Linear Analysis

Michael Li, Leo Lai

March 29, 2017

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

1 Normed spaces and bounded linear operators

Let X be a real or complex vector space. A *norm* on X is a function $\|\cdot\| : X \to \mathbb{R}$ such that the following three conditions hold

(i) $\forall x,\ \|x\| \geq 0$ and $\|x\| = 0 \iff x = 0$ (Positivity)
(ii) $\forall x \in X$, λ scalar, $\|\lambda x\| = |\lambda|\|x\|$ (Homogeneity)
(iii) $\forall x, y \in X,\ \|x + y\| \leq \|x\| + \|y\|$ (Triangle inequality)

The pair $(X, \|\cdot\|)$ is a *normed space*.

Example.

(1) $l_2^n = (\mathbb{R}^n, \|\cdot\|_2)$ or $(\mathbb{C}^n, \|\cdot\|_2)$, where $\|x\|_2 = \sum_{i=1}^n |x_i|^2$.

(2) $l_1^n = (\mathbb{R}^n, \|\cdot\|_1)$, where $\|x\|_1 = \sum_{i=1}^n |x_i|$.

(3) $l_\infty^n = (\mathbb{R}^n, \|\cdot\|_2 = \infty)$, where $\|x\|_\infty = \max_{1\leq i\leq n} |x_i|$.

As a convention, from now on the scalar field will be $\mathbb{R}$ unless otherwise stated.

A norm induces a metric $d(x, y) = \|x - y\|$, which then induces a topology called the *norm topology*. In the norm topology, the following algebraic operations are continuous:

$$+ : X \times X \to X, \quad \cdot : \mathbb{R} \times X \to X, \quad \|\cdot\| : X \to \mathbb{R}$$

This follows from $x_n \to x, y_n \to y \;\Rightarrow\; x_n + y_n \to x + y$, $\lambda_n \to \lambda, x_n \to x \;\Rightarrow\; \lambda_n x_n \to \lambda x$, and $\big|\|x\| - \|y\|\big| \leq \|x - y\|$. Note that the last fact shows that taking norm is Lipschitz with constant 1.

A *Banach space* is a complete normed space. For example, l_2^n, l_1^n, l_∞^n are Banach spaces since convergence in every one of them is coordinate-wise convergence.

It is often useful to consider the *unit ball* B_X and the *unit sphere* S_X for a normed space X, where

$$B_X = \{x \in X : \|x\| \leq 1\}, \qquad S_X = \{x \in X : \|x\| = 1\}$$

In l_2^2 and l_∞^2, B_X are the unit disk and the square $[-1, 1]^2$ respectively.

Remark.

(i) B_X completely determines the norm by $\|x\| = \inf\{t > 0 : x \in tB_X\}$.

(ii) B_X is *convex* ($x, y \in B_X,\ t \in [0, 1] \;\Rightarrow\; (1 - t)x + ty \in B_X$), *symmetric* ($x \in B_X \;\Rightarrow\; -x \in B_X$), and closed (since it is the pre-image of $[0, 1]$ under the norm map). In the case of complex scalars, the symmetric property is replaced by *balanced*, i.e. $x \in B_X,\ |\lambda| = 1 \;\Rightarrow\; \lambda x \in B_X$).

Example.

(4) $l_p^n = (\mathbb{R}^n, \|\cdot\|_p)$, $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ for $1 < p < \infty$. The triangle inequality is known as Minkowski's inequality and will be proven later.

More generally, given $B \subseteq \mathbb{R}^n$. If B is bounded, closed, convex, symmetric, and a neighborhood of 0, then there is a norm on $\mathbb{R}^n$ with unit ball B.

Note that in this course, subspace means linear subspaces.

1.1 The Inequalities of Hölder and Minkowski

A function $f : (0, \infty) \to \mathbb{R}$ is *convex* if for all $x, y \in (0, \infty), t \in [0, 1]$,

$$f((1-t)x + ty) \le (1-t)f(x) + tf(y)$$

and *concave* if $\le$ is replaced with $\ge$.

Lemma. Let $1 < p < \infty$, then the function $x \mapsto x^p$, $(0, \infty) \to \mathbb{R}$ is convex.

Proof. Fix $x > 0$ and $t \in [0, 1]$. Define

$$g(y) = \big((1-t)x + ty)\big)^p - \big((1-t)x^p + ty^p\big)$$

for $y > 0$. Then $g(x) = 0$ and $g'(y) = pt\big((1-t)x + ty\big)^{p-1} - pty^{p-1}$. If $y > x$, then $g'(y) \le 0$. If $y < x$, then $g'(y) \ge 0$. Therefore, $g(y) \le 0$ for all $y > 0$, as required. □

Theorem (Minkowski's Inequality)**.** Let $1 \le p \le \infty$, $x, y \in l_p$, then $x + y \in l_p$, and $\|x + y\|_p \le \|x\|_p + \|y\|_p$.

Proof. WLOG $p \ne 1, \infty$, since they are clear. First assume $\|x\|_p, \|y\|_p \le 1$. We show that for any $t \in [0, 1]$, $(1-t)x + ty \in l_p$, and $\|(1-t)x + ty\|_p \le 1$. To see this, write $x = (x_n),\ y = (y_n)$. For $n \in \mathbb{N}$, we have

$$|(1-t)x_n + ty_n|^p \le \big((1-t)|x_n| + t|y_n|\big)^p \le (1-t)|x_n|^p + t|y_n|^p$$

by lemma 1.1. For $N \in \mathbb{N}$, sum n from 1 to N,

$$\sum_{n=1}^{N} |(1-t)x_n + ty_n|^p \le (1-t)\sum_{n=1}^{N} |x_n|^p + t\sum_{n=1}^{N} |y_n|^p \le (1-t)\|x\|_p^p + t\|y\|_p^p \le 1$$

Letting $N \to \infty$, we obtain the claim.

In general, given $x, y \in l_p$. WLOG they are both non-zero. By the claim,

$$\frac{\|x\|_p}{\|x\|_p + \|y\|_p} \cdot \frac{x}{\|x\|_p} + \frac{\|y\|_p}{\|x\|_p + \|y\|_p} \cdot \frac{y}{\|y\|_p} \in l_p \text{ and has norm } \le 1$$

This rearranges to $x + y \in l_p$, and $\|x + y\|_p \le \|x\|_p + \|y\|_p$. □

Remark. We now know that l_p^n and l_p are normed spaces for all $1 \le p \le \infty$.

Note that if $x = (x_n) \in l_1$, and $y = (y_n) \in l_\infty$, then $|x_n y_n| \le |x_n| \|y\|_\infty$, so $(x_n y_n) \in l_1$, and $\sum |x_n y_n| \le \|x\|_1 \|y\|_\infty$. Hölder's inequality generalizes this observation. Given $p \in (1, \infty)$, the *conjugate index* of p is the unique q such that $\frac{1}{p} + \frac{1}{q} = 1$. E.g. 2 is the conjugate index of 2, 1 is the conjugate index to ∞.

Lemma. Given $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$, we have $ab \le \frac{a^p}{p} + \frac{b^q}{q}$ for all $a, b \ge 0$.

Proof. We may assume $a, b > 0$. The function $\log : (0, \infty) \to \mathbb{R}$ is concave (proof similar to lemma 1.1). Therefore,

$$\log\left(\frac{a^p}{p} + \frac{b^q}{q}\right) \ge \frac{1}{p}\log a^p + \frac{1}{q}\log b^q = \log(ab)$$

The result follows by rearrangement. □

Theorem (Hölder's Inequality)**.** Let $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then for $x = (x_n) \in l_p,\ y = (y_n) \in l_q$, we have $(x_n y_n) \in l_1$, and

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \|x\|_p \cdot \|y\|_q$$

Proof. Assume that $x, y \neq 0$. By homogeneity, further assume that $\|x\|_p = \|y\|_q = 1$. By lemma 1.1, $|x_n y_n| \leq \frac{1}{p}|x_n|^p + \frac{1}{q}|y_n|^q$ for all $n \in \mathbb{N}$. Hence

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \frac{1}{p} \sum_{n=1}^{\infty} |x_n|^p + \frac{1}{q} \sum_{n=1}^{\infty} |y_n|^q = \frac{1}{p}\|x\|_p^p + \frac{1}{q}\|y\|_q^q = 1 = \|x\|_p \cdot \|y\|_q$$

□

Remark. For the case $p = q = 2$, this reduces to the Cauchy-Schwarz inequality.

1.2 Norm topology

Recall that for a subset A of a topological space X, $\bar{A}$ denotes the closure of A. A is *dense* in X if $\bar{A} = X$. X is *separable* if it has a countable dense subset. E.g. $\bar{\mathbb{Q}} = \mathbb{R}$, so $\mathbb{R}$ is separable.

If Y is a subspace of a normed space X, then $\bar{Y}$ is also a subspace: given $x, y \in \bar{Y}$, scalars λ, μ, there exists $(x_n), (y_n) \in Y$ with $x_n \to x,\ y_n \to y$, so $\lambda x + \mu y = \lim_{n\to\infty}(\lambda x_n + \mu y_n) \in \bar{Y}$. Similarly, if $A \subseteq X$ is convex, then so is $\bar{A}$.

Given a subset A of a normed space X, the *closed linear span* of A, denoted by $\text{cls}\, A$ is the closure of the linear span of A (i.e. $\overline{\text{span}\, A}$). This is a closed subspace of X. Note that if A is countable, then $\text{cls}\, A$ is separable, since the set of all finite rational linear combinations of elements of A is a countable dense set.

Example. l_p for $1 \leq p < \infty$ is separable: consider $e_n = (0, \cdots, 0, 1, 0, \cdots)$, with 1 at the n-th location, for each $n \in \mathbb{N}$. Given $x = (x_n) \in l_p$,

$$\|x - \sum_{k=1}^{n} x_k e_k\|_p = \|(0, \cdots, 0, x_{n+1}, x_{n+2}, \cdots)\|_p = \left(\sum_{k=n+1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} \to 0$$

so $l_p = \text{cls}\{e_n : n \in \mathbb{N}\}$.

Similarly c_0 and c are also separable, but l_∞ is not separable.

1.3 Bounded linear operators

Theorem. Let X and Y be normed spaces, $T : X \to Y$ linear. The following are equivalent:

(i) T is continuous at 0
(ii) T is continuous
(iii) T is Lipschitz
(iv) T is *bounded*, i.e. there exists $C \geq 0$ such that $\|Tx\| \leq C\|x\|$ for all $x \in X$.

Proof. (iv) $\Rightarrow$ (iii): For $x, y \in X$,

$$d(x,y) = \|Tx - Ty\| = \|T(x-y)\| \leq C\|x-y\| = d(x,y)$$

so T is Lipschitz with constant C.

(iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) is immediate

(i) $\Rightarrow$ (iv): there exists $\delta > 0$ such that $\|x\| \leq \delta \;\Rightarrow\; \|Tx\| \leq 1$. For any non-zero x, $\|\frac{\delta x}{\|x\|}\| = \delta$, so $\|T\left(\frac{\delta x}{\|x\|}\right)\| \leq 1$. By homogeneity and linearity, $\|Tx\| \leq \frac{1}{\delta}\|x\|$. This also holds if $x = 0$, so (iv) holds with $C = \frac{1}{\delta}$. ☐

The smallest C in (iv) is the *operator norm* of T, denoted by $\|T\|$ or $\|T\|_{\text{op}}$. Note that $\|Tx\| \leq \|T\| \cdot \|x\|$, and $\|T\| = \sup\{\|Tx\| : x \in B_X\}$. We write

$$\mathbb{B}(X,Y) = \{T : X \to Y : T \text{ is bounded linear}\}$$

This is a normed space in the operator norm. For example, if $S, T \in \mathbb{B}(X,Y)$, then $\|(S+T)x\| = \|Sx+Tx\| \leq \|Sx\|+\|Tx\| \leq \|S\|\cdot\|x\|+\|T\|\cdot\|x\| \leq (\|S\|+\|T\|)x$. Taking supremum over $x \in B_X$ gives $S+T \in \mathbb{B}(X,Y)$ and $\|S+T\| \leq \|S\|+\|T\|$. If $X = Y$, then write $\mathbb{B}(X)$ for $\mathbb{B}(X,X)$.

Proposition. Let X, Y, Z be normed spaces, $S \in \mathbb{B}(X,Y)$, $T \in \mathbb{B}(Y,Z)$, then $TS \in \mathbb{B}(X,Z)$, and $\|TS\| \leq \|T\| \cdot \|S\|$.

Proof. The assertion $TS \in \mathbb{B}(X,Z)$ is clear since it is a composition of continuous linear maps. For $x \in X$, $\|TSx\| \leq \|T\| \cdot \|Sx\| \leq \|T\| \cdot \|S\| \cdot \|x\|$. Taking supremum over $x \in B_X$ yields the result. ☐

Example.

(1) For any normed space, the identity map $X \to X$ is in $\mathbb{B}(X)$ and has norm 1. It is denoted by $\mathbb{I}_X$ or Id.

(2) Let X, Y be normed spaces. Consider $X \oplus Y$ with norm $\|(x,y)\|_1 = \|x\|+\|y\|$. If X and Y are both complete, then $X \oplus Y$ is complete under the norm. The projection map $p_X : X \to Y$, $(x,y) \mapsto x$ is a bounded linear map with $\|p_X\| \leq 1$.

(3) Consider $T : (c_{00}, \|\cdot\|_1) \to \mathbb{R}$, $Tx = \sum_{n=1}^{\infty} nx_n$. Here c_{00} is the space of all sequences which are eventually 0. For $e_n = (0, \cdots, 0, 1, 0 \cdots, 0)$ with 1 at the n-th position, $|Te_n| = n$, but $\|e_n\|_1 = 1$, so T is linear but unbounded.

For X, Y normed spaces, $T : X \to Y$ is an *isomorphism* if T is a linear bijection, and T, T^{-1} are both bounded. This is equivalent to stating that T is a linear bijection, and there exists $a, b > 0$ such that $a\|x\| \leq \|Tx\| \leq b\|x\|$ for all $x \in X$. If in addition T is an isometry ($a = b = 1$), then T is an *isometric isomorphism.*

$T : X \to Y$ is an *isomorphic embedding* if $T : X \to T(X)$ is an isomorphism, or equivalently there exists $a, b > 0$ such that $\forall x \in X \; a\|x\| \leq \|Tx\| \leq b\|x\|$, i.e. remove the surjectivity requirement from an isomorphism. If such T exists, then write $X \hookrightarrow Y$.

Two norms $\|\cdot\|$ and $\|\cdot\|'$ on a vector space X are *equivalent* if Id $: (X, \|\cdot\|) \to (X, \|\cdot\|')$ is an isomorphism, which is equivalent to there exists $a, b > 0$ such that for all $x \in X$, $a\|x\| \leq \|x\|' \leq b\|x\|$, which is also equivalent to $aB'_X \subseteq B_X \subseteq bB'_X$. We then write $\|\cdot\| \sim \|\cdot\|'$. For example, $\|\cdot\|_1 \sim \|\cdot\|_2$ on $\mathbb{R}^2$.

Remark.

(1) For X, Y normed spaces, if $X \sim Y$, then X is complete $\iff$ Y is complete. Hence, if $\|\cdot\|$ and $\|\cdot\|'$ are equivalent norms on a vector space X, then $(X, \|\cdot\|)$ is complete $\iff$ $(X, \|\cdot\|')$ is complete.

(2) $(C[0,1], \|\cdot\|_\infty)$ is complete, and $(C[0,1], \|\cdot\|_1)$ is incomplete. Therefore, $\|\cdot\|_\infty \nsim \|\cdot\|_1$. This can also been seen directly by considering a function with a small support and a sharp spike, which has a fixed 1-norm but arbitrarily high ∞-norm.

However, $\|f\|_1 = \int_0^1 |f(t)|dt \leq \|f\|_\infty$, so $\mathbb{I} : (C[0,1], \|\cdot\|_\infty) \to (C[0,1], \|\cdot\|_1)$ is a continuous linear bijection whose inverse is not continuous.

(3) On $X \oplus Y$, $\|(x,y)\|_1 = \|x\| + \|y\|$ is equivalent to $\|(x,y)\|_p = \left(\|x\|^p + \|y\|^p\right)^{\frac{1}{p}}$ for $1 \leq p < \infty$, and to $\|(x,y)\|_\infty = \max(\|x\|, \|y\|)$.

Suppose X, Y are normed spaces, $T_n \to T$ in $\mathbb{B}(X,Y)$ with the operator norm, then for $x \in X$, $\|T_n x - Tx\| = \|(T_n - T)x\| \leq \|T_n - T\| \cdot \|x\| \to 0$ as $n \to \infty$, so $T_n \to T$ pointwise. The converse is false: take $T_n : l_1 \to \mathbb{R}$, $T_n x = x_n$, the projection onto the n-th component. Then $T_n \in \mathbb{B}(l_1, \mathbb{R})$, and $\|T_n\| = 1$. Also, $T_n x \to 0$ as $n \to \infty$ for all $x \in l_1$ since the sequence is summable. Therefore, $T_n \to 0$ pointwise, but $T_n \nrightarrow 0$ in operator norm since all T_n are on the unit sphere.

Theorem. Let X, Y be normed spaces. If Y is complete, then $\mathbb{B}(X,Y)$ is complete.

Proof. Let $(T_n)_{n=1}^\infty$ be a Cauchy sequence in $\mathbb{B}(X,Y)$. For $x \in X$, $\|T_m x - T_n x\| \leq \|T_m - T_n\| \cdot \|x\| \to 0$ as $m, n \to \infty$, so $(T_n x)_{n=1}^\infty$ is a Cauchy sequence in Y, and hence it converges. Call the limit Tx. This defines a function $T : X \to Y$.

T is linear: Let $x, y \in X$, λ, μ scalars,then

$$T(\lambda x + \mu y) = \lim_{n\to\infty} T_n(\lambda x + \mu y) = \lim_{n\to\infty} (\lambda T_n(x) + \mu T_n(y)) = \lambda T(x) + \mu T(y)$$

T is bounded: Since $(T_n)_{n=1}^\infty$ is Cauchy, $\exists n \in \mathbb{N}$ s.t. $\|T_m - T_n\| \leq 1$ for all $m, n \geq N$, so for all $n \geq N$, $\|T_n\| \leq \|T_n - T_N\| + \|T_N\| \leq 1 + \|T_N\|$, so $\|T_n\| \leq M$ for all $n \in \mathbb{N}$, where $M = \max(\|T_1\|, \|T_2\|, \cdots, \|T_{N-1}\|, \|T_N\| + 1)$. It follows that for all $x \in X$, $n \in \mathbb{N}$, $\|T_n x\| \leq M\|x\|$. Take $n \to \infty$ gives $\|Tx\| \leq M\|x\|$, so T is bounded with $\|T\| \leq M$.

$T_n \to T$ in $\mathbb{B}(X,Y)$: Let $\epsilon > 0$. There exists $n \in N$ such that $\|T_m - T_n\| \leq \epsilon$ for all $m, n \geq N$. For $x \in X$, and $m, n \geq N$, $\|T_m x - T_n x\| \leq \epsilon\|x\|$. Fix $x \in X$, $n \geq N$, and let $m \to \infty$ to get $\|Tx - T_n x\| \leq \epsilon\|x\|$. Taking supremum over $x \in B_X$ with a fixed $n \geq N$ to get $\|T - T_n\| \leq \epsilon$. This is true for all $n \geq N$, so $T_n \to T$ is the operator norm. □

2 Dual spaces and the Hahn-Banach theorem

Let X be a normed space. A *functional* on X is a mpa $X \to$ scalar field ($\mathbb{R}$ or $\mathbb{C}$). The *dual space* X^* of X is the space of all bounded linear functionals on X with the operator norm, i.e. $X^* = \mathbb{B}(X, \mathbb{R})$ or $\mathbb{B}(X, \mathbb{C})$. For $f \in X^*$, $\|f\| = \sup\{|f(x)| : x \in X, \|x\| \leq 1\}$, so $|f(x)| \leq \|f\| \cdot \|x\|$ for all $f \in X^*, x \in X$. An alternative notation is $f(x) = \langle x, f \rangle$.

Since $\mathbb{R}$ and $\mathbb{C}$ are complete, theorem 1.3 immediately gives

Theorem. X^* is a Banach space for any normed space X.

To construct elements in the dual space, we will state and prove the Hahn-Banach theorem, which is not examinable because of its reliance on Zorn's lemma.

Theorem (Hahn-Banach Theorem)**.** Let X be a normed space, Y a subspace, $g \in Y^*$. Then there exists $f \in X^*$ such that $f|_Y = g$ and $\|f\| = \|g\|$.

Proof. First assume X is a real normed space. WLOG assume $\|g\| = 1$. We first consider extension by one element. If $Y = X$, then we are done, otherwise pick $x_1 \in X \backslash Y$, and let $Y_1 = Y \oplus \mathbb{R}x_1$. For $y \in Y$ and $\lambda \in \mathbb{R}$, define

$$f_1 : Y \to \mathbb{R},\ f_1(y + \lambda x_1) = g(y) + \lambda\alpha$$

where $\alpha \in \mathbb{R}$ is to be determined. Clearly $f_1|_Y = g$, so we need to choose α such that $\|f_1\| = 1$. Since f_1 take all values g takes on B_Y, $\|f_1\| \geq 1$, so we only need $\|f_1\| \leq 1$. This is equivalent to the following

$$|g(y) + \lambda\alpha| \leq \|y + \lambda x_1\| \quad \forall y \in Y, \lambda \in \mathbb{R}$$

Divide by λ and apply homogeneity, this is equivalent to $|g(y) + \alpha| \leq \|y + x_1\|\ \forall y \in Y$, which is equivalent to

$$-g(z) - \|z + x_1\| \leq \alpha \leq \|y + x_1\| - g(y) \quad \forall y, z \in Y$$

Such α exists if and only if the LHS is at most the RHS for all $y, z \in Y$, for then one can take supremum over all z of the LHS and infimum over all y of the RHS to get a non-empty interval of $\mathbb{R}$. We have

$$-g(z) + g(y) = g(y - z) \leq \|y - z\| \leq \|y + x_1\| + \|z + x_1\|$$

for all $y, z \in Y$, so we can do a one-step extension.

We apply Zorn's lemma on $P = \{(Z, h) : Y \leq Z \leq X,\ h \in Z^*,\ h|_Y = g,\ \|h\| = 1\}$. Define $(Z_1, h_1) \leq (Z_2, h_2)$ if $Z_1 \leq Z_2$ and $h_2|_z = h_1$. This makes $(P, \leq)$ a poset. $P \neq \emptyset$ since $(Y, g) \in P$. If $\{(Z_\alpha, h_\alpha) : \alpha \in I\}$ is a non-empty chain (a totally ordered set under $\leq$), then $Z = \bigcup_{\alpha \in I} Z_\alpha$, $h(z) = h_\alpha(z)$ for $\alpha \in I$ and $z \in Z_\alpha$. The total ordering implies both that Z is a subspace of X and that h is well-defined. This is an upper bound of the chain. Now Zorn's lemma gives a maximal element (W, f) of P (it is not strictly smaller than any elemnt of P). If $W \neq X$, then it can be extended by the above one-step extension, contradicting maximality. Hence $W = X$, and we are done.

Let X be a complex normed space, and let $X_\mathbb{R}$ be the same X viewed as a real normed space. For $f \in X^*$, let $\mathbb{R}e(f)$ be defined by $x \mapsto \mathbb{R}e(f(x))$. Clearly $\mathbb{R}e(f)$ is $\mathbb{R}$-linear and $|\mathbb{R}e(f)(x)| = |\mathbb{R}e(f(x))| \leq |f(x)| \leq \|f\| \cdot \|x\|$. Hence $\mathbb{R}e(f) \in (X_{\mathbb{R}})^*$ with $\|\mathbb{R}e(f)\| \leq \|f\|$. We show that $f \mapsto \mathbb{R}e(f)$ is a $\mathbb{R}$-linear isometric isomorphism $X^* \to (X_{\mathbb{R}})^*$.

For any $f \in X^*$ and $x \in X$, choose $\lambda \in \mathbb{C}$, $|\lambda| = 1$, such that $|f(x)| = \lambda f(x)$. Then

$$|f(x)| = f(\lambda x) = \mathbb{R}e(f(\lambda x)) \leq \|\mathbb{R}e(f)\| \cdot \|\lambda x\| = \|\mathbb{R}e(f)\| \cdot \|x\|$$

Taking supremum over all $\|x\| \leq 1$ gives $\|f\| \leq \|\mathbb{R}e(f)\|$, so they are equal. It remains to show that $\mathbb{R}e(f)$ uniquely determines f. Suppose $\mathbb{R}e(f) = h$ and $\Im(f) = k$, then

$$\left.\begin{aligned} f(x) = h(x) + ik(x) \\ f(x) = -if(ix) = -ih(ix) + k(ix) \end{aligned}\right\} \Rightarrow \; k(x) = -h(ix),\; f(x) = h(x) - ih(ix)$$

Therefore, $X^* \cong (X_{\mathbb{R}})^*$. The real case of Hahn-Banach theorem then finishes the proof. □

Corollary. If X is a real or complex normed space, then $\forall x_0 \in X$, $x_0 \neq 0$, there exists a functional $f \in X^*$ with $\|f\| = 1$ such that $f(x_0) = \|x_0\|$.

Remark.

(1) This shows that X^* separates points, i.e. if $x \neq y$ in X, then $\exists f \in X^*$ such that $f(x) \neq f(y)$ by taking $x_0 = x - y$ in the corollary.

(2) The function constructed in corollary 2 is called a *norming function* for x_0: If $g \in B_{X^*}$, then $|g(x_0)| \leq \|x_0\|$, so $\|x_0\| \geq \sup\{|g(x_0)| : g \in B_{X^*}\}$. Corollary 2 states that the supremum is attained, i.e. $\|x_0\| = \max\{|g(x_0)| : g \in B_{X^*}\}$

(3) Another name for f is the *support functional*: assuming $\|x_0\| = 1$, the ball B_X lies on one side of $\{x \in X : f(x) = 1\}$ and touches it at x_0.

Note. Two applications of the Hahn-Banach theorem:

(1) X is separable $\Rightarrow$ $X \hookrightarrow l_\infty$ isometrically.

(2) X^* is separable $\Rightarrow$ X is separable.

2.1 Dual of l_p spaces

Fix $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, we aim to show that $l_p^* \cong l_q$. For $y = (y_n) \in l_q$, define $\phi_y : l_q \to \mathbb{R}$, $\phi_y(x) = \sum_{n=1}^{\infty} x_n y_n$ for $x = (x_n) \in l_1$. This is well-defined since Hölder's inequality gives $|\phi_y(x)| \leq \|x_p\| \cdot \|y_q\|$. ϕ_y is clearly linear. Furthermore, the inequality gives $\phi_y \in l_p^*$ with $\|\phi_y\| \leq \|y\|_q$.

Theorem. Let $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Define $\phi : l_q \to l_p^*$ by $y \mapsto \phi_y$ defined as above. Then ϕ is an isometric isomorphism.

Proof. We have already seen that $\phi_y \in l_q^*$ with $\|\phi_y\| \leq \|y\|_q$. We first show that $\|\phi y\| = \|y\|_q$. Set $x_n = \lambda_n |y_n|^{q-1}$, where λ_n is a scalar with $|\lambda_n| = 1$ such that

$\lambda_n y_n = |y_n|$. Then $x = (x_n) \in l_p$ since $\|x\|_p^p = \sum |x_n|^p = \sum |y_n|^{pq-p} = \sum |y_n|^q = \|y\|_q^q < \infty$. Using this,

$$\phi_y(x) = \sum_{n=1}^{\infty} x_n y_n = \sum_{n=1}^{\infty} |y_n|^q = \|y\|_q^q = \|x\|_p \|y\|_q^{q-\frac{q}{p}} = \|x\|_p \|y\|_q$$

It follows that $\|\phi_y\| = \|y\|_q$, and so ϕ is a linear isomorphism. We need to show that ϕ is onto. Let $f \in l_p^*$. For $n \in \mathbb{N}$, set $y_n = f(e_n)$, where e_n is the vector with 1 at index n and 0 elsewhere. Let $y = (y_n)$, then we need to check that $y \in l_p$ and $\phi_y = f$.

$y \in l_p$: Fix $N \in \mathbb{N}$. For $1 \leq n \leq N$, set $x_n = \lambda_n |y_n|^{q-1}$, where $\lambda_n = 1$, $\lambda_n y_n = |y_n|$, and set $x_n = 0$ for $n > N$. Consider $x = (x_n) = \sum_{n=1}^{N} x_n e_n$. $\|x\|_p = \left(\sum_{n=1}^{N} |x_n|^p\right)^{\frac{1}{p}} = \left(\sum_{n=1}^{N} |y_n|^q\right)^{\frac{1}{p}}$, so

$$\sum_{n=1}^{N} |y_n|^q = \sum_{n=1}^{N} x_n y_n = f\left(\sum_{n=1}^{N} x_n e_n\right) \leq \|f\| \cdot \|x\|_p = \|f\| \cdot \left(\sum_{n=1}^{N} |y_n|^q\right)^{\frac{1}{p}}$$

Using $\frac{1}{p} + \frac{1}{q} = 1$, we have $\left(\sum_{n=1}^{N} |y_n|^q\right)^{\frac{1}{q}} \leq \|f\|$. Letting $N \to \infty$ gives $\|y\|_q \leq \|f\|$.

$\phi_y = f$: $\phi_y(e_n) = y_n = f(e_n)$ for all $n \Rightarrow \phi_y = f$ on $\operatorname{span}\{e_n : n \in \mathbb{N}\}$ since they are both linear $\Rightarrow \phi_y = f$ on $\operatorname{cls}\{e_n : n \in \mathbb{N}\} = l_p$ since they are both continuous. □

2.2 Second dual

The *second dual* or *bi-dual* of a normed space X is $X^{**} = (X^*)^* = \mathbb{B}(X^*, \mathbb{R})$. For $x \in X$, define $\hat{x} : X^* \to \mathbb{R}$ by $\hat{x}(f) = f(x)$ for $f \in X^*$, called the evaluation at x. $\hat{x}$ is linear. $|\hat{x}(f)| = |f(x)| \leq \|f\| \cdot \|x\|$, so taking supremum over all $f \in B_{X^*}$ gives $\hat{x} \in X^{**}$ with $\|\hat{x}\| \leq \|x\|$. The map $X \to X^{**}$, $x \mapsto \hat{x}$ is the *canonical embedding*. In the alternative notation, $\langle f, \hat{x} \rangle = \langle x, f \rangle$ for $x \in X, f \in X^*$.

Hahn-Banach theorem gives that this is an isomorphic embedding of X into X^{**} since

$$\|\hat{x}\| = \sup\{|f(x)| : f \in X^*, \|f\| \leq 1\} = \|x\|$$

by corollary 2. If this map is onto, then X is *reflexive*. l_p is reflextive for $1 < p < \infty$.

Remark. We have shown that for $\frac{1}{p} + \frac{1}{q} = 1$, $l_p^* \cong l_q$, and so $l_p^{**} \cong l_q^* \cong l_p$, but this is not enough, since there exists non-reflexive spaces X with $X \cong X^{**}$. We need to show that the map constructed in the proof is the canonical embedding.

2.3 Dual operators

Let $T : X \to Y$ be a bounded linear map between normed spaces. The *dual operator* of T is $T^* : Y^* \to X^*$ defined by $T^*(g) = g \circ T$, so $T^*(g)(x) = g(Tx)$ or $\langle x, T^*g \rangle = \langle Tx, g \rangle$ for $x \in X$ and $g \in Y^*$. By proposition 1.3, $T^*g = g \circ T \in X^*$ with $\|T^*g\| \leq \|g\| \cdot \|T\|$, so T^* is well-defined and $T^* \in \mathbb{B}(Y^*, X^*)$ with $\|T^*\| \leq \|T\|$. In fact,

$$\|T^*\| = \sup_{g \in B_{Y^*}} \|T^*g\| = \sup_{g \in B_{Y^*}} \sup_{x \in B_X} |T^*g(x)| = \sup_{x \in B_X} \sup_{g \in B_{Y^*}} |g(Tx)| = \sup_{x \in B_X} \|Tx\|$$

by corollary 2. Therefore, $\|T^*\| = \|T\|$.

Example. For $1 \leq p < \infty$, let $T : l_p \to l_p$ be the right shift, then $T^* : l_p^* \to l_p^*$ is the left shift when l_p^* is identified with l_q, where $\frac{1}{p} + \frac{1}{q} = 1$.

The dual operator has the following properties

(1) $(\mathbb{I}_X)^* = \mathbb{I}_{X^*}$

(2) $(\lambda S + \mu T)^* = \lambda S^* + \mu T^*$ for all $S, T \in \mathbb{B}(X, Y)$, λ, μ scalars since $\langle x, (\lambda S + \mu T)^* g\rangle = \langle (\lambda S + \mu T)x, g\rangle = \lambda\langle Sx, g\rangle + \mu\langle Tx, g\rangle = \lambda\langle x, S^* g\rangle + \mu\langle x, T^* g\rangle = \langle x, (\lambda S^* + \mu T^*)g\rangle$ for all $x \in X$ and $g \in X^*$.

This shows that $T \mapsto T^* : \mathbb{B}(X, Y) \hookrightarrow \mathbb{B}(Y^*, X^*)$ is an isometric embedding.

(3) For $S \in \mathbb{B}(X, Y)$ and $T \in \mathbb{B}(Y, Z)$, $(TS)^* = S^* T^*$. In particular, if $X \sim Y$, then $X^* \sim Y^*$ by dualizing an isomorphism $X \to Y$ and its inverse.

(4) Given $T \in \mathbb{B}(X, Y)$, the following square commutes, i.e. $\widehat{Tx} = T^{**}\hat{x}$ for all $x \in X$:

$$\begin{array}{ccc} X & \xrightarrow{T} & Y \\ \big\downarrow{\scriptstyle \hat{x}} & & \big\downarrow{\scriptstyle \hat{y}} \\ X^{**} & \xrightarrow{T^{**}} & Y^{**} \end{array}$$

This follows from $\langle g, \widehat{Tx}\rangle = \langle Tx, g\rangle = \langle x, T^* g\rangle = \langle T^* g, \hat{x}\rangle = \langle g, T^{**}\hat{x}\rangle$.

3 Finite-dimensional normed spaces

Recall that two norms $\|\cdot\|$ and $\|\cdot\|'$ on a vector space are equivalent if there exists $a, b > 0$ auch that $a\|x\| \leq \|x\|' \leq b\|x\|$ for all $x \in X$. E.g. on $\mathbb{R}^n$, $\|\cdot\|_1 \sim \|\cdot\|_2$, since

$$\left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}} \leq \sum_{i=1}^n |x_i| \leq \sqrt{n}\left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}}$$

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

Proof. Let $\|\cdot\|$ and $\|\cdot\|'$ be two norms on X. Choose basis $x_1, \cdots, x_n \in X$, then there is a linear isomorphism $T : \mathbb{R}^n \to X$, $(a_i)_{i=1}^n \mapsto \sum_{i=1}^n a_i x_i$. We define two norms $\|x\|_{\mathbb{R}} = \|Tx\|$ and $\|x\|'_{\mathbb{R}} = \|Tx\|'$ on $\mathbb{R}^n$. If $\|\cdot\|_{\mathbb{R}} \sim \|\cdot\|'_{\mathbb{R}}$, then $\|\cdot\| \sim \|\cdot\|'$ on X, so WLOG let $X = \mathbb{R}^n$. Further assume $\|\cdot\|' = \|\cdot\|_2$, since norm equivalence is an equivalence relation.

Let $e_1, \cdots, e_n$ be the standard basis, then

$$\begin{aligned} \left|\left\|\sum_{i=1}^n x_i e_i\right\| - \left\|\sum_{i=1}^n y_i e_i\right\|\right| \leq \|\sum_{i=1}^n (x_i - y_i)e_i\| &\leq \sum_{i=1}^n |x_i - y_i| \cdot \|e_i\| \\ &\leq \left(\sum_{i=1}^n |x_i - y_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^n \|e_i\|^2\right)^{\frac{1}{2}} = C\|x - y\|_2 \end{aligned}$$

Therefore, $\|\cdot\| : (\mathbb{R}^n, \|\cdot\|_2) \to \mathbb{R}$ is Lipschitz, and in particular continuous.

Let $S = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$ be the l_2^n unit sphere. Under the Euclidean metric, S is compact. Suppose $a = \inf_{x \in S} \|x\|$ is attained at x_0, then $a = \|x_0\| > 0$. Hence $\|x\| \geq a$ for all $x \in S$. By homogeneity, $a\|x\|_2 \leq \|x\|$ for all $x \in \mathbb{R}^n$. From above, $\|x\| \leq C\|x_2\|$ for all $x \in \mathbb{R}^n$, so $\|\cdot\| \sim \|\cdot\|_2$. ☐

Corollary. Let X, Y be normed spaces, $T : X \to Y$ linear. If $\dim X < \infty$, then T is continuous.

Proof. Let $\|\cdot\|$ denote the norm on X and Y (which may be different). Define a new norm on X by $\|x\|' = \|x\| + \|Tx\|$. It is easy to check that this is a norm. By theorem 3, $\|\cdot\| \sim \|\cdot\|'$ on X, so there is a $b > 0$ such that $\|x\|' \leq b\|x\|$. In particular, $\|Tx\| \leq b\|x\|$ for all $x \in X$, so T is continuous. ☐

Corollary. If $\dim X = \dim Y$, then $X \sim Y$

Proof. Let $T : X \to Y$ be a linear bijection. By corollary 3, T and T^{-1} are both continuous, so they are continuous isomorphisms. ☐

Observe that if (X, d) be a metric space, $Y \subseteq X$, then we have

(i) Y is complete $\Rightarrow$ Y is closed in X.
(ii) X is complete, Y closed in X $\Rightarrow$ Y is complete.

Corollary. (i) If $\dim X < \infty$, then X is complete.
(ii) If X is a normed space and Y is a finite dimensional subspace, then Y is closed.

Proof. (i) If $n = \dim X$, then by corollary 3, $X \sim l_2^n$. l_2^n is complete, so X is complete.
(ii) This follows from (i) and the observations above. ☐

Corollary. If $\dim X < \infty$ and $A \subset X$ is closed and bounded, then A is compact.

Proof. Let $n = \dim X$ and $T : X \to l_2^n$ be an isomorphism, then $T(A)$ is closed and bounded in l_2^n, so it is compact by the Heine-Borel theorem. Apply T^{-1} to get that A is compact. ☐

Remark. In l_p for $1 \leq p < \infty$, let e_n be the vector with 1 at index n and 0 everywhere else. Then $\|e_m - e_n\| = 2^{\frac{1}{p}} \geq 1$ for all $m \neq n$, so $(e_n)_{n=1}^{\infty}$ has no convergent subsequence. In particular, B_{l_p} is not compact, despite being closed and bounded.

Proposition (Riesz's Lemma)**.** Let Y be a proper, closed subspace of a normed space X, then for all $\epsilon > 0$, there is an $x \in B_X$ such that $d(x, Y) = \inf\{\|x - y\| : y \in Y\} > 1 - \epsilon$.

Proof. Let $z \in X \backslash Y$. Since Y is closed, there is a postive r such that $B(z, r) \subseteq X \backslash Y$, so $d(z, Y) \geq r > 0$. There exists $y \in Y$ such that $d(z, Y) \leq \|z - y\| < \frac{1}{1-\epsilon} d(z, Y)$. Set $x = \frac{z-y}{\|z-y\|} \in B_X$, then $d(x, Y) = \frac{d(z-y,Y)}{\|z-y\|} = \frac{d(z,Y)}{\|z-y\|} < 1 - \epsilon$. ☐

Theorem. Let X be a normed space. If B_X is compact, then $\dim X < \infty$.

Proof. Assume $\dim X = \infty$. We will inductively construct a sequence (x_n) in B_X such that $\|x_m - x_n\| > \frac{1}{2}$ for all $m \neq n$. Then (x_n) has no convergent subsequence, so B_X is not sequentially compact, and hence not compact.

Pick $x_1 \in B_X$ arbitrarily. Assume $x_1, \cdots, x_n$ for some $n \in \mathbb{N}$ has been constructed, let $Y = \operatorname{span}\{x_1, \cdots, x_n\}$, then Y is a proper, closed subspace of X by corollary 3. By Riesz's lemma, there exists $x_{n+1} \in B_X$ such that $d(x_{n+1}, Y) > \frac{1}{2}$. In particular, $\|x_{n+1}, x_i\| > \frac{1}{2}$ for $i = 1, \cdots, n$, so we are done. □

4 Baire category theorem and applications

Recall that a subset A of a topological space X is *dense* in X if $\bar{A} = X$, i.e. for all open, non-empty $U \subseteq X$, $U \cap A \neq \emptyset$. For example, $\mathbb{Q}$ is dense in $\mathbb{R}$. Note that $\mathbb{Q}$ and $\mathbb{Q} + \sqrt{2}$ are both dense in $\mathbb{R}$, but their intersection is empty. The Baire category theorem states that open dense sets are "large" in a complete metric space.

Theorem (Baire Category Theorem)**.** Let (X, d) be a complete metric space. If for each $n \in \mathbb{N}$, we have a dense open set U_n in X, then $\bigcap_{n=1}^{\infty} U_n$ is dense in X.

Before we prove this, we fix some notations. In a metric space (X, d), for $x \in X$ and $r > 0$, define the open ball $B(x, r) = \{y \in X : d(y, x) < r\}$ and the closed ball $\bar{B}(x, r) = \{y \in X : d(x, y) \leq r\}$. Note that $\overline{B(x, r)} \subseteq \bar{B}(x, r)$, but equality may not hold in general. If X is a normed space, then $\bar{B}(x, r) = \overline{B(x, r)} = x + rB_X$. Since the open balls form a base for the metric topology, a subset A of X is dense if and only if $A \cap B(x, r) \neq \emptyset$ for all $x \in X, r > 0$.

Proof of theorem 4. Let $x \in X$ and $r > 0$. Need to show that $B(X, r) \cap \left(\bigcap_{n=1}^{\infty} U_n\right) \neq \emptyset$.

U_1 is dense, so $B(x, r) \cap U_1 \neq \emptyset$. Let $x_1 \in B(x, r) \cap U_1$. The set is also open, so let $0 < r_1 < 1$ be such that $\bar{B}(x_1, r_1) \subseteq B(x, r) \cap U_1$.

U_2 is dense, so $B(x_1, r_1) \cap U_2 \neq \emptyset$. Let $x_2 \in B(x_1, r_1) \cap U_2$, then there exists $0 < r_2 < \frac{1}{2}$ such that $\bar{B}(x_2, r_2) \subseteq B(x_1, r_1) \cap U_2 \subseteq B(x, r) \cap U_1 \cap U_2$.

Continue the process inductively to create sequences (x_n) in X and (r_n) in $(0, \infty)$ satisfying the three conditions

(i) $r_n < \frac{1}{n}$ for all $n \geq 1$.
(ii) $\bar{B}(x_n, r_n) \subseteq \bar{B}(x_{n-1}, r_{n-1})$ for all $n > 1$.
(iii) $\bar{B}(x_n, r_n) \subseteq B(x, r) \cap U_1 \cap \cdots \cap U_n$ for all $n \geq 1$.

Let $N \in \mathbb{N}$. If $m, n \geq N$, then by (ii), $x_m, x_n \in \bar{B}(x_N, r_N)$, so $d(x_m, x_n) < 2r_N < \frac{2}{N}$. Therefore (x_n) is Cauchy. (X, d) is complte, so $x_n \to z$ as $n \to \infty$ for some $z \in X$. Fix m, (ii) implies that $x_n \in \bar{B}(x_m, r_m)$ for all $n \geq m$, so $z \in \bar{B}(x_m, r_m)$. By (iii), $z \in B(x, r) \cap \left(\bigcap_{n=1}^{m} U_n\right)$. This holds for all m, so $z \in B(x, r) \cap \left(\bigcap_{n=1}^{\infty} U_n\right)$. □

Example. $\mathbb{R}$ is uncountable: Suppose $f : \mathbb{N} \to \mathbb{R}$ is a function. Let $U_n = \mathbb{R} \backslash \{f(n)\}$, then U_n is open and dense in $\mathbb{R}$. By theorem 4, $\bigcap_{n=1}^{\infty} U_n$ is dense, so in particular non-empty. The intersection is $\mathbb{R} \backslash f(\mathbb{N})$, so f is not surjective.

In a topological space X, a countable intersection of open set is called a G_δ set. A countable union of closed set is called a F_σ set. Therefore, theorem 4 states that countable intersection of dense open sets is a dense G_δ set.

Note that if $U \subseteq X$ is open and dense in X, then $F = X \backslash U$ is closed with $\mathbb{Z}F = X \backslash (\overline{X \backslash F}) = \emptyset$. In general, a subset A of a topological space is *nowhere dense* in X if $\mathbb{Z}\overline{A} = \emptyset$. Being nowhere dense is relative to the larger space. Given $A \subseteq Y \subseteq X$, it is possible that A is nowhere dense in X but dense in Y, by taking e.g. $A = Y \neq \emptyset$.

Note. A is nowhere dense in $X \iff U \nsubseteq \overline{U \cap A}$ for any non-empty, open $U \subseteq X$. A is dense in $X \iff U \subseteq \overline{U \cap A}$ for all non-empty, open $U \subseteq X$.

We can now give an alternative formulation of the Baire category theorem.

Theorem. Let (X, d) be a non-empty, complete metric space. If $X = \bigcup_{n=1}^{\infty} A_n$, then for some $N \in \mathbb{N}$, $\mathbb{Z}\overline{A_N} \neq \emptyset$.

Proof. Let $U_n = X \backslash \overline{A_n}$, then U_n is open for all n, and $\bigcap_{n=1}^{\infty} U_n = X \backslash \bigcup_{n=1}^{\infty} A_n = \emptyset$. Since $X \neq \emptyset$, $\bigcap_{n=1}^{\infty}$ is not dense in X. By theorem 4, for some $N \in \mathbb{N}$, U_N is not dense in X, so $\mathbb{Z}\overline{A_N} = X \backslash (\overline{X \backslash \overline{A_N}}) = X \backslash \overline{U_N} \neq \emptyset$. □

A subset A of X is *meagre* in X (or is of *first category*) if $A = \bigcup_{n=1}^{\infty} A_n$, where each A_n is nowhere dense in X. A is of *second category* in X if it is not of the first category in X. Meagre sets are considered to be small, similar to the role of sets of measure zero in measure theory. E.g. a countable union of meagre sets is meagre.

Theorem (Banach-Steinhaus Theorem or Principle of Uniform Boundedness)**.** Let X be a Banach space, Y a normed space, and $\mathcal{T} \subseteq \mathbb{B}(X, Y)$. Assume that $\mathcal{T}$ is *pointwise bounded*, i.e. for every $x \in X$, $\sup_{T \in \mathcal{T}} \|Tx\| < \infty$. Then $\mathcal{T}$ is *uniformly bounded*, i.e. $\sup_{T \in \mathcal{T}} \|T\| < \infty$.

Proof. For each $n \in \mathbb{N}$, let $F_n = \{x \in X : \|Tx\| \leq n \text{ for all } T \in \mathcal{T}\}$, then by pointwise boundedness, $X = \bigcup_{n \in \mathbb{N}} F_n$. $F_n = \bigcap_{T \in \mathcal{T}} \{x \in X : \|Tx\| \leq n\}$, so it is closed. By theorem 4, there exists $n \in \mathbb{N}$ such that $\mathbb{Z}F_n \neq \emptyset$, so there exists $x_0 \in X$ and $r > 0$ such that $x_0 + rB_X \subseteq F_n$. Given $x \in B_X$, $x_0 \pm rx \in F_n$, so for all $T \in \mathcal{T}$,

$$\|Tx\| = \frac{1}{2r} \|T(x_0 + rx) - T(x_0 - rx)\| \leq \frac{n}{r}$$

Thus $\|T\| \leq \frac{n}{r}$ for all $T \in \mathcal{T}$. □

Example. Let X be a normed space. Say $A \subseteq X$ is *weakly bounded* if for all $f \in X^*$, $\sup_{x \in A} |f(x)| < \infty$. By theorem 4, if A is weakly bounded, then A is norm bounded, i.e. $\sup_{x \in A} \|x\| < \infty$: Consider $\hat{A} = \{\hat{x} : x \in A\} \subseteq X^{**} = \mathbb{B}(X^*, \mathbb{R})$, where $x \mapsto \hat{x}$ is the canonical isometric embedding. A is weakly bounded implies $\hat{A}$ is pointwise bounded, so theorem 4 shows that $\hat{A}$ is uniformly bounded. But $\|x\| = \|\hat{x}\|$, so A is norm bounded.

Corollary. Let X, Y be as in theorem 4. Let (T_n) be a sequence in $\mathbb{B}(X, Y)$ such that $\lim_{n \to \infty} T_n x$ exists for all $x \in X$. Then this defines a bounded linear map $X \to Y$.

Proof. Let $Tx = \lim_{n \to \infty} T_n x$ for each $x \in X$. Then this defines a function $X \to Y$. By the proof of theorem 1.3. Since $(T_n x)_{n=1}^{\infty}$ is convergent, $\{T_n x : n \in \mathbb{N}\}$ is bounded, so by theorem 4, there exists C such that $\|T_n\| \leq C$ for all $n \in \mathbb{N}$. Given $x \in X$, $\|T_n x\| \leq \|T_n\| \cdot \|x\| \leq C\|x\|$ for all $n \in \mathbb{N}$. Letting $n \to \infty$ gives $\|Tx\| \leq C\|x\|$, so $T \in \mathbb{B}(X, Y)$ with $\|T\| \leq C$. □

Given subsets A and B of a metric space (X, d) and $\delta > 0$, say A is ***δ-dense*** in B if for all $b \in B$, there exists $a \in A$ such that $d(a, b) \leq \delta$. E.g. A is dense in B iff A is δ-dense in B for all δ.

Lemma (Open Mapping Lemma)**.** Let X be a Banach space, Y a normed space, and $T \in \mathbb{B}(X, Y)$. Assume that for some $M \geq 0$ and $0 < \delta < 1$, we have $T(MB_X)$ is δ-dense in B_Y, then $T\big(\frac{M}{1-\delta}B_X\big) \supseteq B_Y$.

Proof. Let $y \in B_Y$. Pick $x_1 \in MB_X$ such that $\|y - Tx_1\| \leq \delta$. Then $\frac{y-Tx_1}{\delta} \in B_Y$, so there exists $x_2 \in MB_X$ such that $\left\|\frac{y-Tx_1}{\delta} - Tx_2\right\| \leq \delta$, which can be rearranged to $\|y - T(x_1 + \delta x_2)\| \leq \delta^2$. Continue inductively, we obtain (x_n) in MB_X such that

$$\|y - T(x_1 + \delta x_2 + \cdots \delta^{n-1}x_n)\| \leq \delta^n$$

Let $x = \sum_{k=1}^{\infty} \delta^{k-1}x_k$, then this sum converges since it converges absolutely, with

$$\|x\| \leq \sum_{k=1}^{\infty} \|\delta^{k-1}x_k\| \leq \sum_{k=1}^{\infty} \delta^{k-1}M = \frac{M}{1-\delta}$$

We further have $Tx = \lim_{N\to\infty} \sum_{k=1}^{N} \delta^{k-1}Tx_k = y$. Since y is arbitrary, we are done. □

Remark.

(i) $T(MB_X)$ is always 1-dense in B_Y since $0 \in T(MB_X)$. The lemma shows that a slight strenghthening of the hypothesis is enough for $Tx = y$ to always have a solution with an upper bound on the norm of x.

(ii) In particular, T is surjective, but this is not important.

(iii) There is a sharper form which is less useful: if $\overline{T(MB_X)} \supseteq B_Y$, then $T(MD_X) \supseteq D_Y$, where $D_X = \mathbb{Z}B_X$.

Theorem (Open Mapping Theorem)**.** Let X, Y be Banach spaces, $T : X \to Y$ a surjective bounded linear map. Then T is an *open map*, i.e. for every open set $U \subseteq X$, $T(U)$ is open in Y.

Proof. $Y = T(X) = T\big(\bigcup_{n=1}^{\infty} nB_X\big) = \bigcup_{n=1}^{\infty} T(nB_X)$. By theorem 4, there exists $N \in \mathbb{N}$ such that $\mathbb{Z}\overline{T(NB_X)} \neq \emptyset$, so for some $y_0 \in Y$ and $r > 0$, $\bar{B}(y_0, r) \subseteq \overline{T(NB_X)}$. NB_X is symmetric and convex, so $\overline{T(NB_X)}$ is also symmetric and convex. Given $y \in B_Y$, $y_0 \pm ry \in \overline{T(NB_X)}$, so $-y_0 \pm ry \in \overline{T(NB_X)}$ by symmetry. Hence

$$\frac{1}{2}(y_0 + ry) + \frac{1}{2}(-y_0 + ry) = ry \in \overline{T(NB_X)}$$

This hols for all $y \in B_Y$, so $rB_Y \subseteq \overline{T(NB_X)}$. Therefore, $T(MB_X)$ is $\frac{1}{2}$-dense in B_Y, where $M = N/r$. By lemma 4, $T(2MB_X) \supseteq B_Y$.

Let $U \subseteq X$ be open. Let $y = Tx \in T(U)$. U is open, so there exists $r > 0$ such that $x + rB_X \subseteq U$. Then $T(x + rB_X) = y + rT(B_X) \subseteq T(U)$. Hence, by the previous statement, $y + \frac{r}{2M}B_Y \subseteq y + \frac{r}{2M}T(2MB_X) \subseteq T(U)$, so $T(U)$ is open. □

Theorem (Inversion Theorem)**.** Let X, Y be Banach spaces and $T : X \to Y$ a continuous linear bijection. Then T^{-1} is also continuous.

Proof. By theorem 4, T is an open map, so for open $U \subseteq X$, $(T^{-1})^{-1}(U) = T(U)$ is open in Y, so T^{-1} is continuous. □

Alternative Proof. Follow the proof of the open mapping theorem up to $T(2MB_X) \supseteq B_Y$. Given $y \in B_Y$, there exists $x \in X$ with $\|x\| \leq 2M$ such that $Tx = y$. Then $x = T^{-1}y$, so $\|T^{-1}y\| = \|x\| \leq 2M$. Taking supremum over $y \in B_Y$ gives $\|T^{-1}\| \leq 2M$. □

Remark. This is the linear analogue of the topological inverse function theorem: If X is compact, Y is Hausdorff, and $f : X \to Y$ is a continuous bijection, then f^{-1} is continuous.

Corollary. Let $\|\cdot\|, \|\cdot\|'$ be two complete norms on a vector space X. If there exists C such that $\|x\|' \leq C\|x\|$ for all $x \in X$, then $\|\cdot\| \sim \|\cdot\|'$.

Proof. $\mathbb{I} : (X, \|\cdot\|) \to (X, \|\cdot\|')$ is a continuous linear bijection. Apply theorem 4, the inverse is also bounded, so the norms are equivalent. □

Example. On $C[0,1]$, $\|f\|_1 \leq \|f\|_\infty$. We know that $\|\cdot\|_1 \nsim \|\cdot\|_\infty$, so $\|\cdot\|_1$ is not complete (though the direct proof is much easier).

For a function $f : X \to Y$ between sets, the *graph* of f is the set $\Gamma(f) = \{(x,y) \in X \times Y : y = f(x)\}$. If X, Y are topological spaces, f is continuous, and Y is Hausdorff, then $\Gamma(f)$ is closed in $X \times Y$ with the product topology. The converse is false in general but true if Y is compact.

Suppose now that X and Y are Banach spaces and $T : X \to Y$ is linear.

Theorem (Closed Graph Theorem)**.** Let X, Y be Banach spaces, $T : X \to Y$ linear. If $\Gamma(T)$ is closed in $X \times Y$, then T is continuous

Proof. Recall that $X \times Y$ is a Banach space with norm $\|(x,y)\| = \|x\| + \|y\|$, which also induces the product topology. $U = \Gamma(T)$ is a closed subspace, so U is a Banach space. Consider the projection $P : U \to X$, $P(x,y) = x$. P is bounded since $\|P(x,y)\| = \|x\| \leq \|x\| + \|y\| = \|(x,y)\|$. P is also a linear bijection, with P^{-1} given by $x \mapsto (x, Tx)$. By theorem 4, P^{-1} is bounded, so there exists $C \geq 0$ such that $\|P^{-1}x\| = \|(x,Tx)\| = \|x\| + \|Tx\| \leq C\|x\|$. In particular, $\|Tx\| \leq C\|x\|$ for all $x \in X$. □

This is often used in the following form:

Theorem. Let X, Y, T be as above. Assume that whenever $x_n \to 0$ in X and $Tx_n \to y$ in Y, we have $y = 0$. Then T is continuous.

Proof. Suppose $x_n \to x$ and $Tx_n \to y$ for a sequence $\big((x_n, Tx_n)\big)$ in $\Gamma(f)$. Then $x_n - x \to 0$ and $T(x_n - x) \to y - Tx$. By the hypothesis, $y - Tx = 0$, so $(x,y) \in \Gamma(T)$. Hence $\Gamma(f)$ is closed, and theorem 4 gives that T is continuos. □

Example. Here are some applications of the above theorems

(1) Let X be a closed subspace of l_2 which is also a subset of $l_1 \subseteq l_2$, then there exists $C \geq 0$ such that $\|x\|_1 \leq C\|x\|_2$ for all $x \in X$.

Proof. Define $T : X \to l_1, Tx = x$. Since X is closed, it is complete. Both X and l_1 are Banach spaces, and T is linear. Suppose $x^{(n)} \to 0$ in X and $x^{(n)} \to y$ in l_1, then for a fixed i as $n \to \infty$, $|x_i^{(n)}| \le \|x^{(n)}\|_2 \to 0$ and $|x_i^{(n)} - y_i| \le \|x^{(n)} - y\| \to 0$. Therefore $y_i = 0$ for all i, i.e. $y = 0$. Hence T is continuous by theorem 4. □

(2) Let X be a normed sapce. Let (x_n) be a sequence in X. Assume that $\sum |f(x_n)| < \infty$ for all $f \in X^*$ (such (x_n) is called *weakly unconditionally Cauchy*). Prove that there exists $M \ge 0$ such that $\sum |f(x_n)| \le M\|f\|$ for all $f \in X^*$.

Proof. Define $T : X^* \to l_1$, $T(f) = \big(f(x_n)\big)_{n=1}^{\infty}$. X^* and l_1 are Banach spaces, and T is linear. Suppose $f_n \to 0$ in X^* and $T(f_n) \to y$ in l_1. Write $y = (y_i)$.

$$\left.\begin{array}{ll} |f_n(x_i) - y_i| \le \|T(f_n) - y\|_1 \to 0 & \text{as } n \to \infty \\ |f_n(x_i)| \le \|f\| \cdot \|x_i\| \to 0 & \text{as } n \to \infty \end{array}\right\} \Rightarrow \; y_i = 0 \; \forall i \in \mathbb{N}$$

By the closed graph theorem, T is continuous. □

Alternative Proof. For $n \in \mathbb{N}$, define $T_n : X^* \to l_1$ by

$$T_n(f) = (f(x_1), \cdots, f(x_n), 0, 0, \cdots)$$

This is linear. $\|T_n(f)\|_1 = \sum_{k=1}^{N} |f(x_k)| \le \|f\| \sum_{k=1}^{N} \|x_k\|$, so T_n is bounded. Next, $T_n(f) \to T(f)$ as $n \to \infty$ since $\|T_n(f) - T(f)\|_1 = \sum_{k>n} |f(x_k)| \to 0$ as $n \to \infty$. Thus $T_n \to T$ pointwise. By the Banach-Steinhaus theorem, T is bounded. □

(3) Suppose X is a Banach space and $B \subseteq X$ is an algebraic basis for X, then B cannot be countably infinite.

Proof. Suppose not. List the elements of B as $x_1, x_2, \cdots$. Let $F_n = \text{span}\{x_1, \cdots, x_n\}$, then $X = \bigcup_{n=1}^{\infty} F_n$. F_n is closed since it is finite dimensional. $\mathbb{Z}F_n = \emptyset$ since given $x \in F_n$ and $r > 0$, $x + r\frac{x_{n+1}}{2\|x_{n+1}\|} \in B(x,r) \backslash F_n$, i.e. $B(x,r) \nsubseteq F_n$. This contradicts the Baire category theorem. □

(4) Projections: Let X be a vector space and let Y, Z be subspaces such that $Y \cap Z = \{0\}$ and $Y + Z = X$, i.e. X is the algebraic direct sum of Y and Z. Then $T : Y \times Z \to X$, $(y,z) \mapsto y + z$ is a linear bijection. We also have $P : X \to X$, $y + z \mapsto y$ for $y \in Y,\ z \in Z$. P is linear, $P^2 = P$, $\Im P = Y$, and $\ker P = Z$. P is called the *projection* of X onto Y along Z. Note that $(I - P)(y + z) = z$, so $I - P$ is the projection of X onto Z along Y.
Now assume X is a Banach space. T is continuous since $\|T(y,z)\| = \|y + z\| \le \|y\| + \|z\| = \|(y,z)\|$. If T is an isomorphism, then we say X is the *topological direct sum* of Y and Z, and write $X = Y \oplus Z$. The following are equivalent:

(i) Y, Z are closed.

(ii) $X = Y \oplus Z$.

(iii) P is continuous.

Proof. (i) $\Rightarrow$ (ii): Y, Z closed $\Rightarrow$ Y, Z are complete $\Rightarrow$ $Y \times Z$ is complete. Therefore, T is a continuous linear bijection between Banach spaces. By theorem 4, T^{-1} is continuous.
(ii) $\Rightarrow$ (iii): $\|P(y + Z)\| = \|y\| \le \|y\| + \|z\| = \|(y,z)\| \le \|T^{-1}\|\|y + z\|$.
(iii) $\Rightarrow$ (i): P continuous $\Rightarrow$ $I - P$ is continuous, so $Y = \ker(I - P)$ and $Z = \ker P$ are closed. □

We can also give a direct proof of (i) ⇒ (iii): Assume $x_n \to 0$ in X and $P(x_n) \to y$ in X. Write $x_n = y_n + z_n$ for $y_n \in Y$ and $z_n \in Z$. Then $P(x_n) = y_n \to y$. Since Y is closed, $y \in Y$. In addition, $z_n = x_n - y_n \to -y \in Z$ as Z is closed, so $y \in Y \cap Z = \{0\}$, so $y = 0$. By the closed graph theorem, P is continuous.

5 $C(K)$ spaces

First let K be a set. Let $l_\infty(K)$ be the space of scalar-valued bounded function on K. This is a Banach space in the norm $\|f\|_\infty = \sup_{x \in K} |f(x)|$. For example, $l_\infty = l_\infty(\mathbb{N})$.

Now let K be a topological space. Take $C_b(K) = \{f \in l_\infty(K) : f \text{ continuous}\}$, then this is a closed subspace of $l_\infty(K)$, so it is a Banach space.

Finally, take K to be a compact topological space, then $C_b(K) = C(K)$ =the set of continuous functions $K \to \mathbb{R}$. This chapter is about the Banach space $(C(K), \|\cdot\|_\infty)$.

5.1 Completions

Let (M, d) be a metric space. A *completion* of (M, d) is a complete metric space $(\tilde{M}, \tilde{d})$ together with an isometric map $j : M \to \tilde{M}$ such that $\tilde{M} = \overline{j(M)}$.

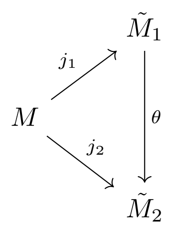
Theorem. Every metric space (M, d) has a completion.

Proof. WLOG $M \neq \emptyset$. Fix $x_0 \in M$. For $x \in M$, define $f_x : M \to \mathbb{R}$, $f_x(y) = d(y, x) - d(y, x_0)$. Then $|f_x(y)| \leq d(x, x_0)$ for all y, so $f_x \in l_\infty(M)$, which is a Banach space with $\|\cdot\|_\infty$. Define $j : M \to l_\infty(M)$, $x \mapsto f_x$. Given $x, z \in M$, and for $y \in M$,

$$|f_x(y) - f_z(y)| = |d(y, x) - d(y, z)| \leq d(x, z)$$

with equality achieved if $y = x$. Therfore, $\|f_x - f_z\|_\infty = d(x, z)$, and so j is an isometry. Put $\tilde{M} = \overline{j(M)}$. Then $\tilde{M}$ is a complete metric space with metric $d(f, g) = \|f - g\|_\infty$. □

Remark. Suppose $(\tilde{M}_k, \tilde{d}_k),\ k = 1, 2$ are both completions of M with corresponding isometric embeddings $j_k : M \to \tilde{M}_k,\ k = 1, 2$. Then there exists a unique isometry $\theta : \tilde{M}_1 \to \tilde{M}_2$ such that the diagram commutes:



Hence completions are unique up to isometry.

Theorem. The completion $(\tilde{X}, \tilde{d})$ of a normed space X is a Banach space.

Proof. Let $j : X \to \tilde{X}$ be an isometry such that $j(X)$ is dense in $\tilde{X}$. Given $x, y \in \tilde{X}$, choose sequences (x_n), (y_n) in X such that $j(x_n) \to x$ and $j(y_n) \to y$ as $n \to \infty$. Define

$$x + y = \lim_{n\to\infty} j(x_n + y_n), \ \lambda x = \lim_{n\to\infty} j(\lambda x), \ \|x\| = \lim_{n\to\infty} \|x_n\|$$

Then it is routine to verify that the operations are well-defined, $\tilde{X}$ becomes a vector space, $\|\cdot\|$ is a norm on $\tilde{X}$, $\tilde{d}(x, y) = \|x - y\|$, and j is linear. □

Alternative Proof. Recall that the canonical map $X \to X^{**}$, $x \mapsto \hat{x}$ is an isometric embedding of X into X^{**}, which is complete. Take $\tilde{X} = \overline{\{\hat{x} : x \in X\}}$, then this is a Banach space containing X as a dense subset. □

5.2 Urysohn's lemma

We now go back to studying $C(K)$ spaces. A topological space K is *normal* if given disjoint closed subsets $E, F \subseteq K$, there exists disjoint open sets $U, V \subseteq K$ such that $E \subseteq U$, $F \subseteq V$.

Proposition. If K is a metric space or a compact Hausdorff space, then K is normal.

Proof. Metric spaces: If $A \subseteq K$, $A \neq \emptyset$, define $d(x, A) = \inf\{d(x, a) : a \in A\}$. This is a continuous function on K, and $d(x, A) = 0 \iff x \in \overline{A}$. Given closed subsets E and F, take $U = \{x \in K : d(x, E) < d(x, F)\}$, and $V = \{x \in K : d(x, E) > d(x, F)\}$, then U and V are open and disjoint. If $x \in E$, then $d(x, E) = 0$, and in addition $x \notin F = \bar{F}$, so $d(x, F) \neq 0$. It follows that $E \subseteq U$. Similarly, $F \subseteq V$.

Compact Hausdorff spaces: Fix $x \in E$. For all $y \in F$, there exists disjoint open sets U_y, V_y such that $x \in U_y$ and $y \in V_y$. The family $\{V_y : y \in F\}$ is an open cover for F, which is compact since it is closed. Therefore, there exists $y_1, \cdots, y_m$ in F such that $F \subseteq \bigcup_{i=1}^m V_{y_i} = V_x$. Take $U_x = \bigcap_{i=1}^m U_{y_i}$, then U_x and V_x are open disjoint, and $x \in U_x$, $F \subseteq V_x$, so the space is *regular* (a point and a closed set not containing the point can be separated by two open sets).

Now, $\{U_x : x \in E\}$ is an open cover for compact set E, so there exists a $x_1, \cdots, x_n \in E$ such that $E \subseteq \bigcup_{j=1}^n U_{x_j}$. Take $U = \bigcup_{j=1}^n U_{x_j}$, and $V = \bigcap_{j=1}^n V_{x_j}$, then they are disjoint open sets with $E \subseteq U$ and $F \subseteq V$. □

Theorem (Urysohn's Lemma)**.** Let K be a normal topological space. If E and F are disjoint closed sets in K, then there exists a continuous function $f : K \to [0, 1]$ such that $f = 0$ on E and $f = 1$ on F.

Idea of Proof: If such f exists, then for $t > 0$, define $F_t = \{x \in K : f(x) < t\}$. This family satisfies three properties

(i) F_t is open.

(ii) $\bigcup_{t>0} F_t = K$.

(iii) $s < t \Rightarrow \overline{F_s} \subseteq F_t$.

Furthermore, $\{t > 0 : x \in F_t\} = (f(t), \infty)$, so f is determined by such family of sets.

Lemma. Let $\mathbb{T} = \mathbb{Q} \cap [0, \infty)$. Let $F_t, t \in \mathbb{T}$ be open sets in a topological space K such that (1) $\bigcup_{t \in \mathbb{T}} F_t = K$ (2) $s < t \Rightarrow \overline{F_s} \subseteq F_t$. Then $f(x) = \inf\{t \in \mathbb{T} : x \in F_t\}$ defines a continuous function $K \to [0, \infty)$.

Proof. By condition (1), f is well-defined. Let $b > 0$, then the pre-image

$$\{x \in K : f(x) < b\} = \{x \in K : \exists t \in \mathbb{T}, t < b \text{ s.t. } x \in F_t\} = \bigcup_{\substack{t \in \mathbb{T} \\ t < b}} F_t$$

is open. Now consider $\{x \in K : f(x) \leq b\}$. If $f(x) \leq b$, then for all $t \in \mathbb{T}$ with $t > b$, we have $t > f(x)$, so there exists $s \in \mathbb{T}, s < t$ such that $x \in F_s \subseteq F_t$. Conversely, if $x \in F_t$ for all $t \in \mathbb{T}$ with $t > b$, then $f(x) \leq t$ for all such t, so $f(x) \leq b$ since T is dense. Therefore

$$\{x \in K : f(x) \leq b\} = \bigcap_{\substack{t \in \mathbb{T} \\ t > b}} F_t = \bigcap_{\substack{t \in \mathbb{T} \\ t > b}} \overline{F_t}$$

is closed. Thus, the pre-image of $(-\infty, b)$ is open. The families (b, ∞) and $(-\infty, b)$ form a subbase of $[0, \infty)$, so f is continuous. □

Remark. Suppose K is normal, $E \subseteq U \subseteq K$, E is closed, and U is open. Then there exists an open set V such that $E \subseteq V \subseteq \overline{V} \subseteq U$. To see this, note that E and $K \backslash U$ are disjoint closed sets, so there exists disjoint open sets V and W such that $E \subseteq V$, $K \backslash U \subseteq W$. Then $E \subseteq V \subseteq K \backslash W \subseteq U$. Since $K \backslash W$ is closed, $\overline{V} \subseteq K \backslash W \subseteq U$.

Proof of Urysohn's Lemma. Enumerate $\mathbb{T}$ as $q_0 = 0, q_1 = 1, q_2, q_3, \cdots$. We will define sets F_{q_n} for $n \geq 0$ inductively. Set $F_0 = F_{q_0} = E$ and $F_1 = F_{q_1} = K \backslash F$. Assume for some $n \geq 1$, we have $F_{q_0}, F_{q_1}, \cdots, F_{q_n}$ satisfying F_{q_j} is open for $j = 1, 2, \cdots, n$ and $q_i < q_j \Rightarrow \overline{F_{q_i}} \subseteq F_{q_j}$. Choose permutation π of $\{0, 1, \cdots, n\}$ such that $q_{\pi(0)} < q_{\pi 1} < \cdots < q_{\pi(n)}$. There exists $i \in \{1, \cdots, n\}$ satisfying $q_{\pi(i-1)} < q_{n+1} < q_{\pi_i}$. By the previous remark, we can choose an open set $F_{q_{n+1}}$ such that $\overline{F_{q_{\pi(i-1)}}} \subseteq F_{q_{n+1}} \subseteq \overline{F_{q_{n+1}}} \subseteq F_{q_{\pi(i)}}$. This completes the inductive construction.
For $t \in \mathbb{T}, t > 1$, set $F_t = K$. By lemma 5.2, $f(x) = \inf\{t \in \mathbb{T} : x \in F_t\}$ defines a continuous function $K \to [0, \infty)$. For $x \in K$, $x \in F_t$ for every $t \in \mathbb{T}$ with $t > 1$, so $f(x) \leq 1$. If $x \in E = F_0$, then $x \in F_t$ for every $t \in T$, so $f(x) = 0$. If $x \in F$, then $x \notin F_1 = K \backslash F$, so $x \notin F_t$ for $t < 1$. Hence $f(x) = 1$. □

Remark.

(i) The normality condition is necessary, since if such f exists for any disjoint closed E and F, then $f^{-1}((0, 1/2))$ and $f^{-1}((1/2, 1))$ are open sets separating E and F.

(ii) If K is compact Hausdoff, then the theorem shows that $C(K)$ separates points of K, i.e. for all $x \neq y$ in K, there exists $f \in C(K)$ such that $f(x) \neq f(y)$.

(iii) For metric spaces, the following function provides a direct proof of the theorem:

$$f(x) = \frac{d(x, E)}{d(x, E) + d(x, F)}$$

Corollary. Let K be a normal space, and $U_1, \cdots, U_n$ open sets in K such that $K = \bigcup_{i=1}^n U_i$. Then there exists continuous functions $\phi_i : K \to [0,1]$, $i = 1, 2, \cdots, n$ such that $\phi_i = 0$ on $K \backslash U_i$, and $\sum_{i=1}^n \phi_i = 1$.

Proof. First we observe that if $E \subseteq K$ is closed, $U_1, \cdots, U_n$ are open sets which cover E, then there exists open sets $V_1, \cdots, V_n$ such that $\overline{V_i} = U_i$ and $E \subseteq \bigcup_{i=1}^n V_i$. This will be proven by induction: for $n = 1$, $E \subseteq U_1$, so by normality there exists an open set V such that $E \subseteq V \subseteq \overline{V} \subseteq U_1$. Let $n > 1$, then $E \backslash U_n \subseteq \bigcup_{i=1}^{n-1} U_i$, so there exists open sets $V_1, \cdots, V_{n-1}$ such that $E \backslash U_n \subseteq \bigcup_{i=1}^{n-1} V_i$ and $\overline{V_i} \subseteq U_i$ for $i = 1, \cdots, n-1$. Now, $E \backslash \bigcup_{i=1}^{n-1} V_i \subseteq U_n$, so there exists an open set V_n such that $E \backslash \bigcup_{i=1}^{n-1} V_i \subseteq V_n \subseteq \overline{V_n} \subseteq U_n$, so $V_1, \cdots, V_{n-1}, V_n$ satisfy the conditions.

Apply the result to $E = K$ with $U_1, \cdots, U_n$ given in the statement. We get open sets $V_1, \cdots, V_n$ such that $K = \bigcup_{i=1}^n V_i$ and $\overline{V_i} \subseteq U_i$ for $i = 1, 2, \cdots, n$. Apply Urysohn's lemma, for each i, there exists continuous functions $\psi_i : K \to [0,1]$ such that $\psi_i = 1$ on $\overline{V_i}$ and $\psi_i = 0$ on $K \backslash U_i$. Let $\psi = \sum_{i=1}^n \psi_i$, then since $K = \bigcup_{i=1}^n V_i$, $\psi \geq 1$ on K. Now $\phi_i = \frac{\psi_i}{\psi}$ for $i = 1, \cdots, n$ satisfy the conditions. □

Remark. The functions $\phi_1, \cdots, \phi_n$ form a *partition of unity* subordinate to the open cover $\{U_i\}$. They can be used to deduce global information from local information, by writing each function as a sum of several functions with small supports.

Theorem (Tietze's Extension Theorem)**.** Let K be a normal space and $L \subseteq K$ closed. Then for all $g \in C_b(L)$, there exists $f \in C_b(K)$ such that $f|_L = g$ and $\|f\|_\infty = \|g\|_\infty$.

Proof. First assume that g can be extended to an $f_1 \in C_b(K)$, we will show that the conclusion of the theorem holds: WLOG $\|g\|_\infty = 1$. Then for any scalar t, define

$$\varphi(t) = \begin{cases} t & \text{if } |t| \leq 1 \\ \frac{t}{|t|} & \text{if } |t| \geq 1 \end{cases}$$

Then φ is continuous. Set $f = \varphi \circ f_1$, then f is continuous, and $|f(x)| = |\varphi(f_1(x))| \leq 1$. In addition, for any x, $|g(x)| \leq 1$, so $f|_L = f_1|_L = g$. Therefore, $\|f\|_\infty = \|g\|_\infty = 1$, and f extends g.

Now assume the scalar field is $\mathbb{R}$, since otherwise one can extend real and imaginary parts separately and apply the above argument. Set $X = C_b(K)$ and $Y = C_b(L)$, then they are real Banach spaces. Define $R : X \to Y$, $R(f) = f|_L$, then R is a bounded linear map. We need to show R is surjective. Let $g \in B_Y$. Define

$$E = \{x \in L : g(x) \in [-1, -1/3]\}, F = \{x \in L : g(x) \in [1/3, 1]\}$$

Then E and F are disjoint closed subset of L, and hence of K. By theorem 5.2, there exists a continuous function $f : K \to \left[-\frac{1}{3}, \frac{1}{3}\right]$ such that $f|_E = -\frac{1}{3}$ and $f|_F = \frac{1}{3}$. Then $\|R(f) - g\|_\infty \leq \frac{2}{3}$ by considering their behaviour on E, F, and $L \backslash (E \cup F)$. Therefore, $R\left(\frac{1}{3} B_X\right)$ is $\frac{2}{3}$-dense in B_Y. By the open mapping lemma, $R(B_X) \supseteq B_Y$. □

Alternative Proof. We first show that $R(X)$ is complete. Suppose $\sum y_n$ in $R(X)$ is an absolutely convergent series, then for all n, there exists $x_n \in X$ such that $R(x_n) = y_n$. WLOG $\|x_n\| = \|y_n\|$ by composing with a suitable φ. Then $\sum x_n$ is absolutely convergent and so converges since X is complete. R is continuous, so $\sum y_n$ is also a convergent series. This shows that $R(X)$ is complete, and hence closed in Y. By Riesz's lemma, if $R(X) \neq Y$, then there exists $g \in B_Y$ such that $d(g, R(X)) > \frac{2}{3}$. This contradicts Urysohn's lemma, as used in the proof above. $\square$

5.3 Stone-Weierstrass theorem

Fix K to be a compact space. $C(K)$ is a real or complex Banach space, with $\|f\| = \sup_{x \in K} |f(x)|$. It has multiplication defined pointwise as $(f \cdot g)(x) = f(x)g(x)$. This makes $C(K)$ into an *algebra*, i.e. a vector space with a multiplication satisfying

$$(fg)h = f(gh),\ f(g+h) = fg + fh, (f+g)h = fh + gh,\ \lambda(fg) = (\lambda f)g = f(\lambda g)$$

for all $f, g, h \in C(K)$ and scalar λ.

Call an algebra A *unital* if there exists $u \in A$ such that $au = ua = a$ for all $a \in A$. It is *commutative* if $ab = ba$ for all $a, b \in A$. Then $C(K)$ is unital and commutative, with unit being the constant function 1.

A *normed algebra* is a normed space $(A, \|\cdot\|)$ with multiplication such that A is an algebra and $\|ab\| \leq \|a\| \cdot \|b\|$ for all $a, b \in A$. It is a *Banach algebra* if in addition, A is complete. Note that multiplication is always continuous in a normed algebra, since $a_n \to a$ and $b_n \to b$ implies

$$\|a_n b_n - ab\| \leq \|a_n - a\| \cdot \|b_n - b\| + \|a\| \cdot \|b_n - b\| + \|b\| \cdot \|a_n - a\| \to 0$$

Let $C^{\mathbb{R}}(K)$ be the space of real continuous functions from K, and $C^{\mathbb{C}}(K)$ be the space of complex continuous functions, then they are Banach algebras.

Recall that a subset $A \subseteq C(K)$ *separates points* of K if for all $x \neq y$, there exists $f \in A$ such that $f(x) \neq f(y)$. When K is compact Hausdorff, Urysohn's lemma implies that $C(K)$ seaprates points. Conversely, if $A \subseteq C(K)$ separates points for some A, then K must be Hausdorff: given $x \neq y$ in K, there exists $f \in C(K)$ such that $f(x) \neq f(y)$. Let U and V be disjoint open sets around $f(x)$ and $f(y)$ respectively, then $f^{-1}(U)$ and $f^{-1}(V)$ are disjoint open sets containing x and y respectively. Set A *strongly separates points* of K if it separates points, and for all $x \in K$, there exists $f \in A$ with $f(x) \neq 0$.

A subset $A \subseteq C(K)$ is a *subalgebra* if it is a subspace and it is closed under multiplication. It is a *unital subalgebra* if in addition, $1 \in A$. Note that if A is a subalgebra, then so is $\overline{A}$ since multiplication is continuous.

Theorem (Stone-Weierstrass Theorem, real version)**.** Let K be a compact space, and let A be a subalgebra of $C^{\mathbb{R}}(K)$ that strongly separates points of K, then A is dense in $C^{\mathbb{R}}(K)$, i.e. $\overline{A} = C^{\mathbb{R}}(K)$.

Lemma. For all $\epsilon > 0$, there exists a real polynomial p such that $p(0) = 0$ and $|p(t) - |t|| < \epsilon$ for all $t \in [-1, 1]$.

Proof. Let $\epsilon > 0$. Let f be an analytic square root of $z + \epsilon^2$ on $\mathbb{C} \backslash (-\infty, -\epsilon^2]$ with $f(x) = \sqrt{x + \epsilon^2} > 0$ for all $x \in \mathbb{R},\ x > -\epsilon^2$. f has a power series expansion $\sum_{n=0}^{\infty} a_n (z-1)^n$ around 1, which has radius of convergence greater than 1, so it converges uniformly on the compact set $[0, 1]$. Choose $N \in \mathbb{N}$ such that $q(x) = \sum_{n=0}^{N} a_n (x-1)^n$ satisfies $|q(x) - f(x)| < \epsilon$ for all $x \in [0, 1]$. $\sqrt{x} \leq f(x) =$

$\sqrt{x+\epsilon^2} \le \sqrt{x}+\epsilon$, so $|f(x)-\sqrt{x}| \le \epsilon$ for all $x \in [0,1]$. Thus, $|q(x)-\sqrt{x}| < 2\epsilon$. In particular, $|q(0)| < 2\epsilon$. Put $p(t) = q(t^2) - q(0)$, then $p(0) = 0$ and

$$|p(t) - |t|| = |q(t^2) - q(0) - \sqrt{t^2}| \le |q(t^2) - \sqrt{t^2}| + |q(0)| < 4\epsilon$$

Furthermore p is a real polynomial since $a_n = \frac{1}{n!} f^{(n)}(1) \in \mathbb{R}$, so p satisfies the conditions. ☐

Corollary. Let A be a closed subalgebra of $C^{\mathbb{R}}(K)$, then A is a *lattice*: for all $f, g \in A$, $\min(f,g),\ \max(f,g) \in A$.

Proof. It is enough to show that $f \in A \ \Rightarrow\ |f| \in A$, since

$$\max(f,g) = \frac{1}{2}(f+g+|f-g|),\ \min(f,g) = \frac{1}{2}(f+g-|f-g|)$$

By scaling, WLOG $\|f\| \le 1$. Let $\epsilon > 0$, and let p be as in lemma 5.3. Write $p(t) = \sum_{k=1}^{n} a_k t^k$, with $a_k \in \mathbb{R}$ for $k = 1, 2, \cdots, n$. Then $p(f) = \sum_{k=1}^{n} a_k f^k \in A$. Now, for all $x \in K$, $|f(x)| \in [-1,1]$, so $|p(f(x)) - |f(x)|| < \epsilon$. Hence $\|p(f) - |f|\| < \epsilon$. Since ϵ is arbitrary, $|f| \in \overline{A} = A$. ☐

Proof of Theorem 5.3. Replace A by $\overline{A}$. We need to show that $A = C^{\mathbb{R}}(K)$. By corollary 5.3, A is a lattice. We will first prove that for all $x \neq y$ in K and $a, b \in \mathbb{R}$, there exists $f \in A$ with $f(x) = a$ and $f(y) = b$. Define $T : A \to \mathbb{R}^2, Tf = (f(a), f(b))$. It will be shown that $T(A)$ contains two linearly independent vectors, so it is surjective. Let $f \in A$ be chosen such that $f(x) \neq f(y)$. WLOG $f(x) \neq 0$. There are two cases.

Case 1: $f(y) \neq 0$. Then $T(f)$ and $T(f^2)$ are linearly independent.

Case 2: $f(y) = 0$. Pick $g \in A$ such that $g(y) \neq 0$. Then there exists λ such that $(f+\lambda g)(x) \neq 0$ and $(f+\lambda g)(y) \neq 0$. Then $T(f)$ and $T(f+\lambda g)$ are independent.

Now fix $f \in C^{\mathbb{R}}(K)$ and $\epsilon > 0$. Fix $x \in K$. For all $y \in K$, choose $g_{x,y} \in A$ such that $g_{x,y}(x) = f(x)$ and $g_{x,y}(y) = f(y)$. By continuity, there exists an open neighbourhood U_y of y such that $|f(z) - g_{x,y}(z)| < \epsilon$ for all $z \in U_y$. The family $\{U_y : y \in K\}$ is an open covering of K, and K is compact, so there exists $y_1, \cdots, y_m \in K$ such that $K = \bigcup_{i=1}^{m} U_{y_i}$. Then $g_x = \min_{1 \le i \le m} g_{x,y_i} \in A$, with $g_x(x) = x$ and $g_x < f + \epsilon$ on K.

For all x, ther exists an open neighbourhood V_x such that $|g_x(z) - f(z)| < \epsilon$ for all $z \in V_x$. The family $\{V_x : x \in K\}$ is an open covering of K, so there exists $x_1, \cdots, x_n$ such that $K = \bigcup_{j=1}^{n} V_{x_j}$. Let $h = \max_{1 \le j \le n} g_{x_j} \in A$, then $f - \epsilon < h < f + \epsilon$ on K, so $\|h - f\| < \epsilon$. Since $h \in A$ and ϵ is arbitrary, we proved that $f \in \overline{A} = A$. ☐

This result fails for scalar field $\mathbb{C}$: take $\Delta = \{z \in \mathbb{C} : |z| \le 1\}$, then it is a compact set. Let $A = \{f : \Delta \to \mathbb{C} : f \text{ continuous on } \Delta,\ f \text{ analytic on } \mathbb{Z}\Delta\}$, then it is a subalgebra of $C(\Delta)$ containing 1 and the identity, so A strongly separates points of Δ. But A is closed in $\|\cdot\|_\infty$, so $\overline{A} = A \neq C(\Delta)$.

Theorem (Stone-Weierstrass Theorem, complex version)**.** Let K be a compact space, and let A be a subalgebra of $C^{\mathbb{C}}(K)$ that strongly separates the points of K and is closed under complex conjugation, then A is dense in $C^{\mathbb{C}}(K)$.

Proof. Let $A^{\mathbb{R}} = A \cap C^{\mathbb{R}}(K)$, then it is a real subalgebra of $C^{\mathbb{R}}(K)$. For $f \in A$, $\mathbb{R}ef = \frac{1}{2}(f + \bar{f})$, and $\Im f = \frac{1}{2i}(f - \bar{f})$ are both in $A^{\mathbb{R}}$, so $A^{\mathbb{R}}$ strongly separates the points of K. By theorem 5.3, $\overline{A^{\mathbb{R}}} = C^{\mathbb{R}}(K)$. It follows that $\overline{A} \supseteq C^{\mathbb{R}}(K) + iC^{\mathbb{R}}(K) = C^{\mathbb{C}}(K)$. ☐

Example. Some applications of the theorem:

(1) The polynomials are dense in $C[a, b]$ (Weierstrass approximation theorem).

(2) Given a compact set $K \subseteq \mathbb{R}^d$, the polynomials on K are dense in $C(K)$, where polynomials are functions of the form

$$t = (t_1, \cdots, t_d) \mapsto \sum_{\alpha=(\alpha_1,\cdots,\alpha_d)\in F} a_\alpha t_1^{\alpha_1} t_2^{\alpha_2} \cdots t_d^{\alpha_d}$$

for F a finite subset of $(\mathbb{Z}_{\geq 0})^d$.

(3) Let $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. The trigonometric polynomials are dense on $C(\mathbb{T})$, where trigonometric polynomials are functions of the form $z \mapsto \sum_{k=-n}^{n} a_k z^k$ for $a_{-n}, \cdots, a_n \in \mathbb{C}$.

(4) Let K, L be compact Hausdorff spaces. The functions on $K \times L$ of the form $(x, y) \mapsto \sum_{i=1}^{n} f_i(x) g_i(y)$, for $f_i \in C(K)$, $g_i \in C(L)$, are dense in $C(K \times L)$ (it separates points by Urysohn's lemma). It follows that for all $f \in C\big([0, 1]^2\big)$,

$$\int_0^1 \int_0^1 f(x, y) dx dy = \int_0^1 \int_0^1 f(x, y) dy dx$$

since the identity holds for all $h(x, y) = \sum_{i=1}^{n} f_i(x) g_i(y)$, $f_i, g_i \in C[0, 1]$.

(5) If K is a compact metric space, then $C(K)$ is separable: let $\{x_1, x_2, \cdots\}$ be a countable dense subset of K. Let $f_0(x) = 1$, and for $n \in \mathbb{N}$, $f_n(x) = d(x, x_n)$. Let A be the subalgebra of $C(K)$ generated by $f_0, f_1, \cdots$, then

$$A = \operatorname{span} \left\{ \prod_{k \in I} f_k : I \text{ finite subset of } \mathbb{Z}_{\geq 0} \right\}$$

which is separable since it is spanned by a countable set. If $x \neq y$ in K, then there exists $n \in \mathbb{N}$ such that $d(x, x_n) < \frac{1}{3} d(x, y)$, then $d(y, x_n) > \frac{2}{3} d(x, y)$, so $f_n(x) \neq f_n(y)$. In addition, $f_0 = 1 \in A$, so A strongly separates the points of K. Since f_n are real valued, A is also closed under complex conjugation in the complex case. Hence $\overline{A} = C(K)$.

Remark. The following converse is also true: if $C(K)$ is separable, and K is ompact Hausdorff, then K is metrizable.

5.4 Fourier series

Recall that $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. We will often identify $\mathbb{T}$ with $\mathbb{R}/2\pi\mathbb{Z}$, and think of $f \in C(\mathbb{T})$ as a 2π-period function on $\mathbb{R}$, and write $f(z) = f(e^{it}) = f(t)$, for $z \in \mathbb{T}$, $t \in \mathbb{R}$.

Given $f \in C(\mathbb{T})$ and integer n, we define the n-th *Fourier coefficient* of f to be

$$\hat{f}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta}) e^{-in\theta} d\theta$$

The *Fourier series* of f is the series $\sum_{n=-\infty}^{\infty} \hat{f}_n z^n = \sum_{n=-\infty}^{\infty} \hat{f}_n e^{int}$. We will try to study when the series converges, and in which mode of convergence.

For $N \in \mathbb{Z}_{\geq 0}$, let $s_N(f)(z) = \sum_{n=-N}^{N} \hat{f}_n z^n$, then this is a trigonometric polynomial. For a trigonometric polynomial $g = \sum_{k=-m}^{m} a_k z^k$, $s_N(g) = g$ for all $N \geq m$, so in particular, $s_N(g) \to g$ uniformly. In general,

$$\begin{aligned} s_N(f)(t) = \sum_{n=-N}^{N} \hat{f}_n e^{int} &= \sum_{n=-N}^{N} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-in\theta} d\theta e^{int} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) \sum_{n=-N}^{N} e^{in(t-\theta)} d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) D_N(t-\theta) d\theta = (f * D_N)(t) \end{aligned}$$

where $D_N(t) = \sum_{n=-N}^{N} e^{int} \in C(\mathbb{T})$ is the *Dirichlet kernel*, and for $f, g \in C(\mathbb{T})$, we define their *convolution* by

$$(f * g)(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) g(t-\theta) d\theta \in C(\mathbb{T})$$

It can be checked that this is commutative, associative, and $(\widehat{f * g})_n = \hat{f}_n \hat{g}_n$. To further analyze $D_N(t)$, we can rewrite it as

$$\begin{aligned} D_N(t) &= \sum_{n=-N}^{N} e^{int} = e^{-iNt} \frac{e^{(2N+1)it} - 1}{e^{it} - 1} \\ &= \frac{e^{i\left(N+\frac{1}{2}\right)t} - e^{-i\left(N+\frac{1}{2}\right)t}}{e^{\frac{it}{2}} - e^{-\frac{it}{2}}} = \frac{\sin\left(N+\frac{1}{2}\right)t}{\sin\frac{t}{2}} \quad (t \neq 0) \\ D_N(0) &= 2N+1 \end{aligned}$$

We have $\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(t) dt = 1$ and $D_N(0) = 2N+1 = \|D_N\|_\infty$. For large N, "$D_N \approx \delta$", the Dirac δ-function. Define linear operators $T_N : C(\mathbb{T}) \to \mathbb{C}$, $T_N(f) = s_N(f)(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) D_N(\theta) d\theta$, then we have the following result.

Lemma. $T_N \in C(\mathbb{T})^*$, with $\|T_N\| = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(t)| dt \to \infty$ as $N \to \infty$.

Proof. T_N is clearly linear, and

$$T_N(f) \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(\theta)||D_N(\theta)| d\theta \leq \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(\theta)| d\theta \right) \cdot \|f\|_\infty$$

so T_N is bounded with $\|T_N\| < \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(\theta)| d\theta$. To prove the reverse inequality, denote the zeroes of D_N over $[-\pi, \pi]$ by $z_1, \cdots, z_{2N}$ in increasing order. Choose $\delta > 0$ such that $(z_j - \delta, z_j + \delta)$ are pairwise disjoint and contained in $[-\pi, \pi]$. Let $f = \operatorname{sign}(D_N)$ on $I = [-\pi, \pi] \setminus \bigcup_{j=1}^{2N} (z_j - \delta, z_j + \delta)$, and extend it to $[-\pi, \pi]$ linearly, then $f \in C(\mathbb{T})$, $\|f\|_\infty = 1$, and

$$\|T_N\| \geq |T_N(f)| \geq \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(\theta)| d\theta - \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) D_N(\theta) d\theta - \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(\theta)| d\theta \right|$$

The error term can be estimated by

$$
\begin{aligned}
|\frac{1}{2\pi}\int_{-\pi}^{\pi}(|D_N(\theta)|-f(\theta)D_N(\theta))\,d\theta| &= |\frac{1}{2\pi}\int_{[-\pi,\pi]\setminus I}(|D_N(\theta)|-f(\theta)D_N(\theta))\,d\theta| \\
&\leq \frac{1}{2\pi}2\delta\cdot 2\|D_N\|_\infty\cdot 2N = \frac{4N(2N+1)}{\pi}\delta
\end{aligned}
$$

Therefore, $\|T_N\| \geq \frac{1}{2\pi}\int_{-\pi}^{\pi}|D_N(\theta)|d\theta - \frac{4N(2N+1)}{\pi}\delta$ for all sufficiently small $\delta > 0$. Hence $\|T_N\| \geq \frac{1}{2\pi}\int_{-\pi}^{\pi}|D_N(\theta)|d\theta$ as required.

Now, over the interval $I_k = \left[\frac{k\pi+\pi/4}{N+1/2}, \frac{k\pi+3\pi/4}{N+1/2}\right]$, we have

$$
|D_N(\theta)| = |\frac{\sin\left(N+\frac{1}{2}\right)\theta}{\sin\frac{\theta}{2}}| \geq \frac{\sqrt{2}}{\theta} \geq \frac{\sqrt{2}\left(N+\frac{1}{2}\right)}{k\pi+\frac{3\pi}{4}} \geq \frac{N+\frac{1}{2}}{(k+1)\pi}
$$

Therefore,

$$
\begin{aligned}
\|T_N\| = \frac{1}{2\pi}\int_{-\pi}^{\pi}|D_N(\theta)|d\theta &\geq \sum_{k=0}^{N-1}\frac{1}{2\pi}\int_{I_k}|D_N(\theta)|d\theta \geq \sum_{k=0}^{N-1}\left(\frac{1}{2\pi}\cdot\frac{\pi/2}{N+\frac{1}{2}}\cdot\frac{N+\frac{1}{2}}{(k+1)\pi}\right) \\
&= \frac{1}{4\pi}\sum_{k=0}^{N-1}\frac{1}{k+1} \geq \frac{1}{4\pi}\int_1^N\frac{dx}{x} = \frac{1}{4\pi}\log n
\end{aligned}
$$

Hence $\|T_N\| \to \infty$ as $N \to \infty$. □

Corollary. There exists $f \in C(\mathbb{T})$ such that $T_N(f) \not\to f(0)$ as $N \to \infty$. In particular, $s_N(f) \not\to f$ pointwise or uniformly.

Proof. $\sup_N \|T_N\| = \infty$, so $\{T_N : N \in \mathbb{N}\}$ is not uniformly bounded. By the principle of uniform boundedness (theorem 4), $\{T_N : N \in \mathbb{N}\}$ is not pointwise bounded, so there exists $f \in C(\mathbb{T})$ such that $(T_N(f))_{N=1}^{\infty}$ is not bounded. □

To remedy the situation, take averages. Define $\sigma_N(f) = \frac{1}{N+1}\sum_{n=0}^{N}s_N(f) = f * K_N$, where $K_N = \frac{1}{N+1}\sum_{n=0}^{N}D_N$ is the N-th *Fejér kernel*. It will be shown that $\sigma_N(f) \to f$ uniformly for all $f \in C(\mathbb{T})$, which gives a new proof that the trigonometric polynomials are dense in $C(\mathbb{T})$. We will first compute $K_N(t)$.

$$
\begin{aligned}
K_N(t) &= \frac{1}{N+1}\sum_{n=0}^{N}e^{-int}\frac{e^{(2n+1)it}-1}{e^{it}-1} = \frac{1}{N+1}\cdot\frac{1}{e^{it}-1}\sum_{n=0}^{N}\left(e^{i(n+1)s}-e^{ins}\right) \\
&= \frac{1}{N+1}\cdot\frac{1}{\left(e^{it}-1\right)^2}\cdot\left(e^{i(N+2)t}-2e^{it}+e^{-iNt}\right) \\
&= \frac{1}{N+1}\cdot\frac{e^{i(N+1)t}-2+e^{-i(N+1)t}}{\left(e^{\frac{it}{2}}-e^{-\frac{it}{2}}\right)^2} \\
&= \frac{1}{N+1}\left(\frac{e^{\frac{i(N+1)t}{2}}-e^{-\frac{i(N+1)t}{2}}}{e^{\frac{it}{2}}-e^{-\frac{it}{2}}}\right)^2 = \frac{1}{N+1}\left(\frac{\sin\left(\frac{N+1}{2}s\right)}{\sin\left(\frac{s}{2}\right)}\right)^2
\end{aligned}
$$

This has the following three properties

(i) $\frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(t)dt = 1$

(ii) $K_N \geq 0$ on $[-\pi, \pi]$

(iii) $K_N \to 0$ uniformly on $[-\pi,\pi]\backslash(-\delta,\delta)$ for all $\delta > 0$, since

$$K_N(t) \leq \frac{1}{N+1} \cdot \frac{1}{\left(\frac{2}{\pi}\cdot\frac{t}{2}\right)^2} \leq \frac{\pi^2}{(N+1)\delta^2} \to 0 \text{ as } N \to \infty$$

Theorem. Let $K_N \in C(\mathbb{T})$ for $N \geq 0$ be functions satisfying the above three properties, then $f * K_N \to f$ uniformly for all $f \in C(\mathbb{T})$.

Proof. Let $f \in C(\mathbb{T})$ and $\epsilon > 0$. f is uniformly continuous on $[-2\pi, \pi]$ and has period 2π, so there exists $\delta > 0$ such that for all $s, t \in \mathbb{R}$, $|s-t| < \delta$ implies $|f(s)-f(t)| < \epsilon$. By property (3), there exists $N_0 \in \mathbb{N}$ such that $K_N(t) < \epsilon$ for all $s \in [-\pi,\pi]\backslash(-\delta,\delta)$ and $N \geq N_0$. Then for those N,

$$\begin{aligned} |f * K_N(t) - f(t)| &= |\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t-s)K_N(s)ds - f(t)| \\ &= |\frac{1}{2\pi}\int_{-\pi}^{\pi} (f(t-s) - f(t))K_N(s)ds| && \text{by (1)} \\ &\leq \left(\frac{1}{2\pi}\int_{|s|\leq\delta} + \frac{1}{2\pi}\int_{|s|>\delta}\right) |f(t-s) - f(t)|K_N(s)ds && \text{by (2)} \\ &\leq \frac{1}{2\pi}\int_{|s|\leq\delta} \epsilon K_N(s)ds + \frac{1}{2\pi}\int_{|s|>\delta} 2\|f\|_\infty \epsilon ds \\ &\leq \epsilon + 2\|f\|_\infty\epsilon = (1+2\|f\|_\infty)\epsilon \end{aligned}$$

This holds for all $t \in \mathbb{T}$, so $\|f * K_N - f\|_\infty \leq (1+2\|f\|_\infty)\epsilon$ for all $N \geq N_0$. □

Note that for $f \in C(\mathbb{T})$, $|\hat{f}_n| = |\frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})e^{in\theta}d\theta| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(e^{i\theta})|d\theta \leq \|f\|_\infty$, so $(\hat{f}_n)_{n\in\mathbb{Z}} \in l_\infty(\mathbb{Z})$, the space of bounded functions on $\mathbb{Z}$.

Proposition (Riemann-Lebesgue Lemma)**.** $\hat{f}_n \to 0$ as $|n| \to \infty$ for every $f \in C(\mathbb{T})$.

Proof. If $f = \sum_{n=-N}^{N} a_n z^n$, then $\hat{f}_n = 0$ for all $|n| > N$, so the result holds. For a general $f \in C(\mathbb{T})$, and given $\epsilon > 0$, there exists trigonometric polynomial $g(z)$ such that $\|f-g\|_\infty < \epsilon$, so for $|n| > N$, $|\hat{f}_n| = |\hat{f}_n - \hat{g}_n| = |\widehat{f-g}_n| \leq \|f-g\|_\infty < \epsilon$. □

Remark. The same result holds for $f \in L^1(\mathbb{T})$, since $|\hat{f}_n| \leq \|f\|_1$ for all $n \in \mathbb{Z}$, and $C(\mathbb{T})$ is dense in $L^1(\mathbb{T})$.

For $f \in C(\mathbb{T})$, $(\hat{f}_n)_{n\in\mathbb{Z}} \in c_0(\mathbb{Z}) = \{(a_n)_{n\in\mathbb{Z}} : a_n \to 0 \text{ as } n \to \infty\}$. This is a closed subspace of $l_\infty(\mathbb{Z})$, and hence a Banach space with $\|\cdot\|_\infty$. Define $F : C(\mathbb{T}) \to c_0(\mathbb{Z})$ by $F(f) = (\hat{f}_n)_{n\in\mathbb{Z}}$. F is called the *Fourier transform*.

Proposition. The Fourier transform $F : C(\mathbb{T}) \to c_0(\mathbb{Z})$ is an injective bounded linear map onto a dense subspace of $c_0(\mathbb{Z})$, but is not surjective.

Proof. Clearly F is linear. $\|F(f)\|_\infty = \sup_{n\in\mathbb{Z}} |\hat{f}_n| \le \|f\|_\infty$, so $\|F\| \le 1$.

Injectivity: Suppose $\hat{f}_n = 0$ for all $n \in \mathbb{Z}$, then $\frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})e^{in\theta}d\theta = 0$ for all $n \in \mathbb{Z}$. Hence $\frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})g(e^{i\theta})d\theta = 0$ for all trigonometric polynomial g. Choose a sequence (g_n) of trigonometric polynomials such that $g_n \to \bar{f}$ uniformly, then

$$\int_{-\pi}^{\pi} |f|^2 = |\int_{-\pi}^{\pi} f(\bar{f} - g_n)| \le \int_{-\pi}^{\pi} |f| \cdot |\bar{f} - g_n| \le 2\pi\|f\|_\infty \cdot \|\bar{f} - g\|_\infty \to 0$$

as $n \to \infty$, so $\int_{-\pi}^{\pi} |f|^2 = 0$. Since f is continuous, $f = 0$.

$\Im F$ is dense: Suppose $(a_n)_{n\in\mathbb{Z}} \in c_0(\mathbb{Z})$ is eventually 0, i.e. there exists $N \in \mathbb{N}$ such that $a_n = 0$ for all $|n| > N$, then $f(z) = \sum_{n=-N}^{N} a_n z^n \in C(\mathbb{T})$ has $F(f) = (a_n)_{n\in\mathbb{Z}}$. The set of sequences which are eventually 0 is dense in $c_0(\mathbb{Z})$, so $\Im F$ is dense.

F is not surjective: Suppose not, then F is a continuous linear bijection between Banach spaces, so by the inversion theorem, it is an isomorphism, i.e. there exists $C > 0$ such that $\|F(f)\|_\infty \ge C\|f\|_\infty$ for all f. Recall that the N-th Dirichlet kernel $D_N(z) = \sum_{n=-N}^{N} z^n$ has $\|D_N\|_\infty = 2N+1$, but $F(D_n) = I(\{n : |n| \le N\})$ (an indicator function). This is a contradiction for large N. □

Remark. The same holds for $f \in L^1(\mathbb{T})$ and $F : L^1(\mathbb{T}) \to c_0(\mathbb{Z})$, e.g. for the last part, $\|D_N\|_1 = \Omega(\log N) \to \infty$ as $N \to \infty$.

5.5 Arzelà-Ascoli Theorem

Let (M, d) be a metric space. For a non-empty subset A of M, define the *diameter* of A to be $\operatorname{diam} A = \sup\{d(x, y) : x, y \in A\} \in [0, \infty]$. Note that $\operatorname{diam} \overline{A} = \operatorname{diam} A$.

Given $\epsilon > 0$, a subset F of M is an ***ϵ-net*** for M if for all $x \in M$, there exists $y \in F$ such that $d(x, y) < \epsilon$, i.e. F is ϵ-dense in M. Equivalently, $M = \bigcup_{x\in F} \overline{B}(x, \epsilon)$. M is *totally bounded* if for all $\epsilon > 0$, there is a finite ϵ-net for M. Equivalently, for all $\epsilon > 0$, there exists finitely many subsets $A_1, \cdots, A_n$ of M with $M = \bigcup_{i=1}^{n} A_i$, and $\operatorname{diam} A_i \le \epsilon$ for all i. Furthermore, then A_i can be chosen to be closed by taking closure.

Lemma. Let M be a metric space and $N \subseteq M$, then

(i) If M is totally bounded, then N is totally bounded.
(ii) If N is totally bounded, then $\overline{N}$ is totally bounded.

Proof. (i): Given $\epsilon > 0$, there exists sets $A_1, \cdots, A_n$ such that $M = \bigcup_{i=1}^{n} A_i$, and $\operatorname{diam} A_i \le \epsilon$ for all i. Then $N = \bigcup_{i=1}^{n}(N \cap A_i)$, and $\operatorname{diam}(N \cap A_i) \le \operatorname{diam} A_i \le \epsilon$ for all i, after discarding those A_i with $N \cap A_i = \emptyset$.

(ii): Given $\epsilon > 0$, there exists sets $A_1, \cdots, A_n$ of N such that $N = \bigcup_{i=1}^{n} A_i$, and $\operatorname{diam} A_i \le \epsilon$ for all i. Then $\overline{N} \subseteq \bigcup_{i=1}^{N} \overline{A_i}$, and $\operatorname{diam} \overline{A_i} = \operatorname{diam} A_i \le \epsilon$ for all i. □

Proposition. For a metric space (M, d), the following are equivalent

(i) M is compact.
(ii) M is sequentially compact.
(iii) M is complete and totally bounded.

Proof. (i) $\Rightarrow$ (ii): Let (x_n) be a sequence in M. Suppose M is not sequentially compact, then for all $x \in M$, there exists $\epsilon_x > 0$ and $N = N_x$ such that $d(x_n, x) \geq \epsilon_x$ for all $n > N$, so $\{n \in \mathbb{N} : x_n \in B(x, \epsilon_x)\}$ is finite. The collection $\{B(x, \epsilon_x) : x \in M\}$ is an open cover of the compact space M, so there exists $z_1, \cdots, z_m \in M$ such that $M = \bigcup_{i=1}^m B(z_i, \epsilon_{z_i})$, then $\mathbb{N} = \{n \in \mathbb{N} : x_n \in M\} = \bigcup_{i=1}^m \{n \in \mathbb{N} : x_n \in B(z_i, \epsilon_{z_i})\}$ is finite, which is a contradiction.

(ii) $\Rightarrow$ (iii): M is complete: given a Cauchy sequence (x_n) in M, by assumption, there exists $k_1 < k_2 < \cdots$ and $x \in M$ such that $x_{k_n} \to x$ as $n \to \infty$. Let $\epsilon > 0$. There exists N such that $d(x_m, x_n) < \frac{\epsilon}{2}$ for all $m, n > N$. Choose a large $m > N$ such that $d(x, x_{k_m}) < \frac{\epsilon}{2}$, then for all $n > N$, $d(x, x_n) \leq d(x, x_{k_m}) + d(x_{k_m}, x_n) < \epsilon$.

M is totally bounded: suppose not, then there exists $\epsilon > 0$ such that M does not have a finite ϵ-net. Inductively construct a sequence (x_n) in M as follow: pick $x_1 \in M$. Given $x_1, \cdots, x_n \in M$ such that $d(x_i, x_j) \geq \epsilon$ for all $i \neq j$, pick $x_{n+1} \in M \setminus \bigcup_{i=1}^N B(x_i, \epsilon)$, which is non-empty since there is no finite ϵ-net. By construction, (x_n) has no Cauchy subsequence. Contradiction.

(iii) $\Rightarrow$ (i): Let $\mathcal{U}$ be an open cover for M without a finite subcover. Given $A \subset M$ such that A cannot be covered by finitely many elements of $\mathcal{U}$ and $\epsilon > 0$, then A is totally bounded by lemma 5.5, so there exists $B_1, \cdots, B_n$ with $A = \bigcup_{i=1}^n B_i$ and $\operatorname{diam} B_i \leq \epsilon$ for all i. If each B_i is covered by finitely many members of $\mathcal{U}$, then so can A, which is a contradiction. Therefore, there exists $B \subseteq A$, $\operatorname{diam} B \leq \epsilon$ such that B is not covered by finitely members of $\mathcal{U}$. If A is closed, then B can be chosen to be closed. By repeating this procedure, inductively construct a sequence of closed sets $M = A_0 \supseteq A_1 \supseteq A_2 \supseteq \cdots$ such that for all $n \in \mathbb{N}$, $\operatorname{diam} A_n < \frac{1}{n}$ and A_n cannot be covered by finitely many sets from $\mathcal{U}$. In particular, none of the A_n is empty, so for each $n \in \mathbb{N}$, choose $x_n \in A_n$. For $m, n \in \mathbb{N}$, $x_m, x_n \in A_{\min(m,n)}$, so $d(x_m, x_n) \leq \operatorname{diam} A_{\min(m,n)} < \frac{1}{\min(m,n)} \to 0$ as $m, n \to \infty$, so (x_n) is Cauchy. Since M is complete, suppose $x_n \to z$ for $z \in M$.

There exists $U \in \mathcal{U}$ such that $z \in U$. U is open, so there exists $r > 0$ such that $B(z, r) \subseteq U$. Pick any n such that $r > \frac{1}{n}$. A_n is closed, and $x_m \in A_n$ for all $m \geq n$, so $z \in A_n$. For all $w \in A_n$, $d(z, w) \leq \operatorname{diam} A_n \leq \frac{1}{n} < r$, so $w \in B(z, r) \subseteq U$. Therefore, $A_n \subseteq U$, contradicting A_n cannot be finitely subcovered by $\mathcal{U}$. □

Call a subset A of a topoligcal space X *relatively compact* in X if $\overline{A}$ is compact.

Corollary. Let (M, d) be a complete metric space and $N \subseteq M$, the following are equivalent

(i) N is relatively compact in M.
(ii) Every sequence in N has a convergent subsequence in M.
(iii) N is totally bounded.

Proof. (i) $\iff$ $\overline{N}$ is compact by definition. Clearly $\overline{N}$ is sequentially compact implies (ii). For the reverse implication, given (x_n) in $\overline{N}$, choose a sequence (y_n) in N such that $d(y_n, x_n) < \frac{1}{n}$ for each $n \in \mathbb{N}$. Then there exists a convergent subsequence (y_{n_k}) of (y_n) tending to $z \in M$ for some z. Then $x_{k_n} \to z$, and $z \in \overline{N}$. Therefore (ii) implies $\overline{N}$ is sequentially compact. Note that (iii) implies $\overline{N}$ is totally bounded. It is also complete, being a closed subset of a compact space. The converse is clear, so (iii) $\iff$ $\overline{N}$ is complete and totally bounded. The result now follows from proposition 5.5. □

Let K be a compact topological space, and $\mathbb{F} \subseteq C(K)$. $\mathbb{F}$ is *totally bounded* if there exists M such that $\|f\|_\infty \leq M$ for all $f \in \mathbb{F}$. $\mathbb{F}$ is *equicontinuous* if for all $x \in K,\ \epsilon > 0$, there exists an open neighbourhood U of x such that $|f(y) - f(x)| < \epsilon$ for all $y \in U,\ f \in \mathbb{F}$.

Theorem (Arzelà-Ascoli Theorem)**.** Let K be a compact topological space, and $\mathbb{F} \subseteq C(K)$, then $\mathbb{F}$ is relatively compact in $C(K)$ if and only if $\mathbb{F}$ is uniformly bounded and equicontinuous.

Proof. $\Longrightarrow$: $\mathbb{F}$ is uniformly bounded: For $n \in \mathbb{N}$, let $U_n = \{f \in C(K) : \|f\|_\infty < n\}$, then $\{U_n : n \in \mathbb{N}\}$ is an open cover of $C(K)$. Since $\overline{\mathbb{F}}$ is compact, there exists $N \in \mathbb{N}$ such that $\mathbb{F} \subseteq U_N$. In particular, $\|f\|_\infty \leq N$ for all $f \in \mathbb{F}$.

$\mathbb{F}$ is equicontinuous: Let $x \in K,\ \epsilon > 0$. By corollary 5.5, $\mathbb{F}$ is totally bounded, so there is a fintie ϵ-net $f_1, \cdots, f_n$ for $\mathbb{F}$. For each $i = 1, \cdots, n$, there exists an open neighbourhood U_i of x such that $|f_i(y) - f_i(x)| < \epsilon$ for all $y \in U_i$. Let $U = \bigcap_{i=1}^n U_i$, then U is an open neighbourhood of x. Given $f \in \mathbb{F}$, there exists $i \in \{1, \cdots, n\}$ such that $\|f - f_i\|_\infty \leq \epsilon$. For all $y \in U \subseteq U_i$,

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

Therefore $\mathbb{F}$ is equicontinuous.

$\Longleftarrow$: Since $(C(K), \|\cdot\|_\infty)$ is complete, it sufficies to show that $\mathbb{F}$ is totally bounded by corollary 5.5. Let $\epsilon > 0$. For all $x \in K$, there exists an open neighbourhood U_x of x such that $|f(y) - f(x)| < \epsilon$ for all $y \in U_x$ and $f \in \mathbb{F}$. The family $\{U_x : x \in K\}$ is an open cover for K, so ther eexists $x_1, \cdots, x_m \in K$ such that $K = \bigcup_{i=1}^m U_{x_i}$. Let $S = \{(f(x_1), \cdots, f(x_m)) : f \in \mathbb{F}\}$. Since $\mathbb{F}$ is uniformly bounded, S is bounded with respect to $\|\cdot\|_\infty$ in $\mathbb{R}^m$ (or $\mathbb{C}^m$), and hence totally bounded by the Heine-Borel theorem. S has a finite ϵ-net, i.e. there exists $f_1, \cdots, f_n \in \mathbb{F}$ such that for all $f \in \mathbb{F}$, there exists $j \in \{1, \cdots, n\}$ such that $|f(x_i) - f_j(x_i)| < \epsilon$ for all $1 \leq i \leq m$. Given $f \in \mathbb{F}$, and choose one such j, then for $y \in K$. there exists $i \in \{1, \cdots, m\}$ such that $y \in U_{x_i}$, and so

$$|f(y) - f_j(y)| \leq |f(y) - f(x_i)| + |f(x_i) - f_j(x_i)| + |f_j(x_i) - f_j(y)| < 3\epsilon$$

by the choice of j and definition of U_{x_i}. Therefore $\|f - f_j\|_\infty < 3\epsilon$, so $f_1, \cdots, f_n$ is a (3ϵ)-net of $\mathbb{F}$. □

Example. Given a bounded linear map $T : X \to Y$ between Banach spaces, T is *compact* if $\overline{T(B_X)}$ is compact in Y. We give an example of a compact operator for $X = Y = C[0,1]$. Let $K : [0,1]^2 \to \mathbb{R}$ be continuous. Define $T : C[0,1] \to C[0,1]$ by

$$(Tf)(x) = \int_0^1 f(y)K(x,y)dy$$

It is easy to check that T is well-defined, linear, and bounded. Let $\mathbb{F} = T(B_{C[0,1]})$, then we need to show that $\mathbb{F}$ is relatively compact. It is uniformly bounded since T is bounded. Given $\epsilon > 0$, K is uniformly continuous, so there exists $\delta > 0$ such that for all $(x_1, y_1), (x_2, y_2) \in [0,1]^2$, $\max(|x_1 - x_2|, |y_1 - y_2|) < \delta$ implies $|K(x_1, y_1) - K(x_2, y_2)| < \epsilon$. Then for $f \in B_{C[0,1]}, x_1, x_2, \in [0,1]$, if $|x_1 - x_2| < \delta$, then

$$|Tf(x_1) - Tf(x_2)| \leq \int_0^1 |f(y)| \cdot |K(x_1, y) - K(x_2, y)|dy \leq \epsilon$$

Therefore, $\mathbb{F}$ is equicontinuous, and so T is compact.

6 Hilbert spaces

Let X be a real or complex vector space. An *inner product* on X is a function $\langle\cdot,\cdot\rangle$: $X \times X \to$ scalars such that

(i) $\forall x, y, z \in X$, λ, μ scalars, $\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle$ (Linearity in first variable)
(ii) $\forall x, y \in X$, $\langle y, x\rangle = \overline{\langle x, y\rangle}$ (Symmetry for $\mathbb{R}$, conjugate symmetry for $\mathbb{C}$)
(iii) $\forall x \in X$, $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0 \iff x = 0$ (Positive definite)

The pair $(X, \langle\cdot,\cdot\rangle)$ is an *inner product space.* Note that for all $x, y, z \in X$ and scalars λ, μ, we have conjugate linearity in the second variable

$$\langle x, \lambda y + \mu z\rangle = \overline{\langle \lambda y + \mu z, x\rangle} = \overline{\lambda\langle y, x\rangle + \mu\langle z, x\rangle} = \bar{\lambda}\langle x, y\rangle + \bar{\mu}\langle x, z\rangle$$

We define the norm on X by $\|x\| = \langle x, x, \rangle^{1/2}$ for all $x \in X$.

Example.

(1) $l_2^n = (\mathbb{R}^n, \|\cdot\|_2)$ with $\langle x, y\rangle = \sum_{i=1}^n x_i\bar{y}_i$, $\|x\| = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2} = \|x\|_2$.

(2) l_2 with $\langle x, y\rangle = \sum_{i=1}^\infty x_i\bar{y}_i$. This converges by the Cauchy-Schwarz inequality. The associated norm is $\|x\| = \|x\|_2$.

(3) $C[0,1]$: $\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}dt$, with $\|f\| = \|f\|_2$.

(4) $C(\mathbb{T})$: $\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})\overline{g(e^{i\theta})}d\theta$

(5) Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Then the function space $L^2(\mu)$ has inner product $\langle f, g\rangle = \int_\Omega f(u)\overline{g(u)}d\mu(u)$, with $\|f\| = \|f\|_2$.

Theorem. Let X be an inner product space, $x, y \in X$. Then

(i) $|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$ (Cauchy-Schwarz inequality)
(ii) $\|x + y\| \leq \|x\| + \|y\|$ (Minkowski's inequality)

Proof. (i): For all scalars λ, $0 \leq \|x + \lambda y\|^2 = \|x\|^2 + \|y\|^2 + 2\mathbb{R}e(\bar{\lambda}\langle x, y\rangle)$. If $\|y\| = 0$, then the result holds. Otherwise, let $\lambda = -\frac{\langle x,y\rangle}{\|y\|^2}$ to get $0 \leq \|x\|^2 + \frac{|\langle x,y\rangle|^2}{\|y\|^2} - 2\frac{|\langle x,y\rangle|^2}{\|y\|^2}$, which rearranges to the desired inequality.

(ii): For all $x, y \in X$, $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\mathbb{R}e\langle x, y\rangle \leq \|x\|^2 + \|y\|^2 + 2|\langle x, y\rangle| \leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = (\|x\| + \|y\|)^2$. □

Remark. It is now easy to check that the norm of an inner product defines a norm. With respect to this norm topology, the inner product is continuous, which can be proven by the Cauchy-Schwarz inequality.

A *Hilbert space* is a complete inner product space. E.g. l_2^n, l_2, $L^2(\mu)$ are Hilbert spaces, but $C[0,1]$, $C(\mathbb{T})$ are not.

Proposition (Polarization Identities)**.** Let X be an inner product space, $x, y \in X$.

(i) In the real case, $\langle x, y\rangle = \frac{1}{4}(\|x + y\|^2 + \|x - y\|^2)$.
(ii) In the complex case, $\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 + i\|x+iy\|^2 - \|x-y\|^2 - i\|x-iy\|^2) = \frac{1}{2\pi}\int_0^{2\pi} e^{i\theta}\|x + e^{i\theta}y\|^2 d\theta$.

Proof. Expand the RHS □

Theorem. The completion $\tilde{X}$ of an inner product space in a Hilbert space.

Proof. Given $x, y \in \tilde{X}$, choose sequences (x_n), (y_n) in X such that $x_n \to x$ and $y_n \to y$. Define $\langle x, y\rangle = \lim_{n\to\infty}\langle x, y\rangle$. It is routine to verify that this is well-defined, gives an inner product on $\tilde{X}$, and induces the same norm. □

Proposition (Parallelogram Law)**.** Let X be an inner product space, $x, y \in X$, then $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$.

Proof. Expand the LHS. □

Remark.

(1) There is a converse: if the norm of a normed space satisfies the parallelogram rule, then it is an inner product space. For the proof, use the polarization identity to define the inner product and check the necessary properties. This gives a new proof of theorem 6.

(2) Generalized parallelogram law: given $x_1, \cdots, x_n$ in an inner product space X,

$$\frac{1}{2^n}\sum_{\epsilon_i=\pm 1}\|\sum_{i=1}^{n}\epsilon_i x_i\|^2 = \sum_{i=1}^{n}\|x_i\|^2$$

For x, y in an inner product space X, we say x is *orthogonal* to y, written $x \perp y$, if $\langle x, y\rangle = 0$. This is a symmetric relation.

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

Proof. Expand the LHS and use $\langle x, y\rangle = 0$. □

More generally, given n vectors $x_1, \cdots, x_n$ with $x_i \perp x_j$ for all $i \neq j$, then $\|\sum x_i\|^2 = \sum \|x_i\|^2$, which can be proven by induction.

Recall that for a non-empty subset A of a metric space (M, d) and a point $x \in M$, $d(x, A) = \inf\{d(x, a) : a \in A\}$. This is not necessarily attained. For example, in l_2, $A = \{(1 + 1/n)e_n : n \in \mathbb{N}\}$ has $d(0, A) = 1$, but $d(0, a) > 1$ for all $a \in A$.

Theorem (Closest Point Theorem)**.** Let Y be a closed subspace of a Hilbert space H. Let $x \in H$. Then there exists a unique $z \in Y$ such that $\|x - z\| = d(x, Y)$.

Proof. Choose a sequence (y_n) in Y such that $\|x - y_n\| \to d(x, Y)$. By proposition 6, for all $m, n \in \mathbb{N}$, we have

$$\begin{aligned} 2(\|x-y_m\|^2 + \|x-y_n\|^2) &= \|y_m - y_n\|^2 + \|2x - y_m - y_n\|^2 \\ &= \|y_m - y_n\|^2 + 4\|x - \frac{y_m + y_n}{2}\|^2 \\ &\geq \|y_m - y_n\|^2 + 4d(x, Y) \end{aligned}$$

since $\frac{y_m + y_n}{2} \in Y$. Rearrange the inequality to get $0 \leq \|y_m - y_n\|^2 \leq 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - 4d(x, Y)^2 \to 0$ as $m, n \to \infty$. Therefore (y_n) is Cauchy. Since H is complete and Y is closed, $y_n \to z$ for some $z \in Y$. We have $\|x - z\| = \lim_{n\to\infty} \|x - y_n\| = d(x, Y)$.

Suppose $w \in Y$ also satisfies $\|x - w\| = d(x, Y)$, then by the calculation above,

$$0 \le \|z - w\|^2 \le 2\|x - z\|^2 + 2\|x - w\|^2 - 4d(x, Y)^2 = 0$$

Hence $\|z - w\|^2 = 0$, and so $z = w$. □

Remark. The same holds if Y is non-empty, closed, and convex.

Let X be an inner product space, then for $y \in X$, define $\theta_y : X \to \text{scalar}$ by $x \mapsto \langle x, y\rangle$, for all $x \in X$. Then θ_y is linear, $|\theta_y(x)| \le \|x\| \cdot \|y\|$, so θ_y is bounded, with $\|\theta_y\| \le \|y\|$. The reverse inequality also holds since $\theta_y(y) = \|y\|^2$. Therefore, $\theta : X \to X^*,\ y \mapsto \theta_y$ is a conjugate linear isometric embedding. We aim to show that θ is onto if X is complete.

For $y \in X$, define $y^\perp = \{x \in X : x \perp y\} = \ker\theta_y$, so it is aclosed subspace of X. For a subset $S \subseteq X$, similarly define $S^\perp = \{x \in X : \forall y \in S\ x \perp y\} = \bigcap_{y\in S} y^\perp$.

Theorem. If Y is a closed subspace of Hilbert space H, then $H = Y \oplus Y^\perp$.

Proof. Recall that for Banach spaces, $X = Y \perp Z$ means $X = Y + Z, Y \cap Z = \{0\}$, $\|y + z\| \sim \|y\| + \|z\|$, the last being equivalent to Y and Z are closed subspaces. Since Y and $Y^\perp$ are both closed subspaces. It suffices to show that $Y \cap Y^\perp = \{0\}$ and $H = Y + Y^\perp$.

$Y \cap Y^\perp = \{0\}$: Let $x \in Y \cap Y^\perp$, then $\|x\|^2 = \langle x, x\rangle = 0$, so $x = 0$.

$H = Y + Y^\perp$: Given $x \in H$, let $y \in Y$ be the unique closest point of Y to x (theorem 6). Let $z = x - y$, then $x = y + z$, so it remains to show that $z \in Y^\perp$. Suppose not, then there exists $w \in Y$ such that $\langle z, w\rangle \ne 0$. By multiplication by a scalar, suppose $\langle z, w\rangle > 0$. Then for $t \in \mathbb{R}, t > 0$,

$$\|z\|^2 = \|x - y\|^2 < \|x - (y + tw)\|^2 = \|z - tw\|^2 = \|z\|^2 + t^2\|w\|^2 - 2t\langle z, w\rangle$$

so $2\langle z, w\rangle < t\|w\|^2$ for all $t > 0$. This is a contradiction as $t \to 0$. □

We call $Y^\perp$ the *orthogonal complement* of Y in H. We say H is the *orthogonal direct sum* of Y and $Y^\perp$. From now on, for Hilbert spaces, $\oplus$ will always mean orthogonal direct sum. The projection $P : H \to H, y + z \mapsto y$ for $y \in Y$, $z \in Y^\perp$ is called the *orthogonal projection* of H onto Y.

For $x \in H$, $x - Px \in Y^\perp$, so $x - Px \perp Px$. By proposition 6, $\|Px\|^2 + \|x - Px\|^2 = \|x\|^2$, so $\|Px\| \le \|x\|$. Hence P is bounded. For $y \in Y$, $x - Px \perp Px - y$, so $\|x - Px\|^2 \le \|x - Px\|^2 + \|Px - y\|^2 = \|x - y\|^2$, with equality if and only if $y = Px$, so Px is the unique closest point of Y to x.

Theorem (Riesz Representation Theorem). **If H is a Hilbert space, and $\theta : H \to H^*$ is defined by $\theta_y(x) = \langle x, y\rangle$, then θ is onto.**

Proof. Let $f \in H^*$. If $f = 0$, then it is represented by θ_0. Now suppose $f \ne 0$, and let $Y = \ker f$, then Y is a closed proper subspace of H, so $H = Y \oplus Y^\perp$ (theorem 6). In particular, $Y^\perp \ne \{0\}$. Pick $z \in Y^\perp \backslash \{0\}$. By rescaling, let $f(z) = 1$. Given $x \in Y^\perp$,

$f(x - f(x)z) = 0$, so $x - f(x)z \in Y \cap Y^\perp = \{0\}$, so $x = f(x)z$. Therefore, $Y^\perp = \text{span}\{z\}$. Let $x \in H$, choose $y \in Y$ and λ scalar such that $x = y + \lambda z$, then

$$f(x) = f(y + \lambda z) = \lambda, \ \langle x, \frac{z}{\|z\|^2}\rangle = \langle y + \lambda z, \frac{z}{\|z\|^2}\rangle = \lambda$$

Hence f is represented by $\frac{z}{\|z\|^2}$. □

Corollary. For a Hilbert space $H, \theta : H \to H^*$ defined as above is a conjugate linear isometric isomorphism onto H^*, i.e. H is self-dual.

6.1 Orthonormal Basis

An *orthonormal sequence* in an inner product space X is a possibly finite sequence (x_n) consisting of pairwise orthogonal unit vectors, i.e. $\langle x_m, x_n\rangle = \delta_{mn}$. If in addition, $\text{cls}\{x_n : n \in \mathbb{N}\} = X$, then (x_n) is an *orthonormal basis* of X.

Example.

(1) In l_2, the unit vectors (e_n) is an orthonormal basis.

(2) In l_2, the set (e_{2n}) is an orthonormal sequence, but not a basis.

(3) In $C(\mathbb{T})$, with inner product $\frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})\overline{g(e^{i\theta})}d\theta$, take $e_n(z) = z^n$, then (e_n) is an orthonormal basis.

Remark. An orthonormal sequence (x_n) is linearly independent, since $\sum_{i=1}^n \lambda_i x_i = 0$ for scalars λ_i implies

$$0 = \langle \sum_{i=1}^n \lambda_i x_i, x_j\rangle = \sum_{i=1}^n \lambda_i \langle x_i, x_j\rangle = \lambda_j$$

for all j. Furthermore, if X has an orthonormal basis, then X is separable.

Theorem (Gram-Schmidt Process). Let X be an inner product space, and $x_1, x_2, \cdots$ be a linearly independent sequence in X. Then there exists orthonormal sequence $e_1, e_2, \cdots$ such that $\text{span}\{e_1, \cdots, e_n\} = \text{span}\{x_1, \cdots, x_n\}$ for all n.

Proof. Construc the sequence inductively. Since $x_1 \neq 0$, take $e_1 = \frac{x_1}{\|x_1\|}$. Assume we have an orthonormal sequence $e_1, \cdots, e_n$ with $\text{span}\{e_1, \cdots, e_m\} = \text{span}\{x_1, \cdots, x_m\}$ for $1 \le m \le n$. Set $e'_{m+1} = x_{n+1} - \sum_{i=1}^n \langle x_{n+1}, e_i\rangle e_i$, then for $1 \le j < n$, $\langle e'_{n+1}, e_j\rangle = 0$. $e_{n+1} \neq 0$ since otherwise $x_{n+1} \in \text{span}\{e_1, \cdots, e_n\} = \text{span}\{x_1, \cdots, x_n\}$. Let $e_{n+1} = \frac{e'_{n+1}}{\|e'_{n+1}\|}$, then $e_1, \cdots, e_{n+1}$ is an orthonormal sequence with $\text{span}\{e_1, \cdots, e_{n+1}\} = \text{span}\{x_1, \cdots, x_{n+1}\}$. □

Corollary. Every separable inner product space X has an orthonormal basis.

Proof. There exists $y_1, \cdots \in X$ such that X is the closure of $\{y_n : n \in \mathbb{N}\}$. Remove y_i if $y_i \in \text{span}\{y_1, \cdots, y_{i-1}\}$, then what remains is a linearly independent subsequence (x_n) with $\text{cls}\{x_1, \cdots, x_n\} = X$. Apply Gram-Schmit to this sequence to get an orthonormal basis for X. □

Example. In $C[-1,1]$ with inner product $\langle f,g\rangle = \int_{-1}^{1} f(t)\overline{g(t)}dt$, the sequence $1, t, t^2, \cdots$ is linearly independent and has closed linear span $C[-1,1]$ by the Weierstrass approximation theorem. We obtain an orthogonal basis through the Gram-Schmidt process which starts with

$$\frac{1}{\sqrt{2}},\ \sqrt{\frac{3}{2}}t,\ \sqrt{\frac{3}{8}}(3t^2-1), \cdots$$

These are the Legendre polynomials.

Corollary. If X is an inner product space, and $\dim X = n$, then $X \cong l_2^n$.

Proof. X is separable, so there exists an orthonormal basis $e_1, \cdots, e_m$, which is finite since they are linearly independent. In addition, $\operatorname{cls}\{e_1, \cdots, e_m\} = \operatorname{span}\{e_1, \cdots, e_m\} = X$, so $m = n$.
Define $T : X \to l_2^n$, $\sum_{i=1}^n \lambda_i e_i \mapsto (\lambda_i)_{i=1}^n$, then

$$\|\sum_{i=1}^n \lambda_i e_i\|^2 = \left\langle \sum_{i=1}^n \lambda_i e_i, \sum_{i=1}^n \lambda_i e_i \right\rangle = \sum_{i=1}^n |\lambda_i|^2 = \|(\lambda_i)_{i=1}^n\|^2$$

so T is an isometric isomorphism. □

Theorem. Let X be an inner product space, and (e_n) an orthonormal basis for X, then for all $x \in X$, $x = \sum_{n=1}^\infty \langle x, e_n\rangle e_n$.

Proof. For $n \in \mathbb{N}$, let $E_n = \operatorname{span}\{e_1, \cdots, e_n\}$. Let $P_n x = \sum_{i=1}^n \langle x, e_i\rangle e_i$. We need to show that $P_n x \to x$ as $n \to \infty$. Note that $\langle x - P_n x, e_i\rangle = 0$ for $i = 1, 2, \cdots, n$, so $x - P_n x \perp y$ for all $y \in E_n$. Let $\epsilon > 0$, then there exists $y \in \operatorname{span}\{e_1, \cdots\}$ such that $\|x - y\| < \epsilon$. There exists $N \in \mathbb{N}$ such that $y \in E_N$, then for all $n \geq N$, $x - P_n x \perp P_n x - y$. By Pythagoras' theorem, we have $\|x - P_n x\|^2 + \|P_n x - y\|^2 = \|x - y\|^2 < \epsilon^2$, so $\|x - P_n x\| < \epsilon$. □

The n-th coefficient $\langle x, e_n\rangle$ is called the n-th *Fourier coefficient* of x with respect to (e_n). For example, if $X = C(\mathbb{T})$ and $e_n(z) = z^n$, $n \in \mathbb{Z}$, then $\langle f, e_n\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\theta})e^{-in\theta}d\theta = \hat{f}_n$. By above, the Fourier series $\sum_{n=-\infty}^{\infty} \hat{f}_n e^{int}$ converges to f in the L^2 norm.

Corollary (Parseval's Identities)**.** Let X be an inner product space, and (e_n) an orthonormal basis, then

(i) $\|x\|^2 = \sum_{n=1}^\infty |\langle x, e_n\rangle|^2$ for all $x \in X$.

(ii) $\langle x, y\rangle = \sum_{n=1}^\infty \langle x, e_n\rangle\overline{\langle y, e_n\rangle}$ for all $x, y \in X$.

Proof. Setting $x = y$ in (ii) gives (i). To prove (ii), use theorem 6.1 and the continuity of inner product

$$\begin{aligned}\langle x, y\rangle &= \left\langle \sum_{m=1}^\infty \langle x, e_m\rangle e_m, \sum_{n=1}^\infty \langle y, e_n\rangle e_n \right\rangle = \lim_{N\to\infty} \left\langle \sum_{m=1}^N \langle x, e_m\rangle e_m, \sum_{n=1}^N \langle y, e_n\rangle e_n \right\rangle \\ &= \lim_{N\to\infty} \sum_{m,n=1}^\infty \langle x, e_m\rangle\overline{\langle y, e_m\rangle}\langle e_m, e_n\rangle = \lim_{N\to\infty}\sum_{n=1}^N \langle x, e_n\rangle\overline{\langle y, e_n\rangle}\end{aligned}$$

□

In particular, this shows that $(\langle x, e_n\rangle)_{n=1}^{\infty} \in l_2$. We now prove the converse.

Theorem (Riesz-Fischer Theorem)**.** Let (e_n) be an infinite orthonormal basis for a Hilbert space H, then for every $c = (c_n) \in l_2$, there exists $x \in H$ such that $\langle x, e_n\rangle = c_n$ for all $n \in \mathbb{N}$.

Proof. Let $s_n = \sum_{n=1}^{n} c_i e_i$, $n \in \mathbb{N}$. For $n > m$,

$$\|s_n - s_m\|^2 = \|\sum_{k=m+1}^{n} c_k e_k\|^2 = \sum_{k=m+1}^{n} |c_k|^2 \to 0 \text{ as } m, n \to \infty$$

Therefore, (s_n) is Cauchy and tends to a limit $x \in H$. By the continuity of inner product, $\langle x, e_n\rangle = \lim_{m\to\infty}\langle s_m, e_n\rangle = c_n$ for all $n \in \mathbb{N}$, as required. ☐

Corollary. Let H be a separable infinite dimensional Hilbert space, then $H \cong l_2$.

Proof. Since H is separable, there exists an orthonormal basis $e_1, e_2, \cdots$ by corollary 6.1. Since $\dim H = \infty$, this sequence is infinite. Define $T : H \to l_2$ by $Tx = (\langle x, e_n\rangle)_{n=1}^{\infty}$. Parseval's identity implies T is well-defined and isometric. T is linear by the properties of inner product. T is surjective by theorem 6.1. ☐

6.2 Operators on Hilbert Spaces

Suppose H is a Hilbert space with orthonormal basis (e_n), then for $T \in \mathbb{B}(H)$, its matrix $A = (a_{ij})$ is defined by $a_{ij} = \langle Te_j, e_i\rangle$. If $x = \sum_{i=1}^{\infty} x_i e_i \in H$, and $y = Tx = \sum_{i=1}^{\infty} y_i e_i$, then $y_i = \langle y, e_i\rangle = \sum_{j=1}^{\infty} a_{ij} x_j$.
Note that not every matrix comes from a bounded linear map. A necessary condition is that the rows and columns are in l_2, but this is not sufficient. However, if a matrix does come from $T \in \mathbb{B}(H)$, then T is uniquely determined.

Theorem. Let H be a Hilbert space, $T \in \mathbb{B}(H)$, then there exists a unique map $T^* : H \to H$ such that $\langle Tx, y\rangle = \langle x, T^*y\rangle$ for all $x, y \in H$. Moreover, $T^* \in \mathbb{B}(H)$ with $\|T^*\| \leq \|T\|$. T^* is called the *adjoint* of T.

Proof. Uniqueness: Suppose $S_1, S_2 : H \to H$ are such that $\langle Tx, y\rangle = \langle x, S_i y\rangle$ for all $x, y \in H,\ i = 1, 2$. Then $\langle x, (S_1 - S_2)y\rangle = 0$ for all $x, y \in H$. Taking $x = (S_1 - S_2)y$ gives $\|(S_1 - S_2)y\|^2 = 0$ for all y, so $S_1 - S_2 = 0$, i.e. $S_1 = S_2$.
Existence: Fix $x, y \in H$, the map $x \mapsto \langle Tx, y\rangle$ is linear. Furthermore,

$$|\langle Tx, y\rangle| \leq \|Tx\| \cdot \|y\| \leq \|T\| \cdot \|x\| \cdot \|y\|$$

so it is also bounded. Since $H^* \cong H$ via a conjugate linear map, there exists $z \in H$ such that $\langle Tx, y\rangle = \langle x, z\rangle$. Define $T^*y = z$ to get a map $T^* : H \to H$ satisfying $\langle Tx, y\rangle = \langle x, T^*y\rangle$ for all $x, y \in H$.
Linearity: For all $x, y, z \in H$, λ, μ scalars,

$$\begin{aligned}\langle x, T^*(\lambda y + \mu z)\rangle &= \langle Tx, \lambda x + \mu z\rangle = \bar{\lambda}\langle Tx, y\rangle + \bar{\mu}\langle Tx, z\rangle \\ &= \bar{\lambda}\langle x, T^*y\rangle + \bar{\mu}\langle x, T^*z\rangle = \langle x, \lambda T^*y + \mu T^*z\rangle\end{aligned}$$

Therefore $T^*(\lambda y + \mu z) = \lambda T^*y + \mu T^*z$.
Continuity: $\|T^*y\|^2 = \langle T^*y, T^*y\rangle = \langle TT^*y, y\rangle \leq \|TT^*y\| \cdot \|y\| \leq \|T\| \cdot \|T^*y\| \cdot \|y\|$, so $\|T^*y\| \leq \|T\| \cdot \|y\|$, as required. ☐

Remark.

(1) Let S be the dual operator of T, i.e. $S : H^* \to H^*$, $f \mapsto f \circ T$. Recall that we have a conjugate linear isometric isomorphism $\theta : H^* \to H$. It is easy to check that $\theta^{-1}S\theta = T^*$, which justifies using $*$ to denote the dual operator and the adjoint.

(2) If H has an orthonormal basis (e_n), and T has matrix A with respect to the basis, then T^* has matrix $\bar{A}^T$, since $\langle T^*e_j, e_i\rangle = \overline{\langle e_i, T^*e_j\rangle} = \overline{\langle Te_i, e_j\rangle} = \overline{a_{ji}}$.

Example. $I^* = I$, $\lambda I^* = \bar{\lambda}I$, (right-shift)* = left-shift.

Proposition. Let H be a Hilbert space, $S, T \in \mathbb{B}(H)$, λ, μ scalars, then

(i) $(\lambda S + \mu T)^* = \bar{\lambda}S^* + \bar{\mu}T^*$

(ii) $(ST)^* = T^*S^*$

(iii) $T^{**} = T$

(iv) $\|T^*\| = \|T\|$

(v) $\|T^*T\| = \|T\|^2$ (C^* identity)

Proof. (i) For all $x, y \in H$,

$$\begin{aligned}\langle x, (\lambda S + \mu T)^* y\rangle &= \langle (\lambda S + \mu T)x, y\rangle = \lambda\langle Sx, y\rangle + \mu\langle Tx, y\rangle \\ &= \lambda\langle x, S^*y\rangle + \mu\langle x + T^*y\rangle = \langle x, (\bar{\lambda}S^* + \bar{\mu}T^*)y\rangle\end{aligned}$$

(ii) $\langle x, (ST)^*y\rangle = \langle STx, y\rangle = \langle Tx, S^*y\rangle = \langle x, T^*S^*y\rangle$.

(iii) $\langle x, T^{**}y\rangle = \langle T^*x, y\rangle = \overline{\langle y, T^*x\rangle} = \overline{\langle Ty, x\rangle} = \langle x, Ty\rangle$.

(iv) By theorem 6.2, $\|T^*\| \le \|T\|$. By (iii), $\|T\| = \|T^{**}\| \le \|T^*\|$, so $\|T\| = \|T^*\|$.

(v) $\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle \le \|x\| \cdot \|T^*Tx\| \le \|x\|^2 \cdot \|T^*T\|$. Taking supremum over all $x \in B_H$ gives $\|T\|^2 \le \|T^*T\|$. From (iv), $\|T^*T\| \le \|T^*\| \cdot \|T\| = \|T\|^2$, so $\|T^2\| = \|T^*T\|$.

□

Let H be a Hilbert space, $T \in \mathbb{B}(H)$, then T is *Hermitian* (or *self-adjoint*) if $T^* = T$, i.e. if $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in H$.

Example.

(i) The identity map.

(ii) Any orthogonal projection: Suppose $Y \subseteq H$ is a closed subspace, $H = Y \oplus Y^\perp$, $P : H \to H$, $p(y + z) = y$ for $y \in Y$, $z \in Y^\perp$. Let $x = y + z, x' = y' + z'$ for $y, y' \in Y$, $z, z' \in Y^\perp$, then $\langle Px, x'\rangle = \langle y, y' + z'\rangle = \langle y, y'\rangle = \langle x, Px'\rangle$.

(iii) If H has orthonormal basis (e_n), $T : H \to H$, $\sum_{i=1}^\infty x_i e_i \mapsto \sum_{i=1}^\infty \lambda_i x_i e_i$, then $T \in \mathbb{B}(H)$ if and only if $(\lambda_i) \in l_\infty$, and T is Hermitian if and only if $\lambda_i \in \mathbb{R}$ for all i.

Proposition. Let H be a complex Hilbert space, then for all $T \in \mathbb{B}(H)$, there exists unique Hermitian operators $T_1,\ T_2$ such that $T = T_1 + iT_2$.

Proof. Uniqueness: $T = T_1 + iT_2 \ \Rightarrow\ T^* = T_1 - iT_2 \ \Rightarrow\ T_1 = \frac{1}{2}(T + T^*), T_2 = \frac{1}{2i}(T - T^*)$.
Existence: Define T_1, T_2 as above, then it is easy to show that they work. □

$T \in \mathbb{B}(H)$ is *unitary* if $T^*T = TT^* = I$, i.e. T is invertible, and $T^* = T^{-1}$. Let $U(H)$ be the set of all unitary operators on H.

Example.

(i) The identity map.

(ii) Given $H = Y \oplus Y^\perp$, $y + z \mapsto y - z$, where $y \in Y, z \in Y^\perp$.

(iii) If H has orthonormal basis $(e_n), T : H \to H, \sum_{i=1}^\infty x_i e_i \mapsto \sum_{i=1}^\infty \lambda_i x_i e_i$, then $T \in U(H)$ if and only if $|\lambda_i| = 1$ for all i.

Proposition. $T \in \mathbb{B}(H)$ is unitary if and only if T is isometric and surjective.

Proof. $\Longrightarrow$: T is invertible, so it is surjective. $\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle = \|x\|^2$ for all $x \in H$, so T is isometric.
$\Longleftarrow$: $\|Tx\| = \|x\|$ for all $x \in H$ implies $\langle Tx, Ty\rangle = \langle x, y\rangle$ for all $x, y \in H$ by the polarization identities, so $\langle x, T^*Ty\rangle = \langle Tx, Ty\rangle = \langle x, y\rangle$. Hence $T^*T = I$. T isometric and surjective implies T^{-1} exists and is also isometric, so $T^* = (T^*T)T^{-1} = T^{-1}$. □

$T \in \mathbb{B}(H)$ is *normal* if $T^*T = TT^*$. Hermitian and unitary operators are normal.

7 Spectral theory

Let X be a Banach space, then $\mathbb{B}(X)$ is also a Banach space in the operator norm (theorem 1.3). In fact, $\mathbb{B}(X)$ is a Banach algebra, with composition as multiplication, i.e. for all $S, T, U \in \mathbb{B}(X)$, λ scalar,

$$(ST)U = S(TU),\ S(T+U) = ST + SU, (S+T)U = SU + TU,$$
$$\lambda(ST) = (\lambda S)T = S(\lambda T),\ \|ST\| \leq \|S\| \cdot \|T\|$$

$T \in \mathbb{B}(X)$ is *invertible* if there exists $S \in \mathbb{B}(X)$ such that $ST = TS = I$, i.e. T is an isomorphism, or equivalently T is a linear bijection (continuity of inverse follows from the inversion theorem). Let $\mathcal{G}(X) = \{T \in \mathbb{B}(X) : T \text{ invertible}\}$, then it is easy to check that $\mathcal{G}(X)$ is a group under composition.

Theorem. Let $T \in \mathbb{B}(X)$, $\|T\| < 1$, then $I - T$ is invertible, with $\|I - T\|^{-1} \leq \frac{1}{1 - \|T\|}$.

Proof. Let $S_n = \sum_{k=0}^n T_k$, then for $n > m$,

$$\|s_n - s_m\| = \|\sum_{k=m+1}^n T_k\| \leq \sum_{k=m+1}^n \|T\|^k \to 0 \text{ as } n, m \to \infty$$

Therefore, (S_n) is a Cauchy sequence, and so converges to $S \in \mathbb{B}(X)$. We have

$$(I-T)S = \lim_{n\to\infty}(I-T)\sum_{k=0}^{n} T^k = \lim_{n\to\infty}(I-T^{n+1}) = I$$

since $\|T^{n+1}\| \leq \|T\|^{n+1} \to 0$ as $n \to \infty$. Similarly, $S(I-T) = I$. Furthermore, $\|S\| = \|\sum_{k=0}^{\infty} T^k\| \leq \sum_{k=0}^{\infty} \|T\|^k = \frac{1}{1-\|T\|}$. □

Theorem. Let X be a Banach space, then

(i) $\mathcal{G}(X)$ is an open subset of $\mathbb{B}(X)$.

(ii) $T \mapsto T^{-1}$ on $\mathcal{G}(X)$ is continuous.

(iii) If $T_n \in \mathcal{G}(X)$ for all $n \in \mathbb{N}$, and $T_n \to T \notin \mathcal{G}(X)$, then $\|T_n^{-1}\| \to \infty$ as $n \to \infty$.

Proof. (i) Fix $T \in \mathcal{G}(X)$. Suppose $S \in \mathbb{B}(X)$, $\|S - T\| < \frac{1}{\|T^{-1}\|}$, then $\|I - ST^{-1}\| = \|(T-S)T^{-1}\| \leq 1$. By theorem 7, $I - (I - ST^{-1}) = ST^{-1} \in \mathcal{G}(X)$, so $S \in \mathcal{G}(X)$.
(iii) $B(T_n, 1/\|T_n^{-1}\|) \subseteq \mathcal{G}(X)$ by above, so $\|T - T_n\| \geq 1/\|T_n^{-1}\|$. Therefore $\|T_n^{-1}\| \geq 1/\|T - T_n\| \to \infty$ as $n \to \infty$.
(ii) Fix $T \in \mathcal{G}(X)$. Suppose $S \in \mathbb{B}(X)$, $\|S - T\| < \frac{1}{\|T^{-1}\|}$. Write $H = T - S$, then $\|HT^{-1}\| < 1$, so by theorem 7, $(I - HT^{-1})^{-1} \in \mathcal{G}(X)$. $S^{-1} = T^{-1}(I - HT^{-1})^{-1}$, so using the power series expansion in theorem 7,

$$\|S^{-1} - T^{-1}\| = \|\sum_{n=1}^{\infty} T^{-1}(HT^{-1})^n\| \leq \sum_{n=1}^{\infty} \|T^{-1}\|^{n+1}\|H\|^n = \frac{\|T^{-1}\|^2\|H\|}{1 - \|H\| \cdot \|T^{-1}\|} \to 0$$

as $H \to 0$, i.e. as $S \to T$. □

Let X, Y be Banach spaces. The *rank* of $T \in \mathbb{B}(X,Y)$ is $\operatorname{rank}(T) = \dim T(X)$. T is of finite rank if $\operatorname{rank}(T) < \infty$. Let $\mathbb{F}(X,Y) = \{T \in \mathbb{B}(X,Y) : \operatorname{rank}(T) < \infty\}$, then this is a subspace of $\mathbb{B}(X,Y)$. As usual, let $\mathbb{F}(X) = \mathbb{F}(X,X)$.

Example.

(1) Let $y \in Y$, $f \in X^*$, then $x \mapsto f(x)y$ has rank 1 if $f \neq 0$ and $y \neq 0$. This is usually denoted by $y \otimes f$, i.e. $(y \otimes f)(x) = \langle x, f\rangle y$.

(2) In general, if $f_1, \cdots, f_n \in X^*$ are linearly independent and $y_1, \cdots, y_n \in Y$ are linearly independent, then $\sum_{i=1}^{n} y_i \otimes f_i$ is a rank n operator.

(3) Let H be a Hilbert space, $(e_i)_{i=1}^{\infty}$ be an orthonormal basis, and $P_n = \sum_{i=1}^{n} e_i \otimes e_i$, then $P_n(x) = \sum_{i=1}^{n} \langle x, e_i\rangle e_i$, so P_n is the orthogonal projection onto $\operatorname{span}\{e_1, \cdots, e_n\}$.

For X, Y Banach spaces, $T \in \mathbb{B}(X,Y)$ is *compact* if $\overline{T(B_X)}$ is compact in Y, i.e. $T(B_X)$ is relatively compact in Y. Since Y is complete, by corollary 5.5, $T \in \mathbb{B}(X,Y)$ is compact $\iff$ $T(B_X)$ is totally bounded $\iff$ for all sequences (x_n) in B_X, (Tx_n) has a convergent subsequence in Y $\iff$ for all bounded sequence (x_n) in X, (Tx_n) has a convergent subsequence in Y. Let $\mathcal{K}(X,Y) = \{T \in \mathbb{B}(X,Y) : T \text{ is compact}\}$, $\mathcal{K}(X) = \mathcal{K}(X,Y)$.

Proposition. Let X, Y, Z be Banach spaces, then

(i) $\mathcal{K}(X,Y)$ is a closed subspace of $\mathbb{B}(X,Y)$.

(ii) Given $T \in \mathbb{B}(X,Y)$, $S \in \mathbb{B}(Y,Z)$, if one of S and T is compact, then ST is compact, i.e. $\mathcal{K}(X)$ is an ideal of $\mathbb{B}(X)$.

Proof. (i) Let $S, T \in \mathcal{K}(X,Y)$. Let (x_n) be bounded in X. S is compact, so there exists a subsequence (y_n) of (x_n) such that (Sy_n) converges. T is compact, so there exists a subsequence (z_n) of (y_n) such that (Tz_n) converges. Now $((S+T)z_n)$ converges, so $S+T$ is compact. Clearly, for scalar λ, T is compact implies λT is compact, so $\mathcal{K}(X,Y)$ is a linear subspace.

Assume $T_n \to T$ in $\mathbb{B}(X,Y)$, and $T_n \in \mathcal{K}(X,Y)$ for all n. Let $\epsilon > 0$. Choose $N \in \mathbb{N}$ such that $\|T - T_n\| < \epsilon$. T_N is compact, so $T_N(B_X)$ has a finite ϵ-net, i.e. there exists $x_1, \cdots, x_k \in B_X$ such that $T_N x_1, \cdots, T_N x_k$ form an ϵ-net of $T_N(B_X)$. Given $x \in B_X$, there exists $j \in \{1, \cdots, k\}$ such that $\|T_N x - T_N x_j\| < \epsilon$, then

$$\|Tx - Tx_j\| \leq \|Tx - T_N x\| + \|T_N x - T_N x_j\| + \|T_N x_j - T_N x\| < 3\epsilon$$

i.e. $Tx_1, \cdots, Tx_k$ is a (3ϵ)-net of $T(B_X)$. □

Example.

(1) If $T \in \mathbb{F}(X,Y)$, then T is compact, since $T(B_X) \subseteq \|T\| \cdot B_{T(X)}$, which is totally bounded because it is finite dimensional. Therefore $\overline{\mathbb{F}(X,Y)} \subseteq \mathcal{K}(X,Y)$. For "most" spaces equality holds.

(2) $I \in \mathcal{K}(X)$ if and only if $\dim X < \infty$. More generally, if $T : X \to Y$ is surjective, and $\dim Y = \infty$, then by the open mapping theorem, $T(B_X) \supseteq \delta B_Y$ for some $\delta > 0$. B_Y is not compact, so $T(B_X)$ is not compact.

(3) $T : l_2 \to l_2$, $\sum_{i=1}^{\infty} x_i e_i \mapsto \sum_{i=1}^{\infty} \lambda_i x_i e_i$ is compact if and only if $\lambda_i \to 0$. The proof is left as an exercise.

(4) For $K : [0,1] \times [0,1] \to \mathbb{R}$ continuous, $T : C[0,1] \to C[0,1]$,

$$(Tf)(x) = \int_0^1 f(y)K(x,y)dy$$

is compact by the Arzelà-Ascoli theorem.

Proposition. Let X, Y be Banach spaces, $T \in \mathbb{B}(X,Y)$, then T is compact if and only if T^* is compact.

In particular, if H is a Hilbert space, and $T \in \mathbb{B}(H)$, then T is compact if and only if T^* is compact, where T^* is the adjoint.

Proof. Hilbert space case: $\Longrightarrow$: Given a bounded sequence (x_n) in H, $(T^* x_n)$ is also bounded, so $(TT^* x_n)$ has a convergent subsequence, i.e. (x_n) has a subsequence (y_n) such that $(TT^* y_n)$ converges.

$$\begin{aligned}\|T^* y_m - T^* y_n\|^2 &= \langle T^*(y_m - y_n), T^*(y_m - y_n)\rangle = \langle TT^*(y_m - y_n), y_m - y_n\rangle \\ &\leq \|TT^*(y_m - y_n)\| \cdot \|y_m - y_n\| \leq 2 \sup \|y_n\| \cdot \|TT^* y_m - TT^* y_n\| \to 0\end{aligned}$$

Hence (T^*y_n) is Cauchy, and so convergent.

$\Longleftarrow$: T^* is compact implies $T^{**} = T$ is compact.

General case sketch: $\Longrightarrow$: Let $K = \overline{T(B_X)}$, then K is a compact Hausdorff space. Define $\theta : T^*Y^* \to C(K)$, $T^*y^* \mapsto y^*|_K$. θ is well-defined since $y^*|_K = y^*|_{\overline{T(B_X)}}$ is determined by the behaviour of $y^* \circ T = T^*y^*$ on B_X. In addition,

$$\|y^*|_K\|_\infty = \sup_{x \in B_X} |y^*(Tx)| = \sup_{x \in B_X} |T^*y^*(x)| = \|T^*y^*\|$$

so θ is isometric. It is clearly linear. Therefore, $T^*B_{Y^*}$ is totally bounded if and only if $\theta(T^*B_{Y^*})$ is totally bounded. Now apply Arzelà-Ascoli theorem.

$\Longleftarrow$: T^* is compact implies T^{**} is compact, but $T = T^{**}|_X$, so T is compact. □

Convention: From now on, the scalar field will be $\mathbb{C}$, and all Banach spaces are not $\{0\}$.

Let X be a Banach space, $T \in \mathbb{B}(X)$. The *spectrum* of T is the set

$$\sigma(T) = \{\lambda \in \mathbb{C} : \lambda I - T \text{ is not invertible}\}$$

Example.

(1) If $\dim X$ is finite, $\sigma(T) =$ set of eigenvalues of T. In particular, $\sigma(T) \neq \emptyset$.

(2) Let $T : l_2 \to l_2$ be the right shift, then T has no eigenvalues, but it will be shown later that $\sigma(T) = \Delta = \{z \in \mathbb{C} : |z| \leq 1\}$.

(3) $\sigma(I) = \{1\}$, $\sigma(\lambda I) = \{\lambda\}$.

Proposition. $\sigma(K)$ is a closed subset of $\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}$. Hence $\sigma(T)$ is compact.

Proof. $\lambda \mapsto \lambda I - T : \mathbb{C} \to \mathbb{B}(X)$ is continuous. $\sigma(T)$ is the inverse image of $\mathbb{B}(X) \backslash \mathcal{G}(X)$, which is closed by theorem 7. If $|\lambda| > \|T\|$, then $\|T/\lambda\| < 1$, so by theorem 7, $I - T/\lambda \in \mathcal{G}(X)$, so $\lambda I - T \in \mathcal{G}(X)$. □

Example. Let $T : l_2 \to l_2$ be the left shift. If $|\lambda| < 1$, then $T(1, \lambda, \lambda^2, \cdots) = (\lambda, \lambda^2, \cdots) = \lambda(1, \lambda, \cdots)$, so λ is an eigenvalue of T. Therefore, $\lambda I - T$ is not injective, so $\lambda \in \sigma(T)$. By the proposition, $\Delta \subseteq \sigma(T) \subseteq \Delta$. Thus $\sigma(T) = \Delta$.

Theorem. Let X be a Banach space, $T \in \mathbb{B}(X)$, then $\sigma(T) \neq \emptyset$

Proof. Let $\rho(T) = \mathbb{C} \backslash \sigma(T)$, the *resolvent set* of T. Let $R_T : \rho(T) \to \mathbb{B}(X)$, $R_T(\lambda) = (\lambda I - T)^{-1}$, the *resolvent map*. R_T is continuous by theorem 7.

$$\begin{aligned} R_T(\lambda) - R_T(\mu) &= (\lambda I - T)^{-1} - (\mu I - T)^{-1} \\ &= (\lambda I - T)^{-1}[(\mu I - T) - (\lambda I - T)](\mu I - T)^{-1} \\ &= (\mu - \lambda) \cdot R_T(\lambda) R_T(\mu) \\ \frac{R_T(\lambda) - R_T(\mu)}{\lambda - \mu} &= -R_T(\lambda) R_T(\mu) \to -R_T(\lambda)^2 \text{ as } \mu \to \lambda \end{aligned}$$

In addition, if $|\lambda| > \|T\|$, then

$$\|(\lambda I - T)^{-1}\| = \|\frac{1}{\lambda}(I - T/\lambda)^{-1}\| \leq \frac{1}{|\lambda|} \cdot \frac{1}{1 - \|T/\lambda\|} = \frac{1}{|\lambda| - \|T\|} \to 0$$

as $|\lambda| \to \infty$. Suppose $\sigma(T) = \emptyset$, then $\rho(T) = \mathbb{C}$.

Hilbert space case: Fix $x, y \in X$, and consider $f : \mathbb{C} \to \mathbb{C}$, $f(\lambda) = \langle R_T(\lambda)(x), y \rangle$, then

$$\frac{f(\lambda) - f(\mu)}{\lambda - \mu} = \left\langle \frac{R_T(\lambda) - R_T(\mu)}{\lambda - \mu}(x), y \right\rangle \to \langle -R_T(\lambda)^2(x), y \rangle$$

as $\mu \to \lambda$. Therefore, f is an entire function, with $|f(\lambda)| \leq \|R_T(\lambda)(x)\| \cdot \|y\| \leq \|R_T(\lambda)\| \cdot \|x\| \cdot \|y\| \to 0$ as $|\lambda| \to \infty$. By Liouville's theorem, $f = 0$. Fix λ, $\langle R_T(\lambda)(x), y \rangle = 0$ for all $x, y \in X$, so $R_T(\lambda) = 0$. Contradiction.

General Banach space case sketch: For $\phi \in \mathbb{B}(X)^*$, consider $f(\lambda) = \phi(R_T(\lambda))$. Apply the above derivation to get $f = 0$. Hahn-Banach theorem then gives $R_T(\lambda) = 0$. □

Example. Let $T : l_2 \to l_2$, $\sum x_i e_i \mapsto \sum 2^{-i} x_i e_i$. T is injective, so 0 is not an eigenvalue. $\|e_n\| = 1$ for all n, but $Te_n = 2^{-n} e_n \to 0$ as $n \to \infty$.

Let X be a Banach space, $T \in \mathbb{B}(X)$. $\lambda \in \mathbb{C}$ is an *approximate eigenvalue* of T if there exists (x_n) in X, $\|x_n\| = 1$ for all n, such that $(\lambda I - T)x_n \to 0$. Such (x_n) is called an *approximate eigenvector* of T.

Proposition. (i) If λ is an eigenvalue, then λ is an approximate eigenvalue.
(ii) If λ is an approximate eigenvalue, then $\lambda \in \sigma(T)$.

Proof. (i)$(\lambda I - T)x = 0$ for some $x \neq 0$, so take $x_n = \frac{x}{\|x\|}$.

(ii)There exists (x_n), $\|x_n\| = 1$ for all n, such that $(\lambda I - T)x_n \to 0$. Suppose there exists $S \in \mathbb{B}(X)$ with $S(\lambda I - T) = I$, then $x_n = S(\lambda I - T)x_n \to 0$. Contradiction □

Remark.

(1) In the example above, 0 is not an eigenvalue, but it is an approximate eigenvalue.

(2) If T is the right shift, then $0 \in \sigma(T)$ since T is not surjective, but $\|Tx\| = \|x\|$ for all x, so 0 is not an approximate eigenvalue.

Let $\sigma_p(T)$ = the set of eigenvalues of T (the *point spectrum* of T), and let $\sigma_{ap}(T)$ = the set of approximate eigenvalues of T (the *approximate point spectrum* of T), then

$$\sigma_p(T) \subseteq \sigma_{ap}(T) \subseteq \sigma(T) (\neq \emptyset)$$

Theorem. $\partial\sigma(T) \subseteq \sigma_{ap}(T)$, where in a topological space, the boundary of a set A is $\partial A = \overline{A} \backslash \mathbb{Z}A$. In particular, $\sigma_{ap}(T) \neq \emptyset$.

Proof. Let $\lambda \in \partial\sigma(T) = \overline{\sigma(T)} \backslash \mathbb{Z}\sigma(T) = \sigma(T) \backslash \mathbb{Z}\sigma(T)$. Then there exists (λ_n) in $\mathbb{C} \backslash \sigma(T)$ such that $\lambda_n \to \lambda$ as $n \to \infty$, so $\lambda_n I - T \to \lambda I - T$. $\lambda_n I - T \in \mathcal{G}(X)$, $\lambda I - T \notin \mathcal{G}(X)$, so by theorem 7, $\|(\lambda_n I - T)^{-1}\| \to \infty$. Then there exists (y_n) in X, $y_n \to 0$, and $\|(\lambda_n I - T)^{-1} y_n\| = 1$ for all n. Set $x_n = (\lambda_n I - T)^{-1} y_n$, then $\|x_n\| = 1$ for all n, and

$$(\lambda I - T)x_n = (\lambda_n I - T)x_n + (\lambda - \lambda_n)x_n = y_n + (\lambda - \lambda_n)x_n \to 0$$

as $n \to \infty$. Therefore, λ is an approximate eigenvalue. □

Example. Let $T : l_2 \to l_2$ be the right shift. If $|\lambda| < 1$, then $\|(\lambda I - T)x\| \geq \|Tx\| - |\lambda| \cdot \|x\|$=$(1 - |\lambda|) \cdot \|x\|$, so $\lambda \notin \sigma_{ap}(T)$. By theorem 7, $\lambda \notin \partial\sigma(T)$. We know that $0 \in \sigma(T) \subseteq \Delta$. If $z \in \Delta \backslash \sigma(T)$, then $\sigma(T)$ must have a boundary on the line segment connecting 0 to z, which is a contradiction. Therefore, $\sigma(T) = \Delta$.

For $T \in \mathbb{B}(X)$, the *spectral radius* of T is

$$r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$$

Proposition 7 and theorem 7 implies that $r(T)$ is well-defined, $r(T) \leq \|T\|$, and there exists $\lambda \in \sigma(T)$ such that $r(T) = |\lambda|$.

Observe that if $A, B \in \mathbb{B}(X)$ satisfies $AB = BA$, and suppose that AB is invertible, then for some $S \in \mathbb{B}(X)$, $SAB = ABS = I$. $A(BS) = I$, and $(BS)A = BSAABS = BSABAS = BAS = I$, so A is invertible, and so B is invertible. We have therefore shown that AB is invertible if and only if A and B are both invertible.

Theorem (Spectral Mapping Theorem)**.** For $T \in \mathbb{B}(X)$ and a polynomial $p(z) = a_0 + a_1 z + \cdots + a_n z^n$, we have $\sigma(p(T)) = \{p(\lambda) : \lambda \in \sigma(T)\}$.

Proof. For $\mu \in \mathbb{C}$, we can factorize $\mu - P(z) = c \prod_{k=1}^{n}(c_k - z)$, where $c, c_1, \cdots, c_n \in \mathbb{C}$, and $c \neq 0$, so $\mu I - p(T) = c \prod_{k=1}^{n}(c_k I - T)$. Note that the factors $c_k I - T$ pairwise commute, so by the observation above, $\mu \in \sigma(p(T))$ if and only if $c_k \in \sigma(T)$ for some k. Since $\{\lambda \in \mathbb{C} : p(\lambda) = \mu\} = \{c_1, \cdots, c_n\}$, the result follows. □

Corollary. $r(T) \leq \inf\limits_{n \in \mathbb{N}} \|T^n\|^{1/n}$

Proof. If $\lambda \in \sigma(T)$, then $\lambda^n \in \sigma(T^n)$ by the spectral mapping theorem, so $|\lambda^n| \leq \|T^n\|$, i.e. $|\lambda| \leq \|T^n\|^{1/n}$. Taking supremum over $\lambda \in \sigma(T)$ gives $r(T) \leq \|T^n\|^{1/n}$. Taking infimum over $n \in \mathbb{N}$ gives the desired result. □

Example. If T is *nilpotent* (there exists $n \in \mathbb{N}$ such that $T^n = 0$), then $r(T) = 0$, so $\sigma(T) = \{0\}$.

Theorem (Gelfand Spectral Radius Formula)**.**

$$r(T) = \inf_{n \in \mathbb{N}} \|T^n\|^{1/n} = \lim_{n \to \infty} \|T^n\|^{1/n}$$

Proof Sketch. Using the proof of theorem 7, we have the series expansion

$$R_T(\lambda) = \lambda^{-1} \mathbb{B}ig(I - \frac{T}{\lambda} \mathbb{B}ig)^{-1} = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}$$

In addition, $\rho(T) \supseteq \{z \in \mathbb{C} : |z| \geq r(T)\}$ by definition, so the series converges over the range. For any $\phi \in \mathbb{B}(X)^*$, apply ϕ to this equation to get a Laurent series with coefficients $\phi(T^n)$. Apply standard results from complex analysis and Hahn-Banach theorem to get the bound $r(T) \geq \limsup \|T^n\|^{1/n}$. □

Theorem. Let X be a Banach space, $T \in \mathcal{K}(X)$, $\lambda \in \sigma_{ap}(T)$, $\lambda \neq 0$, then $\lambda \in \sigma_p(T)$.

Proof. By assumption, there exists a sequence (x_n) in X with $\|x_n\| = 1$ for all n such that $(\lambda I - T)x_n \to 0$. T is compact, so there exists a subsequence (y_n) of (x_n) such that $Ty_n \to z$ as $n \to \infty$, then

$$y_n = \frac{1}{\lambda}[(\lambda I - T)y_n + Ty_n] \to \frac{1}{\lambda}z$$

T is continuous, so $Ty_n \to \frac{1}{\lambda}Tz$. By the uniqueness of limit, $Tz = \lambda z$. In addition, $\|z_n\| = |\lambda| \lim_{n\to\infty} \|y_n\| = |\lambda| \neq 0$, so $z \neq 0$. □

Remark. This shows that for compact T, either $\sigma(T) = \{0\}$ or T has an eigvenvalue.

For a Banach space X and $T \in \mathbb{B}(X)$, a subspace Y of X is an *invariant subspace* for T (or T acts on Y) if $T(Y) \subseteq Y$.

Example.

(1) X and $\{0\}$ are always invariant subspaces.

(2) If x is an eigenvector of T, then $\operatorname{span}\{x\}$ is invariant.

(3) If λ is an eigenvalue of T, then the eigenspace $E_T(\lambda) = \{x \in X : Tx = \lambda x\} = \ker(\lambda I - T)$ is a closed linear subspace of X which is invariant under T.

(4) The right shift on l_2 has no eigenvalues, but $\{(x_n) : x_1 = 0\}$ is a closed invariant subspace which is not X or $\{0\}$.

Invariant subspace problem: Does every $T \in \mathbb{B}(X)$ for every X with $\dim X > 1$ have a closed invariant subspace which is not X or $\{0\}$. For example, if T is compact, and $\sigma(T) \neq \{0\}$, then the answer is yes. In general, the answer is no, but the problem is open for Hilbert spaces.

From now on, we work in a Hilbert space H which is complex and non-zero.

Theorem. $\sigma(T^*) = \{\bar{\lambda} : \lambda \in \sigma(T)\}$.

Proof. Note that T is invertible if and only if T^* is invertible. For $\lambda \in \mathbb{C}$, $\lambda \in \sigma(T) \iff \lambda I - T \notin \mathcal{G}(H) \iff (\lambda I - T)^* = \bar{\lambda} I - T^* \notin \mathcal{G}(H) \iff \bar{\lambda} \in \sigma(T^*)$. □

Remark. Let T be the left shift on l_2, then $\sigma(T) = \Delta$, so $\sigma(T^*) = \sigma(\text{right shift}) = \Delta$.

Theorem. If T is Hermitian, then $\sigma(T) \subseteq \mathbb{R}$.

Proof. Let $\lambda \in \partial\sigma(T) \subseteq \sigma_{ap}(T)$, then there exists (x_n) in H such that $\|x_n\| = 1$ for all n and $(\lambda I - T)x_n \to 0$. For all $z \in H$, $\langle Tz, z\rangle = \langle z, Tz\rangle = \overline{\langle Tz, z\rangle}$, so $\langle Tz, z\rangle \in \mathbb{R}$. We have

$$\lambda = \lambda\langle x_n, x_n\rangle = \underbrace{\langle(\lambda I - T)x_n, x_n\rangle}_{\to 0} + \underbrace{\langle Tx_n, x_n\rangle}_{\in \mathbb{R}}$$

Hence $\lambda \in \mathbb{R}$, and so $\partial\sigma(T) \subseteq \mathbb{R}$. By theorem 7, $\sigma(T)$ is symmetric across the real line, so $\sigma(T) \subseteq \mathbb{R}$. □

Corollary. If T is Hermitian, then $\sigma(T) = \sigma_{ap}(T)$.

Proof. By theorem 7, $\sigma(T) \subseteq \mathbb{R}$, so $\sigma(T) = \partial\sigma(T) \subseteq \sigma_{ap}(T) \subseteq \sigma(T)$. $\square$

Corollary. If T is Hermitian, x and y are eigenvectors with eigenvalues $\lambda \neq \mu$, then $x \perp y$.

Proof. By theorem 7, $\lambda, \mu \in \mathbb{R}$, so

$$\lambda\langle x, y\rangle = \langle Tx, y\rangle = \langle x, Ty\rangle = \langle x, \mu y\rangle = \mu\langle x, y\rangle$$

Since $\lambda \neq \mu$, $\langle x, y\rangle = 0$. $\square$

Theorem. If T is Hermitian, then $r(T) = \|T\|$.

Proof. WLOG $\|T\| = 1$. Since $1 = \|T\| = \sup\{\|Tx\| : \|x\| = 1\}$, there exists a sequence (x_n) in H such that $\|x_n\| = 1$ for all n, and $\|Tx_n\| \to 1$. We have

$$\begin{aligned}\|(I - T^2)x_n\|^2 &= \langle (I - T^2)x_n, (I - T^2)x_n\rangle \\ &= \langle x_n, x_n\rangle - \langle T^2x_n, x_n\rangle - \langle x_n, T^2x_n\rangle + \langle T^2x_n, T^2x_n\rangle \\ &= \|x_n\|^2 + \|T^2x_n\|^2 - 2\|Tx_n\|^2 \\ &\leq \|x_n\|^2 + (\|T\|^2\|x_n\|)^2 - 2\|Tx_n\|^2 \leq 2 - 2\|Tx_n\|^2 \to 0\end{aligned}$$

as $n \to \infty$. Therefore, $(I - T)(I + T)x_n \to 0$. If $(I + T)x_n \to 0$, then -1 is an approximate eigenvalue. Otherwise, for some $\delta > 0$ and passing to a subsequence, $\|(I + T)x_n\| \geq \delta$ for all n. Set $y_n = \frac{(I+T)x_n}{\|(I+T)x_n\|}$, then $\|y_n\| = 1$ for all n, and $\|(I - T)y_n\| \to 0$, so 1 is an approximate eigenvalue of T. Therefore, $r(T) \geq 1$. However, $r(T) \leq 1$ by proposition 7, so $r(T) = 1$. $\square$

Theorem. Let T be compact Hermitian, then

(i) If $\lambda \in \sigma(T),\ \lambda \neq 0$, then λ is an eigenvalue of T, and $\dim E_T(\lambda) < \infty$.

(ii) $\sigma(T)$ is countable.

Proof. (i) Theorem 7 and corollary 7 gives $\lambda \in \sigma_p(T)$. Suppose $\dim E_T(\lambda) = \infty$, then there exists an infinite orthonormal sequence (x_n) in $E_T(\lambda)$. For $m \neq n$,

$$\|Tx_m - Tx_n\|^2 = \|\lambda x_m - \lambda x_n\|^2 = 2|\lambda|^2 > 0$$

so (Tx_n) has no Cauchy subsequence, contradicting the compactness of T.
(ii) For $\delta > 0$, let $S_\delta = \{\lambda \in \sigma(T) : |\lambda| > \delta\}$, so $\sigma(T)\backslash\{0\} = \bigcup_{n=1}^{\infty} S_{1/n}$. It remains to show that each S_δ is finite. If not, let $\lambda_1, \lambda_2, \cdots$ be an infinite sequence of distinct elemnts of S_δ. Let $x_1, x_2, \cdots$ be the corresponding eigenveectors with $\|x_n\| = 1$ for all n. For $m \neq n$, corollary 7 gives $x_m \perp x_n$, so

$$\|Tx_m - Tx_n\|^2 = \|\lambda_m x_m - \lambda_n x_n\|^2 = |\lambda_m|^2 + |\lambda_n|^2 \geq 2\delta^2 > 0$$

so (Tx_n) does not have a Cauchy subsequence, which is a contradiction. $\square$

Remark.

(1) This shows that either $\sigma(T)$ is finite, or $\sigma(T)$ consists of a non-zero sequence $\lambda_1, \lambda_2, \cdots$ tending to 0, together with 0.

(2) The result holds for any compact operators in a Banach space.

Lemma. Let Y be a closed subspace of H, and suppose Y is an invariant subspace of T, then $Y^\perp$ is an invariant subspace of T^*. In particular, if T is Hermitian, then T also acts on $Y^\perp$.

Proof. Let $y \in Y$, $z \in Y^\perp$, then $\langle y, T^*z \rangle = \langle Ty, z \rangle = 0$. Fixing z, this holds for all y, so $T^*z \in Y^\perp$, as required. □

Theorem (Spectral Theorem for Compact Hermitian Operators)**.** Let T be a compact Hermitian operator on H, then we have $H = Y \oplus Y^\perp$, where $Y^\perp = \ker T$, and Y has an orthonormal basis $x_1, x_2, \cdots$ of eigenvectors with corresponding non-zero eigenvalues $\lambda_1, \lambda_2, \cdots$. Moreover, $\sigma(T)$ consists of $\lambda_1, \lambda_2, \cdots$, and 0 if $Y^\perp \neq \{0\}$.

In particular, for each $x \in H$, $x = \sum_{n \geq 1} a_n x_n + z$, where $z \in \ker T$, and $Tx = \sum_{n \geq 1} \lambda_n a_n x_n$, so T is diagonalized.

Proof. For each $\lambda \in \sigma(T) \backslash \{0\}$, we select an orthonormal basis for $E_T(\lambda)$, which is finite by theorem 7. Put them together in order to obtain an orthonormal sequence $x_1, x_2, \cdots$ of eigenvectors. Set $Y = \{x_1, \cdots\}$. For each n, write $Tx_n = \lambda_n x_n$, where $\lambda_n \neq 0$. If $\lambda \in \sigma(T) \backslash \{0\}$, then $E_T(\lambda) \subseteq Y$.

Since $Tx_n \in Y$ for all n, T acts on Y. By lemma 7, T acts on $Y^\perp$. Let $S = T|_{Y^\perp}$, then $S \in \mathbb{B}(Y^\perp)$ and S is compact and Hermitian. By theorem 7, $\|S\| = r(S)$. If $\lambda \in \sigma(S) \backslash \{0\}$, then $\lambda \in \sigma_p(S)$, so λ is an eigenvalue for S, i.e. there exists $x \in Y^\perp$ such that $Tx = Sx = \lambda x$. This is a contradiction since $E_T(\lambda \subseteq Y$. Therefore $S = 0$, i.e. $\ker T \supseteq Y^\perp$. The converse is clear, so $Y^\perp = \ker T$. □

Remark.

(1) This result also holds for compact normal operators. To prove this, take $S = T^*T$, then S is compact Hermitian. By the spectral theorem $H = E_S(\lambda_1) \oplus E_S(\lambda_2) \oplus \cdots \oplus \ker S$. $\|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, Sx \rangle$, so $\ker S = \ker T$. Furthermore, $ST = TS$, so T acts on $E_S(\lambda)$, which is finite dimensional. Diagonalize T on each eigenspace to get the result.

(2) $T = \sum_{n \geq 1} \lambda_n P_n$, where P_n is the orthogonal projection onto $E_T(\lambda_n)$. More generally, treating $P(\lambda)$ as an operator-valued measure, the equation $T = \int_{\sigma(T)} \lambda dP(\lambda)$ holds for all bounded normal operators.