

# Functional Analysis

Alex Rutar\*  
University of Waterloo

Fall 2019<sup>†</sup>

---

\*[arutar@uwaterloo.ca](mailto:arutar@uwaterloo.ca)

<sup>†</sup>Last updated: November 10, 2019



---

# Contents

---

<b>Chapter I</b>	<b>Fundamentals of Functional Analysis</b>	
1	Basic Elements of Functional Analysis . . . . .	1
2	Linear operators and linear functionals . . . . .	5
3	Axiom of Choice and the Hahn-Banach Theorem . . . . .	8
4	Some Applications of Baire Category Theorem . . . . .	14
5	On Compactness of the Unit Ball . . . . .	20
6	More Topology . . . . .	21
7	Nets . . . . .	26
8	Euclidean and Hilbert Spaces . . . . .	33



---

# I. Fundamentals of Functional Analysis

---

## 1 BASIC ELEMENTS OF FUNCTIONAL ANALYSIS

Throughout, we denote by  $\mathbb{F}$  either the field  $\mathbb{R}$  or the field  $\mathbb{C}$ .

### BANACH SPACES

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

- (*non-negative*)  $\|x\| \geq 0$  for any  $x \in X$
- (*non-degenerate*)  $\|x\| = 0$  if and only if  $x = 0$
- (*subadditivity*)  $\|x + y\| \leq \|x\| + \|y\|$  for  $x, y \in X$
- ( *$|\cdot|$ -homogeneity*)  $\|\alpha x\| = |\alpha| \|x\|$  for  $\alpha \in \mathbb{F}$ ,  $x \in X$ .

We call the pair  $(X, \|\cdot\|)$  a **normed vector space**. Furthermore, we say that  $(X, \|\cdot\|)$  is a **Banach space** provided that  $X$  is complete with respect to the metric  $\rho(x, y) = \|x - y\|$ .

*Example.* (i)  $(\mathbb{F}, |\cdot|)$  is a Banach space.

(ii)  $(\mathbb{F}^b, \|\cdot\|_p)$ ,  $x = (x_j)_{j=1}^n$ ,

$$\|x\|_p = \begin{cases} \left( \sum_{j=1}^n |x_j|^p \right)^{1/p} & 1 \leq p < \infty \\ \max_{j=1, \dots, n} |x_j| & p = \infty \end{cases}$$

(iii) Consider the space

$$L_p^{\mathbb{F}} = \left\{ f : [0, 1] \rightarrow \mathbb{F} \mid f \text{ is Lebesgue measurable, } \left( \int_0^1 |f|^p \right)^{1/p} < \infty \right\} \Big/ \sim_{\text{a.e.}}$$

where  $1 \leq p < \infty$ .

(iv)  $L_{\infty}^{\mathbb{F}}[0, 1]$ ,  $\|f\|_{\infty} = \text{ess sup}_{t \in [0, 1]} |f(t)|$ .

(v) Let  $(X, d)$  be a metric space. Then

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

is a Banach space.

Here is a more interesting example:

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

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

We note that for  $f : X \rightarrow \mathbb{F}$  that

$$f \in \text{Lip}^{\mathbb{F}}(X, d) \Leftrightarrow \text{there is } L \geq 0 \text{ s.t. } |f(x) - f(y)| \leq Ld(x, y) \text{ for all } x, y \in X \quad (1.1)$$

It is easy to verify that  $L(f) = \min\{L \geq 0 : (1.1) \text{ holds for } f\}$ . It is an easy exercise to see that  $\text{Lip}^{\mathbb{F}}$  is a vector space, and that  $L : \text{Lip}^{\mathbb{F}}(X, d) \rightarrow \mathbb{R}$  is a **semi-norm** (non-negative, subadditive,  $|\cdot|$ -homogeneous). However, we do not have non-degeneracy (for example, constants are taken to 0). We define the Lipschitz norm

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

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

**PROOF** Let  $(f_n)_{n=1}^{\infty}$  be a Cauchy sequence in  $(\text{Lip}^{\mathbb{F}}(X, d), \|\cdot\|_{\text{Lip}})$ . Since  $\|\cdot\|_{\infty} \leq \|\cdot\|_{\text{Lip}}$  on  $\text{Lip}^{\mathbb{F}}(X, d)$ , we see that  $(f_n)_{n=1}^{\infty}$  is uniformly Cauchy (and bounded), and hence there is  $f = \lim_{n \rightarrow \infty} f_n$  in  $C_b^{\mathbb{F}}(X)$ , where the limit is taken with respect to  $\|\cdot\|_{\infty}$ , since  $(C_b^{\mathbb{F}}(X), \|\cdot\|_{\infty})$  is a Banach space. If  $x, y \in X$ , then

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

Since Cauchy sequences are bounded, we see that  $|f(x) - f(y)| \leq Ld(x, y)$ , where  $L = \sup_{n \in \mathbb{N}} \|f_n\|_{\text{Lip}} < \infty$ . Thus by (1.1),  $f \in \text{Lip}^{\mathbb{F}}(X, d)$ . Exercise: one may verify that  $\|f - f_n\|_{\text{Lip}} \rightarrow 0$ . ■

Another collection of basic examples are given by the sequence spaces. We can define

$$\ell_1^{\mathbb{F}} = \left\{ x = (x_j)_{j=1}^{\infty} \in \mathbb{F}^{\mathbb{N}} \left| \|x\|_1 = \sum_{j=1}^{\infty} |x_j| < \infty \right. \right\}$$

It is easy to see that  $(\ell_1, \|\cdot\|_1)$  is a normed vector space.

For  $1 < p < \infty$ , and write

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

Note that  $0 \in \ell_p$ ,  $\alpha x \in \ell_p$  if  $x \in \ell_p$ . Let  $q = p/(p-1)$  so that  $1/p + 1/q = 1$ . Then  $q$  is called the **conjugate index**. We have

**1.2 Proposition. (Young's Inequality)** If  $a, b \geq 0$  in  $\mathbb{R}$ , then  $ab \leq a^p/p + b^q/q$ , with equality only if  $a^p = b^q$ .

and

**1.3 Proposition. (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

$$\sum_{i=1}^{\infty} |x_i y_i| \leq \|x\|_p \|y\|_q$$

with equality exactly when  $\text{sgn}(x_i y_i) = \text{sgn}(x_k y_k)$  for all  $j, k \in \mathbb{N}$  where  $x_i y_i \neq 0 \neq x_k y_k$ , and  $|x|^p = (|x_j|^p)_{j=1}^\infty$  and  $|y|^q$  are linearly dependent in  $\ell_1$ .

and finally

**1.4 Proposition. (Minkowski's Inequality)** If  $x, y \in \ell_p$ , then  $\|x + y\|_p \leq \|x\|_p + \|y\|_p$  with equality exactly when one of  $x$  or  $y$  is a non-negative scalar combination of the other.

## REVIEW OF TOPOLOGY

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

The metric topology on a metric space  $(X, d)$  is the topology

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

*Example.* (i) Given two metrics  $d, \rho$  on  $X$ , we say that  $d \sim \rho$  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. An example of this are the metrics on  $X = \mathbb{R}$  given by  $d(x, y)$  and  $\rho(x, y) = \frac{|x-y|}{1+|x-y|}$ . Then  $d \sim \rho$  but  $\tau_d = \tau_\rho$ .

(ii) "Sorgenfrey 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 an exercise to verify that  $\tau_{|\cdot|} \subsetneq \sigma$ . We say that  $\sigma$  is **finer** than  $\tau_{|\cdot|}$ .

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

**Definition.** Let  $(X, \tau)$  and  $(Y, \sigma)$  be topological spaces, and  $f : X \rightarrow Y$ . We say that  $f$  is  $(\tau - \sigma)$ -**continuous** at  $x_0$  in  $X$  if,

- given  $V \in \sigma$  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  $X$ .

**SPACE OF BOUNDED CONTINUOUS FUNCTIONS INTO A NORMED SPACE**

Let  $(Y, \|\cdot\|)$  denote a normed space. We let  $\tau_{\|\cdot\|}$  denote the topology given by the metric  $\rho(x, y) = \|x - y\|$ . Let  $(X, \tau)$  denote any topological space. Then we write

$$C_b^Y(X) = \left\{ f : X \rightarrow Y \mid f \text{ is bounded and } \tau - \tau_{\|\cdot\|} \text{ - continuous} \right\}$$

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

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

**PROOF** Let  $(f_n)_{n=1}^\infty$  be a Cauchy sequence in  $(C_b^Y(X), \|\cdot\|_\infty)$ . Then for any  $x \in X$ , we have that  $(f_n(x))_{n=1}^\infty$  is Cauchy in  $(Y, \|\cdot\|)$  since  $\|f_n(x) - f_m(x)\| \leq \|f_n - f_m\|_\infty$ , and hence admits a limit  $f(x)$ . In particular,  $x \mapsto f(x)$  defines a function from  $X$  to  $Y$ . We shall fix  $x_0 \in X$  and show that  $f$  is continuous at  $x_0$ . Given  $\epsilon > 0$ , we let

- $n_1$  be so  $n, m \geq n_1$  so that  $\|f_n - f_m\|_\infty < \epsilon/4$ .
- $n_2$  be so  $n \geq n_2$  so that  $\|f_n(x_0) - f(x_0)\| < \epsilon/4$ .
- $N = \max\{n_1, n_2\}$ .
- $U \in \tau, x_0 \in U$  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  $n_x \geq n_1$  and  $n \geq n_x$ , so that  $\|f_{n_x}(x) - f(x)\| < \epsilon/4$ . We then have

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

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

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

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

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

Notice that if  $\epsilon, n_1$  are as above, and further  $x_0, N$  are as above, we have for  $n \geq n_1$

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

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

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

Let's first note the following general principle: let  $(X, d), (Y, \rho)$  be metric spaces, where  $(X, d)$  is complete. If  $\psi : X \rightarrow Y$  is a  $(d - \rho)$ -isometry, then  $(\psi(X), \rho|_{\psi(X)})$  is a complete metric space.

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

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



With pointwise operations,  $(\ell_\infty, \|\cdot\|_\infty)$  is a normed space. In fact, it is a Banach space. Let us note that

$$f \mapsto (f(t))_{t \in T} : C_b(T, \mathcal{P}(T)) \rightarrow \ell_\infty(T)$$

is a surjective linear isometry, and the result follows.

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

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

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

- (iii)  $c_0 = \{x \in \mathbb{F}^\mathbb{N} \mid \lim_{n \rightarrow \infty} x_n = 0\} \subseteq c \subseteq \ell_\infty$ .

**1.7 Lemma.** *If  $x_0 \in X$  where  $(X, \tau)$  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  $f = \lim_{n \rightarrow \infty} f_n$  with respect to  $\|\cdot\|_\infty$  in  $C_b(X)$ , then  $f(x_0) = \lim_{n \rightarrow \infty} f_n(x_0) = 0$ . Thus  $f \in \mathcal{I}(x_0)$ , and closed subsets of complete spaces are themselves complete. ■

Now,  $f \mapsto (f(n))_{n=1}^\infty : \mathcal{I}(\infty) \rightarrow c_0$  is a (linear) surjective isometry.

- (iv) Consider the Sorgenfity line  $(\mathbb{R}, \sigma)$ : verify that

$$c_b(\mathbb{R}, \sigma) = \left\{ f : \mathbb{R} \rightarrow \mathbb{F} \mid f \text{ is bounded and } \lim_{t \rightarrow t_0^+} f(t) = f(t_0) \text{ for } t \in \mathbb{R} \right\}$$

## 2 LINEAR OPERATORS AND LINEAR FUNCTIONALS

Let  $X, Y$  be vector spaces. We let  $\mathcal{L}(X, Y) = \{S : X \rightarrow Y \mid S \text{ is linear}\}$ ; this is itself a vector space with pointwise operations. Let  $(X, \|\cdot\|)$  be a normed space. We denote

$$D(X) = \{x \in X : \|x\| < 1\}$$

$$S(X) = \{x \in X : \|x\| = 1\}$$

$$B(X) = \{x \in X : \|x\| \leq 1\}$$

(Yes, this notation is confusion. No, I didn't choose it.)

**2.1 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

$$\begin{aligned} \|S\| &= \min\{L > 0 : \|Sx\| \leq L\|x\| \text{ for } x \in X\} \\ &= \sup_{x \in S(X)} \|Sx\| = \sup_{x \in B(X)} \|Sx\| \end{aligned}$$

PROOF ( $i \Rightarrow ii$ ) Obvious

( $ii \Rightarrow iii$ ) Note that

$$Sx_0 + D(Y) = \{Sx_0 + y : y \in D(Y)\} = \{y \in Y : \|Sx_0 - y'\| < 1\}$$

is a neighbourhood of  $Sx_0$ . By the definition of metric continuity, there is  $\delta > 0$  such that

$$x_0 + \delta D(X) = \{x_0 + \delta x : x \in D(X)\} = \{x' \in X : \|x_0 - x'\| < \delta\}$$

such that

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

which implies that  $\delta S(D(X)) \subseteq D(Y)$  and  $S(D(X)) \subseteq D(Y)/\delta$ , in other words that  $\|Sx\| \leq 1/\delta$  for  $x \in D(X)$ .

( $iii \Rightarrow i$ ) If  $x \in X$  and  $\epsilon > 0$ , then

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

Then, letting  $\epsilon \rightarrow 0^+$ , we see that

$$\|Sx\| \leq \|x\| \|S\| = \|S\| \|x\|$$

If  $x, x' \in X$ , then  $\|Sx - Sx'\| \leq \|S\| \|x - x'\|$  is  $S$  is Lipschitz, hence continuous.

To complete the proof, the content of (iii) implies (i) tell us that the Lipschitz constant  $L(S) \leq \|S\|$ . Furthermore, if  $\|x\| = 1$ , the preceding proof gives us that  $\|S\|_{S(X)}$ .

Conversely,

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

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  $\mathcal{B}(X, Y)$ .

**2.2 Theorem.** *If  $X, Y$  are normed spaces, then  $(\mathcal{B}(X, Y), \|\cdot\|)$  is a normed space. Furthermore, if  $Y$  is a Banach spaces, then so to is  $(\mathcal{B}(X, Y), \|\cdot\|)$ .*

PROOF Define

$$\Gamma : \mathcal{B}(X, Y) \rightarrow C_b^Y(B(X))$$

given by  $\Gamma(S) = S|_{B(X)}$ . Then, by definition,  $\Gamma$  is linear, with

$$\|\Gamma(S)\|_\infty = \sup_{x \in B(X)} \|Sx\| = \|S\|$$

Thus  $\|\cdot\|$  is a norm: if  $S, T \in \mathcal{B}(X, Y)$ ,  $\alpha \in \mathbb{F}$ ,

$$\begin{aligned} \|S + T\| &= \|\Gamma(S + T)\|_\infty = \|\Gamma(S) + \Gamma(T)\|_\infty \leq \|\Gamma(S)\|_\infty + \|\Gamma(T)\|_\infty = \|S\| + \|T\| \\ \|\alpha S\| &= \|\Gamma(\alpha S)\|_\infty = |\alpha| \|\Gamma(S)\|_\infty = |\alpha| \|S\|. \end{aligned}$$

Furthermore,  $\Gamma : \mathcal{B}(X, Y) \rightarrow C_b^Y(\mathcal{B}(X))$  is an isometry.

Now suppose that  $Y$  is a Banach space. We will show that  $\Gamma(\mathcal{B}(X, Y))$  is closed in  $C_b^Y(\mathcal{B}(X))$ , and hence  $\mathcal{B}(X, Y) = \Gamma^{-1}(\Gamma(\mathcal{B}(X, Y)))$  is complete. Let  $(S_n)_{n=1}^\infty \subset \mathcal{B}(X, Y)$  be  $\|\cdot\|$ -Cauchy. Then  $(\Gamma(S_n))_{n=1}^\infty$  is  $\|\cdot\|_\infty$ -Cauchy in  $C_b^Y(\mathcal{B}(X))$ , and hence there is  $f \in C_b^Y(\mathcal{B}(X))$  such that  $\lim_{n \rightarrow \infty} \|\Gamma(S_n) - f\|_\infty = 0$ . Then we let  $S : X \rightarrow Y$  be given by

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

If  $x, x' \in X$  and  $\alpha \in \mathbb{F}$  are all such that  $x, x', x + \alpha x' \neq 0$ , then

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

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

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

goes to 0 as  $n$  goes to infinity. ■

**Definition.** Given a vector space  $X$ , let  $X' = \mathcal{L}(X, \mathbb{F})$  denote the **algebraic dual**. If further  $X$  is a normed space, we let  $X^* = \mathcal{B}(X, \mathbb{F})$  denote the (continuous) dual.

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

**2.4 Theorem.** Let for  $x \in \ell_1$ ,  $f_x : c_0 \rightarrow \mathbb{F}$  be given by  $f_x(y) = \sum_{j=1}^\infty x_j y_j$ . Then  $f_x \in c_0^*$  with  $\|f_x\| = \|x\|_1$ . Furthermore, every element of  $c_0^*$  arises as above.

**PROOF** If  $x \in \ell_1$  and  $y \in c_0 \subseteq \ell_\infty$ , then

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

so  $f_x(y) = \sum_{j=1}^\infty x_j y_j$  is well-defined. It is obvious that  $f_x$  is linear:  $f_x(y + \alpha y') = f_x(y) + \alpha f_x(y')$  for  $y, y' \in c_0$  and  $\alpha \in \mathbb{F}$ . Also,  $\|f_x\| \leq \|x\|_1$ . We let  $y^n = (\overline{\text{sgn } x_1}, \dots, \overline{\text{sgn } x_n}, 0, 0, \dots) \in c_0$ , with  $\|y^n\| = 1$ . Then

$$\|f_x\| \geq |f_x(y^n)| = \sum_{j=1}^n x_j \overline{\text{sgn } x_j} = \sum_{j=1}^n |x_j|$$

so that  $\|f_x\| \geq \|x\|_1$ , and hence equality holds.

Now let  $f \in c_0^*$ , and write  $e_n = (0, \dots, 0, 1, 0, 0, \dots) \in c_0$ , and let  $x_n = f(e_n)$ . Then, let  $y \in c_0$  and  $y^n = (y_1, \dots, y_n, 0, 0, \dots)$  and we have

$$\|y - y^n\|_\infty = \sup_{j \geq n+1} |y_j|$$

which goes to 0 as  $n$  goes to infinity. Then since  $f$  is continuous, we have

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

We use sequence  $(y^n)_{n=1}^\infty$  as in  $y^n \in c_0$ , to see that

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

so  $x \in \ell_1$ . Thus  $f = f_x$ , as desired. ■

**2.5 Corollary.**  $\ell_1 \cong c^*$  isometrically isomorphically.

**PROOF** For  $y \in c$ , let  $L(y) = \lim_{n \rightarrow \infty} y_n$ . Given  $y \in c$ , let  $y^n = (y_1, \dots, y_n, L(y), L(y), \dots) \in c$ . Notice that  $\|y - y^n\|_\infty \rightarrow 0$  similarly as above.

We let  $1 = (1, 1, \dots)$ , and  $1_n = (0, \dots, 0, 1, 1, \dots)$ . If  $m < n$ , then  $1_n - 1_m \in c_0$ , so

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

so that  $(f(1_n))_{n=1}^\infty$  is Cauchy in  $\mathbb{F}$ . Let  $x_0 = \lim_{n \rightarrow \infty} f(1_n)$ . Let  $\tilde{x} = (x_0, x_1, \dots) \in \ell_1$ . Then letting  $x_j = f(e_j)$ , we see that

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

Similarly as above, we may show that  $\|f\| = \|\tilde{x}\|_1$ .

*Remark.* We write  $c_0^* \cong \ell_1$  isometrically.

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

### 3 AXIOM OF CHOICE AND THE HAHN-BANACH THEOREM

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

- (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.

**3.1 Theorem.** *The following are equivalent:*

- (i) (Axiom of Choice 1): For any  $x \neq \emptyset$ , there is a function  $\gamma : \mathcal{P}(X) \setminus \{\emptyset\} \rightarrow X$  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  $\lambda$ ,

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

- (iii) (Zorn's Lemma): In a poset  $(S, \leq)$ , 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 Exercise. ■

**Definition.** Let  $X$  be a vector space (over  $k$ ). A subset  $S \subseteq X$  is called

- **linearly independent** if for any distinct  $x_1, \dots, x_n \in S$ , the equation  $0 = \alpha_1 x_1 + \dots + \alpha_n x_n = 0$  where  $\alpha_i \in k$  implies  $\alpha_1 = \dots = \alpha_n = 0$ .
- **spanning** if each  $x \in X$  admits  $x_i \in S$  and  $\alpha_i \in k$  such that  $x = \alpha_1 x_1 + \dots + \alpha_n x_n$ .
- **Hamel basis** if it is both linearly independent and spanning

**3.2 Proposition.** *Any vector space  $X$  admits a Hamel basis.*

PROOF Let  $\mathcal{L} = \{L \subseteq X : L \text{ is linearly independent}\}$ . Then  $(\mathcal{L}, \subseteq)$  is a poset. Verify that for any chain  $\mathcal{C} \subseteq \mathcal{L}$ , that  $U = \bigcup_{L \in \mathcal{C}} L \in \mathcal{L}$  and is an upper bound for  $\mathcal{C}$ . Apply Zorn to find a maximal element  $M$  in  $(\mathcal{L}, \subseteq)$ . Verify that  $M$  is spanning for  $X$ . ■

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

PROOF Our assumption provides  $\{e_n\}_{n=1}^\infty$  which is linearly independent. By normalizing each element, we may and will suppose that each  $\|e_n\| = 1$ . Let

$$\text{span}\{e_n\}_{n=1}^\infty = \left\{ \sum_{j=1}^m \alpha_j e_{n_j} : m \in \mathbb{N}, \alpha_j \in \mathbb{F}, n_1 < \dots < n_m \right\}$$

and let  $B$  be any linearly independent set containing  $\{e_n\}_{n=1}^\infty$ . Define  $f : X = \text{span } B \rightarrow \mathbb{F}$  be given for  $x = \sum_{b \in B \setminus \{e_n\}_{n=1}^\infty} \alpha_b b + \sum_{j=1}^n \alpha_j e_{n_j}$  by  $f(x) = \sum_{j=1}^n \alpha_j n_j$ . The point is that  $f(e_n) = n$  and  $f(e) = 0$  for any other  $e \in B$ . Notice that

$$\|f\| = \sup_{x \in B(X)} |f(x)| \geq \sup_{n \in \mathbb{N}} |f(e_n)| = \sup_{n \in \mathbb{N}} n = \infty$$

■

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

- (non-negative homogeneity)  $\rho(tx) = t\rho(x)$  for  $t \geq 0, x \in X$ .
- (subadditivity)  $\rho(x+y) \leq \rho(x) + \rho(y)$  for  $x, y \in X$ .

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

**PROOF** We first do this for extensions by a single point  $x \in X \setminus Y$ . We wish to find  $c \in \mathbb{R}$  such that

$$f(y) + \alpha c \leq \rho(y + \alpha x)$$

for  $y \in Y$  and  $\alpha \in \mathbb{R}$ . In this case, we let  $F : \text{span } Y \cup \{x\} \rightarrow \mathbb{R}$  be given by  $F(y + \alpha x) = f(y) + \alpha c$ , and we have that  $F$  is linear and satisfies  $F \leq \rho$  on  $\text{span } Y \cup \{x\}$ . To do this, let  $y_+, y_-$  in  $Y$  and observe that  $f(y_+) + f(y_-) = f(y_+ + y_-) \leq \rho(y_+ + y_-) \leq \rho(y_+ + x) + \rho(y_- - x)$  so that  $f(y_-) - \rho(y_- - x) \leq \rho(y_+ + x) - f(y_+)$ . It thus follows that

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

so we may find  $c \in \mathbb{R}$  for which

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

If  $t > 0$ , then for  $y \in Y$ ,

$$c \leq \rho\left(\frac{1}{t}y + x\right) - f\left(\frac{1}{t}y\right) \Rightarrow tc \leq \rho(y + tx) - f(y) \Rightarrow f(y) + tc \leq \rho(y + tx)$$

and if  $s > 0$ , then for  $y \in Y$ ,

$$f\left(\frac{1}{s}y\right) - \rho\left(\frac{1}{s}y - x\right) \leq c \Rightarrow sc \leq f(y) - \rho(y - sx) \Rightarrow f(y) - sc \leq \rho(y - sx)$$

Clearly,  $f(y) + 0 \leq \rho(y + 0x)$ . Hence, we have our desired inequality.

We now use Zorn's lemma to lift this result to the whole space. Consider the set of “ $p$ -extensions” of  $f$ ,

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

Define a partial order on  $\mathcal{E}$  by

$$(\mathcal{M}, \psi) \leq (\mathcal{N}, \phi) \text{ iff } \mathcal{M} \subseteq \mathcal{N}, \phi|_{\mathcal{M}} = \psi$$

Suppose  $\mathcal{C} \subseteq \mathcal{E}$  is a chain with respect to  $\leq$ . We let

- $\mathcal{U} = \bigcup_{(\mathcal{M}, \psi) \in \mathcal{C}} \mathcal{M}$  which is a subspace, since  $\mathcal{C}$  is a chain.
- and define  $\phi : \mathcal{U} \rightarrow \mathbb{R}$  by  $\phi(x) = \psi(x)$  whenever  $x \in \mathcal{M}$ , which is again well-defined since  $\mathcal{C}$  is a chain.

Furthermore, we see that  $\phi \in \mathcal{U}'$ , since if  $x, y \in \mathcal{U}$ , get  $x \in \mathcal{M}, y \in \mathcal{N}$  for some  $(\mathcal{M}, \psi) \leq (\mathcal{N}, \psi') \in \mathcal{C}$ . Then  $\phi(x + y) = \psi'(x + y) = \psi'(x) + \psi'(y) = \phi(x) + \phi(y)$ , etc. Likewise,  $\psi \leq \rho|_{\mathcal{U}}$ . Thus by Zorn's lemma,  $\mathcal{E}$  admits a maximal element  $\mathcal{M}, F$ . Then  $\mathcal{M} = X$ , for if not, then we would find  $x \in X \setminus \mathcal{M}$  and we apply step one to  $\text{span } \mathcal{M} \cup \{x\}$  to get  $F'$ , a strictly larger element violating maximality. ■

Trivially, any  $\mathbb{C}$ -vector space is a  $\mathbb{R}$ -vector space.

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

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

$$f \in X'_{\mathbb{R}} \text{ if and only if } f_{\mathbb{C}} \in X^* = X'_{\mathbb{C}} \text{ with } \|f\| = \|f_{\mathbb{C}}\|$$

**PROOF** (i) and (ii) are straightforward exercises; let's see (iii). We let for  $x \in X$ ,  $z = \operatorname{sgn} f_{\mathbb{C}}(x)$ . Then

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

so we see that  $\|f_{\mathbb{C}}\| \leq \|f\|$ . Conversely,

$$|f(x)| = |\operatorname{Re} f_{\mathbb{C}}(x)| \leq |f_{\mathbb{C}}(x)| \leq \|f_{\mathbb{C}}\| \|x\| \text{ so that } \|f\| \leq \|f_{\mathbb{C}}\| \quad \blacksquare$$

**3.6 Corollary.** *If  $X$  is a normed space,  $Y \subseteq X$  a subspace and  $f \in Y^*$ , then there exists  $F \in X^*$  such that  $F|_Y = f$  and  $\|F\| = \|f\|$ .*

**PROOF** Define  $\rho : X \rightarrow \mathbb{R}$  be given by  $\rho(x) = \|f\| \cdot \|x\|$ , so  $\rho$  is sublinear and  $\operatorname{Re} f \leq \rho|_Y$ . Apply Hahn-banach to this data and get  $\tilde{F} \in X'_{\mathbb{R}}$  such that  $\tilde{F}|_Y = \operatorname{Re} f$  and  $\tilde{F} \leq \rho$ , and let  $F = \tilde{F}_{\mathbb{C}}$ . ■

**3.7 Corollary.** *If  $X$  is a normed space,  $x \in X$ , then there is  $f \in X^*$  such that*

$$\|x\| = f(x) = |f(x)| \text{ and } \|f\| = 1$$

**PROOF** Let  $f_0 : X \rightarrow \mathbb{R}$  be given by  $f_0(\alpha x) = \alpha \|x\|$ . If  $x \neq 0$ , then

$$\|f_0\| = \sup_{\|\alpha x\| \leq 1} |f_0(\alpha x)| = \sup_{\|\alpha x\| \leq 1} |\alpha| \|x\| = 1$$

and apply the previous corollary. If  $x = 0$ , this is trivial. ■

**3.8 Theorem.** *Let  $X$  be a normed space and  $X^{**}$  denote the bidual. For  $x \in X$ , define  $\hat{x} : X^* \rightarrow \mathbb{R}$  by  $\hat{x}(f) = f(x)$ . Then  $\hat{x} \in X^{**}$  with  $\|\hat{x}\| = \|x\|$ , so that  $x \mapsto \hat{x} : X \rightarrow X^{**}$  is a linear isometry.*

**PROOF** Notice that  $|\hat{x}(f)| = |f(x)| \leq \|f\| \|x\|$  so  $\|\hat{x}\| \leq \|x\|$ . The last corollary provides for  $x \in X$  an  $f_x \in X^*$  with  $|f_x(x)| = \|x\|$ . Then  $\|\hat{x}\| \leq |\hat{x}(f_x)| = \|x\|$ . Hence  $\|\hat{x}\| = \|x\|$ . Clearly  $x \mapsto \hat{x}$  is linear. ■

*Remark.* Since  $X^{**}$ , being a dual space, is complete, we have that  $\hat{X} = \{\hat{x} : x \in X\}$  satisfies that its closure  $\overline{\hat{X}} \subseteq X^{**}$  is complete. Hence  $\overline{\hat{X}}$  is a Banach space containing a dense copy of  $X$ . Often, we will simply write  $\hat{X} = \overline{\hat{X}}$  and call it the **completion** of  $X$ .

### GEOMETRIC HAHN-BANACH

If  $A, B \subset X$  with  $A \cap B = \emptyset$  (and 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  $x_0 + \ker \operatorname{Re} f$ .

**3.9 Proposition.** Let  $X$  be a normed space.

(i) If  $f \in X^* \setminus \{0\}$ , then  $\overline{\ker f}$  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** To see (i),  $\ker f = f^{-1}(\{0\})$  is a closed set since  $f$  is continuous. Furthermore, if  $Y \subsetneq X$  is a proper closed subspace, then it is nowhere dense. If not, then there would exist  $y_0 \in Y$  and  $\delta > 0$  such that  $y_0 + \delta D(X) \subseteq Y$ . But then  $D(X) \subseteq \frac{1}{\delta}(Y - y_0) = Y$ , so  $X = \operatorname{span} D(X) \subseteq Y$ , a contradiction.

To see (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

$$0 \notin f(x_0 + \delta D(X)) = f(x_0) + \delta f(D(X)) \implies \frac{1}{\delta}f(x_0) \notin -f(D(X)) = f(D(X)) \quad (3.1)$$

But then  $\|f\| \leq \frac{1}{\delta}f(x_0)$ , for if  $\|f\| > \frac{1}{\delta}f(x_0)$ , there would be  $x \in D(X)$  such that  $|f(x)| > \frac{1}{\delta}|f(x_0)|$ . Thus

$$\left| \frac{f(x_0)}{\delta f(x)} \right| < 1 \implies \frac{f(x_0)}{\delta f(x)} = \frac{1}{\delta}f(x)$$

contradicting the statement in (3.1). ■

**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  $\epsilon(a_0, x) > 0$  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.

**3.10 Lemma. (Minkowski Functional)** Let  $A \subset X$  be a convex set containing 0 and absorbing at 0. Define  $p : X \rightarrow \mathbb{R}$  by  $p(x) = \inf\{t > 0 : x \in tA\}$ . Then  $p$  is a sublinear functional. Moreover, we have that

- (i)  $\{x \in X : p(x) < 1\} \subseteq A \subseteq \{x \in X : p(x) \leq 1\}$ ; and
- (ii) if  $X$  is normed and  $A$  is a neighbourhood of 0, then there is  $N > 0$  such that  $p(x) \leq N\|x\|$  for  $x \in X$ .

**PROOF** First note, for any  $x \in X$ , if  $A$  is absorbing at 0, there is  $s > 0$  such that  $sx \in A$ , so  $x \in \frac{1}{s}A$  and hence  $0 \leq p(x) < \infty$ .

Let's see non-negative homogeneity. Clearly  $p(0) = 0$ . If  $s > 0$  and  $x \in X$ , then

$$p(sx) = \inf\{t > 0 : sx \in tA\} = \inf\left\{t > 0 : x \in \frac{t}{s}A\right\} = s \cdot \inf\left\{\frac{t}{s} > 0 : x \in \frac{t}{s}A\right\} = sp(x)$$



We also have subadditivity. First, note that if  $s, t > 0$  and  $a, b \in A$ , then

$$sa + tb = (s + t) \left( \frac{s}{s+t}a + \frac{t}{s+t}b \right) \in (s + t)A \implies sA + tA \subseteq (s + t)A$$

by convexity, and also  $(s + t)A = \{(s + t)a : a \in A\} \subseteq \{sa + tb : a, b \in A\} = sA + tA$ . Thus  $sA + tA = (s + t)A$ . Now for  $x, y \in X$ , we have

$$\begin{aligned} p(x) + p(y) &= \inf\{s > 0 : x \in sA\} + \inf\{t > 0 : y \in tA\} \\ &= \inf\{s + t : s > 0, t > 0, x \in sA, y \in tA\} \\ &\geq \inf\{s + t : s > 0, t > 0, x + y \in sA + tA = (s + t)A\} \\ &= \inf\{r > 0 : x + y \in rA\} = p(x + y) \end{aligned}$$

so that  $p$  is a sublinear functional. Then

- (i) If  $p(x) < 1$ , then there is  $0 < t < 1$  so  $x \in tA$ ; i.e.  $\frac{1}{t}x \in A$  and  $x = (1 - t)x + t\frac{1}{t}x \in A$ . The second inclusion is obvious.
- (ii) The assumptions provide  $\delta > 0$  so  $\delta D(X) \subseteq A$ . Then for  $x \in X$  and  $\epsilon > 0$ ,

$$x \in (\|x\| + \epsilon)D(X) = \frac{\|x\| + \epsilon}{\delta} \delta D(X) \subseteq \frac{\|x\| + \epsilon}{\delta} A$$

so  $p(x) \leq \frac{\|x\| + \epsilon}{\delta}$  so  $p(x) \leq \frac{1}{\delta} \|x\|$ ; the result follows with  $N = 1/\delta$ . ■

**3.11 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  $a_0$ , we have  $f \in X^*$ ; and
- if  $A$  is absorbing around each of its points (for example if  $A$  is open), then we have  $\operatorname{Re} f(a) > \alpha \geq \operatorname{Re} f(b)$ .

**PROOF** We first re-centre at 0. Let  $A - B = \{a - b : a \in A, b \in B\}$ . Then it is easy to verify that

- (i)  $A - B$  is absorbing at any  $a_0 - b, b \in B$
- (ii)  $A - B$  is convex
- (iii) 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 (resp. open), then  $A - B$  is absorbing around any of its points (resp. open).

Let  $x_0 = a_0 - b_0$  for some  $b_0 \in B$ , and set  $C = x_0 - (A - B)$ , so we have  $0 = x_0 - x_0 \in C$ . 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 (resp.  $A$  is open), then  $C$  is absorbing at each of its points (resp. open).

Let  $p$  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 the lemma,  $p(x_0) > 1$ .

Let us find  $f$  and  $\alpha$ . Let  $f_0 : \mathbb{R}x_0 \rightarrow \mathbb{R}$ , by  $f_0(sx) = sp(x_0)$ . Hence  $f_0$  is linear and  $f_0 \leq p|_{\mathbb{R}x_0}$ , so by Hahn-Banach, get  $f \in X'_\mathbb{R}$  such that  $f \leq p$  on  $X$ . If  $a \in A$  and  $b \in B$ , then

$x_0 - (a - b) \in C$ , so by (i) of the lemma, since  $p(x_0) \geq 1$ , we have  $f(x_0 - (a - b)) \leq p(x_0 - (a - b)) \leq 1$ . Thus  $f(x_0) + f(b) \leq 1 + f(a)$  so in fact  $f(b) \leq f(a)$ . 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, we shall replace  $f$  by  $f_{\mathbb{C}}$

For the remainder of the proof, we suppose  $X$  is a normed space, and  $A$  is a neighbourhood of  $a_0$ . Then part (ii) of the lemma provides  $N > 0$  so that  $p(x) \leq N \|x\|$ . Then for  $x \in X$ ,  $f(x) \leq p(x) \leq N \|x\|$  and  $-f(x) = p(-x) \leq N \|-x\| = N \|x\|$  so  $|f(x)| \leq N \|x\|$ , in other words that  $\|f\| \leq N$  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  $\alpha \leq f(a - tx_0) = f(a) - tf(x_0) < \alpha$ , a contradiction. ■

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

$$\text{conv}(S) = \left\{ \sum_{j=1}^n \lambda_j x_j : n \in \mathbb{N}, x_1, \dots, x_n \in S \text{ and } \lambda_1, \dots, \lambda_n \geq 0 \text{ with } \sum_{j=1}^n \lambda_j = 1 \right\}$$

One can verify that  $\text{conv}(S)$  is in fact convex, and is the smallest convex set containing  $S$ , i.e.

$$\text{conv}(S) = \bigcap \{C : S \subseteq C \subseteq X, C \text{ convex}\}$$

If  $X$  is normed, we let  $\overline{\text{conv}}(S)$  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  $H = \{x \in X : \text{Re } f(x) \leq \alpha\}$  for some  $f \in X'$ ,  $\alpha \in \mathbb{R}$ .

If  $X$  is normed, then the last proposition shows  $H$  is closed if and only if  $f$  is bounded.

**3.12 Theorem.** If  $X$  is a normed vector space and  $\emptyset \neq S \subset X$ , then  $\overline{\text{conv}}(S) = \bigcap \{H : S \subseteq H \subset X, H \text{ a closed half space}\}$ .

**PROOF** It is immediate that  $\overline{\text{conv}}(S) \subseteq \bigcap \{H : S \subseteq H \subset X, H \text{ a closed half space}\}$ . Thus suppose  $x_0 \notin \overline{\text{conv}}(S)$ . Then there is  $\delta > 0$  such that  $(x_0 + \delta D(X)) \cap \overline{\text{conv}}(S) = \emptyset$ . Since  $x_0 + \delta D(X)$  is open and convex, hyperplane separation gives provides  $f \in X^*$  and  $\alpha \in \mathbb{R}$  so  $\text{Re } f(a) > \alpha \geq \text{Re } f(b)$  for  $a \in x_0 + \delta D(X)$  and  $b \in \overline{\text{conv}}(S)$ . Then  $S \subset H = \{y \in X : \text{Re } f(y) \leq \alpha\}$  but  $x_0 \notin H$ . ■

## 4 SOME APPLICATIONS OF BAIRE CATEGORY THEOREM

**4.1 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-trivial 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 nowhere dense; and a set is **non-meagre** (2nd category) otherwise.

**4.2 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  $\blacksquare$

**4.3 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  $x \in U$ ,  $\sup\{\|Tx\| : T \in \mathcal{F}\} < \infty$  (pointwise bounded). Then  $\mathcal{F}$  is uniformly bounded, i.e.  $\sup\{\|T\| : T \in \mathcal{F}\} < \infty$ .

**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  $F_{n_0}$  admits 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  $T$ .  $\blacksquare$

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

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

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

By construction,  $y_{n+1} - y_n \in \delta^n A$ , so  $\|y_{n+1} - y_n\| \leq \delta^n r$  and there is  $x_n \in \delta^n D(X)$  such that  $y_{n+1} - y_n = Tx_n$ . 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  $\|y_n - y\| \leq \delta^n r \rightarrow 0$ . Since  $X$  is complete, we let  $x = \sum_{n=0}^{\infty} x_n$ , and by construction

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

Then by linearity and continuity of  $T$ , we have

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

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

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

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

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,  $\frac{\epsilon}{n} D(Y) \subseteq T(D(X))$ .  $\blacksquare$

**4.6 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.  $\blacksquare$

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  $X \oplus_1 Y$ .

**4.7 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  $T \in \mathcal{B}(X, Y)$ . If  $(x_n) \rightarrow x$  in  $X$ , then  $Tx_n \rightarrow Tx$  in  $Y$ . Thus if  $(x, y) \in \overline{\Gamma(T)}$ , then  $(x, y) = \lim(x_n, Tx_n)$  where  $(x_n, Tx_n) \in \Gamma(T)$ . But then

$$\|y - Tx\| \leq \|y - Tx_n\| + \|Tx_n - Tx\| \leq \|x - x_n\| + \|y - Tx_n\| + \|Tx_n - Tx\| = \|(x - y) - (x_n, Tx_n)\|_1$$

so in fact  $y = Tx$  so  $(x, y) = (x, Tx)$ .

Conversely, if  $\Gamma(T)$  is closed in  $X \oplus_1 Y$ , then  $\Gamma(T)$  is a Banach space. Define  $S : \Gamma(T) \rightarrow X$  by  $S(x, Tx) = x$ . Notice that  $S$  is linear, and

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

so  $\|S\| \leq 1$ , so  $S$  is bounded. It is also clear that  $S$  is bijective, with  $S^{-1} : X \rightarrow \Gamma(T)$  given by  $S^{-1}(x) = (x, Tx)$ . Thus the inverse mapping theorem gives that  $S^{-1}$  is also bounded. Hence for any  $x \in X$ ,

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

so that  $T$  is in fact bounded. ■

**4.8 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  $x_n \rightarrow 0$  for which we may assume that  $Tx_n$  converges in  $Y$ , say  $y = \lim Tx_n$ , then  $y = 0$  too.*

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

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

- (ii) Take  $X = C_b(0, 1)$ ,  $X_0 = \{f \in X : f \text{ is differentiable on } (0, 1), f' \in C_b(0, 1)\}$ . We have  $X_0 \subseteq X$ , and we put the uniform norm  $\|\cdot\|_{\infty}$  on both spaces. We let  $D : X_0 \rightarrow X$ ,  $Df = f'$ . 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  $(f_n, f'_n) \rightarrow (0, g)$  in  $X_0 \oplus_1 X$ . Notice that  $\|f'_n\|_{\infty} < \infty$ , so  $f_n$  is Lipschitz on  $(0, 1)$ , so  $f_n$  is uniformly continuous on  $(0, 1)$ , so  $f_n(0^+) = \lim_{t \rightarrow 0^+} f(t)$  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,

- $f_n \rightarrow 0$  uniformly, so  $f_n(0^+) \rightarrow 0$

- $f'_n \rightarrow g$  uniformly, so for each  $t \in (0, 1)$ ,

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

and again, by the FT of  $C$ ,  $g(t) = 0$ . Thus  $g = 0$ , so  $\Gamma(D)$  is closed. We say that  $D : X_0 \rightarrow X$  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  $J : X \rightarrow X_0$  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  $g \in X$ ,

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

so  $\|J\| \leq 1$ . 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.

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  $Z = \text{span } B'$ .

**4.9 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  $\text{im } P = P(Y) = X$ .*

*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  $y = Py + (I - P)y$  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  $Pz = P^2y = Py = z$ , but  $z \in \ker P$ , so  $z = Pz = 0$ .

( $\Rightarrow$ ) Let  $S : X \oplus Z \rightarrow Y$  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,

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

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  $P_1 : X \oplus Z \rightarrow X$  be given by  $P_1(x, z) = x$ ; and  $J : X \rightarrow Y$  by  $Jx = x$ . Notice that  $\|P_1\| = 1$  and  $\|J\| = 1$ . Define  $P : Y \rightarrow Y$  by  $Py = JP_1S^{-1}y$ . Then

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

**4.10 Theorem.**  $c_0$  is not boundedly complemented in  $\ell_\infty$ .

**PROOF** Let us assume otherwise; hence, there is  $P = P^2 \in \mathcal{B}(\ell_\infty)$  such that  $\text{im } P = c_0$ . 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

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

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

$$\left\| \sum_{i=1}^m \alpha_i y_{F_i} \right\| = \|(I - P)z\| \leq \|I - P\| \|z\| = \|I - P\| \max_{k=1, \dots, m} |\alpha_k| \quad (4.1)$$

Now, let for  $n, k \in \mathbb{N}$ ,  $\mathcal{F}_{n,k} = \{F \in \mathcal{F} : |\delta_k(y_F)| \geq \frac{1}{n}\}$  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  $\alpha_i = \text{sgn } \delta_k(y_{F_i})$ . Then we have each  $|\alpha_i| = 1$ , so by (4.1), 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 follows that  $\mathcal{F}_{n,k}$  is finite. Since each  $y_F \neq 0$  for  $F \in \mathcal{F}$ , we see that  $\mathcal{F} = \bigcup_{n=1}^\infty \bigcup_{k=1}^\infty \mathcal{F}_{n,k}$ , which contradicts that  $|\mathcal{F}| = \mathfrak{c}$ . Hence such a  $P$  must not exist. ■

**4.11 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  $x = \sum_{k=1}^n x_k e_k$ ,  $x_k \in \mathbb{F}$ ,  $\|x\|_\infty = \max_{k=1, \dots, n} |x_k|$ . 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. ■

**4.12 Corollary.** *Let  $(X, \|\cdot\|)$  be a finite dimensional normed space. Then*

- (i)  $K \subseteq X$  is compact if and 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  $B_\infty$ , which is compact.  
(ii) Cauchy sequences are bounded, and thus contained in some scaled copy of  $B_\infty$ , 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. ■

**4.13 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  $(e_1, \dots, e_n)$  for  $Y$ . We may assume that each  $\|e_k\| = 1$ . We define  $f_1, \dots, f_n \in Y' = Y^*$  by

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

By Hahn-Banach, get  $F_1, \dots, F_n \in X^*$  such that  $F_k|_Y = f_k$  and  $\|F_k\| = \|f_k\|$ . Define  $P : X \rightarrow X$  by  $Px = \sum_{k=1}^n F_k(x)e_k$ . Notice that  $\text{im } P \subseteq Y$  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,  $Y = \ker(I - P)$ . ■

## 5 ON COMPACTNESS OF THE UNIT BALL

**5.1 Lemma.** *Let  $X$  be a normed space and  $Y \subsetneq X$  a closed subspace. Then given  $\epsilon \in (0, 1)$  there is  $x_0 \in D(X) \subseteq B(X)$  such that  $d(x_0, Y) > 1 - \epsilon$ .*

**PROOF** Let  $x \in X \setminus Y$  and let  $f : Y + \mathbb{F}x \rightarrow \mathbb{F}$  be given by  $f(y + \alpha x) = \alpha$ ,  $y \in Y$ ,  $\alpha \in \mathbb{F}$ . Then  $f$  is linear and  $\ker f = Y$  is closed,  $Y \subsetneq Y + \mathbb{F}x$ , so  $f$  is bounded. Let  $F \in X^*$  be any Hahn-Banach extension of  $f$  with  $\|F\| = \|f\|$ .

Now, we find  $x_0 \in D(X)$  such that  $|F(x_0)| > (1 - \epsilon)\|F\|$ . Since  $Y \subseteq \ker F$ , we have for  $y \in Y$  that  $\|F\| \|x_0 - y\| \geq |f(x_0 - y)| = |F(x_0)| > (1 - \epsilon)\|F\|$ , so  $\|x_0 - y\| > 1 - \epsilon$ . Hence  $d(x_0, Y) = \inf_{y \in Y} \|x_0 - y\| \geq 1 - \epsilon$ . ■

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

**PROOF** The reverse implication is standard. Thus suppose  $X$  is not finite dimensional. Let  $\epsilon \in (0, 1)$  and let  $x_1 \in B(X) \setminus \{0\}$ . Inductively,

- Find  $x_2 \in B(X)$  such that  $\text{dist}(x_2, \mathbb{F}x_1) \geq 1 - \epsilon$
- Find  $x_3 \in B(X)$  such that  $\text{dist}(x_3, \text{span}\{x_1, x_2\}) \geq 1 - \epsilon$
- Find  $x_{n+1} \in B(X)$  such that  $\text{dist}(x_{n+1}, \text{span}\{x_1, \dots, x_n\}) \geq 1 - \epsilon$

Hence we have  $\{x_n\}_{n=1}^\infty \subset B(X)$  such that for  $m < n$ ,

$$\|x_n - x_m\| \geq d(x_n, \text{span}\{x_1, \dots, x_{n-1}\}) \geq 1 - \epsilon$$

so the sequence admits no converging subsequence and  $B(X)$  is not compact. ■



## 6 MORE TOPOLOGY

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

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

$$\tau_\beta = \left\{ \bigcup_{i \in I} B_i : \{B_i\}_{i \in I} \subset \beta, I \text{ any index set with } |I| \leq |\beta| \right\}$$

is 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  $f_\alpha : X \rightarrow X_\alpha$

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

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

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

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

*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  $(Y, \tau)$ -topological space,  $\emptyset \neq X \subseteq Y$ , and  $i : X \rightarrow Y$  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  $\alpha \in A$ ,  $p_\alpha : X \rightarrow X_\alpha$  denote the projection map onto the component  $\alpha$ . Then the product topology  $\pi = \sigma(X, \{p_\alpha\}_{\alpha \in A})$ . Hence,  $V \in \pi$  if and only if for any  $x \in V$ , there is  $\alpha_1, \dots, \alpha_n \in A$  and  $U_{\alpha_k} \in \tau_{\alpha_k}$  such that  $x_{\alpha_k} = p_{\alpha_k}(x) \in U_{\alpha_k}$  and  $x \in \bigcap_{k=1}^n p_{\alpha_k}^{-1}(U_{\alpha_k}) \subseteq V$ .

Note that if  $X = \prod_{n=1}^\infty X_n$ , each  $(X_n, \tau_n)$  is a topological space, then the basic open sets look like  $U_1 \times U_2 \times \dots \times U_m \times X_{m+1} \times X_{m+2} \times \dots$ .

(iv) **Linear topology:** 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,  $f : X \rightarrow \mathbb{F}$ . 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

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

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

- (v) Now let  $X$  be 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  $\omega^* \subseteq \omega = \sigma(X^*, X^{**}) \subseteq \tau_{\|\cdot\|}$ .

Let  $(X, \tau)$  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  $U_1, \dots, U_n$  covering  $K$ . If  $X$  itself is  $\tau$ -compact, we call  $(X, \tau)$  a compact 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  $\bar{S} = \bigcap \{F \subseteq X : S \subseteq F, X \setminus F \in \tau\}$ .

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

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

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

**PROOF** Suppose  $X$  is compact and  $\mathcal{F} \subseteq \mathcal{P}(X)$  has the finite intersection property but with  $\bigcap_{F \in \mathcal{F}} \bar{F} = \emptyset$ , then  $\{X \setminus \bar{F}\}_{F \in \mathcal{F}}$  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  $\bigcap_{F \in \mathcal{F}} F = \emptyset$ , so there is  $F_1, \dots, F_n \in \mathcal{F}$  with  $\bigcap_{k=1}^n F_k = \emptyset$ . Then  $\{X \setminus F_i\}_{i=1}^n$  is a finite subcover. ■

**Definition.** Let  $X$  be a non-empty set. An **ultrafilter** is a family  $\mathcal{U} \subseteq \mathcal{P}(X)$  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  $\mathcal{U}_{x_0} = \{U \subseteq X : x_0 \in U\}$ .

**6.2 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  $\mathcal{F} \subseteq \mathcal{U}$ .

**PROOF** Let  $\Phi = \{\mathcal{G} \subseteq \mathcal{P}(X) : \mathcal{F} \subseteq \mathcal{G}, \mathcal{G} \text{ has f.i.p.}\}$ . Then  $\Phi$  is partially ordered by inclusion. If  $\Gamma \subseteq \Phi$  is a chain, then  $\mathcal{G}_\Gamma = \bigcup_{\mathcal{G} \in \Gamma} \mathcal{G}$  contains  $\mathcal{F}$  and has the finite intersection property. Hence  $\Phi$  admits a maximal element  $\mathcal{U}$ . Let  $A \in \mathcal{P}(X) \setminus \mathcal{U}$ . Then  $\mathcal{U} \cup \{A\} \not\supseteq \mathcal{U}$ , so  $\mathcal{U} \cup \{A\}$  fails the finite intersection property. Hence get  $U_1, \dots, U_n$  so  $A \cap \bigcap_{k=1}^n U_k = \emptyset$ . Now if  $V_1, \dots, V_m \in \mathcal{U}$ , then  $\bigcap_{j=1}^n V_j \cap \bigcap_{k=1}^n U_k \subseteq \bigcap_{k=1}^n U_k \subseteq X \setminus A$ , 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. ■

**6.3 Corollary.** If  $\mathcal{U} \subseteq \mathcal{P}(X)$  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. ■

**6.4 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. ■

**6.5 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  $\delta_{\mathcal{U}} : \mathcal{P}(\mathbb{N}) \rightarrow \{0, 1\} \subset \mathbb{R}$  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  $\delta_{\mathcal{U}}(A \cup B) = \delta_{\mathcal{U}}(A) + \delta_{\mathcal{U}}(B)$ . If  $E_1, \dots, E_n \subseteq \mathbb{N}$  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)  $L_{\mathcal{U}}|_{\mathfrak{c}_0} = 0$ , so if  $x \in \ell_{\infty}^{\mathbb{R}}$ , then  $\liminf_{n \rightarrow \infty} x_n \leq L_{\mathcal{U}} \leq \limsup_{n \rightarrow \infty} x_n$
- (iii) Exactly one of  $2\mathbb{N}$  and  $2\mathbb{N}-1$  is in  $\mathcal{U}$ , so  $L(\chi_{2\mathbb{N}}) \neq L(\chi_{2\mathbb{N}-1})$ , so  $L_{\mathcal{U}}$  is not translation invariant.
- (iv) Let  $S \in \mathcal{B}(\ell_{\infty})$  be given by  $Sx = \left(\frac{x_1 + \dots + x_n}{n}\right)_{n=1}^{\infty}$ . Then  $L_{\mathcal{U}} \circ S$  is a Banach limit.

**Definition.** If  $(X, \tau)$  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  $U \in \mathcal{U}$ .

**6.6 Proposition.** *Let  $(X, \tau)$  be a topological space. Then  $(X, \tau)$  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

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

If  $(X, \tau)$  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  $\bigcap_{F \in \mathcal{F}} \overline{F} \supseteq \bigcap_{V \in \mathcal{U}} \overline{V} \neq \emptyset$ .

**6.7 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  $\pi$ . Then  $(X, \pi)$  is compact.*

**PROOF** Let  $\mathcal{U}$  be an ultrafilter on  $X$ ; we will show that it admits a  $\pi$ -limit point. Fix  $\alpha \in A$  and let  $\mathcal{U}_{\alpha} = \{p_{\alpha}(V) : V \in \mathcal{U}\}$ , where  $p_{\alpha}$  is the coordinate projection onto  $\alpha$ . 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  $p_{\alpha}^{-1}(S_{\alpha}) \in \mathcal{U}$ , and since  $p_{\alpha}^{-1}$  commutes with complementation,  $\mathcal{U}_{\alpha}$  is an ultrafilter. The last proposition provides a  $\tau_{\alpha}$ -limit point  $x_{\alpha}$  for  $\mathcal{U}_{\alpha}$ . Now let  $x = (x_{\alpha})_{\alpha \in A}$ , where  $x_{\alpha}$  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  $x_{\alpha_k}$  is a  $\tau_{\alpha_k}$ -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  $\pi$ -limit point of  $\mathcal{U}$ . ■

- 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  $X_\alpha$ , 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,  $(Y, \pi)$  is compact. Given  $\alpha_1, \dots, \alpha_n \in A$ , then  $\bigcup_{k=1}^n p_{\alpha_k}^{-1}(\{y\})$ . Since  $\prod_{k=1}^n X_{\alpha_k} \neq \emptyset$ , 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  $\prod_{\alpha \in A} X_\alpha = Y \setminus \bigcup_{\alpha \in A} p_\alpha^{-1}(\{y\}) \neq \emptyset$ .
- (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.

**6.8 Proposition.** *Let  $(X, \tau)$  be a compact space.*

- (i) *If  $K \subseteq X$  is closed, then  $K$  is compact.*  
 (ii) *If  $(Y, \sigma)$  is a topological space and  $f : X \rightarrow Y$  is continuous, then  $f(X)$  is compact.*

PROOF Immediate. ■

*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,  $A^B \cong \{f : B \rightarrow A\}$ .

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

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

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

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

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

and since  $\|f\| \leq 1$ , we have  $|g(x)| \leq |g(x) - f(x)| + |f(x)| < \epsilon/3 + \|x\|$ . Then since  $\epsilon > 0$  is arbitrary, get  $g \in X'$  and  $|g(x)| \geq \|x\|$ , 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  $r > 0$ , 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.

**Definition.** A topological space  $(X, \tau)$  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  $U_x \cap V_y = \emptyset$ .

- 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  $\pi$  the product topology. Then  $(X, \pi)$  is Hausdorff if and only if all  $(X_\alpha, \tau_\alpha)$  are Hausdorff. (Straightforward exercise).

**6.10 Proposition.** *Let  $(X, \tau)$  be a Hausdorff space,  $K \subseteq X$   $\tau$ -compact. Then  $K$  is  $\tau$ -closed.*

PROOF Straightforward exercise. ■

**6.11 Proposition.** *Let  $(X, \tau)$  be a compact space.*

- (i) *If  $(Y, \sigma)$  is a Hausdorff space and  $\phi : X \rightarrow Y$  is continuous and bijective, then  $\phi^{-1} : Y \rightarrow X$  is continuous.*  
 (ii) *If  $\tau' \subseteq \tau$  is a Hausdorff topology on  $X$ , so  $\tau' = \tau$ .*

PROOF (i) If  $F \subseteq X$  is  $\tau$ -closed, then it is  $\tau$ -compact. Hence  $(\phi^{-1})^{-1}(F) = \phi(F)$  is  $\sigma$ -closed, so by A1Q1,  $\phi^{-1}$  is continuous.  
 (ii)  $\text{id} : X \rightarrow X$  is continuous, so if  $U \in \tau'$ , then  $\text{id}^{-1}(U) = U \in \tau$ , so  $\text{id}$  is continuous. Hence by (1)  $\text{id}^{-1}$  is continuous so  $\tau \subseteq \tau'$ . ■

**6.12 Theorem. (Metrization)** *If  $X$  is a separable normed space, then  $B(X^*)$  is  $w^*$ -metrizable, i.e. there exists a metric  $\rho$  on  $B(X^*)$  such that  $w^*|_{B(X^*)} = \tau_\rho$ .*

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  $\rho$  be given by

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

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  $f \in V$ ,

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

so  $f_0 \in V \subset B_{\rho, \epsilon}(f_0)$ . Since  $f_0$  is arbitrary, we have  $\tau_\rho \subseteq w^*|_{B(X^*)}$ , but since  $w^*$  is compact and  $\tau_\rho$  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  $\rho$  above extends to all of  $X^* \times X^*$ . However,  $X^*$  with the weak\* topology is not metrizable if  $X$  is infinite dimensional.  
 (iii)  $X^* = \bigcup_{n=1}^{\infty} nB(X^*)$ , so each  $nB(X^*)$  is metrizable and compact, and thus  $w^*$ -separable. Thus if  $X$  is separable, then  $X^*$  is itself separable.

## 7 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  $x : N \rightarrow X$ .

**Definition.** If  $(x_v)_{v \in N}$  is a net in  $X$ ,  $A \subseteq X$ , we say that  $(x_v)_{v \in N}$  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  $v' \geq v$  so  $x_{v'} \in A$ .

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

**Definition.** We call  $\phi : M \rightarrow N$  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

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

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**.

### NETS AND TOPOLOGY

Now, suppose  $(X, \tau)$  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  $v_U$  such that  $x_v \in U$  whenever  $v \geq v_U$ . We write  $x_0 = \lim_{v \in N} x_v$ , the  $\tau$ -limit of  $(x_v)_{v \in N}$ . Note that this is an abuse of notation, since limit points need not be unique (when  $(X, \tau)$  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  $U$ .

**7.1 Proposition.** *If  $(x_v)_{v \in N}$  is a net in  $(X, \tau)$  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  $\tau$ -limit point of  $x_{v_\mu}$  for some subnet  $(x_{v_\mu})_{\mu \in M}$  of  $(x_v)_{v \in N}$ .*

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

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

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

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

Let's see that  $\mathcal{F}$  is filtering: suppose  $F_{v,U}$  and  $F_{v',U'}$  are in  $\mathcal{F}$ . Get  $\mu \geq v$  and  $\mu \geq v'$  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  $F_{\mu',V} \subseteq F_{v,U} \cap F_{v',U'}$ . We then let  $M = N_{\mathcal{F}}$  be the net construction from the filtering family and set  $v_{(v,F)} = v$ .

Now set  $N_{\mathcal{F}} = \{(v, F) : v \in F, F \in \mathcal{F}\}$  with the standard preorder and  $v_{(v,F)} = v$ . Then the map  $(v, F) \mapsto v$  from  $N_{\mathcal{F}} \rightarrow N$  is eventually cofinal: if  $v_0 \in N$  is arbitrary, take any  $F_0 = F_{v_0, U} \in \mathcal{F}$ . Then  $F_0 = \{v \in N : v \geq v_0, x_v \in U\}$ , so if  $F_{\mu, V} \in \mathcal{F}$  with  $F_{\mu, V} \subseteq F_0$ , 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:  $(v, F) \mapsto v : N_{\mathcal{F}} \rightarrow N$  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. ■

**7.2 Proposition.** *If  $(Y, \sigma)$  is another topological space, then  $f : X \rightarrow Y$  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  $f(x_0) \in V$ . The assumptions on  $f$  tell us there is  $v \in N_{\tau_{x_0}}$  such that for  $v \geq v_V$ , we have  $f(x_v) \in V$ . 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:

- Given topologies  $\tau, \tau'$  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  $x_0 \in X$ .
- (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  $\pi$ . 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  $\pi$  is the coarsest topology making each  $\mu_\alpha$  continuous.
- 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  $x \in X$ .

## ROLES OF WEAK AND WEAK\* TOPOLOGIES IN CONVEXITY

**7.3 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

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

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

**PROOF** The separation theorem and the fact that  $B$  is  $\|\cdot\|$ -open (i.e.  $w^* \subseteq \tau_{\|\cdot\|}$ ) provides  $F \in X^{**}$  and  $\alpha \in \mathbb{R}$  such that  $\operatorname{Re} F(f) \leq \alpha \operatorname{Re} F(g)$  for  $f \in A$ ,  $g \in B$ . Since  $B \in w^* = \sigma(X^*, \hat{X})$ , if  $f_0 \in B$ , then there are  $x_1, \dots, x_n$  in  $X$  such that

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

Let  $Y = \bigcap_{i=1}^n \ker \hat{x}_i \subseteq X^*$ . Then for  $i = 1, \dots, n$ ,  $\hat{x}_i(f_0 + Y) = \{f_0(x_i)\} \subset f_0(x_i) + \mathbb{D}$ , so that  $f_0 + Y \subseteq U \subseteq B$ . Thus if  $f \in Y$ , then  $\operatorname{Re} F(f_0 + f) > \alpha$  and hence  $\operatorname{Re} F(f) > \alpha - \operatorname{Re} F(f_0)$  which implies that  $f \in \ker \operatorname{Re} F$ , so  $f \in \ker F$ . That is,  $Y \subseteq \ker F$ . The next lemma shows that  $F \in \operatorname{span}\{\hat{x}_1, \dots, \hat{x}_n\} \subseteq \hat{X}$ , i.e.  $F = \hat{x}$  for some  $x \in X$ . ■

**7.4 Lemma.** In an  $\mathbb{F}$ -vector space, if  $f_0, f_1, \dots, f_n \in X'$  with  $\ker f_0 \supseteq \bigcap_{i=1}^n \ker f_i$ , then  $f \in \operatorname{span}\{f_1, \dots, f_n\}$ .

**PROOF** Define  $T : X \rightarrow \mathbb{F}^n$  by  $Tx = (f_1(x), \dots, f_n(x))$ . Then  $\ker T = \bigcap_{i=1}^n \ker f_i$ . Let  $\mathcal{R} = \operatorname{im} T \subseteq \mathbb{F}^n$  and  $g_0 \in \mathcal{R}'$  by  $g_0(Tx) = f_0(x)$ . Then  $g_0$  is well-defined: if  $Tx = Ty$ , then  $x - y \in \ker T \subseteq \ker f_0$ , so  $f_0(x - y) = 0$  so  $f_0(x) = f_0(y)$ . Also  $g_0$  is linear. Let  $g \in (\mathbb{F}^n)'$  such that  $g|_{\mathcal{R}} = g_0$ . Hence there are  $\alpha_1, \dots, \alpha_n \in \mathbb{F}$  such 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  $f_0 = \sum_{j=1}^n \alpha_j f_j$ . ■

**7.5 Theorem. ( $w^*$ -Closed Convex Hull)** If  $S \subset X^*$ , then

$$\overline{\operatorname{co}}^{w^*} S = \bigcap \{f \in X^* : \operatorname{Re} f(x) \leq \alpha\} \supseteq S : x \in X, \alpha \in \mathbb{R}$$

**PROOF** The set on the right is  $w^*$ -closed and convex being the intersection of such. Conversely, if  $f \in X^* \setminus \overline{\operatorname{co}}^{w^*} S$ , which is open, then there is a basic  $w^*$ -open neighbourhood

$$B = \bigcap_{j=1}^n \hat{x}_j^{-1}(f(x_j) + \mathbb{D}) \subseteq X^* \setminus \overline{\operatorname{co}}^{w^*} S$$

so that  $B \cap \overline{\operatorname{co}}^{w^*} S = \emptyset$ . Also,  $B$  is convex. ■

**Remark.** If  $X$  is a normed space, a closed half space  $H = \{x \in X : \operatorname{Re} f(x) \leq \alpha\}$  for some  $f$  in  $X^*$ ,  $\alpha \in \mathbb{R}$ . Hence,  $H$  is weakly closed  $(\operatorname{Re} f)^{-1}([\alpha, \infty)) = f^{-1}(\{z \in \mathbb{C} : \operatorname{Re} z \geq \alpha\})$  is  $w$ -closed. Thus if  $S \subset X$ , we have  $\overline{\operatorname{co}} S \in w = \sigma(X, X^*) \subseteq \tau_{\|\cdot\|}$ , so  $\overline{\operatorname{co}} S$  is automatically weakly closed. Hence if  $C \subseteq X$  is convex, then  $C$  is norm closed if and only if  $C$  is  $w$ -closed.



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

$$\begin{aligned} E^\circ &= \{f \in X^* : \operatorname{Re} f(x) \leq 1 \text{ for all } x \text{ in } E\} \subseteq X^* \\ &= \bigcap_{x \in E} \{f \in X^* : \operatorname{Re} \hat{x}(f) \leq 1\} \end{aligned}$$

so  $E^\circ$  is convex and  $w^*$ -closed in  $X^*$ , and  $0 \in E^\circ$ .

If  $F \subseteq X^*$  (non-empty), let the **pre-polar** of  $F$  be given by

$$F_\circ = \{x \in X : \operatorname{Re} f(x) \leq 1 \text{ for all } f \text{ in } F\}$$

so, like above,  $F_\circ$  is convex,  $(w-)$ closed, and  $0 \in F_\circ$ .

**7.6 Theorem. (Bipolar)** (i) If  $\emptyset \neq E \subseteq X$ , then  $(E^\circ)_\circ = \overline{\operatorname{co}}(E \cup \{0\})$ .

(ii) If  $\emptyset \neq F \subseteq X^*$ , then  $(F_\circ)^\circ = \overline{\operatorname{co}}^{w^*}(F \cup \{0\})$ .

**PROOF** (i) Note that  $E \cup \{0\} \subseteq (E^\circ)_\circ$ , so  $\overline{\operatorname{co}}(E \cup \{0\}) \subseteq (E^\circ)_\circ$ . If  $x_0 \in X \setminus \overline{\operatorname{co}}(E \cup \{0\})$ , then the separation theorem provides  $f \in X^*$ ,  $\alpha \in \mathbb{R}$  so  $\operatorname{Re} f(x_0) > \alpha \geq \operatorname{Re} f(x)$  for  $x \in E \cup \{0\}$ . Notice that  $\alpha \geq \operatorname{Re} f(0) = 0$ , and we let  $\beta = \frac{1}{2}[\operatorname{Re} f(x_0) + \alpha] > 0$ , so  $\operatorname{Re} f(x_0) > \beta \geq \operatorname{Re} f(x)$  for  $x \in E \cup \{0\}$ ,  $\beta > 0$ . Let  $g = \frac{1}{\beta}f$  and we see that  $g \in E^\circ$  and as  $\operatorname{Re} g(x_0) > 1$ ,  $x_0 \notin (E^\circ)_\circ$ .

(ii) Similar, use  $w^*$ -separation. ■

**Remark.** Let  $Y \subseteq X$  be a subspace. If  $f \in Y^0$ , then  $\operatorname{Re} f(y) \leq 1$  for  $y \in Y$  implies that  $f(y) = 0$  for all  $y \in Y$ . We write  $Y^a = Y^0$ , and  $Y^a = \{f \in X^* : f|_Y = 0\}$  is called the **annihilator** of  $Y$ . Likewise, if  $Z \subseteq X^*$  is a subspace, then  $Z_a = Z_0$  where  $Z_a = \{x \in X : f(x) = 0 \text{ for each } f \in Z\}$  is called the **pre-annihilator**. Notice that  $Y^a$  and  $Z_a$  are subspaces.

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

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

**7.8 Lemma.** If  $X$  is a normed space, then  $B(X)^0 = B(X^*)$  and  $B(X^*)_0 = B(X)$ .

**PROOF** If  $f \in B(X)^0$ , then  $\operatorname{Re} f(x) \leq 1$  for  $x \in B(X)$ . Thus for  $x \in B(X)$ ,  $|f(x)| = \overline{\operatorname{sgn} f(x)} f(x) = f(\operatorname{sgn} f(x)x) \leq 1$ , so  $\|f\| \leq 1$  and  $f \in B(X^*)$ . Conversely, if  $f \in B(X^*)$ ,  $x \in B(X)$ , then  $\operatorname{Re} f(x) \leq |f(x)| \leq 1$  so  $f \in B(X)^0$ . Then use the Bipolar theorem. ■

**7.9 Theorem. (Goldstine)** If  $X$  is a normed space, then  $\overline{B(\hat{X})}^{w^*} = B(X^{**})$ . Note that  $w^* = \sigma(X^{**}, \hat{X}^*)$ .

**PROOF** The Bipolar theorem provides  $\overline{B(\hat{X})}^{w^*} = \overline{\operatorname{co}}^{w^*} B(\hat{x}) = (B(\hat{X})_\circ)^\circ$ . But, in  $X^*$ ,

$$\begin{aligned} B(X)^\circ &= \{f \in X^* : \operatorname{Re} f(x) \leq 1 \text{ for } x \text{ in } B(X)\} \\ &= \{f \in \hat{X}^* : \operatorname{Re} \hat{x}(f) \leq 1 \text{ for } x \text{ in } B(X)\} \\ &= B(\hat{X})_\circ \end{aligned}$$

Hence we have, using the lemma,

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

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

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

**7.10 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  $x_\nu = \|F\|y_\nu$ . ■

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  $x \mapsto \hat{x}$  is an isometry, so a reflexive space is always complete.

**7.11 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** The map  $\iota : x \mapsto \hat{x}$  is a  $w - w^*|_{\hat{X}}$ -homeomorphism. Recall  $w^* = \sigma(X^{**}, \hat{X}^*)$ , and  $w^*|_{\hat{X}} = \sigma(\hat{X}, (\hat{X})^*|_{\hat{X}})$  and we have for  $x_0 \in X$ , net  $(x_\nu)_{\nu \in N}$  in  $X$ ,

$$\begin{aligned} w - \lim_{\nu \in N} x_\nu = x_0 &\iff \lim_{\nu \in N} f(x_\nu) = f(x_0) \forall f \in X^* \\ &\iff \lim_{\nu \in N} \hat{x}_\nu(f) = \hat{x}_0(f) \forall f \in X^* \\ &\iff \lim_{\nu \in N} \hat{f}(\hat{x}_\nu) = \hat{f}(\hat{x}_0) \end{aligned}$$

and having the same convergent nets means that the topologies are the same.

(i  $\Rightarrow$  ii) By assumption,  $\widehat{B(\hat{X})} = B(\hat{X}) = B(X^{**})$ . Since  $B(X^{**})$  is  $w^*$ -compact, and hence  $\iota^{-1}(B(X^{**})) = B(X)$  is  $w$ -compact

(ii  $\Rightarrow$  i) If  $B(X)$  is  $w$ -compact, then since  $x \mapsto \hat{x} : X \rightarrow X^{**}$  is continuous, we see that  $B(\hat{X}) = \widehat{B(X)}$  is  $w^*$ -compact.

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

(iii  $\Rightarrow$  iv)  $B(X^*)$  is compact, hence  $w$ -compact, so by (ii) implies (i) applied to  $X^*$ , we have that  $X^*$  is reflexive.

(iv  $\Rightarrow$  i) We assume  $\widehat{X^*} = X^{***}$ . Thus on  $X^{***}$ , we have  $w = \sigma(X^{**}, X^{***}) = \sigma(X^{**}, \widehat{X^*}) = w^*$ . Now  $B(\hat{X}) = B(X^{**}) \cap \hat{X}$  is norm-closed and convex, hence  $w$ -closed, by Closed Convex Hull

theorem. Thus from above,  $B(\hat{X})$  is  $w^*$ -closed, so  $B(\hat{X}) = \overline{B(\hat{X})}^{w^*} = B(X^{**})$  by Goldstine, so  $\hat{X} = X^{**}$ . ■

**7.12 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. ■

### 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.  $\text{ext } C = \{x \in C : \{x\} \text{ is a face of } C\}$ . Hence  $x \in \text{ext } C$  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)  $\text{ext } F \subseteq \text{ext } C$ .

(iv) If  $f \in X'$  and  $\text{Re } f(C) = [a, b]$ , then  $(\text{Re } f)^{-1}(\{b\})$  is itself a face of  $C$ .

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

**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  $F_1, \dots, F_n \in \mathcal{C}$ , we may assume  $F_1 \supseteq \dots \supseteq F_n$  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  $x \in X$ ,  $\text{Re } \hat{x} : X^* \rightarrow \mathbb{R}$  is  $w^*$ -continuous, and hence  $\text{Re } \hat{x}(M) = [a_x, b_x]$  since the only compact convex subsets of  $\mathbb{R}$  are compact intervals. But then  $F_x = (\text{Re } \hat{x})^{-1}(\{b_x\}) \cap M$  is a  $w^*$ -closed convex face in  $M$ , so that  $F_x = M$ . If  $f, g \in M$ , then  $\text{Re } f(x) = \text{Re } \hat{x}(f) = b_x = \text{Re } \hat{x}(g) = \text{Re } g(x)$ , so  $f = g$  and hence  $M = \{f\}$  and  $f \in \text{ext } F$ .

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

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

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

$$\widehat{\overline{co}^w \text{ext } C} = \overline{co}^{w^*} \text{ext } \hat{C} = \hat{C}$$

so that  $C = \overline{co}^w \text{ext } C = \overline{co} \text{ext } C$  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  $\text{ext } B(X) \subseteq S(X)$ .

**7.15 Proposition.**  $\text{ext } B(\ell_p) = S(\ell_p)$

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

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

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

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

**7.17 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  $\text{ext } B(c_0) \neq \emptyset$ . ■

**Definition.** Let  $(X, \tau)$  be a compact Hausdorff space, and let

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

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

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

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

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

$$|\mu_f(g)| = |\mu_f(g^+ - g^-)| = |\mu(fg^+) + \mu(fg^-)| \leq \mu(fg^+) + \mu(fg^-) = \mu(fg) \leq \mu(f\|g\|_{\infty}) = \mu(f)\|g\|_{\infty} \quad (7.1)$$

and, with  $f = \mu$ , we have

$$|\mu(g)| \leq \mu(|g|) \quad (7.2)$$

(II) Let  $\delta \in \text{ext } P(X)$ . We first show for  $h, g$  in  $C$  that  $\delta(hg) = \delta(h)\delta(g)$ . To see this, since  $\delta \neq 0$ , we may find  $0 \leq f \leq 1$  such that  $0 < \delta(f) < 1$ . Now let  $\mu = \frac{1}{\delta(f)}\delta_f$  so, for  $g \in C$ , (7.1) provides

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

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

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

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

Suppose now for each  $x \in X$ , there exists some  $f_x \in \ker \delta$  so that  $f_x(x) \neq 0$ . 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  $x_1, \dots, x_n$  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

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

since each  $f_{x_j} \in \ker \delta$ , which is absurd. Hence there is  $x \in X$  so  $f(x) = 0$  whenever  $f \in \ker \delta$ , so  $\ker \delta \supseteq \ker \hat{x}$ , so  $\delta \in \mathbb{R}\hat{x}$  and since  $\delta(\#) = 1 = \hat{x}(\#)$ , 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 (7.2),

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

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  $\text{ext } B(C^{\mathbb{F}}(X, \tau)^*) = \{z\hat{x} : z \in \mathbb{F}, |z| = 1, x \in X^*\}$ .

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

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

**PROOF** If  $\delta \in \text{ext } PA(\mathbb{N})$ , 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  $[\cdot, \cdot] : X \rightarrow \mathbb{F}$  is called **Hermitian** if for  $x, x', y$  in  $X$ ,  $\alpha \in \mathbb{F}$ , we have

$$(i) \quad [x + \alpha x', y] = [x, y] + \alpha [x', y]$$

$$(ii) \quad [\overline{y}, x] = [x, y]$$

and furthermore **positive** if

3.  $[x, x] \geq 0$  for all  $x \in X$

and **non-degenerate** if

4.  $[x, y] = 0$  for all  $y \in X$  implies  $x = 0$ .

**8.1 Proposition.** Let  $[\cdot, \cdot]$  be a positive Hermitian form. Let  $p(x) = [x, x]^{1/2}$ , so  $p : X \rightarrow [0, \infty)$ . Then for  $x, y \in X$  and  $\alpha \in \mathbb{F}$ , we have

- (i)  $p(\alpha x) = |\alpha|p(x)$
- (ii)  $[x, y] \leq p(x)p(y)$
- (iii)  $p(x + y) \leq p(x) + p(y)$
- (iv)  $[\cdot, \cdot]$  is non-degenerate if and only if  $[x, x] > 0$  for  $x \in X \setminus \{0\}$ .

Furthermore, in this case, we have

- Equality in (ii) if and only if  $x, y$  are linearly dependent
- $[x, y] = p(x)p(y)$  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)  $p(\alpha x) = (\alpha \bar{\alpha} [x, x])^{1/2} = |\alpha|p(x)$

(ii) If  $\alpha \in \mathbb{F}$ , then

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

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

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

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

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

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

(iii)

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

(iv) We see, by (iii), if  $p(x)^2 = [x, x] = 0$ , then  $[x, y] = 0$  for all  $y$ . Hence  $[\cdot, \cdot]$  is non-degenerate if and only if  $[x, x] > 0$  for  $x \in X \setminus \{0\}$ . If  $x, y$  are linearly dependent, then equality holds in (ii) by direct computation. If  $x, y$  are not linearly dependent, then the choice of  $\alpha = \operatorname{sgn} [x, y]$  in (ii) gives strict inequality. The condition  $[x, y] = p(x)p(y)$  requires non-negativity of  $[x, y]$ , showing one is a  $R_{\geq 0}$  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  $(f, g) = \int_0^1 f \bar{g}$
- (ii) (Euclidean Space) Recall  $\ell = \{x \in \mathbb{F}^{\mathbb{N}} : x_n = 0 \text{ for all but finitely many } n\}$ , and  $(\ell, \langle \cdot, \cdot \rangle)$  with  $\langle x, y \rangle = \sum_{j=1}^{\infty} x_j \bar{y}_j$
- (iii) (Hilbert Space)  $(L_2[0, 1], (\cdot, \cdot))$ ,  $(f, g) = \int_{[0, 1]} f \bar{g}$ .
- (iv) (Hilbert Space)  $(\ell_2, (\cdot, \cdot))$ ,  $(x, y) = \sum_{j=1}^{\infty} x_j \bar{y}_j$  (convergence by Hölder's inequality)
- (v) (Non-separable Hilbert Space) Let  $\Gamma$  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  $\Gamma_n = \{\gamma \in \Gamma : a_\gamma \geq 1/n\}$  and we have

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

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  $\ell_2$  says that

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

Thus,  $\sum_{k=1}^{\infty} x_{\gamma_k} \bar{y}_{\gamma_k}$  is absolutely converging. Write  $(x, y) = \sum_{\gamma \in \Gamma} x_\gamma \bar{y}_\gamma = \sum_{k=1}^{\infty} x_{\gamma_k} \bar{y}_{\gamma_k}$ . 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  $\Gamma'$ ), 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  $w : \mathbb{N} \rightarrow (0, \infty)$ . Let  $\ell_2^w = \{x \in \mathbb{F}^{\mathbb{N}} : \sum_{k=1}^{\infty} |x_k|^2 w(k) < \infty\}$ . Notice that if  $x, y \in \ell_2^w$ , then  $(x_k w(k)^{1/2})_{k=1}^{\infty}, (y_k w(k)^{1/2})_{k=1}^{\infty} \in \ell_2$ , so it follows that

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

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

## VARIOUS IDENTITIES

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

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  $\left\| \sum_{j=1}^n x_j \right\|^2 = \sum_{j=1}^n \|x_j\|^2$ .
- *Parallelogram law*:  $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$ .

Note that if  $\mathbb{F} = \mathbb{C}$ ,  $(x, y) = \frac{1}{4} \sum_{k=0}^3 i^k \|x + i^k y\|^2$  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  $f_y : \mathcal{E} \rightarrow \mathbb{F}$  by  $f_y(x) = (x, y)$  is linear with  $\|f_y\| = \|y\|$ . Furthermore,  $|f_y(x)| = \|y\|$  for  $y \neq 0$ ,  $x \in B(\mathcal{E})$  if and only if  $x = \frac{\zeta}{\|y\|} y$  where  $|\zeta| = 1$ .*

**PROOF** Linearity is from an assumption on  $(\cdot, \cdot)$ . Furthermore, Cauchy-Schwarz tells us that

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

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

**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  $\mathcal{M} = \text{span}\{e_1, \dots, e_n\}$ . Then for  $x \in \mathcal{E}$  we have that*

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

**PROOF** (i) If  $1 \leq k \leq n$ , we have

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

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  $\alpha \in \mathbb{F}^n$ ,

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

so that

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



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

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

which is minimized when  $s = 1$ . ■

- Remark.* (i) The proof above shows that  $P_{\mathcal{M}}x$  is the unique element of  $\mathcal{M}$  satisfying  $\text{dist}(x, \mathcal{M}) = \|x - P_{\mathcal{M}}x\|$ .  
 (ii) It may be shown that  $P_{\mathcal{M}} : \mathcal{E} \rightarrow \mathcal{E}$  is linear with  $\text{im } P_{\mathcal{M}} = \mathcal{M}$ ,  $P_{\mathcal{M}}^2 = P_{\mathcal{M}}$ , 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)  $\overline{\text{span}} E = \mathcal{E}$
- (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) 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  $\mathcal{E}_F = \text{span } F$ , so that  $\mathcal{E}_F \subseteq \mathcal{E}_{F'}$  if  $F \subseteq F'$  in  $\mathcal{F}$  and  $\text{span } E = \bigcup_{F \in \mathcal{F}} \mathcal{E}_F$ . Hence for  $x \in \mathcal{E}$ , we have

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

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

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

( $ii \Leftrightarrow iii$ ) We have

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

( $ii \Rightarrow iv$ ) 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;  $x = \lim_{F \in \mathcal{F}} \sum_{e \in F} (x, e) e \in \overline{\text{span } E}$ , i.e.  $\mathcal{E} \subseteq \overline{\text{span } E} \subseteq \mathcal{E}$ . ■

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

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

**PROOF** Let  $\mathcal{E}_n = \text{span}\{x_1, \dots, x_n\}$ . We set

$$\begin{aligned} z_1 &= x_1 & e_1 &= \frac{z_1}{\|z_1\|} \\ z_2 &= x_2 - P_{\mathcal{E}_1} x_2 & e_2 &= \frac{z_2}{\|z_2\|} \\ &\vdots & & \\ z_{n+1} &= x_{n+1} - P_{\mathcal{E}_n} x_{n+1} & e_{n+1} &= \frac{z_{n+1}}{\|z_{n+1}\|} \end{aligned}$$

where  $P_{\mathcal{E}_n} x = \sum_{j=1}^n (x, e_j) e_j$ . Inductively,  $z_n \in \mathcal{E}_n$  and  $z_n \perp \mathcal{E}_k$  for  $k = 1, \dots, n-1$ . Hence each set  $\{z_1, \dots, z_n\}$  is orthonormal and  $\text{span}\{z_1, \dots, z_n\} \subseteq \text{span}\{x_1, \dots, x_n\}$  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  $n_{k+1} = \min\{n : x_n \notin \text{span}\{x_{n_1}, \dots, x_{n_k}\}\}$ . Then  $\{x_{n_1}, x_{n_2}, \dots\}$  and normalize to get an orthonormal set  $E = \{e_1, e_2, \dots\}$  which satisfies  $\text{span } E = \overline{\text{span}\{x_n\}_{n=1}^\infty} = \mathcal{E}$ . ■

**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 an  $\alpha = (\alpha_e)_{e \in E} \in \ell_2(E)$ , we have that  $\sum_{e \in E} \alpha_e e \in \mathcal{E}$ .

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

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

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

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

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

( $\Leftarrow$ ) Let  $(x^{(n)})_{n=1}^\infty$  be Cauchy in  $\mathcal{E}$ . Let  $\mathcal{M} = \overline{\text{span}\{x^{(n)}\}_{n=1}^\infty} \subset \mathcal{E}$  so  $\mathcal{M}$  is separable and admits a countable orthonormal basis  $E = (e_1, e_2, \dots)$ . 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  $U : \ell_2(E) \rightarrow \mathcal{M}$  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  $\|\alpha^{(n)} - \alpha^{(m)}\|_2 = \|U_\alpha^{(n)} - U_\alpha^{(m)}\| = \|x^{(n)} - x^{(m)}\|$  so  $(\alpha^{(n)})_{n=1}^{\infty}$  is Cauchy and in  $\ell_2(E)$  and hence admits a limit  $\alpha$ . Furthermore,

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

as required. ■

**Definition.** If  $\emptyset \neq S \subset \mathcal{E}$ ,  $(\mathcal{E}, (\cdot, \cdot))$  a Euclidean space, we define its **perpendicular** 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)  $(\overline{\text{span}} S)^\perp = S^\perp$ . Notice that  $(\text{span } S)^\perp = S^\perp$  (easy test) and use (iii)

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