Functional Analysis Review
MA 69300
Spring 2021

This document is intended to give an overview of the main results from functional analysis that we will encounter in this course. We assume the reader has some familiarity with these concepts, or at least concepts from Analysis and Topology such as convergence, continuity, open sets, and closed sets. We will consider vector spaces X over a field $\mathbb{F}$ where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. In some instances, we will specify $\mathbb{F}$ when the distinction is important. If we are working with two or more vector spaces, we will assume they are all over the same field unless specified otherwise.

1 Banach Spaces

Definition 1. A norm on X is a function $\|\cdot\| : X \to [0, \infty)$ satisfying the following properties:
(i) (*Definiteness*) $\|x\| = 0 \iff x = 0$,
(ii) (*Absolute Homogeneity*) $\|\alpha x\| = |\alpha| \, \|x\| \quad \forall \, (\alpha, x) \in \mathbb{F} \times X$,
(iii) (*Triangle Inequality*) $\|x + y\| \leq \|x\| + \|y\| \quad \forall \, x, y \in X$.

A vector space X together with a norm $\|\cdot\|$ is called a normed vector space. The Triangle Inequality in Definition 1 can be used to establish the inequality

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

This inequality implies that a norm defines a (uniformly) continuous function.

Definition 2. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in X. We say the sequence $\{x_n\}_{n=1}^{\infty}$ converges to a point $x \in X$ if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$n \geq N \implies \|x_n - x\| \leq \varepsilon$$

If $\{x_n\}_{n=1}^{\infty}$ converges to x, we usually write $x_n \to x$ or $\lim\limits_{n \to \infty} x_n = x$. Limits in normed vector spaces enjoy familiar properties. Suppose $\alpha \in \mathbb{F}$ and $\{x_n\}_{n=1}^{\infty}, \{y_n\}_{n=1}^{\infty}$ are both convergent sequences in X. Then the sequences $\{\alpha x_n\}_{n=1}^{\infty}$ and $\{x_n + y_n\}_{n=1}^{\infty}$ both converge in X and

$$\lim_{n \to \infty} \alpha x_n = \alpha \lim_{n \to \infty} x_n,$$

$$\lim_{n \to \infty} (x_n + y_n) = \lim_{n \to \infty} x_n + \lim_{n \to \infty} y_n.$$

Furthermore, if $\{x_n\}_{n=1}^{\infty}$ is a convergent sequence in X, the continuity of the norm also gives us

$$\lim_{n \to \infty} \|x_n\| = \left\| \lim_{n \to \infty} x_n \right\| .$$

Definition 3. A sequence of points $\{x_n\}_{n=1}^{\infty}$ in X is called a Cauchy sequence if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$n, m \geq N \implies \|x_n - x_m\| \leq \varepsilon.$$

If every Cauchy sequence in X converges to a point $x \in X$, then we say X is complete. A complete normed vector space is called a Banach space.

Definition 4. A function $T : X \to Y$ between vector spaces is called a linear map if T satisfies the following properties:
(i) (*Homogeneity*) $T(\alpha x) = \alpha T(x) \quad \forall \, (\alpha, x) \in \mathbb{F} \times X$,
(ii) (*Additivity*) $T(x + y) = T(x) + T(y) \quad \forall \, x, y \in X$.

From either of these properties, we see linear maps always satisfy $T(0) = 0$. For the minimalists, one can check that $T : X \to Y$ is linear if and only if $T(\alpha u + v) = \alpha T(u) + T(v)$ for all $\alpha \in \mathbb{F}$ and $u, v \in X$. One can also verify that the set of all linear maps from X to Y with the addition $(T + S)(x) = T(x) + S(x)$ and scalar multiplication $(\alpha T)(x) = \alpha T(x)$ is a vector space over $\mathbb{F}$. We denote this space by $L(X, Y)$.

When the spaces X and Y are normed, we can talk about continuous linear maps $T : X \to Y$. It turns out linear maps have a useful characterization of continuity.

Definition 5. Let X and Y be normed vector spaces. Define a function $\|\cdot\| : L(X, Y) \to [0, \infty]$ by $\|T\| = \sup\{\|T(x)\| : x \in X, \|x\| \leq 1\}$. A map $T \in L(X, Y)$ is said to be bounded if $\|T\| < \infty$.

Let $B(X, Y)$ denote the set of all bounded linear maps from X to Y. That is, $B(X, Y) = \{T \in L(X, Y) : \|T\| < \infty\}$. Note that $0 \in B(X, Y)$. If $T \in B(X, Y)$ and $\alpha \in \mathbb{F}$, then $\|\alpha T(x)\| = |\alpha| \|T(x)\| \leq |\alpha| \|T\|$ for all $x \in X$ with $\|x\| \leq 1$, so it follows that

$$\|\alpha T\| \leq |\alpha| \|T\| < \infty.$$

Furthermore, if $T, S \in B(X, Y)$ then $\|T(x) + S(x)\| \leq \|T(x)\| + \|S(x)\| \leq \|T\| + \|S\|$ for all $x \in X$ with $\|x\| \leq 1$. Therefore,

$$\|T + S\| \leq \|T\| + \|S\| < \infty.$$

This implies that $B(X, Y)$ is a vector subspace of $L(X, Y)$. In fact, the above work leads us towards the following:

Proposition 1. *$B(X, Y)$ is a normed vector space with the norm*

$$\|T\| = \sup\{\|T(x)\| : x \in X, \|x\| \leq 1\}.$$

Proof. We have already established that $B(X, Y)$ is a vector space. Our previous work shows that the function $\|\cdot\|$ on $B(X, Y)$ satisfies the triangle inequality:

$$\|T + S\| \leq \|T\| + \|S\| \quad \forall\, T, S \in B(X, Y).$$

Now, let $\alpha \in \mathbb{F}$ and $T \in B(X, Y)$. We have already established

$$\|\alpha T\| \leq |\alpha| \|T\|.$$

Fix $\varepsilon > 0$. By definition of $\|T\|$, there exists $x \in X$ with $\|x\| \leq 1$ such that

$$\|T\| - \frac{\varepsilon}{|\alpha| + 1} < \|T(x)\| \iff \|T\| < \|T(x)\| + \frac{\varepsilon}{|\alpha| + 1}.$$

Multiplying this inequality by $|\alpha|$ gives

$$|\alpha| \|T\| \leq |\alpha| \|T(x)\| + \frac{|\alpha| \varepsilon}{|\alpha| + 1} = \|\alpha T(x)\| + \frac{|\alpha| \varepsilon}{|\alpha| + 1} \leq \|\alpha T\| + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we conclude $|\alpha| \|T\| \leq \|\alpha T\|$. Therefore, we have equality

$$\|\alpha T\| = |\alpha| \|T\|.$$

Evidently, if $T \in B(X, Y)$ and $T = 0$ then $\|T\| = 0$. Lastly, suppose $T \in B(X, Y)$ and $\|T\| = 0$. Let $x \in X \setminus \{0\}$. Then,

$$\frac{1}{\|x\|} \|T(x)\| = \left\| \frac{1}{\|x\|} T(x) \right\| = \left\| T\left(\frac{x}{\|x\|} \right) \right\| \leq \|T\| = 0.$$

Therefore, $\|T(x)\| = 0 \implies T(x) = 0$. We also have $T(0) = 0$. Hence, we conclude $T = 0$. The function $\|\cdot\|$ satisfies the necessary properties in Definition 1. We conclude that $\|\cdot\|$ defines a norm on $B(X, Y)$. ☐

It follows from the above proof that if $T \in B(X, Y)$ and $x \in X \setminus \{0\}$, then

$$\frac{1}{\|x\|} \|T(x)\| \leq \|T\|.$$

Hence, we have the important inequality

$$\|T(x)\| \leq \|T\| \|x\| \quad \forall\, T \in B(X, Y),\ x \in X. \tag{1}$$

We can now prove the following:

Proposition 2. *$T \in B(X, Y)$ if and only if $T : X \to Y$ is continuous.*

Proof. ($\Longrightarrow$) Suppose $T \in B(X, Y)$. Fix $\varepsilon > 0$ and let $\delta = \dfrac{\varepsilon}{\|T\| + 1}$. Then, if $x, y \in X$ and $\|x - y\| < \delta$ equation (1) implies

$$\|T(x) - T(y)\| = \|T(x - y)\| \leq \|T\| \|x - y\| \leq \|T\| \frac{\varepsilon}{\|T\| + 1} < \varepsilon.$$

Therefore, T is (uniformly) continuous.

($\Longleftarrow$) Suppose $T : X \to Y$ is continuous. Then T is continuous at $x = 0$. Hence, there exists $\delta > 0$ such that

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

Let $x \in X$ with $0 < \|x\| \leq 1$. Then, $\left\| \dfrac{\delta x}{2 \|x\|} \right\| = \dfrac{\delta}{2} < \delta$. Hence,

$$\frac{\delta}{2 \|x\|} \|T(x)\| = \left\| T\left(\frac{\delta x}{2 \|x\|} \right) \right\| < 1,$$

$$\implies \|T(x)\| < \frac{2\|x\|}{\delta} \leq \frac{2}{\delta}.$$

Therefore, we conclude $\|T(x)\| \leq \frac{2}{\delta}$ for all $x \in X$ with $\|x\| \leq 1$. Hence, $\|T\| < \infty$ and we conclude $T \in B(X, Y)$.

□

If we impose additional restrictions on the target space Y, we can say something more about $B(X, Y)$.

Theorem 3. *Let X be a normed vector space and let Y be a Banach space. Then $B(X, Y)$ is a Banach space.*

Proof. Let $\{T_n\}_{n=1}^{\infty}$ be a Cauchy sequence of maps in $B(X, Y)$. For each $x \in X$, we have the inequality

$$\|T_n(x) - T_m(x)\| = \|(T_n - T_m)(x)\| \leq \|T_n - T_m\| \|x\| \quad \forall\, n, m \geq 1.$$

This implies that, for each $x \in X$, the sequence $\{T_n(x)\}_{n=1}^{\infty}$ is a Cauchy sequence in Y. Since Y is Banach, the sequence $\{T_n(x)\}_{n=1}^{\infty}$ converges to a point in Y. By the uniqueness of limits in Banach spaces, we can define a function $T : X \to Y$ by

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

The linearity of T now follows from the linearity of limits in normed vector spaces and the linearity of the maps T_n:

$$T(\alpha x) = \lim_{n\to\infty} T_n(\alpha x) = \lim_{n\to\infty} \alpha T_n(x) = \alpha \lim_{n\to\infty} T_n(x) = \alpha T(x) \quad \forall\, (\alpha, x) \in \mathbb{F} \times X,$$

and

$$T(x + y) = \lim_{n\to\infty} T_n(x + y) = \lim_{n\to\infty} (T_n(x) + T_n(y))$$
$$\lim_{n\to\infty} T_n(x) + \lim_{n\to\infty} T_n(y) = T(x) + T(y) \quad \forall\, x, y \in X.$$

To show that $T \in B(X, Y)$, let $x \in X$ with $\|x\| \leq 1$. The inequality

$$\big|\, \|T_n\| - \|T_m\| \,\big| \leq \|T_n - T_m\| \quad \forall\, m, n \geq 1$$

implies that $\{\|T_n\|\}_{n=1}^{\infty}$ is a Cauchy sequence of real numbers. Therefore, there is $L \in \mathbb{R}$ such that $\lim_{n\to\infty} \|T_n\| = L$. Then, we have

$$\|T(x)\| \leq \|T(x) - T_n(x)\| + \|T_n(x)\| \leq \|T(x) - T_n(x)\| + \|T_n\| \quad \forall\, n \geq 1.$$

Letting n tend to infinity in the above inequality implies

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

Since $x \in X$ with $\|x\| \leq 1$ is arbitrary, we conclude $\|T\| < \infty$. Thus, $T \in B(X, Y)$.
It remains to show that $T_n \to T$. Fix $\varepsilon > 0$. Let $x \in X$ with $\|x\| \leq 1$. There exists $N \in \mathbb{N}$ such that

$$m, n \geq N \implies \|T_n - T_m\| \leq \varepsilon.$$

For $m, n \geq N$, we have

$$\|T_n(x) - T_m(x)\| \leq \|T_n - T_m\| \leq \varepsilon.$$

Then, using the continuity of norms, letting m approach infinity in the first part of the above inequality gives

$$\|T_n(x) - T(x)\| \leq \varepsilon \quad \forall\ n \geq N.$$

Since $x \in X$ with $\|x\| \leq 1$ is arbitrary, the above inequality implies

$$\|T_n - T\| \leq \varepsilon \quad \forall\ n \geq N.$$

Since $\varepsilon > 0$ is arbitrary, we conclude $T_n \to T$. Therefore, $B(X, Y)$ is a Banach space.

□

Let X be a vector space. An element $\varphi \in L(X, \mathbb{F})$ is called a linear functional. If we further assume that X is a normed vector space, an element $\varphi \in B(X, \mathbb{F})$ is called a bounded linear functional. We use the notation $X^* = B(X, \mathbb{F})$, and we say X^* is the dual space of X. Since we are considering $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, Theorem 1 implies that for any normed vector space X the dual space X^* is a Banach space.

We are going to establish one of the most important theorems in functional analysis. In order to do so, we need to recall the following lemma (which is equivalent to the Axiom of Choice):

Lemma 4 (Zorn's Lemma). *Let A be a nonempty partially ordered set with partial order $\leq$. That is, $\leq$ is a relation on A satisfying*
(i) (Reflexivity) $a \leq a \quad \forall\ a \in A$,
(ii) (Antisymmetry) $\forall\ a, b \in A$, $a \leq b$ and $b \leq a$ implies $a = b$,
(iii) (Transitivity) $\forall\ a, b, c \in A$, $a \leq b$ and $b \leq c$ implies $a \leq c$.
Suppose that every totally ordered subset of A has an upper bound in A. That is, for every $B \subset A$ with the property that $\forall\ a, b \in B$ either $a \leq b$ or $b \leq a$, there exists $L \in A$ such that $a \leq L \quad \forall\ a \in B$. Then A has a maximal element. That is, there exists $M \in A$ such that if $a \in A$ and $M \leq a$, then $M = a$.

We can now establish the important theorem:

Theorem 5 (Hahn-Banach Extension Theorem)**.** *Let X be a vector space over $\mathbb{R}$. Suppose there is a function $m : X \to \mathbb{R}$ such that*
(i) $m(tx) = tm(x) \quad \forall\ (t, x) \in [0, \infty) \times X$,
(ii) $m(x + y) \leq m(x) + m(y) \quad \forall\ x, y \in X$.

Let $V \subset X$ be a vector subspace of X, and suppose there is a linear functional $\phi : V \to \mathbb{R}$ such that

$$\phi(x) \leq m(x) \quad \forall\ x \in V.$$

Then there exists a linear functional $\varphi : X \to \mathbb{R}$ such that

$$\varphi(x) = \phi(x) \quad \forall\ x \in V,$$

and

$$\varphi(x) \leq m(x) \quad \forall\ x \in X.$$

Proof. We suppose $V \neq X$. Let $x_0 \in X \setminus V$. We consider the linear span $S = \text{span}\{x_0\} = \{\alpha x_0 : \alpha \in \mathbb{R}\}$. We will show that we can extend ϕ to a linear functional $\varphi : W \to \mathbb{R}$ where $W = V + S = \{v + \alpha x_0 : (v, \alpha) \in V \times \mathbb{R}\}$ and $\varphi(w) \leq m(w)$ for all $w \in W$. Note that since $x_0 \notin V$, it follows that for each $w \in W$ there exist *unique* $v \in V$ and $\alpha \in \mathbb{R}$ such that $w = v + \alpha x_0$.

To figure out how we should define φ, we temporarily assume that we have such an extension $\varphi : W \to \mathbb{R}$. Linearity of φ and $\varphi = \phi$ on V imply

$$\varphi(v + \alpha x_0) = \varphi(v) + \alpha\varphi(x_0) = \phi(v) + \alpha\varphi(x_0).$$

Thus, φ will be completely determined by $\varphi(x_0)$. To determine an appropriate value for $\varphi(x_0)$, we recall that φ needs to satisfy

$$\varphi(w) \leq m(w) \quad \forall\ w \in W.$$

This imposes the conditions

$$\phi(v) + \varphi(x_0) = \varphi(v + x_0) \leq m(v + x_0)$$

$$\iff \varphi(x_0) \leq m(v + x_0) - \phi(v) \quad \forall\ v \in V,$$

and

$$\phi(v) - \varphi(x_0) = \varphi(v - x_0) \leq m(v - x_0)$$

$$\iff \varphi(x_0) \geq \phi(v) - m(v - x_0) \quad \forall\ v \in V.$$

These inequalities suggest a way to define $\varphi(x_0)$. Note that

$$m(v+x_0)-\phi(v)\geq m(v+x_0)-m(v)=m(v+x_0)-m(v+x_0+(-x_0))$$

$$\geq m(v+x_0)-(m(v+x_0)+m(-x_0))=-m(-x_0)\quad\forall\, v\in V,$$

so $L=\inf\{m(v+x_0)-\phi(v):v\in V\}\in\mathbb{R}$ and $L\geq -m(-x_0)$.

Similarly,

$$\phi(v)-m(v-x_0)\leq m(v)-m(v-x_0)=m(v-x_0+x_0)-m(v-x_0)$$

$$\leq(m(v-x_0)+m(x_0))-m(v-x_0)=m(x_0)\quad\forall\, v\in V,$$

and so $M=\sup\{\phi(v)-m(v-x_0):v\in V\}\in\mathbb{R}$ and $M\leq m(x_0)$.

The conditions on $\varphi(x_0)$ require that $\varphi(x_0)\leq L$ and $\varphi(x_0)\geq M$. We show that it is possible to choose a value satisfying these conditions.
Indeed, let $x,y\in V$. We have

$$m(x+x_0)-\phi(x)-(\phi(y)-m(y-x_0))=m(x+x_0)+m(y-x_0)-\phi(x)-\phi(y)\geq m(x+y)-\phi(x+y)\geq 0$$

$$\iff \phi(y)-m(y-x_0)\leq m(x+x_0)-\phi(x).$$

Since $x\in V$ is arbitrary, this implies that for each $y\in V$ we have

$$\phi(y)-m(y-x_0)\leq L.$$

Since $y\in V$ is arbitrary, this implies

$$M\leq L.$$

Let $r=\frac{L+M}{2}$ so that $M\leq r\leq L$. Then, we can define a linear functional $\varphi:W\to\mathbb{R}$ by the rule

$$\varphi(v+\alpha x_0)=\phi(v)+\alpha r\quad\forall\, v+\alpha x_0\in W.$$

By definition, we have

$$\varphi(v)=\phi(v)\quad\forall\, v\in V.$$

Now, if we consider $w=v+\alpha x_0$ with $v\in V$ and $\alpha>0$, we have $\frac{v}{\alpha}\in V$, and so $L\leq m\left(\frac{v}{\alpha}+x_0\right)-\phi\left(\frac{v}{\alpha}\right)$. Then

$$\varphi(w)=\phi(v)+\alpha r\leq\phi(v)+\alpha\left(m\left(\frac{v}{\alpha}+x_0\right)-\phi\left(\frac{v}{\alpha}\right)\right)$$

$$=\phi(v)-\alpha\phi\left(\frac{v}{\alpha}\right)+\alpha m\left(\frac{v}{\alpha}+x_0\right)=m(v+\alpha x_0)=m(w).$$

If we consider $w = v + \alpha x_0$ with $v \in V$ and $\alpha < 0$, we have $-\frac{v}{\alpha} \in V$, and so $-M \leq m\left(-\frac{v}{\alpha} - x_0\right) - \phi\left(-\frac{v}{\alpha}\right)$. Then

$$\varphi(w) = \phi(v) + (-\alpha)(-r) \leq \phi(v) + (-\alpha)\left(m\left(-\frac{v}{\alpha} - x_0\right) - \phi\left(-\frac{v}{\alpha}\right)\right)$$

$$= \phi(v) + \alpha\phi\left(-\frac{v}{\alpha}\right) + (-\alpha)m\left(-\frac{v}{\alpha} - x_0\right) = m(v + \alpha x_0) = m(w).$$

Of course, if $w = v \in V$, then

$$\varphi(w) = \phi(w) \leq m(w).$$

Therefore, we conclude

$$\varphi(w) \leq m(w) \quad \forall\, w \in W.$$

Hence, we have extended ϕ to a linear functional $\varphi : W \to \mathbb{R}$ satisfying the necessary properties.

To finish the proof, we define the set A to be the collection of all pairs (ψ, U) where U is a subspace of X with $V \subset U$, and $\psi : U \to \mathbb{R}$ is a linear functional with $\psi(x) = \phi(x)$ for all $x \in V$ and $\psi(x) \leq m(x)$ for all $x \in U$. Note that $(\phi, V) \in A$, so $A \neq \emptyset$. We define a partial order $\leq$ on A such that $(\psi_1, U_1) \leq (\psi_2, U_2)$ if and only if $U_1 \subset U_2$ and $\psi_2(x) = \psi_1(x)$ for all $x \in U_1$. In particular, note that $(\phi, V) \leq (\psi, U)$ for all $(\psi, U) \in A$.

Let $B = \{(\psi_j, U_j)\}_{j \in J}$ be a (nonempty) totally ordered subset of A. Let $U = \bigcup_{j \in J} U_j$. The total order on B implies that U is a vector subspace of X, and by definition of the set A, we have $V \subset U$. We define a linear functional $\psi : U \to \mathbb{R}$ by the rule

$$\psi(x) = \psi_j(x), \text{ if } x \in U_j.$$

Note that ψ is well-defined since if $x \in U_j \cap U_i$, then either $(\psi_j, U_j) \leq (\psi_i, U_i)$ or $(\psi_i, U_i) \leq (\psi_j, U_j)$ and so, by definition of the order $\leq$, we have $\psi_i(x) = \psi_j(x)$. Since $B \neq \emptyset$, there exists $j_0 \in J$. Since $(\psi_{j_0}, U_{j_0}) \in A$, it follows that $(\phi, V) \leq (\psi_{j_0}, U_{j_0})$. Therefore,

$$\psi(x) = \psi_{j_0}(x) = \phi(x) \quad \forall\, x \in V.$$

Furthermore, let $x \in U$. Then there is $j \in J$ such that $x \in U_j$.
Then

$$\psi(x) = \psi_j(x) \leq m(x).$$

Hence, we conclude $(\psi, U) \in A$. The definition of the pair (ψ, U) also implies $(\psi_j, U_j) \leq (\psi, U)$ for all $j \in J$. That is, (ψ, U) is an upper bound for B. By Zorn's Lemma, there exists a maximal element $(\varphi, \widetilde{X})$ of A.

By definition of A, we have

$$\varphi(x) = \phi(x) \quad \forall\ x \in V,$$

and

$$\varphi(x) \leq m(x) \quad \forall\ x \in \widetilde{X}.$$

Thus, the proof will be finished if we show $\widetilde{X} = X$. Assume for sake of contradiction that $\widetilde{X} \neq X$. Then there exists $x_0 \in X \setminus \widetilde{X}$. Let $W = \widetilde{X} + \text{span}\{x_0\}$. By the first part of the proof, there exists a linear functional $\widetilde{\varphi} : W \to \mathbb{R}$ such that

$$\widetilde{\varphi}(x) = \varphi(x) \quad \forall\ x \in \widetilde{X},$$

and

$$\widetilde{\varphi}(x) \leq m(x) \quad \forall\ x \in W.$$

Moreover, we have

$$(\phi, V) \leq (\varphi, \widetilde{X}) \leq (\widetilde{\varphi}, W).$$

Hence, we conclude $(\widetilde{\varphi}, W) \in A$, and the maximality of $(\varphi, \widetilde{X})$ implies we must have equality $(\varphi, \widetilde{X}) = (\widetilde{\varphi}, W)$, which implies $\widetilde{X} = W$, which contradicts that $x_0 \in W \setminus \widetilde{X}$. Therefore, to avoid contradiction, we conclude $\widetilde{X} = X$.

□

Of course, we have a complex version of the Hahn-Banach Theorem:

Theorem 6. *Let X be a vector space over $\mathbb{C}$. Suppose there is a function $m : X \to \mathbb{R}$ such that*
(i) $m(tx) = tm(x) \quad \forall\ (t, x) \in [0, \infty) \times X$,
(ii) $m(x + y) \leq m(x) + m(y) \quad \forall\ x, y \in X$.

Let $V \subset X$ be a vector subspace of X, and suppose there is a linear functional $\phi : V \to \mathbb{C}$ such that

$$Re\,\phi(x) \leq m(x) \quad \forall\ x \in V.$$

Then there exists a linear functional $\varphi : X \to \mathbb{C}$ such that

$$\varphi(x) = \phi(x) \quad \forall\ x \in V,$$

and

$$Re\,\varphi(x) \leq m(x) \quad \forall\ x \in X.$$

Proof. By assumption, $\operatorname{Re}\phi : V \to \mathbb{R}$ is an $\mathbb{R}$-linear functional satisfying

$$\operatorname{Re}\phi(x) \leq m(x) \quad \forall\, x \in V.$$

By Theorem 5, there exists an $\mathbb{R}$-linear functional $\psi : X \to \mathbb{R}$ such that

$$\psi(x) = \operatorname{Re}\phi(x) \quad \forall\, x \in V,$$

and

$$\psi(x) \leq m(x) \quad \forall\, x \in X.$$

Now, define a map $\varphi : X \to \mathbb{C}$ by

$$\varphi(x) = \psi(x) - i\psi(ix), \qquad x \in X.$$

From the $\mathbb{R}$-linearity of ψ, we immediately have

$$\varphi(x+y) = \psi(x+y) - i\psi(i(x+y)) = \psi(x) + \psi(y) - i(\psi(ix) + \psi(iy))$$

$$= \psi(x) - i\psi(ix) + \psi(y) - i\psi(iy) = \varphi(x) + \varphi(y) \quad \forall\, x, y \in X,$$

and if $a + ib \in \mathbb{C}$, then

$$\varphi((a+ib)x) = \psi(ax + ibx) - i\psi(iax - bx) = a\psi(x) + b\psi(ix) - ia\psi(ix) + ib\psi(x)$$

$$= a(\psi(x) - i\psi(ix)) + ib(\psi(x) - i\psi(ix)) = (a+ib)(\psi(x) - i\psi(ix)) = (a+ib)\varphi(x) \quad \forall\, x \in X.$$

Thus, φ is a $\mathbb{C}$-linear functional on X. Moreover,

$$\varphi(x) = \psi(x) - i\psi(ix) = \operatorname{Re}\phi(x) - i\operatorname{Re}\phi(ix)$$

$$= \operatorname{Re}\phi(x) - i\operatorname{Re} i\phi(x) = \operatorname{Re}\phi(x) + i\operatorname{Im}\phi(x) = \phi(x) \quad \forall\, x \in V,$$

and

$$\operatorname{Re}\varphi(x) = \psi(x) \leq m(x) \quad \forall\, x \in X.$$

□

The Hahn-Banach Theorem leads to an important consequence.

Corollary 7. *Let X be a normed vector space over $\mathbb{F}$ and let $Y \subset X$ be a closed subspace of X. For every $x_0 \in X \setminus Y$ there exists $\varphi \in X^*$ such that $\varphi(x_0) = \inf_{y \in Y} \|x_0 - y\| > 0$, $\|\varphi\| \leq 1$, and $\varphi(y) = 0$ for all $y \in Y$. In particular, if $Y = \{0\}$, then $\varphi(x_0) = \|x_0\|$.*

Proof. Let $m : X \to \mathbb{R}$ be the function defined by

$$m(x) = \inf_{y \in Y} \|x - y\|, \qquad x \in X.$$

Note that $0 \in Y$, so $m(0) = 0$. If $\alpha \in \mathbb{F} \setminus \{0\}$ and $x \in X$, since Y is a vector space, we have $y \in Y \iff \frac{y}{\alpha} \in Y$. Hence,

$$m(\alpha x) = \inf_{y \in Y} \|\alpha x - y\| = \inf_{y \in Y} |\alpha| \left\| x - \frac{y}{\alpha} \right\| = |\alpha| \inf_{z \in Y} \|x - z\| = |\alpha| \, m(x).$$

Fix $x_1, x_2 \in X$, and let $y, z \in Y$ be arbitrary. Then $y + z \in Y$, so we have

$$m(x_1 + x_2) \leq \|x_1 + x_2 - (y + z)\| \leq \|x_1 - y\| + \|x_2 - z\|$$

$$\iff m(x_1 + x_2) - \|x_1 - y\| \leq \|x_2 - z\|.$$

Since $y, z \in Y$ are arbitrary, we conclude that for every $y \in Y$

$$m(x_1 + x_2) - \|x_1 - y\| \leq m(x_2)$$

$$\iff m(x_1 + x_2) - m(x_2) \leq \|x_1 - y\|.$$

Since $y \in Y$ is arbitrary, we conclude

$$m(x_1 + x_2) - m(x_2) \leq m(x_1)$$

$$\iff m(x_1 + x_2) \leq m(x_1) + m(x_2).$$

Thus, m satisfies the conditions (i) and (ii) in Theorems 5 and 6.

Let $x_0 \in X \setminus Y$. We claim $m(x_0) \neq 0$. Indeed, if $m(x_0) = 0$, then for every $n \in \mathbb{N}$ we have

$$0 = m(x_0) = \inf_{y \in Y} \|x_0 - y\| < \frac{1}{n}$$

which implies there exists $y_n \in Y$ such that

$$\|x_0 - y_n\| < \frac{1}{n}.$$

The above inequality then implies that the sequence $\{y_n\}_{n=1}^{\infty}$ of points in Y converges to x_0. Since Y is closed, this implies $x_0 \in Y$, which contradicts that $x_0 \in X \setminus Y$. Therefore, $m(x_0) \neq 0$. In particular, $m(x_0) > 0$.

Let $V = \{\alpha x_0 : \alpha \in \mathbb{F}\}$. Define a linear functional $\phi : V \to \mathbb{F}$ by $\phi(\alpha x_0) = \alpha m(x_0)$. Then for each $\alpha \in \mathbb{F}$ we have

$$\operatorname{Re} \phi(\alpha x_0) = \operatorname{Re}(\alpha m(x_0)) = \operatorname{Re}(\alpha) m(x_0) \leq |\alpha| \, m(x_0) = m(\alpha x_0).$$

Hence,

$$\operatorname{Re}\phi(x) \leq m(x) \quad \forall\, x \in V.$$

Therefore, by Theorem 5 for $\mathbb{F} = \mathbb{R}$ and Theorem 6 for $\mathbb{F} = \mathbb{C}$, there exists a linear functional $\varphi : X \to \mathbb{F}$ such that

$$\varphi(x) = \phi(x) \quad \forall\, x \in V,$$

and

$$\operatorname{Re}\varphi(x) \leq m(x) \quad \forall\, x \in X.$$

Now, for each $x \in X$ with $\varphi(x) \neq 0$, taking $\alpha = \dfrac{\overline{\varphi(x)}}{|\varphi(x)|}$, we first recall that

$$m(\alpha x) = |\alpha|\, m(x) = m(x),$$

and so

$$|\varphi(x)| = \alpha\varphi(x) = \varphi(\alpha x) = \operatorname{Re}\varphi(\alpha x) \leq m(\alpha x) = m(x).$$

Therefore, we conclude

$$|\varphi(x)| \leq m(x) \quad \forall\, x \in X.$$

By definition, $m(y) = 0$ for all $y \in Y$, so the above inequality implies $\varphi(y) = 0$ for all $y \in Y$. Moreover, since $m(x) \leq \|x\|$ for all $x \in X$, the previous inequality implies

$$|\varphi(x)| \leq \|x\| \quad \forall\, x \in X.$$

In particular, for all $x \in X$ with $\|x\| \leq 1$ we have $|\varphi(x)| \leq 1$. Thus, $\varphi \in X^*$ and $\|\varphi\| \leq 1$.

Lastly, we observe

$$\varphi(x_0) = \phi(x_0) = m(x_0) = \inf_{y \in Y} \|x_0 - y\| > 0.$$

□

There is a variation of the contrapositive of Corollary 7 that is quite important, so we state it as a theorem below.

Theorem 8. *Let X be a normed vector space over $\mathbb{F}$, and let $S \subset X$. Let $Y = \overline{span(S)}$ be the closure of the linear span of S. Let $x_0 \in X$. Suppose that for every $\varphi \in X^*$ such that $\varphi(s) = 0$ for all $s \in S$ we have $\varphi(x_0) = 0$. Then $x_0 \in Y$.*

Proof. Assume for sake of contradiction that $x_0 \in X \setminus Y$. By Corollary 6 there exists $\varphi \in X^*$ such that $\varphi(x_0) \neq 0$ and $\varphi(y) = 0$ for all $y \in Y$. Since $S \subset Y$, it follows that $\varphi(s) = 0$ for all $s \in S$. By our assumption, we must have $\varphi(x_0) = 0$, a contradiction. Therefor, we conclude $x_0 \in Y$.

□

We can take it one step further and consider the double dual of X, which is the space $X^{**} = (X^*)^* = B(X^*, \mathbb{F})$. This may seem a bit unnecessary, but this gives us a way to embed any normed vector space X into a Banach space. To do this, we define a canonical embedding $\iota : X \to X^{**}$ where for each $x \in X$ the map $\iota(x) \in X^{**}$ is defined by

$$\iota(x)(\varphi) = \varphi(x) \quad \forall\, \varphi \in X^*.$$

Let $x \in X$. Then, for all $\varphi \in X^*$ such that $\|\varphi\| \leq 1$, we have

$$|\iota(x)(\varphi)| = |\varphi(x)| \leq \|\varphi\| \|x\| \leq \|x\|,$$

$$\implies \|\iota(x)\| \leq \|x\|.$$

Now, for each $x \in X \setminus \{0\}$, by Corollary 7, there is $\varphi \in X^*$ with $\|\varphi\| \leq 1$ and $\varphi(x) = \|x\|$. Hence

$$|\iota(x)(\varphi)| = |\varphi(x)| = \|x\|,$$

and so we conclude

$$\|\iota(x)\| = \|x\| \quad \forall\, x \in X.$$

That is, ι is an isometry (i.e. norm preserving). In particular, ι is injective, so the map $\iota : X \to \iota(X)$ is a bijection. Now, there is an inverse linear map $\iota^{-1} : \iota(X) \to X$ which is also an isometry. We say that X is isometrically isomorphic to the space $\iota(X)$.
The advantage here is that $\iota(X)$ is a subspace of the Banach space X^{**}, hence, $\overline{\iota(X)}$ is a Banach space (it is a closed subspace of a Banach space). Thus, by identifying X with $\iota(X)$, the space $\overline{\iota(X)}$ is a Banach space containing X as a dense subset. We call $\overline{\iota(X)}$ the completion of X.

Definition 6. Let X be a normed vector space. We say a sequence $\{x_n\}_{n=1}^{\infty}$ in X converges weakly to $x \in X$ if

$$\lim_{n\to\infty} \varphi(x_n) = \varphi(x) \quad \forall\, \varphi \in X^*.$$

We say a sequence $\{\varphi_n\}_{n=1}^{\infty}$ in X^* converges weak-* (read "weak star") to $\varphi \in X^*$ if

$$\lim_{n\to\infty} \varphi_n(x) = \varphi(x) \quad \forall\, x \in X.$$

Recall the Bolzano-Weierstrass Theorem which states that every bounded sequence $\{\alpha_n\}_{n=1}^{\infty}$ in $\mathbb{F}$ has a subsequence which converges in $\mathbb{F}$. The next result is a variation of that theorem in normed vector spaces. The special case we will establish is suitable for our purposes. Those who are familiar with the Arzelà-Ascoli Theorem might recognize the diagonalization argument.

Theorem 9 (Banach-Alaoglu Theorem)**.** *Let X be a normed vector space and suppose there is a countable subset $S = \{s_1, s_2, s_3, \dots\}$ of X such that $\overline{S} = X$ (we say X is separable). Then every bounded sequence in X^* has a subsequence which converges weak-* in X^*.*

Proof. Let $\{\varphi_n\}_{n=1}^{\infty}$ be a bounded sequence in X^*. There exists $M > 0$ such that

$$\|\varphi_n\| \leq M \quad \forall\ n \in \mathbb{N}.$$

Then we have

$$|\varphi_n(s_1)| \leq \|\varphi_n\| \, \|s_1\| \leq M \, \|s_1\| \quad \forall\ n \in \mathbb{N}.$$

By the Bolzano-Weierstrass Theorem, there is a subsequence $\left\{\varphi_{n_k^1}\right\}_{k=1}^{\infty}$ of $\{\varphi_n\}_{n=1}^{\infty}$ and a point $\varphi(s_1) \in \mathbb{F}$ such that

$$\lim_{k\to\infty} \varphi_{n_k^1}(s_1) = \varphi(s_1).$$

Similarly, we have

$$\left|\varphi_{n_k^1}(s_2)\right| \leq M \, \|s_2\| \quad \forall\ k \geq 1.$$

By the Bolzano-Weierstrass Theorem, there is a subsequence $\left\{\varphi_{n_k^2}\right\}_{k=1}^{\infty}$ of $\left\{\varphi_{n_k^1}\right\}_{k=1}^{\infty}$ and a point $\varphi(s_2) \in \mathbb{F}$ such that

$$\lim_{k\to\infty} \varphi_{n_k^2}(s_2) = \varphi(s_2).$$

Assume that we have constructed sequences $\left\{\varphi_{n_k^1}\right\}_{k=1}^{\infty}, \left\{\varphi_{n_k^2}\right\}_{k=1}^{\infty}, \dots, \left\{\varphi_{n_k^m}\right\}_{k=1}^{\infty}$ such that $\left\{\varphi_{n_k^1}\right\}_{k=1}^{\infty}$ is a subsequence of $\{\varphi_n\}_{n=1}^{\infty}$, $\left\{\varphi_{n_k^{j+1}}\right\}_{k=1}^{\infty}$ is a subsequence of $\left\{\varphi_{n_k^j}\right\}_{k=1}^{\infty}$ for all $j \in \{1, 2, \dots, m-1\}$, and there exist points $\varphi(s_1), \varphi(s_2), \dots, \varphi(s_m) \in \mathbb{F}$ such that

$$\lim_{k\to\infty} \varphi_{n_k^j}(s_j) = \varphi(s_j) \quad \forall\ j \in \{1, 2, \dots, m\}.$$

Then we have

$$\left|\varphi_{n_k^m}(s_{m+1})\right| \leq M \, \|s_{m+1}\| \quad \forall\ k \geq 1.$$

By the Bolzano-Weierstrass Theorem, there is a subsequence $\left\{\varphi_{n_k^{m+1}}\right\}_{k=1}^{\infty}$ of $\left\{\varphi_{n_k^m}\right\}_{k=1}^{\infty}$ and a point $\varphi(s_{m+1}) \in \mathbb{F}$ such that

$$\lim_{k\to\infty} \varphi_{n_k^{m+1}}(s_{m+1}) = \varphi(s_{m+1}).$$

By recursion, there exists a collection of sequences $\left\{\varphi_{n_k^j}\right\}_{k=1}^{\infty}$, $j \in \mathbb{N}$, and a sequence of points $\{\varphi(s_j)\}_{j=1}^{\infty}$ in $\mathbb{F}$ such that $\left\{\varphi_{n_k^1}\right\}_{k=1}^{\infty}$ is a subsequence of $\{\varphi_n\}_{n=1}^{\infty}$, $\left\{\varphi_{n_k^{j+1}}\right\}_{k=1}^{\infty}$ is a subsequence of $\left\{\varphi_{n_k^j}\right\}_{k=1}^{\infty}$ for all $j \geq 1$ and

$$\lim_{k\to\infty} \varphi_{n_k^j}(s_j) = \varphi(s_j) \quad \forall\, j \in \mathbb{N}.$$

For each $k \in \mathbb{N}$, let $\varphi_{n_k} = \varphi_{n_k^k}$. One can show that the sequence $\{\varphi_{n_k}\}_{k=1}^{\infty}$ is a subsequence of $\{\varphi_n\}_{n=1}^{\infty}$. Moreover, for each $j \in \mathbb{N}$, the sequence $\{\varphi_{n_k}\}_{k=j}^{\infty}$ is a subsequence of $\left\{\varphi_{n_k^j}\right\}_{k=1}^{\infty}$, and so we have

$$\lim_{k\to\infty} \varphi_{n_k}(s_j) = \varphi(s_j).$$

Let $x \in X$. Fix $\varepsilon > 0$. There exists $j \geq 1$ such that

$$\|x - s_j\| \leq \frac{\varepsilon M}{3}.$$

Choose $N \in \mathbb{N}$ such that

$$k, l \geq N \implies |\varphi_{n_k}(s_j) - \varphi_{n_l}(s_j)| \leq \frac{\varepsilon}{3}.$$

Then, for all $k, l \geq N$ we have

$$\begin{aligned} |\varphi_{n_k}(x) - \varphi_{n_l}(x)| &\leq |\varphi_{n_k}(x) - \varphi_{n_k}(s_j)| + |\varphi_{n_k}(s_j) - \varphi_{n_l}(s_j)| + |\varphi_{n_l}(s_j) - \varphi_{n_l}(x)| \\ &\leq \|\varphi_{n_k}\|\,\|x - s_j\| + \frac{\varepsilon}{3} + \|\varphi_{n_l}\|\,\|s_j - x\| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Thus, $\{\varphi_{n_k}(x)\}_{k=1}^{\infty}$ is a Cauchy sequence in $\mathbb{F}$. Hence, $\{\varphi_{n_k}(x)\}_{k=1}^{\infty}$ converges in $\mathbb{F}$. Therefore, we can define a function $\varphi : X \to \mathbb{F}$ by the rule

$$\varphi(x) = \lim_{k\to\infty} \varphi_{n_k}(x), \quad x \in X.$$

Just as in the proof of Theorem 3, it follows that φ is linear since it is the pointwise limit of the linear maps φ_{n_k}. To show that φ is bounded, let $x \in X$ be such that $\|x\| \leq 1$. Then we have

$$|\varphi(x)| \leq |\varphi(x) - \varphi_{n_k}(x)| + |\varphi_{n_k}(x)| \leq |\varphi(x) - \varphi_{n_k}(x)| + M \quad \forall\, k \geq 1.$$

Sending k to infinity in the above inequality implies

$$|\varphi(x)| \leq M.$$

Since $x \in X$ with $\|x\| \leq 1$ is arbitrary, we conclude $\|\varphi(x)\| \leq M$ and so $\varphi \in X^*$. By the definition of φ, it follows that $\{\varphi_{n_k}\}_{k=1}^{\infty}$ converges weak-* to φ.

□

2 Hilbert Spaces

Definition 7. Let X be a vector space over $\mathbb{F}$. An inner product on X is a function $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{F}$ which satisfies the following properties:
(i) $\langle x, x\rangle \geq 0$ for all $x \in X$,
(ii) $\langle x, x\rangle = 0$ if and only if $x = 0$,
(iii) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ for all $x, y, z \in X$,
(iv) $\langle \alpha x, y\rangle = \alpha \langle x, y\rangle$ for all $x, y \in X$ and $\alpha \in \mathbb{F}$,
(v) $\langle x, y\rangle = \overline{\langle y, x\rangle}$ for all $x, y \in X$.

Note that if $\langle\cdot,\cdot\rangle$ is an inner product on X, then for each $x \in X$, the function $\langle\cdot, x\rangle : X \to \mathbb{F}$ is a linear functional on X. Furthermore, for each $x, y, z \in X$ and $\alpha \in \mathbb{F}$, we have *conjugate* linearity in the second argument:

$$\langle x, \alpha y + z\rangle = \overline{\langle \alpha y + z, x\rangle} = \overline{\alpha \langle y, x\rangle + \langle z, x\rangle} = \overline{\alpha}\overline{\langle y, x\rangle} + \overline{\langle z, x\rangle} = \overline{\alpha} \langle x, y\rangle + \langle x, z\rangle .$$

In particular, these linearity relationships imply

$$\langle 0, x\rangle = 0 = \langle x, 0\rangle \quad \forall\, x \in X.$$

A vector space X together with an inner product $\langle\cdot,\cdot\rangle$ is called an inner product space.

Let X be an inner product space. Condition (i) in the definition of an inner product implies that we can define a function $\|\cdot\| : X \to [0, \infty)$ by the rule

$$\|x\| = \sqrt{\langle x, x\rangle}, \qquad x \in X.$$

As suggested by the notation, we will show that $\|\cdot\|$ defines a norm on X. Condition (ii) in the definition of an inner product implies

$$\|x\| = 0 \iff x = 0.$$

Let $(\alpha, x) \in \mathbb{F} \times X$. The conjugate linearity of $\langle\cdot,\cdot\rangle$ implies

$$\langle \alpha x, \alpha x\rangle = \alpha \langle x, \alpha x\rangle = \alpha\overline{\alpha} \langle x, x\rangle = |\alpha|^2 \langle x, x\rangle .$$

Hence,

$$\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{|\alpha|^2 \langle x, x\rangle} = |\alpha| \sqrt{\langle x, x\rangle} = |\alpha| \, \|x\| .$$

It remains to establish the triangle inequality. To do this, we will establish another important relationship between the functions $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$.

Proposition 10 (Cauchy-Schwarz Inequality)**.** *Let X be an inner product space. Then*

$$|\langle x, y\rangle| \leq \|x\| \, \|y\| \quad \forall\, x, y \in X.$$

Furthermore, we have equality if and only if $y = 0$ or there is $\alpha \in \mathbb{F}$ such that $x = \alpha y$.

Proof. Let $x, y \in X$. If $y = 0$, we observe that $\langle x, y\rangle = 0 = \|x\|\,\|y\|$, so we assume $y \neq 0$. For $\alpha \in \mathbb{F}$, we have:

$$\langle x - \alpha y, x - \alpha y\rangle = \langle x, x - \alpha y\rangle - \alpha\,\langle y, x - \alpha y\rangle = \langle x, x\rangle - \overline{\alpha}\,\langle x, y\rangle - \alpha\overline{\langle x, y\rangle} + |\alpha|^2\,\langle y, y\rangle$$

$$\implies 0 \leq \langle x - \alpha y, x - \alpha y\rangle = \|x\|^2 - 2\mathrm{Re}\left(\overline{\alpha}\,\langle x, y\rangle\right) + |\alpha|^2\,\|y\|^2.$$

Choosing the special value $\alpha = \dfrac{\langle x, y\rangle}{\|y\|^2}$ gives

$$0 \leq \langle x - \alpha y, x - \alpha y\rangle = \|x\|^2 - 2\frac{|\langle x, y\rangle|^2}{\|y\|^2} + \frac{|\langle x, y\rangle|^2}{\|y\|^2} = \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2}$$

$$\iff \frac{|\langle x, y\rangle|^2}{\|y\|^2} \leq \|x\|^2$$

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

$$\iff |\langle x, y\rangle| \leq \|x\|\,\|y\|.$$

This proves the desired inequality.

Now, if $|\langle x, y\rangle| = \|x\|\,\|y\|$, then the above chain of inequalities become equality, and we have $0 = \langle x - \alpha y, x - \alpha y\rangle$, which implies $x = \alpha y$.

Conversely, if $x = \alpha y$ for some $\alpha \in \mathbb{F}$, we have (recall we are assuming $y \neq 0$)

$$\alpha\,\|y\|^2 = \alpha\,\langle y, y\rangle = \langle \alpha y, y\rangle = \langle x, y\rangle$$

$$\implies \alpha = \frac{\langle x, y\rangle}{\|y\|^2}.$$

Hence,

$$\|x\|^2 = \langle x, x\rangle = \langle \alpha y, \alpha y\rangle = |\alpha|^2\,\|y\|^2 = \frac{|\langle x, y\rangle|^2}{\|y\|^2}$$

$$\implies |\langle x, y\rangle|^2 = \|x\|^2\,\|y\|^2$$

$$\implies |\langle x, y\rangle| = \|x\|\,\|y\|.$$

□

Proposition 11 (Triangle Inequality)**.** *Let X be an inner product space. Then*

$$\|x + y\| \leq \|x\| + \|y\| \quad \forall\, x, y \in X.$$

Furthermore, we have equality if and only if $y = 0$ or there exists $\alpha \geq 0$ such that $x = \alpha y$.

Proof. Let $x, y \in X$. Using the Cauchy-Schwarz inequality, we have

$$\text{Re } \langle x, y\rangle \le |\langle x, y\rangle| \le \|x\| \, \|y\| \, .$$

Hence,

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + \langle y, x\rangle + \langle x, y\rangle + \langle y, y\rangle$$

$$= \|x\|^2 + 2\text{Re } \langle x, y\rangle + \|y\|^2 \le \|x\|^2 + 2\|x\| \, \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

$$\implies \|x+y\| \le \|x\| + \|y\| \, .$$

Now, if $\|x+y\| = \|x\| + \|y\|$, then the above inequalities imply $\text{Re } \langle x, y\rangle = |\langle x, y\rangle| = \|x\| \, \|y\|$. Hence, by Proposition 10, if $y \neq 0$ then there is $\alpha \in \mathbb{F}$ such that $x = \alpha y$. The proof of Proposition 10 shows that $\alpha = \dfrac{\langle x, y\rangle}{\|y\|^2}$, and the equality $\text{Re } \langle x, y\rangle = |\langle x, y\rangle|$ implies $\langle x, y\rangle = |\langle x, y\rangle| \ge 0$. Therefore, we conclude $\alpha \ge 0$.

Conversely, if there exists $\alpha \ge 0$ such that $x = \alpha y$, then $\alpha + 1 \ge 0$ and so

$$\|x+y\| = \|\alpha y + y\| = \|(\alpha+1)y\| = (\alpha+1)\|y\| = \alpha\|y\| + \|y\| = \|\alpha y\| + \|y\| = \|x\| + \|y\| \, .$$

□

Now that we have the triangle inequality, every inner product space X has a norm given by the function $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$. An inner product space which is complete with respect to the norm induced by its inner product is called a Hilbert space.

Let X be an inner product space and let $y \in X$. We already mentioned that the map $\langle \cdot, y\rangle : X \to \mathbb{F}$ is a linear functional on X. With the Cauchy-Schwarz inequality, for every $x \in X$ with $\|x\| \le 1$ we have $|\langle x, y\rangle| \le \|y\|$. Moreover, if $y \neq 0$, then $x = \dfrac{y}{\|y\|}$ satisfies $\|x\| = 1$ and $\langle x, y\rangle = \|y\|$. Therefore, we conclude $\langle \cdot, y\rangle \in X^*$ and $\|\langle \cdot, y\rangle\| = \|y\|$ (the same conclusion is true if $y = 0$).

Proposition 12 (Pythagorean Theorem)**.** *Let X be an inner product space. For every $x, y \in X$ such that $\langle x, y\rangle = 0$, we have*

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 \, .$$

Proof. Let $x, y \in X$ be such that $\langle x, y\rangle = 0$. Then we have

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + \langle y, x\rangle + \langle x, y\rangle + \langle y, y\rangle = \|x\|^2 + \|y\|^2 \, .$$

□

Proposition 13 (Parallelogram Law). *Let X be an inner product space. Then*

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right) \quad \forall\, x, y \in X.$$

Proof. Let $x, y \in X$. Then

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + 2\mathrm{Re}\,\langle x, y\rangle + \|y\|^2,$$

and

$$\|x-y\|^2 = \langle x-y, x-y\rangle = \|x\|^2 - 2\mathrm{Re}\,\langle x, y\rangle + \|y\|^2.$$

Hence, we have

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

□

Definition 8. Let X be an inner product space. We say $x, y \in X$ are orthogonal if $\langle x, y\rangle = 0$. If $S \subset X$, then the orthogonal complement of S, denoted $S^\perp$, is the set of all elements of X which are orthogonal to every element of S. That is, $S^\perp = \{x \in X : \langle x, s\rangle = 0 \quad \forall\, s \in S\}$.

One can verify that $S^\perp$ is a closed subspace of X for any subset S of X, (in particular, S itself does not need to be a vector space) and $S^\perp = \overline{\mathrm{span}(S)}^\perp$. We have $S \cap S^\perp \subset \{0\}$ for any subset S of X, and we have equality if and only if $0 \in S$. Moreover, since every element of S is orthogonal to every element of $S^\perp$, it follows that $S \subset (S^\perp)^\perp$. We will say more on this later.

Lemma 14. *Let X be an inner product space and let Y be a subspace of X. For every $x \in X$, there is at most one $y \in Y$ such that $x - y \in Y^\perp$.*

Proof. Let $x \in X$. Suppose there exist y_1, $y_2 \in Y$ such that $x - y_1$, $x - y_2 \in Y^\perp$. Then

$$y_2 - y_1 = (x - y_1) - (x - y_2) \in Y \cap Y^\perp = \{0\}.$$

Hence, $y_1 = y_2$. □

A subset C of a vector space X is said to be convex if for every $x, y \in C$ and $t \in [0,1]$ we have $tx + (1-t)y \in C$. Although we will mainly use the following lemma for vector spaces, the more general statement yields itself with the same amount of effort.

Lemma 15 (Best Approximations). *Let X be a Hilbert space and let $C \subset X$ be a nonempty closed convex set. For every $x_0 \in X$ there exists a unique $y_0 \in C$ such that*

$$\|x_0 - y_0\| = \inf_{y \in C} \|x_0 - y\|.$$

Proof. Let $x_0 \in X$. By definition of infimum, there exists a sequence $\{y_n\}_{n=1}^{\infty}$ of points in C such that

$$\lim_{n\to\infty} \|x_0 - y_n\| = \inf_{y\in C} \|x_0 - y\| .$$

Now, for $n, m \geq 1$, using the Parallelogram Law with $x = \dfrac{x_0 - y_n}{2}$ and $y = \dfrac{x_0 - y_m}{2}$ we have

$$2\left(\left\|\frac{x_0 - y_n}{2}\right\|^2 + \left\|\frac{x_0 - y_m}{2}\right\|^2\right) = \left\|x_0 - \left(\frac{y_n + y_m}{2}\right)\right\|^2 + \left\|\frac{y_n - y_m}{2}\right\|^2$$

$$\implies \|y_n - y_m\|^2 = 2\left(\|x_0 - y_n\|^2 + \|x_0 - y_m\|^2\right) - 4\left\|x_0 - \left(\frac{y_n + y_m}{2}\right)\right\|^2 .$$

By convexity, $\dfrac{y_n + y_m}{2} \in C$ for all $n, m \geq 1$. Hence, the above equality gives

$$\|y_n - y_m\|^2 \leq 2\left(\|x_0 - y_n\|^2 + \|x_0 - y_m\|^2\right) - 4\left(\inf_{y\in C} \|x_0 - y\|\right)^2 \quad \forall\, n, m \geq 1.$$

Fix $\varepsilon > 0$, and choose $N \in \mathbb{N}$ such that

$$\|x_0 - y_n\|^2 - \left(\inf_{y\in C} \|x_0 - y\|\right)^2 \leq \frac{\varepsilon^2}{4} \quad \forall\, n \geq N.$$

Then,

$$2\left(\|x_0 - y_n\|^2 + \|x_0 - y_m\|^2\right) - 4\left(\inf_{y\in C} \|x_0 - y\|\right)^2$$

$$= 2\left(\|x_0 - y_n\|^2 - \left(\inf_{y\in C} \|x_0 - y\|\right)^2\right) + 2\left(\|x_0 - y_m\|^2 - \left(\inf_{y\in C} \|x_0 - y\|\right)^2\right)$$

$$\leq 2\left(\frac{\varepsilon^2}{4}\right) + 2\left(\frac{\varepsilon^2}{4}\right) = \varepsilon^2 \quad \forall\, n, m \geq N,$$

and so

$$\|y_n - y_m\|^2 \leq \varepsilon^2 \quad \forall\, n, m \geq N$$

$$\implies \|y_n - y_m\| \leq \varepsilon \quad \forall\, n, m \geq N.$$

Therefore, $\{y_n\}_{n=1}^{\infty}$ is a Cauchy sequence in X. Since X is a Hilbert space, there exists $y_0 \in X$ such that $y_n \to y_0$. Since C is closed, we conclude $y_0 \in C$. Then, by the continuity of the norm, we conclude

$$\|x_0 - y_0\| = \lim_{n\to\infty} \|x_0 - y_n\| = \inf_{y\in C} \|x_0 - y\|.$$

To show that y_0 is unique, let $z_0 \in C$ be such that $\|x_0 - z_0\| = \inf_{y\in C} \|x_0 - y\|$. By another application of the Parallelogram Law with $x = \frac{x_0 - y_0}{2}$ and $y = \frac{x_0 - z_0}{2}$ we have

$$2\left(\left\|\frac{x_0 - y_0}{2}\right\|^2 + \left\|\frac{x_0 - z_0}{2}\right\|^2\right) = \left\|x_0 - \left(\frac{y_0 + z_0}{2}\right)\right\|^2 + \left\|\frac{y_0 - z_0}{2}\right\|^2$$

$$\implies \|y_0 - z_0\|^2 = 2\left(\|x_0 - y_0\|^2 + \|x_0 - z_0\|^2\right) - 4\left\|x_0 - \left(\frac{y_0 + z_0}{2}\right)\right\|^2.$$

By convexity, $\frac{y_0 + z_0}{2} \in C$. Together with $\|x_0 - y_0\| = \inf_{y\in C} \|x_0 - y\| = \|x_0 - z_0\|$, the above equality gives

$$\|y_0 - z_0\|^2 \leq 2\left(\|x_0 - y_0\|^2 + \|x_0 - z_0\|^2\right) - 4\left(\inf_{y\in C} \|x_0 - y\|\right)^2 = 0$$

$$\implies \|y_0 - z_0\| = 0.$$

Thus, $y_0 = z_0$.

□

Theorem 16 (Orthogonal Projection)**.** *Let X be a Hilbert space and let Y be a closed subspace of X. Let $P : X \to Y$ denote the map such that for every $x \in X$ the point $P(x)$ is the unique point in Y such that $\|x - P(x)\| = \inf_{y\in Y} \|x - y\|$. Then $x - P(x) \in Y^\perp$ for all $x \in X$. If $x = y + z$ with $y \in Y$ and $z \in Y^\perp$, then $y = P(x)$. Furthermore, P is linear.*

Proof. By Lemma 15, P is a well-defined function. Let $x \in X$ and $y \in Y$. By definition of $P(x)$ we have

$$\|x - P(x)\|^2 = \left(\inf_{z\in Y} \|x - z\|\right)^2 \leq \|x - P(x) - y\|^2 = \langle x - P(x) - y, x - P(x) - y\rangle$$

$$= \langle x - P(x), x - P(x)\rangle - \langle x - P(x), y\rangle - \langle y, x - P(x)\rangle + \langle y, y\rangle$$

$$= \|x - P(x)\|^2 - 2\text{Re}\left(\langle x - P(x), y\rangle\right) + \|y\|^2.$$

Hence, we conclude

$$\text{Re}\left(\langle x - P(x), y\rangle\right) \leq \frac{\|y\|^2}{2} \quad \forall\, y \in Y.$$

Since $-y \in Y$ for every $y \in Y$, this inequality also implies

$$-\text{Re}\left(\langle x - P(x), y\rangle\right) = \text{Re}\left(\langle x - P(x), -y\rangle\right) \le \frac{\|-y\|^2}{2} = \frac{\|y\|^2}{2} \quad \forall\, y \in Y.$$

Therefore, we conclude

$$|\text{Re}\left(\langle x - P(x), y\rangle\right)| \le \frac{\|y\|^2}{2} \quad \forall\, y \in Y.$$

If $\mathbb{F} = \mathbb{C}$ then $iy \in Y$ for every $y \in Y$, and so

$$|\text{Im}\left(\langle x - P(x), y\rangle\right)| = |\text{Re}\left(-i\,\langle x - P(x), y\rangle\right)| = |\text{Re}\left(\langle x - P(x), iy\rangle\right)| \le \frac{\|iy\|^2}{2} = \frac{\|y\|^2}{2}$$

for every $y \in Y$. Therefore, these inequalities imply

$$|\langle x - P(x), y\rangle| \le |\text{Re}\left(\langle x - P(x), y\rangle\right)| + |\text{Im}\left(\langle x - P(x), y\rangle\right)| \le \|y\|^2 \quad \forall\, y \in Y.$$

Fix $y \in Y$, and let $\varepsilon > 0$. Let $\delta = \dfrac{\varepsilon}{\|y\|^2 + 1}$, so that $\delta\,\|y\|^2 \le \varepsilon$. The above inequality implies

$$\delta\,|\langle x - P(x), y\rangle| = |\delta\,\langle x - P(x), y\rangle| = |\langle x - P(x), \delta y\rangle| \le \|\delta y\|^2 = \delta^2\,\|y\|^2,$$

$$\implies |\langle x - P(x), y\rangle| \le \delta\,\|y\|^2 \le \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we conclude $\langle x - P(x), y\rangle = 0$. Since $y \in Y$ is arbitrary, we conclude $x - P(x) \in Y^\perp$.

Now, if $x = y + z$ with $y \in Y$ and $z \in Y^\perp$, then $x - y = z \in Y^\perp$. Since $x - P(x) \in Y^\perp$, Lemma 14 implies $y = P(x)$. To show that P is linear, let $x_1, x_2 \in X$ and $\alpha \in \mathbb{F}$. Then

$$\begin{aligned}\alpha x_1 + x_2 &= \alpha(P(x_1) + x_1 - P(x_1)) + P(x_2) + (x_2 - P(x_2)) \\ &= \alpha P(x_1) + P(x_2) + (\alpha(x_1 - P(x_1)) + (x_2 - P(x_2))),\end{aligned}$$

where $\alpha P(x_1) + P(x_2) \in Y$ and $\alpha(x_1 - P(x_1)) + (x_2 - P(x_2)) \in Y^\perp$. Therefore, we conclude $P(\alpha x_1 + x_2) = \alpha P(x_1) + P(x_2)$.

□

The map $P : X \to Y$ in Theorem 16 is called the orthogonal projection of X onto Y. Since $Y^\perp$ is also a closed subspace of X, we have the orthogonal projection $P^\perp : X \to Y^\perp$ of X onto $Y^\perp$. These maps give us a nice way to decompose points in a Hilbert space.

Corollary 17 (Orthogonal Decomposition)**.** *Let X be a Hilbert space and let Y be a closed subspace of X. Let $P : X \to Y$ be the orthogonal projection onto Y, and $P^\perp : X \to Y^\perp$ be the orthogonal projection onto $Y^\perp$. Then $x = P(x) + P^\perp(x)$ for all $x \in X$. If $x = y + z$ with $y \in Y$ and $z \in Y^\perp$ then $y = P(x)$ and $z = P^\perp(x)$. Moreover, P and $P^\perp$ are bounded linear maps with $\|P\| \le 1$ and $\|P^\perp\| \le 1$.*

Proof. The linearity of P and $P^{\perp}$ follow from Theorem 16.
Let $x \in X$. Since $Y \subset (Y^{\perp})^{\perp}$, we have

$$x - P^{\perp}(x),\ P(x) \in (Y^{\perp})^{\perp},$$

and

$$x - P(x),\ P^{\perp}(x) \in Y^{\perp}.$$

Hence

$$P^{\perp}(x) - (x - P(x)) = P(x) - (x - P^{\perp}(x)) \in Y^{\perp} \cap (Y^{\perp})^{\perp} = \{0\}.$$

Therefore, we have the decomposition $x = P(x) + P^{\perp}(x)$.

Now, suppose that $x = y + z$ with $y \in Y$ and $z \in Y^{\perp}$. By Theorem 16 using the orthogonal projection P, we conclude $y = P(x)$. We also observe $x = z + y$ with $z \in Y^{\perp}$ and $y \in (Y^{\perp})^{\perp}$. By Theorem 16 using the orthogonal projection $P^{\perp}$, we conclude $z = P^{\perp}(x)$.

It remains to show P and $P^{\perp}$ are bounded. For every $x \in X$ we have $P(x) = x - P^{\perp}(x)$, so

$$\|P(x)\| = \left\|x - P^{\perp}(x)\right\| = \inf_{z \in Y^{\perp}} \|x - z\| \leq \|x\| \quad \forall\, x \in X.$$

The above inequality now implies that $\|P(x)\| \leq 1$ for all $x \in X$ with $\|x\| \leq 1$. Therefore, P is bounded with $\|P\| \leq 1$.
Similarly, for every $x \in X$ we have $P^{\perp}(x) = x - P(x)$, so

$$\left\|P^{\perp}(x)\right\| = \|x - P(x)\| = \inf_{z \in Y} \|x - z\| \leq \|x\| \quad \forall\, x \in X.$$

Thus, we conclude $P^{\perp}$ is bounded with $\left\|P^{\perp}\right\| \leq 1$.

□

Using the orthogonal decomposition of points in Hilbert spaces, we can establish an interesting result regarding orthogonal complements.

Lemma 18. *Let X be a Hilbert space and $S \subset X$. Then $\overline{\text{span}(S)} = (S^{\perp})^{\perp}$.*

Proof. Since $S \subset (S^{\perp})^{\perp}$, and $(S^{\perp})^{\perp}$ is a closed subspace of X, we immediately have $\overline{\text{span}(S)} \subset (S^{\perp})^{\perp}$.
Let P be the orthogonal projection onto $\overline{\text{span}(S)}$. Let $x \in (S^{\perp})^{\perp}$. By Theorem 16, we have $x - P(x) \in \overline{\text{span}(S)}^{\perp} = S^{\perp}$, and $P(x) \in \overline{\text{span}(S)} \subset (S^{\perp})^{\perp}$, and so

$$x - P(x) \in S^{\perp} \cap (S^{\perp})^{\perp} = \{0\}.$$

Hence, $x = P(x) \in \overline{\text{span}(S)}$. Thus, $(S^{\perp})^{\perp} \subset \overline{\text{span}(S)}$.

Therefore, we conclude $\overline{\text{span}(S)} = (S^{\perp})^{\perp}$.

□

In particular, if Y is a closed subspace of a Hilbert space X, then $Y = (Y^\perp)^\perp$.

Earlier, we discussed that for every y in an inner product space X, the map $\langle \cdot, y\rangle : X \to \mathbb{F}$ is a bounded linear functional. On a Hilbert space, this gives a complete characterization of all bounded linear functionals.

Theorem 19 (Riesz Representation Theorem)**.** *Let X be a Hilbert space. For every $\varphi \in X^*$ there exists a unique $y \in X$ such that $\varphi(x) = \langle x, y\rangle$ for all $x \in X$.*

Proof. We first establish uniqueness. If $y, z \in X$ and $\langle x, y\rangle = \langle x, z\rangle$ for all $x \in X$. Then $\langle x, y - z\rangle = 0$ for all $x \in X$. In particular, $\langle y - z, y - z\rangle = 0 \implies y - z = 0 \implies y = z$.

Now, let $\varphi \in X^*$. If $\varphi = 0$, then $\varphi(x) = \langle x, 0\rangle$ for all $x \in X$.

Suppose $\varphi \neq 0$. There exists $x_0 \in X$ such that $\varphi(x_0) \neq 0$. Let $Y = \{x \in X : \varphi(x) = 0\}$. The continuity of φ implies that Y is a closed subspace of X. By Corollary 17, there exist $y_0 \in Y$ and $z_0 \in Y^\perp$ such that $x_0 = y_0 + z_0$. Then we have

$$\varphi(x_0) = \varphi(y_0) + \varphi(z_0) = \varphi(z_0).$$

In particular, $\varphi(z_0) \neq 0$ and $z_0 \neq 0$. Let $y = \dfrac{\overline{\varphi(z_0)} z_0}{\|z_0\|^2}$. Note that $y \in Y^\perp$, and

$$\langle z_0, y\rangle = \frac{\varphi(z_0)}{\|z_0\|^2} \langle z_0, z_0\rangle = \varphi(z_0).$$

Let $x \in X$. Then we have

$$\varphi\left(x - \frac{\varphi(x) z_0}{\varphi(z_0)}\right) = \varphi(x) - \varphi\left(\frac{\varphi(x) z_0}{\varphi(z_0)}\right) = \varphi(x) - \varphi(x) = 0.$$

Hence, $x - \dfrac{\varphi(x) z_0}{\varphi(z_0)} \in Y$. Since $y \in Y^\perp$, we have

$$\left\langle x - \frac{\varphi(x) z_0}{\varphi(z_0)}, y\right\rangle = 0,$$

$$\implies \langle x, y\rangle = \left\langle \frac{\varphi(x) z_0}{\varphi(z_0)}, y\right\rangle = \frac{\varphi(x)}{\varphi(z_0)} \langle z_0, y\rangle = \varphi(x).$$

Since $x \in X$ is arbitrary, we conclude $\varphi(x) = \langle x, y\rangle$ for all $x \in X$.

□

Definition 9. Let X be an inner product space. A subset S of X is said to be orthogonal if for every $s_1, s_2 \in S$ with $s_1 \neq s_2$ we have $\langle s_1, s_2\rangle = 0$. An orthogonal set S is said to be orthonormal if $\langle s, s\rangle = 1$ for all $s \in S$.

For our purposes, we will usually deal with countable orthogonal or orthonormal sets. In this case, we will usually refer to a countable set $S = \{e_j\}_{j=1}^{\infty}$ as being orthogonal if $\langle e_j, e_k \rangle = 0$ for all $j \neq k$, and orthonormal if S is orthogonal and $\langle e_j, e_j \rangle = 1$ for all $j \geq 1$. In particular, when we enumerate a countable orthogonal or orthonormal set $S = \{e_1, e_2, \dots\}$ (whether finite or infinite) we assume $e_j = e_k \iff j = k$.

Lemma 20 (Generalized Pythagorean Theorem)**.** *Let X be an inner product space and let $S = \{e_1, e_2, \dots, e_n\}$ be a finite orthogonal set in X. Suppose $\{\alpha_1, \alpha_2, \dots, \alpha_n\} \subset \mathbb{F}$. Then*

$$\left\| \sum_{j=1}^{n} \alpha_j e_j \right\|^2 = \sum_{j=1}^{n} |\alpha_j|^2 \|e_j\|^2 .$$

In particular, if S is orthonormal we have

$$\left\| \sum_{j=1}^{n} \alpha_j e_j \right\|^2 = \sum_{j=1}^{n} |\alpha_j|^2 .$$

Proof. We argue by induction on n. The case $n = 1$ is immediate. Suppose that for some $n \in \mathbb{N}$ we have proved the result for all orthogonal sets of size n and all subsets of $\mathbb{F}$ of size n. Let $\{e_1, e_2, \dots, e_{n+1}\} \subset X$ be orthogonal and let $\{\alpha_1, \alpha_2, \dots, \alpha_{n+1}\} \subset \mathbb{F}$. Then we have

$$\left\langle \sum_{j=1}^{n} \alpha_j e_j, \alpha_{n+1} e_{n+1} \right\rangle = \sum_{j=1}^{n} \alpha_j \overline{\alpha_{n+1}} \langle e_j, e_{n+1} \rangle = 0.$$

Hence, by the Pythagorean Theorem and our induction hypothesis, we have

$$\left\| \sum_{j=1}^{n+1} \alpha_j e_j \right\|^2 = \left\| \left(\sum_{j=1}^{n} \alpha_j e_j \right) + \alpha_{n+1} e_{n+1} \right\|^2$$

$$= \left\| \sum_{j=1}^{n} \alpha_j e_j \right\|^2 + \|\alpha_{n+1} e_{n+1}\|^2 = \left(\sum_{j=1}^{n} |\alpha_j|^2 \|e_j\|^2 \right) + |\alpha_{n+1}|^2 \|e_{n+1}\|^2 = \sum_{j=1}^{n+1} |\alpha_j|^2 \|e_j\|^2 .$$

Therefore, by induction, we have the generalized Pythagorean Theorem. □

Orthonormal sets have useful properties in Hilbert spaces.

Lemma 21. *Let X be a Hilbert space and let $S = \{e_j\}_{j=1}^{\infty}$ be an orthonormal subset of X. Let $P : X \to \overline{span(S)}$ be the orthogonal projection onto $\overline{span(S)}$. Then*

$$P(x) = \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j \quad \forall\ x \in X.$$

Proof. Let $x \in X$ and let $n \in \mathbb{N}$. Observe that for every $k \geq 1$ we have

$$\left\langle x - \sum_{j=1}^{n} \langle x, e_j \rangle \, e_j, e_k \right\rangle = \langle x, e_k \rangle - \left\langle \sum_{j=1}^{n} \langle x, e_j \rangle \, e_j, e_k \right\rangle$$

$$= \langle x, e_k \rangle - \sum_{j=1}^{n} \langle x, e_j \rangle \langle e_j, e_k \rangle = \langle x, e_k \rangle - \langle x, e_k \rangle = 0.$$

Therefore, it follows that $x - \sum_{j=1}^{n} \langle x, e_j \rangle \, e_j \in S^{\perp} = \overline{\text{span}(S)}^{\perp}$.
Hence, by the generalized Pythagorean Theorem we have

$$\|x\|^2 = \left\| \left(x - \sum_{j=1}^{n} \langle x, e_j \rangle \, e_j \right) + \sum_{j=1}^{n} \langle x, e_j \rangle \, e_j \right\|^2$$

$$= \left\| x - \sum_{j=1}^{n} \langle x, e_j \rangle \, e_j \right\|^2 + \left\| \sum_{j=1}^{n} \langle x, e_j \rangle \, e_j \right\|^2 \geq \left\| \sum_{j=1}^{n} \langle x, e_j \rangle \, e_j \right\|^2 = \sum_{j=1}^{n} |\langle x, e_j \rangle|^2 .$$

Since $n \in \mathbb{N}$ is arbitrary, it follows by the monotone convergence theorem that $\sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2$ converges and

$$\sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2 \leq \|x\|^2 .$$

Fix $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that

$$m > n \geq N \implies \sum_{j=n+1}^{m} |\langle x, e_j \rangle|^2 \leq \varepsilon^2.$$

Then for all $m > n \geq N$ we have

$$\left\| \sum_{j=1}^{m} \langle x, e_j \rangle \, e_j - \sum_{j=1}^{n} \langle x, e_j \rangle \, e_j \right\|^2 = \left\| \sum_{j=n+1}^{m} \langle x, e_j \rangle \, e_j \right\|^2 = \sum_{j=n+1}^{m} |\langle x, e_j \rangle|^2 \leq \varepsilon^2.$$

Hence, by interchanging the roles of m and n, it follows that

$$\left\| \sum_{j=1}^{m} \langle x, e_j \rangle \, e_j - \sum_{j=1}^{n} \langle x, e_j \rangle \, e_j \right\| \leq \varepsilon \quad \forall \; n, m \geq N.$$

Thus, $\left\{\sum_{j=1}^{n} \langle x, e_j \rangle e_j \right\}_{n=1}^{\infty}$ is a Cauchy sequence in $\overline{\text{span}(S)}$. Since $\overline{\text{span}(S)}$ is a Hilbert space (it is a closed subspace of a Hilbert space) it follows that $\lim_{n\to\infty} \sum_{j=1}^{n} \langle x, e_j \rangle e_j = \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j$ defines a point in $\overline{\text{span}(S)}$. We showed that $x - \sum_{j=1}^{n} \langle x, e_j \rangle e_j \in \overline{\text{span}(S)}^{\perp}$ for all $n \in \mathbb{N}$. Since $\overline{\text{span}(S)}^{\perp}$ is closed, it follows that $x - \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j = \lim_{n\to\infty} \left(x - \sum_{j=1}^{n} \langle x, e_j \rangle e_j \right) \in \overline{\text{span}(S)}^{\perp}$. Then we have the decomposition

$$x = \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j + \left(x - \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j \right).$$

with $\sum_{j=1}^{\infty} \langle x, e_j \rangle e_j \in \overline{\text{span}(S)}$ and $x - \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j \in \overline{\text{span}(S)}^{\perp}$. By Theorem 16, it follows that $P(x) = \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j$.

$\square$

There are a few more important properties of orthonormal sets that are useful to keep in mind.

Proposition 22. *Let X be a Hilbert space and let $S = \{e_j\}_{j=1}^{\infty}$ be an orthonormal subset of X. Let $Y = \overline{span(S)}$ and let $P : X \to Y$ be the orthogonal projection of X onto Y. The following are equivalent:*

(i) $X = Y$,
(ii) For all $x \in X$, if $\langle x, e_j \rangle = 0$ for all $j \geq 1$ then $x = 0$,
(iii) For all $x \in X$, $\|x\|^2 = \sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2$.

Proof. (i) $\Longrightarrow$ (ii)
Suppose $X = Y$. Let $x \in X$ and suppose $\langle x, e_j \rangle = 0$ for all $j \geq 1$. Since $x \in Y$, we have $x = P(x)$. Then by Lemma 21

$$x = P(x) = \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j = 0.$$

(ii) $\Longrightarrow$ (iii)

Suppose that for every $x \in X$, if $\langle x, e_j \rangle = 0$ for all $j \geq 1$ then $x = 0$. Let $x \in X$. By Theorem 16 we have $x - P(x) \in Y^{\perp}$. In particular, $\langle x - P(x), e_j \rangle = 0$ for all $j \geq 1$. By assumption, we have $x - P(x) = 0$. Thus

$$x = P(x) = \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j.$$

In particular, since the norm is a continuous function, we have by the generalized Pythagorean Theorem

$$\sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2 = \lim_{n\to\infty} \sum_{j=1}^{n} |\langle x, e_j \rangle|^2 = \lim_{n\to\infty} \left\| \sum_{j=1}^{n} \langle x, e_j \rangle e_j \right\|^2 = \left\| \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j \right\|^2 = \|x\|^2 .$$

(iii) $\Longrightarrow$ (i)

Suppose for all $x \in X$ we have $\|x\|^2 = \sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2$. Let $x \in X$. Lemma 21 implies

$$P(x) = \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j.$$

By continuity of the norm, we have

$$\sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2 = \lim_{n\to\infty} \sum_{j=1}^{n} |\langle x, e_j \rangle|^2 = \lim_{n\to\infty} \left\| \sum_{j=1}^{n} \langle x, e_j \rangle e_j \right\|^2 = \left\| \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j \right\|^2 = \|P(x)\|^2 .$$

In particular, $\|x\|^2 = \|P(x)\|^2$. By Theorem 16, $x - P(x) \in Y^{\perp}$, so $\langle x - P(x), P(x) \rangle = 0$. By the Pythagorean Theorem we have

$$\|x\|^2 = \|(x - P(x)) + P(x)\|^2 = \|x - P(x)\|^2 + \|P(x)\|^2$$
$$\Longrightarrow \|x - P(x)\|^2 = 0.$$

Thus, $x = P(x) \in Y$. Since $x \in X$ is arbitrary, we conclude $X \subset Y$. Since $Y \subset X$, we conclude $X = Y$.

$\square$

For our last result, we look at what the Banach-Alaoglu Theorem implies for Hilbert spaces.

Theorem 23. *Let X be a Hilbert space such that there exists a countable set $S \subset X$ with $\overline{S} = X$. Every bounded sequence in X has a weakly convergent subsequence.*

Proof. Let $\{x_n\}_{n=1}^{\infty}$ be a bounded sequence in X. Since $\|\langle \cdot, x_n \rangle\| = \|x_n\|$ for all $n \in \mathbb{N}$, it follows that $\{\langle \cdot, x_n \rangle\}_{n=1}^{\infty}$ is a bounded sequence in X^*. By Theorem 9, there exists a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ and $\varphi_0 \in X^*$ such that

$$\lim_{k \to \infty} \langle x, x_{n_k} \rangle = \varphi_0(x) \quad \forall\, x \in X.$$

By the Riesz Representation Theorem, there exists $x_0 \in X$ such that

$$\varphi_0(x) = \langle x, x_0 \rangle \quad \forall\, x \in X.$$

Let $\varphi \in X^*$. By the Riesz Representation Theorem, there is $y \in X$ such that $\varphi(\cdot) = \langle \cdot, y \rangle$. Then

$$\lim_{k \to \infty} \varphi(x_{n_k}) = \lim_{k \to \infty} \langle x_{n_k}, y \rangle = \overline{\varphi_0(y)} = \langle x_0, y \rangle = \varphi(x_0).$$

Since $\varphi \in X^*$ is arbitrary, we conclude $\{x_{n_k}\}_{k=1}^{\infty}$ converges weakly to x_0.

□