CHAPTER 3

FUNCTIONAL ANALYSIS

3.1 Function spaces

For $p \in [1, \infty)$, we define

$$L^p[0,1] = \{f : [0,1] \to \mathbb{R} \cup \{\pm\infty\} \mid |f|^p \text{ (Lebesgue) integrable}\}.$$

Moreover, we have the norm

$$\|f\|_p = \left(\int_0^1 |f|^p \, dm\right)^{1/p}.$$

For the L_∞ norm, we define the concept of essential supremum.

Definition 3.1.1. Let $f : [0,1] \to \mathbb{R} \cup \{\pm\infty\}$ be a function. Then, the *essential supremum* of f is the set

$$\text{ess sup}(f) = \inf\{a \in [0,1] \mid m(f(x) > a) = 0\}.$$

Then, we define

$$L^\infty[0,1] = \{f : [0,1] \to \mathbb{R} \cup \{\pm\infty\} \mid \text{ess sup}\,|f| \text{ exists}\}.$$

Moreover, we have the norm

$$\|f\|_\infty = \text{ess sup}\,|f|.$$

We will now prove Holder's Inequality. First, we start with a lemma.

Lemma 3.1.2. *Let $a, b \in [0, \infty)$, and $\lambda \in (0,1)$. Then,*

$$a^\lambda b^{1-\lambda} \le \lambda a + (1-\lambda) b.$$

Moreover, we have equality if and only if $a = b$.

Proof. If $b = 0$, then

$$a^\lambda b^{1-\lambda} = 0 \le \lambda a.$$

Instead, assume that $b \neq 0$. In that case, define the function $f : (0, \infty) \to \mathbb{R}$ given by

$$f(t) = \lambda t + 1 - \lambda - t^\lambda.$$

We have

$$f'(t) = \lambda - \lambda t^{\lambda - 1}, \qquad f''(t) = -(\lambda - 1)\lambda t^{\lambda - 2}.$$

So, $f'(t) = 0$ if and only if $t = 1$, with $f''(1) = -(\lambda - 1)\lambda > 0$. In that case, f has a global minimum at $t = 1$. Therefore, for all $t \in (0, \infty)$, $f(t) \ge f(1)$. Setting $t = \frac{a}{b}$, we find that

$$\lambda \frac{a}{b} + 1 - \lambda - \frac{a^\lambda}{b^\lambda} \ge 0 \implies a^\lambda b^{1-\lambda} \le \lambda a + (1-\lambda) b.$$

□

Proposition 3.1.3 (Holder's Inequality). *Let $p \in (1, \infty)$ and $q \in (1, \infty)$ such that*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Let $f, g : [0, 1] \to \mathbb{R} \cup \{\pm\infty\}$ be measurable functions. Then,

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

In particular, if $f \in L^p[0, 1]$ and $g \in L^q[0, 1]$, then $fg \in L^1[0, 1]$. Moreover, we have equality if and only if $\alpha|f|^p = \beta|g|^q$ almost everywhere in $[0, 1]$, for some $\alpha, \beta \in \mathbb{R}^\times$.

Proof. If $\|f\|_p = 0$ or $\|g\|_q = 0$, then the function is equal to 0 almost everywhere in $[0, 1]$. In that case, $fg = 0$ almost everywhere, meaning that $\|fg\|_1 = 0$.
Instead, if $\|f\|_p = \infty$ or $\|g\|_q = \infty$, then we must have $\|fg\|_1 \leq \|f\|_p \|g\|_q$.
Otherwise, assume that $\|f\|_p$ and $\|g\|_p$ are finite are non-zero. Without loss of generality, assume that $\|f\|_p = 1$ and $\|g\|_q = 1$.[1] In that case, set $a = |f(x)|^p$, $b = |g(x)|^q$ and $\lambda = \frac{1}{p}$. Then, the lemma above tells us that

$$\begin{aligned} f(x)g(x) &= (|f(x)|^p)^{1/p}(|g(x)|^q)^{1/q} \\ &= a^\lambda b^{1-\lambda} \\ &\leq \lambda a + (1 - \lambda)b \\ &= \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q. \end{aligned}$$

Integrating the inequality, we find that

$$\begin{aligned} \|fg\|_1 &= \int_0^1 |fg| \; dm \\ &\leq \int_0^1 \frac{1}{p}|f|^p + \frac{1}{q}|g|^q \; dm \\ &= \frac{1}{p}\|f\|_p^p + \frac{1}{q}\|g\|_q^q \\ &= \frac{1}{p} + \frac{1}{q} \\ &= 1 = \|f\|_p \|g\|_q. \end{aligned}$$

□

Now, we show that L^p are truly norms, by proving it satisfies the Triangle Inequality.

Proposition 3.1.4 (Minkowski Inequality). *Let $p \in [1, \infty)$ and $f, g \in L^p$. Then,*

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

Proof. If $p = 1$, then the result follows from the Triangle Inequality in $\mathbb{R}$. Otherwise, we have $p \in (1, \infty)$. In that case,

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

[1] This is possible since a (non-negative) scalar multiple does not affect the inequality.

using the Triangle Inequality on $\mathbb{R}$. Let $q \in (1, \infty)$ such that

$$\frac{1}{p} + \frac{1}{q} = 1 \iff (p-1)q = p.$$

Integrating the inequality, we find that

$$\int |f+g|^p \ dm \leq \int |f||f+g|^{p-1} \ dm + \int |g||f+g|^{p-1} \ dm.$$

Using Holder's Inequality now, we find that

$$\int |f||f+g|^{p-1} \ dm = \||f| \cdot |f+g|^{p-1}\|_1 \leq \|f\|_p \|(f+g)^{p-1}\|_q$$
$$\int |g||f+g|^{p-1} \ dm = \||g| \cdot |f+g|^{p-1}\|_1 \leq \|g\|_p \|(f+g)^{p-1}\|_q.$$

We have

$$\|(f+g)^{p-1}\|_q = \left(\int (|f+g|^{p-1})^q \ dm\right)^{1/q} = \left(\int |f+g|^p \ dm\right)^{1/q}.$$

Putting this together, we find that

$$\int |f+g|^p \ dm \leq (\|f\|_p + \|g\|_p) \left(\int |f+g|^p \ dm\right)^{1/q}.$$

This implies that

$$\left(\int |f+g|^p \ dm\right)^{1-1/q} \leq \|f\|_p + \|g\|_p.$$

We have $1 - \frac{1}{q} = \frac{1}{p}$. So,

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

□

3.2 Completeness of L^p

Definition 3.2.1. Let V be a normed vector space over $\mathbb{R}$, and let $(x_n)_{n=1}^{\infty}$ be a sequence in V. We say that the series

$$\sum_{k=1}^{\infty} x_k$$

converges if the sequence of partial sums $(s_n)_{n=1}^{\infty}$ in V given by

$$s_n = \sum_{k=1}^{n} x_k$$

converges. Further, we say that the series converges absolutely if the sequence of partial sums $(y_n)_{n=1}^{\infty}$ in $\mathbb{R}$ given by

$$y_n = \sum_{k=1}^{n} \|x_k\|$$

converges.

Proposition 3.2.2. *Let V be a normed vector space over $\mathbb{R}$. Then, V is complete if and only if every absolutely convergent series is convergent.*

Proof.

- First, assume that V is complete. Let $(x_k)_{k=1}^{\infty}$ be a sequence such that the series $\sum x_k$ is absolutely convergent. In that case, the series $\sum \|x_k\|$ is Cauchy. Now, define the partial sums $(s_n)_{n=1}^{\infty}$ and $(t_n)_{n=1}^{\infty}$ by

$$s_n = \sum_{k=1}^{n} x_k, \qquad t_n = \sum_{k=1}^{n} \|x_k\|.$$

  Next, let $\varepsilon > 0$. We can find an $N \in \mathbb{Z}_{\geq 1}$ such that for $m, n \in \mathbb{Z}_{\geq 1}$ with $m \geq n$, if $m, n \geq N$, then

$$|t_m - t_n| = \sum_{k=n+1}^{m} \|x_k\| < \varepsilon.$$

  In that case, for $m, n \in \mathbb{Z}_{\geq 1}$ with $m \geq n$, if $m, n \geq N$, then

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

  Therefore, the series $\sum x_k$ is Cauchy. Since V is complete, this implies that $\sum x_k$ is convergent. So, every absolutely convergent series is convergent.

- Now, assume that every absolutely convergent series is convergent. Let $(x_n)_{n=1}^{\infty}$ be a Cauchy sequence in V. In that case, for each $k \in \mathbb{Z}_{\geq 1}$, we can find an $N_k \in \mathbb{Z}_{\geq 1}$, with $N_k \geq N_{k-1}$ for $k \geq 2$, such that for

$m, n \in \mathbb{Z}_{\geq 1}$, if $m, n \geq N$, then $\|x_m - x_n\| < \frac{1}{2^{k-1}}$. Now, define the sequence $(y_n)_{n=1}^\infty$ by $y_1 = x_{N_1}$ and $y_n = x_{N_n} - x_{N_{n-1}}$ for $n \geq 2$. We show that $\sum y_n$ is absolutely convergent. Let $\varepsilon > 0$. We find that

$$\sum_{j=1}^\infty \|y_j\| \leq \|x_{N_1}\| + \sum_{j=2}^\infty \|x_{N_j} - x_{N_{j-1}}\| \|x_{N_1}\| + \sum_{j=2}^\infty \frac{1}{2^j} = x_{N_1} + 1.$$

This implies that $\sum y_n$ is absolutely convergent. Since every absolutely convergent series is convergent, we find that $\sum y_n$ is convergent. In that case, the sequence of partial sums $(s_n)_{n=1}^\infty$

$$s_n = \sum_{k=1}^n y_n = x_{N_n}$$

converges. So, the subsequence $(x_{N_n})_{n=1}^\infty$ converges. Since (x_n) is Cauchy, this implies that $(x_n)_{n=1}^\infty$ converges. Therefore, V is complete.

□

Proposition 3.2.3. *Let $p \in [1, \infty)$. Then, $L^p[0,1]$ is a Banach space, i.e. the function space is complete.*

Proof. Let $(f_n)_{n=1}^\infty$ be a sequence in $L^p[0,1]$ such that $\sum f_n$ is absolutely convergent. □

3.3 Linear Operators

Definition 3.3.1. Let V, W be normed vector spaces, and let $T : V \to W$ be a function. We say that T is a *linear operator* if

- $T\lambda \boldsymbol{v} = \lambda T(\boldsymbol{v})$ for all $\lambda \in \mathbb{R}$ and $\boldsymbol{v} \in V$;
- $T(\boldsymbol{v} + \boldsymbol{w}) = T(\boldsymbol{v}) + t(\boldsymbol{w})$ for all $\boldsymbol{v}, \boldsymbol{w} \in V$.

Definition 3.3.2. Let V, W be normed vector spaces, and let $T : V \to W$ be a linear operator. Then, T is *bounded* if there exists a $c > 0$ such that for all $\boldsymbol{v} \in V$, $\|T(\boldsymbol{v})\|_W \leq c\|\boldsymbol{v}\|_V$.

Dimensional and linear operators

If V and W are both finite dimensional, then T can be represented by a matrix. Let the basis of V be

$$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\},$$

and the basis for W be

$$\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_m\}.$$

Then, for all $i \in \{1, 2, \dots, n\}$,

$$T(\boldsymbol{v}_i) = a_{1i}\boldsymbol{w}_1 + a_{2i}\boldsymbol{w}_2 + \dots + a_{mi}\boldsymbol{w}_m.$$

Now, define

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Let $\boldsymbol{x} \in V$. We have

$$\boldsymbol{x} = b_1\boldsymbol{v}_1 + b_2\boldsymbol{v}_2 + \dots b_n\boldsymbol{v}_n.$$

Then,

$$\begin{aligned} T(\boldsymbol{x}) &= b_1(a_{11}\boldsymbol{w}_1 + a_{21}\boldsymbol{w}_2 + \dots + a_{m1}\boldsymbol{w}_m) \\ &\quad + b_2(a_{12}\boldsymbol{w}_1 + a_{22}\boldsymbol{w}_2 + \dots + a_{m2}\boldsymbol{w}_m) + \dots \\ &\quad + b_n(a_{1n}\boldsymbol{w}_1 + a_{2n}\boldsymbol{w}_2 + \dots + a_{mn}\boldsymbol{w}_m) \\ &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\ &= A\boldsymbol{x}. \end{aligned}$$

Moreover, linear operators in finite dimensions are always bounded. To see this, define the set

$$S^n = \{\boldsymbol{v} \in V \mid \|\boldsymbol{v}\|_V = 1\}.$$

Since V is finite-dimensional, Heine-Borel theorem tells us that the set S^n is compact. Moreover, the function $f : S^n \to \mathbb{R}$ given by $f(\boldsymbol{v}) = \|T(\boldsymbol{v})\|_V$

is continuous. So, by the Extreme Value Theorem, the function attains a maximum. That is, for all $\boldsymbol{v} \in V$, if $\|\boldsymbol{v}\|_V = 1$, then $\|T(\boldsymbol{v})\|_W \leq K$, for some $K > 0$. Now, let $\boldsymbol{v} \in V$. If $\boldsymbol{v} = 0$, then we have

$$\|\boldsymbol{v}\|_V = 0 \leq 0 = K\|T(\boldsymbol{v})\|_W.$$

Otherwise, if $\boldsymbol{v} \neq \mathbf{0}$, then we have

$$\left\| \frac{1}{\|\boldsymbol{v}\|_V} \boldsymbol{v} \right\|_V = \frac{1}{\|\boldsymbol{v}\|_V} \|\boldsymbol{v}\|_V = 1.$$

So, we have $\frac{1}{\|\boldsymbol{v}\|_V} \boldsymbol{v} \in S^n$. In that case,

$$\left\| T\left(\frac{1}{\|\boldsymbol{v}\|_V} \boldsymbol{v} \right) \right\|_W \leq K.$$

This implies that

$$\|T(\boldsymbol{v})\|_W \leq K\|\boldsymbol{v}\|_V.$$

So, for all $\boldsymbol{v} \in V$,

$$\|T(\boldsymbol{v})\|_W \leq K\|\boldsymbol{v}\|_V.$$

Therefore, the linear operator is bounded.

In infinite dimensions, linear operators need not be bounded. To see this, define the function $T : \ell^\infty \to \ell^\infty$ by

$$T(x_1, x_2, x_3, \dots) = (x_1, 2x_2, 3x_3, \dots).$$

This is a linear function. But, it is not bounded- for all $n \in \mathbb{Z}_{\geq 1}$, let

$$e_n^{(k)} = \begin{cases} 1 & n = k \\ 0 & n \neq k \end{cases}.$$

We have $\|e^{(k)}\|_\infty = 1$, but $\|T(e^{(k)})\|_\infty = n$. So, for all $n \in \mathbb{Z}_{\geq 1}$,

$$\|T(e^{(k)})\|_\infty \geq n\|e^{(k)}\|_\infty.$$

Therefore, the operator cannot be bounded.

Bounded linear operator space

Proposition 3.3.3. *Let V and W be normed vector spaces, and let $T : V \to W$ be a linear operator. Then, the following are equivalent:*

1. *T is continuous;*
2. *T is continuous at $\mathbf{0}$;*
3. *T is bounded.*

Proof. We show (1) $\implies$ (2) $\implies$ (3) $\implies$ (1).

- We know that (1) $\implies$ (2).

- We show that (2) $\implies$ (3). Since T is continuous at $\mathbf{0}$, there exists a $\delta > 0$ such that for $\boldsymbol{x} \in V$, if $\|\boldsymbol{x}\|_V < \delta$, then $\|T(\boldsymbol{x})\|_W < 1$. Now, let $\boldsymbol{x} \in V$. If $\boldsymbol{x} = \mathbf{0}$, then
$$\|T(\boldsymbol{x})\|_W = 0 \leq 0 = \frac{2}{\delta}\|\boldsymbol{x}\|_V.$$
Now, if $\boldsymbol{x} \neq \mathbf{0}$, then we have
$$\left\| \frac{\delta}{2\|\boldsymbol{x}\|_V}\boldsymbol{x} \right\|_V = \frac{\delta}{2\|\boldsymbol{x}\|_V}\|\boldsymbol{x}\|_V = \frac{\delta}{2} < \delta.$$
So,
$$\left\| T\left(\frac{\delta}{2\|\boldsymbol{x}\|_V}\boldsymbol{x} \right) \right\|_W < 1.$$
We have
$$\left\| T\left(\frac{\delta}{2\|\boldsymbol{x}\|_V}\boldsymbol{x} \right) \right\|_W = \left\| \frac{\delta}{2\|\boldsymbol{x}\|_V} T(\boldsymbol{x}) \right\|_W = \frac{\delta}{2\|\boldsymbol{x}\|_V} \|T(\boldsymbol{x})\|_W .$$
This implies that
$$\|T(\boldsymbol{x})\|_W < \frac{2}{\delta}\|\boldsymbol{x}\|_V.$$
So, for all $\boldsymbol{x} \in V$,
$$\|T(\boldsymbol{x})\|_W \leq \frac{2}{\delta}\|\boldsymbol{x}\|_V.$$
In that case, T is bounded.

- We show that (3) $\implies$ (1). Since T is bounded, there exists a $c > 0$ such that for all $\boldsymbol{v} \in V$, $\|T(\boldsymbol{v})\|_W \leq c\|\boldsymbol{v}\|_V$. Let $\boldsymbol{v} \in V$. Let $\boldsymbol{u} \in V$ and $\varepsilon > 0$. Set $\delta = \frac{\varepsilon}{c}$. In that case, for $\boldsymbol{v} \in V$, if $\|\boldsymbol{u} - \boldsymbol{v}\|_V < \delta$, then
$$\begin{aligned} \|T(\boldsymbol{u}) - T(\boldsymbol{v})\|_W &= \|T(\boldsymbol{u} - \boldsymbol{v})\|_W \\ &\leq c\|\boldsymbol{u} - \boldsymbol{v}\|_V \\ &< c \cdot \delta = \varepsilon. \end{aligned}$$
In that case, T is continuous.

So, the statements are equivalent. □

Using properties of continuity, this implies that:

- if $T : V \to W$ is linear and $\lambda \in \mathbb{R}$, then λT is bounded; and
- if $T, U : V \to W$ are linear, then $T + U$ is bounded.

So, the set of bounded linear operators from T to V is a vector space. It is denoted by $L(V, W)$. Moreover, we can define a norm on the vector space.

Definition 3.3.4. Let V and W be normed vector spaces, and let $T : V \to W$ be a bounded linear operator. We define the *operator norm* of T to be:
$$\|T\| = \sup_{\substack{\boldsymbol{v} \in V \\ \|\boldsymbol{v}\|_V = 1}} \|T(\boldsymbol{v})\|_W.$$

The value $\|T\|$ is the smallest bound $c \geq 0$ that satisfies $\|T(\boldsymbol{x})\|_W \leq c\|\boldsymbol{x}\|_V$.

We will prove the inequality relating $\|T\|$ and $\|T(\boldsymbol{v})\|_W$.

Proposition 3.3.5. *Let V and W be normed vector spaces, and let $T : V \to W$ be a bounded linear operator. Then,*

$$\|T(\boldsymbol{x})\|_W \leq \|T\|\|\boldsymbol{x}\|_V.$$

Proof. Let $\boldsymbol{v} \in V$. If $\boldsymbol{v} = \mathbf{0}$, then

$$\|T(\boldsymbol{v})\|_W = 0 \leq 0 = \|T\|\|\boldsymbol{v}\|_V.$$

Otherwise, we have $\boldsymbol{v} \neq \mathbf{0}$. We know that

$$\left\| \frac{1}{\|\boldsymbol{v}\|_V}\boldsymbol{v} \right\|_V = \frac{1}{\|\boldsymbol{v}\|_V}\|\boldsymbol{v}\|_V = 1.$$

So,

$$\left\| T\left(\frac{1}{\|\boldsymbol{v}\|_V}\boldsymbol{v} \right) \right\|_W \leq \|T\|.$$

This implies that $\|T(\boldsymbol{v})\|_W \leq \|T\|\|\boldsymbol{v}\|_V$. □

Using this result, we characterise the operator norm as the supremum of all non-zero values.

Proposition 3.3.6. *Let V and W be normed vector spaces, and let $T : V \to W$ be a bounded linear operator. Then,*

$$\|T\| = \sup_{\substack{\boldsymbol{v} \in V \\ \boldsymbol{v} \neq \mathbf{0}}} \frac{\|T(\boldsymbol{v})\|_W}{\|\boldsymbol{v}\|_V}.$$

Proof. Let $\boldsymbol{v} \in V$ with $\boldsymbol{v} \neq \mathbf{0}$. We know that $\|T(\boldsymbol{v})\|_W \leq \|T\|\|\boldsymbol{v}\|_V$. So,

$$\|T\| \geq \sup_{\substack{\boldsymbol{v} \in V \\ \boldsymbol{v} \neq \mathbf{0}}} \frac{\|T(\boldsymbol{v})\|_W}{\|\boldsymbol{v}\|_V}.$$

Now, let $\varepsilon > 0$. Since

$$\|T\| = \sup_{\substack{\boldsymbol{v} \in V \\ \|\boldsymbol{v}\|_V = 1}} \|T(\boldsymbol{v})\|_W,$$

we can find a $\boldsymbol{v} \in V$ with $\|\boldsymbol{v}\|_V = 1$ such that $\|T(\boldsymbol{v})\|_W > \|T\| + \varepsilon$. In that case, $\frac{\|T(\boldsymbol{v})\|_W}{\|\boldsymbol{v}\|_V} > \|T\| + \varepsilon$. So, we have

$$\|T\| = \sup_{\substack{\boldsymbol{v} \in V \\ \boldsymbol{v} \neq \mathbf{0}}} \frac{\|T(\boldsymbol{v})\|_W}{\|\boldsymbol{v}\|_V}.$$

□

Next, we prove that the value $\|T\|$ is the smallest bound $c \geq 0$ that satisfies $\|T(\boldsymbol{x})\|_W \leq c\|\boldsymbol{x}\|_V$.

Proposition 3.3.7. *Let V and W be normed vector spaces, and let $T : V \to W$ be a bounded linear operator. Define*

$$S = \{c \geq 0 \mid \|T(\boldsymbol{v})\|_W \leq c\|\boldsymbol{v}\|_V \ \forall \boldsymbol{v} \in V\}.$$

Then,

$$\|T\| = \inf(S).$$

Proof. We know that for all $\boldsymbol{v} \in V$, $\|T(\boldsymbol{v})\| \leq \|T\|\|\boldsymbol{v}\|_V$. So, $\|T\| \in S$. Now, let $c \in S$. We know that for all $\boldsymbol{v} \in V$, $c\|T(\boldsymbol{v})\|_W \leq c\|\boldsymbol{v}\|_V$. In that case, for all $\boldsymbol{v} \in V$ with $\|\boldsymbol{v}\|_V$, we find that $\|T(\boldsymbol{v})\|_W \leq c$. Since

$$\|T\| = \sup_{\substack{\boldsymbol{v} \in V \\ \|\boldsymbol{v}\|_V = 1}} \|T(\boldsymbol{v})\|_W,$$

we find that $c \geq \|T\|$. Since $\|T\| \in S$, we must have $\|T\| = \inf(S)$. □

Now, we show that $L(V, W)$ is complete if W is complete.

Proposition 3.3.8. *Let V and W be normed vector spaces, and let W be complete. Then, $L(V, W)$ is complete.*

Proof. Let $(T_n)_{n=1}^\infty$ be a Cauchy sequence in $L(V, W)$. For all $\boldsymbol{v} \in V$, we have the sequence $(T_n(\boldsymbol{v}))_{n=1}^\infty$ in W. We show that $(T_n(\boldsymbol{v}))$ is Cauchy. If $\boldsymbol{v} = \mathbf{0}$, then we have $T_n(\boldsymbol{v}) = \mathbf{0}$ for all $n \in \mathbb{Z}_{\geq 1}$. So, the sequence $(T_n(\boldsymbol{v}))$ is Cauchy. Otherwise, we have $\boldsymbol{v} \neq \mathbf{0}$. Then, let $\varepsilon > 0$. Since (T_n) is Cauchy, there exists an $N \in \mathbb{Z}_{\geq 1}$ such that for $m, n \in \mathbb{Z}_{\geq 1}$ such that if $m, n \geq N$, then $\|T_m - T_n\| < \frac{\varepsilon}{\|\boldsymbol{v}\|_V}$. In that case, for $m, n \in \mathbb{Z}_{\geq 1}$, such that if $m, n \geq N$, then

$$\begin{aligned}
\|T_m(\boldsymbol{v}) - T_n(\boldsymbol{v})\|_W &= \|(T_m - T_n)(\boldsymbol{v})\|_W \\
&\leq \|T_m - T_n\|\|\boldsymbol{v}\|_V \\
&< \frac{\varepsilon}{\|\boldsymbol{v}\|_V}\|\boldsymbol{v}\|_V = \varepsilon.
\end{aligned}$$

This implies that $(T_n(\boldsymbol{v}))$ is Cauchy. Since W is complete, for each $\boldsymbol{v} \in V$, we can find a $\boldsymbol{w}_{\boldsymbol{v}}$ such that $T_n(\boldsymbol{v}) \to \boldsymbol{w}_{\boldsymbol{v}}$. Now, define the function $T : V \to W$ by $T(\boldsymbol{v}) = \boldsymbol{w}_{\boldsymbol{v}}$.

First, we claim that T is linear. For all $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ and $n \in \mathbb{Z}_{\geq 1}$,

$$T_n(\boldsymbol{v}_1 + \boldsymbol{v}_2) = T_n(\boldsymbol{v}_1) + T_n(\boldsymbol{v}_2).$$

We have $T_n(\boldsymbol{v}_1 + \boldsymbol{v}_2) \to T(\boldsymbol{v}_1 + \boldsymbol{v}_2)$, $T_n(\boldsymbol{v}_1) \to T(\boldsymbol{v}_1)$ and $T_n(\boldsymbol{v}_{\mathbf{2}}) \to T(\boldsymbol{v}_2)$. Since limits are unique, this implies that $T(\boldsymbol{v}_1 + \boldsymbol{v}_2) = T(\boldsymbol{v}_1) + T(\boldsymbol{v}_2)$. Now, for $\boldsymbol{v} \in V$, $\lambda \in \mathbb{R}$ and $n \in \mathbb{Z}_{\geq 1}$,

$$T_n(\lambda\boldsymbol{v}) = \lambda T_n(\boldsymbol{v}).$$

Since $T_n(\lambda\boldsymbol{v}) \to T(\lambda\boldsymbol{v})$ and $T_n(\boldsymbol{v}) \to T(\boldsymbol{v})$, we find that $T(\lambda\boldsymbol{v}) = \lambda T(\boldsymbol{v})$. This implies that T is linear.

Next, we claim that T is bounded. Since (T_n) is Cauchy, there exists an $N \in \mathbb{Z}_{\geq 1}$ such that for all $m, n \in \mathbb{Z}_{\geq 1}$, if $m, n \geq N$, then $\|T_m - T_n\| < 1$. In particular, for all $m, n \in \mathbb{Z}_{\geq 1}$, if $m, n \geq N$, then $\|T_m(\boldsymbol{v}) - T_n(\boldsymbol{v})\|_W \leq \|\boldsymbol{v}\|_V$

for all $\boldsymbol{v} \in V$. We know that T_N is bounded. So, there exists a $c > 0$ such that for all $\boldsymbol{v} \in V$, $\|T_n(\boldsymbol{v})\|_W \leq c\|\boldsymbol{v}\|_V$. Now, let $\boldsymbol{v} \in V$. Since $T_n(\boldsymbol{v}) \to T(\boldsymbol{v})$, there exists an $N' \in \mathbb{Z}_{\geq 1}$ such that for all $n \in \mathbb{Z}_{\geq 1}$, if $n \geq N'$, then $\|T_n(\boldsymbol{v}) - T(\boldsymbol{v})\|_W \leq \|\boldsymbol{v}\|_V$. Now, fix $n = \max(N, N')$. In that case,

$$\begin{aligned}\|T(\boldsymbol{v})\|_V &\leq \|T(\boldsymbol{v}) - T_n(\boldsymbol{v})\|_W + \|T_n(\boldsymbol{v}) - T_N(\boldsymbol{v})\|_W + \|T_N(\boldsymbol{v})\|_W \\ &\leq (2 + c)\|\boldsymbol{v}\|_V.\end{aligned}$$

This implies that T is bounded.

Finally, we show that $T_n \to T$. Let $\varepsilon > 0$. Since (T_n) is Cauchy, we can find an $N \in \mathbb{Z}_{\geq 1}$ such that for all $m, n \in \mathbb{Z}_{\geq 1}$, if $m, n \geq N$, then $\|T_m - T_n\| < \frac{\varepsilon}{3}$. In particular, for all $m, n \in \mathbb{Z}_{\geq 1}$, if $m, n \geq N$, then $\|T_m(\boldsymbol{v}) - T_n(\boldsymbol{v})\|_W \leq \frac{\varepsilon}{3}\|\boldsymbol{v}\|_V$ for all $\boldsymbol{v} \in V$. Now, let $n \in \mathbb{Z}_{\geq 1}$ such that $n \geq N$. Moreover, let $\boldsymbol{v} \in V$ with $\|\boldsymbol{v}\|_V = 1$. Since $T_n(\boldsymbol{v}) \to T(\boldsymbol{v})$, we can find an $N' \in \mathbb{Z}_{\geq 1}$ such that for $m \in \mathbb{Z}_{\geq 1}$, if $m \geq N'$, then $\|T_m(\boldsymbol{v}) - T(\boldsymbol{v})\|_W < \frac{\varepsilon}{3}$. Now, fix $m = \max(N, N')$. In that case,

$$\begin{aligned}\|T_n(\boldsymbol{v}) - T(\boldsymbol{v})\|_W &\leq \|T_n(\boldsymbol{v}) - T_m(\boldsymbol{v})|_W + \|T_m(\boldsymbol{v}) - T(\boldsymbol{v})\|_W \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2}{3}\varepsilon.\end{aligned}$$

That is, for $n \in \mathbb{Z}_{\geq 1}$, if $n \geq N$, then for all $\boldsymbol{v} \in V$ with $\|\boldsymbol{v}\|_V = 1$,

$$\|T_n(\boldsymbol{v}) - T(\boldsymbol{v})\|_W < \frac{2}{3}\varepsilon.$$

This implies that for $n \in \mathbb{Z}_{\geq 1}$, if $n \geq N$, then

$$\|T_n - T\| = \sup_{\substack{\boldsymbol{v} \in V \\ \|\boldsymbol{v}\|_V = 1}} \|T_n(\boldsymbol{v}) - T(\boldsymbol{v})\|_W \leq \frac{2}{3}\varepsilon < \varepsilon.$$

So, $T_n \to T$. In that case, $L(V, W)$ is complete. $\square$

3.4 Banach theory

Definition 3.4.1. Let V be a normed vector space. Then, a bounded linear map $T : V \to \mathbb{R}$ is a *linear functional*.

Definition 3.4.2. Let V be a normed vector space, and let $p : V \to \mathbb{R}$ be a function. Then, p is a *sub-linear functional* if:

- $p(\lambda \boldsymbol{v}) = \lambda p(\boldsymbol{v})$ for all $\boldsymbol{v} \in V$, $\lambda \in \mathbb{R}$, and
- $p(\boldsymbol{u} + \boldsymbol{v}) \leq p(\boldsymbol{u}) + p(\boldsymbol{v})$ for all $\boldsymbol{u}, \boldsymbol{v} \in V$.

Theorem 3.4.3 (The Hahn-Banach Theorem)**.** *Let V be a normed vector space and $p : V \to \mathbb{R}$ be a sub-linear functional. Let H be a subspace of V, and $f : H \to \mathbb{R}$ be a linear such that $f(\boldsymbol{x}) \leq p(\boldsymbol{x})$ for all $\boldsymbol{x} \in H$. Then, there exists a linear functional $F : V \to \mathbb{R}$ such that $F(\boldsymbol{x}) = f(\boldsymbol{x})$ for all $\boldsymbol{x} \in H$ and $F(\boldsymbol{x}) \leq p(\boldsymbol{x})$ for all $\boldsymbol{x} \in V$.*

Proof. Let $\boldsymbol{y} \in V \setminus H$. We show that there exists a linear functional $G : S \to \mathbb{R}$ such that $G(\boldsymbol{x}) = f(\boldsymbol{x})$ for all $\boldsymbol{x} \in H$ and $G(\boldsymbol{x}) \leq p(\boldsymbol{x})$ for all $\boldsymbol{x} \in S$, where

$$S = \{\boldsymbol{x} + \lambda \boldsymbol{y} \mid \boldsymbol{x} \in H, \lambda \in \mathbb{R}\}.$$

Now, let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in H$. We find that

$$\begin{aligned} f(\boldsymbol{x}_1) + f(\boldsymbol{x}_2) &= f(\boldsymbol{x}_1 + \boldsymbol{x}_2) \\ &\leq p(\boldsymbol{x}_1 + \boldsymbol{x}_2) \\ &\leq p(\boldsymbol{x}_1 - \boldsymbol{y}) + p(\boldsymbol{x}_2 + \boldsymbol{y}). \end{aligned}$$

So, we find that for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in H$,

$$f(\boldsymbol{x}_1) - p(\boldsymbol{x}_1 - \boldsymbol{y}) \leq p(\boldsymbol{x}_2 + \boldsymbol{y}) - f(\boldsymbol{x}_2).$$

This implies that

$$\sup_{\boldsymbol{x} \in H} f(\boldsymbol{x}) - p(\boldsymbol{x} - \boldsymbol{y}) \leq \inf_{\boldsymbol{x} \in H} p(\boldsymbol{x} + \boldsymbol{y}) - f(\boldsymbol{x}).$$

Next, let $\alpha \in \mathbb{R}$ such that

$$\sup_{\boldsymbol{x} \in H} f(\boldsymbol{x}) - p(\boldsymbol{x} - \boldsymbol{y}) \leq \alpha \leq \inf_{\boldsymbol{x} \in H} p(\boldsymbol{x} + \boldsymbol{y}) - f(\boldsymbol{x}).$$

Define the function $G : S \to \mathbb{R}$ by $G(\boldsymbol{x} + \lambda \boldsymbol{y}) = f(\boldsymbol{x}) + \alpha\lambda$. Then, G is linear, with $G(\boldsymbol{x}) = f(\boldsymbol{x})$ for all $\boldsymbol{x} \in H$. Now, we show that $G(\boldsymbol{x}) \leq p(\boldsymbol{x})$ for all $\boldsymbol{x} \in S$. So, let $\boldsymbol{x} + \lambda \boldsymbol{y} \in S$. If $\lambda = 0$, then we have

$$G(\boldsymbol{x} + \lambda \boldsymbol{y}) = f(\boldsymbol{x}) \leq p(\boldsymbol{x}) = p(\boldsymbol{x} + \lambda \boldsymbol{y}).$$

Next, assume that $\lambda > 0$. Then,

$$\begin{aligned} G(\boldsymbol{x} + \lambda \boldsymbol{y}) &= f(\boldsymbol{x}) + \lambda\alpha \\ &\leq f(\boldsymbol{x}) + \lambda(p(\boldsymbol{x}/\lambda + \boldsymbol{y}) - f(\boldsymbol{x}/\lambda)) \\ &= f(\boldsymbol{x}) + p(\boldsymbol{x} + \lambda \boldsymbol{y}) - f(\boldsymbol{x}) \\ &= p(\boldsymbol{x} + \lambda \boldsymbol{y}). \end{aligned}$$

Otherwise, we have $\lambda < 0$. Then,

$$\begin{aligned} G(\boldsymbol{x} + \lambda \boldsymbol{y}) &= f(\boldsymbol{x}) + \lambda\alpha \\ &\leq f(\boldsymbol{x}) + \lambda(f(-\boldsymbol{x}/\lambda) - p(-\boldsymbol{x}/\lambda - y)) \\ &= f(\boldsymbol{x}) - f(\boldsymbol{x}) + p(\boldsymbol{x} + \lambda \boldsymbol{y}) \\ &= p(\boldsymbol{x} + \lambda \boldsymbol{y}). \end{aligned}$$

So, for all $\boldsymbol{x} + \lambda \boldsymbol{y} \in S$, $G(\boldsymbol{x} + \lambda \boldsymbol{y}) \leq p(\boldsymbol{x} + \lambda \boldsymbol{y})$. Now, we can keep extending the function G from S up to V, and get the required result. □

Corollary 3.4.4. *Let V be a normed vector space, $H \subseteq V$ be a closed subspace, $\boldsymbol{y} \in V \setminus H$, and let*

$$\delta = \inf_{\boldsymbol{x} \in H} \|\boldsymbol{y} - \boldsymbol{x}\|.$$

Then, we can find a non-trivial linear functional $F : V \to \mathbb{R}$ such that $H \subseteq \ker F$, $\|F\| = 1$ and $F(\boldsymbol{y}) = \delta$.

Proof. Let

$$S = \{\boldsymbol{x} + \lambda \boldsymbol{y} \mid \boldsymbol{y} \in H, \lambda \in \mathbb{R}\}.$$

Define the function $f : S \to \mathbb{R}$ by $f(\boldsymbol{x} + \lambda \boldsymbol{y}) = \lambda\delta$. By definition, it is a non-trivial linear functional. We find that for all $\boldsymbol{x} + \lambda \boldsymbol{y} \in S$,

$$|f(\boldsymbol{x} + \lambda \boldsymbol{y})| = |\lambda\delta| \leq |\lambda| \cdot \left\| \boldsymbol{y} - \left(-\frac{\boldsymbol{x}}{\lambda}\right) \right\| = \|\boldsymbol{x} + \lambda \boldsymbol{y}\|.$$

So, the Hahn-Banach theorem tells us that there exists an extension of f, given by $F : V \to \mathbb{R}$.
We know that for all $\boldsymbol{x} \in H \subseteq S$,

$$F(\boldsymbol{x} + \lambda \boldsymbol{y}) = f(\boldsymbol{x}) = \boldsymbol{0}.$$

So, $H \subseteq \ker F$. Moreover, since $\boldsymbol{y} \in S$, we find that

$$F(\boldsymbol{y}) = f(\boldsymbol{y}) = f(\boldsymbol{0} + 1 \cdot \boldsymbol{y}) = \delta.$$

By the Hahn-Banach theorem, we know that for all $\boldsymbol{x} \in V$, $\|F(\boldsymbol{x})\| \leq \|\boldsymbol{x}\|$. In that case, $\|F\| \leq 1$. □

Now, let V be a normed vector space and let $\boldsymbol{x} \in V$ with $\boldsymbol{x} \neq \boldsymbol{0}$. Then, $\{\boldsymbol{0}\} \subseteq V$ is a closed subspace with $\boldsymbol{x} \notin \{\boldsymbol{0}\}$. So, the result above tells us that there exists a non-trivial linear functional $f : V \to \mathbb{R}$ such that $\|f\| = 1$ and $\|f(\boldsymbol{x})\| = \|\boldsymbol{x}\|$.

Corollary 3.4.5. *Let V be a normed vector space, and let $\boldsymbol{x}, \boldsymbol{y} \in V$ with $\boldsymbol{x} \neq \boldsymbol{y}$. Then, there exists a linear functional $f : V \to \mathbb{R}$ such that $f(\boldsymbol{x}) \neq f(\boldsymbol{y})$.*

Proof. Since $\boldsymbol{x} \neq \boldsymbol{y}$, the result above tells us that there exists a linear functional $f : V \to \mathbb{R}$ such that

$$\|f(\boldsymbol{x} - \boldsymbol{y})\| = \|\boldsymbol{x} - \boldsymbol{y}\| > 0.$$

This implies that $f(\boldsymbol{x} - \boldsymbol{y}) \neq \boldsymbol{0}$, and so $f(\boldsymbol{x}) \neq f(\boldsymbol{y})$. □

Corollary 3.4.6. *Let V be a normed vector space. For each $\boldsymbol{v} \in V$, define the map $ev_{\boldsymbol{v}} : V^* \to \mathbb{R}$ by $ev_{\boldsymbol{v}}(f) = f(\boldsymbol{v})$. Then, the map $\iota : V \to V^{**}$ given by $\iota(\boldsymbol{v}) = ev_{\boldsymbol{v}}$ is a linear isometry.*

Proof. We find that for $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ and $f \in V^*$,

$$\begin{aligned}\iota(\boldsymbol{v}_1 + \boldsymbol{v}_2)(f) &= ev_{\boldsymbol{v_1}+\boldsymbol{v_2}}(f) \\ &= f(\boldsymbol{v}_1 + \boldsymbol{v}_2) \\ &= f(\boldsymbol{v}_1) + f(\boldsymbol{v}_2) \\ &= ev_{\boldsymbol{v}_1}(f) + ev_{\boldsymbol{v}_2}(f) \\ &= \iota(\boldsymbol{v}_1)(f) + \iota(\boldsymbol{v}_2)(f).\end{aligned}$$

This implies that $\iota(\boldsymbol{v}_1 + \boldsymbol{v}_2) = \iota(\boldsymbol{v}_1) + \iota(\boldsymbol{v}_2)$. Now, for $\boldsymbol{v} \in V$ and $f \in V^*$,

$$\begin{aligned}\iota(\lambda\boldsymbol{v})(f) &= ev_{\lambda\boldsymbol{v}}(f) \\ &= f(\lambda\boldsymbol{v}) \\ &= \lambda f(\boldsymbol{v}) \\ &= \lambda ev_{\boldsymbol{v}}(f) \\ &= \lambda\iota(\boldsymbol{v})(f).\end{aligned}$$

So, $\iota(\lambda\boldsymbol{v}) = \lambda\iota(\boldsymbol{v})$. This implies that ι is a linear map.
Now, for $\boldsymbol{v} \in V$ and $f \in V^*$,

$$\|\iota(\boldsymbol{v})(f)\| = \|f(\boldsymbol{v})\| \leq \|f\|\|\boldsymbol{v}\|.$$

This implies that $\|\iota\| = 1$, using the result above. So, ι is an isometry. □

Open maps

Theorem 3.4.7 (Baire-Category Theorem)**.** *Let X be a complete metric space, and let $(U_n)_{n=1}^{\infty}$ be a collection of dense open sets in X. Then,*

- *the intersection*
$$\bigcap_{n=1}^{\infty} U_n$$
is dense in X;
- *X cannot be written as a countable union of nowhere dense sets.*

Proof.

- Let $W \subseteq X$ be open and non-empty. We show that
$$W \cap \bigcap_{n=1}^{\infty} U_n$$
is not empty. So, let $V_0 = W$. We know that U_1 is dense, so $V_0 \cap U_1$ is not empty. Moreover, since V_0 and U_1 are open, we find that the intersection is open. So, there exists an $x_1 \in V_0 \cap U_1$ such that $B_X(x_n, r_n) \subseteq V_0 \cap U_1$. So, set $V_1 = B_X(x_1, r_1)$. We can iteratively define V_n for $n \in \mathbb{Z}_{\geq 1}$. Without loss of generality, we may assume that $r_n < \frac{1}{2n}$ and
$$\overline{B}_X(x_n, r_n) \subseteq V_{n-1} \cap U_n.$$

We claim that the sequence $(x_n)_{n=1}^\infty$ is Cauchy. Let $\varepsilon > 0$. Choose a natural number $N \in \mathbb{Z}_{\geq 1}$ such that $N > \frac{1}{\varepsilon}$. Let $m, n \in \mathbb{Z}_{\geq 1}$ with $m, n \geq N$. Then, we have $x_m, x_n \in V_N = B_X(x_N, \frac{1}{2N})$. This implies that

$$d(x_m, x_n) \leq d(x_m, x_N) + d(x_N, x_n) < \frac{1}{2N} + \frac{1}{2N} = \frac{1}{N} < \varepsilon.$$

So, (x_n) is Cauchy. Since X is complete, there exists an $x \in X$ such that $x_n \to x$.

Finally, we claim that $x \in V_n$ for all $n \in \mathbb{Z}_{\geq 1}$. Let $n \in \mathbb{Z}_{\geq 1}$. For all $m \in \mathbb{Z}_{\geq 1}$, if $m > n$, then we know that $x_m \in V_{n+1}$. In that case, the sequence $(x_m)_{m=n+1}^\infty$ is in V_{n+1}. So, the limit x must be in the closure, i.e.

$$x \in \overline{V_{n+1}} \subseteq V_n \cap U_{n+1} \subseteq V_n.$$

This implies that $x \in V_n$. Therefore,

$$x \in W \cap \bigcap_{n=1}^\infty U_n.$$

This implies that the intersection

$$\bigcap_{n=1}^\infty U_n$$

is dense.

- Assume, for a contradiction, that X can be written as a countable union of nowhere dense sets. So, let $(V_n)_{n=1}^\infty$ be a collection of nowhere dense sets in X such that

$$X = \bigcup_{n=1}^\infty V_n.$$

In that case,

$$\varnothing = X \setminus \bigcup_{n=1}^\infty \overline{V_n} = \bigcap_{n=1}^\infty X \setminus \overline{V_n}.$$

We know that for each $n \in \mathbb{Z}_{\geq 1}$, $X \setminus \overline{V_n}$ is open and dense. From the result above, we know that the intersection is dense, i.e. it is not empty. This is a contradiction. So, V cannot be written as a countable union of nowhere dense sets.

□

Definition 3.4.8. Let X and Y be topological spaces, and let $f : X \to Y$ be a function. Then, f is an *open map* if for all $U \subseteq X$, $f(U) \subseteq Y$ is open.

Proposition 3.4.9. *Let X and Y be metric spaces and let $f : X \to Y$ be a function. Then, f is an open map if and only if for every $x \in X$ and $\delta > 0$, there exists an $\varepsilon > 0$ such that for $y \in X$, if $d_Y(f(x), f(y)) < \varepsilon$, then there exists a $z \in X$ such that $d_X(x, z) < \delta$ with $f(y) = f(z)$.*

Proof.

- Assume that f is an open map. Let $x \in X$ and $\delta > 0$. Let
$$U = B_X(x, \delta).$$
Since f is an open map, we know that $f(U)$ is open. We have $f(x) \in f(U)$. In that case, there exists an $\varepsilon > 0$ such that for $y \in X$, if $d_Y(f(x), f(y)) < \varepsilon$, then $f(y) \in f(B_X(x, \delta))$. So, there exists a $z \in X$ with $d_X(x, z) < \delta$ with $f(y) = f(z)$. That is, for all $x \in X$ and $\delta > 0$, there exists an $\varepsilon > 0$ such that for $y \in X$, if $d_Y(f(x), f(y)) < \varepsilon$, then there exists a $z \in X$ such that $d_X(x, z) < \delta$ with $f(y) = f(z)$.

- Assume that for every $x \in X$ and $\delta > 0$, there exists an $\varepsilon > 0$ such that for $y \in X$, if $d_Y(f(x), f(y)) < \varepsilon$, then there exists a $z \in X$ such that $d_X(x, z) < \delta$ with $f(y) = f(z)$. Let $U \subseteq X$ be open. We show that $f(U)$ is open. So, let $x \in U$. Since U is open, there exists a $\delta_x > 0$ such that for all $y \in X$, if $d_X(x, y) < \delta_x$, then $y \in U$. Now, we can find an $\varepsilon_x > 0$ such that for $y \in X$, if $d_Y(f(x), f(y)) < \varepsilon_x$, then there exists a $z \in X$ such that $d_X(x, z) < \delta_x$ with $f(y) = f(z)$. This implies that $z \in U$. So, $f(y) = f(z) \in f(U)$. In that case, for all $x \in X$, there exists an ε_x such that for all $y \in X$, if $d_Y(f(x), f(y)) < \varepsilon_x$, then $f(y) \in f(U)$. This implies that $f(U)$ is open. In other words, f is an open map.

□

Another way of writing this is the following- $f : X \to Y$ is an open map if and only if for every $x \in X$ and $\delta > 0$, there exists an $\varepsilon > 0$ such that
$$B_Y(f(x), \varepsilon) \subseteq f(B_X(x, \delta)).$$

Proposition 3.4.10. *Let V and W be normed vector spaces, and let $T : V \to W$ be a linear operator. Then, T is an open map if and only if there exists an $\varepsilon > 0$ such that for $\boldsymbol{v} \in V$, if $\|T(\boldsymbol{v})\|_W < \varepsilon$, then there exists a $\boldsymbol{u} \in V$ such that $\|\boldsymbol{u}\|_V < 1$ with $T(\boldsymbol{u}) = T(\boldsymbol{v})$.*

Proof.

- Assume that T is an open map. Set $\boldsymbol{x} = \boldsymbol{0}$ and $\delta = 1$. In that case, there exists an $\varepsilon > 0$ such that for $\boldsymbol{v} \in V$, if $\|T(\boldsymbol{v}) - T(\boldsymbol{x})\|_W < \varepsilon$, then there exists a $\boldsymbol{u} \in V$ such that $\|\boldsymbol{u} - \boldsymbol{x}\|_V < \delta$ with $T(\boldsymbol{u}) = T(\boldsymbol{v})$. That is, there exists an $\varepsilon > 0$ such that for $\boldsymbol{v} \in V$, if $\|T(\boldsymbol{v})\|_W < \varepsilon$, then there exists a $\boldsymbol{u} \in V$ such that $\|\boldsymbol{u}\|_V < 1$ with $T(\boldsymbol{u}) = T(\boldsymbol{v})$.

- Assume that there exists an $\varepsilon > 0$ such that for $\boldsymbol{v} \in V$, if $\|T(\boldsymbol{v})\|_W < \varepsilon$, then there exists a $\boldsymbol{u} \in V$ such that $\|\boldsymbol{u}\|_V < 1$ with $T(\boldsymbol{u}) = T(\boldsymbol{v})$. We show that T is an open map. Let $\boldsymbol{x} \in V$ and $\delta > 0$. We can find an $\varepsilon > 0$ such that for $\boldsymbol{v} \in V$, if $\|T(\boldsymbol{v} - \boldsymbol{x})\|_W < \frac{\varepsilon}{\delta}$, then there exists a $\boldsymbol{u} \in V$ such that $\|\frac{1}{\delta}\boldsymbol{u}\|_V < 1$ with $T(\boldsymbol{u}) = T(\boldsymbol{v} - \boldsymbol{x})$. Using linearity and replacing $\boldsymbol{u}$ with $\boldsymbol{u} + \delta\boldsymbol{x}$; $\boldsymbol{v}$ with $\delta\boldsymbol{v}$; $\boldsymbol{x}$ with $\delta\boldsymbol{x}$, we find that for $\boldsymbol{v} \in V$, if $\|T(\boldsymbol{v}) - T(\boldsymbol{x})\|_W < \varepsilon$, then there exists a $\boldsymbol{u} \in V$ such that $\|\boldsymbol{u} - \boldsymbol{x}\|_V < \delta$ with $T(\boldsymbol{u}) = T(\boldsymbol{v})$. So, T is an open map.

□

Another way of writing this is the following- $T : V \to W$ is an open map if and only if there exists an $\varepsilon > 0$ such that

$$B_W(0, \varepsilon) \subseteq f(B_V(0, 1)).$$

Theorem 3.4.11 (Open Mapping Theorem). *Let V and W be Banach spaces, and let $T : V \to W$ be a surjective bounded linear operator. Then, T is an open map.*

Proof. □

Corollary 3.4.12. *Let V and W be Banach spaces, and let $T : V \to W$ be a bijective linear bounded operator. Then, T is an isomorphism.*

Proof. Since T is surjective, the open mapping theorem tells us that T is an open map. In that case, for all $U \subseteq V$ open, $f(U) \subseteq W$ is open. That is, for all $U \subseteq V$ open, the preimage of the inverse $(f^{-1})^{-1}(U) = f(U)$ is open. So, T^{-1} is a bounded (linear) operator. Therefore, T is an isomorphism. □

Graphs

Definition 3.4.13. Let X and Y be sets. Then, the *graph of f* is the set

$$\text{Graph}(f) = \{(x, f(x)) \mid x \in X\}.$$

For normed vector spaces V and W, we will use the following norm on $V \times W$:

$$\|(\boldsymbol{v}, \boldsymbol{w})\|_\infty = \max(\|\boldsymbol{v}\|_V, \|\boldsymbol{w}\|_W).$$

So, $V \times W$ is a normed vector space as well. By this definition, we can preserve properties that V and W both have in $V \times W$.

Proposition 3.4.14. *Let V and W be normed vector spaces, let $(\boldsymbol{v}_n, \boldsymbol{w}_n)_{n=1}^\infty$ be a sequence in $V \times W$, and let $(\boldsymbol{v}, \boldsymbol{w}) \in V \times W$. Then, $(\boldsymbol{v}_n, \boldsymbol{w}_n) \to (\boldsymbol{v}, \boldsymbol{w})$ if and only if $\boldsymbol{v}_n \to \boldsymbol{v}$ and $\boldsymbol{w}_n \to \boldsymbol{w}$.*

Proof.

- Assume that $(\boldsymbol{v}_n, \boldsymbol{w}_n) \to (\boldsymbol{v}, \boldsymbol{w})$. Let $\varepsilon > 0$. We can find an $N \in \mathbb{Z}_{\geq 1}$ such that for $n \in \mathbb{Z}_{\geq 1}$, if $n \geq N$, then

  $$\|(\boldsymbol{v}_n, \boldsymbol{w}_n) - (\boldsymbol{v}, \boldsymbol{w})\|_\infty < \varepsilon.$$

  In that case, we have

  $$\begin{aligned} \max(\|\boldsymbol{v}_n - \boldsymbol{v}\|_V, \|\boldsymbol{w}_n - \boldsymbol{w}\|_W) &= \|(\boldsymbol{v}_n - \boldsymbol{v}, \boldsymbol{w}_n - \boldsymbol{w})\|_\infty \\ &= \|(\boldsymbol{v}_n, \boldsymbol{w}_n) - (\boldsymbol{v}, \boldsymbol{w})\|_\infty < \varepsilon. \end{aligned}$$

  So, we have both $\|\boldsymbol{v}_n - \boldsymbol{v}\|_V < \varepsilon$ and $\|\boldsymbol{w}_n - \boldsymbol{w}\|_W < \varepsilon$. This implies that $\boldsymbol{v}_n \to \boldsymbol{v}$ and $\boldsymbol{w}_n \to \boldsymbol{w}$.

- Assume that $\boldsymbol{v}_n \to \boldsymbol{v}$ and $\boldsymbol{w}_n \to \boldsymbol{w}$. Let $\varepsilon > 0$. We can find $N_1, N_2 \in \mathbb{Z}_{\geq 1}$ such that for $n \in \mathbb{Z}_{\geq 1}$, if $n \geq N_1$, then

$$\|\boldsymbol{v}_n - \boldsymbol{v}\|_V < \varepsilon,$$

and if $n \geq N_2$, then

$$\|\boldsymbol{w}_n - \boldsymbol{w}\|_W < \varepsilon.$$

In that case, set $N = \max(N_1, N_2)$. Then, for $n \in \mathbb{Z}_{\geq 1}$, if $n \geq N$, then

$$\begin{aligned}\|(\boldsymbol{v}_n, \boldsymbol{w}_n) - (\boldsymbol{v}, \boldsymbol{w})\|_\infty &= \|(\boldsymbol{v}_n - \boldsymbol{v}, \boldsymbol{w}_n - \boldsymbol{w})\|_\infty \\ &= \max(\|\boldsymbol{v}_n - \boldsymbol{v}\|_V, \|\boldsymbol{w}_n - \boldsymbol{w}\|_W) < \varepsilon.\end{aligned}$$

Therefore, $(\boldsymbol{v}_n, \boldsymbol{w}_n) \to (\boldsymbol{v}, \boldsymbol{w})$.

□

Proposition 3.4.15. *Let V and W be normed vector spaces, let $(\boldsymbol{v}_n, \boldsymbol{w}_n)_{n=1}^\infty$ be a sequence in $V \times W$. Then, $(\boldsymbol{v}_n, \boldsymbol{w}_n)$ is Cauchy if and only if $(\boldsymbol{v}_n)$ and $(\boldsymbol{w}_n)$ are Cauchy.*

Proof.

- Assume that $(\boldsymbol{v}_n, \boldsymbol{w}_n)$ is Cauchy. Let $\varepsilon > 0$. We can find an $N \in \mathbb{Z}_{\geq 1}$ such that for $m, n \in \mathbb{Z}_{\geq 1}$, if $m, n \geq N$, then

$$\|(\boldsymbol{v}_n, \boldsymbol{w}_n) - (\boldsymbol{v}_m, \boldsymbol{w}_m)\|_\infty < \varepsilon.$$

In that case, we have

$$\begin{aligned}\max(\|\boldsymbol{v}_n - \boldsymbol{v}_m\|_V, \|\boldsymbol{w}_n - \boldsymbol{w}_m\|_W) &= \|(\boldsymbol{v}_n - \boldsymbol{v}_m, \boldsymbol{w}_n - \boldsymbol{w}_m)\|_\infty \\ &= \|(\boldsymbol{v}_n, \boldsymbol{w}_n) - (\boldsymbol{v}_m, \boldsymbol{w}_m)\|_\infty < \varepsilon.\end{aligned}$$

So, we have both $\|\boldsymbol{v}_n - \boldsymbol{v}_m\|_V < \varepsilon$ and $\|\boldsymbol{w}_n - \boldsymbol{w}_m\|_W < \varepsilon$. This implies that $(\boldsymbol{v}_n)$ and $(\boldsymbol{v}_n)$ are Cauchy.

- Assume that $(\boldsymbol{v}_n)$ and $(\boldsymbol{w}_n)$ are Cauchy. Let $\varepsilon > 0$. We can find $N_1, N_2 \in \mathbb{Z}_{\geq 1}$ such that for $m, n \in \mathbb{Z}_{\geq 1}$, if $m, n \geq N_1$, then

$$\|\boldsymbol{v}_n - \boldsymbol{v}_m\|_V < \varepsilon,$$

and if $m, n \geq N_2$, then

$$\|\boldsymbol{w}_n - \boldsymbol{w}_m\|_W < \varepsilon.$$

In that case, set $N = \max(N_1, N_2)$. Then, for $m, n \in \mathbb{Z}_{\geq 1}$, if $m, n \geq N$, then

$$\begin{aligned}\|(\boldsymbol{v}_n, \boldsymbol{w}_n) - (\boldsymbol{v}_m, \boldsymbol{w}_m)\|_\infty &= \|(\boldsymbol{v}_n - \boldsymbol{v}_m, \boldsymbol{w}_n - \boldsymbol{w}_m)\|_\infty \\ &= \max(\|\boldsymbol{v}_n - \boldsymbol{v}_m\|_V, \|\boldsymbol{w}_n - \boldsymbol{w}_m\|_W) < \varepsilon.\end{aligned}$$

Therefore, $(\boldsymbol{v}_n, \boldsymbol{w}_n)$ is Cauchy.

□

Proposition 3.4.16. *Let V and W be normed vector spaces. Then, $V \times W$ is a Banach space if and only if V and W are both Banach spaces.*

Proof.

- Assume that $V \times W$ is a Banach space. Let $(\boldsymbol{v}_n)_{n=1}^{\infty}$ and $(\boldsymbol{w}_n)_{n=1}^{\infty}$ be Cauchy sequences in V and W respectively. In that case, the sequence $(\boldsymbol{v}_n, \boldsymbol{w}_n)$ is a Cauchy sequence. Since $V \times W$ is complete, there exists a $(\boldsymbol{v}, \boldsymbol{w}) \in V \times W$ such that $(\boldsymbol{v}_n, \boldsymbol{w}_n) \to (\boldsymbol{v}, \boldsymbol{w})$. So, we must have $\boldsymbol{v}_n \to \boldsymbol{v}$ and $\boldsymbol{w}_n \to \boldsymbol{w}$. This implies that V and W are complete.

- Assume that V and W are Banach spaces. Let $(\boldsymbol{v}_n, \boldsymbol{w}_n)_{n=1}^{\infty}$ be a Cauchy sequence in $V \times W$. In that case, the sequences $(\boldsymbol{v}_n)$ and $(\boldsymbol{w}_n)$ are Cauchy sequences in V and W respectively. Since V and W are complete, there exists a $\boldsymbol{v} \in V$ and a $\boldsymbol{w} \in W$ such that $\boldsymbol{v}_n \to \boldsymbol{v}$ and $\boldsymbol{w}_n \to \boldsymbol{w}$.

□

Definition 3.4.17. Let V, W be normed vector spaces and let $T : V \to W$ be a linear operator. We say that T is *closed* if $\text{Graph}(T) \subseteq V \times W$ is closed.

Proposition 3.4.18. *Let V, W be normed vector spaces and let $T : V \to W$ be a bounded linear operator. Then, T is closed.*

Proof. Let $(\boldsymbol{v}_n, T(\boldsymbol{v}_n))_{n=1}^{\infty}$ be a sequence in $\text{Graph}(T)$, with $(\boldsymbol{v}_n, T(\boldsymbol{v}_n)) \to (\boldsymbol{v}, \boldsymbol{w})$, for some $(\boldsymbol{v}, \boldsymbol{w}) \in V \times W$. Under the $\|.\|_\infty$ norm, this implies that $\boldsymbol{v}_n \to \boldsymbol{v}$ and $T(\boldsymbol{v}_n) \to \boldsymbol{w}$. Since T is bounded, it is continuous. So, $T(\boldsymbol{v}) \to T(\boldsymbol{v})$ as well. In that case, $T(\boldsymbol{v}) = \boldsymbol{w}$. This implies that $(\boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{v}, T(\boldsymbol{v})) \in \text{Grap}(T)$. So, $\text{Graph}(T) \subseteq V \times W$ is closed. □

Theorem 3.4.19 (Closed Graph Theorem)**.** *Let V, W be Banach spaces and let $T : V \to W$ be a closed linear operator. Then, T is bounded.*

Proof. Since T is closed, we know that $\text{Graph}(T) \subseteq V \times W$ is closed. The spaces V and W are complete, so the product space $V \times W$ is also complete. So, the closed subspace $\text{Graph}(T)$ is complete. In particular, it is a Banach space under the $\|.\|_\infty$ norm.
Now, define the projection maps $\pi_V : V \times W \to V$ and $\pi_W : V \times W \to W$ by $\pi_V(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v}$ and $\pi_W(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{w}$. We know that for all $(\boldsymbol{v}, \boldsymbol{w}) \in V \times W$,

$$\|\pi_V(\boldsymbol{v}, \boldsymbol{w})\|_V = \|\boldsymbol{v}\|_V \leq \|\boldsymbol{v}, \boldsymbol{w}\|_\infty.$$

This implies that π_V is a bounded operator. Similarly, π_W is also bounded.
Next, let $\pi : \text{Graph}(T) \to V$ be the restriction of π_V. We claim that π_V is bijective. Let $(\boldsymbol{v}_1, T(\boldsymbol{v}_1)), (\boldsymbol{v}_2, T(\boldsymbol{v}_2)) \in V \times W$ with $\pi(\boldsymbol{v}_1, T(\boldsymbol{v}_1)) = \pi(\boldsymbol{v}_2, T(\boldsymbol{v}_2))$. In that case,

$$\boldsymbol{v}_1 = \pi(\boldsymbol{v}_1, T(\boldsymbol{v}_1)) = \pi(\boldsymbol{v}_2, T(\boldsymbol{v}_2)) = \boldsymbol{v}_2.$$

So, π is injective. Further, let $\boldsymbol{v} \in V$. We have $(\boldsymbol{v}, T(\boldsymbol{v})) \in \text{Graph}(T)$. In that case,

$$\pi(\boldsymbol{v}, T(\boldsymbol{v})) = \boldsymbol{v}.$$

So, π is surjective as well. This implies that π is a bijective linear bounded operator. Since $\text{Graph}(T)$ and V are Banach spaces, this implies that the inverse map π^{-1} is a linear bounded operator as well.

Finally, we claim that $T = \pi_W \circ \pi^{-1}$. Let $\boldsymbol{v} \in V$. We find that

$$\begin{aligned}(\pi_W \circ \pi^{-1})(\boldsymbol{v}) &= \pi_W(\pi^{-1}(\boldsymbol{v})) \\ &= \pi_W(\boldsymbol{v}, T(\boldsymbol{v})) \\ &= T(\boldsymbol{v}).\end{aligned}$$

So, $T = \pi_W \circ \pi^{-1}$. We know that both π^{-1} and π_W are bounded, so their composition T is also bounded. □

Compact Operators

Definition 3.4.20. Let V, W be normed vector spaces and let $T : V \to W$ be a linear operator. Then, T is *compact* if for any bounded sequence $(\boldsymbol{v}_n)_{n=1}^{\infty}$ in V, the sequence $(T(v)_n)$ in W has a convergent subsequence.

We can characterise the concept of a compact operator with a compact space.

Proposition 3.4.21. *Let V, W be normed vector spaces and let $T : V \to W$ be a linear operator. Then, T is compact if and only if $T(\overline{B}_V(\mathbf{0}, 1))$ is compact.*

Proof.

- First, assume that T is compact. Let $(\boldsymbol{x}_n)_{n=1}^{\infty}$ be a sequence in $\overline{B}_V(\mathbf{0}, 1)$. We know that $(\boldsymbol{x}_n)$ is bounded. So, we know that $(T(\boldsymbol{x}_n))$ has a convergent subsequence. This implies that $T(\overline{B}_V(\mathbf{0}, 1))$ is compact.
- Now, assume that $T(\overline{B}_V(\mathbf{0}, 1))$ is compact. Let $(\boldsymbol{x}_n)_{n=1}^{\infty}$ be a bounded sequence. Without loss of generality, assume that for all $m, n \in \mathbb{Z}_{\geq 1}$, $\|\boldsymbol{x}_m - \boldsymbol{x}_n\| \leq 1$.[2] Next, define the sequence $(\boldsymbol{y}_n)_{n=1}^{\infty}$ in V given by $\boldsymbol{y}_n = \boldsymbol{x}_n - \boldsymbol{x}_1$. We know that for all $n \in \mathbb{Z}_{\geq 1}$, $\|\boldsymbol{x}_m - \boldsymbol{x}_1\| \leq 1$. So, the sequence $(\boldsymbol{y}_n)$ is in $\overline{B}_V(\mathbf{0}, 1)$. This implies that $(\boldsymbol{y}_n)$ has a convergent subsequence $(\boldsymbol{y}_{n_k})_{k=1}^{\infty}$. In that case, $(\boldsymbol{x}_{n_k})$ is also a convergent subsequence. This implies that T is compact.

□

We will now consider an example of a compact operator. Define the map $T : \ell^{\infty} \to \ell^{\infty}$ by

$$T(x_1, x_2, x_3, x_4, \dots) = (x_1, x_2, 0, 0, \dots).$$

This is a compact operator since for any $(x_n)_{n=1}^{\infty}$ in ℓ^{∞}, $T(\boldsymbol{x}_n)_{n=3}^{\infty}$ is a convergent subsequence of $(T(\boldsymbol{x}_n))_{n=1}^{\infty}$.

Now, we show that a compact operator must be bounded.

Proposition 3.4.22. *Let V, W be normed vector spaces and let $T : V \to W$ be a compact operator. Then, T is bounded.*

Proof. Assume that T is not bounded. So, there exists a sequence $(\boldsymbol{v}_n)_{n=1}^{\infty}$ in V such that for all $n \in \mathbb{Z}_{\geq 1}$, $\|\boldsymbol{v}_n\| \leq 1$ and $\|T(\boldsymbol{v}_n)\|_W \geq n$. Then, $(T(\boldsymbol{v}_n))$ cannot have a convergent subsequence, since the subsequence would need to be bounded. Therefore, T cannot be compact. So, if T is compact, then T is bounded. □

[2] Since the sequence is bounded, we can find a $K > 0$ such that for all $n \in \mathbb{Z}_{\geq 1}$, $\|\boldsymbol{x}_n\| \leq K$. In that case, we have $\|\boldsymbol{x}_m - \boldsymbol{x}_n\| \leq 2K$. So, we can scale by $\frac{1}{2K}$ to achieve this.

The converse of this statement is however false. Let $T : \ell^\infty \to \ell^\infty$ be the identity function. This is clearly a bounded function. Now, define the sequence $(\boldsymbol{e}^{(k)})_{k=1}^\infty$ given by

$$\boldsymbol{e}_n^{(k)} = \begin{cases} 1 & n = k \\ 0 & \text{otherwise.} \end{cases}$$

Then, for all $k \in \mathbb{Z}_{\geq 1}$, $\|\boldsymbol{e}_k\|_\infty = 1$. So, $(\boldsymbol{e}^{(k)})$ is bounded. However, it does not have a convergent subsequence- for all $k, m \in \mathbb{Z}_{\geq 1}$, if $k \neq m$, then

$$\|\boldsymbol{e}^{(m)} - \boldsymbol{e}^{(k)}\|_\infty = 1.$$

In that case, any subsequence will not be Cauchy. This implies that $(T(\boldsymbol{e}^{(k)})) = (\boldsymbol{e}^{(k)})$ does not have a convergent subsequence. So, $T : V \to W$ is a bounded operator, but not compact.

Proposition 3.4.23. *Let V, W be Banach spaces, and let $(T_n)_{n=1}^\infty$ be a sequence of compact operators from V to W, with $T_n \to T$. Then, T is a compact operator.*

Proof. Let $(\boldsymbol{v}_n)_{n=1}^\infty$ be a bounded sequence in V. Since T_1 is compact, we can find a subsequence $(\boldsymbol{v}_{1n})_{n=1}^\infty$ of $(\boldsymbol{v}_n)$ such that $(T_1(\boldsymbol{v}_{1n}))$ is convergent. Moreover, since T_2 is compact, we can find a subsequence $(\boldsymbol{v}_{2n})_{n=1}^\infty$ of $(\boldsymbol{v}_{1n})$ such that $(T_2(\boldsymbol{v}_{2n}))$ is convergent. Using this approach, we can inductively define the sequences $(\boldsymbol{v}_{kn})_{n=1}^\infty$ for all $k \in \mathbb{Z}_{\geq 1}$.
Now, define the sequence $(\boldsymbol{x}_n)_{n=1}^\infty$ by $\boldsymbol{x}_n = \boldsymbol{v}_{nn}$. By definition, this is a subsequence of $(\boldsymbol{v}_n)$. Without loss of generality, assume that for all $n \in \mathbb{Z}_{\geq 1}$, $\|\boldsymbol{x}_n\|_V \leq 1$.[3] Next, we claim that $(T(\boldsymbol{x}_n))$ is Cauchy. So, let $\varepsilon > 0$. Since $T_n \to T$, we can find an $K \in \mathbb{Z}_{\geq 1}$ such that for $n \in \mathbb{Z}_{\geq 1}$, if $n \geq K$, then $\|T_n - T\| < \frac{\varepsilon}{3}$. In that case, for $n \in \mathbb{Z}_{\geq 1}$, if $n \geq K$, then $\|T_n(\boldsymbol{x}_n) - T(\boldsymbol{x}_n)\|_W < \frac{\varepsilon}{3}$ for all $n \in \mathbb{Z}_{\geq 1}$. By definition, $(\boldsymbol{x}_n)$ is a subsequence of $(\boldsymbol{v}_{Kn})$. We know that $(T_K(\boldsymbol{v}_{Kn}))$ is Cauchy, so $(T_K(\boldsymbol{x}_n))$ is also Cauchy. In that case, there exists an $N' \in \mathbb{Z}_{\geq 1}$ such that for $m, n \in \mathbb{Z}_{\geq 1}$, if $m, n \geq N'$, then $\|T_K(\boldsymbol{x}_m) - T_K(\boldsymbol{x}_n)\|_W < \frac{\varepsilon}{3}$. Now, set $N = \max(K, N')$. Then, for all $m, n \in \mathbb{Z}_{\geq 1}$, if $m, n \geq N$, then

$$\begin{aligned} \|T(\boldsymbol{x}_m) - T(\boldsymbol{x}_n)\|_W &\leq \|T(\boldsymbol{x}_n) - T_K(\boldsymbol{x}_n)\|_W + \|T_K(\boldsymbol{x}_n) - T_K(\boldsymbol{x}_m)\|_W \\ &\quad + \|T_K(\boldsymbol{x}_m) - T(\boldsymbol{x}_m)\|_W \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

This implies that $(T(\boldsymbol{x}_n))$ is Cauchy. Since W is complete, $(T(\boldsymbol{x}_n))$ is convergent. So, T is a compact operator. □

[3] We have to scale the vectors $(\boldsymbol{x}_n)$ so that the norm is always less than 1. Then, the original $(\boldsymbol{x}_n)$ is convergent if and only if the scaled $(\boldsymbol{x}_n)$ is convergent.

3.5 Hilbert Theory

Proposition 3.5.1. *Let V be an inner product spaces and let $(\boldsymbol{v}_n)_{n=1}^{\infty}$ and $(\boldsymbol{w}_n)_{n=1}^{\infty}$ be sequences in V, with $\boldsymbol{v}_n \to \boldsymbol{v}$ and $\boldsymbol{w}_n \to \boldsymbol{w}$, for some $\boldsymbol{v}, \boldsymbol{w} \in W$. Then, $\langle \boldsymbol{v}_n, \boldsymbol{w}_n \rangle \to \langle \boldsymbol{v}, \boldsymbol{w} \rangle$.*

Proof. We can find a $K_1 > 0$ such that $\|\boldsymbol{w}\| \leq K_1$. Since $\boldsymbol{v}_n \to \boldsymbol{v}$, the sequence $(\boldsymbol{v}_n)$ is bounded. So, there exists a $K_2 > 0$ such that for all $n \in \mathbb{Z}_{\geq 1}$, $\|\boldsymbol{x}_n\| \leq K_2$.
Now, let $\varepsilon > 0$. Since $\boldsymbol{v}_n \to \boldsymbol{v}$, we can find an $N_1 \in \mathbb{Z}_{\geq 1}$ such that for $n \in \mathbb{Z}_{\geq 1}$, if $n \geq N_1$, then $\|\boldsymbol{v} - \boldsymbol{v}_n\| < \frac{\varepsilon}{2K_1}$. Similarly, since $\boldsymbol{w}_n \to \boldsymbol{w}$, we can find an $N_2 \in \mathbb{Z}_{\geq 1}$ such that for $n \in \mathbb{Z}_{\geq 1}$, if $n \geq N_2$, then $\|\boldsymbol{w} - \boldsymbol{w}_n\| < \frac{\varepsilon}{2K_2}$. Now, set $N = \max(N_1, N_2)$. In that case, for $n \in \mathbb{Z}_{\geq 1}$, if $n \geq N$, then

$$\begin{aligned}
|\langle \boldsymbol{v}, \boldsymbol{w} \rangle - \langle \boldsymbol{v}_n, \boldsymbol{w}_n \rangle| &= |\langle \boldsymbol{v}, \boldsymbol{w} \rangle - \langle \boldsymbol{v}_n, \boldsymbol{w} \rangle + \langle \boldsymbol{v}_n, \boldsymbol{w} \rangle - \langle \boldsymbol{v}_n, \boldsymbol{w}_n \rangle| \\
&= |\langle \boldsymbol{v} - \boldsymbol{v}_n, \boldsymbol{w} \rangle + \langle \boldsymbol{v}_n, \boldsymbol{w} - \boldsymbol{w}_n \rangle| \\
&\leq |\langle \boldsymbol{v} - \boldsymbol{v}_n, \boldsymbol{w} \rangle| + |\langle \boldsymbol{v}_n, \boldsymbol{w} - \boldsymbol{w}_n \rangle| \\
&\leq \|\boldsymbol{v} - \boldsymbol{v}_n\| \|\boldsymbol{w}\| + \|\boldsymbol{v}_n\| \|\boldsymbol{w} - \boldsymbol{w}_n\| \\
&< \frac{\varepsilon}{2K_1} \cdot K_1 + K_2 \cdot \frac{\varepsilon}{2K_2} = \varepsilon.
\end{aligned}$$

This implies that $\langle \boldsymbol{v}_n, \boldsymbol{w}_n \rangle \to \langle \boldsymbol{v}, \boldsymbol{w} \rangle$. □

Definition 3.5.2. Let V be an inner product space and let $E \subseteq V$ be a subspace. We define the *orthogonal complement of* E by the set

$$E^{\perp} = \{\boldsymbol{v} \in V \mid \langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0 \ \forall \boldsymbol{w} \in E\}.$$

The set $E^{\perp}$ is always closed.

Proposition 3.5.3. *Let V be an inner product space and let $E \subseteq V$ be a subspace. Then, $E^{\perp}$ is closed.*

Proof. Let $(\boldsymbol{v}_n)_{n=1}^{\infty}$ be a sequence in $E^{\perp}$ with $\boldsymbol{v}_n \to \boldsymbol{v}$, for some $\boldsymbol{v} \in V$. Let $\boldsymbol{w} \in E$. We know that for all $n \in \mathbb{Z}_{\geq 1}$, $\langle \boldsymbol{v}_n, \boldsymbol{w} \rangle = 0$. In that case,

$$\langle \boldsymbol{v}_n, \boldsymbol{w} \rangle \to 0.$$

Moreover, since $\boldsymbol{v}_n \to \boldsymbol{v}$, we know that $\langle \boldsymbol{v}_n, \boldsymbol{w} \rangle \to \langle \boldsymbol{v}, \boldsymbol{w} \rangle$. This implies that $\boldsymbol{v} \in E^{\perp}$. So, $E^{\perp}$ is closed. □

Theorem 3.5.4 (Pythagoras' Theorem)**.** *Let V be an inner product space, and let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ are vectors in V that are mutually orthogonal. Then,*

$$\left\| \sum_{k=1}^{n} \boldsymbol{x}_k \right\|^2 = \sum_{k=1}^{n} \|\boldsymbol{x}_k\|^2.$$

Proof. We find that

$$\begin{aligned}\left\|\sum_{k=1}^{n}\boldsymbol{x}_k\right\|^2 &= \left\langle\sum_{k=1}^{n}\boldsymbol{x}_k, \sum_{k=1}^{n}\boldsymbol{x}_k\right\rangle \\ &= \sum_{j=1}^{n}\sum_{k=1}^{n}\langle\boldsymbol{x}_j, \boldsymbol{x}_k\rangle \\ &= \sum_{k=1}^{n}\langle\boldsymbol{x}_k, \boldsymbol{x}_k\rangle \\ &= \sum_{k=1}^{n}\|\boldsymbol{x}_k\|^2.\end{aligned}$$

□

Proposition 3.5.5. *Let V be a Hilbert space and let $H \subseteq V$ be a closed subspace. Then, for every $\boldsymbol{v} \in V$, there exists unique $\boldsymbol{x} \in H$ and $\boldsymbol{y} \in H^{\perp}$ such that $\boldsymbol{v} = \boldsymbol{x} + \boldsymbol{y}$.*[4]

Proof. Let

$$\delta = \inf_{\boldsymbol{x}\in H}\|\boldsymbol{v} - \boldsymbol{x}\|.$$

We know that there exists a sequence $(\boldsymbol{x}_n)_{n=1}^{\infty}$ in H such that

$$\|\boldsymbol{v} - \boldsymbol{x}_n\| \to \delta.$$

We claim that $(\boldsymbol{x}_n)$ is Cauchy. Let $\varepsilon > 0$. Fix $K \geq 1$ such that $K \geq \frac{\delta}{\varepsilon}$. Since $\|\boldsymbol{v} - \boldsymbol{x}_n\| \to \delta$, we can find an $N \in \mathbb{Z}_{\geq 1}$ such that for $n \in \mathbb{Z}_{\geq 1}$, if $n \geq N$, then

$$\delta \leq \|\boldsymbol{v} - \boldsymbol{x}_n\| < \delta + \frac{\varepsilon}{16K}.$$

Now, let $m, n \in \mathbb{Z}_{\geq 1}$ with $m, n \geq N$. By the Parallelogram Law, we know that

$$\|\boldsymbol{x}_m - \boldsymbol{x}_n\|^2 = 2\|\boldsymbol{x}_m - \boldsymbol{v}\|^2 + 2\|\boldsymbol{x}_n - \boldsymbol{v}\|^2 - 4\|\tfrac{1}{2}(\boldsymbol{x}_m + \boldsymbol{x}_n) - \boldsymbol{x}\|^2.$$

Since $\boldsymbol{x}_m, \boldsymbol{x}_n \in H$, we find that $\frac{1}{2}(\boldsymbol{x}_m + \boldsymbol{x}_n) \in H$. Therefore,

$$\|\tfrac{1}{2}(\boldsymbol{x}_m + \boldsymbol{x}_n) - \boldsymbol{v}\| \geq \delta.$$

In that case,

$$\begin{aligned}\|\boldsymbol{x}_m - \boldsymbol{x}_n\|^2 &\leq 2\|\boldsymbol{x}_m - \boldsymbol{v}\|^2 + 2\|\boldsymbol{x}_n - \boldsymbol{v}\|^2 - 4\delta^2 \\ &< 4\left(\delta + \frac{\varepsilon}{16K}\right)^2 - 4\delta^2 \\ &= \frac{\delta\varepsilon}{2K} + \frac{1}{256K^2}\varepsilon^2 \\ &\leq \frac{1}{2}\varepsilon^2 + \frac{1}{256K^2}\varepsilon^2 \\ &= \left(\frac{1}{2} + \frac{1}{256K^2}\right)\varepsilon^2 \leq \varepsilon^2.\end{aligned}$$

[4]This is denoted by $V = H \oplus H^{\perp}$.

So, for all $m, n \in \mathbb{Z}_{\geq 1}$, if $m, n \geq N$, then

$$\|\boldsymbol{x}_m - \boldsymbol{x}_n\| < \varepsilon.$$

This implies that $(\boldsymbol{x}_n)$ is Cauchy. Since $H \subseteq V$ is closed, we find that H is complete. In that case, there exists an $\boldsymbol{x} \in H$ such that $\boldsymbol{x}_n \to \boldsymbol{x}$. Then, we have $\|\boldsymbol{v} - \boldsymbol{x}\| = \delta$.
Next, we show that the value $\boldsymbol{x} \in H$ is unique. So, let $\boldsymbol{z} \in H$ such that $\|\boldsymbol{v} - \boldsymbol{z}\| = \delta$. By the Parallelogram Law, we know that

$$\|\boldsymbol{x} - \boldsymbol{z}\|^2 = 2\|\boldsymbol{x} - \boldsymbol{v}\|^2 + 2\|\boldsymbol{z} - \boldsymbol{v}\|^2 - 4\|\tfrac{1}{2}(\boldsymbol{x} + \boldsymbol{z}) - \boldsymbol{v}\|^2.$$

By construction, we have $\|\boldsymbol{x} - \boldsymbol{v}\| = \delta = \|\boldsymbol{v} - \boldsymbol{z}\|$. Moreover, since $\boldsymbol{x}, \boldsymbol{z} \in H$, we must find that

$$\|\boldsymbol{x} - \boldsymbol{z}\|^2 \leq 2\delta^2 + 2\delta^2 - 4\delta^2 = 0.$$

So, we have $\boldsymbol{x} = \boldsymbol{z}$.
Finally, set $\boldsymbol{y} = \boldsymbol{v} - \boldsymbol{x}$. We claim that $\boldsymbol{y} \in H^\perp$. So, let $\boldsymbol{w} \in H$. If $\boldsymbol{w} = \boldsymbol{0}$, then we know that $\langle \boldsymbol{w}, \boldsymbol{y} \rangle = 0$. Otherwise, we have $\boldsymbol{w} \neq \boldsymbol{0}$. In that case, define the function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(t) = \|\boldsymbol{y} + t\boldsymbol{w}\|^2 = \|\boldsymbol{y}\|^2 + 2t\langle \boldsymbol{w}, \boldsymbol{y} \rangle + t^2\|\boldsymbol{w}\|^2.$$

Also,

$$f(t) = \|\boldsymbol{y} + t\boldsymbol{w}\|^2 = \|\boldsymbol{v} - (\boldsymbol{x} - t\boldsymbol{w})\|^2.$$

For $t \in \mathbb{R}$, we have $\boldsymbol{x}, \boldsymbol{w} \in H$, so $\boldsymbol{x} - t\boldsymbol{w} \in H$ as well. So, $f(t) \geq \delta^2$. Moreover, we saw that only $\boldsymbol{x} \in H$ satisfies $\|\boldsymbol{x} - \boldsymbol{v}\| = \delta$. So, we have $f(0) = \delta^2$, and $f(t) > \delta^2$ for $t \in \mathbb{R}^\times$. This implies that $f'(0) = 0$. We have

$$f'(t) = 2\langle \boldsymbol{w}, \boldsymbol{y} \rangle + 2t\|\boldsymbol{w}\|.$$

So,

$$f'(0) = 2\langle \boldsymbol{w}, \boldsymbol{y} \rangle = 0.$$

This implies that $\langle \boldsymbol{w}, \boldsymbol{y} \rangle = 0$. In that case, $\boldsymbol{y} \in H^\perp$. By the uniqueness of $\boldsymbol{x}$, we find that $\boldsymbol{y} \in H^\perp$ is unique. □

Linear functionals

Proposition 3.5.6. *Let V be an inner product space, and let $\boldsymbol{v} \in V$. Then, the function $f_{\boldsymbol{v}} : V \to \mathbb{R}$ given by $f_{\boldsymbol{v}}(\boldsymbol{w}) = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$ is a linear functional.*

Proof. We know that for $\boldsymbol{w}_1, \boldsymbol{w}_2 \in V$,

$$f_{\boldsymbol{v}}(\boldsymbol{w}_1 + \boldsymbol{w}_2) = \langle \boldsymbol{w}_1 + \boldsymbol{w}_2, \boldsymbol{v} \rangle = \langle \boldsymbol{w}_1, \boldsymbol{v} \rangle + \langle \boldsymbol{w}_2, \boldsymbol{v} \rangle = f_{\boldsymbol{v}}(\boldsymbol{w}_1) + f_{\boldsymbol{v}}(\boldsymbol{w}_2).$$

Moreover, for $\boldsymbol{w} \in V$ and $\lambda \in \mathbb{R}$,

$$f_{\boldsymbol{v}}(\lambda\boldsymbol{w}) = \langle \lambda\boldsymbol{w}, \boldsymbol{v} \rangle = \lambda\langle \boldsymbol{w}, \boldsymbol{v} \rangle = \lambda f_{\boldsymbol{v}}(\boldsymbol{w}).$$

So, f is a linear map. Furthermore, for $\boldsymbol{w} \in V$,

$$f_{\boldsymbol{v}}(\boldsymbol{w}) = \langle \boldsymbol{v}, \boldsymbol{w} \rangle \leq \|\boldsymbol{v}\|\|\boldsymbol{w}\|.$$

In that case, f is a linear functional, with $\|f_{\boldsymbol{v}}\| \leq \|\boldsymbol{v}\|$. □

Proposition 3.5.7. *Let V be a Hilbert space, and let $f : V \to \mathbb{R}$ be a linear functional. Then, there exists a unique $\boldsymbol{v} \in V$ such that for all $\boldsymbol{w} \in V$, $f(\boldsymbol{w}) = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$.*

Proof. If f is the trivial map, then we know that for all $\boldsymbol{w} \in V$,

$$f(\boldsymbol{w}) = 0 = \langle \boldsymbol{w}, \boldsymbol{0} \rangle.$$

Now, assume that f is a non-trivial map. Since f is bounded, we find that

$$\ker(f) = f^{-1}(0)$$

is a closed subspace. Moreover, since $V \setminus \ker(f)$ is non-empty, there exists a $\boldsymbol{z} \in \ker(f)^{\perp}$ such that $\|\boldsymbol{z}\| = 1$.[5] Now, let $\boldsymbol{v} \in V$. Define

$$\boldsymbol{u} = f(\boldsymbol{v})\boldsymbol{z} - \boldsymbol{v} f(\boldsymbol{z}).$$

We have

$$f(\boldsymbol{u}) = f(\boldsymbol{v})f(\boldsymbol{z}) - f(\boldsymbol{v})f(\boldsymbol{z}) = 0,$$

so $\boldsymbol{u} \in \ker(f)$. This implies that $\langle \boldsymbol{z}, \boldsymbol{u} \rangle = 0$. In that case,

$$f(\boldsymbol{z})\langle \boldsymbol{v}, \boldsymbol{z} \rangle = f(\boldsymbol{v})\langle \boldsymbol{z}, \boldsymbol{z} \rangle = f(\boldsymbol{v}).$$

So, for all $\boldsymbol{v} \in V$,

$$f(\boldsymbol{v}) = \langle \boldsymbol{v}, f(\boldsymbol{z})\boldsymbol{z} \rangle.$$

Now, we show that the value $\boldsymbol{v} \in V$ is unique. So, assume that there exist $\boldsymbol{x}, \boldsymbol{y} \in V$ such that for all $\boldsymbol{w} \in V$,

$$\langle \boldsymbol{x}, \boldsymbol{w} \rangle = \langle \boldsymbol{y}, \boldsymbol{w} \rangle.$$

In that case, $\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{w} \rangle = 0$ for all $\boldsymbol{w} \in V$. In particular,

$$\|\boldsymbol{x} - \boldsymbol{y}\|^2 = \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle = 0.$$

Therefore, $\boldsymbol{x} = \boldsymbol{y}$. So, the value $\boldsymbol{v} \in V$ is unique. □

Linear algebra

Definition 3.5.8. Let V be an inner product space. Then, a subset $E \subseteq V$ is an *orthogonal set* if $\boldsymbol{0} \notin E$ and for all $\boldsymbol{u}, \boldsymbol{v} \in V$, if $\boldsymbol{u} \neq \boldsymbol{v}$, then $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$.

Definition 3.5.9. Let V be an inner product space. Then, a subset $E \subseteq V$ is an *orthonormal set* if $\boldsymbol{0} \notin E$ and for all $\boldsymbol{u}, \boldsymbol{v} \in V$, if $\boldsymbol{u} \neq \boldsymbol{v}$, then $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$, and if $\boldsymbol{u} = \boldsymbol{v}$, then $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 1$.

Lemma 3.5.10. *Let V be an inner product space and let $E \subseteq V$ be an orthogonal set. Then, E is linearly independent.*

Proof. Let

$$c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + \cdots + c_n\boldsymbol{x}_n = \boldsymbol{0},$$

for $c_1, c_2, \ldots, c_n \in \mathbb{R}$ and $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in E$. We find that for all $i \in \{1, 2, \ldots, n\}$,

$$c_i\|\boldsymbol{x}_i\|^2 = c_i\langle \boldsymbol{x}_i, \boldsymbol{x}_i \rangle = \sum_{j=1}^{n} c_j\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle = \left\langle \boldsymbol{x}_i, \sum_{j=1}^{n} c_j\boldsymbol{x}_j \right\rangle = \langle \boldsymbol{x}_i, \boldsymbol{0} \rangle = 0.$$

Since $\boldsymbol{x}_i \neq \boldsymbol{0}$, we must have $c_i = 0$. So, E is linearly independent. □

[5] There exists a $\boldsymbol{v} \in V \setminus \ker(f)$. So, $\boldsymbol{v} = \boldsymbol{x} + \boldsymbol{y}$ for $\boldsymbol{x} \in \ker(f)$ and $\boldsymbol{y} \in \ker(f)^{\perp}$. Then, we must have $\boldsymbol{y} \neq \boldsymbol{0}$, so we can normalise $\boldsymbol{y}$ to get $\boldsymbol{z}$.

We can use the Gram-Schmidt process to convert a (countable) collection of linearly independent vectors into an orthonormal set of vectors with the same span. Let V be an inner product space, and let

$$E = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots\} \subseteq V$$

be a linearly independent set. We will construct the set

$$F = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \dots\} \subseteq V$$

of orthonormal sets such that $\operatorname{span}(E) = \operatorname{span}(F)$. For $n \in \mathbb{Z}_{\geq 1}$, we define

$$\boldsymbol{y}_n' = \boldsymbol{x}_n - \sum_{k=1}^{n} \langle \boldsymbol{x}_n, \boldsymbol{y}_k \rangle \boldsymbol{y}_k,$$

and $\boldsymbol{y}_n$ is the normalised vector of $\boldsymbol{y}_n'$.

It also has the same span since there is a bijection from the linearly independent set and the orthonormal set, and we can write all the $\boldsymbol{x}_n$ vectors as a linear combination of $\boldsymbol{y}_i$, for $i \in \{1, 2, \dots, n\}$.

We illustrate the Gram-Schmidt process in $\mathbb{R}^3$. Consider the set

$$\{(1, 0, 0), (0, 1, 1), (1, 0, 1)\}.$$

This is a linearly independent set, but not orthonormal. To construct the set of orthonormal vectors, we first set $\boldsymbol{y}_1' = \boldsymbol{x}_1$. Since $\boldsymbol{y}_1'$ is a unit vector, we set $\boldsymbol{y}_1 = \boldsymbol{x}_1$. Now, we set

$$\boldsymbol{y}_2' = (0, 1, 1) - \langle (1, 0, 0), (0, 1, 1) \rangle (1, 0, 0) = (0, 1, 1).$$

Therefore, we normalise $\boldsymbol{y}_2'$ to get:

$$\boldsymbol{y}_2 = \frac{1}{\|(0, 1, 1)\|_2}(0, 1, 1) = \frac{1}{\sqrt{2}}(0, 1, 1).$$

Next, we set

$$\begin{aligned} \boldsymbol{y}_3' &= (1, 0, 1) - \langle (1, 0, 0), (1, 0, 1) \rangle (1, 0, 0) - \frac{1}{2} \langle (0, 1, 1), (1, 0, 1) \rangle (0, 1, 1) \\ &= \frac{1}{2}(0, -1, 1). \end{aligned}$$

Finally, we normalise $\boldsymbol{y}_3'$ to get:

$$\boldsymbol{y}_3 = \frac{1}{\|(0, -1, 1)\|_2}(0, -1, 1) = \frac{1}{\sqrt{2}}(0, -1, 1).$$

Then, the set

$$\{\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3\}$$

has equal span and is orthonormal.

Proposition 3.5.11 (Bessel's Inequality). *Let V be a Hilbert space and let $(\boldsymbol{x}_\alpha)_{\alpha \in A}$ be an orthonormal collection in V, for some indexing set A. Then, for all $\boldsymbol{v} \in V$,*

$$\sum_{\alpha \in A} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 \leq \|\boldsymbol{v}\|^2.$$

In particular, the set of α such that $\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \neq 0$ is countable.

Proof. Let $F \subseteq A$ be finite. Then, for all $\boldsymbol{v} \in V$,

$$\begin{aligned} 0 &\leq \left\| \boldsymbol{v} - \sum_{\alpha \in F} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \boldsymbol{x}_\alpha \right\|^2 \\ &= \|\boldsymbol{v}\|^2 - 2 \sum_{\alpha \in F} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \cdot \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle + \sum_{\alpha \in F} \sum_{\beta \in F} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 \cdot |\langle \boldsymbol{v}_\alpha, \boldsymbol{x}_\beta \rangle|^2 \\ &= \|\boldsymbol{v}\|^2 - 2 \sum_{\alpha \in F} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 + \sum_{\alpha \in F} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 \\ &= \|\boldsymbol{v}\|^2 - \sum_{\alpha \in F} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2. \end{aligned}$$

Therefore,

$$\sum_{\alpha \in F} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 \leq \|\boldsymbol{v}\|^2.$$

Now, let

$$A^\times = \{\alpha \in A \mid \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \neq 0\},$$

and for $n \in \mathbb{Z}_{\geq 1}$, let

$$A_n = \{\alpha \in A \mid \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle > 1/n\}.$$

We claim that $|A_n| \leq n^2 \|\boldsymbol{y}\|^2$. Assume, for a contradiction, that $|A_n| > n^2 \|\boldsymbol{y}\|^2$. In that case, for a finite subset $A_n' \subseteq A_n$ with $|A_n'| > n^2 \|\boldsymbol{y}\|^2$,

$$\sum_{\alpha \in A_n'} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 > \sum_{\alpha \in A_n} \frac{1}{n^2} \geq \frac{1}{n^2} \cdot n^2 \|\boldsymbol{y}\|^2 = \|\boldsymbol{y}\|^2.$$

This is a contradiction. So, we must have that A_n is finite. Therefore, the countable union

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

is also countable.
Finally, enumerate

$$A = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots\}.$$

We have shown that for all $n \in \mathbb{Z}_{\geq 1}$,

$$\sum_{\alpha=1}^{n} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 \leq \|\boldsymbol{v}\|^2.$$

Therefore,

$$\sum_{\alpha=1}^{\infty} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 \leq \|\boldsymbol{v}\|^2.$$

This implies that

$$\sum_{\alpha \in A} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 = \sum_{\alpha=1}^{\infty} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 \leq \|\boldsymbol{v}\|^2.$$

□

Proposition 3.5.12. *Let V be a Hilbert space and let $(\boldsymbol{x}_\alpha)_{\alpha \in A}$ be an orthonormal collection in V, for some indexing set A. Then, the following are equivalent:*

1. *for all $\boldsymbol{v} \in V$, if $\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle = 0$ for all $\alpha \in A$, then $\boldsymbol{v} = \boldsymbol{0}$;*
2. *for all $\boldsymbol{v} \in V$,*
$$\|\boldsymbol{v}\|^2 = \sum_{\alpha \in A} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2;$$
3. *for all $\boldsymbol{v} \in V$,*
$$\boldsymbol{v} = \sum_{\alpha \in A} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \boldsymbol{x}_\alpha.$$

If the collection $(\boldsymbol{x}_\alpha)$ satisfies any of the conditions above, then we say that it forms an orthonormal basis for V.

Proof. We show (1) $\implies$ (3) $\implies$ (2) $\implies$ (1).

- Assume that for all $\boldsymbol{v} \in V$, if $\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle = 0$ for all $\alpha \in A$, then $\boldsymbol{v} = \boldsymbol{0}$. Let $\boldsymbol{v} \in V$. By Bessel's inequality, we know that
$$\sum_{\alpha \in A} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 \le \|\boldsymbol{v}\|^2.$$
So, the series
$$\sum_{\alpha \in A} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2$$
is Cauchy. Moreover, we know that the set of $\alpha \in A$ such that $\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \neq 0$ is countable. So, let $(\boldsymbol{x}_\alpha)_{\alpha=1}^\infty$ be a collection of such values. Now, let $\varepsilon > 0$. Since the series is Cauchy, we can find an $N \in \mathbb{Z}_{\ge 1}$ such that for $m, n \in \mathbb{Z}_{\ge 1}$, if $m \ge n \ge N$, then
$$\sum_{\alpha=n+1}^{m} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 = \left| \sum_{\alpha=1}^{m} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 - \sum_{\alpha=1}^{n} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 \right| < \varepsilon^2.$$
By Pythagoras' Theorem, we know that for $m, n \in \mathbb{Z}$ with $m \ge n$,
$$\left\| \sum_{\alpha=n+1}^{m} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \boldsymbol{x}_\alpha \right\|^2 = \sum_{\alpha=n+1}^{m} \|\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \boldsymbol{x}_\alpha\|^2 = \sum_{\alpha=n+1}^{m} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2.$$
Therefore, for $m, n \in \mathbb{Z}_{\ge 1}$, if $m \ge n \ge N$, then
$$\begin{aligned} \left\| \sum_{\alpha=1}^{m} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \boldsymbol{x}_\alpha - \sum_{\alpha=1}^{n} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \boldsymbol{x}_\alpha \right\| &= \left\| \sum_{\alpha=n+1}^{m} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \boldsymbol{x}_\alpha \right\| \\ &= \sqrt{\sum_{\alpha=n+1}^{m} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2} < \varepsilon. \end{aligned}$$

This implies that the series

$$\sum_{\alpha \in A} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \boldsymbol{x}_\alpha \sum_{\alpha=1}^{\infty} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \boldsymbol{x}_\alpha$$

converges. Now, let

$$\boldsymbol{z} = \boldsymbol{v} - \sum_{\alpha \in A} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \boldsymbol{x}_\alpha.$$

We find that for all $\beta \in A$,

$$\begin{aligned} \langle \boldsymbol{z}, \boldsymbol{x}_\beta \rangle &= \langle \boldsymbol{v}, \boldsymbol{x}_\beta \rangle - \sum_{\alpha \in A} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \cdot \langle \boldsymbol{x}_\alpha, \boldsymbol{x}_\beta \rangle \\ &= \langle \boldsymbol{v}, \boldsymbol{x}_\beta \rangle - \langle \boldsymbol{v}, \boldsymbol{x}_\beta \rangle = 0. \end{aligned}$$

This implies that $\boldsymbol{z} = 0$. So, we find that

$$\boldsymbol{v} = \sum_{\alpha \in A} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \boldsymbol{x}_\alpha.$$

- Assume that for all $\boldsymbol{v} \in V$,

$$\boldsymbol{v} = \sum_{\alpha \in A} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \boldsymbol{x}_\alpha.$$

In that case, for all $\boldsymbol{v} \in V$,

$$\begin{aligned} \|\boldsymbol{v}\|^2 &= \langle \boldsymbol{v}, \boldsymbol{v} \rangle \\ &= \sum_{\alpha \in A} \sum_{\beta \in A} \langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle \cdot \langle \boldsymbol{v}, \boldsymbol{x}_\beta \rangle \cdot |\langle \boldsymbol{x}_\alpha, \boldsymbol{x}_\beta \rangle|^2 \\ &= \sum_{\alpha \in A} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 \cdot |\langle \boldsymbol{x}_\alpha, \boldsymbol{x}_\alpha \rangle|^2 \\ &= \sum_{\alpha \in A} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2. \end{aligned}$$

- Assume that for all $\boldsymbol{v} \in V$,

$$\|\boldsymbol{v}\|^2 = \sum_{\alpha \in A} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2.$$

Now, fix $\boldsymbol{v} \in V$ such that $\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle = 0$ for all $\alpha \in A$. Then,

$$\|\boldsymbol{v}\|^2 = \sum_{\alpha \in A} |\langle \boldsymbol{v}, \boldsymbol{x}_\alpha \rangle|^2 = \sum_{\alpha \in A} 0 = 0.$$

This implies that $\boldsymbol{v} = \mathbf{0}$.

□

Proposition 3.5.13. *Let V be a Hilbert space. Then, it is separable if and only if it has a countable orthonormal basis. Moreover, if V has a countable orthonormal basis, then any orthonormal basis of V is countable.*

Proof. □

Unitary operators

Definition 3.5.14. Let V and W be inner product spaces, and let $T : V \to W$ be a linear operator. Then, T is *unitary* if it is surjective and for all $\boldsymbol{u}, \boldsymbol{v} \in V$,

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle_V = \langle T(\boldsymbol{u}), T(\boldsymbol{v}) \rangle_W.$$

Proposition 3.5.15. *Let V and W be inner product spaces, and let $T : V \to W$ be a linear operator. Then, T is unitary if and only if T is a surjective isometry.*

Proof.

- Assume that T is unitary. In that case, for all $\boldsymbol{v} \in V$,

  $$\|T(\boldsymbol{v})\|^2 = \langle T(\boldsymbol{v}), T(\boldsymbol{v}) \rangle = \langle \boldsymbol{v}, \boldsymbol{v} \rangle = \|\boldsymbol{v}\|^2.$$

  So, $\|T(\boldsymbol{v})\| = \|\boldsymbol{v}\|$. This implies that T is a (surjective) isometry.

- Assume that T is a surjective isometry. Let $\boldsymbol{u}, \boldsymbol{v} \in V$. We find that

  $$\begin{aligned} \langle T(\boldsymbol{u}), T(\boldsymbol{v}) \rangle &= \frac{1}{4}\|T(\boldsymbol{u}) + T(\boldsymbol{v})\|^2 - \frac{1}{4}\|T(\boldsymbol{u}) - T(\boldsymbol{v})\|^2 \\ &= \frac{1}{4}\|\boldsymbol{u} + \boldsymbol{v}\|^2 - \frac{1}{4}\|\boldsymbol{u} - \boldsymbol{v}\|^2 \\ &= \langle \boldsymbol{u}, \boldsymbol{v} \rangle. \end{aligned}$$

  Therefore, T is unitary.

$\square$

Proposition 3.5.16. *Let V be a separable Hilbert space, with a countable orthonormal basis $(\boldsymbol{x}_k)_{k=1}^{\infty}$. For each $\boldsymbol{v} \in V$, define the sequence $(v_k)_{k=1}^{\infty}$ in $\mathbb{R}$ by $v_k = \langle \boldsymbol{v}, \boldsymbol{x}_k \rangle$, and define the function $f : V \to \ell^2$ by $f(\boldsymbol{v}) = (\boldsymbol{v}_k)$. Then, f is a unitary map.*

Proof. Since $(\boldsymbol{x}_k)$ forms an orthonormal basis, we know that

$$\|(\boldsymbol{v}_k)\|_2^2 = \sum_{k=1}^{\infty} |\langle \boldsymbol{v}, \boldsymbol{x}_k \rangle|^2 = \|\boldsymbol{v}\|_V^2.$$

This implies that $(\boldsymbol{v}_k)$ is in ℓ^2- the function f is well-defined. Moreover, it is an isometry.
Next, we show that f is linear. Let $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$. We find that for all $k \in \mathbb{Z}_{\geq 1}$,

$$f(\boldsymbol{v}_1 + \boldsymbol{v}_2)_k = \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{x}_k \rangle = \langle \boldsymbol{v}_1, \boldsymbol{x}_k \rangle + \langle \boldsymbol{v}_2, \boldsymbol{x}_k \rangle = f(\boldsymbol{v}_1)_k + f(\boldsymbol{v}_2)_k.$$

So, $f(\boldsymbol{v}_1 + \boldsymbol{v}_2) = f(\boldsymbol{v}_1) + f(\boldsymbol{v}_2)$. Now, let $\boldsymbol{v} \in V$ and $\lambda \in \mathbb{R}$. We find that for all $k \in \mathbb{Z}_{\geq 1}$,

$$f(\lambda \boldsymbol{v})_k = \langle \lambda \boldsymbol{v}, \boldsymbol{x}_k \rangle = \lambda \langle \boldsymbol{v}, \boldsymbol{x}_k \rangle = \lambda f(\boldsymbol{v})_k.$$

So, $f(\lambda \boldsymbol{v}) = \lambda f(\boldsymbol{v})$. This implies that f is linear.

Finally, we show that f is surjective. Let $(x_n)_{n=1}^{\infty}$ be a sequence in ℓ^2. Consider the series

$$\sum_{n=1}^{\infty} a_n \boldsymbol{x}_n.$$

We claim that the series is Cauchy. Since $(x_n)_{n=1}^{\infty}$ is in ℓ^2, we can find an $N \in \mathbb{Z}_{\geq 1}$ such that for $m, n \in \mathbb{Z}_{\geq 1}$, if $m \geq n \geq N$, then

$$\sum_{k=n+1}^{m} |a_k|^2 = \left| \sum_{k=1}^{m} |a_k|^2 - \sum_{k=1}^{n} |a_k|^2 \right| < \varepsilon.$$

Since $(\boldsymbol{x}_n)$ is an orthonormal collection, Pythagoras' Theorem tells us that for $m, n \in \mathbb{Z}_{\geq 1}$ with $m \geq n$,

$$\left\| \sum_{k=n+1}^{m} a_k \boldsymbol{x}_k \right\|^2 = \sum_{k=n+1}^{m} \|a_k \boldsymbol{x}_k\|^2 = \sum_{k=n+1}^{m} |a_k|^2.$$

In that case, for $m, n \in \mathbb{Z}_{\geq 1}$, if $m \geq n \geq N$, then

$$\left| \sum_{k=1}^{m} \|a_k \boldsymbol{x}_k\|^2 - \sum_{k=1}^{n} \|a_k \boldsymbol{x}_k\|^2 \right| = \sum_{k=n+1}^{m} |a_k|^2 < \varepsilon.$$

Therefore, the series is Cauchy. Since V is complete, we can find a $\boldsymbol{v} \in V$ such that

$$\sum_{n=1}^{\infty} a_n \boldsymbol{x}_n = \boldsymbol{v}.$$

In that case, for all $n \in \mathbb{Z}_{\geq 1}$,

$$f(\boldsymbol{v})_k = \langle \boldsymbol{v}, \boldsymbol{x}_k \rangle = \sum_{n=1}^{\infty} a_n \langle \boldsymbol{x}_n, \boldsymbol{x}_k \rangle = a_k \langle \boldsymbol{x}_k, \boldsymbol{x}_k \rangle = a_k.$$

This implies that $f(\boldsymbol{v}) = (\boldsymbol{x}_n)$. So, f is a surjective isometry- f is unitary. $\square$