MTH503: Functional Analysis

Semester 1, 2022-2023

Prahlad Vaidyanathan

Contents

I. Preliminaries

1. Review of Linear Algebra

Note: All vector spaces in this course will be over $\mathbb{R}$ or $\mathbb{C}$ ($= \mathbb{K}$), and will be assumed to be *non-zero*.

Definition 1.1. A Hamel basis for a vector space $\mathbf{E}$ is a set $\Lambda \subset \mathbf{E}$ such that every element of $\mathbf{E}$ can be expressed uniquely as a finite linear combination of elements in Λ.

Theorem 1.2 (Zorn's Lemma)**.** *Let $(\mathcal{F}, \leq)$ be a partially ordered set such that every totally ordered subset has an upper bound. Then $\mathcal{F}$ has a maximal element.*

Theorem 1.3. *Every vector space has a Hamel basis. In fact, if $\Lambda_0 \subset \mathbf{E}$ is any linearly independent set, then there exists a Hamel basis Λ of $\mathbf{E}$ such that $\Lambda_0 \subset \Lambda$.*

Example 1.4.

(i) For $\mathbf{E} = \mathbb{K}^n$, we write $e_i := (0, 0, \ldots, 0, 1, 0, \ldots, 0)$ (with 1 in the i^{th} position). The set $\{e_i : 1 \leq i \leq n\}$ is called the standard basis for $\mathbb{K}^n$.

(ii) Define

$$c_{00} := \{(x_n)_{n=1}^\infty : x_i \in \mathbb{K}, \text{ and there exists } N \in \mathbb{N} \text{ such that } x_i = 0 \text{ for all } i \geq N\}$$

It is a vector space over $\mathbb{K}$ where the vector space operators are defined componentwise. Write e_i for the sequence

$$(e_i)_j = \delta_{i,j} = \begin{cases} 1 & : \text{ if } i = j, \\ 0 & : \text{ otherwise.} \end{cases}$$

Then, $\{e_i : i \in \mathbb{N}\}$ is a basis for c_{00}.

(iii) Define

$$c_0 = \{(x_n)_{n=1}^\infty : x_i \in \mathbb{K}, \text{ and } \lim_{i \to \infty} x_i = 0\}$$

Note that $\{e_i : i \in \mathbb{N}\}$ as above is a linearly independent set, but *not* a basis for c_0 (give an example of an element in c_0 that cannot be expressed as a linear combination of the $\{e_i\}$).

(iv) Let $a, b \in \mathbb{R}$ with $a < b$, and define

$$C[a, b] := \{f : [a, b] \to \mathbb{K} \text{ continuous}\}.$$

This is a vector space over $\mathbb{K}$ under pointwise addition and scalar multiplication. For $n \geq 0$, let $e_n(x) := x^n$, then $\{e_n : n \geq 0\}$ is a linearly independent set, but it is not a basis for $C[a, b]$ (once again, do verify this).

More generally, if X is a compact, Hausdorff space, we set $C(X)$ to denote the space of continuous, $\mathbb{K}$-valued functions on X. This is a vector space under pointwise operations as well.

For the most part, all examples will fall into three 'types': Finite dimensional vector spaces, sequence spaces, and function spaces.

Theorem 1.5. *If $\mathbf{E}$ is a vector space, then any two Hamel bases of $\mathbf{E}$ have the same cardinality. This common number is called the dimension of $\mathbf{E}$.*

Definition 1.6. Let $\mathbf{E}$ and $\mathbf{F}$ be two vector spaces.

(i) A function $T : \mathbf{E} \to \mathbf{F}$ is said to be a linear transformation or an operator if

$$T(\alpha x + y) = \alpha T(x) + T(y)$$

for all $x, y \in \mathbf{E}$ and $\alpha \in \mathbb{K}$.

(ii) We write $L(\mathbf{E}, \mathbf{F})$ for the set of all linear operators from $\mathbf{E}$ to $\mathbf{F}$. If $S, T \in L(\mathbf{E}, \mathbf{F})$ and $\alpha \in \mathbb{K}$, we define the operators $(S + T)$ and αS by

$$(S + T)(x) := S(x) + T(x), \text{ and } (\alpha S)(x) = \alpha S(x).$$

Clearly, this makes $L(\mathbf{E}, \mathbf{F})$ a $\mathbb{K}$-vector space.

(iii) If $\mathbf{F} = \mathbb{K}$, then a linear transformation $T : \mathbf{E} \to \mathbb{K}$ is called a linear functional.

(iv) Given a linear transformation $T : \mathbf{E} \to \mathbf{F}$, define

$$\ker(T) := \{x \in \mathbf{E} : T(x) = 0\}, \text{ and } \operatorname{Range}(T) := \{T(x) : x \in \mathbf{E}\}.$$

Then, $\ker(T)$ and $\operatorname{Range}(T)$ are subspaces of $\mathbf{E}$ and $\mathbf{F}$ respectively.

(v) A linear transformation $T : \mathbf{E} \to \mathbf{F}$ is said to be an isomorphism if T is bijective. If such a map exists, we write $\mathbf{E} \cong \mathbf{F}$.

(End of Day 1)

Example 1.7.

(i) Let $\mathbf{E} = \mathbb{K}^n, \mathbf{F} = \mathbb{K}^m$, then any $m \times n$ matrix A with entries in $\mathbb{K}$ defines a linear transformation $T_A : \mathbf{E} \to \mathbf{F}$ given by $x \mapsto A(x)$. Conversely, if $T \in L(\mathbf{E}, \mathbf{F})$, then the matrix whose columns are $\{T(e_i) : 1 \leq i \leq n\}$ defines an $m \times n$ matrix A such that $T = T_A$. If $M_{m\times n}(\mathbb{K})$ denotes the vector space of all such matrices, then there is an isomorphism of vector spaces

$$L(\mathbf{E}, \mathbf{F}) \cong M_{m\times n}(\mathbb{K})$$

given by $T_A \mapsto A$. If we replace the standard basis $\{e_1, e_2, \ldots, e_n\}$ by another basis Λ of $\mathbf{E}$, we get another isomorphism from $L(\mathbf{E}, \mathbf{F}) \to M_{m\times n}(\mathbb{K})$. Thus, the isomorphism is not canonical (it depends on the choice of basis).

(ii) Let $\mathbf{E} = c_{00}$ and define $\varphi : \mathbf{E} \to \mathbb{K}$ by

$$\varphi((x_j)) := \sum_{n=1}^{\infty} x_n.$$

Note that φ is well-defined and linear. Thus, $\varphi \in L(c_{00}, \mathbb{K})$.

(iii) Let $\mathbf{E} = C[a,b]$ and define $\varphi : \mathbf{E} \to \mathbb{K}$ by

$$\varphi(f) := \int_a^b f(t)dt.$$

Then, $\varphi \in L(C[a,b], \mathbb{K})$.

(iv) Let $\mathbf{E} = \mathbf{F} = C[0,1]$. Define $T : \mathbf{E} \to \mathbf{F}$ by

$$T(f)(x) := \int_0^x f(t)dt.$$

Note that T is well-defined (from Calculus) and linear. Thus, $T \in L(\mathbf{E}, \mathbf{F})$.

Definition 1.8. Let $\mathbf{E}$ be a vector space, and $\mathbf{F}$ be a subspace of $\mathbf{E}$.

(i) The quotient space, denoted by $\mathbf{E}/\mathbf{F}$, is the quotient group, viewing $\mathbf{E}$ as an abelian group under addition, and $\mathbf{F}$ as a (normal) subgroup. Note that $\mathbf{E}/\mathbf{F}$ has a natural vector space structure, with addition given by

$$(x + \mathbf{F}) + (y + \mathbf{F}) := (x + y) + \mathbf{F},$$

and scalar multiplication given by $\alpha(x + \mathbf{F}) := \alpha x + \mathbf{F}$ for $\alpha \in \mathbb{K}$ and $x, y \in \mathbf{E}$.

(ii) The quotient map, denoted by $\pi : \mathbf{E} \to \mathbf{E}/\mathbf{F}$, is given by $x \mapsto x + \mathbf{F}$. It is a surjective linear transformation such that $\ker(\pi) = \mathbf{F}$.

(iii) Furthermore, we define the codimension of $\mathbf{F}$ by $\text{codim}(\mathbf{F}) := \dim(\mathbf{E}/\mathbf{F})$.

(iv) If $\text{codim}(\mathbf{F}) = 1$, then we say that $\mathbf{F}$ is a hyperplane of $\mathbf{E}$.

Given a non-zero linear functional $\varphi : \mathbf{E} \to \mathbb{K}$, the subspace $\ker(\varphi)$ is a hyperplane in $\mathbf{E}$; and conversely, every hyperplane is of this form. Henceforth, we will write '$\mathbf{F} < \mathbf{E}$' to indicate that $\mathbf{F}$ is a subspace of $\mathbf{E}$.

Proposition 1.9. *Let $\mathbf{E}$ be a finite dimensional vector space and $\mathbf{F} < \mathbf{E}$. Then codim$(\mathbf{F}) = \dim(\mathbf{E}) - \dim(\mathbf{F})$*

Theorem 1.10 (First Isomorphism Theorem)**.** *Let $T : \mathbf{E} \to \mathbf{F}$ be a linear transformation. Then,*

(i) $\ker(T) < \mathbf{E}$ *and* $\text{Range}(T) < \mathbf{F}$.

(ii) Furthermore, the map $\widehat{T} : \mathbf{E}/\ker(T) \to \text{Range}(T)$ given by

$$x + \ker(T) \mapsto T(x)$$

is an isomorphism.

Theorem 1.11 (Rank-Nullity Theorem)**.** *If $T : \mathbf{E} \to \mathbf{F}$ is a linear transformation and $\mathbf{E}$ is finite dimensional, then* $\dim(\ker(T)) + \dim(\text{Range}(T)) = \dim(\mathbf{E})$.

2. Review of Measure Theory

Definition 2.1. Let X be a set. A σ-algebra on X is a collection $\mathfrak{M}$ of subsets of X satisfying the following axioms:

(a) $\emptyset \in \mathfrak{M}$.

(b) If $E \in \mathfrak{M}$, then $E^c := X \setminus E \in \mathfrak{M}$.

(c) If $\{E_1, E_2, \ldots\}$ is a sequence of sets in $\mathfrak{M}$, then $\bigcup_{n=1}^{\infty} E_n \in \mathfrak{M}$.

The pair $(X, \mathfrak{M})$ is called a measurable space, and the members of $\mathfrak{M}$ are called measurable sets.

If $\{\mathfrak{M}_\alpha : \alpha \in J\}$ is a family of σ-algebras on a set X, then the intersection $\bigcap_{\alpha \in J} \mathfrak{M}_\alpha$ is also a σ-algebra. In particular, if $\mathcal{S}$ is a collection of subsets of X, then there is a unique smallest σ-algebra on X that contains $\mathcal{S}$. This is called the σ-algebra generated by $\mathcal{S}$.

Definition 2.2. Let X be a topological space. The σ-algebra generated by the topology on X is called the Borel σ-algebra on X, and is denoted by $\mathfrak{B}_X$. The members of this σ-algebra are called Borel sets.

Important examples of Borel sets are the following: A countable union of closed sets is called an F_σ-set, and the countable intersection of open sets is called a G_δ-set.

Definition 2.3. Let $(X, \mathfrak{M})$ be a measurable space, and Y be a topological space. A function $f : X \to Y$ is said to be measurable if $f^{-1}(U) \in \mathfrak{M}$ for every open set $U \subset Y$.

For the most part, measurable functions in this book will take values in $\mathbb{K}$ ($= \mathbb{R}$ or $\mathbb{C}$), where the latter is equipped with the usual topology. When it is important to make a distinction, we will refer to such functions as *real-measurable* or *complex-measurable*, as the case may be.

Example 2.4.

(i) Given a subset $E \subset X$, the characteristic function of E is the map $\chi_E : X \to \mathbb{R}$ given by

$$\chi_E(x) = \begin{cases} 1 & : \text{ if } x \in E, \\ 0 & : \text{ otherwise.} \end{cases}$$

Clearly, χ_E is a measurable function if and only if E is a measurable set.

(ii) More generally, a linear combination of characteristic functions of measurable sets is measurable. Such a function is called a simple function. Alternatively, a simple function is a measurable function whose range is a finite set.

(iii) If X and Y are both topological spaces, and we take $\mathfrak{M} = \mathfrak{B}_X$, then any measurable function $f : X \to Y$ is said to be Borel measurable. Notice that every continuous function is Borel measurable (however, there are Borel measurable functions that are not continuous).

(End of Day 2)

Proposition 2.5. *Let $(X, \mathfrak{M})$ be a measurable space.*

(i) If $f : X \to \mathbb{K}$, and $g : X \to \mathbb{K}$ are measurable functions, and $\alpha \in \mathbb{K}$, then $\alpha f + g$ is also measurable. So is the pointwise product $fg : X \to \mathbb{K}$, which is given by $x \mapsto f(x)g(x)$.

(ii) If $u : X \to \mathbb{R}$ and $v : X \to \mathbb{R}$ are real-measurable functions, then $f := u + iv$ is complex-measurable. Conversely, if $f : X \to \mathbb{C}$ is complex-measurable, then its real and imaginary parts are real-measurable functions.

(iii) If $f, g : X \to \mathbb{R}$ are measurable, then so are $\max\{f, g\}$ and $\min\{f, g\}$, which are defined by $\max\{f, g\}(x) := \max\{f(x), g(x)\}$, and $\min\{f, g\}(x) := \min\{f(x), g(x)\}$. In particular,

$$f^+ := \max\{f, 0\}, \text{ and } f^- := -\min\{f, 0\}$$

are both measurable.

(iv) If $f : X \to \mathbb{R}$ is measurable, then so is $|f| = f^+ + f^-$.

(v) If $\{f_n\}$ are a sequence of $\mathbb{K}$-valued measurable functions, then $\limsup_{n\to\infty} f_n$ and $\liminf_{n\to\infty} f_n$ are both measurable. In particular, the pointwise limit of measurable functions (if it exists) is measurable.

Theorem 2.6. *Let $f : X \to \mathbb{R}_+$ be a non-negative measurable function. Then, there is a sequence (s_n) of simple functions such that, for each $x \in X$, $(s_n(x))$ is an increasing sequence of non-negative real numbers with $\lim_{n\to\infty} s_n(x) = f(x)$.*

Definition 2.7. Let $(X, \mathfrak{M})$ be a measurable space. A positive measure on $(X, \mathfrak{M})$ is a function $\mu : \mathfrak{M} \to [0, \infty]$ satisfying the following axioms.

(a) $\mu(\emptyset) = 0$.

(b) μ is *countably additive*: If $\{E_1, E_2, \ldots\}$ is a sequence of mutually disjoint sets in $\mathfrak{M}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n).$$

The triple $(X, \mathfrak{M}, \mu)$ is called a measure space.

Example 2.8.

(i) Let X be any set, and $x_0 \in X$ be a fixed point. Let $\mathfrak{M} := 2^X$ be the set of all subsets of X, and let $\mu : \mathfrak{M} \to \mathbb{R}$ be the function

$$\mu(E) := \begin{cases} 1 & : \text{ if } x_0 \in E, \\ 0 & : \text{ if } x_0 \notin E. \end{cases}$$

This is called the Dirac measure at x_0, and is denoted by δ_{x_0}.

(ii) Let X be any set, and $\mathfrak{M} := 2^X$ as above. Define $\mu : \mathfrak{M} \to [0,\infty]$ by

$$\mu(E) := \begin{cases} |E| & : \text{ if } E \text{ is finite,} \\ \infty & : \text{ otherwise.} \end{cases}$$

(where $|\cdot|$ denotes the cardinality function). It is clear that this is a measure on $(X,\mathfrak{M})$, and is called the counting measure.

(iii) If X is a topological space, a measure on X is called a Borel measure if its domain contains $\mathfrak{B}_X$. Note that the domain of the measure may be larger than $\mathfrak{B}_X$ as well.

(iv) A measure μ on a measurable space $(X,\mathfrak{M})$ is said to be a finite measure if $\mu(X) < \infty$, and it is said to be σ-finite if X can be expressed as a countable union of sets of finite measure.

Example 2.9.

(i) Consider $\mathbb{R}$, equipped with the usual topology. Then, there is a σ-algebra $\mathfrak{L}$, which contains $\mathfrak{B}_{\mathbb{R}}$, and a positive measure $m : \mathfrak{L} \to [0,\infty]$ satisfying the following properties:

(a) If $a, b \in \mathbb{R}$ with $a < b$, then $m([a,b)) = (b-a)$.

(b) If $E \in \mathfrak{L}$ and $x \in \mathbb{R}$, then $E + x \in \mathfrak{L}$ and $m(E+x) = m(E)$. This property is called translation invariance of the measure m (here, $E + x$ is the set $\{y + x : y \in E\}$).

(c) If $E \in \mathfrak{L}$, then

$$m(E) = \inf\{m(U) : U \text{ open}, E \subset U\} = \sup\{m(K) : K \text{ compact}, K \subset E\}.$$

This property is called regularity of the measure m.

(d) If $E \in \mathfrak{L}$ is such that $m(E) = 0$, and $F \subset E$, then $F \in \mathfrak{L}$ (and hence, $m(F) = 0$). This property is called completeness of the measure m.

The members of $\mathfrak{L}$ are called Lebesgue measurable sets. By construction, every Borel set is Lebesgue measurable. There do exist subsets of $\mathbb{R}$ which are not Lebesgue measurable .

(ii) We may do the same for $\mathbb{R}^n$ when $n \geq 2$. There is a σ-algebra $\mathfrak{L}_n$ on $\mathbb{R}^n$, which contains the Borel σ-algebra $\mathfrak{B}_{\mathbb{R}^n}$, and a measure $m = m_n$ on $\mathfrak{L}_n$ with the property that

$$m\left(\prod_{i=1}^n [a_i, b_i)\right) = \prod_{i=1}^n (b_i - a_i),$$

and satisfying properties (b)-(d) exactly as above. Since any such rectangle has finite measure, it follows that m is a σ-finite measure on $\mathbb{R}^n$.

(iii) Finally, if $X \subset \mathbb{R}$ is a measurable set, then we may define a σ-algebra on E by

$$\mathfrak{L}_X := \{E \cap X : E \in \mathfrak{L}\}.$$

Then, $\mathfrak{L}_X \subset \mathfrak{L}$, and therefore, we may restrict the Lebesgue measure to $\mathfrak{L}_X$ to obtain a measure on X. We will simply refer to this as the Lebesgue measure on X. We will almost entirely focus on the case when $X = [a, b]$ is a compact interval in $\mathbb{R}$, so as to have a finite measure to work with.

Given a measure space $(X, \mathfrak{M}, \mu)$ and a measurable function $f : X \to \mathbb{K}$, we would like to make sense of the symbol

$$\int_X f d\mu = \int_X f(x) d\mu(x)$$

Now, if $f = \chi_E$ is a characteristic function, then it makes sense to define

$$\int_X \chi_E d\mu := \mu(E)$$

More generally, if $s = \sum_{i=1}^n \alpha_i \chi_{E_i}$ is a non-negative simple function, and the sets $\{E_1, E_2, \ldots, E_n\}$ are mutually disjoint, then we may define

$$\int_X s d\mu := \sum_{i=1}^n \alpha_i \mu(E_i).$$

(we require s to be non-negative, because μ is allowed to take the value ∞, and we want to avoid potential landmines such as '$\infty - \infty$'). If $f : X \to \mathbb{R}_+$ is a non-negative measurable function, we define

$$\int_X f d\mu := \sup \left\{ \int_X s d\mu : s \text{ is simple, and } 0 \leq s \leq f \right\}.$$

This definition has the important property that it is monotone: if $0 \leq g \leq f$ are both measurable functions, then $\int_X g d\mu \leq \int_X f d\mu$. Suppose that $f : X \to \mathbb{R}$ is real-measurable, so that we may define $f^+ := \max\{f, 0\}$ and $f^- := -\min\{f, 0\}$ as we did above. Then, $|f| = f^+ + f^-$ is a non-negative function, which allows us to make the following definition.

Definition 2.10. A function $f : X \to \mathbb{R}$ is said to be integrable if

$$\int_X |f| d\mu < \infty.$$

By the monotonicity of the integral, it follows that if f is integrable, then $\int_X f^+ d\mu < \infty$ and $\int_X f^- d\mu < \infty$. Therefore, we may *define*

$$\int_X f d\mu := \int_X f^+ d\mu - \int_X f^- d\mu.$$

Finally, if f is a complex-valued measurable function, then we may define integrability exactly as in Definition 2.10. Furthermore, if we write $f = u + iv$, where u and v are real-measurable functions, then integrability of f implies the integrability of u and v. Hence, we may define

$$\int_X f d\mu := \int_X u d\mu + i \int_X v d\mu = \int_X u^+ d\mu - \int_X u^- d\mu + i \int_X v^+ d\mu - i \int_X v^- d\mu.$$

Note: If $f : X \to \mathbb{C}$ and $g : X \to \mathbb{C}$ are integrable functions, and $\alpha \in \mathbb{C}$ is a scalar, then $(\alpha f + g)$ is integrable, and

$$\int_X (\alpha f + g) d\mu = \alpha \int_X f d\mu + \int_X g d\mu.$$

(End of Day 3)

Example 2.11.

(i) Let X be any set, $x_0 \in X$, and $\mu = \delta_{x_0}$ be the Dirac measure at x_0 as in Example 2.8. Then, any function $f : X \to \mathbb{C}$ is measurable, integrable, and

$$\int_X f d\delta_{x_0} = f(x_0).$$

(ii) If $X = \mathbb{N}$ and μ denotes the counting measure as in Example 2.8, then a function $f : X \to \mathbb{C}$ corresponds to a sequence $(f(n))_{n=1}^{\infty}$ of complex numbers. Also, f is integrable if and only if the corresponding series is absolutely convergent, and, in that case,

$$\int_{\mathbb{N}} f d\mu = \sum_{n=1}^{\infty} f(n).$$

(iii) If $f : [a,b] \to \mathbb{K}$ is a bounded, Riemann integrable function, then f is Lebesgue measurable, and its Lebesgue integral coincides with its Riemann integral (see [2, Section 4.2]). Therefore, whenever $f : [a,b] \to \mathbb{K}$ is a measurable function, we have the liberty to write

$$\int_a^b f = \int_a^b f(t) dt := \int_{[a,b]} f dm.$$

Theorem 2.12 (Fatou's Lemma)**.** *Let $(X, \mathfrak{M}, \mu)$ be a measure space, and (f_n) be a sequence of non-negative measurable functions. Then,*

$$\int_X \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu.$$

Theorem 2.13 (Monotone Convergence Theorem)**.** *Let $(X, \mathfrak{M}, \mu)$ be a measure space, and (f_n) be a sequence of non-negative measurable functions such that*

(i) $0 \leq f_1(x) \leq f_2(x) \leq \ldots$ *for all* $x \in X$.

(ii) $\lim_{n\to\infty} f_n(x) = f(x)$ *for all* $x \in X$.

Then, f is measurable and $\int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu$.

An immediate consequence of the Monotone Convergence Theorem is the following result.

Proposition 2.14. *Let $(X, \mathfrak{M}, \mu)$ be a measure space, and (f_n) be a sequence of non-negative measurable functions. Then, $\sum_{n=1}^{\infty} f_n$ is measurable, and*

$$\int_X \left(\sum_{n=1}^{\infty} f_n \right) d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu$$

Theorem 2.15 (Dominated Convergence Theorem)**.** *Let $(X, \mathfrak{M}, \mu)$ be a measure space, and (f_n) be a sequence of $\mathbb{K}$-valued measurable functions. Suppose that*

(i) $\lim_{n\to\infty} f_n(x) = f(x)$ *for all $x \in X$, and*

(ii) there is an integrable functions $g : X \to [0, \infty]$ such that $|f_n(x)| \leq g(x)$ for all $x \in X$ and $n \in \mathbb{N}$.

Then, f is integrable and $\int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu$.

Now, a property is said to hold almost everywhere (in symbols, we write 'a.e.' or 'a.e. $[\mu]$') if the set on which it fails to hold is contained in a set of measure zero. For instance, we would write "$f = \lim_{n\to\infty} f_n$ a.e." if the set $\{x \in X : \lim_{n\to\infty} f_n(x) \neq f(x)\}$ is contained in a set of measure zero. It is a fact that the convergence theorems mentioned above hold if we assume that the sequence converges almost everywhere (in other words, it need not converge *everywhere* for the conclusions to hold).

Let $(X, \mathfrak{M}, \mu)$ and $(Y, \mathfrak{N}, \nu)$ be two measure spaces. We wish to construct a σ-algebra and a measure on $X \times Y$. A measurable rectangle is a set of the form $E \times F$, where $E \in \mathfrak{M}$ and $F \in \mathfrak{N}$. The σ-algebra on $X \times Y$ generated by all such measurable rectangles is called the product σ-algebra, and is denoted by $\mathfrak{M} \otimes \mathfrak{N}$. Now, there is a measure λ on $\mathfrak{M} \otimes \mathfrak{N}$ such that

$$\lambda(E \times F) = \mu(E)\nu(F)$$

for any $E \in \mathfrak{M}$ and $F \in \mathfrak{N}$. Furthermore, if both μ and ν are σ-finite measures, then there is exactly one measure on $\mathfrak{M} \otimes \mathfrak{N}$ satisfying this property. This is called the product measure on $X \times Y$, and is denoted by $\mu \times \nu$. The Fubini-Tonelli Theorem now tells us how one may integrate functions on $X \times Y$ with respect to this product measure.

Theorem 2.16 (Fubini-Tonelli Theorem)**.** *Let $(X, \mathfrak{M}, \mu)$ and $(Y, \mathfrak{N}, \nu)$ be two σ-finite measure spaces, and let $f : X \times Y \to \mathbb{C}$ be a measurable function.*

(i) (Tonelli, 1909) If f is a non-negative function, then the functions

$$g(x) := \int_Y f(x, y) d\nu(y), \text{ and } h(y) := \int_X f(x, y) d\mu(x)$$

are both non-negative measurable functions on $(X, \mathfrak{M})$ *and* $(Y, \mathfrak{N})$ *respectively, and*

$$\int_{X \times Y} f d(\mu \times \nu) = \int_X \left(\int_Y f(x,y) d\nu(y) \right) d\mu(x) \\ = \int_Y \left(\int_X f(x,y) d\mu(x) \right) d\nu(y). \tag{I.1}$$

(ii) (Fubini, 1907) If f *is integrable over* $X \times Y$*, then the functions* g *and* h *as above are defined almost everywhere, both* g *and* h *are integrable over* X *and* Y *respectively, and Equation I.1 holds.*

The most important example of this phenomenon is, once again, the Lebesgue measure. If $\mathfrak{B}_{\mathbb{R}^n}$ denotes the Borel σ-algebra on $\mathbb{R}^n$, then $\mathfrak{B}_{\mathbb{R}^n} \otimes \mathfrak{B}_{\mathbb{R}^k} = \mathfrak{B}_{\mathbb{R}^{n+k}}$. If we restrict the Lebesgue measure m_n to $\mathfrak{B}_{\mathbb{R}^n}$, the product measure $m_n \times m_k$ is precisely the restriction of m_{n+k} to $\mathfrak{B}_{\mathbb{R}^{n+k}}$. The measure m_{n+k} is thus an extension of the product measure $m_n \times m_k$.

II. Normed Linear Spaces

1. Definitions and Examples

Definition 1.1. A norm on a $\mathbb{K}$-vector space $\mathbf{E}$ is a function

$$\|\cdot\| : \mathbf{E} \to \mathbb{R}_+$$

which satisfies the following properties for all $x, y \in \mathbf{E}$, and all $\alpha \in \mathbb{K}$.

(a) $\|x\| \geq 0$, and $\|x\| = 0$ if and only if $x = 0$.

(b) $\|\alpha x\| = |\alpha| \|x\|$.

(c) (Triangle inequality) $\|x + y\| \leq \|x\| + \|y\|$.

The pair $(\mathbf{E}, \|\cdot\|)$ is called a normed linear space.

Remark 1.2. Let $(\mathbf{E}, \|\cdot\|)$ be a normed linear space.

(i) The function $d(x, y) := \|x - y\|$ defines a metric on $\mathbf{E}$, called the metric induced by the norm. This makes $\mathbf{E}$ a topological space.

(ii) A sequence $(x_n) \subset \mathbf{E}$ converges to $x \in \mathbf{E}$ if and only if $\lim_{n\to\infty} \|x_n - x\| = 0$. When this happens, we will write $x_n \to x$.

(iii) By the triangle inequality, vector space addition is a continuous map from $\mathbf{E} \times \mathbf{E} \to \mathbf{E}$. Therefore, if $x_n \to x$ and $y_n \to y$, then $(x_n + y_n) \to (x + y)$. Similarly, scalar multiplication is also a continuous map from $\mathbb{K} \times \mathbf{E} \to \mathbf{E}$. (Here, $\mathbf{E} \times \mathbf{E}$ and $\mathbb{K} \times \mathbf{E}$ may both be equipped with a product metric).

(iv) For any $x, y \in \mathbf{E}$, the triangle inequality implies that

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

Thus, the norm function $\mathbf{E} \to \mathbb{R}_+$ is continuous. Hence, if $x_n \to x$ in $\mathbf{E}$ then $\|x_n\| \to \|x\|$ in $\mathbb{R}$.

(End of Day 4)

Now, let us revisit the examples from Example 1.4, and equip those spaces with norms.

Example 1.3.

(i) $\mathbb{K}(= \mathbb{R}$ or $\mathbb{C})$ is a normed linear space with the absolute value norm.

(ii) $\mathbb{K}^n$ may be equipped with many different norms. We give two such, and we will give more later on.

   (i) The 1-norm is given by $\|(x_1, x_2, \dots, x_n)\|_1 := \sum_{i=1}^n |x_i|$.

   (ii) The supremum norm given by $\|(x_1, x_2, \dots, x_n)\|_\infty := \sup_{1 \leq i \leq n} |x_i|$.

(iii) c_{00} is a normed linear space with a variety of norms. In fact, the 1-norm and supremum norm may be defined exactly as above (except that we need to take an infinite sum, in principle).

(iv) c_0 (the space of sequences 'vanishing at infinity') is a normed linear space with the supremum norm. Note that the 1-norm no longer makes sense on c_0.

(v) $C[a, b]$ may also be equipped with many norms. The definitions of the 1-norm and supremum norm as similar to the case of $\mathbb{K}^n$ and c_{00} above.

   (i) The 1-norm is given by

$$\|f\|_1 := \int_a^b |f(t)| dt.$$

   Note that it is a norm because if $\|f\|_1 = 0$ and f is continuous, then $f \equiv 0$ (This is no longer true if we replace $C[a, b]$ by the larger class of Riemann-integrable functions.).

   (ii) The supremum norm given by

$$\|f\|_\infty := \sup_{x \in [a,b]} |f(x)|.$$

Definition 1.4. An inner product on a vector space $\mathbf{E}$ is a function

$$\langle \cdot, \cdot \rangle : \mathbf{E} \times \mathbf{E} \to \mathbb{K}$$

satisfying the following properties for all $x, y, z \in \mathbf{E}$, and $\alpha, \beta \in \mathbb{K}$.

(a) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$,

(b) $\langle x, y \rangle = \overline{\langle y, x \rangle}$,

(c) $\langle x, x \rangle \geq 0$, and $\langle x, x \rangle = 0$ if and only if $x = 0$.

The pair $(E, \langle \cdot, \cdot \rangle)$ is called an inner product space

Lemma 1.5 (Cauchy-Schwarz Inequality (Cauchy, 1821, Bunyakovsky, 1859, and Schwarz, 1888))**.** *If* $\mathbf{E}$ *is an inner product space, and* $x, y \in \mathbf{E}$*, then*

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle.$$

Moreover, equality holds if and only if the set $\{x, y\}$ *is linearly dependent.*

Proof. The inequality is clearly true if $x = 0$, so we may assume that $x \neq 0$. Now, set $z := y - \frac{\langle y,x\rangle}{\langle x,x\rangle}x$, and compute

$$\begin{aligned}\langle z, y\rangle &= \left\langle y - \frac{\langle y,x\rangle}{\langle x,x\rangle}x, y\right\rangle \\ &= \langle y,y\rangle - \frac{\langle y,x\rangle\langle x,y\rangle}{\langle x,x\rangle} \\ &= \langle y,y\rangle - \frac{|\langle x,y\rangle|^2}{\langle x,x\rangle}.\end{aligned}$$

Since $\langle z, x\rangle = 0$, we see that $\langle z, y\rangle = \langle z, z\rangle \geq 0$. This gives us the required inequality. The second statement concerning equality is left as an exercise. □

Corollary 1.6. *If* **E** *is an inner product space, then the function*

$$\|x\| := \sqrt{\langle x,x\rangle}$$

defines a norm on **E**. *This is called the norm* <u>*induced by the inner product*</u>.

Proof. We only check the triangle inequality, since the other axioms are obvious. For $x, y \in \mathbf{E}$, consider

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2 \\ &= \|x\|^2 + 2\operatorname{Re}(\langle x,y\rangle) + \|y\|^2 \\ &\leq \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2.\end{aligned}$$

Note that the penultimate step follows from the Cauchy-Schwarz Inequality. Taking square roots now gives us the triangle inequality. □

One important consequence of the Cauchy-Schwarz Inequality is the fact that the inner product is a continuous map from $\mathbf{E} \times \mathbf{E}$ to $\mathbb{K}$, when **E** is equipped with this norm. In other words, if $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$ in $\mathbb{K}$.

Let us now look at the most basic examples of inner product spaces.

Example 1.7.

(i) $\mathbb{K}^n$ with the Euclidean inner product given by

$$\langle (x_j), (y_j)\rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

The induced norm is denoted by $\|\cdot\|_2$.

(ii) c_{00} with the Euclidean inner product given by

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

Note that this is a finite sum for any two vectors in c_{00}, and is thus well-defined.

(iii)

$$\ell^2 := \left\{ (x_n)_{n=1}^{\infty} : x_i \in \mathbb{K} \text{ for all } i \geq 1, \text{ and } \sum_{i=1}^{\infty} |x_i|^2 < \infty \right\}.$$

This is the space of *square-summable* sequences, and is equipped with the Euclidean inner product

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

It is not obvious that this series converges, so let's prove that. Fix $(x_n), (y_n) \in \ell^2$, and define $s_m := \sum_{i=1}^{m} x_i \overline{y_i}$. Fix integers $n > m$, and observe that

$$|s_n - s_m| = \left| \sum_{i=m+1}^{n} x_i \overline{y_i} \right| \leq \left(\sum_{i=m+1}^{n} |x_i|^2 \right)^{1/2} \left(\sum_{i=m+1}^{n} |y_i|^2 \right)^{1/2}.$$

by the Cauchy-Schwarz Inequality in $\mathbb{K}^n$. For a fixed $\epsilon > 0$, we may choose $M \in \mathbb{N}$ such that $\sum_{i=M+1}^{\infty} |x_i|^2 < \epsilon$ and $\sum_{i=M+1}^{\infty} |y_i|^2 < \epsilon$. If $n > m \geq M$, it follows that $|s_n - s_m| < \epsilon$. Thus, the sequence (s_n) is a Cauchy sequence in $\mathbb{K}$, and the series $\sum_{i=1}^{\infty} x_i \overline{y_i}$ converges. Once the inner product is well-defined, the other axioms are trivial to check.

Definition 1.8. Fix $0 < p < \infty$.

(i) A <u>$p-$integrable</u> function $f : [a, b] \to \mathbb{K}$ is a Lebesgue measurable function such that

$$\int_a^b |f(t)|^p dt < \infty.$$

(ii) Let $\mathcal{L}^p[a, b]$ be the set of all p-integrable measurable functions. Observe that if $f, g \in \mathcal{L}^p[a, b]$, then

$$|f + g|^p \leq [2 \max\{|f|, |g|\}]^p \leq 2^p [|f|^p + |g|^p]. \tag{II.1}$$

Hence, $f + g \in \mathcal{L}^p[a, b]$, making $\mathcal{L}^p[a, b]$ a vector space.

(iii) Define $\mu_p : \mathcal{L}^p[a, b] \to \mathbb{R}_+$ by

$$\mu_p(f) := \left(\int_a^b |f|^p(t) dt \right)^{1/p}$$

Then, μ_p enjoys the following properties:

(i) $\mu_p(f) \geq 0$ for all $f \in \mathcal{L}^p[a,b]$.

(ii) $\mu_p(\alpha f) = |\alpha|\mu_p(f)$ for all $\alpha \in \mathbb{K}$.

(iii) Note that $\mu_p(f) = 0$ implies that $f \equiv 0$ a.e. (not that $f = 0$).

We do not know, as yet, if μ_p satisfies the triangle inequality.

(iv) Define $\mathbf{N} := \{f \in \mathcal{L}^p[a,b] : \mu_p(f) = 0\} = \{f \in \mathcal{L}^p[a,b] : f \equiv 0 \text{ a.e.}\}$. Then, $\mathbf{N}$ is a subspace of $\mathcal{L}^p[a,b]$ by Equation II.1, so we may define

$$L^p[a,b] := \mathcal{L}^p[a,b]/\mathbf{N}.$$

Then, $L^p[a,b]$ is a vector space.

(End of Day 5)

(v) For $f + \mathbf{N} \in L^p[a,b]$, we write

$$\|f + \mathbf{N}\|_p := \mu_p(f).$$

Note that if $f + \mathbf{N} = g + \mathbf{N}$, then $f \equiv g$ a.e., and so $\mu_p(f) = \mu_p(g)$. Hence, $\|\cdot\|_p : L^p[a,b] \to \mathbb{R}_+$ is well-defined, and clearly satisfies the first two axioms of a norm.

Henceforth, we identify two functions that are equal a.e., and merely write $\|f\|_p$ for $\|f + \mathbf{N}\|_p$.

Example 1.9. For $0 < p < 1$, the triangle inequality fails. Take $f = \chi_{(0,1/2)}, g = \chi_{(1/2,1)} \in \mathcal{L}^p[0,1]$, then $\|f\|_p = \|g\|_p = 2^{-1/p}$, $\|f+g\|_p = 1$, and

$$\|f\|_p + \|g\|_p = 2^{-1/p} + 2^{-1/p} = 2^{1-1/p} < 1.$$

Lemma 1.10. *If $a, b \geq 0$ and $0 < \lambda < 1$, then $a^\lambda b^{1-\lambda} \leq \lambda a + (1-\lambda)b$. Moreover, equality holds if and only if $a = b$.*

Proof. If $b = 0$, there is nothing to prove, so assume $b \neq 0$ and set $t = a/b$. Then, we wish to prove that

$$t^\lambda \leq \lambda t + (1-\lambda)$$

with equality if and only if $t = 1$. The function $f : [0,\infty) \to \mathbb{R}$ given by $t \mapsto t^\lambda - \lambda t$ satisfies $f'(t) = \lambda t^{\lambda-1} - \lambda$. Since $0 < \lambda < 1$, f is increasing for $t < 1$ and decreasing for $t > 1$. Hence, the global maximum of f occurs at $t = 1$. The result now follows from the fact that $f(1) = 1 - \lambda$. □

Theorem 1.11 (Hölder's Inequality (Rogers, 1888 and Hölder, 1889))**.** *Let $1 < p < \infty$ and $q \in \mathbb{R}$ such that $1/p + 1/q = 1$. If $f, g : [a,b] \to \mathbb{K}$ are measurable functions, then*

$$\int_a^b |fg| \leq \|f\|_p \|g\|_q$$

Furthermore, equality holds if and only if there exist constants $\alpha, \beta \in \mathbb{K}$ such that $\alpha\beta \neq 0$, and $\alpha|f|^p \equiv \beta|g|^q$ a.e.

Proof. If either term on the right hand side is 0 or ∞, there is nothing to prove. Furthermore, if the inequality holds for any pair f, g, then it also holds for all pairs $\alpha f, \beta g$ for $\alpha, \beta \in \mathbb{K}$. Therefore, replacing f by $f/\|f\|_p$ and g by $g/\|g\|_q$, it suffices to assume that

$$\|f\|_p = \|g\|_q = 1$$

So fix $x \in [a, b]$ and let $a = |f(x)|^p, b = |g(x)|^q$, and $\lambda = 1/p$ in Lemma 1.10, so that

$$|f(x)g(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q}.$$

Integrating both sides, we get

$$\int_a^b |fg| \leq \frac{1}{p}\int_a^b |f|^p + \frac{1}{q}\int_a^b |g|^q = \frac{1}{p} + \frac{1}{q} = 1 = \|f\|_p\|g\|_q.$$

Furthermore, equality holds if and only if $|f(x)|^p = |g(x)|^q$ a.e. □

Given $1 \leq p \leq \infty$, the real number q satisfying the relation $\frac{1}{p} + \frac{1}{q} = 1$ is called the conjugate exponent of p. Note that if $p = 1$, then $q = \infty$, and vice-versa.

Theorem 1.12 (Minkowski's Inequality). *If $1 \leq p < \infty$ and $f, g \in L^p[a, b]$, then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_q.$$

Thus, $(L^p[a, b], \|\cdot\|_p)$ is a normed linear space.

Proof. The result is obvious if $p = 1$ or $f + g = 0$ a.e. Otherwise,

$$|f + g|^p \leq (|f| + |g|)|f + g|^{p-1}.$$

Let q be the conjugate exponent of p. By Hölder's Inequality,

$$\begin{aligned}\int_a^b |f + g|^p &\leq \|f\|_p \||f + g|^{p-1}\|_q + \|g\|_p \||f + g|^{p-1}\|_q \\ &= [\|f\|_p + \|g\|_p]\left[\int_a^b |f + g|^{(p-1)q}\right]^{1/q}.\end{aligned}$$

Now $(p - 1)q = p$ and Equation II.1 tells us that $\int_a^b |f + g|^p < \infty$. Thus, we may divide by $\left[\int_a^b |f + g|^p\right]^{1/q}$ on both sides to obtain

$$\|f + g\|_p = \left[\int_a^b |f + g|^p\right]^{1-1/q} \leq \|f\|_p + \|g\|_p.$$

□

Definition 1.13. In what follows, we will write 'm' to denote the Lebesgue measure.

(i) A function $f : [a,b] \to \mathbb{K}$ is said to be essentially bounded if there exists $M \in \mathbb{R}$ such that
$$m(\{x \in [a,b] : |f(x)| > M\}) = 0.$$
A number M satisfying this condition is called an essential bound of f.

(ii) Let $\mathcal{L}^\infty[a,b]$ be the set of all essentially bounded measurable functions. If $f,g \in \mathcal{L}^\infty[a,b]$, then $f+g \in \mathcal{L}^\infty[a,b]$ (this is because the union of two sets of measure zero has measure zero). Therefore, $\mathcal{L}^\infty[a,b]$ is a vector space.

(iii) For $f \in \mathcal{L}^\infty[a,b]$, define
$$\mu_\infty(f) := \inf\{M > 0 : M \text{ is an essential bound for } f\}.$$
Predictably, the quantity $\mu_\infty(f)$ is called the essential supremum of f.

Lemma 1.14. *If $f \in \mathcal{L}^\infty[a,b]$, then $|f| \leq \mu_\infty(f)$ a.e.*

Proof. We wish to prove that $\mu_\infty(f)$ is an essential bound for f. If $n \in \mathbb{N}$, then $\mu_\infty(f) + 1/n$ is not a lower bound for the set $A_f := \{M > 0 : M \text{ is an essential bound for } f\}$. Hence, there exists $M_n \in A_f$ such that $M_n \leq \mu_\infty(f) + 1/n$. Now,
$$\begin{aligned}\{x \in [a,b] : |f(x)| > \mu_\infty(f)\} &= \bigcup_{n=1}^\infty \{x \in [a,b] : |f(x)| > \mu_\infty(f) + 1/n\} \\ &\subset \bigcup_{n=1}^\infty \{x \in [a,b] : |f(x)| > M_n\}.\end{aligned}$$
But each set $\{x \in [a,b] : |f(x)| > M_n\}$ has measure zero, and thus $m(\{x \in [a,b] : |f(x)| > \mu_\infty(f)\}) = 0$. Therefore, $|f| \leq \mu_\infty(f)$ a.e. □

Definition 1.15. Consider $\mathbf{N} := \{f \in \mathcal{L}^\infty[a,b] : \mu_\infty(f) = 0\}$. Then, $f \in \mathbf{N}$ if and only if $f = 0$ a.e. Therefore, $\mathbf{N}$ is a subspace of $\mathcal{L}^\infty[a,b]$. We define
$$L^\infty[a,b] := \mathcal{L}^\infty[a,b]/\mathbf{N}$$
and for any $f + \mathbf{N} \in L^\infty[a,b]$, we write
$$\|f + \mathbf{N}\|_\infty := \mu_\infty(f).$$
Then, $\|\cdot\|_\infty$ is well-defined and a norm on $L^\infty[a,b]$

(End of Day 6)

Definition 1.16.

(i) If $X = \{1, 2, \ldots, n\}$ is a finite set equipped with the counting measure, then a function $f : X \to \mathbb{K}$ is determined by a tuple $(f(1), f(2), \ldots, f(n))$ of scalars. Therefore, we may identify
$$\mathbb{K}^n = L^p(X, \mu),$$
equipped with the norm
$$\|(x_i)\|_p := \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} & : \text{if } 1 \leq p < \infty, \\ \max_{1 \leq i \leq n} |x_i| & : \text{if } p = \infty \end{cases}$$

(ii) If $X = \mathbb{N}$, equipped with the counting measure, then a function $f : X \to \mathbb{K}$ is determined by a sequence $(f(n))_{n=1}^{\infty}$. Therefore, elements in $L^p(X, \mu)$ are thought of as sequences. For $1 \leq p < \infty$, we define the *little ℓ^p* spaces by

$$\ell^p := L^p(X, \mu) = \left\{ (x_n)_{n=1}^{\infty} : x_i \in \mathbb{K} \text{ for all } i \in \mathbb{N}, \text{ and } \sum_{i=1}^{\infty} |x_i|^p < \infty \right\}.$$

equipped with the norm given by

$$\|(x_n)\|_p := \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p}.$$

For $p = \infty$, we define

$$\ell^\infty := L^\infty(X, \mu) = \{(x_n)_{n=1}^{\infty} : x_i \in \mathbb{K} \text{ for all } i \in \mathbb{N}, \text{ and } (x_n) \text{ is bounded}\}.$$

equipped with the supremum norm $\|(x_n)\|_\infty := \sup_{n \in \mathbb{N}} |x_n|$.

2. Bounded Linear Operators

Let $\mathbf{E}$ be a normed linear space. For $x \in \mathbf{E}$ and $r > 0$, we write $B(x, r) := \{y \in \mathbf{E} : \|y - x\| < r\}$, and $B[x, r] := \{y \in \mathbf{E} : \|y - x\| \leq r\}$. Note that $B(x, r)$ is open and $B[x, r]$ is closed. When it is necessary to do so, we will write $B_{\mathbf{E}}(x, r)$ instead of $B(x, r)$ to emphasize that this is a subset of $\mathbf{E}$. The closed unit ball is the set $B[0, 1]$ (which we will also denote by $B_{\mathbf{E}}$) and the open unit ball is $B(0, 1)$. The unit sphere is the set $S_{\mathbf{E}} = \{x \in \mathbf{E} : \|x\| = 1\}$.

Definition 2.1. Let $\mathbf{E}$ and $\mathbf{F}$ be normed linear spaces. A linear operator $T : \mathbf{E} \to \mathbf{F}$ is said to be

(i) continuous if it is continuous with respect to the norm topologies on $\mathbf{E}$ and $\mathbf{F}$.

(ii) bounded if there exists $M \geq 0$ such that $\|T(x)\| \leq M\|x\|$ for all $x \in \mathbf{E}$.

Theorem 2.2. *For a linear operator $T : \mathbf{E} \to \mathbf{F}$ between normed linear spaces, the following are equivalent:*

(i) T is continuous.

(ii) T is continuous at one point of $\mathbf{E}$.

(iii) T is continuous at $0 \in \mathbf{E}$.

(iv) T is bounded.

(v) T is uniformly continuous.

Proof. Observe that (i) $\Rightarrow$ (ii) and (v) $\Rightarrow$ (i) hold by definition.

(ii) $\Rightarrow$ (iii): If T is continuous at a point $x_0 \in \mathbf{E}$, then for any $\epsilon > 0$, choose $\delta > 0$ such that

$$\|x - x_0\| < \delta \Rightarrow \|T(x) - T(x_0)\| < \epsilon.$$

So if $\|x\| < \delta$, then let $z := x + x_0$, so that $\|z - x_0\| < \delta$. Then, $\|T(z) - T(x_0)\| < \epsilon$, which implies that $\|T(x)\| < \epsilon$. Therefore, T is continuous at 0.

(iii) $\Rightarrow$ (iv): Suppose T is continuous at 0, then for $\epsilon = 1$, there exists $\delta > 0$ such that

$$\|x\| < \delta \Rightarrow \|T(x)\| < 1.$$

So for any non-zero vector $y \in \mathbf{E}$, let $x := \frac{\delta}{2}\frac{y}{\|y\|}$. Then $\|x\| < \delta$, and so $\|T(x)\| < 1$. Therefore,

$$\|T(y)\| < \frac{2}{\delta}\|y\|.$$

Since this holds for any $y \in \mathbf{E}$, T is bounded.

(iv) $\Rightarrow$ (v): Suppose that there exists $M > 0$ such that $\|T(x)\| \leq M\|x\|$ for all $x \in \mathbf{E}$, then for any $\epsilon > 0$, choose $\delta := \frac{\epsilon}{2M}$. If $\|x - y\| < \delta$, then

$$\|T(x) - T(y)\| = \|T(x - y)\| \leq M\|x - y\| \leq \frac{\epsilon}{2} < \epsilon.$$

Therefore, T is continuous on $\mathbf{E}$.

$\square$

Example 2.3.

(i) Let $\mathbf{E}$ be any inner product space, and $y \in \mathbf{E}$ be fixed. Define $\varphi : \mathbf{E} \to \mathbb{K}$ by

$$\varphi(x) := \langle x, y\rangle.$$

Then $|\varphi(x)| \leq \|x\|\|y\|$ by the Cauchy-Schwarz Inequality, and so φ is bounded.

(ii) Let $T : \mathbb{K}^n \to \mathbf{E}$ be any operator, where $\mathbb{K}^n$ is endowed with the supremum norm, and $\mathbf{E}$ is any normed linear space. Then, for any $x = (x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$, $x = \sum_{i=1}^n x_i e_i$, where $\{e_1, e_2, \ldots, e_n\}$ is the standard basis for $\mathbb{K}^n$. Then,

$$\|T(x)\| = \|\sum_{i=1}^n x_i T(e_i)\| \leq \sum_{i=1}^n |x_i|\|T(e_i)\| \leq \|x\| \left(\sum_{i=1}^n \|T(e_i)\|\right).$$

If $M := \sum_{i=1}^n \|T(e_i)\|$, then $\|T(x)\| \leq M\|x\|$. We have thus proved that any linear operator $T : \mathbb{K}^n \to \mathbf{E}$ is continuous.

(iii) Let $\mathbf{E} = c_{00}$ and $\varphi : \mathbf{E} \to \mathbb{K}$ be given by

$$\varphi((x_j)) = \sum_{n=1}^\infty x_n.$$

(i) If $\mathbf{E}$ has the 1-norm, then φ is continuous since $|\varphi(x)| \leq \|x\|_1$.

(ii) If $\mathbf{E}$ has the supremum norm, let $x^k = (1, 1, \ldots, 1, 0, 0, \ldots)$, where the 1 appears k times. Then, $\|x^k\|_\infty = 1$ for all $k \in \mathbb{N}$, but $|\varphi(x^k)| = k$. Hence, there is no $M > 0$ such that $|\varphi(x)| \leq M\|x\|_\infty$ for all $x \in \mathbf{E}$, and so φ cannot be continuous.

(iv) Let $\mathbf{E} = \mathbf{F} = (C[0,1], \|\cdot\|_\infty)$, and $T : \mathbf{E} \to \mathbf{F}$ be the operator

$$T(f)(x) = \int_0^x f(t)dt.$$

Then, for any $x \in [0,1]$,

$$|T(f)(x)| \leq \int_0^x |f(t)|dt \leq x\|f\|_\infty \leq \|f\|_\infty.$$

Hence, $\|T(f)\|_\infty \leq \|f\|_\infty$, and T is continuous.

(v) Let $\mathbf{E} = C[0,1]$, and define $\varphi : \mathbf{E} \to \mathbb{K}$ by $\varphi(f) := f(0)$.

(i) If $\mathbf{E}$ has the supremum norm, then $|\varphi(f)| \leq \|f\|_\infty$, so φ is continuous.

(ii) If $\mathbf{E}$ has the 1-norm, then consider a sequence (f_k) of non-negative continuous functions such that $f_k(0) = k$, and $\int_0^1 f_k(t)dt = 1$ (triangles of large height but area 1). Thus, $\|f_k\| = 1$ for all $k \in \mathbb{N}$, but $|\varphi(f_k)| = k$. As before, this implies that φ is not continuous.

(End of Day 7)

Definition 2.4. Let $\mathbf{E}$ and $\mathbf{F}$ be two normed linear spaces

(i) Write $\mathcal{B}(\mathbf{E}, \mathbf{F})$ for the set of all bounded linear operators from $\mathbf{E}$ to $\mathbf{F}$. Note that $\mathcal{B}(\mathbf{E}, \mathbf{F})$ is a subset of $L(\mathbf{E}, \mathbf{F})$. Furthermore, if $S, T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$, then $S + T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$ because addition is a continuous operation on $\mathbf{F}$. Similarly, $\alpha T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$ for any $\alpha \in \mathbb{K}$. Thus, $\mathcal{B}(\mathbf{E}, \mathbf{F})$ is a vector space.

(ii) Write $\mathcal{B}(\mathbf{E})$ for the space $\mathcal{B}(\mathbf{E}, \mathbf{E})$.

(iii) Write $\mathbf{E}^*$ for the space $\mathcal{B}(\mathbf{E}, \mathbb{K})$. This is called the (continuous) dual space of $\mathbf{E}$.

For any $T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$, we write

$$\nu(T) := \inf\{M > 0 : \|T(x)\| \leq M\|x\| \text{ for all } x \in \mathbf{E}\}$$

Note that the set on the right hand side is not empty, and thus $\nu(T) < \infty$.

Lemma 2.5. *For any $T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$ and any $x \in \mathbf{E}$, $\|T(x)\| \leq \nu(T)\|x\|$*

Proof. For each $n \in \mathbb{N}$, $\nu(T) + 1/n$ is not a lower bound for the set $A_T := \{M > 0 : \|T(x)\| \leq M\|x\| \text{ for all } x \in \mathbf{E}\}$. Hence, there exists $M_n \in A_T$ such that $\nu(T) \leq M_n < \nu(T) + 1/n$, and

$$\|T(x)\| \leq M_n\|x\|$$

for all $x \in \mathbf{E}$ and $n \in \mathbb{N}$. Fixing $x \in \mathbf{E}$, we let $n \to \infty$ to obtain $\|T(x)\| \leq \nu(T)\|x\|$. $\square$

Proposition 2.6. *The function $\nu : \mathcal{B}(\mathbf{E}, \mathbf{F}) \to \mathbb{R}_+$ defined above is a norm on $\mathcal{B}(\mathbf{E}, \mathbf{F})$.*

Proof.

(i) Clearly, $\nu(T) \geq 0$ and $\nu(0) = 0$. If $\nu(T) = 0$, then the fact that $T = 0$ follows from Lemma 2.5.

(ii) Fix $T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$ and $0 \neq \lambda \in \mathbb{K}$, and define $A_1 := \{M > 0 : \|T(x)\| \leq M\|x\| \text{ for all } x \in \mathbf{E}\}$ and $A_2 := \{K > 0 : \|(\lambda T)(y)\| \leq K\|y\| \text{ for all } y \in \mathbf{E}\}$. For any $M \in A_1$,

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

for all $x \in \mathbf{E}$. Therefore, $M|\lambda| \in A_2$, and $\nu(\lambda T) = \inf A_2 \leq M|\lambda|$. This is true for any $M \in A_1$, so

$$\nu(\lambda T) \leq |\lambda| \inf A_1 = |\lambda|\nu(T).$$

Replacing λ with $1/\lambda$ and T by λT, we conclude that

$$\nu(T) = \nu\left(\frac{1}{\lambda}\lambda T\right) \leq \frac{1}{|\lambda|}\nu(\lambda T).$$

Hence, $\nu(\lambda T) \geq |\lambda|\nu(T)$ as well.

(iii) Fix $S, T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$, then for any $x \in \mathbf{E}$, we have

$$\begin{aligned}\|(S+T)(x)\| &= \|S(x) + T(x)\| \leq \|S(x)\| + \|T(x)\| \\ &\leq \nu(S)\|x\| + \nu(T)\|x\| = (\nu(S) + \nu(T))\|x\|.\end{aligned}$$

By definition, this implies that $\nu(S+T) \leq \nu(S) + \nu(T)$.

□

As is customary, we write $\|T\| := \nu(T)$ from here on. An important inequality that merits attention here is that, for any $T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$ and any $x \in \mathbf{E}$, one has

$$\|T(x)\| \leq \|T\|\|x\|$$

Proposition 2.7. *If $T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$, then*

$$\begin{aligned}\|T\| &= \sup\{\|T(x)\| : x \in \mathbf{E}, \|x\| = 1\} \\ &= \sup\{\|T(x)\| : x \in \mathbf{E}, \|x\| \leq 1\}\end{aligned}$$

Proof. Let $\alpha := \sup\{\|T(x)\| : x \in \mathbf{E}, \|x\| = 1\}$ and $\beta := \sup\{\|T(x)\| : x \in \mathbf{E}, \|x\| \leq 1\}$, then clearly $\alpha \leq \beta$.

For any $x \in \mathbf{E}$ with $\|x\| \leq 1$, $\|T(x)\| \leq \|T\|\|x\| \leq \|T\|$. Hence, $\beta \leq \|T\|$. To complete the proof, it suffices to show that $\alpha \geq \|T\|$. To that end, set

$$A := \{M > 0 : \|T(x)\| \leq M\|x\| \text{ for all } x \in \mathbf{E}\}.$$

For any $n \in \mathbb{N}$, $\|T\| - 1/n \notin A$, so there exists $x_n \in \mathbf{E}$ such that

$$\|T(x_n)\| > (\|T\| - 1/n)\|x_n\|.$$

In particular, $x_n \neq 0$, so if $y_n := x_n/\|x_n\|$, then $\|y_n\| = 1$ and $\|T(y_n)\| > \|T\| - 1/n$. Thus, $\alpha > \|T\| - 1/n$ for each $n \in \mathbb{N}$. Hence, $\alpha \geq \|T\|$ as well. □

Example 2.8.

(i) Let $\mathbf{E}$ be an inner product space, $y \in \mathbf{E}$ and define $\varphi : \mathbf{E} \to \mathbb{K}$ by $x \mapsto \langle x, y \rangle$. Then by the Cauchy-Schwarz Inequality, $|\varphi(x)| \leq \|x\|\|y\|$ for all $x \in \mathbf{E}$. Hence, $\|\varphi\| \leq \|y\|$. Furthermore, if $x = y$, then $\|y\|^2 = |\varphi(y)| \leq \|\varphi\|\|y\|$, and so $\|\varphi\| = \|y\|$.

(ii) Let $\mathbf{E} = \mathbb{K}^n$ with the 1-norm and let $\{e_1, e_2, \ldots, e_n\}$ be the standard basis for $\mathbf{E}$. Let $\mathbf{F}$ be a normed linear space and $T : \mathbf{E} \to \mathbf{F}$ be a linear operator. Then for $x = \sum_{i=1}^n x_i e_i$, we have

$$\|T(x)\| \leq \sum_{i=1}^n |x_i| \|T(e_i)\|$$

and so T is continuous with $\|T\| \leq \max_{1 \leq i \leq n} \|T(e_i)\|$. If $x = e_i$, then $\|x\|_1 = 1$ and $\|T(x)\| = \|T(e_i)\|$, and so by Proposition 2.7, $\|T\| \geq \|T(e_i)\|$. This is true for all $1 \leq i \leq n$, so $\|T\| = \max_{1 \leq i \leq n} \|T(e_i)\|$.

(iii) Let $\mathbf{E} = c_{00}$ with the 1-norm and $\varphi : \mathbf{E} \to \mathbb{K}$ be given by

$$\varphi((x_j)) := \sum_{n=1}^\infty x_n,$$

then $\|\varphi\| \leq 1$. Also, for $x = e_1$, we have $\|x\| = 1$ and $|\varphi(x)| = 1$, so that $\|\varphi\| = 1$.

(iv) Define $T : L^1[0,1] \to L^1[0,1]$ by

$$T(f)(x) = \int_0^x f(t)dt.$$

Note that T is well-defined because

$$\begin{aligned} \int_0^1 |T(f)(x)|dx &= \int_0^1 \left| \int_0^x f(t)dt \right| dx \\ &\leq \int_0^1 \int_0^x |f(t)|dtdx \\ &\leq \int_0^1 \int_0^1 |f(t)|dtdx = \|f\|_1. \end{aligned} \tag{II.2}$$

This also proves that T is bounded with $\|T\| \leq 1$. To prove that $\|T\| = 1$, we set $f_n = n\chi_{[0,1/n]}$. Then, $\|f_n\|_1 = 1$, and

$$T(f_n)(x) = \int_0^x n\chi_{[0,1/n]}(t)dt = \begin{cases} 1 & : \text{if } x \geq 1/n \\ nx & : \text{if } x < 1/n. \end{cases}$$

Hence,

$$\|T(f_n)\|_1 = \int_0^{1/n} nt dt + \int_{1/n}^1 dt = n\frac{1}{2n^2} + 1 - \frac{1}{n} = 1 - \frac{1}{2n}.$$

Thus, $\|T(f_n)\|_1 \to 1$ as $n \to \infty$, and therefore $\|T\| = 1$.

(End of Day 8)

3. Banach Spaces

Definition 3.1.

(i) A normed linear space that is complete with respect to the induced metric is called a Banach Space.

(ii) An inner product space that is complete with respect to the norm induced by the inner product is called a Hilbert space.

Proposition 3.2. *For $1 \leq p \leq \infty$, $(\mathbb{K}^n, \|\cdot\|_p)$ is a Banach space.*

Proof. Assume $p = \infty$, as the case when $p < \infty$ is similar. Suppose (x^m) is a Cauchy sequence in $(\mathbb{K}^n, \|\cdot\|_p)$ with $x^m = (x_1^m, x_2^m, \ldots, x_n^m)$. Then for any $1 \leq i \leq n$, the sequence (x_i^m) is Cauchy in $\mathbb{K}$. Since $\mathbb{K}$ is complete, there exists $y_i \in \mathbb{K}$ such that $\lim_{m\to\infty} x_i^m = y_i$. Thus, for any $\epsilon > 0$, there exists $N_i \in \mathbb{N}$ such that

$$|x_i^m - y_i| < \epsilon$$

for all $m \geq N_i$. Let $N_0 = \max\{N_1, N_2, \ldots, N_n\}$. Then, for all $m \geq N_0$,

$$\|x^m - y\|_\infty = \sup_{1\leq i\leq n} |x_i^m - y_i| < \epsilon.$$

Thus, $x^m \to y$ in norm. $\square$

Proposition 3.3. *For $1 \leq p \leq \infty$, ℓ^p is a Banach space.*

Proof. We prove this if $p < \infty$ as the $p = \infty$ case is similar. Suppose (x^k) is a Cauchy sequence in ℓ^p with $x^k = (x_1^k, x_2^k, \ldots, x_n^k, \ldots)$. We prove that (x^k) converges in the following steps.

(i) For any $n \in \mathbb{N}$, $|x_n^k - x_n^m| \leq \|x^k - x^m\|_p$, so (x_n^k) is Cauchy in $\mathbb{K}$. Since $\mathbb{K}$ is complete, there exists $y_n \in \mathbb{K}$ such that $\lim_{n\to\infty} x_n^k = y_n$.

(ii) We wish to prove that $y = (y_n) \in \ell^p$. Since (x^k) is Cauchy, it is bounded, so there exists $R > 0$ such that $\|x^m\|_p \leq R$ for all $m \in \mathbb{N}$. For any fixed $j \in \mathbb{N}$, this implies

$$\left(\sum_{n=1}^{j} |x_n^m|^p\right)^{1/p} \leq R.$$

Now let $m \to \infty$ in the finite sum to conclude that $\left(\sum_{n=1}^{j} |y_n|^p\right)^{1/p} \leq R$. This is true for all $j \in \mathbb{N}$, so

$$\left(\sum_{n=1}^{\infty} |y_n|^p\right)^{1/p} \leq R.$$

Hence, $y \in \ell^p$ as required.

(iii) We wish to prove that $x^k \to y$ in $\|\cdot\|_p$. For any $\epsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that $\|x^k - x^m\|_p < \epsilon$ for all $k, m \geq N_0$. Now if $j \in \mathbb{N}$ is fixed, consider

$$\left(\sum_{n=1}^{j} |x_n^k - x_n^m|^p\right)^{1/p} \leq \|x^k - x^m\|_p < \epsilon.$$

Let $m \to \infty$ in the finite sum to obtain $\left(\sum_{n=1}^{j} |x_n^k - y_n|^p\right)^{1/p} \leq \epsilon$. This is true for all $j \in \mathbb{N}$, so

$$\left(\sum_{n=1}^{\infty} |x_n^k - y_n|^p\right)^{1/p} \leq \epsilon$$

for all $k \geq N_0$. Thus, $\|x^k - y\|_p \to 0$ as desired.

□

Proposition 3.4. *For $1 \leq p < \infty$, c_{00} is dense in ℓ^p. In particular, $(c_{00}, \|\cdot\|_p)$ is not complete.*

Proof. Fix $x = (x_n) \in \ell^p$, and $\epsilon > 0$. Then there exists $N_0 \in \mathbb{N}$ such that $\sum_{n=N_0}^{\infty} |x_n|^p < \epsilon$. Hence, $y := (x_1, x_2, \ldots, x_{N_0}, 0, 0, \ldots) \in c_{00}$ and $\|x - y\|_p^p < \epsilon$. □

Note that, in contrast to Proposition 3.4, c_{00} is *not* dense in ℓ^∞. Even so, $(c_{00}, \|\cdot\|_\infty)$ is not complete (Try to prove these statements).

Proposition 3.5. *$L^\infty[a, b]$ is a Banach space.*

Proof. Let $(f_n) \subset L^\infty[a, b]$ be a Cauchy sequence, then define

$$A_k := \{x \in [a, b] : |f_k(x)| > \|f_k\|_\infty\}, \text{ and}$$
$$B_{k,m} := \{x \in [a, b] : |f_k(x) - f_m(x)| > \|f_k - f_m\|_\infty\}$$

By Lemma 1.14, each of these sets has measure zero, so

$$C := \left(\bigcup_{k=1}^{\infty} A_k\right) \cup \left(\bigcup_{k,m=1}^{\infty} B_{k,m}\right)$$

also has measure zero. Furthermore, on $D := [a, b] \setminus C$, each f_k is bounded and uniformly Cauchy. So for any $x \in D$, the inequality

$$|f_k(x) - f_m(x)| \leq \|f_k - f_m\|_\infty$$

implies that $(f_m(x))_{m=1}^{\infty}$ is a Cauchy sequence in $\mathbb{K}$. Hence, we may define $f : D \to \mathbb{K}$ by

$$f(x) = \lim_{m\to\infty} f_m(x)$$

We may extend f to all of $[a, b]$ by defining $f \equiv 0$ on C. Furthermore, for $\epsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that $\|f_k - f_m\|_\infty < \epsilon$ for all $k, m \geq N_0$. For any fixed $x \in D$, and $k \geq N_0$, $|f_k(x) - f_m(x)| < \epsilon$, so that

$$|f_k(x) - f(x)| \leq \epsilon.$$

Hence, $f - f_k$ is bounded on D and so $f - f_k \in L^\infty[a,b]$. Therefore, $f = (f - f_k) + f_k \in L^\infty[a,b]$. Finally, the above inequality also proves that $\|f_k - f\|_\infty \leq \epsilon$ for all $k \geq N_0$. Therefore, $f_k \to f$ in $L^\infty[a,b]$. $\square$

(End of Day 9)

Notice that the essential supremum of a continuous function is the same as its supremum. Therefore, we may think of $(C[a,b], \|\cdot\|_\infty)$ as a subspace of $L^\infty[a,b]$.

Proposition 3.6. *$(C[a,b], \|\cdot\|_\infty)$ is closed in $L^\infty[a,b]$. In particular, $(C[a,b], \|\cdot\|_\infty)$ is a Banach space.*

Definition 3.7. Let $\mathbf{E}$ be a normed linear space and $(x_n) \subset \mathbf{E}$ be a sequence.

(i) We say that the series $\sum_{n=1}^\infty x_n$ is convergent if the sequence (s_n) of partial sums defined by $s_n := \sum_{k=1}^n x_k$ converges to a point in $\mathbf{E}$. In other words, there exists $s \in \mathbf{E}$ such that for any $\epsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that

$$\left\| \left(\sum_{k=1}^n x_k \right) - s \right\| < \epsilon$$

for all $n \geq N_0$.

(ii) We say that the series $\sum_{n=1}^\infty x_n$ is absolutely convergent if $\sum_{n=1}^\infty \|x_n\| < \infty$. Note that this is a series of non-negative real numbers, so to say that the series $\sum_{n=1}^\infty x_n$ is absolutely convergent is the same as saying that there is a real number $M \geq 0$ such that $\sum_{i=1}^n \|x_i\| \leq M$ for all $n \in \mathbb{N}$.

Proposition 3.8. *A normed linear space $\mathbf{E}$ is a Banach space if and only if every absolutely convergent series is convergent in $\mathbf{E}$.*

Proof. Let $\mathbf{E}$ be a Banach space and $(x_n) \subset \mathbf{E}$ such that $\sum_{n=1}^\infty \|x_n\| < \infty$. Let $s_n := \sum_{j=1}^n x_j$, then it suffices to show that (s_n) is a Cauchy sequence. If $\epsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that $\sum_{n=N_0}^\infty \|x_n\| < \epsilon$. Hence if $n, m \geq N_0$ with $n > m$, then

$$\|s_n - s_m\| = \left\| \sum_{k=m+1}^n x_k \right\| \leq \sum_{k=m+1}^n \|x_n\| \leq \sum_{k=N_0}^\infty \|x_n\| < \epsilon.$$

Thus, the series $\sum_{n=1}^\infty x_n$ is convergent.

Conversely, suppose every absolutely convergent series is convergent in $\mathbf{E}$, choose a Cauchy sequence $(x_n) \subset \mathbf{E}$. Since (x_n) is Cauchy, it suffices (Does it?) to prove that (x_n) has a convergent subsequence. Now, for each $j \in \mathbb{N}$, there exists $N_j \in \mathbb{N}$ such that $\|x_k - x_l\| < 2^{-j}$ for all $k, l \geq N_j$. By induction, we may choose $N_1 < N_2 < \ldots$, and so we obtain a subsequence (x_{N_j}) such that $\|x_{N_{j+1}} - x_{N_j}\| < 2^{-j}$ for all $j \in \mathbb{N}$. Thus,

$$\sum_{j=1}^\infty \|x_{N_{j+1}} - x_{N_j}\| < \infty.$$

By hypothesis, the series $\sum_{j=1}^{\infty}(x_{N_{j+1}} - x_{N_j})$ converges in $\mathbf{E}$. But this is a telescoping series. Consider the partial sum, and it collapses as $\sum_{j=1}^{n}(x_{N_{j+1}} - x_{N_j}) = x_{N_{n+1}} - x_{N_1}$. So if the partial sums converge, so does (x_{N_j}). $\square$

Theorem 3.9 (F. Riesz and Fischer, 1907)**.** *For $1 \le p < \infty$, $L^p[a,b]$ is a Banach space.*

Proof. If $(f_n) \in L^p[a,b]$ is such that $M := \sum_{n=1}^{\infty} \|f_n\|_p < \infty$. Define $g_k := \sum_{n=1}^{k} |f_n|$ and $g := \sum_{n=1}^{\infty} |f_n|$. By Minkowski's Inequality, $\|g_k\|_p \le M$, and so by Fatou's Lemma,

$$\int_a^b g^p \le \liminf \int_a^b g_k^p \le M^p.$$

In particular, $g(x) < \infty$ a.e. Hence, we may define

$$f := \sum_{n=1}^{\infty} f_n,$$

and this converges a.e. (we may define $f \equiv 0$ on the set of measure zero where the series does not converge). Then, f is measurable and $|f| \le g$, so $f \in L^p[a,b]$ by the above inequality. Now define $s_k := \sum_{n=1}^{k} f_n$, then $s_k \in L^p[a,b]$ and we want to prove that $\|s_k - f\|_p \to 0$. Observe that $s_k \to f$ pointwise and $|f - s_k|^p \le (2g)^p \in L^1[a,b]$. Therefore, by the Dominated Convergence Theorem,

$$\|f - s_k\|_p^p = \int_a^b |f - s_k|^p \to 0.$$

Thus, we have verified that any absolutely convergent series in $L^p[a,b]$ converges. By Proposition 3.8, $L^p[a,b]$ is complete. $\square$

Remark 3.10. Let (X,d) be a metric space.

(i) If $A \subset X$ is a set, then for any $x \in X$, define $d(x,A) := \inf\{d(x,y) : y \in A\}$. The function $x \mapsto d(x,A)$ is continuous, and $d(x,A) = 0$ if and only if $x \in \overline{A}$.

(ii) If $A, B \subset X$ are two disjoint closed sets, then define

$$f(x) := \frac{d(x,A)}{d(x,A) + d(x,B)}.$$

Then, $f : X \to [0,1]$ is continuous, $f \equiv 0$ on A, and $f \equiv 1$ on B.

Lemma 3.11. *Let $K \subset [a,b]$ be a compact set. Then, there exists $g \in C[a,b]$ such that $g \equiv 1$ on K and $g < 1$ on $[a,b] \setminus K$.*

Proof. Let d denote the usual metric on $[a,b]$. For each $n \in \mathbb{N}$, consider $G_n = \{x \in [a,b] : d(x,K) \ge 1/n\}$. Then, G_n is closed, and $G_n \cap K = \emptyset$. Hence, by Remark 3.10, there exists $g_n \in C[a,b]$ such that $0 \le g_n \le 1$, $g_n \equiv 1$ on K, and $g_n \equiv 0$ on G_n. Now, the series

$$g := \sum_{n=1}^{\infty} \frac{1}{2^n} g_n$$

converges in $C[a,b]$ (since $C[a,b]$ is a Banach space and the series is absolutely convergent). The function g satisfies the required properties. $\square$

(End of Day 10)

Proposition 3.12. *If $1 \leq p < \infty$, then $C[a,b]$ is dense in $L^p[a,b]$. In particular, $(C[a,b], \|\cdot\|_p)$ is not complete.*

Proof. Given $f \in L^p[a,b]$, and $\epsilon > 0$, we want to prove that there exists $g \in C[a,b]$ such that $\|f - g\|_p < \epsilon$.

(i) Suppose $f = \chi_K$ where $K \subset [a,b]$ is compact: Let $g \in C[a,b]$ be as in Lemma 3.11. For $n \in \mathbb{N}$, define $g^n \in C[a,b]$ by the pointwise product $g^n(x) = g(x)^n$. Then, $g^n \to f$ pointwise. Furthermore, $|g^n - f|^p \leq 2^p \in L^1[a,b]$ for all $n \in \mathbb{N}$. By the Dominated Convergence Theorem, $\|g^n - f\|_p^p = \int_a^b |g^n - f|^p \to 0$. Hence, there exists $N \in \mathbb{N}$ such that $\|g^N - f\|_p < \epsilon$.

(ii) If $f = \chi_E$ where $E \subset [a,b]$ measurable, then for $\epsilon > 0$, there exists a compact set $K \subset E$ such that $m(E \setminus K) < \epsilon$. Hence, $\|\chi_K - \chi_E\|_p^p < \epsilon$. Now apply part (i).

(iii) If $f = \sum_{i=1}^n \alpha_i \chi_{E_i} \in L^1[a,b]$ is a simple function, then apply part (ii) to each E_i and take a linear combination.

(iv) If $f \in L^p[a,b]$ is non-negative, then choose a sequence of simple functions (s_n) such that $0 \leq s_n \leq s_{n+1} \to f$ pointwise (from Theorem 2.6). Since $|s_n - f|^p \leq (2f)^p \in L^1[a,b]$, the Dominated Convergence Theorem implies that $\|s_n - f\|_p \to 0$. Now, apply part (iii) to s_N for N large enough.

(v) If $f \in L^p[a,b]$ is real-valued, then write it as $f = f^+ - f^-$ and apply part (iv) to each of f^+ and f^-.

(vi) If $f \in L^p[a,b]$ is complex-valued, then apply part (v) to the real and imaginary parts of f.

□

Theorem 3.13. *Let* $\mathbf{E}$ *be a normed linear space.*

(i) If $\mathbf{F}$ *is complete, then* $\mathcal{B}(\mathbf{E}, \mathbf{F})$ *is complete.*

(ii) In particular, $\mathbf{E}^*$ *is a Banach space.*

Proof. Suppose $(T_n) \subset \mathcal{B}(\mathbf{E}, \mathbf{F})$ is a Cauchy sequence, then for any $x \in \mathbf{E}$, the inequality $\|T_n(x) - T_m(x)\| \leq \|T_n - T_m\|\|x\|$ implies that $(T_n(x)) \subset \mathbf{F}$ is a Cauchy sequence. Since $\mathbf{F}$ is complete, this sequence converges in $\mathbf{F}$. Hence we may define $T : \mathbf{E} \to \mathbf{F}$ by

$$T(x) = \lim_{n \to \infty} T_n(x).$$

It is clear that T is linear. Furthermore, since (T_n) is Cauchy, there exists $M > 0$ such that $\|T_n\| \leq M$ for all $n \in \mathbb{N}$. Hence $\|T(x)\| \leq M\|x\|$ for all $x \in \mathbf{E}$, so $T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$. We now want to prove that $\|T_n - T\| \to 0$. To this end, choose $\epsilon > 0$ and $N_0 \in \mathbb{N}$ such that $\|T_n - T_m\| < \epsilon$ for all $n, m \geq N_0$. Then, for any $x \in \mathbf{E}$ and $n \geq N_0$ fixed,

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

Therefore, $\|T - T_n\| \leq \epsilon$ for all $n \geq N_0$. Thus, $T_n \to T$ in $\mathcal{B}(\mathbf{E}, \mathbf{F})$ as required. □

Definition 3.14. A topological space is said to be separable if it has a countable dense subset.

Remark 3.15. Let $\mathbf{E}$ be an normed linear space.

(i) If $\mathbf{E}$ has a dense, separable subspace, then $\mathbf{E}$ is separable.

(ii) If $\mathbf{E}$ contains an uncountable family of disjoint open sets, then $\mathbf{E}$ is not separable.

Example 3.16.

(i) Depending on whether $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, we write $\mathbb{K}_0 := \mathbb{Q}$ or $\mathbb{K}_0 = \mathbb{Q} \times \mathbb{Q}$. In either case, $\mathbb{K}_0$ is countable and dense in $\mathbb{K}$, which makes $\mathbb{K}$ separable. Furthermore, it allows us to prove separability for a number of other spaces.

(ii) $(\mathbb{K}^n, \|\cdot\|_p)$ is separable (for any $1 \leq p \leq \infty$) since $\mathbb{K}_0^n$ is dense in $\mathbb{K}^n$. In fact, we will soon see (Theorem 4.5) that the norm is irrelevant; $\mathbb{K}^n$ is separable with respect to any norm.

(iii) $(c_{00}, \|\cdot\|_p)$ is separable (for any $1 \leq p \leq \infty$) since we may choose sequences with entries from $\mathbb{K}_0$. This would give a subset which has the same cardinality as

$$\bigcup_{n=1}^{\infty} \mathbb{K}_0^n$$

which is countable and dense in c_{00}.

(iv) By Proposition 3.4, and Example (iii), ℓ^p is separable if $1 \leq p < \infty$.

(v) ℓ^∞ is not separable.

Proof. For each subset $A \subset \mathbb{N}$, choose $\chi_A \in \ell^\infty$. Then if $A \neq B$, then $\|\chi_A - \chi_B\|_\infty = 1$. Thus, $\{B(\chi_A; 1/3) : A \subset \mathbb{N}\}$ forms an uncountable family of disjoint open sets (because the power set of an infinite set is necessarily uncountable). □

(End of Day 11)

(vi) By the Weierstrass Approximation Theorem, polynomials with coefficients in $\mathbb{K}_0$ form a dense subset of $(C[a,b], \|\cdot\|_\infty)$. For any $1 \leq p < \infty$ and any $f \in C[a,b]$,

$$\|f\|_p \leq \|f\|_\infty (b-a)^{1/p}.$$

Therefore, this set is also dense in $C[a,b]$ with respect to $\|\cdot\|_p$. Hence, $(C[a,b], \|\cdot\|_p)$ is separable for all $1 \leq p \leq \infty$.

(vii) By Proposition 3.12 and Example (vi), $L^p[a,b]$ is separable for $1 \leq p < \infty$.

(viii) $L^\infty[a,b]$ is not separable.

Proof. For each $t \in [a,b]$, consider $f_t = \chi_{[a,t]}$, then if $s \neq t$, $\|f_s - f_t\|_\infty = 1$. Once again, we obtain an uncountable family of disjoint open sets. □

4. Finite Dimensional Spaces

Definition 4.1. Let $\mathbf{E}$ be a vector space and $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on $\mathbf{E}$. We say that these norms are equivalent (In symbols, $\|\cdot\|_1 \sim \|\cdot\|_2$) if they induce the same metric topologies on $\mathbf{E}$ (See Remark 1.2).

Note that this is an equivalence relation on the class of norms on $\mathbf{E}$.

Proposition 4.2. *Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if and only if there exist two constants $\alpha, \beta > 0$ such that*

$$\alpha\|x\|_1 \leq \|x\|_2 \leq \beta\|x\|_1 \tag{II.3}$$

for all $x \in \mathbf{E}$.

Proof. Let τ_1 and τ_2 be the topologies generated by $\|\cdot\|_1$ and $\|\cdot\|_2$ respectively.

(i) Suppose $\|\cdot\|_1 \sim \|\cdot\|_2$, then consider $B_1 := \{x \in \mathbf{E} : \|x\|_1 < 1\}$. By hypothesis, $B_1 \in \tau_2$. In particular, since $0 \in B_1$, there exists $\delta > 0$ such that

$$B_2 := \{x \in \mathbf{E} : \|x\|_2 < \delta\} \subset B_1.$$

Now, we apply the scaling trick. For any $0 \neq y \in \mathbf{E}$, consider $z := \frac{\delta}{2}\frac{y}{\|y\|_2}$. Then $\|z\|_2 < \delta$, so $\|z\|_1 < 1$, and hence

$$\frac{\delta}{2}\|y\|_1 < \|y\|_2$$

for all $y \in \mathbf{E}$. Thus $\alpha := \delta/2$ satisfies the first inequality of Equation II.3. By symmetry, there exists $\beta \geq 0$ such that $\|y\|_2 \leq \beta\|y\|_1$ for all $y \in \mathbf{E}$.

(ii) Now suppose that there exist constants $\alpha, \beta > 0$ such that Equation II.3 holds. Then choose $U \in \tau_1$. We want to prove that $U \in \tau_2$. For any $x \in U$, there exists $r > 0$ such that

$$U_1 := \{y \in \mathbf{E} : \|y - x\|_1 < r\} \subset U.$$

Let $U_2 := \{y \in \mathbf{E} : \|y - x\|_2 < \alpha r\}$, then for any $y \in U_2$, $\|y - x\|_1 \leq \frac{\|y-x\|_2}{\alpha} < r$. Therefore, $U_2 \subset U_1 \subset U$. This is true for every $x \in U$, and so $U \in \tau_2$. Hence, $\tau_1 \subset \tau_2$. By symmetry, $\tau_2 \subset \tau_1$ as well.

□

Example 4.3.

(i) Let $\mathbf{E} = \mathbb{K}^n$, and consider $\|\cdot\|_2$ and $\|\cdot\|_\infty$ on $\mathbf{E}$. Then, for any $x \in \mathbb{K}^n$,

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty,$$

and so $\|\cdot\|_2 \sim \|\cdot\|_\infty$.

(ii) Let $\mathbf{E} = c_{00}$ and consider $\|\cdot\|_1$ and $\|\cdot\|_\infty$ on $\mathbf{E}$. Then $\|x\|_\infty \leq \|x\|_1$ for all $x \in \mathbf{E}$, but if $x^k = (1, 1, \ldots, 1, 0, 0, \ldots)$, then $\|x^k\|_1 = k$ and $\|x^k\|_\infty = 1$. Hence, there is no constant $\beta > 0$ satisfying $\|x\|_1 \leq \beta\|x\|_\infty$ for all $x \in \mathbf{E}$. Thus, $\|\cdot\|_1 \nsim \|\cdot\|_\infty$. Notice that we knew this fact already, from a different point of view: In item (iii), we had constructed a linear functional $\varphi : \mathbf{E} \to \mathbb{K}$ which is continuous with respect to $\|\cdot\|_1$, but not with respect to $\|\cdot\|_\infty$. Therefore, the two topologies could not have been the same!

(iii) Suppose $\mathbf{E}$ is a vector space with two equivalent norms $\|\cdot\|_1$ and $\|\cdot\|_2$. If $\mathbf{E}$ is complete with respect to $\|\cdot\|_1$, then it is complete with respect to $\|\cdot\|_2$ (do check this fact!).

(iv) If $\mathbf{E} = C[a, b]$ then for any $1 \leq p < \infty$, $\|\cdot\|_p \nsim \|\cdot\|_\infty$ (by Proposition 3.12 and Proposition 3.6).

Lemma 4.4 (Heine-Borel Theorem (Borel, 1894))**.** *Every closed and bounded subset of* $(\mathbb{K}^n, \|\cdot\|_1)$ *is compact.*

Theorem 4.5. *Any two norms on a finite dimensional vector space are equivalent.*

Proof. Let $(\mathbf{E}, \|\cdot\|_\mathbf{E})$ be a finite dimensional normed linear space with basis $\{e_1, e_2, \ldots, e_n\}$. For any $x = \sum_{i=1}^n x_i e_i \in \mathbf{E}$, define

$$\|x\|_1 := \sum_{i=1}^n |x_i|,$$

and note that $\|\cdot\|_1$ is a norm on $\mathbf{E}$. Since the equivalence of norms is an equivalence relation, it suffices to show that $\|\cdot\|_\mathbf{E} \sim \|\cdot\|_1$.

If $D := \max\{\|e_j\|_\mathbf{E} : 1 \leq j \leq n\}$, then $\|x\|_\mathbf{E} \leq D\|x\|_1$. This implies that, for any $x, y \in \mathbf{E}$,

$$|\|x\|_\mathbf{E} - \|y\|_\mathbf{E}| \leq \|x - y\|_\mathbf{E} \leq D\|x - y\|_1.$$

Hence the function $f : (\mathbf{E}, \|\cdot\|_1) \to \mathbb{R}_+$ given by $x \mapsto \|x\|_\mathbf{E}$ is continuous. Note that the unit sphere $S = \{x \in \mathbf{E} : \|x\|_1 = 1\}$ is a closed and bounded set, and hence compact by Lemma 4.4. Thus, $f : S \to \mathbb{R}_+$ attains its minimum on S. Therefore, there exists $x_0 \in S$ and $C \in \mathbb{R}_+$ such that

$$C = \|x_0\|_\mathbf{E} \leq \|x\|_\mathbf{E}$$

for all $x \in S$. If $C = 0$, then $x_0 = 0$, contradicting the fact that $x_0 \in S$. Hence, $C > 0$. Furthermore, for any non-zero $x \in \mathbf{E}$, $y := x/\|x\|_1 \in S$, so $C \leq \|y\|_\mathbf{E}$. Unwrapping this, we see that

$$C\|x\|_1 \leq \|x\|_\mathbf{E}$$

for all $x \in \mathbf{E}$. We conclude that $\|\cdot\|_\mathbf{E} \sim \|\cdot\|_1$. □

(End of Day 12)

Corollary 4.6. *Let* $\mathbf{E}$ *be finite dimensional, and* $\mathbf{F}$ *be any normed linear space. Then, any linear operator* $T : \mathbf{E} \to \mathbf{F}$ *is continuous.*

Proof. Define a norm on $\mathbf{E}$ by $\|x\|_G := \|x\| + \|T(x)\|$. It is easy to see that this is a norm. By Theorem 4.5, there exists $M > 0$ such that $\|T(x)\| \le \|x\|_G \le M\|x\|$ for all $x \in \mathbf{E}$. Therefore, T is continuous by Theorem 2.2. $\square$

Definition 4.7. A linear operator $T : \mathbf{E} \to \mathbf{F}$ is said to be a

(i) topological isomorphism if T is an isomorphism of vector spaces and a homeomorphism (i.e. T and T^{-1} are both continuous).

(ii) isometric isomorphism if it is a topological isomorphism that is isometric.

We write $\mathbf{E} \cong \mathbf{F}$ if they are isometrically isomorphic.

Corollary 4.8. *Let $\mathbf{E}$ be a finite dimensional normed linear space with $\dim(\mathbf{E}) = n$, then there is a topological isomorphism $T : (\mathbb{K}^n, \|\cdot\|_1) \to \mathbf{E}$.*

Proof. Choose any linear bijection $T : \mathbb{K}^n \to \mathbf{E}$ and apply Corollary 4.6 to both T and T^{-1}. $\square$

The next result follows directly from Corollary 4.8 and the fact that $(\mathbb{K}^n, \|\cdot\|_1)$ is complete.

Corollary 4.9. *Every finite dimensional normed linear space is a Banach space.*

Corollary 4.10. *Let $\mathbf{E}$ be an normed linear space and $\mathbf{F} < \mathbf{E}$ be a finite dimensional subspace. Then, $\mathbf{F}$ is closed in $\mathbf{E}$.*

Lemma 4.11 (F. Riesz, 1918). *Let $\mathbf{E}$ be an normed linear space and $\mathbf{F} < \mathbf{E}$ be a proper, closed subspace. Then for any $0 < t < 1$, there is a vector $x_t \in \mathbf{E}$ such that $\|x_t\| = 1$ and $d(x_t, \mathbf{F}) \ge t$.*

Proof. Since $\mathbf{F} < \mathbf{E}$ is a proper closed subspace, there exists $x \in \mathbf{E} \setminus \mathbf{F}$, whence $d := d(x, \mathbf{F}) > 0$. If $0 < t < 1$, then $d/t > d$ so there exists $y \in \mathbf{F}$ such that $\|x - y\| \le d/t$. Now, $x_t := \frac{x-y}{\|x-y\|}$ satisfies $\|x_t\| = 1$ and

$$d(x_t, \mathbf{F}) = \frac{1}{\|x-y\|} d(x, F) \ge t.$$

$\square$

Definition 4.12. A topological space is said to be locally compact if every point has an open neighbourhood with compact closure.

Theorem 4.13. *For a normed linear space $\mathbf{E}$, the following are equivalent:*

(i) $\mathbf{E}$ is finite dimensional.

(ii) Every closed and bounded subset of $\mathbf{E}$ is compact.

(iii) $\mathbf{E}$ is locally compact.

(iv) The closed unit ball $B[0,1]$ is compact.

(v) The unit sphere $S_{\mathbf{E}} = \{x \in \mathbf{E} : \|x\| = 1\}$ is compact.

Proof.

(i) $\Rightarrow$ (ii): If $\dim(\mathbf{E}) = n$, let $T : (\mathbb{K}^n, \|\cdot\|_1) \to \mathbf{E}$ be a topological isomorphism (which exists by Corollary 4.8). Let $B \subset \mathbf{E}$ be a closed bounded set, then $T^{-1}(B)$ is closed and bounded in $\mathbb{K}^n$. Hence, $T^{-1}(B)$ is compact by the Heine-Borel Theorem. Since T is continuous, $B = T(T^{-1}(B))$ is compact.

(ii) $\Rightarrow$ (iii): If every closed and bounded subset of $\mathbf{E}$ is compact, then $B[x,1]$ is compact for all $x \in \mathbf{E}$. Since $B[x,1] = \overline{B(x,1)}$, it follows that $\mathbf{E}$ is locally compact.

(iii) $\Rightarrow$ (iv): If $\mathbf{E}$ is locally compact, then there exists an open set U such that $0 \in U$ and $\overline{U}$ is compact. Hence, there exists $r > 0$ such that $B(0,r) \subset U$. Therefore, $B[0,r] \subset \overline{U}$, and so $B[0,r]$ is compact. But $B[0,r] = rB[0,1]$ and the map $x \mapsto rx$ is a homeomorphism. We conclude that $B[0,1]$ is compact.

(iv) $\Rightarrow$ (v): Obvious, since $S_{\mathbf{E}}$ is closed subset of $B[0,1]$.

(v) $\Rightarrow$ (i): Suppose $S_{\mathbf{E}}$ is compact and $\mathbf{E}$ is infinite dimensional, then we repeatedly apply Riesz' Lemma to arrive at a contradiction. Choose $0 \neq x_1 \in \mathbf{E}$ with $\|x_1\| = 1$, and let $\mathbf{F}_1 := \text{span}\{x_1\}$. Then, $\mathbf{F}_1 < \mathbf{E}$ is a closed proper subspace (by Corollary 4.10), and hence there exists $x_2 \in \mathbf{E}$ such that $\|x_2\| = 1$, and $d(x_2, \mathbf{F}_1) \geq 1/2$. Let $\mathbf{F}_2 := \text{span}\{x_1, x_2\}$, which is again a proper, closed subspace of $\mathbf{E}$. Once again, there exists $x_3 \in \mathbf{E}$ such that $\|x_3\| = 1$, and $d(x_3, \mathbf{F}_2) \geq 1/2$. Thus proceeding, we get a sequence (x_n) such that $\|x_n\| = 1$, and

$$d(x_n, \mathbf{F}_{n-1}) \geq 1/2,$$

where $\mathbf{F}_k := \text{span}\{x_1, x_2, \ldots, x_k\}$. In particular,

$$\|x_n - x_m\| \geq 1/2$$

for all $n \geq m$. Hence, $(x_n) \subset S_{\mathbf{E}}$ cannot have a convergent subsequence. This contradicts the assumption that $S_{\mathbf{E}}$ is compact. Hence, $\mathbf{E}$ must have been finite dimensional to begin with.

□

(End of Day 13)

III. Hilbert Spaces

1. Orthogonality

Throughout this section, we will use the letter $\mathbf{H}$ to denote a Hilbert space, and $\langle \cdot, \cdot \rangle$ to denote the inner product on $\mathbf{H}$.

Definition 1.1.

(i) We say that two elements $x, y \in \mathbf{H}$ are orthogonal if $\langle x, y \rangle = 0$. If this happens, we write $x \perp y$.

(ii) For two subsets $A, B \subset \mathbf{H}$, we write $A \perp B$ if $x \perp y$ for all $x \in A$ and $y \in B$.

(iii) For any set $A \subset \mathbf{H}$, write $A^\perp := \{x \in \mathbf{H} : x \perp y \text{ for all } y \in A\}$. If $A = \{x\}$, then we simply write $x^\perp$ instead of $\{x\}^\perp$.

Remark 1.2. Let A be any subset of $\mathbf{H}$.

(i) For each $y \in \mathbf{H}$, the linear functional $\varphi_y : \mathbf{H} \to \mathbb{K}$ given by $\varphi_y(x) := \langle x, y \rangle$ is continuous (see Example 2.3), and

$$A^\perp = \bigcap_{y \in A} \ker(\varphi_y).$$

Thus, regardless of what A is, $A^\perp$ is always closed, and a *subspace* of $\mathbf{H}$.

(ii) For any $A \subset \mathbf{H}$, it is easy to check that $A \cap A^\perp \subset \{0\}$, and that $A \subset (A^\perp)^\perp$.

(iii) Note that $(A^\perp)^\perp$ is always a closed subspace of $\mathbf{H}$, so one cannot expect the equality $A = (A^\perp)^\perp$ unless A were also a closed subspace (see Proposition 1.10).

Proposition 1.3. *Let $x, y \in \mathbf{H}$, then*

(i) (Polarization Identity): $\|x+y\|^2 = \|x\|^2 + 2\operatorname{Re}\langle x, y \rangle + \|y\|^2$.

(ii) (Pythagoras' Theorem): If $x \perp y$, then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

(iii) (Parallelogram law): $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$.

Example 1.4. Fix $1 \leq p \leq \infty$.

(i) Let $\mathbf{E} = \ell^p$, and set $x := e_1 + e_2$ and $y := e_1 - e_2$. Then, $\|x+y\|_p = \|2e_1\|_p = 2$ and $\|x-y\|_p = \|2e_2\|_p = 2$, and

$$\|x\|_p = \|y\|_p = \begin{cases} 2^{1/p} & : \text{ if } 1 \leq p < \infty \\ 1 & : \text{ if } p = \infty \end{cases}$$

Hence, the parallelogram law holds if and only if $8 = 4 \times 2^{2/p}$, which happens only when $p = 2$.

(ii) Let $\mathbf{E} = L^p[0,1]$, and $f(t) := t$ and $g(t) := 1 - t$. Then, $\|f + g\|_p = 1$ and

$$\|f - g\|_p = \|f\|_p = \|g\|_p = \begin{cases} \frac{1}{(1+p)^{1/p}} & : \text{ if } 1 \leq p < \infty \\ 1 & : \text{ if } p = \infty \end{cases}.$$

Once again, parallelogram law holds if and only if $p = 2$.

Definition 1.5. Let $\mathbf{E}$ be a vector space. A subset $A \subset \mathbf{E}$ is said to be convex if, for any $x, y \in A$, the set $[x, y] := \{tx + (1-t)y : 0 \leq t \leq 1\}$ is contained in A.

There are two kinds of convex sets we will be most interested in: The closed (or open) unit ball in a normed linear space $\mathbf{E}$, and any vector subspace of $\mathbf{E}$. Also, if A is a convex set, then $A + x$ is also a convex set for any $x \in \mathbf{E}$. Therefore, when dealing with (non-empty) convex sets, we may often assume that $0 \in A$ by translating.

Theorem 1.6 (Best Approximation Property). *Let $A \subset \mathbf{H}$ be a non-empty, closed, convex set, and $x \in \mathbf{H}$. Then, there exists a unique vector $x_0 \in A$ such that*

$$\|x - x_0\| = d(x, A) = \inf\{\|x - y\| : y \in A\}.$$

This vector x_0 is called the best approximation of x in A.

Proof. Since $d(x, A) = d(0, A - x)$, replacing A by $A - x$ (which is also convex), we may assume without loss of generality that $x = 0$.

(i) Existence: By definition, there exists a sequence $(y_n) \subset A$ such that $d := d(0, A) = \lim_{n\to\infty} \|y_n\|$. We wish to prove that (y_n) is Cauchy. By the parallelogram law,

$$\left\|\frac{y_n - y_m}{2}\right\|^2 = \frac{1}{2}(\|y_n\|^2 + \|y_m\|^2) - \left\|\frac{y_n + y_m}{2}\right\|^2.$$

Since A is convex, $(y_n + y_m)/2 \in A$, so $\|(y_n + y_m)/2\|^2 \geq d^2$. For $\epsilon > 0$, choose $N_0 \in \mathbb{N}$ such that $\|y_n\|^2 < d^2 + \epsilon$ for all $n \geq N_0$. Then, for $n, m \geq N_0$, we have

$$\left\|\frac{y_n - y_m}{2}\right\|^2 < \frac{1}{2}(2d^2 + 2\epsilon) - d^2 = \epsilon,$$

and so $\|y_n - y_m\| < 2\sqrt{\epsilon}$ for all $n, m \geq N_0$. Thus, (y_n) is Cauchy, and hence convergent in $\mathbf{H}$. Since A is closed, there exists $x_0 \in A$ such that $y_n \to x_0$. By the continuity of the norm (Remark 1.2), $d = \lim_{n\to\infty} \|y_n\| = \|x_0\|$.

(ii) Uniqueness: Suppose $x_0, x_1 \in A$ are such that $\|x_0\| = \|x_1\| = d$. Then, $(x_0 + x_1)/2 \in A$, and hence

$$d \leq \left\|\frac{1}{2}(x_0 + x_1)\right\| \leq \frac{1}{2}(\|x_0\| + \|x_1\|) \leq d,$$

and so $\|\frac{1}{2}(x_0 + x_1)\| = d$. The parallelogram law then implies that

$$d^2 = \left\|\frac{x_0 + x_1}{2}\right\|^2 = d^2 - \left\|\frac{x_0 - x_1}{2}\right\|^2.$$

Therefore, $x_0 = x_1$.

$\square$

Proposition 1.7. *Let* $\mathbf{M} < \mathbf{H}$ *be a closed subspace and* $x \in \mathbf{H}$*. Then* $x_0 \in \mathbf{M}$ *is the best approximation of* x *in* $\mathbf{M}$ *if and only if* $(x - x_0) \perp \mathbf{M}$*.*

Proof. If x_0 is the best approximation of x in $\mathbf{M}$, then we wish to prove that $\langle x - x_0, y\rangle = 0$ for all $y \in \mathbf{M}$. It suffices to prove this when $\|y\| = 1$, so fix $y \in \mathbf{M}$ with $\|y\| = 1$ and let $\alpha := \langle x - x_0, y\rangle$. Then, $z := x_0 + \alpha y \in \mathbf{M}$, so

$$\begin{aligned}
\|x - x_0\|^2 \leq \|x - z\|^2 &= \|(x - x_0) - \alpha y\|^2 \\
&= \|x - x_0\|^2 + \|\alpha y\|^2 - 2\operatorname{Re}\langle x - x_0, \alpha y\rangle \\
&= \|x - x_0\|^2 + |\alpha|^2\|y\|^2 - 2|\alpha|^2 \\
&= \|x - x_0\|^2 - |\alpha|^2.
\end{aligned}$$

Hence, $|\alpha|^2 = 0$, which implies that $x - x_0 \perp y$.

Conversely, suppose $(x - x_0) \perp \mathbf{M}$, then we wish to prove that $\|x - x_0\| = d(x, \mathbf{M})$. In other words, we wish to prove that $\|x - x_0\| \leq \|x - y\|$ for all $y \in \mathbf{M}$. For any $y \in \mathbf{M}$, we have $(x_0 - y) \in \mathbf{M}$, so $(x - x_0) \perp (x_0 - y)$. By Pythagoras' Theorem,

$$\|x - y\|^2 = \|(x - x_0) + (x_0 - y)\|^2 = \|x - x_0\|^2 + \|x_0 - y\|^2 \geq \|x - x_0\|^2.$$

Therefore, $\|x - x_0\| \leq \|x - y\|$ for all $y \in \mathbf{M}$. $\square$

Definition 1.8. Let $\mathbf{M} < \mathbf{H}$ be a closed subspace. For $x \in \mathbf{H}$, let $P_{\mathbf{M}}(x) \in \mathbf{M}$ denote the best approximation of x in $\mathbf{M}$. In other words, $P_{\mathbf{M}}(x)$ is the unique vector in $\mathbf{M}$ such that $\|x - P_{\mathbf{M}}(x)\| = d(x, \mathbf{M})$. By Proposition 1.7, this is equivalent to requiring that $x - P_{\mathbf{M}}(x) \in \mathbf{M}^{\perp}$. The map $P_{\mathbf{M}} : \mathbf{H} \to \mathbf{M}$ is called the orthogonal projection of $\mathbf{H}$ onto $\mathbf{M}$ (and we will often think of it as a map from $\mathbf{H}$ to itself, and denote it by P when the subspace is implicit).

(End of Day 14)

Proposition 1.9. *Let* $P : \mathbf{H} \to \mathbf{M}$ *be the orthogonal projection onto a closed subspace* $\mathbf{M} < \mathbf{H}$*. Then*

(i) P *is a linear transformation.*

(ii) P *is bounded and* $\|P\| \leq 1$*. If* $\mathbf{M} \neq \{0\}$*, then* $\|P\| = 1$*.*

(iii) $P \circ P = P$*.*

(iv) $\ker(P) = \mathbf{M}^{\perp}$ *and* $\operatorname{Range}(P) = \mathbf{M}$*.*

Proof.

(i) Let $x_1, x_2 \in \mathbf{H}$ and $\alpha \in \mathbb{K}$. Set $z = x_1 + \alpha x_2$ and $z_0 = P(x_1) + \alpha P(x_2)$. We wish to prove that $P(z) = z_0$. For any $y \in \mathbf{M}$,

$$\langle z - z_0, y\rangle = \langle x_1 - P(x_1), y\rangle + \alpha\langle x_2 - P(x_2), y\rangle = 0.$$

Hence, $z - z_0 \in \mathbf{M}^{\perp}$. Since $z_0 \in \mathbf{M}$, it follows from Proposition 1.7 that $P(z) = z_0$.

(ii) For any $x \in \mathbf{H}$, $x = (x - P(x)) + P(x)$ and $(x - P(x)) \perp P(x)$. Therefore, by Pythagoras' Theorem,

$$\|x\|^2 = \|x - P(x)\|^2 + \|P(x)\|^2 \geq \|P(x)\|^2.$$

Hence, P is continuous and $\|P\| \leq 1$. Moreover, if $\mathbf{M} \neq \{0\}$, then choose a non-zero vector $y \in \mathbf{M}$. Then, $P(y) = y$ (by definition of the best approximation), and thus $\|P(y)\| = \|y\|$. This proves that $\|P\| = 1$.

(iii) Now, if $y \in \mathbf{M}$ then $P(y) = y$ (as above). If $x \in \mathbf{M}$ then $y = P(x) \in \mathbf{M}$, so $P(P(x)) = P(x)$.

(iv) If $P(x) = 0$, then $x = x - P(x) \in \mathbf{M}^\perp$. Hence $\ker(P) \subset \mathbf{M}^\perp$. Conversely, if $x \in \mathbf{M}^\perp$, then $0 \in \mathbf{M}$ satisfies the conditions of Proposition 1.7. By uniqueness of the best approximation, it follows that $P(x) = 0$. Hence, $\ker(P) = \mathbf{M}^\perp$. That $\text{Range}(P) = \mathbf{M}$ is evident from part (iii) and the definition.

□

Proposition 1.10. *Let $\mathbf{M} < \mathbf{H}$ be a subspace, then $(\mathbf{M}^\perp)^\perp = \overline{\mathbf{M}}$.*

Proof. If $x \in \mathbf{M}$, then for any $y \in \mathbf{M}^\perp, \langle x, y\rangle = 0$. Hence, $x \in (\mathbf{M}^\perp)^\perp$. Thus, $\mathbf{M} \subset (\mathbf{M}^\perp)^\perp$. However, $(\mathbf{M}^\perp)^\perp$ is closed, so $\overline{\mathbf{M}} \subset (\mathbf{M}^\perp)^\perp$.

Conversely, if $x \in (\mathbf{M}^\perp)^\perp$, then let $x_0 = P_{\overline{\mathbf{M}}}(x)$ denote the best approximation of x in $\overline{\mathbf{M}}$. Then $x_0 \in \overline{\mathbf{M}}$ and $x - x_0 \in \overline{\mathbf{M}}^\perp$. However, $\mathbf{M}^\perp = \overline{\mathbf{M}}^\perp$ (by continuity of the inner product), so it follows that $\langle x, x - x_0\rangle = 0$ and $\langle x_0, x - x_0\rangle = 0$. Thus, $\|x - x_0\|^2 = \langle x - x_0, x - x_0\rangle = 0$, whence $x = x_0 \in \overline{\mathbf{M}}$. □

From now on, when S and T are two operators, 'ST' will be used to denote the composition $S \circ T$.

Proposition 1.11. *Let $\mathbf{M} < \mathbf{H}$ be a closed subspace, and $P = P_{\mathbf{M}}$. Since $\mathbf{M}^\perp$ is a closed subspace, we may set $Q = P_{\mathbf{M}^\perp}$. Then*

(i) $PQ = QP = 0$.

(ii) $P + Q = I$ where $I : \mathbf{H} \to \mathbf{H}$ denotes the identity map.

Proof.

(i) If $x \in \mathbf{H}$, then $Q(x) \in \mathbf{M}^\perp$. By Proposition 1.9, $\ker(P) = \mathbf{M}^\perp$, and hence $PQ(x) = 0$. Since $(\mathbf{M}^\perp)^\perp = \mathbf{M}$ (by Proposition 1.10), the fact that $QP = 0$ follows by symmetry.

(ii) If $x \in \mathbf{H}$, then $(x - P(x)) \in \mathbf{M}^\perp$ and $Q(x) \in \mathbf{M}^\perp$, so

$$x - P(x) - Q(x) \in \mathbf{M}^\perp$$

However, $x - Q(x) \in (\mathbf{M}^\perp)^\perp = \mathbf{M}$ by Proposition 1.10, and $P(x) \in \mathbf{M}$, so $x - Q(x) - P(x) \in \mathbf{M}$. Since $\mathbf{M} \cap \mathbf{M}^\perp = \{0\}$, it follows that $x - P(x) - Q(x) = 0$. Hence, $x = (P + Q)(x)$. This is true for any $x \in \mathbf{H}$, and thus $(P + Q) = I$.

$\square$

Definition 1.12. Given two Hilbert spaces $(\mathbf{H}_1, \langle\cdot,\cdot\rangle_{\mathbf{H}_1})$ and $(\mathbf{H}_2, \langle\cdot,\cdot\rangle_{\mathbf{H}_2})$, there is a natural inner product on the direct product $\mathbf{H}_1 \times \mathbf{H}_2$, given by

$$\langle (x_1, y_1), (x_2, y_2)\rangle := \langle x_1, x_2\rangle_{\mathbf{H}_1} + \langle x_2, y_2\rangle_{\mathbf{H}_2}.$$

Under this inner product, one can prove that $\mathbf{H}_1 \times \mathbf{H}_2$ is a Hilbert space (see **??**). This is called the Hilbert space direct sum of $\mathbf{H}_1$ and $\mathbf{H}_2$, and is denoted by $\mathbf{H}_1 \oplus \mathbf{H}_2$.

Corollary 1.13. *Let $\mathbf{M} < \mathbf{H}$ be a closed subspace, then the map $T : \mathbf{M} \oplus \mathbf{M}^\perp \to \mathbf{H}$ given by $T(x, y) := x + y$ is an isometric isomorphism.*

Proof. Let $P : \mathbf{H} \to \mathbf{M}$ and $Q : \mathbf{H} \to \mathbf{M}^\perp$ be the orthogonal projections onto $\mathbf{M}$ and $\mathbf{M}^\perp$ respectively. Then, by Proposition 1.11,

$$\mathbf{H} = (P + Q)(\mathbf{H}) = P(\mathbf{H}) + Q(\mathbf{H}) = \mathbf{M} + \mathbf{M}^\perp.$$

Furthermore, $\mathbf{M} \cap \mathbf{M}^\perp = \{0\}$. From this, it is clear that T is a linear bijection. Finally, the parallelogram law shows that, for any $x \in \mathbf{M}$ and $y \in \mathbf{M}^\perp$,

$$\|T(x, y)\|^2 = \|x + y\|^2 = \|x\|^2 + \|y\|^2 = \|(x, y)\|^2_{\mathbf{M}\oplus\mathbf{M}^\perp}.$$

Thus, T is isometric as well. $\square$

The proof of the next corollary is an easy consequence, and is relegated to the exercises.

Corollary 1.14. *If $\mathbf{M} < \mathbf{H}$ is any subspace, then $\mathbf{M}$ is dense in $\mathbf{H}$ if and only if $\mathbf{M}^\perp = \{0\}$.*

2. The Riesz Representation Theorem

Definition 2.1. If $y \in \mathbf{H}$, define $\varphi_y : \mathbf{H} \to \mathbb{K}$ by $x \mapsto \langle x, y\rangle$. Then, $\varphi_y \in \mathbf{H}^*$ and $\|\varphi_y\| = \|y\|$ (see Example 2.8). Hence, we get a map $\Delta : \mathbf{H} \to \mathbf{H}^*$ given by

$$\Delta(y) := \varphi_y.$$

Note that $\Delta(\alpha y) = \overline{\alpha}\Delta(y)$ and $\Delta(y_1 + y_2) = \Delta(y_1) + \Delta(y_2)$. In other words, Δ is a conjugate-linear isometry.

(End of Day 15)

Theorem 2.2 (Riesz Representation Theorem (F. Riesz and M. Fréchet, 1907))**.** *For any $\varphi \in \mathbf{H}^*$, there exists a unique vector $y \in \mathbf{H}$ such that*

$$\varphi(x) = \langle x, y\rangle$$

for all $x \in \mathbf{H}$. Hence, $\Delta : \mathbf{H} \to \mathbf{H}^*$ is a conjugate-linear isometric isomorphism.

Proof. Fix $0 \neq \varphi \in \mathbf{H}^*$, then $\mathbf{M} := \ker(\varphi)$ is a closed subspace of $\mathbf{H}$. Since $\varphi \neq 0$, $\mathbf{M} \neq \mathbf{H}$, and thus $\mathbf{M}^\perp \neq \{0\}$ by Corollary 1.14. Choose $y_0 \in \mathbf{M}^\perp$ such that $\varphi(y_0) = 1$. Now for any $x \in \mathbf{H}$, if we set $\alpha := \varphi(x)$, then $x - \alpha y_0 \in \mathbf{M}$. Hence,

$$0 = \langle x - \alpha y_0, y_0 \rangle = \langle x, y_0 \rangle - \varphi(x)\|y_0\|^2.$$

Note that $y_0 \neq 0$ since $\varphi(y_0) = 1$. Hence, if $y = y_0/\|y_0\|^2$, then for any $x \in \mathbf{H}$, $\varphi(x) = \langle x, y \rangle$. This completes the proof of existence. As for uniqueness, this follows from the fact that Δ is both linear and isometric (do check this!). $\square$

Remark 2.3. If $\mathbf{H}$ is a Hilbert space, then $\mathbf{H}^*$ may be equipped with an inner product

$$(\varphi_y, \varphi_z) := \langle z, y \rangle.$$

Since $\mathbf{H}^*$ is complete, it is a Hilbert space.

Proposition 2.4. *Let $\mathbf{E}$ be a normed linear space, $\mathbf{F}$ a Banach space, and $\mathbf{E}_0 < \mathbf{E}$ be a dense subspace. If $T_0 \in \mathcal{B}(\mathbf{E}_0, \mathbf{F})$, then there exists a unique linear operator $T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$ such that*

$$T|_{\mathbf{E}_0} = T_0.$$

Moreover, $\|T\| = \|T_0\|$, so T is called the norm-preserving extension of T_0.

Proof. For any $x \in \mathbf{E}$, choose a sequence $(x_n) \subset \mathbf{E}_0$ such that $x_n \to x$. Then, (x_n) is Cauchy. Since T_0 is bounded, it follows that $(T_0(x_n))$ is Cauchy in $\mathbf{F}$. Since $\mathbf{F}$ is complete, there exists $y \in \mathbf{F}$ such that $T_0(x_n) \to y$. Define $T : \mathbf{E} \to \mathbf{F}$ by

$$T(x) := \lim_{n\to\infty} T_0(x_n).$$

(i) We have to prove that T is well-defined: Suppose $(z_n) \subset \mathbf{E}$ is another sequence such that $z_n \to x$, then $\|T_0(z_n) - T_0(x_n)\| \leq \|T_0\|\|z_n - x_n\| \to 0$. Hence, $\lim_{n\to\infty} T_0(z_n) = \lim_{n\to\infty} T_0(x_n)$.

(ii) T is linear: If $x_n \to x$ and $y_n \to y$, then $x_n + y_n \to x + y$ (see Remark 1.2). Since T_0 is linear, it follows that

$$T(x+y) = \lim_{n\to\infty} T_0(x_n + y_n) = \lim_{n\to\infty} T_0(x_n) + \lim_{n\to\infty} T_0(y_n) = T(x) + T(y).$$

Similarly, $T(\alpha x) = \alpha T(x)$ for all $\alpha \in \mathbb{K}$ and $x \in \mathbf{E}$.

(iii) T is bounded: If $x_n \to x$, then $\|x_n\| \to \|x\|$, and so

$$\|T(x)\| = \lim_{n\to\infty} \|T_0(x_n)\| \leq \|T_0\| \lim_{n\to\infty} \|x_n\| = \|T_0\|\|x\|.$$

Therefore, T is bounded with $\|T\| \leq \|T_0\|$. However, T is an extension of T_0, and so $\|T\| \geq \|T_0\|$ holds trivially. Thus, $\|T\| = \|T_0\|$.

(iv) Uniqueness: Let $T_1, T_2 \in \mathcal{B}(\mathbf{E}, \mathbf{F})$ be two bounded linear operators such that $T_1|_{\mathbf{E}_0} = T_0 = T_2|_{\mathbf{E}_0}$. Then, for any $x \in \mathbf{E}$, choose $(x_n) \subset \mathbf{E}_0$ such that $x_n \to x$, so that
$$T_1(x) = \lim_{n\to\infty} T_1(x_n) = \lim_{n\to\infty} T_0(x_n) = \lim_{n\to\infty} T_2(x_n) = T_2(x).$$
Therefore, $T_1 = T_2$.

□

Corollary 2.5. *Let $\mathbf{E}$ be a normed linear space and $\mathbf{E}_0 < \mathbf{E}$ be a dense subspace of $\mathbf{E}$. Then, the map $\mathbf{E}^* \to \mathbf{E}_0^*$ given by*
$$\varphi \mapsto \varphi|_{\mathbf{E}_0}$$
is an isometric isomorphism of Banach spaces.

Proof. The map $S : \varphi \mapsto \varphi|_{\mathbf{E}_0}$ is clearly well-defined and linear. Furthermore, by Proposition 2.4, for any $\psi \in \mathbf{E}_0^*$, there exists a unique element $\varphi \in \mathbf{E}^*$ such that
$$\varphi|_{\mathbf{E}_0} = \psi.$$
Hence, S is surjective. Also, since $\|\varphi\| = \|\psi\|$, it follows that S is isometric and thus injective. □

Corollary 2.6. *Let $\mathbf{M} < \mathbf{H}$ be a subspace of $\mathbf{H}$, and let $\varphi : \mathbf{M} \to \mathbb{K}$ be a bounded linear functional. Then, there exists $\psi \in \mathbf{H}^*$ such that*
$$\psi|_{\mathbf{M}} = \varphi \text{ and } \|\psi\| = \|\varphi\|.$$

We say that ψ is a norm-preserving extension of φ. It may not be unique (see **??**).

Proof. Let $\varphi \in \mathbf{M}^*$, then by Corollary 2.5, there exists $\psi_0 : \overline{\mathbf{M}} \to \mathbb{K}$ linear such that $\psi_0|_{\mathbf{M}} = \varphi$ and $\|\psi_0\| = \|\varphi\|$. Since $\overline{\mathbf{M}} < \mathbf{H}$, it is a Hilbert space. By the Riesz Representation Theorem, there exists $y \in \overline{\mathbf{M}}$ such that $\psi_0(x) = \langle x, y\rangle$ for all $x \in \overline{\mathbf{M}}$. Now simply define $\psi : \mathbf{H} \to \mathbb{K}$ by
$$\psi(x) := \langle x, y\rangle.$$
Then, ψ is clearly an extension of ψ_0, and hence of φ. Furthermore, $\|\psi\| = \|y\| = \|\psi_0\| = \|\varphi\|$. □

(End of Day 16)

3. Orthonormal Bases

Definition 3.1. Let $\mathbf{H}$ be a Hilbert space, and Λ be a subset of $\mathbf{H}$.

(i) Λ is said to be orthogonal if $x \perp y$ for all distinct $x, y \in \Lambda$.

(ii) Λ is said to be orthonormal if it is orthogonal and $\|x\| = 1$ for all $x \in \Lambda$.

(iii) A maximal orthonormal set is called an orthonormal basis of $\mathbf{H}$.

Warning: An orthonormal basis may not be a Hamel basis for $\mathbf{H}$

Lemma 3.2. *Every orthonormal set is linearly independent.*

Proof. Let Λ be an orthonormal set, and $\{x_1, x_2, \ldots, x_n\} \subset \Lambda$ satisfy $\sum_{i=1}^{n} \alpha_i x_i = 0$. Then, for any fixed $j \in \{1, 2, \ldots, n\}$, $\alpha_j = \langle \sum_{i=1}^{n} \alpha_i x_i, x_j \rangle = 0$. Hence, $\{x_1, x_2, \ldots, x_n\}$ is linearly independent. □

Lemma 3.3. *Let $\Lambda \subset \mathbf{H}$ be an orthonormal set, then the following are equivalent:*

(i) *Λ is an orthonormal basis of $\mathbf{H}$.*

(ii) $\Lambda^{\perp} = \{0\}$.

(iii) $\text{span}(\Lambda)$ *is dense in* $\mathbf{H}$.

Proof. Let $\Lambda \subset \mathbf{H}$ be an orthonormal set.

(i) $\Rightarrow$ (ii): Suppose Λ is an orthonormal basis, and $\Lambda^{\perp} \neq \{0\}$, then choose $x \in \Lambda^{\perp}$ such that $\|x\| = 1$, then $\Lambda \cup \{x\}$ is an orthonormal set. This contradicts the maximality of Λ. Therefore, it must happen that $\Lambda^{\perp} = \{0\}$.

(ii) $\Rightarrow$ (iii): Suppose $\Lambda^{\perp} = \{0\}$. Since $\text{span}(\Lambda)^{\perp} \subset \Lambda^{\perp}$, it follows that $\text{span}(\Lambda)^{\perp} = \{0\}$. By Corollary 1.14, $\text{span}(\Lambda)$ is dense in $\mathbf{H}$.

(iii) $\Rightarrow$ (i): Suppose $\overline{\text{span}(\Lambda)} = \mathbf{H}$, and let Λ' be an orthonormal set that contains Λ. We wish to prove that $\Lambda' = \Lambda$. Suppose not, then there is a vector $x \in \Lambda' \setminus \Lambda$. Since Λ' is orthonormal, $\|x\| = 1$, and $x \perp \Lambda$. This implies that $x \perp \text{span}(\Lambda)$. By continuity of the inner product (Example 2.3), it follows that

$$x \perp \overline{\text{span}(\Lambda)}.$$

Hence, $x \perp \mathbf{H}$ and so $x = 0$. This contradicts the assumption that $\|x\| = 1$. We must conclude that $\Lambda' = \Lambda$.

□

The next result follows directly as in Theorem 1.3 by the use of Zorn's Lemma.

Theorem 3.4. *If $\Lambda_0 \subset \mathbf{H}$ is any orthonormal set, then there is an orthonormal basis $\Lambda \subset \mathbf{H}$ that contains Λ_0.*

Example 3.5.

(i) Let $\mathbf{H} = (\mathbb{K}^n, \|\cdot\|_2)$, then the standard basis is an orthonormal basis.

(ii) Let $\mathbf{H} = \ell^2$, then $\{e_1, e_2, \ldots\}$ is an orthonormal basis.

Proof. Clearly, $\Lambda = \{e_1, e_2, \ldots\}$ is orthonormal. Furthermore, if $x = (x_n) \perp \Lambda$, then $x_n = \langle x, e_n \rangle = 0$ for all n, and hence $\Lambda^{\perp} = \{0\}$. □

(iii) Let $\mathbf{H} = L^2[-\pi, \pi]$ and $\mathbb{K} = \mathbb{C}$. For $n \in \mathbb{Z}$, define

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

Then

$$\langle e_n, e_m \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)t} dt = \begin{cases} 1 & : \text{ if } n = m, \\ 0 & : \text{ otherwise.} \end{cases}$$

It is a fact $\{e_n : n \in \mathbb{Z}\}$ forms an orthonormal basis (without proof).

Theorem 3.6 (Gram-Schmidt Orthogonalization (Gram, 1883 and Schmidt, 1907))**.** *Let $\{x_1, x_2, \dots, x_n\} \subset \mathbf{H}$ be linearly independent. Define $\{u_1, u_2, \dots, u_n\}$ inductively by $u_1 := x_1$, and*

$$u_j = x_j - \sum_{i=1}^{j-1} \frac{\langle x_j, u_i \rangle}{\langle u_i, u_i \rangle} u_i \tag{III.1}$$

for $j \geq 2$. Then, $\{u_1, u_2, \dots, u_n\}$ is an orthogonal set, and $\text{span}(\{u_1, u_2, \dots, u_n\}) = \text{span}(\{x_1, x_2, \dots, x_n\})$.

Proof. We proceed by induction on n, since this is clearly true if $n = 1$. If $n > 1$, suppose $\{u_1, u_2, \dots, u_{n-1}\}$ is an orthogonal set such that $\text{span}(\{u_1, u_2, \dots, u_{n-1}\}) = \text{span}(\{x_1, x_2, \dots, x_{n-1}\})$. Then, if u_n is given by Equation III.1, then clearly $u_n \in \text{span}(\{x_1, x_2, \dots, x_n\})$ and $u_n \neq 0$ since $\{x_1, x_2, \dots, x_n\}$ is linearly independent. Also, $\langle u_n, u_j \rangle = 0$ for all $j < n$, so $\{u_1, u_2, \dots, u_n\}$ is orthogonal. In particular, $\{u_1, u_2, \dots, u_n\}$ is linearly independent (by Lemma 3.2). Hence,

$$\text{span}(\{u_1, u_2, \dots, u_n\}) = \text{span}(\{x_1, x_2, \dots, x_n\}),$$

since both spaces have the same dimension. □

Corollary 3.7. *If $\mathbf{H}$ is a Hilbert space and $\{x_1, x_2, \dots\}$ is a linearly independent set, then there exists an orthonormal set $\{e_1, e_2, \dots\}$ such that* $\text{span}(\{e_1, e_2, \dots, e_n\}) = \text{span}(\{x_1, x_2, \dots, x_n\})$ *for all $n \in \mathbb{N}$.*

Example 3.8. Let $\mathbf{H} = L^2[-1, 1]$ and $x_n(t) = t^n$, then $\{x_0, x_1, \dots\}$ is a linearly independent subset of $\mathbf{H}$. The Gram-Schmidt process gives us an orthogonal set $\{P_0, P_1, \dots\}$ which is given explicitly as

$$P_n(t) = \frac{1}{2^n n!} \left(\frac{d}{dt} \right)^n (t^2 - 1)^n.$$

These polynomials are called Legendre polynomials. Since $\text{span}(\{P_n\}) = \text{span}(\{x_n\})$, it follows by Weierstrass' Approximation Theorem and Proposition 3.12 that

$$\overline{\text{span}}(\{P_1, P_2, \dots\}) = \mathbf{H}.$$

Therefore, if we divide by their norms, the set

$$\left\{ \sqrt{\frac{2n+1}{2}} P_n : n = 0, 1, \dots \right\}$$

forms an orthonormal basis for $\mathbf{H}$.

(End of Day 17)

Proposition 3.9. *Let $\{e_1, e_2, \dots, e_n\}$ be an orthonormal set in $\mathbf{H}$, and let $\mathbf{M} := \operatorname{span}\{e_1, e_2, \dots, e_n\}$. If $P : \mathbf{H} \to \mathbf{M}$ denotes the orthogonal projection onto $\mathbf{M}$, then*

$$P(x) = \sum_{k=1}^{n} \langle x, e_k \rangle e_k.$$

for any $x \in \mathbf{H}$.

Proof. Let $x \in \mathbf{H}$, and set $x_0 := \sum_{k=1}^{n} \langle x, e_k \rangle e_k$. Then, $x_0 \in \mathbf{M}$, and $\langle x - x_0, e_j \rangle = 0$ for any $1 \leq j \leq n$. Hence, $x - x_0 \in \mathbf{M}^{\perp}$. Proposition 1.7 then implies that $P(x) = x_0$. □

Theorem 3.10 (Bessel's Inequality). *If $\{e_1, e_2, \dots\}$ is an orthonormal set and $x \in \mathbf{H}$, then $\sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 \leq \|x\|^2$.*

Proof. For each $n \in \mathbb{N}$, write $x_n := x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i$. Then, $x_n \perp e_i$ for all $1 \leq i \leq n$, so $x_n \perp \sum_{i=1}^{n} \langle x, e_i \rangle e_i$. Thus, by Pythagoras' Theorem,

$$\|x\|^2 = \|x_n\|^2 + \left\| \sum_{i=1}^{n} \langle x, e_i \rangle e_i \right\|^2 = \|x_n\|^2 + \sum_{i=1}^{n} |\langle x, e_i \rangle|^2 \geq \sum_{i=1}^{n} |\langle x, e_i \rangle|^2.$$

This is true for all $n \in \mathbb{N}$, which implies the result. □

If a series of complex numbers converge, then the n^{th} term goes to zero as $n \to \infty$. This observation gives us an important consequence of Bessel's Inequality.

Corollary 3.11 (Riemann-Lebesgue Lemma). *Let $\{e_1, e_2, \dots\}$ be an orthonormal set, and $x \in \mathbf{H}$. Then, $\lim_{n\to\infty} \langle x, e_n \rangle = 0$.*

Corollary 3.12. *Let Λ be an orthonormal set in $\mathbf{H}$ and $x \in \mathbf{H}$. Then, $\{e \in \Lambda : \langle x, e \rangle \neq 0\}$ is a countable set.*

Proof. For each $n \in \mathbb{N}$, define $\Lambda_n := \{e \in \Lambda : |\langle x, e \rangle| \geq 1/n\}$. If $\{e_1, e_2, \dots, e_N\} \subset \Lambda_n$, then by Bessel's Inequality,

$$\frac{N}{n^2} = \sum_{k=1}^{N} \frac{1}{n^2} \leq \sum_{k=1}^{N} |\langle x, e_k \rangle|^2 \leq \|x\|^2.$$

Thus, Λ_n must be finite, and $\{e \in \Lambda : \langle x, e \rangle \neq 0\} = \bigcup_{n=1}^{\infty} \Lambda_n$ must be countable. □

Remark 3.13. Let $\{x_\alpha : \alpha \in J\}$ be a (possibly uncountable) set in an normed linear space $\mathbf{E}$. We want to make sense of an expression of the form

$$\sum_{\alpha \in J} x_\alpha. \tag{III.2}$$

If J is countable, we can do this by enumerating $J = \mathbb{N}$ and treating Equation III.2 as a series in the usual sense. However, if J is uncountable, we need to make some changes. Define $\mathcal{F} := \{F \subset J : F \text{ is finite}\}$. For each $F \in \mathcal{F}$, we define a partial sum

$$s_F := \sum_{\alpha \in F} x_\alpha.$$

We say that the expression in Equation III.2 exists if there is a vector $s \in \mathbf{E}$ with the property that, for all $\epsilon > 0$, there exists a finite set $F_0 \in \mathcal{F}$ such that, for any $F \in \mathcal{F}$ containing F_0, one has $\|s_F - s\| < \epsilon$.

In other words, $\mathcal{F}$ is a (directed) partially ordered set under inclusion and $\{s_F : F \in \mathcal{F}\}$ forms a net. Our requirement is that this net be convergent in $\mathbf{E}$.

Corollary 3.14. *Let Λ be an orthonormal set in $\mathbf{H}$ and $x \in \mathbf{H}$, then $\sum_{e \in \Lambda} |\langle x, e\rangle|^2 \leq \|x\|^2$.*

Proof. Clearly, the sum is the same as the sum over the set $\{e \in \Lambda : \langle x, e\rangle \neq 0\}$. By Corollary 3.12, this set is countable, and so it reduces to a countable sum. The result now follows from Bessel's Inequality. □

Lemma 3.15. *Let Λ be an orthonormal set in $\mathbf{H}$. Then, for any $x \in \mathbf{H}$, the series*

$$\sum_{e \in \Lambda} \langle x, e\rangle e$$

converges in $\mathbf{H}$.

Proof. By Corollary 3.12, the set $\{e \in \Lambda : \langle x, e\rangle \neq 0\}$ is countable. Denote this set by $\{e_n : n \in \mathbb{N}\}$, then

$$\sum_{e \in \Lambda} \langle x, e\rangle e = \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n.$$

Define $x_n := \sum_{i=1}^{n} \langle x, e_i\rangle e_i$, and it now suffices to prove that (x_n) is Cauchy. By Bessel's Inequality, $\sum_{i=1}^{\infty} |\langle x, e_i\rangle|^2 \leq \|x\|^2 < \infty$. Hence if $\epsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that $\sum_{i=N_0}^{\infty} |\langle x, e_i\rangle|^2 < \epsilon$. If $n > m \geq N_0$, Pythagoras' Theorem tells us that

$$\|x_n - x_m\|^2 = \left\| \sum_{i=m+1}^{n} \langle x, e_i\rangle e_i \right\|^2 = \sum_{i=m+1}^{n} |\langle x, e_i\rangle|^2 \leq \sum_{i=N_0}^{\infty} |\langle x, e_i\rangle|^2 < \epsilon.$$

Hence, (x_n) is Cauchy and converges in $\mathbf{H}$. □

Theorem 3.16. *Let Λ be an orthonormal basis of $\mathbf{H}$. Then, for each $x, y \in \mathbf{H}$, we have*

(i) Fourier Expansion: $x = \sum_{e \in \Lambda} \langle x, e\rangle e$.

(ii) $\langle x, y\rangle = \sum_{e \in \Lambda} \langle x, e\rangle \overline{\langle y, e\rangle}$.

(iii) Parseval's identity: $\|x\|^2 = \sum_{e \in \Lambda} |\langle x, e\rangle|^2$.

Proof. Let $x, y \in \mathbf{H}$ be fixed, then by Corollary 3.12, the set $\{e \in \Lambda : \langle x, e\rangle \neq 0\} \cup \{e \in \Lambda : \langle y, e\rangle \neq 0\}$ is countable. Replacing Λ by this set, we may assume that $\Lambda = \{e_n : n \in \mathbb{N}\}$ is countable.

(i) Write $z := \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n$, which exists by Lemma 3.15. By continuity and linearity of the inner product, it follows that $\langle z, e_j\rangle = \langle x, e_j\rangle$ for any $j \in \mathbb{N}$. Hence, $x - z \in \Lambda^{\perp}$. Since $\Lambda^{\perp} = \{0\}$, $x = z$.

(ii) By part (i), write $x = \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n$, and $y = \sum_{m=1}^{\infty} \langle y, e_m\rangle e_m$. Then, by the continuity of the inner product in both variables, we see that

$$\langle x, y\rangle = \sum_{n,m=1}^{\infty} \langle \langle x, e_n\rangle e_n, \langle y, e_m\rangle e_m\rangle .$$

Since $e_n \perp e_m$ if $n \neq m$, we get $\langle x, y\rangle = \sum_{n=1}^{\infty} \langle x, e_n\rangle \overline{\langle y, e_n\rangle}$.

(iii) Follows directly from part (ii).

□

(End of Day 18)

4. Isomorphisms of Hilbert Spaces

Lemma 4.1. *Any two orthonormal bases of a Hilbert space have the same cardinality.*

This common number is called the dimension of the Hilbert space.

Proof. If $\mathbf{H}$ is finite dimensional, then this statement is familiar from linear algebra (see [1, Section I.8], for instance). Therefore, we will assume $\mathbf{H}$ is infinite dimensional. Let Λ_1 and Λ_2 be two orthonormal bases of $\mathbf{H}$. If $f \in \Lambda_1$, then there exists $e \in \Lambda_2$ such that $\langle f, e\rangle \neq 0$. Furthermore, $\Lambda_e := \{f \in \Lambda_1 : \langle f, e\rangle \neq 0\}$ is a countable set by Corollary 3.12. Since

$$\Lambda_1 \subset \bigcup_{e \in \Lambda_2} \Lambda_e,$$

we conclude that $|\Lambda_1| \leq |\Lambda_2 \times \mathbb{N}| = |\Lambda_2|$. By symmetry, $|\Lambda_2| \leq |\Lambda_1|$ holds as well. □

There is a clash of terminology here that you need to be aware of. The Hilbert space dimension of $\mathbf{H}$ is *not necessarily* the same as the dimension of $\mathbf{H}$ as a vector space. For instance, $\dim(\ell^2) = \aleph_0$ (countable infinity), but its vector space dimension is uncountable. Proceed with caution!

Definition 4.2. Let I be any set, and $1 \leq p < \infty$

(i) A function $f : I \to \mathbb{K}$ is said to be p-summable if $\operatorname{supp}(f) := \{i \in I : f(i) \neq 0\}$ is countable, and $\sum_{i \in I} |f(i)|^p < \infty$ (where the series is defined in the sense of Remark 3.13).

(ii) Let $\ell^p(I)$ denote the set of all p-summable functions on I. Then the inequality

$$|f+g|^p \leq [2\max\{|f|,|g|\}]^p \leq 2^p[|f|^p + |g|^p]$$

shows that $\ell^p(I)$ is a vector space. If $f \in \ell^p(I)$, we define

$$\|f\|_p := \left(\sum_{i\in I} |f(i)|^p\right)^{1/p},$$

and this satisfies Minkowski's Inequality since the verification only requires a countable sum. Hence, $\ell^p(I)$ is a normed linear space. Furthermore, $\ell^p(I)$ is a Banach space as before.

(iii) Also, $\ell^2(I)$ has an inner product given by

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

Once again this is well-defined since $\operatorname{supp}(f) \cup \operatorname{supp}(g)$ is countable. Hence, $\ell^2(I)$ is a Hilbert space.

Lemma 4.3. *Let I be any set. Then,* $\dim(\ell^2(I)) = |I|$.

Proof. For each $i \in I$, define $e_i : I \to \mathbb{K}$ by

$$e_i(j) = \delta_{i,j} = \begin{cases} 1 & : \text{if } i = j \\ 0 & : \text{if } i \neq j \end{cases}$$

Then, the set $\Lambda := \{e_i : i \in I\}$ forms an orthonormal set in $\ell^2(I)$. If $f \in \ell^2(I)$ satisfies $f \perp \Lambda$, then for any $i \in I$,

$$f(i) = \langle f, e_i\rangle = 0,$$

and so $f \equiv 0$. Hence, $\Lambda^\perp = \{0\}$ and we conclude that Λ is an orthonormal basis for $\ell^2(I)$. In particular, $\dim(\ell^2(I)) = |\Lambda| = |I|$. □

The basis Λ constructed as above is called the standard orthonormal basis for $\ell^2(I)$.

Lemma 4.4. *Let* $\mathbf{H}$ *and* $\mathbf{K}$ *be Hilbert spaces. A linear map* $T : \mathbf{H} \to \mathbf{K}$ *is an isometry if and only if* $\langle Tx, Ty\rangle_{\mathbf{K}} = \langle x, y\rangle_{\mathbf{H}}$ *for all* $x, y \in \mathbf{H}$.

A simple example of an isometry that is not an isomorphism is the right shift operator $S : \ell^2 \to \ell^2$ given by

$$S(x_1, x_2, x_3, \ldots) := (0, x_1, x_2, x_3, \ldots).$$

Clearly, the only thing that prevents an isometry from being an isomorphism is surjectivity.

Definition 4.5. An operator $U : \mathbf{H} \to \mathbf{K}$ is called a unitary operator if U is a surjective isometry. $\mathbf{H}$ is said to be isomorphic to $\mathbf{K}$ if there is a unitary map $U : \mathbf{H} \to \mathbf{K}$. If such a map exists, we write $\mathbf{H} \cong \mathbf{K}$.

Clearly, If U is a unitary, then U^{-1} is also a unitary. Also, if U and V are unitaries (with appropriate range and domain), then UV is a unitary. Therefore, the notion of isomorphism is indeed an equivalence relation on the collection of Hilbert spaces.

Theorem 4.6. *Let $\mathbf{H}$ be a Hilbert space and Λ be an orthonormal basis of $\mathbf{H}$, then $\mathbf{H} \cong \ell^2(\Lambda)$.*

Proof. For any $x \in \mathbf{H}$, define $\hat{x} : \Lambda \to \mathbb{K}$ by

$$\hat{x}(e) := \langle x, e \rangle$$

By Corollary 3.12, $\text{supp}(\hat{x})$ is countable and by Bessel's Inequality, $\hat{x} \in \ell^2(\Lambda)$. Thus, we define $U : \mathbf{H} \to \ell^2(\Lambda)$ given by $x \mapsto \hat{x}$. Note that U is linear by the axioms of the inner product. Furthermore, by Theorem 3.16, we see that

$$\langle U(x), U(y) \rangle = \sum_{e \in \Lambda} \langle x, e \rangle \overline{\langle y, e \rangle} = \langle x, y \rangle$$

and so U is an isometry by Lemma 4.4. Finally, for each $f \in \Lambda$, $\widehat{f}(e) = \langle f, e \rangle = \delta_{e,f}$. Therefore, $\{U(f) : f \in \Lambda\}$ is the standard orthonormal basis for $\ell^2(\Lambda)$ and hence $\text{Range}(U)^{\perp} = \{0\}$. By Corollary 1.14, $\text{Range}(U)$ is dense in $\ell^2(\Lambda)$. However, U is an isometry, and so $\text{Range}(U)$ is a complete subspace of $\ell^2(\Lambda)$. Thus, $\text{Range}(U)$ is closed, and U is surjective. □

Corollary 4.7. *Two Hilbert spaces are isomorphic if and only if they have the same dimension.*

Proof. Suppose $U : \mathbf{H} \to \mathbf{K}$ is an isomorphism, and $\Lambda_{\mathbf{H}} \subset \mathbf{H}$ is any orthonormal basis for $\mathbf{H}$, then $U(\Lambda_{\mathbf{H}}) \subset \mathbf{K}$ is an orthonormal set. Hence, by Theorem 3.4, there is an orthonormal basis $\Lambda_{\mathbf{K}}$ of $\mathbf{K}$ such that $U(\Lambda_{\mathbf{H}}) \subset \Lambda_{\mathbf{K}}$. In particular,

$$\dim(\mathbf{H}) = |\Lambda_{\mathbf{H}}| = |U(\Lambda_{\mathbf{H}})| \leq |\Lambda_{\mathbf{K}}| = \dim(\mathbf{K}).$$

By symmetry, $\dim(\mathbf{K}) \leq \dim(\mathbf{H})$ as well.

Conversely, suppose $\dim(\mathbf{H}) = \dim(\mathbf{K})$, let $\Lambda_{\mathbf{H}}$ and $\Lambda_{\mathbf{K}}$ denote two orthonormal bases of $\mathbf{H}$ and $\mathbf{K}$ respectively. Then $|\Lambda_{\mathbf{H}}| = |\Lambda_{\mathbf{K}}|$, and hence there is a natural isomorphism $\ell^2(\Lambda_{\mathbf{H}}) \xrightarrow{\cong} \ell^2(\Lambda_{\mathbf{K}})$. Now apply Theorem 4.6 to conclude that $\mathbf{H} \cong \mathbf{K}$. □

(End of Day 19)

Theorem 4.8. *A Hilbert space $\mathbf{H}$ is separable if and only if $\dim(\mathbf{H})$ is countable.*

Proof. If $\mathbf{H}$ is separable, and $\Lambda \subset \mathbf{H}$ is an orthonormal set, then for any pair of distinct vectors $e, f \in \Lambda$, we have

$$\|e - f\| = \sqrt{2}.$$

Hence, the balls $\{B(e; \sqrt{2}/2) : e \in \Lambda\}$ form a family of disjoint open sets in $\mathbf{H}$. Since $\mathbf{H}$ is separable, this family must be countable, and hence Λ must be countable.

Conversely, if $\mathbf{H}$ has a countable orthonormal basis $\Lambda := \{e_n : n \in \mathbb{N}\}$, then $\text{span}(\Lambda)$ is dense in $\mathbf{H}$. Now, set $\mathbb{K}_0 = \mathbb{Q}$ or $\mathbb{K}_0 = \mathbb{Q} \times \mathbb{Q}$, according as $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. Then, $A := \text{span}_{\mathbb{K}_0}(\Lambda)$ is countable, and $\overline{A} = \mathbf{H}$. Therefore, $\mathbf{H}$ is separable. □

Corollary 4.9. *Any two separable, infinite dimensional Hilbert spaces are isomorphic.*

5. Fourier Series of L^2 Functions

We will assume that $\mathbb{K} = \mathbb{C}$ throughout this section. Back in Example 3.5, we had defined an orthonormal set $\{e_n : n \in \mathbb{Z}\}$ in $L^2[-\pi, \pi]$ by

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

Remark 5.1. Let X be a compact Hausdorff space, and let $C(X)$ denote the space of all complex-valued, continuous functions on X. Given $f, g \in C(X)$, we may define the *product* of f and g by

$$(f \cdot g)(x) := f(x)g(x).$$

Clearly, $C(X)$ is closed under this operation and has the structure of a ring. Moreover,

$$(f + g) \cdot h = (f \cdot h) + (g \cdot h), \text{ and } \alpha(f \cdot g) = (\alpha f) \cdot g$$

for all $f, g, h \in C(X)$ and $\alpha \in \mathbb{C}$. This gives $C(X)$ the structure of an algebra. Therefore, we may now speak of a subalgebra of $C(X)$ - namely, a vector subspace of $C(X)$ that is closed under this multiplication operation.

Theorem 5.2 (Stone-Weierstrass). *Let X be a compact Hausdorff space, and let $C(X)$ be the algebra of continuous, complex-valued functions on X, equipped with the supremum norm. Let $\mathcal{A}$ be a subalgebra of $C(X)$ satisfying the following properties:*

(P1) $\mathcal{A}$ contains the constant function $\mathbf{1}$.

(P2) For any pair of distinct points $x, y \in X$, there is a function $f \in \mathcal{A}$ such that $f(x) \neq f(y)$. In other words, $\mathcal{A}$ separates points of X.

(P3) If $f \in \mathcal{A}$, then $\overline{f} \in \mathcal{A}$, where

$$\overline{f}(x) := \overline{f(x)}.$$

Then, $\mathcal{A}$ is dense in $C(X)$.

Proposition 5.3. *Let $S^1 := \{z \in \mathbb{C} : |z| = 1\}$, and let $C(S^1)$ be equipped with the supremum norm. Let $\mathbf{F}$ denote the space of all polynomials in z and $\overline{z}$, thought of as a subspace of $C(S^1)$. Then, $\mathbf{F}$ is dense in $C(S^1)$.*

Fix $\mathbf{E} := \{f \in C[-\pi, \pi] : f(-\pi) = f(\pi)\}$. For $n \in \mathbb{Z}$, define $e_n \in \mathbf{E}$ as above, and set $\mathcal{A} := \text{span}(\{e_n\}) \subset \mathbf{E}$. Note that

$$\mathcal{A} = \text{span}(\{\cos(nt), \sin(nt) : n \in \mathbb{Z}\}).$$

Hence, an element of $\mathcal{A}$ is called a trigonometric polynomial.

Now, if $f \in C[-\pi, \pi]$ and $1 \leq p < \infty$, then $\|f\|_p \leq (2\pi)^{1/p}\|f\|_\infty$. Hence, if a subset of $C[-\pi, \pi]$ is dense with respect to the supremum norm, then it is dense with respect to any p-norms. We wish to prove that the set $\mathcal{A}$ defined above is dense in $C[-\pi, \pi]$

with respect to $\|\cdot\|_2$. In order to do this, we need one fact about S^1. It is the quotient space of the interval $[-\pi, \pi]$ via the identification $-\pi \sim \pi$. Specifically, the map $q : [-\pi, \pi] \to S^1$ given by

$$t \mapsto e^{it}$$

is the quotient map. Therefore, a function $F : S^1 \to \mathbb{C}$ is continuous if and only if $F \circ q : [-\pi, \pi] \to \mathbb{C}$ is continuous. We use this fact below.

Proposition 5.4. *$\mathcal{A}$ is dense in $\mathbf{E}$ with respect to the supremum norm. Therefore, it is dense with respect to $\|\cdot\|_p$ for all $1 \leq p \leq \infty$.*

Proof. Define $\theta : C(S^1) \to \mathbf{E}$ by

$$\theta(f)(t) := f(e^{it}).$$

Then θ is linear, and, since the function $q : [-\pi, \pi] \to S^1$ is surjective, it follows that $\|\theta(f)\|_\infty = \|f\|_\infty$. Therefore, θ is isometric, and hence injective.

Given $f \in \mathbf{E}$, define $F : S^1 \to \mathbb{C}$ by $F(e^{it}) := f(t)$. Since $f(-\pi) = f(\pi)$, F is well-defined and continuous (by the discussion preceding this proof). Hence, $F \in C(S^1)$ and $\theta(F) = f$. Therefore, θ is an isometric isomorphism.

If $\zeta \in C(S^1)$ denotes the identity function $\zeta(z) = z$, we have $\theta(\zeta)(t) = e^{it}$ and $\theta(\overline{\zeta})(t) = e^{-it}$. Therefore, if $\mathbf{F}$ is the subspace of $C(S^1)$ defined in Proposition 5.3, then $\theta(\mathbf{F}) = \mathcal{A}$. Since $\mathbf{F}$ is dense in $C(S^1)$, $\mathcal{A}$ is dense in $\mathbf{E}$. $\square$

(End of Day 20)

Lemma 5.5. *$\mathbf{E}$ is dense in $L^p[-\pi, \pi]$ with respect to $\|\cdot\|_p$ for all $1 \leq p < \infty$.*

Proof. For $n \in \mathbb{N}$, define

$$f_n(t) = \begin{cases} 1 & : \text{ if } -\pi + 1/n \leq t \leq \pi - 1/n, \\ 0 & : \text{ if } t \in \{-\pi, \pi\}, \\ \text{linear} & : \text{ otherwise} \end{cases}$$

Then $\|f_n - 1\|_p \leq \frac{2}{n}$. For any $g \in C[-\pi, \pi]$, note that $f_n g \in \mathbf{E}$ and $\|f_n g - g\|_p \leq \frac{2}{n}\|g\|_\infty \to 0$. Hence, $C[-\pi, \pi] \subset \overline{\mathbf{E}}^{\|\cdot\|_p}$. Now, the result follows from the fact that $C[-\pi, \pi]$ is dense in $L^p[-\pi, \pi]$ (Proposition 3.12). $\square$

Theorem 5.6. *If*

$$e_n(t) := \frac{1}{\sqrt{2\pi}} e^{int},$$

then the set $\{e_n : n \in \mathbb{Z}\}$ forms an orthonormal basis of $L^2[-\pi, \pi]$.

Proof. By Example 3.5, we know that it is an orthonormal set. By Proposition 5.4 and Lemma 5.5, $\overline{\text{span}(\{e_n\})} = L^2[-\pi, \pi]$. By Lemma 3.3, it is an orthonormal basis. $\square$

Theorem 5.7. *Let $f \in L^2[-\pi, \pi]$. For $n \in \mathbb{Z}$, the $\underline{n^{th} \text{ Fourier coefficient}}$ of f is defined as*

$$\widehat{f}(n) := \langle f, e_n \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt.$$

Then, we have

(i) *Fourier Expansion: $f(t) \sim \sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{int}$.*

   Note: It is traditional to use the symbol $\sim$ here to indicate that the convergence is in the L^2 norm, but not necessarily pointwise.

(ii) *Fourier Series: The map $U : L^2[-\pi, \pi] \to \ell^2(\mathbb{Z})$ given by $f \mapsto \widehat{f}$ is an isomorphism of Hilbert spaces.*

(iii) *Parseval's identity: $\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^2$.*

(iv) *Riemann-Lebesgue Lemma:*

$$\lim_{n \to \pm\infty} \int_{-\pi}^{\pi} f(t) e^{int} dt = \lim_{n \to \pm\infty} \int_{-\pi}^{\pi} f(t) \cos(nt) dt = \lim_{n \to \pm\infty} \int_{-\pi}^{\pi} f(t) \sin(nt) dt = 0.$$

(v) *Riesz-Fischer Theorem: For any $(c_n) \in \ell^2(\mathbb{Z})$, there exists $f \in L^2[-\pi, \pi]$ such that $\widehat{f}(n) = c_n$ for all $n \in \mathbb{N}$.*

IV. Dual Spaces

1. The Duals of L^p Spaces

Definition 1.1. Throughout this section, fix $1 \leq p, q \leq \infty$, where q is the conjugate exponent of p (In other words, $\frac{1}{p} + \frac{1}{q} = 1$ if $1 < p < \infty$, $q = \infty$ when $p = 1$, and vice-versa).

(i) For each $y \in \ell^q$, define $\varphi_y : \ell^p \to \mathbb{K}$ by

$$\varphi_y((x_j)) := \sum_{n=1}^{\infty} x_n y_n.$$

For $1 < p < \infty$, φ_y is well-defined by Hölder's Inequality (Theorem 1.11), and

$$|\varphi_y(x)| \leq \|x\|_p \|y\|_q.$$

If $p = 1$ or $p = \infty$, the inequality is obvious. Hence, $\varphi_y \in (\ell^p)^*$ and $\|\varphi_y\| \leq \|y\|_q$. Furthermore, for any $y, z \in \ell^q$, and $\alpha \in \mathbb{K}$, we have

$$\varphi_{y+z} = \varphi_y + \varphi_z, \text{ and } \varphi_{\alpha y} = \alpha \varphi_y.$$

Therefore, we obtain a linear operator $\Delta : \ell^q \to (\ell^p)^*$ given by

$$y \mapsto \varphi_y,$$

satisfying $\|\Delta(y)\| \leq \|y\|_q$ for all $y \in \ell^q$.

(ii) Similarly, for each $g \in L^q[a,b]$, we define $\varphi_g : L^p[a,b] \to \mathbb{K}$ by

$$f \mapsto \int_a^b f(t)g(t)dt$$

Once again, φ_g is well-defined by Hölder's Inequality, and we get a linear operator

$$\Delta : L^q[a,b] \to (L^p[a,b])^*,$$

satisfying $\|\Delta(g)\| \leq \|g\|_q$ for all $g \in L^q[a,b]$.

For $x \in \mathbb{K}$, we write

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

so that $\operatorname{sgn}(x)x = |x|$ for all $x \in \mathbb{K}$. For $j \in \mathbb{N}$, we will write $e_j = (0, 0, \ldots, 0, 1, 0, \ldots)$ (with 1 in the j^{th} position) as in Example 1.4. Note that $e_j \in \ell^p$ for all $1 \leq p \leq \infty$.

Proposition 1.2. *For $1 \leq p \leq \infty$, the map $\Delta : \ell^q \to (\ell^p)^*$ is isometric.*

Proof. In all cases, we know that $\|\varphi_y\| \leq \|y\|_q$, so it suffices to prove the reverse inequality. We break the proof into three cases.

(i) If $p = 1$ and $q = \infty$: Let $y = (y_1, y_2, \ldots) \in \ell^\infty$, then

$$|y_j| = |\varphi_y(e_j)| \leq \|\varphi_y\|\|e_j\|_1 = \|\varphi_y\|.$$

Hence, $\|y\|_\infty \leq \|\varphi_y\|$ as required.

(ii) If $p = \infty$ and $q = 1$: Let $y = (y_1, y_2, \ldots) \in \ell^1$, then for each $n \in \mathbb{N}$, define $x^n = (x_j^n)$ by

$$x_j^n = \begin{cases} \operatorname{sgn}(y_j) & : \text{ if } 1 \leq j \leq n, \\ 0 & : \text{otherwise.} \end{cases}$$

Then, $x^n \in \ell^\infty$ and $\|x^n\|_\infty \leq 1$. Furthermore,

$$\sum_{j=1}^{n} |y_j| = \sum_{j=1}^{n} x_j^n y_j = \varphi_y(x^n) \leq \|\varphi_y\|\|x^n\|_\infty \leq \|\varphi_y\|.$$

Hence, $\|y\|_1 \leq \|\varphi_y\|$.

(iii) If $1 < p < \infty$: Let $y = (y_1, y_2, \ldots) \in \ell^q$, then for each $n \in \mathbb{N}$, define $x^n = (x_j^n)$ by

$$x_j^n = \begin{cases} \operatorname{sgn}(y_j)|y_j|^{q-1} & : \text{ if } 1 \leq j \leq n, \\ 0 & : \text{otherwise.} \end{cases}$$

Then, $x^n \in \ell^p$ and

$$\sum_{j=1}^{n} |y_j|^q = \sum_{j=1}^{n} x_j^n y_j = \sum_{j=1}^{\infty} x_j^n y_j = \varphi_y(x^n).$$

Also,

$$|\varphi_y(x^n)| \leq \|\varphi_y\|\|x^n\|_p = \|\varphi_y\| \left(\sum_{j=1}^{n} |y_j|^{qp-p} \right)^{1/p} = \|\varphi_y\| \left(\sum_{j=1}^{n} |y_j|^q \right)^{1/p},$$

because $qp - p = p(q-1) = q$. Therefore, we conclude that

$$\left(\sum_{j=1}^{n} |y_j|^q \right)^{1/q} \leq \|\varphi_y\|.$$

This is true for all $n \in \mathbb{N}$, and so $\|y\|_q \leq \|\varphi_y\|$.

□

(End of Day 21)

Theorem 1.3. *If $1 \leq p < \infty$, then the map $\Delta : \ell^q \to (\ell^p)^*$ is an isometric isomorphism.*

Proof. By Proposition 1.2, it suffices to show that Δ is surjective. So fix $\varphi \in (\ell^p)^*$, and we want to construct $y \in \ell^q$ such that $\varphi = \varphi_y$. If $e_j \in \ell^p$ are defined as above, we set

$$y := (\varphi(e_1), \varphi(e_2), \ldots).$$

(i) If $p = 1$, then $q = \infty$: For each $i \in \mathbb{N}$,

$$|y_i| = |\varphi(e_i)| \leq \|\varphi\| \|e_i\|_1 = \|\varphi\|.$$

Hence, $y \in \ell^\infty$, and so $\varphi_y \in (\ell^1)^*$. Now if $x \in c_{00}$, write $x = (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$, then

$$\varphi(x) = \sum_{i=1}^n x_i \varphi(e_i) = \sum_{i=1}^n x_i y_i = \varphi_y(x).$$

Now, φ and φ_y are both continuous functions that agree on c_{00}, which is dense in ℓ^1 (Proposition 3.4). It follows that $\varphi = \varphi_y$.

(ii) If $1 < p < \infty$, then $1 < q < \infty$: For $n \in \mathbb{N}$, write

$$y^n := (y_1, y_2, \ldots, y_n, 0, 0, \ldots).$$

Then, $y^n \in c_{00} \subset \ell^q$, so we may consider the corresponding linear functional $\varphi_{y^n} \in (\ell^p)^*$. For each $n \in \mathbb{N}$, and $x \in \ell^p$, let $x^n := (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$. Then,

$$\varphi_{y^n}(x) = \sum_{j=1}^n x_j y_j = \varphi\left(\sum_{j=1}^n x_j e_j\right) = \varphi(x^n).$$

Hence, for any $x \in \ell^p$, we have

$$|\varphi_{y^n}(x)| \leq \|\varphi\| \|x^n\|_p \leq \|\varphi\| \|x\|_p,$$

and therefore $\|\varphi_{y^n}\| \leq \|\varphi\|$ for all $n \in \mathbb{N}$. Since Δ is an isometry, it follows that

$$\left(\sum_{i=1}^n |y_i|^q\right)^{1/q} = \|y^n\|_q \leq \|\varphi\|$$

for all $n \in \mathbb{N}$. We conclude that $y \in \ell^q$, and that $\|y\|_q \leq \|\varphi\|$. Now that $y \in \ell^q$, we know that $\varphi_y \in (\ell^p)^*$. Furthermore, for any $x \in c_{00}$, write $x = (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$, then

$$\varphi(x) = \sum_{i=1}^n x_i \varphi(e_i) = \varphi_y(x)$$

and so $\varphi = \varphi_y$ on c_{00}. Once again, since c_{00} is dense in ℓ^p, it follows that $\varphi = \varphi_y$.

$\square$

Corollary 1.4. *Let $1 \leq p \leq \infty$, and let q be the conjugate exponent of p. If $\mathbf{E} = (\mathbb{K}^n, \|\cdot\|_p)$, then $\mathbf{E}^* \cong (\mathbb{K}^n, \|\cdot\|_q)$.*

Proposition 1.5. *For $1 \leq p \leq \infty$, the map $\Delta : L^q[a,b] \to (L^p[a,b])^*$ is isometric.*

Proof. Fix $g \in L^q[a,b]$. We know that $\|\varphi_g\| \leq \|g\|_q$, so it suffices to prove the reverse inequality. Once again, we break it into three cases.

(i) If $p = 1$ and $q = \infty$: For $n \in \mathbb{N}$, define

$$E_n = \{t \in [a,b] : |g(t)| > \|\varphi_g\| + 1/n\},$$

and set $f_n = \operatorname{sgn}(g)\chi_{E_n}$. Then, $\|f_n\|_1 = m(E_n) < \infty$, and

$$|\varphi_g(f_n)| = \int_{E_n} |g(t)|dt > \left(\|\varphi_g\| + 1/n\right) m(E_n).$$

However, $|\varphi_g(f_n)| \leq \|\varphi_g\| m(E_n)$, and so $m(E_n) = 0$ for each $n \in \mathbb{N}$. Therefore,

$$m(\{t \in [a,b] : |g(t)| > \|\varphi_g\|\}) = m\left(\bigcup_{n=1}^{\infty} E_n\right) = 0.$$

Thus, $\|g\|_\infty \leq \|\varphi_g\|$ as required.

(ii) If $p = \infty$ and $q = 1$: Let $f = \operatorname{sgn}(g)$, then $\|f\|_\infty \leq 1$, and

$$\|g\|_1 = \int_a^b |g(t)|dt = \varphi_g(f) = |\varphi_g(f)| \leq \|\varphi_g\| \|f\|_\infty.$$

Therefore, $\|g\|_1 \leq \|\varphi_g\|$.

(iii) If $1 < p < \infty$: If we set $f = \operatorname{sgn}(g)|g|^{q-1}$, then

$$\int_a^b |f(t)|^p dt = \int_a^b |g(t)|^{(q-1)p} dt = (\|g\|_q)^q = \int_a^b |g(t)|^q dt = \varphi_g(f).$$

In particular, $f \in L^p[a,b]$ and $\|f\|_p = \|g\|_q^{q/p}$. Also, $(\|g\|_q)^q \leq \|\varphi_g\| \|f\|_p$, which implies that

$$\|g\|_q = (\|g\|_q)^{q - q/p} \leq \|\varphi_g\|.$$

$\square$

(i) Let $f \in C[a,b]$, and define

$$F(x) = \int_a^x f(t)dt.$$

Then, F is differentiable on (a,b), and $F' = f$.

(ii) If $F : [a,b] \to \mathbb{K}$ is continuously differentiable on $[a,b]$, then F' is Riemann-integrable, and for all $x \in [a,b]$,

$$F(x) = F(a) + \int_a^x F'(t)dt.$$

Lebesgue's Fundamental Theorem of Calculus is an answer to the the corresponding questions for L^1 functions:

(i) Suppose $f \in L^1[a,b]$, then we may define

$$F(x) = \int_a^x f(t)dt.$$

Is F differentiable, and, if so, is it true that $F' = f$?

(ii) Suppose $F : [a,b] \to \mathbb{C}$ is a function, under what conditions is F differentiable, and $F' \in L^1[a,b]$? Furthermore, if this is true, then does it follow that

$$F(x) = F(a) + \int_a^x F'(t)dt$$

holds for all $x \in [a,b]$?

The answer to the first of these questions is the Lebesgue Differentiation Theorem (see [2, Section 5.3] for the proof).

Theorem 1.6 (Lebesgue's Differentiation Theorem)**.** *Let $f \in L^1[a,b]$ and define*

$$F(x) = \int_a^x f(t)dt$$

Then F is differentiable a.e., and $F' = f$ a.e.

(End of Day 22)

Lemma 1.7. *Let $f \in L^1[a,b]$. Then, for any $\epsilon > 0$, there exists $\delta > 0$ such that, for any measurable set $E \subset [a,b]$,*

$$m(E) < \delta \Rightarrow \int_E |f| < \epsilon.$$

Definition 1.8. A function $F : [a,b] \to \mathbb{K}$ is said to be absolutely continuous if for any $\epsilon > 0$, there exists $\delta > 0$ such that, for any finite collection $\{[a_i, b_i] : 1 \leq i \leq n\}$ of non-overlapping subintervals of $[a,b]$,

$$\sum_{i=1}^n (b_i - a_i) < \delta \Rightarrow \sum_{i=1}^n |F(b_i) - F(a_i)| < \epsilon.$$

Example 1.9. Let $F : [a,b] \to \mathbb{K}$ be a function.

(i) If F is Lipschitz continuous, then F is absolutely continuous. This is because if $L > 0$ such that $|F(x) - F(y)| \leq L|x - y|$ for all $x, y \in [a, b]$, then for $\epsilon > 0$, we may choose $\delta := \epsilon/2L$.

(ii) If F is continuously differentiable on $[a, b]$, then F is absolutely continuous. This is because, by the Mean-Value Theorem, F is Lipschitz continuous with Lipschitz constant $L := \sup_{x\in[0,1]} |F'(x)|$.

(iii) If there exists $f \in L^1[a, b]$ such that

$$F(x) = \int_a^x f(t)dt.$$

Then, it follows from Lemma 1.7 that F is absolutely continuous.

Theorem 1.10 (Lebesgue's Fundamental Theorem of Calculus)**.** *Let $F : [a, b] \to \mathbb{K}$ be an absolutely continuous function. Then*

(i) F is differentiable a.e. and $F' \in L^1[a, b]$.

(ii) Furthermore, for almost every $x \in [a, b]$,

$$F(x) = F(a) + \int_a^x F'(t)dt.$$

Lemma 1.11. *Let $1 \leq p < \infty$, and let $g \in L^1[a, b]$. Suppose that there exists $M > 0$ such that*

$$\left| \int_a^b f(t)g(t)dt \right| \leq M\|f\|_p$$

for all bounded functions $f \in L^p[a, b]$. Then, $g \in L^q[a, b]$ and $\|g\|_q \leq M$.

Proof. Once again, we break the proof into two cases.

(i) If $p = 1$: Let $n \in \mathbb{N}$, and consider

$$E_n = \{x \in [a, b] : |g(x)| > M + 1/n\}.$$

We wish to prove that $m(E_n) = 0$. If we set let $f_n = \text{sgn}(g)\chi_{E_n}$, then $\|f_n\|_1 = m(E_n)$, and

$$\left| \int_a^b f_n(t)g(t)dt \right| = \int_{E_n} |g(t)|dt \geq (M + 1/n)m(E_n).$$

However,

$$\left| \int_a^b f_n(t)g(t)dt \right| \leq M\|f_n\|_1 = Mm(E_n),$$

and therefore $m(E_n) = 0$. This is true for all $n \in \mathbb{N}$, and hence

$$m(\{t \in [a, b] : |g(t)| > M\}) = m\left(\bigcup_{n=1}^{\infty} E_n \right) = 0.$$

Thus, $g \in L^\infty[a, b]$ and $\|g\|_\infty \leq M$ as required.

(ii) If $1 < p < \infty$: Since $g \in L^1[a,b]$, it follows that $|g(x)| < \infty$ a.e. For $n \in \mathbb{N}$, define

$$g_n(x) = \begin{cases} g(x) & : \text{ if } |g(x)| \leq n \\ 0 & : \text{ otherwise.} \end{cases}$$

Then, each g_n is measurable, bounded, and $g_n \to g$ pointwise a.e. If we set $f_n := \text{sgn}(g_n)|g_n|^{q-1}$, then

$$\|f_n\|_p^p = \int_a^b |g_n(t)|^{(q-1)p} dt = \int_a^b |g_n(t)|^q dt = \|g_n\|_q^q.$$

In particular, $f_n \in L^p[a,b]$ and $\|f_n\|_p = \|g_n\|_q^{q/p}$. Also, $|g_n|^q = f_n g_n = f_n g$, and hence

$$\|g_n\|_q^q = \int_a^b f_n(t)g(t)dt \leq M\|f_n\|_p = M\|g_n\|_q^{q/p}.$$

Therefore, $\|g_n\|_q \leq M$ for all $n \in \mathbb{N}$. Finally, by Fatou's Lemma,

$$\int_a^b |g(t)|^q dt \leq \liminf \int_a^b |g_n(t)|^q dt \leq M^q.$$

We conclude that $g \in L^q[a,b]$ and $\|g\|_q \leq M$.

□

Lemma 1.12. *Let $1 \leq p < \infty$ and $\varphi \in (L^p[a,b])^*$. Define $F : [a,b] \to \mathbb{K}$ by*

$$F(x) = \varphi(\chi_{[a,x]}).$$

Then, F is absolutely continuous.

Proof. We wish to show that, for any $\epsilon > 0$, there exists $\delta > 0$ such that, for any finite collection $\{[a_i, b_i] : 1 \leq i \leq n\}$ of non-overlapping subintervals of $[a,b]$,

$$\sum_{i=1}^n (b_i - a_i) < \delta \Rightarrow \sum_{i=1}^n |F(b_i) - F(a_i)| < \epsilon.$$

So assume φ is non-zero, fix $\epsilon > 0$, and let $\{[a_i, b_i] : 1 \leq i \leq n\}$ be a finite collection of non-overlapping intervals in $[a,b]$. Define

$$f = \sum_{i=1}^n \text{sgn}(F(b_i) - F(a_i))\chi_{[a_i,b_i]}.$$

Then,

$$\int_a^b |f(t)|^p dt = \sum_{i=1}^n |b_i - a_i|,$$

and

$$\varphi(f) = \sum_{i=1}^n \text{sgn}(F(b_i) - F(a_i))\varphi(\chi_{[b_i,a_i]}) = \sum_{i=1}^n |F(b_i) - F(a_i)|,$$

since $\varphi(\chi_{[b_i,a_i]}) = F(b_i) - F(a_i)$. Therefore,

$$\sum_{i=1}^n |F(b_i) - F(a_i)| \leq \|\varphi\| \|f\|_p = \|\varphi\| (\sum_{i=1}^n |b_i - a_i|)^{1/p}.$$

We conclude that $\delta := \epsilon^p / 2\|\varphi\|^p$ works. $\square$

Recall that a function $f : [a,b] \to \mathbb{K}$ is said to be a step function if it is a finite linear combination of characteristic functions of sub-intervals of $[a,b]$. In particular, a step function is a simple function. Let $\mathcal{S}$ denote the set of all step functions on $[a,b]$, then it is clear that $\mathcal{S} \subset L^p[a,b]$ for any $1 \leq p \leq \infty$. The next lemma is a small piece of the puzzle we need to complete our proof, and we leave it as an exercise for the reader.

Lemma 1.13. *Let $\mathcal{S}$ denote the set of all step functions on an interval $[a,b]$. If $1 \leq p < \infty$, then $\mathcal{S}$ is dense in $L^p[a,b]$. Furthermore, if $f \in L^p[a,b]$ is a bounded function, then there is a sequence $(f_n) \subset \mathcal{S}$, and $M > 0$ such that $f_n \to f$ in $L^p[a,b]$, and $|f_n| \leq M$ for all $n \in \mathbb{N}$.*

Theorem 1.14 (Riesz Representation Theorem (F. Riesz, 1909))**.** *If $1 \leq p < \infty$, then the map $\Delta : L^q[a,b] \to (L^p[a,b])^*$ is an isometric isomorphism.*

Proof. By Proposition 1.5, it suffices to prove that Δ is surjective. So fix $\varphi \in (L^p[a,b])^*$, and we want to construct $g \in L^q[a,b]$ such that $\varphi = \varphi_g$. As you might expect, we start with the function $F : [a,b] \to \mathbb{C}$ given by

$$F(x) := \varphi(\chi_{[a,x]}).$$

By Lemma 1.12, F is absolutely continuous. Hence, by Lebesgue's Fundamental Theorem of Calculus, there exists $g \in L^1[a,b]$ such that

$$\varphi(\chi_{[a,x]}) = F(x) = \int_a^x g(t)dt = \int_a^b \chi_{[a,x]}(t)g(t)dt$$

Hence, if $f \in L^p[a,b]$ is a step function, then by linearity, it follows that

$$\varphi(f) = \int_a^b f(t)g(t)dt.$$

Now, let $\mathcal{S} \subset L^p[a,b]$ denote the subspace of step functions. If $f \in L^p[a,b]$ is a bounded function, then, by Lemma 1.13, there is a sequence $(f_n) \in \mathcal{S}$, and $M > 0$ such that $f_n \to f$ in $L^p[a,b]$, and $|f_n| \leq M$ for all $n \in \mathbb{N}$. Replacing (f_n) by a subsequence if necessary, we may assume that $f_n \to f$ pointwise a.e. Hence, $f_n g \to fg$ pointwise a.e., and $|f_n g| \leq M|g| \in L^1[a,b]$ for all $n \in \mathbb{N}$. By the Dominated Convergence Theorem,

$$\int_a^b f(t)g(t)dt = \lim_{n\to\infty} \int_a^b f_n(t)g(t)dt = \lim_{n\to\infty} \varphi(f_n).$$

Since $\|f_n\|_p \to \|f\|_p$ (by Remark 1.2), it follows that

$$\left| \int_a^b f(t)g(t)dt \right| = \lim_{n\to\infty} |\varphi(f_n)| \leq \lim_{n\to\infty} \|\varphi\| \|f_n\|_p = \|\varphi\| \|f\|_p.$$

Therefore, the inequality

$$\left| \int_a^b f(t)g(t)dt \right| \le \|\varphi\| \|f\|_p$$

holds for all bounded measurable functions $f \in L^p[a,b]$. By Lemma 1.11, we conclude that $g \in L^q[a,b]$ and $\|g\|_q \le \|\varphi\|$.

Now consider the bounded linear functional $\varphi_g \in (L^p[a,b])^*$, and observe that $\varphi(f) = \varphi_g(f)$ for all $f \in \mathcal{S}$. Since $\mathcal{S}$ is dense in $L^p[a,b]$, it follows that $\varphi = \varphi_g$. This concludes the proof. $\square$

(End of Day 23)

2. The Hahn-Banach Extension Theorem

Suppose $\mathbf{E}$ is a normed linear space, and $\mathbf{F} < \mathbf{E}$ a subspace. Given a bounded linear functional $\varphi : \mathbf{F} \to \mathbb{K}$, we would like to construct a bounded linear functional $\psi : \mathbf{E} \to \mathbb{K}$ such that

$$\psi|_{\mathbf{F}} = \varphi.$$

In other words, ψ would be a continuous extension of φ. Furthermore, we would like $\|\psi\| = \|\varphi\|$. Hence, ψ would be a norm-preserving extension of φ. We have already seen two situations in which this is possible: If $\overline{\mathbf{F}} = \mathbf{E}$, then there is a unique continuous extension of φ by Proposition 2.4; and if $\mathbf{E}$ is a Hilbert space, then such is a continuous, norm-preserving extension exists by Corollary 2.6.

Definition 2.1. A seminorm is a function $p : \mathbf{E} \to \mathbb{R}_+$ such that

$$p(\alpha x) = |\alpha| p(x) \text{ and } p(x+y) \le p(x) + p(y)$$

for all $x, y \in \mathbf{E}$ and $\alpha \in \mathbb{K}$.

Lemma 2.2. *Let* $\mathbf{E}$ *be a vector space over* $\mathbb{R}$*, and* $p : \mathbf{E} \to \mathbb{R}$ *be a seminorm on* $\mathbf{E}$*. Let* $\mathbf{F} < \mathbf{E}$ *be a subspace of* $\mathbf{E}$ *of codimension one, and let* $\varphi : \mathbf{F} \to \mathbb{R}$ *be a linear functional on* $\mathbf{F}$ *satisfying*

$$\varphi(x) \le p(x)$$

for all $x \in \mathbf{F}$*. Then, there exists a linear functional* $\psi : \mathbf{E} \to \mathbb{R}$ *such that* $\psi|_{\mathbf{F}} = \varphi$*, and*

$$\psi(x) \le p(x)$$

for all $x \in \mathbf{E}$*.*

Proof. Since $\operatorname{codim}(\mathbf{F}) = 1$, there exists $e \in \mathbf{E} \setminus \mathbf{F}$ such that every $z \in \mathbf{E}$ can be expressed in the form $z = x + \alpha e$ for some (unique) $x \in \mathbf{F}$ and $\alpha \in \mathbb{R}$. As mentioned above, we need to determine a real number

$$t := \psi(e)$$

so that $\psi(z) \le p(z)$ may hold for all $z \in \mathbf{E}$.

(i) Suppose $\psi : \mathbf{E} \to \mathbb{R}$ exists such that $\psi|_{\mathbf{F}} = \varphi$ and $\psi(z) \leq p(z)$ for all $z \in \mathbf{E}$. Then set $t := \psi(e)$, so that $\psi(x + \alpha e) = \varphi(x) + \alpha t \leq p(x + \alpha e)$. Consider the two cases:

 (i) If $\alpha > 0$, then t must satisfy the inequality

$$t \leq \frac{p(x + \alpha e) - \varphi(x)}{\alpha}$$

 for all $x \in \mathbf{F}$. Hence,

$$t \leq t_1 := \inf \left\{ \frac{p(x + \alpha e) - \varphi(x)}{\alpha} : x \in \mathbf{F}, \alpha \in \mathbb{R}_{>0} \right\}.$$

 (ii) If $\alpha < 0$, we write $\beta = -\alpha > 0$. Then, for any $y \in \mathbf{F}$,

$$\psi(y - \beta e) = \varphi(y) - \beta t \leq p(y - \beta e).$$

 Hence, t must satisfy the inequality

$$t \geq \frac{\varphi(y) - p(y - \beta e)}{\beta}$$

 for all $y \in \mathbf{F}$. Therefore,

$$t \geq t_2 := \sup \left\{ \frac{\varphi(y) - p(y - \beta e)}{\beta} : y \in \mathbf{F}, \beta \in \mathbb{R}_{>0} \right\}.$$

 However, this is only possible if $t_2 \leq t_1$, so let us verify this fact.

(ii) We want to show that, for all $x, y \in \mathbf{F}$, and all $\alpha, \beta > 0$, the inequality

$$\frac{\varphi(y) - p(y - \beta e)}{\beta} \leq \frac{p(x + \alpha e) - \varphi(x)}{\alpha}$$

holds. Cross-multiplying, this amounts to proving that

$$\varphi(\alpha y) + \varphi(\beta x) \leq \beta p(x + \alpha e) + \alpha p(y - \beta e).$$

Since p is a seminorm, this reduces to proving that

$$\varphi(\alpha y + \beta x) \leq p(\beta x + \alpha\beta e) + p(\alpha y - \alpha\beta e).$$

Now, this last inequality follows from a calculation:

$$\begin{aligned} \varphi(\alpha y + \beta x) &= \varphi(\alpha y - \alpha\beta e + \alpha\beta e + \beta x) \\ &= \varphi(\alpha y - \alpha\beta e) + \varphi(\beta x + \alpha\beta e) \\ &\leq p(\alpha y - \alpha\beta e) + p(\beta x + \alpha\beta e). \end{aligned}$$

Therefore, we conclude that $t_2 \leq t_1$ as desired.

(iii) Having discussed the necessary condition in Step (i), we now construct the extension. Consider t_1 and t_2 as above, so that $t_2 \leq t_1$. We choose $t \in \mathbb{R}$ such that $t_2 \leq t \leq t_1$, and *define* $\psi : \mathbf{E} \to \mathbb{R}$ by

$$\psi(x + \alpha e) := \varphi(x) + \alpha t.$$

Then, this map ψ is linear, and satisfies the condition that $\psi(z) \leq p(z)$ for all $z \in \mathbf{E}$.

□

Theorem 2.3 (Hahn-Banach Theorem - Real Case (Hahn, 1927 and Banach, 1929))**.** *Let* $\mathbf{E}$ *be a vector space over* $\mathbb{R}$*, and* $p : \mathbf{E} \to \mathbb{R}$ *be a seminorm on* $\mathbf{E}$*. Let* $\mathbf{F} < \mathbf{E}$ *be a subspace, and* $\varphi : \mathbf{F} \to \mathbb{R}$ *a linear functional such that*

$$\varphi(x) \leq p(x)$$

for all $x \in \mathbf{F}$*. Then, there exists a linear functional* $\psi : \mathbf{E} \to \mathbb{R}$ *such that*

$$\psi|_{\mathbf{F}} = \varphi,$$

and $\psi(x) \leq p(x)$ *for all* $x \in \mathbf{E}$*.*

Proof. If $n := \operatorname{codim}(\mathbf{F}) < \infty$, then we repeat Lemma 2.2 inductively.

If not, then we appeal to Zorn's Lemma. Define $\mathcal{F}$ to be the set of all pairs $(\mathbf{W}, \psi_{\mathbf{W}})$, where $\mathbf{W} < \mathbf{E}$ is a subspace containing $\mathbf{F}$, and $\psi_{\mathbf{W}} : \mathbf{W} \to \mathbb{R}$ is a linear functional on $\mathbf{W}$ such that

$$\psi_{\mathbf{W}}|_{\mathbf{F}} = \varphi,$$

and $\psi_{\mathbf{W}}(x) \leq p(x)$ for all $x \in \mathbf{W}$. The set $\mathcal{F}$ is clearly non-empty since $(\mathbf{F}, \varphi) \in \mathcal{F}$. Define a partial order $\leq$ on $\mathcal{F}$ by setting $(\mathbf{W}_1, \psi_{\mathbf{W}_1}) \leq (\mathbf{W}_2, \psi_{\mathbf{W}_2})$ if and only if

$$\mathbf{W}_1 \subset \mathbf{W}_2 \text{ and } \psi_{\mathbf{W}_2}|_{\mathbf{W}_1} = \psi_{\mathbf{W}_1}.$$

It is easy to check that $\mathcal{F}$ now becomes a partially ordered set. We now verify that Zorn's Lemma is applicable: Let $\mathcal{C}$ be a totally ordered subset of $\mathcal{F}$. Define

$$\mathbf{W}_0 := \bigcup_{(\mathbf{W}, \psi_{\mathbf{W}}) \in \mathcal{C}} \mathbf{W}.$$

Then, since $\mathcal{C}$ is totally ordered, $\mathbf{W}_0$ is a subspace. Define $\psi_0 : \mathbf{W}_0 \to \mathbb{R}$ by

$$\psi_0(x) := \psi_{\mathbf{W}}(x)$$

if $x \in \mathbf{W}$. Then, one can check that ψ_0 is well-defined (again, since $\mathcal{C}$ is totally ordered). Furthermore, $\psi_0(x) \leq p(x)$ holds for all $x \in \mathbf{W}_0$. Hence, $(\mathbf{W}_0, \psi_0) \in \mathcal{F}$ and it is clearly an upper bound for $\mathcal{C}$.

Zorn's Lemma now tells us that $\mathcal{F}$ has a maximal element, which we denote by $(\mathbf{F}_0, \varphi_0)$. Now, we claim that $\mathbf{F}_0 = \mathbf{E}$, and this is where we need the previous lemma. Suppose not, then there exists $e \in \mathbf{E} \setminus \mathbf{F}_0$. Define $\mathbf{F}_1 := \text{span}(\mathbf{F}_0 \cup \{e\})$. Then, by Lemma 2.2, we may extend φ_0 to a linear map $\varphi_1 : \mathbf{F}_1 \to \mathbb{R}$ satisfying

$$\varphi_1(x) \leq p(x)$$

for all $x \in \mathbf{F}_1$. This would mean that $(\mathbf{F}_1, \varphi_1) \in \mathcal{F}$, which would contradict the maximality of $(\mathbf{F}_0, \varphi_0)$. Hence, $\mathbf{F}_0 = \mathbf{E}$, and $\psi = \varphi_0$ is the required linear functional. □

(End of Day 24)

Lemma 2.4. *Let* $\mathbf{E}$ *be a complex vector space.*

(i) If $\varphi : \mathbf{E} \to \mathbb{R}$ *is an* $\mathbb{R}$*-linear functional, then*

$$\widehat{\varphi}(x) := \varphi(x) - i\varphi(ix)$$

is a $\mathbb{C}$*-linear functional*

(ii) If $\psi : \mathbf{E} \to \mathbb{C}$ *is a* $\mathbb{C}$*-linear functional, then* $\varphi := \text{Re}(\psi)$ *is a* $\mathbb{R}$*-linear functional, and* $\widehat{\varphi} = \psi$.

(iii) If p *is a seminorm and* φ *and* $\widehat{\varphi}$ *are as in part (i), then the inequality*

$$|\varphi(x)| \leq p(x)$$

holds for all $x \in \mathbf{E}$ *if and only if*

$$|\widehat{\varphi}(x)| \leq p(x)$$

holds for all $x \in \mathbf{E}$.

(iv) If $\mathbf{E}$ *is a normed linear space and* φ *and* $\widehat{\varphi}$ *are as above, then* $\|\varphi\| = \|\widehat{\varphi}\|$.

Proof. The proofs of the first two statements are left as an exercise.

(iii) Suppose that the inequality

$$|\widehat{\varphi}(x)| \leq p(x)$$

holds for all $x \in \mathbf{E}$. Then, clearly, the same inequality holds with φ. Conversely, suppose

$$|\varphi(x)| \leq p(x)$$

holds for all $x \in \mathbf{E}$, then fix $x \in \mathbf{E}$ and choose $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ such that $|\widehat{\varphi}(x)| = \lambda\widehat{\varphi}(x)$. Then,

$$|\widehat{\varphi}(x)| = \widehat{\varphi}(\lambda x) = \text{Re}(\widehat{\varphi}(\lambda x)) = \varphi(\lambda x) \leq p(\lambda x) = |\lambda| p(x) = p(x).$$

Thus, $|\widehat{\varphi}(x)| \leq p(x)$ for all $x \in \mathbf{E}$.

(iv) Let $p(x) := \|\varphi\|\|x\|$, then $|\varphi(x)| \leq p(x)$ for all $x \in \mathbf{E}$. By part (iii), $|\widehat{\varphi}(x)| \leq p(x)$ for all $x \in \mathbf{E}$, and so $\|\widehat{\varphi}\| \leq \|\varphi\|$. The reverse inequality is obvious since $|\varphi(x)| = |\operatorname{Re}(\widehat{\varphi}(x))| \leq |\widehat{\varphi}(x)|$ for all $x \in \mathbf{E}$.

□

Theorem 2.5 (Hahn-Banach Theorem - General Case)**.** *Let* $\mathbf{E}$ *be a normed linear space, and* $p : \mathbf{E} \to \mathbb{R}_+$ *be a seminorm on* $\mathbf{E}$*. Let* $\mathbf{F} < \mathbf{E}$ *be a subspace, and* $\varphi : \mathbf{F} \to \mathbb{K}$ *a linear functional such that*

$$|\varphi(x)| \leq p(x)$$

for all $x \in \mathbf{F}$*. Then, there exists a linear functional* $\psi : \mathbf{E} \to \mathbb{K}$ *such that* $\psi|_{\mathbf{F}} = \varphi$*, and*

$$|\psi(x)| \leq p(x)$$

for all $x \in \mathbf{E}$*.*

Proof.

(i) Suppose $\mathbf{E}$ is an $\mathbb{R}$-vector space: By Theorem 2.3, there exists a linear functional $\psi : \mathbf{E} \to \mathbb{R}$ such that

$$\psi(x) \leq p(x)$$

for all $x \in \mathbf{E}$. Then,

$$-\psi(x) = \psi(-x) \leq p(-x) = p(x)$$

holds for all $x \in \mathbf{E}$, and so $|\psi(x)| \leq p(x)$ holds.

(ii) Suppose $\mathbf{E}$ is a $\mathbb{C}$-vector space: Define $f : \mathbf{F} \to \mathbb{R}$ by $f = \operatorname{Re}(\varphi)$. Then f is linear and

$$|f(x)| \leq p(x)$$

for all $x \in \mathbf{F}$. By part (i), there is a linear functional $g : \mathbf{E} \to \mathbb{R}$ such that

$$g|_{\mathbf{F}} = f,$$

and $|g(x)| \leq p(x)$ for all $x \in \mathbf{E}$. Define $\psi : \mathbf{E} \to \mathbb{C}$ by $\psi := \widehat{g}$ as in Lemma 2.4, then ψ is $\mathbb{C}$-linear, and satisfies

$$|\psi(x)| \leq p(x)$$

for all $x \in \mathbf{E}$. This completes the proof.

□

Corollary 2.6. *Let* $\mathbf{E}$ *be a normed linear space, and* $\mathbf{F} < \mathbf{E}$ *a subspace. Then, for any bounded linear functional* $\varphi \in \mathbf{F}^*$*, there exists* $\psi \in \mathbf{E}^*$ *such that*

$$\psi|_{\mathbf{F}} = \varphi,$$

and $\|\psi\| = \|\varphi\|$*.*

Proof. Apply the Hahn-Banach Theorem with $p(x) := \|\varphi\|\|x\|$ $\square$

Corollary 2.7. *Let* $\mathbf{E}$ *be a normed linear space,* $\{e_1, e_2, \ldots, e_n\} \subset \mathbf{E}$ *be a finite, linearly independent set, and* $\{\alpha_1, \alpha_2, \ldots, \alpha_n\} \subset \mathbb{K}$ *be arbitrary scalars. Then there exists a bounded linear functional* $\psi \in \mathbf{E}^*$ *such that*

$$\psi(e_i) = \alpha_i$$

for all $1 \leq i \leq n$.

Proof. Take $\mathbf{F} := \operatorname{span}(\{e_1, e_2, \ldots, e_n\})$ and define $\varphi : \mathbf{F} \to \mathbb{K}$ by

$$\varphi(e_i) = \alpha_i,$$

extended linearly to all of $\mathbf{F}$. Since $\mathbf{F}$ is finite dimensional, φ is a bounded linear functional on $\mathbf{F}$ (by Corollary 4.6). We may now apply Corollary 2.6 to get a linear functional $\psi \in \mathbf{E}^*$ satisfying the required conditions. $\square$

Corollary 2.8. *Let* $\mathbf{E}$ *be a normed linear space, and* $x \in \mathbf{E}$*. Then,*

$$\|x\| = \sup\{|\psi(x)| : \psi \in \mathbf{E}^*, \|\psi\| \leq 1\}.$$

Furthermore, this supremum is attained.

Proof. Set $\alpha := \sup\{|\psi(x)| : \psi \in \mathbf{E}^*, \|\psi\| \leq 1\}$. Then, is is easy to see that $\alpha \leq \|x\|$. Conversely, set $\mathbf{F} := \operatorname{span}(x)$. Define $\varphi : \mathbf{F} \to \mathbb{K}$ by

$$\beta x \mapsto \beta\|x\|.$$

Then, $\varphi \in \mathbf{F}^*$; indeed, it is easy to check that $\|\varphi\| = 1$. By Corollary 2.6, there exists $\psi \in \mathbf{E}^*$ such that $\psi(x) = \|x\|$, and $\|\psi\| = 1$. Hence, $\|x\| = \alpha$ as desired. $\square$

If $\mathbf{E}$ is a normed linear space, and $\varphi \in \mathbf{E}^*$, then we know that

$$\|\varphi\| = \sup\{|\varphi(y)| : y \in \mathbf{E}, \|y\| \leq 1\}. \tag{IV.1}$$

Corollary 2.8 may be thought of as a mirror image of this fact, except that the norm in Equation IV.1 may not be attained.

Corollary 2.9. *If* $x, y \in \mathbf{E}$ *are two vectors such that* $\psi(x) = \psi(y)$ *for all* $\psi \in \mathbf{E}^*$*, then* $x = y$.

3. Quotient Spaces

Let us fix some notation we will use throughout this section: $\mathbf{E}$ will denote a fixed normed linear space, $\mathbf{F} < \mathbf{E}$ will be a subspace, and $\pi : \mathbf{E} \to \mathbf{E}/\mathbf{F}$ is the natural quotient map $x \mapsto x + \mathbf{F}$.

(End of Day 25)

Proposition 3.1. *Define* $p : \mathbf{E} \to \mathbb{R}_+$ *by*

$$p(x) := d(x, \mathbf{F}) = \inf\{\|x - y\| : y \in \mathbf{F}\}.$$

Then,

(i) *p defines a seminorm on* $\mathbf{E}$.

(ii) *If* $\mathbf{F}$ *is closed, p induces a norm on* $\mathbf{E}/\mathbf{F}$ *given by* $\|x + \mathbf{F}\| := p(x)$.

Proof. By Homework 1.3, it suffices to prove (i).

(a) Clearly, $p(x) \geq 0$, and $p(x) = 0$ if and only if $x \in \overline{\mathbf{F}}$.

(b) Now if $\alpha \in \mathbb{K}$, then consider $d(\alpha x, \mathbf{F}) = \inf\{\|\alpha x - y\| : y \in \mathbf{F}\}$. If $\alpha = 0$, then $\alpha x = 0 \in \mathbf{F}$, so $p(\alpha x) = 0$. If $\alpha \neq 0$, then the map $y \mapsto \alpha y$ is a bijection on $\mathbf{F}$, and hence
$$p(\alpha x) = \inf\{\|\alpha x - \alpha y\| : y \in \mathbf{F}\} = |\alpha| p(x).$$

(c) Finally, if $x_1, x_2 \in \mathbf{E}$, then for any $y_1, y_2 \in \mathbf{F}$, we have $\|(x_1 + x_2) - (y_1 + y_2)\| \leq \|x_1 - y_1\| + \|x_2 - y_2\|$. Since $y_1 + y_2 \in \mathbf{F}$, it follows that
$$p(x_1 + x_2) \leq \|x_1 - y_1\| + \|x_2 - y_2\|.$$
Taking infimums independently, we conclude that $p(x_1 + x_2) \leq p(x_1) + p(x_2)$.

This proves that p is a seminorm. □

Proposition 3.2. *If* $\mathbf{F}$ *is a proper, closed subspace of* $\mathbf{E}$, *then the quotient map* $\pi : \mathbf{E} \to \mathbf{E}/\mathbf{F}$ *is continuous, and* $\|\pi\| = 1$.

Proof. If $x \in \mathbf{E}$, we have

$$\|\pi(x)\| = \|x + \mathbf{F}\| = d(x, \mathbf{F}) \leq \|x\|,$$

since $0 \in \mathbf{F}$. Hence π is continuous and $\|\pi\| \leq 1$. Furthermore, by Riesz' Lemma (Lemma 4.11), for each $0 < t < 1$, there exists $x_t \in \mathbf{E}$ such that $\|x_t\| = 1$, and $\|\pi(x_t)\| \geq t$. Hence, $\|\pi\| \geq 1$. □

Proposition 3.3. *Let* $\mathbf{E}$ *be an normed linear space,* $\mathbf{F} < \mathbf{E}$ *a subspace,* $x_0 \in \mathbf{E} \setminus \mathbf{F}$ *and suppose that*
$$d := d(x_0, \mathbf{F}) > 0.$$
Then, there exists $\varphi \in \mathbf{E}^*$ *such that*

(a) $\varphi(x) = 0$ *for all* $x \in \mathbf{F}$,

(b) $\varphi(x_0) = 1$, *and*

(c) $\|\varphi\| = d^{-1}$.

Proof. Since $d(x_0, \mathbf{F}) = d(x_0, \overline{\mathbf{F}}) > 0$, we may assume without loss of generality that $\mathbf{F}$ is closed. If $\pi : \mathbf{E} \to \mathbf{E}/\mathbf{F}$ denotes the natural quotient map, then $\|\pi(x_0)\| = d$. Hence, by Corollary 2.8, there exists $\psi \in (\mathbf{E}/\mathbf{F})^*$ such that $\|\psi\| = 1$, and

$$\psi(x_0 + \mathbf{F}) = \|x_0 + \mathbf{F}\| = d.$$

Define $\varphi : \mathbf{E} \to \mathbb{K}$ by

$$\varphi := d^{-1}\psi \circ \pi.$$

Then, $\varphi \in \mathbf{E}^*$ satisfies conditions (a) and (b). To verify condition (c), observe that $\|\pi\| \leq 1$, and therefore

$$|\varphi(y)| \leq d^{-1}\|\psi\|\|\pi(y)\| \leq d^{-1}\|y\|$$

for all $y \in \mathbf{E}$. Hence, $\|\varphi\| \leq d^{-1}$. However, $\|\psi\| = 1$, and hence

$$\sup\{|\psi(z + \mathbf{F})| : \|z + \mathbf{F}\| = 1\} = 1.$$

Therefore, there exists a sequence $(x_n + \mathbf{F}) \in \mathbf{E}/\mathbf{F}$ such that $\|x_n + \mathbf{F}\| = 1$ for all $n \in \mathbb{N}$, and

$$\lim_{n\to\infty} |\psi(x_n + \mathbf{F})| = 1.$$

Let $y_n \in \mathbf{F}$ such that $\|x_n + y_n\| < 1 + 1/n$ for each $n \in \mathbb{N}$, so that

$$d^{-1}|\psi(x_n + \mathbf{F})| = |\varphi(x_n + y_n)| \leq \|\varphi\|\|x_n + y_n\| \leq \|\varphi\|(1 + 1/n).$$

Letting $n \to \infty$, we conclude that $d^{-1} \leq \|\varphi\|$. □

For any bounded linear functional $\varphi \in \mathbf{E}^*$, it is clear that its restriction $\varphi|_{\mathbf{F}}$ is a bounded linear functional on $\mathbf{F}$, and that $\|\varphi|_{\mathbf{F}}\| \leq \|\varphi\|$. This gives us a bounded linear map $R : \mathbf{E}^* \to \mathbf{F}^*$ given by

$$\varphi \mapsto \varphi|_{\mathbf{F}}.$$

The Hahn-Banach Theorem merely says that this map is surjective.

Definition 3.4. The space $\ker(R) := \{\varphi \in \mathbf{E}^* : \varphi|_{\mathbf{F}} = 0\}$ is called the <u>annihilator</u> of $\mathbf{F}$, and is denoted by $\mathbf{F}^{\perp}$. Note that $\mathbf{F}^{\perp}$ is a subspace of $\mathbf{E}^*$, and that it is closed.

Corollary 3.5. *Let $\mathbf{E}$ be an normed linear space, and $\mathbf{F} < \mathbf{E}$ be a subspace. Then,*

$$\overline{\mathbf{F}} = \bigcap_{\varphi \in \mathbf{F}^{\perp}} \ker(\varphi).$$

In other words, a vector $x \in \mathbf{E}$ belongs to $\overline{\mathbf{F}}$ if and only if $\varphi(x) = 0$ for all $\varphi \in \mathbf{F}^{\perp}$.

Proof. Clearly,

$$\mathbf{W} := \bigcap_{\varphi \in \mathbf{F}^{\perp}} \ker(\varphi)$$

is a closed subspace containing $\mathbf{F}$, and hence $\overline{\mathbf{F}} \subset \mathbf{W}$. Conversely, if $x_0 \notin \overline{\mathbf{F}}$, then $d(x_0, \mathbf{F}) > 0$. By Proposition 3.3, there exists $\varphi \in \mathbf{E}^*$ such that $\mathbf{F} \subset \ker(\varphi)$, and

$$\varphi(x_0) \neq 0.$$

Hence, $x_0 \notin \mathbf{W}$. We conclude that $\mathbf{W} \subset \overline{\mathbf{F}}$ as well. □

Corollary 3.6. *Let* $\mathbf{E}$ *be an normed linear space, and* $\mathbf{F} < \mathbf{E}$ *a subspace. Then,* $\mathbf{F}$ *is dense in* $\mathbf{E}$ *if and only if* $\mathbf{F}^{\perp} = \{0\}$.

(End of Day 26)

4. Separability and Reflexivity

Let us now return to an unresolved question from the first section of this chapter. Recall that for each $y = (y_n) \in \ell^1$, the map $\varphi_y : \ell^\infty \to \mathbb{K}$, given by $(x_k) \mapsto \sum_{n=1}^{\infty} x_n y_n$ is a bounded linear functional. Furthermore, the map

$$\Delta : \ell^1 \to (\ell^\infty)^*$$

given by $\Delta(y) := \varphi_y$, is an isometry by Proposition 1.2. Similarly, we have an isometry

$$\Delta : L^1[a,b] \to (L^\infty[a,b])^*.$$

We now show that these maps may not be surjective.

Proposition 4.1. *Let* $\mathbf{E}$ *be an normed linear space. If* $\mathbf{E}^*$ *is separable, then* $\mathbf{E}$ *is separable.*

Proof. If $\mathbf{E}^*$ is separable, then it is easy to see that the unit sphere

$$S := \{\varphi \in \mathbf{E}^* : \|\varphi\| = 1\}$$

is separable. We choose a countable set $\{\varphi_1, \varphi_2, \ldots\} \subset S$ that is dense in S. For each $n \in \mathbb{N}$, $\|\varphi_n\| = 1$. Therefore, there exist vectors $x_n \in \mathbf{E}$ such that $\|x_n\| = 1$ and

$$|\varphi_n(x_n)| > 1/2.$$

Let $\mathbf{F} = \operatorname{span}\{x_n : n \geq 1\}$. If we set $\mathbb{K}_0 = \mathbb{Q}$ or $\mathbb{Q} \times \mathbb{Q}$ according as $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, then

$$D := \operatorname{span}_{\mathbb{K}_0}\{x_n : n \geq 1\}$$

is countable, and dense in $\mathbf{F}$. Hence, $\mathbf{F}$ is separable. We wish to prove that $\overline{\mathbf{F}} = \mathbf{E}$, and in order to do that, we would like to appeal to Corollary 3.6. Suppose $\varphi \in \mathbf{F}^{\perp}$ is a non-zero bounded linear operator, then we may assume that $\|\varphi\| = 1$. Since $\varphi(x_n) = 0$ for all $n \in \mathbb{N}$, it follows that

$$\|\varphi - \varphi_n\| \geq |\varphi(x_n) - \varphi_n(x_n)| > 1/2$$

for all $n \in \mathbb{N}$. This contradicts the fact that $\{\varphi_n : n \in \mathbb{N}\}$ is dense in S. Thus, $\mathbf{F}^{\perp} = \{0\}$, and $\mathbf{F}$ is dense in $\mathbf{E}$ by Corollary 3.6. We conclude that $\mathbf{E}$ is separable. $\square$

Corollary 4.2. *The maps*

$$\begin{aligned} &\Delta : \ell^1 \to (\ell^\infty)^*, \text{ and} \\ &\Delta : L^1[a,b] \to (L^\infty[a,b])^* \end{aligned}$$

are not surjective

Proof. If Δ were an isomorphism, then $(\ell^\infty)^*$ would be isomorphic to ℓ^1. However, ℓ^1 is separable, so this would imply ℓ^∞ was separable, which contradicts Example 3.16. Similarly, $(L^\infty[a,b])^* \not\cong L^1[a,b]$. $\square$

Let c denote the vector space of all convergent sequences in $\mathbb{K}$ (see Homework 2.5), thought of as a closed subspace of ℓ^∞. Define $\varphi : c \to \mathbb{K}$ by

$$\varphi((x_j)) := \lim_{n\to\infty} x_n$$

This is well-defined, bounded, and linear. Therefore, by the Hahn-Banach Theorem, we may extend this to a bounded linear functional on ℓ^∞.

Example 4.3.

(i) Let $\psi \in (\ell^\infty)^*$ be a linear functional such that

$$\psi((x_n)) = \lim_{n\to\infty} x_n$$

whenever $(x_n) \in c$. We claim that $\psi \notin \Delta(\ell^1)$: Suppose there exists $y = (y_j)$ such that $\psi = \Delta(y) = \varphi_y$, then

$$\lim_{i\to\infty} x_i = \sum_{n=1}^{\infty} x_n y_n$$

for all $(x_j) \in c$. In particular, for $j \in \mathbb{N}$ fixed, consider $e_j \in \ell^\infty$ (as defined in Example 1.4). Then, $e_j \in c_0 \subset c$, so

$$y_j = \psi(e_j) = 0.$$

Thus, $y = 0$. But if this were true, it would imply that $\psi = 0$. However, $\psi \neq 0$ since $\psi(\mathbf{1}) = 1$. Therefore, there is no $y \in \ell^1$ such that $\varphi = \Delta(y)$.

(ii) Consider $C[a,b] \subset L^\infty[a,b]$ and define $\varphi : C[a,b] \to \mathbb{K}$ by

$$\varphi(f) := f(b).$$

Then, φ is a bounded linear functional, so by the Hahn-Banach Theorem, there exists a bounded linear functional $\psi : L^\infty[a,b] \to \mathbb{K}$ such that

$$\psi(f) = f(b)$$

for all $f \in C[a,b]$. We claim that $\psi \notin \Delta(L^1[a,b])$. Suppose there exists $g \in L^1[a,b]$ such that $\psi = \varphi_g$, then g would satisfy

$$f(b) = \int_a^b f(t)g(t)dt$$

for all $f \in C[a,b]$. For $n \in \mathbb{N}$, let $f_n \in C[a,b]$ be given by

$$f_n(t) = \begin{cases} 0 & : \text{if } a \le t \le b - 1/n \\ 1 & : \text{if } t = b \\ \text{linear} & : \text{otherwise} \end{cases}$$

Then, for any $t \in [a, b)$,
$$\lim_{n\to\infty} f_n(t) = 0$$
Hence, $f_n \to 0$ pointwise a.e. Since $f_n \in L^\infty[a, b]$, and $g \in L^1[a, b]$, we may apply Dominated Convergence Theorem to conclude that
$$\lim_{n\to\infty} \psi(f_n) = \lim_{n\to\infty} \int_a^b f_n(t)g(t)dt = 0.$$
However, $\psi(f_n) = f_n(b) = 1$ for all $n \in \mathbb{N}$. This contradiction shows that $\psi \neq \varphi_g$ for any $g \in L^1[a, b]$.

Definition 4.4. Let $\mathbf{E}$ be an normed linear space.

(i) The <u>double dual</u> of $\mathbf{E}$ is the dual of $\mathbf{E}^*$, and is denoted by $\mathbf{E}^{**} := (\mathbf{E}^*)^*$.

(ii) For each $x \in \mathbf{E}$, define $\widehat{x} : \mathbf{E}^* \to \mathbb{K}$ by
$$\widehat{x}(\varphi) := \varphi(x).$$
Note that $\widehat{x}$ is a linear functional on $\mathbf{E}^*$, and
$$\|\widehat{x}\| = \sup\{|\varphi(x)| : \varphi \in \mathbf{E}^*, \|\varphi\| = 1\} = \|x\| \tag{IV.2}$$
by Corollary 2.8. Therefore, $\widehat{x} \in \mathbf{E}^{**}$.

(iii) Define $J : \mathbf{E} \to \mathbf{E}^{**}$ by
$$J(x) := \widehat{x}.$$
Then, J is a linear transformation, which is isometric by Equation IV.2.

(iv) $\mathbf{E}$ is said to be <u>reflexive</u> if J is an isomorphism from $\mathbf{E}$ to $\mathbf{E}^{**}$.

Example 4.5.

(i) If $\mathbf{E}$ is finite dimensional, then it is reflexive.

Proof. If $\mathbf{E}$ is finite dimensional, then
$$\dim(\mathbf{E}) = \dim(\mathbf{E}^*) = \dim(\mathbf{E}^{**}).$$
Since J is injective, it must be surjective. □

(End of Day 27)

(ii) Every Hilbert space is reflexive.

Proof. Given the Riesz Representation Theorem, this proof is almost tautological. Let $\mathbf{H}$ be a Hilbert space, and define $\Delta : \mathbf{H} \to \mathbf{H}^*$ by $\Delta(y) := \varphi_y$ where
$$\varphi_y(x) := \langle x, y\rangle.$$

Then Δ is an conjugate-linear isomorphism of normed linear spaces. In particular, $\mathbf{H}^*$ is a Hilbert space under the inner product

$$(\varphi_y, \varphi_z) := \langle z, y \rangle.$$

Hence, if $T \in \mathbf{H}^{**}$, then, by the Riesz Representation Theorem applied to $\mathbf{H}^*$, there exists $\varphi \in \mathbf{H}^*$ such that

$$T(\psi) = (\psi, \varphi)$$

for all $\psi \in \mathbf{H}^*$. By the Riesz Representation Theorem, there exists $y \in \mathbf{H}$ such that $\varphi = \varphi_y$. For any $z \in \mathbf{H}$, taking $\psi = \varphi_z$ in the above equation gives

$$T(\varphi_z) = (\varphi_z, \varphi_y) = \langle y, z \rangle.$$

Now consider $\widehat{y} \in \mathbf{H}^{**}$, and observe that

$$\widehat{y}(\varphi_z) = \varphi_z(y) = \langle y, z \rangle.$$

Therefore, $T = \widehat{y}$ and so J is surjective. □

(iii) For $1 < p < \infty$, ℓ^p is reflexive.

Proof. Let $q \in (1, \infty)$ such that $1/p + 1/q = 1$, and let $\Delta_p : \ell^q \to (\ell^p)^*$ be the map from Definition 1.1. Similarly, let $\Delta_q : \ell^p \to (\ell^q)^*$ denote the corresponding map obtained by letting ℓ^p act on ℓ^q. Then, by Theorem 1.3, both maps induce isomorphisms

$$\Delta_p : \ell^q \cong (\ell^p)^* \text{ and } \Delta_q : \ell^p \cong (\ell^q)^*.$$

Now, suppose $T \in (\ell^p)^{**}$, then $T : (\ell^p)^* \to \mathbb{K}$ is bounded and linear. Hence, $T \circ \Delta_p : \ell^q \to \mathbb{K}$ is bounded and linear. Since Δ_q is surjective, there exists $x \in \ell^p$ such that

$$\Delta_q(x) = T \circ \Delta_p.$$

Hence, for any $y \in \ell^q$, we have

$$\sum_{n=1}^{\infty} x_n y_n = \Delta_q(x)(y) = T(\Delta(y)).$$

However, we observe that

$$\widehat{x}(\Delta(y)) = \widehat{x}(\varphi_y) = \varphi_y(x) = \sum_{n=1}^{\infty} x_n y_n.$$

Since Δ is surjective, we conclude that $\widehat{x} = T$. Therefore, $J : \ell^p \to (\ell^p)^{**}$ is surjective. □

(iv) A similar argument shows that $L^p[a, b]$ is reflexive, provided $1 < p < \infty$.

Proposition 4.6. *Let* $\mathbf{E}$ *be a reflexive space, then* $\mathbf{E}$ *is separable if and only if* $\mathbf{E}^*$ *is separable.*

Proof. If $\mathbf{E}^*$ is separable, then $\mathbf{E}$ is separable by Proposition 4.1. Conversely, if $\mathbf{E}$ is separable and reflexive then $\mathbf{E}^{**} = (\mathbf{E}^*)^*$ is separable. Hence, $\mathbf{E}^*$ is separable by Proposition 4.1. $\square$

Proposition 4.7. *Let $\mathbf{E}$ be a reflexive space and $\mathbf{F} < \mathbf{E}$ a closed subspace, then $\mathbf{F}$ is reflexive.*

Proof. We have a isometric maps $J_{\mathbf{F}} : \mathbf{F} \to \mathbf{F}^{**}$ and $J_{\mathbf{E}} : \mathbf{E} \to \mathbf{E}^{**}$. Assuming $J_{\mathbf{E}}$ is surjective, we want to show that $J_{\mathbf{F}}$ is surjective. So suppose $T \in \mathbf{F}^{**}$, then we wish to show that there exists $x \in \mathbf{F}$ such that $T = J_{\mathbf{F}}(x)$. Consider $T : \mathbf{F}^* \to \mathbb{K}$ as a bounded linear functional, and define $S : \mathbf{E}^* \to \mathbb{K}$ by

$$S(\varphi) := T(\varphi|_{\mathbf{F}}).$$

Then, S is clearly a linear functional, and

$$|S(\varphi)| = |T(\varphi|_{\mathbf{F}})| \leq \|T\|\|\varphi|_{\mathbf{F}}\| \leq \|T\|\|\varphi\|$$

Hence, S is a bounded, and therefore in $\mathbf{E}^{**}$. Since $\mathbf{E}$ is reflexive, there exists $x \in \mathbf{E}$ such that

$$S = J_{\mathbf{E}}(x).$$

We claim that $x \in \mathbf{F}$, and that $T = J_{\mathbf{F}}(x)$. Suppose $x \notin \mathbf{F}$, then, by Proposition 3.3, there exists $\varphi \in \mathbf{E}^*$ such that $\varphi|_{\mathbf{F}}= 0$, and $\varphi(x) = 1$. However, this would imply that

$$1 = \varphi(x) = J_{\mathbf{E}}(x)(\varphi) = S(\varphi) = T(\varphi|_{\mathbf{F}}) = T(0) = 0.$$

This is a contradiction, and therefore it must happen that $x \in \mathbf{F}$. Now, for any $\varphi \in \mathbf{F}^*$, we choose a Hahn-Banach extension $\psi \in \mathbf{E}^*$ of φ. Then, $\psi(x) = \varphi(x)$ since $x \in \mathbf{F}$. Therefore,

$$T(\varphi) = T(\psi|_{\mathbf{F}}) = S(\psi) = J_{\mathbf{E}}(x)(\psi) = \psi(x) = \varphi(x) = J_{\mathbf{F}}(x)(\varphi).$$

We conclude that $J_{\mathbf{F}}$ is surjective. $\square$

Example 4.8.

(i) Any reflexive space is necessarily complete, because the dual space of any normed linear space is complete. Therefore, $(c_{00}, \|\cdot\|_p)$ is not reflexive for any $1 \leq p \leq \infty$ (see Proposition 3.4 and Mid-Sem Exam Q1).

(ii) For the same reason, $(C[a,b], \|\cdot\|_p)$ is not reflexive, if $1 \leq p < \infty$ (by Proposition 3.12).

(iii) ℓ^1 is separable, but its dual space ℓ^∞ is not separable (by Example 3.16). Therefore, ℓ^1 is not reflexive. Similarly, $L^1[a,b]$ is not reflexive either.

(iv) $C[a,b]$ (equipped with the supremum norm) is not reflexive because $C[a,b]^*$ is not separable (Homework 6.1), while $C[a,b]$ itself is separable.

(v) $L^\infty[a,b]$ is not reflexive since it has a closed subspace $C[a,b]$ that is not reflexive.

(vi) From Homework 6.3, there is an isometric isomorphism $\ell^1 \to (c_0)^*$. Hence,

$$(c_0)^{**} \cong \ell^\infty.$$

Now, c_{00} is dense in c_0 (by Mid-Sem Exam Q1), which tells us that c_0 is separable (by Remark 3.15). Since ℓ^∞ is not separable, it follows that c_0 is not reflexive.

(vii) Since c_0 is a closed subspace of ℓ^∞, it follows from Proposition 4.7 that ℓ^∞ is not reflexive.

(End of Day 28)

V. Operators on Banach Spaces

1. The Principle of Uniform Boundedness

Theorem 1.1 (Baire Category Theorem (Baire, 1899)). *Let (X, d) be a complete metric space and $\{V_n\}_{n=1}^{\infty}$ be a countable collection of open, dense subsets of X. Then,*

$$G := \bigcap_{n=1}^{\infty} V_n$$

is dense in X.

Proof. As is usual, we write $B(x, r) := \{y \in X : d(y, x) < r\}$ and $B[x, r] := \{y \in X : d(y, x) \leq r\}$.

Let U be a non-empty open set. We wish to prove that $G \cap U \neq \emptyset$. We do this by inductively constructing a sequence whose limit point lies in this intersection. To begin with, $U \cap V_1$ is non-empty, and open. Hence, there exists $x_1 \in U \cap V_1$ and $r_1 > 0$ such that $B(x_1, r_1) \subset U \cap V_1$. By shrinking r_1 if need be, we may assume that $r_1 < 1$, and

$$\overline{B(x_1, r_1)} = B[x_1, r_1] \subset U \cap V_1.$$

Since V_2 is dense in X, $B(x_1, r_1) \cap V_2 \neq \emptyset$. Once again, there exists $x_2 \in X$, and $r_2 > 0$ such that $r_2 < \frac{1}{2}$, and

$$B[x_2, r_2] \subset B(x_1, r_1) \cap V_2.$$

Observe that $B[x_2, r_2] \subset U \cap (V_1 \cap V_2)$. Thus proceeding, we obtain a sequence $(x_n) \subset X$ and $(r_n) \subset \mathbb{R}_+$ such that

(i) $B[x_n, r_n] \subset B(x_{n-1}, r_{n-1})$,

(ii) $r_n < 1/n$, and

(iii) $B[x_n, r_n] \subset U \bigcap \left[\bigcap_{i=1}^{n-1} V_i\right]$.

Now, for $m \in \mathbb{N}$ fixed, and all $n > m$, we have $d(x_n, x_m) < r_m < 1/m$, and hence (x_n) is Cauchy. Since X is complete, there exists $x_0 \in X$ such that $x_n \to x_0$. We claim that $x_0 \in U \cap G$. To see this, note that, for all $n \geq m$,

$$x_n \in B[x_m, r_m].$$

Since this set is closed, it follows that

$$x_0 \in B[x_m, r_m] \subset U \bigcap \left[\cap_{i=1}^{m-1} V_i\right].$$

This is true for all $m \in \mathbb{N}$ and hence $x_0 \in U \cap G$. $\square$

Let (X,d) be a metric space and $A \subset X$. We say that A is nowhere dense if $\overline{A}$ has empty interior (equivalently, if $X \setminus \overline{A}$ is a dense open set). Note that A is nowhere dense if and only if $\overline{A}$ is nowhere dense. The Baire Category Theorem may now be restated as

Corollary 1.2. *A complete metric space cannot be written as a countable union of nowhere dense sets.*

Lemma 1.3. *Every closed, proper subspace of a normed linear space is nowhere dense.*

Corollary 1.4. *If $\mathbf{E}$ is an infinite dimensional Banach space, then $\mathbf{E}$ cannot have a countable Hamel basis.*

Proof. Suppose $A := \{e_n : n \in \mathbb{N}\}$ is a countable subset of $\mathbf{E}$, then define $\mathbf{F}_n := \text{span}\{e_1, e_2, \ldots, e_n\}$. Since $\mathbf{F}_n$ is finite dimensional, $\mathbf{F}_n$ is both closed, and a proper subspace of $\mathbf{E}$. Therefore, it is nowhere dense by Lemma 1.3. By Corollary 1.2,

$$\mathbf{E} \neq \bigcup_{n=1}^{\infty} \mathbf{F}_n.$$

In particular, $\mathbf{E}$ contains at least one element that cannot be expressed as a finite linear combination of elements in A. $\square$

Example 1.5.

(i) Let $\mathbf{H}$ be an infinite dimensional separable Hilbert space, and Λ an orthonormal basis for $\mathbf{H}$, then Λ cannot be a Hamel basis for $\mathbf{H}$ (see **??**).

(ii) There is no norm on c_{00} that makes it a Banach space.

Theorem 1.6 (Principle of Uniform Boundedness (Hahn, 1922, Banach and Steinhaus, 1927))**.** *Let $\mathbf{E}$ be a Banach space, and $\mathbf{F}$ be any normed linear space. Let $\mathcal{G} \subset \mathcal{B}(\mathbf{E}, \mathbf{F})$ be a collection of bounded linear operators such that*

$$\sup_{T \in \mathcal{G}} \|T(x)\| < \infty$$

for each $x \in \mathbf{E}$. Then, $\sup_{T \in \mathcal{G}} \|T\| < \infty$.

Proof. For $n \in \mathbb{N}$, define

$$B_n := \bigcap_{T \in \mathcal{G}} \{x \in \mathbf{E} : \|T(x)\| \leq n\}.$$

Then, each B_n is closed, because it is the intersection of a family of closed sets. Furthermore, by hypothesis,

$$\mathbf{E} = \bigcup_{n=1}^{\infty} B_n.$$

By the Baire Category Theorem, there exists $N \in \mathbb{N}$ such that B_N has non-empty interior. Therefore, there exists $x_0 \in \mathbf{E}$, and $r > 0$ such that $B(x_0, r_0) \subset B_N$. Now, we apply the scaling trick: For any $x \in \mathbf{E}$, set

$$z := \frac{r_0 x}{2\|x\|} + x_0 \in B(x_0, r_0).$$

Then, for any $T \in \mathcal{G}$, $\|T(z)\| \leq N$. Unwrapping this, we get

$$\|T(x)\| = \frac{2\|x\|}{r_0}\|T(z) - T(x_0)\| \leq \frac{2\|x\|}{r_0}(N + \|T(x_0)\|).$$

Hence, if $M := \frac{2}{r_0}(N + \|T(x_0)\|)$, then for any $x \in \mathbf{E}$ and any $T \in \mathcal{G}$, we have $\|T(x)\| \leq M\|x\|$. Thus, $\sup\{\|T\| : T \in \mathcal{G}\} \leq M < \infty$. $\square$

Theorem 1.7 (Banach-Steinhaus Theorem, 1927)**.** *Let* $\mathbf{E}$ *be a Banach space, and* $\mathbf{F}$ *be any normed linear space. Suppose* $(T_n)_{n=1}^\infty \subset \mathcal{B}(\mathbf{E}, \mathbf{F})$ *is a sequence of bounded operators such that, for each* $x \in \mathbf{E}$*, the sequence* $(T_n(x))_{n=1}^\infty$ *is convergent in* $\mathbf{F}$*. Then, the map* $T : \mathbf{E} \to \mathbf{F}$ *defined by*

$$T(x) := \lim_{n\to\infty} T_n(x)$$

is a bounded linear map. Furthermore, $\|T\| \leq \liminf_{n\to\infty} \|T_n\|$.

Proof. It is easy to see that the map T defined above is linear; we need only prove that it is bounded. For each $x \in \mathbf{E}$, $(T_n(x))_{n=1}^\infty$ is convergent, and hence bounded. By the Principle of Uniform Boundedness, there exists $M > 0$ such that $\|T_n\| \leq M$ for all $n \in \mathbb{N}$. Therefore, for each $x \in \mathbf{E}$,

$$\|T(x)\| = \lim_{n\to\infty} \|T_n(x)\| \leq M\|x\|.$$

Hence, $T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$ and $\|T\| \leq M$. The norm inequality is a refinement of this argument, and is left as an exercise. $\square$

(End of Day 29)

2. The Open Mapping and Closed Graph Theorems

For a normed linear space $\mathbf{E}$, subsets $A, B \subset \mathbf{E}$ and a scalar $\lambda \in \mathbb{K}$, we define $A + B := \{a + b : a \in A, b \in B\}$, and $\lambda A := \{\lambda a : a \in A\}$. We begin with a small observation.

Lemma 2.1. *If* A *is convex, then* $A + A = 2A$.

Proof. It is clear that $2A \subset A + A$. For the reverse inclusion, let $x, y \in A$. Then, since A is convex, $\frac{x+y}{2} \in A$. Therefore,

$$x + y = 2\frac{x+y}{2} \in 2A.$$

This proves the reverse inclusion as well. $\square$

Lemma 2.2. *Let* $\mathbf{E}$ *be an normed linear space,* $\mathbf{F}$ *be a Banach space and* $T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$ *be a surjective, bounded linear map. Then, for every* $r > 0$*, there exists* $s > 0$ *such that*

$$B_{\mathbf{F}}(0, s) \subset \overline{T(B_{\mathbf{E}}(0, r))}$$

Proof. Fix $r > 0$. For $n \in \mathbb{N}$, define

$$B_n := n\overline{T(B_{\mathbf{E}}(0, r))}.$$

Then, B_n is closed, and

$$\mathbf{F} = T(\mathbf{E}) = \bigcup_{n=1}^{\infty} B_n.$$

Since $\mathbf{F}$ is complete, by the Baire Category Theorem (Corollary 1.2), there exists $N \in \mathbb{N}$ such that B_N has non-empty interior. Since the map $y \mapsto Ny$ is a homeomorphism of $\mathbf{E}$, B_1 must contain an open set. Thus, there exists $y_0 \in \mathbf{F}$, and $s_0 > 0$ such that

$$B_{\mathbf{F}}(y_0, s_0) \subset B_1.$$

In particular, $y_0 \in B_1$ and so $-y_0 \in B_1$ as well. Now, for any $y \in B_{\mathbf{F}}(0, s_0)$, we have

$$y = (y + y_0) + (-y_0) \in B_{\mathbf{F}}(y_0, s_0) + B_1 \subset B_1 + B_1$$

Since B_1 is convex, Lemma 2.1 implies that $B_{\mathbf{F}}(0, s_0) \subset 2B_1$. Therefore, $s = s_0/2$ works. ☐

Lemma 2.3. *Let* $\mathbf{E}$ *and* $\mathbf{F}$ *be Banach spaces, and* $T \in \mathcal{B}(\mathbf{E}, \mathbf{F})$ *be a surjective, bounded linear map. Then, for every* $r > 0$*, there exists* $s > 0$ *such that*

$$B_{\mathbf{F}}(0, s) \subset T(B_{\mathbf{E}}(0, r)).$$

Proof. For each natural number $n \in \mathbb{N}$, set $r_n := r/2^{n+1}$. By Lemma 2.2, there exists $s_n > 0$ such that

$$B_{\mathbf{F}}(0, s_n) \subset \overline{T(B_{\mathbf{E}}(0, r_n))}. \tag{V.1}$$

Furthermore, we may choose s_n inductively in such a way that $\lim_{n\to\infty} s_n = 0$. We claim that

$$B_{\mathbf{F}}(0, s_1) \subset T(B_{\mathbf{E}}(0, r)).$$

Fix $y \in B_{\mathbf{F}}(0, s_1)$, then $y_1 \in \overline{T(B_{\mathbf{E}}(0, r_1))}$ by Equation V.1. Therefore, there exists $x_1 \in B_{\mathbf{E}}(0, r_1)$ such that

$$\|T(x_1) - y\| < s_2.$$

Now, $T(x_1) - y \in B_{\mathbf{F}}(0, s_2)$, and $B_{\mathbf{F}}(0, s_2) \subset \overline{T(B_{\mathbf{E}}(0, r_2))}$. Therefore, there exists $x_2 \in B_{\mathbf{E}}(0, r_2)$ such that

$$\|T(x_2) + T(x_1) - y\| < s_3.$$

Thus proceeding, we obtain a sequence $(x_n) \subset \mathbf{E}$ such that, for each $n \in \mathbb{N}$, $x_n \in B_{\mathbf{E}}(0, r_n)$, and

$$\|T(x_n) + T(x_{n-1}) + \ldots + T(x_1) - y\| < s_{n+1} \tag{V.2}$$

Now, note that
$$\sum_{n=1}^{\infty} \|x_n\| \le \sum_{n=1}^{\infty} \frac{r}{2^{n+1}} = \frac{r}{2} < \infty.$$
Since $\mathbf{E}$ is a Banach space (by Proposition 3.8), there exists $z \in \mathbf{E}$ such that
$$z = \sum_{n=1}^{\infty} x_n.$$

By Equation V.2, we have that
$$\left\| T\left(\sum_{i=1}^{n} x_i\right) - y \right\| < s_{n+1}.$$

By assumption, $\lim_{n\to\infty} s_n = 0$. Since T is continuous, it follows that $T(z) = y$. Note that $\|z\| \le \frac{r}{2} < r$, so $y \in T(B_{\mathbf{E}}(0,r))$. This proves that $B_{\mathbf{F}}(0,s_1) \subset T(B_{\mathbf{E}}(0,r))$. □

Theorem 2.4 (Open Mapping Theorem (Banach, 1932))**.** *Let* $\mathbf{E}$ *and* $\mathbf{F}$ *be Banach spaces, and* $T \in \mathcal{B}(\mathbf{E},\mathbf{F})$ *be a surjective, bounded linear map. Then* T *is an open map.*

Proof. Let $V \subset \mathbf{E}$ be an open set. We wish to prove that $T(V)$ is open. So let $y \in T(V)$, and write $y = T(x)$ for some $x \in V$. Since V is open, there exists $r > 0$ such that $B_{\mathbf{E}}(x,r) \subset V$. By Lemma 2.3, there exists $s > 0$ such that
$$B_{\mathbf{F}}(0,s) \subset T(B_{\mathbf{E}}(0,r)).$$

Hence,
$$B_{\mathbf{F}}(y,s) = y + B_{\mathbf{F}}(0,s) \subset y + T(B_{\mathbf{E}}(0,r)) = T(x + B_{\mathbf{E}}(0,r)) = T(B_{\mathbf{E}}(x,r)) \subset T(V).$$

Thus, for each $y \in T(V)$, there exists $s > 0$ such that $B_{\mathbf{F}}(y,s) \subset T(V)$. In other words,$T(V)$ is an open set. □

Theorem 2.5 (Bounded Inverse Theorem)**.** *Let* $\mathbf{E}$ *and* $\mathbf{F}$ *be Banach spaces. If* $T \in \mathcal{B}(\mathbf{E},\mathbf{F})$ *is bijective, then* T^{-1} *is continuous.*

Proof. Since T is bijective, T^{-1} is a well-defined linear operator. That T^{-1} is continuous follows from the Open Mapping Theorem. □

Corollary 2.6 (First Isomorphism Theorem)**.** *Let* $\mathbf{E}$ *and* $\mathbf{F}$ *be Banach spaces, and* $T \in \mathcal{B}(\mathbf{E},\mathbf{F})$ *be a surjective, bounded linear operator. Then, the map* $\widehat{T} : \mathbf{E}/\ker(T) \to \mathbf{F}$ *given by*
$$x + \ker(T) \mapsto T(x)$$

is a topological isomorphism.

(End of Day 30)

Example 2.7. The examples given below show that the completeness assumptions of the Open Mapping Theorem cannot be avoided:

(i) In Question 3 of the Mid-Sem Exam, we constructed an example of an injective map $T : \ell^2 \to \ell^2$, whose inverse (defined on $\text{Range}(T)$) is not continuous. The reason for this, of course, is that $\text{Range}(T)$ is not complete.

(ii) Let $\mathbf{E} := (C[0,1], \|\cdot\|_\infty)$ and $\mathbf{F}$ be the subspace

$$\mathbf{F} := \{f \in C^1[0,1] : f(0) = 0\},$$

equipped with the supremum norm. Note that $\mathbf{F}$ is not closed (and hence not complete). Define $T : \mathbf{E} \to \mathbf{F}$ by

$$T(f)(x) := \int_0^x f(t)dt$$

Then T is well-defined, bounded, and bijective. However, $T^{-1} : \mathbf{F} \to \mathbf{E}$ is the map

$$f \mapsto f',$$

which is not bounded (See Question 2 on Quiz 1).

(iii) If $\iota : (c_{00}, \|\cdot\|_1) \to (c_{00}, \|\cdot\|_\infty)$ to be the identity map, then ι is clearly bijective and bounded. However, the two norms are not equivalent (see Example 4.3). Therefore, the inverse map is not continuous.

Definition 2.8. Let X and Y be topological spaces and $f : X \to Y$ be a function. The graph of f is the set

$$G(f) := \{(x, f(x)) : x \in X\} \subset X \times Y$$

Lemma 2.9. *Let X and Y be two metric spaces, and $f : X \to Y$ be a continuous map. Then, $G(f)$ is closed in $X \times Y$ (where $X \times Y$ is equipped with a product metric).*

Proof. Choose a sequence $(x_n, f(x_n)) \in G(f)$ such that $(x_n, f(x_n)) \to (x, y)$ in $X \times Y$. Then, by definition of the product topology, $x_n \to x$ in X, and $f(x_n) \to y$ in Y. Since f is continuous, $f(x_n) \to f(x)$ as well, and since Y is Hausdorff, it follows that

$$y = f(x).$$

Therefore, $(x, y) \in G(f)$. □

It is quite easy to construct a discontinuous (non-linear) function $f : \mathbb{R} \to \mathbb{R}$ whose graph is closed. However, the Closed Graph Theorem does provide a converse of Lemma 2.9 in the context of linear maps.

Definition 2.10. Let $\mathbf{E}$ and $\mathbf{F}$ be normed linear spaces, and $T : \mathbf{E} \to \mathbf{F}$ be a linear operator (not necessarily bounded). The function $\|\cdot\|_G : \mathbf{E} \to \mathbb{R}_+$

$$\|x\|_G := \|x\|_{\mathbf{E}} + \|T(x)\|_{\mathbf{F}}$$

defines a norm on $\mathbf{E}$, and is called the graph norm on $\mathbf{E}$ with respect to T. Note that

(i) $\|x\|_{\mathbf{E}} \leq \|x\|_G$ for all $x \in \mathbf{E}$, and

(ii) $\|\cdot\|_{\mathbf{E}} \sim \|\cdot\|_G$ if and only if T is bounded.

Lemma 2.11. *Let* $\mathbf{E}$ *be a vector space that is a Banach space with respect to two norms* $\|\cdot\|_1$ *and* $\|\cdot\|_2$. *Suppose that there exists* $c > 0$ *such that*

$$\|x\|_1 \leq c\|x\|_2$$

for all $x \in \mathbf{E}$. *Then,* $\|\cdot\|_1 \sim \|\cdot\|_2$.

There is an important caveat to Lemma 2.11: The space $\mathbf{E}$ must be complete with respect to each norm independently. For instance, if we take $\mathbf{E} := C[0,1]$, and equip it with $\|\cdot\|_\infty$ and $\|\cdot\|_1$, then $(C[0,1], \|\cdot\|_\infty)$ is complete, and $\|f\|_1 \leq \|f\|_\infty$ for all $f \in \mathbf{E}$. However, the two norms are not equivalent.

Theorem 2.12 (Closed Graph Theorem (Banach, 1932))**.** *Let* $\mathbf{E}$ *and* $\mathbf{F}$ *be Banach spaces, and* $T : \mathbf{E} \to \mathbf{F}$ *be a linear operator. If* $G(T)$ *is closed in* $\mathbf{E} \times \mathbf{F}$, *then* T *is continuous.*

Proof. Consider the Graph norm $\|\cdot\|_G$ on $\mathbf{E}$ with respect to T, as above. By Lemma 2.11, it suffices to show that $(\mathbf{E}, \|\cdot\|_G)$ is a Banach space. So, suppose $(x_n) \subset \mathbf{E}$ is Cauchy with respect to $\|\cdot\|_G$. Then, $\|x_n - x_m\|_{\mathbf{E}} \leq \|x_n - x_m\|_G$, and

$$\|T(x_n) - T(x_m)\|_{\mathbf{F}} \leq \|x_n - x_m\|_G.$$

Hence, (x_n) is Cauchy in $\mathbf{E}$, and $(T(x_n))$ is Cauchy in $\mathbf{F}$. Since both spaces are complete, there exist $x \in \mathbf{E}$ and $y \in \mathbf{F}$ such that $\lim_{n\to\infty} x_n = x$, and

$$\lim_{n\to\infty} T(x_n) = y.$$

Since $G(T)$ is closed, this implies that $T(x) = y$. We conclude that $\|x_n - x\|_G \to 0$. Hence, $(\mathbf{E}, \|\cdot\|_G)$ is a Banach space, and Lemma 2.11 implies that

$$\|\cdot\|_G \sim \|\cdot\|_{\mathbf{E}}.$$

Therefore, T is bounded. □

VI. Weak Topologies

1. Weak Convergence

Definition 1.1. Let $\mathbf{E}$ be an normed linear space. A sequence $(x_n) \subset \mathbf{E}$ is said to converge weakly to $x \in \mathbf{E}$ if

$$\varphi(x_n) \to \varphi(x)$$

for all $\varphi \in \mathbf{E}^*$. If this happens, we write $x_n \xrightarrow{w} x$.

In this chapter, if $x_n \to x$ in the norm, then we say that $x_n \to x$ strongly, and we write $x_n \xrightarrow{s} x$.

Example 1.2.

(i) If $x_n \xrightarrow{w} x$ and $x_n \xrightarrow{w} y$, then $x = y$. This is because $\mathbf{E}^*$ separates points of $\mathbf{E}$ (Corollary 2.9). Furthermore, if $x_n \xrightarrow{w} x$ then any subsequence (x_{n_k}) of (x_n) also converges weakly to x.

(ii) If $x_n \xrightarrow{s} x$, then $x \xrightarrow{w} x$. This is because every element of $\mathbf{E}^*$ is continuous with respect to the norm topology.

(iii) If $\mathbf{E}$ is a finite dimensional normed linear space and $x_n \xrightarrow{w} x$, then $x_n \xrightarrow{s} x$.

Proof. By Corollary 4.8, we may assume without loss of generality that $\mathbf{E} = (\mathbb{K}^m, \|\cdot\|_1)$ for some $m \in \mathbb{N}$. For each $1 \leq i \leq m$, the projection maps $\pi_i : \mathbf{E} \to \mathbb{K}$ are bounded linear functionals. Hence, $\pi_i(x_n) \to \pi_i(x)$. Then,

$$\lim_{n\to\infty} \|x_n - x\|_1 = \lim_{n\to\infty} \sum_{i=1}^{m} |\pi_i(x_n) - \pi_i(x)| = 0.$$

Hence, $x_n \xrightarrow{s} x$. $\square$

(iv) If $\mathbf{H}$ is an infinite dimensional Hilbert space and $(e_n) \subset \mathbf{H}$ an orthonormal sequence, then for any $x \in \mathbf{H}$,

$$\lim_{n\to\infty} \langle e_n, x\rangle = 0$$

by the Riemann-Lebesgue Lemma (Corollary 3.11). By the Riesz Representation Theorem, $e_n \xrightarrow{w} 0$.

However, we claim that (e_n) does not converge strongly to *any* point in $\mathbf{H}$. If $x \in \mathbf{H}$ were such that $e_n \xrightarrow{s} x$, then $e_n \xrightarrow{w} x$, whence $x = 0$. But $\|e_n\| = 1$ for all $n \in \mathbb{N}$, so the continuity of the norm (Remark 1.2) implies that $\|x\| = 1$. This contradiction shows that (e_n) is not strongly convergent.

Lemma 1.3. *Let* $\mathbf{E}$ *be an normed linear space and* $x_n \xrightarrow{w} x$*, then* (x_n) *is bounded, and*

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|.$$

Proof. Consider the map $J : \mathbf{E} \to \mathbf{E}^{**}$ given by $J(x) := \widehat{x}$, where $\widehat{x} : \mathbf{E}^* \to \mathbb{K}$ is given by

$$\widehat{x}(\varphi) := \varphi(x).$$

By hypothesis, $\widehat{x_n}(\varphi) \to \widehat{x}(\varphi)$ for all $\varphi \in \mathbf{E}^*$. By the Banach-Steinhaus Theorem applied to the Banach space $\mathbf{E}^*$, $(\|\widehat{x_n}\|)$ is a bounded sequence, and

$$\|\widehat{x}\| \leq \liminf \|\widehat{x_n}\|.$$

Now the result follows from the fact that J is an isometry. □

Proposition 1.4. *Let* $\mathbf{E}$ *be an normed linear space,* $(x_n) \subset \mathbf{E}$ *be a bounded sequence, and* $\mathcal{G} \subset \mathbf{E}^*$ *be such that* $\text{span}(\mathcal{G})$ *is a norm dense subset of* $\mathbf{E}^*$*. Suppose* $x \in \mathbf{E}$ *is a vector such that*

$$\varphi(x_n) \to \varphi(x)$$

for all $\varphi \in \mathcal{G}$*. Then,* $x_n \xrightarrow{w} x$

Proof. By assumption, $\varphi(x_n) \to \varphi(x)$ for all $\varphi \in \mathbf{F} := \text{span}(\mathcal{G})$. Now, if $\psi \in \mathbf{E}^*$, and $\epsilon > 0$, then there exists $\varphi \in \mathbf{F}$ such that $\|\psi - \varphi\| < \epsilon$. Then, we look to exploit the inequality

$$|\psi(x_n) - \psi(x)| \leq |\psi(x_n) - \varphi(x_n)| + |\varphi(x_n) - \varphi(x)| + |\varphi(x) - \psi(x)| \tag{VI.1}$$

By hypothesis, there exists $M > 0$ such that $\|x_n\| \leq M$ for all $n \in \mathbb{N}$. Also, there exists $N \in \mathbb{N}$ such that $|\varphi(x_n) - \varphi(x)| < \epsilon$ for all $n \geq N$. Plugging all this back in Equation VI.1, we get

$$|\psi(x_n) - \psi(x)| \leq M\epsilon + \epsilon + \epsilon\|x\|,$$

which holds for all $n \geq N$. Since this is true for any $\epsilon > 0$, $\psi(x_n) \to \psi(x)$. □

Corollary 1.5. *Let* $\mathbf{E} = c_0$ *or* ℓ^p *with* $1 < p < \infty$*, and let* $(x^n) \subset \mathbf{E}$ *be a bounded sequence such that*

$$x_j^n \to x_j$$

for each $j \in \mathbb{N}$*. Then* $x^n \xrightarrow{w} x$*.*

Proof. We first assume $\mathbf{E} = \ell^p$ for $1 < p < \infty$, since the other case is similar. For each $j \in \mathbb{N}$, the evaluation map $\varphi_j : \mathbf{E} \to \mathbb{K}$ is given by

$$\varphi_j((y_n)) := y_j,$$

and let $\mathcal{G} := \{\varphi_j : j \in \mathbb{N}\}$. By assumption, $\varphi(x^n) \to \varphi(x)$ for each $\varphi \in \mathcal{G}$, so we look to apply Proposition 1.4. Recall that we have an isomorphism $\Delta : \ell^q \to (\ell^p)^*$, where $1/p + 1/q = 1$. Under this isomorphism,

$$e_j \mapsto \varphi_j.$$

Since $1 < q < \infty$, c_{00} is dense in ℓ^q. Therefore, $\operatorname{span}(\mathcal{G}) = \Delta(c_{00})$ is dense in $\mathbf{E}^*$. The conclusion now follows from Proposition 1.4.

For the case of $\mathbf{E} = c_0$, the argument is identical, except we use the isomorphism $\Delta : \ell^1 \to (c_0)^*$ proved in Homework 6.3. $\square$

Example 1.6. Let $\mathbf{E} = \ell^\infty$ and

$$x^k := (\underbrace{1, 1, 1, \ldots, 1}_{k \text{ times}}, 0, 0, \ldots)$$

For $j \in \mathbb{N}$, consider the evaluation linear functional $\varphi_j \in \mathbf{E}^*$ given by $\varphi_j((y_n)) := y_j$. Then, $\lim_{k\to\infty} \varphi_j(x^k) = 1$. Therefore, if $x^k \xrightarrow{w} x$, then it follows that

$$x = (1, 1, 1, \ldots) = \mathbf{1}.$$

However, let $\psi \in (\ell^\infty)^*$ be a Banach limit (Example 4.3), so that

$$\psi((x_j)) = \lim_{n\to\infty} x_n$$

for all $(x_n) \in c$. Then, in particular, $\psi(x^k) = 0$ for all $n \in \mathbb{N}$. Since $\psi(\mathbf{1}) = 1$, it follows that (x^k) is not weakly convergent.

However, (x^k) is bounded in ℓ^∞. Thus, the conclusion of Corollary 1.5 does not hold for ℓ^∞. Once again, this failure is down to the fact that c_{00} is not dense in $(\ell^\infty)^*$, since the latter is not separable.

(End of Day 31)

2. The Hahn-Banach Separation Theorem

In Chapter IV, we proved the Hahn-Banach Extension Theorem. The version for real vector spaces relied on Lemma 2.2, where we extended the linear functional from a subspace of codimension one to the whole space. Let us now revisit this argument, this time paying attention to the seminorm $p : \mathbf{E} \to \mathbb{R}$. A closer look at the proof tells us that we needed two properties of p:

(a) $p(\alpha x) = \alpha p(x)$ for all $x \in \mathbf{E}$, and $\alpha \geq 0$, and

(b) $p(x+y) \leq p(x) + p(y)$ for all $x, y \in \mathbf{E}$.

We never used the fact that $p(\alpha x) = |\alpha| p(x)$ for all $\alpha \in \mathbb{R}$ (The analogous property *was* needed, however, for complex vector spaces).

Definition 2.1. Let $\mathbf{E}$ be a vector space. A function $p : \mathbf{E} \to \mathbb{R}$ is said to be a sublinear functional if it satisfies the properties (a) and (b) given above.

Theorem 2.2. *Let C be a non-empty, convex, open subset of a normed linear space $\mathbf{E}$ such that $0 \in C$. For $x \in \mathbf{E}$, define*

$$p(x) := \inf\{t > 0 : t^{-1}x \in C\}.$$

Then,

(i) There exists $M > 0$ such that $0 \leq p(x) \leq M\|x\|$ for all $x \in \mathbf{E}$. In particular, $p(x) < \infty$.

(ii) p is a sublinear functional.

(iii) For any $x \in \mathbf{E}$, $p(x) < 1$ if and only if $x \in C$.

This function p is called the Minkowski functional (or gauge) of C.

Proof.

(i) Since $0 \in C$ and C is open, there exists $r > 0$ such that $B(0, r) \subset C$. Thus, for any $x \in \mathbf{E}$,
$$\frac{r}{2\|x\|}x \in C.$$
Hence, $p(x) \leq \frac{2\|x\|}{r}$, so $M := 2/r$ works.

(ii) If $x \in \mathbf{E}$, and $\alpha > 0$, then $t^{-1}x \in C$ if and only if $(t\alpha)^{-1}\alpha x \in C$. From this, it follows that
$$p(\alpha x) = \alpha p(x).$$
Now, if $x, y \in \mathbf{E}$, then we wish to prove that $p(x+y) \leq p(x) + p(y)$. To that end, fix $\epsilon > 0$. Then, there exist $s, t > 0$ such that $s^{-1}x \in C$, and $t^{-1}y \in C$ such that $s < p(x) + \epsilon$, and $t < p(y) + \epsilon$. Define
$$r := \frac{s}{s+t}.$$
Then, $0 < r < 1$. Since C is convex,
$$(s+t)^{-1}(x+y) = r(s^{-1}x) + (1-r)(t^{-1}y) \in C.$$
Hence, $p(x+y) \leq s + t < p(x) + p(y) + 2\epsilon$. This is true for all $\epsilon > 0$, and thus $p(x+y) \leq p(x) + p(y)$.

(iii) If $x \in \mathbf{E}$ such that $p(x) < 1$, then there exists $0 < t < 1$ such that $t^{-1}x \in C$. Since C is convex and $0 \in C$,

$$x = (1-t)0 + t(t^{-1}x) \in C.$$

Conversely, if $x \in C$, then, since C is open, there exists $r > 0$ such that $B(x,r) \subset C$. In particular,

$$x + \frac{r}{2\|x\|}x \in C.$$

Therefore, if we set $t := 1 + \frac{r}{2\|x\|}$, then $t > 1$ and $p(x) \leq t^{-1}$. In particular, $p(x) < 1$. This completes the proof.

□

Proposition 2.3. *Let $\mathbf{E}$ be a normed linear space over $\mathbb{R}$, and let $C \subset \mathbf{E}$ be a non-empty, convex, open set. If $x_0 \notin C$, then there exists $\psi \in \mathbf{E}^*$ such that*

$$\psi(x) < \psi(x_0)$$

for all $x \in C$.

Proof. We first assume that $0 \in C$. Let $\mathbf{F} := \text{span}(x_0)$, and let p denote the Minkowski functional of C. Define $\varphi : \mathbf{F} \to \mathbb{R}$ by $\varphi(\alpha x_0) = \alpha$. Since $x_0 \notin C$, $\alpha^{-1}(\alpha x_0) \notin C$ for any $\alpha > 0$, and hence

$$p(\alpha x_0) > \alpha = \varphi(\alpha x_0).$$

If $\alpha < 0$, then this equation holds trivially since $p(\alpha x_0) \geq 0$. Thus, by the Hahn-Banach Theorem, there exists $\psi : \mathbf{E} \to \mathbb{R}$ such that $\psi|_{\mathbf{F}} = \varphi$, and

$$\psi(x) \leq p(x)$$

for all $x \in \mathbf{E}$. By Theorem 2.2, there exists $M > 0$ such that $p(x) \leq M\|x\|$ for all $x \in \mathbf{E}$. Hence,

$$\psi(x) \leq M\|x\|$$

for all $x \in \mathbf{E}$. Replacing x by $-x$, the same inequality holds, so we conclude that ψ is a bounded linear functional. Now, if $x \in C$, then

$$\psi(x) \leq p(x) < 1 = \psi(x_0).$$

Thus, ψ is the desired linear functional.

Now, we consider the case when $0 \notin C$. Fix $x_1 \in C$, and consider $D := C - x_1$, then D is also open and convex, $0 \in D$ and $x_0 - x_1 \notin D$. By the first part of the theorem, there exists $\psi \in \mathbf{E}^*$ such that

$$\psi(y) < \psi(x_0 - x_1)$$

for all $y \in D$. Since ψ is linear, we conclude that $\psi(x) < \psi(x_0)$ for all $x \in C$. □

(End of Day 32)

It is evident that convexity is necessary for Proposition 2.3 to work (simply visualize a counterexample in $\mathbb{R}^2$). The next example shows that openness is also unavoidable.

Example 2.4. Let $\mathbf{E} := c_0$, and define

$$C = \{(x_n) \in c_0 : \text{there exists } N \in \mathbb{N} \text{ such that } x_N > 0, \text{ and } x_n = 0 \text{ for all } n > N\}.$$

Observe that C is convex, and $0 \notin C$. Now, suppose that there exists $\psi \in \mathbf{E}^*$ such that $\psi(x) < \psi(0) = 0$ for all $x \in C$. By Homework 6.3, there exists $y = (y_n) \in \ell^1$ such that

$$\psi((x_n)) = \sum_{n=1}^{\infty} x_n y_n.$$

for all $(x_n) \in c_0$. Therefore, for each $j \in \mathbb{N}$, $y_j = \psi(e_j) < 0$. However, if $x := (y_2, -y_1, 0, 0, \ldots) \in C$, then $\psi(x) = y_1 y_2 - y_2 y_1 = 0$. This contradicts our choice of ψ.

Definition 2.5. Let $\mathbf{E}$ be a vector space over $\mathbb{R}$, $\varphi : \mathbf{E} \to \mathbb{R}$ a non-zero linear functional, and $\alpha \in \mathbb{R}$. The set

$$[\varphi = \alpha] := \{x \in \mathbf{E} : \varphi(x) = \alpha\}$$

is called an affine hyperplane of $\mathbf{E}$.

Theorem 2.6 (Hahn-Banach Separation Theorem - I (Hahn, 1927, and Banach, 1929))**.** *Let $\mathbf{E}$ be a normed linear space over $\mathbb{R}$, and let A and B be two non-empty, disjoint, convex subsets of $\mathbf{E}$. If A is open, then there exists $\psi \in \mathbf{E}^*$, and $\alpha \in \mathbb{R}$, such that*

$$\psi(a) \leq \alpha \leq \psi(b)$$

for all $a \in A$ and $b \in B$. In other words, the closed hyperplane $[\psi = \alpha]$ separates A from B.

Proof. Define $C := A - B$, then it is easy to see that C is convex because A and B are. Furthermore, C is open because

$$C = \bigcup_{b \in B} (A - b),$$

which is a union of sets which are homeomorphic to A. Finally, since A and B are disjoint, $0 \notin C$. By Proposition 2.3, there exists $\psi \in \mathbf{E}^*$ such that $\psi(x) < \psi(0) = 0$ for all $x \in C$. This implies that $\psi(a) < \psi(b)$ for all $a \in A$ and $b \in B$. Therefore, any $\alpha \in \mathbb{R}$ satisfying

$$\sup_{a \in A} \psi(a) \leq \alpha \leq \inf_{b \in B} \psi(b)$$

will do the job. $\square$

Theorem 2.7 (Hahn-Banach Separation Theorem - II)**.** *Let $\mathbf{E}$ be a normed linear space over $\mathbb{R}$, and let A and B be two non-empty, disjoint, convex subsets on $\mathbf{E}$. If A is closed, and B is compact, then there exists $\psi \in \mathbf{E}^*$, $\alpha \in \mathbb{R}$, and $\epsilon > 0$, such that*

$$\psi(a) \leq \alpha - \epsilon < \alpha + \epsilon \leq \psi(b)$$

for all $a \in A$ and $b \in B$. In other words, the closed hyperplane $[\psi = \alpha]$ strictly separates A from B.

Proof. To begin with, we claim that there is a number $r > 0$ such that

$$[A + B(0, r)] \cap [B + B(0, r)] = \emptyset$$

Geometrically, one may think of these sets as *thickening* both A and B, while still keeping them disjoint. Suppose not, then for all $n \in \mathbb{N}$, there exists $u_n \in [A + B(0, 1/n)] \cap [B + B(0, 1/n)]$. Write

$$u_n = a_n + x_n, \text{ and } b_n + y_n,$$

where $a_n \in A, b_n \in B, \|x_n\| < 1/n$, and $\|y_n\| < 1/n$. Hence, $\|a_n - b_n\| \leq \frac{2}{n}$. Since B is compact, there is a subsequence (b_{n_k}) of (b_n) and a point $b \in B$ such that $\lim_{k\to\infty} b_{n_k} = b$. Hence, $\lim_{k\to\infty} a_{n_k} = b$ as well. Since A is closed,

$$b \in A \cap B.$$

This contradicts the fact that A and B are disjoint, thus proving the claim.

Now, choose $r > 0$ such that $\tilde{A} := A + B(0, r)$ and $\tilde{B} := B + B(0, r)$ are disjoint. Note that both $\tilde{A}$ and $\tilde{B}$ are convex and open (as in the proof of the previous theorem). Hence, by Theorem 2.6, there exists $\psi \in \mathbf{E}^*$, and $\alpha \in \mathbb{R}$, such that

$$\psi(u) \leq \alpha \leq \psi(v)$$

for all $u \in \widetilde{A}$ and $v \in \widetilde{B}$. Now, for $a \in A$, and $z \in B[0, 1]$, $a + \frac{r}{2}z \in \tilde{A}$. Therefore,

$$\psi(a) + \frac{r}{2}\psi(z) \leq \alpha.$$

This is true for every $z \in B[0, 1]$, so we conclude that

$$\psi(a) + \frac{r}{2}\|\psi\| \leq \alpha.$$

Similarly, for any $b \in B$, $\alpha \leq \psi(b) - \frac{r}{2}\|\psi\|$. Therefore, $\epsilon := \frac{r}{2}\|\psi\|$ works. $\square$

Theorem 2.8 (Hahn-Banach Separation Theorem - Complex Case)**.** *Let* $\mathbf{E}$ *be a normed linear space over* $\mathbb{C}$*, and let* A *and* B *be two non-empty, disjoint, convex subsets of* $\mathbf{E}$*.*

(i) If A *is open, then there exists* $\psi \in \mathbf{E}^*$*, and* $\alpha \in \mathbb{R}$ *such that*

$$\mathrm{Re}(\psi)(a) \leq \alpha \leq \mathrm{Re}(\psi)(b)$$

for all $a \in A$*, and* $b \in B$*.*

(ii) If A *is closed, and* B *is compact, then there exists* $\psi \in \mathbf{E}^*$*,* $\alpha \in \mathbb{R}$*, and* $\epsilon > 0$ *such that*

$$\mathrm{Re}(\psi)(a) \leq \alpha - \epsilon < \alpha + \epsilon \leq \mathrm{Re}(\psi)(b)$$

for all $a \in A$*, and* $b \in B$*.*

(End of Day 33)

3. The Weak Topology

Lemma 3.1. *Let X be a set and $\mathcal{B} \subset 2^X$ be a collection of subsets of X satisfying two conditions:*

- $\bigcup_{B \in \mathcal{B}} B = X$.
- *For every $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, there exists $B_3 \in \mathcal{B}$ such that $x \in B_3$, and $B_3 \subset B_1 \cap B_2$.*

Then, there is a unique topology $\tau_{\mathcal{B}}$ on X such that

(i) $\mathcal{B} \subset \tau_{\mathcal{B}}$, *and*

(ii) if σ is any other topology such that $\mathcal{B} \subset \sigma$, then $\tau_{\mathcal{B}} \subset \sigma$.

Furthermore, $\mathcal{B}$ is a basis for the topology $\tau_{\mathcal{B}}$.

Proof. Let $\mathcal{U}$ be the collection of all topologies on X which contain $\mathcal{B}$. Then, $2^X \in \mathcal{U}$, so $\mathcal{U} \neq \emptyset$. Therefore, we may define

$$\tau_{\mathcal{B}} = \bigcap_{\sigma \in \mathcal{U}} \sigma.$$

Then, it is easy to see that $\tau_{\mathcal{B}}$ is a topology and it satisfies (i) and (ii). That $\mathcal{B}$ is a basis for $\tau_{\mathcal{B}}$ follows by the assumptions made on it. □

Proposition 3.2. *Let X be any set, (Y, τ_Y) be a topological space, and let $\mathcal{G}$ denote a collection of functions from X to Y. Then, there is a unique topology $\tau_{\mathcal{G}}$ on X such that*

(i) Each $f \in \mathcal{G}$ is continuous with respect to $\tau_{\mathcal{G}}$,

(ii) If σ is any other topology on X such that $f : (X, \sigma) \to (Y, \tau_Y)$ is continuous for all $f \in \mathcal{G}$, then $\tau_{\mathcal{G}} \subset \sigma$.

In other words, $\tau_{\mathcal{G}}$ is the smallest topology that makes every member of $\mathcal{G}$ continuous. This topology is called the weak topology on X defined by $\mathcal{G}$.

Proof. Define $\mathcal{B} \subset 2^X$ to be the collection of all finite intersections of sets of the form $f^{-1}(U)$, for some $U \in \tau_Y$, and $f \in \mathcal{G}$. In other words, $B \in \mathcal{B}$ if and only if there exist finitely many functions $f_1, f_2, \ldots, f_n \in \mathcal{G}$, and open sets $U_1, U_2, \ldots, U_n \in \tau_Y$ such that

$$B = \bigcap_{i=1}^{n} f_i^{-1}(U_i).$$

Note that $f^{-1}(Y) = X$ for any $f \in \mathcal{G}$. Also, if $B_1, B_2 \in \mathcal{B}$, then $B_1 \cap B_2 \in \mathcal{B}$. Therefore, Lemma 3.1 applies, and we take $\tau_{\mathcal{G}} := \tau_{\mathcal{B}}$.

It is now clear that each $f \in \mathcal{G}$ is continuous with respect to $\tau_{\mathcal{G}}$ (by construction). Furthermore, if σ is any other topology on X such that $f : (X, \sigma) \to (Y, \tau_Y)$ is continuous for each $f \in \mathcal{G}$, then σ will necessarily contain every member of $\mathcal{B}$. But $\mathcal{B}$ is a basis for $\tau_{\mathcal{G}}$, so $\tau_{\mathcal{G}} \subset \sigma$ must hold. □

Before we begin our journey, let us pause to prove a very useful result in this context; one that will be used repeatedly throughout the book.

Proposition 3.3. *Let X be a set, (Y, τ_Y) be a topological space, and let $\mathcal{G}$ be a collection of functions from X to Y. Let $\tau_{\mathcal{G}}$ be the weak topology on X defined by $\mathcal{G}$. If (Z, τ_Z) is any topological space, then a function*

$$g : (Z, \tau_Z) \to (X, \tau_{\mathcal{G}})$$

is continuous if and only if $f \circ g : (Z, \tau_Z) \to (Y, \tau_Y)$ is continuous for each $f \in \mathcal{G}$.

Proof. If $g : (Z, \tau_Z) \to (X, \tau_{\mathcal{G}})$ is continuous, then clearly $f \circ g : (Z, \tau_Z) \to (Y, \tau_Y)$ is continuous because $f : (X, \tau_{\mathcal{G}}) \to (Y, \tau_Y)$ is continuous by construction.

Now suppose $g : Z \to X$ is function such that $f \circ g : (Z, \tau_Z) \to (Y, \tau_Y)$ is continuous for each $f \in \mathcal{G}$. To prove that g is continuous, we choose an open set $U \in \tau_{\mathcal{G}}$ and prove that $g^{-1}(U) \in \tau_Z$. We may assume that U is a basic open set, so there would exist $\{f_1, f_2, \ldots, f_n\} \subset \mathcal{G}$ and open sets $\{U_1, U_2, \ldots, U_n\} \subset \tau_Y$ such that

$$U = \bigcap_{i=1}^{n} f_i^{-1}(U_i).$$

Then,

$$g^{-1}(U) = \bigcap_{i=1}^{n} g^{-1}(f_i^{-1}(U_i)) = \bigcap_{i=1}^{n} (f_i \circ g)^{-1}(U_i)$$

and this set belongs to τ_Z by hypothesis. Therefore, g is continuous. □

Definition 3.4. Let $\mathbf{E}$ be a normed linear space, and $Y = \mathbb{K}$ equipped with the usual topology. Let $\mathcal{G} := \mathbf{E}^*$, the set of all bounded linear functionals on $\mathbf{E}$. The weak topology on $\mathbf{E}$ is the topology defined by $\mathcal{G}$ by means of Proposition 3.2. This is denoted by

$$\sigma(\mathbf{E}, \mathbf{E}^*)$$

to indicate that it is the topology on $\mathbf{E}$ inherited from $\mathbf{E}^*$. Elements of $\sigma(\mathbf{E}, \mathbf{E}^*)$ are called weakly open sets, and their complements are called weakly closed sets. A subset of $\mathbf{E}$ is said to be weakly compact if it is compact with respect to $\sigma(\mathbf{E}, \mathbf{E}^*)$.

For clarity, we will henceforth refer to the norm topology on $\mathbf{E}$ by $\sigma(\mathbf{E}, \|\cdot\|)$, members of which will be referred to as norm-open (or strongly open) sets. The terms norm-closed and norm-compact are defined analogously.

Remark 3.5. Some remarks are in order before we proceed:

(i) Since each $\varphi \in \mathbf{E}^*$ is continuous with respect to the norm, it follows that

$$\sigma(\mathbf{E}, \mathbf{E}^*) \subset \sigma(\mathbf{E}, \|\cdot\|),$$

since $\sigma(\mathbf{E}, \mathbf{E}^*)$ is the smallest topology with this property. In other words, we have the following implications:

$$\begin{aligned} \text{weakly open} &\Rightarrow \text{norm-open,} \\ \text{weakly closed} &\Rightarrow \text{norm-closed, and} \\ \text{weakly compact} &\Leftarrow \text{norm-compact} \end{aligned}$$

(ii) For a set $A \subset \mathbf{E}$, we write

$$\overline{A}^w$$

to denote the intersection of all weakly closed set containing A. It is called the <u>weak closure</u> of A. We denote the norm-closure of A by $\overline{A}^{\|\cdot\|}$, if the context demands it. Observe that $\overline{A}^{\|\cdot\|} \subset \overline{A}^w$, since $\overline{A}^w$ is a norm-closed set that contains A.

(iii) We now describe basic open sets in $\sigma(\mathbf{E}, \mathbf{E}^*)$.

(a) If $\varphi \in \mathbf{E}^*$ and $\epsilon > 0$, then

$$\{x \in \mathbf{E} : |\varphi(x)| < \epsilon\}$$

is a (sub-)basic open neighbourhood of 0, since it is inverse image under φ of an open set in $\mathbb{K}$.

(b) In general, if $\varphi_1, \varphi_2, \ldots, \varphi_n \in \mathbf{E}^*$ and $\epsilon_i > 0$, then

$$\{x \in \mathbf{E} : |\varphi_i(x)| < \epsilon_i \text{ for all } 1 \leq i \leq n\}$$

is a basic open neighbourhood of 0. If $\epsilon = \min\{\epsilon_i : 1 \leq i \leq n\}$, then this neighbourhood contains

$$\{x \in \mathbf{E} : |\varphi_i(x)| < \epsilon \text{ for all } 1 \leq i \leq n\}.$$

(c) If $x_0 \in \mathbf{E}, \varphi \in \mathbf{E}^*, \epsilon > 0$, then

$$\{x \in \mathbf{E} : |\varphi(x) - \varphi(x_0)| < \epsilon\}$$

is a (sub-)basic open neighbourhood of x_0. As argued above, every basic open neighbourhood of x_0 will contain one of the form

$$\{x \in \mathbf{E} : |\varphi_i(x) - \varphi_i(x_0)| < \epsilon \text{ for all } 1 \leq i \leq n\},$$

for some $\varphi_1, \varphi_2, \ldots, \varphi_n \in \mathbf{E}^*$, and $\epsilon > 0$ fixed.

(iv) Finally, one may use this description of open sets to prove that, for a sequence (x_n) in $\mathbf{E}$, $x_n \xrightarrow{w} x$ if and only if $x_n \to x$ with respect to $\sigma(\mathbf{E}, \mathbf{E}^*)$.

(End of Day 34)

Proposition 3.6. *The weak topology $\sigma(\mathbf{E}, \mathbf{E}^*)$ is Hausdorff.*

Proof. If $x, y \in \mathbf{E}$ are distinct, then by Corollary 2.9, there exists $\varphi \in \mathbf{E}^*$ such that $\varphi(x) \neq \varphi(y)$. Define

$$\epsilon := \frac{|\varphi(x) - \varphi(y)|}{3},$$

and set $U := B_{\mathbb{K}}(\varphi(x), \epsilon)$, and $V := B_{\mathbb{K}}(\varphi(y), \epsilon)$. Then, $\varphi^{-1}(U)$ and $\varphi^{-1}(V)$ are disjoint, weakly open neighbourhoods of x and y respectively. $\square$

Definition 3.7. A <u>topological vector space</u> is a vector space $\mathbf{E}$ equipped with a Hausdorff topology, such that the addition map $a : \mathbf{E} \times \mathbf{E} \to \mathbf{E}$, and the scalar multiplication map $s : \mathbb{K} \times \mathbf{E} \to \mathbf{E}$ are both continuous (here, $\mathbb{K}$ is equipped with the usual norm topology, and the product spaces are equipped with the product topologies).

Proposition 3.8. *The space $(\mathbf{E}, \sigma(\mathbf{E}, \mathbf{E}^*))$ is a topological vector space.*

Proof. We only prove that a is continuous, as the proof for s is similar. Let $W \in \sigma(\mathbf{E}, \mathbf{E}^*)$ denote a weakly open set, and a tuple $(x, y) \in a^{-1}(W)$ so that $z = a(x, y) = x + y \in W$. Then, there exist $\varphi_1, \varphi_2, \ldots, \varphi_n \in \mathbf{E}^*$, and $\epsilon > 0$ such that

$$W' := \{w \in \mathbf{E} : |\varphi_i(w) - \varphi_i(z)| < \epsilon \text{ for all } 1 \leq i \leq n\} \subset W.$$

Define weakly open neighbourhoods U and V of x and y respectively by

$$U := \{u \in \mathbf{E} : |\varphi_i(u) - \varphi_i(x)| < \epsilon/2 \text{ for all } 1 \leq i \leq n\}$$
$$V := \{v \in \mathbf{E} : |\varphi_i(v) - \varphi_i(y)| < \epsilon/2 \text{ for all } 1 \leq i \leq n\}.$$

If $(u, v) \in U \times V$, then $|\varphi_i(u + v) - \varphi_i(z)| < \epsilon$ for all $1 \leq i \leq n$. Hence, $a(u, v) = u + v \in W'$. Thus, $U \times V \subset a^{-1}(W)$ and $(x, y) \in U \times V$. This is true for any $(x, y) \in a^{-1}(W)$, proving that $a^{-1}(W)$ is open. $\square$

Theorem 3.9. *If $\mathbf{E}$ is finite dimensional, then the weak and norm topologies coincide.*

Proof. Let $\sigma(\mathbf{E}, \mathbf{E}^*)$ and $\sigma(\mathbf{E}, \|\cdot\|)$ denote the weak and norm topologies respectively. By definition, $\sigma(\mathbf{E}, \mathbf{E}^*) \subset \sigma(\mathbf{E}, \|\cdot\|)$. If $\mathbf{E}$ is finite dimensional, we wish to prove that $\sigma(\mathbf{E}, \|\cdot\|) \subset \sigma(\mathbf{E}, \mathbf{E}^*)$.

We may assume without loss of generality that $\mathbf{E} = (\mathbb{K}^n, \|\cdot\|_\infty)$. For $1 \leq i \leq n$, let $\pi_i : \mathbf{E} \to \mathbb{K}$ denote the coordinate projections. Then, for any $x \in \mathbf{E}$ and $r > 0$,

$$B_{\mathbf{E}}(x, r) = \{y \in \mathbf{E} : |\pi_i(y) - \pi_i(x)| < r \text{ for all } 1 \leq i \leq n\}.$$

Thus, $B_{\mathbf{E}}(x, r) \in \sigma(\mathbf{E}, \mathbf{E}^*)$. This is true for any basic open set $B_{\mathbf{E}}(x, r) \in \sigma(\mathbf{E}, \|\cdot\|)$, and hence $\sigma(\mathbf{E}, \|\cdot\|) \subset \sigma(\mathbf{E}, \mathbf{E}^*)$ as well. $\square$

Theorem 3.10. *Let $\mathbf{E}$ be an normed linear space, and $C \subset \mathbf{E}$ be a convex set. Then, C is weakly closed if and only if it is norm-closed.*

Proof. If C is weakly closed, then it is norm-closed by Remark 3.5. Conversely, if C is convex and norm-closed, then we wish to prove that C is weakly closed. Suppose $x \notin C$, then, by the Hahn-Banach Separation Theorem, there exists $\psi \in \mathbf{E}^*, \alpha \in \mathbb{R}$, and $\epsilon > 0$ such that $\text{Re}(\psi)(x) < \alpha - \epsilon$, and

$$\text{Re}(\psi)(y) > \alpha + \epsilon$$

for all $y \in C$. Note that $\psi : (\mathbf{E}, \sigma(\mathbf{E}, \mathbf{E}^*)) \to \mathbb{K}$ is continuous by the very definition of the weak topology, and $\text{Re} : \mathbb{K} \to \mathbb{R}$ is continuous (it is the identity map if $\mathbb{K} = \mathbb{R}$). Hence,

$$U = \{u \in \mathbf{E} : \text{Re}(\psi(u)) < \alpha - \epsilon\}$$

is weakly open, $x \in U$ and $U \cap C = \emptyset$. This is true for any $x \notin C$, which proves that $\mathbf{E} \setminus C$ is weakly open, which is what we wanted to prove. □

Theorem 3.11. *Let* $\mathbf{E}$ *be an infinite dimensional subspace. If* $S_{\mathbf{E}}$ *and* $B_{\mathbf{E}}$ *denote the unit sphere and closed unit ball respectively, then*

$$\overline{S_{\mathbf{E}}}^w = B_{\mathbf{E}}.$$

In particular, $S_{\mathbf{E}}$ *is not weakly closed.*

Proof. Since $B_{\mathbf{E}}$ is convex and norm-closed, $B_{\mathbf{E}}$ is weakly closed by Theorem 3.10. Since $S_{\mathbf{E}} \subset B_{\mathbf{E}}$, it follows that

$$\overline{S_{\mathbf{E}}}^w \subset B_{\mathbf{E}}.$$

To prove the converse, it suffices to choose $x_0 \in \mathbf{E}$ such that $\|x_0\| < 1$, and prove that $x_0 \in \overline{S_{\mathbf{E}}}^w$. So, let $U \in \sigma(\mathbf{E}, \mathbf{E}^*)$ be any weakly open neighbourhood of x_0. We wish to prove that $U \cap S_{\mathbf{E}} \neq \emptyset$.

We may assume without loss of generality that

$$U = \{x \in \mathbf{E} : |\varphi_i(x) - \varphi_i(x_0)| < \epsilon \text{ for all } 1 \leq i \leq n\}$$

for some $\varphi_1, \varphi_2, \ldots, \varphi_n \in \mathbf{E}^*$ and $\epsilon > 0$. Now, define $T : \mathbf{E} \to \mathbb{K}^n$ by

$$T(x) := (\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)).$$

Then, T cannot be injective since $\dim(\mathbf{E}) = \infty$. Hence, there exists a non-zero vector $y_0 \in \mathbf{E}$ such that $\varphi_i(y_0) = 0$ for all $1 \leq i \leq n$. This implies that $x_0 + ty_0 \in U$ for all $t \in \mathbb{R}$. Now, consider $g : \mathbf{E} \to \mathbb{R}_+$ by

$$g(t) := \|x_0 + ty_0\|.$$

Then, g is norm continuous (by Remark 1.2) and $g(0) = \|x_0\| < 1$. Finally, since

$$t\|y_0\| \leq \|x_0\| + \|x_0 + ty_0\|,$$

it follows that $\lim_{t\to\infty} g(t) = \infty$. By the Intermediate Value Theorem, there exists $t_0 \in \mathbb{R}$ such that $\|x_0 + t_0y_0\| = 1$. Thus, $x_0 + t_0y_0 \in S_{\mathbf{E}} \cap U$ and $S_{\mathbf{E}} \cap U \neq \emptyset$. □

Corollary 3.12. *If* $\mathbf{E}$ *is an infinite dimensional normed linear space, then* $\sigma(\mathbf{E}, \mathbf{E}^*) \neq \sigma(\mathbf{E}, \|\cdot\|)$.

(End of Day 35)

VII. Instructor Notes

(i) The course design remained the same, but more emphasis was placed on reviewing earlier material. This was to compensate for learning losses due to Covid. Therefore, the total amount of material covered was much less than earlier iterations.

(ii) Even so, I underestimated how much the students have missed out on due to the two year break (and rampant cheating). Despite the repeated reviews, they were still lost.

(iii) Finally, the attendance and interest of the students was never really there. Most days, there was 5/20 students in class. This was reflected in their grades as well.

Bibliography

[1] Paul R. Halmos. *Finite-dimensional vector spaces*. second. Undergraduate Texts in Mathematics. Springer-Verlag, New York-Heidelberg, 1974, pp. viii+200.

[2] H. L. Royden. *Real analysis*. Third. Macmillan Publishing Company, New York, 1988, pp. xx+444. ISBN: 0-02-404151-3.