UNIVERSITY OF CALIFORNIA

LET THERE BE LIGHT

SAN DIEGO

MATH 240B LECTURE NOTES: TOPOLOGY AND FUNCTIONAL ANALYSIS

BRUCE K. DRIVER

Abstract. These are lecture notes from Math 240B on point set topology and functional analysis following Folland's Book.

Contents

Date: March 16, 2001 *File:top5.tex*.

This research was partially supported by NSF Grant DMS 96-12651.

Department of Mathematics, 0112.
University of California, San Diego .
La Jolla, CA 92093-0112 .

1. Point Set Topology Basics

As usual we will let X and Y be sets and $\mathcal{P}(X)$ denote the power set of X.

Definition 1.1. A collection of set $\tau \subset \mathcal{P}(X)$ is a **topology** on X if

1. $\emptyset, X \in \tau$,
2. τ is closed under finite intersections, and
3. τ is closed under arbitrary unions.

The members of τ are called **open** sets and the sets $C \subset X$ such that $C^c \in \tau$ are called **closed** sets. We will often write $V \subset_o X$ to denote that $V \in \tau$ and $C \sqsubset X$ to indicate that C is a closed subset of X.

Example 1.2. Suppose that $X = \{1, 2\}$, then

1. $\tau_0 = \{\emptyset, X\}$ (Trivial topology on X), $\tau_1 = \{\emptyset, X, \{1\}\}$, $\tau_2 = \{\emptyset, X, \{2\}\}$, $\tau_3 = \{\emptyset, X, \{1\}, \{2\}\}$ - Discrete topology are all topologies on X.
2. Suppose (X, τ) is a topology space and $Y \subset X$. Then $\tau_Y \equiv \{U \cap Y : U \in \tau\}$ is a **topology** on Y called the **relative topology** on Y. (Notice that the closed sets in Y relative to τ_Y are precisely those sets of the form $C \cap Y$ where C is close in X. Indeed, $B \subset Y$ is closed iff $Y \setminus B = Y \cap V$ for some $V \in \tau$ which is equivalent to $B = Y \setminus (Y \cap V) = Y \cap V^c$ for some $V \in \tau$.)
3. Suppose that X is a set and $\mathcal{E} \subset \mathcal{P}(X)$ be a collection of subsets of X. We let $\tau(\mathcal{E})$ be the **smallest topology** on X containing $\mathcal{E}$. If $\tau = \tau(\mathcal{E})$, we call $\mathcal{E}$ a **subbase** for the topology τ.

Proposition 1.3. *Given $\mathcal{E} \subset \mathcal{P}(X)$, $\tau(\mathcal{E})$ consists of arbitrary unions of finite intersections of elements from $\mathcal{E} \cup \{X, \emptyset\}$.*

Proof. Let $\tau \subset \mathcal{P}(X)$ denote the collection of sets consisting of arbitrary unions of finite intersections of elements from $\mathcal{E} \cup \{X, \emptyset\}$. Then $\tau \subset \tau(\mathcal{E})$ so to finish the proof it suffices to show that τ is a topology.

Suppose $\{A_i : i \in I\}$ and $\{B_j : j \in J\}$ are collection's of sets which are finite intersections of the elements in $\mathcal{E} \cup \{X, \emptyset\}$. Set

$$A = \bigcup_{i \in I} A_i \text{ and } B = \bigcup_{j \in J} B_j.$$

Then

$$A \cap B = \bigcup \{A_i \cap B_j : i \in I \text{ and } j \in J\} \in \tau$$

So τ is closed under finite intersections, it is obviously closed under unions and contains $X, \emptyset$. That is τ is a topology. ■

Definition 1.4. We say $\mathcal{E}$ is a **base** for a topology if $\emptyset \in \mathcal{E}$, $X = \cup\mathcal{E}$ and for all $U, V \in \mathcal{E}$,

$$U \cap V = \bigcup\{W : W \in \mathcal{E} \text{ and } W \subset U \cap V\}.$$

Suppose that $\mathcal{E}$ is a base for a topology and $\{U_i\}_{i=1}^n \subset \mathcal{E}$ and $\cap_{i=1}^n U_i \neq \emptyset$. Then for $x \in \cap_{i=1}^n U_i$, there exists $W_1 \in \mathcal{E}$ such that $x \in W_1 \subset U_1 \cap U_2$. Similarly, there exists $W_2 \in \mathcal{E}$ such that $x \in W_2 \in W_1 \cap U_3 \subset U_1 \cap U_2 \cap U_3$. Continuing inductively this way we find there exists $W_{n-1} \in \mathcal{E}$ such that $x \in W_{n-1} \subset \cap_{i=1}^n U_i$. This shows that

$$\cap_{i=1}^n U_i = \bigcup\{W : W \in \mathcal{E} \text{ and } W \subset \cap_{i=1}^n U_i\}. \tag{1.1}$$

The proof of the following proposition is an easy consequence of Eq. (1.1) and Proposition 1.3.

Proposition 1.5. *If $\mathcal{E}$ is a base for a topology then the general element of $\tau(\mathcal{E})$ is of the form $V = \cup\mathcal{B}$ where $\mathcal{B} \subset \mathcal{E}$.*

Example 1.6. Suppose that $X = \{1, 2, 3, 4\}$ and $\mathcal{E} = \{A := \{1, 2\}, B := \{2, 3\}\}$ in which case

$$\tau(\mathcal{E}) = \{A \cap B, A, B, A \cup B, X, \emptyset\}.$$

If $\widetilde{\mathcal{E}} = \{\{1, 2\}, \{2\}, \{2, 3\}, \emptyset, X\}$ then $\tilde{\mathcal{E}}$ is a base for $\tau(\mathcal{E})$.

Example 1.7. Let (X, ρ) be a metric space and let

$$\mathcal{E}_\rho = \{B_x(r) : x \in X \text{ and } r > 0\} \cup \{X, \emptyset\}$$

where

$$B_x(r) = \{y \in X : \rho(x, y) < r\}.$$

Then $\mathcal{E}_\rho$ is a base for a topology which we denote by τ_ρ and call the induced topology on (X, ρ). Notice that V is open in this topology iff for all $x \in V$, there exists $\epsilon > 0$ such that $B_x(\epsilon) \subset V$.

Proof. To prove that $\mathcal{E}_\rho$ is a base for a topology we must show: for any $x, y \in X$ and $\epsilon, \delta > 0$ such that $B_x(\epsilon) \cap B_y(\delta) \neq \emptyset$, then for all for all $z \in B_x(\epsilon) \cap B_y(\delta)$ there exists $r > 0$ such that $B_z(r) \subset B_x(\epsilon) \cap B_y(\delta)$. This is fairly clear from Figure 1. The formal proof is as follows. If $\omega \in B_z(r)$, then by the triangle inequality,

$$\rho(x, \omega) \leq \rho(x, z) + \rho(\omega, z) = \rho(x, z) + r$$

and similarly

$$\rho(y, \omega) \leq \rho(y, z) + r.$$

Therefore if we choose $0 < r < \min(\epsilon - \rho(x, z), \delta - \rho(y, z))$, then $\rho(x, \omega) < \epsilon$ and $\rho(y, w) < \delta$, i.e. $\omega \in B_x(\epsilon) \cap B_y(\delta)$ showing that $B_z(r) \subset B_x(\epsilon) \cap B_y(\delta)$. ■

If (X, ρ) is a metric space we will also define the closed ball at x of radius r by

$$C_x(r) := \{y \in X : \rho(x, y) \leq r\}.$$

Let us now show that $C = C_x(r)$ is indeed closed. Recall that in a metric space that $|\rho(x, y) - \rho(x, z)| \leq \rho(y, z)$. If $y \notin C$, then $\delta := \rho(x, y) - \epsilon > 0$ and if $z \in B_y(\delta)$ we have

$$\epsilon < \epsilon + \delta - \rho(y, z) \leq \rho(x, y) - \rho(y, z) \leq \rho(x, z)$$

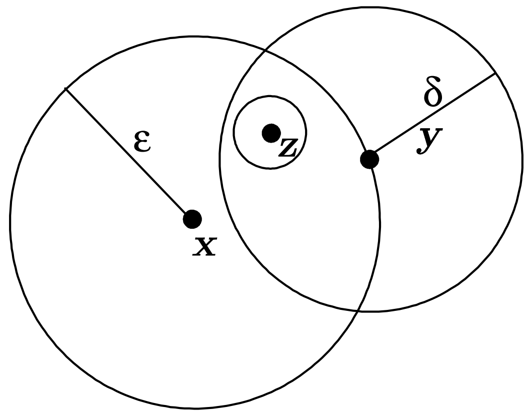


FIGURE 1. Squeezing a ball in the intersection.

which shows that $B_\delta(y) \cap C = \emptyset$. This shows that C^c is open and hence that C is closed.

Definition 1.8. Let $A \subset X$ be a set. The interior, A^0, of A is the largest open set contained in A, i.e.

$$A^0 = \bigcup\{V \in \tau : V \subset A\}.$$

The closure $\bar{A}$, of A is the smallest closed set which contains A, i.e.

$$\bar{A} = \bigcap\{C \supset A : C \sqsubset X\}.$$

The boundary of A is the set

$$\partial A \equiv \bar{A} \setminus A^0.$$

We also define the set of **accumulation points** of A to be

$$\text{acc}(A) = \{x \in X : V \cap A \setminus \{x\} \neq \emptyset \text{ for all } V \in \tau \text{ such that } x \in V\}.$$

A set $E \subset X$ is said to be **nowhere dense** if $\bar{E}^0 = \emptyset$.

Let $A \subset X$, then

$$(A^0)^c = \bigcap\{V^c : V \in \tau \text{ and } V \subset A\} = \bigcap\{C : C \text{ is closed } C \supset A^c\} = \overline{A^c}.$$

Similarly $(\bar{A})^c = (A^c)^0$. Hence the boundary of A may be written as

$$\partial A \equiv \bar{A} \setminus A^0 = \bar{A} \cap (A^0)^c = \bar{A} \cap \overline{A^c}, \tag{1.2}$$

which is to say ∂A consists of the points in both the closure of A and A^c.

Definition 1.9. A set $E \subset X$ is a **neighborhood** of a point $x \in X$ if $x \in E^0 \subset E$.

Notation 1.10. Let $\tau_x = \{V \in \tau : x \in V\}$. So τ_x consists of all the open neighborhoods of x. A collection $\eta \subset \tau_x$ is called a **neighborhood base** at $x \in X$ if for all $V \in \tau_x$ there exists $W \in \eta$ such that $W \subset V$.

Proposition 1.11. *Let $A \subset X$ and $x \in X$.*

1. *If $V \subset_o X$ and $A \cap V = \emptyset$ then $\bar{A} \cap V = \emptyset$.*

2. *$x \in \bar{A}$ iff $V \cap A \neq \emptyset$ for all $V \in \tau_x$.*
3. *$x \in \partial A$ iff $V \cap A \neq \emptyset$ and $V \cap A^c \neq \emptyset$ for all $V \in \tau_x$.*
4. *$\bar{A} = A \cup \mathrm{acc}(A)$.*

Proof. 1. Since $A \cap V = \emptyset$, $A \subset V^c$ and since V^c is closed, $\bar{A} \subset V^c$. That is to say $\bar{A} \cap V = \emptyset$.

2. We will prove $x \notin \bar{A}$ iff there exists $V \in \tau_x$ such that $V \cap A = \emptyset$. If $x \notin \bar{A}$ then $V = \overline{A}^c \in \tau_x$ and $V \cap A \subset V \cap \bar{A} = \emptyset$. Conversely if there exists $V \in \tau_x$ such that $V \cap A = \emptyset$ then by 1. $\bar{A} \cap V = \emptyset$. The third assertion easily follows from the second and Eq. (1.2). Item 4. is an easy consequence of the definition of $\mathrm{acc}(A)$ and item 2. ∎

Lemma 1.12. *Let $A \subset Y \subset X$, let $\bar{A}^Y$ denote the closure of A in Y with its relative topology and $\bar{A} = \bar{A}^X$ be the closure of A in X, then $\bar{A}^Y = \bar{A}^X \cap Y$.*

Proof. Let $x \in Y$ then $x \in \bar{A}^Y$ iff for all $V \in \tau_x^Y$, $V \cap A \neq \emptyset$. This happens iff for all $U \in \tau_x^X$, $U \cap Y \cap A = U \cap A \neq \emptyset$ which happens iff $x \in \bar{A}^X$. That is to say $\bar{A}^Y = \bar{A}^X \cap Y$.

An alternative proof may be given as follows:

$$\begin{aligned}\bar{A}^Y &= \cap\{B \sqsubset Y : A \subset B\} = \cap\{C \cap Y : A \subset C \sqsubset X\} \\ &= Y \cap (\cap\{C : A \subset C \sqsubset X\}) = Y \cap \bar{A}^X\end{aligned}$$

wherein we have made use of the comments in Item (2) of Example 1.2. ∎

Definition 1.13. Let $\{x_n\} \subset X$, we say $x_n \to x$ as $n \to \infty$ or $\lim_{n\to\infty} x_n = x$ if for all $V \in \tau_x$, $x_n \in V$ almost always (abbreviated a.a.) by which we mean that $\{n : x_n \notin V\}$ is finite.

Remark 1.14. (1) If $\tau = \{X, \emptyset\}$ then given $\{x_n\} \subset X$, $x_n \to x$ for all $x \in X$, i.e. **all** sequences converge to **all** points $x \in X$.

(2)We say that a topology τ on X is **Hausdorff** or T_2 if for all $x \neq y \in X$ there exists $V \in \tau_x$ and $W \in \tau_y$ such that $V \cap W = \emptyset$. When τ is Hausdorff the limits of convergent sequences are unique. Indeed if $x_n \to x \in X$ and $y \neq x$ we may choose $V \in \tau_x$ and $W \in \tau_y$ such that $V \cap W = \emptyset$. Then $x_n \in V$ a.a. implies $x_n \notin W$ for all but a finite number of n and hence $x_n \nrightarrow y$.

Definition 1.15. Let (X, τ) be a topological space. We say that X is **first countable** iff every point $x \in X$ has a countable neighborhood base and we say that X is **second countable** iff there exists a countable base for τ.

Example 1.16. Every metric space is first countable.

When τ is first countable, we may formulate many topological notions in terms of sequences. The next Lemma is one such example.

Lemma 1.17. *Suppose there exists $\{x_n\}_{n=1}^\infty \subset A$ such that $\lim_{n\to\infty} x_n = x$, then $x \in \bar{A}$. Conversely if (X, τ) is a first countable space (like a metric space) then if $x \in \bar{A}$ there exists $\{x_n\}_{n=1}^\infty \subset A$ such that $\lim_{n\to\infty} x_n = x$.*

Proof. As we have already seen $x \in \bar{A}$ iff $V \cap A \neq \emptyset$ for all $V \in \tau_x$. If there exists $\{x_n\}_{n=1}^\infty \subset A$ such that $\lim_{n\to\infty} x_n = x$, then for all $V \in \tau_x$ we have $x_n \in V$ a.a. so that $V \cap A \neq \emptyset$. This shows that $x \in \bar{A}$.

For the converse we now assume that (X, τ) is first countable and that $\{V_n\}_{n=1}^\infty$ is a countable neighborhood base at x such that $V_1 \supset V_2 \supset V_3 \supset \ldots$. Then if $x \in \bar{A}$, there exists $x_n \in V_n \cap A$ for all n. It is now easily seen that $x_n \to x$ as $n \to \infty$. ∎

Example 1.18. Let (X, ρ) be a metric space and $\tau = \tau_\rho$ be the induced topology on X. Then $\overline{B_x(\epsilon)} \subset C_x(\epsilon)$. It is not generally true that $\overline{B_x(\epsilon)} = C_x(\epsilon)$. For example let $X = \{1, 2\}$ and $\rho(1, 2) = 1$, then $B_1(1) = \{1\}$, $\overline{B_1(1)} = \{1\}$ while $C_1(1) = X$. Another counter example is to take

$$X = \left\{(x, y) \in \mathbb{R}^2 : x = 0 \text{ or } x = 1\right\}$$

with the usually Euclidean metric coming from the plane. Then

$$\begin{aligned} B_{(0,0)}(1) &= \left\{(0, y) \in \mathbb{R}^2 : |y| < 1\right\}, \\ \overline{B_{(0,0)}(1)} &= \left\{(0, y) \in \mathbb{R}^2 : |y| \le 1\right\}, \text{ while} \\ C_{(0,0)}(1) &= \overline{B_{(0,0)}(1)} \cup \{(0, 1)\}. \end{aligned}$$

In spite of the above examples, Lemmas 1.19 and 1.40 below shows that for certain metric spaces of interest it is true that $\overline{B_x(\epsilon)} = C_x(\epsilon)$.

Lemma 1.19. *Suppose that $(X, |\cdot|)$ is a normed vector space and ρ is the metric on X defined by $\rho(x, y) = |x - y|$. Then*

$$\begin{aligned} \overline{B_x(\epsilon)} &= C_x(\epsilon) \text{ and} \\ \partial B_x(\epsilon) &= \{y \in X : \rho(x, y) = \epsilon\}. \end{aligned}$$

Proof. We must show that $C := C_x(\epsilon) \subset \overline{B_x(\epsilon)} =: \bar{B}$. For $y \in C$, let $v = y - x$, then

$$|v| = |y - x| = \rho(x, y) \le \epsilon.$$

Let $\alpha_n = 1 - 1/n$ so that $\alpha_n \uparrow 1$ as $n \to \infty$. Let $y_n = x + \alpha_n v$, then $\rho(x, y_n) = \alpha_n \rho(x, y) < \epsilon$, so that $y_n \in B_x(\epsilon)$ and $\rho(y, y_n) = 1 - \alpha_n \to 0$ as $n \to \infty$. This shows that $y_n \to y$ as $n \to \infty$ and hence that $y \in \bar{B}$. ■

Definition 1.20. Let $f : X \to Y$ be a function between two topological spaces. Then f is **continuous** if $f^{-1}(V)$ is open in X for all V open in Y or equivalently $f^{-1}(C)$ is closed in X for all C closed in Y. We also say that f is **continuous at** $x \in X$ if for all W open in Y such that $f(x) \in W$ there exists $V \in \tau_x$ such that $f(V) \subset W$, i.e. $x \in f^{-1}(W)^0$ for all W open in Y such that $f(x) \in W$.

Definition 1.21. A map $f : X \to Y$ between topological spaces is called a **homeomorphism** provided that f is bijective, f is continuous and $f^{-1} : Y \to X$ is continuous. If there exists $f : X \to Y$ which is a homeomorphism, we say that X and Y are homeomorphic. (As topological spaces X and Y are essentially the same.)

Lemma 1.22. *Let $f : X \to Y$ be a function between two topological spaces. Then*

1. *f is continuous iff f is continuous at x for all $x \in X$*
2. *If $\mathcal{E} \subset \mathcal{P}(Y)$ is a subbase for the topology on Y then f is continuous iff $f^{-1}(V)$ is open in X for all $V \in \mathcal{E}$.*
3. *If $g : Y \to Z$ is another continuous function then $g \circ f$ is also continuous.*

Proof.

1. $(\Rightarrow)$ Let $V \subset_o Y$. Then for all $x \in f^{-1}(V)$ there exists $W_x \subset_o X$ such that $f(W_x) \subset V$. and therefore

$$f^{-1}(V) = \bigcup_{x \in f^{-1}(V)} W_x \subset_o X.$$

($\Leftarrow$) Given $x \in X$ and $V \subset_o Y$ such that $f(x) \in V$ take $W \equiv f^{-1}(V) \subset_o X$ then $x \in W$ and $f(W) \subset V$.

2. ($\Rightarrow$) Clear. ($\Leftarrow$) Recall that $f^{-1}(\tau(\mathcal{E})) = \tau(f^{-1}(\mathcal{E}))$ where $f^{-1}(\mathcal{E}) = \{f^{-1}(V) : V \in \mathcal{E}\}$. Thus if $f^{-1}(\mathcal{E})$ consists of open sets then $f^{-1}(\tau(\mathcal{E})) = \tau(f^{-1}(\mathcal{E})) \subset \tau_X$, i.e. $f^{-1}(\tau_Y) \subset \tau_X$.
3. Suppose that $V \subset_o Z$, then $(g \circ f)^{-1}(V) = f^{-1}(g^{-1}(V)) \subset_o X$ since $g^{-1}(V) \subset_o Y$.

■

Lemma 1.23. *Suppose that $f : X \to Y$ is a map between topological spaces. Then the following are equivalent:*

1. *f is continuous.*
2. *$f(\bar{A}) \subset \overline{f(A)}$ for all $A \subset X$*
3. *$\overline{f^{-1}(B)} \subset f^{-1}(\bar{B})$ for all $B \sqsubset X$.*

Proof. If f is continuous, then $f^{-1}\left(\overline{f(A)}\right)$ is closed and since $A \subset f^{-1}(f(A)) \subset f^{-1}\left(\overline{f(A)}\right)$ it follows that $\bar{A} \subset f^{-1}\left(\overline{f(A)}\right)$. From this equation we learn that $f(\bar{A}) \subset \overline{f(A)}$ so that (1) implies (2) Now assume (2), then for $B \subset Y$ (taking $A = f^{-1}(\bar{B})$) we have

$$f(\overline{f^{-1}(B)}) \subset f(\overline{f^{-1}(\bar{B})}) \subset \overline{f(f^{-1}(\bar{B}))} \subset \bar{B}$$

and therefore

$$\overline{f^{-1}(B)} \subset f^{-1}(\bar{B}). \tag{1.3}$$

This shows that (2) implies (3) Finally if Eq. (1.3) holds for all B, then when B is closed this shows that

$$\overline{f^{-1}(B)} \subset f^{-1}(\bar{B}) = f^{-1}(B) \subset \overline{f^{-1}(B)}$$

which shows that

$$f^{-1}(B) = \overline{f^{-1}(B)}.$$

Therefore $f^{-1}(B)$ is closed whenever B is closed which implies that f is continuous.

■

Exercise 1.24. Suppose that A and B are closed subsets of a topological space X and $f \in C(A)$ and $g \in C(B)$ such that $f = g$ on $A \cap B$. Show

$$F(x) = \begin{cases} f(x) & \text{if} \quad x \in A \\ g(x) & \text{if} \quad x \in B \end{cases}$$

defines a continuous function on $A \cup B$.

Solution. Let C be a closed set $\mathbb{R}$ or $\mathbb{C}$ depending on the context, then

$$F^{-1}(C) = f^{-1}(C) \cup g^{-1}(C).$$

By the continuity of f and g, $f^{-1}(C)$ and $g^{-1}(C)$ are relatively closed sets in A and B respectively and since A and B are closed, it follows that $f^{-1}(C)$ and $g^{-1}(C)$ are closed in X as well. Therefore $F^{-1}(C)$ is closed in X and hence closed in $A \cup B$, showing the F is continuous. ■

Proposition 1.25. *If $f : X \to Y$ is continuous at $x \in X$ and $\lim_{n\to\infty} x_n = x \in X$, then $\lim_{n\to\infty} f(x_n) = f(x) \in Y$. Moreover, if there exists a countable neighborhood base $\eta \subset \tau_x$, then f is continuous at x iff $\lim_{n\to\infty} f(x_n) = f(x)$ for all sequences $\{x_n\} \subset X$ which are convergent to x.*

Proof. If $f : X \to Y$ is continuous and $W \in \tau_{f(x)} \subset \mathcal{P}(Y)$, then there exists $V \in \tau_x$ such that $f(V) \subset W$. Since $x_n \to x$, $x_n \in V$ a.a. and therefore $f(x_n) \in f(V) \subset W$ a.a., i.e. $f(x_n) \to f(x)$ as $n \to \infty$.

Conversely suppose that $\eta \equiv \{W_n\}_{n=1}^{\infty} \subset \tau_x$ is a countable base and $\lim_{n\to\infty} f(x_n) = f(x)$ for all sequences $\{x_n\} \subset X$. By replacing W_n by $W_1 \cap \cdots \cap W_n$ if necessary, we may assume that $\{W_n\}_{n=1}^{\infty}$ is a decreasing sequence of sets. If f were **not** continuous at x then there exists $V \in \tau_{f(x)}$ such that $x \notin f^{-1}(V)^0$. Therefore, W_n is not a subset of $f^{-1}(V)$ for all n. Hence for each n, we may choose $x_n \in W_n \setminus f^{-1}(V)$. This sequence then has the property that $x_n \to x$ as $n \to \infty$ while $f(x_n) \notin V$ for all n and hence $\lim_{n\to\infty} f(x_n) \neq f(x)$. ■

1.1. Connectedness.

Definition 1.26. (X, τ) is **disconnected** if there exists non-empty open sets U and V of X such that $U \cap V = \emptyset$ and $X = U \cup V$. We say $\{U, V\}$ is a **disconnection** of X. The topological space (X, τ) is called **connected** if it is not disconnected, i.e. if there are no disconnection of X. If $A \subset X$ we say A is connected iff (A, τ_A) is connected where τ_A is the relative topology on A. Explicitly, A is disconnected in (X, τ) iff there exists $U, V \in \tau$ such that $U \cap A \neq \emptyset$, $U \cap A \neq \emptyset$, $A \cap U \cap V = \emptyset$ and $A \subset U \cup V$.

The reader should check that the following statement is an equivalent definition of connectivity. A topological space (X, τ) is connected iff the only sets $A \subset X$ which are both open and closed are the sets X and $\emptyset$.

Remark 1.27. Let $A \subset Y \subset X$. Then A is connected in X iff A is connected in Y.

Proof. Since

$$\tau_A \equiv \{V \cap A : V \subset X\} = \{V \cap A \cap Y : V \subset X\} = \{U \cap A : U \subset_o Y\},$$

the relative topology on A inherited from X is the same as the relative topology on A inherited from Y. Since connectivity is a statement about the relative topologies on A, A is connected in X iff A is connected in Y. ■

The following elementary but important lemma is left as an exercise to the reader.

Lemma 1.28. *Suppose that $f : X \to Y$ is a continuous map between topological spaces. Then $f(X) \subset Y$ is connected if X is connected.*

Proposition 1.29. *Let (X, τ) be a topological space.*

1. *If $B \subset X$ is a connected set and X is the disjoint union of two open sets U and V, then either $B \subset U$ or $B \subset V$.*
2. *a) If $A \subset X$ is connected, then $\bar{A}$ is connected.*
   b) More generally, if A is connected and $B \subset \mathrm{acc}(A)$, then $A \cup B$ is connected as well.
3. *If $\{E_\alpha\}_{\alpha\in A}$ is a collection of connected sets such that $\bigcap_{\alpha\in A} E_\alpha \neq \emptyset$, then $Y := \bigcup_{\alpha\in A} E_\alpha$ is connected as well.*
4. *Suppose $A, B \subset X$ are non-empty connected subsets of X such that $\bar{A} \cap B \neq \emptyset$, then $A \cup B$ is connected in X.*

5. *Every point $x \in X$ is contained in a unique maximal connected subset C_x of X and this subset is closed. The set C_x is called the* ***connected component*** *of x.*

Proof.

1. Since B is the disjoint union of the relatively open sets $B \cap U$ and $B \cap V$, we must have $B \cap U = B$ or $B \cap V = B$ for otherwise $\{B \cap U, B \cap V\}$ would be a disconnection of B.
2. a. Let $Y = \bar{A}$ equipped with the relative topology from X. Suppose that $U, V \subset_o Y$ form a disconnection of $Y = \bar{A}$. Then by 1. either $A \subset U$ or $A \subset V$. Say that $A \subset U$. Since U is both open an closed in Y, it follows that $Y = \bar{A} \subset U$. Therefore $V = \emptyset$ and we have a contradiction to the assumption that $\{U, V\}$ is a disconnection of $Y = \bar{A}$. Hence we must conclude that $Y = \bar{A}$ is connected as well.

   b. Now let $Y = A \cup B$ with $B \subset \mathrm{acc}(A)$, then
$$\bar{A}^Y = \bar{A} \cap Y = (A \cup \mathrm{acc}(A)) \cap Y = A \cup B.$$
   Because A is connected in Y, by (2) b. $Y = A \cup B = \bar{A}^Y$ is also connected.
3. Let $Y := \bigcup_{\alpha \in A} E_\alpha$. By Remark 1.27, we know that E_α is connected in Y for each $\alpha \in A$. If $\{U, V\}$ were a disconnection of Y, By item (1), either $E_\alpha \subset U$ or $E_\alpha \subset V$ for all α. Let $\Lambda = \{\alpha \in A : E_\alpha \subset U\}$ then $U = \cup_{\alpha \in \Lambda} E_\alpha$ and $V = \cup_{\alpha \in A \setminus \Lambda} E_\alpha$. (Notice that neither Λ or $A \setminus \Lambda$ can be empty since U and V are not empty.) Since
$$\emptyset = U \cap V = \bigcup\nolimits_{\alpha \in \Lambda, \beta \in \Lambda^c} (E_\alpha \cap E_\beta) \supset \bigcap_{\alpha \in A} E_\alpha \neq \emptyset.$$
   we have reached a contradiction and hence no such disconnection exists.
4. Let $Y = A \cup B$ and, for sake of contraction, suppose that Y were disconnected. Since A and B are connected, it follows from item (1) that $\{A, B\}$ is the only possible disconnection of Y. In particular it follows that $\bar{A}^Y = A$. On the other hand we have seen that $\bar{A}^Y = \bar{A} \cap Y = \bar{A} \cap B$. Therefore $A = \bar{A} \cap B$. But since $\{A, B\}$ is a disconnection, $\emptyset = A \cap B = \bar{A} \cap B \neq \emptyset$ which is a contradiction.
5. Let $\mathcal{C}$ denote the collection of connected subsets $C \subset X$ such that $x \in C$. Then by item 3., the set $C_x := \cup \mathcal{C}$ is also a connected subset of X which contains x and clearly this is the unique maximal connected set containing x. Since $\bar{C}_x$ is also connected by item (2) and C_x is maximal, $C_x = \bar{C}_x$, i.e. C_x is closed.

■

Example 1.30. The connected subsets of $\mathbb{R}$ are intervals.

Proof. Suppose that $A \subset \mathbb{R}$ is a connected subset and that $a, b \in A$ with $a < b$. If there exists $c \in (a, b)$ such that $c \notin A$, then $U := (-\infty, c) \cap A$ and $V := (c, \infty) \cap A$ would form a disconnection of A. Hence $(a, b) \subset A$. Let $\alpha := \inf(A)$ and $\beta := \sup(A)$ and choose $\alpha_n, \beta_n \in A$ such that $\alpha_n < \beta_n$ and $\alpha_n \downarrow \alpha$ and $\beta_n \uparrow \beta$ as $n \to \infty$. By what we have just shown, $(\alpha_n, \beta_n) \subset A$ for all n and hence $(\alpha, \beta) = \cup_{n=1}^\infty (\alpha_n, \beta_n) \subset A$. From this it follows that $A = (\alpha, \beta)$, $[\alpha, \beta)$, $(\alpha, \beta]$ or $[\alpha, \beta]$, i.e. A is an interval.

Conversely[1] suppose that A is an interval, and for sake of contradiction, suppose that $\{U, V\}$ is a disconnection of A with $a \in U$, $b \in V$. After relabeling U and V if necessary we may assume that $a < b$. Since A is an interval $[a,b] \subset A$. Let $p = \sup([a,b] \cap U)$, then because U and V are open, $a < p < b$. Now p can not be in U for otherwise $\sup([a,b] \cap U) > p$ and p can not be in V for otherwise $p < \sup([a,b] \cap U)$. From this it follows that $p \notin U \cup V$ and hence $A \neq U \cup V$ contradicting the assumption that $\{U, V\}$ is a disconnection. ∎

1.2. **Separable Spaces.**

Definition 1.31. A set $A \subset X$ is said to be **dense** if $\bar{A} = X$. A topological space X is **separable** if there exists a countable dense subset $A \subset X$.

Example 1.32. he following are example of countable dense sets.

1. The rational number $\mathbb{Q}$ are dense in $\mathbb{R}$ equipped with the usual topology.
2. More generally, $\mathbb{Q}^d$ is a countable dense subset of $\mathbb{R}^d$ for any $d \in \mathbb{N}$.
3. If (X, ρ) is a metric space which is separable then every subset $Y \subset X$ is also separable in the induced topology.

Proof. (3) Let $A \subset X$ be a countable dense set and let $A = \{x_n\}_{n=1}^\infty$. Set $\rho(x, Y) = \inf\{\rho(x, y) : y \in Y\}$ the distance from x to Y. Recall that $\rho(\cdot, Y) : X \to [0, \infty)$ is continuous. Indeed, if $x, z \in X$ and $y \in Y$ then

$$\rho(x, Y) \leq \rho(x, y) \leq \rho(x, z) + \rho(z, y). \tag{1.4}$$

Taking the inf over $y \in Y$ in (1.4) implies

$$\rho(x, Y) \leq \rho(x, z) + \rho(z, Y) \text{ or}$$
$$\rho(x, Y) - \rho(z, Y) \leq \rho(x, z). \tag{1.5}$$

Equation (1.5) along with Eq. (1.5) with x and z interchanged shows that

$$|\rho(x, Y) - \rho(z, Y)| \leq \rho(x, z)$$

which certainly implies that $\rho(\cdot, Y)$ is continuous.

Let $\epsilon_n = \rho(x_n, Y) \geq 0$ and for each n let $y_n \in B_{x_n}(\frac{1}{n}) \cap Y$ if $\epsilon_n = 0$ otherwise choose $y_n \in B_{x_n}(2\epsilon_n) \cap Y$. Then if $y \in Y$ and $\epsilon > 0$ we may choose $n \in \mathbb{N}$ such that $\rho(y, x_n) \leq \epsilon_n < \epsilon/3$ and $\frac{1}{n} < \epsilon/3$. If $\epsilon_n > 0$, $\rho(y_n, x_n) \leq 2\epsilon_n < 2\epsilon/3$ and if $\epsilon_n = 0$, $\rho(y_n, x_n) < \epsilon/3$ and therefore

$$\rho(y, y_n) \leq \rho(y, x_n) + \rho(x_n, y_n) < \epsilon.$$

This shows that $B \equiv \{y_n\}_{n=1}^\infty$ is a countable dense subset of Y. ∎

Proposition 1.33. *Every separable metric space is second countable.*

---

[1](An Old proof.) Conversely suppose that A is an interval which for sake of contradiction is not connected. Let $U, V \subset_o \mathbb{R}$ such that $A \cap U$ and $A \cap V$ is a disconnection of A. Let $a \in A \cap U$ and $b \in A \cap V$ and notice that $a \neq b$ because $A \cap U \cap V = \emptyset$. With out loss of generality we may assume that $a < b$. Since A is an interval, we know that $[a, b] \subset A$.

Now let $p \equiv \sup\{[a, b] \cap U\}$. We will now finish the proof by showing that $p \in A \cap U \cap V$ which will contradict the assumption that $A \cap U \cap V = \emptyset$. By Lemma 1.17 $p \in \overline{U \cap A}^{\mathbb{R}}$ and by (1) of Lemma 1.12, $\overline{U \cap A}^{\mathbb{R}} = \overline{U \cap A}^{A}$. Since $U \cap A$ is closed in A, it follows that $p \in \overline{U \cap A}^{A} = U \cap A$. From this it follows that $p = b$ for otherwise $\sup\{[a, b] \cap U\} > p$. But then $p = b \in V \cap A$ and hence $p \in A \cap U \cap V$.

Proof. Let $\{x_n\}_{n=1}^{\infty}$ be a countable dense subset of X. Let $\mathcal{E} \equiv \{X, \emptyset\} \bigcup_{m,n=1}^{\infty} \{B_{x_n}(r_m)\} \subset \tau_\rho$, where $\{r_m\}_{m=1}^{\infty}$ is dense in $(0, \infty)$. Then $\mathcal{E}$ is a countable base for τ_ρ. To see this let $V \subset X$ be open and $x \in V$. Choose $\epsilon > 0$ such that $B_x(\epsilon) \subset V$ and then choose $x_n \in B_x(\epsilon/3)$. Choose r_m near $\epsilon/3$ such that $\rho(x, x_n) < r_m < \epsilon/3$ so that $x \in B_{x_n}(r_m) \subset V$. This shows $V = \bigcup \{B_{x_n}(r_m) : B_{x_n}(r_m) \subset V\}$. ■

1.3. Bounded Functions as Metric Spaces.

Definition 1.34. Let $B(X, \mathbb{C})$ denote the bounded functions in $\mathbb{C}^X$ and $B(X, \mathbb{R})$ denote the bounded functions in $\mathbb{R}^X$. We may equip these spaces with the supremum (or uniform) norm:

$$\|f\|_\infty \equiv \|f\|_u \equiv \sup\{|f(x)| : x \in X\}$$

and let ρ be the associated "uniform metric,"

$$\rho(f, g) = \|f - g\|_\infty \text{ for all } f, g \in B(X)$$

where $B(X)$ is either $B(X, \mathbb{C})$. or $B(X, \mathbb{R})$.

Lemma 1.35. *The metric space $(B(X), \rho)$ as defined above is complete.*

Proof. It is easily checked that ρ is a metric. For completeness, let $\{f_n\}_{n=1}^{\infty}$ be a Cauchy sequence in $B(X)$ then $\{f_n(x)\}_{n=1}^{\infty} \subset \mathbb{C}$ is Cauchy for all $x \in X$. Hence by completeness of $\mathbb{C}$ or $\mathbb{R}$, $f(x) \equiv \lim_{n\to\infty} f_n(x)$ exists for all $x \in X$. Noting that

$$\begin{aligned} |f(x) - f_n(x)| &\le |f(x) - f_m(x)| + |f_m(x) - f_n(x)| \\ &\le |f(x) - f_m(x)| + \|f_m - f_n\|_\infty, \end{aligned}$$

we have

$$|f(x) - f_n(x)| \le \overline{\lim_{m\to\infty}} |f(x) - f_m(x)| + \overline{\lim_{m\to\infty}} \|f_m - f_n\|_\infty = \overline{\lim_{m\to\infty}} \|f_m - f_n\|_\infty.$$

Taking the sup of the left member of this equation over x and the letting $n \to \infty$ gives

$$\|f - f_n\|_\infty \le \overline{\lim_{m\to\infty}} \|f_m - f_n\|_\infty \to 0 \text{ as } n \to \infty$$

wherein we have used $\{f_n\}$ is Cauchy in $B(X)$. Therefore, $f_n \to f$ uniformly and

$$\|f\|_\infty \le \|f - f_n\|_\infty + \|f_n\|_\infty < \infty$$

so that $f \in B(X)$. ■

Notation 1.36. Given a topological space X, let $C(X)$ denote the continuous functions from $X \to \mathbb{R}$ or $X \to \mathbb{C}$. Also let $BC(X) = C(X) \cap B(X)$, i.e. the space of bounded continuous functions on X.

Proposition 1.37. *The space $BC(X)$ is a closed subspace of $(B(X), \rho)$, where ρ is the sup-norm metric.*

Proof. We will prove this by showing that $\overline{BC(X)} \subset BC(X)$. So let $f \in \overline{BC(X)}$ and choose $f_n \in BC(X)$ such that $f_n \xrightarrow{\rho} f$ as $n \to \infty$, i.e. $\rho(f, f_n) \to 0$. We will shows that $f \in C(X)$ and hence $f \in BC(X)$ showing $\overline{BC(X)} \subset BC(X)$.

Let $x \in X$ and $\epsilon > 0$ be given and choose $N \in \mathbb{N}$ so large that $\|f - f_n\|_\infty \le \epsilon/3$ for all $n \ge N$. Then for any in X we have

$$\begin{aligned}|f(x) - f(y)| &\le |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f(y) - f_n(y)| \\ &\le 2\|f - f_n\|_\infty + |f_n(x) - f_n(y)| \\ &\le \frac{2}{3}\epsilon + |f_n(x) - f_n(y)|.\end{aligned}$$

Choose V open in X such that $x \in V$ and $f_n(V) \subset B_{f_n(x)}(\epsilon/3)$, i.e. so that $|f_n(x) - f_n(y)| < \epsilon/3$ for all $y \in V$. Then

$$|f(x) - f(y)| < \frac{2\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \text{ for all } y \in V.$$

Since $\epsilon > 0$ is arbitrary, we shown that f is continuous at x. ■

Remark 1.38. Notice that $BC(X)$ is complete since it is easily verified that a closed subset of a complete metric space is also a complete metric space.

1.3.1. ***Application: An ODE Existence Theorem.*** In this section suppose that $\mathbb{R}$ and $\mathbb{R}^d$ are equipped with the standard Euclidean topologies that $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is a continuous function. Further assume there is a constant $K < \infty$ such that

$$|f(t, x) - f(t, y)| \le K|x - y| \text{ for all } t \in \mathbb{R} \text{ and } x, y \in \mathbb{R}^d.$$

We wish to find a C^1 function $y : \mathbb{R} \to \mathbb{R}^d$ solving the ordinary differential equation,

$$\dot{y}(t) = f(t, y(t)) \text{ with } y(0) = y_0 \tag{1.6}$$

where y_0 is a given point $\mathbb{R}^d$. It is easily checked that solving Eq. (1.6) is equivalent to finding a continuous function $y : \mathbb{R} \to \mathbb{R}^d$ such that

$$y(t) = y_0 + \int_0^t f(\tau, y(\tau))d\tau. \tag{1.7}$$

In general no such solution exists. However, if we restrict t in some neighborhood of 0, and only require y to be defined on this neighborhood, solutions do exist. In particular we have the following theorem.

Theorem 1.39. *Keeping the notation and assumptions above. For all $0 < T < K^{-1}$, there exists a unique solution $y : (-T, T) \to \mathbb{R}^d$ to Eq. (1.6).*

Proof. Let X denote the complete metric space of bounded functions from $(-T, T) \to \mathbb{R}^d$ equipped with the uniform metric:

$$\rho(x, y) = \sup_{t \in (-T,T)} |x(t) - y(t)|.$$

For $x \in X$, let

$$S(x)(t) = y_0 + \int_0^t f(\tau, x(\tau))d\tau$$

and notice that for $x, y \in X$ and $t \in (-T, T)$ that

$$\begin{aligned}|S(x)(t) - S(y)(t)| &= \left|\int_0^t [f(\tau, x(\tau)) - f(\tau, y(\tau))]d\tau\right| \\ &\le \left|\int_0^t K|x(\tau) - y(\tau)|d\tau\right| \\ &\le TK\|x - y\|_\infty.\end{aligned}$$

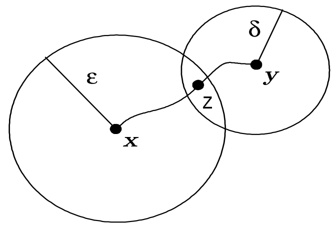


FIGURE 2. An almost length minimizing curve joining x to y.

Taking the sup of this equation over t shows that

$$\|S(x) - S(y)\|_\infty \leq \alpha \|x - y\|_\infty \tag{1.8}$$

where $\alpha = KT < 1$. This equation with $y \equiv 0$ implies that

$$\|S(x)\|_\infty \leq \|S(0)\|_\infty + \|S(x) - S(0)\|_\infty \leq \|S(0)\|_\infty + \alpha \|x\|_\infty < \infty.$$

Therefore $S : X \to X$ is a contraction and therefore by the contraction mapping principle (see your homework) has a unique fixed point, y. This is to say

$$y(t) = S(y)(t) = y_0 + \int_0^t f(\tau), y(\tau))d\tau \text{ for all } t \in (-T, T).$$

■

1.4. **Appendix on Riemannian Metrics.** This subsection is not completely self contained and may safely be skipped.

Lemma 1.40. *Suppose that X is a Riemannian (or sub-Riemannian) manifold and ρ is the metric on X defined by*

$$\rho(x, y) = \inf \{\ell(\sigma) : \sigma(0) = x \text{ and } \sigma(1) = y\}$$

where $\ell(\sigma)$ is the length of the curve σ. We define $\ell(\sigma) = \infty$ if σ is not piecewise smooth.

Then

$$\overline{B_x(\epsilon)} = C_x(\epsilon) \text{ and}$$
$$\partial B_x(\epsilon) = \{y \in X : \rho(x, y) = \epsilon\}.$$

Proof. Let $C := C_x(\epsilon) \subset \overline{B_x(\epsilon)} =: \bar{B}$. We will show that $C \subset \bar{B}$ by showing $\bar{B}^c \subset C^c$. Suppose that $y \in \bar{B}^c$ and choose $\delta > 0$ such that $B_y(\delta) \cap \bar{B} = \emptyset$. In particular this implies that

$$B_y(\delta) \cap B_x(\epsilon) = \emptyset.$$

We will finish the proof by showing that $\rho(x, y) \geq \epsilon + \delta > \epsilon$ and hence that $y \in C^c$. This will be accomplished by showing: if $\rho(x, y) < \epsilon + \delta$ then $B_y(\delta) \cap B_x(\epsilon) \neq \emptyset$.

If $\rho(x, y) < \max(\epsilon, \delta)$ then either $x \in B_y(\delta)$ or $y \in B_x(\epsilon)$. In either case $B_y(\delta) \cap B_x(\epsilon) \neq \emptyset$. Hence we may assume that $\max(\epsilon, \delta) \leq \rho(x, y) < \epsilon + \delta$. Let $\alpha > 0$ be a number such that

$$\max(\epsilon, \delta) \leq \rho(x, y) < \alpha < \epsilon + \delta$$

and choose a curve σ from x to y such that $\ell(\sigma) < \alpha$. Also choose $0 < \delta' < \delta$ such that $0 < \alpha - \delta' < \epsilon$ which can be done since $\alpha - \delta < \epsilon$. Let $k(t) = \rho(y, \sigma(t))$ a continuous function on $[0,1]$ and therefore $k([0,1]) \subset \mathbb{R}$ is a connected set which contains 0 and $\rho(x,y)$. Therefore there exists $t_0 \in [0,1]$ such that $\rho(y,\sigma(t_0)) = k(t_0) = \delta'$. Let $z = \sigma(t_0) \in B_y(\delta)$ then

$$\rho(x,z) \le \ell(\sigma|_{[0,t_0]}) = \ell(\sigma) - \ell(\sigma|_{[t_0,1]}) < \alpha - \rho(z,y) = \alpha - \delta' < \epsilon$$

and therefore $z \in B_x(\epsilon) \cap B_x(\delta) \neq \emptyset$. ■

Remark 1.41. Suppose again that X is a Riemannian (or sub-Riemannian) manifold and

$$\rho(x,y) = \inf\{\ell(\sigma) : \sigma(0) = x \text{ and } \sigma(1) = y\}.$$

Let σ be a curve from x to y and let $\epsilon = \ell(\sigma) - \rho(x,y)$. Then for all $0 \le u < v \le 1$,

$$\rho(\sigma(u),\sigma(v)) \le \ell(\sigma|_{[u,v]}) + \epsilon.$$

So if σ is within ϵ of a length minimizing curve from x to y that $\sigma|_{[u,v]}$ is within ϵ of a length minimizing curve from $\sigma(u)$ to $\sigma(v)$. In particular if $\rho(x,y) = \ell(\sigma)$ then $\rho(\sigma(u),\sigma(v)) = \ell(\sigma|_{[u,v]})$ for all $0 \le u < v \le 1$, i.e. if σ is a length minimizing curve from x to y that $\sigma|_{[u,v]}$ is a length minimizing curve from $\sigma(u)$ to $\sigma(v)$.

To prove these assertions notice that

$$\begin{aligned}\rho(x,y) + \epsilon = \ell(\sigma) &= \ell(\sigma|_{[0,u]}) + \ell(\sigma|_{[u,v]}) + \ell(\sigma|_{[v,1]}) \\ &\ge \rho(x,\sigma(u)) + \ell(\sigma|_{[u,v]}) + \rho(\sigma(v),y)\end{aligned}$$

and therefore

$$\begin{aligned}\ell(\sigma|_{[u,v]}) &\le \rho(x,y) + \epsilon - \rho(x,\sigma(u)) - \rho(\sigma(v),y) \\ &\le \rho(\sigma(u),\sigma(v)) + \epsilon.\end{aligned}$$

2. Normal Spaces

Definition 2.1. A topological space (X,τ) is said to be **normal** or T_4 if:

1. X is Hausdorff and
2. if for any two closed disjoint subsets $A, B \subset X$ there exists disjoint open sets $V, W \subset X$ such that $A \subset V$ and $B \subset W$.

Remark 2.2. Suppose that X is normal and $C \subset W \subset_0 X$ and C is closed. Then there exists $U \subset_o X$ such that

$$C \subset U \subset \bar{U} \subset W.$$

Indeed, Since W^c is closed and $C \cap W^c = \emptyset$, there exists disjoint open sets U and V such that $C \subset U$ and $W^c \subset V$. Therefore $C \subset U \subset V^c \subset W$ and since V^c is closed, we may conclude that $C \subset U \subset \bar{U} \subset V^c \subset W$.

The converse of the above remark holds as well. Namely if for all $C \subset W \subset_o X$ with C closed, there exists $U \subset_o X$ such that $C \subset U \subset \bar{U} \subset W$, then X is normal. To prove this, if A and B are disjoint closed set in X, then $A \subset B^c$ and B^c is open, hence there exists $U \subset_o X$ such that

$$A \subset U \subset \bar{U} \subset B^c$$

and by the same token there exists $W \subset_o X$ such that $\bar{U} \subset W \subset \bar{W} \subset B^c$. Taking complements of the last expression implies

$$B \subset \bar{W}^c \subset W^c \subset \bar{U}^c.$$

Let $V = \bar{W}^c$. Then $A \subset U \subset_o X$, $B \subset V \subset_o X$ and $U \cap V \subset U \cap W^c = \emptyset$.

Lemma 2.3 (Urysohn's Lemma for Metric Spaces)**.** *Every metric space, (X, ρ), is normal. Moreover if A and B are disjoint closed subsets of X, then there exists $f \in C(X, [0, 1])$ such that $f = 0$ on A and $f = 1$ on B.*

Proof. Let $\rho_A(x) = \rho(x, A)$ and $\rho_B(x) = \rho(x, B)$ denote the distance from x to A and B respectively. We will show that

$$f(x) = \frac{\rho_A(x)}{\rho_A(x) + \rho_B(x)}$$

is the desired function. Since B^c is open, if $x \notin B$ there exists an $\epsilon > 0$ such that $B_x(\epsilon) \subset B^c$ and hence $\rho_B(x) \geq \epsilon > 0$. Similarly if $x \notin A$ than $\rho_A(x) > 0$ and therefore because $A^c \cup B^c = X$,

$$\rho_A(x) + \rho_B(x) > 0 \text{ for all } x \in X.$$

Therefore f is well defined and being the composition of continuous functions is continuous so $f \in C(X, [0, 1])$. It is now clear that and $f(x) = 0$ if $x \in A$ and $f(x) = 1$ if $x \in B$.

The open sets, $V = f^{-1}(-\infty, 1/2)$ and $W = f^{-1}(1/2, \infty)$, are disjoint and $A \subset V$ and $B \subset W$. ■

Theorem 2.4 (Urysohn's Lemma for Normal Spaces)**.** *Let X be a normal space. Assume A, B are disjoint closed subsets of X. Then there exists $f \in C(X, [0, 1])$ such that $f = 0$ on A and $f = 1$ on B.*

We will prove this theorem after the next Lemma. The idea of the proof is to define f by its level sets. For motivational purposes, suppose that $f \in C(X, [0, 1])$ such that $f = 0$ on A and $f = 1$ on B. For $r > 0$, let $U_r = \{f < r\}$. Then for $r < s$, $U_r \subset \{f \leq r\} \subset U_s$ and since $\{f \leq r\}$ is closed this implies that

$$A \subset U_r \subset \bar{U}_r \subset \{f \leq r\} \subset U_s \subset B^c$$

for all $0 < r < s \leq 1$. Therefore associated to the function f is the collection open set $\{U_r\}_{r>0}$ with the property that $A \subset U_r \subset \bar{U}_r \subset U_s \subset B^c$ for all $0 < r < s \leq 1$ and $U_r = X$ if $r > 1$. Finally let us notice that we may recover the function f from the sequence $\{U_r\}_{r>0}$ by the formula

$$f(x) = \inf\{r > 0 : x \in U_r\}.$$

Hopefully these remarks will help the reader understand the motivation for the proof of Theorem 2.4. For the remainder of this section let

$$\mathbb{D} \equiv \left\{k 2^{-n} : k = 1, 2, \ldots, 2^{-1}, n = 1, 2, \ldots\right\}$$

be the dyadic rationales in $(0, 1]$.

Lemma 2.5. *Suppose that (X, τ) is normal and A, B are disjoint closed subsets of X. Then there exists $\{U_r\}_{r \in \mathbb{D}} \subset \tau$ such that for all $r < s$ in $\mathbb{D}$,*

$$A \subset U_r \subset \overline{U}_r \subset U_s \subset U_1 = B^c.$$

Proof. Let $U_1 = B^c$. By Remark 2.2 there exists $U_{1/2} \subset_o X$ such that

$$A \subset U_{1/2} \subset \bar{U}_{1/2} \subset U_1.$$

Similarly we may construct $U_{1/2}, U_{3/4} \subset_o X$ such that

$$A \subset U_{1/4} \subset \bar{U}_{1/4} \subset U_{1/2} \subset \bar{U}_{1/2} \subset U_{3/4} \subset \bar{U}_{3/4} \subset U_1$$

and then we may construct $U_{1/8}, U_{3/8}, U_{5/8}, U_{7/8} \subset_o X$ such that

$$\begin{aligned} A &\subset U_{1/8} \subset \bar{U}_{1/8} \subset U_{1/4} \subset \bar{U}_{1/4} \subset U_{3/8} \subset \bar{U}_{3/8} \subset U_{1/2} \\ &\subset \bar{U}_{1/2} \subset U_{5/8} \subset \bar{U}_{5/8} \subset U_{3/4} \subset \bar{U}_{3/4} \subset U_{7/8} \subset \bar{U}_{7/8} \subset U_1. \end{aligned}$$

It is now easy to continue (inductively) in this fashion to construct the desired sequence $\{U_r\}_{r\in\mathbb{D}} \subset \tau$. The details are left to the reader. ■

Proof. (Urysohn's Lemma for Normal Spaces) Let $\{U_r\}_{r\in\mathbb{D}} \subset \tau$ be as in Lemma 2.5, $U_r \equiv X$ if $r > 1$ and define

$$f(x) = \inf\{r \in \mathbb{D} \cup [1, \infty) : x \in U_r\}.$$

Notice that $f(x) \in [0, 1]$ for all $x \in X$, if $x \in A$ then $f(x) = 0$ since $x \in U_r$ for all $r \in \mathbb{D}$ and if $x \in B$, then $x \notin U_r$ for all $r \in \mathbb{D}$ and hence $f(x) = 1$. So it only remains to show f is continuous. We will do this by showing that $\{f < \alpha\}$ and $\{f > \alpha\}$ are open sets for all $\alpha \in \mathbb{R}$. This will shows f is continuous since $\mathcal{E} = \{(\alpha, \infty), (-\infty, \alpha) : \alpha \in \mathbb{R}\}$ is a subbase for the topology on $\mathbb{R}$.

If $x \in X$, then $f(x) < \alpha$ iff there exists $r < \alpha$ such that $x \in U_r$ so that

$$\{f < \alpha\} = \bigcup_{r<\alpha} U_r \subset_o X.$$

Similarly, $f(x) > \alpha$ iff there exists $r > \alpha$ with $x \notin U_r$. Now if $r > \alpha$ and $x \notin U_r$ then for $\alpha < s < r$, $x \notin \bar{U}_s \subset U_r$. Thus we have shown that

$$\{f > \alpha\} = \bigcup_{s>\alpha} \overline{U}_s^c \subset_o X.$$

■

Theorem 2.6 (Tietze Extension Theorem)**.** *Let X be a normal space, $A \sqsubset X$, $-\infty < a < b < \infty$ and $f \in C(A, [a, b])$. Then there exists $F \in C(X, [a, b])$ such that $F|_A = f$.*

Proof. By translating and scaling we may assume $a = 0$ and $b = 1$. We will now construct F by a sequence of approximations. Firstly let $B, C \subset A$ be the closed subsets of X defined by $B = f^{-1}([0, \frac{1}{3}])$ and $C = f^{-1}([\frac{2}{3}, 1])$. By Urysohn's Lemma choose $g_1 \in C(X, [0, \frac{1}{3}])$, such that $g_1 = 0$ on B and $g_1 = \frac{1}{3}$ on C. Let $f_1 = f - g_1|_A$, then $f_1(x) = f(x)$ for $x \in B$, $0 \le f_1(x) \le f(x) \le 1/3$ for $x \in f^{-1}([1/3, 2/3])$ and $0 \le f_1(x) \le 2/3$ for $x \in C$ and thus

$$0 \le f_1 \le 2/3 \text{ on } A.$$

Applying the same construction we may find $g_2 \in C(X, [0, \frac{1}{3}])$ such that $f_2 := \frac{3}{2} f_1 - g_2|_A$ satisfies

$$0 \le f_2 \le 2/3 \text{ on } A.$$

Continuing this way inductively we may find $g_n \in C(X, [0, \frac{1}{3}])$ and $f_n := \frac{3}{2} f_{n-1} - g_n|_A$ such that

$$0 \le f_n \le 2/3 \text{ on } A.$$

Now on A,

$$\begin{aligned} f &= f_1 + g_1 = \frac{2}{3}(f_2 + g_2) + g_1 = \frac{2}{3}f_2 + \frac{2}{3}g_2 + g_1 \\ &= \frac{2}{3}\frac{2}{3}(f_3 + g_3) + \frac{2}{3}g_2 + g_1 = \left(\frac{2}{3}\right)^2 f_3 + \left(\frac{2}{3}\right)^2 g_3 + \frac{2}{3}g_2 + g_1 \\ &= \dots \\ &= \left(\frac{2}{3}\right)^{n-1} f_n + \sum_{k=1}^{n-1} \left(\frac{2}{3}\right)^{k-1} g_k. \end{aligned}$$

Hence if we define $F_n(x) := \sum_{k=1}^{n-1} \left(\frac{2}{3}\right)^{k-1} g_k$, then F_n converges uniformly to the continuous function

$$F = \sum_{k=1}^{\infty} \left(\frac{2}{3}\right)^{k-1} g_k.$$

Since $|f_n| \le 2/3$ on A,

$$|f - F| = \lim_{n\to\infty} |f - F_n| = \lim_{n\to\infty} \left(\frac{2}{3}\right)^{n-1} |f_n| = 0 \text{ on } A,$$

i.e. $f = F|_A$. To finish the proof notice that, on X,

$$0 \le F = \sum_{k=1}^{\infty} \left(\frac{2}{3}\right)^{k-1} g_k \le \frac{2}{3} \sum_{k=1}^{\infty} \left(\frac{2}{3}\right)^{k-1} = \frac{2}{3}\frac{1}{1 - 2/3} = 1.$$

■

Corollary 2.7. *Suppose that X is a normal topological space, $A \subset X$ is closed, $F \in C(A, \mathbb{R})$. Then there exists $F \in C(X)$ such that $F|_A = f$.*

Proof. Let $g = \arctan(f) \in C(A, (-\frac{\pi}{2}, \frac{\pi}{2}))$. Then by the Tietze extension theorem, there exists $G \in C(X, [-\frac{\pi}{2}, \frac{\pi}{2}])$ such that $G|_A = g$. Let $B \equiv G^{-1}(\{-\frac{\pi}{2}, \frac{\pi}{2}\}) \sqsubset X$, then $B \cap A = \emptyset$. By Urysohn's lemma there exists $h \in C(X, [0,1])$ such that $h \equiv 1$ on A and $h = 0$ on B and in particular $hG \in C(A, (-\frac{\pi}{2}, \frac{\pi}{2}))$ and $(hG)|_A = g$. The function $F \equiv \tan(hG) \in C(X)$ is an extension of f. ■

3. Compact Spaces

Definition 3.1. Let (X, τ) be a topological space and $A \subset X$. We say a subset $\mathcal{U} \subset \tau$ is an **open cover** of A if $A \subset \cup\mathcal{U}$. The set A is said to be **compact** if every open cover of A has finite a sub-cover, i.e. if $\mathcal{U}$ is an open cover of A there exists $\mathcal{U}_0 \subset\subset \mathcal{U}$ such that $\mathcal{U}_0$ is a cover of A. (We will write $A \sqsubset\sqsubset X$ to denote that $A \subset X$ and A is compact.) A subset $A \subset X$ is **precompact** if $\bar{A}$ is compact. As usual, the reader should notice that $A \subset X$ is compact iff (A, τ_A) is compact.

Definition 3.2. We say a collection $\mathcal{F}$ of closed subsets of a topological space (X, τ) has the **finite intersection property if** $\cap\mathcal{F}_0 \neq \emptyset$ for all $\mathcal{F}_0 \subset\subset \mathcal{F}$.

The notion of compactness may be expressed in terms of closed sets as follows.

Proposition 3.3. *A topological space X is compact iff every family of closed sets $\mathcal{F} \subset \mathcal{P}(X)$ with the **finite intersection property** satisfies $\bigcap \mathcal{F} \neq \emptyset$.*

Proof. ($\Rightarrow$) Suppose that X is compact and $\mathcal{F} \subset \mathcal{P}(X)$ is a collection of closed sets such that $\bigcap \mathcal{F} = \emptyset$. Let

$$\mathcal{U} = \mathcal{F}^c := \{C^c : C \in \mathcal{F}\} \subset \tau,$$

then $\mathcal{U}$ is a cover of X and hence has a finite subcover, $\mathcal{U}_0$. Let $\mathcal{F}_0 = \mathcal{U}_0^c \subset\subset \mathcal{F}$, then $\cap \mathcal{F}_0 = \emptyset$ so that $\mathcal{F}$ does not have the finite intersection property.

($\Leftarrow$) If X is not compact, there exists an open cover $\mathcal{U}$ of X with no finite subcover. Let $\mathcal{F} = \mathcal{U}^c$, then $\mathcal{F}$ is a collection of closed sets with the finite intersection property while $\bigcap \mathcal{F} = \emptyset$. ■

Proposition 3.4. *Closed subsets of compact spaces are compact.*

Proof. Let $F \subset X$ be closed and $\mathcal{U}$ be an open cover of F then $\mathcal{U} \cup \{F^c\}$ is an open cover of X. Therefore there exists $\mathcal{U}_0 \subset\subset \mathcal{U}$ such that $\mathcal{U}_0 \cup \{F^c\}$ covers X. The finite collection of open sets $\mathcal{U}_0$ is a finite subcover of F. ■

Proposition 3.5. *Suppose that (X, τ) is a Hausdorff space, $K \sqsubset\sqsubset X$ and $x \notin K$. Then there exists $U, V \in \tau$ such that $U \cap V = \emptyset$, $x \in U$ and $K \subset V$. In particular K is closed.*

Proof. Because X is Hausdorff, for all $y \in K$ there exists $V_y \in \tau_y$ and $U_y \in \tau_x$ such that $V_y \cap U_y = \emptyset$. The cover $\{V_y\}_{y \in K}$ of K has a finite subcover, $\{V_y\}_{y \in \Lambda}$ for some $\Lambda \subset\subset K$. Let $V = \cup\{V_y : y \in \Lambda\}$ and $U = \cap\{U_y : y \in \Lambda\}$, then $U, V \in \tau$ satisfy $x \in U$, $K \subset V$ and $U \cap V = \emptyset$. This shows that K^c is open and hence that K is closed. ■

Proposition 3.6. *If (X, τ) is Hausdorff and compact then X is normal.*

Proof. Let A and B be closed disjoint subsets of X. By Proposition 3.4, both A and B are compact. By Proposition 3.5, for all $x \in B$ there exist $V_x \in \tau_x$, $U_x \subset_o X$ such that $A \subset U_x$ and $U_x \cap V_x = \emptyset$. By Compactness of A, there is a finite set $\Lambda \subset\subset B$ such that $V = \cup_{x \in \Lambda} V_x$ contains B. Let $U := \cap_{x \in \Lambda} U_x \in \tau$, then $U \cap V = \emptyset$ while $A \subset U$ and $B \subset V$. ■

The next Proposition contains some easily verified facts about compact sets. The proof is left to the reader

Proposition 3.7. *Suppose that X and Y are topological spaces.*

1. *Suppose that $A \subset X$ compact and $f : X \to Y$ is a continuous, then $f(A)$ is compact in Y.*
2. *The finite union of compact sets is compact.*
3. *If X is compact and $f \in C(X)$, then f is bounded, i.e. $C(X) = BC(X)$.*
4. *If X is compact and Y is Hausdorff and $f : X \to Y$ is a continuous bijection then f is a homeomorphism, i.e. $f^{-1} : Y \to X$ is continuous as well. (Just show that $f(C) \sqsubset Y$ for all $C \sqsubset X$.)*

3.1. **Compactness in Metric Spaces.** Let (X, ρ) be a metric space and let $B'_x(\epsilon) = B_x(\epsilon) \setminus \{x\}$. Let us start with the following elementary lemma which is left as an exercise to the reader.

Lemma 3.8. *Let $E \subset X$ be a subset of a metric space (X, ρ). Then the following are equivalent:*

1. *$x \in X$ is an accumulation point of E.*
2. *$B'_x(\epsilon) \cap E \neq \emptyset$ for all $\epsilon > 0$.*

3. $B_x(\epsilon) \cap E$ *is an infinite set for all* $\epsilon > 0$.
4. *There exists* $\{x_n\}_{n=1}^{\infty} \subset E \setminus \{x\}$ *with* $\lim_{n\to\infty} x_n = x$.

Definition 3.9. A metric space (X, ρ) is said to be $\epsilon -$ **bounded** $(\epsilon > 0)$ provided there exists a finite cover of X by balls of radius ϵ. The metric space is **totally bounded** if it is $\epsilon -$ bounded for all $\epsilon > 0$.

Theorem 3.10. *Let X be a metric space. The following are equivalent.*

(a) *X is compact.*
(b) *Every infinite subset of X has an accumulation point.*
(c) *X is totally bounded and complete.*

Proof. The proof will consist of showing that $a \Rightarrow b \Rightarrow c \Rightarrow a$.

($a \Rightarrow b$) We will show that **not** $b \Rightarrow$ **not** a. Suppose there exists $E \subset X$,such that $\#(E) = \infty$ and E has no accumulation points. Then for all $x \in X$ there exists $V_x \in \tau_x$ such that $(V_x \setminus \{x\}) \cap E = \emptyset$. Clearly $\mathcal{V} = \{V_x\}_{x\in X}$ is a cover of X, yet $\mathcal{V}$ has no finite sub cover. Indeed, for each $x \in X$, $V_x \cap E$ consists of at most one point, therefore if $\Lambda \subset\subset X$, $\cup_{x\in\Lambda} V_x$ can only contain a finite number of points from E, in particular $X \neq \cup_{x\in\Lambda} V_x$.

($b \Rightarrow c$) Let $\epsilon > 0$ be given and choose $x_1 \in X$. If possible choose $x_2 \in X$ such that $d(x_2, x_1) \geq \epsilon$, then if possible choose $x_3 \in X$ such that $d(x_3, \{x_1, x_2\}) \geq \epsilon$ and continue inductively choosing points $\{x_j\}_{j=1}^{n} \subset X$ such that $d(x_n, \{x_1, \dots, x_{n-1}\}) \geq \epsilon$. This process must terminate, for otherwise we could choose $E = \{x_j\}_{j=1}^{\infty}$ and infinite number of distinct points such that $d(x_j, \{x_1, \dots, x_{j-1}\}) \geq \epsilon$ for all $j = 2, 3, 4, \dots$. Since for all $x \in X$ the $B_x(\epsilon/3) \cap E$ can contain at most one point, no point $x \in X$ is an accumulation point of E.

($c \Rightarrow a$) For sake of contradiction, assume there exists a cover an open cover $\mathcal{V} = \{V_\alpha\}_{\alpha\in A}$ of X with no finite subcover. Since X is totally bounded for each $n \in \mathbb{N}$ there exists $\Lambda_n \subset\subset X$ such that

$$X = \bigcup_{x\in\Lambda_n} \overline{B_x(1/n)}.$$

Choose $x_1 \in \Lambda_1$ such that no finite subset of $\mathcal{V}$ covers $K_1 := \overline{B_{x_1}(1)}$. Since $K_1 = \cap_{x\in\Lambda_2} K_1 \cap \overline{B_x(1/2)}$, there exists $x_2 \in \Lambda_2$ such that $K_2 := K_1 \cap \overline{B_{x_2}(1/2)}$ can not be covered by a finite subset of $\mathcal{V}$. Continuing this way inductively, we construct sets $K_n = K_{n-1} \cap \overline{B_{x_n}(1/n)}$ with $x_n \in \Lambda_n$ such no K_n can be covered by a finite subset of $\mathcal{V}$. Now choose $y_n \in K_n$ for each n. Since $\{K_n\}_{n=1}^{\infty}$ is a decreasing sequence of closed sets such that $\operatorname{diam}(K_n) \leq 2/n$, it follows that $\{y_n\}$ is a Cauchy and hence convergent sequence and

$$y = \lim_{n\to\infty} y_n \in \cap_{m=1}^{\infty} K_m.$$

Since $\mathcal{V}$ is a cover of X, there exists $V \in \mathcal{V}$ such that $x \in V$. Since $K_n \downarrow \{y\}$ and $\operatorname{diam}(K_n) \to 0$, it now follows that $K_n \subset V$ for some n large. But this violates the assertion that K_n can not be covered by a finite subset of $\mathcal{V}$. ■

Corollary 3.11. *Let X be a metric space then X is compact iff* ***all*** *sequences $\{x_n\} \subset X$ have convergent subsequences.*

Proof. If X is compact and $\{x_n\} \subset X$

1. If $\#(\{x_n : n = 1, 2, \dots\}) < \infty$ then choose $x \in X$ such that $x_n = x$ i.o. let $\{n_k\} \subset \{n\}$ such that $x_{n_k} = x$ for all k. Then $x_{n_k} \to x$

2. If $\#(\{x_n : n = 1, 2, \dots\}) = \infty$. We know $E = \{x_n\}$ has an accumulation point $\{x\}$, hence there exists $x_{n_k} \to x$.

Conversely if E is an infinite set let $\{x_n\}_{n=1}^\infty \subset E$ be a sequence of distinct elements of E. We may, by passing to a subsequence, assume $x_n \to x \in X$ as $n \to \infty$. Now $x \in X$ is an accumulation point of E by Theorem 3.10 and hence X is compact. ■

The following is an application of the previous theorem.

Proposition 3.12. *Suppose that (X, ρ) is a complete separable metric space and μ is a probability measure on $\mathcal{B} = \sigma(\tau_\rho)$. The for all $\epsilon > 0$, there exists $K_\epsilon \sqsubset\sqsubset X$ such that $\mu(K_\epsilon) \geq 1 - \epsilon$.*

Proof. Let $\{x_k\}_{k=1}^\infty$ be a countable dense subset of X. Then $X = \cup_k \overline{B_{x_k}(1/n)}$ for all $n \in \mathbb{N}$. Hence by continuity of μ, there exists, for all $n \in \mathbb{N}$, $N_n < \infty$ such that $\mu(F_n) \geq 1 - \epsilon 2^{-n}$ where $F_n := \cup_{k=1}^{N_n} \overline{B_{x_k}(1/n)}$. Let $K := \cap_{n=1}^\infty F_n$ then

$$\mu(X \setminus K) = \mu(\cup_{n=1}^\infty F_n^c) \leq \sum_{n=1}^\infty \mu(F_n^c) = \sum_{n=1}^\infty (1 - \mu(F_n)) \leq \sum_{n=1}^\infty \epsilon 2^{-n} = \epsilon$$

so that $\mu(K) \geq 1 - \epsilon$. Moreover K is compact since K is closed and totally bounded; $K \subset F_n$ for all n and each F_n is $1/n$ – bounded. ■

3.2. Locally compact spaces.

Definition 3.13. A topological space X is locally compact if for all $x \in X$ there exists a compact neighborhood N_x of x, i.e. $N_x \sqsubset\sqsubset X$ and $x \in N_x^0$.

Remark 3.14. If X is Hausdorff, then X is locally compact iff for every $x \in X$ there exists $V \in \tau_x$ which is precompact. To verify this assertion, suppose that N_x is as in Definition 3.13. Let $V = N_x^0 \in \tau$, then, since N_x is closed by Proposition 3.5, $\bar{V} \subset N_x$ and hence $\bar{V}$ is compact by Proposition 3.4. Conversely if $V \in \tau_x$ is precompact, $N_x = \bar{V}$ is a compact neighborhood of x because $x \in V \subset N_x^0$.

Finite dimensional Euclidean spaces, $\mathbb{R}^n$, are typical examples of locally compact spaces. Also any subset of a locally compact space with the relative topology is locally compact.

Proposition 3.15. *Suppose X is a locally compact Hausdorff space (LCH for short) and $U \subset_o X$. For all compact subset K of X such that $K \subset U$, there exists a precompact open set V such that $K \subset V \subset \overline{V} \subset U \subset X$.*

Proof. By local compactness, for all $x \in K$, there exists $U_x \in \tau_x$ such that $\bar{U}_x$ is compact. Since K is compact, there exists $\Lambda \subset\subset K$ such that $\{U_x\}_{x \in \Lambda}$ is a cover of K. The set $O = \cup_{x \in \Lambda} U_x$ is an open set which contains K and is precompact since $\bar{O} = \cup_{x \in \Lambda} \bar{U}_x$ – the finite union of compact sets. So by replacing U by $U \cap O$ if necessary, we may assume that $\bar{U}$ is compact.

Let $Y = \bar{U}$, a compact Hausdorff space and hence a normal space. Since ∂U is closed in X and contained in Y, ∂U is close in Y.[2] Now K and ∂U are closed

[2]Notice that the boundary of U in Y, $\partial_Y U$, is given by the same as the boundary of U in X because

$$\partial_Y U = \bar{U}^Y \setminus U = \bar{U} \cap Y \setminus U = \bar{U} \setminus U = \partial U.$$

disjoint subsets of Y and therefore, there exists disjoint relatively open sets V and W in Y such that $K \subset V$ and $\partial U \subset W$. Since $V \subset Y \setminus W \sqsubset Y$,

$$\bar{V} = \bar{V} \cap Y = \bar{V}^Y \subset Y \setminus W \subset Y \setminus \partial U = U,$$

wherein we have used $V \subset Y$ with Y closed in X. Finally V is open in X since $V \subset U \subset_o X$ and V is relatively open in $\bar{U}$. (So $V = W \cap \bar{U}$ for some $W \subset_0 X$ and hence $V = V \cap U = W \cap U \subset_o X$.) ∎

Definition 3.16. Suppose that $f : X \to \mathbb{C}$ is a function. The support of f, $\mathrm{supp}(f)$, is the smallest closed set outside of which f is 0, i.e.

$$\mathrm{supp}(f) = \overline{\{f \neq 0\}} = \cap \{C : C \sqsubset X \text{ and } f|_{C^c} \equiv 0\}.$$

Alternatively

$$\mathrm{supp}(f)^c = \{f = 0\}^0 = \cup \{V : V \in \tau \text{ and } f|_V \equiv 0\}.$$

If $U \subset_0 X$, we will write $C_c(U)$ for those function $f \in C(U)$ with compact support in U. By abuse of notation, we will consider $C_c(U)$ as a subspace of $C_c(X)$ by identifying $f \in C_c(U)$ with its extension by zero to a function on X.

For example, let $f(x) = \sin(x) 1_{[0,4\pi]}(x)$, then

$$\{f \neq 0\} = (0, 4\pi) \setminus \{\pi, 2\pi, 3\pi\}$$

and therefore $\mathrm{supp}(f) = [0, 4\pi]$.

Lemma 3.17 (Locally Compact Version of Urysohn's Lemma)**.** *Let X be a locally compact Hausdorff space and $K \sqsubset\sqsubset U \subset_o X$. Then there exists $f \in C(X, [0,1])$ such that $f = 1$ on K and $\mathrm{supp}(f) \sqsubset\sqsubset U$. Alternatively put, if K is compact and C is closed in X such that $K \cap C = \emptyset$, then there exists $f \in C_c(X, [0,1])$ such that $f = 1$ on K and $f = 0$ on C.*

Proof. Let V be an precompact open set as in Proposition 3.15, then $\bar{V}$ is a compact Hausdorff space and hence is normal. Since K and ∂V are closed subsets of $\bar{V}$, Urysohn's lemma implies there exists $g \in C(\overline{V}, [0,1])$ such that $g = 1$ on K and $g = 0$ on ∂V. We may now define $f = g$ on $\bar{V}$ and $f \equiv 0$ on $\overline{V}^c$, then $f \in C(X, [0,1])$ is the desired function since $\mathrm{supp}(f) \subset \bar{V}$ and $\bar{V}$ is compact. ∎

Theorem 3.18 (Locally Compact Version Tietze's Extension Theorem)**.** *Let X be a locally compact Hausdorff space, K be a compact subset of X and $f \in C(K, [0,1])$. Then there exists $F \in C(X, [0,1])$ such that $F|_K = f$. Moreover F may be taken to vanish outside a compact set, i.e. $\mathrm{supp}(F)$ is compact.*

Proof. By Proposition 3.15 there exists a precompact open set, V, such that $K \subset V \subset \overline{V} \subset X$. Let $g \in C(K \cup \partial V)$ be defined by

$$g(x) = \begin{cases} f(x) & x \in K \\ 0 & x \in \partial V. \end{cases}$$

Then by the Tietze extension theorem, there exists $G \in C(\overline{V}, [0,1])$ such that $G|_K = f$ and $G|_{\partial V} \equiv 0$. The desired function $F \in C(X)$ may now be defined by

$$F(x) = \begin{cases} G(x) & \text{for } x \in \overline{V} \\ 0 & x \notin \overline{V}. \end{cases}.$$

∎

Definition 3.19. Let X be a topological space. We say that a function $f \in C(X)$ vanishes at infinity if for all $\epsilon > 0$, $\{|f| \geq \epsilon\}$ is compact in X. We will denote the function $f \in C(X)$ vanishing at infinity by $C_0(X)$.

Notice that $C_0(X) \subset BC(X)$.

Proposition 3.20. *Let X be a topological space, $BC(X)$ be the space of bounded continuous functions on X with the supremum norm topology. Then*

1. *$C_0(X)$ is a closed subspace of $BC(X)$.*
2. *If we further assume that X is a locally compact Hausdorff space, then $C_0(X) = \overline{C_c(X)}$.*

Proof.

1. Suppose that $\{f_n\}_{n=1}^{\infty} \subset C_0(X)$ and $f \in BC(X)$ such that $\|f - f_n\| \to 0$ as $n \to \infty$. Since
$$|f| \leq |f_n| + |f - f_n| \leq |f_n| + \|f - f_n\|,$$
we have for all $\epsilon > 0$ that
$$\{|f| \geq \epsilon\} \subset \{|f_n| + \|f - f_n\| \geq \epsilon\} = \{|f_n| \geq \epsilon - \|f - f_n\|\} \sqsubset\sqsubset X$$
provided that n is large enough that $\|f - f_n\| < \epsilon$. Since $\{|f| \geq \epsilon\}$ is a closed subset of a compact set, $\{|f| \geq \epsilon\}$ is compact as well. Hence $f \in C_0(X)$.
2. Since $C_0(X)$ is a closed subspace of $BC(X)$ and $C_c(X) \subset C_0(X)$, we always have $\overline{C_c(X)} \subset C_0(X)$. Now suppose that $f \in C_0(X)$ and let $K_n \equiv \{|f| \geq \frac{1}{n}\} \sqsubset\sqsubset X$. By Lemma 3.17 we may choose $\phi_n \in C_c(X, [0,1])$ such that $\phi_n \equiv 1$ on K_n. Define $f_n \equiv \phi_n f \in C_c(X)$. Then
$$\|f - f_n\|_\infty = \|(1 - \phi_n) f\|_\infty \leq \frac{1}{n} \to 0 \text{ as } n \to \infty.$$
This shows that $f \in \overline{C_c(X)}$.

■

Proposition 3.21 (Alexanderov Compactification)**.** *Suppose that X is a non-compact Locally compact Hausdorff space. Let $X^* = X \cup \{\infty\}$, where $\{\infty\}$ is a some point disjoint form X. Let*
$$\tau^* = \tau \cup \{(X \setminus K) \cup \{\infty\} : K \sqsubset\sqsubset X\}.$$
The τ^* is a topology on X^* and (X^*, τ^*) is a compact Hausdorff space. Moreover if $f \in C(X)$ extends to continuously to X^* iff $f = g + c$ with $g \in C_0(X)$ and $c \in \mathbb{C}$. The extension is given by $f(\infty) = c$.

Proof. Suppose that $U, V \in \tau^*$. If U or $V \in \tau$, then $U \cap V \in \tau \subset \tau^*$ and $U = (X \setminus K) \cup \{\infty\}$ and $V = (X \setminus L) \cup \{\infty\}$ where K and L are compact sets then
$$U \cap V = (X \setminus K) \cap (X \setminus L) \cup \{\infty\} = (X \setminus (K \cup L)) \cup \{\infty\} \in \tau^*$$
since $K \cup L$ is compact. Therefore τ^* is closed under intersections. Now if $U = (X \setminus K) \cup \{\infty\}$ with $K \sqsubset\sqsubset X$ and $V \in \tau$, then
$$X^* \setminus (U \cup V) = X \setminus ((X \setminus K) \cup V) = X \cap (K \cap V^c) = K \cap V^c$$
a compact set and hence
$$U \cup V = (X \setminus (K \cap V^c)) \cup \{\infty\} \in \tau^*.$$

Similarly if $U_\alpha = (X \setminus K_\alpha) \cup \{\infty\}$ with $K_\alpha \sqsubset\sqsubset X$ for all $\alpha \in A$, then

$$\cup_{\alpha\in A} U_\alpha = (X \setminus \cap_{\alpha\in A} K_\alpha) \cup \{\infty\} \in \tau^*.$$

The other possible cases are now easily checked so the τ^* is closed under arbitrary unions. Since $\emptyset$ and X^* are in τ^* (because $\emptyset$ is compact), it follows that τ^* is a topology.

Let $i : X \to X^*$ be the inclusion map. Then i is continuous and open, i.e. $i(V)$ is open in X^* for all V open in X. If if $f \in C(X^*)$, then $g = f|_X - f(\infty) = f \circ i - f(\infty)$ is continuous on X. Moreover, for all $\epsilon > 0$ there exists $V \in \tau^*_\infty$ such that

$$|g(x)| = |f(x) - f(\infty)| < \epsilon \text{ for all } x \in V \setminus \{\infty\}.$$

Since $V = (X \setminus K) \cup \{\infty\}$, it follows that g vanishes at ∞ since $\{|g| \geq \epsilon\} \subset K$. Conversely if $g \in C_0(X)$ extend g to X^* by setting $g(\infty) = 0$. Given $\epsilon > 0$, the set $K = \{|g| \geq \epsilon\}$ is compact, hence $g(X^* \setminus K) \subset (-\epsilon, \epsilon)$, which shows that g is continuous at ∞ and so g is continuous on X^*. Now it $f = g + c$ with $c \in \mathbb{C}$ and $g \in C_0(X)$, it follows by what we just proved that f extended to X^* by $f(\infty) = c$ is continuous on X^*. ■

Lemma 3.22. *Suppose that X is a locally compact Hausdorff space and $E \subset X$. Then E is closed iff $E \cap K$ is closed for all $K \sqsubset\sqsubset X$.*

Proof. Since compact subsets of Hausdorff spaces are closed, $E \cap K$ is closed if E is closed and K is compact. If E were not closed there exists $x \in \overline{E} \setminus E$. By Proposition 3.15 there exists a compact set $K \subset X$ such that $x \in K^0 \subset K$. I now claim that $x \in \overline{K \cap E}$. Indeed, if $V \in \tau_x$ then $V \cap K \cap E \supset V \cap K^0 \cap E \neq \emptyset$

since $V \cap K^0 \in \tau_x$ and $x \in \bar{E}$. This shows that $x \in \overline{K \cap E}$ and since $x \notin E$ we see that $x \in \overline{K \cap E} \setminus (K \cap E)$. In particular this shows that $K \cap E \neq \overline{K \cap E}$ and hence $K \cap E$ is not closed. ■

3.3. **Partitions of Unity.**

Definition 3.23. Let $\{U_\alpha\}_{\alpha\in A}$ be an open cover of $E \subset X$. We say $\{h_\alpha \in C(X, [0, 1])\}_{\alpha\in A}$ is a partition of unity of E subordinate to $\{U_\alpha\}$ if

1. $\mathrm{supp}(h_\alpha) \subset U_\alpha$ for all $\alpha \in A$,
2. for all $x \in X$ there exists neighborhood N_x of x such that $h_\alpha|_{N_x} \equiv 0$ for all but finitely many α.
3. $\sum\limits_{\alpha\ \in A} h_\alpha(x) = 1$ for all $x \in E$.

Proposition 3.24. *Suppose that X is a locally compact Hausdorff space, $K \subset X$ is a compact set, and $\mathcal{U} = \{U_j\}_{j=1}^n$ is an open cover of K. Then there exists a partition of unity $\{h_j\}_{j=1}^n$ of K subordinate to $\mathcal{U}$ such that* $\mathrm{supp}(h_j)$ *is compact for all j.*

Proof. For all $x \in K$ choose $V_x \in \tau_x$such that $\overline{V}_x \subset U_j$ for some j and $\overline{V}_x$ is compact. Choose $\Lambda \subset\subset K$ such that $K \subset \bigcup\limits_{x\in\Lambda} V_x$. Let $F_j = \cup\left\{\bar{V}_x : x \in \Lambda \text{ and } \overline{V}_x \subset U_j\right\}$, then F_j is compact, $F_j \subset U_j$ for all j, and $K \subset \cup_{j=1}^n F_j$. Choose $f_j \in C_c(X, [0, 1])$ such that $f_j = 1$ on F_j and $\mathrm{supp}\,(f_j) \subset U_j$. Then

$$g = \sum_{j=1}^n f_j \in C_c(X)$$

and $g \geq 1$ on K and hence $K \subset \{g > \frac{1}{2}\}$. Choose $v \in C_c(X, [0,1])$ such that $v = 1$ on K and $\mathrm{supp}(v) \subset \{g > \frac{1}{2}\}$ and define $f_0 \equiv 1 - v$. Then $f_0 = 0$ on K, $f_0 = 1$ if $g \leq \frac{1}{2}$ and therefore,

$$f_0 + f_1 + \cdots + f_n = f_0 + g > 0$$

on X. The desired partition of unity may be constructed as

$$h_j(x) = \frac{f_j(x)}{f_0(x) + \cdots + f_n(x)}.$$

Indeed $\mathrm{supp}\,(h_j) = \mathrm{supp}\,(f_j) \subset U_j$, $h_j \in C_c(X, [0,1])$ and on K,

$$h_1 + \cdots + h_n = \frac{f_1 + \cdots + f_n}{f_0 + f_1 + \cdots + f_n} = \frac{f_1 + \cdots + f_n}{f_1 + \cdots + f_n} = 1.$$

■

4. Compactness in Function spaces

In this section, let X be a topological space.

Definition 4.1. Let $\mathcal{F} \subset C(X)$.

1. $\mathcal{F}$ is equicontinuous at $x \in X$ iff for all $\epsilon > 0$ there exists $U \in \tau_x$ such that $|f(y) - f(x)| < \epsilon$ for all $y \in U$ and $f \in \mathcal{F}$.
2. $\mathcal{F}$ is equicontinuous if $\mathcal{F}$ is equicontinuous at all points $x \in X$.
3. $\mathcal{F}$ is pointwise bounded if $\sup\{f(x) \in \mathbb{C} | f \in \mathcal{F}\} < \infty$ for all $x \in X$.

Theorem 4.2 (Ascoli-Arzela Theorem)**.** *Let X be a compact Hausdorff space and $\mathcal{F} \subset C(X)$. Then $\mathcal{F}$ is precompact in $C(X)$ iff $\mathcal{F}$ is equicontinuous and point-wise bounded.*

Proof. ($\Leftarrow$) Since $B(X)$ is a complete metric space, we must show $\mathcal{F}$ is totally bounded. Let $\epsilon > 0$ be given. By equicontinuity there exists $V_x \in \tau_X$ for all $x \in X$ such that $|f(y) - f(x)| < \epsilon/2$ if $y \in V_x$ and $f \in \mathcal{F}$. Since X is compact we may choose $\Lambda \subset\subset X$ such that $X = \cup_{x\in\Lambda} V_x$. We have now decomposed X into "blocks" $\{V_x\}_{x\in\Lambda}$ such that each $f \in \mathcal{F}$ is constant to within ϵ on V_x. Since $\sup\{f(x) : x \in \Lambda \text{ and } f \in \mathcal{F}\} < \infty$, it is now evident that

$$M \equiv \sup\{f(x) : x \in X \text{ and } f \in \mathcal{F}\} \leq \sup\{f(x) : x \in \Lambda \text{ and } f \in \mathcal{F}\} + \epsilon < \infty.$$

Let $\mathbb{D} \equiv \{k\epsilon/2 : k \in \mathbb{Z}\} \cap [-M, M]$ and notice that if $f \in \mathcal{F}$ and $\phi \in \mathbb{D}^\Lambda$ (i.e. $\phi : \Lambda \to \mathbb{D}$ is a function) is chosen so that $|\phi(x) - f(x)| \leq \epsilon/2$ for all $x \in \Lambda$ then for $x \in \Lambda$ and $y \in V_x$,

$$|f(y) - \phi(x)| \leq |f(y) - f(x)| + |f(x) - \phi(x)| < \epsilon.$$

Hence if we set, for $\phi \in \mathbb{D}^\Lambda$,

$$\mathcal{F}_\phi \equiv \{f \in \mathcal{F} : |f(y) - \phi(x)| < \epsilon \text{ for } y \in V_x \text{ and } x \in \Lambda\}$$

then $\mathcal{F} = \bigcup \left\{\mathcal{F}_\phi : \phi \in \mathbb{D}^\Lambda\right\}$.

Let $\Gamma := \left\{\phi \in \mathbb{D}^\Lambda : \mathcal{F}_\phi \neq \emptyset\right\}$ and for each $\phi \in \Gamma$ choose $f_\phi \in \mathcal{F}_\phi \cap \mathcal{F}$. For $f \in \mathcal{F}_\phi$, $x \in \Lambda$ and $y \in V_x$ we have

$$|f(y) - f_\phi(y)| \leq |f(y) - \phi(x))| + |\phi(x) - f_\phi(y)| < 2\epsilon.$$

So $\|f - f_\phi\| < 2\epsilon$ for all $f \in \mathcal{F}_\phi$ showing that $\mathcal{F}_\phi \subset B_{f_\phi}(2\epsilon)$. Therefore,

$$\mathcal{F} = \cup_{\phi\in\Gamma} \mathcal{F}_\phi \subset \cup_{\phi\in\Gamma} B_{f_\phi}(2\epsilon)$$

and because $\epsilon > 0$ was arbitrary we have shown that $\mathcal{F}$ is totally bounded.

($\Rightarrow$) Since $\|\cdot\| : C(X) \to [0, \infty)$ is a continuous function on $C(X)$ it is bounded on any compact subset $\mathcal{F} \subset C(X)$. This shows that $\sup\{\|f\| : f \in \mathcal{F}\} < \infty$ which clearly implies that $\mathcal{F}$ is pointwise bounded.[3] Suppose $\mathcal{F}$ were **not** equicontinuous at some point $x \in X$ that is to say there exists $\epsilon > 0$ such that for all $V \in \tau_x$, $\sup_{y\in V} \sup_{f\in\mathcal{F}} |f(y) - f(x)| > \epsilon$.[4] Equivalently put, to each $V \in \tau_x$ we may choose

$$f_V \in \mathcal{F} \text{ and } x_V \in V \quad \text{such that} \quad |f_V(x) - f_V(x_V)| \geq \epsilon.^5 \tag{4.1}$$

Set $\mathcal{C}_V = \overline{\{f_W : W \in \tau_x \text{ and } W \subset V\}}^{\|\cdot\|_u} \subset \mathcal{F}$ and notice for any $\mathcal{V} \subset\subset \tau_x$ that

$$\cap_{V\in\mathcal{V}} \mathcal{C}_V \supseteq \mathcal{C}_{\cap\mathcal{V}} \neq \emptyset,$$

so that $\{\mathcal{C}_V\}_V \in \tau_x \subset \mathcal{F}$ has the finite intersection property. Since $\mathcal{F}$ is compact, it follows that there exists some

$$f \in \bigcap_{V\in\tau_x} \mathcal{C}_V \neq \emptyset.$$

Since f is continuous, there exists $V \in \tau_x$ such that $|f(x) - f(y)| < \epsilon/3$ for all $y \in V$. Since $f \in \mathcal{C}_V$, there exists $W \subset V$ such that $\|f - f_W\| < \epsilon/3$. We now arrive at a contradiction since

$$\begin{aligned}\epsilon &\leq |f_W(x) - f_W(x_W)| \leq |f_W(x) - f(x)| + |f(x) - f(x_W)| + |f(x_W) - f_W(x_W)| \\ &< \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon\end{aligned}$$

which is a contradiction. ■

Definition 4.3. A topological space (X, τ) is said to be σ − compact if there exists compact set $K_n \subset X$ such that $X = \cup_{n=1}^{\infty} K_n$. Notice that we may assume, by replacing K_n by $K_1 \cup K_2 \cup \cdots \cup K_n$ if necessary, that $K_n \uparrow X$.

Lemma 4.4. *Suppose X is a σ-compact and locally compact Hausdorff space. Then there exists precompact open set $S_n \subset X$ such that $S_n \subset \bar{S}_n \subset S_{n+1}$ for all n.*

Proof. Choose $K_n \uparrow X$ as in Definition 4.3. Let $\mathcal{U}_1 \subset \tau$ be a finite cover of K_1 by precompact open sets. Take $S_1 = \cup\mathcal{U}_1$. Then S_1 is open and precompact. We may now inductively construct S_n as in the lemma with the added property that $K_n \subset S_n$. Indeed if $S_1, \ldots, S_n$ have been chosen as described such that $K_i \subset S_i$ for all $i = 1, 2, \ldots, n$, then let $\mathcal{U}_{n+1}$ a finite cover of $K_n \cup \bar{S}_n$ by precompact open sets and define $S_{n+1} = \cup\mathcal{U}_{n+1}$. ■

Corollary 4.5. *Let $\{f_n\} \subset C(X)$ be a pointwise bounded sequence of functions which is equicontinuous on compact subsets of X. Then there exists a subsequence*

---

[3]One could also prove that $\mathcal{F}$ is pointwise bounded by considering the continuous evaluation maps $e_x : C(X) \to \mathbb{R}$ given by $e_x(f) = f(x)$ for all $x \in X$.

[4]If X is first countable we could finish the proof with the following argument. Let $\{V_n\}_{n=1}^{\infty}$ be a neighborhood base at x such that $V_1 \supset V_2 \supset V_3 \supset \ldots$. By the assumption that $\mathcal{F}$ is not equicontinuous at x, there exist $f_n \in \mathcal{F}$ and $x_n \in V_n$ such that $|f_n(x) - f_n(x_n)| \geq \epsilon \ \forall\ n$. Since $\mathcal{F}$ is a compact metric space by passing to a subsequence if necessary we may assume that f_n converges uniformly to some $f \in \mathcal{F}$. Because $x_n \to x$ as $n \to \infty$ we learn that

$$\begin{aligned}\epsilon &\leq |f_n(x) - f_n(x_n)| \leq |f_n(x) - f(x)| + |f(x) - f(x_n)| + |f(x_n) - f_n(x_n)| \\ &\leq 2\|f_n - f\| + |f(x) - f(x_n)| \to 0 \text{ as } n \to \infty\end{aligned}$$

which is a contradiction.

$\{n_k\} \subset \{n\}$ such that $g_k := f_{n_k}$ is a sequence which is convergent uniformly on compact subsets of X.

Proof. Let K_n's be compact subsets of X and use Proposition 3.15 to find precompact open sets V_n containing K_n. So by replacing K_n by $\bar{V}_n$ if necessary, we may find compact sets K_n such that $X = \cup_{n=1}^{\infty} K_n^0$.

We may now apply Theorem 4.2 repeatedly to find a nested family of subsequences

$$\{f_n\} \supset \{g_n^1\} \supset \{g_n^2\} \supset \{g_n^3\} \supset \dots$$

such that $\left\{g_n^k\right\}_{n=1}^{\infty}$ is a sequence of continuous functions uniformly convergent on K_k. Using Cantor's trick, define the subsequence $\{h_k\}$ of $\{f_n\}$ by $h_k \equiv g_k^k$. Then $\{h_k\}$ is uniformly convergent on each K_n. Now if $K \subset X$ is an arbitrary compact set, there exists $M < \infty$ such that $K \subset \cup_{n=1}^{M} K_n^0 \subset \cup_{n=1}^{M} K_n$ and therefore $\{h_k\}$ is uniformly convergent on K as well. ∎

Theorem 4.6 (Peano's Existence Theorem)**.** *Suppose $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is a bounded continuous function. Then there exists a solution to the differential equation*

$$\begin{aligned} &\dot{x}(t) = f(t, x(t)) \text{ for } t \geq 0 \text{ with} \\ &x(0) = x_0. \end{aligned}$$

Proof. Given $\epsilon > 0$, there exists a unique function $x_\epsilon \in C([-\epsilon, \infty) \to \mathbb{R}^d)$ such that $x_\epsilon(t) \equiv x_0$ for $-\epsilon \leq t \leq 0$ and

$$x_\epsilon(t) = x_0 + \int_0^t f(\tau, x_\epsilon(\tau - \epsilon)) d\tau \text{ for all } t \geq 0. \tag{4.2}$$

Indeed if $t \in [0, \epsilon]$, define $x_\epsilon(t)$ by Eq. (4.2), then use Eq. (4.2) to define x_ϵ on $[\epsilon, 2\epsilon]$, etc. Let $M = \sup |f(t, x)| < \infty$, then

$$|x_\epsilon(t)| \leq |x_0| + \int_0^t |f(\tau, x_\epsilon(\tau - \epsilon))| \, d\tau \leq |x_0| + Mt$$

and

$$|x_\epsilon(t) - x_\epsilon(s)| = \left| \int_s^t f(\tau, x_\epsilon(\tau - \epsilon)) d\tau \right| \leq M|t - s|$$

for all $t \geq s \geq 0$ and $\epsilon > 0$. Therefore $\{x_\epsilon\}_{\epsilon > 0}$ is an equicontinuous pointwise bounded family in $C([0, \infty), \mathbb{R}^d)$ and hence there exists $\epsilon_k \downarrow 0$ such that x_{ϵ_k} uniformly converges on compact subintervals of $[0, \infty)$.to some $x \in C([0, \infty), \mathbb{R}^d)$. Passing to the limit in Eq. (4.2) implies that

$$x(t) \equiv \lim_{k \to \infty} x_{\epsilon_k}(t) = x_0 + \int_0^t f(\tau, x(\tau)) d\tau.$$

∎

We could give another proof of this theorem as follows. Let $f_\epsilon(t, \cdot) = f(t, \cdot) * \delta_\epsilon$ where δ_ϵ is an approximate δ – function as described in Lemma 5.2 and 5.3 below. Then let $x_\epsilon(t)$ solve

$$\begin{aligned} &\dot{x}_\epsilon(t) = f_\epsilon(t, x_\epsilon(t)) \text{ for } t \geq 0 \text{ with} \\ &x_\epsilon(0) = x_0 \end{aligned}$$

which we know has solutions by the Lipschitz case proved on the homework. Moreover $\{x_\epsilon\}_{\epsilon>0}$ has uniformly bounded derivatives and hence are equicontinuous. They are also pointwise bounded, so there exists $\epsilon_k \downarrow 0$ such that x_{ϵ_k} uniformly converges on compact subintervals of $\mathbb{R}$ to some $x \in C(\mathbb{R}, \mathbb{R}^d)$. Passing to the limit in the integral equation

$$x_{\epsilon_k}(t) = x_0 + \int_0^t f_{\epsilon_k}(\tau, x_{\epsilon_k}(\tau))d\tau$$

then shows that x solves the desired ordinary differential equation.

5. Approximation Theorems

5.1. Classical Weierstrass Approximation Theorem.

Theorem 5.1 (Weierstrass Approximation Theorem). *Let $f \in C([a,b], \mathbb{R})$, then there exists polynomials $P_n(x)$ such that $P_n \to f$ uniformly.*

We will give the proof after the next two lemmas. For this lemma let $\{Q_n\}_{n=1}^\infty$ be the following sequence of "approximate δ – functions:"

$$Q_n(x) \equiv \frac{1}{c_n}(1-x^2)^n 1_{|x|\le 1} \text{ where } c_n := \int_{-1}^1 (1-x^2)^n dx. \tag{5.1}$$

We will give two proof of this theorem. The first is based on approximate δ – functions and the second is based on the weak law of large numbers.

Lemma 5.2. *Let $\{Q_n\}_{n=1}^\infty$ be the sequence of functions defined in Eq. (5.1), then*

$$\int_{\mathbb{R}} Q_n(x)\ dx = \int_{-1}^1 Q_n(x)\ dx = 1$$

and for any $\epsilon > 0$

$$\lim_{n\to\infty} \int_{|x|\ge\epsilon} Q_n(x)dx = 0.$$

Proof. The first assertion is obvious from the definition of c_n. For the second we have (using symmetry)

$$\begin{aligned}\int_{|x|\ge\epsilon} Q_n(x)dx &= \frac{2\int_\epsilon^1 (1-x^2)^n dx}{2\int_0^\epsilon (1-x^2)^n dx + 2\int_\epsilon^1 (1-x^2)^n dx} \\ &\le \frac{\int_\epsilon^1 \frac{x}{\epsilon}(1-x^2)^n dx}{\int_0^\epsilon \frac{x}{\epsilon}(1-x^2)^n dx} = \frac{(1-x^2)^{n+1}|_\epsilon^1}{(1-x^2)^{n+1}|_0^\epsilon} \\ &= \frac{(1-\epsilon^2)^{n+1}}{1-(1-\epsilon^2)^{n+1}} \to 0 \text{ as } n\to\infty.\end{aligned}$$

$$\begin{aligned}\int_{|x|\ge\epsilon} Q_n(x)dx &= \frac{2\int_\epsilon^1 (1-x^2)^n dx}{2\int_0^\epsilon (1-x^2)^n dx + 2\int_\epsilon^1 (1-x^2)^n dx} \\ &\le \frac{\int_\epsilon^1 \frac{x}{\epsilon}(1-x^2)^n dx}{\int_0^\epsilon (1-x^2)^n dx} \le \frac{\epsilon^{-1}\int_\epsilon^1 x(1-x^2)^n dx}{(1-\epsilon^2)^n \epsilon}.\end{aligned}$$

Combining this equation with

$$\int_{\epsilon}^{1} x(1-x^2)^n dx = \frac{-(1-x^2)^{n+1}}{2(n+1)}|_{\epsilon}^{1} = \frac{(1-\epsilon^2)^{(n+1)}}{2(n+1)}$$

shows that

$$\int_{|x|\geq\epsilon} Q_n(x)dx \leq \frac{(1-\epsilon^2)^{(n+1)}}{(1-\epsilon^2)^n \epsilon^2 2(n+1)} = \frac{(1-\epsilon^2)}{\epsilon^2 2(n+1)} \to 0 \text{ as } n \to \infty.$$

■

Lemma 5.3. *Suppose that* $\{Q_n\}_{n=1}^{\infty}$ *is a sequence of positive functions on* $\mathbb{R}$ *such that*

$$\int_{\mathbb{R}} Q_n(x)\ dx = 1 \text{ and} \tag{5.2}$$

$$\lim_{n\to\infty} \int_{|x|\geq\epsilon} Q_n(x)dx = 0 \tag{5.3}$$

for all $\epsilon > 0$. *For* $f \in BC(\mathbb{R})$, *let*

$$Q_n * f(x) = \int_{\mathbb{R}} Q_n(x-y)f(y)dy = \int_{\mathbb{R}} Q_n(y)f(x-y)dy.$$

Then $Q_n * f$ *converges to* f *uniformly on compact subsets of* $\mathbb{R}$.

Proof. Let $x \in \mathbb{R}$, then because of Eq. (5.2),

$$|Q_n * f(x) - f(x)| = \left| \int_{\mathbb{R}} Q_n(y) \left(f(x-y) - f(x)\right) dy \right| \leq \int_{\mathbb{R}} Q_n(y) \left|f(x-y) - f(x)\right| dy.$$

Let $M = \sup\{|f(x)| : x \in \mathbb{R}\}$ and $\epsilon > 0$, then by and Eq. (5.2)

$$\begin{aligned} |Q_n * f(x) - f(x)| &\leq \int_{|y|\leq\epsilon} Q_n(y) \left|f(x-y) - f(x)\right| dy \\ &\quad + \int_{|y|>\epsilon} Q_n(y) \left|f(x-y) - f(x)\right| dy \\ &\leq \sup_{|z|\leq\epsilon} |f(x+z) - f(x)| + 2M \int_{|y|>\epsilon} Q_n(y)dy. \end{aligned}$$

Let K be a compact subset of $\mathbb{R}$, then

$$\sup_{x\in K} |Q_n * f(x) - f(x)| \leq \sup_{|z|\leq\epsilon, x\in K} |f(x+z) - f(x)| + 2M \int_{|y|>\epsilon} Q_n(y)dy$$

and hence by Eq. (5.3),

$$\limsup_{n\to\infty} \sup_{x\in K} |Q_n * f(x) - f(x)| \leq \sup_{|z|\leq\epsilon, x\in K} |f(x+z) - f(x)|.$$

This finishes the proof since the right member of this equation tends to 0 as $\epsilon \downarrow 0$ by uniform continuity. ■

Proof. (Weierstrass Approximation Theorem) We begin with two reductions:

1. We may reduce the problem to the case where $a = 0$ and $b = 1$ by considering the function

$$g(x) = f(a + x(b-a)) \text{ for } x \in [0, 1].$$

2. Moreover, by considering the function $g_2(x) = g(x) - g(0) - x(g(1) - g(0))$ we may assume that $f \in C([0,1], \mathbb{R})$ and $f(0) = f(1) = 0$.

So suppose that $f \in C([0,1], \mathbb{R})$ and $f(0) = f(1) = 0$ and that f has been extended to all $\mathbb{R}$ by setting $f = 0$ on $\mathbb{R} \backslash [0,1]$. Let $Q_n(x)$ be defined as in Eq. (5.1). Then by Lemma 5.2 and 5.3, $p_n(x) := (Q_n * f)(x) \to f(x)$ uniformly for $x \in [0,1]$ as $n \to \infty$. So to finish the proof it only remains to show $p_n(x)$ is a polynomial when $x \in [0,1]$. This follows for $x \in [0,1]$ since

$$(1 - (x-y)^2)^n = \sum_{k=0}^{2n} a_k(y) x^k$$

where $a_k(y)$ are polynomials in y and therefore

$$\begin{aligned} p_n(x) &= \int_{\mathbb{R}} Q_n(x-y) f(y) dy = \frac{1}{c_n} \int_0^1 f(y)(1 - (x-y)^2)^n 1_{|x-y| \le 1} dy \\ &= \frac{1}{c_n} \int_0^1 f(y)(1 - (x-y)^2)^n dy \\ &= \frac{1}{c_n} \sum_{k=0}^{2n} \left(\int_0^1 f(y) a_k(y) dy \right) x^k. \end{aligned}$$

■

Proof. (The second proof of the Weierstrass Approximation Theorem) As in the first proof it suffices to assume that $f \in C([0,1])$. For $x \in [0,1]$ let μ_x be the measure on $\{0,1\}$ given by $\mu_x(\{0\}) = 1 - x$ and $\mu_x(\{1\}) = x$. For $n \in \mathbb{N}$ and $x \in [0,1]$, let μ_x^n denote the $n -$fold product of μ_x with itself on $\Omega_n := \{0,1\}^n$ and let $X_i(\omega) = \omega_i$ for $\omega \in \{0,1\}^n$. We also let $S_n = (X_1 + X_2 + \cdots + X_n)/n$. The law of large numbers then states that S_n should be close to

$$\int_{\Omega_n} X_i d\mu_x^n = \int_{\{0,1\}} \omega d\mu_x(\omega) = 1 \cdot x + 0 \cdot (1 - x) = x$$

when n is large Let us define

$$p_n(x) := \int_{\Omega_n} f(S_n) d\mu_x^n = \sum_{\omega \in \Omega_n} f\left(\frac{\omega_1 + \omega_2 + \cdots + \omega_n}{n} \right) \prod_{i=1}^n x^{\omega_i} (1-x)^{1-\omega_i}.$$

The later shows that $p_n(x)$ is a polynomial in x of degree at most n. By the law of large numbers we expect that $p_n(x)$ is close to $f(x)$, a fact which we will now verify. Let $\epsilon > 0$ be given, then since $\mu_x^n(\Omega_n) = 1$,

$$\begin{aligned} |f(x) - p_n(x)| &= \left| \int_{\Omega_n} (f(x) - f(S_n)) d\mu_x^n \right| \le \int_{\Omega_n} |f(x) - f(S_n)| d\mu_x^n \\ &\le \int_{\{|S_n - x| > \epsilon\}} |f(x) - f(S_n)| d\mu_x^n + \int_{\{|S_n - x| \le \epsilon\}} |f(x) - f(S_n)| d\mu_x^n \\ &\le 2M \mu_x^n (|S_n - x| > \epsilon) \\ &\quad + \sup \{ |f(y) - f(x)| : y \in [0,1] \text{ and } |y - x| \le \epsilon \} \\ &\le 2M \mu_x^n (|S_n - x| > \epsilon) + \delta_\epsilon \end{aligned} \tag{5.4}$$

where $M = \sup \{ |f(x)| : x \in [0,1] \}$ and

$$\delta_\epsilon = \sup \{ |f(y) - f(x)| : x, y \in [0,1] \text{ and } |y - x| \le \epsilon \}.$$

By uniform continuity, $\delta_\epsilon \to 0$ as $\epsilon \to 0$ and by Chebyshev's inequality,

$$\mu_x^n\left(|S_n - x| > \epsilon\right) \le \frac{1}{\epsilon^2}\int_{\Omega_n}(S_n - x)^2 d\mu_x^n = \frac{1}{\epsilon^2}\int_{\Omega_n}(\frac{1}{n}\sum_{k=1}^n (X_k - x))^2 d\mu_x^n$$
$$= \frac{1}{n^2\epsilon^2}\int_{\Omega_n}\sum_{k,j=1}^n (X_k - x)(X_j - x) d\mu_x^n.$$

By Fubini's theorem, it follows that for $k \ne j$ that

$$\int_{\Omega_n}(X_k - x)(X_j - x)d\mu_x^n = \left[\int_{\{0,1\}}(\omega - x)\,d\mu_x(\omega)\right]^2 = 0$$

and

$$\int_{\Omega_n}(X_k - x)^2 d\mu_x^n = \int_{\{0,1\}}(\omega - x)^2\,d\mu_x(\omega) = (1-x)^2 x + x^2(1-x) \le 2.$$

Combining the last three displayed equations shows that

$$\mu_x^n\left(|S_n - x| > \epsilon\right) \le \frac{1}{n^2\epsilon^2}2n = \frac{2}{n\epsilon^2}$$

which combined with Eq. (5.4) implies that

$$\sup_{x\in[0,1]}|f(x) - p_n(x)| \le \frac{4M}{n\epsilon^2} + \delta_\epsilon$$

and therefore

$$\limsup_{n\to\infty}\sup_{x\in[0,1]}|f(x) - p_n(x)| \le \delta_\epsilon \to 0 \text{ as } \epsilon \to 0.$$

■

5.2. **The Stone-Weierstrass Theorem.** We now wish to generalize Theorem 5.1 to more general topological spaces. We will first need some definitions.

Definition 5.4. Let X be a topological space and $\mathcal{A} \subset C(X) = C(X,\mathbb{R})$ or $C(X,\mathbb{C})$ be a collection of functions. Then

1. $\mathcal{A}$ is said to **separate points** if for all distinct points $x, y \in X$ there exists $f \in \mathcal{A}$ such that $f(x) \ne f(y)$.
2. $\mathcal{A}$ is an **algebra** if $\mathcal{A}$ is a vector subspace of $C(X)$ which is closed under pointwise multiplication.
3. $\mathcal{A}$ is called a **lattice** if $f \vee g := \max(f,g)$ and $f \wedge g = \min(f,g) \in \mathcal{A}$ for all $f, g \in \mathcal{A}$.
4. $\mathcal{A} \subset C(X)$ is closed under conjugation if $\bar{f} \in \mathcal{A}$ whenever $f \in \mathcal{A}$.[6]

Remark 5.5. If X is a topological space such that $C(X,\mathbb{R})$ separates points then X is Hausdorff. Indeed if $x, y \in X$ and $f \in C(X,\mathbb{R})$ such that $f(x) \ne f(y)$, then $f^{-1}(J)$ and $f^{-1}(I)$ are disjoint open sets containing x and y respectively when I and J are disjoint intervals containing $f(x)$ and $f(y)$ respectively.

Lemma 5.6. *If $\mathcal{A} \subset C(X,\mathbb{R})$ is a closed algebra then $|f| \in \mathcal{A}$ for all $f \in \mathcal{A}$ and $\mathcal{A}$ is a lattice.*

[6]This is of course no restriction when $C(X) = C(X,\mathbb{R})$.

Proof. Let $f \in \mathcal{A}$ and let $M = \sup_{x\in X} |f(x)|$. Using Theorem 5.1, there are polynomials $P_n(t)$ such that

$$\lim_{n\to\infty} \sup_{|t|\le M} ||t| - P_n(t)| = 0.$$

By replacing P_n by $P_n - P_n(0)$ if necessary we may assume that $P_n(0) = 0$. Sine $\mathcal{A}$ is an algebra, it follows that $f_n = P_n(f) \in \mathcal{A}$ and $|f| \in \mathcal{A}$, $|f|$ being the uniform limit of the f_n's. This also shows that $\mathcal{A}$ is a lattice since

$$f \vee g = \frac{1}{2}\,(f + g + |f - g|)$$
$$f \wedge g = \frac{1}{2}\,(f + g - |f - g|).$$

■

Lemma 5.7. *Let $\mathcal{A} \subset C(X,\mathbb{R})$ be an algebra which separates points and $x, y \in X$ be distinct points such that*

$$\exists\, f, g \in \mathcal{A} \text{ such that } f(x) \neq 0 \textit{ and } g(y) \neq 0. \tag{5.5}$$

Then

$$V := \{(f(x), f(y)) : f \in \mathcal{A}\} = \mathbb{R}^2. \tag{5.6}$$

Proof. It is clear that V is a non-zero subspace of $\mathbb{R}^{2.}$ If $\dim(V) = 1$, then $V = \mathrm{span}(a,b)$ with $a \neq 0$ and $b \neq 0$ by the assumption in Eq. (5.5). Since $(a,b) = (f(x), f(y))$ for some $f \in \mathcal{A}$ and $f^2 \in \mathcal{A}$, it follows that $(a^2, b^2) = (f^2(x), f^2(y)) \in V$ as well. Since $\dim V = 1$, (a,b) and (a^2, b^2) are linearly dependent and therefore

$$0 = \det \begin{pmatrix} a & a^2 \\ b & b^2 \end{pmatrix} = ab^2 - ba^2 = ab(b-a)$$

which implies that $a = b$. But this the implies that $f(x) = f(y)$ for all $f \in \mathcal{A}$, violating the assumption that $\mathcal{A}$ separates points. Therefore we conclude that $\dim(V) = 2$, i.e. that $V = \mathbb{R}^2$. ■

Theorem 5.8 (Stone-Weierstrass Theorem)**.** *Suppose X is a compact Hausdorff space. $\mathcal{A} \subset C(X,\mathbb{R})$ is a **closed** subalgebra which separates points and for $x \in X$ let*

$$\mathcal{A}_x \equiv \{f(x) : f \in \mathcal{A}\} \textit{ and}$$
$$\mathcal{I}_x = \{f \in C(X,\mathbb{R}) : f(x) = 0\}.$$

Then either one of the following two cases hold.

1. *$\mathcal{A}_x = \mathbb{R}$ for all $x \in X$, i.e. for all $x \in X$ there exists $f \in \mathcal{A}$ such that $f(x) \neq 0$.*[7]
2. *There exists a unique point $x_0 \in X$ such that $\mathcal{A}_{x_0} = \{0\}$.*

Moreover in case (1) $\mathcal{A} = C(X,\mathbb{R})$ and in case (2) $\mathcal{A} = \mathcal{I}_{x_0} = \{f \in C(X,\mathbb{R}) : f(x_0) = 0\}$.

Proof. If there exists x_0 such that $\mathcal{A}_{x_0} = \{0\}$ (x_0 is unique since $\mathcal{A}$ separates points) then $\mathcal{A}_{x_0} \subset \mathcal{I}_{x_0}$. If such an x_0 exists let $\mathcal{C} = \mathcal{I}_{x_0}$ and if $\mathcal{A}_x = \mathbb{R}$ for all x,

[7] If $\mathcal{A}$ contains the constant function 1, then this hypothesis holds.

set $\mathcal{C} = C(X, \mathbb{R})$. Let $f \in \mathcal{C}$, then by Lemma 5.7, for all $x, y \in X$ such that $x \neq y$ there exists $g_{xy} \in \mathcal{A}$ such that $f = g_{xy}$ on $\{x, y\}$.[8]

Claim 5.9. *Given $\epsilon > 0$ and $x \in X$ there exists $g_x \in \mathcal{A}$ such that $g_x(x) = f(x)$ and $f < g_x + \epsilon$ on X.*

To prove the claim, let V_y be an open neighborhood of y such that $|f - g_{xy}| < \epsilon$ on V_y so in particular $f < \epsilon + g_{xy}$ on V_y. By compactness, there exists $\Lambda \subset\subset X$ such that $X = \bigcup_{y \in \Lambda} V_y$. Set

$$g_x(z) = \max\{g_{xy}(z) : y \in \Lambda\},$$

then for any $y \in \Lambda$, $f < \epsilon + g_{xy} < \epsilon + g_x$ on V_y and therefore $f < \epsilon + g_x$ on X. Moreover, by construction $f(x) = g_x(x)$.

We now will finish the proof of the theorem. For each $x \in X$, let U_x be a neighborhood of x such that $|f - g_x| < \epsilon$ on U_x. Choose $\Gamma \subset\subset X$ such that $X = \bigcup_{x \in \Gamma} U_x$ and define

$$g = \min\{g_x : x \in \Gamma\} \in \mathcal{A}.$$

Then $f < g + \epsilon$ on X and for $x \in \Gamma$, $g_x < f + \epsilon$ on U_x and hence $g < f + \epsilon$ on U_x. Since $X = \bigcup_{x \in \Gamma} U_x$, we conclude that

$$f < g + \epsilon \text{ and } g < f + \epsilon \text{ on } X,$$

i.e. that $|f - g| < \epsilon$ on X. Since $\epsilon > 0$ is arbitrary it follows that $f \in \bar{\mathcal{A}} = \mathcal{A}$. ■

Theorem 5.10 (Complex Stone-Weierstrass Theorem)**.** *Let X be a compact Hausdorff space. Suppose $\mathcal{A} \subset C(X, \mathbb{C})$ is closed in the uniform topology, separates points, and is closed under conjugation. Then either $\mathcal{A} = C(X)$ or $\mathcal{A} = \mathcal{I}_{x_0}$ for some $x_0 \in X$.*

Proof. Since

$$\text{Re } f = \frac{f + \bar{f}}{2} \text{ and } \text{Im } f = \frac{f - \bar{f}}{2i},$$

Re f and Im f are both in $\mathcal{A}$.Therefore

$$\mathcal{A}_{\mathbb{R}} = \{\text{Re } f, \text{Im } f : f \in \mathcal{A}\}$$

is a real sub-algebra of $C(X, \mathbb{R})$ which separates points. Therefore either $\mathcal{A}_{\mathbb{R}} = C(X, \mathbb{R})$ or $\mathcal{A}_{\mathbb{R}} = \mathcal{I}_{x_0} \cap C(X, \mathbb{R})$ for some x_0 and hence $\mathcal{A} = C(X, \mathbb{C})$ or $\mathcal{I}_{x_0}$ respectively. ■

Example 5.11. Let $X = S^1 = \{z \in \mathbb{C} : |z| = 1\}$ and $\mathcal{A}$ be the set of polynomials in z and $\bar{z}$, i.e. $\mathcal{A}$ is the algebra generated by $\{1, z, \overline{z}\}$. Then $\mathcal{A}$ separates points, $1 \in \mathcal{A}$ and $\mathcal{A}$ is closed under conjugation allows us to conclude from Theorem 5.10 that $\bar{\mathcal{A}} = C(X)$.

Corollary 5.12. *Given a continuous function $f : \mathbb{R} \to \mathbb{C}$ which is 2π -periodic and $\epsilon > 0$ there exists $p(\theta) = \sum_{n=-N}^{n} \alpha_n e^{in\theta}$ such that $|f(\theta) - P(\theta)| < \epsilon$ for all $\theta \in \mathbb{R}$.*

[8] If we are in the case whrere $\mathcal{A}_{x_0} = \{0\}$ and $x = x_0$ or $y = x_0$, then g_{xy} exists merely by the fact that $\mathcal{A}$ separates points.

Theorem 5.13. *Let X be non-compact locally compact Hausdorff space, space. If $\mathcal{A}$ is a closed subalgebra of $C_0(X,\mathbb{R})$ which separates points. Then either $\mathcal{A} = C_0(X,\mathbb{R})$ or there exists $x_0 \in X$ such that $\mathcal{A} = \{f \in C_0(X,\mathbb{R}) : f(x_0) = 0\}$.*

Proof. If there exists $x_0 \in X$ such $\mathcal{A} \subset \{f \in C_0(X,\mathbb{R}) : f(x_0) = 0\}$ let X^* be the one point compactification of $X \setminus \{x_0\}$ otherwise let X^* be the one point compactification of X. Now apply Theorem 5.8 to learn that $\mathcal{A} \subset C(X^*,\mathbb{R})$ is $\mathcal{A} = \mathcal{I}_\infty = C_0(X^* \setminus \{\infty\}, \mathbb{R})$. ■

Example 5.14. Let $X = [0,\infty)$, $\lambda > 0$ be fixed, $\mathcal{A}$ be the algebra generate by $t \to e^{-\lambda t}$. So the general element $f \in \mathcal{A}$ is of the form $f(t) = p(e^{-\lambda t})$, where $p(x)$ is a polynomial. Since $\mathcal{A} \subset C_0(X,\mathbb{R})$ separates points and is pointwise positive, $\bar{\mathcal{A}} = C_0(X,\mathbb{R})$.

As an application of this example, we will shows that the Laplace transform is injective.

Theorem 5.15. *For $f \in L^1([0,\infty), dx)$, let the Laplace transform is defined by*

$$\mathcal{L}f(\lambda) \equiv \int_0^\infty e^{-\lambda x} f(x)dx \text{ for all } \lambda > 0.$$

If $\mathcal{L}f(\lambda) \equiv 0$ then $f(x) = 0$ for m -a.e. x.

Proof. Suppose that $f \in L^1([0,\infty), dx)$ such that $\mathcal{L}f(\lambda) \equiv 0$. Let $g \in C_0([0,\infty),\mathbb{R})$ and $\epsilon > 0$ be given. Choose $\{a_\lambda\}_{\lambda>0}$ such that $\{\lambda > 0 : a_\lambda \neq 0\}$ is a finite set and

$$|g(x) - \sum_{\lambda>0} a_\lambda e^{-\lambda x}| < \epsilon \text{ for all } x \geq 0.$$

Then

$$\begin{aligned}
\left|\int_0^\infty g(x)f(x)dx\right| &= \left|\int_0^\infty \left(g(x) - \sum_{\lambda>0} a_\lambda e^{-\lambda x}\right) f(x)dx\right| \\
&\leq \int_0^\infty \left|g(x) - \sum_{\lambda>0} a_\lambda e^{-\lambda x}\right| |f(x)|\, dx \\
&\leq \epsilon \|f\|_1.
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, it follows that $\int_0^\infty g(x)f(x)dx = 0$ for all $g \in C_0([0,\infty),\mathbb{R})$. Let

$$\operatorname{sgn} f(x) = \begin{cases} \frac{f(x)}{|f(x)|} & \text{if } f(x) \neq 0 \\ 0 & \text{if } f(x) = 0 \end{cases}$$

and define $d\mu(x) \equiv |f(x)|dm(x)$. Then μ is a measure on $[0,\infty)$ such that $\mu([0,\infty)) < \infty$. By our regularity theorems, we know that $C_c([0,\infty))$ is dense in $L^1(\mu)$. Therefore there exists $g_n \in C_c([0,\infty))$ such that $g_n \to \operatorname{sgn} f$ in $L^1(\mu)$. Therefore,

$$\begin{aligned}
0 &= \int_0^\infty g_n(x)f(x)dx \\
&= \int_0^\infty g_n(x)\operatorname{sgn} f(x)d\mu(x) \overset{n\to\infty}{\to} \int_0^\infty |\operatorname{sgn} f(x)|^2 d\mu(x) = \int_0^\infty |f(x)|dm(x).
\end{aligned}$$

■

6. Product spaces and Tychonoff's Theorem

6.1. **Product Spaces.** Let $\{(X_\alpha, \tau_\alpha)\}_{\alpha \in A}$ be a collection of topological spaces (we assume $X_\alpha \neq \emptyset$) and let $X_A = \prod_{\alpha \in A} X_\alpha$. Recall that $x \in X_A$ is a function

$$x : A \to \coprod_{\alpha \in A} X_\alpha$$

such that $x_\alpha := x(\alpha) \in X_\alpha$ for all $\alpha \in A$. An element $x \in X_A$ is called a choice function and the **axiom of choice** states that $X_A \neq \emptyset$ provided that $X_\alpha \neq \emptyset$ for each $\alpha \in A$. If each X_α above is the same set X, we will denote $X_A = \prod_{\alpha \in A} X_\alpha$ by X^A. So $x \in X^A$ is a function from A to X.

Notation 6.1. For $\alpha \in A$, let $\pi_\alpha : X_A \to X_\alpha$ be the canonical projection map, $\pi_\alpha(x) = x_\alpha$. The **product topology** $\tau = \otimes_{\alpha \in A} \tau_\alpha$ is the smallest topology on X_A such that each projection π_α is continuous. Explicitly, τ is the topology generated by

$$\mathcal{E} = \{\pi_\alpha^{-1}(V_\alpha) : \alpha \in A, V_\alpha \in \tau_\alpha\}. \tag{6.1}$$

A "basic" open set in this topology is of the form

$$V = \{x \in X_A : \pi_\alpha(x) \in V_\alpha \text{ for } \alpha \in \Lambda\} \tag{6.2}$$

where Λ is a finite subset of A and $V_\alpha \in \tau_\alpha$ for all $\alpha \in \Lambda$. We will sometimes write V above as

$$V = \prod_{\alpha \in \Lambda} V_\alpha \times \prod_{\alpha \notin \Lambda} X_\alpha = V_\Lambda \times X_{A \setminus \Lambda}.$$

Proposition 6.2. *Suppose Y is a topological space and $f : Y \to X_A$ is a map. Then f is continuous iff $\pi_\alpha \circ f : Y \to X_\alpha$ is continuous for all $\alpha \in A$.*

Proof. If f is continuous then $\pi_\alpha \circ f$ is the composition of two continuous functions and hence is continuous. Conversely if $\pi_\alpha \circ f$ is continuous for all $\alpha \in A$, the $(\pi_\alpha \circ f)^{-1}(V_\alpha) = f^{-1}(\pi_\alpha^{-1}(V_\alpha))$ is open in Y for all $\alpha \in A$ and $V_\alpha \subset_o X_\alpha$. That is to say, $f^{-1}(\mathcal{E})$ consists of open sets, and therefore f is continuous since $\mathcal{E}$ is a subbase for the product topology. ∎

Proposition 6.3. *Suppose that (X, τ) is a topological space and $\{f_n\} \subset X^A$ is a sequence. Then $f_n \to f$ in the product topology of X^A iff $f_n(\alpha) \to f(\alpha)$ for all $\alpha \in A$.*

Proof. Since π_α is continuous, if $f_n \to f$ then $f_n(\alpha) = \pi_\alpha(f_n) \to \pi_\alpha(f) = f(\alpha)$ for all $\alpha \in A$. Conversely, $f_n(\alpha) \to f(\alpha)$ for all $\alpha \in A$ iff $\pi_\alpha(f_n) \to \pi_\alpha(f)$ for all $\alpha \in A$. Therefore if $V = \pi_\alpha^{-1}(V_\alpha) \in \mathcal{E}$ and $f \in V$, then $\pi_\alpha(f) \in V_\alpha$ and $\pi_\alpha(f_n) \in V_\alpha$ a.a and hence $f_n \in V$ a.a. This shows that $f_n \to f$ as $n \to \infty$. ∎

Proposition 6.4. *Let (X_α, τ_α) be topological spaces and X_A be the product space with the product topology.*

1. *If X_α is Hausdorff for all $\alpha \in A$, then so is X_A.*
2. *If each X_α is connected for all $\alpha \in A$, then so is X_A.*

Proof.

1. Let $x, y \in X_A$ be distinct points. Then there exists $\alpha \in A$ such that $\pi_\alpha(x) = x_\alpha \neq y_\alpha = \pi_\alpha(y)$. Since X_α is Hausdorff, there exists disjoint open sets $U, V \subset X_\alpha$ such $\pi_\alpha(x) \in U$ and $\pi_\alpha(y) \in V$. Then $\pi_\alpha^{-1}(U)$ and $\pi_\alpha^{-1}(V)$ are disjoint open sets in X_A containing x and y respectively.
2. Let us begin with the case of two factors, namely assume that X and Y are connected topological spaces, then we will show that $X \times Y$ is connected as well. To do this let $p = (x_0, y_0) \in X \times Y$ and E denote the connected component of p. Since $\{x_0\} \times Y$ is homeomorphic to Y, $\{x_0\} \times Y$ is connected in $X \times Y$ and therefore $\{x_0\} \times Y \subset E$, i.e. $(x_0, y) \in E$ for all $y \in Y$. A similar argument now shows that $X \times \{y\} \subset E$ for any $y \in Y$, that is to $X \times Y = E$. By induction the theorem holds whenever A is a finite set.

   For the general case, again choose a point $p \in X_A = X^A$ and let $C = C_p$ be the connected component of p in X_A. Recall that C_p is closed and therefore if C_p is a proper subset of X_A, then $X_A \setminus C_p$ is a non-empty open set. By the definition of the product topology, this would imply that $X_A \setminus C_p$ contains an open set of the form
$$V := \cap_{\alpha \in \Lambda} \pi_\alpha^{-1}(V_\alpha) = V_\Lambda \times X_{A \setminus \Lambda}$$
   where $\Lambda \subset\subset A$ and $V_\alpha \in \tau_\alpha$ for all $\alpha \in \Lambda$. We will now show that no such V can exist and hence $X_A = C_p$, i.e. X_A is connected.

   Define $\phi : X_\Lambda \to X_A$ by $\phi(y) = x$ where
$$x_\alpha = \begin{cases} y_\alpha & \text{if} \quad \alpha \in \Lambda \\ p_\alpha & \text{if} \quad \alpha \notin \Lambda. \end{cases}$$
   If $\alpha \in \Lambda$, $\pi_\alpha \circ \phi(y) = y_\alpha = \pi_\alpha(y)$ and if $\alpha \in A \setminus \Lambda$ then $\pi_\alpha \circ \phi(y) = p_\alpha$ so that in every case $\pi_\alpha \circ \phi : X_\Lambda \to X_\alpha$ is continuous and therefore ϕ is continuous.

   Since X_Λ is a product of a finite number of connected spaces it is connected by step 1. above. Hence so is the continuous image, $\phi(X_\Lambda) = X_\Lambda \times \{p_\alpha\}_{\alpha \in A \setminus \Lambda}$, of X_Λ. Now $p \in \phi(X_\Lambda)$ and $\phi(X_\Lambda)$ is connected implies that $\phi(X_\Lambda) \subset C$. On the other hand one easily sees that
$$\emptyset \neq V \cap \phi(X_\Lambda) \subset V \cap C$$
   contradicting the assumption that $V \subset C^c$.

■

6.2. Tychonoff's Theorem.

The main theorem of this subsection is that the product of compact spaces is compact.

Theorem 6.5 (Tychonoff's Theorem). *Let $\{X_\alpha\}_{\alpha \in A}$ be a collection of non-empty compact spaces. Then $X_A = \prod_{\alpha \in A} X_\alpha$ is compact in the product space topology.*

The proof of this theorem requires Zorn's lemma which is equivalent to the axiom of choice, (see Theorem 8.7 below). Before going into the details of the proof, let us first prove the following special case.

Proposition 6.6. *Suppose that X and Y are non-empty compact topological spaces, then $X \times Y$ is compact in the product topology.*

Proof. Let $\mathcal{U}$ be an open cover of $X \times Y$. Then for each $(x, y) \in X \times Y$ there exist $U \in \mathcal{U}$ such that $(x, y) \in U$. By definition of the product topology, there also exist $V_x \in \tau_x^X$ and $W_y \in \tau_y^Y$ such that $V_x \times W_y \subset U$. Therefore $\mathcal{V} :=$

$\{V_x \times W_y : (x,y) \in X \times Y\}$ is also an open cover of $X \times Y$. We will now show that $\mathcal{V}$ has a finite sub-cover, say $\mathcal{V}_0 \subset\subset \mathcal{V}$. Assuming this is proved for the moment, this implies that $\mathcal{U}$ also has a finite subcover because each $V \in \mathcal{V}_0$ is contained in some $U_V \in \mathcal{U}$. So to complete the proof it suffices to show that every cover $\mathcal{V}$ of the form $\mathcal{V} = \{V_\alpha \times W_\alpha : \alpha \in A\}$ where $V_\alpha \subset_o X$ and $W_\alpha \subset_o Y$ has a finite subcover.

Given $x \in X$, let $f_x : Y \to X \times Y$ be the map $f_x(y) = (x,y)$ and notice that f_x is continuous since $\pi_X \circ f_x(y) = x$ and $\pi_Y \circ f_x(y) = y$ are continuous maps. From this we conclude that $\{x\} \times Y = f_x(Y)$ is compact. Similarly, it follows that $X \times \{y\}$ is compact for all $y \in Y$.

Since $\mathcal{V}$ is a cover of $\{x\} \times Y$, there exist $\Gamma_x \subset\subset A$ such that $\{x\} \times Y \subset \bigcup_{\alpha\in\Gamma_x} (V_\alpha \times W_\alpha)$ without loss of generality we may assume that Γ_x is chosen so that $x \in V_\alpha$ for all $\alpha \in \Gamma_x$. Let $U_x \equiv \bigcap_{\alpha\in\Gamma_x} V_\alpha \subset_o X$. Then $\{U_x\}_{x\in X}$ is an open cover of X which is compact, hence there exists $\Lambda \subset\subset X$ such that $X = \cup_{x\in\Lambda} U_x$. The proof is completed by showing that $\mathcal{V}_0 := \cup_{x\in\Lambda} \cup_{\alpha\in\Gamma_x} \{V_\alpha \times W_\alpha\}$ is a cover of $X \times Y$,

$$\cup\mathcal{V}_0 = \cup_{x\in\Lambda} \cup_{\alpha\in\Gamma_x} (V_\alpha \times W_\alpha) \supset \cup_{x\in\Lambda} \cup_{\alpha\in\Gamma_x} (U_x \times W_\alpha) = \cup_{x\in\Lambda} (U_x \times Y) = X \times Y.$$

■

The next lemma is the crux of the proof of the general case of Theorem 6.5.

Lemma 6.7 (Alexander's Lemma). *Let (X,τ) be a topological space. Let $\mathcal{E} \subset \tau$ be a subbase for the topology of τ, $\tau = \tau(\mathcal{E})$. Then X is compact iff every open cover $\mathcal{B} \subset \mathcal{E}$ of X has a finite subcover:*

Proof. It is clear that if X is compact and $\mathcal{B} \subset \mathcal{E}$ is a cover of X, then $\mathcal{B}$ has a finite subcover. So it suffices to show that if X is not compact, there exists a cover $\mathcal{B} \subset \mathcal{E}$ with not finite subcover. We will construct $\mathcal{B}$ in two steps.

1. There is a maximal open cover $\mathcal{A}$ of X with no finite subcover. Let $\mathcal{U}$ be the collection of open covers of X with **no** finite subcovers and partially ordered $\mathcal{U}$ by inclusion. ($\mathcal{U}$ is not empty by definition since X is not compact.) Suppose that $\{\mathcal{A}_\beta\}_{\beta\in B} \subset \mathcal{U}$ is a linearly ordered collection. Then if $U_1, \ldots, U_n \in \bigcup_{\beta\in B} \mathcal{A}_\beta$, by linear order of $\{\mathcal{A}_\beta\}_{\beta\in B}$, there exists $\beta_0 \in B$ such that $U_1, \ldots U_n \in \mathcal{A}_{\beta_0}$ for some $\beta_0 \in B$. Since $\mathcal{A}_{\beta_0} \in \mathcal{U}$ is an open cover of X with no finite subcover, it follows that $\bigcup_{i=1}^n U_i \neq X$ and hence $\bigcup_{\beta\in B} \mathcal{A}_\beta$ is a cover of X with no finite subcover. Hence $\bigcup_{\beta\in\mathcal{B}} \mathcal{A}_\beta \in \mathcal{U}$ is an upper bound for $\{\mathcal{A}_\beta\}_{\beta\in\mathcal{B}}$, so we may apply Zorn's Lemma to conclude that $\mathcal{U}$ has a maximal element which we denote by $\mathcal{A}$.
2. We will finish that proof by showing that $\mathcal{B} \equiv \mathcal{A} \cap \mathcal{E}$ is a cover of X. (Notice that $\mathcal{B}$ can not have a finite subcover, since if it did $\mathcal{A}$ would also have finite subcover.) For sake of contradiction, suppose that there exists $x \in X \setminus \cup\mathcal{B}$. Since $\mathcal{A}$ covers X there exists $U \in \mathcal{A}$ such that $x \in U$ and because $\tau = \tau(\mathcal{E})$, there exists $V_1, \ldots, V_n \in \mathcal{E}$ such that
$$x \in V_1 \cap \cdots \cap V_n \subset U.$$
Notice that no V_j is in $\mathcal{A}$, for if one were (say V_j), we would have $V_j \in \mathcal{B}$ and hence $x \in \cup\mathcal{B}$. Because $\mathcal{A}$ is maximal, $\mathcal{A} \cup \{V_j\}$ must have a finite subcover for each j. Hence for each $j = 1, 2, \ldots, n$ there exists $\mathcal{A}_j \subset\subset \mathcal{A}$ such that if

$W_j = \cup \mathcal{A}_j$ then $V_j \cup W_j = X$. Now

$$U \cup \bigcup_1^n W_j \supseteq \left(\bigcap_1^n V_j\right) \cup \left(\bigcup_1^n W_j\right) = X$$

which implies that $\mathcal{A}' := \{U\} \cup \cup_{j=1}^n \mathcal{A}_j$ is a finite subcover of $\mathcal{A}$.[9] This is a contraction to the fact that $\mathcal{A}$ had no finite subcover and hence $\cup\mathcal{B} = X$.

■

We now give the proof of Tychonoff's Theorem.

Proof. (Tychonoff's Theorem) By Alexander's Lemma 6.7, we must shows that every open cover $\mathcal{V}$, which is contained in the subbase $\mathcal{E} = \{\pi_\alpha^{-1}(U) : U \subset_0 X_\alpha, \alpha \in A\}$, has a finite subcover. If $\mathcal{V}$ is such an open cover and $\alpha \in A$, let

$$\mathcal{V}_\alpha \equiv \{U \subset_0 X_\alpha : \pi_\alpha^{-1}(U) \in \mathcal{V}\}.$$

Claim: There exists some $\beta \in A$ such that $\mathcal{V}_\beta$ is a cover of X_β.

If there were no such β, then $V_\beta := \cup\mathcal{V}_\beta \subsetneq X_\beta$ for all $\beta \in A$. So by the Axiom of choice there would exists some $x \in \prod_{\alpha\in A} (X_\beta \setminus V_\beta)$. On the other hand

$$\mathcal{V} = \bigcup_{\alpha\in A} \pi_\alpha^{-1}(\mathcal{V}_\alpha) \equiv \bigcup_{\alpha\in A} \{\pi_\alpha^{-1}(U) : U \in \mathcal{V}_\alpha\}$$

and $x \in \cup\mathcal{V}$ ($\mathcal{V}$ is a cover of X_A), so there must exist some $\beta \in A$ and $U_\beta \in \mathcal{V}_\beta$ such that $\pi_\beta(x) \in U_\beta$. But this is precluded by choice of $x \in \prod_{\alpha\in A} (X_\beta \setminus V_\beta)$. This proves the claim.

Let $\beta \in A$ such that $\mathcal{V}_\beta$ is a cover of X_β. Using the compactness of X_β, there exists $\{U_i\}_{i=1}^n \subset \mathcal{V}_\beta$ ($n < \infty$) such that $\{U_i\}_{i=1}^n$ covers X_β. Then $\left\{\pi_\beta^{-1}(U_i)\right\}_{i=1}^n \subset \mathcal{V}$ is now the desired subcover of X_A. ■

7. Urysohn's metrization Theorem

Definition 7.1. Let $I = [0,1]$ and A be a non-empty set, then I^A with the product topology is called a cube.

Definition 7.2. Let X be a topological space. A subset $\mathcal{F} \subset C(X, I)$ is said to **separates points and closed sets** if for all $x \in X$ and closed sets $F \subset X$ such that $x \notin F$, there exists $f \in \mathcal{F}$ such that $f(x) \notin \overline{f(F)}$.

Remark 7.3. If f is as in Definition 7.2, we may find $h \in C(I, I)$ such that $h = 0$ on $\overline{f(F)}$ and $h(f(x)) = 1$. Then $g = h \circ f \in C(X, I)$ has the property that $g(x) = 1$ and $g(F) = \{0\}$.

Notation 7.4. Given $\mathcal{F} \subset C(X, I)$, define $e : X \to I^{\mathcal{F}}$ by $x \to e_x$ where $e_x(f) = f(x)$. The function $e_x \in I^{\mathcal{F}}$ is called the evaluation map at $x \in X$.

Proposition 7.5. *Let X be a topological space and $\mathcal{F} \subset C(X, I)$. Then*

1. *The map $e : X \to I^{\mathcal{F}}$ is continuous.*
2. *If $\mathcal{F}$ separates points then e is injective.*

---

[9] Let $\widetilde{X} \equiv (\bigcap_1^n V_j) \cup (\bigcup_1^n W_j)$ and notice that for all $x \in X = V_j \cup W_j$, $x \in V_j$ or $x \in W_j$. Hence if $x \in \bigcup_j W_j$ then $x \in \widetilde{X}$ and if $x \notin \bigcup_j W_j$ then $x \in V_j$ for all j i.e. $x \in \cap V_j$ which again shows that $x \in \widetilde{X}$. That is to say $X \subseteq \widetilde{X}$ and hence $X = \tilde{X}$.

3. *If $\mathcal{F}$ separates points and closed sets; then e is a homeomorphism onto its range.*

Proof. (1) For $f \in \mathcal{F}$, let $\pi_f : I^{\mathcal{F}} \to I$ be the projection map, $\pi_f(z) = z(f)$. Then

$$\pi_f \circ e(x) = \pi_f(e_x) = e_x(f) = f(x)$$

which shows that $\pi_f \circ e = f$ is continuous for all $f \in \mathcal{F}$ and therefore e is continuous.

(2) Since $e_x = e_y$ iff $f(x) = e_x(f) = e_y(f) = f(y)$ for all $f \in \mathcal{F}$, e is injective iff $\mathcal{F}$ separates points.

(3)We must show that e is an open map, i.e. if $U \subset_0 X$ then $e(U)$ is open in $e(X)$ in the relative topology. Let $U \subset_0 X$ and $x \in U$. Since $x \notin U^c$ there exists $f \in \mathcal{F}$ such that $f(x) \notin \overline{f(U^c)}$. Let

$$V \equiv \pi_f^{-1}([\overline{f(U^c)}]^c) = \{z \in I^{\mathcal{F}} : z(f) \notin \overline{f(U^c)}\} \subset I^{\mathcal{F}}.$$

Then

$$V \cap e(X) = \{e_y : f(y) \notin \overline{f(U^c)}\}$$

is a relatively open subset of $e(X)$ and by construction $x \in V \cap e(X)$. We will now finish the proof by showing $V \cap e(X) \subset e(U)$. This will be accomplished by showing if $y \notin U$ then $e_y \notin V$. Now if $y \in U^c$, then $e_y(f) = f(y) \in f(U^c) \subset \overline{f(U^c)}$ and therefore $e_y \notin V$. ■

Theorem 7.6 (Urysohn Metrization Theorem)**.** *Every second countable normal space X is metrizable. Moreover, there is a metric S compatible with the topology such that X is totally bounded and hence the completion of X is compact.*

Proof. The proof of the theorem will be broken into four steps.

1. By Folland problem # 4.76 there exists a countable subset $\mathcal{F} \subset C(X, [0, 1])$ such that for all $x \in X$ and $E \sqsubset X$ there exists $f \in \mathcal{F}$ such that $f(x) = 1$ and $f|_E \equiv 0$.
2. By Folland problem # 4.77 the product topology and $Y \equiv [0, 1]^{\mathcal{F}}$ is metrizable. Let d be a metric on Y compatible with the product topology.
3. For $x \in X$, let $e_x : X \to Y$ be defined by $e_x(f) = f(x)$. Then by Proposition 7.5 the map $e : X \to Y$ define by $x \to e_x$ is a homeomorphism onto $e(Y) = \{e_x : x \in \mathcal{F}\} \subset Y$ when $e(Y)$ is equipped with the relative topology from Y.
4. Define $\rho(x, y) = d(e_x, e_y)$, then since $e : X \to e(X) \subset Y$ is a homeomorphism, ρ is a metric on $e(X)$ compatible with the topology on X. By Tychonoff's Theorem, Y is compact and therefore the closure $\overline{e(X)}^d$, of $e(X)$ in Y is compact as well. Since the completion, $\overline{X}^\rho$, of X is homeomorphic to the $\overline{e(X)}^d$ we conclude that $\overline{X}^\rho$ is compact and hence totally bounded. It now easily follows that X is totally bounded as well.

■

8. Zorn's Lemma and the Hausdorff Maximal Principle

Definition 8.1. A partial order $\leq$ on X is a relation with following properties

(i) If $x \leq y$ and $y \leq z$ then $x \leq z$.
(ii) If $x \leq y$ and $y \leq x$ then $x = y$.
(iii) $x \leq x$ for all $x \in X$.

Example 8.2. Let Y be a set and $X = \mathcal{P}(Y)$. There are two natural partial orders on X.

1. Ordered by inclusion, $A \le B$ is $A \subset B$ and
2. Ordered by reverse inclusion, $A \le B$ if $B \subset A$.

Definition 8.3. Let $(X, \le)$ be a partially ordered set we say X is **linearly** a **totally** ordered if for all $x, y \in X$ either $x \le y$ or $y \le x$. The real numbers $\mathbb{R}$ with the usual order $\le$ is a typical example.

Definition 8.4. Let $(X, \le)$ be a partial ordered set. We say $x \in X$ is a **maximal** element if for all $y \in X$ such that $y \ge x$ implies $y = x$, i.e. there is no element larger than x. An **upper bound** for a subset E of X is an element $x \in X$ such that $x \ge y$ for all $y \in E$.

Example 8.5. Let

$$X = \left\{ \ a = \{1\} \quad b = \{1, 2\} \quad c = \{3\} \quad d = \{2, 4\} \quad e = \{2\} \ \right\}$$

ordered by set inclusion. Then b and d are maximal elements despite that fact that $b \not\le a$ and $a \not\le b$. We also have

- If $E = \{a, e, c\}$, then E has **no** upper bound.

Definition 8.6.
- If $E = \{a, e\}$, then b is an upper bound.
- $E = \{e\}$, then b and d are upper bounds.

Theorem 8.7. *The following are equivalent.*

1. ***The axiom of choice.***
2. ***The Hausdorff Maximal Principle:*** *Every partially ordered set has a* ***maximal*** *(relative to the inclusion order) linearly ordered subset.*
3. ***Zorn's Lemma:*** *If X is partially ordered set such that every linearly ordered subset of X has an upper bound, then X has a maximal element.*[10]

Proof. $(2 \Rightarrow 3)$ Let X be a partially ordered subset as in 3 and let $\mathcal{F} = \{E \subset X : E \text{ is linearly ordered}\}$ which we equip with the inclusion partial ordering. By 2. there exist a maximal element $E \in \mathcal{F}$. By assumption, the linearly ordered set E has an upper bound $x \in X$. The element x is maximal, for if $y \in Y$ and $y \ge x$, then $E \cup \{y\}$ is still an linearly ordered set containing E. So by maximality of E, $E = E \cup \{y\}$, i.e. $y \in E$ and therefore $y \le x$ showing which combined with $y \ge x$ implies that $y = x$.[11]

---

[10] If X is a countable set we may prove Zorn's Lemma by induction. Let $\{x_n\}_{n=1}^\infty$ be an ennumeration of X, and define $E_n \subset X$ inductively as follows. For $n = 1$ let $E_1 = \{x_1\}$, and if E_n have been choosen, let $E_{n+1} = E_n \cup \{x_{n+1}\}$ if x_{n+1} is an upper bound for E_n otherwise let $E_{n+1} = E_n$. The set $E = \cup_{n=1}^\infty E_n$ is a linearly ordered (you check) subset of X and hence by assumtion E has an upper bound, $x \in X$. I claim that his element is maximal, for if there exists $y = x_m \in X$ such that $y \ge x$, then x_m would be an upper bound for E_{m-1} and therefore $y = x_m \in E_m \subset E$. That is to say if $y \ge x$, then $y \in E$ and hence $y \le x$, so $y = x$. (Hence we may view Zorn's lemma as a " jazzed" up version of induction.)

[11] Similalry one may show that $3 \Rightarrow 2$. Let $\mathcal{F} = \{E \subset X : E \text{ is linearly ordered}\}$ and order $\mathcal{F}$ by inclusion. If $\mathcal{M} \subset \mathcal{F}$ is linearly ordered, let $E = \cup\mathcal{M} = \bigcup_{A \in \mathcal{M}} A$. If $x, y \in E$ then $x \in A$ and $y \in B$ for some $A, B \subset \mathcal{M}$. Now $\mathcal{M}$ is linearly ordered by set inclusion so $A \subset B$ or $B \subset A$ i.e. $x, y \in A$ or $x, y \in B$. Sinse A and B are linearly order we must have either $x \le y$ or $y \le x$, that is to say E is linearly ordered. Hence by 3. there exists a maximal element $E \in \mathcal{F}$ which is the assertion in 2.

$(3 \Rightarrow 1)$ Let $\{X_\alpha\}_{\alpha \in A}$ be a collection of non-empty sets, we must show $\prod_{\alpha \in A} X_\alpha$ is not empty. Let $\mathcal{G}$ denote the collection of functions $g : D(g) \to \coprod_{\alpha \in A} X_\alpha$ such that $D(g)$ is a subset of A, and for all $\alpha \in D(g)$, $g(\alpha) \in X_\alpha$. Notice that $\mathcal{G}$ is not empty, for we may let $\alpha_0 \in A$ and $x_0 \in X_\alpha$ and then set $D(g) = \{\alpha_0\}$ and $g(\alpha_0) = x_0$ to construct an element of $\mathcal{G}$. We now put a partial order on $\mathcal{G}$ as follows. We say that $f \le g$ for $f, g \in \mathcal{G}$ provided that $D(f) \subset D(g)$ and $f = g|_{D(f)}$. If $\Phi \subset \mathcal{G}$ is a linearly ordered set, let $D(h) = \cup_{g \in \Phi} D(g)$ and for $\alpha \in D(g)$ let $h(\alpha) = g(\alpha)$. Then $h \in \mathcal{G}$ is an upper bound for Φ. So by Zorn's Lemma there exists a maximal element $h \in \mathcal{G}$. To finish the proof we need only show that $D(h) = A$. If this were not the case, then let $\alpha_0 \in A \setminus D(h)$ and $x_0 \in X_{\alpha_0}$. We may now define $D(\tilde{h}) = D(h) \cup \{\alpha_0\}$ and

$$\tilde{h}(\alpha) = \begin{cases} h(\alpha) & \text{if} \quad \alpha \in D(h) \\ x_0 & \text{if} \quad \alpha = \alpha_0. \end{cases}$$

Then $h \le \tilde{h}$ while $h \ne \tilde{h}$ violating the fact that h was a maximal element.

$(1 \Rightarrow 2)$ Let $(X, \le)$ be a partially ordered set. Let $\mathcal{F}$ be the collection of linearly ordered subsets of X which we order by set inclusion. Given $x_0 \in X$, $\{x_0\} \in \mathcal{F}$ is linearly ordered set so that $\mathcal{F} \ne \emptyset$.

Fix an element $P_0 \in \mathcal{F}$. If P_0 is not maximal there exists $P_1 \in \mathcal{F}$ such that $P_0 \subsetneq P_1$. In particular we may choose $x \notin P_0$ such that $P_0 \cup \{x\} \in \mathcal{F}$. The idea now is to keep repeating this process of adding points $x \in X$ until we construct a maximal element P of $\mathcal{F}$. We now have to take care of some details.

We may assume with out loss of generality that $\tilde{\mathcal{F}} = \{P \in \mathcal{F} : P \text{ is not maximal}\}$ is a non-empty set. For $P \in \tilde{\mathcal{F}}$, let $P^* = \{x \in X : P \cup \{x\} \in \mathcal{F}\}$. As the above argument shows, $P^* \ne \emptyset$ for all $P \in \tilde{\mathcal{F}}$. Using the axiom of choice, there exists $f \in \prod_{P \in \tilde{\mathcal{F}}} P^*$. We now define $g : \mathcal{F} \to \mathcal{F}$ by

$$g(P) = \begin{cases} P & \text{if} \quad P \text{ is maximal} \\ P \cup \{f(x)\} & \text{if} \quad P \text{ is not maximal.} \end{cases} \tag{8.1}$$

The proof is completed by the next lemma which shows that g must have a fixed point $P \in \mathcal{F}$. This fixed point is maximal by construction of g. ■

Lemma 8.8. *The function $g : \mathcal{F} \to \mathcal{F}$ defined in Eq. (8.1) has a fixed point.*[12]

Proof. The **idea of the proof** is as follows. Let $P_0 \in \mathcal{F}$ be chosen arbitrarily. Notice that $\Phi = \left\{g^{(n)}(P_0)\right\}_{n=0}^{\infty} \subset \mathcal{F}$ is a linearly ordered set and it is therefore easily verified that $P_1 = \bigcup_{n=0}^{\infty} g^{(n)}(P_0) \in \mathcal{F}$. Similarly we may repeat the process to construct $P_2 = \bigcup_{n=0}^{\infty} g^{(n)}(P_1) \in \mathcal{F}$ and $P_3 = \bigcup_{n=0}^{\infty} g^{(n)}(P_2) \in \mathcal{F}$, etc. etc. Then take $P_\infty = \cup_{n=0}^\infty P_n$ and start again with P_0 replaced by P_∞. Then keep going this way until eventually the sets stop increasing in size, in which case we have found our fixed point. The problem with this strategy is that we may never win. (This is very reminiscent of constructing measurable sets and the way out is to use measure theoretic like arguments.)

Let us now start the **formal proof.** Again let $P_0 \in \mathcal{F}$ and let $\mathcal{F}_1 = \{P \in \mathcal{F} : P_0 \subset P\}$. Notice that $\mathcal{F}_1$ has the following properties:

[12] Here is an easy proof if the elements of $\mathcal{F}$ happened to all be finite sets and there existed a set $P \in \mathcal{F}$ with a maximal number of elments. In this case the condition that $P \subset g(P)$ would imply that $P = g(P)$, otherwise $g(P)$ would have more elements than P.

1. $P_0 \in \mathcal{F}_1$.
2. If $\Phi \subset \mathcal{F}_1$ is a totally ordered (by set inclusion) subset then $\cup\Phi \in \mathcal{F}_1$.
3. If $P \in \mathcal{F}_1$ then $g(P) \in \mathcal{F}_1$.

Let us call a general subset $\mathcal{F}' \subset \mathcal{F}$ satisfying these three conditions a tower and let

$$\mathcal{F}_0 = \cap \{\mathcal{F}' : \mathcal{F}' \text{ is a tower}\}.$$

Standard arguments show that $\mathcal{F}_0$ is still a tower and clearly is the smallest tower containing P_0. (Morally speaking $\mathcal{F}_0$ consists of all of the sets we were trying to constructed in the "idea section" of the proof.)

We now claim that $\mathcal{F}_0$ is a linearly ordered subset of $\mathcal{F}$. To prove this let $\Gamma \subset \mathcal{F}_0$ be the linearly ordered set

$$\Gamma = \{C \in \mathcal{F}_0 : \text{ for all } A \in \mathcal{F}_0 \text{ either } A \subset C \text{ or } C \subset A\}.$$

Shortly we will show that $\Gamma \subset \mathcal{F}_0$ is a tower and hence that $\mathcal{F}_0 = \Gamma$. That is to say $\mathcal{F}_0$ is linearly ordered. Assuming this for the moment let us finish the proof. Let $P \equiv \cup\mathcal{F}_0$ which is in $\mathcal{F}_0$ by property 2 and is clearly the largest element in $\mathcal{F}_0$. By 3. it now follows that $P \subset g(P) \in \mathcal{F}_0$ and by maximality of P, we have $g(P) = P$, the desired fixed point. So to finish the proof, we must show that Γ is a tower.

First off it is clear that $P_0 \in \Gamma$ so in particular Γ is not empty. For each $C \in \Gamma$ let

$$\Phi_C := \{A \in \mathcal{F}_0 : \text{either } A \subset C \text{ or } g(C) \subset A\}.$$

We will begin by showing that $\Phi_C \subset \mathcal{F}_0$ is a tower and therefore that $\Phi_C = \mathcal{F}_0$.

1. $P_0 \in \Phi_C$ since $P_0 \subset C$ for all $C \in \Gamma \subset \mathcal{F}_0$. 2. If $\Phi \subset \Phi_C \subset \mathcal{F}_0$ is totally ordered by set inclusion, then $A_\Phi := \cup\Phi \in \mathcal{F}_0$. We must show $A_\Phi \in \Phi_C$, that is that $A_\Phi \subset C$ or $C \subset A_\Phi$. Now if $A \subset C$ for all $A \in \Phi$, then $A_\Phi \subset C$ and hence $A_\Phi \in \Phi_C$. On the other hand if there is some $A \in \Phi$ such that $g(C) \subset A$ then clearly $g(C) \subset A_\Phi$ and again $A_\Phi \in \Phi_C$.

3. Given $A \in \Phi_C$ we must show $g(A) \in \Phi_C$, i.e. that

$$g(A) \subset C \text{ or } g(C) \subset g(A). \tag{8.2}$$

There are three cases to consider: either $A \subsetneq C$, $A = C$, or $g(C) \subset A$. In the case $A = C$, $g(C) = g(A) \subset g(A)$ and if $g(C) \subset A$ then $g(C) \subset A \subset g(A)$ and Eq. (8.2) holds in either of these cases. So assume that $A \subsetneq C$. Since $C \in \Gamma$, either $g(A) \subset C$ (in which case we are done) or $C \subset g(A)$. Hence we may assume that

$$A \subsetneq C \subset g(A).$$

Now if C were a proper subset of $g(A)$ it would then follow that $g(A) \setminus A$ would consist of at least two points which contradicts the definition of g. Hence we must have $g(A) = C \subset C$ and again Eq. (8.2) holds, so Φ_C is a tower.

It is now easy to show that Γ is a tower. It is again clear that $P_0 \in \Gamma$ and Property 2. may be checked for Γ in the same way as it was done for Φ_C above. For Property 3., if $C \in \Gamma$ we may use $\Phi_C = \mathcal{F}_0$ to conclude that for all $A \in \mathcal{F}_0$ that either $A \subset C \subset g(C)$ or $g(C) \subset A$, i.e. $g(C) \in \Gamma$. Thus Γ is a tower and we are done. ■

9. Nets

CAUTION: this section is unedited, hence very rough.

In this section (which may be skipped) we develop the notion of nets. Nets are generalization of sequences. Let us begin by showing that for general topological spaces, sequences are not always adequate.

We start by considering $C(\mathbb{R}) \subset \mathbb{C}^{\mathbb{R}}$. If $\{f_n\} \subset C(\mathbb{R})$ and $f_n \to f$ pointwise (which is the notion of convergence in $\mathbb{C}^{\mathbb{R}}$) then f is a Borel measurable function. Hence the sequential limits of elements in $C(\mathbb{R})$ is contained in the Borel measurable functions which is properly contained in $\mathbb{C}^{\mathbb{R}}$. On the other hand we have the

Claim 9.1. $\overline{C(\mathbb{R})} = \mathbb{C}^{\mathbb{R}}$.

Proof. If $f \in \mathbb{C}^{\mathbb{R}}$, a typical neighborhood of f is

$$N = \{g \in \mathbb{C}^{\mathbb{R}} : |g(x_i) - f(x_i)| < \epsilon \text{ for } i = 1, \dots, n\},$$

where $\epsilon > 0$ and $\{x_i\}_{i=1}^n$ is a finite subset of $\mathbb{R}$. Clearly $N \cap C(\mathbb{R}) \neq \emptyset$ so that $f \in \overline{C(\mathbb{R})}$. ■

Definition 9.2. A **directed set** $(A, \leq)$ is a set with a relation such that

1. $\alpha \leq \alpha$
2. $\alpha \leq \beta, \beta \leq \gamma$ implies $\alpha \leq \gamma$ and
3. if $\alpha, \beta \in A$ there exists $\gamma \in A$ such that $\alpha \leq \gamma$ and $\beta \leq \gamma$.

A **net** is function $x : A \to X$ where A is a directed set. We will often denote x by $\{x_\alpha\}_{\alpha \in A}$.

Example 9.3 (Directed sets)**.**

1. $A = 2^X : \alpha \leq \beta$ if $\alpha \subset \beta$. Note $\alpha \leq \beta, \beta \leq \gamma$ implies $\alpha \subset \beta \subset \gamma$ implies $\alpha \subset \gamma$ and if $\alpha \leq \gamma$ $\alpha, \beta \in 2^X$ then $\alpha \leq \alpha \cup \beta$ and $\beta \leq \alpha \cup \beta$.
2. $A = 2^X : \alpha \leq \beta$ if $\beta \subset \alpha$ reverse inclusion. Say $\alpha \leq \beta, \beta \leq \gamma$ implies $\alpha \supseteq \beta \supseteq \gamma$ implies $\alpha \supseteq \gamma$ on $\alpha \leq \gamma$. $\alpha, \beta \in A \quad \alpha \supseteq \alpha \cap \beta, \beta \supseteq \alpha \cap \beta$ implies $\alpha, \beta \leq \alpha \cap \beta$.

Definition 9.4. Let $\{x_\alpha\}_{\alpha \in A} \subset X$ be a net then

- $x_\alpha \to x$, as iff for all $V \in \tau_x$ $x_\alpha \in V$ **eventually** i.e. there exists $\beta = \beta_V \in A$ such that for all $\alpha \geq \beta$ $x_\alpha \in V$.
- x is a cluster point of $\{x_\alpha\}_{\alpha \in A}$ if for all $V \in \tau_x$, $x_\alpha \in V$ **frequently**, i.e. for all $\beta \in A$ there exists $\alpha \geq \beta$ such that $x_\alpha \in V$.

Proposition 9.5. *Let X be a topological space and $E \subset X$. Then*

- *x is an accumulation point of E iff there exists net $\{x_\alpha\} \subset E \setminus \{x\}$such that $x_\alpha \to x$.*
- *$x \in \bar{E}$ iff there exists $\{x_\alpha\} \subset E$ such that $x_\alpha \to x$.*

Proof.

- Say x is an accumulation point of E.
  $A = \tau_x$ by reverse set inclusion for all $\alpha \in \tau_x$ choose $x_\alpha \in (\alpha \setminus \{x\}) \cap E$. Then given $V \in \tau_x$ for all $\alpha \geq V$ i.e. and $\alpha \subset V, x_\alpha \in V$ implies $x_\alpha \to x$.
- **Conversely** If $\{x_\alpha\}_{\alpha \in A} \subset E \backslash \{x\}$ and $x_\alpha \to x$ then for all $V \in \tau_x$ there exists $\beta \in A$ such that $x_\alpha \in V$ for all $\alpha \geq \beta$ in particular $x_\alpha \in (E \setminus \{x\}) \cap V \neq \emptyset$. So $x \in A_c < (E)$.

- **Recall** $\overline{E} = E \cup \operatorname{acc}(E), \operatorname{acc}(E) =$ Accumulation points of E. For clearly $\operatorname{acc}(E) \subset \overline{E}$ so $E \cup \operatorname{acc}(E) \subset E$. Conversely if $x \notin (E \cup \operatorname{acc}(E))$ implies $x \notin E$ and $x \notin \operatorname{acc}(E)$ implies there exists $V \in \tau_x$ such that $V \cap (E \setminus \{x\}) = V \cap E = \emptyset$. Showing $E \cup \operatorname{acc}(E)$ is closed. Thus $E \subset E \cup \operatorname{acc}(E)$. Therefore if $x \in \overline{E} = E \cup \operatorname{acc}(E)$ then if $x \in E$ take $x_\alpha = x$ if $x \in \operatorname{acc}(E)$ take x_α as above. One easily sees if $\{x_\alpha\}_{\alpha \in A} \subset E$ and $x_\alpha \to x$ then $x \in \overline{E}$.

■

Proposition 9.6. *$f : X \to Y$ is continuous at x iff $f(x_\alpha) \to f(x)$ for all $x_\alpha \to x$.*

Proof. f is continuous, $x_\alpha \to x$ then given $V \in \tau_{f(x)}$ there exists $W \in \tau_x$ such that $f(w) \subset V$. So $x_\alpha \in W$ eventually implies $f(x_\alpha) \in V$ eventually implies $f(x_\alpha) \to f(x)$.

Conversely If f is **not** continuous at x. Then there exists $W \in \tau_{f(x)}$ such that for all $V \in \tau_x f(V) \subset W$. For all $V \in \tau_x$ choose $x_V \in W \setminus f(V)$ (Axiom of choice). Then $x_V \to x$ as we have seen above. While $f(x_V) \in W^c$ so $f(x_V) \not\to f(x)$. ■

Definition 9.7 (Subnet)**.** $\langle x_\alpha \rangle_{\alpha \in A}$ is a net $\langle y_\beta \rangle_{\beta \in B}$ is a subnet if there exists $\beta \to \alpha\beta$ such that

(i) $y_\beta = x_{\alpha\beta}$ for all $\beta \in B$

(ii) for all $\alpha_0 \in A$ there exists $\beta_0 \in B$ such that for all $\beta \geq \beta_0$ $\alpha\beta \geq \alpha_0$.

Proposition 9.8. *$\langle x_\alpha \rangle_{\alpha \in A}$ is a net, x is a **cluster** point of $\langle x_\alpha \rangle_{\alpha \in A}$ iff there exists a subnet $\langle y_\beta \rangle_{\beta \in B}$ such that $y_\beta \to x$.*

Proof. Suppose $y_\beta = x_{\alpha\beta} \to x$ and $W \in \tau_x$ and $\alpha_0 \in A$ there exists $\beta_0 \in B$such that $y_\beta \in W$ for all $\beta \geq \beta_0$. i.e. $x_{\alpha\beta} \in W$ for all $\beta \geq \beta_0$. Choose $\beta_1 \in B$ such that $\alpha_\beta \geq \alpha_0$ for all $\beta \geq \beta_1$ then choose $\beta_3 \in B$ such that $\beta_3 \geq \beta_1$ and $\beta_3 \geq \beta_2$ then $\alpha\beta \geq \alpha_0$ and $x_{\alpha\beta} \in W$ for all $\beta \geq \beta_3$ implies $x_\alpha \in W$ frequently.

Conversely: Assume x is a cluster point of $\langle x_\alpha \rangle_{\alpha \in A}$. Consider $\tau_x x A$ with $(U, \alpha) \leq (U', \alpha')$ iff $\alpha \leq \alpha'$ and $U \supseteq U'$. For all $(U, \gamma) \in \tau_x x A$. Choose $\alpha_{(U,r)} \geq \gamma$ in A such that $y_{(y,\gamma)} = x_{\alpha_{(U,\gamma)}} \in U$. Then if $\alpha_0 \in A$ for all $(U', \gamma') \geq (U, \alpha_0)$ i.e. $\gamma' \geq \alpha_0$ and $U' \subset U$ $\alpha_{(U',\gamma')} \geq \gamma' \geq \alpha_0$ implies $\alpha_{(U',\gamma')} \geq \alpha_0$.

Moreover, Given $W \in \tau_x$ then $y_{(U,\gamma)} \in U \subset W$ for all $U \subset W$. Hence fixing $\alpha \in A$ we see if $(U, \gamma) \geq (W, \alpha)$ then $y_{(U,\gamma)} = x_{\alpha_{(U,\gamma)}} \in U \subset W$ showing that $y_{(U,\gamma)} \to x$. ■

Exercise 9.9. [#34, p. 121]

Let $\langle x_\alpha \rangle_{\alpha \in A}$ be a net in a topological space and for each $\alpha \in A$ let $E_\alpha \equiv \{x_\beta : \beta \geq \alpha\}$. Then x is a cluster point of $\langle x_\alpha \rangle$ iff $x \in \bigcap\limits_{\alpha \in A} \overline{E}_\alpha$.

Proof. If x is a cluster point, then given $W \in \tau_x$ we know $E_\alpha \cap W \neq \emptyset$ for all $\alpha \in E$ since $x_\beta \in W$ frequently thus $x \in \overline{E}_\alpha$ implies $x \in \bigcap\limits_{\alpha \in A} \overline{E}_\alpha$.

Conversely: If x is **not** a **cluster point** of $\langle x_\alpha \rangle$ then there exists $W \in \tau_x$ and $\alpha \in A$ such that $x_\beta \notin W$ for all $\beta \geq \alpha$ i.e. $W \cap E_\alpha = \emptyset$ i.e. $x \notin \overline{E}_\alpha$ implies $x \notin \bigcap\limits_{\alpha \in A} \overline{E}_\alpha$ ■

Theorem 9.10. *X is compact iff Every net has a cluster point.*

Proof. Let X be compact and $\langle x_\alpha \rangle_{\alpha \in A} \subset X$. Set $F_\alpha = \overline{\{x_\beta : \beta \geq \alpha\}}$. Now F_α is closed, $F_{\alpha_1} \cap \cdots \cap F_{\alpha_n} \supseteq F_\gamma$ provided $\gamma \geq \alpha_i, \ldots i = 1, \ldots, n$ which exists since A

FIGURE 3. (cirx)

is directed. Therefore $F_{\alpha_1} \cap \cdots \cap F_{\alpha_n} \neq \emptyset$ i.e. $\{F_\alpha\}_{\alpha \in A}$ has the finite intersection property. X is not compact implies there exists $x \in \bigcap_{\alpha \in a} F_\alpha$. By the above problem implies x is a cluster point of $\langle x_\alpha \rangle_{\alpha \in A}$. If X is not compact let $\{U_\beta\}_{\beta \in B}$ be an infinite cover with no finite subcover. Let $A = \{\alpha \subset B : \#\alpha < \infty\}$ ordered by inclusion. for $\alpha \in A$ choose $x_\alpha \in X \setminus \left(\bigcup_{\beta \in \alpha} U_b \right) \neq \emptyset$. Then $\langle x_\alpha \rangle_{\alpha \in A}$ is a net in X.

Claim 9.11. *it has no cluster point. For if $x \in X$ choose β such that $x \in U_\beta$. Then for all $\alpha \geq \{\beta\}$ i.e. $\beta \in \alpha$ $x_\alpha \notin \bigcup_{\gamma \in \alpha} U_\alpha \supseteq U_\beta$ i.e. $x_\alpha \notin U_\beta$ implies x is not a cluster point of $\langle x_\alpha \rangle$.*

■

10. Banach Spaces

Let $\mathbb{F}$ denote either $\mathbb{C}$ or $\mathbb{R}$.

Definition 10.1. A norm on a vector space X is a function $\|\cdot\| : X \to [0, \infty)$ such that

1. (Homogeneity) $\|\lambda x\| = |\lambda| \|x\|$ for all $\lambda \in \mathbb{C}$ and $x \in X$.
2. (Triangle inequality) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.
3. (Positive definite) $\|x\| = 0$ implies $x = 0$.

A pair $(X, \|\cdot\|)$ where X is a vector space and $\|\cdot\|$ is a norm on X is called a normed vector space.

Definition 10.2. If $(X, \|\cdot\|)$ is a normed vector space, then we say $\{x_n\}_{n=1}^\infty \subset X$ is a Cauchy sequence if $\lim_{m,n\to\infty} \|x_m - x_n\| = 0$. The normed vector space is a **Banach space** if it is complete, i.e. if every $\{x_n\}_{n=1}^\infty \subset X$ which is Cauchy is convergent where $\{x_n\}_{n=1}^\infty \subset X$ is convergent iff there exists $x \in X$ such that $\lim_{n\to\infty} \|x_n - x\| = 0$. As usual we will abbreviate this last statement by writing $\lim_{n\to\infty} x_n = x$.

Theorem 10.3. *A normed space $(X, \|\cdot\|)$ is a Banach space iff for every sequence $\{x_n\}_{n=1}^\infty$ such that $\sum_{n=1}^\infty \|x_n\| < \infty$ then $\lim_{N\to\infty} \sum_{n=1}^N x_n = S$ exists in X (that is to say every absolutely convergent series is a convergent series in X). As usual we will denote S by $\sum_{n=1}^\infty x_n$.*

Proof. ($\Rightarrow$)If X is complete and $\sum_{n=1}^{\infty} \|x_n\| < \infty$ then sequence $S_N \equiv \sum_{n=1}^{N} x_n$ for $N \in \mathbb{N}$ is Cauchy because (for $N > M$)

$$\|S_N - S_M\| \leq \sum_{n=M+1}^{N} \|x_n\| \to 0 \text{ as } M, N \to \infty.$$

Therefore$S = \sum_{n=1}^{\infty} x_n := \lim_{N\to\infty} \sum_{n=1}^{N} x_n$ exists in X.

($\Longleftarrow$) Suppose that $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence and let $\{y_k = x_{n_k}\}_{k=1}^{\infty}$ be a subsequence of $\{x_n\}_{n=1}^{\infty}$ such that $\sum_{n=1}^{\infty} \|y_{n+1} - y_n\| < \infty$. By assumption

$$y_{N+1} - y_1 = \sum_{n=1}^{N} (y_{n+1} - y_n) \to S = \sum_{n=1}^{\infty} (y_{n+1} - y_n) \in X \text{ as } N \to \infty.$$

This shows that $\lim_{N\to\infty} y_N$ exists and is equal to $x := y_1 + S$. Since $\{x_n\}_{n=1}^{\infty}$ is Cauchy,

$$\|x - x_n\| \leq \|x - y_k\| + \|y_k - x_n\| \to 0 \text{ as } k, n \to \infty$$

showing that $\lim_{n\to\infty} x_n$ exists and is equal to x. ∎

Example 10.4. We have the following examples of Banach spaces. Suppose that X is a set then

1. The bounded functions $B(X)$ on X is a Banach space with the norm
$$\|f\| = \sup_{x\in X} |f(x)|$$
and
2. if X is a topological space the subspace $BC(X) \subset B(X)$ is closed and hence also a Banach space in the above norm.
3. If $(X, \mathcal{M}, \mu)$ is a measurable space then $L^1(X, \mathcal{M}, d\mu)$ is a Banach space with
$$\|f\|_1 := \int_X |f|\, d\mu$$
provided that we agree to identify functions f of g which agree μ – a.e.

Definition 10.5. Let X and Y be normed spaces and $T : X \to Y$ be a linear map. Then T is said to be bounded provided there exists $C < \infty$ such that $\|T(x)\| \leq C\|x\|_X$ for all $x \in X$. We denote the best constant by $\|T\|$, i.e.

$$\|T\| = \sup_{x\neq 0} \frac{\|T(x)\|}{\|x\|} = \sup_{x\neq 0} \{\|T(x)\| : \|x\| = 1\}.$$

The number $\|T\|$ is called the operator norm of T.

Notation 10.6. Let $L(X, Y)$ denote the bounded linear operators from X to Y.

Remark 10.7. (1) The operator norm is a norm on the vector space $L(X, Y)$. (2) Moreover if $T : X \to Y$ and $S : Y \to Z$ are linear maps between normed vector spaces then $\|ST\| \leq \|S\|\|T\|$, where $ST := S \circ T$. For example the triangle inequality is verified as follows, if $A, B \in L(X, Y)$ then

$$\|A + B\| = \sup_{x\neq 0} \frac{\|Ax + Bx\|}{\|x\|} \leq \sup_{x\neq 0} \frac{\|Ax\| + \|Bx\|}{\|x\|} \leq \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} + \sup_{x\neq 0} \frac{\|Bx\|}{\|x\|} = \|A\| + \|B\|.$$

For the second remark we have for $x \in X$, $\|STx\| \le \|S\|\|Tx\| \le \|S\|\|T\|\|x\|$ from which it follows that $\|ST\| \le \|S\|\|T\|$.

Proposition 10.8. *Suppose that X and Y are normed spaces and $T : X \to Y$ is a linear map. The the following are equivalent:*
(a) *T is continuous.*
(b) *T is continuous at 0.*
(c) *T is bounded.*

Proof. (a) $\Rightarrow$ (b) trivial. (b) $\Rightarrow$ (c) If T continuous at 0 then there exist $\delta > 0$ such that $\|T(x)\| \le 1$ if $\|x\| \le \delta$. Therefore for any $x \in X$, $\|T(\delta x/\|x\|)\| \le 1$ which implies that $\|T(x)\| \le \frac{1}{\delta}\|x\|$ and hence $\|T\| \le \frac{1}{\delta} < \infty$. (c) $\Rightarrow$ (a) Let $x \in X$ and $\epsilon > 0$ be given. Then

$$\|T(y) - T(x)\| = \|T(y - x)\| \le \|T\| \, \|y - x\| < \epsilon$$

provided $\|y - x\| < \epsilon/\|T\| \equiv \delta$. ∎

The following simple theorem is often useful for defining bounded linear transformations.

Theorem 10.9 (B.L.T. Theorem)**.** *Suppose that X is a normed space, Y is a Banach space, and $D \subset X$ is a dense linear subspace of X. Suppose that $T : D \to Y$ is a linear transformation and there exists $C < \infty$ such that $\|Tx\| \le C\|x\|$ for all $x \in D$. Then T has a unique extension to an element of $L(X, Y)$.*

Exercise 10.10. Prove Theorem 10.9

Exercise 10.11. Suppose that X is a Banach space, $I = [a, b]$ with $-\infty < a < b < \infty$. For a function $f : I \to X$, let $\|f\| = \sup\{\|f(t)\|_X : t \in I\}$ and define

$$Y = \{f : I \to X : \|f\| < \infty\}.$$

Show $(Y, \|\cdot\|)$ is a Banach space.

Example 10.12. Let X be a Banach space, $I = [a, b]$ with $-\infty < a < b < \infty$ and $(Y, \|\cdot\|)$ be the Banach space as in Exercise 10.11. Let $D \subset Y$ denote those functions $f \in Y$ of the form $f(t) = \sum_{j=1}^n x_j 1_{A_j}(t)$, where $A_j \in \mathcal{B}$ (the Borel σ – algebra on I) and $x_j \in X$. Alternatively put, $f \in D$ provided that $f : I \to X$ is a bounded simple function. For $f \in D$, let

$$I(f) = \sum_{x \in X} x m(f^{-1}(\{x\}).$$

Just as for real valued functions one shows that $I : D \to X$ is linear. Moreover for $f \in D$,

$$\begin{aligned}\|I(f)\| &= \left\|\sum_{x \in X} x m(f^{-1}(\{x\})\right\| \le \sum_{x \in X} \|x\| \, m(\{f = x\}) \\ &= \int_I \|f(t)\| \, dm(t) \le m(I)\,\|f\|.\end{aligned}$$

This shows that I is bounded and therefore by Theorem 10.9, I has a unique bounded extension to $\bar{D}$ – the closure of D in Y.

Exercise 10.13. Show that $C(I \to X) \subset \bar{D}$. Hence the above example constructs the integral of $f \in C(I \to X)$.

Exercise 10.14. Let $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$ and $T : X \to Y$ be a linear transformation so that T is given by matrix multiplication by an $m \times n$ matrix. Let us identify the linear transformation T with this matrix.

1. (*) Assume the norms on X and Y are the ℓ^1 – norms, i.e. for $x \in \mathbb{R}^n$, $\|x\| = \sum_{j=1}^n |x_j|$. Then the operator norm of T is given by
$$\|T\| = \max_{1 \le j \le n} \sum_{i=1}^m |T_{ij}|.$$
2. Assume the norms on X and Y are the ℓ^∞ – norms, i.e. for $x \in \mathbb{R}^n$, $\|x\| = \max_{1 \le j \le n} |x_j|$. Then the operator norm of T is given by
$$\|T\| = \max_{1 \le i \le m} \sum_{j=1}^n |T_{ij}|.$$
3. (*) Assume the norms on X and Y are the ℓ^2 – norms, i.e. for $x \in \mathbb{R}^n$, $\|x\|^2 = \sum_{j=1}^n x_j^2$. Show $\|T\|^2$ is the largest eigenvalue of the matrix $T^{tr}T : \mathbb{R}^n \to \mathbb{R}^n$.

Example 10.15. If X is finite dimensional normed space then all linear maps are bounded.

Example 10.16. Suppose that $K : [0,1] \times [0,1] \to \mathbb{C}$ is a continuous function. Let $T : L^1([0,1], dm) \to C([0,1])$ be defined by
$$(Tf)(x) = \int_0^1 K(x,y) f(y) dy.$$
It is easily checked that this map is linear and maps to $C([0,1])$ as advertised. (Use the dominated convergence theorem.) If M is a bound for $|K|$, then
$$|(Tf)(x)| \le \int_0^1 |K(x,y)f(y)|\, dy \le M \|f\|_1$$
which shows that $\|Tf\|_\infty \le M \|f\|_1$ and hence that
$$\|T\|_{L^1 \to C} \le \max\{|K(x,y)| : x, y \in [0,1]\} < \infty.$$
We can in fact show that $\|T\| = M$ as follows. Let $(x_0, y_0) \in [0,1]^2$ such that $|K(x_0,y_0)| = M$. Then given $\epsilon > 0$, there exists a neighborhood $U = I \times J$ of (x_0, y_0) such that $|K(x,y) - K(x_0,y_0)| < \epsilon$ for all $(x,y) \in U$. Let $f \in C_c(I, [0,\infty))$ such that $\int_0^1 f(x)dx = 1$. Choose $\alpha \in \mathbb{C}$ such that $|\alpha| = 1$ and $\alpha K(x_0,y_0) = M$, then
$$\begin{aligned}|(T\alpha f)(x_0)| &= \left|\int_0^1 K(x_0,y)\alpha f(y)dy\right| = \left|\int_I K(x_0,y)\alpha f(y)dy\right| \\ &\ge \operatorname{Re}\int_I \alpha K(x_0,y) f(y)dy \ge \int_I (M-\epsilon) f(y)dy = (M-\epsilon)\|\alpha f\|_{L^1}\end{aligned}$$
and hence
$$\|T\alpha f\|_C \ge (M-\epsilon)\|\alpha f\|_{L^1}$$
showing that $\|T\| \ge M - \epsilon$. Since $\epsilon > 0$ is arbitrary, we learn that $\|T\| \ge M$ and hence $\|T\| = M$.

Similarly one easily shows that $T|_{C([0,1])} : C([0,1]) \to C([0,1])$ is bounded and

$$\|T\|_{C\to C} \le \sup\left\{\int_0^1 |K(x,y)|\,dy : x \in [0,1]\right\} < \infty.$$

One may also vies $T : C([0,1]) \to L^1([0,1])$ in which case

$$\|T\|_{L^1\to C} \le \int_0^1 \max_y |K(x,y)|\,dx < \infty.$$

Proposition 10.17. *Suppose that X is a normed vector space and Y is a Banach space. Then $(L(X,Y), \|\cdot\|_{op})$ is a Banach space.*

Proof. We must show $(L(X,Y), \|\cdot\|_{op})$ is complete. Suppose that $T_n \in L(X,Y)$ is a sequence of operators such that $\sum_{n=1}^{\infty} \|T_n\| < \infty$. Then

$$\sum_{n=1}^{\infty} \|T_n x\| \le \sum_{n=1}^{\infty} \|T_n\|\,\|x\| < \infty$$

and therefore by the completeness of Y, $Sx := \sum_{n=1}^{\infty} T_n x = \lim_{N\to\infty} S_N$ exists in Y, where $S_N := \sum_{n=1}^{N} T_n$. The reader should check that $S : X \to Y$ so defined in linear. Since,

$$\|Sx\| = \lim_{N\to\infty} \|S_N x\| \le \lim_{N\to\infty} \sum_{n=1}^{N} \|T_n x\| \le \sum_{n=1}^{\infty} \|T_n\|\,\|x\|,$$

S is bounded and

$$\|S\| \le \sum_{n=1}^{\infty} \|T_n\|. \tag{10.1}$$

Similarly,

$$\|Sx - S_M x\| = \lim_{N\to\infty} \|S_N x - S_M x\| \le \lim_{N\to\infty} \sum_{n=M+1}^{N} \|T_n\|\,\|x\| = \sum_{n=M+1}^{\infty} \|T_n\|\,\|x\|$$

and therefore,

$$\|S - S_M\| \le \sum_{n=M}^{\infty} \|T_n\| \to 0 \text{ as } M \to \infty.$$

■

Definition 10.18. Let $T : X \to Y$ be a linear map between normed spaces X and Y. Then $T : X \to Y$ is an **isometry** if $\|Tx\| = \|x\|$ for all $x \in X$ and $T : X \to Y$ is **invertible** if T is a bijection and T^{-1} is bounded.

Proposition 10.19. *Suppose X is a Banach space and $T \in L(X) \equiv L(X,X)$ satisfies $\sum_{n=0}^{\infty} \|T^n\| < \infty$. Then $I - T$ is invertible and*

$$(I-T)^{-1} = \text{“}\frac{1}{I-T}\text{”} = \sum_{n=0}^{\infty} T^n \text{ and } \left\|(I-T)^{-1}\right\| \le \sum_{n=0}^{\infty} \|T^n\|.$$

In particular if $\|T\| < 1$ *then the above formula holds and*

$$\left\|(I-T)^{-1}\right\| \leq \frac{1}{1-\|T\|}.$$

Proof. Let $S_N := \sum_{n=0}^{N} T^n$. By Proposition 10.17, $S = \sum_{n=0}^{\infty} T^n := \lim_{N\to\infty} S_N$ exists in $L(X)$. Moreover,

$$(I-T)S_N = S_N(I-T) = I - T^{N+1} \to I \in L(X) \text{ as } N \to \infty. \tag{10.2}$$

Since for any $A \in L(X)$, $\|AS - AS_N\| \leq \|A\| \|S - S_N\| \to 0$ as $N \to \infty$ and $\|SA - S_N A\| \leq \|A\| \|S - S_N\| \to 0$ as $N \to \infty$, it follows from Eq. (10.2) that $(I-T)S = S(I-T) = I$. This shows that$(I-T)^{-1} = S$ and the bound on $(I-T)^{-1}$ follows from Eq. (10.1). Furthermore, if $\|T\| < 1$, then $\|T^n\| \leq \|T\|^n$ and

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

proving the last assertion in the statement. ■

Corollary 10.20. *Let* $L^\times(X)$ *denote the invertible elements in* $L(X)$. *Then* $L^\times(X)$ *is an open subset of* $L(X)$. *More specifically, if* $A \in L^\times(X)$ *and* $B \in L(X)$ *satisfies*

$$\|B - A\| < \|A^{-1}\|^{-1} \tag{10.3}$$

then $B \in L^\times(X)$.

Proof. Let A and B be as above, then

$$B = A - (A - B) = A\left[I - A^{-1}(A-B))\right]$$

and

$$\left\|A^{-1}(A-B))\right\| \leq \|A^{-1}\| \|A - B\| < \|A^{-1}\| \|A^{-1}\|^{-1} = 1.$$

Therefore $\left[I - A^{-1}(A-B))\right]$ is invertible and hence so is B with

$$B^{-1} = \left[I - A^{-1}(A-B))\right]^{-1} A^{-1}.$$

Notice that

$$\begin{aligned}\left\|B^{-1}\right\| &\leq \left\|\left[I - A^{-1}(A-B))\right]^{-1}\right\| \|A^{-1}\| \leq \|A^{-1}\| \frac{1}{1 - \|A^{-1}(A-B))\|} \\ &\leq \|A^{-1}\| \frac{1}{1 - \|A^{-1}\| \|A - B\|}.\end{aligned}$$

■

10.0.1. *Application.* Consider the linear differential equation

$$\dot{x}(t) = A(t)x(t) \text{ where } x(0) = x_0 \in \mathbb{R}^n. \tag{10.4}$$

Here $A \in C(\mathbb{R} \to L(\mathbb{R}^n))$ and $x \in C^1(\mathbb{R} \to \mathbb{R}^n)$. As usual this equation may be written in its equivalent integral form, i.e. we are looking for $x \in C(\mathbb{R}, \mathbb{R}^n)$ such that

$$x(t) = x_0 + \int_0^t A(\tau)x(\tau)d\tau. \tag{10.5}$$

Theorem 10.21. *Let $\phi \in C([0,T],\mathbb{R}^n)$, then the integral equation*

$$x(t) = \phi(t) + \int_0^t A(\tau)x(\tau)d\tau \tag{10.6}$$

has a unique solution given by

$$x(t) = \phi(t) + \sum_{n=1}^{\infty} \int_{\Delta_n(t)} A(\tau_n)\dots A(\tau_1)\phi(t_n)d\tau_1\dots d\tau_n$$

where

$$\Delta_n(t) = \{0 \le \tau_1 \le \cdots \le \tau_n \le t\}.$$

Moreover,

$$|x(t)| \le \|\phi\|\, e^{\int_0^T \|A(\tau)\|d\tau}.$$

Proof. Define $\mathcal{A} : C([0,T],\mathbb{R}^n) \to C([0,T],\mathbb{R}^n)$ by

$$(\mathcal{A}x)(t) = \int_0^t A(\tau)x(\tau)d\tau.$$

Then x solves Eq. (10.5) iff $x = \phi + \mathcal{A}x$ or equivalently iff $(I - \mathcal{A})x = \phi$. The theorem will be proved by show $(I-\mathcal{A})^{-1}$ exists and $\sum_{n=1}^{\infty} \|\mathcal{A}^n\| < \infty$. An induction argument shows

$$\begin{aligned}
(\mathcal{A}^n\phi)(t) &= \int_0^t d\tau_n A(\tau_n)(\mathcal{A}^{n-1}\phi)(\tau_n) \\
&= \int_0^t d\tau_n \int_0^{\tau_n} d\tau_{n-1} A(\tau_n)A(\tau_{n-1})(\mathcal{A}^{n-2}\phi)(\tau_{n-1}) \\
&\vdots \\
&= \int\limits_{0\le\tau_1\le\cdots\le\tau_n\le\tau} A(\tau_n)\dots A(\tau_1)\phi(\tau_1)d\tau_1\dots d\tau_n \\
&= \int_{\Delta_n(t)} A(\tau_n)\dots A(\tau_1)\phi(t_n)d\tau_1\dots d\tau_n.
\end{aligned}$$

Hence

$$|(\mathcal{A}^n\phi)(t)|_{\mathbb{R}^n} \le \left\{ \int\limits_{0\le\tau_1\le\cdots\le\tau_n\le\tau} \|A(\tau_n)\|\dots\|A(\tau_1)\|d\tau_1\dots d\tau_n \right\} \|\phi\|_u.$$

Therefore

$$\begin{aligned}
\|\mathcal{A}^n\| &\le \int\limits_{0\le\tau_1\le\cdots\le\tau_n\le T} \|A(\tau_n)\|\dots\|A(\tau_1)\|d\tau_1\dots d\tau_n \\
&= \frac{1}{n!} \int\limits_{[0,T]^n} \|A(\tau_n)\|\dots\|A(\tau_1)\|d\tau_1\dots d\tau_n \\
&= \frac{1}{n!}\left(\int_0^T \|A(\tau)\|d\tau\right)^n. \qquad (10.7)
\end{aligned}$$

Alternatively, one can prove this last equality by induction on n. Namely let

$$F(t) = \int_0^t \|A(\tau)\| d\tau$$

then by induction one shows that

$$I_n(t) := \int_{0 \le \tau_1 \le \cdots \le \tau_n \le T} \|A(\tau_n)\| \ldots \|A(\tau_1)\| d\tau_1 \ldots d\tau_n = \frac{1}{n!} F^n(t).$$

Indeed,

$$I_{n+1}(t) = \int_0^t \frac{1}{n!} F^n(\tau) \dot{F}(\tau) d\tau = \int_0^t \frac{1}{(n+1)!} \frac{d}{d\tau} F^{n+1}(\tau) d\tau = \frac{1}{(n+1)!} F^{n+1}(t)$$

proving Eq. (10.7) again. Using this estimate we then have

$$\sum_{n=0}^{\infty} \|\mathcal{A}^n\| \le e^{\int_0^T \|A(\tau)\| d\tau} < \infty.$$

Therefore $(I - \mathcal{A})^{-1}$ exists and $(I - \mathcal{A})^{-1} = \sum_{n=0}^{\infty} \mathcal{A}^n$ and

$$\left\|(I - \mathcal{A})^{-1}\right\| \le e^{\int_0^T \|A(\tau)\| d\tau}.$$

■

10.1. **More about sums in Banach spaces.**

Definition 10.22. Suppose that X is a Normed space and $\{v_\alpha \in X : \alpha \in A\}$ is a given collection of vectors in X. We say that $s = \sum_{\alpha \in A} v_\alpha \in X$ if for all $\epsilon > 0$ there exists a finite set $\Gamma_\epsilon \subset A$ such that $\left\|s - \sum_{\alpha \in \Lambda} v_\alpha\right\| < \epsilon$ for all $\Lambda \subset\subset A$ such that $\Gamma_\epsilon \subset \Lambda$. (Unlike the case of real valued sums, this does not imply that $\sum_{\alpha \in \Lambda} \|v_\alpha\| < \infty$. See Proposition 12.16, from which one may manufacture counter-examples to this false premise.)

Lemma 10.23. *(1) When X is a Banach space, $\sum_{\alpha \in A} v_\alpha$ exists in X iff for all $\epsilon > 0$ there exists $\Gamma_\epsilon \subset\subset A$ such that $\left\|\sum_{\alpha \in \Lambda} v_\alpha\right\| < \epsilon$ for all $\Lambda \subset\subset A \setminus \Gamma_\epsilon$. Also if $\sum_{\alpha \in A} v_\alpha$ exists in X then $\{\alpha \in A : v_a \neq 0\}$ is at most countable. (2) If $s = \sum_{\alpha \in A} v_\alpha \in X$ exists and $T : X \to Y$ is a bounded linear map between normed spaces, then $\sum_{\alpha \in A} T v_\alpha$ exists in Y and*

$$Ts = T \sum_{\alpha \in A} v_\alpha = \sum_{\alpha \in A} T v_\alpha.$$

Proof. (1) Suppose that $s = \sum_{\alpha \in A} v_\alpha$ exists and $\epsilon > 0$. Let $\Gamma_\epsilon \subset\subset A$ be as in Definition 10.22. Then for $\Lambda \subset\subset A \setminus \Gamma_\epsilon$,

$$\begin{aligned}\left\|\sum_{\alpha \in \Lambda} v_\alpha\right\| &\le \left\|\sum_{\alpha \in \Lambda} v_\alpha + \sum_{\alpha \in \Gamma_\epsilon} v_\alpha - s\right\| + \left\|\sum_{\alpha \in \Gamma_\epsilon} v_\alpha - s\right\| \\ &= \left\|\sum_{\alpha \in \Gamma_\epsilon \cup \Lambda} v_\alpha - s\right\| + \epsilon < 2\epsilon.\end{aligned}$$

Conversely, suppose for all $\epsilon > 0$ there exists $\Gamma_\epsilon \subset\subset A$ such that $\left\|\sum_{\alpha\in\Lambda} v_\alpha\right\| < \epsilon$ for all $\Lambda \subset\subset A \setminus \Gamma_\epsilon$. Let $\gamma_n := \cup_{k=1}^n \Gamma_{1/k} \subset A$ and set $s_n := \sum_{\alpha\in\gamma_n} v_\alpha$. Then for $m > n$,

$$\|s_m - s_n\| = \left\| \sum_{\alpha\in\gamma_m\setminus\gamma_n} v_\alpha \right\| \le 1/n \to 0 \text{ as } m, n \to \infty.$$

Therefore $\{s_n\}_{n=1}^\infty$ is Cauchy and hence convergent in X. Let $s := \lim_{n\to\infty} s_n$, then for $\Lambda \subset\subset A$ such that $\gamma_n \subset A$, we have

$$\left\| s - \sum_{\alpha\in\Lambda} v_\alpha \right\| \le \|s - s_n\| + \left\| \sum_{\alpha\in\Lambda\setminus\gamma_n} v_\alpha \right\| \le \|s - s_n\| + \frac{1}{n}.$$

Since the right member of this equation goes to zero as $n \to \infty$, it follows that $\sum_{\alpha\in A} v_\alpha$ exists and is equal to s.

Let $\gamma := \cup_{n=1}^\infty \gamma_n$ – a countable subset of A. Then for $\alpha \notin \gamma$, $\{\alpha\} \subset A \setminus \gamma_n$ for all n and hence

$$\|v_a\| = \left\| \sum_{\beta\in\{\alpha\}} v_\beta \right\| \le 1/n \to 0 \text{ as } n \to \infty.$$

Therefore $v_\alpha = 0$ for all $\alpha \in A \setminus \gamma$.

(2) Let Γ_ϵ be as in Definition 10.22 and $\Lambda \subset\subset A$ such that $\Gamma_\epsilon \subset \Lambda$. Then

$$\left\| Ts - \sum_{\alpha\in\Lambda} Tv_\alpha \right\| \le \|T\| \left\| s - \sum_{\alpha\in\Lambda} v_\alpha \right\| < \|T\|\,\epsilon$$

which shows that $\sum_{\alpha\in\Lambda} Tv_\alpha$ exists and is equal to Ts. ■

11. Dual Spaces X^*

Notation 11.1. If X is a normed space we denote the Banach space $L(X, \mathbb{F})$ by X^* and refer to X^* as the (**continuous**) dual space of X.

Proposition 11.2. *Let X be a complex vector space over $\mathbb{C}$. If $f \in X^*$ and $u = \mathrm{Re} f \in X^*_{\mathbb{R}}$ then*

$$f(x) = u(x) - iu(ix). \tag{11.1}$$

Conversely if $u \in X^*_{\mathbb{R}}$ and f is defined by Eq. (11.1), then $f \in X^*$ and $\|u\|_{X^*_{\mathbb{R}}} = \|f\|_{X^*}$. More generally if p is a semi-norm on X, then

$$|f| \le p \text{ iff } u \le p.$$

Proof. Let $v(x) = \mathrm{Im}\ f(x)$, then

$$v(ix) = \mathrm{Im}\ f(ix) = \mathrm{Im}(if(x)) = \mathrm{Re} f(x) = u(x).$$

Therefore

$$f(x) = u(x) + iv(x) = u(x) + iu(-ix) = u(x) - iu(ix).$$

Conversely for $u \in X^*_{\mathbb{R}}$ let $f(x) = u(x) - iu(ix)$. Then

$$f((a + ib)x) = u(ax + ibx) - iu(iax - bx) = au(x) + bu(ix) - i(au(ix) - bu(x))$$

while

$$(a + ib)f(x) = au(x) + bu(ix) + i(bu(x) - au(ix)).$$

So f is complex linear.

Because $|u(x)| = |\mathrm{Re} f(x)| \le |f(x)|$, it follows that $\|u\| \le \|f\|$. For $x \in X$ choose $\lambda \in S^1 \subset \mathbb{C}$ such that $|f(x)| = \lambda f(x)$ so

$$|f(x)| = f(\lambda x) = u(\lambda x) \le \|u\| \, \|\lambda x\| = \|u\| \|x\|.$$

Since $x \in X$ is arbitrary, this shows that $\|f\| \le \|u\|$ so $\|f\| = \|u\|$.[13]

For the last assertion, it is clear that $|f| \le p$ implies that $u \le |u| \le |f| \le p$. Conversely if $u \le p$ and $x \in X$, choose $\lambda \in S^1 \subset \mathbb{C}$ such that $|f(x)| = \lambda f(x)$. Then

$$|f(x)| = \lambda f(x) = f(\lambda x) = u(\lambda x) \le p(\lambda x) = p(x)$$

holds for all $x \in X$. ■

Definition 11.3 (Minkowski functional)**.** $p : X \to \mathbb{R}$ is a Minkowski functional if

1. $p(x + y) \le p(x) + p(y)$ for all $x, y \in X$ and
2. $p(cx) = cp(x)$ for all $c \ge 0$ and $x \in X$.

Example 11.4. Suppose that $X = \mathbb{R}$ and

$$p(x) = \inf \{\lambda \ge 0 : x \in \lambda[-1, 2] = [-\lambda, 2\lambda]\}.$$

Notice that if $x \ge 0$, then $p(x) = x/2$ and if $x \le 0$ then $p(x) = -x$, i.e.

$$p(x) = \begin{cases} x/2 & \text{if } \; x \ge 0 \\ |x| & \text{if } \; x \le 0. \end{cases}$$

From this formula it is clear that $p(cx) = cp(x)$ for all $c \ge 0$ but not for $c < 0$. Moreover, p satisfies the triangle inequality, indeed if $p(x) = \lambda$ and $p(y) = \mu$, then $x \in \lambda[-1, 2]$ and $y \in \mu[-1, 2]$ so that

$$x + y \in \lambda[-1, 2] + \mu[-1, 2] \subset (\lambda + \mu)[-1, 2]$$

which shows that $p(x + y) \le \lambda + \mu = p(x) + p(y)$. To check the last set inclusion let $a, b \in [-1, 2]$, then

$$\lambda a + \mu b = (\lambda + \mu)\left(\frac{\lambda}{\lambda + \mu} a + \frac{\mu}{\lambda + \mu} b\right) \in (\lambda + \mu)[-1, 2]$$

since $[-1, 2]$ is a convex set and $\frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} = 1$.

Theorem 11.5 (Hahn-Banach)**.** *Let X be a real vector space, $M \subset X$ be a subspace $f : M \to \mathbb{R}$ be a linear functional such that $f \le p$ on M. Then there exists a linear functional $F : X \to \mathbb{R}$ such that $F|_M = f$ and $F \le p$.*

13

Proof. To understand better why $\|f\| = \|u\|$, notice that

$$\|f\|^2 = \sup_{\|x\|=1} |f(x)|^2 = \sup_{\|x\|=1} (|u(x)|^2 + |u(ix)|^2).$$

Supppose that $M = \sup_{\|x\|=1} |u(x)|$ and this supremum is attained at $x_0 \in X$ with $\|x_0\| = 1$. Replacing x_0 by $-x_0$ if necessary, we may assume that $u(x_0) = M$. Since u has a maximum at x_0,

$$\begin{aligned} 0 &= \frac{d}{dt}\Big|_0 u\left(\frac{x_0 + itx_0}{\|x_0 + itx_0\|}\right) \\ &= \frac{d}{dt}\Big|_0 \left\{\frac{1}{|1 + it|} \; (u(x_0) + tu(ix_0))\right\} = u(ix_0) \end{aligned}$$

since $\frac{d}{dt}|_0 |1 + it| = \frac{d}{dt}|_0 \sqrt{1 + t^2} = 0$.This explains why $\|f\| = \|u\|$. ■

Proof. Step (1) We show for all $x \in X \setminus M$ there exists and extension F to $M \oplus \mathbb{R}x$ with the desired properties. If F exists and $\alpha = F(x)$, then for all $y \in M$ and $\lambda \in \mathbb{R}$ we must have $f(y)+\lambda\alpha = F(y+\lambda x) \leq p(y+\lambda x)$ i.e. $\lambda\alpha \leq p(y+\lambda x) - f(y)$. Equivalently put we must find $\alpha \in \mathbb{R}$ such that

$$\alpha \leq \frac{p(y+\lambda x) - f(y)}{\lambda} \text{ for all } y \in M \text{ and } \lambda > 0$$
$$\alpha \geq \frac{p(z-\mu x) - f(z)}{\mu} \text{ for all } z \in M \text{ and } \mu > 0.$$

So if $\alpha \in \mathbb{R}$ is going to exist, we have to prove, for all $y, z \in M$ and $\lambda, \mu > 0$ that

$$\frac{f(z) - p(z-\mu x)}{\mu} \leq \frac{p(y+\lambda x) - f(y)}{\lambda}$$

or equivalently

$$\begin{aligned} f(\lambda z + \mu y) &\leq \mu p(y+\lambda x) + \lambda p(z - \mu x) \\ &= p(\mu y + \mu\lambda x) + p(\lambda z - \lambda\mu x). \end{aligned} \tag{11.2}$$

But

$$\begin{aligned} f(\lambda z + \mu y) &= f(\lambda z + \mu\lambda x) + f(\lambda z - \lambda\mu x) \\ &\leq p(\lambda z + \mu\lambda x) + p(\lambda z - \lambda\mu x) \end{aligned}$$

which shows that Eq. (11.2) is true and by working backwards, there exist an $\alpha \in \mathbb{R}$ such that $f(y) + \lambda\alpha \leq p(y + \lambda x)$. Therefore $F(y + \lambda x) := f(y) + \lambda\alpha$ is the desired extension.

Step (2) Let us now write $F : X \to \mathbb{R}$ to mean F is defined on a linear subspace $D(F) \subset X$ and $F : D(F) \to \mathbb{R}$ is linear. For $F, G : X \to \mathbb{R}$ we will say $F < G$ if $D(F) \subset D(G)$ and $F = G|_{D(F)}$, that is G is an extension of F. Let

$$\mathcal{F} = \{F : X \to \mathbb{R} : M \subset D(F),\ F \leq p \text{ on } D(F)\}.$$

Then $(\mathcal{F}, <)$ is a partially ordered set. If $\Phi \subset \mathcal{F}$ is a chain (i.e. a linearly ordered subset of $\mathcal{F}$) then Φ has an upper bound $G \in \mathcal{F}$ defined by $D(G) = \bigcup_{F \in \Phi} D(F)$ and $G(x) = F(x)$ for $x \in D(F)$. Then it is easily checked that $D(G)$ is a linear subspace, $G \in \mathcal{F}$, and $F < G$ for all $F \in \Phi$. We may now apply Zorn's Lemma to conclude there exists a maximal element $F \in \mathcal{F}$. Necessarily, $D(F) = X$ for otherwise we could extend F by step (1), violating the maximality of F. Thus F is the desired extension of f. ■

Corollary 11.6. *Suppose that X is a complex vector space, $p : X \to [0, \infty)$ is a semi-norm, $M \subset X$ is a linear subspace, and $f : M \to \mathbb{C}$ is linear functional such that $|f(x)| \leq p(x)$ for all $x \in M$. Then there exists $F \in X'$ (X' is the **algebraic** dual of X) such that $F|_M = f$ and $|F| \leq p$.*

Proof. Let $u = \mathrm{Re} f$ then $u \leq p$ on M and hence by Theorem 11.5, there exists $U \in X'_{\mathbb{R}}$ such that $U|_M = u$ and $U \leq p$ on M. Define $F(x) = U(x) - iU(ix)$ then as in Proposition 11.2, $F = f$ on M and $|F| \leq p$. ■

Theorem 11.7. *Let X be a normed space $M \subset X$ be a closed subspace and $x \in X \setminus M$. Then there exists $f \in X^*$ such that $\|f\| = 1, f(x) = \delta = d(x, M)$ and $f = 0$ on M.*

Proof. Define $f : M \oplus \mathbb{C}x \to \mathbb{C}$ by $f(m+\lambda x) \equiv \lambda\delta$ for all $m \in M$ and $\lambda \in \mathbb{C}$. Notice that

$$\|m + \lambda x\| = |\lambda| \|x + m/\lambda\| \geq |\lambda|\delta$$

and hence

$$|f(m + \lambda x)| = |\lambda|\delta \leq \|m + \lambda x\|$$

which shows $\|f\| \leq 1$. In fact, since $|f(m + x)| = \delta = \inf_{m \in M} \|x + m\|$, $\|f\| = 1$. By Hahn-Banach theorem there exists $F \in X^*$ such that $F|_{M\oplus\mathbb{C}x} = f$ and $|F(x)| \leq \|x\|$ for all $x \in X$, i.e. $\|F\| \leq 1$. Since $1 = \|f\| \leq \|F\| \leq 1$ we see $\|F\| = \|f\|$. ■

Corollary 11.8. *The linear map $x \in X \to \hat{x} \in X^{**}$ where $\hat{x}(f) = f(x)$ for all $x \in X$ is an isometry. (This isometry need not be surjective.)*

Proof. Since $|\hat{x}(f)| = |f(x)| \leq \|f\|_{X^*} \|x\|_X$ for all $f \in X^*$, it follows that $\|\hat{x}\|_{X^{**}} \leq \|x\|_X$. Now applying Theorem 11.7 with $M = \{0\}$, there exists $f \in X^*$ such that $\|f\| = 1$ and $|\hat{x}(f)| = f(x) = \|x\|$, which shows that $\|\hat{x}\|_{X^{**}} \geq \|x\|_X$. This shows that $x \in X \to \hat{x} \in X^{**}$ is an isometry. Since isometries are necessarily injective, we are done. ■

Definition 11.9. A Banach space X is reflexive if the map $x \in X \to \hat{x} \in X^{**}$ is surjective.

11.1. **Weak Topology.**

Definition 11.10. (1) Weak topology on X is the topology generated by X^*. i.e. sets of the form

$$N = \cap_{i=1}^{n} \{x \in X : |f_i(x) - f_i(x_0)| < \epsilon\}$$

where $f_i \in X^*$ and $\epsilon > 0$ form a neighborhood base for the weak topology on X at x_0.

(2) The Weak-$*$ topology on X^* is the topology generated by X, i.e.

$$N \equiv \cap_{i=1}^{n} \{g \in X^* : |f(x_i) - g(x_i)| < \epsilon\}$$

where $x_i \in X$ and $\epsilon > 0$ forms a neighborhood base for the weak-$*$ topology on X^* at $f \in X^*$.

Theorem 11.11 (Alaoglu's Theorem)**.** *If X is a normed space the unit ball in X^* is weak - $*$ compact.*

Proof. For all $x \in X$ let $D_x = \{z \in \mathbb{C} : |z| \leq \|x\|\}$. Then $D_x \subset \mathbb{C}$ is a compact set and so by Tychonoff's Theorem $\Omega \equiv \prod_{x \in X} D_x$ is compact in the product topology. If $f \in \overline{B} := \{f \in X^* : \|f\| \leq 1\}$, $|f(x)| \leq \|f\| \|x\| \leq \|x\|$ which implies that $f(x) \in D_x$ for all $x \in X$, i.e. $\overline{B} \subset \Omega$. The topology on $\bar{B}$ inherited from the weak-$*$ topology on X^* is the same as that relative topology coming from the product topology on Ω. So to finish the proof it suffices to show $\bar{B}$ is a closed subset of the compact space Ω. To prove this let $\pi_x(f) = f(x)$ be the projection maps.

Then

$$\begin{aligned}\overline{B} &= \{f \in \Omega : f \text{ is linear}\} \\ &= \{f \in \Omega : f(x+cy) - f(x) - cf(y) = 0 \text{ for all } x, y \in X \text{ and } c \in \mathbb{C}\} \\ &= \bigcap_{x,y \in X} \bigcap_{c \in \mathbb{C}} \{f \in \Omega : f(x+cy) - f(x) - cf(y) = 0\} \\ &= \bigcap_{x,y \in X} \bigcap_{c \in \mathbb{C}} (\pi_{x+cy} - \pi_x - c\pi_y)^{-1}(\{0\})\end{aligned}$$

which is closed because $(\pi_{x+cy} - \pi_x - c\pi_y) : \Omega \to \mathbb{C}$ is continuous. ■

Definition 11.12. Strong and weak operator topologies on $L(X, Y)$ are the smallest topologies such that

1. **Strong** $T \in L(X, Y) \longrightarrow Tx \in Y$ is continuous for all $x \in X$.
2. **Weak** $T \in L(X, Y) \longrightarrow f(Tx) \in \mathbb{C}$ is continuous for all $x \in X$ and $f \in Y^*$.

12. Hilbert Spaces

Definition 12.1. Let H be a complex vector space. An inner product on H is a function on $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$ such that

1. $\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle$ i.e. $x \to \langle x, z\rangle$ is linear.
2. $\overline{\langle x, y\rangle} = \langle y, x\rangle$.
3. $\|x\|^2 \equiv \langle x, x\rangle \geq 0$ with equality $\|x\|^2 = 0$ iff $x = 0$.

Notice that combining properties (1) and (2) that $x \to \langle z, x\rangle$ is anti-linear for fixed $z \in H$, i.e.

$$\langle z, ax + by\rangle = \bar{a}\langle z, x\rangle + \bar{b}\langle z, y\rangle.$$

We will often find the following formula useful:

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y\rangle = \|x\|^2 + \|y\|^2 + \langle x, y\rangle + \langle y, x\rangle \\ &= \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\langle x, y\rangle \qquad (12.1)\end{aligned}$$

Theorem 12.2 (Schwarz Inequality)**.** *Let $(H, \langle \cdot, \cdot \rangle)$ be an inner product space, then for all $x, y \in H$*

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

and equality holds iff x and y are linearly dependent.

Proof. If $y = 0$, the result holds trivially. So assume that $y \neq 0$. First off notice that if $x = \alpha y$ for some $\alpha \in \mathbb{C}$, then $\langle x, y\rangle = \alpha \|y\|^2$ and hence

$$|\langle x, y\rangle| = |\alpha| \|y\|^2 = \|x\|\|y\|.$$

Moreover, in this case $\alpha := \frac{\langle x, y\rangle}{\|y\|^2}$.

Now suppose that $x \in H$ is arbitrary, let

$$z \equiv x - \frac{\langle x, y\rangle}{\|y\|^2} y.$$

(So z is the "orthogonal projection" of x onto y.) Then

$$\begin{aligned}0 \le \|z\|^2 &= \left\| x - \frac{\langle x, y\rangle}{\|y\|^2} y \right\|^2 = \|x\|^2 + \frac{|\langle x, y\rangle|^2}{\|y\|^4}\|y\|^2 - 2\mathrm{Re}\langle x, \frac{\langle x, y\rangle}{\|y\|^2} y\rangle \\ &= \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2}\end{aligned}$$

from which it follows that

$$0 \le \|y\|^2\|x\|^2 - |\langle x, y\rangle|^2$$

with equality iff $z = 0$ or equivalently iff

$$x = \frac{\langle x, y\rangle y}{\|y\|^2}.$$

■

Corollary 12.3. *Let $(H, \langle\cdot,\cdot\rangle)$ be an inner product space and $\|x\| := \sqrt{\langle x, x\rangle}$. Then $\|\cdot\|$ is a norm on H. Moreover $\langle\cdot,\cdot\rangle$ is continuous on $H \times H$, where H is viewed as the normed space $(H, \|\cdot\|)$.*

Proof. The only non-trivial thing to verify that $\|\cdot\|$ is a norm is the triangle inequality:

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

where we have made use of Schwarz's inequality. Taking the square root of this inequality shows $\|x + y\| \le \|x\| + \|y\|$. For the continuity assertion:

$$\begin{aligned}|\langle x, y\rangle - \langle x', y'\rangle| &= |\langle x - x', y\rangle + \langle x', y - y'\rangle| \\ &\le \|y\|\|x - x'\| + \|x'\|\|y - y'\| \\ &\le \|y\|\|x - x'\| + (\|x\| + \|x - x'\|)\,\|y - y'\| \\ &= \|y\|\|x - x'\| + \|x\|\|y - y'\| + \|x - x'\|\|y - y'\|\end{aligned}$$

from which it follows that $\langle\cdot,\cdot\rangle$ is continuous. ■

Definition 12.4. Let $(H, \langle\cdot,\cdot\rangle)$ be an inner product space, we say $x, y \in H$ are **orthogonal** and write $x \perp y$ iff $\langle x, y\rangle = 0$. More generally if $A \subset H$ is a set we say x is **orthogonal to** A and write $x \perp A$ iff $\langle x, y\rangle = 0$ for all $y \in A$. We also introduce the set

$$A^\perp = \{x \in H : x \perp A\}.$$

We also say that a set $S \subset H$ is **orthogonal** if $x \perp y$ for all $x, y \in S$ such that $x \neq y$ and if S also satisfies $\|x\| = 1$ for all $x \in S$, then S is said to be **orthonormal.**

Proposition 12.5. *Let $(H, \langle\cdot,\cdot\rangle)$ be an inner product space then*

1. ***(Parallelogram Law)***

(12.2)
$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

for all $x, y \in H$.

2. ***(Pythagorean Theorem)*** *If $S \subset H$ is a finite orthonormal set, then*

(12.3)
$$\|\sum_{x \in S} x\|^2 = \sum_{x \in S} \|x\|^2.$$

3. *If $A \subset H$ is a set, then $A^{\perp}$ is a* ***closed*** *linear subspace of H.*

Remark 12.6. See Proposition 12.28 in the appendix below for the "converse" of the parallelogram law.

Proof. I will assume that H is a complex Hilbert space, the real case being easier.

(1) For $x, y \in H$,

$$\begin{aligned}\|x+y\|^2 + \|x-y\|^2 &= \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\langle x, y\rangle \\ &\quad + \|x\|^2 + \|y\|^2 - 2\mathrm{Re}\langle x, y\rangle \\ &= 2\|x\|^2 + 2\|y\|^2.\end{aligned}$$

(2) This is a simple computation

$$\begin{aligned}\|\sum_{x\in S} x\|^2 = \langle \sum_{x\in S} x, \sum_{y\in S} y\rangle &= \sum_{x,y\in S} \langle x, y\rangle \\ = \sum_{x\in S}\langle x, x\rangle &= \sum_{x\in S} \|x\|^2.\end{aligned}$$

(3) This is a consequence of the continuity of $\langle\cdot,\cdot\rangle$ and the fact that

$$A^{\perp} = \cap_{x\in A} \ker(\langle \cdot, x\rangle)$$

where $\ker(\langle\cdot, x\rangle) = \{y \in H : \langle y, x\rangle = 0\}$ – a closed subspace of H. ■

Definition 12.7. A Hilbert space is an inner product space $(H, \langle\cdot,\cdot\rangle)$ such that the induced Hilbertian norm is complete.

Definition 12.8. A subset C of a vector space X is said to be convex if for all $x, y \in C$ the line segment $[x, y] := \{tx + (1-t)y : 0 \le t \le 1\}$ joining x to y is contained in C as well. (Notice that any vector subspace of X is convex.)

Theorem 12.9. *Suppose that H is a Hilbert space and $M \subset H$ be a closed convex subset of H. Then for any $x \in H$ there exists a unique $y \in M$ such that*

$$\|x - y\| = d(x, M) = \inf_{z\in M} \|x - z\|.$$

Moreover, if M is a vector subspace of H, then the point y may also be characterized as the unique point in M such that $(x - y) \perp M$.

Proof. Let $y_n \in M$ such that $\|x - y_n\| = \delta_n \to \delta \equiv d(x, M)$. Then by the parallelogram law,

$$\begin{aligned}2\|x - y_n\|^2 + 2\|x - y_m\|^2 &= \|2x - (y_n + y_m)\|^2 + \|y_n - y_m\|^2 \\ &= 4\|x - \frac{(y_n + y_m)}{2}\|^2 + \|y_n - y_m\|^2 \\ &\ge 4\delta^2 + \|y_n - y_m\|^2 \qquad (12.4)\end{aligned}$$

where we have used the fact that M is convex so that $(y_n + y_m)/2 \in M$. Letting $m, n \to \infty$ in the previous inequality implies

$$2\delta^2 + 2\delta^2 \ge 4\delta^2 + \lim\sup_{m,n\to\infty} \|y_n - y_m\|^2,$$

i.e. $\limsup_{m,n\to\infty} \|y_n - y_m\|^2 = 0$. Therefore $\{y_n\}_{n=1}^{\infty}$ is Cauchy and hence convergent. Because M is closed, $y := \lim_{n\to\infty} y_n \in M$. Also

$$\|x - y\| = \lim_{n\to\infty} \|x - y_n\| = \lim_{n\to\infty} \delta_n = \delta = d(x, M)$$

and therefore y is a desired closest element in M to x.

To show that y is unique, suppose that $z \in M$ were another point such that $\|x - z\| = \delta = d(x, M)$. Then using the parallelogram law as in Eq. (12.4) we find

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

from which we learn $\|y - z\|^2$, i.e. $y = z$.

Now suppose that M is a subspace of H and $y \in M$ is the closest point in M to x. Then for $w \in M$, the function

$$g(t) \equiv \|x - (y + tw)\|^2 = \|x - y\|^2 - 2t\mathrm{Re}\langle x - y, w\rangle + t^2\|w\|^2$$

has a minimum at $t = 0$. Therefore $0 = g'(0) = -2\mathrm{Re}\langle x - y, w\rangle$. Since $w \in M$ is arbitrary, this implies that $(x - y) \perp M$. Finally suppose $y \in M$ is any point such that $(x - y) \perp M$. Then for $z \in M$, by Pythagorean's theorem,

$$\begin{aligned} \|x - z\|^2 &= \|x - y + y - z\|^2 \\ &= \|x - y\|^2 + \|y - z\|^2 \\ &\geq \|x - y\|^2 \end{aligned}$$

which shows $d(x, M)^2 \geq \|x - y\|^2$. That is to say y is the point in M closest to x.
■

Definition 12.10. Let H be a Hilbert space and $M \subset H$ be a closed subspace. The orthogonal projection of H onto M is the function $P_M : H \to H$ such that for $x \in H$, $P_M(x)$ is the unique element in M such that $(x - P_M(x)) \perp M$.

Proposition 12.11. *Let H be a Hilbert space and $M \subset H$ be a closed subspace. The orthogonal projection P_M satisfies:*

1. *P_M is linear (and hence we will write $P_M x$ rather than $P_M(x)$.*
2. *$P_M^2 = P_M$ (P_M is a projection.)*
3. *$P_M^* = P_M$, i.e. $\langle P_M x, y\rangle = \langle x, P_M y\rangle = \langle P_M x, P_M y\rangle$ for all $x, y \in H$.*
4. *$\mathrm{ran}(P_M) = M$ and $\ker(P_M) = M^\perp$*

Proof.

1. Let $x_1, x_2 \in H$ and $\alpha \in \mathbb{F}$, then $P_M x_1 + \alpha P_M x_2 \in M$ and
$$P_M x_1 + \alpha P_M x_2 - (x_1 + \alpha x_2) = [P_M x_1 - x_1 + \alpha(P_M x_2 - x_2)] \in M^\perp$$
showing
$$P_M x_1 + \alpha P_M x_2 = P_M(x_1 + \alpha x_2).$$
Then P_M is linear.
2. Obviously $\mathrm{ran}(P_M) = M$ and $Px = x$ for all $x \in M$. Therefore $P_M^2 = P_M$.
3. Let $x, y \in H$, then since $(x - P_M x)$ and $(y - P_M y)$ are in $M^\perp$,
$$\begin{aligned} \langle P_M x, y\rangle &= \langle P_M x, P_M y + y - P_M y\rangle \\ &= \langle P_M x, P_M y\rangle \\ &= \langle P_M x + (x - P_M), P_M y\rangle \\ &= \langle x, P_M y\rangle. \end{aligned}$$

4. It is clear that $\operatorname{ran}(P_M) \subset M$. Moreover, if $x \in M$, then $P_M x = x$ implies that $\operatorname{ran}(P_M) = M$. Now $x \in \ker(P_M)$ iff $P_M x = 0$ iff $x = x - 0 \in M^\perp$.

■

Corollary 12.12. *Suppose that $M \subset H$ is a proper closed subspace of a Hilbert space H, then $H = M \oplus M^\perp$.*

Proof. Given $x \in H$, let $y = P_M x$ so that $x - y \in M^\perp$. Then $x = y + (x - y) \in M + M^\perp$. If $x \in M \cap M^\perp$, then $x \perp x$, i.e. $\|x\|^2 = \langle x, x\rangle = 0$. So $M \cap M^\perp = \{0\}$.[14] ■

Proposition 12.13. *The map*

$$z \in H \xrightarrow{j} \langle \cdot, z\rangle \in H^* \tag{12.5}$$

is a conjugate linear isometric isomorphism.

Proof. The map j is conjugate linear by the axioms of the inner products. Moreover, for $x, z \in H$,

$$|\langle x, z\rangle| \le \|x\| \, \|z\| \text{ for all } x \in H$$

with equality when $x = z$. This implies that $\|jz\|_{H^*} = \|\langle \cdot, z\rangle\|_{H^*} = \|z\|$. Therefore j is isometric and this shows that j is injective. To finish the proof we must show that j is surjective. So let $f \in H^*$ which we assume with out loss of generality is non-zero. Then $M = \ker(f)$ – a closed proper subspace of H. Since, by Corollary 12.12, $H = M \oplus M^\perp$, $f : H/M \cong M^\perp \to \mathbb{F}$ is a linear isomorphism. This shows that $\dim(M^\perp) = 1$ and hence $H = M \oplus \mathbb{F}x_0$ where $x_0 \in M^\perp \setminus \{0\}$.[15] Choose $z = \lambda x_0 \in M^\perp$ such that $f(x_0) = \langle x_0, z\rangle$. (So $\lambda = \bar{f}(x_0)/\|x_0\|^2$.) Then for $x = m + \lambda x_0$ with $m \in M$ and $\lambda \in \mathbb{F}$ we have

$$f(x) = \lambda f(x_0) = \lambda \langle x_0, z\rangle = \langle \lambda x_0, z\rangle = \langle m + \lambda x_0, z\rangle = \langle x, z\rangle$$

which shows that $f = jz$. ■

Definition 12.14. $\{u_\alpha\}_{\alpha \in A} \subset H$ is an orthonormal set if $u_\alpha \perp u_\beta$ for all $\alpha \ne \beta$ and $\|u_\alpha\| = 1$.

Proposition 12.15 (Bessel's Inequality)**.** *Let $\{u_\alpha\}_{\alpha \in A}$ be an orthonormal set, then*

$$\sum_{\alpha \in A} |\langle x, u_\alpha\rangle|^2 \le \|x\|^2 \text{ for all } x \in H. \tag{12.6}$$

In particular the set $\{\alpha \in A : \langle x, u_\alpha\rangle \ne 0\}$ is at most countable for all $x \in H$.

---

[14] Old Proof follows.

Let P be the projection $P = P_M$ and Q be the be the complementary projection, $Q = 1 - P$. It is straightforward to check that $Q^2 = Q$ and $PQ = QP = 0$. Since for all $x \in H$, $x = Px + Qx$ and if $Px = Qy$ for some $x, y \in H$, then $0 = QPx = Q^2 y = Qy$. Therefore $H = \operatorname{ran}(P) \oplus \operatorname{ran}(Q)$. Since $\operatorname{ran}(P) = M$, to finish the proof it suffices to show that $\operatorname{ran}(Q) = M^\perp$. By definition of P, $Qx = x - Px \in M^\perp$ for all $x \in H$ so that $\operatorname{ran}(Q) \subset M^\perp$. Conversely, if $x \in M^\perp = \ker(P)$, $x = (1 - P)x = Qx \in \operatorname{ran}(Q)$.

[15] Alternatively, choose $x_0 \in M^\perp \setminus \{0\}$ such that $f(x_0) = 1$. For $x \in M^\perp$ we have $f(x - \lambda x_0) = 0$ provided that $\lambda := f(x)$. Therefore $x - \lambda x_0 \in M \cap M^\perp = \{0\}$, i.e. $x = \lambda x_0$. This again shows that $M^\perp$ is spanned by x_0.

Proof. Let $\Gamma \subset A$ be any finite set. Then

$$\begin{aligned} 0 \le \|x - \sum_{\alpha\in\Gamma} \langle x, u_\alpha\rangle u_\alpha\|^2 &= \|x\|^2 - 2\mathrm{Re}\sum_{\alpha\in\Gamma}\langle x, u_\alpha\rangle\ \langle u_\alpha, x\rangle + \sum_{\alpha\in\Gamma}|\langle x, u_\alpha\rangle|^2 \\ &= \|x\|^2 - \sum_{\alpha\in\Gamma}|\langle x, u_\alpha\rangle|^2 \end{aligned}$$

showing that

$$\sum_{\alpha\in\Gamma}|\langle x, u_\alpha\rangle|^2 \le \|x\|^2.$$

Taking the supremum of this equation of $\Gamma \subset\subset A$ then proves Eq. (12.6). ■

Proposition 12.16. *Suppose $A \subset H$ is an orthogonal set. Then $s = \sum_{v\in A} v$ exists in H iff $\sum_{v\in A}\|v\|^2 < \infty$. (In particular A must be at most a countable set.) Moreover, if $\sum_{v\in A}\|v\|^2 < \infty$, then*

1. $\|s\|^2 = \sum_{v\in A}\|v\|^2$ *and*
2. $\langle s, x\rangle = \sum_{v\in A}\langle v, x\rangle$ *for all* $x \in H$.

Similarly if $\{v_n\}_{n=1}^\infty$ be an orthogonal set, then $s = \sum\limits_{n=1}^{\infty} v_n$ exists in H iff $\sum\limits_{n=1}^{\infty}\|v_n\|^2 < \infty$. In particular if $\sum\limits_{n=1}^{\infty} v_n$ exists, then it is independent of rearrangements of $\{v_n\}_{n=1}^\infty$.

Proof. Suppose $s = \sum_{v\in A} v$ exists. Then there exists $\Gamma \subset\subset A$ such that

$$\sum_{v\in\Lambda}\|v\|^2 = \left\|\sum_{v\in\Lambda} v\right\|^2 \le 1$$

for all $\Lambda \subset\subset A\backslash\Gamma$,wherein the first inequality we have used Pythagorean's theorem. Taking the supremum over such Λ shows that $\sum_{v\in A\backslash\Gamma}\|v\|^2 \le 1$ and therefore

$$\sum_{v\in A}\|v\|^2 \le 1 + \sum_{v\in\Gamma}\|v\|^2 < \infty.$$

Conversely, suppose that $\sum_{v\in A}\|v\|^2 < \infty$. Then for all $\epsilon > 0$ there exists $\Gamma_\epsilon \subset\subset A$ such that if $\Lambda \subset\subset A \setminus \Gamma_\epsilon$,

$$\left\|\sum_{v\in\Lambda} v\right\|^2 = \sum_{v\in\Lambda}\|v\|^2 < \epsilon^2.$$

Hence by Lemma 10.23, $\sum_{v\in A} v$ exists.

For item 1, let $s_\epsilon := \sum_{v\in\Gamma_\epsilon} v$, then

$$|\|s\| - \|s_\epsilon\|| \le \|s - s_\epsilon\| < \epsilon$$

and

$$0 \le \sum_{v\in A}\|v\|^2 - \|s_\epsilon\|^2 < \epsilon^2.$$

Letting $\epsilon \to 0$ we deduce from the previous two equations that $\|s\|^2 = \sum_{v\in A}\|v\|^2$. Item 2. is a special case of Lemma 10.23.

For the final assertion, let $s_N \equiv \sum_{n=1}^{N} v_n$ and suppose that $\lim_{N\to\infty} s_N = s$ exists in H. Then in particular $\{s_N\}_{N=1}^{\infty}$ is Cauchy so for $N > M$.

$$\sum_{n=M+1}^{N} \|v_n\|^2 = \|s_N - s_M\|^2 \to 0 \text{ as } M, N \to \infty$$

which shows that $\sum_{n=1}^{\infty} \|v_n\|^2$ is convergent, i.e. $\sum_{n=1}^{\infty} \|v_n\|^2 < \infty$. ∎

Corollary 12.17. *Suppose H is a Hilbert space, $\beta \subset H$ is an orthonormal set and $M = \overline{\operatorname{span} \beta}$. Then*

$$P_M x = \sum_{u\in\beta} \langle x, u\rangle u, \tag{12.7}$$

$$\sum_{u\in\beta} |\langle x, u\rangle|^2 = \|P_M x\|^2 \quad \text{and} \tag{12.8}$$

$$\sum_{u\in\beta} \langle x, u\rangle\langle u, y\rangle = \langle P_M x, y\rangle \tag{12.9}$$

for all $x, y \in H$.

Proof. By Bessel's inequality, $\sum_{u\in\beta} |\langle x, u\rangle|^2 \le \|x\|^2$ for all $x \in H$ and hence by Proposition 12.15, $Px := \sum_{u\in\beta} \langle x, u\rangle u$ exists in H for all $x \in H$ and

$$\langle Px, y\rangle = \sum_{u\in\beta} \langle \langle x, u\rangle u, y\rangle = \sum_{u\in\beta} \langle x, u\rangle\langle u, y\rangle \tag{12.10}$$

for all $y \in H$. Taking $y \in \beta$ in this expression shows that $\langle Px, y\rangle = \langle x, y\rangle$, i.e. that $\langle x - Px, y\rangle = 0$. Since $y \in \beta$ is arbitrary, we learn that $(x - Px) \perp \operatorname{span} \beta$ and by continuity we also have $(x - Px) \perp M = \overline{\operatorname{span} \beta}$. Since Px is also in M, it follows from the definition of P_M that $Px = P_M x$ proving Eq. (12.7). Equations (12.8) and (12.9) now follow from (12.10), Proposition 12.16 and the fact that $\langle P_M x, y\rangle = \langle P_M x, P_M y\rangle$ for all $x, y \in H$. For example,

$$\begin{aligned}\langle P_M x, y\rangle = \langle P_M x, P_M y\rangle &= \langle \sum_{u\in\beta} \langle x, u\rangle u, P_M y\rangle = \sum_{u\in\beta} \langle x, u\rangle\langle u, P_M y\rangle \\ &= \sum_{u\in\beta} \langle x, u\rangle\langle P_M u, y\rangle = \sum_{u\in\beta} \langle x, u\rangle\langle u, y\rangle.\end{aligned}$$

∎

12.1. **Hilbert Space Basis.**

Definition 12.18 (Basis)**.** Let H be a Hilbert space. A **basis** β of H is a maximal orthonormal subset $\beta \subset H$.

Proposition 12.19. *Every Hilbert space has an orthonormal basis.*

Proof. Let $\mathcal{F}$ be the collection of all orthonormal subsets of H ordered by inclusion. If $\Phi \subset \mathcal{F}$ is linearly ordered then $\cup\Phi$ is an upper bound. By Zorn's Lemma there exists a maximal element $\beta \in \mathcal{F}$. ∎

An orthonormal set $\beta \subset H$ is said to be **complete** if $\beta^\perp = \{0\}$. That is to say if $\langle x, u\rangle = 0$ for all $u \in \beta$ then $x = 0$.

Remark 12.20. An orthonormal set $\beta \subset H$ is a basis iff β is complete.

Proof. Suppose that $\beta \subset H$. If β is not complete, then there exists a unit vector $x \in \beta^\perp \setminus \{0\}$. The set $\beta \cup \{x\}$ is an orthonormal set properly containing β, so β is not maximal. Conversely, if β is not maximal, there exists an orthonormal set $\beta_1 \subset H$ such that $\beta \subsetneqq \beta_1$. Then if $x \in \beta_1 \setminus \beta$, we have $\langle x, u\rangle = 0$ for all $u \in \beta$ showing β is not complete. ■

Theorem 12.21. *Let $\beta \subset H$ be an orthonormal set. Then the following are equivalent:*

1. *β is complete or equivalently a basis.*
2. *$x = \sum_{u\in\beta} \langle x, u\rangle u$ for all $x \in H$.*
3. *$\langle x, y\rangle = \sum_{u\in\beta} \langle x, u\rangle \ \langle u, y\rangle$ for all $x, y \in H$.*
4. *$\|x\|^2 = \sum_{u\in\beta} |\langle x, u\rangle|^2$ for all $x \in H$.*

Proof. Let $M = \overline{\text{span}\ \beta}$ and $P = P_M$.

(1) $\Rightarrow$ (2) By Corollary 12.17, $\sum_{u\in\beta} \langle x, u\rangle u = P_M x$. Therefore

$$x - \sum_{u\in\beta} \langle x, u\rangle u = x - P_M x \in M^\perp = \beta^\perp = \{0\}.$$

(2) $\Rightarrow$ (3) is a consequence of Proposition 12.16.

(3) $\Rightarrow$ (4) is obvious, just take $y = x$.

(4) $\Rightarrow$ (1) If $x \in \beta^\perp$, then by 4), $\|x\| = 0$, i.e. $x = 0$. This shows that β is maximal. ■

Proposition 12.22. *A Hilbert space H is separable iff H has a countable orthonormal basis $\beta \subset H$. Moreover, if H is separable, all orthonormal bases of H are countable.*

Proof. Let $\mathbb{D} \subset H$ be a countable dense set $\mathbb{D} = \{u_n\}_{n=1}^\infty$. By Gram-Schmidt process there exists $\beta = \{v_n\}_{n=1}^\infty$ an orthonormal set such that $\text{span}\{v_n : n = 1, 2 \ldots, N\} \supseteq \text{span}\{u_n : n = 1, 2 \ldots, N\}$. So if $\langle x, v_n\rangle = 0$ for all n then $\langle x, u_n\rangle = 0$ for all n. Since $\mathbb{D} \subset H$ is dense we may choose $\{w_k\} \subset \mathbb{D}$ such that $x = \lim_{k\to\infty} w_k$ and therefore $\langle x, x\rangle = \lim_{k\to\infty} \langle x, w_k\rangle = 0$. That is to say $x = 0$ and β is complete.

Conversely if $\beta \subset H$ is a countable orthonormal basis, then the countable set

$$\mathbb{D} = \left\{ \sum_{u\in\beta} a_u u : a_u \in \mathbb{Q} + i\mathbb{Q} : \#\{u : a_u \neq 0\} < \infty \right\}$$

is dense in H.

Finally let $\beta = \{u_n\}_{n=1}^\infty$ be an basis and $\beta_1 \subset H$ be another orthonormal basis. Then the sets

$$A_n = \{v \in \beta_1 : \langle v, u_n\rangle \neq 0\}$$

are countable for each $n \in \mathbb{N}$ and hence $B := \bigcup_{n=1}^\infty A_n$ is a countable subset of A. The proof will be finished by showing B is complete and hence maximal, so that $A = B$. To see that B is complete, suppose that $x \in B^\perp$ and there exists $v \in A \setminus B$, so $\langle u_n, v\rangle = 0$ for all $n \in \mathbb{N}$. Then

$$\langle x, v\rangle = \sum_{n=1}^\infty \langle x, u_n\rangle \langle u_n, v\rangle = 0.$$

Since by assumption $\langle x, v\rangle = 0$ for all $v \in B$, it follows that $\langle x, v\rangle = 0$ for all $v \in A$ and because A is complete, $x = 0$. ■

Remark 12.23. Suppose that $\{u_n\}_{n=1}^{\infty}$ is a **total** subset of H, i.e. $\overline{\operatorname{span}\{u_n\}} = H$. Let $\{v_n\}_{n=1}^{\infty}$ be the vectors found by performing Gram-Schmidt on the set $\{u_n\}_{n=1}^{\infty}$. Then $\{v_n\}_{n=1}^{\infty}$ is an orthonormal basis for H.

Example 12.24.

1. $H = L^2([-\pi,\pi], dm)$, then by the Stone-Weierstrass theorem, $\{e^{in\theta}\}_{n=-\infty}^{\infty}$ is total and therefore $e_n(\theta) = \frac{1}{\sqrt{2\pi}} e^{in\theta}$ for $n \in \mathbb{Z}$ is an orthonormal basis. Indeed, we may identify H with $L^2(S^1, d\theta)$ by the map, $f \in H \to (\theta \to f(e^{i\theta})) \in L^2(S^1, d\theta)$. Under this identification, $e^{in\theta}$ corresponds to z^n and we have seen by the Stone-Weierstrass theorem that the algebra generated by z and z^{-1} is dense in $C(S^1)$. Since $C(S^1)$ is dense in $L^2(S^1, d\theta)$, it follows that the algebra generated by z and z^{-1} is dense in $L^2(S^1, d\theta)$ as well.
2. Let $H = L^2([-1,1], dm)$ and $A := \{1, x, x^2, x^3 \dots\}$. Then A is total in H by the Stone-Weierstrass theorem and a similar argument as in the first example. The result of doing Gram-Schmidt on this set is the Legendre Polynomials.
3. Let $H = L^2(\mathbb{R}, e^{-\frac{1}{2}x^2} dx)$. Fact $A := \{1, x, x^2, x^3 \dots\}$ is total in H and the result of doing Gram-Schmidt on A now gives the **Hermite Polynomials.**

Remark 12.25 (An Interesting Phenomena). Let $H = L^2([-1,1], dm)$ and $B := \{1, x^3, x^6, x^9, \dots\}$. Then again A is total in H by the same argument as in item 2. Example 12.24. This is true even though B is a proper subset of A. Notice that A is an algebraic basis for the polynomials on $[-1,1]$ while B is not! The following computations may help relieve some of the reader's anxiety. Let $f \in L^2([-1,1], dm)$, then, making the change of variables $x = y^{1/3}$, shows that

$$\int_{-1}^{1} |f(x)|^2 \, dx = \int_{-1}^{1} \left| f(y^{1/3}) \right|^2 \frac{1}{3} y^{-2/3} dy = \int_{-1}^{1} \left| f(y^{1/3}) \right|^2 d\mu(y) \tag{12.11}$$

where $d\mu(y) = \frac{1}{3} y^{-2/3} dy$. Since $\mu([-1,1]) = m([-1,1]) = 2$, μ is a finite measure on $[-1,1]$ and hence regular and hence $C([-1,1])$ is dense in $L^2([-1,1], d\mu)$. Thus the usual Stone Weierstrass argument shows as above that $A := \{1, x, x^2, x^3 \dots\}$ is a total in $L^2([-1,1], d\mu)$. In particular for any $\epsilon > 0$ there exists a polynomial $p(y)$ such that

$$\int_{-1}^{1} \left| f(y^{1/3}) - p(y) \right|^2 d\mu(y) < \epsilon^2.$$

However, by Eq. (12.11) we have

$$\epsilon^2 > \int_{-1}^{1} \left| f(y^{1/3}) - p(y) \right|^2 d\mu(y) = \int_{-1}^{1} \left| f(x) - p(x^3) \right|^2 dx.$$

Alternatively, if $f \in C([-1,1])$, then $g(y) = f(y^{1/3})$ is back in $C([-1,1])$. Therefore for any $\epsilon > 0$, there exists a polynomial $p(y)$ such that

$$\begin{aligned} \epsilon > \|g - p\|_u &= \sup\{|g(y) - p(y)| : y \in [-1,1]\} \\ &= \sup\left\{ \left| g(x^3) - p(x^3) \right| : x \in [-1,1] \right\} = \sup\left\{ \left| f(x) - p(x^3) \right| : x \in [-1,1] \right\}. \end{aligned}$$

This gives another proof the polynomials in x^3 are dense in $C([-1,1])$ and hence in $L^2([-1,1])$.

12.1.1. *Non-complete inner product spaces.* Part of Theorem 12.21 goes through when H is a not necessarily complete inner product space. We have the following proposition.

Proposition 12.26. *Let $(H, \langle\cdot,\cdot\rangle)$ be a not necessarily complete inner product space and $\beta \subset H$ be an orthonormal set. Then the following two conditions are equivalent:*

Theorem 12.27. 1. $x = \sum_{u\in\beta} \langle x, u\rangle u$ *for all* $x \in H$.

2. $\|x\|^2 = \sum_{u\in\beta} |\langle x, u\rangle|^2$ *for all* $x \in H$.

Moreover, either of these two conditions implies that $\beta \subset H$ is a maximal orthonormal set. However $\beta \subset H$ being a maximal orthonormal set is not sufficient to conditions for 1) and 2) hold!

Proof. As in the proof of Theorem 12.21, 1) implies 2). For 2) implies 1) let $\Lambda \subset\subset \beta$ and consider

$$\left\|x - \sum_{u\in\Lambda}\langle x, u\rangle u\right\|^2 = \|x\|^2 - 2\sum_{u\in\Lambda}|\langle x, u\rangle|^2 + \sum_{u\in\Lambda}|\langle x, u\rangle|^2$$
$$= \|x\|^2 - \sum_{u\in\Lambda}|\langle x, u\rangle|^2.$$

Since $\|x\|^2 = \sum_{u\in\beta} |\langle x, u\rangle|^2$, it follows that for every $\epsilon > 0$ there exists $\Lambda_\epsilon \subset\subset \beta$ such that for all $\Lambda \subset\subset \beta$ such that $\Lambda_\epsilon \subset \Lambda$,

$$\left\|x - \sum_{u\in\Lambda}\langle x, u\rangle u\right\|^2 = \|x\|^2 - \sum_{u\in\Lambda}|\langle x, u\rangle|^2 < \epsilon$$

showing that $x = \sum_{u\in\beta} \langle x, u\rangle u$.

Suppose $x \in \beta^\perp$. If 2) is valid then $\|x\|^2 = 0$, i.e. $x = 0$. So β is maximal. Let us now construct a counter example to prove the last assertion.

Take $H = \mathrm{Span}\{e_i\}_{i=1}^\infty \subset \ell^2$ and let $\tilde{u}_n = e_1 - (n+1)e_{n+1}$ for $n = 1, 2 \ldots$. Applying Gramn-Schmidt to $\{\tilde{u}_n\}_{n=1}^\infty$ we construct an orthonormal set $\beta = \{u_n\}_{n=1}^\infty \subset H$. I now claim that $\beta \subset H$ is maximal. Indeed if $x \in \beta^\perp$ then $x \perp u_n$ for all n which then implies that for all $n \in \mathbb{N}$,

$$0 = (x, \tilde{u}_n) = x_1 - (n+1)x_{n+1}.$$

Therefore $x_{n+1} = (n+1)^{-1} x_1$ for all n. Since $x \in \mathrm{Span}\{e_i\}_{i=1}^\infty$, $x_N = 0$ for some N sufficiently large and therefore $x_1 = 0$ which in turn implies that $x_n = 0$ for all n. So $x = 0$ and hence β is maximal in H. On the other hand, β is not maximal in ℓ^2, sine the above argument shows that $\beta^\perp$ in ℓ^2 is given is the span of $v = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots)$. Let P be the orthogonal projection of ℓ^2 onto $\mathrm{Span}\beta = v^\perp$. Then

$$\sum_{i=1}^\infty \langle x, u_n\rangle u_n = Px,$$

so that $\sum_{i=1}^{\infty}\langle x,u_n\rangle u_n = x$ iff $x\in \mathrm{Span}\beta = v^{\perp}\subset \ell^2$. For example if $x=(1,0,0,\dots)\in H$ or $x=e_i$ for any i, then $x\notin v^{\perp}$ and hence

$$\sum_{i=1}^{\infty}\langle x,u_n\rangle u_n \neq x.$$

■

12.2. Appendix: Converse of the Parallelogram Law.

Proposition 12.28 (Parallelogram Law Converse)**.** *If $(X,\|\cdot\|)$ is a normed space such that Eq. (12.2) holds for all $x,y\in X$, then there exists a unique inner product on $\langle\cdot,\cdot\rangle$ such that $\|x\|:=\sqrt{\langle x,x\rangle}$ for all $x\in X$. In this case we say that $\|\cdot\|$ is a Hilbertian norm.*

Proof. If $\|\cdot\|$ is going to come from an inner product $\langle\cdot,\cdot\rangle$, it follows from Eq. (12.1) that

$$2\mathrm{Re}\langle x,y\rangle = \|x+y\|^2-\|x\|^2-\|y\|^2$$

and

$$-2\mathrm{Re}\langle x,y\rangle = \|x-y\|^2-\|x\|^2-\|y\|^2.$$

Subtracting these two equations gives the "polarization identity,"

$$4\mathrm{Re}\langle x,y\rangle = \|x+y\|^2-\|x-y\|^2.$$

Replacing y by iy in this equation then implies that

$$4\mathrm{Im}\langle x,y\rangle = \|x+iy\|^2-\|x-iy\|^2$$

from which we find

$$\langle x,y\rangle = \frac{1}{4}\sum_{\epsilon\in G}\epsilon\|x+\epsilon y\|^2 \tag{12.12}$$

where $G=\{\pm1,\pm i\}$ – a cyclic subgroup of $S^1\subset\mathbb{C}$. Hence if $\langle\cdot,\cdot\rangle$ is going to exists we must define it by Eq. (12.12).

Notice that

$$\begin{aligned}\langle x,x\rangle &= \frac{1}{4}\sum_{\epsilon\in G}\epsilon\|x+\epsilon x\|^2 = \|x\|^2+i\|x+ix\|^2-i\|x-ix\|^2\\ &= \|x\|^2+i\left|1+i\right|^2\|x\|^2-i\left|1-i\right|^2\|x\|^2=\|x\|^2.\end{aligned}$$

So to finish the proof of (4) we must show that $\langle x,y\rangle$ in Eq. (12.12) is an inner product. Since

$$\begin{aligned}4\langle y,x\rangle &= \sum_{\epsilon\in G}\epsilon\|y+\epsilon x\|^2=\sum_{\epsilon\in G}\epsilon\|\epsilon\left(y+\epsilon x\right)\|^2\\ &=\sum_{\epsilon\in G}\epsilon\|\epsilon y+\epsilon^2 x\|^2\\ &=\|y+x\|^2+\|-y+x\|^2+i\|iy-x\|^2-i\|-iy-x\|^2\\ &=\|x+y\|^2+\|x-y\|^2+i\|x-iy\|^2-i\|x+iy\|^2\\ &=4\overline{\langle x,y\rangle}\end{aligned}$$

it suffices to show that $x \to \langle x, y\rangle$ is linear for all $y \in H$. (The rest of this proof may safely be skipped by the reader.) For this we will need to derive an identity from Eq. (12.2). To do this we make use of Eq. (12.2) three times to find

$$\begin{aligned}\|x+y+z\|^2 &= -\|x+y-z\|^2 + 2\|x+y\|^2 + 2\|z\|^2\\ &= \|x-y-z\|^2 - 2\|x-z\|^2 - 2\|y\|^2 + 2\|x+y\|^2 + 2\|z\|^2\\ &= \|y+z-x\|^2 - 2\|x-z\|^2 - 2\|y\|^2 + 2\|x+y\|^2 + 2\|z\|^2\\ &= -\|y+z+x\|^2 + 2\|y+z\|^2 + 2\|x\|^2 - 2\|x-z\|^2 - 2\|y\|^2 + 2\|x+y\|^2 + 2\|z\|^2.\end{aligned}$$

Solving this equation for $\|x+y+z\|^2$ gives

$$\|x+y+z\|^2 = \|y+z\|^2 + \|x+y\|^2 - \|x-z\|^2 + \|x\|^2 + \|z\|^2 - \|y\|^2. \tag{12.13}$$

Using Eq. (12.13), for $x, y, z \in H$,

$$\begin{aligned}4\operatorname{Re}\langle x+z, y\rangle &= \|x+z+y\|^2 - \|x+z-y\|^2\\ &= \|y+z\|^2 + \|x+y\|^2 - \|x-z\|^2 + \|x\|^2 + \|z\|^2 - \|y\|^2\\ &- \left(\|z-y\|^2 + \|x-y\|^2 - \|x-z\|^2 + \|x\|^2 + \|z\|^2 - \|y\|^2\right)\\ &= \|z+y\|^2 - \|z-y\|^2 + \|x+y\|^2 - \|x-y\|^2\\ &= 4\operatorname{Re}\langle x, y\rangle + 4\operatorname{Re}\langle z, y\rangle. \qquad (12.14)\end{aligned}$$

Now suppose that $\delta \in G$, then since $|\delta| = 1$,

$$\begin{aligned}4\langle \delta x, y\rangle &= \frac{1}{4}\sum_{\epsilon\in G}\epsilon\|\delta x + \epsilon y\|^2 = \frac{1}{4}\sum_{\epsilon\in G}\epsilon\|x + \delta^{-1}\epsilon y\|^2\\ &= \frac{1}{4}\sum_{\epsilon\in G}\epsilon\delta\|x + \delta\epsilon y\|^2 = 4\delta\langle x, y\rangle \qquad (12.15)\end{aligned}$$

where in the third inequality, the substitution $\epsilon \to \epsilon\delta$ was made in the sum. Since

$$\operatorname{Im}\langle x, y\rangle = \operatorname{Re}\langle -ix, y\rangle$$

it follows from Eq. (12.14) and (12.15) that

$$\begin{aligned}4\operatorname{Im}\langle x+z, y\rangle &= 4\operatorname{Re}\langle -ix - iz, y\rangle = 4\operatorname{Re}\langle -ix, y\rangle + 4\operatorname{Re}\langle -iz, y\rangle\\ &= 4\operatorname{Im}\langle x, y\rangle + 4\operatorname{Im}\langle z, y\rangle\end{aligned}$$

which combined with Eq. (12.14) shows

$$\langle x+z, y\rangle = \langle x, y\rangle + \langle z, y\rangle.$$

Because of this equation and Eq. (12.15) to finish the proof that $x \to \langle x, y\rangle$ is linear, it suffices to show $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$ for all $\lambda > 0$. Now if $\lambda = m \in \mathbb{N}$, then

$$\langle mx, y\rangle = \langle x + (m-1)x, y\rangle = \langle x, y\rangle + \langle (m-1)x, y\rangle$$

so that by induction $\langle mx, y\rangle = m\langle x, y\rangle$. Replacing x by x/m then shows that $\langle x, y\rangle = m\langle m^{-1}x, y\rangle$ so that $\langle m^{-1}x, y\rangle = m^{-1}\langle x, y\rangle$ and so if $n \in \mathbb{N}$ we find

$$\langle \frac{n}{m}x, y\rangle = n\langle \frac{1}{m}x, y\rangle = \frac{n}{m}\langle x, y\rangle$$

so that $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$ for all $\lambda > 0$ and $\lambda \in \mathbb{Q}$. By continuity, it now follows that $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$ for all $\lambda > 0$. (Question, could this have been carried out algebraically?) ■

12.3. **Appendix: Proofs via orthonormal bases.** In this appendix, let us give some proofs of the previous theorems making use of that fact that every Hilbert space has an orthonormal basis. (This appendix may safely be skipped.) As above, let H be a Hilbert space.

Let $\beta \subset H$ be an orthonormal set and $M = \overline{\text{span}\ \beta} \subset H$. Then M is a closed subspace of H and hence a Hilbert space. (Moreover $\beta \subset M$ is an orthonormal basis for M since if $x \in \beta^\perp$ then $x \in \overline{\text{span}\ \beta}^\perp = M^\perp$. Therefore if $x \in M$ and $x \in \beta^\perp$, then $x = 0$.) For $x \in H$, let $Px = \sum\limits_{u\in\beta} \langle x, u\rangle u$ then $P : H \to H$ is orthogonal projection of H onto M. Indeed, if $y = Px = \sum\limits_{u\in\beta} \langle x, u\rangle u$, then for $v \in \beta$,

$$\langle x - y, v\rangle = \langle x, v\rangle - \sum_{u\in\beta} \langle x, u\rangle\langle u, v\rangle = \langle x, v\rangle - \langle x, v\rangle = 0$$

showing $x - y \in \beta^\perp = M^\perp$. So if $m \in M$,

$$\|x - m\|^2 = \|(x - y) + y - m\|^2 = \|x - y\|^2 + \|y - m\|^2 \geq \|x - y\|^2$$

with equality iff $m = y$ which shows that y is the unique element in M minimizing the distance of x to M.

Let us also show using an orthonormal basis that the map

$$\begin{aligned} H &\longrightarrow H^* \\ x &\longrightarrow \langle\cdot, x\rangle = \ell_x \end{aligned}$$

is a conjugate linear isometric isomorphism. First off by the Schwarz's inequality,

$$\|\ell_x\| = \sup_{\|y\|=1} |\ell_x(y)| = \sup_{\|y\|=1} |\langle x, y\rangle| \leq \|x\|.$$

Moreover taking $y = x/\|x\|$, $|l_x(y)| = \|x\|$ showing that $\|\ell_x\| \geq \|x\|$ and therefore that $\|\ell_x\| = \|x\|$. So $x \to \ell_x$ is an isometric map which is easily seen to be conjugate linear. Hence we need only show that to every $f \in H^* \setminus \{0\}$ there exists $x \in H$ such that $f = \ell_x = \langle\cdot, x\rangle$. Let $\beta \subset H$ be an orthonormal basis for H, then if x is going to exist we would have

$$x = \sum_{u\in\beta} \langle x, u\rangle u = \sum_{u\in\beta} \bar{f}(u)u. \tag{12.16}$$

In order to make use of this formula, we need to show $\sum_{u\in\beta} |f(u)|^2 < \infty$. Suppose that $\Lambda \subset\subset \beta$ and $y \in \text{Span}\beta$, then

$$f(y) = f(\sum_{u\in\Lambda} \langle y, u\rangle u) = \sum_{u\in\Lambda} \langle y, u\rangle f(u) = \langle y, \sum_{u\in\Lambda} \bar{f}(u)u\rangle.$$

Taking $y = \sum_{u\in\Lambda} \bar{f}(u)u$ in this expression then shows that

$$f(y) = \|y\|^2 = \|y\|\,\|y\|$$

showing $\|y\|_H \leq \|f\|_{H^*}$. Squaring this inequality then shows that

$$\sum_{u\in\Lambda} |f(u)|^2 = \|y\|_H^2 \leq \|f\|_{H^*}^2 < \infty$$

and since $\Lambda \subset\subset \beta$ is arbitrary, it follows that $\sum_{u\in\beta} |f(u)|^2 < \infty$ as desired. Therefore we may define x by Eq. (12.16). Then for $y \in H$,

$$f(y) = f\left(\sum_{u\in\beta}\langle y,u\rangle u\right) = \sum_{u\in\beta}\langle y,u\rangle f(u) = \sum_{u\in\beta}\langle y,\bar{f}(u)u\rangle = \langle y,\sum_{u\in\beta}\bar{f}(u)u\rangle = \langle y,x\rangle.$$

13. Baire Category Theorem and its consequences

Recall that a set E is said to be nowhere dense iff $\left(\bar{E}\right)^o = \emptyset$, i.e. $\bar{E}$ has empty interior. Also notice that E is nowhere dense is equivalent to

$$X = \left(\left(\bar{E}\right)^o\right)^c = \overline{\left(\bar{E}\right)^c} = \overline{(E^c)^o}.$$

That is to say E is nowhere dense iff E^c has dense interior.

13.1. Baire Category Theorem.

Theorem 13.1 (Baire Category Theorem)**.** *Let (X,ρ) be a complete metric space.*

1. *If $\{V_n\}_{n=1}^\infty$ is a sequence of dense open sets, then $G := \bigcap\limits_{n=1}^\infty V_n$ is dense in X.*
2. *If $\{E_n\}_{n=1}^\infty$ is a sequence of nowhere dense sets, then $X \neq \bigcup_{n=1}^\infty E_n$.*

Proof. 1) We must shows that $\bar{G} = X$ which is equivalent to showing that $W \cap G \neq \emptyset$ for all non-empty open sets $W \subset X$. Since V_1 is dense, $W \cap V_1 \neq \emptyset$ and hence there exists $x_1 \in X$ and $\epsilon_1 > 0$ such that

$$\overline{B(x_1,\epsilon_1)} \subset W \cap V_1.$$

Since V_2 is dense, $B(x_1,\epsilon_1) \cap V_2 \neq \emptyset$ and hence there exists $x_2 \in X$ and $\epsilon_2 > 0$ such that

$$\overline{B(x_2,\epsilon_2)} \subset B(x_1,\epsilon_1) \cap V_2.$$

Continuing this way inductively, we may choose $\{x_n \in X \text{ and } \epsilon_n > 0\}_{n=1}^\infty$ such that

$$\overline{B(x_n,\epsilon_n)} \subset B(x_{n-1},\epsilon_{n-1}) \cap V_n\ \forall n.$$

Furthermore we can clearly do this construction in such a way that $\epsilon_n \downarrow 0$ as $n \uparrow \infty$. Hence $\{x_n\}_{n=1}^\infty$ is Cauchy sequence and $x = \lim\limits_{n\to\infty} x_n$ exists in X since X is complete. Since $\overline{B(x_n,\epsilon_n)}$ is closed, $x \in \overline{B(x_n,\epsilon_n)} \subset V_n$ so that $x \in V_n$ for all n and hence $x \in G$. Moreover, $x \in \overline{B(x_1,\epsilon_1)} \subset W \cap V_1$ implies $x \in W$ and hence $x \in W \cap G$ showing $W \cap G \neq \emptyset$.

2) For the second assertion, since $\bigcup_{n=1}^\infty E_n \subset \bigcup_{n=1}^\infty \bar{E}_n$, it suffices to shows that $X \neq \bigcup_{n=1}^\infty \bar{E}_n$ or equivalently that $\emptyset \neq \bigcap_{n=1}^\infty \left(\bar{E}_n\right)^c = \bigcap_{n=1}^\infty (E_n^c)^o$. As we have observed, E_n is nowhere dense is equivalent to $(E_n^c)^o$ being a dense open set, hence by part 1), $\bigcap_{n=1}^\infty (E_n^c)^o$ is dense in X and hence not empty. ■

Definition 13.2. A subset $E \subset X$ is **meager** or of the **first category** if $E = \bigcup\limits_{n=1}^\infty E_n$ where each E_n is nowhere dense. And a set $F \subset X$ is called **residual** if F^c is meager.

Remarks 13.3. The reader should think of meager as being the topological analogue of sets of measure 0 and residual as being the topological analogue of sets of full measure.

1. If F is a residual set, then there exists nowhere dense sets $\{E_n\}$ such that
$$F^c = \cup_{n=1}^{\infty} E_n \subset \cup_{n=1}^{\infty} \bar{E}_n.$$
Taking complements of this equation shows that
$$\cap_{n=1}^{\infty} \bar{E}_n^c \subset F,$$
i.e. F contains a set of the form $\cap_{n=1}^{\infty} V_n$ with each V_n being an open dense subset of X.

   Conversely, if $\cap_{n=1}^{\infty} V_n \subset F$ with each V_n being an open dense subset of X, then $F^c \subset \cup_{n=1}^{\infty} V_n^c$ and hence $F^c = \cup_{n=1}^{\infty} E_n$ where $E_n = F^c \cap V_n^c$, a nowhere dense subset of X. Therefore F is residual iff F contains a countable intersection of dense open sets.
2. A countable union of meager sets is meager and a subsets of a meager set is meager.
3. A countable intersection of residual sets is residual.

The Baire Category Theorem may be stated as follows. If X is a complete metric space, then (1) all residual sets are dense in X and 2) X is not meager.

13.2. Application to Banach Spaces.

Theorem 13.4 (Open Mapping Theorem)**.** *Let X, Y be Banach spaces, $T \in L(X,Y)$. If T is surjective then T is an open mapping.*

Proof. For all $\alpha > 0$ let $B_\alpha = \{x \in X : \|x\|_X < \alpha\} \subset X$, $E_\alpha = T(B_\alpha)$ and $B(0,\alpha) = \{y \in Y : \|y\|_Y < \alpha\}$.

Claim 1. For all $\alpha > 0$ there exists $\delta > 0$ such that $B(0,\delta) \subset \overline{E_\alpha}$.

Since $Y = \bigcup_{n-1}^{\infty} E_n$, the Baire category theorem implies there exists n such that $\overline{E}_n^0 \neq \emptyset$, i.e. there exists $y \in \overline{E}_n$ and $\epsilon > 0$ such that $\overline{B(y,\epsilon)} \subset \overline{E}_n$. Suppose $\|y'\| < \epsilon$ then y and $y+y'$ are in $B(y,\epsilon) \subset \overline{E}_n$ hence there exists $x', x \in B_n$ such that $\|Tx' - (y+y')\|$ and $\|Tx - y\|$ may be made as small as we please, which we abbreviate as follows
$$\|Tx' - (y+y')\| \approx 0 \text{ and } \|Tx - y\| \approx 0.$$
Hence by the triangle inequality,
$$\begin{aligned}\|T(x'-x) - y'\| &= \|Tx' - (y+y') - (Tx - y)\| \\ &\le \|Tx' - (y+y')\| + \|Tx - y\| \approx 0\end{aligned}$$
with $x' - x \in B_{2n}$. This shows that $y' \in \overline{E_{2n}}$ which implies $B(0,\epsilon) \subset \overline{E_{2n}}$. Since the map $\phi_\alpha : Y \to Y$ given by $\phi_\alpha(y) = \frac{\alpha}{2n} y$ is a homeomorphism, $\phi_\alpha(E_{2n}) = E_\alpha$ and $\phi_\alpha(B(0,\epsilon)) = B(0, \frac{\alpha\epsilon}{2n})$, it follows that $B(0,\delta) \subset \overline{E_\alpha}$ where $\delta \equiv \frac{\alpha\epsilon}{2n} > 0$.

Claim 2. There exists $\epsilon > 0$ such that $B(0,\epsilon) \subset E_1$, i.e. Claim 1. holds with out the closure.

By Claim 1, there exists $\delta > 0$ such that $B(0,\delta) \subset \overline{E}_1$. As in the proof of Claim 1. we also have $B(0,\alpha\delta) \subset \overline{E}_\alpha$ for all $\alpha > 0$. Now let $\epsilon := \delta/2$ and $y \in Y$ be such that $\|y\| < \epsilon = \delta/2$. Then $y \in B(0, \frac{1}{2}\delta) \subset \overline{E}_{\frac{1}{2}}$ so there exists There exists $x_1 \in B_{1/2}$ such $\|Tx_1 - y\| \approx 0$ and in particular we may assume that
$$\|y - Tx_1\| < 2^{-1}\epsilon = \delta/4 = 2^{-2}\delta.$$

Similarly, $y - Tx_1 \in B(0, \frac{1}{4}\delta) \subset \overline{E}_{\frac{1}{4}}$ implies there exists $x_2 \in B_{1/4}$ such that

$$\|y - Tx_1 - Tx_2\| < 2^{-2}\epsilon = 2^{-3}\delta$$

and hence an $x_3 \in B_{2^{-3}}$ such that

$$\|y - Tx_1 - Tx_2 - Tx_3\| < 2^{-3}\epsilon = 2^{-4}\delta.$$

So by induction, we find $x_n \in B_{2^{-n}}$ such that

$$\|y - \sum_{k=1}^{n} Tx_k\| = \|y - Tx_1 - Tx_2 - \cdots - Tx_n\| < 2^{-n}\epsilon = 2^{-(n+1)}\delta. \tag{13.1}$$

Since

$$\sum_{n=1}^{\infty} \|x_n\| < \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n = \frac{1}{2}\left(\frac{1}{1-\frac{1}{2}}\right) = 1,$$

$x \equiv \sum_{n=1}^{\infty} x_n$ exists and $\|x\| < 1$. Passing to the limit in Eq. (13.1) shows, $\|y - Tx\| = 0$. Thus we have shown if $\|y\| < \epsilon$ then $y \in T(B_1) = E_1$ which is the content of Claim 2.

We now show that T is open. If $x \in V \subset_o X$ and $y = Tx \in TV$ we must show that TV contains a ball $B(y, \delta)$ for some $\delta > 0$. Now the following statements are easily seen to be equivalent:

$$\begin{aligned} B(y,\delta) &= B(Tx,\delta) \subset TV \\ B(0,\delta) &\subset TV - Tx = T(V - x) \\ B(0,\alpha\delta) &\subset T\left[\alpha(V - x)\right] \end{aligned} \tag{13.2}$$

for some $\alpha > 0$. But since $V - x$ is a neighborhood of 0, there exists $\alpha > 0$ such that $B_1 \subset \alpha(V - x)$ and hence by Claim 2. there exists an $\epsilon > 0$ such that

$$B(0,\epsilon) \subset TB_1 \subset \alpha(V - x).$$

Therefore we see that Eq. (13.2) holds provided we choose $\delta = \epsilon/\alpha > 0$. ■

Corollary 13.5. *If X, Y are Banach spaces and $T \in L(X,Y)$ is invertible (i.e. a bijective linear transformation) then the inverse map, T^{-1}, is* ***bounded****, i.e. $T^{-1} \in L(Y,X)$. (Note that T^{-1} is automatically linear.)*

Theorem 13.6 (Closed Graph Theorem)**.** *Let X and Y be Banach space $T : X \to Y$ linear is continuous iff T is closed i.e. $\Gamma(T) \subset X \times Y$ is closed.*

Proof. If T continuous and $(x_n, Tx_n) \to (x,y) \in X \times Y$ as $n \to \infty$ then $Tx_n \to Tx = y$ which implies $(x,y) = (x,Tx) \in \Gamma(T)$.

Conversely: If T is **closed** then the following diagram commutes

$$\begin{tikzcd} & \Gamma(T) \arrow[dr, "\pi_2"] & \\ X \arrow[ur, "\Gamma"] \arrow[rr, "T"'] & & Y \end{tikzcd}$$

where $\Gamma(x) := (x, Tx)$.

The map $\pi_2 : X \times Y \to X$ is continuous and $\pi_1|_{\Gamma(T)} : \Gamma(T) \to X$ is continuous bijection which implies $\pi_1|_{\Gamma(T)}^{-1}$ is bounded by the open mapping Theorem 13.4.

Hence $T = \pi_2 \circ \pi_1|_{\Gamma(T)}^{-1}$ is bounded, being the composition of bounded operators. ■

As an application we have the following proposition.

Proposition 13.7. *Let H be a Hilbert space. Suppose that $T : H \to H$ is a linear (not necessarily bounded) map such that there exists $T^* : H \to H$ such that*

$$\langle Tx, Y\rangle = \langle x, T^*Y\rangle \ \forall \ x, y \in H.$$

Then T is bounded.

Proof. It suffices to show that T is closed. To prove this suppose that $x_n \in H$ such that $(x_n, Tx_n) \to (x, y) \in H \times H$. Then for any $z \in H$,

$$\langle Tx_n, z\rangle \quad = \quad \langle x_n, T^*z\rangle \quad \longrightarrow \quad \langle x, T^*z\rangle = \langle Tx, z\rangle \text{ as } n \to \infty.$$

On the other hand $\lim_{n\to\infty}\langle Tx_n, z\rangle = \langle y, z\rangle$ as well and therefore $\langle Tx, z\rangle = \langle y, z\rangle$ for all $z \in H$. This shows that $Tx = y$ and proves that T is closed. ■

Here is another example.

Example 13.8. Suppose that $\mathcal{M} \subset L^2([0,1], m)$ is a closed subspace such that each element of $\mathcal{M}$ has a representative in $C([0,1])$. We will abuse notation and simply write $\mathcal{M} \subset C([0,1])$. Then

1. There exists $A \in (0, \infty)$ such that $\|f\|_u \le A\|f\|_{L^2}$ for all $f \in \mathcal{M}$.
2. For all $x \in [0,1]$ there exists $g_x \in \mathcal{M}$ such that
$$f(x) = \langle f, g_x\rangle \text{ for all } f \in \mathcal{M}.$$
Moreover we have $\|g_x\| \le A$.
3. The subspace $\mathcal{M}$ is finite dimensional and $\dim(\mathcal{M}) \le A^2$.

Proof. 1) I will give a two proofs of part 1. Each proof requires that we first show that $(\mathcal{M}, \|\cdot\|_u)$ is a complete space. To prove this it suffices to shows that $\mathcal{M}$ is a closed subspace of $C([0,1])$. So let $\{f_n\} \subset \mathcal{M}$ and $f \in C([0,1])$ such that $\|f_n - f\|_u \to 0$ as $n \to \infty$. Then $\|f_n - f_m\|_{L^2} \le \|f_n - f_m\|_u \to 0$ as $m, n \to \infty$, and since $\mathcal{M}$ is closed in $L^2([0,1])$, $L^2 - \lim_{n\to\infty} f_n = g \in \mathcal{M}$. By passing to a subsequence if necessary we know that g $g(x) = \lim_{n\to\infty} f_n(x) = f(x)$ for m - a.e. x. So $f = g \in \mathcal{M}$.

i)Let $i : (\mathcal{M}, \|\cdot\|_u) \to (\mathcal{M}, \|\cdot\|_2)$ be the identity map. Then i is bounded and bijective. By the open mapping theorem, $j = i^{-1}$ is bounded as well. Hence there exists $A < \infty$ such that $\|f\|_u = \|j(f)\| \le A\|f\|_2$ for all $f \in \mathcal{M}$.

ii) Let $j : (\mathcal{M}, \|\cdot\|_2) \to (\mathcal{M}, \|\cdot\|_u)$ be the identity map. We will shows that j is a closed operator and hence bounded by the closed graph theorem. Suppose that $f_n \in \mathcal{M}$ such that $f_n \to f$ in L^2 and $f_n = j(f_n) \to g$ in $C([0,1])$. Then as in the first paragraph, we conclude that $g = f = j(f)$ a.e. showing j is closed. Now finish as in last line of proof i).

2) For $x \in [0,1]$, let $e_x : \mathcal{M} \to \mathbb{C}$ be the evaluation map $e_x(f) = f(x)$. Then

$$|e_x(f)| \le |f(x)| \le \|f\|_u \le A\|f\|_{L^2}$$

which shows that $e_x \in \mathcal{M}^*$. Hence there exists a unique element $g_x \in \mathcal{M}^*$ such that

$$f(x) = e_x(f) = \langle f, g_x\rangle \text{ for all } f \in \mathcal{M}.$$

Moreover $\|g_x\|_{L^2} = \|e_x\|_{\mathcal{M}^*} \le A$.

3) Let $\{f_j\}_{j=1}^n$ be an orthonormal subset of $\mathcal{M}$. Then

$$A^2 \geq \|e_x\|_{\mathcal{M}^*}^2 = \|g_x\|_{L^2}^2 \geq \sum_{j=1}^n |\langle f_j, g_x\rangle|^2 = \sum_{j=1}^n |f_j(x)|^2$$

and integrating this equation over $x \in [0,1]$ implies that

$$A^2 \geq \sum_{j=1}^n \int_0^1 |f_j(x)|^2 dx = \sum_{j=1}^n 1 = n$$

which shows that $n \leq A^2$. Hence $\dim(\mathcal{M}) \leq A^2$. ■

Remark 13.9. Keeping the notation in Example 13.8, $G(x,y) = g_x(y)$ for all $x, y \in [0,1]$. Then

$$f(x) = e_x(f) = \int_0^1 f(y)\overline{G(x,y)}dy \text{ for all } f \in \mathcal{M}.$$

The function G is called the reproducing kernel for $\mathcal{M}$.

The above example generalizes as follows.

Proposition 13.10. *Suppose that $(X, \mathcal{M}, \mu)$ is a finite measure space, $p \in [1,\infty)$ and W is a closed subspace of $L^p(\mu)$ such that $W \subset L^p(\mu) \cap L^\infty(\mu)$. Then* $\dim(W) < \infty$.

Proof. With out loss of generality we may assume that $\mu(X) = 1$. As in Example 13.8, we shows that W is a closed subspace of $L^\infty(\mu)$ and hence by the open mapping theorem, there exists a constant $A < \infty$ such that $\|f\|_\infty \leq A\|f\|_p$ for all $f \in W$. Now if $1 \leq p \leq 2$, then

$$\|f\|_\infty \leq A\|f\|_p \leq A\|f\|_2$$

and if $p \in (2,\infty)$, then $\|f\|_p^p \leq \|f\|_2^2 \|f\|_\infty^{p-2}$ or equivalently,

$$\|f\|_p \leq \|f\|_2^{2/p} \|f\|_\infty^{1-2/p} \leq \|f\|_2^{2/p} \left(A\|f\|_p\right)^{1-2/p}$$

from which we learn that $\|f\|_p \leq A^{1-2/p}\|f\|_2$ and therefore that $\|f\|_\infty \leq AA^{1-2/p}\|f\|_2$ so that in any case there exists a constant $B < \infty$ such that $\|f\|_\infty \leq B\|f\|_2$.

Let $\{f_n\}_{n=1}^N$ be an orthonormal subset of W and $f = \sum_{n=1}^N c_n f_n$ with $c_n \in \mathbb{C}$, then

$$\left\|\sum_{n=1}^N c_n f_n\right\|_\infty^2 \leq B^2 \sum_{n=1}^N |c_n|^2 \leq B^2 |c|^2$$

where $|c|^2 := \sum_{n=1}^N |c_n|^2$. For each $c \in \mathbb{C}^N$, there is an exception set E_c such that for $x \notin E_c$,

$$\left|\sum_{n=1}^N c_n f_n(x)\right|^2 \leq B^2 |c|^2.$$

Let $\mathbb{D} := (\mathbb{Q}+i\mathbb{Q})^N$ and $E = \cap_{c\in\mathbb{D}} E_c$. Then $\mu(E) = 0$ and for $x \notin E$, $\left|\sum_{n=1}^N c_n f_n(x)\right| \le B^2 |c|^2$ for all $c \in \mathbb{D}$. By continuity it then follows for $x \notin E$ that

$$\left|\sum_{n=1}^N c_n f_n(x)\right|^2 \le B^2 |c|^2 \text{ for all } c \in \mathbb{C}^N.$$

Taking $c_n = f_n(x)$ in this inequality implies that

$$\left|\sum_{n=1}^N |f_n(x)|^2\right|^2 \le B^2 \sum_{n=1}^N |f_n(x)|^2 \text{ for all } x \notin E$$

and therefore that

$$\sum_{n=1}^N |f_n(x)|^2 \le B^2 \text{ for all } x \notin E.$$

Integrating this equation over x then implies that $N \le B^2$, i.e. $\dim(W) \le B^2$. ■

Theorem 13.11 (Uniform Boundedness Principle)**.** *Let X and Y be a normed vector spaces. Suppose $\mathcal{A} \subset L(X,Y)$ and let*

$$R = R_{\mathcal{A}} := \{x \in X : \sup_{A\in\mathcal{A}} \|Ax\| = \infty\}. \tag{13.3}$$

Then $\sup_{A\in\mathcal{A}} \|A\| < \infty$ iff R is not residual. In particular if X is a Banach space and $\sup_{A\in\mathcal{A}} \|Ax\| < \infty$ for all $x \in X$ then $\sup_{A\in\mathcal{A}} \|A\| < \infty$.

Proof. If $M := \sup_{A\in\mathcal{A}} \|A\| < \infty$, then $\sup_{A\in\mathcal{A}} \|Ax\| \le M\|x\| < \infty$ for all $x \in X$, so that $R = \emptyset$ and R is not residual. Conversely, if R is not residual, then $R^c = \{x \in X : \sup_{A\in\mathcal{A}} \|Ax\| < \infty\}$ is not meager. For each $n \in \mathbb{N}$, let $E_n \subset X$ be the closed sets given by

$$E_n = \{x : \sup_{A\in\mathcal{A}} \|Ax\| \le n\} = \bigcap_{A\in\mathcal{A}} \{x : \|Ax\| \le n\}.$$

Then $R^c = \cup_{n=1}^\infty E_n$ and since R^c is not meager, there exists an $n \in \mathbb{N}$ such that $E_n^0 \ne 0$. Let $B(x,\delta)$ be a ball such that $\overline{B(x,\delta)} \subset E_n$. Then for $y \in X$ with $\|y\| = \delta$ we know $x - y \in \overline{B(x,\delta)} \subset E_n$, so that $Ay = Ax - A(x-y)$ and hence for any $A \in \mathcal{A}$,

$$\|Ay\| \le \|Ax\| + \|A(x-y)\| \le n + n = 2n.$$

Hence it follows that $\|A\| \le 2n/\delta$ for all $A \in \mathcal{A}$, i.e. $\sup_{A\in\mathcal{A}} \|A\| \le 2n/\delta < \infty$.

If $\sup_{A\in\mathcal{A}} \|Ax\| < \infty$, then $R = \emptyset$. If X is a Banach space, then residual sets are dense and hence $R = \emptyset$ is not a Residual set. Therefore $\sup_{A\in\mathcal{A}} \|A\| < \infty$ by the first part of the Theorem. ■

Example 13.12. Suppose that $\{c_n\}_{n=1}^\infty \subset \mathbb{C}$ is a sequence of numbers such that

$$\lim_{N\to\infty} \sum_{n=1}^N a_n c_n \text{ exists in } \mathbb{C} \text{ for all } a \in \ell^1.$$

Then $c \in \ell^\infty$.

Proof. Let $f_N \in \left(\ell^1\right)^*$ be given by $f_N(a) = \sum_{n=1}^N a_n c_n$ and set $M_N := \max\{|c_n| : n = 1, \dots, N\}$. Then

$$|f_N(a)| \le M_N \|a\|_{\ell^1}$$

and by taking $a = e_k$ with k such $M_N = |c_k|$, we learn that $\|f_N\| = M_N$. Now by assumption, $\lim_{N\to\infty} f_N(a)$ exists for all $a \in \ell^1$ and in particular,

$$\sup_N |f_N(a)| < \infty \text{ for all } a \in \ell^1.$$

So by the Theorem 13.11,

$$\infty > \sup_N \|f_N\| = \sup_N M_N = \sup\{|c_n| : n = 1, 2, 3, \dots\}.$$

■

13.3. **Applications to Fourier Series.** Let $T = S^1$ be the unit circle in S^1 and m denote the normalized arc length measure on T. So if $f : T \to [0, \infty)$ is measurable, then

$$\int_T f(w)dw := \int_T f dm := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})d\theta.$$

Also let $\phi_n(z) = z^n$ for all $n \in \mathbb{Z}$. Recall that $\{\phi_n\}_{n\in\mathbb{Z}}$ is an orthonormal basis for $L^2(T)$. For $n \in \mathbb{N}$ let

$$\begin{aligned} s_n(f, z) &:= \sum_{k=-n}^{n} \langle f, \phi_n\rangle \phi_k(z) = \sum_{k=-n}^{n} \langle f, \phi_n\rangle z^k = \sum_{k=-n}^{n} \left(\int_T f(w)\bar{w}^k dw\right) z^k \\ &= \int_T f(w)\left(\sum_{k=-n}^{n} \bar{w}^k z^k\right) dw = \int_T f(w) d_n(z\bar{w})dw \end{aligned}$$

where $d_n(\alpha) := \sum_{k=-n}^{n} \alpha^k$. Now $\alpha d_n(\alpha) - d_n(\alpha) = \alpha^{n+1} - \alpha^{-n}$, so that

$$d_n(\alpha) := \sum_{k=-n}^{n} \alpha^k = \frac{\alpha^{n+1} - \alpha^{-n}}{\alpha - 1}$$

with the convention that

$$\frac{\alpha^{n+1} - \alpha^{-n}}{\alpha - 1}|_{\alpha=1} = \lim_{\alpha\to 1} \frac{\alpha^{n+1} - \alpha^{-n}}{\alpha - 1} = 2n + 1 = \sum_{k=-n}^{n} 1^k.$$

Writing $\alpha = e^{i\theta}$, we find

$$\begin{aligned} D_n(\theta) := d_n(e^{i\theta}) &= \frac{e^{i\theta(n+1)} - e^{-i\theta n}}{e^{i\theta} - 1} = \frac{e^{i\theta(n+1/2)} - e^{-i\theta(n+1/2)}}{e^{i\theta/2} - e^{-i\theta/2}} \\ &= \frac{\sin(n + \frac{1}{2})\theta}{\sin \frac{1}{2}\theta}. \end{aligned}$$

Recall by Hilbert space theory, $L^2(T) - \lim_{n\to\infty} s_n(f, \cdot) = f$ for all $f \in L^2(T)$. We will now show that the convergence is not pointwise for all $f \in C(T) \subset L^2(T)$.

Proposition 13.13. *h $z \in T$, there exists a residual set $R_z \subset C(T)$ such that $\sup_n |s_n(f, z)| = \infty$ for all $f \in R_z$. Recall that $C(T)$ is a complete metric space, hence R_z is a dense subset of $C(T)$.*

Proof. By symmetry considerations, it suffices to take $z = 1 \in T$. Let $\Lambda_n f := s_n(f,1)$. Then

$$|\Lambda_n f| = \left|\int_T f(w) d_n(\bar{w}) dw\right| \le \int_T |f(w) d_n(\bar{w})|\, dw \le \|f\|_u \int_T |d_n(\bar{w})|\, dw$$

showing

$$\|\Lambda_n\| \le \int_T |d_n(\bar{w})|\, dw.$$

Since $C(T)$ is dense in $L^1(T)$, there exists $f_k \in C(T,\mathbb{R})$ such that $f_k(w) \to \operatorname{sgn} d_k(\bar{w})$ in L^1. By replacing f_k by $(f_k \wedge 1) \vee (-1)$ we may assume that $\|f_k\|_u \le 1$. It now follows that

$$\|\Lambda_n\| \ge \frac{|\Lambda_n f_k|}{\|f_k\|_u} \ge \left|\int_T f_k(w) d_n(\bar{w}) dw\right|$$

and passing to the limit as $k \to \infty$ implies that $\|\Lambda_n\| \ge \int_T |d_n(\bar{w})|\, dw$. Hence we have shown that

$$\|\Lambda_n\| = \int_T |d_n(\bar{w})|\, dw = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|d_n(e^{-i\theta})\right| d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}\right| d\theta. \tag{13.4}$$

Since

$$\sin x = \int_0^x \cos y dy \le \int_0^x |\cos y|\, dy \le x$$

for all $x \ge 0$. Since $\sin x$ is even, $|\sin x| \le |x|$ for all $x \in \mathbb{R}$. Using this in Eq. (13.4) implies that

$$\begin{aligned}\|\Lambda_n\| &\ge \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\frac{\sin(n+\frac{1}{2})\theta}{\frac{1}{2}\theta}\right| d\theta = \frac{2}{\pi}\int_0^{\pi} \left|\sin(n+\frac{1}{2})\theta\right| \frac{d\theta}{\theta}. \\ &= \frac{2}{\pi}\int_0^{\pi} \left|\sin(n+\frac{1}{2})\theta\right| \frac{d\theta}{\theta}\end{aligned}$$

Since

$$\int_0^{\pi} \left|\sin(n+\frac{1}{2})\theta\right| \frac{d\theta}{\theta} = \int_0^{(n+\frac{1}{2})\pi} |\sin y| \frac{dy}{y} \to \infty \text{ as } n \to \infty,$$

we learn that $\sup_n \|\Lambda_n\| = \infty$. So by Theorem 13.11,

$$R_1 = \{f \in C(T) : \sup_n |\Lambda_n f| = \infty\}$$

is a residual set. ■

See Rudin Chapter 5 for more details.

Lemma 13.14. *For $f \in L^1(T)$, let*

$$\hat{f}(n) := \langle f, \phi_n\rangle = \int_T f(w)\bar{w}^n dw.$$

Then $\hat{f} \in c_0$ and the map $f \in L^1(T) \to c_0$ is a one to one bounded linear transformation into but not onto c_0.

Proof. By Bessel's inequality, $\sum_{n\in\mathbb{Z}}\left|\hat{f}(n)\right|^2 < \infty$ for all $f \in L^2(T)$ and in particular $\lim_{|n|\to\infty}\left|\hat{f}(n)\right| = 0$. Given $f \in L^1(T)$ and $g \in L^2(T)$ we have

$$\left|\hat{f}(n) - \hat{g}(n)\right| = \left|\int_T [f(w) - g(w)]\,\bar{w}^n dw\right| \le \|f - g\|_1$$

and hence

$$\limsup_{n\to\infty}\left|\hat{f}(n)\right| = \limsup_{n\to\infty}\left|\hat{f}(n) - \hat{g}(n)\right| \le \|f - g\|_1$$

for all $g \in L^2(T)$. Since $L^2(T)$ is dense in $L^1(T)$, it follows that $\limsup_{n\to\infty}\left|\hat{f}(n)\right| = 0$ for all $f \in L^1$, i.e. $\hat{f} \in c_0$.

Since $\left|\hat{f}(n)\right| \le \|f\|_1$, we have $\left\|\hat{f}\right\|_{c_0} \le \|f\|_1$ showing that $\Lambda f := \hat{f}$ is a bounded linear transformation from $L^1(T)$ to c_0.

To see that Λ is injective, suppose $\hat{f} = \Lambda f \equiv 0$, then $\int_T f(w)p(w,\bar{w})dw = 0$ for all polynomials p in w and $\bar{w}$. By the Stone - Wierstrass theorem, this implies that

$$\int_T f(w)g(w)dw = 0$$

for all $g \in C(T)$. Since $C(T)$ is dense in $L^1(T, |f(w)|\,dw)$, there exists bounded functions $g_n \in C(T)$ such that $g_n \to \overline{\operatorname{sgn} f} \in L^1(T, |f(w)|\,dw)$. Thus

$$0 = \lim_{n\to\infty}\int_T f(w)g_n(w)dw = \lim_{n\to\infty}\int_T g_n(w)\operatorname{sgn} f(w)\,|f(w)|\,dw = \int_T |f(w)|\,dw$$

which shows that $f = 0$ a.e.

If Λ were surjective, the open mapping theorem would imply that $\Lambda^{-1} : c_0 \to L^1(T)$ is bounded. In particular this implies there exists $C < \infty$ such that

$$\|f\|_{L^1} \le C\left\|\hat{f}\right\|_{c_0} \text{ for all } f \in L^1(T). \tag{13.5}$$

Taking $f = d_n$, we find $\left\|\hat{d}_n\right\|_{c_0} = 1$ while $\lim_{n\to\infty}\|d_n\|_{L^1} = \infty$ contradicting Eq. (13.5). Therefore $\operatorname{ran}(\Lambda) \ne c_0$. ■

14. L^p-spaces

Let $(X, \mathcal{M}, \mu)$ be a measure space and for $0 < p < \infty$ and a measurable function $f : X \to \mathbb{C}$ let

$$\|f\|_p \equiv (\int |f|^p d\mu)^{1/p}.$$

When $p = \infty$, let

$$\|f\|_\infty = \inf\{a \ge 0 : \mu(|f| > a) = 0\}$$

For $1 \le p \le \infty$, let

$$L^p(X, \mathcal{M}, \mu) = \{f : X \to \mathbb{C} : f \text{ is measurable and } \|f\|_p < \infty\}/\sim$$

where $f \sim g$ iff $f = g$ a.e. Notice that $\|f - g\|_p = 0$ iff $f \sim g$ and if $f \sim g$ then if $f \sim g$ then $\|f\|_p = \|g\|_p$. In general we will continue to write f for the equivalence class containing f.

Remark 14.1. We have $\|f\|_\infty \le M$ iff $|f(x)| \le M$ for μ - a.e. x. To see this, suppose that $\|f\|_\infty \le M$, then for all $a > M$, $\mu(|f| > a) = 0$ and therefore $\mu(|f| > M) = \lim_{n\to\infty} \mu(|f| > M + 1/n) = 0$, so that $|f(x)| \le M$ for μ - a.e. x. Conversely, if $|f| \le M$ a.e., then for $a > M$, $\mu(|f| > a) = 0$ and hence $\|f\|_\infty \le M$. So in conclusion we have shown

$$\|f\|_\infty = \inf \{a \ge 0 : |f(x)| \le a \text{ for } \mu - \text{a.e. } x\}.$$

Our first goal in this section is to show $\left(L^p(X, \mathcal{M}, \mu), \|\cdot\|_p\right)$ is a Banach space. We will start with the case $p = \infty$.

Theorem 14.2. *The function $\|\cdot\|_\infty$ is a norm on L^∞ and $(L^\infty(X, \mathcal{M}, \mu), \|\cdot\|_\infty)$ is a Banach space. A sequence $\{f_n\}_{n=1}^\infty \subset L^\infty$ converges to $f \in L^\infty$ iff there exists $E \in \mathcal{M}$ such that $\mu(E) = 0$ and $f_n \to f$ uniformly on E^c. Moreover, bounded simple function are dense in L^∞.*

Proof. Suppose that $f, g \in L^\infty$, then $|f| \le \|f\|_\infty$ a.e. and $|g| \le \|g\|_\infty$ a.e. so that $|f + g| \le |f| + |g| \le \|f\|_\infty + \|g\|_\infty$ a.e. and therefore

$$\|f + g\|_\infty \le \|f\|_\infty + \|g\|_\infty.$$

Therefore $\|\cdot\|_\infty$ satisfies the triangle inequality. The reader may easily check the remaining conditions that ensure $\|\cdot\|_\infty$ is a norm.

Suppose that $\{f_n\}_{n=1}^\infty \subset L^\infty$ is a sequence such $f_n \to f \in L^\infty$, i.e. $\|f - f_n\|_\infty \to 0$ as $n \to \infty$. Then for all $k \in \mathbb{N}$, there exists $N_k < \infty$ such that

$$\mu\left(|f - f_n| > k^{-1}\right) = 0 \text{ for all } n \ge N_k.$$

Let

$$E = \cup_{k=1}^\infty \cup_{n \ge N_k} \left\{|f - f_n| > k^{-1}\right\}.$$

Then $\mu(E) = 0$ and for $x \in E^c$, we have $|f(x) - f_n(x)| \le k^{-1}$ for all $n \ge N_k$. This shows that $f_n \to f$ uniformly on E^c. Conversely, if there exists $E \in \mathcal{M}$ such that $\mu(E) = 0$ and $f_n \to f$ uniformly on E^c, then for any $\epsilon > 0$,

$$\mu(|f - f_n| \ge \epsilon) = \mu(\{|f - f_n| \ge \epsilon\} \cap E^c) = 0$$

for all n sufficiently large. That is to say $\limsup_{n\to\infty} \|f - f_n\|_\infty \le \epsilon$ for all $\epsilon > 0$. The density of simple functions will be left as an exercise to the reader.

So the last thing to prove is the completeness of L^∞ for which we will use Theorem 10.3. Suppose that $\{f_n\}_{n=1}^\infty \subset L^\infty$ is a sequence such that $\sum_{n=1}^\infty \|f_n\|_\infty < \infty$. Let $M_n := \|f_n\|_\infty$, $E_n := \{|f_n| > M_n\}$, and $E := \cup_{n=1}^\infty E_n$ so that $\mu(E) = 0$. Then for $x \in E^c$ we have

$$\sum_{n=1}^\infty \sup_{x \in E^c} |f_n(x)| \le \sum_{n=1}^\infty M_n < \infty$$

which shows that $S_N(x) = \sum_{n=1}^N f_n(x)$ converges uniformly to $S(x) := \sum_{n=1}^\infty f_n(x)$ on E^c, that is to say $\|S - S_n\|_\infty \to 0$ as $n \to \infty$. ■

14.1. **Some inequalities.**

Proposition 14.3. *Let $f : [0,\infty) \to [0,\infty)$ be a continuous strictly increasing function such that $f(0) = 0$ (for simplicity) and $\lim_{s\to\infty} f(s) = \infty$. Let $g = f^{-1}$ and for $s, t \geq 0$ let*

$$F(s) = \int_0^s f(s')ds' \text{ and } G(t) = \int_0^t g(t')dt'.$$

Then for all $s, t \geq 0$,

$$st \leq F(s) + G(t)$$

and equality holds iff $t = f(s)$.

Proof. Let

$$A_s := \{(\sigma,\tau) : 0 \leq \tau \leq f(\sigma) \text{ for } 0 \leq \sigma \leq s\} \text{ and}$$
$$B_t := \{(\sigma,\tau) : 0 \leq \sigma \leq g(\tau) \text{ for } 0 \leq \tau \leq t\}$$

then as one sees from Figure 4, $[0,s]\times[0,t] \subset A_s \cup B_t$. (In the figure: $s = 3$, $t = 1$, A_3 is the region under $t = f(s)$ for $0 \leq s \leq 3$ and B_1 is the region to the right of the curve $s = g(t)$ for $0 \leq t \leq 1$.) Hence if m is Lebesgue measure on the plane,

$$st = m([0,s]\times[0,t]) \leq m(A_s) + m(B_t) = F(s) + G(t).$$

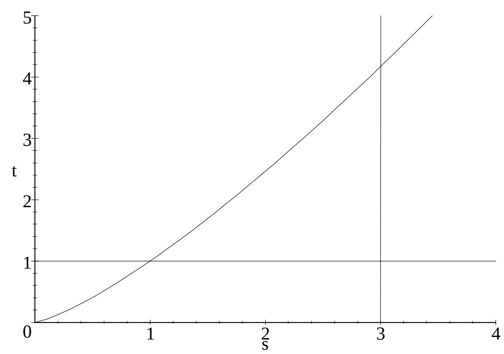


FIGURE 4. A picture proof of Proposition 14.3.

■

Definition 14.4. The conjugate exponent $q \in [1,\infty]$ to $p \in [1,\infty]$ is $q := \frac{p}{p-1}$ with the convention that $q = \infty$ if $p = 1$. Notice that q satisfies

$$\frac{1}{p} + \frac{1}{q} = 1,\ 1 + \frac{q}{p} = q,\ p - \frac{p}{q} = 1 \text{ and } q(p-1) = p. \tag{14.1}$$

Lemma 14.5. *If $s, t \geq 0$ then*

$$st \leq \frac{s^q}{q} + \frac{t^p}{p}$$

with equality if and only if $s^q = t^p$.

Proof. Let $F(s) = \frac{s^p}{p}$ for $p > 1$,

$$f(s) = s^{p-1} = t$$

or

$$g(t) = t^{\frac{1}{p-1}} = t^{q-1}$$

because $q \equiv 1/(p-1) + 1$ or $1/(p-1) = q - 1$. Therefore $G(t) = t^q/q$ and hence

$$st \leq \frac{s^p}{p} + \frac{t^q}{q}$$

with equality iff $t = s^{p-1}$. ■

Theorem 14.6 (Hölder's inequality)**.** *Suppose that* $1 < p < \infty$ *and* $q := \left(\frac{p}{p-1}\right)$, *or equivalently*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

If f, g *are measurable functions then* $\|fg\|_1 \leq \|f\|_p \cdot \|g\|_q$ *with equality iff*

$$\left(\frac{|f|}{\|f\|_p}\right)^p = \left(\frac{|g|}{\|g\|_q}\right)^q \text{ a.e.}$$

Proof. The cases where $\|f\|_q = 0$ or ∞ or $\|g\|_p = 0$ or ∞ are easy to deal with and are left to the reader. So we will assume now that $0 < \|f\|_q, \|g\|_p < \infty$. Let $s = |f|/\|f\|_p$ and $t = |g|/\|g\|_q$ then Lemma 14.5 implies

$$\frac{|fg|}{\|f\|_p \|g\|_q} \leq \frac{1}{p} \frac{|f|^p}{\|f\|_p} + \frac{1}{q} \frac{|g|^q}{\|g\|^q}.$$

Integrating this equation then gives

$$\frac{\|fg\|_1}{\|f\|_p \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1$$

with equality iff

$$\frac{|g|}{\|g\|_q} = \frac{|f|^{p-1}}{\|f\|_p^{(p-1)}} \text{ a.e.} \iff \frac{|g|}{\|g\|_q} = \frac{|f|^{p/q}}{\|f\|_p^{p/q}} \text{ a.e.} \iff |g|^q \|f\|_p^p = \|g\|_q^q |f|^p \text{ a.e.}$$

■

Theorem 14.7 (Minikowski's Inequality)**.** *If* $1 \leq p < \infty$ *and* $f, g \in L^p$ *then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p,$$

with equality iff

$$\operatorname{sgn}(f) = \operatorname{sgn}(g) \text{ when } p = 1 \text{ and}$$
$$f = cg \text{ for some } c > 0 \text{ when } p > 1.$$

Proof. Since

$$|f + g|^p \leq (2 \max(|f|, |g|))^p = 2^p \max(|f|^p, |g|^p) \leq 2^p (|f|^p + |g|^p)$$

it follows that

$$\|f + g\|_p^p \leq 2^p \left(\|f\|_p^p + \|g\|_p^p\right) < \infty.$$

In the case that $p = 1$,

$$\|f + g\|_1 = \int_X |f + g| d\mu \leq \int_X |f| d\mu + \int_X |g| d\mu$$

with equality iff

$$|f| + |g| = |f + g| \text{ a.e.} \quad \Longleftrightarrow \text{ sgn}(f) = \text{sgn}(g) \text{ a.e.}$$

Now assume that $p > 1$. We may assume $\|f + g\|_p > 0$ since if $\|f + g\|_p = 0$ the theorem is easily verified. Now

$$|f + g|^p = |f + g||f + g|^{p-1} \le (|f| + |g|)|f + g|^{p-1}$$

with equality iff $\text{sgn}(f) = \text{sgn}(g)$ a.e. Integrating this equation and applying Holder's inequality gives

$$\begin{aligned}\int_X |f + g|^p d\mu &\le \int_X |f| \, |f + g|^{p-1} d\mu + \int_X |g| \, |f + g|^{p-1} d\mu \\ &\le (\|f\|_p + \|g\|_p) \| \, |f + g|^{p-1} \, \|_q \end{aligned} \tag{14.2}$$

with equality iff

$$\left(\frac{|f|}{\|f\|_p}\right)^p = \left(\frac{|f + g|^{p-1}}{\||f + g|^{p-1}\|_q}\right)^q = \left(\frac{|g|}{\|g\|_p}\right)^p \text{ a.e.}$$
$$\text{and } \text{sgn}(f) = \text{sgn}(g) \text{ a.e.}$$

Now

$$\||f + g|^{p-1}\|_q^q = \int_X (|f + g|^{p-1})^q d\mu = \int_X |f + g|^p d\mu \tag{14.3}$$

wherein we have used Eq. (14.1), $q(p - 1) = p$. Combining Eqs. (14.2) and (14.3) implies

$$\|f + g\|_p^p \le \|f\|_p \|f + g\|_p^{p/q} + \|g\|_p \|f + g\|_p^{p/q} \tag{14.4}$$

with equality iff

$$\begin{aligned} &\text{sgn}(f) = \text{sgn}(g) \text{ and} \\ &\left(\frac{|f|}{\|f\|_p}\right)^p = \frac{|f + g|^p}{\|f + g\|_p^p} = \left(\frac{|g|}{\|g\|_p}\right)^p \text{ a.e.}\end{aligned} \tag{14.5}$$

Solving for $\|f + g\|_p$ in Eq. (14.4) with the aid of Eq. (14.1) shows that $\|f + g\|_p \le \|f\|_p + \|g\|_p$ with equality iff Eq. (14.5) holds which happens iff $f = cg$ a.e. with $c > 0$. ■

Theorem 14.8 (Completeness of $L^p(\mu)$). *Suppose that $\{f_n\} \subset L^p(\mu)$ is Cauchy, then there exists $f \in L^p(\mu)$ such that $f_n \xrightarrow{L^p} f$. Moreover f is unique modulo the equivalence relation of being equal off sets of measure zero.*

Proof. Write

$$\|f\| = \left(\int_X |f|^p \, d\mu\right)^{1/p}.$$

By Chebyschev's inequality,

$$\begin{aligned}\mu\left(|f_n - f_m| \ge \epsilon\right) &= \mu\left(|f_n - f_m|^p \ge \epsilon^p\right) \\ &\le \frac{1}{\epsilon^p} \int_X |f_n - f_m|^p \, d\mu = \frac{1}{\epsilon^p} \|f_n - f_m\|^p \to 0 \text{ as } m, n \to \infty\end{aligned}$$

for all $\epsilon > 0$. This shows that $\{f_n\}$ is L^0-Cauchy (i.e. Cauchy in measure) so there exists $\{g_j\} \subset \{f_n\}$ such that $g_j \to f$ a.e. Now by Fatou's Lemma,

$$\begin{aligned}\|g_j - f\|^p &= \int \lim_{k\to\infty} \inf |g_j - g_k|^p \le \lim_{k\to\infty} \inf \int |g_j - g_k|^p d\mu \\ &= \lim_{k\to\infty} \inf \|g_j - g_k\|^p \to 0 \text{ as } j \to \infty.\end{aligned}$$

In particular, $\|f\| \le \|g_j - f\| + \|g_j\| < \infty$ so the $f \in L^p$ and $g_j \xrightarrow{L^p} f$. The proof is finished because,

$$\|f_n - f\| \le \|f_n - g_j\| + \|g_j - f\| \to 0 \text{ as } j, n \to \infty.$$

■

14.2. **Corollaries of Hölder's Inequality.** The $L^p(\mu)$ – norm controls two types of behaviors of f, namely the "behavior at infinity" and the behavior of local singularities So in particular, if f is blowing up at a point $x_0 \in X$, then locally near x_0 it is harder for $f \in L^p(\mu)$ as p increases. On the other hand a function $f \in L^p(\mu)$ is allowed to decay at infinity slower and slower as p increases. With these "insights" in mind, we should not in general expect $L^p(\mu) \subset L^q(\mu)$ or $L^q(\mu) \subset L^p(\mu)$. However, there are two notable exceptions that we shall prove below. (1) If $\mu(X) < \infty$, then there is no behavior at infinity to worry about and we expect that $L^q(\mu) \subset L^p(\mu)$ for all $q \le p$. See Corollary 14.10 below. (2) If μ is counting measure, i.e. $\mu(A) = \#(A)$, then all functions in $L^p(\mu)$ for any p can not blow up on a set of positive measure, so there are no local singularities. Hence we expect in this case that $L^p(\mu) \subset L^q(\mu)$ for all $q \le p$, see Corollary 14.13 below.

Corollary 14.9. *Suppose that $r, p, q \in (0, \infty]$ are numbers such that*

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{q}, \tag{14.6}$$

then for measurable functions $f, g : X \to \mathbb{C}$,

$$\|fg\|_r \le \|f\|_p \cdot \|g\|_q.$$

Proof. The case $r = \infty$ is easy and is left to the reader. So assume $r \in (0, \infty)$. By Eq. (14.6), $a = p/r$ and $b = q/r$ are conjugate exponents. Thus by Hölder's inequality,

$$\|fg\|_r^r = \||f|^r |g|^r\|_1 \le \| |f|^r \|_a \cdot \| |g|^r \|_b = \|f\|_p^{r/p} \cdot \|g\|_q^{r/q}$$

from which the desired result follows. ■

Corollary 14.10. *If $\mu(X) < \infty$, then $L^p(\mu) \subset L^q(\mu)$ for all $0 < p < q \le \infty$ and the inclusion map is bounded.*

Proof. Choose $a \in [1, \infty]$ such that

$$\frac{1}{p} = \frac{1}{a} + \frac{1}{q}, \text{ i.e. } a = \frac{q-p}{pq}.$$

Then by Corollary 14.9,

$$\|f\|_p = \|f \cdot 1\|_p \le \|f\|_q \cdot \|1\|_a = \mu(X)^{1/a} \|f\|_q = \mu(X)^{(\frac{1}{p} - \frac{1}{q})} \|f\|_q.$$

The reader may easily check this final formula is correct even when $q = \infty$ provided we interpret $(1/p - 1/q) = 1/p$ in this case. ■

Proposition 14.11. *Suppose that $0 < p < q < r \le \infty$, then $L^q \subset L^p + L^r$, i..e every function $f \in L^q$ may be written as $f = g + h$ with $g \in L^p$ and $h \in L^r$.*

Proof. The blow up points of f are contained in the set $E := \{|f| > 1\}$ and the behavior of f at infinity is solely determined by f on E^c. Hence let $g = f1_E$ and $h = f1_{E^c}$ so that $f = g + h$. By our discussion at the beginning of this section we expect that $g \in L^p$ and $h \in L^r$. Indeed,

$$|g|^p = |f|^p 1_E = |f|^p 1_{\{|f|>1\}} \le |f|^q 1_{\{|f|>1\}} \le |f|^q$$

which after integrating shows that $\|g\|_p^p \le \|f\|_q^q$. Similarly

$$|h|^r = |f|^r 1_{E^c} = |f|^r 1_{\{|f|\le 1\}} \le |f|^q 1_{\{|f|\le 1\}} \le |f|^q$$

so that $\|h\|_r^r \le \|f\|_q^q$. ■

Corollary 14.12. *Suppose that $0 < p < q < r \le \infty$, then $L^p \cap L^r \subset L^q$ and we have*

$$\|f\|_q \le \|f\|_p^\lambda \|f\|_r^{1-\lambda}$$

where $\lambda \in (0,1)$ is determined so that

$$\frac{1}{q} = \frac{\lambda}{p} + \frac{1-\lambda}{r} \text{ with } \lambda = p/q \text{ if } r = \infty.$$

Proof. Let λ be determined as above, $a = p/\lambda$ and $b = r/(1-\lambda)$, then by Corollary 14.9,

$$\|f\|_q = \left\| |f|^\lambda |f|^{1-\lambda} \right\|_q \le \left\| |f|^\lambda \right\|_a \left\| |f|^{1-\lambda} \right\|_b = \|f\|_p^\lambda \|f\|_r^{1-\lambda}.$$

■

Again, the heuristic explanation of this corollary is that if $f \in L^p \cap L^r$, then f has local singularities no worse than an L^r function and behavior at infinity no worse than an L^p function. Hence $f \in L^p$ for any q between p and r.

Corollary 14.13. *Suppose now that μ is counting measure. Then $L^p(\mu) \subset L^q(\mu)$ for all $0 < p \le q \le \infty$.*

Proof. Suppose that $0 < p < r = \infty$, then

$$\|f\|_\infty^p = \sup\{|f(x)|^p : x \in X\} \le \sum_{x \in X} |f(x)|^p = \|f\|_p^p.$$

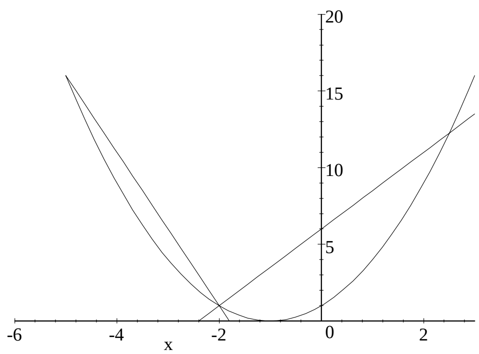
Now apply Corollary 14.12 with $r = \infty$ to find

$$\|f\|_q \le \|f\|_p^{p/q} \|f\|_\infty^{1-p/q} \le \|f\|_p^{p/q} \|f\|_p^{1-p/q} = \|f\|_p.$$

■

14.2.1. *Jensen's Inequality.*

Definition 14.14. A function $\phi : (a,b) \to \mathbb{R}$ is convex if for all $a < x_0 < x_1 < b$ and $t \in [0,1]$ $\phi(x_t) \le t\phi(x_1) + (1-t)\phi(x_0)$ where $x_t = tx_1 + (1-t)x_0$.



A convex function with along with two cords corresponding to $x_0 = -2$ and $x_1 = 4$ and $x_0 = -5$ and $x_1 = -2$.

The following Proposition is clearly motivated by Figure 14.14.

Proposition 14.15. *Suppose that $\phi : (a, b) \to \mathbb{R}$ is a convex function, then*

1. *For all $u, v, w, z \in (a, b)$ such that $u < v$, $w < z$, $u \le w$ and $v \le z$ we have*

$$\frac{\phi(v) - \phi(u)}{v - u} \le \frac{\phi(z) - \phi(w)}{z - w}. \tag{14.7}$$

2. *For each $c \in (a, b)$, the right and left sided derivatives $\phi'_{\pm}(c)$ exists in $\mathbb{R}$ and if $a < u < v < b$, then $\phi'_+(u) \le \phi'_-(v) \le \phi'_+(v)$.*
3. *The function ϕ is continuous.*
4. *For all $t \in (a, b)$ there exists $\beta \in \mathbb{R}$ such that $\phi(x) \ge \phi(t) + \beta(x - t)$ for all $x \in (a, b)$.*

Proof. 1a) Suppose first that $u < v = w < z$, in which case Eq. (14.7) is equivalent to

$$(\phi(v) - \phi(u))(z - v) \le (\phi(z) - \phi(v))(v - u)$$

which after solving for $\phi(v)$ is equivalent to the following equations holding:

$$\phi(v) \le \phi(z)\frac{v - u}{z - u} + \phi(u)\frac{z - v}{z - u}.$$

But this last equation states that $\phi(v) \le \phi(z)t + \phi(u)(1 - t)$ where $t = \frac{v-u}{z-u}$ and $v = tz + (1 - t)u$ and hence is valid by the definition of ϕ being convex.

1b) Now assume that Suppose first that $u = w < v < z$, in which case Eq. (14.7) is equivalent to

$$(\phi(v) - \phi(u))(z - u) \le (\phi(z) - \phi(u))(v - u)$$

which after solving for $\phi(v)$ is equivalent to

$$\phi(v)(z - u) \le \phi(z)(v - u) + \phi(u)(z - v)$$

which is equivalent to

$$\phi(v) \le \phi(z)\frac{v - u}{z - u} + \phi(u)\frac{z - v}{z - u}.$$

Again this equation is valid by the convexity of ϕ.

1c) $u < w < v = z$, in which case Eq. (14.7) is equivalent to

$$(\phi(z) - \phi(u))(z - w) \le (\phi(z) - \phi(w))(z - u)$$

and this is equivalent to the inequality,

$$\phi(w) \le \phi(z)\frac{w-u}{z-u} + \phi(u)\frac{z-w}{z-u}$$

which again is true by the convexity of ϕ.

1) General case. If $u < w < v < z$, then by 1a-1c)

$$\frac{\phi(z)-\phi(w)}{z-w} \ge \frac{\phi(v)-\phi(w)}{v-w} \ge \frac{\phi(v)-\phi(u)}{v-u}$$

and if $u < v < w < z$

$$\frac{\phi(z)-\phi(w)}{z-w} \ge \frac{\phi(w)-\phi(v)}{w-v} \ge \frac{\phi(w)-\phi(u)}{w-u}.$$

We have now taken care of all possible cases.

2) On the set $a < w < z < b$, Eq. (14.7) shows that $(\phi(z)-\phi(w))\,/\,(z-w)$ is a decreasing function in w and an increasing function in z and therefore $\phi'_{\pm}(x)$ exists for all $x \in (a,b)$. Also from Eq. (14.7) we learn that

$$\phi'_{+}(u) \le \frac{\phi(z)-\phi(w)}{z-w} \text{ for all } a < u < w < z < b, \tag{14.8}$$

$$\frac{\phi(v)-\phi(u)}{v-u} \le \phi'_{-}(z) \text{ for all } a < u < v < z < b, \tag{14.9}$$

and letting $w \uparrow z$ in the first equation also implies that

$$\phi'_{+}(u) \le \phi'_{-}(z) \text{ for all } a < u < z < b.$$

The inequality, $\phi'_{-}(z) \le \phi'_{+}(z)$, is also an easy consequence of Eq. (14.7).

3) Since $\phi(x)$ has both left and right finite derivatives, it follows that ϕ is continuous. (For an alternative proof, see Rudin.)

4) Given t, let $\beta \in [\phi'_{-}(t), \phi'_{+}(t)]$, then by Eqs. (14.8) and (14.9),

$$\frac{\phi(t)-\phi(u)}{t-u} \le \phi'_{-}(t) \le \beta \le \phi'_{+}(t) \le \frac{\phi(z)-\phi(t)}{z-t}$$

for all $a < u < t < z < b$. Item 4. now follows. ■

Corollary 14.16. *Suppose $\phi : (a,b) \to \mathbb{R}$ is differential then ϕ is convex iff ϕ' is non decreasing. In particular if $\phi \in C^2(a,b)$ then ϕ is convex iff $\phi'' \ge 0$.*

Proof. By Proposition 14.15, if ϕ is convex then ϕ' is non-decreasing. Conversely if ϕ' is increasing then

$$\frac{\phi(x_1)-\phi(c)}{x_1-c} = \phi'(\xi_1) \text{ for some } \xi_1 \in (c, x_1)$$

and

$$\frac{\phi(c)-\phi(x_0)}{c-x_0} = \phi'(\xi_2) \text{ for some } \xi_2 \in (x_0, c).$$

Hence

$$\frac{\phi(x_1)-\phi(c)}{x_1-c} \ge \frac{\phi(c)-\phi(x_0)}{c-x_0}$$

for all $x_0 < c < x_1$ from which it follows that ϕ is convex. ■

Example 14.17. The function $\exp(x)$ is convex , x^p is convex iff $p \ge 1$ and $-\log(x)$ is convex.

Theorem 14.18 (Jensen's Inequality)**.** *Suppose that $(X, \mathcal{M}, \mu)$ is a probability space, i.e. μ is a positive measure and $\mu(X) = 1$. Also suppose that $f \in L^1(\mu)$, $f : X \to (a, b)$, and $\phi : (a, b) \to \mathbb{R}$ is a convex function. Then*

$$\phi\left(\int_X f d\mu\right) \le \int_X \phi(f) d\mu$$

where if $\phi \circ f \notin L^1(\mu)$, then $\phi \circ f$ is integrable in the extended sense and $\int_X \phi(f) d\mu = \infty$.

Proof. Let $t = \int_X f d\mu \in (a, b)$ and let $\beta \in \mathbb{R}$ such that $\phi(s) - \phi(t) \ge \beta(s - t)$ for all $s \in (a, b)$. Then integrating the inequality, $\phi(f) - \phi(t) \ge \beta(f - t)$, implies that

$$0 \le \int_X \phi(f) d\mu - \phi(t) = \int_X \phi(f) d\mu - \phi(\int_X f d\mu).$$

Moreover, if $\phi(f)$ is not integrable, then $\phi(f) \ge \phi(t) + \beta(f - t)$ which shows that negative part of $\phi(f)$ is integrable. Therefore, $\int_X \phi(f) d\mu = \infty$ in this case. ■

As an application, we may use Corollary 14.16 and Theorem 14.18 to give another proof of Lemma 14.5. Let $a = \ln s$ and $b = \ln t$. Then by Jensen's inequality,

$$st = e^{(a+b)} = e^{\left(\frac{1}{q} qa + \frac{1}{p} pa\right)} \le \frac{1}{q} e^{qa} + \frac{1}{p} e^{pa} = \frac{1}{q} s^q + \frac{1}{p} t^p$$

with equality iff $qa = pa$ iff $s^q = t^p$. As a check

$$st = ss^{q/p} = s^{1+q/p} = s^q$$

so by Eq. (14.1)

$$st = s^q = \frac{s^q}{q} + \frac{s^q}{p} = \frac{s^q}{q} + \frac{t^p}{p}.$$

14.3. **The Dual of L^p spaces.** Throughout this section we assume $(X, \mathcal{M}, \mu)$ is a σ-finite measure space, $q \in [1, \infty]$ and $p \in [1, \infty]$ are conjugate exponents, i.e. $p^{-1} + q^{-1} = 1$. For $g \in L^q$, let $\phi_g \in (L^p)^*$ be given by

$$\phi_g(f) = \int gf \ d\mu.$$

By Hölder's inequality

$$|\phi_g(f)| \le \int |gf| d\mu \le \|g\|_q \|f\|_p \tag{14.10}$$

which implies that

$$\|\phi_g\|_{(L^p)^*} \le \|g\|_q. \tag{14.11}$$

We now show that the reverse inequality.

Proposition 14.19. *Keeping the notation above, for all $g \in L^q$,*

$$\|\phi_g\|_{(L^p)^*} = \|g\|_q. \tag{14.12}$$

Proof. Assume first that $q < \infty$ so $p > 1$. We want equality to hold in (14.10) which happens iff

$$|gf| = gf \text{ and } \left(\frac{|g|}{\|g\|_q}\right)^q = \left(\frac{|f|}{\|f\|_p}\right)^p \text{ a.e.}$$

Therefore we should take $|f| = \|f\|_p |g|^{q/p}/\|g\|_q^{q/p}$ and $\mathrm{sgn}(f)\mathrm{sgn}(g) = 1$ so $\mathrm{sgn}(f) = \overline{\mathrm{sgn}(g)}$. Therefore

$$f = \mathrm{sgn}(f)|f| = \overline{\mathrm{sgn}(g)} \frac{\|f\|_p}{\|g\|_q^{q/p}} |g|^{q/p}.$$

and since the constants are irrelevant for our purposes, let $f = \overline{\mathrm{sgn}(g)}|g|^{q/p}$, where $|g|^{q/p} \equiv 1$ if $q = 1$ and $p = \infty$. (Alternatively, just try f of the form $f = \overline{\mathrm{sgn}(g)}|g|^\lambda$. In order for $f \in L^p$ we must choose λ such that $\lambda p = q$, i.e. $\lambda = q/p$ as above.) If $p = \infty$, we find

$$|\phi_g(f)| = \int_X g\, \overline{\mathrm{sgn}(g)} d\mu = \|g\|_{L^1(\mu)} = \|g\|_1 \|f\|_\infty$$

which shows that $\|\phi_g\|_{(L^\infty)^*} \ge \|g\|_1$. If $p < \infty$, then

$$\|f\|_p^p = \int |f|^p = \int |g|^q = \|g\|_q^q$$

while

$$\phi_g(f) = \int gf d\mu = \int |g||g|^{q/p} d\mu = \int |g|^q d\mu = \|g\|_q^q.$$

Hence

$$\frac{|\phi_g(f)|}{\|f\|_p} = \frac{\|g\|_q^q}{\|g\|_q^{q/p}} = \|g\|_q^{q(1-\frac{1}{p})} = \|g\|_q.$$

This shows that $||\phi_g|| \ge \|g\|_q$ which combined with Eq. (14.11) implies Eq. (14.12).

Now assume that $p = 1$ and $q = \infty$ and let $\|g\|_\infty = M$. Choose $X_n \in \mathcal{M}$ such that $X_n \uparrow X$ as $n \to \infty$ and $\mu(X_n) < \infty$ for all n. For any $\epsilon > 0$, $\mu(|g| \ge M - \epsilon) > 0$ and $X_n \cap \{|g| \ge M - \epsilon\} \uparrow \{|g| \ge M - \epsilon\}$. Therefore, $\mu(X_n \cap \{|g| \ge M - \epsilon\}) > 0$ for n sufficiently large. Let

$$f = \overline{\mathrm{sgn}(g)} 1_{X_n \cap \{|g| \ge M - \epsilon\}},$$

then

$$\|f\|_1 = \mu(X_n \cap \{|g| \ge M - \epsilon\}) \in (0, \infty).$$

Moreover,

$$\begin{aligned} |\phi_g(f)| &= \int_{X_n \cap \{|g| \ge M-\epsilon\}} \overline{\mathrm{sgn}(g)} g d\mu = \int_{X_n \cap \{|g| \ge M-\epsilon\}} |g| d\mu \\ &\ge (M - \epsilon)\mu(X_n \cap \{|g| \ge M - \epsilon\}) = (M - \epsilon)\|f\|_1 \end{aligned}$$

and this implies that $\|\phi_g\|_{(L^1)^*} \ge M - \epsilon$. Since $\epsilon > 0$ is arbitrary, it follows that $\|\phi_g\|_{(L^1)^*} \ge M = \|g\|_\infty$. ■

Theorem 14.20. *Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space and suppose that $p, q \in [1, \infty]$ are conjugate exponents. Then for $p \in [1, \infty)$, the map $g \in L^q \to \phi_g \in (L^p)^*$ is an isometric isomorphism of Banach space. We summarize this by writing $(L^p)^* = L^q$ for all $1 \le p < \infty$.*

Proof. The only point that we have not yet proved is the surjectivity of the map $g \in L^q \to \phi_g \in (L^p)^*$. (When $p = 2$ this follows from the general Hilbert space theory.)

Case 1. Assume $\mu(X) < \infty$. Let $\phi \in (L^p)^*$. If $\phi = \phi_g$ for some $g \in L^q$ then $g \in L^1$ (by Holder's inequality) and

$$\phi(1_A) = \int_A g d\mu \tag{14.13}$$

for all $A \in \mathcal{M}$. This suggests that we let $\lambda(A) \equiv \phi(1_A)$. Let us now show that λ is a complex measure and $\lambda \ll \mu$.

Suppose that $A = \coprod_{n=1}^\infty A_n$ with $A_n \in \mathcal{M}$, then[16]

$$\|1_A - \sum_{n=1}^N 1_{A_n}\|_{L^p} = \|1_{\cup_{n=N+1}^\infty A_n}\|_{L^p} = \left[\mu(\cup_{n=N+1}^\infty A_n)\right]^{\frac{1}{p}} \to 0 \text{ as } N \to \infty.$$

Therefore

$$\lambda(A) = \phi(1_A) = \sum_1^\infty \phi(1_{A_n}) = \sum_1^\infty \lambda(A_n).$$

Combining the last equation with the fact that $\lambda(\emptyset) = 0$ implies that λ is a complex measure. Moreover if $\mu(A) = 0$, $1_A = 0$ in L^p and thus, $\lambda(A) = \phi(1_A) = \phi(0) = 0$ showing $\lambda \ll \mu$.

The Radon-Nikodym theorem now implies there exists $g = d\lambda/d\mu \in L^1(\mu)$. Then for this g, Eq. (14.13) is valid for all $A \in \mathcal{M}$ and then by linearity we find

$$\phi(f) = \int_X fg d\mu \text{ for all } f \in S \tag{14.14}$$

where S is the space of complex valued simple functions on X. Given a bounded measurable function g on X, let $f_n \in S$ such that $|f_n| \le |f|$ and $f_n \to f$ pointwise as $n \to \infty$. By the dominated convergence theorem, $f_n \to f$ in $L^p(\mu)$ and therefore,

$$\phi(f) = \lim_{n\to\infty} \phi(f_n) = \lim_{n\to\infty} \int_X f_n g d\mu = \int_X fg d\mu \tag{14.15}$$

and Eq. (14.14) holds for all bounded measurable functions f. We will now show that $g \in L^q(\mu)$.

Given $M < \infty$, let $\phi_M(f) := \phi(1_{|g| \le M} f)$. Then for $f \in L^\infty$, we have by Eq. (14.15) that

$$\phi_M(f) = \int_X fg 1_{|g|\le M} d\mu = \phi_{g_M}(f)$$

where $g_M := g1_{|g|\le M} \in L^\infty \subset L^p$ for all $p \ge 1$. Moreover,

$$|\phi_M(f)| \le \left|\phi(1_{|g|\le M} f)\right| \le \|\phi\| \left\|1_{|g|\le M} f\right\|_{L^p} \le \|\phi\| \, \|f\|_{L^p}$$

so that $\|\phi_M\|_{(L^p)^*} \le \|\phi\|_{(L^p)^*}$. Since $L^\infty(\mu)$ is dense in L^p, it follows that $\phi_M = \phi_{g_M}$ for all $M < \infty$. By Proposition 14.19, we learn

$$\|g_M\|_q = \|\phi_M\|_{(L^p)^*} \le \|\phi\|_{(L^p)^*}.$$

Fatou's lemma or the dominated convergence theorem now allows us to let $M \to \infty$ to conclude that $\|g\|_q \le \|\phi\|_{(L^p)^*}$[17]. Once we know this, it again follows from Eq. (14.15) and the density of $L^\infty \subset L^p$ that $\phi = \phi_g$.

[16]It is at this point that the proof breaks down when $p = \infty$.

[17]The argument leading to this conclusion may be replaced by the reverse Holder inequality of Theorem 14.22 below.

Case 2. Now suppose that μ is σ-finite and $X_n \in \mathcal{M}$ are sets such that $0 < \mu(X_n) < \infty$ and $X = \coprod_{n=1}^{\infty} X_n$. Identify $L^p(X_n, \mathcal{M}_n, \mu_n)$ with

$$\{f \in L^p(X, \mathcal{M}, \mu)) : f = 0 \text{ on } E_n^c\}.$$

So $\phi|_{L^p(X_n)}$ is still bounded on $L^p(X_n)$ and hence by Case 1. there exists $g_n \in L^q(X_n)$ such that $\phi|_{L^p(X_n,\mu_n)} = \phi_{g_n}$ on $L^p(X_n, \mu_n)$ for all n. Define $g \equiv \sum_{n=1}^{\infty} g_n$. Notice that

$$\phi|_{L^p(\cup_{n=1}^N X_n)} = \phi_{\sum_{n=1}^{N} g_n} \text{ on } L^p(\bigcup_{n=1}^{N} X_n)$$

and this implies that

$$\|\sum_{n=1}^{N} g_n\|_q = \|\phi|_{L^p(\cup_{n=1}^N X_n)}\| \leq \|\phi\| \tag{14.16}$$

for all N. Since $\{g_n\}_{n=1}^{\infty}$ have disjoint supports,

$$|\sum_{n=1}^{N} g_n| = \sum_{n=1}^{N} |g_n| \uparrow |g| \text{ as } N \to \infty$$

and so we may pass to the limit in Eq. (14.16) using the monotone convergence theorem to show

$$\|g\|_q = \|\sum_{n=1}^{\infty} g_n\|_q = \lim_{N\to\infty} \|\sum_{n=1}^{N} g_n\|_q \leq \|\phi\|_{(L^p)^*} < \infty.$$

Each $f \in L^p(X)$ decomposes as $f = \sum_{n=1}^{\infty} f\, 1_{X_n}$ and by the dominated convergence theorem implies,

$$\|f - \sum_{1}^{N} f 1_{X_n}\|_p = \|f\left(1 - 1_{\cup_{n=N+1}^{\infty} X_n}\right)\|_p \to 0 \text{ as } N \to \infty.$$

Hence

$$\phi(f) = \sum_{n=1}^{\infty} \phi(f 1_{X_n}) = \sum_{n=1}^{\infty} \phi_g(f 1_{X_n}) = \phi_g(f)$$

since $\phi(f) = \phi_g(f)$ for all $f \in L^p(X_n)$ and ϕ and ϕ_g are continuous on $L^p(X)$. ■

Theorem 14.22 fails in general when $p = \infty$.

Example 14.21. Consider $X = [0, 1]$, $\mathcal{M} = \mathcal{B}$, and $\mu = m$. Then $(L^\infty)^* \neq L^1$.

Proof. Let $M := C([0, 1])\text{“} \subset \text{”} L^\infty([0, 1], dm)$. It is easily seen for $f \in M$, that $\|f\|_\infty = \sup\{f(x) : x \in [0, 1]\}$ for all $f \in M$. Therefore M is a closed subspace of L^∞. Define $\ell(f) = f(0)$ for all $f \in M$. Then $\ell \in M^*$ with norm 1. By the Hahn-Banach theorem there exists an extension $L \in (L^\infty)^*$ such that $L = \ell$ on M and $\|L\| = 1$. Suppose that there exists $g \in L^1$ such that

$$L(f) = \phi_g(f) = \int_{[0,1]} fg dm \text{ for all } f \in L^\infty.$$

Let $f_n(x) = (1 - nx) 1_{x \leq n^{-1}}$, then $L(f_n) = 1$ for all n and hence

$$1 = \lim_{n\to\infty} L(f_n) = \lim_{n\to\infty} \int_{[0,1]} f_n g dm = \int_{\{0\}} g dm = 0,$$

wherein we have made use of the Dominated convergence theorem in the second to last equality. From this contradiction, we conclude that $L \neq \phi_g$ for any $g \in L^1$. ■

14.4. Converse of Hölder's Inequality.

Theorem 14.22 (Converse of Hölder's Inequality)**.** *Assume that* $(X, \mathcal{M}, \mu)$ *is a* σ *– compact measure space,* $q \in [1, \infty]$ *and* p *be the corresponding conjugate exponent. Also let* S *denote the set of simply functions* f *on* X *such that* $\mu(f \neq 0) < \infty$*. For* $g : X \to \mathbb{C}$ *measurable such that* $fg \in L^1$ *for all* $f \in S$,[18] *let*

$$M_q(g) = \sup\left\{\left|\int_X fg d\mu\right| : f \in S \text{ with } \|f\|_p = 1\right\}. \tag{14.17}$$

If $M_q(g) < \infty$ *then* $g \in L^q$ *and* $M_q(g) = \|g\|_q$.

Proof. As above, let $X_n \in \mathcal{M}$ be sets such that $\mu(X_n) < \infty$ and $X_n \uparrow X$ as $n \uparrow \infty$. Suppose that $q = 1$ and hence $p = \infty$. Choose simple functions f_n on X such that $|f_n| \leq 1$ and $\overline{\operatorname{sgn}(g)} = \lim_{n\to\infty} f_n$ in the pointwise sense. Then $1_{X_m} f_n \in S$ and therefore

$$\left|\int_X 1_{X_m} f_n g d\mu\right| \leq M_q(g)$$

for all m, n. By assumption $1_{X_m} g \in L^1(\mu)$ and therefore by the dominated convergence theorem we may let $n \to \infty$ in this equation to find

$$\int_X 1_{X_m} |g| \, d\mu \leq M_q(g)$$

for all m. The monotone convergence theorem then implies that

$$\int_X |g| \, d\mu = \lim_{n\to\infty} \int_X 1_{X_m} |g| \, d\mu \leq M_q(g)$$

showing $g \in L^1(\mu)$ and $\|g\|_1 \leq M_q(g)$. Since Holder's inequality implies that $M_q(g) \leq \|g\|_1$, we have proved the theorem in case $q = 1$.

For $q > 1$, we will begin by assuming that $g \in L^q(\mu)$. Since $p \in [1, \infty)$ we know that S is a dense subspace of $L^p(\mu)$ and therefore

$$M_q(g) = \sup\left\{\left|\int_X fg d\mu\right| : f \in L^p(\mu) \text{ with } \|f\|_p = 1\right\} = \|g\|_q$$

where the last equality follows by Proposition 14.19.

So it remains to show that if $fg \in L^1$ for all $f \in S$ and $M_q(g) < \infty$ then $g \in L^q(\mu)$. For $n \in \mathbb{N}$, let $g_n \equiv 1_{X_n} 1_{|g| \leq n} g$. Then $g_n \in L^q(\mu)$, in fact $\|g_n\|_q \leq n\mu(X_n)^{1/q} < \infty$. So by the previous paragraph,

$$\begin{aligned}\|g_n\|_q = M_q(g_n) &= \sup\left\{\left|\int_X f 1_{X_n} 1_{|g| \leq n} g d\mu\right| : f \in L^p(\mu) \text{ with } \|f\|_p = 1\right\} \\ &\leq M_q(g) \left\|f 1_{X_n} 1_{|g| \leq n}\right\|_p \leq M_q(g)\end{aligned}$$

[18] This is equivalent to requiring $1_A g \in L^1(\mu)$ for all $A \in \mathcal{M}$ such that $\mu(A) < \infty$.

wherein the second to last inequality we have made use of the definition of $M_q(g)$ and the fact that $f 1_{X_n} 1_{|g| \leq n} \in S$. If $q \in (1, \infty)$, an application of the monotone convergence theorem (or Fatou's Lemma) now shows that

$$\|g\|_q = \lim_{n \to \infty} \|g_n\|_q \leq M_q(g) < \infty.$$

And if $q = \infty$, then $\|g_n\|_\infty \leq M_q(g) < \infty$ for all n implies $|g_n| \leq M_q(g)$ a.e. which then implies that $|g| \leq M_q(g)$ a.e. since $|g| = \lim_{n \to \infty} |g_n|$. That is $g \in L^\infty(\mu)$ and $\|g\|_\infty \leq M_\infty(g)$. ■