Part II — Linear Analysis

Theorems with proof

Based on lectures by J. W. Luk

Notes taken by Dexter Chua

Michaelmas 2015

Part IB Linear Algebra, Analysis II and Metric and Topological Spaces are essential

Normed and Banach spaces. Linear mappings, continuity, boundedness, and norms. Finite-dimensional normed spaces. [4]

The Baire category theorem. The principle of uniform boundedness, the closed graph theorem and the inversion theorem; other applications. [5]

The normality of compact Hausdorff spaces. Urysohn's lemma and Tiezte's extension theorem. Spaces of continuous functions. The Stone-Weierstrass theorem and applications. Equicontinuity: the Ascoli-Arzelà theorem. [5]

Inner product spaces and Hilbert spaces; examples and elementary properties. Orthonormal systems, and the orthogonalization process. Bessel's inequality, the Parseval equation, and the Riesz-Fischer theorem. Duality; the self duality of Hilbert space. [5]

Bounded linear operations, invariant subspaces, eigenvectors; the spectrum and resolvent set. Compact operators on Hilbert space; discreteness of spectrum. Spectral theorem for compact Hermitian operators. [5]

Contents

0 Introduction

1 Normed vector spaces

Proposition. Addition $+ : V \times V \to V$, and scalar multiplication $\cdot : \mathbb{F} \times V \to V$ are continuous with respect to the topology induced by the norm (and the usual product topology).

Proof. Let U be open in V. We want to show that $(+)^{-1}(U)$ is open. Let $(\mathbf{v}_1, \mathbf{v}_2) \in (+)^{-1}(U)$, i.e. $\mathbf{v}_1 + \mathbf{v}_2 \in U$. Since $\mathbf{v}_1 + \mathbf{v}_2 \in U$, there exists ε such that $B(\mathbf{v}_1 + \mathbf{v}_2, \varepsilon) \subseteq U$. By the triangle inequality, we know that $B(\mathbf{v}_1, \frac{\varepsilon}{2}) + B(\mathbf{v}_2, \frac{\varepsilon}{2}) \subseteq U$. Hence we have $(\mathbf{v}_1, \mathbf{v}_2) \in B\left((\mathbf{v}_1, \mathbf{v}_2), \frac{\varepsilon}{2}\right) \subseteq (+)^{-1}(U)$. So $(+)^{-1}(U)$ is open.

Scalar multiplication can be done in a very similar way. □

Proposition. If $(V, \|\cdot\|)$ is a normed vector space, then $B(t) = B(\mathbf{0}, t) = \{\mathbf{v} : \|\mathbf{v}\| < t\}$ is absolutely convex.

Proof. By triangle inequality. □

Proposition. A topological vector space $(V, \mathcal{U})$ is normable if and only if there exists an absolutely convex, bounded open neighbourhood of $\mathbf{0}$.

Proof. One direction is obvious — if V is normable, then $B(t)$ is an absolutely convex, bounded open neighbourhood of $\mathbf{0}$.

The other direction is not too difficult as well. We define the *Minkowski functional* $\mu : V \to \mathbb{R}$ by

$$\mu_C(\mathbf{v}) = \inf\{t > 0 : \mathbf{v} \in tC\},$$

where C is our absolutely convex, bounded open neighbourhood.

Note that by definition, for any $t < \mu_C(\mathbf{v})$, $\mathbf{v} \notin tC$. On the other hand, by absolute convexity, for any $t > \mu_C(\mathbf{v})$, we have $\mathbf{v} \in tC$.

We now show that this is a norm on V:

(i) If $\mathbf{v} = \mathbf{0}$, then $\mathbf{v} \in 0C$. So $\mu_C(\mathbf{0}) = 0$. On the other hand, suppose $\mathbf{v} \neq \mathbf{0}$. Since a singleton point is closed, $U = V \setminus \{\mathbf{v}\}$ is an open neighbourhood of 0. Hence there is some t such that $C \subseteq tU$. Alternatively, $\frac{1}{t}C \subseteq U$. Hence, $\mathbf{v} \notin \frac{1}{t}C$. So $\mu_C(\mathbf{v}) \geq \frac{1}{t} > 0$. So $\mu_C(\mathbf{v}) = \mathbf{0}$ iff $\mathbf{v} = \mathbf{0}$.

(ii) We have

$$\mu_C(\lambda\mathbf{v}) = \inf\{t > 0 : \lambda\mathbf{v} \in tC\} = \lambda \inf\{t > 0 : \mathbf{v} \in tC\} = \lambda\mu_C(\mathbf{v}).$$

(iii) We want to show that

$$\mu_C(\mathbf{v} + \mathbf{w}) \leq \mu_C(\mathbf{v}) + \mu_C(\mathbf{w}).$$

This is equivalent to showing that

$$\inf\{t > 0 : \mathbf{v} + \mathbf{w} \in tC\} \leq \inf\{t > 0 : \mathbf{v} \in tC\} + \inf\{r > 0 : \mathbf{w} \in rC\}.$$

This is, in turn equivalent to proving that if $\mathbf{v} \in tC$ and $\mathbf{w} \in rC$, then $(\mathbf{v} + \mathbf{w}) \in (t + r)C$.

Let $\mathbf{v}' = \mathbf{v}/t, \mathbf{w}' = \mathbf{w}/r$. Then we want to show that if $\mathbf{v}' \in C$ and $\mathbf{w}' \in C$, then $\frac{1}{(t+r)}(t\mathbf{v}' + r\mathbf{w}') \in C$. This is exactly what is required by convexity. So done. □

1.1 Bounded linear maps

Proposition. Let X, Y be normed vector spaces, $T : X \to Y$ a linear map. Then the following are equivalent:

(i) T is continuous.

(ii) T is continuous at $\mathbf{0}$.

(iii) T is bounded.

Proof. (i) $\Rightarrow$ (ii) is obvious.

(ii) $\Rightarrow$ (iii): Consider $B_Y(1) \subseteq Y$, the unit open ball. Since T is continuous at 0, $T^{-1}(B_Y(1)) \subseteq X$ is open. Hence there exists $\varepsilon > 0$ such that $B_X(\varepsilon) \subseteq T^{-1}(B_Y(1))$. So $T(B_{\mathbf{x}}(\varepsilon)) \subseteq B_Y(1)$. So $T(B_X(1)) \subseteq B_Y\left(\frac{1}{\varepsilon}\right)$. So T is bounded.

(iii) $\Rightarrow$ (i): Let $\varepsilon > 0$. Then $\|T\mathbf{x}_1 - T\mathbf{x}_2\|_Y = \|T(\mathbf{x}_1 - \mathbf{x}_2)\|_Y \leq C\|\mathbf{x}_1 - \mathbf{x}_2\|_X$. This is less than ε if $\|\mathbf{x}_1 - \mathbf{x}_2\| < C^{-1}\varepsilon$. So done. □

1.2 Dual spaces

Proposition. Let V be a normed vector space. Then V^* is a Banach space.

Proof. Suppose $\{T_i\} \in V^*$ is a Cauchy sequence. We define T as follows: for any $\mathbf{v} \in V$, $\{T_i(\mathbf{v})\} \subseteq \mathbb{F}$ is Cauchy sequence. Since $\mathbb{F}$ is complete (it is either $\mathbb{R}$ or $\mathbb{C}$), we can define $T : V \to \mathbb{R}$ by

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

Our objective is to show that $T_i \to T$. The first step is to show that we indeed have $T \in V^*$, i.e. T is a bounded map.

Let $\|\mathbf{v}\| \leq 1$. Pick $\varepsilon = 1$. Then there is some N such that for all $i > N$, we have

$$|T_i(\mathbf{v}) - T(\mathbf{v})| < 1.$$

Then we have

$$\begin{aligned} |T(\mathbf{v})| &\leq |T_i(\mathbf{v}) - T(\mathbf{v})| + |T_i(\mathbf{v})| \\ &< 1 + \|T_i\|_{V^*}\|\mathbf{v}\|_V \\ &\leq 1 + \|T_i\|_{V^*} \\ &\leq 1 + \sup_i \|T_i\|_{V^*} \end{aligned}$$

Since T_i is Cauchy, $\sup_i \|T_i\|_{V^*}$ is bounded. Since this bound does not depend on $\mathbf{v}$ (and N), we get that T is bounded.

Now we want to show that $\|T_i - T\|_{V^*} \to 0$ as $n \to \infty$.

For arbitrary $\varepsilon > 0$, there is some N such that for all $i, j > N$, we have

$$\|T_i - T_j\|_{V^*} < \varepsilon.$$

In particular, for any $\mathbf{v}$ such that $\|\mathbf{v}\| \leq 1$, we have

$$|T_i(\mathbf{v}) - T_j(\mathbf{v})| < \varepsilon.$$

Taking the limit as $j \to \infty$, we obtain

$$|T_i(\mathbf{v}) - T(\mathbf{v})| \leq \varepsilon.$$

Since this is true for any $\mathbf{v}$, we have

$$\|T_i - T\|_{V^*} \leq \varepsilon.$$

for all $i > N$. So $T_i \to T$. □

1.3 Adjoint

Proposition. T^* is bounded.

Proof. We want to show that $\|T^*\|_{\mathcal{B}(Y^*,X^*)}$ is finite. For simplicity of notation, the supremum is assumed to be taken over non-zero elements of the space. We have

$$\begin{aligned}
\|T^*\|_{\mathcal{B}(Y^*,X^*)} &= \sup_{g\in Y^*} \frac{\|T^*(g)\|_{X^*}}{\|g\|_{Y^*}}\\
&= \sup_{g\in Y^*} \sup_{\mathbf{x}\in X} \frac{|T^*(g)(\mathbf{x})|/\|\mathbf{x}\|_X}{\|g\|_{Y^*}}\\
&= \sup_{g\in Y^*} \sup_{\mathbf{x}\in X} \frac{|g(T\mathbf{x})|}{\|g\|_{Y^*}\|\mathbf{x}\|_X}\\
&\leq \sup_{g\in Y^*} \sup_{\mathbf{x}\in X} \frac{\|g\|_{Y^*}\|T\mathbf{x}\|_Y}{\|g\|_{Y^*}\|\mathbf{x}\|_X}\\
&\leq \sup_{\mathbf{x}\in X} \frac{\|T\|_{\mathcal{B}(X,Y)}\|\mathbf{x}\|_X}{\|\mathbf{x}\|_X}\\
&= \|T\|_{\mathcal{B}(X,Y)}
\end{aligned}$$

So it is finite. □

1.4 The double dual

Proposition. Let $\phi : V \to V^{**}$ be defined by $\phi(\mathbf{v})(g) = g(\mathbf{v})$. Then ϕ is a bounded linear map and $\|\phi\|_{\mathcal{B}(V,V^*)} \leq 1$

Proof. Again, we are taking supremum over non-zero elements. We have

$$\begin{aligned}
\|\phi\|_{\mathcal{B}(V,V^*)} &= \sup_{\mathbf{v}\in V} \frac{\|\phi(\mathbf{v})\|_{V^{**}}}{\|\mathbf{v}\|_V}\\
&= \sup_{\mathbf{v}\in V} \sup_{g\in V^*} \frac{|\phi(\mathbf{v})(g)|}{\|\mathbf{v}\|_V\|g\|_{V^*}}\\
&= \sup_{\mathbf{v}\in V} \sup_{g\in V^*} \frac{|g(\mathbf{v})|}{\|\mathbf{v}\|_V\|g\|_{V^*}}\\
&\leq 1.
\end{aligned}$$

□

1.5 Isomorphism

1.6 Finite-dimensional normed vector spaces

Proposition. Let V be an n-dimensional vector space. Then all norms on V are equivalent to the norm $\|\cdot\|_{\ell_1^n}$.

Corollary. All norms on a finite-dimensional vector space are equivalent.

Proof. Let $\|\cdot\|$ be a norm on V.

Let $\mathbf{v} = (v_1, \cdots, v_n) = \sum v_i \mathbf{e}_i \in V$. Then we have

$$\begin{aligned}\|\mathbf{v}\| &= \left\|\sum v_i \mathbf{e}_i\right\| \\ &\leq \sum_{i=1}^n |v_i| \|\mathbf{e}_i\| \\ &\leq \left(\sup_i \|\mathbf{e}_i\|\right) \sum_{i=1}^n |v_i| \\ &\leq C \|\mathbf{v}\|_{\ell_1^n},\end{aligned}$$

where $C = \sup \|\mathbf{e}_i\| < \infty$ since we are taking a finite supremum.

For the other way round, let $S_1 = \{\mathbf{v} \in V : \|\mathbf{v}\|_{\ell_1^n} = 1\}$. We will show the two following results:

(i) $\|\cdot\| : (S_1, \|\cdot\|_{\ell_1^n}) \to \mathbb{R}$ is continuous.

(ii) S_1 is a compact set.

We first see why this gives what we want. We know that for any continuous map from a compact set to $\mathbb{R}$, the image is bounded and the infimum is achieved. So there is some $\mathbf{v}_* \in S_1$ such that

$$\|\mathbf{v}_*\| = \inf_{\mathbf{v} \in S_1} \|\mathbf{v}\|.$$

Since $\mathbf{v}_* \neq 0$, there is some c' such that $\|\mathbf{v}\| \geq c'$ for all $\mathbf{v} \in S_1$.

Now take an arbitrary non-zero $\mathbf{v} \in V$, since $\frac{\mathbf{v}}{\|\mathbf{v}\|_{\ell_1^n}} \in S_1$, we know that

$$\left\|\frac{\mathbf{v}}{\|\mathbf{v}\|_{\ell_1^n}}\right\| \geq c',$$

which is to say that

$$\|\mathbf{v}\| \geq c' \|\mathbf{v}\|_{\ell_1^n}.$$

Since we have found $c, c' > 0$ such that

$$c' \|\mathbf{v}\|_{\ell_1^n} \leq \|\mathbf{v}\| \leq c \|\mathbf{v}\|_{\ell_1^n},$$

now let $C = \max\left\{c, \frac{1}{c'}\right\} > 0$. Then

$$C^{-1} \|\mathbf{v}\|_2 \leq \|\mathbf{v}\|_1 \leq C \|\mathbf{v}\|_2.$$

So the norms are equivalent. Now we can start to prove (i) and (ii).

First, let $\mathbf{v}, \mathbf{w} \in V$. We have

$$\big|\|\mathbf{v}\| - \|\mathbf{w}\|\big| \leq \|\mathbf{v} - \mathbf{w}\| \leq C\|\mathbf{v} - \mathbf{w}\|_{\ell_1^n}.$$

Hence when $\mathbf{v}$ is close to $\mathbf{w}$ under ℓ_1^n, then $\|\mathbf{v}\|$ is close to $\|\mathbf{w}\|$. So it is continuous.

To show (ii), it suffices to show that the unit ball $B = \{\mathbf{v} \in V : \|\mathbf{v}\|_{\ell_1^n} \leq 1\}$ is compact, since S_1 is a closed subset of B. We will do so by showing it is sequentially compact.

Let $\{\mathbf{v}^{(k)}\}_{k=1}^\infty$ be a sequence in B. Write

$$\mathbf{v}^{(k)} = \sum_{i=1}^n \lambda_i^{(k)} \mathbf{e}_i.$$

Since $\mathbf{v}^{(k)} \in B$, we have

$$\sum_{i=1}^n |\lambda_i^{(k)}| \leq 1.$$

Consider the sequence $\lambda_1^{(k)}$, which is a sequence in $\mathbb{F}$.

We know that $|\lambda_1^{(k)}| \leq 1$. So by Bolzano-Weierstrass, there is a convergent subsequence $\lambda_1^{(k_{j_1})}$.

Now look at $\lambda_2^{(k_{j_1})}$. Since this is bounded, there is a convergent subsequence $\lambda_2^{(k_{j_2})}$.

Iterate this for all n to obtain a sequence k_{j_n} such that $\lambda_i^{(k_{j_n})}$ is convergent for all i. So $\mathbf{v}^{(k_{j_n})}$ is a convergent subsequence. □

Proposition. Let V be a finite-dimensional normed vector space. Then the closed unit ball

$$\bar{B}(1) = \{\mathbf{v} \in V : \|\mathbf{v}\| \leq 1\}$$

is compact.

Proof. This follows from the proof above. □

Proposition. Let V be a finite-dimensional normed vector space. Then V is a Banach space.

Proof. Let $\{\mathbf{v}_i\} \in V$ be a Cauchy sequence. Since $\{\mathbf{v}_i\}$ is Cauchy, it is bounded, i.e. $\{\mathbf{v}_i\} \subseteq \bar{B}(R)$ for some $R > 0$. By above, $\bar{B}(R)$ is compact. So $\{\mathbf{v}_i\}$ has a convergent subsequence $\mathbf{v}_{i_k} \to \mathbf{v}$. Since $\{\mathbf{v}_i\}$ is Cauchy, we must have $\mathbf{v}_i \to \mathbf{v}$. So $\mathbf{v}_i$ converges. □

Proposition. Let V, W be normed vector spaces, V be finite-dimensional. Also, let $T : V \to W$ be a linear map. Then T is bounded.

Proof. Recall discussions last time about V^* for finite-dimensional V. We will do a similar proof.

Note that since V is finite-dimensional, $\operatorname{im} T$ finite dimensional. So wlog W is finite-dimensional. Since all norms are equivalent, it suffices to consider the case where the vector spaces have ℓ_1^n and ℓ_1^m norm. This can be represented by a matrix T_{ij} such that

$$T(x_1, \cdots, x_n) = \left(\sum T_{1i} x_i, \cdots, \sum T_{mi} x_i\right).$$

We can bound this by

$$\|T(x_1, \cdots, x_n)\| \leq \sum_{j=1}^{m} \sum_{i=1}^{n} |T_{ji}||x_i| \leq m \left(\sup_{i,j} |T_{ij}|\right) \sum_{i=1}^{n} |x_i| \leq C\|\mathbf{x}\|_{\ell_1^n}$$

for some $C > 0$, since we are taking the supremum over a finite set. This implies that $\|T\|_{\mathcal{B}(\ell_1^n, \ell_1^m)} \leq C$. □

Proof. (alternative) Let $T : V \to W$ be a linear map. We define a norm on V by $\|\mathbf{v}\|' = \|\mathbf{v}\|_V + \|T\mathbf{v}\|_W$. It is easy to show that this is a norm.

Since V is finite dimensional, all norms are equivalent. So there is a constant $C > 0$ such that for all $\mathbf{v}$, we have

$$\|\mathbf{v}\|' \leq C\|\mathbf{v}\|_V.$$

In particular, we have

$$\|T\mathbf{v}\| \leq C\|\mathbf{v}\|_V.$$

So done. □

Proposition. Let V be a normed vector space. Suppose that the closed unit ball $\bar{B}(1)$ is compact. Then V is finite dimensional.

Proof. Consider the following open cover of $\bar{B}(1)$:

$$\bar{B}(1) \subseteq \bigcup_{\mathbf{y} \in \bar{B}(1)} B\left(\mathbf{y}, \frac{1}{2}\right).$$

Since $\bar{B}(1)$ is compact, this has a finite subcover. So there is some $\mathbf{y}_1, \cdots, \mathbf{y}_n$ such that

$$\bar{B}(1) \subseteq \bigcup_{i=1}^{n} B\left(\mathbf{y}_i, \frac{1}{2}\right).$$

Now let $Y = \operatorname{span}\{\mathbf{y}_1, \cdots, \mathbf{y}_n\}$, which is a finite-dimensional subspace of V. We want to show that in fact we have $Y = V$.

Clearly, by definition of Y, the unit ball

$$B(1) \subseteq Y + B\left(\frac{1}{2}\right),$$

i.e. for every $\mathbf{v} \in B(1)$, there is some $\mathbf{y} \in Y, \mathbf{w} \in B(\frac{1}{2})$ such that $\mathbf{v} = \mathbf{y} + \mathbf{w}$. Multiplying everything by $\frac{1}{2}$, we get

$$B\left(\frac{1}{2}\right) \subseteq Y + B\left(\frac{1}{4}\right).$$

Hence we also have

$$B(1) \subseteq Y + B\left(\frac{1}{4}\right).$$

By induction, for every n, we have

$$B(1) \subseteq Y + B\left(\frac{1}{2^n}\right).$$

As a consequence,
$$B(1) \subseteq \bar{Y}.$$
Since Y is finite-dimensional, we know that Y is complete. So Y is a closed subspace of V. So $\bar{Y} = Y$. So in fact
$$B(1) \subseteq Y.$$
Since every element in V can be rescaled to an element of $B(1)$, we know that $V = Y$. Hence V is finite dimensional. □

1.7 Hahn–Banach Theorem

Proposition. Let V be a real normed vector space, and $W \subseteq V$ has co-dimension 1. Assume we have the following two items:

– $p : V \to \mathbb{R}$ (not necessarily linear), which is positive homogeneous, i.e.
$$p(\lambda \mathbf{v}) = \lambda p(\mathbf{v})$$
for all $\mathbf{v} \in V, \lambda > 0$, and subadditive, i.e.
$$p(\mathbf{v}_1 + \mathbf{v}_2) \leq p(\mathbf{v}_1) + p(\mathbf{v}_2)$$
for all $\mathbf{v}_1, \mathbf{v}_2 \in V$. We can think of something like a norm, but more general.

– $f : W \to \mathbb{R}$ a linear map such that $f(\mathbf{w}) \leq p(\mathbf{w})$ for all $\mathbf{w} \in W$.

Then there exists an extension $\tilde{f} : V \to \mathbb{R}$ which is linear such that $\tilde{f}|_W = f$ and $\tilde{f}(\mathbf{v}) \leq p(\mathbf{v})$ for all $\mathbf{v} \in V$.

Proof. Let $\mathbf{v}_0 \in V \setminus W$. Since W has co-dimension 1, every element $\mathbf{v} \in V$ can be written uniquely as $\mathbf{v} = \mathbf{w} + a\mathbf{v}_0$, for some $\mathbf{w} \in W, a \in \mathbb{R}$. Therefore it suffices to define $\tilde{f}(\mathbf{v}_0)$ and then extend linearly to V.

The condition we want to meet is
$$\tilde{f}(\mathbf{w} + a\mathbf{v}_0) \leq p(\mathbf{w} + a\mathbf{v}_0) \quad (*)$$
for all $\mathbf{w} \in W, a \in \mathbb{R}$. If $a = 0$, then this is satisfied since $\tilde{f}$ restricts to f on W.

If $a > 0$ then $(*)$ is equivalent to
$$\tilde{f}(\mathbf{w}) + a\tilde{f}(\mathbf{v}_0) \leq p(\mathbf{w} + a\mathbf{v}_0).$$
We can divide by a to obtain
$$\tilde{f}(a^{-1}\mathbf{w}) + \tilde{f}(v_0) \leq p(a^{-1}\mathbf{w} + \mathbf{v}_0).$$
We let $\mathbf{w}' = a^{-1}\mathbf{w}$. So we can write this as
$$\tilde{f}(\mathbf{v}_0) \leq p(\mathbf{w}' + \mathbf{v}_0) - f(\mathbf{w}'),$$
for all $\mathbf{w}' \in W$.

If $a < 0$, then $(*)$ is equivalent to
$$\tilde{f}(\mathbf{w}) + a\tilde{f}(\mathbf{v}_0) \leq p(\mathbf{w} + a\mathbf{v}_0).$$

We now divide by a and flip the sign of the equality. So we have

$$\tilde{f}(a^{-1}\mathbf{w}) + \tilde{f}(\mathbf{v}_0) \geq -(-a^{-1})p(\mathbf{w} + a\mathbf{v}_0).$$

In other words, we want

$$\tilde{f}(\mathbf{v}_0) \geq -p(-a^{-1}\mathbf{w} - v_0) - f(a^{-1}\mathbf{w}).$$

We let $\mathbf{w}' = -a^{-1}\mathbf{w}$. Then we are left with

$$\tilde{f}(\mathbf{v}_0) \geq -p(\mathbf{w}' - v_0) + f(\mathbf{w}').$$

for all $\mathbf{w}' \in W$.

Hence we are done if we can define a $\tilde{f}(\mathbf{v}_0)$ that satisfies these two conditions. This is possible if and only if

$$-p(\mathbf{w}_1 - \mathbf{v}_0) + f(\mathbf{w}_1) \leq p(\mathbf{w}_2 + \mathbf{v}_0) - f(\mathbf{w}_2)$$

for all $\mathbf{w}_1, \mathbf{w}_2$. This holds since

$$\begin{aligned} f(\mathbf{w}_1) + f(\mathbf{w}_2) &= f(\mathbf{w}_1 + \mathbf{w}_2) \\ &\leq p(\mathbf{w}_1 + \mathbf{w}_2) \\ &= p(\mathbf{w}_1 - \mathbf{v}_0 + \mathbf{w}_2 + \mathbf{v}_0) \\ &\leq p(\mathbf{w}_1 - \mathbf{v}_0) + p(\mathbf{w}_2 + \mathbf{v}_0). \end{aligned}$$

So the result follows. □

Lemma (Zorn's lemma)**.** Let $(S, \leq)$ be a non-empty partially ordered set such that every totally-ordered subset S' has an upper bound in S. Then S has a maximal element.

Theorem (Hahn–Banach theorem*)**.** Let V be a real normed vector space, and $W \subseteq V$ a subspace. Assume we have the following two items:

- $p : V \to \mathbb{R}$ (not necessarily linear), which is positive homogeneous and subadditive;
- $f : W \to \mathbb{R}$ a linear map such that $f(\mathbf{w}) \leq p(\mathbf{w})$ for all $\mathbf{w} \in W$.

Then there exists an extension $\tilde{f} : V \to \mathbb{R}$ which is linear such that $\tilde{f}|_W = f$ and $\tilde{f}(\mathbf{v}) \leq p(\mathbf{v})$ for all $\mathbf{v} \in V$.

Proof. Let S be the set of all pairs $(\tilde{V}, \tilde{f})$ such that

(i) $W \subseteq \tilde{V} \subseteq V$

(ii) $\tilde{f} : \tilde{V} \to \mathbb{R}$ is linear

(iii) $\tilde{f}|_W = f$

(iv) $\tilde{f}(\tilde{\mathbf{v}}) \leq p(\tilde{\mathbf{v}})$ for all $\tilde{\mathbf{v}} \in V$

We introduce a partial order $\leq$ on S by $(\tilde{V}_1, \tilde{f}_1) \leq (\tilde{V}_2, \tilde{f}_2)$ if $\tilde{V}_1 \subseteq \tilde{V}_2$ and $\tilde{f}_2|_{\tilde{V}_1} = \tilde{f}_1$. It is easy to see that this is indeed a partial order.

We now check that this satisfies the assumptions of Zorn's lemma. Let $\{(\tilde{V}_\alpha, \tilde{f}_\alpha)\}_{\alpha \in A} \subseteq S$ be a totally ordered set. Define $(\tilde{V}, \tilde{f})$ by

$$\tilde{V} = \bigcup_{\alpha \in A} \tilde{V}_\alpha, \quad \tilde{f}(\mathbf{x}) = \tilde{f}_\alpha(\mathbf{x}) \text{ for } \mathbf{x} \in \tilde{V}_\alpha.$$

This is well-defined because $\{(\tilde{V}, \tilde{f}_\alpha)\}_{\alpha \in A}$ is totally ordered. So if $\mathbf{x} \in \tilde{V}_{\alpha_1}$ and $\mathbf{x} \in \tilde{V}_{\alpha_2}$, wlog assume $(\tilde{V}_{\alpha_1}, \tilde{f}_{\alpha_1}) \leq (\tilde{V}_{\alpha_2}, \tilde{f}_{\alpha_2})$. So $\tilde{f}_{\alpha_2}|_{\tilde{V}_{\alpha_2}} = \tilde{f}_{\alpha_1}$. So $\tilde{f}_{\alpha_1}(\mathbf{x}) = \tilde{f}_{\alpha_2}(\mathbf{x})$.

It should be clear that $(\tilde{V}, \tilde{f}) \in S$ and $(\tilde{V}, \tilde{f})$ is indeed an upper bound of $\{(\tilde{V}_\alpha, \tilde{f}_\alpha)\}_{\alpha \in A}$. So the conditions of Zorn's lemma are satisfied.

Hence by Zorn's lemma, there is an maximal element $(\tilde{W}, \tilde{f}) \in S$. Then by definition, $\tilde{f}$ is linear, restricts to f on W, and bounded by p. We now show that $\tilde{W} = V$.

Suppose not. Then there is some $\mathbf{v}_0 \in V \setminus \tilde{W}$. Define $\tilde{V} = \operatorname{span}\{\tilde{W}, \mathbf{v}_0\}$. Now $\tilde{W}$ is a co-dimensional 1 subspace of $\tilde{V}$. By our previous result, we know that there is some $\tilde{\tilde{f}} : \tilde{V} \to \mathbb{R}$ linear such that $\tilde{\tilde{f}}|_{\tilde{W}} = \tilde{f}$ and $\tilde{\tilde{f}}(\mathbf{v}) \leq p(\mathbf{v})$ for all $\mathbf{v} \in \tilde{V}$.

Hence we have $(\tilde{W}, \tilde{\tilde{f}}) \in S$ but $(\tilde{W}, \tilde{f}) < (\tilde{V}, \tilde{\tilde{f}})$. This contradicts the maximality of $(\tilde{W}, \tilde{f})$. □

Corollary (Hahn-Banach theorem 2.0)**.** Let $W \subseteq V$ be real normed vector spaces. Given $f \in W^*$, there exists a $\tilde{f} \in V^*$ such that $\tilde{f}|_W = f$ and $\|\tilde{f}\|_{V^*} = \|f\|_{W^*}$.

Proof. Use the Hahn-Banach theorem with $p(\mathbf{x}) = \|f\|_{W^*}\|\mathbf{x}\|_V$ for all $\mathbf{x} \in V$. Positive homogeneity and subadditivity follow directly from the axioms of the norm. Then by definition $f(\mathbf{w}) \leq p(\mathbf{w})$ for all $\mathbf{w} \in W$. So Hahn-Banach theorem says that there is $\tilde{f} : V \to \mathbb{R}$ linear such that $\tilde{f}|_W = f$ and $\tilde{f}(\mathbf{v}) \leq p(\mathbf{w}) = \|f\|_{W^*}\|\mathbf{v}\|_V$.

Now notice that

$$\tilde{f}(\mathbf{v}) \leq \|f\|_{W^*}\|\mathbf{v}\|_V, \quad -\tilde{f}(\mathbf{v}) = \tilde{f}(-\mathbf{v}) \leq \|f\|_{W^*}\|\mathbf{v}\|_V$$

implies that $|\tilde{f}(\mathbf{v})| \leq \|f\|_{W^*}\|\mathbf{v}\|_V$ for all $\mathbf{v} \in V$.

On the other hand, we have (again taking supremum over non-zero $\mathbf{v}$)

$$\|\tilde{f}\|_{V^*} = \sup_{\mathbf{v} \in V} \frac{|\tilde{f}(\mathbf{v})|}{\|\mathbf{v}\|_V} \geq \sup_{\mathbf{w} \in W} \frac{|f(\mathbf{w})|}{\|\mathbf{w}\|_W} = \|f\|_{W^*}.$$

So indeed we have $\|\tilde{f}\|_{V^*} = \|f\|_{W^*}$. □

Proposition. Let V be a real normed vector space. For every $\mathbf{v} \in V \setminus \{0\}$, there is some $f_\mathbf{v} \in V^*$ such that $f_\mathbf{v}(\mathbf{v}) = \|\mathbf{v}\|_V$ and $\|f_\mathbf{v}\|_{V^*} = 1$.

Proof. Apply Hahn-Banach theorem (2.0) with $W = \operatorname{span}\{\mathbf{v}\}$, $f'_\mathbf{v}(\mathbf{v}) = \|\mathbf{v}\|_V$. □

Corollary. Let V be a real normed vector space. Then $\mathbf{v} = \mathbf{0}$ if and only if $f(\mathbf{v}) = 0$ for all $f \in V^*$.

Corollary. Let V be a non-trivial real normed vector space, $\mathbf{v}, \mathbf{w} \in V$ with $\mathbf{v} \neq \mathbf{w}$. Then there is some $f \in V^*$ such that $f(\mathbf{v}) \neq f(\mathbf{w})$.

Corollary. If V is a non-trivial real normed vector space, then V^* is non-trivial.

Proposition. The map $\phi : V \to V^{**}$ is an isometry, i.e. $\|\phi(\mathbf{v})\|_{V^{**}} = \|\mathbf{v}\|_V$.

Proof. We have previously shown that

$$\|\phi\|_{\mathcal{B}(V,V^{**})} \leq 1.$$

It thus suffices to show that the norm is greater than 1, or that

$$\|\phi(\mathbf{v})\|_{V^{**}} \geq \|\mathbf{v}\|_V.$$

We can assume $\mathbf{v} \neq \mathbf{0}$, for which the inequality is trivial. We have

$$\|\phi(\mathbf{v})\|_{V^{**}} = \sup_{f \in V^*} \frac{|\phi(\mathbf{v})(f)|}{\|f\|_{V^*}} \geq \frac{|\phi(\mathbf{v})(f_\mathbf{v})|}{\|f_\mathbf{v}\|_{V^*}} = |f_\mathbf{v}(\mathbf{v})| = \|\mathbf{v}\|_V,$$

where $f_\mathbf{v}$ is the function such that $f_\mathbf{v}(\mathbf{v}) = \|\mathbf{v}\|_V, \|f_\mathbf{v}\|_{V^*} = 1$ as we have previously defined.

So done. □

Proposition.

$$\|T^*\|_{\mathcal{B}(W^*,V^*)} = \|T\|_{\mathcal{B}(V,W)}.$$

Proof. We have already shown that

$$\|T^*\|_{\mathcal{B}(W^*,V^*)} \leq \|T\|_{\mathcal{B}(V,W)}.$$

For the other inequality, first let $\varepsilon > 0$. Since

$$\|T\|_{\mathcal{B}(V,W)} = \sup_{\mathbf{v} \in V} \frac{\|T\mathbf{v}\|_W}{\|\mathbf{v}\|_V}$$

by definition, there is some $\mathbf{v} \in V$ such that $\|T\mathbf{v}\|_W \geq \|T\|_{\mathcal{B}(V,W)}\|\mathbf{v}\|_V - \varepsilon$. wlog, assume $\|\mathbf{v}\|_V = 1$. So

$$\|T\mathbf{v}\|_W \geq \|T\|_{\mathcal{B}(V,W)} - \varepsilon.$$

Therefore, we get that

$$\begin{aligned}
\|T^*\|_{\mathcal{B}(W^*,V^*)} &= \sup_{f \in W^*} \frac{\|T^*(f)\|_{V^*}}{\|f\|_{W^*}} \\
&\geq \|T^*(f_{T\mathbf{v}})\|_{V^*} \\
&\geq |T^*(f_{T\mathbf{v}})(\mathbf{v})| \\
&= |f_{T\mathbf{v}}(T\mathbf{v})| \\
&= \|T\mathbf{v}\|_W \\
&\geq \|T\|_{\mathcal{B}(V,W)} - \varepsilon,
\end{aligned}$$

where we used the fact that $\|f_{T\mathbf{v}}\|_{W^*}$ and $\|\mathbf{v}\|_V$ are both 1. Since ε is arbitrary, we are done. □

2 Baire category theorem

2.1 The Baire category theorem

Theorem (Baire category theorem)**.** Let X be a complete metric space. Then X is of second category.

Proof. We will prove that the intersection of a countable collection of open dense sets is non-empty. Let U_n be a countable collection of open dense set.

The key to proving this is completeness, since that is the only information we have. The idea is to construct a sequence, show that it is Cauchy, and prove that the limit is in the intersection.

Construct a sequence $x_n \in X$ and $\varepsilon_n > 0$ as follows: let x_1, ε_1 be defined such that $\overline{B(x_1, \varepsilon_1)} \subseteq U_1$. This exists U_1 is open and dense. By density, there is some $x_1 \in U_1$, and ε_1 exists by openness.

We define the x_n iteratively. Suppose we already have x_n and ε_n. Define $x_{n+1}, \varepsilon_{n+1}$ such that $\overline{B(x_{n+1}, \varepsilon_{n+1})} \subseteq \overline{B(x_n, \varepsilon_n)} \cap U_{n+1}$. Again, this is possible because U_{n+1} is open and dense. Moreover, we choose our ε_{n+1} such that $\varepsilon_{n+1} < \frac{1}{n}$ so that $\varepsilon_n \to 0$.

Since $\varepsilon_n \to 0$, we know that x_n is a Cauchy sequence. By completeness of X, we can find an $x \in X$ such that $x_n \to x$. Since x is the limit of x_n, we know that $x \in \overline{B(x_n, \varepsilon_n)}$ for all n. In particular, $x \in U_n$ for all n. So done. □

2.2 Some applications

Proposition. $\mathbb{R} \setminus \mathbb{Q} \neq \emptyset$, i.e. there is an irrational number.

Proof. Recall that we defined $\mathbb{R}$ to be the completion of $\mathbb{Q}$. So we just have to show that $\mathbb{Q}$ is not complete.

First, note that $\mathbb{Q}$ is countable. Also, for all $q \in \mathbb{Q}$, $\{q\}$ is closed and has empty interior. Hence

$$\mathbb{Q} = \bigcup_{q \in \mathbb{Q}} \{q\}$$

is the countable union of nowhere dense sets. So it is not complete by the Baire category theorem. □

Proposition. Let $\hat{\ell}_1$ be a normed vector space defined by the vector space

$$V = \{(x_1, x_2, \cdots) : x_i \in \mathbb{R}, \exists I \in \mathbb{N} \text{ such that } i > I \Rightarrow x_i = 0\},$$

with componentwise addition and scalar multiplication. This is the space of all sequences that are eventually zero.

We define the norm by

$$\|x\|_{\hat{\ell}_1} = \sum_{i=1}^{\infty} |x_i|.$$

Then $\hat{\ell}_1$ is not a Banach space.

Proof. Let

$$E_n = \{x \in \hat{\ell}_1 : x_i = 0, \forall i \geq n\}.$$

By definition,

$$\hat{\ell}_1 = \bigcup_{n=1}^{\infty} E_n.$$

We now show that E_n is nowhere dense. We first show that E_n is closed. If $x_j \to x$ in $\hat{\ell}_1$ with $x_j \in E_n$, then since x_j is 0 from the nth component onwards, x is also 0 from the nth component onwards. So we must have $x \in E_n$. So E_n is closed.

We now show that E_n has empty interior. We need to show that for all $x \in E_n$ and $\varepsilon > 0$, there is some $y \in \hat{\ell}_1$ such that $\|y - x\| < \varepsilon$ but $y \notin E_n$. This is also easy. Given $x = (x_1, \cdots, x_{n-1}, 0, 0, \cdots)$, we consider

$$y = (x_1, \cdots, x_{n-1}, \varepsilon/2, 0, 0, \cdots).$$

Then $\|y - x\|_{\hat{\ell}_1} < \varepsilon$ but $y \notin E_n$. Hence by the Baire category theorem, $\hat{\ell}_1$ is not complete. □

Proposition. There exists an $f \in C([0, 1])$ which is nowhere differentiable.

Proof. (sketch) We want to show that the set of all continuous functions which are differentiable at at least one point is contained in a meagre subset of $C([0, 1])$. Then this set cannot be all of $C([0, 1])$ since $C([0, 1])$ is complete.

Let $E_{m,n}$ be the set of all $f \in C([0, 1])$ such that

$$(\exists x)(\forall y)\ 0 < |y - x| < \frac{1}{m} \Rightarrow |f(y) - f(x)| < n|y - x|.$$

(where the quantifiers range over $[0, 1]$).

We now show that

$$\{f \in C([0, 1]) : f \text{ is differentiable somewhere}\} \subseteq \bigcup_{n,m=1}^{\infty} E_{m,n}.$$

This is easy from definition. Suppose f is differentiable at x_0. Then by definition,

$$\lim_{y \to x_0} \frac{f(y) - f(x_0)}{y - x_0} = f'(x_0).$$

Let $n \in \mathbb{N}$ be such that $|f'(x_0)| < n$. Then by definition of the limit, there is some m such that whenever $0 < |y - x| < \frac{1}{m}$, we have $\frac{|f(y) - f(x)|}{|y - x_0|} < n$. So $f \in E_{m,n}$.

Finally, we need to show that each $E_{m,n}$ is closed and has empty interior. This is left as an exercise for the reader. □

Theorem (Banach-Steinhaus theorem/uniform boundedness principle)**.** Let V be a Banach space and W be a normed vector space. Suppose T_α is a collection of bounded linear maps $T_\alpha : V \to W$ such that for each fixed $\mathbf{v} \in V$,

$$\sup_\alpha \|T_\alpha(\mathbf{v})\|_W < \infty.$$

Then

$$\sup_\alpha \|T_\alpha\|_{\mathcal{B}(V,W)} < \infty.$$

Proof. Let
$$E_n = \{\mathbf{v} \in V : \sup_\alpha \|T_\alpha(\mathbf{v})\|_W \leq n\}.$$
Then by our conditions,
$$V = \bigcup_{n=1}^\infty E_n.$$
We can write each E_n as
$$E_n = \bigcap_\alpha \{\mathbf{v} \in V : \|T_\alpha(\mathbf{v})\|_W \leq n\}.$$
Since T_α is bounded and hence continuous, so $\{\mathbf{v} \in V : \|T_\alpha(\mathbf{v})\|_W \leq n\}$ is the continuous preimage of a closed set, and is hence closed. So E_n, being the intersection of closed sets, is closed.

By the Baire category theorem, there is some n such that E_n has non-empty interior. In particular, $(\exists n)(\exists \varepsilon > 0)(\exists \mathbf{v}_0 \in V)$ such that for all $\mathbf{v} \in B(\mathbf{v}_0, \varepsilon)$, we have
$$\sup_\alpha \|T_\alpha(\mathbf{v})\|_W \leq n.$$
Now consider arbitrary $\|\mathbf{v}'\|_V \leq 1$. Then
$$\mathbf{v}_0 + \frac{\varepsilon}{2}\mathbf{v}' \in B(\mathbf{v}_0, \varepsilon).$$
So
$$\sup_\alpha \left\| T_\alpha \left(\mathbf{v}_0 + \frac{\varepsilon \mathbf{v}'}{2}\right)\right\|_W \leq n.$$
Therefore
$$\sup_\alpha \|T_\alpha \mathbf{v}'\|_W \leq \frac{2}{\varepsilon}\left(n + \sup_\alpha \|T_\alpha \mathbf{v}_0\|\right).$$
Note that the right hand side is independent of $\mathbf{v}'$. So
$$\sup_{\|\mathbf{v}'\| \leq 1} \sup_\alpha \|T_\alpha \mathbf{v}'\|_W \leq \infty. \qquad \square$$

Theorem (Osgood)**.** Let $f_n : [0, 1] \to \mathbb{R}$ be a sequence of continuous functions such that for all $x \in [0, 1]$
$$\sup_n |f_n(x)| < \infty$$
Then there are some a, b with $0 \leq a < b \leq 1$ such that
$$\sup_n \sup_{x \in [a,b]} |f_n(x)| < \infty.$$

Proof. See example sheet. $\square$

Theorem (Open mapping theorem)**.** Let V and W be Banach spaces and $T : V \to W$ be a bounded surjective linear map. Then T is an open map, i.e. $T(U)$ is an open subset of W whenever U is an open subset of V.

Proof. We can break our proof into three parts:

(i) We first want an easy way to check if a map is an open map. We want to show that T is open if and only if $T(B_V(1)) \supseteq B_W(\varepsilon)$ for some $\varepsilon > 0$. Note that one direction is easy — if T is open, then by definition $T(B_V(1))$ is open, and hence we can find the epsilon required. So we are going to prove the other direction.

(ii) We show that $\overline{T(B_V(1))} \supseteq B_W(\varepsilon)$ for some $\varepsilon > 0$

(iii) By rescaling the norm in W, we may wlog the ε obtained above is in fact 1. We then show that if $\overline{T(B_V(1))} \supseteq B_W(1)$, then $T(B_V(1)) \supseteq B_W(\frac{1}{2})$.

We now prove them one by one.

(i) Suppose $T(B_V(1)) \supseteq B_W(\varepsilon)$ for some $\varepsilon > 0$. Let $U \subseteq V$ be an open set. We want to show that $T(U)$ is open. So let $\mathbf{p} \in U, \mathbf{q} = T\mathbf{p}$.

Since U is open, there is some $\delta > 0$ such that $B_V(\mathbf{p}, \delta) \subseteq U$. We can also write the ball as $B_V(\mathbf{p}, \delta) = \mathbf{p} + B_V(\delta)$. Then we have

$$\begin{aligned} T(U) &\supseteq T(\mathbf{p} + B_V(\delta)) \\ &= T\mathbf{p} + T(B_V(\delta)) \\ &= T\mathbf{p} + \delta T(B_V(1)) \\ &\supseteq \mathbf{q} + \delta B_W(\varepsilon) \\ &= \mathbf{q} + B_W(\delta\varepsilon) \\ &= B_W(\mathbf{q}, \delta\varepsilon). \end{aligned}$$

So done.

(ii) This is the step where we use the Baire category theorem.

Since T is surjective, we can write W as

$$W = \bigcup_{n=1}^{\infty} T(B_V(n)) = \bigcup_{n=1}^{\infty} T(nB_V(1)) = \bigcup_{n=1}^{\infty} \overline{T(nB_V(1))}.$$

We have written W as a countable union of closed sets. Since W is a Banach space, by the Baire category theorem, there is some $n \geq 1$ such that $\overline{T(nB_V(1))}$ has non-empty interior. But since $\overline{T(nB_V(1))} = n\overline{T(B_V(1))}$, and multiplication by n is a homeomorphism, it follows that $\overline{T(B_V(1))}$ has non-empty interior. So there is some $\varepsilon > 0$ and $\mathbf{w}_0 \in W$ such that

$$\overline{T(B_V(1))} \supseteq B_W(\mathbf{w}_0, \varepsilon).$$

We have now found an open ball in the neighbourhood, but we want a ball centered at the origin. We will use linearity in two ways. Firstly, since if $\mathbf{v} \in B_V(1)$, then $-\mathbf{v} \in B_V(1)$. By linearly of T, we know that

$$\overline{T(B_V(1))} \supseteq B_W(-\mathbf{w}_0, \varepsilon).$$

Then by linearity, intuitively, since the image contains the balls $B_W(\mathbf{w}_0, \varepsilon)$ and $B_W(-\mathbf{w}_0, \varepsilon)$, it must contain everything in between. In particular, it must contain $B_W(\varepsilon)$.

To prove this properly, we need some additional work. This is easy if we had $T(B_V(1)) \supseteq B_W(\mathbf{w}_0, \varepsilon)$ instead of the closure of it — for any $\mathbf{w} \in B_W(\varepsilon)$, we let $\mathbf{v}_1, \mathbf{v}_2 \in B_V(1)$ be such that $T(\mathbf{v}_1) = \mathbf{w}_0 + \mathbf{w}$, $T(\mathbf{v}_2) = -\mathbf{w}_0 + \mathbf{w}$. Then $\mathbf{v} = \frac{\mathbf{v}_1 + \mathbf{v}_2}{2}$ satisfies $\|\mathbf{v}\|_V < 1$ and $T(\mathbf{v}) = \mathbf{w}$.

Since we now have the closure instead, we need to mess with sequences. Since $\overline{T(B_V(1))} \supseteq \pm\mathbf{w}_0 + B_W(\varepsilon)$, for any $\mathbf{w} \in B_W(\varepsilon)$, we can find sequences $(\mathbf{v}_i)$ and $(\mathbf{u}_i)$ such that $\|\mathbf{v}_i\|_V, \|\mathbf{u}_i\|_V < 1$ for all i and $T(\mathbf{v}_i) \to \mathbf{w}_0 + \mathbf{w}$, $T(\mathbf{u}_i) \to -\mathbf{w}_0 + \mathbf{w}$.

Now by the triangle inequality, we get

$$\left\| \frac{\mathbf{v}_i + \mathbf{u}_i}{2} \right\| < 1,$$

and we also have

$$\frac{\mathbf{v}_i + \mathbf{u}_i}{2} \to \frac{\mathbf{w}_0 + \mathbf{w}}{2} + \frac{-\mathbf{w}_0 + \mathbf{w}}{2} = \mathbf{w}.$$

So $\mathbf{w} \in \overline{T(B_V(1))}$. So $\overline{T(B_V(1))} \supseteq B_W(\varepsilon)$.

(iii) Let $\mathbf{w} \in B_W(\frac{1}{2})$. For any δ, we know

$$\overline{T(B_V(\delta))} \supseteq B_W(\delta).$$

Thus, picking $\delta = \frac{1}{2}$, we can find some $\mathbf{v}_1 \in V$ such that

$$\|\mathbf{v}_1\|_V < \frac{1}{2}, \quad \|T\mathbf{v}_1 - \mathbf{w}\| < \frac{1}{4}.$$

Suppose we have recursively found $\mathbf{v}_n$ such that

$$\|\mathbf{v}_n\|_V < \frac{1}{2^n}, \quad \|T(\mathbf{v}_1 + \cdots + \mathbf{v}_n) - \mathbf{w}\| < \frac{1}{2^{n+1}}.$$

Then picking $\delta = \frac{1}{2^{n+1}}$, we can find $\mathbf{v}_{n+1}$ satsifying the properties listed above. Then $\sum_{n=1}^\infty \mathbf{v}_n$ is Cauchy, hence convergent by completeness. Let $\mathbf{v}$ be the limit. Then

$$\|\mathbf{v}\|_V \leq \sum_{i=1}^\infty \|\mathbf{v}_i\|_V < 1.$$

Moreover, by continuity of T, we know $T\mathbf{v} = \mathbf{w}$. So we are done. □

Theorem (Inverse mapping theorem)**.** Let V, W be Banach spaces, and $T : V \to W$ be a bounded linear map which is both injective and surjective. Then T^{-1} exists and is a bounded linear map.

Proof. We know that T^{-1} as a function of sets exists. It is also easy to show that it is linear since T is linear. By the open mapping theorem, since $T(U)$ is open for all $U \subseteq V$ open. So $(T^{-1})^{-1}(U)$ is open for all $U \subseteq V$. By definition, T^{-1} is continuous. Hence T^{-1} is bounded since boundedness and continuity are equivalent. □

Theorem (Closed graph theorem)**.** Let V, W be Banach spaces, and $T : V \to W$ a linear map. If the graph of T is closed, i.e.

$$\Gamma(T) = \{(\mathbf{v}, T(\mathbf{v})) : \mathbf{v} \in V\} \subseteq V \times W$$

is a closed subset of the product space (using the norm $\|(\mathbf{v}, \mathbf{w})\|_{V \times W} = \max\{\|\mathbf{v}\|_V, \|\mathbf{w}\|_W\}$), then T is bounded.

Proof. Consider $\phi : \Gamma(T) \to V$ defined by $\phi(\mathbf{v}, T(\mathbf{v})) = \mathbf{v}$. We want to apply the inverse mapping theorem to this. To do so, we need to show a few things. First we need to show that the spaces are Banach spaces. This is easy — $\Gamma(T)$ is a Banach space since it is a closed subset of a complete space, and we are already given that V is Banach.

Now we need to show surjectivity and injectivity. This is surjective since for any $\mathbf{v} \in V$, we have $\phi(\mathbf{v}, T(\mathbf{v})) = \mathbf{v}$. It is also injective since the function T is single-valued.

Finally, we want to show ϕ is bounded. This is since

$$\|\mathbf{v}\|_V \leq \max\{\|\mathbf{v}\|, \|T(\mathbf{v})\|\} = \|(\mathbf{v}, T(\mathbf{v}))\|_{\Gamma(T)}.$$

By the inverse mapping theorem, ϕ^{-1} is bounded, i.e. there is some $C > 0$ such that

$$\max\{\|\mathbf{v}\|_V, \|T(\mathbf{v})\|\} \leq C\|\mathbf{v}\|_V$$

In particular, $\|T(\mathbf{v})\| \leq C\|\mathbf{v}\|_V$. So T is bounded. □

3 The topology of $C(K)$

3.1 Normality of compact Hausdorff spaces

Theorem. Let X be a Hausdorff space. If $C_1, C_2 \subseteq X$ are *compact* disjoint subsets, then there are some $U_1, U_2 \subseteq X$ disjoint open such that $C_1 \subseteq U_1, C_2, \subseteq U_2$.

In particular, if X is a compact Hausdorff space, then X is normal (since closed subsets of compact spaces are compact).

Proof. Since C_1 and C_2 are disjoint, by the Hausdorff property, for every $p \in C_1$ and $q \in C_2$, there is some $U_{p,q}, V_{p,q} \subseteq X$ disjoint open with $p \in U_{p,q}, q \in V_{p,q}$.

Now fix a p. Then $\bigcup_{q \in C_2} V_{p,q} \supseteq C_2$ is an open cover. Since C_2 is compact, there is a finite subcover, say

$$C_2 \subseteq \bigcup_{i=1}^{n} V_{p,q_i} \text{ for some } \{q_1, \cdots, q_n\} \subseteq C_2.$$

Note that n and q_i depends on which p we picked at the beginning.

Define

$$U_p = \bigcap_{i=1}^{n} U_{p,q_i}, \quad V_p = \bigcup_{i=1}^{n} V_{p,q_i}.$$

Since these are finite intersections and unions, U_p and V_p are open. Also, U_p and V_p are disjoint. We also know that $C_2 \subseteq V_p$.

Now note that $\bigcup_{p \in C_1} U_p \supseteq C_1$ is an open cover. By compactness of C_1, there is a finite subcover, say

$$C_1 \subseteq \bigcup_{j=1}^{m} U_{p_j} \text{ for some } \{p_1, \cdots, p_m\} \subseteq C_1.$$

Now define

$$U = \bigcup_{j=1}^{m} U_{p_j}, \quad V = \bigcap_{j=1}^{m} V_{p_j}.$$

Then U and V are disjoint open with $C_1 \subseteq U$, $C_2 \subseteq V$. So done. □

3.2 Tietze-Urysohn extension theorem

Lemma (Urysohn's lemma). Let X be normal and C_0, C_1 be disjoint closed subsets of X. Then there is a $f \in C(X)$ such that $f|_{C_0} = 0$ and $f|_{C_1} = 1$, and $0 \leq f(x) \leq 1$ for all X.

Proof. In this proof, all subsets labeled C are closed, and all subsets labeled U are open.

First note that normality is equivalent to the following: suppose $C \subseteq U \subseteq X$, where U is open and C is closed. Then there is some $\tilde{C}$ closed, $\tilde{U}$ open such that $C \subseteq \tilde{U} \subseteq \tilde{C} \subseteq U$.

We start by defining $U_1 = X \setminus C_1$. Since C_0 and C_1 are disjoint, we know that $C_0 \subseteq U_1$. By normality, there exists $C_{\frac{1}{2}}$ and $U_{\frac{1}{2}}$ such that

$$C_0 \subseteq U_{\frac{1}{2}} \subseteq C_{\frac{1}{2}} \subseteq U_1.$$

Then we can find $C_{\frac{1}{4}}, C_{\frac{3}{4}}, U_{\frac{1}{4}}, U_{\frac{3}{4}}$ such that

$$C_0 \subseteq U_{\frac{1}{4}} \subseteq C_{\frac{1}{4}} \subseteq U_{\frac{1}{2}} \subseteq C_{\frac{1}{2}} \subseteq U_{\frac{3}{4}} \subseteq C_{\frac{3}{4}} \subseteq U_1.$$

Iterating this, we get that for all dyadic rationals $q = \frac{a}{2^n}$, $a, n \in \mathbb{N}, 0 < a < 2^n$, there are some U_q open, C_q closed such that $U_q \subseteq C_q$, with $C_q \subseteq U_{q'}$ if $q < q'$.

We now define f by

$$f(x) = \inf \{q \in (0,1] \text{ dyadic rational} : x \in U_q\},$$

with the understanding that $\inf \emptyset = 1$. We now check the properties desired.

– By definition, we have $0 \leq f(x) \leq 1$.

– If $x \in C_0$, then $x \in U_q$ for all q. So $f(x) = 0$.

– If $x \in C_1$, then $x \notin U_q$ for all q. So $f(x) = 1$.

– To show f is continuous, it suffices to check that $\{x : f(x) > \alpha\}$ and $\{x : f(x) < \alpha\}$ are open for all $\alpha \in \mathbb{R}$, as this shows that the pre-images of all open intervals in $\mathbb{R}$ are open. We know that

$$\begin{aligned} f(x) < \alpha &\Leftrightarrow \inf\{q \in (0,1) \text{ dyadic rational} : x \in U_q\} < \alpha \\ &\Leftrightarrow (\exists q)\ q < \alpha \text{ and } x \in U_q \\ &\Leftrightarrow x \in \bigcup_{q<\alpha} U_q. \end{aligned}$$

Hence we have

$$\{x : f(x) < \alpha\} = \bigcup_{q<\alpha} U_q.$$

which is open, since each U_q is open for all q. Similarly we know that

$$\begin{aligned} f(x) > \alpha &\Leftrightarrow \inf\{q : x \in U_q\} > \alpha \\ &\Leftrightarrow (\exists q > \alpha)\ x \notin C_q \\ &\Leftrightarrow x \in \bigcup_{q>\alpha} X \setminus C_q. \end{aligned}$$

Since this is a union of complement of closed sets, this is open. □

Theorem (Tietze-Urysohn extension theorem)**.** Let X be a normal topological space, and $C \subseteq X$ be a closed subset. Suppose $f : C \to \mathbb{R}$ is a continuous function. Then there exists an extension $\tilde{f} : X \to \mathbb{R}$ which is continuous and satisfies $\tilde{f}|_C = f$ and $\|\tilde{f}\|_{C(X)} = \|f\|_{C(C)}$.

Proof. The idea is to repeatedly use Urysohn's lemma to get better and better approximations. We can assume wlog that $0 \leq f(x) \leq 1$ for all $x \in C$. Otherwise, we just translate and rescale our function. Moreover, we can assume that the $\sup_{x \in C} f(x) = 1$. It suffices to find $\tilde{f} : X \to \mathbb{R}$ with $\tilde{f}|_C = f$ with $0 \leq \tilde{f}(x) \leq 1$ for all $x \in X$.

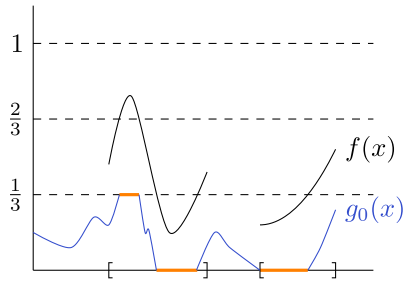
We define the sequences of continuous functions $f_i : C \to \mathbb{R}$ and $g_i : X \to \mathbb{R}$ for $i \in \mathbb{N}$. We want to think of the sum $\sum_{i=0}^n g_i$ to be the approximations, and f_{n+1} the error on C.

Let $f_0 = f$. This is the error we have when we approximate with the zero function.

We first define g_0 on a subset of X by

$$g_0(x) = \begin{cases} 0 & x \in f_0^{-1}\left(\left[0, \frac{1}{3}\right]\right) \\ \frac{1}{3} & x \in f_0^{-1}\left(\left[\frac{2}{3}, 1\right]\right) \end{cases}.$$

We can then extend this to the whole of X with $0 \leq g_0(x) \leq \frac{1}{3}$ for all x by Urysohn's lemma.



We define

$$f_1 = f_0 - g_0|_C.$$

By construction, we know that $0 \leq f_1 \leq \frac{2}{3}$. This is our first approximation. Note that we have now lowered our maximum error from 1 to $\frac{2}{3}$. We now repeat this.

Given $f_i : C \to \mathbb{R}$ with $0 \leq f_i \leq \left(\frac{2}{3}\right)^i$, we define g_i by requiring

$$g_i(x) = \begin{cases} 0 & x \in f_i^{-1}\left(\left[0, \frac{1}{3}\left(\frac{2}{3}\right)^i\right]\right) \\ \frac{1}{3}\left(\frac{2}{3}\right)^i & x \in f_i^{-1}\left(\left[\left(\frac{2}{3}\right)^{i+1}, \left(\frac{2}{3}\right)^i\right]\right) \end{cases},$$

and then extending to the whole of X with $0 \leq g_i \leq \frac{1}{3}\left(\frac{2}{3}\right)^i$ and g_i continuous. Again, this exists by Urysohn's lemma. We then define $f_{i+1} = f_i - g_i|_C$.

We then have

$$\sum_{i=0}^{n} g_i = (f_0 - f_1) + (f_1 - f_2) + \cdots + (f_n - f_{n+1}) = f - f_{n+1}.$$

We also know that

$$0 \leq f_{i+1} \leq \left(\frac{2}{3}\right)^{i+1}.$$

We conclude by letting

$$\tilde{f} = \sum_{i=0}^{\infty} g_i.$$

This exists because we have the bounds

$$0 \leq g_i \leq \frac{1}{3}\left(\frac{2}{3}\right)^i,$$

and hence $\sum_{i=0}^{n} g_i$ is Cauchy. So the limit exists and is continuous by the completeness of $C(X)$.

Now we check that

$$\sum_{i=0}^{n} g_i|_C - f = -f_{n+1}.$$

Since we know that $\|f_{n+1}\|_{C(C)} \to 0$. Therefore, we know that

$$\left.\sum_{i=0}^{\infty} g_i\right|_C = \tilde{f}|_C = f.$$

Finally, we check the bounds. We need to show that $0 \leq \tilde{f}(x) \leq 1$. This is true since $g_i \geq 0$ for all i, and also

$$|\tilde{f}(x)| \leq \sum_{i=0}^{\infty} g_i(x) \leq \sum_{i=0}^{n} \frac{1}{3}\left(\frac{2}{3}\right)^i = 1.$$

So done. □

3.3 Arzelà-Ascoli theorem

Theorem (Arzelà-Ascoli theorem)**.** Let K be a compact topological space. Then $F \subseteq C(K)$ is pre-compact, i.e. $\bar{F}$ is compact, if and only if F is bounded and equicontinuous.

Proposition. Let X be a complete metric space. Then $E \subseteq X$ is totally bounded if and only if for every sequence $\{y_i\}_{i=1}^{\infty} \subseteq E$, there is a subsequence which is Cauchy.

Corollary. Let X be a complete metric space. Then $E \subseteq X$ is totally bounded if and only if $\bar{E}$ is compact.

Theorem (Arzelà-Ascoli theorem)**.** Let K be a compact topological space. Then $F \subseteq C(K)$ is pre-compact, i.e. $\bar{F}$ is compact, if and only if F is bounded and equicontinuous.

Proof. By the previous corollary, it suffices to prove that F is totally bounded if and only if F is bounded and equicontinuous. We first do the boring direction.

($\Rightarrow$) Suppose F is totally bounded. First notice that F is obviously bounded, since F can be written as the finite union of ε-balls, which must be bounded.

Now we show F is equicontinuous. Let $\varepsilon > 0$. Since F is totally bounded, there exists a finite ε-net for F, i.e. there is some $\{f_1, \cdots, f_n\} \subseteq F$ such that for every $f \in F$, there exists an $i \in \{1, \cdots, n\}$ such that $\|f - f_i\|_{C(K)} < \varepsilon$.

Consider a point $x \in K$. Since $\{f_1, \cdots, f_n\}$ are continuous, for each i, there exists a neighbourhood U_i of x such that $|f_i(y) - f_i(x)| < \varepsilon$ for all $y \in U_i$.

Let

$$U = \bigcap_{i=1}^{n} U_i.$$

Since this is a finite intersection, U is open. Then for any $f \in F$, $y \in U$, we can find some i such that $\|f - f_i\|_{C(K)} < \varepsilon$. So

$$|f(y) - f(x)| \leq |f(y) - f_i(y)| + |f_i(y) - f_i(x)| + |f_i(x) - f(x)| < 3\varepsilon.$$

So F is equicontinuous at x. Since x was arbitrary, F is equicontinuous.

($\Leftarrow$) Suppose F is bounded and equicontinuous. Let $\varepsilon > 0$. By equicontinuity, for every $x \in K$, there is some neighbourhood U_x of x such that $|f(y) - f(x)| < \varepsilon$ for all $y \in U_x, f \in F$. Obviously, we have

$$\bigcup_{x \in K} U_x = K.$$

By the compactness of K, there are some $\{x_1, \cdots, x_n\}$ such that

$$\bigcup_{i=1}^{n} U_{x_i} \supseteq K.$$

Consider the restriction of functions in F to these points. This can be viewed as a bounded subset of ℓ_∞^n, the n-dimensional normed vector space with the supremum norm. Since this is finite-dimensional, boundedness implies total boundedness (due to, say, the compactness of the closed unit ball). In other words, there is a finite ε-net $\{f_1, \cdots, f_n\}$ such that for every $f \in F$, there is a $j \in \{1, \cdots, m\}$ such that

$$\max_i |f(x_i) - f_j(x_i)| < \varepsilon.$$

Then for every $f \in F$, pick an f_j such that the above holds. Then

$$\|f - f_j\|_{C(K)} = \sup_y |f(y) - f_j(y)|$$

Since $\{U_{x_i}\}$ covers K, we can write this as

$$\begin{aligned} &= \max_i \sup_{y \in U_{x_i}} |f(y) - f_j(y)| \\ &\leq \max_i \sup_{y \in U_{x_i}} \big(|f(y) - f(x_i)| + |f(x_i) - f_j(x_i)| + |f_j(x_i) - f_j(y)|\big) \\ &< \varepsilon + \varepsilon + \varepsilon = 3\varepsilon. \end{aligned}$$

So done. □

Proposition. Let X be a (complete) metric space. Then $E \subseteq X$ is totally bounded if and only if for every sequence $\{y_i\}_{i=1}^\infty \subseteq E$, there is a subsequence which is Cauchy.

Proof. ($\Rightarrow$) Let $E \subseteq X$ be totally bounded, $\{y_i\} \in E$. For every $j \in \mathbb{N}$, there exists a finite $\frac{1}{j}$-net, call it N_j.

Now since N_1 is finite, there is some x_1 such that there are infinitely many y_i's in $B(x_1, 1)$. Pick the first y_i in $B(x_1, 1)$ and call it y_{i_1}.

Now there is some $x_2 \in N_2$ such that there are infinitely many y_i's in $B(x_1, 1) \cap B(x_2, \frac{1}{2})$. Pick the one with smallest value of $i > i_1$, and call this y_{i_2}. Continue till infinity.

This procedure gives a sequence $x_i \in N_i$ and a subsequence $\{y_{i_k}\}$, and also

$$y_{i_n} \in \bigcap_{j=1}^{n} B\left(x_j, \frac{1}{j}\right).$$

It is easy to see that $\{y_{i_n}\}$ is Cauchy since if $m > n$, then $d(y_{i_m}, y_{i_n}) < \frac{2}{n}$.

($\Leftarrow$) Suppose E is not totally bounded. So there is no finite ε-net. Pick any y_1. Pick y_2 such that $d(y_1, y_2) \geq \varepsilon$. This exists because there is no finite ε-net.

Now given $y_1, \cdots, y_n$ such that $d(y_i, y_j) \geq \varepsilon$ for all $i, j = 1, \cdots, n$, $i \neq j$, we pick y_{n+1} such that $d(y_{n+1}, y_j) \geq \varepsilon$ for all $j = 1, \cdots, n$. Again, this exists because there is no finite ε-net. Then clearly any subsequence of $\{y_n\}$ is not Cauchy. □

Theorem (Peano*)**.** Given f continuous, then there is some $\varepsilon > 0$ such that $x' = f(x)$ with boundary condition $x(0) = x_0 \in \mathbb{R}$ has a solution in $(-\varepsilon, \varepsilon)$.

Proof. (sketch) We approximate f by a sequence of continuously differentiable functions f_n such that $\|f - f_n\|_{C(K)} \to 0$ for some $K \subseteq \mathbb{R}$. We use Picard-Lindelöf to get a solution for all n. Then we use the ODE to get estimates for the solution. Finally, we can use Arzelà-Ascoli to extract a limit as $n \to \infty$. We can then show it is indeed a solution. □

3.4 Stone–Weierstrass theorem

Theorem (Weierstrass approximation theorem)**.** The set of polynomials are dense in $C([0, 1])$.

Theorem (Stone-Weierstrass theorem)**.** Let K be compact, and $\mathcal{A} \subseteq C_{\mathbb{R}}(K)$ be a subalgebra (i.e. it is a subset that is closed under the operations) with the property that it separates points, i.e. for every $x, y \in K$ distinct, there exists some $f \in \mathcal{A}$ such that $f(x) \neq f(y)$. Then either $\bar{\mathcal{A}} = C_{\mathbb{R}}(K)$ or there is some $x_0 \in K$ such that

$$\bar{\mathcal{A}} = \{f \in C_{\mathbb{R}(K)} : f(x_0) = 0\}.$$

Lemma. Let K compact, $\mathcal{L} \subseteq C_{\mathbb{R}}(K)$ be a subset which is closed under taking maximum and minimum, i.e. if $f, g \in \mathcal{L}$, then $\max\{f, g\} \in \mathcal{L}$ and $\min\{f, g\} \in \mathcal{L}$ (with $\max\{f, g\}$ defined as $\max\{f, g\}(x) = \max\{f(x), g(x)\}$, and similarly for minimum).

Given $g \in C_{\mathbb{R}}(K)$, assume further that for any $\varepsilon > 0$ and $x, y \in K$, there exists $f_{x,y} \in \mathcal{L}$ such that

$$|f_{x,y}(x) - g(x)| < \varepsilon, \quad |f_{x,y}(y) - g(y)| < \varepsilon.$$

Then there exists some $f \in \mathcal{L}$ such that

$$\|f - g\|_{C_{\mathbb{R}}(K)} < \varepsilon,$$

i.e. $g \in \bar{\mathcal{L}}$.

Proof. Let $g \in C_{\mathbb{R}}(K)$ and $\varepsilon > 0$ be given. So for every $x, y \in K$, there is some $f_{x,y} \in \mathcal{L}$ such that

$$|f_{x,y}(x) - g(x)| < \varepsilon, \quad |f_{x,y}(y) - g(y)| < \varepsilon.$$

Claim. For each $x \in K$, there exists $f_x \in \mathcal{L}$ such that $|f_x(x) - g(x)| < \varepsilon$ and $f_x(z) < g(z) + \varepsilon$ for all $z \in K$.

Since $f_{x,y}$ is continuous, there is some $U_{x,y}$ containing y open such that

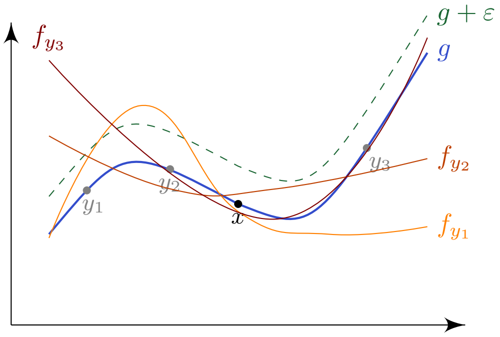
$$|f_{x,y}(z) - g(z)| < \varepsilon$$

for all $z \in U_{x,y}$. Since

$$\bigcup_{y \in K} U_{x,y} \supseteq K,$$

by compactness of K, there exists a some $y_1, \cdots, y_n$ such that

$$\bigcup_{i=1}^{n} U_{x,y_i} \supseteq K.$$



We then let

$$f_x(z) = \min\{f_{x,y_1}(z), \cdots, f_{x,y_n}(z)\}$$

for every $z \in K$. We then see that this works. Indeed, by assumption, $f_x \in \mathcal{L}$. If $z \in K$ is some arbitrary point, then $z \in U_{x,y_i}$ for some i. Then

$$f_{x,y_i}(z) < g(z) + \varepsilon.$$

Hence, since f_x is the minimum of all such f_{x,y_i}, for any z, we have

$$f_x(z) < g(z) + \varepsilon.$$

The property at x is also clear.

Claim. There exists $f \in \mathcal{L}$ such that $|f(z) - g(z)| < \varepsilon$ for all $z \in K$.

We are going to play the same game with this. By continuity of f_x, there is V_x containing x open such that

$$|f_x(w) - g(w)| < \varepsilon$$

for all $w \in V_x$. Since

$$\bigcup_{x \in K} V_x \supseteq K,$$

by compactness of K, there is some $\{x_1, \cdots, x_m\}$ such that

$$\bigcup_{j=1}^{m} V_{x_j} \supseteq K.$$

Define

$$f(z) = \max\{f_{x_1}(z), \cdots, f_{x_m}(z)\}.$$

Again, by assumption, $f \in \mathcal{L}$. Then we know that

$$f(z) > g(z) - \varepsilon.$$

We still have our first bound

$$f(z) < g(z) + \varepsilon.$$

Therefore we have

$$\|f - g\|_{C_{\mathbb{R}}(K)} < \varepsilon.$$

□

Lemma. Let $\mathcal{A} \subseteq C_{\mathbb{R}}(K)$ be a subalgebra that is a closed subset in the topology of $C_{\mathbb{R}}(K)$. Then $\mathcal{A}$ is closed under taking maximum and minimum.

Proof. First note that

$$\begin{aligned}\max\{f(x), g(x)\} &= \frac{1}{2}(f(x) + g(x)) + \frac{1}{2}|f(x) - g(x)|,\\ \min\{f(x), g(x)\} &= \frac{1}{2}(f(x) + g(x)) - \frac{1}{2}|f(x) - g(x)|.\end{aligned}$$

Since $\mathcal{A}$ is an algebra, it suffices to show that $f \in A$ implies $|f| \in A$ for every f such that $\|f\|_{C_{\mathbb{R}}(K)} \leq 1$.

The key observation is the following: consider the function $h(x) = \sqrt{x + \varepsilon^2}$. Then $h(x^2)$ approximates $|x|$. This has the property that the Taylor expansion of $h(x)$ centered at $x = \frac{1}{2}$ is uniformly convergent for $x \in [0, 1]$. Therefore there exists a polynomial $S(x)$ such that

$$|S(x) - \sqrt{x + \varepsilon^2}| < \varepsilon.$$

Now note that $S(x) - S(0)$ is a polynomial with no constant term. Therefore, since $\mathcal{A}$ is an algebra, if $f \in \mathcal{A}$, then $S(f^2) - S(0) \in \mathcal{A}$ by closure.

Now look at

$$\||f| - (S(f^2) - S(0))\|_{C_{\mathbb{R}}(K)} \leq \||f| - \sqrt{f^2 + \varepsilon^2}\| + \|\sqrt{f^2 + \varepsilon^2} - S(f^2)\| + \|S(0)\|.$$

We will make each individual term small. For the first term, note that

$$\sup_{x \in [0,1]} |x - \sqrt{x^2 + \varepsilon^2}| = \sup_{x \in [0,1]} \frac{\varepsilon^2}{|x + \sqrt{x^2 + \varepsilon^2}|} = \varepsilon.$$

So the first term is at most ε. The second term is also easy, since S is chosen such that $|S(x) - \sqrt{x + \varepsilon^2}| < 1$ for $x \in [0, 1]$, and $|f(x)^2| \leq 1$ for all $x \in [0, 1]$. So it is again bounded by ε.

By the same formula, $|S(0) - \sqrt{0 + \varepsilon^2}| < \varepsilon$. So $|S(0)| < 2\varepsilon$. So

$$\||f| - (S(f^2) - S(0))\|_{C_{\mathbb{R}}(K)} < 4\varepsilon.$$

Since $\varepsilon > 0$ and $\mathcal{A}$ is closed in the topology of $C_{\mathbb{R}}(K)$, $f \in \mathcal{A}$ and $\|f\|_{C_{\mathbb{R}}(K)} \leq 1$ implies that $|f| \in \mathcal{A}$. □

Theorem (Stone-Weierstrass theorem)**.** Let K be compact, and $\mathcal{A} \subseteq C_\mathbb{R}(K)$ be a subalgebra (i.e. it is a subset that is closed under the operations) with the property that it separates points, i.e. for every $x, y \in K$ distinct, there exists some $f \in \mathcal{A}$ such that $f(x) \neq f(y)$. Then either $\bar{\mathcal{A}} = C_\mathbb{R}(K)$ or there is some $x_0 \in K$ such that

$$\bar{\mathcal{A}} = \{f \in C_{\mathbb{R}(K)} : f(x_0) = 0\}.$$

Proof. Note that there are two possible outcomes. We will first look at the first possibility.

Consider the case where for all $x \in K$, there is some $f \in \mathcal{A}$ such that $f(x) \neq 0$. Let $g \in C_\mathbb{R}(K)$ be given. By our previous lemmas, to approximate g in $\bar{\mathcal{A}}$, we just need to show that we can approximate g at two points. So given any $\varepsilon > 0$, $x, y \in K$, we want to find $f_{x,y} \in \mathcal{A}$ such that

$$|f_{x,y}(x) - g(x)| < \varepsilon, \quad |f_{x,y}(y) - g(y)| < \varepsilon. \tag{$*$}$$

For every $x, y \in K$, $x \neq y$, we first show that there exists $h_{x,y} \in \mathcal{A}$ such that $h_{x,y}(x) \neq 0$, and $h_{x,y}(x) \neq h_{x,y}(y)$. This is easy to see. By our assumptions, we can find the following functions:

(i) There exists $h_{x,y}^{(1)}$ such that $h_{x,y}^{(1)} \neq h_{x,y}^{(1)}(y)$

(ii) There exists $h_{x,y}^{(2)}$ such that $h_{x,y}^{(2)}(x) \neq 0$.

(iii) There exists $h_{x,y}^{(3)}$ such that $h_{x,y}^{(3)}(y) \neq 0$.

Then it is an easy exercise to show that some linear combination of $h_{x,y}^{(1)}$ and $h_{x,y}^{(2)}$ and $h_{x,y}^{(3)}$ works, say $h_{x,y}$.

We will want to find our $f_{x,y}$ that satisfies $(*)$. But we will do better. We will make it equal g on x and y. The idea is to take linear combinations of $h_{x,y}$ and $h_{x,y}^2$. Instead of doing the messy algebra to show that we can find a working linear combination, just notice that $(h_{x,y}(x), h_{x,y}(y))$ and $(h_{x,y}(x)^2, h_{x,y}(y)^2)$ are linearly independent vectors in $\mathbb{R}^2$. Therefore there exists $\alpha, \beta \in \mathbb{R}$ such that

$$\alpha(h_{x,y}(x), h_{x,y}(y)) + \beta(h_{x,y}(x)^2, h_{x,y}(y)^2) = (g(x), g(y)).$$

So done.

In the other case, given $\mathcal{A}$, suppose there is $x_0 \in K$ such that $f(x_0) = 0$ for all $f \in \mathcal{A}$. Consider the algebra

$$\mathcal{A}' = \mathcal{A} + \lambda 1 = \{f + \lambda 1 : f \in \mathcal{A}, \lambda \in \mathbb{R}\}$$

Since $\mathcal{A}$ separates points, and for any $x \in K$, there is some $f \in \mathcal{A}'$ such that $f(x) \neq 0$ (e.g. $f = 1$), by the previous part, we know that $\bar{\mathcal{A}}' = C_\mathbb{R}(K)$.

Now note that

$$\bar{\mathcal{A}} \subseteq \{f \in C_\mathbb{R}(K) : f(x_0) = 0\} = B.$$

So we suffices to show that we have equality, i.e. for any $g \in B$ and $\varepsilon > 0$, there is some $f \in \mathcal{A}$ such that

$$\|f - g\|_{C_\mathbb{R}(K)} < \varepsilon.$$

Since $\bar{\mathcal{A}}' = C_{\mathbb{R}}(K)$, given such g and ε, there is some $f \in \mathcal{A}$ and $\lambda_0 \in \mathbb{R}$ such that

$$\|g - (f + \lambda 1)\|_{C_{\mathbb{R}}(K)} < \varepsilon.$$

But $g(x_0) = f(x_0) = 0$, which implies that $|\lambda| < \varepsilon$. Therefore $\|g - f\|_{C_{\mathbb{R}}(K)} < 2\varepsilon$. So done. □

Theorem (Complex version of Stone-Weierstrass theorem)**.** Let K be compact and $\mathcal{A} \subseteq C_{\mathbb{C}}(K)$ be a subalgebra over $\mathbb{C}$ which separates points and is closed under complex conjugation (i.e. if $f \in \mathcal{A}$, then $\bar{f} = \mathcal{A}$). Then either $\bar{\mathcal{A}} = C_{\mathbb{C}}(K)$ or these is an x_0 such that $\bar{\mathcal{A}} = \{f \in C_{\mathbb{C}}(K) : f(x_0) = 0\}$.

Proof. It suffices to show that either $\bar{\mathcal{A}} \supseteq C_{\mathbb{R}}(K)$ or there exists a point x_0 such that $\bar{\mathcal{A}} \supseteq \{f \in C_{\mathbb{R}}(K) : f(x_0) = 0\}$, since we can always break a complex function up into its real and imaginary parts.

Now consider

$$\mathcal{A}' = \left\{\frac{f + \bar{f}}{2} : f \in \mathcal{A}\right\} \cup \left\{\frac{f - \bar{f}}{2i} : f \in \mathcal{A}\right\}.$$

Now note that by closure of $\mathcal{A}$, we know that $\mathcal{A}'$ is a subset of $\mathcal{A}$ and is a subalgebra of $C_{\mathbb{R}}(K)$ over $\mathbb{R}$, which separates points. Hence by the real version of Stone-Weierstrass, either $\bar{\mathcal{A}}' = C_{\mathbb{R}}(K)$ or there is some x_0 such that $\bar{\mathcal{A}}' = \{f \in C_{\mathbb{R}}(K) : f(x_0) = 0\}$. So done. □

4 Hilbert spaces

4.1 Inner product spaces

Proposition. Let $f \in C(S^1)$. Then

$$\lim_{N \to \infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - S_N(f)(x)|^2 \, \mathrm{d}x = 0.$$

Proposition (Cauchy-Schwarz inequality)**.** Let $(V, \langle \cdot\,, \cdot \rangle)$ be an inner product space. Then for all $\mathbf{v}, \mathbf{w} \in V$,

$$|\langle \mathbf{v}, \mathbf{w}\rangle| \leq \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle \langle \mathbf{w}, \mathbf{w}\rangle},$$

with equality iff there is some $\lambda \in \mathbb{R}$ or $\mathbb{C}$ such that $\mathbf{v} = \lambda \mathbf{w}$ or $\mathbf{w} = \lambda \mathbf{v}$.

Proof. wlog, we can assume $\mathbf{w} \neq 0$. Otherwise, this is trivial. Moreover, assume $\langle \mathbf{v}, \mathbf{w}\rangle \in \mathbb{R}$. Otherwise, we can just multiply $\mathbf{w}$ by some $e^{i\alpha}$.

By non-negativity, we know that for all t, we have

$$\begin{aligned} 0 &\leq \langle \mathbf{v} + t\mathbf{w}, \mathbf{v} + t\mathbf{w}\rangle \\ &= \langle \mathbf{v}, \mathbf{v}\rangle + 2t\langle \mathbf{v}, \mathbf{w}\rangle + t^2\langle \mathbf{w}, \mathbf{w}\rangle. \end{aligned}$$

Therefore, the discriminant of this quadratic polynomial in t is non-positive, i.e.

$$4(\langle \mathbf{v}, \mathbf{w}\rangle)^2 - 4\langle \mathbf{v}, \mathbf{v}\rangle\langle \mathbf{w}, \mathbf{w}\rangle \leq 0,$$

from which the result follows.

Finally, note that if equality holds, then the discriminant is 0. So the quadratic has exactly one root. So there exists t such that $\mathbf{v} + t\mathbf{w} = 0$, which of course implies $\mathbf{v} = -t\mathbf{w}$. □

Proposition. Let $(V, \langle \cdot\,, \cdot \rangle)$ be an inner product space. Then

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$$

defines a norm.

Proof. The first two axioms of the norm are easy to check, since it follows directly from definition of the inner product that $\|\mathbf{v}\| \geq 0$ with equality iff $\mathbf{v} = \mathbf{0}$, and $\|\lambda \mathbf{v}\| = |\lambda|\|\mathbf{v}\|$.

The only non-trivial thing to check is the triangle inequality. We have

$$\begin{aligned} \|\mathbf{v} + \mathbf{w}\|^2 &= \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w}\rangle \\ &= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + |\langle \mathbf{v}, \mathbf{w}\rangle| + |\langle \mathbf{w}, \mathbf{v}\rangle| \\ &\leq \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\|\mathbf{v}\|\|\mathbf{w}\| \\ &= (\|\mathbf{v}\| + \|\mathbf{w}\|)^2 \end{aligned}$$

Hence we know that $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$. □

Proposition. Let $(E, \|\cdot\|)$ be a Euclidean space. Then there is a *unique* inner product $\langle \cdot\,, \cdot \rangle$ such that $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$.

Proof. The real and complex cases are slightly different.

First suppose E is a vector space over $\mathbb{R}$, and suppose also that we have an inner product $\langle \cdot, \cdot \rangle$ such that $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$. Then

$$\langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w}\rangle = \|\mathbf{v}\|^2 + 2\langle \mathbf{v}, \mathbf{w}\rangle + \|\mathbf{w}\|^2.$$

So we get

$$\langle \mathbf{v}, \mathbf{w}\rangle = \frac{1}{2}(\|\mathbf{v}+\mathbf{w}\|^2 - \|\mathbf{v}\|^2 - \|\mathbf{w}\|^2). \tag{$*$}$$

In particular, the inner product is completely determined by the norm. So this must be unique.

Now suppose E is a vector space over $\mathbb{C}$. We have

$$\langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w}\rangle = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + \langle \mathbf{v}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{v}\rangle \tag{1}$$
$$\langle \mathbf{v}-\mathbf{w}, \mathbf{v}-\mathbf{w}\rangle = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - \langle \mathbf{v}, \mathbf{w}\rangle - \langle \mathbf{w}, \mathbf{v}\rangle \tag{2}$$
$$\langle \mathbf{v}+i\mathbf{w}, \mathbf{v}+i\mathbf{w}\rangle = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - i\langle \mathbf{v}, \mathbf{w}\rangle + i\langle \mathbf{w}, \mathbf{v}\rangle \tag{3}$$
$$\langle \mathbf{v}-i\mathbf{w}, \mathbf{v}-i\mathbf{w}\rangle = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + i\langle \mathbf{v}, \mathbf{w}\rangle - i\langle \mathbf{w}, \mathbf{v}\rangle \tag{4}$$

Now consider $(1) - (2) + i(3) - i(4)$. Then we obtain

$$\|\mathbf{v}+\mathbf{w}\|^2 - \|\mathbf{v}-\mathbf{w}\|^2 + i\|\mathbf{v}+i\mathbf{w}\|^2 - i\|\mathbf{v}-i\mathbf{w}\|^2 = 4\langle \mathbf{v}, \mathbf{w}\rangle. \tag{$\dagger$}$$

So again $\langle \mathbf{v}, \mathbf{w}\rangle$ is again determined by the norm. □

Proposition (Parallelogram law)**.** Let $(E, \|\cdot\|)$ be a Euclidean space. Then for $\mathbf{v}, \mathbf{w} \in E$, we have

$$\|\mathbf{v}-\mathbf{w}\|^2 + \|\mathbf{v}+\mathbf{w}\|^2 = 2\|\mathbf{v}\|^2 + 2\|\mathbf{w}\|^2.$$

Proof. This is just simple algebraic manipulation. We have

$$\begin{aligned}\|\mathbf{v}-\mathbf{w}\|^2 + \|\mathbf{v}+\mathbf{w}\|^2 &= \langle \mathbf{v}-\mathbf{w}, \mathbf{v}-\mathbf{w}\rangle + \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w}\rangle\\ &= \langle \mathbf{v}, \mathbf{v}\rangle - \langle \mathbf{v}, \mathbf{w}\rangle - \langle \mathbf{w}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle\\ &+ \langle \mathbf{v}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle\\ &= 2\langle \mathbf{v}, \mathbf{v}\rangle + 2\langle \mathbf{w}, \mathbf{w}\rangle.\end{aligned}$$ □

Proposition (Pythagoras theorem)**.** Let $(E, \|\cdot\|)$ be a Euclidean space, and let $\mathbf{v}, \mathbf{w} \in E$ be orthogonal. Then

$$\|\mathbf{v}+\mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2.$$

Proof.

$$\begin{aligned}\|\mathbf{v}+\mathbf{w}\|^2 &= \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w}\rangle\\ &= \langle \mathbf{v}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle\\ &= \langle \mathbf{v}, \mathbf{v}\rangle + 0 + 0 + \langle \mathbf{w}, \mathbf{w}\rangle\\ &= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2.\end{aligned}$$ □

Proposition. Let $(E, \|\cdot\|)$ be a Euclidean space. Then $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{C}$ is continuous.

Proof. Let $(\mathbf{v}, \mathbf{w}) \in E \times E$, and $(\tilde{\mathbf{v}}, \tilde{\mathbf{w}}) \in E \times E$. We have

$$\begin{aligned} \|\langle \mathbf{v}, \mathbf{w}\rangle - \langle \tilde{\mathbf{v}}, \tilde{\mathbf{w}}\rangle\| &= \|\langle \mathbf{v}, \mathbf{w}\rangle - \langle \mathbf{v}, \tilde{\mathbf{w}}\rangle + \langle \mathbf{v}, \tilde{\mathbf{w}}\rangle - \langle \tilde{\mathbf{v}}, \tilde{\mathbf{w}}\rangle\| \\ &\leq \|\langle \mathbf{v}, \mathbf{w}\rangle - \langle \mathbf{v}, \tilde{\mathbf{w}}\rangle\| + \|\langle \mathbf{v}, \tilde{\mathbf{w}}\rangle - \langle \tilde{\mathbf{v}}, \tilde{\mathbf{w}}\rangle\| \\ &= \|\langle \mathbf{v}, \mathbf{w} - \tilde{\mathbf{w}}\rangle\| + \|\langle \mathbf{v} - \tilde{\mathbf{v}}, \tilde{\mathbf{w}}\rangle\| \\ &\leq \|\mathbf{v}\|\|\mathbf{w} - \tilde{\mathbf{w}}\| + \|\mathbf{v} - \tilde{\mathbf{v}}\|\|\tilde{\mathbf{w}}\| \end{aligned}$$

Hence for $\mathbf{v}, \mathbf{w}$ sufficiently closed to $\tilde{\mathbf{v}}, \tilde{\mathbf{w}}$, we can get $\|\langle \mathbf{v}, \mathbf{w}\rangle - \langle \tilde{\mathbf{v}}, \tilde{\mathbf{w}}\rangle\|$ arbitrarily small. So it is continuous. ☐

Proposition. Let $(E, \|\cdot\|)$ denote a Euclidean space, and $\bar{E}$ its completion. Then the inner product extends to an inner product on $\bar{E}$, turning $\bar{E}$ into a Hilbert space.

Proof. Recall we constructed the completion of a space as the equivalence classes of Cauchy sequences (where two Cauchy sequences (x_n) and (x'_n) are equivalent if $|x_n - x'_n| \to 0$). Let $(\mathbf{x}_n), (\mathbf{y}_n)$ be two Cauchy sequences in E, and let $\tilde{x}, \tilde{y} \in \bar{E}$ denote their equivalence classes. We define the inner product as

$$\langle \tilde{\mathbf{x}}, \tilde{\mathbf{y}}\rangle = \lim_{n\to\infty} \langle \mathbf{x}_n, \mathbf{y}_n\rangle. \tag{$*$}$$

We want to show this is well-defined. Firstly, we need to make sure the limit exists. We can show this by showing that this is a Cauchy sequence. We have

$$\begin{aligned} \|\langle \mathbf{x}_n, \mathbf{y}_n\rangle - \langle \mathbf{x}_m, \mathbf{y}_m\rangle\| &= \|\langle \mathbf{x}_n, \mathbf{y}_n\rangle - \langle \mathbf{x}_m, \mathbf{y}_n\rangle + \langle \mathbf{x}_m, \mathbf{y}_n\rangle - \langle \mathbf{x}_m, \mathbf{y}_m\rangle\| \\ &\leq \|\langle \mathbf{x}_n, \mathbf{y}_n\rangle - \langle \mathbf{x}_m, \mathbf{y}_n\rangle\| + \|\langle \mathbf{x}_m, \mathbf{y}_n\rangle - \langle \mathbf{x}_m, \mathbf{y}_m\rangle\| \\ &\leq \|\langle \mathbf{x}_n, \mathbf{x}_m, \mathbf{y}_n\rangle\| + \|\langle \mathbf{x}_m, \mathbf{y}_n - \mathbf{y}_m\rangle\| \\ &\leq \|\mathbf{x}_n - \mathbf{x}_m\|\|\mathbf{y}_n\| + \|\mathbf{x}\|\|\mathbf{y}_n - \mathbf{y}_n\| \end{aligned}$$

So $\langle \mathbf{x}_n, \mathbf{y}_n\rangle$ is a Cauchy sequence since $(\mathbf{x}_n)$ and $(\mathbf{y}_n)$ are.

We also need to show that $(*)$ does not depend on the representatives for $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$. This is left as an exercise for the reader

We also need to show that $\langle\,\cdot\,,\,\cdot\,\rangle_{\bar{E}}$ define the norm of $\|\cdot\|_{\bar{E}}$, which is yet another exercise. ☐

Proposition. Let E be a Euclidean space and $S \subseteq E$. Then $S^\perp$ is a closed subspace of E, and moreover

$$S^\perp = (\overline{\operatorname{span} S})^\perp.$$

Proof. We first show it is a subspace. Let $\mathbf{u}, \mathbf{v} \in S^\perp$ and $\lambda, \mu \in \mathbb{C}$. We want to show $\lambda\mathbf{u} + \mu\mathbf{v} \in S^\perp$. Let $\mathbf{w} \in S$. Then

$$\langle \lambda\mathbf{u} + \mu\mathbf{v}, \mathbf{w}\rangle = \lambda\langle \mathbf{u}, \mathbf{w}\rangle + \mu\langle \mathbf{v}, \mathbf{w}\rangle = 0.$$

To show it is closed, let $\mathbf{u}_n \in S^\perp$ be a sequence such that $\mathbf{u}_n \to \mathbf{u} \in E$. Let $\mathbf{w} \in S$. Then we know that

$$\langle \mathbf{u}_n, \mathbf{w}\rangle = 0.$$

Hence, by the continuity of the inner product, we have

$$0 = \lim_{n\to\infty} \langle \mathbf{u}_n, \mathbf{w}\rangle = \langle \lim \mathbf{u}_n, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle.$$

The remaining part is left as an exercise. ☐

Theorem. Let $(E, \|\cdot\|)$ be a Euclidean space, and $F \subseteq E$ a *complete* subspace. Then $F \oplus F^\perp = E$.

Hence, by definition of the direct sum, for $\mathbf{x} \in E$, we can write $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, where $\mathbf{x}_1 \in F$ and $\mathbf{x}_2 \in F^\perp$. Moreover, $\mathbf{x}_1$ is uniquely characterized by

$$\|\mathbf{x}_1 - \mathbf{x}\| = \inf_{\mathbf{y} \in F} \|\mathbf{y} - \mathbf{x}\|.$$

Proof. We already know that $F \oplus F^\perp$ is a direct sum. It thus suffices to show that the sum is the whole of E.

Let $\mathbf{y}_i \in F$ be a sequence with

$$\lim_{i \to \infty} \|\mathbf{y}_i - \mathbf{x}\| = \inf_{\mathbf{y} \in F} \|\mathbf{y} - \mathbf{x}\| = d.$$

We want to show that $\mathbf{y}$ is a Cauchy sequence. Let $\varepsilon > 0$ be given. Let $n_0 \in \mathbb{N}$ such that for all $i \geq n_0$, we have

$$\|\mathbf{y}_i - \mathbf{x}\|^2 \leq d^2 + \varepsilon.$$

We now use the parallelogram law for $\mathbf{v} = \mathbf{x} - \mathbf{y}_i$, $\mathbf{w} = \mathbf{x} - \mathbf{y}_j$ with $i, j \geq n_0$. Then the parallelogram law says:

$$\|\mathbf{v} + \mathbf{w}\|^2 + \|\mathbf{v} - \mathbf{w}\|^2 = 2\|\mathbf{v}\|^2 + 2\|\mathbf{w}\|^2,$$

or

$$\|\mathbf{y}_j - \mathbf{y}_i\|^2 + \|2\mathbf{x} - \mathbf{y}_i - \mathbf{y}_j\|^2 = 2\|\mathbf{y}_i - \mathbf{x}\|^2 + 2\|\mathbf{y}_j - \mathbf{x}\|^2.$$

Hence we know that

$$\begin{aligned} \|\mathbf{y}_i - \mathbf{y}_j\|^2 &\leq 2\|\mathbf{y}_i - \mathbf{x}\|^2 + 2\|\mathbf{y}_j - \mathbf{x}\|^2 - 4\left\|\mathbf{x} - \frac{\mathbf{y}_i + \mathbf{y}_j}{2}\right\|^2 \\ &\leq 2(d^2 + \varepsilon) + 2(d^2 + \varepsilon) - 4d^2 \\ &\leq 4\varepsilon. \end{aligned}$$

So $\mathbf{y}_i$ is a Cauchy sequence. Since F is complete, $\mathbf{y}_i \to \mathbf{y} \in F$ for some F. Moreover, by continuity, of $\|\cdot\|$, we know that

$$d = \lim_{i \to \infty} \|\mathbf{y}_i - \mathbf{x}\| = \|\mathbf{y} - \mathbf{x}\|.$$

Now let $\mathbf{x}_1 = \mathbf{y}$ and $\mathbf{x}_2 = \mathbf{x} - \mathbf{y}$. The only thing left over is to show $\mathbf{x}_2 \in F^\perp$. Suppose not. Then there is some $\tilde{\mathbf{y}} \in F$ such that

$$\langle \tilde{\mathbf{y}}, \mathbf{x}_2 \rangle \neq 0.$$

The idea is that we can perturbe $\mathbf{y}$ by a little bit to get a point even closer to $\mathbf{x}$. By multiplying $\tilde{\mathbf{y}}$ with a scalar, we can assume

$$\langle \tilde{\mathbf{y}}, \mathbf{x}_2 \rangle > 0.$$

Then for $t > 0$, we have

$$\begin{aligned} \|(\mathbf{y} + t\tilde{\mathbf{y}}) - \mathbf{x}\|^2 &= \langle \mathbf{y} + t\tilde{\mathbf{y}} - \mathbf{x}, \mathbf{y} + t\tilde{\mathbf{y}} - \mathbf{x} \rangle \\ &= \langle \mathbf{y} - \mathbf{x}, \mathbf{y} - \mathbf{x} \rangle + \langle t\tilde{\mathbf{y}}, \mathbf{y} - \mathbf{x} \rangle + \langle \mathbf{y} - \mathbf{x}, t\tilde{\mathbf{y}} \rangle + t^2 \langle \tilde{\mathbf{y}}, \tilde{\mathbf{y}} \rangle \\ &= d^2 - 2t\langle \tilde{\mathbf{y}}, \mathbf{x}_2 \rangle + t^2 \|\tilde{\mathbf{y}}\|^2. \end{aligned}$$

Hence for sufficiently small t, the t^2 term is negligible, and we can make this less that d^2. This is a contradiction since $\mathbf{y} + t\tilde{\mathbf{y}} \in F$. □

Corollary. Let E be a Euclidean space and $F \subseteq E$ a complete subspace. Then there exists a projection map $P : E \to E$ defined by $P(\mathbf{x}) = \mathbf{x_1}$, where $\mathbf{x_1} \in F$ is as defined in the theorem above. Moreover, P satisfies the following properties:

(i) $P(E) = F$ and $P(F^\perp) = \{0\}$, and $P^2 = P$. In other words, $F^\perp \leq \ker P$.

(ii) $(I - P)(E) = F^\perp$, $(I - P)(F) = \{0\}$, $(I - P)^2 = (I - P)$.

(iii) $\|P\|_{\mathcal{B}(E,E)} \leq 1$ and $\|I - P\|_{\mathcal{B}(E,E)} \leq 1$, with equality if and only if $F \neq \{0\}$ and $F^\perp \neq \{0\}$ respectively.

4.2 Riesz representation theorem

Proposition (Riesz representation theorem)**.** Let H be a Hilbert space. Then $\phi : H \to H^*$ defined by $\mathbf{v} \mapsto \langle\,\cdot\,, \mathbf{v}\rangle$ is an isometric anti-isomorphism, i.e. it is isometric, bijective and

$$\phi(\lambda\mathbf{v} + \mu\mathbf{w}) = \bar{\lambda}\phi(\mathbf{v}) + \bar{\mu}\phi(\mathbf{v}).$$

Proof. We first prove all the easy bits, namely everything but surjectivity.

– To show injectivity, if $\phi_\mathbf{v} = \phi_\mathbf{u}$, then $\langle\mathbf{w}, \mathbf{v}\rangle = \langle\mathbf{w}, \mathbf{u}\rangle$ for all $\mathbf{w}$ by definition. So $\langle\mathbf{w}, \mathbf{v} - \mathbf{u}\rangle = 0$ for all $\mathbf{w}$. In particular, $\langle\mathbf{v} - \mathbf{w}, \mathbf{v} - \mathbf{w}\rangle = 0$. So $\mathbf{v} - \mathbf{w} = 0$.

– To show that it is an anti-homomorphism, let $\mathbf{v}, \mathbf{w}, \mathbf{y} \in H$ and $\lambda, \mu \in \mathbb{F}$. Then

$$\phi_{\lambda\mathbf{v}+\mu\mathbf{w}}(\mathbf{y}) = \langle\mathbf{y}, \lambda\mathbf{v} + \mu\mathbf{w}\rangle = \bar{\lambda}\langle\mathbf{y}, \mathbf{v}\rangle + \bar{\mu}\langle\mathbf{y}, \mathbf{w}\rangle = \bar{\lambda}\phi_\mathbf{v}(\mathbf{y}) + \bar{\mu}\phi_\mathbf{w}(\mathbf{y}).$$

– To show it is isometric, let $\mathbf{v}, \mathbf{w} \in H$ and $\|\mathbf{w}\|_H = 1$. Then

$$|\phi_\mathbf{v}(\mathbf{w})| = |\langle\mathbf{w}, \mathbf{v}\rangle| \leq \|\mathbf{w}\|_H\|\mathbf{v}\|_H = \|\mathbf{v}\|_H.$$

Hence, for all $\mathbf{v}$, $\|\phi_\mathbf{v}\|_{H^*} \leq \|\mathbf{v}\|_H$ for all $\mathbf{v} \in H$. To show $\|\phi_\mathbf{v}\|_{H^*}$ is exactly $\|\mathbf{v}\|_H$, it suffices to note that

$$|\phi_\mathbf{v}(\mathbf{v})| = \langle\mathbf{v}, \mathbf{v}\rangle = \|\mathbf{v}\|_H^2.$$

So $\|\phi_\mathbf{v}\|_{H^*} \geq \|\mathbf{v}\|_H^2/\|\mathbf{v}\|_H = \|\mathbf{v}\|_H$.

Finally, we show surjectivity. Let $\xi \in H^*$. If $\xi = 0$, then $\xi = \phi_\mathbf{0}$.

Otherwise, suppose $\xi \neq 0$. The idea is that $(\ker\xi)^\perp$ is one-dimensional, and then the $\mathbf{v}$ we are looking for will be an element in this complement. So we arbitrarily pick one, and then scale it appropriately.

We now write out the argument carefully. First, we note that since ξ is continuous, $\ker\xi$ is closed, since it is the inverse image of the closed set $\{0\}$. So $\ker\xi$ is complete, and thus we have

$$H = \ker\xi \oplus (\ker\xi)^\perp.$$

The next claim is that $\dim(\ker\xi) = 1$. This is an immediate consequence of the first isomorphism theorem, whose proof is the usual one, but since we didn't prove that, we will run the argument manually.

We pick any two elements $\mathbf{v}_1, \mathbf{v}_2 \in (\ker \xi)^\perp$. Then we can always find some λ, μ not both zero such that

$$\lambda \xi(\mathbf{v}_1) + \mu \xi(\mathbf{v}_2) = 0.$$

So $\lambda \mathbf{v}_1 + \mu \mathbf{v}_2 \in \ker \xi$. But they are also in $(\ker \xi)^\perp$ by linearity. Since $\ker \xi$ and $(\ker \xi)^\perp$ have trivial intersection, we deduce that $\lambda \mathbf{v}_1 + \mu \mathbf{v}_2 = 0$. Thus, any two vectors in $(\ker \xi)^\perp$ are dependent. Since $\xi \neq 0$, we know that $\ker \xi$ has dimension 1.

Now pick any $\mathbf{v} \in (\ker \xi)^\perp$ such that $\xi(\mathbf{v}) \neq 0$. By scaling it appropriately, we can obtain a $\mathbf{v}$ such that

$$\xi(\mathbf{v}) = \langle \mathbf{v}, \mathbf{v} \rangle.$$

Finally, we show that $\xi = \phi_{\mathbf{v}}$. To prove this, let $\mathbf{w} \in H$. We decompose $\mathbf{w}$ using the previous theorem to get

$$\mathbf{w} = \alpha \mathbf{v} + \mathbf{w}_0$$

for some $\mathbf{w}_0 \in \ker \xi$ and $\alpha \in \mathbb{F}$. Note that by definition of $(\ker \xi)^\perp$, we know that $\langle \mathbf{w}_0, \mathbf{v} \rangle = 0$. Hence we know that

$$\begin{aligned} \xi(\mathbf{w}) = \xi(\alpha \mathbf{v} + \mathbf{w}_0) = \xi(\alpha \mathbf{v}) &= \alpha \xi(\mathbf{v}) \\ &= \alpha \langle \mathbf{v}, \mathbf{v} \rangle = \langle \alpha \mathbf{v}, \mathbf{v} \rangle = \langle \alpha \mathbf{v} + \mathbf{w}_0, \mathbf{v} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle. \end{aligned}$$

Since $\mathbf{w}$ was arbitrary, we are done. □

Proposition. For $f \in C(S^1)$, defined, for each $k \in \mathbb{Z}$,

$$\hat{f}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikx} f(x)\, \mathrm{d}x.$$

The partial sums are then defined as

$$S_N(f)(x) = \sum_{n=-N}^{N} e^{inx} \hat{f}(k).$$

Then we have

$$\lim_{N \to \infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - S_N(f)(x)|^2\, \mathrm{d}x = 0.$$

Proof. Consider the following Hilbert space $L^2(S^1)$ defined as the completion of $C_{\mathbb{C}}(S^1)$ under the inner product

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \bar{g}(x)\, \mathrm{d}x,$$

Consider the closed subspace

$$U_N = \operatorname{span}\{e^{inx} : |n| \leq N\}.$$

Then in fact S_N defined above by

$$S_N(f)(x) = \sum_{n=-N}^{N} e^{-inx} \hat{f}(k)$$

is the projection operator onto U_N. This is since we have the orthonormal condition

$$\langle e^{inx}, e^{-imx}\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}e^{-imx}\,\mathrm{d}x = \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases}$$

Hence it is easy to check that if $f \in U_N$, say $f = \sum_{n=-N}^{N} a_n e^{inx}$, then $S_N f = f$ since

$$S_N(f) = \sum_{n--N}^{N} \hat{f}(k)e^{-inx} = \sum_{n=-N}^{N} \langle f, e^{inx}\rangle e^{-inx} = \sum_{n=-N}^{N} a_n e^{-inx} = f$$

using the orthogonality relation. But if $f \in U_N^\perp$, then

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-inx} f(x)\,\mathrm{d}x = 0$$

for all $|n| < N$. So $S_N(f) = 0$. So this is indeed a projection map.

In particular, we will use the fact that projection maps have norms ≤ 1. Hence for any $P(x)$, we have

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |S_N(f)(x) - S_N(P)(x)|^2\,\mathrm{d}x \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - P(x)|^2\,\mathrm{d}x$$

Now consider the *algebra* $\mathcal{A}$ generated $\{e^{inx} : n \in \mathbb{Z}\}$. Notice that $\mathcal{A}$ separates points and is closed under complex conjugation. Also, for every $x \in S^1$, there exists $f \in \mathcal{A}$ such that $f(x) \neq 0$ (using, say $f(x) = e^{ix}$). Hence, by Stone-Weierstrass theorem, $\bar{\mathcal{A}} = C_{\mathbb{C}}(S^1)$, i.e. for every $f \in C_{\mathbb{C}}(S^1)$ and $\varepsilon > 0$, there exists a polynomial P of e^{ix} and e^{-ix} such that

$$\|P - f\| < \varepsilon.$$

We are almost done. We now let $N > \deg P$ be a large number. Then in particular, we have $S_N(P) = P$. Then

$$\begin{aligned}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |S_N(f) - f|^2\,\mathrm{d}x\right)^{\frac{1}{2}} &\leq \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |S_N(f) - S_N(P)|^2\,\mathrm{d}x\right)^{\frac{1}{2}} \\ &\quad + \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |S_N(P) - P|^2\,\mathrm{d}x\right)^{\frac{1}{2}} \\ &\quad + \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |P - f|^2\,\mathrm{d}x\right)^{\frac{1}{2}} \\ &\leq \varepsilon + 0 + \varepsilon \\ &= 2\varepsilon.\end{aligned}$$

So done. □

4.3 Orthonormal systems and basis

Proposition. Let H be a Hilbert space. Let S be a maximal orthonormal system. Then $\overline{\operatorname{span} S} = H$.

Proof. Recall that $S^\perp = (\overline{\operatorname{span} S})^\perp$. Since H is a Hilbert space, we have

$$H = \overline{\operatorname{span} S} \oplus (\overline{\operatorname{span} S})^\perp = \overline{\operatorname{span} S} \oplus S^\perp.$$

Since S is maximal, $S^\perp = \{0\}$. So done. □

Proposition. Let E be Euclidean, and let S be an orthonormal system. If $\overline{\operatorname{span} S} = E$, then S is maximal.

Proof.

$$S^\perp = (\overline{\operatorname{span} S})^\perp = E^\perp = \{0\}.$$ □

Proposition. Let $\{\mathbf{x}_i\}_{i=1}^n$, $n \in \mathbb{N}$ be linearly independent. Then there exists $\{\mathbf{e}_i\}_{i=1}^n$ such that $\{\mathbf{e}_i\}_{i=1}^n$ is an orthonormal system and

$$\operatorname{span}\{\mathbf{x}_1, \cdots, \mathbf{x}_j\} = \operatorname{span}\{\mathbf{e}_1, \cdots, \mathbf{e}_j\}$$

for all $j \leq n$.

Proof. Define $\mathbf{e}_1$ by

$$\mathbf{e}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|}.$$

Assume we have defined $\{\mathbf{e}_i\}_{i=1}^j$ orthonormal such that

$$\operatorname{span}\{\mathbf{x}_1, \cdots, \mathbf{x}_j\} = \operatorname{span}\{\mathbf{e}_1, \cdots, \mathbf{e}_j\}.$$

Then by linear independence, we know that

$$\mathbf{x}_{j+1} \notin \operatorname{span}\{\mathbf{x}_1, \cdots, \mathbf{x}_j\} = \operatorname{span}\{\mathbf{e}_1, \cdots, \mathbf{e}_j\} = F_j.$$

We now define

$$\tilde{\mathbf{x}}_{j+1} = \mathbf{x}_{j+1} - P_{F_j}(\mathbf{x}_{j+1}),$$

where P_{F_j} is the projection onto F_j given by

$$P_{F_j} = \sum_{i=1}^j \langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i.$$

Since F_j is a closed, finite subspace, we know that

$$\mathbf{x}_{j+1} - P_{F_j} \mathbf{x}_{j+1} \perp F_j.$$

Thus

$$\mathbf{e}_{j+1} = \frac{\tilde{\mathbf{x}}_{j+1}}{\|\tilde{\mathbf{x}}_{j+1}\|}$$

is the right choice. We can also write this in full as

$$\mathbf{e}_{j+1} = \frac{\mathbf{x}_{j+1} - \sum_{i=1}^j \langle \mathbf{x}_j \mathbf{e}_j \rangle \mathbf{e}_i}{\|\mathbf{x}_{j+1} - \sum_{i=1}^j \langle \mathbf{x}_j \mathbf{e}_j \rangle \mathbf{e}_i\|}.$$

So done. □

Proposition. Let H be separable, i.e. there is an infinite set $\{\mathbf{y}_i\}_{i \in \mathbb{N}}$ such that

$$\overline{\operatorname{span}\{\mathbf{y}_i\}} = H.$$

Then there exists a countable basis for $\operatorname{span}\{\mathbf{y}_i\}$.

Proof. We find a subset $\{\mathbf{y}_{i_j}\}$ such that $\operatorname{span}\{\mathbf{y}_i\} = \operatorname{span}\{\mathbf{y}_{i_j}\}$ and $\{\mathbf{y}_{i_j}\}$ are independent. This is easy to do since we can just throw away the useless dependent stuff. At this point, we do Gram-Schmidt, and done. □

4.4 The isomorphism with ℓ_2

Lemma (Bessel's inequality)**.** Let E be Euclidean and $\{\mathbf{e}_i\}_{i=1}^N$ with $N \in \mathbb{N} \cup \{\infty\}$ an orthonormal system. For any $\mathbf{x} \in E$, define $x_i = \langle \mathbf{x}, \mathbf{e}_i\rangle$. Then for any $j \leq N$, we have

$$\sum_{i=1}^j |x_i|^2 \leq \|\mathbf{x}\|^2.$$

Proof. Consider the case where j is finite first. Define

$$F_j = \operatorname{span}\{\mathbf{e}_1, \cdots, \mathbf{e}_j\}.$$

This is a finite dimensional subspace of E. Hence an orthogonal projection P_{F_j} exists. Moreover, we have an explicit formula for this:

$$P_{F_j} = \sum_{i=1}^j \langle \mathbf{x}, \mathbf{e}_i\rangle \mathbf{e}_i.$$

Thus

$$\sum_{i=1}^j |x_i|^2 = \|P_{F_j}\mathbf{x}\|^2 \leq \|\mathbf{x}\|^2$$

since we know that $\|P_{F_j}\| \leq 1$. Taking the limit as $j \to \infty$ proves the case for infinite j. □

Proposition. Let H be a separable Hilbert space, with a countable basis $\{\mathbf{e}_i\}_{i=1}^N$, where $N \in \mathbb{N} \cup \{\infty\}$. Let $\mathbf{x}, \mathbf{y} \in H$ and

$$x_i = \langle \mathbf{x}, \mathbf{e}_i\rangle, \quad y_i = \langle \mathbf{y}, \mathbf{e}_i\rangle.$$

Then

$$\mathbf{x} = \sum_{i=1}^N x_i \mathbf{e}_i, \quad \mathbf{y} = \sum_{i=1}^N y_i \mathbf{e}_i,$$

and

$$\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{i=1}^N x_i \bar{y}_i.$$

Moreover, the sum converges absolutely.

Proof. We only need to consider the case $N = \infty$. Otherwise, it is just finite-dimensional linear algebra.

First, note that our expression is written as an infinite sum. So we need to make sure it converges. We define the partial sums to be

$$\mathbf{s}_n = \sum_{i=1}^n x_i \mathbf{e}_i.$$

We want to show $\mathbf{s}_n \to \mathbf{x}$. By Bessel's inequality, we know that

$$\sum_{i=1}^\infty |x_i|^2 \leq \|\mathbf{x}\|^2.$$

In particular, the sum is bounded, and hence converges.

For any $m < n$, we have

$$\|\mathbf{s}_n - \mathbf{s}_m\| = \sum_{i=m+1}^{n} |x_i|^2 \leq \sum_{i=m+1}^{\infty} |x_i|^2.$$

As $m \to \infty$, the series must go to 0. Thus $\{\mathbf{s}_n\}$ is Cauchy. Since H is Hilbert, $\mathbf{s}_n$ converges, say

$$\mathbf{s}_n \to \mathbf{s} = \sum_{i=1}^{\infty} x_i \mathbf{e}_i.$$

Now we want to prove that this sum is indeed $\mathbf{x}$ itself. Note that so far in the proof, we have *not* used the fact that $\{\mathbf{e}_i\}$ is a basis. We just used the fact that it is orthogonal. Hence we should use this now. We notice that

$$\langle \mathbf{s}, \mathbf{e}_i \rangle = \lim_{n\to\infty} \langle \mathbf{s}_n, \mathbf{e}_i \rangle = \lim_{n\to\infty} \sum_{j=1}^{n} x_j \langle \mathbf{e}_j, \mathbf{e}_i \rangle = x_i.$$

Hence we know that

$$\langle \mathbf{x} - \mathbf{s}, \mathbf{e}_i \rangle = 0.$$

for all i. So $\mathbf{x} - \mathbf{s}$ is perpendicular to all $\mathbf{e}_i$. Since $\{\mathbf{e}_i\}$ is a basis, we must have $\mathbf{x} - \mathbf{s} = 0$, i.e. $\mathbf{x} = \mathbf{s}$.

To show our formula for the inner product, we can compute

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y} \rangle &= \lim_{n\to\infty} \left\langle \sum_{i=1}^{n} x_i \mathbf{e}_i, \sum_{j=1}^{n} y_j \mathbf{e}_j \right\rangle \\
&= \lim_{n\to\infty} \sum_{i,j=1}^{n} x_i \bar{y}_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle \\
&= \lim_{n\to\infty} \sum_{i,j=1}^{n} x_i \bar{y}_j \delta_{ij} \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} x_i \bar{y}_i \\
&= \sum_{i=1}^{\infty} x_i \bar{y}_i.
\end{aligned}$$

Note that we *know* the limit exists, since the continuity of the inner product ensures the first line is always valid.

Finally, to show absolute convergence, note that for all finite j, we have

$$\sum_{i=1}^{j} |x_i \bar{y}_i| \leq \sqrt{\sum_{i=1}^{n} |x_i|^2} \sqrt{\sum_{i=1}^{n} |y_i|^2} \leq \|\mathbf{x}\| \|\mathbf{y}\|.$$

Since this is a uniform bound for any j, the sum converges absolutely. □

Proposition. Let H be a separable Hilbert space with orthonormal basis $\{\mathbf{e}_i\}_{i\in\mathbb{N}}$. Let $\{a_i\}_{i\in\mathbb{N}} \in \ell_2(\mathbb{C})$. Then there exists an $\mathbf{x} \in H$ with $\langle \mathbf{x}, \mathbf{e}_i\rangle = a_i$. Moreover, this $\mathbf{x}$ is exactly

$$\mathbf{x} = \sum_{i=1}^{\infty} x_i \mathbf{e}_i.$$

Proof. The only thing we need to show is that this sum converges. For any $n \in \mathbb{N}$, define

$$\mathbf{s}_n = \sum_{i=1}^{n} a_i \mathbf{e}_i \in H.$$

For $m < n$, we have

$$\|\mathbf{s}_n - \mathbf{s}_m\|^2 = \sum_{m+1}^{n} |a_i|^2 \to 0$$

as $m \to \infty$ because $\{a_i\} \in \ell^2$. Hence $\mathbf{s}_n$ is Cauchy and as such converges to $\mathbf{x}$. Obviously, we have

$$\langle \mathbf{x}, \mathbf{e}_i\rangle = \lim_{n\to\infty} \sum_{j=1}^{n} a_j \langle \mathbf{e}_j, \mathbf{e}_i\rangle = a_i.$$

So done. □

4.5 Operators

Theorem. Let X be a Banach space, $T \in B(X)$. Then $\sigma(T)$ is a non-empty, closed subset of

$$\{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|_{\mathcal{B}(X)}\}.$$

Lemma. Let X be a Banach space, $T \in \mathcal{B}(X)$ and $\|T\|_{\mathcal{B}(X)} < 1$. Then $I - T$ is invertible.

Proof. To prove it is invertible, we construct an explicit inverse. We want to show

$$(I - T)^{-1} = \sum_{i=0}^{\infty} T^i.$$

First, we check the right hand side is absolutely convergent. This is since

$$\sum_{i=0}^{\infty} \|T^i\|_{\mathcal{B}(X)} \leq \sum_{i=0}^{\infty} \|T\|^i_{\mathcal{B}(X)} \leq \frac{1}{1 - \|T\|_{\mathcal{B}(X)}} < \infty.$$

Since X is Banach, and hence $\mathcal{B}(X)$ is Banach, the limit is well-defined. Now it is easy to check that

$$\begin{aligned}(I - T)\sum_{i=1}^{\infty} T^i &= (I - T)(I + T + T^2 + \cdots)\\ &= I + (T - T) + (T^2 - T^2) + \cdots\\ &= I.\end{aligned}$$

Similarly, we have

$$\left(\sum_{i=1}^{\infty} T^i\right)(I - T) = I.$$ □

Lemma. Let X be a Banach space, $S_1 \in \mathcal{B}(X)$ be invertible. Then for all $S_2 \in \mathcal{B}(X)$ such that

$$\|S_1^{-1}\|_{\mathcal{B}(X)}\|S_1 - S_2\|_{\mathcal{B}(X)} < 1,$$

S_2 is invertible.

Proof. We can write

$$S_2 = S_1(I - S_1^{-1}(S_1 - S_2)).$$

Since

$$\|S_1^{-1}(S_1 - S_2)\|_{\mathcal{B}(X)} \leq \|S_1^{-1}\|_{\mathcal{B}(X)}\|S_1 - S_2\|_{\mathcal{B}(X)} < 1$$

by assumption, by the previous lemma, $(I - S_1^{-1}(S_1 - S_2))^{-1}$ exists. Therefore the inverse of S_2 is

$$S_2^{-1} = (I - S_1^{-1}(S_1 - S_2))^{-1}S_1^{-1}. \qquad \square$$

Theorem. Let X be a Banach space, $T \in \mathcal{B}(X)$. Then $\sigma(T)$ is a non-empty, closed subset of

$$\{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|_{\mathcal{B}(X)}\}.$$

Proof. We first prove the closedness of the spectrum. It suffices to prove that the resolvent set $\rho(T) = \mathbb{C} \setminus \sigma(T)$ is open, by the definition of closedness.

Let $\lambda \in \rho(T)$. By definition, $S_1 = T - \lambda I$ is invertible. Define $S_2 = T - \mu I$. Then

$$\|S_1 - S_2\|_{\mathcal{B}(X)} = \|(T - \lambda I) - (T - \mu I)\|_{\mathcal{B}(X)} = |\lambda - \mu|.$$

Hence if $|\lambda - \mu|$ is sufficiently small, then $T - \mu I$ is invertible by the above lemma. Hence $\mu \in \rho(T)$. So $\rho(T)$ is open.

To show $\sigma(T) \subseteq \{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|_{\mathcal{B}(X)}\}$ is equivalent to showing

$$\{\lambda \in \mathbb{C} : |\lambda| > \|T\|_{\mathcal{B}(X)}\} \subseteq \mathbb{C} \setminus \sigma(T) = \rho(T).$$

Suppose $|\lambda| > \|T\|$. Then $I - \lambda^{-1}T$ is invertible since

$$\|\lambda^{-1}T\|_{\mathcal{B}(X)} = \lambda^{-1}\|T\|_{\mathcal{B}(X)} < 1.$$

Therefore, $(I - \lambda^{-1}T)^{-1}$ exists, and hence

$$(\lambda I - T)^{-1} = \lambda^{-1}(I - \lambda T)^{-1}$$

is well-defined. Therefore $\lambda I - T$, and hence $T - \lambda I$ is invertible. So $\lambda \in \rho(T)$.

Finally, we need to show it is non-empty. How did we prove it in the case of finite-dimensional vector spaces? In that case, it ultimately boiled down to the fundamental theorem of algebra. And how did we prove the fundamental theorem of algebra? We said that if $p(x)$ is a polynomial with no roots, then $\frac{1}{p(x)}$ is bounded and entire, hence constant.

We are going to do the same proof. We look at $\frac{1}{T - \lambda I}$ as a function of λ. If $\sigma(T) = \emptyset$, then this is an everywhere well-defined function. We show that this is entire and bounded, and hence by "Liouville's theorem", it must be constant, which is impossible (in the finite-dimensional case, we would have inserted a det there).

So suppose $\sigma(T) = \emptyset$, and consider the function $R : \mathbb{C} \to \mathcal{B}(X)$, given by $R(\lambda) = (T - \lambda I)^{-1}$.

We first show this is entire. This, by definition, means R is given by a power series near any point $\lambda_0 \in \mathbb{C}$. Fix such a point. Then as before, we can expand

$$\begin{aligned} T - \lambda I &= (T - \lambda_0 I)\Big[I - (T - \lambda_0 I)^{-1}\Big((T - \lambda_0 I) - (T - \lambda I)\Big)\Big] \\ &= (T - \lambda_0 I)\Big[I - (\lambda - \lambda_0)(T - \lambda_0 I)^{-1}\Big]. \end{aligned}$$

Then for $(\lambda - \lambda_0)$ small, we have

$$\begin{aligned} (T - \lambda I)^{-1} &= \left(\sum_{i=0}^{\infty} (\lambda - \lambda_0)^i (T - \lambda_0 I)^{-i}\right) (T - \lambda_0 I)^{-1} \\ &= \sum_{i=0}^{\infty} (\lambda - \lambda_0)^i (T - \lambda_0 I)^{-i-1}. \end{aligned}$$

So this is indeed given by an absolutely convergent power series near λ_0.

Next, we show R is bounded, i.e.

$$\sup_{\lambda \in \mathbb{C}} \|R(\lambda)\|_{\mathcal{B}(X)} < \infty.$$

It suffices to prove this for λ large. Note that we have

$$(T - \lambda I)^{-1} = \lambda^{-1}(\lambda^{-1} T - I)^{-1} = -\lambda^{-1} \sum_{i=0}^{\infty} \lambda^{-i} T^i.$$

Hence we get

$$\begin{aligned} \|(\lambda I - T)^{-1}\|_{\mathcal{B}(X)} &\leq |\lambda|^{-1} \sum_{i=0}^{\infty} |\lambda|^{-i} \|T^i\|_{\mathcal{B}(X)} \\ &\leq |\lambda|^{-1} \sum_{i=0}^{\infty} \left(|\lambda|^{-1} \|T\|_{\mathcal{B}(X)}\right)^i \\ &\leq \frac{1}{|\lambda| - \|T\|_{\mathcal{B}(X)}}, \end{aligned}$$

which tends to 0 as $|\lambda| \to \infty$. So it is bounded.

By "Liouville's theorem", $R(\lambda)$ is constant, which is clearly a contradiction since $R(\lambda) \neq R(\mu)$ for $\lambda \neq \mu$. □

Proposition (Liouville's theorem for Banach space-valued analytic function)**.** Let X be a Banach space, and $F : \mathbb{C} \to X$ be entire (in the sense that F is given by an absolutely convergent power series in some neighbourhood of any point) and norm bounded, i.e.

$$\sup_{z \in \mathbb{C}} \|F(z)\|_X < \infty.$$

Then F is constant.

Proof. Let $f \in X^*$. Then we show $f \circ F : \mathbb{C} \to \mathbb{C}$ is bounded and entire. To see it is bounded, just note that f is a bounded linear map. So

$$\sup_{z \in \mathbb{C}} |f \circ F(z)| \leq \sup_{z \in \mathbb{C}} \|f\|_{X^*} \|F(z)\|_X < \infty.$$

Analyticity can be shown in a similar fashion, exploiting the fact that f^* is linear.

Hence Liouville's theorem implies $f \circ F$ is constant, i.e. $(f \circ F)(z) = (f \circ F)(0)$. In particular, this implies $f(F(z) - F(0)) = 0$. Moreover, this is true for all $f \in X^*$. Hence by (corollary of) Hahn-Banach theorem, we know $F(z) - F(0) = 0$ for all $z \in \mathbb{C}$. Therefore F is constant. □

Theorem. We have

$$\sigma_{ap}(T) \supseteq \partial\sigma(T),$$

where $\partial\sigma(T)$ is the boundary of $\sigma(T)$ in the topology of $\mathbb{C}$. In particular, $\sigma_{ap}(T) \neq \emptyset$.

Proof. Let $\lambda \in \partial\sigma(T)$. Pick sequence $\{\lambda_n\}_{n=1}^\infty \subseteq \rho(T) = \mathbb{C} \setminus \sigma(T)$ such that $\lambda_n \to \lambda$. We claim that $R(\lambda_n) = (T - \lambda_n I)^{-1}$ satisfies

$$\|R(\lambda_n)\|_{\mathcal{B}(X)} \to \infty.$$

If this were the case, then we can pick $\mathbf{y}_n \in X$ such that $\|\mathbf{y}_n\| \to 0$ and $\|R(\lambda_n)(\mathbf{y}_n)\| = 1$. Setting $\mathbf{x}_n = R(\lambda_n)(\mathbf{y}_n)$, we have

$$\begin{aligned}
\|(T - \lambda I)\mathbf{x}_n\| &\leq \|(T - \lambda_n I)\mathbf{x}_n\|_X + \|(\lambda - \lambda_n)\mathbf{x}_n\|_X\\
&= \|(T - \lambda_n I)(T - \lambda_n I)^{-1}\mathbf{y}_n\|_X + \|(\lambda - \lambda_n)\mathbf{x}_n\|\\
&= \|\mathbf{y}_n\|_X + |\lambda - \lambda_n|\\
&\to 0.
\end{aligned}$$

So $\lambda \in \sigma_{ap}(T)$.

Thus, it remains to prove that $\|R(\lambda_n)\|_{\mathcal{B}(X)} \to \infty$. Recall from last time if S_1 is invertible, and

$$\|S_1^{-1}\|_{\mathcal{B}(X)} \|S_1 - S_2\|_{\mathcal{B}(X)} \leq 1, \tag{$*$}$$

then S_2 is invertible. Thus, for any $\mu \in \sigma(T)$, we have

$$\|R(\lambda_n)\|_{\mathcal{B}(X)} |\mu - \lambda_n| = \|R(\lambda_n)\|_{\mathcal{B}(X)} \|(T - \lambda_n I) - (T - \mu I)\|_{B(X)} \geq 1.$$

Thus, it follows that

$$\|R(\lambda_n)\|_{\mathcal{B}(X)} \geq \frac{1}{\inf\{|\mu - \lambda_n| : \mu \in \sigma(T)\}} \to \infty.$$

So we are done. □

Proposition. Let X, Y be Banach spaces. Then $T \in \mathcal{L}(X, Y)$ is compact if and only if $T(B(1))$ is totally bounded if and only if $\overline{T(B(1))}$ is compact.

Proposition. Let X be a Banach space. Then $\mathcal{B}_0(X)$ is a closed subspace of $\mathcal{B}(X)$. Moreover, if $T \in \mathcal{B}_0(X)$ and $S \in \mathcal{B}(X)$, then $TS, ST \in \mathcal{B}_0(X)$.

In a more algebraic language, this means $\mathcal{B}_0(X)$ is a closed ideal of the algebra $\mathcal{B}(X)$.

Proof. There are three things to prove. First, it is obvious that $\mathcal{B}_0(X)$ is a subspace. To check it is closed, suppose $\{T_n\}_{n=1}^\infty \subseteq \mathcal{B}_0(X)$ and $\|T_n - T\|_{\mathcal{B}(X)} \to 0$. We need to show $T \in \mathcal{B}_0(X)$, i.e. $T(B(1))$ is totally bounded.

Let $\varepsilon > 0$. Then there exists N such that

$$\|T - T_n\|_{\mathcal{B}(X)} < \varepsilon$$

whenever $n \geq N$. Take such an n. Then $T_n(B(1))$ is totally bounded. So there exists $\mathbf{x}_1, \cdots, \mathbf{x}_k \in B(1)$ such that $\{T_n \mathbf{x}_i\}_{i=1}^k$ is an ε-net for $T_n(B(1))$. We now claim that $\{T\mathbf{x}_i\}_{i=1}^k$ is an 3ε-net for $T(B(1))$.

This is easy to show. Let $\mathbf{x} \in X$ be such that $\|\mathbf{x}\| \leq 1$. Then by the triangle inequality,

$$\begin{aligned} \|T\mathbf{x} - T\mathbf{x}_i\|_X &\leq \|T\mathbf{x} - T_n\mathbf{x}\| + \|T_n\mathbf{x} - T_n\mathbf{x}_i\| + \|T_n\mathbf{x}_i - T\mathbf{x}_i\| \\ &\leq \varepsilon + \|T_n\mathbf{x} - T_n\mathbf{x}_i\|_X + \varepsilon \\ &= 2\varepsilon + \|T_n\mathbf{x} - T_n\mathbf{x}_i\|_X \end{aligned}$$

Now since $\{T_n\mathbf{x}_i\}$ is an ε-net for $T_n(B(1))$, there is some i such that $\|T_n\mathbf{x} - T_n\mathbf{x}_i\| < \varepsilon$. So this gives

$$\|T\mathbf{x} - T\mathbf{x}_i\|_X \leq 3\varepsilon.$$

Finally, let $T \in \mathcal{B}_0(X)$ and $S \in \mathcal{B}(X)$. Let $\{\mathbf{x}_n\} \subseteq X$ such that $\|\mathbf{x}_n\|_X \leq 1$. Since T is compact, i.e. $\overline{T(B(1))}$ is compact, there exists a convergence subsequence of $\{T\mathbf{x}_i\}$.

Since S is bounded, it maps a convergent sequence to a convergent sequence. So $\{ST\mathbf{x}_n\}$ also has a convergent subsequence. So $\overline{ST(B(1))}$ is compact. So ST is compact.

We also have to show that $TS(B(1))$ is totally bounded. Since S is bounded, $S(B(1))$ is bounded. Since T sends a bounded set to a totally bounded set, it follows that $TS(B(1))$ is totally bounded. So TS is compact. □

Theorem. Let X be an infinite-dimensional Banach space, and $T \in \mathcal{B}(X)$ be a compact operator. Then $\sigma_p(T) = \{\lambda_i\}$ is at most countable. If $\sigma_p(T)$ is infinite, then $\lambda_i \to 0$.

The spectrum is given by $\sigma(T) = \sigma_p(T) \cup \{0\}$. Moreover, for every non-zero $\lambda_i \in \sigma_p(T)$, the eigenspace has finite dimensions.

Proposition. Let H be a Hilbert space, and $T \in \mathcal{B}_0(H)$ a compact operator. Let $a > 0$. Then there are only finitely many linearly independent eigenvectors whose eigenvalue have magnitude $\geq a$.

Proof. Suppose not. There there are infinitely many independent $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots$ such that $T\mathbf{x}_i = \lambda_i \mathbf{x}_i$ with $|\lambda_i| \geq a$.

Define $X_n = \operatorname{span}\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$. Since the $\mathbf{x}_i$'s are linearly independent, there exists $\mathbf{y}_n \in X_n \cap X_{n-1}^\perp$ with $\|\mathbf{y}_n\|_H = 1$.

Now let

$$\mathbf{z}_n = \frac{\mathbf{y}_n}{\lambda_n}.$$

Note that

$$\|\mathbf{z}_n\|_H \leq \frac{1}{a}.$$

Since X_n is spanned by the eigenvectors, we know that T maps X_n into itself. So we have

$$T\mathbf{z}_n \in X_n.$$

Moreover, we claim that $T\mathbf{z}_n - \mathbf{y}_n \in X_{n-1}$. We can check this directly. Let

$$\mathbf{y}_n = \sum_{k=1}^{n} c_k \mathbf{x}_k.$$

Then we have

$$\begin{aligned}
T\mathbf{z}_n - \mathbf{y}_n &= \frac{1}{\lambda_n} T\left(\sum_{k=1}^{n} c_k \mathbf{x}_k\right) - \sum_{k=1}^{n} c_k \mathbf{x}_k \\
&= \sum_{k=1}^{n} c_k \left(\frac{\lambda_k}{\lambda_n} - 1\right) \mathbf{x}_k \\
&= \sum_{k=1}^{n-1} c_k \left(\frac{\lambda_k}{\lambda_n} - 1\right) \mathbf{x}_k \in X_{n-1}.
\end{aligned}$$

We next claim that $\|T\mathbf{z}_n - T\mathbf{z}_m\|_H \geq 1$ whenever $n > m$. If this holds, then T is not compact, since $T\mathbf{z}_n$ does not have a convergent subsequence.

To show this, wlog, assume $n > m$. We have

$$\|T\mathbf{z}_n - T\mathbf{z}_m\|_H^2 = \|(T\mathbf{z}_n - \mathbf{y}_n) - (T\mathbf{z}_m - \mathbf{y}_n)\|_H^2$$

Note that $T\mathbf{z}_n - \mathbf{y}_n \in X_{n-1}$, and since $m < n$, we also have $T\mathbf{z}_m \in X_{n-1}$. By construction, $\mathbf{y}_n \perp X_{n-1}$. So by Pythagorean theorem, we have

$$\begin{aligned}
&= \|T\mathbf{z}_n - \mathbf{y}_n - T\mathbf{z}_m\|_H^2 + \|\mathbf{y}_n\|_H^2 \\
&\geq \|\mathbf{y}_n\|^2 \\
&= 1
\end{aligned}$$

So done. □

Lemma. Let H be a Hilbert space, and $T \in \mathcal{B}(H)$ compact. Then $\operatorname{im}(I - T)$ is closed.

Proof. We let S be the orthogonal complement of $\ker(I - T)$, which is a closed subspace, hence a Hilbert space. We shall consider the restriction $(I - T)|_S$, which has the same image as $I - T$.

To show that $\operatorname{im}(I - T)$ is closed, it suffices to show that $(I - T)|_S$ is bounded *below*, i.e. there is some $C > 0$ such that

$$\|\mathbf{x}\|_H \leq C\|(I - T)\mathbf{x}\|_H$$

for all $\mathbf{x} \in S$. If this were the case, then if $(I - T)\mathbf{x}_n \to \mathbf{y}$ in H, then

$$\|\mathbf{x}_n - \mathbf{x}_m\| \leq C\|(I - T)(\mathbf{x}_n - \mathbf{x}_m)\| \to 0,$$

and so $\{\mathbf{x}_n\}$ is a Cauchy sequence. Write $\mathbf{x}_n \to \mathbf{x}$. Then by continuity, $(I - T)\mathbf{x} = \mathbf{y}$, and so $\mathbf{y} \in \operatorname{im}(I - T)$.

Thus, suppose $(I - T)$ is not bounded below. Pick $\mathbf{x}_n$ such that $\|\mathbf{x}_n\|_H = 1$, but $(I - T)\mathbf{x}_n \to 0$. Since T is compact, we know $T\mathbf{x}_n$ has a convergent subsequence. We may wlog $T\mathbf{x}_n \to \mathbf{y}$. Then since $\|T\mathbf{x}_n - \mathbf{x}_n\|_H \to 0$, it follows that we also have $\mathbf{x}_n \to \mathbf{y}$. In particular, $\|\mathbf{y}\| = 1 \neq 0$, and $\mathbf{y} \in S$.

But $\mathbf{x}_n \to \mathbf{y}$ also implies $T\mathbf{x}_n \to T\mathbf{y}$. So this means we must have $T\mathbf{y} = \mathbf{y}$. But this is a contradiction, since $\mathbf{y}$ does not lie in $\ker(I - T)$. □

Proposition. Let H be a Hilbert space, $T \in \mathcal{B}(H)$ compact. If $\lambda \neq 0$ and $\lambda \in \sigma(T)$, then $\lambda \in \sigma_p(T)$.

Proof. We will prove if $\lambda \neq 0$ and $\lambda \not\in \sigma_p(T)$, then $\lambda \not\in \sigma(T)$. In other words, let $\lambda \neq 0$ and $\ker(T - \lambda I) = \{0\}$. We will show that $T - \lambda I$ is surjective, i.e. $\mathrm{im}(T - \lambda I) = H$.

Suppose this is not the case. Denote $H_0 = H$ and $H_1 = \mathrm{im}(T - \lambda I)$. We know that H_1 is closed and is hence a Hilbert space. Moreover, $H_1 \subsetneq H_0$ by assumption.

We now define the sequence $\{H_n\}$ recursively by

$$H_n = (T - \lambda I)H_{n-1}.$$

We claim that $H_n \subsetneq H_{n-1}$. This must be the case, because the map $(T - \lambda I)^n : H_0 \to H_n$ is an isomorphism (it is injective and surjective). So the inclusion $H_n \subseteq H_{n-1}$ is isomorphic to the inclusion $H_1 \subseteq H_0$, which is strict.

Thus we have a strictly decreasing sequence

$$H_0 \supsetneq H_1 \supsetneq H_2 \supsetneq \cdots$$

Let $\mathbf{y}_n$ be such that $\mathbf{y}_n \in H_n$, $\mathbf{y}_n \perp H_{n+1}$ and $\|\mathbf{y}_n\|_H = 1$. We now claim $\|T\mathbf{y}_n - T\mathbf{y}_m\| \geq |\lambda|$ if $n \neq m$. This then contradicts the compactness of T. To show this, again wlog we can assume that $n > m$. Then we have

$$\begin{aligned}\|T\mathbf{y}_n - T\mathbf{y}_m\|_H^2 &= \|(T\mathbf{y}_n - \lambda\mathbf{y}_n) - (T\mathbf{y}_m - \lambda\mathbf{y}_m) - \lambda\mathbf{y}_m + \lambda\mathbf{y}_n\|^2\\ &= \|(T - \lambda I)\mathbf{y}_n - (T - \lambda I)\mathbf{y}_m - \lambda\mathbf{y}_m + \lambda\mathbf{y}_n\|_H^2\end{aligned}$$

Now note that $(T - \lambda I)\mathbf{y}_n \in H_{n+1} \subseteq H_{m+1}$, while $(T - \lambda I)\mathbf{y}_m$ and $\lambda\mathbf{y}_n$ are both in H_{m+1}. So $\lambda\mathbf{y}_m$ is perpendicular to all of them, and Pythagorean theorem tells

$$\begin{aligned}&= |\lambda|^2\|\mathbf{y}_m\|^2 + \|(T - \lambda I)\mathbf{y}_m - (T - \lambda I)\mathbf{y}_m - \lambda\mathbf{y}_m\|^2\\ &\geq |\lambda|^2\|\mathbf{y}_m\|^2\\ &= |\lambda|^2.\end{aligned}$$

This contradicts the compactness of T. Therefore $\mathrm{im}(T - \lambda I) = H$. □

Theorem. Let H be an infinite-dimensional Hilbert space, and $T \in \mathcal{B}(H)$ be a compact operator. Then $\sigma_p(T) = \{\lambda_i\}$ is at most countable. If $\sigma_p(T)$ is infinite, then $\lambda_i \to 0$.

The spectrum is given by $\sigma(T) = \sigma_p(T) \cup 0$. Moreover, for every non-zero $\lambda_i \in \sigma_p(T)$, the eigenspace has finite dimensions.

Proof. As mentioned, it remains to show that $\sigma(T) = \sigma_p(T) \cup \{0\}$. The previous proposition tells us $\sigma(T) \setminus \{0\} \subseteq \sigma_p(T)$. So it only remains to show that $0 \in \sigma(T)$.

There are two possible cases. The first is if $\{\lambda_i\}$ is infinite. We have already shown that $\lambda_i \to 0$. So $0 \in \sigma(T)$ by the closedness of the spectrum.

Otherwise, if $\{\lambda_i\}$ is finite, let $E_{\lambda_1}, \cdots, E_{\lambda_n}$ be the eigenspaces. Define

$$H' = \operatorname{span}\{E_{\lambda_1}, \cdots, E_{\lambda_n}\}^\perp.$$

This is non-empty, since each E_{λ_i} is finite-dimensional, but H is infinite dimensional. Then T restricts to $T|_{H'} : H' \to H'$.

Now $T|_{H'}$ has no non-zero eigenvalues. By the previous discussion, we know $\sigma(T|_{H'}) \subseteq \{0\}$. By non-emptiness of $\sigma(T|_{H'})$, we know $0 \in \sigma(T|_{H'}) \subseteq \sigma(T)$.

So done. □

4.6 Self-adjoint operators

Theorem (Spectral theorem)**.** Let H be an infinite dimensional Hilbert space and $T : H \to H$ a compact self-adjoint operator.

(i) $\sigma_p(T) = \{\lambda_i\}_{i=1}^N$ is at most countable.

(ii) $\sigma_p(T) \subseteq \mathbb{R}$.

(iii) $\sigma(T) = \{0\} \cup \sigma_p(T)$.

(iv) If E_{λ_i} are the eigenspaces, then $\dim E_{\lambda_i}$ is finite if $\lambda_i \neq 0$.

(v) $E_{\lambda_i} \perp E_{\lambda_j}$ if $\lambda_i \neq \lambda_j$.

(vi) If $\{\lambda_i\}$ is infinite, then $\lambda_i \to 0$.

(vii)
$$T = \sum_{i=1}^N \lambda_i P_{E_{\lambda_i}}.$$

Proposition. Let H be a Hilbert space and $T \in \mathcal{B}(H)$ self-adjoint. Then $\sigma_p(T) \subseteq \mathbb{R}$.

Proof. Let $\lambda \in \sigma_p(T)$ and $\mathbf{v} \in \ker(T - \lambda I) \setminus \{0\}$. Then by definition of $\mathbf{v}$, we have

$$\lambda = \frac{\langle T\mathbf{v}, \mathbf{v}\rangle}{\|\mathbf{v}\|_H^2} = \frac{\langle \mathbf{v}, T\mathbf{v}\rangle}{\|\mathbf{v}\|_H^2} = \bar{\lambda}.$$

So $\lambda \in \mathbb{R}$. □

Proposition. Let H be a Hilbert space and $T \in \mathcal{B}(H)$ self-adjoint. If $\lambda, \mu \in \sigma_p(T)$ and $\lambda \neq \mu$, then $E_\lambda \perp E_\mu$.

Proof. Let $\mathbf{v} \in \ker(T - \lambda I) \setminus \{0\}$ and $\mathbf{w} \in \ker(T - \mu I) \setminus \{0\}$. Then

$$\lambda\langle \mathbf{v}, \mathbf{w}\rangle = \langle T\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, T\mathbf{w}\rangle = \bar{\mu}\langle \mathbf{v}, \mathbf{w}\rangle = \mu\langle \mathbf{v}, \mathbf{w}\rangle,$$

using the fact that eigenvalues are real. Since $\lambda \neq \mu$ by assumption, we must have $\langle \mathbf{v}, \mathbf{w}\rangle = 0$. □

Proposition. Let H be a Hilbert space and $T \in \mathcal{B}(H)$ a *compact* self-adjoint operator. If $T \neq 0$, then T has a non-zero eigenvalue.

Lemma. Let H be a Hilbert space, and $T \in \mathcal{B}(H)$ a compact self-adjoint operator. Then

$$\|T\|_{\mathcal{B}(H)} = \sup_{\|\mathbf{x}\|_H = 1} |\langle \mathbf{x}, T\mathbf{x}\rangle|$$

Proof. Write

$$\lambda = \sup_{\|\mathbf{x}\|_H = 1} |\langle \mathbf{x}, T\mathbf{x}\rangle|.$$

Note that one direction is easy, since for all $\mathbf{x}$, Cauchy-Schwarz gives

$$|\langle \mathbf{x}, T\mathbf{x}\rangle| \leq \|T\mathbf{x}\|_H \|\mathbf{x}\|_H = \|T\|_{\mathcal{B}(H)} \|\mathbf{x}\|_H^2.$$

So it suffices to show the inequality in the other direction. We now claim that

$$\|T\|_{\mathcal{B}(H)} = \sup_{\|\mathbf{x}\|_H = 1, \|\mathbf{y}\|_H = 1} |\langle T\mathbf{x}, \mathbf{y}\rangle|.$$

To show this, recall that $\phi : H \to H^*$ defined by $\mathbf{v} \mapsto \langle \,\cdot\,, \mathbf{v}\rangle$ is an isometry. By definition, we have

$$\|T\|_{\mathcal{B}(H)} = \sup_{\|\mathbf{x}\|_H = 1} \|T\mathbf{x}\|_H = \sup_{\|\mathbf{x}\|_H = 1} \|\phi(T\mathbf{x})\|_{H^*} = \sup_{\|\mathbf{x}\|_H = 1} \sup_{\|\mathbf{y}\|_H = 1} |\langle \mathbf{y}, T\mathbf{x}\rangle|.$$

Hence, it suffices to show that

$$\sup_{\|\mathbf{x}\|_H = 1, \|\mathbf{y}\|_H = 1} |\langle T\mathbf{x}, \mathbf{y}\rangle| \leq \lambda.$$

Take $\mathbf{x}, \mathbf{y} \in H$ such that $\|\mathbf{x}\|_H = \|\mathbf{y}\|_H = 1$. We first perform a trick similar to the polarization identity. First, by multiplying $\mathbf{y}$ by an appropriate scalar, we can wlog assume $\langle T\mathbf{x}, \mathbf{y}\rangle$ is real. Then we have

$$\begin{aligned} |\langle T(\mathbf{x} + \mathbf{y}), \mathbf{x} + \mathbf{y}\rangle - \langle T(\mathbf{x} - \mathbf{y}), \mathbf{x} - \mathbf{y}\rangle| &= 2|\langle T\mathbf{x}, \mathbf{y}\rangle + \langle T\mathbf{y}, \mathbf{x}\rangle| \\ &= 4|\langle T\mathbf{x}, \mathbf{y}\rangle|. \end{aligned}$$

Hence we have

$$\begin{aligned} |\langle T\mathbf{x}, \mathbf{y}\rangle| &= \frac{1}{4}|\langle T(\mathbf{x} + \mathbf{y}), \mathbf{x} + \mathbf{y}\rangle - \langle T(\mathbf{x} - \mathbf{y}), \mathbf{x} - \mathbf{y}\rangle| \\ &\leq \frac{1}{4}(\lambda\|\mathbf{x} + \mathbf{y}\|_H^2 + \lambda\|\mathbf{x} - \mathbf{y}\|_H^2) \\ &= \frac{\lambda}{4}(2\|\mathbf{x}\|_H^2 + 2\|\mathbf{y}\|_H^2) \\ &= \lambda, \end{aligned}$$

where we used the parallelogram law. So we have $\|T\|_{\mathcal{B}(H)} \leq \lambda$. □

Proposition. Let H be a Hilbert space and $T \in \mathcal{B}(H)$ a *compact* self-adjoint operator. If $T \neq 0$, then T has a non-zero eigenvalue.

Proof. Since $T \neq 0$, then $\|T\|_{\mathcal{B}(H)} \neq 0$. Let $\|T\|_{\mathcal{B}(H)} = \lambda$. We now claim that either λ or $-\lambda$ is an eigenvalue of T.

By the previous lemma, there exists a sequence $\{\mathbf{x}_n\}_{n=1}^\infty \subseteq H$ such that $\|\mathbf{x}_n\|_H = 1$ and $\langle \mathbf{x}_n, T\mathbf{x}_n\rangle \to \pm\lambda$.

We consider the two cases separately. Suppose $\langle \mathbf{x}_n, T\mathbf{x}_n\rangle \to \lambda$. Consider $T\mathbf{x}_n - \lambda\mathbf{x}_n$. Since T is compact, there exists a subsequence such that $T\mathbf{x}_{n_k} \to \mathbf{y}$ for some $\mathbf{y} \in H$. For simplicity of notation, we assume $T\mathbf{x}_n \to \mathbf{y}$ itself. We have

$$\begin{aligned}0 &\leq \|T\mathbf{x}_n - \lambda\mathbf{x}_n\|_H^2\\ &= \langle T\mathbf{x}_n - \lambda\mathbf{x}_n, T\mathbf{x}_n - \lambda\mathbf{x}_n\rangle\\ &= \|T\mathbf{x}_n\|_H^2 - 2\lambda\langle T\mathbf{x}_n, \mathbf{x}_n\rangle + \lambda^2\|\mathbf{x}_n\|^2\\ &\to \lambda^2 - 2\lambda^2 + \lambda^2\\ &= 0\end{aligned}$$

as $n \to \infty$. Note that we implicitly used the fact that $\langle T\mathbf{x}_n, \mathbf{x}_n\rangle = \langle \mathbf{x}_n, T\mathbf{x}_n\rangle$ since $\langle T\mathbf{x}_n, \mathbf{x}_n\rangle$ is real. So we must have

$$\|T\mathbf{x}_n - \lambda\mathbf{x}_n\|_H^2 \to 0.$$

In other words,

$$\mathbf{x}_n \to \frac{1}{\lambda}\mathbf{y}.$$

Finally, we show $\mathbf{y}$ is an eigenvector. This is easy, since

$$T\mathbf{y} = \lim_{n\to\infty} T(\lambda\mathbf{x}_n) = \lambda\mathbf{y}.$$

The case where $\mathbf{x}_n \to -\lambda$ is entirely analogous. In this case, $-\lambda$ is an eigenvalue. The proof is exactly the same, apart form some switching of signs. □

Proposition. Let H be an infinite dimensional Hilbert space and $T : H \to H$ a compact self-adjoint operator. Then

$$T = \sum_{i=1}^N \lambda_i P_{E_{\lambda_i}}.$$

Proof. Let

$$U = \operatorname{span}\{E_{\lambda_1}, E_{\lambda_2}, \cdots\}.$$

Firstly, we clearly have

$$T|_U = \sum_{i=1}^N \lambda_i P_{E_{\lambda_i}}.$$

This is since for any $\mathbf{x} \in U$ can be written as

$$\mathbf{x} = \sum_{i=1}^n P_{E_{\lambda_i}}\mathbf{x}.$$

Less trivially, this is also true for $\bar{U}$, i.e.

$$T|_{\bar{U}} = \sum_{i=1}^N \lambda_i P_{E_{\lambda_i}},$$

but this is also clear from definition once we stare at it hard enough.

We also know that

$$H = \bar{U} \oplus U^{\perp}.$$

It thus suffices to show that

$$T|_{U^{\perp}} = 0.$$

But since $T|_{U^{\perp}}$ has no non-zero eigenvalues, this follows from our previous proposition. So done. □