x = \sum_{j=1}^n \left(x, e_j\right) e_j$. Inductively, $z_n \in \mathcal{E}_n$ and $z_n \perp \mathcal{E}_k$ for $k = 1, \ldots, n-1$. Hence each set $\{z_1, \ldots, z_n\}$is orthonormal and $\operatorname{span}\{z_1, \ldots, z_n\} \subseteq \operatorname{span}\{x_1, \ldots, x_n\}$ is of full dimension and hence equal. ∎

8.6 Corollary. *Any separable Euclidean space admits an orthonormal basis.*

Proof Let $\{x_n\}_{n=1}^{\infty}$ be dense in $\mathcal{E}$. Let $n_1 = \min\{n : x_n \neq 0\}$, and $n_{k+1} = \min\{n : x_n \notin \operatorname{span}\{x_{n_1}, \ldots, x_{n_k}\}\}$. Then $\{x_{n_1}, x_{n_2}, \ldots\}$ and normalize to get an orthonormal set $E = \{e_1, e_2, \ldots\}$ which satisfies $\overline{\operatorname{span} E} = \overline{\operatorname{span}}\{x_n\}_{n=1}^{\infty} = \mathcal{E}$. ∎

8.7 Theorem. (Riesz-Fischer) *Let $(\mathcal{E}, (\cdot,\cdot))$ be a Euclidean space. Then $\mathcal{E}$ is a Hilbert space if and only if for any orthonormal set E and $\alpha = (\alpha_e)_{e\in E} \in \ell_2(E)$, we have that $\sum_{e\in E} \alpha_e e \in \mathcal{E}$.*

Proof ($\Longrightarrow$) If $\alpha \in \ell_2(E)$ then $E_\alpha = \{e \in E : \alpha_e \neq 0\}$ is countable, and write $E_\alpha = (e_1, e_2, \ldots)$. If $m < n$, we have

$$\left\| \sum_{k=1}^{n} \alpha_{e_k} e_k - \sum_{k=1}^{m} \alpha_{e_k} e_k \right\|^2 = \sum_{k=m+1}^{n} |\alpha_{e_k}|^2 \leq \sum_{k=n+1}^{\infty} |\alpha_{e_k}|^2 \to 0$$

so $x_\alpha = \sum_{k=1}^{\infty} \alpha_{e_k} e_k = \lim_{n\to\infty} \sum_{k=1}^{n} \alpha_{e_k} e_k$ converges. If $F \in \mathcal{F}$, $F \supseteq \{e_1, \ldots, e_n\}$, then

$$\left\| x_\alpha - \sum_{e\in F} \alpha_e e \right\|^2 = \sum_{e=\{e_1,e_2,\ldots\}\setminus F} |\alpha_e|^2 \leq \sum_{k=n+1}^{\infty} |\alpha_{e_k}|^2 \to 0$$

so $x_\alpha = \sum_{e\in E} \alpha_e e = \lim_{F\in\mathcal{F}} \sum_{e\in F} (x, e)\, e$.

($\Longleftarrow$) Let $(x^{(n)})_{n=1}^{\infty}$ be Cauchy in $\mathcal{E}$. Let $\mathcal{M} = \overline{\operatorname{span}\{x^{(n)}\}_{n=1}^{\infty}} \subset \mathcal{E}$ so $\mathcal{M}$ is separable and admits a countable orthonormal basis $E = (e_1, e_2, \ldots)$. Then we appeal to orthonormal basis to see that for any $x \in \mathcal{M}$, $\sum_{k=1}^{\infty} |(x, e_k)|^2 = \|x\|^2 < \infty$ and $x = \sum_{k=1}^{\infty} (x, e_k)\, e_k$.

Our present assumption show that $U : \ell_2(E) \to \mathcal{M}$ given by $U_\alpha = \sum_{k=1}^{\infty} \alpha_k e_k$ always converges in $\mathcal{M} \subseteq \mathcal{E}$. Then orthonormal basis theorem gives $\|U_\alpha\| = \|\alpha\|_2$ so U is a surjective isometry. We let $\alpha^{(n)} = ((x^{(n)}, e_k))_{k=1}^{\infty} \in \ell_2(E)$, then $\left\|\alpha^{(n)} - \alpha^{(m)}\right\|_2 = \left\|U_\alpha^{(n)} - U_\alpha^{(m)}\right\| = \left\|x^{(n)} - x^{(m)}\right\|$ so $(\alpha^{(n)})_{n=1}^{\infty}$ is Cauchy and in $\ell_2(E)$ and hence admits a limit α. Furthermore,

$$\left\| \sum_{k=1}^{\infty} \alpha_k e_k - x^{(n)} \right\| = \left\| U_\alpha - U_\alpha^{(n)} \right\| = \left\| \alpha - \alpha^{(n)} \right\| \to 0$$

as required. ∎

Definition. If $\emptyset \neq S \subset \mathcal{E}$, $(\mathcal{E}, (\cdot,\cdot))$ a Euclidean space, we define its **perpindicular** by $S^\perp = \{y \in \mathcal{E} : ((x, y) = 0 \text{ for any } x \in S)\}$.

Remark. (i) $S \subseteq T$ implies $T^\perp \subseteq S^\perp$

(ii) $S^\perp = \bigcap_{x\in S} \ker f_x$ and is thus closed.

(iii) $\overline{S}^\perp = S^\perp$, since $\overline{S}^\perp \subseteq \overline{S}^\perp$, and if $y \in S^\perp$ and $x \in \overline{S}$, then $x = \lim x_n$ with $x_n \in S$ so $(x, y) = f_y(x) = f_y \lim x_n = \lim f_y(x_n) = \lim (x_n, y) = 0$.

(iv) $(\overline{\operatorname{span}} S)^\perp = S^\perp$. Notice that $(\operatorname{span} S)^\perp = S^\perp$ (easy test) and use (iii)

(v) $\overline{\operatorname{span}} S \cap S^\perp = \{0\}$.

8.8 Theorem. (Existence of Orthonormal Basis) *Let $(H, (\cdot,\cdot))$ be a Hilbert space.*

(i) Given an orthonormal set $E \subset H$, $P_E : H \to H$, $P_E x = \sum_{e\in E} (x, e)\, e$ satisfies

$$\operatorname{im} P_E \subseteq \overline{\operatorname{span} E} \textit{ for } x \in H, x - P_E x \in E^\perp$$

(ii) H admits an orthonormal basis, i.e. an orthonormal set M such that $\overline{\operatorname{span} M} = H$.

Proof (i) Let $\mathcal{F} = \{F \subseteq E : |F| < \infty\}$ be directed by inclusion, and for $F \in \mathcal{F}$, $\mathcal{E}_F = \operatorname{span} F$. Then as in the proof of OMBT, we have for $x \in H$

$$0 \leq \operatorname{dist}(x, \operatorname{span} E)^2 = \lim_{F \in \mathcal{F}} \operatorname{dist}(x, \mathcal{E}_F)^2 = \|x\|^2 - \sum_{e \in E} |(x,e)|^2$$

so $\sum_{e \in E} |(x,e)|^2 \leq \|x\|^2 < \infty$. Thus appealing to Riesz-Fischer, $P_E x = \sum_{e \in E} (x,e) e$ converges in H. Since $P_E x = \lim_{F \in \mathcal{F}} \sum_{e \in F} (x,e), e$, we see that $P_E x \in \overline{\operatorname{span}} E$, so $\operatorname{im} P_E \subseteq \overline{\operatorname{span}} E$. Moreover, if $e' \in E$, $f_{e'} = (\cdot, e') \in H^*$ so

$$(x - P_E x, e') = (x, e') - f_{e'}\left(\sum_{e \in E} (x,e) e\right) = (x, e') - \sum_{e \in E} (x,e) f_{e'}(e) = -$$

so $x - P_E x \in E^{\perp}$.

(ii) Let $\mathcal{O} = \{E \subseteq H : E \text{ is orthonormal}\}$, which is partially ordered by inclusion. Note that $\emptyset \in \mathcal{O}$ vacuously. If $\mathcal{C} \subseteq \mathcal{O}$ is a chain, then $\bigcup_{E \in \mathcal{C}} \in \mathcal{O}$ is an upper bound for $\mathcal{C}$. By Zorn' get a maximal element M.

Suppose $\overline{\operatorname{span}} M \subsetneq H$, and get $x \in H \setminus \overline{\operatorname{span}} M$ and $y = x - P_M x \in (\overline{\operatorname{span}} M)^{\perp} \setminus \{0\}$. But then $M \subsetneq M \cup \{\frac{1}{\|y\|} y\}$, violating maximality. ∎

8.9 Corollary. *If H is a Hilber space with orthonormal basis E, then the map*

$$U : H \to \ell_2(E), Ux = ((x,e))_{e \in E}$$

is a surjective isometry which respects inner products.

Proof We know $\|x\|^2 = \sum_{e \in E} |(x,e)|^2 = \|Ux\|_2$ from ONBT. It is evident that U is linear and $\operatorname{im} U$ is dense in $\ell_2(E)$ so that U is surjective. Finally, if $x, y \in H$, then

$$(x,y)_H = \sum_{e \in E} (x,e)(e,y) = \sum_{e \in E} (x,e)\overline{(y,e)} = (Ux, Uy)_{\ell_2(E)}$$

as required. ∎

Remark. If each of E, E' is an orthonormal basis for a Euclidean space $(\mathcal{E}, (\cdot,\cdot))$, then $|E| = |E'|$. We let $\mathfrak{k}$ be any countable dense subfield of $\mathbb{F}$. Then $D = \operatorname{span}_{\mathfrak{k}}$, so $|D| = \aleph_0 |E| = |E|$ when $|E|$ is infinite. Since for e', e'' in E', $\|e' - e''\| = \sqrt{2}$, we have that any ball $e' + \frac{1}{\sqrt{2}} D(\mathcal{E})$ contains at least one element of D, and $d_{e'} \neq d_{e''}$ if $e' \neq e''$ in E'. This shows that $|E| \geq |E'|$. Likewise $|E'| \leq |E|$.

8.10 Corollary. (Orthogonal complementation) *Let $(\mathcal{E}, \|\cdot,\cdot\|)$ be a Euclidean space and $\mathcal{M} \subseteq \mathcal{E}$ a subspace which is complete with respect to the norm induced from $(\cdot,\cdot)$, i.e. $(\mathcal{M}, (\cdot,\cdot))$ is a Hilbert space. Then there is a unique operator $P_{\mathcal{M}} = P : \mathcal{E} \to \mathcal{E}$ such that $\operatorname{im} P = \mathcal{M}$ and $\operatorname{im}(I - P) = \mathcal{M}^{\perp}$. Moreover,*

- *P is linear*
- *$\|P\| \leq 1$, $\|P\| = 1$ if $\mathcal{M} \neq \{0\}$*
- *$P^2 = P$*

- *for $x, y \in \mathcal{E}$, $(Px, y) = (Px, Py) = (x, Py)$.*

*Such an operator is called the **orthogonal projection**.*

Proof The theorem above prvides an orthonormal basis E for $\mathcal{M}$. Then P_E, as defined above, satisfies $\operatorname{im} P = \mathcal{M}$ and $\operatorname{im}(I - P) = \mathcal{M}^\perp$. Moreover, if P satisfies those conditions, then for $x \in \mathcal{E}$,

$$Px + x - Px = x = P_E x + x - P_e X$$

so that

$$Px - P_e X = [x - P_E x] - [x - Px] \in \mathcal{M} \cap \mathcal{M}^\perp = \{0\}$$

so $Px = P_e x$. Then if $x, y \in \mathcal{E}$ and $\alpha \in \mathbb{F}$,

$$P(x + \alpha y) + x + \alpha y - P(x + \alpha y) = x + \alpha y = Px + x - Px + \alpha[Py + y - Py]$$

so

$$P(x + \alpha y) - [Px + \alpha py] = x - Px + \alpha[y - Py] - [x + \alpha y - P(x + \alpha y)] \in \mathcal{M} \cap \mathcal{M}^\perp = \{0\}$$

and we have linearity.

If $x \in \mathcal{E}$, Pythagoras tells us that $\|x\|^2 = \|Px\|^2 + \|x - Px\|^2$ so $\|Px\| \leq \|x\|$, i.e. $\|P\| \leq 1$. If $e' \in E$, $Pe' = P_E e' = \sum_{e \in E} (e', e) e = e'$, so $P|_{\operatorname{span} E} = \mathrm{id}$ and by uniqueness of extension of bounded linear functionals (uniformly continuous), we see that $P|_{\mathcal{M}} = \mathrm{id}_{\mathcal{M}}$. This shows that if $\mathcal{M} \neq \{0\}$, $\|P\| = 1$ and $P = P^2$. Furthermore, this also shows that $\operatorname{im} P = \mathcal{M}$. Finally,

$$(Px, y) = (Px, Py + y - Py) = (Px, Py)$$

and likewise $(x, Py) = (Px, Py)$. ■

8.11 Corollary. *Let H be a Hilbert space.*

(i) If $\mathcal{M}$ is a closed subspace, then $(\mathcal{M}^\perp)^\perp = \mathcal{M}$.

(ii) If $\emptyset \neq S \subset H$, then $(S^\perp)^\perp = \overline{\operatorname{span}} S$.

Proof (i) We have $\mathcal{M} \subseteq \mathcal{M}^{\perp\perp}$ and $\mathcal{M}$ is complete and thus admis an orthogonal projection $P_{\mathcal{M}} H \to H$ with $\operatorname{im} P_{\mathcal{M}} = \mathcal{M}$ and $\operatorname{im}(I - P_{\mathcal{M}}) = M^\perp$. Now if $x \in \mathcal{M}^{\perp\perp}$, $P_{\mathcal{M}} x \in \mathcal{M}$ so that $x - P_{\mathcal{M}} x \in \mathcal{M}^{\perp\perp} + \mathcal{M} = \mathcal{M}^{\perp\perp}$ so that $x - P_{\mathcal{M}} x \in \mathcal{M}^\perp$. Thus

$$x - P_{\mathcal{M}} x \in \mathcal{M}^{\perp\perp} \cap \mathcal{M}^\perp = \{0\}$$

so that $x \in P_{\mathcal{M}} x \in \mathcal{M}$. Hence $\mathcal{M}^{\perp\perp} \subseteq \mathcal{M}$.

(ii) We have $(\overline{\operatorname{span}} S)^\perp = S^\perp$ and apply (i). ■

8.12 Theorem. (Riesz-Fréchet) *If H is a Hilbert space and $f \in H^*$, then there is a unique $x_0 \in H$ such that $f = f_{x_0}$; i.e. $f(x) = (x, x_0)$ for all $x \in H$.*

Proof If $f = 0$, let $x_0 = 0$. If $f \neq 0$, $\ker f \subsetneq H$ so $(\ker f)^{\perp\perp} = \ker f$, so $(\ker f)^\perp \neq \{0\}$. If $x_1, x_2 \in (\ker f)^\perp$, then $f(x_2) x_1 - f(x_1) x_2 \in (\ker f)^\perp \cap \ker f = \{0\}$, so that $\dim(\ker f)^\perp = 1$ and $(\ker f)^\perp = \mathbb{F} x_1$. But then $f_{x_1} = (\cdot, x_1)$ has $\ker f_{x_1} = (\mathbb{F} x_1)^\perp = (\ker f)^{\perp\perp} = \ker f$, so there is $\alpha \in \mathbb{F}$ so $f = \alpha f_{x_1} = f_{\overline{\alpha} x_1}$. Set $x_0 = \overline{\alpha} x_1$.

Uniqueness holds since $x \mapsto f_x : H \to H^*$ is an isometry and thus injective. ■

Remark. (i) Many results above may fail in a non-complete Euclidean space. Consider $(\ell, (\cdot,\cdot))$ where ℓ is the space of finitely supported sequences. Define $f : \ell \to \mathbb{F}$ by $f(x) = \sum_{k=1}^{\infty} \frac{1}{k} x_k$. By Hölder, $|f(x)| \le \left(\sum_{k=1}^{\infty} \frac{1}{k^2}\right) \|x\|_2$ so that $f \in \ell^*$. If there were $x^{(0)} \in \ell$ so that $f = f_{x^{(0)}}$ for some $x^{(0)} \in (\ker f)^{\perp} \setminus \{0\}$, we would then have $x_k^{(0)} = (e_k, x^{(0)}) = \frac{1}{k}$, which is non-zero for infinitely many k, giving a contradiction. In fact, $(\ker f)^{\perp} = \{0\}$ so that $(\ker f)^{\perp\perp} = \ell$.

(ii) Let $(\mathcal{E}, (\cdot,\cdot))$ be a Euclidean space. Let $H = \overline{\mathcal{E}}$ be the metrical completion with respect to $\|x\|_2$. If $x, y \in H$, then $x = \lim x_n = \lim x_n'$ with $x_n, x_n' \in \mathcal{E}$, and $y = \lim y_n = \lim y_n'$ similarly. Then

$$\begin{aligned}\left|(x_n, y_n) - (x_n, y_n)\right| &\le |(x_n, y_n) - (x_n, y_m)| + |(x_n, y_m) - (x_n, y_n)| \\ &\le \|x_n\| \left\|y_n - y_m\right\| + \|x_n - x_m\| \left\|y_m\right\|\end{aligned}$$

so that $((x_n, y_n))_{n=1}^{\infty} \subset \mathbb{F}$ is Cauchy, and thus admits a limit. Moreover, $|(x_n, y_n) - (x_n', y_n')| \le \|x_n\| \left\|y_n - y_n'\right\| + \|x_n - x_n'\| \left\|y_n'\right\|$. Thus, $(x, y) = \lim_{n\to\infty} (x_n, y_n) = \lim_{n\to\infty} (x_n', y_n')$ is well-defined on $H \times H$. It is straightforward to verify that this is an inner product, and $\|x\| = \lim_{n\to\infty} \|x_n\| = (x, x)^{1/2}$. Thus the completion of a Euclidean space is a Hilber space.

(iii) As a consequence of (ii), we have $\mathcal{E}^* = \{f_x : x \in H\}$ where $H = \overline{\mathcal{E}}$, as above. Furthermore, $\overline{\mathcal{E}} \cong H^{**}$.

(iv) If H is a Hilbert space, the map $f \mapsto f_x$ from $H \to H^*$ is
- a conjugate linear map: $f_{x+\alpha y} = f_x + \overline{\alpha} f_y$
- an isometry: $\|f_x\| = \|x\|$

8.2 Adjoint Operators

Definition. Let X, Y be vector spaces over $\mathbb{F}$ and $T \in \mathcal{L}(X, Y)$. We define the **adjoint** of T to be the map $T^* \in \mathcal{L}(Y', X')$ given by $T^* f = f \circ T$.

We can interpret the adjoint in terms of the natural bilinear pairing $\langle \cdot, \cdot \rangle : V \times V^* \to \mathbb{F}$ given by $\langle v, f \rangle = f(v)$. In particular, if $T : V \to W$ is linear map, then $T^* : W^* \to V^*$ satisfies

$$\langle Tv, g \rangle_W = g(Tv) = T^* g(v) = \langle v, T^* g \rangle_V.$$

We may also naturally identify $T^{**}|_{\hat{X}} : \hat{X} \to Y^{**}$ with T. Suppose $\hat{x} \in \hat{X}$ is arbitrary: then $T^{**}(\hat{x}) = \hat{x} \circ T^*$, so if $f \in Y^*$, then

$$T^{**}(\hat{x})(f) = \hat{x} \circ T^*(f) = \hat{x}(f \circ T) = f(T(x)).$$

so that $T^{**}(\hat{x}) = \widehat{T(x)}$. In particular, since $x \mapsto \hat{x}$ is an isometry,

$$\begin{aligned}\left\|T^{**}|_{\hat{X}}\right\| &= \sup_{\hat{x} \in B(\hat{X})} \|T^{**} \hat{x}\| = \sup_{x \in B(X)} \left\|\widehat{T(x)}\right\| \\ &= \sup_{x \in B(X)} \|T(x)\| = \|T\|\end{aligned}$$

so that this identification is norm-preserving.

8.13 Proposition. *Let X, Y, Z be normed spaces, $T \in \mathcal{B}(X, Y)$ and $S \in \mathcal{B}(Y, Z)$. Then*

(i) $T^* \in \mathcal{B}(Y^*, X^*)$ *with* $\|T^*\| = \|T\|$
(ii) $T \mapsto T^* : \mathcal{B}(X,Y) \to \mathcal{B}(Y^*, X^*)$ *is a continuous linear map*
(iii) $(S \circ T)^* = T^* \circ S^* \in \mathcal{B}(Z^*, X^*)$.

Proof We prove (i) and (ii); (iii) is easy and left as an exercise.

(i) If $f \in Y^*$, then

$$\|T^* f\| = \sup_{x \in B(X)} |T^* f(x)| \le \|f\| \cdot \sup_{x \in B(X)} \|Tx\| = \|f\| \cdot \|T\|$$

so $\|T^*\| \le \|T\|$. Conversely, by the preceding remarks and applying the above result to T^*, we have $\|T\| = \left\|T^{**}|_{\hat{X}}\right\| \le \|T^{**}\| \le \|T^*\|$ so equality holds.

(ii) Suppose $T = \lim_{\nu \in N} T_\nu$; then

$$\lim_{\nu \in N} \|T_\nu^* - T^*\| = \lim_{\nu \in N} \|(T_\nu - T)^*\| = \lim_{\nu \in N} \|T_\nu - T\| = 0$$

is sequentially continuous, and hence continuous. ■

If H, K are Hilber spaces and $T \in \mathcal{B}(H,K)$, then for each $x \in K$, we define by Riesz-Frechet the element T^*x by the rule $f_{T^*x} = T^* f_x$. Then properties (i) and (iii) still hold, but (ii) is replaced by the observation that $T \mapsto T^*$ is conjugate linear. Note that, by definition, $(Tx, y) = (x, T^*y)$ for $x, y \in H$ and T^* is unique with this property.

8.14 Theorem. (Kernel-Annhilator) *If X and Y are Banach spaces,* $T \in B(X,Y)$, *then* $\ker T = [\mathrm{im}(T^*)]_a$ *and* $\ker(T^*) = (\mathrm{im}\, T)^a$.

Proof We have

$$\ker T = \{x \in X : Tx = 0\} = \{x \in X : T^* g(x) = g(Tx) = 0 \text{ for all } x \in X\} = [\mathrm{im}(T^*)]_a$$

and

$$\ker(T^*) = \{g \in Y^* : T^* g = 0\} = \{g \in Y^* : g(Tx) = T^* g(x) = 0 \text{ for all } x \in X\} = [\mathrm{im}(T)]^a \quad ■$$

Remark. If $T \in \mathcal{B}(H,K)$ where H, K are Hilbert spaces, then $\ker T = (\mathrm{im}\, T^*)^\perp$, identifying $T^{**} = T$ since Hilbert spaces are reflexive.

8.15 Theorem. (Characterization of Invertibility) *Let* X, Y *be Banach spaces,* $T \in B(X,Y)$. *Then TFAE:*
(i) T *is invertible*
(ii) T^* *is invertible*
(iii) $\overline{\mathrm{im}\, T} = Y$ *and* $\inf\{\|Tx\| : x \in S(X)\} > 0$, *we say that* T *is* ***bounded below****, and*
(iv) both T *and* T^* *are bounded below.*

Proof $(i \Rightarrow ii)$ Let $T^{-1} \in \mathcal{B}(Y,X)$, so $I_Y = TT^{-1}$, $I_X = T^{-1}T$. Then $(T^{-1})^* T^* = (TT^{-1})^* = I_Y^* = I_{Y^*}$ and likewise for the reverse.

$(ii \Rightarrow iii)$ By the kernel-annhilator theorem, we have $(\mathrm{im}\, T)^a = \ker(T^*) = \{0\}$ in Y^*, so by annhilator-preannhilator, $\overline{\mathrm{im}\, T} = (\mathrm{im}\, T)^a_a = \{0\}_a = Y$. Now if $x \in S(X)$, find $f \in X^*$ so $f(x) = \|x\| = 1 = \|f\|$ (by Hahn-Banach). Then

$$1 = f(x) = [T^*(T^* {-1}) f](x) = [(T^*)^{-1} f](Tx) \le \left\|(T^*)^{-1} f\right\| \|Tx\| \le \left\|(T^*)^{-1}\right\| \|Tx\|$$

so that $\|Tx\| \geq \frac{1}{\|(T^*)^{-1}\|} > 0$ and T is bounded below.

$(iii \Rightarrow i)$ Let T be bounded below, and set $c = \inf\{\|Tx\|x : x \in S(X)\} > 0$, then for $x \in X \setminus \{0\}$, $\|Tx\| = \|x\| \left\| T\left(\frac{1}{\|x\|}x\right)\right\| \geq c\|x\|$. If $y \in \overline{\operatorname{im} T}$, then $y = \lim y_n$, each $y_n = Tx_n \in \operatorname{im} T$. Then

$$\|x_n - x_m\| \leq \frac{1}{c}\|Tx_n - Tx_m\|$$

so $(x_n)_{n=1}^\infty$ is Cauchy as $(Tx_n)_{n=1}^\infty$ converges. THen $x = \lim x_n \in X$ and by continuity of T, $y = Tx \in \operatorname{im} T$. Notive as well that bounded below implies $\ker T = \{0\}$.

We assume that T is bounded below and $\operatorname{im} T = \overline{\operatorname{im} T} = Y$, so T is bijective, hence invertible.

$(i, ii \Rightarrow iv)$ Use (iii)

$(iv \Rightarrow iii)$ We suppose that T is bounded below, and so is T^*. Then $\{0\} = \ker(T^*)$ in Y^*, so $Y = \{0\}_a = \ker(T^*)_a = \overline{\operatorname{im} T}$ and T is bounded below provides $\operatorname{im} T = \overline{\operatorname{im} T} = Y$, so $\ker T = \{0\}$. ■

Remark. Reasons why $T \in \mathcal{B}(X, Y)$ is not intervible: $\ker T \supsetneq \{0\}$, $\operatorname{im} T \subsetneq Y$, T is not bounded below.

Example. Let $T : \ell_p \to \ell_p$ be given by $T(x_n)_{n=1}^\infty = \left(\frac{1}{n}x_n\right)_{n=1}^\infty$, so $\|T\|-1$. Notice that $\ker T = \{0\}$ and $\overline{\operatorname{im} T} = \ell_p$. However, T is not bounded below.

8.3 Spectral Theory for Bounded operators

Let X be a $\mathbb{C}$−Banach space, and $\mathcal{B}(X) = \mathcal{B}(X, X)$.

Definition. If $T \in \mathcal{V}(X)$, we define the **resolvent** of T by $\rho(T) = \{\lambda \in \mathbb{C} : \lambda I - T \text{ is invertible}\}$. Then the **spectrum** of T, $\sigma(T)$, is given by $\sigma(T) = \mathbb{C} \setminus \rho(T)$. We define the **point spectrum** $\sigma_p(T) = \{\lambda \in \mathbb{C} : \ker(\lambda I - T) \supsetneq \{0\}\}$, so $\sigma_p(T) \subseteq \sigma(T)$.

Example.
(i) If X is finite dimensional, then $\sigma(T) = \sigma_p(T)$.
(ii) Let $1 \leq p < \infty$ and define $S : \ell_p \to \ell_p$ by $S(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots)$. Notice that S is linear and $\|Sx\|_p = \|x\|_p$, so $\|S\| = 1$; as well, $\ker S = \{0\}$ so that $0 \notin \sigma_p(S)$. Suppose for contradiction $0 \neq \lambda \in \sigma_p(S)$, so get some $0 \neq x$ so that $Sx = \lambda x$. Let k be such that $x_k \neq 0$; but then $0 = (S^k x)_k = \lambda^k x_k \neq 0$, a contradiction, so that $\sigma_p(S) = \emptyset$.

For any $T \in \mathcal{B}(X)$, is $\sigma(T) \neq \emptyset$?

Let

$$\mathcal{G}(X) = \{T \in \mathcal{B}(X) : T \text{ is invertible}\}$$

Notive that if $S, T \in \mathcal{G}(X)$, then $(ST)^{-1} = T^{-1}S^{-1}$, so $\mathcal{G}(X)$ is a group in $\mathcal{B}(X)$ with identity I. Note that $\mathcal{B}(X)$ is compelte, and if $S, T \in \mathcal{B}(X)$, then $\|ST\| \leq \|S\|\|T\|$, so that $S \mapsto ST$ and $S \mapsto TS$ for some $T \in \mathcal{B}(X)$ are continuous.

8.16 Theorem. (Inversion) *(i) If $T \in D(X)$, then $\sum_{k=0}^\infty T^k$ converges in $\mathcal{B}(X)$, and $\sum_{k=0}^\infty T^k = (I - T)^{-1}$.*

(ii) If $S, T \in \mathcal{B}(X)$ such that $S \in \mathcal{G}(X)$ and $\|T - S\| < \frac{1}{\|S^{-1}\|}$, then $T \in \mathcal{G}(X)$ with $T^{-1} = S^{-1} + \sum_{k=1}^\infty [S^{-1}(S - T)]^k S$.

Thus, we find that $\mathcal{G}(X)$ is open in $\mathcal{B}(X)$ and $T \mapsto T^{-1}$ on $\mathcal{G}(X)$ is continuous.

Proof (i) Let $S_n = \sum_{k=0}^{\infty} T^k$, so for $m < n$, we have

$$\|S_n - S_m\| \le \sum_{k=m+1}^{\infty} \left\|T^k\right\| \le \sum_{k=m+1}^{n} \|T\|^k = \frac{\|T\|^{m+1}}{1-\|T\|} \to 0$$

since $\|T\| < 1$, so $(S_n)_{n=1}^{\infty}$ is Cauchy, and thus convergent in $\mathcal{B}(X)$. Also,

$$(I-T)S_n = I - T^{n+1} \to U \text{ as } T^{n+1} \to 0$$

since $\|T\| < 1$. Similarly, $S_n(I-T) \to I$, so that $S = \sum_{k=0}^{\infty} T^k$ has $S(I-T) = I = (I-T)S$.

(ii) We have $\left\|S^{-1}S - T)\right\| \le \left\|S^{-1}\right\| \|S-T\| < 1$ so by (i)

$$T = S - (S-T) = S[I - S^{-1}(S-T)] \in \mathcal{G}(X)$$

Furthermore,

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

In particular, we see that for $S \in \mathcal{G}(X)$, $S + \frac{1}{\|S^{-1}\|} D(X) \subseteq \mathcal{G}(X$, so (a) holds. From (ii), we see that

$$\left\|T^{-1} - S^{-1}\right\| \le \sum_{k=1}^{\infty} \left\|[S^{-1}(T-S)]^k S\right\| \le \sum_{k=1}^{\infty} \left\|S^{-1}\right\|^k \|T-S\|^k \left\|S^{-1}\right\| = \frac{\left\|S^{-1}\right\|^2 \|T-S\|}{1 - \left\|S^{-1}\right\| \|T-S\|}$$

so that $\lim_{T\to S} \left\|T^{-1} - S^{-1}\right\| = 0$. ∎

Definition. Suppose $\mathcal{B}$ is a $\mathbb{C}$–Banach space, $U \subseteq \mathbb{C}$ and $F : U \to \mathcal{B}$. We say that F is **holomorphic** if for any $z_0 \in U$ the limit

$$F'(z_0) = \lim_{z\to z_0} \frac{1}{z - z_0}[F(z) - F(z_0)]$$

exists.

Remark. Just as in calculus, a holomorphic funtion is continuous on its domain.

8.17 Proposition. *Let $T \in \mathcal{B}(X)$. Then*

(i) $\rho(T)$ is open in $\mathbb{C}$

*(ii) $R(z) = R_T(z) = (zI - T)^{-1}$ defines a holomorphic function on $\rho(T)$, called the **resolvent function**, and*

(iii) $\sigma(T) \subseteq \|T\| \overline{\mathbb{D}}$, and for $|z| > \|T\|$, $R(z) \le \frac{1}{|z| - \|T\|}$.

Proof (i) Define $F : \mathbb{C} \to \mathcal{B}(X)$ by $F(z) = zI - T$. Then F is continuous and $\rho(T) = F^{-1}(\mathcal{G}(X))$.

(ii) If $z, z_0 \in \rho(T)$, then

$$\begin{aligned} R(z) - R(z_0) &= (zI-T)^{-1} - (z_0I-T)^{-1} = (zI-T)^{-1}[(z_0I-T) - (zI-T)](z_0I-T)^{-1} \\ &= (z_0 - z)(zI-T)^{-1}(z_0I-T)^{-1} \end{aligned}$$

Hence

$$\frac{1}{z - z_0}[R(z) - R(z_0)] = -(zI-T)^{-1}(z_0I-T)^{-1} \to -(z_0I-T)^{-2}$$

by continuity of inversion.

(iii) If $|z| > \|T\|$, then $\left\|\frac{1}{z}T\right\| < 1$ so $zI - T = z(I - \frac{1}{z}T) \in \mathcal{G}(X)$, so $\sigma(T) \subseteq \|T\|\overline{\mathbb{D}}$. Furthermore, for $|z| > \|T\|$, we have

$$R(z) = (zI - T)^{-1} = \frac{1}{z}(I - \frac{1}{z}T)^{-1} = \frac{1}{z}\sum_{k=0}^{\infty} \frac{1}{z^k} T^k$$ ■

8.18 Theorem. (Liouville) *If $f : \mathbb{C} \to \mathbb{C}$ is holomorphic and bounded, then f is constant.*

Proof Apply the Cauchy integral formula. ■

8.19 Theorem. (Liouville for Banach Spaces) *If $F : \mathbb{C} \to \mathcal{B}$ is holomorphic and bounded, then F is constant.*

Proof Let $f \in \mathcal{B}^*$ and let $F_f = f \circ F : \mathbb{C} \to \mathbb{C}$. Notice that for $z, z_0 \in \mathbb{C}$,

$$\frac{F_f(z) - F_f(z_0)}{z - z_0} = f\left(\frac{1}{z - z_0}[F(z) - F(z_0)]\right) \to f(F^1(z_0))$$

by linearity and continuity of f, and hence $F_f' = f \circ F'$. Also, if F is bounded, then for $z \in \mathbb{C}$, $|F_f(z)| = |f(F(z))| \leq \|f\|\|F(z)\|$ shows that F_f is bounded, so by Liouville's theorem, is constant. In particular, if $z, z' \in \mathbb{C}$, $f(F(z) - F(z')) = F_f(z) - F_f(z') = 0$. Thus by Hahn-Banach, we have $F(z) = F(z')$ for any $z, z' \in \mathbb{C}$, so F is constant. ■

8.20 Theorem. *If $T \in \mathcal{B}(X)$, then $\sigma(T) \neq \emptyset$ and compact.*

Proof If $\sigma(T) = \emptyset$, then $R : \mathbb{C} \to \mathcal{B}(X)$ is holomorphic. Hence, as $\|T\|\overline{\mathbb{D}}$ is compact in $\mathbb{C}$, R is bounded on $\|T\|\overline{\mathbb{D}}$; and if $|z| > \|T\|$, we have

$$\|R(z)\| \leq \frac{1}{|z| - \|T\|} \to 0$$

It follows that R is bounded on $\mathcal{B}(X)$, and hence constant, and thus $R = 0$. But $R(z)(zI - T) = I$, a contradiction.

Moreover, $\rho(T) = \mathbb{C} \setminus \sigma(T)$ is open, and $\sigma(T) \subseteq \|T\|\overline{\mathbb{D}} \subset \mathbb{C}$. Thus $\sigma(T)$ is a non-empty compact set. ■

8.21 Corollary. (Joke) *$\mathbb{C}$ is algebraically closed.*

Proof Let $p(x) \in \mathbb{C}[x]$ be an arbitrary irreducible polynomial with $p(x) = (x - r_1)\cdots(x - r_n)$ for some $r_i \in \overline{\mathbb{C}}$. Consider the operator $T : \mathbb{C}^n \to \mathbb{C}^n$ with diagonal $r_1, \ldots, r_n$ and hence characteristic polynomial $p(x)$. Then $\emptyset \neq \sigma(T) = \sigma_p(T) = \{x \in \mathbb{C} : p(x) = 0\}$, so p has some root in $\mathbb{C}$, so that $\deg p = 1$. ■

8.22 Proposition. *(i) If X is a (non-Hilbert) Banach space, then $\sigma(T^*) = \sigma(T)$.*
(ii) If H is a Hilbert space, $T \in \mathcal{B}(H)$, then $\sigma(T^*) = \{\overline{\lambda} : \lambda \in \sigma(T)\}$.

Proof (i) $(\lambda I_X - T)^* = \lambda I_{X^*} - T^*$ and is invertible if and only if $\lambda I_X - T$ is invertible
(ii) Same. ■

Definition. We define the **point spectrum** $\sigma_p(T) = \{\lambda \in \mathbb{C} : \ker(\lambda I - T) \neq \{0\}\}$, the **approximate point spectrum** $\sigma_{ap}(T) = \{\lambda \in \mathbb{C} : \lambda I - T \text{ is not bounded below}\}$, and the **compression spectrum** $\sigma_{com}(T) = \{\lambda \in \mathbb{C} : \overline{\operatorname{im}}(\lambda I - T) \subsetneq X\}$.

Remark. (i) $\sigma_p(T) \subseteq \sigma_{ap}(T)$.

(ii) We have $[\operatorname{im}(\lambda I - T)]^a = \ker(\lambda I - T^*)$ by kernel-annhilator so $\overline{\operatorname{im}}(\lambda I - T) = \ker(\lambda I - T^*)$ by annhilator-preannhilator, so that $\sigma_{com}(T) = \sigma_p(T^*)$.

8.23 Lemma. *If $(T_n)_{n=1}^\infty \subset \mathcal{G}(X)$ satisfies that*

- $T = \lim_{n\to\infty} T_n$
- $M = \sup_{n\in\mathbb{N}} \left\|T_n^{-1}\right\| < \infty$

then $T \in \mathcal{G}(X)$.

Proof Since $M > 0$, for sufficiently large n, we have $\|T - T_n\| \leq \frac{1}{M} \leq \frac{1}{\|T_n^{-1}\|}$, so $T \in \mathcal{G}(X)$ by inversion theorem. ∎

8.24 Proposition. *(i) $\partial\sigma(T) \subseteq \sigma_{ap}(T)$*

(ii) $\sigma_{ap}(T)$ is closed

Hence $\sigma_{ap}(T)$ is always a non-empty closed subset of $\mathbb{C}$.

Proof (i) Let $\lambda \in \partial\sigma(T)$, so there is $(\lambda_n)_{n=1}^\infty \subset \rho(T) = \mathbb{C} \setminus \sigma(T)$ such that $\lambda = \lim_{n\to\infty} \lambda_n$. Then $\|(\lambda_n I - T) - (\lambda I - T)\| = |\lambda_n - \lambda| \to 0$, but $\lambda I - T \notin \mathcal{G}(X)$, so by the lemma, $\sup_{n\in\mathbb{N}} \left\|(\lambda_n I - T)^{-1}\right\| = \infty$. Passing to a subsequence if necessary, we may suppose $\lim_{n\to\infty} \left\|(\lambda_n I - T)^{-1}\right\| = \infty$.

For each index n, let $x_n \in S(X)$ so $\alpha_n = \left\|(\lambda_n I - T)^{-1} x_n\right\| > \left\|(\lambda n - T)^{-1}\right\| - \frac{1}{n}$ so $\lim_{n\to\infty} \alpha_n = \infty$. Then $y_n = \frac{1}{\alpha_n}(\lambda_n I - T)^{-1} x_n$, so $y_n \in S(X)$ and

$$\begin{aligned}(\lambda I - T)y_n &= (\lambda_n I - T)y_n + (\lambda - \lambda_n)y_n \\ &= \frac{1}{\alpha_n} x_n + (\lambda - \lambda_n)y_n \to 0\end{aligned}$$

so $\lambda I - T$ is not bounded below.

(ii) If $\lambda = \lim_{n\to\infty} \lambda_n$, each $\lambda_n \in \sigma_{ap}(T)$, for each n find $x_n \in S(X)$ so $\|(\lambda_n I - T)x_n\| < \frac{1}{n}$. Then

$$\|(\lambda I - T)x_n\| \leq \|(\lambda_n I - T)x_n\| + \|(\lambda - \lambda_n)x_n\| < \frac{1}{n} + |\lambda - \lambda_n| \to 0$$

so $\lambda I - T$ is not bounded below. ∎

Example. Let $S \in B(\ell_p)$, $1 < p < \infty$, where $S(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$, the *unilateral shift* map. It is immediate that $\|Sx\|_p = \|x\|_p$ for $x \in \ell_p$, so $\|S\| = 1$. Recall that $\ell_p^* \cong \ell_q$ where p, q are conjugate. Define a bilinear form on $\ell_p \times \ell_q$ by $\langle x, y\rangle = \sum_{k=1}^\infty x_k y_k$. We compute $\langle x, S^* y\rangle = \langle Sx, y\rangle = \sum_{k=1}^\infty x_k y_{k+1} = \langle x, (y_2, y_3, \dots)\rangle$ so $S^*(y_1, y_2, \dots) = (y_2, y_3, \dots)$. Recall that $\sigma_p(S) = \emptyset$. However, if $\lambda \in \mathbb{D}$, let $y_\lambda = (1, \lambda, \lambda^2, \dots) \in \ell_q$. Then $S^*(y_\lambda) = \lambda \cdot y_\lambda$. Hence $\sigma_p(S^*) \supseteq \mathbb{D}$. Furthermore, if $\lambda \in \sigma_p(S^*)$ and $y \in \ker(\lambda I - S^*)$, then $\lambda^n y = (S^*)^n \to 0$ so $\lambda^n \to 0$, forcing $|\lambda| < 1$. Thus

$$\mathbb{D} = \sigma_p(S^*) \subseteq \sigma(S^*) = \sigma(S) \subseteq \overline{D}$$

since $\|S\| = 1$, and since $\sigma(S)$ is compact, $\sigma(S) = \overline{D}$.

We know that $\sigma_{ap}(S) \supseteq \partial\sigma(S) = \mathbb{S}$. If $\lambda \in \mathbb{D}$, then for $x \in \ell_p$, $\|(S - \lambda I)x\|_p \geq \|Sx\|_p - \|\lambda x\|_p = (1 - |\lambda|)\|x\|_p$, so $S - \lambda I$ is bounded below. Thus $\sigma_{ap}(S) \cap \mathbb{S} = \emptyset$, so $\sigma_{ap}(S) = \mathbb{S}$. In conclusion,

$$
\begin{aligned}
\sigma(S) &= \mathbb{D} & \sigma_p(S) &= \emptyset \\
\sigma_{ap}(S) &= \mathbb{S} = \partial\sigma(S) & \sigma_{com}(S) &= \sigma_p(S^*) = \mathbb{D} \\
\sigma(S^*) &= \overline{\mathbb{D}} & \sigma_p(S^*)&^{=}\mathbb{D} \\
\sigma_{ap}(S^*) &= \partial\sigma(S^*) \cup \sigma_p(S^*) = \overline{\mathbb{D}} & \sigma_{com}(S^*) &= \emptyset
\end{aligned}
$$

Remark. Let $\sigma_p(T)$, $\sigma_{com}(T)$ may be empty, and if non-empty need not be closed.

Remark. If $p = 1$ and $S \in B(\ell_1)$ is the unilateral shift, as above, and $L \in \ell_\infty^*$ be a Banach limit. Then

$$S^{**}L = L \circ S^L$$

so $\sigma_p(S^{**}) \ni 1$. Thus $\sigma_{com}(S^*) = \sigma_p(S^{**}) \neq \emptyset$.

8.25 Theorem. (Spectral Mapping) *Let $T \in \mathcal{B}(X)$, $p \in \mathbb{C}[x]$, then $\sigma(p(T)) = p(\sigma(T))$.*

Proof We may assume that $p \neq 0$. Let $\lambda_0 \in \mathbb{C}$ and write $p(t) - \lambda_0 = \alpha \prod_{k=1}^n (t - \lambda_k)$. Then

$$p(T) - \lambda_0 = \alpha \prod_{k=1}^{n} (T - \lambda_k I)$$

Thus $p(T) - \lambda_0 I \notin \mathcal{G}(X)$ if and only some $T - \lambda_I \notin \mathcal{G}(X)$, so $\lambda_0 \in \sigma(p(T))$ if and only if at least one $\lambda_k \in \sigma(T)$ if and only if $p(\lambda) - \lambda_0 = 0$ for some $\lambda \in \sigma(T)$, i.e. $\lambda_0 = p(\lambda) \in p(\sigma(T))$.■

8.26 Theorem. (Spectral Radius Formula) *If $T \in \mathcal{B}(X)$, let $r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$. Then $r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$.*

Proof By the spectral mapping theorem, $r(T^n) = r(T)^n$. Moreover, $r(T^n) \leq \|T^n\|$ since $\sigma(T^n_{\subseteq}\|T^n\|\overline{\mathbb{D}}$. Thus $r(T) = r(T^n)^{1/n} \leq \|T^n\|^{1/n}$ so that $r(T) \leq \liminf_{n\to\infty} \|T^n\|^{1/n}$.

Now, let $f \in \mathcal{B}(X)^*$. We recall that for $|z| > \|T\|$,

$$
\begin{aligned}
R(z) = (zI - T)^{-1} &= \frac{1}{z}(I - \frac{1}{z}T)^{-1} = \frac{1}{z}\sum_{k=0}^{\infty} \frac{1}{z^k} T^k \\
&= \sum_{k=1}^{\infty} \frac{1}{z^k} T^{k-1}.
\end{aligned}
$$

Thus the Holomorphic funciton $F_f = f \circ R : \rho(T) = \mathbb{C} \setminus \sigma(T) \to \mathbb{C}$ satisfies $F_f(z) = \sum_{k=1}^\infty f(T^{k-1})\frac{1}{z^k}$. From $\mathbb{C}$–analysis, the holomorphic function admits a Laurent series for all z with $|z| > r(T)$, and by uniqueness of Laurent series, $F_f(z) = \sum_{n=1}^\infty f(T^{k-1})\frac{1}{z^k}$ for $|z| > r(T)$. Hence if $|z_0| > r(T)$, we have that $\left\{ f\left(\frac{1}{z_0^k} T^k\right)\right\}_{k=1}^{\infty}$ is bounded in $\mathbb{C}$. Doing this for any $f \in \mathcal{B}(X)^*$, we may apply Banach-Steinhaus to see that $\left\{\frac{1}{z_0}^k T^k\right\}_{k=1}^{\infty}$ is bounded in $\mathcal{B}(X)$, so

$$\sup_{n\in\mathbb{N}} \left\| \frac{1}{z_0^n} T^n \right\| \leq M < \infty$$

Then $\|T^n\| \leq M|z_0|^n$ so $\|T^n\|^{1/n} \leq M^{1/n}|z_0|$, so that

$$\limsup_{n\to\infty} \|T^n\|^{1/n} \leq |z_0|.$$

This applies for any $|z_0| > r(T)$, so the result follows. ∎

8.4 Compact Operators

Definition. Let X, Y be Banach spaces over $\mathbb{F}$. A linear opeartor $K : X \to Y$ is called **compact** if $\overline{K(B(X))}$ is compact in Y. We let

$$\mathcal{K}(X,Y) = \{K \in \mathcal{L}(X,Y) : K \text{ is compact}\}.$$

Since compact sets in Y are bounded, we immediately see that $K(X,Y) \subseteq \mathcal{B}(X,Y)$.

8.27 Proposition. *Let X, Y be Banach spaces. Then*
(i) $\mathcal{K}(X,Y)$ is a closed subspace of $\mathcal{B}(X,Y)$, and
(ii) If W, Z are also Banach spaces, $S \in \mathcal{B}(Y,Z)$, $T \in \mathcal{B}(W,X)$, $K \in \mathcal{K}(X,Y)$, then $SKT \in \mathcal{K}(W,Z)$.

Proof (i) Let $K, L \in \mathcal{K}(X,Y)$, $\alpha \in \mathbb{F}$, and $(x_n)_{n=1}^\infty \subset B(X)$. Then $\overline{K(B(X))} \times \overline{L(B(X))}$ is compact in $Y \times Y \cong Y \oplus_1 Y$, so we have converging

$$\left(\left(Kx_{n_j}, Lx_{n_j}\right)\right)_{j=1}^\infty \subset \overline{K(B(X))} \times \overline{L(B(X))}$$

Now, $(K+\alpha L)(x_{n_j}) = Kx_{n_j} + \alpha L x_{n_j} \to y + \alpha y'$, so $((K+\alpha L)x_n)_{n=1}^\infty$ admits a converging subsequence in Y.

Now let $K \in \overline{\mathcal{K}(X,Y)} \subseteq \mathcal{B}(X,Y)$, so $K = \lim_{n\to\infty} K_n$ with each $K_n \in \mathcal{K}(X,Y)$. Let $\epsilon > 0$. Let $n_0 \in \mathbb{N}$ be so $n \geq n_0$ implies $\|K - K_n\| < \epsilon/3$. Since K_{n_0} is compact, there are $\{x_1, \ldots, x_m\} \subset B(X)$ such that

$$K_{n_0}(B(X)) \subseteq \bigcup_{j=1}^m (K_{n_0}x_j + \frac{\epsilon}{3}B(Y))$$

Then for $x \in B(X)$, find x_j so $K_{n_0}x \in K_{n_0}x_j + \frac{\epsilon}{3}B(X)$ and hence

$$\begin{aligned}\left\|Kx - Kx_j\right\| &\leq \left\|Kx - K_{n_0}x\right\| + \left\|K_{n_0}x - K_{n_0}x_j\right\| + \left\|K_{n_0}x_j - Kx_j\right\| \\ &\leq \|K - K_0\|\|x\| + \frac{\epsilon}{3} + \left\|K - K_{n_0}\right\|\left\|x_j\right\| < \epsilon\end{aligned}$$

and we see that $K(B(X)) \subseteq \bigcup_{j=1}^m (Kx + \epsilon B(Y))$, and hence is totally bounded.

(ii) We have

$$\begin{aligned}\overline{SKT(B(W))} &\subseteq \overline{SK(\|T\|B(X))} = \|T\|\overline{SK(B(X))} \\ &= \|T\|\overline{S(K(B(X)))} \subseteq \|T\|\overline{S(\overline{K(B(X))})} = \|T\|S(\overline{K(B(X))})\end{aligned}$$

is a continuous image of a compact set and hence compact. ∎

Example. (i) Given Banach spaces X, Y, let $\mathcal{F}(X,Y) = \{F \in \mathcal{B}(X,Y) : \operatorname{rank} F < \infty\}$. Note that $\operatorname{rank} F = \dim(\operatorname{im} F)$. If $F \in \mathcal{F}(X,Y)$, then $F(B(X)) \subseteq \|F\| B(\operatorname{im} F)$ is subset of a compact set, so F is compact. Thus $\overline{\mathcal{F}(X,Y)} \subseteq \mathcal{K}(X,Y)$.

(ii) Also W, Z Banach spaces, $S \in \mathcal{B}(Y,Z)$, $T \in \mathcal{B}(W,Z)$, then $SFT \in \mathcal{F}(W,Z)$ for $F \in \mathcal{F}(X,Y)$.

(iii) Let $I = [0,1]$ and $k \in C(I^2)$. Then for $f \in C(I)$, $x \in I$, define

$$Kf(x) = \int_0^1 k(x,y)f(y)dy$$

This defines a compact operator in $\mathcal{K}(C(I), C(I))$.
To see this, let

$$\mathcal{A} = \operatorname{span}\{(x,y) \mapsto \phi(x)\psi(y), \phi, \psi \in C(I)\} \subseteq C(I^2).$$

Then $\mathcal{A}$ is an algebra of functions, $1 \in \mathcal{A}$, $\mathcal{A}$ is point separating, and if $H \in \mathcal{A}$, then $\overline{h} \in \mathcal{A}$. Thus by Stone-Weierstrass, the uniform closure $\overline{\mathcal{A}} = C(I^2)$. Hence, $k = \lim k_n$ where each $k_n(x,y) = \sum_{j=1}^{m_n} \phi_{n_j}(x)\psi_{n_j}(y)$. Let

$$\begin{aligned} K_n f(x) &= \int_0^1 k_n(x,y)f(y)dy = \int_0^1 \sum_{j=1}^{m_n} \phi_{n_j}(x)\psi_{n_j}(y)f(y)dy \\ &= \sum_{j=1}^{m_n} \left[\int_0^1 \psi_{n_j}(y)f(y)dy\right]\phi_{n_j}(x) \end{aligned}$$

so that

$$K_n f = \sum_{j=1}^{m_n} \left[\int_0^1 \psi_{n_j} f\right]\phi_{n_j} \in \operatorname{span}\{\phi_{n_1}, \ldots, \phi_{n_{m_n}}\} \subset C(I)$$

so K has finite rank, and furthermore,

$$|K_n f(x)| \le \int_0^1 |k_n(x,y)| \left\|f(y)\right\| \le \int_0^1 \|k_n\|_\infty \|f\|_\infty = \|k_n\|_\infty \|f\|_\infty$$

so that $\|K_n f\|_\infty \le \|k_n\|_\infty \|f\|_\infty$, so K_n is bounded. Thus each $K_n \in \mathcal{F}(C(I))$. Furthermore, for $f \in C(I)$,

$$\|(K - K_n)f\|_\infty \le \|k - k_n\|_\infty \|f\|_\infty$$

so $\|K - K_n\| \le \|k - k_n\|_\infty \xrightarrow{n\to\infty} 0$ so $K \in \overline{\mathcal{F}(C(I))} \subseteq \mathcal{K}(C(I))$.
Exercise: if $1 \le p < \infty$, then K, as above, defines an operator in $\mathcal{K}(L_p(I))$. Notice that $\|Kf\|_p \le \|k\|_\infty \|f\|_p$, $J : C(I) \to L_p(I)$ "identity" is bounded.

8.28 Theorem. *Let X, Y be Banach spaces, $K \in \mathcal{B}(X,Y)$. Then the following are equivalent:*

(i) $K \in \mathcal{K}(X,Y)$

(ii) $K^*|_{B(Y^*)} : B(Y^*) \to X^*)$ *is w^*–norm continuous*

(iii) $K^* \in \mathcal{K}(Y^*, X^*)$

Proof $(i \Rightarrow ii)$ Let $f_0 = w^*\text{–}\lim_{\nu\in N} f_\nu$ in $B(Y^*)$. Given $\epsilon > 0$, let $\{x_1,\dots,x_n\} \subset B(X)$ such that $K(B(X)) \subseteq \bigcup_{j=1}^n (Kx_j + \frac{\epsilon}{3}B(X))$. Hence if $x \in B(X)$, there is x_j such that $\left\|Kx - Kx_j\right\| < \frac{\epsilon}{3}$. Then for each $\nu \in N$, we have

$$\begin{aligned}
|f_\nu(Kx) - f_0(Kx)| &\leq |f_\nu(Kx) - f_\nu(Kx_j)| + |f_\nu(x_j) - f_0(Kx_j)| + |f_0(Kx_j) - f_0(Kx)| \\
&\leq \|f_\nu\| \left\|Kx - Kx_j\right\| + |f_\nu(Kx_j) - f_0(Kx_j)| + \|f_0\| \left\|Kx_j - Kx\right\| \\
&< \frac{2\epsilon}{3} + |f_\nu(Kx_j) - f_0(Kx_0)|
\end{aligned}$$

so

$$\|K^* f_\nu - K^* f_0\| = \sup_{x\in B(X)} |f_\nu(Kx) - f_0(Kx)| \leq \frac{2\epsilon}{3} + \max_{j=1,\dots,n} |f_\nu(Kx_j) - f_0(Kx_j)|$$

and hence there is $\nu_0 \in N$ so $\nu \geq \nu_0$ implies that $|f_\nu(Kx_j) - f_0(Kx_j)| < \epsilon/3$ for $j = 1,\dots,n$, so $\nu \geq \nu_0$ implies that $\|K^* f_\nu - K^* f_0\| \leq \epsilon$.

$(ii \Rightarrow iii)$ We have that $B(Y^*)$ is w^*–compact, so $K^*(B(Y^*))$ is norm compact.

$(iii \Rightarrow i)$ By the proof above, K^{**} is compact. Hence $\overline{K(\widehat{B(X)})} = K^{**}(B(\hat{X})) \subseteq K^{**}(B(X^{**}))$ is compact, and $\overline{K(B(X))} \cong \overline{K(\widehat{B(X)})}$ is compact. ■

8.29 Corollary. *If X and Y are reflexive, $K \in \mathcal{B}(X,Y)$, then the following are equivalent:*

(i) $K \in K(X,Y)$

(ii) $K(B(X))$ *is compact*

(iii) $K|_{B(X)} : B(X) \to Y$ *is w–norm continuous.*

Proof In a reflexive space, $w = w^*$ and $K = K^{**}$. ■

Remark. Let X, Y be reflexive, $K \in \mathcal{B}(X,Y)$. Then K is w–norm continuous if and only if $K \in \mathcal{F}(X,Y)$.

8.5 Spectral Theory for Compact Operators

8.30 Lemma. *Let $K \in \mathcal{K}(X)$. Suppose there are sequences*

- *of closed subspaces $Y_1 \subsetneq Y_2 \subsetneq \cdots$ of X*
- *scalars $(\alpha_j)_{j=1}^\infty$ such that $(K - \alpha_j I)Y_j \subseteq Y_{j-1}$ for $j = 1,2,3,\dots$.*

Then $\lim_{j\to\infty} \alpha_j = 0$.

Proof Suppose not, i.e. $\limsup_{j\to\infty} |\alpha_j| > 0$. Passing to a subsequence if necessary, we may assume $|\alpha_j| \geq \epsilon > 0$ for all j. By the Riesz lemma, find $x_j \in B(Y_j)$ so $\operatorname{dist}(x_j, Y_{j-1}) > 1/2$. Then $y_j = (K - \alpha_j U)x_j \in Y_{j-1}$, implying that $Kx_j = y_j + \alpha_j x_j \in Y_j$. If $i < j$, we have

$$\left\|Kx_j - Kx_i\right\| = \left\|y_j + \alpha_j x_j - Kx_i\right\| = |\alpha_j| \left\|x_j + \frac{1}{\alpha_j}(y_j - Kx_i)\right\| > \frac{|\alpha_j|}{2} \geq \frac{\epsilon}{2}$$

so that $(Kx_j)_{j=1}^\infty$ admits no Cauchy sequence, a contradiction. ■

Remark. If $T \in \mathcal{B}(X)$, $\{\lambda_1,\dots,\lambda_n\} \subseteq \sigma_p(T)$ distinct. Then $\{\ker(\lambda_i I - T)\}_{i=1}^n$ is a linearly independent sequence of subspace. This is true for $n = 1$, so suppose it holds for $n - 1$. Let

$x_j \in \ker(\lambda_i I - T)\setminus\{0\}$ for each $1 \le j \le n$, and suppose $\alpha_1,\ldots,\alpha_n \in \mathbb{C}$ satisfy $0 = \alpha_1 x_1 + \cdots + \alpha_n x_n$. Suppose $\alpha_{j_0} \neq 0$. Write $x_{j_0} = \sum_{\substack{j=1\\ j\neq j_0}}^n \alpha_j' x_j$ where $\alpha_j' = \alpha_j/\alpha_{j_0}$. Then

$$\sum_{\substack{j=1\\ j\neq j_0}}^n \alpha_j' \lambda_{j_0} x_{j_0} = \lambda_{j_0} x_{j_0} = T x_{j_0} = \sum_{\substack{j=1\\ j\neq j_0}}^n \alpha_j' \lambda_j x_j$$

so that $0 \sum_{\substack{j=1\\ j\neq j_0}}^n \alpha_j'(\lambda_j - \lambda_{j_0}) x_j$, a contradiction.

8.31 Theorem. (Spectral Theorem for Compact Operators) *Let X be an infinite dimensional $\mathbb{C}$−Banach space and $K \in \mathcal{K}(X)$. Then*

(i) *$\lambda \in \mathbb{C}\setminus\{0\}$, then each "generalized eigenspace" $\ker[(\lambda I - K)^n]$ is finite dimensional. Furthermore, there is $n_\lambda \in \mathbb{N}$ such that $\ker[(\lambda I - K)^n] = \ker[(\lambda I - K)^{n_\lambda}]$ for $n \ge n_\lambda$.*

(ii) *$\sigma(K) = \sigma_p(K) \cup \{0\}$, where $\sigma_p(K) = \{\lambda_1, \lambda_2, \ldots\}$ with $\lim_{n\to\infty} \lambda_n = 0$ if infinite.*

Proof (i) We have $\ker(\lambda I - K) \subseteq \ker[(\lambda I - k)^2] \subseteq \cdots$. If n_λ as above does not exist, let $n_1 = 1$ and $n_{k+1} = \min\{n : \ker[(\lambda I - K)^{n_k}]\} \subsetneq \ker[(\lambda I - K)^n]$, and set $Y_0 = \{0\}$ and $Y_k = \ker[(\lambda I - K)^{n_k}]$. Then $Y_0 \subsetneq T_1 \subsetneq \cdots$ and $(\lambda I - K)Y_k \subseteq Y_{k-1}$. But the preceding lemma implies that $\lim_{n\to\infty} \lambda = 0$, a contradiction.

Now suppose $\dim \ker[(\lambda I - K)^n] = \infty$ for some $n = n_\lambda$. Then let $k \in \mathbb{N}$ be so Y_k is infinite dimensional, while Y_{k-1} is finite dimensional. Thus we can find a linearly independent $\{x_j\}_{j=1}^\infty \subset Y_k \setminus Y_{k-1}$. We let $V_0 = Y_{k-1}$ and $V_n = \overline{\operatorname{span}}\{V_0, x_1, \ldots, x_n\}$ for each $N \in \mathbb{N}$. Then $V_0 \subsetneq V_1 \subsetneq \cdots$ and $(\lambda I - K)V_n \subseteq (\lambda I - K)Y_k \subseteq Y_{k-1} = V_0 \subseteq V_{n-1}$ for $N \in \mathbb{N}$. Again by the preceding lemma, we have our contradiction.

(ii) We first show that $\sigma_{ap}(K) \subseteq \sigma_p(K) \cup \{0\}$. If $\lambda \in \sigma_{ap}(K)\setminus\{0\}$, then there is $(x_n)_{n=1}^\infty \subset S(X)$ such that $(\lambda I - K)x_n \xrightarrow{n\to\infty} 0$. Since K is compact, passing to a subseuqnece if necessary, we may assume $y = \lim K x_n$ exists. Then

$$\left\|\lambda x_n - y\right\| \le \|(\lambda I - K)x_n\| + \left\|K x_n - y\right\| \xrightarrow{n\to\infty} 0$$

where each $\|\lambda x_n\| = |\lambda| \neq 0$ so $y \neq 0$. Furthermore, $Ky = \lim_{n\to\infty} K(\lambda x_n) = \lambda \lim_{n\to\infty} K x_n = \lambda y$, so $\lambda \in \sigma_p(K)$.

Now suppose that $\{\lambda_n\}_{n=1}^\infty \subseteq \partial\sigma(K) \setminus \{0\} \subseteq \sigma_{ap}(K) \setminus \{0\} \subseteq \sigma_p(K)$. Let $Y_0 = \{0\}$, $Y_n = \sum_{j=1}^n \ker(\lambda_J I - K)$. By linear independence of eigenspaces, $Y_0 \subsetneq Y_1 \subsetneq \cdots$ and $(\lambda_n I - K)Y_n \subseteq Y_{n-1}$ for $n \in \mathbb{N}$. Then the previous lemma shows that $\lim_{n\to\infty} \lambda_n = 0$. Hence for each $n \in \mathbb{N}$¡ we have that $\partial\sigma(K) \setminus \frac{1}{n}\mathbb{D}$ is finite, and hence $\partial\sigma(K) \subseteq \{0\} \cup \bigcup_{n=1}^\infty (\partial\sigma(K) \setminus \frac{1}{n}\mathbb{D})$ is countable. Thus $\partial\sigma(K)$ is countable, so $\sigma(K)\setminus\{0\}$ is countable. Moreover, note that $0 \in \sigma(K)$, for if not, then $K \in \mathcal{G}(X)$ and $I = KK^{-1} \in \mathcal{K}(X)$, contradicting infinite dimensionality of X. ∎

Example. We may have $\sigma(K) = \{0\}$, $\sigma_p(K) = \emptyset$. Let $K : \ell_p \to \ell_p$ be given by $K(x_1, x_2, \ldots) = (0, x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \ldots)$. It is easy to see that $K \in \mathcal{K}(\ell_p)$ with $\|K\| = 1$. A straightforward induction shows $K^n x = (0, \ldots, 0, \frac{1}{n!}x_1, \frac{1}{(n+1)!}x_2, \ldots)$ so $\|K^n\| = \frac{1}{n!}$. The spectral radius formula shows that $r(K) = \lim_{n\to\infty} \frac{1}{(n!)^{1/n}} = 0$, so $\sigma(K) = \{0\}$. However, as for the unilateral shift, $\ker K = \{0\}$ but $0 \notin \sigma_p(K)$, so $\sigma_p(K) = \emptyset$.

Let $\mathcal{H}$ be a complex Hilbert space. Recall that $\mathcal{H}^* = \{f_y : y \in \mathcal{H}\}$ where $f_y(x) = (x,y)$. If $T \in \mathcal{B}(X)$, implicitly define T^*y as $f_{T^*y} = T^*f_y$. Note that $(\alpha T)^* = \overline{\alpha}T^*$ and $(Tx,y) = f_y(Tx) = T^*f_y(x) = f_{T^*y}(x) = (x, T^*y)$.

Definition. Let $H \in \mathcal{B}(\mathcal{H})$. Then H is called **Hermitian** or **self-adjoint** if $H = H^*$. Let $N \in \mathcal{B}(H)$ then N is **normal** if $N^*N = NN^*$.

Certainly Hermitian operators are normal.

8.32 Proposition. *(i) If $T \in \mathcal{B}(\mathcal{H})$, there are unique Hermitian operators* $\operatorname{Re} T$, $\operatorname{Im} T$, *such that* $T = \operatorname{Re} T + i \operatorname{Im} T$.
(ii) T is normal if and only if $(\operatorname{Re} T)(\operatorname{Im} T) = (\operatorname{Im} T)(\operatorname{Re} T)$.

Proof (i) Let $\operatorname{Re} T = \frac{1}{2}(T + T^*)$ and $\operatorname{Im} T = \frac{1}{2i}(T - T^*)$. If we had $T = H_1 + iH_2$, H_1, H_2 each Hermitian, then

$$H_1 + iH_2 = T = \operatorname{Re} T + i \operatorname{Im} T \Longrightarrow H_2 - \operatorname{Re} T = i(\operatorname{Im} T - H_2) := S$$

where $S^* = S = -S^*$ so that $S^* = 0$ and $S = 0$.

(ii) We have

$$TT^* = (\operatorname{Re} T + i \operatorname{Im} T)(\operatorname{Re} T - i \operatorname{Im} T) = (\operatorname{Re} T)^2 + i(\operatorname{Im} T \operatorname{Re} T + \operatorname{Re} T \operatorname{Im} T) + (\operatorname{Im} T)^2$$

and

$$T^*T = (\operatorname{Re} T - i \operatorname{Im} T)(\operatorname{Re} T + i \operatorname{Im} T) = (\operatorname{Re} T)^2 - i(\operatorname{Im} T \operatorname{Re} T - \operatorname{Re} T \operatorname{Im} T) + (\operatorname{Im} T)^2$$

so that $TT^* = T^*T$ if and only if $2 \operatorname{Im} T \operatorname{Re} T = 2 \operatorname{Re} T \operatorname{Im} T$. ■

8.33 Proposition. (C^*–identity) *If $T \in \mathcal{B}(\mathcal{H})$, then $\|T^*T\| = \|T\|^2$.*

Proof We have

$$\begin{aligned} \|T\|^2 &= \sup_{x \in B(\mathcal{H})} \|Tx\|^2 = \sup_{x \in B(\mathcal{H})} (Tx, Tx) \\ &= \sup_{x \in B(\mathcal{H})} (T^*Tx, x) \le \sup_{x \in B(\mathcal{H})} \|T^*Tx\| = \|T^*T\| \end{aligned}$$

while $\|T^*T\| \le \|T^*\|\|T\| = \|T\|^2$. ■

8.34 Proposition. *If $N \in \mathcal{B}(\mathcal{H})$ is normal, then $r(N) = \|N\|$.*

Proof Suppose first that $H = H^*$ in $\mathcal{B}(\mathcal{H})$. The C^*– identity gives $\left\|H^2\right\| = \|H^*H\| = \|H\|^2$ and hence inductively for $k \in \mathbb{N}$ $\left\|H^{2k}\right\| = \|H\|^{2k}$ since H^k is Hermitian for any k. Now since $N^*N = NN^*$, we see that $(N^*N)^n = (N^*)^n N^n$ for $n \in \mathbb{N}$. Hence if $k \in \mathbb{N}$, we use the C^*–identity and the fact that N^*N is Hermitian to get

$$\|N\|^{2^{k+1}} = \|N^*N\|^{2^k} = \left\|(N^*N)^{2^k}\right\| = \left\|(N^{2^k})^* N^{2^k}\right\| = \left\|N^{2^k}\right\|^2$$

and hence inductively $\|N\|^{2^k} = \left\|N^{2^k}\right\|$. Now, by the spectral radius formula,

$$r(N) = \lim_{n \to \infty} \|N^n\|^{1/n} = \lim_{k \to \infty} \left\|N^{2^k}\right\|^{1/2^k} = \|N\|. \qquad ■$$

Remark. If $T \in \mathcal{B}(X)$, then $\sigma(T) = \sigma_{ap}(T) \cup \sigma_{com}(T)$. Indeed, if $\lambda \in \sigma(T)$, then $\lambda I - T \notin \mathcal{G}(X)$, and hence at least one of the following holds:
- $\lambda I - T$ is not bounded below, i.e. $\lambda \in \sigma_{ap}(T)$
- $im(\lambda I - T)$ is not dense, i.e. $\lambda \in \sigma_{com}(T)$.

8.35 Proposition. *If $H = H^*$ in $\mathcal{B}(\mathcal{H})$, then $\sigma(H) = \sigma_{ap}(H) \subseteq \mathbb{R}$.*

Proof If $\lambda \in \sigma_{ap}(H)$, then there is $(x_n)_{n=1}^{\infty} \subset S(\mathcal{H})$ such that $(\lambda I - H)x_n \to 0$. Then

$$\lambda - (Hx_n, x_n) = \lambda(x_n, x_n) - (Hx_n, x_n) = ((\lambda I - H)x_n, x_n) \to 0$$

where $(Hx_n, x_n) = \overline{(x_n, hx_n)} = \overline{(Hx_n, x_n)} \in \mathbb{R}$ for each n, so $\lambda \in \mathbb{R}$. Now $\sigma_{com}(H) = \sigma_p(H^*) = \sigma_p(H) \subseteq \sigma_{ap}(H)$, so that $\sigma(H) = \sigma_{ap}(H) \cup \sigma(H) = \sigma_{ap}(H)$ as required. ■

Remark. (i) $(Hx, x) \in \mathbb{R}$ if $H = H^*$, $x \in \mathcal{H}$. If we define $[\cdot,\cdot] : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ by $[x, y] = (Hx, y)$, then $[\cdot,\cdot]$ is Hermitian by the Polarization identity.
(ii) If, further, $(Hx, x) \geq 0$ for any $x \in \mathcal{H}$, then $\sigma(H) \subseteq [0, \infty)$. We call such H **positive**
(iii) If $T \in \mathcal{B}(\mathcal{H})$, then $\sigma(T^*T) \subseteq [0, \infty)$ as T^*T is Hermitian, $(T^*Tx, x) = \|Tx\|^2 \geq 0$.

8.36 Theorem. (Continuous Functional Calculus for Bounded Hermitian Operators) *Let $H \in \mathcal{B}(\mathcal{H})$ be Hermitian. Then there is a unique bounded operator $\Phi : C(\sigma(H)) \to \mathcal{B}(\mathcal{H})$ satisfying*
- *$\Phi(1) = I$, $\Phi(\iota) = H$, where ι is the identity function on $\sigma(H)$*
- *$\Phi(fg) = \Phi(f)\Phi(g)$ for f, g in $C(\sigma(H))$.*

Furthermore,
- *Φ is an isometry*
- *$\Phi(f)T = T\Phi(f)$ for any $f \in C(\sigma(H))$, provided that $HT = TH$.*
- *$\Phi(\overline{f}) = \Phi(f)^*$ for $f \in C(\sigma(H))$*

Proof Let $\mathrm{Pol}(\sigma(H))$ denote the set of polynomial functions restricted to $\sigma(H)$. Then we have
(i) $\sigma(p(H)) = p(\sigma(H))$ by the Spectral Mapping Theorem
(ii) $\|p(H)\| = r(p(H))$ since $p(H)$ is normal.

Let $\Phi_0 : \mathrm{Pol}(\sigma(H)) \to \mathcal{B}(\mathcal{H})$ by $\Phi_0(p) = p(H)$. Then Φ_0 is the unique linear operator such that $\Phi_0(1) = I$, $\Phi_0(\iota) = H$, and $\Phi_0(pq) = \Phi_0(p)\Phi_0(q)$ for $p, q \in \mathrm{Pol}(\sigma(H))$. Moreover, by (i) and (ii),

$$\|p\|_\infty = r(p(H)) = \|p(H)\|$$

and $\Phi_0 : \mathrm{Pol}(\sigma(H)) \to \mathcal{B}(\mathcal{H})$ is an isometry.

Now, since $\sigma(H)$ is compact and $\sigma(H) \subseteq \mathbb{R}$, $\sigma(H)$ is conjugacy closed so by Stone-Weierstrass $\overline{\mathrm{Pol}(\sigma(H))} = C(\sigma(H))$, so Φ_0 admits a unique extension $\Phi : C(\sigma(H)) \to \mathcal{B}(\mathcal{H})$. [***TODO: Explain exact statement of Stone-Weierstrass***] Certainly Φ is an isometry; let's verify the non-trivial required properties.

If $f, g \in C(\sigma(H))$, say $f = \lim p_n$, $g = \lim q_n$, then $\|fg - p_n q_n\|_\infty \to 0$, and hence

$$\|\Phi(fg) - \Phi(f)\Phi(g)\| \leq \|\Phi(fg) - \Phi_0(p_n q_n)\| + \|\Phi_0(p_n)\Phi_0(q_n) - \Phi(f)\Phi(g)\|$$

which converges to 0, so $\Phi(fg) = \Phi(f)\Phi(g)$. The remaining properties follow identically.■

8.37 Theorem. (Spectral Theorem for Compact Hermitian Operators) *Let $H = H^*$ in $\mathcal{K}(\mathcal{H})$ and enumerate $\sigma(H) \setminus \{0\} = \{\lambda_1, \lambda_2, \ldots\} \subseteq \mathbb{R}$. Then there exists a sequence $\{P_1, P_2, \ldots\}$ of finite rank orthogonal projections such that*

- *$H = \sum_{n=1}^{\infty} \lambda_n P_n$ converging in norm*
- *If $m \neq n$, $P_n P_m = 0$ so $\operatorname{im} P_n \perp \operatorname{im} P_m$*
- *For $T \in \mathcal{B}(\mathcal{H})$, $TH = HT$ if and only if $TP_n = P_n T$ for each n.*
- *Each $\operatorname{im} P_n$ is finite dimensional.*

Proof We have $\sigma(H) = \{\lambda_1, \lambda_2, \ldots\} \cup \{0\}$ with $\lim_{n\to\infty} \lambda_n = 0$, so for each $n \in \mathbb{N}$, we may define $\phi_n = \chi_{\{\lambda_n\}} \in C(\sigma(H))$ and $\iota = \sum_{n=1}^{\infty} \lambda_n \phi_n$, where the infinite sum converges uniformly. By the continuous functional calculus, let $P_n = \phi_n(H)$, so that

$$H = \iota(H) = \sum_{n=1}^{\infty} \lambda_n \phi_n(H) = \sum_{n=1}^{\infty} \lambda_n P_n \tag{8.1}$$

Also, if $n, m \geq 1$, $\phi_n \phi_m = \phi_n$ if and only if $n = m$, so $P_n^2 = \phi_n^2(H) = P_n$, and $\|P_n\| = \left\|\phi_n\right\|_\infty = 1$ so by A4 P_n is an orthogonal projection. If $n \neq m$, $P_n P_m = \phi_n(H)\phi_(m)H = 0(H) = 0$ so that $\operatorname{im} P_n \perp \operatorname{im} P_m$.

Suppose now that $HT = TH$. Then $P_n T = \phi_n(H)T = T\phi_n(H) = TP_n$ for each n. If $P_n T = TP_n$ for each n, then T commutes with the infinite sum in (8.1) and hence with H.

Finally, multiplicativity of the functional calculus to see that $P_n = \frac{1}{\lambda_n} P_n H$ so $P_n \in \mathcal{K}(\mathcal{H})$. Moreover, $B(\operatorname{im} P_n) = P_n(B(\operatorname{im} P_n)) \subseteq P_n(\mathcal{B}(\mathcal{H})$ is compact, and hence $\operatorname{im} P_n$ is finite dimensional. ■

8.38 Corollary. (Spetral Theorem for Compact Normal Operators) *If $K \in \mathcal{K}(\mathcal{H})$ with $K^*K = KK^*$, then given an enumeration $\{\mu_1, \mu_2, \ldots\}$ of $\sigma(K) \setminus \{0\}$ there is a sequence of finite rank orthogonal projections $\{P_1, P_2, \ldots\}$ such that $K = \sum_{n=1}^{\infty} \mu_n P_n$ and $P_n P_m = 0$ if $N \neq m$.*

Proof Let $H_1 = \operatorname{Re} K$, $H_2 = \operatorname{im} K$, so $H_1 H_2 = H_2 H_1$, each is compact Hermitian, and write

$$H_j = \sum_{n=1}^{\infty} \lambda_{jn} P_{jn}$$

as above. Notive that $P_{1n} P_{2n} = P_{2n} P_{1n}$ for n, m, so each is a norm 1 idempotent and hence an orthogonal projection. Moreover, $\operatorname{im} P_{1n} P_{2m} \subseteq \operatorname{im} P_{1n}$ is finite dimensional. Let $\{P_1, P_2, \ldots\}$ be an enumeration of $\{P_{1n} P_{2m} : m, n \in \mathbb{N} \cup \{0\} \times \mathbb{N} \cup \{0\}\}$ and set $P_j = P_{\ker H_j}$. Then $\{\mu_1, \mu_2, \ldots\}$ is an associated enumeration of $\lambda_{1n} + i\lambda_{2m}$. ■

Definition. Let $\mathcal{H}$ be a Hilbert space. A **partial isometry** is a linear operator $U : \mathcal{H} \to \mathcal{H}$ such that $U|_{(\ker U)^\perp}$ is an isometry. Hence $\|U\| = 1$, or $\|U\| = 0$ if $\ker U = \mathcal{H}$.

8.39 Theorem. (Polar Decomposition) *Let $T \in \mathcal{B}(\mathcal{H})$. Then there is an operator $|T|$ such that $|T|^* = |T|$ and a partial isometry U such that $T = U|T|$.*

Proof We have T^*T is Hermitian with $\sigma(T^*T) \subseteq [0, \infty)$. Thus we may define by the functional calculus $|T| = (T^*T)^{1/2}$ is a bounded operator. Notive that $\sqrt{\cdot} : [0, \infty) \to [0, \infty)$ is $\mathbb{R}$-valued, so $|T|^* = |T|$. Given $x \in \mathcal{H}$,

$$\||T|x\|^2 = (|T|x, |T|x) = \left(|T|^2 x, x\right) = (T^*Tx.x) = (Tx, Tx) = \|Tx\|^2.$$

Define $U_0 : \operatorname{im}|T| \to \mathcal{H}$ by $U_0|T|x = Tx$. If $|T|x = |T|x'$, then $x - x' \in \ker|T| = \ker T$, so $Tx = Tx'$ so U_0 is well-defined. Similarly, U_0 is linear, and we computed above that U_0 is an isometry. Let $U_1 : \overline{\operatorname{im} T} \to \mathcal{H}$ be the unique continuous extension. We note that $\overline{\operatorname{im}|T|} = (\ker T)^\perp$, and define $U = U_1 P_{\overline{\operatorname{im}|T|}}$. Then $T = U|T|$ by construction of U_0 and $\ker U = \ker T$. ■

8.40 Corollary. *Let $K \in \mathcal{K}(\mathcal{H})$. Then there exists a sequence*

- $s_1(K) > s_2(K) > \cdots$ *with* $\lim_{n\to\infty} s_n(K) = 0$ *if infinite*
- *A sequence $\{P_1, P_2, \ldots\}$ of finite rank orthogonal projections with $P_n P_m = 0$ if $n \neq m$*
- *A partial isometry U on $\mathcal{H}$ with* $\ker U = \ker K$ *such that* $K = \sum_{n=1}^\infty s_1(K) U P_n$

Proof We let $\{\lambda_1, \lambda_2, \ldots\} \subset (0,\infty)$ and $\{P_1, P_2, \ldots\}$ arise from the spectral theorem applied to K^*K, so $K^*K = \sum_{n=1}^\infty \lambda_n P_n$. Then $|K| = (K^*K)^{1/2} = \sum_{n=1}^\infty \sqrt{\lambda_n} P_n$. Let $s_n(K) = \sqrt{\lambda_n}$ and apply the polar decomposition. ■

- Note that $\ker U = \ker(K^*K) \perp \operatorname{im} P_n$ so UP_n is a partial isometry of finite rank.
- $\operatorname{im} UP_n \perp \operatorname{im} UP_m$ for $n \neq m$. To see this, f $x, y \in \mathcal{H}$, then by polarization,

$$\begin{aligned}
(UP_n x, UP_m y) &= \frac{1}{4}\sum_{k=0}^{3} i^3 \Big(U(P_n x + i^k P_m y), U(P_n x + i^k P_m y) \Big) \\
&= \frac{1}{4}\sum_{k=0}^{3} i \left\| U(P_n + i P_m y) \right\| = \frac{1}{4}\sum_{k=0}^{3} \left\| P_n + i P_m y \right\| \\
&= (P_n x, P_m y) = 0
\end{aligned}$$

For example, compute U, $|K|$ for $x \in \ell_2$ and $Kx = (0, x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \ldots)$.