

Elements of Functional Analysis

A Series of Lecture Notes Compiled by Matthew R. Gamel

This is an eclectic series of notes compiled by Matthew Gamel. Much of the content comes from the lecture notes of my Fall 2004-Spring 2005 real analysis courses taught by Prof. Roger Smith at Texas A&M University and my lecture notes for my Fall 2006-Spring 2007 functional analysis course taught by Prof. George Androulakis at the University of South Carolina. Some of the content comes from my solutions to homework problems and from other sources as well.

This set of notes evolved out of my desire to organize an eclectic series of notes from various courses. I hope that these notes will prove to be beneficial as a reference or a supplement to a graduate level course in the subject.

This set of notes is outlined as follows (note that these notes are currently incomplete):

- a. Introduction to Banach and Hilbert Spaces
- b. Cornerstone Theorems of Functional Analysis
- c. Weak and Weak* Topologies
- d. Schauder Bases and Other Topics in Banach Space Theory
- e. Function Spaces

I have also included some exercises for the interested reader at the end of each section.

If there are any mathematical or typographical errors, please send them to me at gamel@math.sc.edu.

1 Banach and Hilbert Spaces

Here we discuss the basic theory of Banach and Hilbert spaces. Unless otherwise specified, all vector spaces are over \mathbb{C} .

1.1 Normed Vector Spaces and Inner Product Spaces

Definition 1.1.0

A *normed space* is a vector space X equipped with a function $\|\cdot\|$ from X to \mathbb{R} satisfying the following properties:

- (a) $\|x\| \geq 0$ for all $x \in X$ with equality iff $x = 0$.
- (b) $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{C}$ and $x \in X$
- (c) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$

□

If X is a normed space, we will often write its norm as $\|\cdot\|_X$ to avoid confusion with other spaces. We omit the subscript when the space in question is understood. In addition, if we define $\rho(x, y) := \|x - y\|$ then it is easy to see that (X, ρ) is a metric space whose metric is induced by the norm.

If (x_n) is a sequence in X , we say that x_n converges to $x_0 \in X$, written $x_n \rightarrow x_0$, if we have $\lim_{n \rightarrow \infty} \|x_n - x_0\| = 0$. Moreover, if the normed space X is equipped with two norms $\|\cdot\|_1$ and $\|\cdot\|_2$, we say that these norms are equivalent if there exist constants $C_1 > 0$ and $C_2 > 0$ so that

$$C_1\|x\|_1 \leq \|x\|_2 \leq C_2\|x\|_1 \text{ for all } x \in X$$

It is easy to see that a sequence (x_n) converges in $\|\cdot\|_1$ if and only if it converges in $\|\cdot\|_2$ provided that they are equivalent. Equivalent norms, therefore, generate equivalent norm topologies.

When applying the theory of Banach and Hilbert spaces, the following inequality makes life exceptionally easier. We shall make use of this later so we prove it here.

Proposition 1.1.1 (Young's Inequality)

Let $a, b \in \mathbb{R}^+$ and suppose that $p > 1$ and q are such that $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

Proof. If either a or b is 0 (or both, possibly) then the result is obvious. Assume, then, that they are nonzero. Set $f(t) = \frac{1}{p}t^p + \frac{1}{q} - t$ where $0 \leq t < \infty$ so that $f'(t) = t^{p-1} - 1$. On $(0, 1)$, $f'(t) < 0$ so f is decreasing from $f(0) = \frac{1}{q} > 0$ down to $t = 1$. On $(1, \infty)$, $f'(t) > 0$ so f is increasing and the point $t = 1$ is an absolute minimum. Hence $f(t) \geq 0$ for all $t \geq 0$. Setting $t = \frac{a}{b^{q-1}}$ and simplifying yields the result (verify). □

We now define the basic notion of an inner product space and then prove some other useful results.

Definition 1.1.2

An *inner-product space* is a vector space X equipped with a function (\cdot, \cdot) from $X \times X$ to \mathbb{C} satisfying the following properties:

- (a) $(x, x) \geq 0$ with equality iff $x = 0$ for all $x \in X$.
- (b) $(x, y) = \overline{(y, x)}$ for all $x, y \in X$
- (c) $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$ for all $\alpha, \beta \in \mathbb{C}$ and $x, y, z \in X$.

We say that $x, y \in X$ are *orthogonal*, written $x \perp y$, provided that $(x, y) = 0$. □

Notice that property (b) also implies that (x, x) must be a real number. Hence property (a) makes sense. If our vector space is over \mathbb{R} instead of \mathbb{C} , then property (b) simply reduces to $(x, y) = (y, x)$. Also if we combine property (b) and property (c), we get $(x, \alpha y + \beta z) = \overline{\alpha}(x, y) + \overline{\beta}(x, z)$. This next result is essential.

Proposition 1.1.3 (Cauchy-Schwartz Inequality)

Let X be an inner-product space. Then, for all $x, y \in X$, we have

$$|(x, y)| \leq \sqrt{(x, x)}\sqrt{(y, y)}$$

Proof. Let $\lambda \in \mathbb{C}$ and let $x, y \in X$. Then

$$0 \leq (x + \lambda y, x + \lambda y) = (x, x) + 2\operatorname{Re}[\lambda(x, y)] + |\lambda|^2(y, y)$$

If $x = 0$ or $y = 0$ (or both) then the inequality clearly holds. Now suppose that y is nonzero and set

$$\lambda = -\frac{(y, x)}{(y, y)}$$

Then, after some minor arithmetic, we get

$$0 \leq (x, x) - 2\frac{|(y, x)|^2}{(y, y)} + \frac{|(y, x)|^2}{(y, y)}$$

so the result clearly holds. □

Just as we noted that a normed space is also a metric space, we here demonstrate that an inner-product space is also a normed space.

Proposition 1.1.4

Let X be an inner-product space. Then $\|x\| := \sqrt{(x, x)}$ defines a norm on X .

Proof. The first few properties are obvious so we only verify that the triangle inequality holds. We have

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

□

Definition 1.1.5

Let X be a normed space. We call X a *Banach space* if it is complete in norm; that is, every Cauchy sequence in X converges to a point in X . □

This next definition is really a special case of the previous definition but we elect to state it here separately due to its importance.

Definition 1.1.6

Let H be an inner-product space. We call H a *Hilbert space* if it is complete in the norm induced by the inner-product. □

Essentially, a Hilbert space can be thought of as a generalization of Euclidean space in finite dimensions and consequently there are many nice geometric interpretations associated with Hilbert spaces. We shall see this at a later time.

1.2 Operators, Linear Functionals, and the Hahn-Banach Theorem

Again, we are assuming all normed vector spaces are over \mathbb{C} unless otherwise specified. We begin with a useful equivalent formulation of a Banach space.

Proposition 1.2.0

Let X be a normed space. The following are equivalent:

- (a) X is a Banach space
- (b) For each sequence $x_n \in X$ ($\forall n$) with the property that

$$\sum_{k=1}^{\infty} \|x_k\|_X < \infty$$

we have the partial sums $s_n = \sum_{k=1}^n x_k$ converge in norm.

Proof. Suppose that X is a Banach space and suppose that x_n is a sequence in X with $\sum_{k=1}^{\infty} \|x_k\|_X < \infty$. Clearly

$$\|s_n - s_m\|_X = \left\| \sum_{m+1}^n x_k \right\|_X \leq \sum_{m+1}^n \|x_k\|_X$$

so that the partial sums are Cauchy and, hence, must converge. Conversely, let x_n be Cauchy in norm. By dropping to a subsequence, we may without loss of generality

assume that $\|x_n - x_{n-1}\|_X \leq \frac{1}{2^n}$ when $n \geq 2$. Define $y_1 = x_1$ and $y_n = x_n - x_{n-1}$ when $n \geq 2$. Then

$$\sum_{k=1}^n \|y_k\|_X \leq \|x_1\|_X + \sum_{k=2}^n \frac{1}{2^k} \leq \|x_1\|_X + 1 < \infty$$

Thus, we get $\sum_{k=1}^{\infty} \|y_k\|_X < \infty$. Clearly $s_n := \sum_{k=1}^n y_k = x_n$ so that by hypothesis, $s_n = x_n$ converges and X is a Banach space. □

If X and Y are normed spaces and $T : X \rightarrow Y$, we call T an *operator*. The following equivalent conditions are useful and extensively used.

Proposition 1.2.1

Let X and Y be normed spaces and let $T : X \rightarrow Y$ be a linear operator. Then the following are equivalent:

- (a) T is continuous.
- (b) T is continuous at 0.
- (c) (Boundedness) There is a constant $C > 0$ so that $\|Tx\|_Y \leq C\|x\|_X$ for all $x \in X$.

Proof. It is easy to see that (a) and (b) are equivalent. It is also obvious that (c) implies (a) and (b). We will now show that (b) implies (c). Assume (b) and suppose that no such $C > 0$ exists. Then, by normalizing if necessary, we may find a sequence $y_n \in X$ with $\|y_n\|_X = 1$ such that $\|Ty_n\|_Y \geq n$. Let $x_n = \frac{1}{\sqrt{n}}y_n$ so that $\|x_n\|_X \rightarrow 0$ as $n \rightarrow \infty$. Then $x_n \rightarrow 0$ and $Tx_n = \frac{1}{\sqrt{n}}Ty_n$ so $\|Tx_n\|_Y \geq \frac{n}{\sqrt{n}} = \sqrt{n}$ but Tx_n does not go to 0, a contradiction. □

Definition 1.2.2

Let X and Y be normed spaces. We denote by $\mathcal{L}(X, Y)$ the set of all continuous linear operators from X to Y . If $Y = \mathbb{R}$ (or \mathbb{C}) then $\mathcal{L}(X, \mathbb{R})$ is called the *dual* of X and is denoted by $X^* = \mathcal{L}(X, \mathbb{R})$. If $\phi \in X^*$, we sometimes call ϕ a *linear functional*. □

By the previous proposition, we shall employ the terms *continuous* and *bounded* interchangeably in the context of linear operators between normed spaces. If $T : X \rightarrow Y$ is bounded, we define

$$\|T\| := \sup_{x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\|x\|_X=1} \|Tx\|_Y$$

This is called the *operator norm* and if we equip the space of all continuous linear operators $\mathcal{L}(X, Y)$ with this norm, we get a normed space and in some cases, a Banach space.

Theorem 1.2.3

If X and Y are normed spaces, then so is $\mathcal{L}(X, Y)$. Additionally, if Y is a Banach space, then so is $\mathcal{L}(X, Y)$.

Proof. It is easy to check that $\mathcal{L}(X, Y)$ is a normed vector space. Let $T_n \in \mathcal{L}(X, Y)$ be Cauchy. Then for each $x \in X$, $\|(T_n - T_m)x\|_Y \leq \|T_n - T_m\| \cdot \|x\|_X$ so that $T_n x$ is Cauchy in Y . Since Y is a Banach space, there is a $y \in Y$ so that $T_n x \rightarrow y$. Define $Tx := y$. We will now show that T is linear and that $T_n \rightarrow T$ in the operator norm.

First, if $x_1, x_2 \in X$ then $T_n x_1 \rightarrow Tx_1$ and $T_n x_2 \rightarrow Tx_2$ so that $T_n x_1 + T_n x_2 \rightarrow Tx_1 + Tx_2$. But $T_n x_1 + T_n x_2 = T_n(x_1 + x_2) \rightarrow T(x_1 + x_2)$ so by uniqueness of limits, $Tx_1 + Tx_2 = T(x_1 + x_2)$. Similarly, $T(\lambda x) = \lambda Tx$ whenever $x \in X$ so T is linear.

Let $\epsilon > 0$. Then we may find an N_0 so that $\|T_n - T_m\| < \epsilon$ whenever $n, m \geq N_0$. Then, for each $x \in X$ with $\|x\|_X = 1$, we get $\|T_n x - T_m x\|_Y < \epsilon$ when $n, m \geq N_0$. By continuity of the norm, we fix any $n \geq N_0$ and let $m \rightarrow \infty$ to get $\|T_n x - Tx\|_Y < \epsilon$. This holds for each $n \geq N_0$ and $x \in X$ with $\|x\|_X = 1$. Thus

$$\|T_n - T\| = \sup_{\|x\|_X=1} \|T_n x - Tx\|_Y < \epsilon \text{ whenever } n \geq N_0$$

Finally, because $\|T\| < \|T_{N_0}\| + 1$, we're finished. □

An immediate consequence of this is if $Y = \mathbb{R}$ (or \mathbb{C}) then X^* is a Banach space. We may also define the double dual of a normed space X given by

$$X^{**} := (X^*)^* = \mathcal{L}(X^*, \mathbb{F})$$

where $\mathbb{F} = \mathbb{R}$ or \mathbb{C} . In a likewise manner, X^{**} is a Banach space.

Definition 1.2.4 (Canonical Embeddings)

Let X be a real or complex normed space and for each $x \in X$, define a map $J_x(\phi) := \phi(x)$ for all $\phi \in X^*$. We call J_x the *canonical embedding* of X into X^{**} . □

The canonical embedding satisfies many nice properties; namely, it induces a continuous injective isometry from X into X^{**} (that is, $\|J_x\|_{X^{**}} = \|x\|_X$). The details are left as an exercise for the reader.

Definition 1.2.5 (Reflexivity)

A normed space X is called *reflexive* if the canonical embedding is surjective. □

Not every space is reflexive as we shall later see and it turns out to be fairly tricky to show that many spaces are reflexive. Before we move on, we introduce one more concept relating to dual spaces.

Definition 1.2.6 (Adjoint Operators)

Let X and Y be normed spaces. If $T : X \rightarrow Y$ is linear then the adjoint of T , denoted T^* is given by $T^*\phi := \phi \circ T$ for all $\phi \in Y^*$. □

It is obvious that the adjoint operator $T^* : Y^* \rightarrow X^*$ is linear and we leave it to the reader to show that if T is bounded, then so is its adjoint with $\|T\| = \|T^*\|$.

We will now prepare to prove our first major result. However, we need some definitions before we proceed. At this point, we will now require that our normed space be over \mathbb{R} rather than \mathbb{C} .

Definition 1.2.7 (Sublinear Functionals and Seminorms)

A *sublinear functional* on a real vector space X is a function $p : X \rightarrow \mathbb{R}$ such that

- (a) $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$
- (b) If $\lambda \geq 0$ then $p(\lambda x) = \lambda p(x)$ for all $x \in X$.

On a real or complex vector space, a *seminorm* is a function $p : X \rightarrow [0, \infty)$ satisfying

- (a) $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$

(b) $p(\lambda x) = |\lambda|p(x)$ for all $x \in X$ and $\lambda \in \mathbb{R}$ (or \mathbb{C})

□

The only difference between a seminorm and a norm is that the condition $p(x) = 0$ if and only if $x = 0$ is missing. We briefly digress to give an example.

Example (Minkowski Functionals)

Let $K \subset \mathbb{R}^2$ be convex and contain a neighborhood of the origin. Let $K_\lambda = \{\lambda x : x \in K\}$ and define $p(x) := \inf\{\lambda : x \in K_\lambda\}$. We shall prove that this function, called the Minkowski Functional, is a sublinear functional.

Firstly, observe that $p(cx) = cp(x)$ whenever $c \geq 0$. Let $\epsilon > 0$ and $x, y \in K$. Then there is a λ and a μ so that $p(x) \leq \lambda < p(x) + \epsilon$ and $p(y) \leq \mu < p(y) + \epsilon$. Clearly, $x \in K_\lambda$ and $y \in K_\mu$. Hence, there is a $k_1, k_2 \in K$ so that $x = \lambda k_1$ and $y = \mu k_2$. Then

$$x + y = \lambda k_1 + \mu k_2 = (\lambda + \mu) \left(\frac{\lambda}{\lambda + \mu} k_1 + \frac{\mu}{\lambda + \mu} k_2 \right) = (\lambda + \mu)k$$

for some $k \in K$ by convexity. So, we have

$$p(x + y) \leq \lambda + \mu \leq p(x) + p(y) + 2\epsilon$$

Letting $\epsilon \rightarrow 0^+$ yields the desired result.

◇

We now proceed to prove our first major result, the Hahn-Banach Theorem. This enables us to extend linear functionals on a subspace of a vector space to the whole space.

Theorem 1.2.8 (Hahn-Banach Theorem)

Let X be a real vector space with a sublinear functional p and $Y \subset X$ a subspace. Suppose that $\psi : Y \rightarrow \mathbb{R}$ is a linear functional and that $\psi(y) \leq p(y)$ for all $y \in Y$. Then ψ extends to a linear functional $\phi : X \rightarrow \mathbb{R}$ so that $\phi|_Y = \psi$ and $\phi(x) \leq p(x)$ for all $x \in X$.

Proof. Our strategy will be to demonstrate the theorem in two basic steps. During the first step, we will essentially attempt to extend ψ as far as we are able. Then, if our extension is not defined on all of X , we extend it again to obtain a contradiction.

Denote by \mathcal{M} the collection of all subspaces $M \supseteq Y$ and linear functionals $\theta : M \rightarrow \mathbb{R}$ so that $\theta(m) \leq p(m)$ for $m \in M$ and $\theta|_Y = \psi$. Order this set in the following way:

$(M_1, \theta_1) \prec (M_2, \theta_2)$ iff $M_1 \subseteq M_2$ and $\theta_2|_{M_1} = \theta_1$.

Let $\mathcal{U} = \{(M_\alpha, \theta_\alpha) : \alpha \in A\}$ be a totally ordered subset of \mathcal{M} (i.e. any two elements in \mathcal{U} are comparable). Set

$$M := \bigcup_{\alpha \in A} M_\alpha$$

It is now not difficult to see that M is a subspace of X containing Y . On M , define $\theta(x) = \theta_\alpha(x)$ when $x \in M_\alpha$. It is easy to see that this is well defined. For if $x \in M_\alpha$ and $x \in M_\beta$ then without loss of generality, assume that $M_\alpha \subset M_\beta$. Since $\theta_\beta|_{M_\alpha} = \theta_\alpha$ we get $\theta_\alpha(x) = \theta_\beta(x)$.

We also have $\theta(x) \leq p(x)$ for all $x \in M$ because $\theta_\alpha(x) \leq p(x)$ for all $\alpha \in A$. Moreover, if $x_1, x_2 \in M$ with $x_1 \in M_{\alpha_1}$ and $x_2 \in M_{\alpha_2}$ then without loss of generality, suppose that $M_{\alpha_1} \subseteq M_{\alpha_2}$. Then

$$\theta(x_1 + x_2) = \theta_{\alpha_2}(x_1 + x_2) = \theta_{\alpha_2}(x_1) + \theta_{\alpha_2}(x_2)$$

So θ is a linear functional with $(M, \theta) \in \mathcal{M}$ and $(M, \theta) \succ (M_\alpha, \theta_\alpha)$ for all α . By Zorn's Lemma, there is a maximal element of \mathcal{M} , call it (Z, ω) .

We now have an extension so we now contend that $Z = X$. Suppose the contrary, that Z is strictly contained in X but not equal to X . Fix an element $x_0 \in X$ with $x_0 \notin Z$. We will extend $\omega : Z \rightarrow \mathbb{R}$ to $\phi : Z_0 \rightarrow \mathbb{R}$ satisfying $\phi(x) \leq p(x)$ for $x \in Z_0$ and $\phi|_Z = \omega$ where $Z_0 = \{z + \lambda x_0 : z \in Z \text{ and } \lambda \in \mathbb{R}\}$. This will be a contradiction because (Z, ω) will no longer be maximal.

The functional ϕ , if it exists, is determined by $\gamma = \phi(x_0)$. We must have

$$\phi(z + \lambda x_0) \leq p(z + \lambda x_0) \text{ for all } z \in Z \text{ and } \lambda \in \mathbb{R}$$

This implies that $\omega(z) + \lambda\gamma \leq p(z + \lambda x_0)$ for all $z \in Z$ and $\lambda \in \mathbb{R}$. This is true if $\lambda = 0$ so if we can choose a γ to make this true when $\lambda > 0$ and $\lambda < 0$ then we will have constructed an extension.

If $\lambda > 0$ we have

$$\omega\left(\frac{z}{\lambda}\right) + \gamma \leq p\left(\frac{z}{\lambda} + x_0\right) \quad \forall z \in Z, \lambda > 0$$

Put $z_1 = \frac{z}{\lambda}$ so we must have $z_1 \in Z$. Letting z and λ vary gives all elements in Z . Then

$$\gamma \leq p(z_1 + x_0) - \omega(z_1) \quad \forall z_1 \in Z$$

If $\lambda < 0$ set $\lambda = -\mu$ for $\mu > 0$. We want $\omega(z) - \mu\gamma \leq p(z - \mu x_0)$ or

$$\omega\left(\frac{z}{\mu}\right) - \gamma \leq p\left(\frac{z}{\mu} - x_0\right)$$

Put $z_2 = \frac{z}{\mu}$ to get $\gamma \geq \omega(z_2) - p(z_2 - x_0)$ for all $z_2 \in Z$. Combining this condition with the one we obtained when $\lambda > 0$, we can make the choice if

$$\omega(z_2) - p(z_2 - x_0) \leq p(z_1 + x_0) - \omega(z_1) \quad \forall z_1, z_2 \in Z$$

However, observe that

$$\begin{aligned} \omega(z_1 + z_2) &\leq p(z_1 + z_2) = p((z_1 + x_0) + (z_2 - x_0)) \\ &\leq p(z_1 + x_0) + p(z_2 - x_0) \end{aligned}$$

which is precisely the condition we needed to satisfy. Thus, it is possible to choose such a γ (there is at least one) in the given interval. Hence, $\phi(x) = \omega(x) + \lambda\gamma$ extends ω on Z to all of Z_0 contradicting the maximality of (Z, ω) . It must therefore follow that $Z = X$. \square

We can modify the result for normed spaces as follows:

Corollary 1.2.9 (Hahn-Banach Theorem for Normed Spaces)

Let X be a normed real vector space, $Y \subset X$ a subspace, and $\psi \in Y^*$. Then ψ extends to a bounded linear functional $\phi \in X^*$ with $\|\phi\| = \|\psi\|$.

Proof. Exercise for the reader; use the Hahn-Banach Theorem above. \square

Notice that we do not require any completeness for the Hahn-Banach Theorem to apply; we merely need a vector space or a normed space over \mathbb{R} . With a minimal amount of effort, we may prove a version of the Hahn-Banach Theorem for vector spaces over \mathbb{C} . We do this now:

Corollary 1.2.10 (Hahn-Banach Theorem for Complex Vector Spaces)

Let X be a vector space over \mathbb{C} , $Y \subset X$ be a complex subspace, and p a seminorm on X . Let $\psi : Y \rightarrow \mathbb{C}$ be a linear functional satisfying $|\psi(y)| \leq p(y)$ for all $y \in Y$. Then ψ extends to $\phi : X \rightarrow \mathbb{C}$ such that $|\phi(x)| \leq p(x)$ for all $x \in X$.

Proof. We can interpret X and Y to also be vector spaces over \mathbb{R} , call them \tilde{X} and \tilde{Y} . In this respect, let $\theta : \tilde{Y} \rightarrow \mathbb{R}$ where $\theta(y) = \operatorname{Re}(\psi(y))$. It is not difficult to check and see that θ is linear over \tilde{Y} .

Additionally, we have $|\theta(y)| \leq |\psi(y)| \leq p(y)$ for all $y \in \tilde{Y}$. By the real form of the Hahn-Banach Theorem, θ extends to $\beta : \tilde{X} \rightarrow \mathbb{R}$ with $\beta(x) \leq p(x)$ for $x \in \tilde{X}$. Here β is \mathbb{R} -linear. Moreover, we have

$$\theta(iy) = \operatorname{Re}(\psi(iy)) = \operatorname{Re}(i\psi(y)) = -\operatorname{Im}(\psi(y))$$

Thus $\psi(y) = \theta(y) - i\theta(iy)$. Define $\phi(x) = \beta(x) - i\beta(ix)$. Because β is \mathbb{R} -linear, then so is ϕ . We now need to check that it is \mathbb{C} -linear so we must show that $\phi(ix) = i\phi(x)$. We have

$$\phi(ix) = \beta(ix) - i\beta(-x) = i\beta(x) + \beta(ix) = i\phi(x)$$

Finally, we must show that $|\phi(x)| \leq p(x)$ for $x \in X$. Since $\phi(x)$ is a complex number, we may write it as $|\phi(x)|e^{it_0}$ for some $0 \leq t_0 \leq 2\pi$. Hence $e^{-it_0}\phi(x) = |\phi(x)|$ so that $\phi(e^{-it_0}x) = |\phi(x)|$. This means that $\phi(e^{-it_0}x)$ is a real number so we must have $\phi(e^{-it_0}x) = \beta(e^{-it_0}x)$. Since $\beta(e^{-it_0}x) \leq p(e^{-it_0}x) = |e^{-it_0}|p(x) = p(x)$ so we have $|\phi(x)| \leq p(x)$ for all $x \in X$ and this completes the proof. □

Let us illustrate the use of the Hahn-Banach Theorem with an example.

Example 1.2.11

Let X be a normed space over \mathbb{R} (or \mathbb{C}) and $Y \subset X$ a subspace. Set $x_0 \in X \setminus Y$. We will use the Hahn-Banach Theorem to demonstrate the existence of a $\phi \in X^*$ so that $\phi|_Y = 0$ and $\phi(x_0) = d(x_0, Y)$ ¹

To see this, set $p(x) = d(x, Y)$. It is not difficult to see that this is a sublinear functional on X . Consider the subspace $S = \{\lambda x_0 : \lambda \in \mathbb{R}\}$ of X and put $\psi(\lambda x_0) = \lambda p(x_0)$. If

¹ $d(x_0, Y) := \inf_{y \in Y} \|x_0 - y\|_X$

$\lambda < 0$ then we must have $\psi(\lambda x_0) < 0$ so that $\psi(\lambda x_0) \leq p(\lambda x_0)$. If $\lambda \geq 0$ then $\psi(\lambda x_0) = \lambda p(x_0) = p(\lambda x_0)$ so p dominates ψ on S . Moreover, it is obvious that $\psi(x)$ is linear and continuous. By the Hahn-Banach Theorem, extend ψ to ϕ on X where $\phi(x) \leq p(x)$. Note that because

$$-\phi(x) = \phi(-x) \leq p(-x) = \inf_{y \in Y} \| -x - y \| = \inf_{z \in Y} \| -x + z \| = p(x)$$

we get $|\phi(x)| \leq p(x)$ so that ϕ vanishes on Y . ◇

The following result allows us to separate subspaces of a Banach space by a linear functional.

Theorem 1.2.12 (Hahn-Banach Separation Theorem)

Let X be a normed space over \mathbb{C} and $A, B \subseteq X$ be convex, nonempty, disjoint sets with A open. Then there is a $\phi \in X^*$ so that

$$\operatorname{Re} \phi(a) < \operatorname{Re} \phi(b) \quad \forall a \in A, b \in B$$

Proof. Firstly, assume that X is over \mathbb{R} . Fix an $a_0 \in A$ and $b_0 \in B$ and put

$$C := A - B + (b_0 - a_0) = \{a - b + b_0 - a_0 : a \in A \text{ and } b \in B\}$$

Since A and B are convex, we see that C is also convex. Moreover, $0 \in C$. Since A is open, it follows that C is also open because

$$C = \bigcup_{b \in B} (A - b + b_0 - a_0)$$

Let $p(x)$ be the Minkowski functional on C . That is, define

$$p(x) := \inf\{\lambda \geq 0 : x \in \lambda C\}$$

We now demonstrate that

$$C = \{x \in X : p(x) < 1\}$$

Choose $x \in X$ with $p(x) < 1$. We may firstly assume that $x \neq 0$ so that $p(x) > 0$ since we know that $0 \in C$. Then there is a $0 < \lambda_0 < 1$ so that $x \in \lambda_0 C$. Define $y = \frac{1}{\lambda_0}x$ so that $y \in C$. Since C is convex we have $\lambda_0 y + (1 - \lambda_0) \cdot 0 = x \in C$. We now show reverse inclusion. Fix an $x \in C$ and define a map $\eta_x : \mathbb{R} \rightarrow X$ by $\eta_x(t) := tx$. Clearly η_x is continuous. Because $\eta_x(1) = x$ and C is open, we can find an open ball about x contained entirely in C . By continuity, we may find a $\delta > 0$ so that $\eta_x(t) \in C$ when $1 - \delta \leq t \leq 1 + \delta$. In particular, $\eta_x(1 + \delta) \in C$ so that $x \in \frac{1}{1+\delta}C$. This gives

$$p(x) \leq \frac{1}{1 + \delta}$$

so that $p(x) < 1$ as desired.

Now, put $Y := \text{span}(b_0 - a_0)$. On Y , define $\psi : Y \rightarrow \mathbb{R}$ by $\psi(y) = \psi(\lambda(b_0 - a_0)) := \lambda$. Clearly, ψ is linear and we easily note that $\psi(b_0 - a_0) = 1$. Also, we have $b_0 - a_0 \notin C$ because $0 \notin A - B$ as $A \cap B = \emptyset$. By what we did above, $b_0 - a_0 \notin C$ tells us that $p(b_0 - a_0) \geq 1$. Thus $\psi(b_0 - a_0) = 1 \leq p(b_0 - a_0)$ so that if $\lambda \geq 0$, we have

$$\psi(\lambda(b_0 - a_0)) = \lambda\psi(b_0 - a_0) \leq \lambda p(b_0 - a_0) = p(\lambda(b_0 - a_0))$$

If $\lambda < 0$ then $\psi(\lambda(b_0 - a_0)) < 0$ but $p(\lambda(b_0 - a_0)) \geq 0$ so that $\psi(y) \leq p(y)$ for all $y \in Y$. By Hahn-Banach, ψ extends to a linear map $\phi : \mathbb{R} \rightarrow X$ with $\phi|_Y = \psi$ and $\phi(x) \leq p(x)$ for all $x \in X$. By what we have above, $\phi(x) \leq 1$ for all $x \in C$. Thus $|\phi(x)| \leq 1$ for each $x \in C \cap (-C)$ so that ϕ is bounded in an open neighborhood of 0. Thus $|\phi(x)| \leq M$ for all x in a neighborhood of the unit ball of X . Thus $\phi \in X^*$.

Now, for each $a \in A$ and $b \in B$, we get $\phi(a - b + b_0 - a_0) < 1$ so that $\phi(a - b) + \psi(b_0 - a_0) < 1$ or $\phi(a - b) < 0$ so that $\phi(a) < \phi(b)$.

Now, assume that X is over \mathbb{C} . We may interpret X as a vector space over \mathbb{R} so that there is an \mathbb{R} -linear functional $\phi \in X^*$ satisfying $\phi(a) < \phi(b)$ for all $a \in A$ and $b \in B$. Define $\psi(x) := \phi(x) - i\phi(ix)$. Then $\psi(x)$ is our desired functional (we leave it to the reader to verify).

□

1.3 Finite Dimensional Normed Spaces

In this section, we present some basic theory pertaining to finite dimensional normed spaces. In the study of Banach space theory, finite dimensional spaces are rather uninteresting because they satisfy so many nice properties as we shall see. It is when we begin to study infinite dimensional spaces that entirely new and interesting phenomenon occur.

Theorem 1.3.0

Let X be a finite dimensional normed linear space over \mathbb{R} (or \mathbb{C}). Then every closed, bounded set in X is compact.

Proof. Let X be a finite dimensional normed space. Let $\dim(X) = n$ and select a basis for X , say $\{x_1, x_2, \dots, x_n\}$. Define a map $L : \mathbb{R}^n \rightarrow X$ by

$$L(\xi) = \sum_{k=1}^n \xi_k x_k$$

where $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$. We shall work with $\|\cdot\|_\infty$ in \mathbb{R}^n . We know that L is continuous because

$$\|L\xi - L\zeta\|_X = \left\| \sum_{k=1}^n (\xi_k - \zeta_k) x_k \right\|_X \leq \sum_{k=1}^n |\xi_k - \zeta_k| \|x_k\|_X \leq \|\xi - \zeta\|_\infty \sum_{k=1}^n \|x_k\|_X$$

If $Y \subset X$ is closed and bounded then $A := L^{-1}(Y)$ is also closed by continuity of L . Because L is bijective (check), we see that $Y = L(A)$. If we can show that $A \subset \mathbb{R}^n$ is compact, then Y will also be compact because L is continuous.

Consider the boundary of the ℓ_∞ unit ball in \mathbb{R}^n , $\partial B_1^\infty(0)$. Since $f(\xi) = \|L\xi\|_X$ is continuous and $\partial B_1^\infty(0)$ is compact, we know that f attains a minimum, m . Let ξ_0 be such that $\|\xi_0\|_\infty = 1$ and $f(\xi_0) = \|L\xi_0\|_X = m$. Then

$$m = \left\| \sum_{k=1}^n \xi_k^0 x_k \right\|_X$$

Each x_i is linearly independent and $\xi_0 \neq 0$ so we must have $m > 0$. Since Y is bounded, there is a $M > 0$ with $\|y\|_X \leq M$ for all $y \in Y$. If $\xi \in \mathbb{R}^n$ then $\xi/\|\xi\|_\infty$ is a unit vector in ℓ_∞ so we must have

$$\left\| L \left(\frac{\xi}{\|\xi\|_\infty} \right) \right\|_X \geq m$$

so that $\|L\xi\|_X \geq m\|\xi\|_\infty$ (clearly this also holds if $\xi = 0$). For each $\xi \in A$ we have $L\xi \in Y$ so $m\|\xi\|_\infty \leq \|L\xi\| \leq M$ or $\|\xi\|_\infty \leq M/m$. Thus, A is bounded and we are done. Technically, we proved this result for a real normed space but the proof is identical for a complex normed space. □

It turns out that finite dimensional spaces may be completely characterized by this compactness property. In other words, the converse is also true: if X is a normed space and every closed bounded set in X is compact, then X is finite dimensional. To prove this, we first present a lemma.

Lemma 1.3.1 (Riesz's Lemma)

Let X be a normed space and suppose that K is a closed, proper, nontrivial subspace of X . If $0 < \mu < 1$ then there exists a point $x \in X$ so that $\|x\| = 1$ and $d(x, K) > \mu$.

Proof. Since K is proper, we may find an $x_0 \in K^c \cap X$ so that $d(x_0, K) > 0$. Then there exists a $k \in K$ so that $\|x_0 - k\| < \mu^{-1}d(x_0, K)$. Define $x := (x_0 - k)/\|x_0 - k\|$ so that $\|x\| = 1$ and

$$d(x, K) = \frac{d(x_0 - k, K)}{\|x_0 - k\|} = \frac{d(x_0, K)}{\|x_0 - k\|} > \mu$$

which is what we wished to show. □

We are now ready to prove the converse of Theorem 1.3.0.

Theorem 1.3.2

Suppose that X is a normed space and that every closed, bounded subset of X is compact. Then X is finite dimensional.

Proof. It suffices to prove that if the closed unit ball is compact, then X is finite dimensional. Suppose that X is not finite dimensional. Choose an x_1 with $\|x_1\| = 1$ and let $U_1 := \text{span}(x_1)$. Then U_1 is a closed, proper subspace of X . By Riesz's Lemma, choose an x_2 so that $\|x_2\| = 1$ and $d(x_2, U_1) > 1/2$. Then $\|x_2 - x_1\| > 1/2$. Let

$U_2 := \text{span}(x_1, x_2)$. Again, by Riesz's Lemma, choose an x_3 so that $d(x_3, U_2) > 1/2$ so that $\|x_3 - x_1\| > 1/2$ and $\|x_3 - x_2\| > 1/2$. Continue this process inductively to generate a sequence (x_n) so that $\|x_n - x_j\| > 1/2$ whenever $j < n$ and $\|x_n\| = 1$. But then (x_n) is a sequence in the closed unit ball that has no convergent subsequence, a contradiction. \square

We may therefore reformulate a stronger version of Theorem 1.3.0 as follows:

Let X be a normed space over \mathbb{R} or \mathbb{C} . Then X is finite dimensional if and only if every closed bounded subset of X is compact.

We now deduce some interesting corollaries to this statement.

Corollary 1.3.3

Let X be a finite dimensional normed space over \mathbb{R} (or \mathbb{C}). Then all norms on X are equivalent.

Proof. Let $\|\cdot\|$ be any norm on X . Fix a basis $\mathcal{B} := \{x_1, x_2, \dots, x_n\}$ and without loss of generality, assume that $\|x_j\| = 1$ for each j . Then any $x \in X$ has the unique representation $x = \sum_{k=1}^n c_k x_k$ where each $c_k \in \mathbb{R}$ (or \mathbb{C}). Define the norm

$$\|x\|_{\mathcal{B}} := \sum_{k=1}^n |c_k|$$

By uniqueness of basis representations, this is well defined.

Consider $S := \{x \in X : \|x\|_{\mathcal{B}} = 1\}$ and define $f(x) := \|x\|$ on S . By the previous theorem S is closed and bounded and, hence, compact in the topology induced by $\|\cdot\|_{\mathcal{B}}$. We also note that for each $x, y \in X$, we have

$$|f(x) - f(y)| \leq \|x - y\| \leq \|x - y\|_{\mathcal{B}}$$

Thus, $f(x)$ is continuous in the topology induced by $\|\cdot\|_{\mathcal{B}}$. Because S is compact, there exists positive constants m and M such that

$$m \leq \|x\| \leq M \quad \forall x \in S$$

Then for any $x \in X$, we have $x/\|x\|_{\mathcal{B}} \in S$ so that

$$m\|x\|_{\mathcal{B}} \leq \|x\| \leq M\|x\|_{\mathcal{B}}$$

Thus, any two norms are equivalent to $\|\cdot\|_{\mathcal{B}}$ and, hence, each other. □

Norm equivalence tells us that there is one and only one norm topology on X . We shall later see that something stronger is true: there is precisely one linear Hausdorff topology on a finite dimensional space. Of course, this linear T_2 topology must be the norm topology.

The following is an easy corollary.

Corollary 1.3.4

Let X be a finite dimensional normed space over \mathbb{R} (or \mathbb{C}). Then X is a Banach space in any norm that one may impose on X .

Proof. Exercise for the reader. □

We now make a few comments about the dual of a finite dimensional normed space.

Let X be a finite dimensional space where $\dim(X) = n$. Choose a basis for X , say $\mathcal{B} = \{x_1, x_2, \dots, x_n\}$. By Hahn-Banach, there exists a linear functional $\phi_1 \in X^*$ so that $\phi_1(x_j) = \delta_{1j}$. Continue this process to choose finitely many functionals $\phi_i \in X^*$ ($1 \leq i \leq n$) so that $\phi_i(x_j) = \delta_{ij}$ and put

$$\mathcal{B}^* := \{\phi_1, \dots, \phi_n\}$$

It is clear that the functionals in \mathcal{B}^* are linearly independent. Moreover, if $x = \sum_1^n c_k x_k$, then $\phi_j(x) = c_j$ for $1 \leq j \leq n$. Then if $\psi \in X^*$, we see that

$$\psi(x) = \sum_{k=1}^n c_k \psi(x_k) = \sum_{k=1}^n \phi_k(x) \psi(x_k)$$

That is, \mathcal{B}^* forms a basis for the dual space X^* and $\dim(X) = \dim(X^*)$. We refer to \mathcal{B}^* as the *dual basis* for X^* . Things in general are not always so nice and we will see this later. We conclude this section with a simple observation.

Proposition 1.3.5

Let X be a real or complex finite dimensional normed space. Then X is reflexive.

Proof. It is necessary to demonstrate that the canonical embedding is surjective. Let $\psi \in X^{**}$. Then by the discussion above, we may write $\psi = \sum_{k=1}^n c_k \psi_k$ where $\psi_k(\phi_j) = \delta_{kj}$ and the ψ_j 's and ϕ_j 's form the dual bases for X^{**} and X^* respectively. Define $x := \sum_{k=1}^n c_k x_k$. Then if $\phi \in X^*$, write $\phi = \sum_{k=1}^n b_k \phi_k$ so that

$$\psi(\phi) = \sum_{k=1}^n c_k \psi_k(\phi) = \sum_{k=1}^n \sum_{j=1}^n c_k b_j \psi_k(\phi_j) = \sum_{k=1}^n c_k b_k = \phi(x) = J_x(\phi)$$

This is what we wished to show. □

In general, it is not a trivial task to prove that a space is reflexive despite the simplicity of the previous argument. We will later revisit finite dimensional spaces when we discuss weak topologies.

To conclude, we here present a pithy summary of this section: *all finite dimensional normed spaces are norm-equivalent, reflexive Banach spaces that are topologically the same as \mathbb{R}^n .* If life were only this simple in practice.

1.4 Inner-Product and Hilbert Spaces

Recall that a Hilbert space is a special type of Banach space; namely, an inner-product space induces a natural norm $\|x\| = (x, x)^{1/2}$ so a Hilbert space is a Banach space in this particular norm. Note that all Hilbert and inner-product spaces will be taken over \mathbb{C} unless we explicitly specify that they are over \mathbb{R} .

Let's begin with a basic definition.

Definition 1.4.0

Let H be an inner product space. We say that a pair of vectors x_α, x_β for $\alpha, \beta \in A$ and $\alpha \neq \beta$ are *orthogonal* if $(x_\alpha, x_\beta) = 0$. A set of vectors $\{x_\alpha : \alpha \in A\}$ is called *orthonormal* if each vector is orthogonal and $\|x_\alpha\|_H = 1$ for all $\alpha \in A$. □

If two vectors $x, y \in H$ are orthogonal, we will often write $x \perp y$. We now present some basic results.

Proposition 1.4.1 (The Pythagorean Theorem)

Let H be an inner product space and suppose that $\{x_1, x_2, \dots, x_n\} \subset H$ are orthogonal. Then

$$\left\| \sum_{k=1}^n x_k \right\|_H^2 = \sum_{k=1}^n \|x_k\|_H^2$$

Proof. We have

$$\left\| \sum_{k=1}^n x_k \right\|_H^2 = \left(\sum_{k=1}^n x_k, \sum_{k=1}^n x_k \right) = \sum_{h,k} (x_h, x_k) = \sum_{k=1}^n (x_k, x_k)$$

□

Proposition 1.4.2

Let H be an inner product space and suppose that $\{x_1, x_2, \dots, x_n\} \subset H$ are orthonormal. Then

(a) $\left\| \sum_{k=1}^n a_k x_k \right\|_H^2 = \sum_{k=1}^n |a_k|^2$ for all $a_k \in \mathbb{C}$

(b) $\sum_{k=1}^n |(x, x_k)|^2 \leq \|x\|_H^2$ (Bessel's Inequality)

Proof. For (a) we have

$$\begin{aligned} \left\| \sum_{k=1}^n a_k x_k \right\|_H^2 &= \left(\sum_{k=1}^n a_k x_k, \sum_{k=1}^n a_k x_k \right) \\ &= \sum_{k=1}^n \sum_{h=1}^n a_k \bar{a}_h (x_k, x_h) \\ &= \sum_{k=1}^n |a_k|^2 (x_k, x_k) \\ &= \sum_{k=1}^n |a_k|^2 \end{aligned}$$

To obtain (b), set $y = \sum_1^n (x, x_k)x_k$ so that by (a) we have $\|y\|_H^2 = \sum_1^n |(x, x_k)|^2$. Set $z = y - x$ so that for each j , we have

$$(z, x_j) = (y, x_j) - (x, x_j) = (x, x_j) \cdot (x_j, x_j) - (x, x_j) = 0$$

Since $z \perp x_k$ for each k , we must also have $z \perp a$ for each $a \in \text{span}(x_1, x_2, \dots, x_n)$. Hence $z \perp y$. By the Pythagorean Theorem, $\|y - z\|^2 = \|y\|^2 + \|z\|^2$. Hence

$$\|x\|^2 = \|y\|^2 + \|z\|^2 \geq \sum_{k=1}^n |(x, x_k)|^2$$

□

If we are given a finite collection of linearly independent vectors in an inner product space, we may always construct an orthonormal set of vectors that has the same span as our original set of vectors. This is called the Gram-Schmidt orthonormalization process and we discuss this now.

The Gram-Schmidt Orthonormalization Process

Let $\phi_1, \phi_2, \dots, \phi_n$ be linearly independent vectors in an inner product space V . Then there exists n orthonormal vectors $\psi_1, \psi_2, \dots, \psi_n$ so that

$$\text{span}(\phi_1, \dots, \phi_n) = \text{span}(\psi_1, \dots, \psi_n)$$

Proof. We prove by induction. If $n = 1$ then set $\psi_1 = \frac{\phi_1}{\|\phi_1\|_V}$. Assume this is true for some n and let $\phi_1, \dots, \phi_n, \phi_{n+1}$ be linearly independent. Apply the inductive hypothesis to get $\text{span}(\phi_1, \dots, \phi_n) = \text{span}(\psi_1, \dots, \psi_n)$ where ψ_1, \dots, ψ_n are orthonormal. Set

$$s_n(f) = \sum_{k=1}^n (f, \psi_k)\psi_k$$

Clearly, $\phi_{n+1} \notin \text{span}(\psi_1, \dots, \psi_n)$. Set $\varphi = \phi_{n+1} - s_n(\phi_{n+1}) \neq 0$. Then for $1 \leq j \leq n$ we have

$$(\varphi, \psi_j) = (\phi_{n+1}, \psi_j) - \sum_{k=1}^n (\phi_{n+1}, \psi_k)(\psi_k, \psi_j) = (\phi_{n+1}, \psi_j) - (\phi_{n+1}, \psi_j) = 0$$

So φ is orthogonal to each ψ_j . Now let $\psi_{n+1} = \frac{\varphi}{\|\varphi\|_V}$ and we're finished. □

We will now focus our attention on ways of representing elements in a Hilbert space. We begin with a simple convergence result.

Proposition 1.4.3

Let H be a Hilbert space and $(x_k)_{k=1}^\infty$ be an orthonormal sequence in H and let $\alpha_k \in \mathbb{C}$ for all k . Then the sequence $y_n := \sum_1^n \alpha_k x_k$ converges in $\|\cdot\|_H$ if and only if $\sum_1^\infty |\alpha_k|^2 < \infty$.

Proof. Exercise for the reader. □

If $y_n := \sum_1^n \alpha_k x_k$ converges to a $y \in H$, we shall henceforth *denote* its limit y as

$\sum_{k=1}^\infty \alpha_k x_k$. Thus, convergence will always be referring to norm convergence unless it is otherwise clear from the context.

Definition 1.4.4

An *orthonormal basis* for a Hilbert space H is an orthonormal set $X = \{x_\alpha\}_{\alpha \in A}$ such that the linear span, $\text{span}(X)$, is dense in H . □

We will now demonstrate the existence of such a basis for any Hilbert space but we first make an observation that will prove to be useful.

Remark: Let's consider an arbitrary collection of orthonormal vectors $\{x_\alpha\}_{\alpha \in A}$ in H . If we fix an $x \in H$ and let $S_n = \{x_\alpha : |(x, x_\alpha)| \geq \frac{1}{n}\}$ then it turns out that S_n must be finite. If not, let (x_m) be a sequence of distinct points in S_n . Then

$$\frac{1}{n^2} \leq |(x, x_m)|^2 \text{ for all } m$$

By Bessel's Inequality, it follows that

$$\frac{m}{n^2} = \sum_{k=1}^m \frac{1}{n^2} \leq \sum_{k=1}^m |(x, x_m)|^2 \leq \|x\|_H^2$$

for all m , a contradiction. Hence each S_n must be finite. But then, we must have $\bigcup_{n=1}^{\infty} S_n$ is countable so $(x, x_\alpha) = 0$ for all but a countable number of α 's in A .

◇

Theorem 1.4.5

Let H be a Hilbert space. Then the following hold:

- (a) H has an orthonormal basis
- (b) If K is a closed subspace of H then there is an orthonormal basis for H containing an orthonormal basis for K .

Proof. Let $\mathcal{U} := \{X : X \subset H \text{ is orthonormal}\}$. Order \mathcal{U} by set inclusion and select a totally ordered subset $\mathcal{U}_0 = \{X_\alpha : \alpha \in A \text{ and } X_\alpha \text{ is an orthonormal set}\}$. Let us put

$$X_0 := \bigcup_{\alpha \in A} X_\alpha$$

Observe firstly that X_0 is an orthonormal set and a maximal element of \mathcal{U}_0 . Hence by Zorn's Lemma, there is a maximal orthonormal set in H . Denote this orthonormal set as $X := \{x_\alpha : \alpha \in A\}$. We will now demonstrate that X forms an orthonormal basis for H .

Suppose not. Select an $x_0 \in H \setminus \overline{\text{span}(X)}$. Let $S := \{\alpha : (x_0, x_\alpha) = 0\}$ and $T := \{\alpha : (x_0, x_\alpha) \neq 0\}$. Then T is countable by the previous remark so let $\{x_k\}_{k=1}^{\infty}$ be an enumeration of T . By Bessel's Inequality, we have $\sum_{k=1}^n |(x_0, x_k)|^2 \leq \|x_0\|_H^2$ so by

Proposition 1.4.3, there is a $y_0 \in H$ with $y_0 = \sum_{k=1}^{\infty} (x_0, x_k)x_k$. Set $y := x_0 - y_0$ and put

$$s_n := \sum_{k=1}^n (x_0, x_k)x_k$$

For each $\alpha \in S$, we have

$$(y, x_\alpha) = (x_0, x_\alpha) - \lim_{n \rightarrow \infty} \sum_{k=1}^n (x_0, x_k)(x_k, x_\alpha) = 0$$

Similarly, for each fixed j we have

$$(y, x_j) = (x_0, x_j) - \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n (x_0, x_k) x_k, x_j \right) = (x_0, x_j) - (x_0, x_j) = 0$$

So $y \perp x_\alpha$ for all $\alpha \in A$. If $y = 0$, then $\lim_{n \rightarrow \infty} \|x_0 - s_n\|_H = 0$ so $x_0 \in \overline{\text{span}(X)}$, a contradiction. If $y \neq 0$, then $X \cup \left\{ \frac{y}{\|y\|_H} \right\}$ is a larger orthonormal set, contradicting the maximality of X . Hence, $\overline{\text{span}(X)} = H$.

The proof of part (b) is left as an exercise for the reader. □

We will now present a list of equivalent conditions that are extremely useful, especially in Fourier analysis.

Proposition 1.4.6

Let $\{x_k\}_{k=1}^\infty$ be a countable orthonormal set of vectors in a Hilbert space H . Then the following are equivalent:

- (a) $\{x_k\}_{k=1}^\infty$ forms an orthonormal basis for H .
- (b) (Completeness) If $x \in H$ and $(x, x_k) = 0$ for all k then $x = 0$.
- (c) Every $x \in H$ may be uniquely written as $x = \sum_{k=1}^\infty (x, x_k) x_k$
- (d) (Parseval's Identity) $\|x\|_H^2 = \sum_{k=1}^\infty |(x, x_k)|^2$ for all $x \in H$.

Proof. (a) \implies (b). Fix $x \in H$ and suppose that $x \perp x_k$ for all k . Then clearly we get $x \perp s$ for all $s \in \text{span}(X)$. By assumption, there is a sequence (s_n) in the linear span so that $s_n \rightarrow x$ in norm. Hence $0 = (s_n, x) \rightarrow (x, x) = \|x\|_H^2$ so $x = 0$.

(b) \implies (c). Fix an $x \in H$. By Bessel's Inequality, we see that $\sum_{k=1}^n |(x, x_k)|^2 \leq \|x\|_H^2$ for all n . Thus $\sum_{k=1}^\infty |(x, x_k)|^2 < \infty$ so by Proposition 1.4.3, we see that there is a $y \in H$ so that $y = \sum_{k=1}^\infty (x, x_k) x_k$. It is easy to see that $(x - y) \perp x_k$ for all k so by hypothesis, $x = y$. Additionally, if $x = \sum_{k=1}^\infty a_k x_k$ for $a_k \in \mathbb{C}$ for all k , then taking the inner product of both sides gives $a_k = (x, x_k)$.

(c) \implies (d). It easily follows from the definition of convergence that

$$\left\| \sum_{k=1}^n (x, x_k) x_k \right\|_H^2 \rightarrow \|x\|_H^2 \text{ as } n \rightarrow \infty$$

Now, apply Proposition 1.4.2(a) to obtain (d).

(d) \implies (a). Fix $x \in H$. Then for all n we see that

$$\left\| x - \sum_{k=1}^n (x, x_k) x_k \right\|_H^2 = \|x\|_H^2 - \sum_{k=1}^n |(x, x_k)|^2 \rightarrow 0 \text{ as } n \rightarrow \infty$$

This concludes the proof. □

Note that in the above proof, we could have replaced our countable orthonormal basis with an arbitrary one $\{x_\alpha\}_{\alpha \in A}$ because for each x , we know that $(x, x_\alpha) = 0$ for all but countably many $\alpha \in A$.

Proposition 1.4.7

Let H be a Hilbert space and suppose that H has a countable orthonormal basis. Then any other orthonormal basis must be countable.

Proof. Suppose that we have a countable orthonormal basis for a Hilbert space H . Enumerate our orthonormal basis by the sequence (x_n) in H . Suppose that $\mathcal{B} = \{x_\alpha : \alpha \in A\}$ is another orthonormal basis for H . For each $n \in \mathbb{N}$, define

$$A_n := \{\alpha \in A : (x_n, x_\alpha) \neq 0\}$$

Then each A_n must be countable so that $\bigcup_1^\infty A_n$ is countable. We now claim that

$$A = \bigcup_{n=1}^{\infty} A_n$$

Suppose, by contradiction, that $\bigcup_1^\infty A_n$ is strictly contained in A . Let $\beta \in A \setminus \bigcup_1^\infty A_n$. Then $(x_n, x_\beta) = 0$ for all n . Since all of the x_n 's forms an orthonormal basis, we must have $x_\beta = 0$ and the result then follows. □

If we have a closed subspace of a Hilbert space, we can decompose the space into a direct sum of the closed space with the set of all elements orthogonal to it. If $M \subset H$ where H is an inner product space, we set

$$M^\perp := \{x \in H : (x, y) = 0 \text{ for all } y \in M\}$$

By continuity of the inner product and linearity, it follows that M^\perp is a closed subspace of H . We now recall a definition from linear algebra.

If V is a vector space and W_1 and W_2 are subspaces, then we write $V = W_1 \oplus W_2$ if $W_1 \cap W_2 = \{0\}$ and $v = w_1 + w_2$ for $w_1 \in W_1$ and $w_2 \in W_2$. We say that V is the *direct sum* of W_1 and W_2 .

Theorem 1.4.8 (Orthogonal Decomposition of a Hilbert Space)

Let $M \subset H$ be a (topologically) closed subspace of a Hilbert space H . Then $H = M \oplus M^\perp$ and if $x = y + z$ where $y \in M$ and $z \in M^\perp$ then z and y are unique and are at a minimal distance to x .

Proof. We first observe that for any $x, y \in H$ that

$$\|x + y\|_H^2 + \|x - y\|_H^2 = 2\|x\|_H^2 + 2\|y\|_H^2$$

This is called the *parallelogram law* and the proof is not difficult (see the exercises). Set $x \in H$ and let

$$\delta := \inf_{y \in M} \|x - y\|_H$$

We then know that there is a sequence $y_n \in M$ ($\forall n$) so that $\|x - y_n\|_H \rightarrow \delta$. By the parallelogram law, we have

$$\|y_n - y_m\|_H^2 + \|y_n + y_m - 2x\|_H^2 = 2(\|y_n - x\|_H^2 + \|y_m - x\|_H^2)$$

We then manipulate the above equation so that

$$\|y_n - y_m\|_H^2 = 2(\|y_n - x\|_H^2 + \|y_m - x\|_H^2) - 4\left\|\frac{1}{2}(y_n + y_m) - x\right\|_H^2$$

It is important to note that because M is a subspace, we have $(1/2)(y_n + y_m) \in M$. Hence, we get

$$\|y_m - y_n\|_H^2 \leq 2(\|y_n - x\|_H^2 + \|y_m - x\|_H^2) - 4\delta^2$$

The above inequality tells us that (y_n) is Cauchy in $\|\cdot\|_H$ so it must converge to some $y \in M$ since M is closed. Hence $\|x - y\|_H = \delta$.

Put $z := x - y$. Fix $v \in M$. We can always make (z, v) real by multiplying v by an appropriate constant so without loss of generality, we assume this is real. Consider the function

$$g(t) := \|tv + z\|_H^2 = t^2\|v\|_H^2 + 2t(z, v) + \|z\|_H^2$$

for real valued t . $g(t)$ attains its minimum value of δ^2 when $t = 0$ since $tv + z = x - (y - tv)$ but $y - tv$ is in M for any $t \in \mathbb{R}$. Hence $g'(0) = 2(z, v) = 0$ so that $z \in M^\perp$.

If $v \in M \cap M^\perp$ then $(v, m) = 0$ for all $m \in M$. Since $v \in M$, we get $\|v\|_H^2 = (v, v) = 0$ so that $v = 0$. Hence $M \cap M^\perp = \{0\}$. To get uniqueness, let $x = y_0 + z_0$ where $y_0 \in M$ and $z_0 \in M^\perp$. Then $y - y_0 = z - z_0$ are elements in $M \cap M^\perp$ so $y = y_0$ and $z = z_0$.

Finally, we establish that z and y are the elements closest to x . Let $y_0 \in M$. Then

$$\|x - y_0\|_H^2 = \|(x - y) + (y - y_0)\|_H^2 = \|(x - y)\|_H^2 + \|y - y_0\|_H^2 \geq \|x - y\|_H^2$$

since $x - y = z \in M^\perp$ and $y - y_0 \in M$ (apply the Pythagorean Theorem). Notice that we have equality iff $y = y_0$. Exactly the same argument can be used to show that z is the element of M^\perp closest to x . □

The next theorem is an application of this orthogonal decomposition and allows us to characterize all linear, continuous functionals on a Hilbert space H .

Theorem 1.4.9 (Riesz Representation Theorem)

Let H be a Hilbert space and let $\phi \in H^*$. Then there exists a unique $x_0 \in H$ so that $\phi(x) = (x, x_0)$ with $\|\phi\| = \|x_0\|_H$.

Proof. Let $M := \ker(\phi)$. Then M is a closed subspace of H and by the decomposition theorem above, we get $H = M \oplus M^\perp$. If M^\perp is trivial, then $H = M$ so that $\phi = 0$. In this case, let $x_0 = 0$. Assume, then, that M^\perp is nontrivial. Choose $z \in M^\perp$ with $\|z\|_H = 1$ and define $u := \phi(x)z - \phi(z)x$ for $x \in H$. Clearly $u \in M$. Then

$$0 = (u, z) = \phi(x)\|z\|_H^2 - \phi(z)(x, z) = \phi(x) - (x, \overline{\phi(z)}z)$$

so setting $x_0 = \overline{\phi(z)}z$ gives $\phi(x) = (x, x_0)$. Moreover, $\|x_0\|_H = |\phi(z)| \leq \|\phi\|$ and $|\phi(x)| \leq \|x\|_H \cdot \|x_0\|_H$ giving $\|\phi\| \leq \|x_0\|_H$ yielding equality. Finally, if $(x, x_1) = (x, x_2)$ for all $x \in H$, we get $(x, x_1 - x_2) = 0$ for all x so that $x_1 = x_2$ giving uniqueness. \square

We may now completely characterize the dual of any Hilbert space. We have:

Corollary 1.4.10

Let H be a Hilbert space. Then $H \cong H^*$ (we employ the symbol \cong to mean isometrically isomorphic²).

Proof. Clearly each $x_0 \in H$ gives an element of the dual, namely (x, x_0) . The Riesz Representation Theorem tells us that we may identify elements of the dual with elements of H and that the identification is actually surjective. Moreover, the fact that $\|\phi\| = \|x_0\|_H$ for $\phi \in H^*$ tells us that this relationship is an isometry. \square

Remark: It is somewhat customary to write $H = H^*$ instead of $H \cong H^*$. Throughout these notes, I shall employ the notation $H \cong H^*$. \diamond

Corollary 1.4.11 (Reflexivity of Hilbert Space)

Every Hilbert space is reflexive.

Proof. Define $S : H^* \rightarrow H$ by $S(\phi) = x_0$ where x_0 is the unique element in H so that $\phi(x) = (x, x_0)_H$. It is obvious that H^* is a Banach space but we also contend that it is a Hilbert space. Define $(\phi, \varphi) := (S\varphi, S\phi)_H$. It is clear that (\cdot, \cdot) is an inner product on H^* . This inner product induces the operator norm because

$$(\phi, \phi) = (S\phi, S\phi)_H = \|S\phi\|_H^2 = \|x_0\|_H^2 = \|\phi\|^2$$

²Two spaces X and Y are isometrically isomorphic if there exists a bounded, linear, bijective isometry $T : X \rightarrow Y$.

Let $\psi \in H^{**}$. H^* is a Hilbert space as we have just seen so by the Riesz Representation Theorem, there is a $\phi_0 \in H^*$ so that $\psi(\phi) = (\phi, \phi_0) = (S\phi_0, S\phi)_H$. But $\phi(x) = (x, x_0)_H$ where $x_0 = S\phi$. Set $z_0 := S\phi_0$. Then

$$\psi(\phi) = (S\phi_0, S\phi)_H = (z_0, x_0)_H = \phi(z_0) = J_{z_0}(\phi)$$

Thus the canonical embedding is surjective and we are finished. □

We conclude this section by proving a slight generalization of the Riesz Representation Theorem for real Hilbert spaces that has a wide variety of uses, particularly in the study of partial differential equations³.

Definition 1.4.12

Let X be a real or complex normed space. A *bilinear form* is a map $B : X \times X \rightarrow \mathbb{R}$ that is linear in both variables. A bilinear form is continuous (or bounded) if there is a constant $M \geq 0$ so that

$$|B(u, v)| \leq M\|u\|_X\|v\|_X$$

A bilinear form is called *coercive* on a subspace $Y \subseteq X$ if there is a $C > 0$ so that

$$B(v, v) \geq C\|v\|_X^2$$

Finally, a bilinear form is called *symmetric* if $B(x, y) = B(y, x)$ for all $x, y \in X$. □

At this point, we will attempt to demonstrate that if $\phi \in H^*$ and B is a bounded, coercive bilinear form, then there exists a unique $u_0 \in H$ so that $\phi(v) = B(v, u_0)$ for all $v \in H$. It is not difficult to prove that a unique solution to this problem exists when B is symmetric.

This is because $B(u, v)$ forms an inner product in this case (verify) in which H is complete (again, verify). The main difficulty in solving this problem occurs when our bilinear form fails to be symmetric. This, however, is of little consequence in the long run due to the following result.

³As someone with an interest in PDEs, it seems almost immoral to leave out the Lax-Milgram theorem whilst discussing Hilbert spaces.

Theorem 1.4.13 (Lax-Milgram)

Let H be a real Hilbert space and B be a continuous, coercive, bilinear form on H . Fix a $\phi \in H^*$. Then there is a unique $u_0 \in H$ so that $B(u_0, v) = \phi(v)$ for all $v \in H$.

Proof. If $u \in H$, define a functional $T_u(v) = B(u, v)$. Then for each u , $T_u \in H^*$ (verify). Then by the Riesz Representation Theorem, there is a $u_0 \in H$ so that $T_u(v) = B(u, v) = (u_0, v)$ for all $v \in H$. For each $u \in H$, define an operator $T : H \rightarrow H$ by $T(u) := u_0$. Then by definition, $T_u(v) = (Tu, v) = B(u, v)$ for all $v \in H$. We will now demonstrate that $T \in \mathcal{L}(H)$. Firstly, observe that when $c_1, c_2 \in \mathbb{R}$ and $u_1, u_2 \in H$, we get

$$\begin{aligned} (T(c_1u_1 + c_2u_2), v) &= B(c_1u_1 + c_2u_2, v) = c_1B(u_1, v) + c_2B(u_2, v) \\ &= c_1(Tu_1, v) + c_2(Tu_2, v) = (c_1Tu_1 + c_2Tu_2, v) \end{aligned}$$

This holds for all $v \in H$ to give linearity. By definition, we get

$$\|Tu\|_H^2 = B(u, Tu) \leq M\|u\|_H\|Tu\|_H$$

for some $M \geq 0$ so that $\|Tu\|_H \leq M\|u\|_H$ and T is bounded. Furthermore, by coercivity we get

$$C\|u\|_H^2 \leq B(u, u) = (Tu, u) \leq \|Tu\|_H\|u\|_H$$

so that $C\|u\|_H \leq \|Tu\|_H$ (here, C is the coercivity constant). This obviously tells us that T is injective. Moreover, if (y_n) is a sequence in $R(T)$ (the range of T) where $y_n \rightarrow y_0$ in norm, then there is a sequence (x_n) in H so that $y_n = T(x_n)$. Since y_n converges, it is Cauchy. Then for any $\epsilon > 0$ there is an N_0 so that

$$C\|x_n - x_m\|_H \leq \|T(x_n) - T(x_m)\|_H < \epsilon$$

when $n, m \geq N_0$. That is, (x_n) is Cauchy in H so $x_n \rightarrow x_0$ for some $x_0 \in H$. Then $y_n \rightarrow T(x_0)$ so that $y_0 = T(x_0)$ and $R(T)$ is a closed subspace of H . By the orthogonal decomposition theorem, $H = R(T) \oplus R(T)^\perp$. We claim that $R(T) = H$. To see this, let $u_0 \in R(T)^\perp$. Then $C\|u_0\|_H^2 \leq B(u_0, u_0) = (Tu_0, u_0) = 0$ so that $R(T)^\perp = \{0\}$.

By the Riesz Representation Theorem, there is a $v_0 \in H$ so that $\phi(v) = (v_0, v)$ for all $v \in H$. Since $R(T) = H$, we can find a $u_0 \in H$ so that $T(u_0) = v_0$. Then

$$B(u_0, v) = (Tu_0, v) = (v_0, v) = \phi(v)$$

It remains to show that this u_0 is unique. Suppose that there are two solutions u_1 and u_2 . Then $B(u_1, v) = B(u_2, v)$ for all $v \in H$. Then $B(u_1 - u_2, v) = 0$ for all $v \in H$. In particular,

$$C\|u_1 - u_2\|_H^2 \leq B(u_1 - u_2, u_1 - u_2) = 0$$

so that $u_1 = u_2$. This completes the proof. □

We remark that the solution to the variational problem satisfies

$$\|u_0\|_H \leq \frac{1}{C}\|\phi\|$$

We leave this trivial estimate as an exercise for the interested reader.

1.5 Exercises

All normed linear spaces are over \mathbb{C} unless otherwise specified.

1. Let (X, ρ) be a metric space. Prove that (X, σ) is a metric space where

$$\sigma(x, y) := \frac{\rho(x, y)}{1 + \rho(x, y)}$$

In addition, show that (X, ρ) is complete if and only if (X, σ) is complete.

2. Exhibit a concrete example of a metric space whose metric does not induce a norm.
3. On a normed space X , prove that the norm $\|\cdot\|_X$ is a continuous function.

In exercises 4-6, you will consider the space of k -times continuously differentiable functions $f : [a, b] \rightarrow \mathbb{R}$. We shall denote this space as $C^k[a, b]$ where $k \in \mathbb{Z}^+$. Here, $C^0[a, b] = C[a, b]$ simply denotes the space of continuous real valued functions on $[a, b]$.

4. Consider $C[a, b]$ and let

$$\|f\|_{C[a,b]} := \max_{a \leq x \leq b} |f(x)|$$

- (a) Prove that $\|\cdot\|_{C[a,b]}$ defines a norm and that $C[a, b]$ equipped with this norm forms a separable Banach space.
- (b) Show that this norm does not induce an inner-product so that $C[a, b]$ cannot be a Hilbert space. *Suggestion:* Consider $f(x) \equiv 1$ and $g(x) = \frac{x-a}{b-a}$ and the parallelogram law for inner-product spaces (see Exercise 10).

5. Consider $C^1[a, b]$. Show that

$$\|f\|_{C^1[a,b]} := \|f\|_{C[a,b]} + \|f'\|_{C[a,b]}$$

defines a norm on $C^1[a, b]$ and that $C^1[a, b]$ is a Banach space if equipped with this norm. Here $\|\cdot\|_{C[a,b]}$ is the norm given in the previous exercise. How might we generalize this for $C^k[a, b]$?

6. Consider $C[a, b]$ and define

$$\|f\| := \int_a^b |f(x)| dx$$

Prove that this defines a norm on $C[a, b]$ but that $C[a, b]$ with this norm does not form a Banach space.

7. Let ℓ^∞ be the space of all bounded sequences and put

$$\|s_n\|_{\ell^\infty} := \sup_{n \in \mathbb{N}} |s_n|$$

Prove that this is a non-separable Banach space. *Hint:* Consider the space of all sequences with either a 0 or 1 as an entry.

8. Prove from scratch the Hahn-Banach theorem for finite dimensional real normed spaces without using Zorn's Lemma.

9. Let H be an inner product space and let $x_1, x_2, \dots, x_n \in H$ be nonzero orthonormal vectors. Prove that they are linearly independent.

10. Let H be an inner-product space and let $x, y \in H$. Prove that

$$\|x + y\|_H^2 + \|x - y\|_H^2 = 2 \cdot \|x\|_H^2 + 2 \cdot \|y\|_H^2$$

This is called the *parallelogram law* for inner-product spaces.

11. Let ℓ^2 be the space of all complex sequences s_n with the property that $\sum_{n=1}^{\infty} |s_n|^2 < \infty$. Show that ℓ^2 is a vector space and an inner product space with inner product

$$(z_n, w_n)_{\ell^2} := \sum_{n=1}^{\infty} z_n \overline{w_n}$$

In addition, show that ℓ^2 is a separable Hilbert space.

12. In this exercise, you will demonstrate that every finite dimensional normed space is complete regardless of the norm imposed on the space. There are several different methods of accomplishing this task and you will explore two such ways here.

(a) Prove this by using norm equivalence.

- (b) Prove this by induction on dimension. To do this, you may find it helpful to utilize the example presented after the proof of the Hahn-Banach Theorem.
13. Prove Proposition 1.4.3.
 14. Prove that if H is a Hilbert space and $K \subset H$ is a closed subspace, then H has an orthonormal basis consisting of an orthonormal basis for K .
 15. Prove the Hahn-Banach Theorem for normed spaces.
 16. Let X be a normed space and let $x_0 \in X$ with $x_0 \neq 0$. Then there exists a $\phi \in X^*$ such that $\|\phi\| = 1$ and $\phi(x_0) = \|x_0\|_X$.
 17. Let X and Y be normed spaces and $T \in \mathcal{L}(X, Y)$. Prove that $T^* \in \mathcal{L}(Y^*, X^*)$ with $\|T\| = \|T^*\|$.
 18. Prove that a Hilbert space H is separable if and only if H has a countable orthonormal basis.
 19. Let X be a real or complex normed space. Prove that the canonical embedding of X into X^{**} is a continuous, injective isometry.
 20. Prove the conclusion of the Lax-Milgram Theorem when we make the stronger assumption that $B(x, y)$ is symmetric.

2 The Cornerstone Theorems

In functional analysis, there are a handful of results that are foundational to the subject. We have already seen one such result in the previous section: the Hahn-Banach Theorem. In this section, we prove the remaining results.

We first elaborate on notation. If X is a normed space, let

$$B_X(x_0, r) := \{x \in X : \|x - x_0\|_X < r\}$$

be an open ball centered at x_0 with radius r . If the underlying space X is understood, we will write $B(x_0, r)$. If $x_0 = 0$ and $r = 1$, we will simply write B or B_X .

2.1 The Baire Category Theorem

The first result that we will demonstrate will be needed later and it is not difficult to prove.

Theorem 2.1.0 (Baire Category Theorem)

Let (X, ρ) be a complete metric space and let V_n be open, dense sets in X . Then $\bigcap_{n=1}^{\infty} V_n$ is dense in X .

Proof. Consider an arbitrary open ball $B(x_0, r) := \{x \in X : \rho(x, x_0) < r\}$ in X . We will show that

$$\left(\bigcap_{n=1}^{\infty} V_n \right) \cap B(x_0, r) \neq \emptyset$$

Now V_1 is open and dense so there is a ball $B(x_1, 2r_1) \subset V_1 \cap B(x_0, r)$ with $r_1 < 1$. In a similar manner, V_2 is open and dense so there is a ball $B(x_2, 2r_2) \subset V_2 \cap B(x_1, r_1)$ with $r_2 < \frac{1}{2}$. In general, V_j is open and dense so there is a ball $B(x_j, 2r_j) \subset V_j \cap B(x_{j-1}, r_{j-1})$ with $r_j < \frac{1}{j}$. If $n > m$, then $\rho(x_n, x_m) < r_m < \frac{1}{m}$ and if $n \leq m$ then $\rho(x_n, x_m) < r_n < \frac{1}{n}$ so that (x_n) is Cauchy. Thus, there is a $y \in X$ so that $x_n \rightarrow y$.

Note that $y \in \overline{B(x_n, r_n)}$ for all n because $\overline{B(x_n, r_n)}$ contains all but the first $n - 1$ terms of the sequence. But then

$$y \in \overline{B(x_n, r_n)} \subset B(x_n, 2r_n) \subset \left(\bigcap_{k=1}^n V_k \right) \cap B(x_0, r) \quad \forall n$$

□

We may interpret this in the following way: a set is called *nowhere dense* if $\text{int } \overline{A} = \emptyset$. The Baire Category Theorem then tells us that a complete metric space cannot be written as a countable union of nowhere dense sets. For if $X = \bigcup_1^\infty A_k$ where $\text{int } \overline{A_k} = \emptyset$ for all k , then $\emptyset = \bigcap_1^\infty (\overline{A_k})^c$. But each $(\overline{A_k})^c$ is open and dense which is a contradiction.

Definition 2.1.1

Let (X, \mathcal{T}) be a topological space. A set in X is said to be of *first category* if it can be written as a countable union of nowhere dense sets. If a set is not first category, it is said to be of *second category*.

□

Thus, the Baire Category Theorem says the following: *Every complete metric space is of second category.*

2.2 The Open Mapping Theorem

Most of the cornerstone theorems are simple applications of the Baire Category Theorem. Yet, these theorems are so important and so widely used that we devote an entire section to them. We begin with the open mapping theorem.

If $f : X \rightarrow Y$ where X and Y are topological spaces, we say that f is an *open (closed) mapping* if $f(U)$ is open (closed) in Y whenever U is open (closed) in X . From basic complex function theory, we know that every nonconstant holomorphic function is an open mapping. We prove a similar result for Banach Spaces but first, we need a lemma.

Lemma 2.2.0

Let $T \in \mathcal{L}(X, Y)$ where X is a Banach space and Y is a normed space. Assume that for some $\epsilon > 0$ we get $\epsilon B_Y \subseteq \overline{T(B_X)}$. Then $\overline{T(B_X)} \subseteq 2T(B_X)$.

Proof. Let $y \in \overline{T(B_X)}$. We will find an $x \in 2B_X$ so that $Tx = y$. There exists an $x_1 \in B_X$ so that $\|Tx_1 - y\|_Y < \epsilon/2$. Then $y - Tx_1 \in \frac{\epsilon}{2}B_Y \subseteq \frac{1}{2}\overline{T(B_X)}$ by assumption. Likewise, there is an $x_2 \in \frac{1}{2}B_X$ so that

$$\|y - Tx_1 - Tx_2\|_Y < \frac{\epsilon}{2^2} \quad \text{so that} \quad y - Tx_1 - Tx_2 \in \frac{\epsilon}{2^2}B_Y \subseteq \frac{1}{2^2}\overline{T(B_X)}$$

Continue this construction inductively to obtain a sequence (x_n) with $\|x_n\|_X < \frac{1}{2^{n-1}}$ for all n . Thus $\sum_{k=1}^{\infty} \|x_k\|_X < \infty$ so that $s_n := \sum_{k=1}^n x_k \rightarrow x \in X$ in norm.

Note also that $\|x\|_X \leq \sum \|x_k\|_X \leq 2$. Hence

$$\left\| y - T \left(\sum_{k=1}^n x_k \right) \right\|_X < \frac{\epsilon}{2^n} \rightarrow 0$$

as $n \rightarrow \infty$. Thus $y = Tx$. □

Theorem 2.2.1 (Open Mapping Theorem)

Let X and Y be Banach spaces. If $T : X \rightarrow Y$ is a bounded, linear, surjective operator, then T is open.

Proof. Observe that

$$Y = T(X) = \bigcup_{k=1}^{\infty} T(kB_X) = \bigcup_{k=1}^{\infty} \overline{T(kB_X)}$$

By the Baire Category Theorem, there exists an n so that $\text{int } \overline{T(nB_X)} \neq \emptyset$. Thus, there is an open $W \subset \overline{T(nB_X)}$. Choose $y_0 \in W$ and $\eta > 0$ so that $y_0 + y \in W$ whenever $\|y\|_Y < \eta$. Then, there are sequences (a_j) and (b_j) in nB_X so that $Ta_j \rightarrow y_0$ and $Tb_j \rightarrow y_0 + y$ as $j \rightarrow \infty$. Setting $c_j := b_j - a_j$ gives $\|c_j\|_X < 2n$ with $Tc_j \rightarrow y$. Thus, for all $y \in \eta B_Y$, there is a sequence (c_j) in $2nB_X$ so that $Tc_j \rightarrow y$. In other words, $\eta B_Y \subseteq \overline{T(2nB_X)}$. Then, we may choose an appropriate $\epsilon > 0$ so that $\epsilon B_Y \subseteq \overline{T(B_X)}$.

Let U be an arbitrary open set in X . For $x \in U$, we have $Tx \in T(U)$. Choose a $\delta > 0$ so that $x + \delta B_X \subseteq U$ and $Tx + \delta T(B_X) = T(x + \delta B_X) \subset T(U)$. Then by the Lemma above, we have

$$Tx + \frac{\delta\epsilon}{2}B_Y \subseteq Tx + \frac{\delta}{2}\overline{T(B_X)} \subseteq Tx + \delta T(B_X) \subseteq T(U)$$

This is what we wished to show. □

We obtain the following useful result as an easy corollary.

Corollary 2.2.2 (Inverse Mapping Theorem)

Let X and Y be Banach spaces where $T : X \rightarrow Y$ is a bounded, linear bijection. Then $T^{-1} \in \mathcal{L}(Y, X)$.

Proof. We know an inverse exists and it is not terribly difficult to check to see that T^{-1} is linear. Continuity of T^{-1} is equivalent to openness of T so we are finished. □

2.3 Uniformly Bounded Operators and the Closed Graph Theorem

In this section, we prove the remaining two cornerstone theorems: the uniform boundedness principle and the closed graph theorem. We begin with the uniform boundedness principle.

Theorem 2.3.0 (Uniform Boundedness Principle/Banach-Steinhaus Theorem)

Let X be a Banach space and Y be a normed space. Let $\{T_\alpha\}$ be a collection of bounded, linear operators from X to Y . Then

$$\sup_{\alpha} \|T_\alpha x\|_Y < \infty \quad \forall x \in X \quad \text{implies} \quad \sup_{\alpha} \|T_\alpha\| < \infty$$

Proof. Write

$$\begin{aligned}
X &= \bigcup_{n=1}^{\infty} \Omega_n \quad \text{where } \Omega_n := \{x \in X : \sup_{\alpha} \|T_{\alpha}x\|_Y \leq n\} \\
&= \bigcap_{\alpha} \{x \in X : \|T_{\alpha}x\|_Y \leq n\}
\end{aligned}$$

Clearly each Ω_n is closed. By the Baire Category Theorem, there is an m so that $\text{int } \Omega_m \neq \emptyset$. Then there is an $x_0 \in X$ and an $\epsilon > 0$ so that $B(x_0, \epsilon) \subseteq \Omega_m$. For all $y \in X$ with $\|y\|_X = 1$, we get $x_0 + (\epsilon/2)y \in B(x_0, \epsilon)$ so that

$$\frac{\epsilon}{2} \|T_{\alpha}y\|_Y - \|T_{\alpha}x_0\|_Y \leq \|T_{\alpha}(x_0 + (\epsilon/2)y)\|_Y \leq m \quad \forall \alpha$$

But this implies that

$$\|T_{\alpha}\| \leq \frac{2}{\epsilon} \left(m + \sup_{\beta} \|T_{\beta}x_0\|_Y \right) < \infty \quad \forall \alpha$$

□

We now move on to the closed graph theorem. Firstly, note that if X and Y are normed spaces, then we may make the product space $X \times Y$ a normed space via

$$\|(x, y)\|_p := \begin{cases} (\|x\|_X^p + \|y\|_Y^p)^{1/p} & 1 \leq p < \infty \\ \max(\|x\|_X, \|y\|_Y) & p = \infty \end{cases}$$

It does not matter which norm we choose because they each induce the same topology (verify). Our last cornerstone theorem is useful when attempting to show that an operator is bounded. First, we need a definition.

Definition 2.3.1

Let $T : X \rightarrow Y$ be an operator. Define the *graph of T* to be

$$G(T) := \{(x, Tx) : x \in X\} \subseteq X \times Y$$

□

Theorem 2.3.2 (Closed Graph Theorem)

Let X and Y be Banach spaces and suppose that $T : X \rightarrow Y$ is a linear operator. If the graph of T , $G(T)$, is a closed subset of $X \times Y$, then T is bounded.

Proof. Let $\pi_X : X \times Y \rightarrow X$ be the projection map $\pi_X(x, y) = x$ onto X . Similarly, let π_Y be the projection map from $X \times Y$ onto Y . Now

$$\|\pi_X(x, y)\|_X = \|x\|_X \leq \|(x, y)\|_2$$

so that π_X is bounded and linear. Similarly for π_Y . Consider $p_X := \pi_X|_G$ where $G = G(T)$ is the graph. Since G is closed, it is a Banach space. Clearly, p_X is bijective and since it is continuous, the inverse mapping theorem tells us that p_X^{-1} is a bounded, linear operator. But $p_X^{-1}(x) = (x, Tx)$. Hence $T = \pi_Y \circ p_X^{-1}$ so that T is continuous. \square

2.4 Exercises

1. Let (X, \mathcal{T}) be a topological space. Show that a set $A \subset X$ is nowhere dense if and only if $(\overline{A})^c$ is dense in X .
2. Suppose that $T : X \rightarrow Y$ is linear and that $\phi \circ T \in X^*$ for all $\phi \in Y^*$. Prove that T is bounded.
3. Suppose that X and Y are Banach spaces and that (T_n) is a sequence in $\mathcal{L}(X, Y)$ so that $(T_n x)$ converges in Y for all $x \in X$. If $Tx := \lim T_n x$, then prove that T is a bounded, linear operator.
4. Let X and Y be normed spaces. Show that each $\|(x, y)\|_p$ are equivalent norms on $X \times Y$ for $1 \leq p \leq \infty$. Additionally, if X and Y are Banach spaces, show that $X \times Y$ is a Banach space in $\|(x, y)\|_p$ for $1 \leq p \leq \infty$.
5. Let X be a normed space that is complete in norms $\|\cdot\|_1$ and $\|\cdot\|_2$ with $\|x\|_1 \leq \|x\|_2$ for all $x \in X$. Show that $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.
6. Let X and Y be Banach spaces with $X \cong Y$. Prove that $X^* \cong Y^*$ (*Suggestion:* Consider the adjoint of our isometry from X to Y).

3 Weak and Weak* Topologies

In the last few sections, we have primarily concerned ourselves with the norm topology on a normed space X . However, there are many times when it may be more convenient to consider other topologies on X aside from the norm topology. In this section, we shall introduce the weak and weak* topologies on a normed vector space. For our purposes, weak topologies can tell us quite a few things about reflexivity. It is assumed that the reader is familiar with the basics of general point-set topology.

3.1 Basic Ideas

We first introduce some notation. All vector spaces we consider will either be over \mathbb{R} or \mathbb{C} . We shall, therefore, employ the notation \mathbb{F} to denote either \mathbb{R} or \mathbb{C} . If it is necessary to make the field explicit, we shall do so.

Definition 3.1.0

Let X be a vector space over \mathbb{F} . A *topological vector space* (TVS or linear topology) is a Hausdorff topology on X so that addition and scalar multiplication are continuous on $X \times X$ and $\mathbb{F} \times X$ respectively. □

Some authors require that a TVS be T_1 rather than T_2 but most linear topologies encountered in practice are T_2 .

Definition 3.1.1

A set $V \subset X$ is called *balanced* (or circled) if for any $x \in V$ and $\lambda \in \mathbb{F}$ with $|\lambda| \leq 1$, we get $\lambda x \in V$. □

The following is easy but is quite useful:

Proposition 3.1.2

Let X be a topological vector space. Then every neighborhood V of 0 contains a balanced neighborhood of 0.

Proof. Let V be a neighborhood of 0. By continuity of $(\alpha, x) \mapsto \alpha x$ at $(0, 0)$ we are guaranteed the existence of a neighborhood W of 0 and some $\lambda_0 > 0$ so that $|\lambda| \leq \lambda_0$ and $x \in W$ implies $\lambda x \in V$. Set

$$U := \bigcup_{|\lambda| \leq \lambda_0} \lambda W$$

It is clear that U is a neighborhood of 0 and $U \subset V$. Also, if $\alpha \in \mathbb{R}$ (or \mathbb{C}) with $|\alpha| \leq 1$ and $y \in U$, then $y = \lambda_1 y_1$ where $y_1 \in W$ and $|\lambda_1| \leq \lambda_0$. Then, $\alpha y = \alpha \lambda_1 y_1 \in U$ as $|\alpha \lambda_1| \leq \lambda_0$. □

There is much more we can say about linear topologies but we withhold these things until or unless we need them. The concept of a weak topology is actually fairly general so we here present the general definition.

Definition 3.1.3 (Weak Topologies)

Let X be a set and $(X_\alpha, \mathcal{T}_\alpha)_{\alpha \in A}$ a collection of topological spaces. For each $\alpha \in A$, let $f_\alpha : X \rightarrow X_\alpha$ and set $\mathcal{F} := \{f_\alpha : \alpha \in A\}$. The *weak topology* on X induced by \mathcal{F} is the weakest topology on X making each $f \in \mathcal{F}$ continuous. □

We will now present a fairly standard illustration of the weak topology in a general setting.

Example (Product Topologies)

Let $(X_\alpha, \mathcal{T}_\alpha)$ be a collection of topological spaces. Normally, we define the product topology \mathcal{T}_w on $X := \prod_\alpha X_\alpha$ in the following way: let \mathcal{B} be a base for \mathcal{T}_w consisting of sets of the form $\prod_\alpha U_\alpha$ where each U_α is open in X_α and $U_\alpha = X_\alpha$ for all but finitely many $\alpha \in A$. Note that each

$$\prod_\alpha U_\alpha = \bigcap_{k=1}^n \pi_{\alpha_k}^{-1}(U_{\alpha_k})$$

where each $\pi_\alpha : X \rightarrow X_\alpha$ is a projection map and $U_\alpha = X_\alpha$ except when $\alpha = \alpha_j$ for some $1 \leq j \leq n$. Thus, the weak topology is the topology whose subbase consists of sets of the form $\pi_\alpha^{-1}(U_\alpha)$.

It is clear that in this topology, each π_α is continuous. However, this topology is the weakest topology that makes the projection maps continuous. That is, if \mathcal{T} is another topology that makes the projection maps continuous, then clearly $\mathcal{T}_w \subset \mathcal{T}$.

It is important to realize that this topology is fundamentally different than the *box topology*. The box topology has as a base all sets of the form $\prod_{\alpha} U_{\alpha}$ where each U_{α} is open (i.e. we omit the condition that all but finitely many must be the factor space). The box topology does, however, coincide with the product topology when the product space has only finitely many factors.

◇

In general, it is easily verified that the weak topology is a topology whose subbase is given by $\mathcal{C} := \{f_{\alpha}^{-1}(U_{\alpha}) : U_{\alpha} \in \mathcal{T}_{\alpha}, \alpha \in A\}$.

3.2 Weak Topologies on Normed Spaces

Before we begin our discussion of weak and weak* topologies, it is necessary to recall a few things. For each $x \in X$, let $J_x = Jx$ be the canonical embedding of X into X^{**} (recall that $(Jx)(\phi) := \phi(x)$ for all $\phi \in X^*$). Then we may identify X with $J(X) \subset X^{**}$ so that $X \hookrightarrow J(X) \subset X^{**}$.

Definition 3.2.0

If X is a normed space, let $\sigma(X, X^*)$ be the weak topology on X induced by X^* . We denote by $\sigma(X^*, X)$ the weak* topology, the weakest topology so that the set of all canonical embeddings $\{J_x : x \in X\}$ are continuous on X^* .

□

Let $\epsilon > 0$ and $F \subset X^*$ be finite. Define

$$V_{\epsilon, F}(x_0) := \{x \in X : |\phi(x) - \phi(x_0)| < \epsilon, \phi \in F\}$$

We then leave it to the reader to verify that an open neighborhood base about $x_0 \in X$ is given by

$$\mathcal{B}_{x_0} := \{V_{\epsilon, F}(x_0) : \epsilon > 0, F \subset X^* \text{ is finite}\}$$

Likewise, if

$$V_{\epsilon, F}^*(\phi_0) := \{\phi \in X^* : |\phi(x) - \phi_0(x)| < \epsilon, x \in F\}$$

where $F \subset X$ is finite and $\epsilon > 0$ then an open neighborhood base about $\phi_0 \in X^*$ is given by

$$\mathcal{B}_{\phi_0}^* := \{V_{\epsilon, F}^*(\phi_0) : \epsilon > 0, F \subset X \text{ is finite}\}$$

We now deduce an elementary but important result concerning convergence in the weak and weak* topology.

Proposition 3.2.1

A net (x_λ) in X converges to x_0 in $\sigma(X, X^*)$ if and only if $\phi(x_\lambda) \rightarrow \phi(x_0)$ for all $\phi \in X^*$. Likewise, a net (ϕ_λ) in X^* converges to ϕ_0 in the weak* topology if and only if $\phi_\lambda(x) \rightarrow \phi_0(x)$ for all $x \in X$.

Proof. Suppose that $x_\lambda \rightarrow x_0$ weakly. Then for all $\epsilon > 0$, $V_{\epsilon, \phi}(x_0)$ is an open neighborhood of x_0 (note that in this case $F = \{\phi\}$ so we omit the set brackets). Then, there is a λ_0 so that $x_\lambda \in V_{\epsilon, \phi}(x_0)$ when $\lambda \succ \lambda_0$. That is, $|\phi(x_\lambda) - \phi(x_0)| < \epsilon$ when $\lambda \succ \lambda_0$ giving the result. Conversely, let V be an open neighborhood of x_0 . Then there is an $\epsilon > 0$ and a finite $F \subset X^*$ so that $V_{\epsilon, F} \subset V$. But $\phi(x_\lambda) \rightarrow \phi(x_0)$ for all $\phi \in X^*$ by assumption. Then, there exists λ_j so that $|\phi_j(x_\lambda) - \phi_j(x_0)| < \epsilon$ when $\lambda \succ \lambda_j$ and $F = \{\phi_1, \dots, \phi_n\}$. Since each λ_j is a member of a directed set, we may take the maximal element of all such λ_j ($1 \leq j \leq n$), call it λ_0 . Then $x_\lambda \in V_{\epsilon, F}(x_0) \subset V$ for all $\lambda \succ \lambda_0$. That is, $x_\lambda \rightarrow x_0$ and we're finished. The proof for weak* convergence is similar. □

If (x_n) is a sequence in a normed space X , we say that x_n *converges strongly* to x_0 if it is norm convergent, i.e. $\|x_n - x_0\|_X \rightarrow 0$ as $n \rightarrow \infty$. It is easily verified that strong sequential convergence implies weak sequential convergence. The following example illustrates that the converse is untrue.

Example (Weak Convergence)

If $\psi \in L^2(0, 2\pi)^*$ then by the Riesz Representation Theorem, there is a $g \in L^2(0, 2\pi)$ so that

$$\psi(f) = \int_0^{2\pi} f(x)g(x)dx$$

Consider the sequence $f_n(x) := \cos(nx)$. By the Riemann-Lebesgue Lemma,

$$\psi(\cos(nx)) = \int_0^{2\pi} \cos(nx)g(x)dx \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

so $f_n \rightarrow 0$ weakly. However, one may easily verify that $\int_0^{2\pi} \cos^2(nx) dx = \pi$ for all n so that f_n does not converge to 0 in L^2 . ◇

Weak convergence, as the name seems to imply, is hence generally weaker than strong convergence. Even so, we may nevertheless say some nice things about weakly convergent sequences.

Proposition 3.2.2

Let (x_n) be a sequence in a normed space X that converges weakly to x_0 . Then:

- (a) (x_n) is bounded in norm.
- (b) $\|x_0\|_X \leq \liminf_{n \rightarrow \infty} \|x_n\|_X$

Proof. For (a), we without loss of generality assume that $x_0 = 0$. Apply the canonical embedding from X to X^{**} to each x_n . Then $J_{x_n}(\phi) = \phi(x_n)$ where $\phi \in X^*$ so that $|J_{x_n}(\phi)| \leq M$ where $M = M(\phi)$. By the uniform boundedness principle, $\|J_{x_n}\| \leq C$ for all n . But $\|J_{x_n}\| = \|x_n\|_X$ giving the result. For (b), use the Hahn-Banach Theorem to choose a $\phi \in X^*$ so that $\phi(x_0) = \|x_0\|_X$ and $\|\phi\| = 1$. Then $|\phi(x_n)| \leq \|x_n\|_X$ for all n . Taking the limit infimum of both sides gives the desired result. □

Remark on Notation: If a sequence (x_n) in X converges weakly to x_0 , we will write $x_n \xrightarrow{w} x_0$. Likewise, if a sequence (ϕ_n) in X^* converges to ϕ_0 in the weak* topology, we will write $\phi_n \xrightarrow{w^*} \phi_0$. ◇

We will now turn our attention to some of the more topological properties of weak and weak* topologies.

Proposition 3.2.3

If X is a normed space, then $\sigma(X, X^*)$ and $\sigma(X^*, X)$ are Hausdorff, topological vector spaces.

Proof. We show that addition is continuous and continuity of scalar multiplication will follow similarly. Let (x_λ, y_λ) be a net that converges to (x_0, y_0) in $X \times X$. Then by considering the projection maps, $x_\lambda \rightarrow x_0$ and $y_\lambda \rightarrow y_0$ in X . Then

$$\phi(x_\lambda + y_\lambda) = \phi(x_\lambda) + \phi(y_\lambda) \rightarrow \phi(x_0) + \phi(y_0) = \phi(x_0 + y_0)$$

That is, $x_\lambda + y_\lambda \rightarrow x_0 + y_0$ weakly so addition is continuous.

To see that it is T_2 , let $x, y \in X$ with $x \neq y$. By the Hahn-Banach Theorem, there is a $\phi \in X^*$ so that $\phi(x - y) > 0$ or $\phi(x) \neq \phi(y)$. As the real line with the standard topology is T_2 , we may find disjoint open intervals I_1 and I_2 in \mathbb{R} containing $\phi(x)$ and $\phi(y)$ respectively. Then $\phi^{-1}(I_1)$ and $\phi^{-1}(I_2)$ are open and do the trick. The proof for $\sigma(X^*, X)$ is analogous and left to the reader. □

We will now prepare to prove our first major result in this section. Our goal is to demonstrate that the weak topology of a normed space is metrizable if and only if X is finite dimensional. The next theorem demonstrates one direction of this.

Theorem 3.2.4 (Uniqueness of T_2 Topologies)

Every finite dimensional normed space admits a unique T_2 linear topology.

Proof. Let $\mathcal{T}_{\|\cdot\|}$ be the norm topology on X . Because the norm topology is T_2 and linear, it suffices to show that any other T_2 linear topology \mathcal{T} must be the norm topology. Fix a basis $\mathcal{B} = \{x_1, \dots, x_n\}$ and without loss of generality (by norm equivalence), assume that

$$\|x\| = \sum_{j=1}^n |c_j| \quad \text{where} \quad x = \sum_{j=1}^n c_j x_j$$

Let (x_λ) be a net in X with $x_\lambda \rightarrow 0$. Then $x_\lambda = \sum_{j=1}^n c_\lambda^{(j)} x_j$ so that $c_\lambda^{(j)} \rightarrow 0$ in \mathbb{R} (or \mathbb{C}). Consider the identity map $I : (X, \mathcal{T}_{\|\cdot\|}) \rightarrow (X, \mathcal{T})$. Then

$$Ix_\lambda = x_\lambda = \sum_{j=1}^n c_\lambda^{(j)} x_j \xrightarrow{\text{in } \mathcal{T}} \sum_{j=1}^n 0x_j = 0$$

because \mathcal{T} is a linear topology. Hence I is continuous and $\mathcal{T} \subset \mathcal{T}_{\|\cdot\|}$. Now, consider the unit ball about the origin in X , $B = B(0, 1)$. Let $S := \{x \in X : \|x\| = 1\}$ be the unit sphere in X . Now X is finite dimensional so S is $\mathcal{T}_{\|\cdot\|}$ -compact. But by the inclusion $\mathcal{T} \subset \mathcal{T}_{\|\cdot\|}$, we see that S is also \mathcal{T} -compact. Because X is T_2 , we see that S is \mathcal{T} -closed⁴.

⁴Recall that a compact subset of a T_2 space is closed.

Now $0 \notin S = \overline{S}$ (where the closure is the \mathcal{T} -closure) so we may find a \mathcal{T} -open set V with $0 \in V$ so that $V \cap S = \emptyset$. We may, without loss of generality, assume that V is balanced (if not, choose a balanced subset). We claim that $V \subset B(0, 1)$. If not, there is an $x \in V$ with $\|x\| \geq 1$. But then $x/\|x\| \in V \cap S$, a contradiction. Thus, $B(0, 1)$ is a \mathcal{T} neighborhood of 0.

If U is a $\mathcal{T}_{\|\cdot\|}$ -open set and $x_0 \in U$, then there is an $r > 0$ so that $B(x_0, r) \subset U$. But $rV + x_0 \subset B(x_0, r) \subset U$ where $rV + x_0$ is \mathcal{T} -open. Hence U is \mathcal{T} -open so that $\mathcal{T}_{\|\cdot\|} \subset \mathcal{T}$ yielding equality. □

We are almost ready to prove our metrizable theorem but first, we need a lemma.

Lemma 3.2.5

Let $\phi_1, \phi_2, \dots, \phi_n \in X^*$ where X is a normed space. Then $\phi \in \text{span}(\phi_1, \dots, \phi_n)$ if and only if

$$\bigcap_{j=1}^n \ker \phi_j \subset \ker \phi$$

Proof. The direction \implies is trivial. Assume, then, that $\bigcap_{j=1}^n \ker \phi_j \subset \ker \phi$. Define $\Psi : X \rightarrow \mathbb{R}^n$ by $\Psi(x) := (\phi_1(x), \dots, \phi_n(x))$. On the range $\Psi(X)$, set $\varphi : \Psi(X) \rightarrow \mathbb{R}$ by $\varphi(\phi_1(x), \dots, \phi_n(x)) := \phi(x)$. If $y_j = (\phi_1(x_j), \dots, \phi_n(x_j))$ for $j = 1, 2$ so that $y_1 = y_2$ then $\phi(x_1) = \phi(x_2)$. Thus φ is well defined.

Clearly $\Psi(X) \subset \mathbb{R}^n$ is a subspace. Then φ extends to a $\tilde{\varphi}$ on all of \mathbb{R}^n (in a linear algebra sense by tweaking basis values). But then there is a $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ so that $\tilde{\varphi}(y) = (\lambda, y)$ for all $y \in \mathbb{R}^n$ (here, $(x, y) = \sum_1^n x_k y_k$ is the inner product on \mathbb{R}^n). This holds, in particular, when $y \in \Psi(X)$. □

We are now ready to demonstrate our theorem.

Theorem 3.2.6 (Metrizability of the Weak Topology)

Let X be a normed space. Then the weak topology on X is metrizable if and only if X is finite dimensional.

Proof. Suppose that $\dim(X) < \infty$. Because $\sigma(X, X^*)$ is a T_2 linear topology, the uniqueness theorem for T_2 linear topologies tells us that $\sigma(X, X^*) = \mathcal{T}_{\|\cdot\|}$. Hence, the weak topology is metrizable.

Now, assume that the weak topology is metrizable and that X is not finite dimensional. Then the weak topology is first countable. Let $\mathcal{B}_0 := \{B_i : i \in \mathbb{N}\}$ be a countable neighborhood base at 0. Then there is an $\epsilon_1 > 0$ and finite $F_1 \subset X^*$ so that $V_{\epsilon_1, F_1}(0) \subset B_1$. We may assume that $\epsilon_1 = 1$ by replacing each $\phi \in F_1$ by $(1/\epsilon_1)\phi$. Let $V_1 := V_{\epsilon_1, F_1}(0)$. Likewise, there is an $\epsilon_2 > 0$ and finite $F_2 \subset X^*$ so that $V_{\epsilon_2, F_2}(0) \subset B_2$. Again, we may assume that $\epsilon_2 = 1$. Put $V_2 := V_{\epsilon_2, F_1 \cup F_2}(0) \subset V_{\epsilon_2, F_2}(0)$. Continue this process inductively to obtain a countable collection of open neighborhoods $\{V_n : n \in \mathbb{N}\}$.

By construction, the collection of all such V_n forms a countable neighborhood base at 0 with $V_{n+1} \subset V_n$. Enumerate the union of all such V_k by a sequence (ϕ_k) by first enumerating the elements in F_1, F_2, \dots and so forth.

We first claim that $\bigcap_{j=1}^n \ker \phi_j$ is nontrivial for all n . For if it were trivial for some m , then

$$\bigcap_{j=1}^m \ker \phi_j = \{0\} \subset \ker \phi \text{ for any } \phi \in X^*$$

By our lemma, $\phi \in \text{span}(\phi_1, \dots, \phi_m)$ for all $\phi \in X^*$. Thus, X^* is finite dimensional so that X^{**} is finite dimensional. But X may be embedded as a subspace of X^{**} implying that X is finite dimensional, a contradiction. Thus, $\bigcap_{j=1}^n \ker \phi_j \neq \{0\}$ for all n .

For each n , choose a nonzero

$$x_n \in \bigcap_{j=1}^{s_n} \ker \phi_j$$

where s_n denotes the position in the sequence (ϕ_k) containing all elements up to F_n . Assume also, without loss of generality, that $\|x_n\|_X = n$ for all n . Then $x_n \in V_n$ for all n . Let V be any weakly open neighborhood of 0. Then there is an m so that $V_m \subset V$. But this implies that $x_n \in V$ for $n \geq m$ so that $x_n \rightarrow 0$ weakly. But $\|x_n\|_X \rightarrow \infty$ as $n \rightarrow \infty$, a contradiction because (x_n) must be bounded in norm. □

Note that this theorem did not require any condition of completeness (this property is somewhat superfluous anyway since all finite dimensional normed spaces are Banach spaces). Our goal in the later sections will be to show that in a general Banach space X , weak compactness behaves like sequential compactness. This is a remarkable result because the weak topology is almost never metrizable in practice as our theorem here shows.

We now prove a theorem concerning metrizability in the weak* topology.

Theorem 3.2.7

Let X be a separable normed space and $K \subset X^*$ be weak* compact. Then K is metrizable in the weak* topology.

Proof. Let $D := \{x_n : n \in \mathbb{N}\}$ be a countable dense subset of X and define the following:

$$\rho(\phi, \varphi) := \sum_{k=1}^{\infty} 2^{-k} \min(1, |\phi(x_k) - \varphi(x_k)|) \text{ for all } \phi, \varphi \in X^*$$

The reader may verify that ρ is indeed a metric (one uses the fact that D is norm dense in X to show this). Let \mathcal{T}_ρ be the topology generated by ρ . Firstly, we demonstrate that $\mathcal{T}_\rho \subset \sigma(X^*, X)$. To demonstrate this, we will show that every \mathcal{T}_ρ closed set is also weak* closed.

Let A be closed in \mathcal{T}_ρ and let (ϕ_λ) be a net in A with $\phi_\lambda \xrightarrow{w^*} \phi_0$. It is necessary to show that $\phi_0 \in A$. For any $\epsilon > 0$ and $j \in \mathbb{N}$, consider the open neighborhood of ϕ_0 :

$$V_{j,\epsilon}^* := \{\phi \in X^* : |\phi(x_n) - \phi_0(x_n)| < \epsilon, 1 \leq n \leq j\}$$

Let $\epsilon > 0$ and fix $j \in \mathbb{N}$. Then there is a $\lambda_0 \in \Lambda$ so that $\phi_\lambda \in V_{j,\epsilon}^*$ when $\lambda \succ \lambda_0$. Thus

$$\rho(\phi_\lambda, \phi_0) < \epsilon \sum_{k=1}^j 2^{-k} + \sum_{k=j+1}^{\infty} 2^{-k} = (\epsilon - 1) \sum_{k=1}^j 2^{-k} + 1 < \epsilon$$

whenever $\lambda \succ \lambda_0$. But this, in turn, implies that $\phi_\lambda \rightarrow \phi_0$ in ρ . But A is closed in \mathcal{T}_ρ so that $\phi_0 \in A$.

The reverse inclusion will follow given the validity of the following claim: if \mathcal{T}_1 and \mathcal{T}_2 are two topologies on K with $\mathcal{T}_1 \subset \mathcal{T}_2$, K is \mathcal{T}_2 -compact, and \mathcal{T}_1 is Hausdorff, then

$\mathcal{T}_2 \subset \mathcal{T}_1$. Let A be \mathcal{T}_2 closed and let $x_\lambda \rightarrow x_0$ in \mathcal{T}_1 where $x_\lambda \in A$ for all λ . Since $A \subset K$ and K is \mathcal{T}_2 compact, it follows that A is \mathcal{T}_2 compact as well. Thus, there is a subnet (x_{λ_μ}) with $x_{\lambda_\mu} \rightarrow x_1 \in A$ in \mathcal{T}_2 . But net convergence in \mathcal{T}_2 implies net convergence in \mathcal{T}_1 so that $x_{\lambda_\mu} \rightarrow x_1$ in \mathcal{T}_1 . But \mathcal{T}_1 is Hausdorff so limits are unique and $x_0 = x_1 \in A$ and we're finished. □

This last theorem will play an important role when we later discuss the Banach-Mazur Theorem. We now turn to the issue of compactness and will revisit this idea as we further develop the theory of Banach spaces. We will especially need many of these ideas as we explore the idea of reflexivity.

3.3 Compactness and Reflexivity

Our next theorem in some ways justifies the usefulness of the weak* topology. Before we present this result, we first recall a major result from topology.

Theorem (Tychonoff's Theorem)

A nonempty product space is compact (in the product topology) if and only if each factor space is compact. □

We are now ready to prove the Alaoglu Theorem:

Theorem 3.3.0 (Banach-Alaoglu Theorem)

Let X be a normed space. Then the closed unit ball $B^* := \{\phi \in X^* : \|\phi\| \leq 1\}$ is compact in the weak* topology.

Proof. For each $x \in X$, set $A_x := \{z \in \mathbb{C} : |z| \leq \|x\|_X\}$ and put $A := \prod_{x \in X} A_x$. By the Tychonoff Theorem, A is compact in the product topology. We may interpret the elements of A as the set of all $\phi : X \rightarrow \mathbb{C}$ so that $|\phi(x)| \leq \|x\|_X$ for all $x \in X$. Now B^* contains all such elements that are linear. Moreover, we know that the product topology on A is the weakest topology so that the projection maps $\pi_x : A \rightarrow A_x$ are continuous. But each $\pi_x(\phi) = \phi(x) = J_x(\phi)$ is really the canonical embedding from X into X^{**} so the topology on A is really the weak* topology.

Because $B^* \subset A$ and A is compact in the weak* topology, it suffices to show that B^* is (relatively) weak* closed. Let (ϕ_λ) be a net in B^* with $\phi_\lambda \rightarrow \phi_0$. Then $\|\phi_\lambda\| \leq 1$ for all

λ . By assumption, we know that $\phi_\lambda(x) \rightarrow \phi_0(x)$ for all $x \in X$. Then $|\phi_\lambda(x)| \leq \|x\|$ for all λ and all $x \in X$ so that $|\phi_0(x)| \leq \|x\|$ for all $x \in X$. Hence $\|\phi_0\| \leq 1$ so $\phi_0 \in B^*$ and we're finished. □

Note that the Alaoglu Theorem is true for any closed ball, not just the unit closed ball.

We could use the uniform boundedness principle to demonstrate the following but we treat it as a corollary to the Alaoglu Theorem.

Corollary 3.3.1

Let (ϕ_n) be a sequence in X^* so that $\phi_n \xrightarrow{w^*} \phi_0$. Then $\|\phi_0\| \leq \liminf_{n \rightarrow \infty} \|\phi_n\|$.

Proof. Set $M := \liminf_{n \rightarrow \infty} \|\phi_n\|$. Fix $\epsilon > 0$. Then there exists a subsequence (ϕ_{n_j}) of (ϕ_n) so that $\|\phi_{n_j}\| \leq M + \epsilon$ for all j . Because the norm-ball of radius $M + \epsilon$ is weak* compact and because the weak* topology is T_2 , we see that the norm-ball of radius $M + \epsilon$ is weak* closed. Thus $\|\phi_0\| \leq M + \epsilon$ for all $\epsilon > 0$. Letting $\epsilon \rightarrow 0^+$ yields the desired result. □

If (X, \mathcal{T}) is a topological space and $K \subset X$ recall that K is called sequentially compact if every sequence has a convergent subsequence. If X is metrizable, then sequential compactness is equivalent to compactness but in a general topological space, compactness is typically stronger property. This next theorem asserts that any norm bounded subset of a Hilbert space is weakly sequentially compact. Let us prove this fact now.

Theorem 3.3.2

Every norm bounded sequence in a Hilbert space H contains a weakly convergent subsequence.

Proof. Suppose that $\|x_n\|_H \leq M$ for all n . Denote by Y the closure of the linear span of each term in the sequence. Then the sequence of real numbers (x_1, x_n) has a convergent subsequence, call it $s_n^1 := (x_1, x_n^1)$. Now, (x_2, x_n^1) has a convergent subsequence, call it $s_n^2 := (x_2, x_n^2)$. Continue this process to obtain, for each j , a convergent sequence $s_n^j := (x_j, x_n^j)$. Consider the diagonal sequence (x_m^m) in H . For each j , (x_j, x_m^m) converges because whenever $m > j$, it is a subsequence of the convergent sequence (x_j, x_m^j) . It is clear that (x, x_m^m) converges as $m \rightarrow \infty$ whenever x is a finite sum of the form

$$x = \sum_{k=1}^N c_k x_k$$

But finite sums of this form are dense in Y so that if $y \in Y$, there is a sequence of finite sums (y_n) so that $\|y_n - y\|_H \rightarrow 0$ as $n \rightarrow \infty$. Also, (y_n, x_m^m) converges for all n by our remark above. But

$$(y, x_m^m) = (y_n, x_m^m) + (y - y_n, x_m^m)$$

where $|(y - y_n, x_m^m)| \leq M\|y - y_n\|_H$. But then

$$\limsup_{m \rightarrow \infty} (y, x_m^m) = \lim_{m \rightarrow \infty} (y_n, x_m^m) + \limsup_{m \rightarrow \infty} (y - y_n, x_m^m)$$

We could also take the limit infimum instead so that

$$\limsup_{m \rightarrow \infty} (y, x_m^m) = \liminf_{m \rightarrow \infty} (y, x_m^m) = \limsup_{n \rightarrow \infty} \lim_{m \rightarrow \infty} (y_n, x_m^m)$$

Thus, (y, x_m^m) converges for all $y \in Y$.

By virtue of H being a Hilbert space, we have $H = Y \oplus Y^\perp$. Thus if $x \in H$, then $x = y + \hat{y}$ where $y \in Y$ and $\hat{y} \in Y^\perp$. Then

$$(x, x_m^m) = (y, x_m^m) + (\hat{y}, x_m^m) = (y, x_m^m)$$

as $(\hat{y}, x_m^m) = 0$ for all m . In other words, (x, x_m^m) converges for all $x \in H$. Put $w_m := x_m^m$ and define

$$\varphi_0(x) := \lim_{m \rightarrow \infty} (w_m, x) \text{ for all } x \in H$$

It is clear that $\varphi_0 \in H^*$. By the Riesz Representation Theorem, there exists an $x_0 \in H$ so that $\varphi_0(x) = (x, x_0)$ for all $x \in H$. If $\phi \in H^*$, there is an $x_1 \in H$ so that $\phi(x) = (x, x_1)$ and

$$\phi(w_m) = (w_m, x_1) \rightarrow \varphi_0(x_1) = (x_1, x_0) = \phi(x_0)$$

That is, $w_m \xrightarrow{w} x_0$ which is what we wished to show. □

In the later sections, we will prove the Eberlein-Šmulian theorem which tells us that in a Banach space, weak compactness is equivalent to weak sequential compactness. Thus, given the validity of Eberlein-Šmulian, the theorem we just proved tells us that the unit ball in a Hilbert space is weakly compact. It turns out that this very condition is equivalent to a space being reflexive and we shall explore this now.

To save time, the notation $B := \{x \in X : \|x\|_X \leq 1\}$ and $B^{**} := \{\psi \in X^{**} : \|\psi\| \leq 1\}$ will be employed. As usual, $J = J_x$ will denote the canonical embedding of X into X^{**} .

Theorem 3.3.3 (Goldstein)

$J(B)$ is weak* dense in B^{**} .

Proof. Note that if $x \in B$, then $\|x\|_X \leq 1$ so that $\|J_x\| = \|x\|_X \leq 1$ and $J_x \in B^{**}$. That is, $J(B) \subset B^{**}$. By the Alaoglu Theorem, B^{**} is weak* compact (here, we take X^* as our normed space). Because the weak* topology is T_2 , we know that B^{**} is weak* closed and hence $\overline{J(B)} \subset B^{**}$ where the closure is the weak* closure. We will now demonstrate the reverse inclusion.

Let $\psi_0 \in B^{**}$. It is only necessary to show that for any $\epsilon > 0$ and $\phi_1, \dots, \phi_N \in X^*$ with $\|\phi_k\| = 1$ that the open neighborhood

$$V_\epsilon := \{\psi \in X^{**} : |\psi(\phi_k) - \psi_0(\phi_k)| < \epsilon, 1 \leq k \leq N\}$$

meets $J(B)$. It is enough to show the existence of an $x \in X$ so that $\|x\|_X < 1 + \epsilon$ with $(J_x - \psi_0)(\phi_k) = 0$ for each $1 \leq k \leq N$. Given the validity of this claim for the moment, we would then see that $(1 + \epsilon)^{-1}x$ belongs to B and for each $1 \leq k \leq N$ we would have

$$\left| \frac{1}{1 + \epsilon} J_x(\phi_k) - \psi_0(\phi_k) \right| = \left| \frac{1}{1 + \epsilon} J_x(\phi_k) - J_x(\phi_k) \right| \leq \frac{\epsilon}{1 + \epsilon} \|J_x\| < \epsilon$$

We will now demonstrate the claim. Let $S := \text{span}(\phi_1, \dots, \phi_N)$ so that $\dim(S) < \infty$. It is then easy to see that $J : X \rightarrow S^*$ is surjective. Let $\Psi := \psi_0|_S$. Then $\Psi \in S^*$ so there is an $x \in X$ so that $J_x = \Psi$. It is clear that $(J_x - \psi_0)(\phi_k) = 0$ for each k . Because $\|x\|_X = \|J_x\| = \|\Psi\| \leq 1$, it is clear that $\|x\|_X < 1 + \epsilon$ for all $\epsilon > 0$. □

We are now ready to prove our useful characterisation of reflexivity to which we earlier alluded.

Corollary 3.3.4 (Characterization of Reflexive Spaces)

The closed unit ball of a normed space X is weakly compact if and only if X is reflexive.

Proof. Note that the canonical embedding $J : X \rightarrow X^{**}$ is continuous in the weak topology of X . Thus if B is weakly compact, then $J(B)$ is weak* compact and, hence, weak* closed. By Goldstein’s Theorem, $\overline{J(B)} = J(B) = B^{**}$. This easily implies that the canonical embedding is surjective. On the other hand, suppose that X is reflexive. Then $(X, \text{weak topology})$ and $(X^{**}, \text{weak* topology})$ are homeomorphic via J . But B^{**} is weak* compact by Alaoglu’s Theorem and the unit ball in X is consequently weakly compact. □

We will now conclude this section with an important example due to Robert C. James in 1950. James constructed an example of a Banach space that is *not* reflexive. The space he constructed is given by

$$J := \left\{ (x_n) : \lim_{n \rightarrow \infty} x_n = 0 \quad \text{and} \quad \|(x_n)\|_J < \infty \right\}$$

where

$$\|(x_n)\|_J := \sup \left(\sum_{k=1}^n (x_{p_{2k-1}} - x_{p_{2k}})^2 + (x_{p_{2k+1}})^2 \right)^{1/2}$$

and the supremum is taken over all n and all increasing sequences of positive integers $p_1, p_2, \dots, p_{2n+1}$. In 1951, he utilized a similar non-reflexive space but constructed an isomorphism to its double dual. Thus even if $X \cong X^{**}$ it does not necessarily follow that X is reflexive.

3.4 Exercises

1. Let X be a finite dimensional normed space. Prove that a sequence (x_n) in X converges weakly to x_0 if and only if it converges strongly to x_0 .

2. This exercise helps further familiarize you with the weak and weak* topologies.
 - (a) Convince yourself that \mathcal{B}_{x_0} and $\mathcal{B}_{\phi_0}^*$ form neighborhood bases about x_0 and ϕ_0 respectively in the weak and weak* topology.
 - (b) Prove that the weak* topology is the topology of pointwise convergence; that is, show that a net (ϕ_λ) in X^* converges to ϕ_0 in the weak* topology if and only if $\phi_\lambda(x) \rightarrow \phi_0(x)$ for all $x \in X$.

3. Let X be a normed space. Prove that the norm-closed ball $B = \{x \in X : \|x\| \leq 1\}$ is weakly closed. Additionally, show that if $E \subset X$ is norm bounded, then its weak closure is also norm bounded.

Exercises 4 and 5 are for those who are familiar with duality of L^p and $C(K)$. We will explore these topics later.

4. Let $1 < p < \infty$. Show that a sequence (f_n) in $L^p(0, 1)$ converges weakly to f if and only if:
 - (a) $\|f_n\|_{L^p} \leq M$ for some $M > 0$.
 - (b) $\lim_{n \rightarrow \infty} \int_a^b f_n(t) dt = \int_a^b f(t) dt$ for all $(a, b) \subset (0, 1)$.

5. Let $K \subset \mathbb{R}$ be compact. Prove that a sequence (f_n) in $C(K)$ converges weakly to f if and only if:
 - (a) $|f(x)| \leq M$ for all $x \in K$
 - (b) $f_n(x) \rightarrow f(x)$ pointwise for each $x \in K$

4 Bibliography

This needs to be organized, cleaned up, and formatted.

1. Gerald Folland, *Real Analysis*
2. Albiac and Kalton, *Topics in Banach Space Theory*
3. Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*
4. James, R. C., *Bases and Reflexivity of Banach Spaces*, Ann. Math., 52, 518-527 (1950).
5. James, R. C., *A Non-Reflexive Banach Space Isometric with Its Second Conjugate Space*, Proc. Nat. Acad. Sci., 37, 174-177 (1951).