Contents

Lecture 1: 8/17 Vector Spaces & ℓ^p Spaces

Topics of the course:

(1) Basics of vector spaces & bases.

(2) Banach space.

(3) Hilbert space theory.

(4) Weak topologies; methods of weak convergence.

(5) Fixed-point theory and applications, & spectral theory, if time permits.

Spoiler: we didn't get to fixed-point theory :(No Brouwer's FPT.

Begin of Course

Vector Spaces

Definition 1.1

A **vector space** V over $\mathbb{K}$ (a field) is a set V, along with with notions of addition in V and multiplication by scalars, i.e., for all $\lambda \in \mathbb{K}$ and $x, y \in V$, we have

(1) $x + y \in V$, and

(2) $\lambda x \in V$.

Quick facts following from these properties:

(1) multiplication is (left) distributive: $\alpha(x + y) = \alpha x + \alpha y$,

(2) multiplication is (right) distributive: $(\alpha\beta)x = \alpha(\beta x)$,

(3) multiplication is associative: $1_{\mathbb{K}} \cdot x = x$,

(4) addition is associative: $x + (y + z) = (x + y) + z$,

(5) addition is commutative: $x + y = y + x$.

Remark

If a vector space is over $\mathbb{R}$, we call it a **real space**. If it's over $\mathbb{C}$ we call it a **complex space**.

Example 1.2

(1) $\mathbb{R}^n$ is a vector space over $\mathbb{R}$ but not over $\mathbb{C}$: take $\lambda = i \in \mathbb{C}$ and $x \in \mathbb{R}^n$, then $\lambda x \notin \mathbb{R}^n$ but $\mathbb{C}^n$.

(2) If V_1, V_2 are vector spaces then so is $V_1 \times V_2$, with

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2) \text{ and } \lambda(x_1, y_1) := (\lambda x_1, \lambda y_1).$$

(3) $C([a,b], \mathbb{K}) := \{f : [a,b] \to \mathbb{K} : f \text{ continuous}\}$ is a vector space over $\mathbb{K}$. Clearly the sum of two continuous functions is continuous, and any scalar multiple multiple of a continuous function is continuous.

Remark

It's customary to define the set of continuous functions $[a,b] \to \mathbb{R}$ simply by $C([a,b])$.

(4) $\mathcal{P}(I) := \{p : I \to \mathbb{R} \mid p(x) = \sum_{i=0}^{n} a_i x^i\}$, the set of real polynomials, is a vector space.

(5) For $p \in [1, \infty)$, $\ell^p(\mathbb{K}) := \{x : x_i \in \mathbb{K}, \sum x_i^p < \infty\}$ (all p^{th} power summable sequence) is a vector space.

For $p = \infty$, $\ell^\infty(\mathbb{K}) := \{x : x_i \in \mathbb{K}, \sup|x_i| < \infty\}$ (all bounded sequence) is a vector space. The addition and multiplication are defined by

$$x_y := \{x_i, +y_i\}_{i \geqslant 1} \text{ and } \alpha x := \{\alpha x_i\}_{i \geqslant 1}.$$

Proof. The closure of scalar multiplication is immediate, and so is the addition for $p = \infty$. For $p < \infty$, we have

$$\sum_{i \geqslant 1} |x_i + y_i|^p \leqslant \sum_{i \geqslant 1} (2 \max(|x_i|, |y_i|))^p \leqslant 2^p \left[\sum_{i \geqslant 1} |x_i|^p + \sum_{i \geqslant 1} |y_i|^p\right] < \infty.$$

□

Definition 1.3

We say U is a **subspace** of V (over the same field $\mathbb{K}$) if $U \subset V$ and U is also closed under addition and scalar multiplication, i.e., also a vector space itself.

Example 1.4

(1) $X := \{f \in C([-1,1]) : \int_{-1}^{0} f(x)\, \mathrm{d}x = \int_{0}^{1} f(x)\, \mathrm{d}x = 0\}$. This is a vector space and a subspace of $C([a,b])$.

(2) For any $y \in \mathbb{R}^n$, $U := \{x \in \mathbb{R}^n : x \cdot y = 0\}$ is a subspace of $\mathbb{R}^n$. Clearly if $x_1, x_2 \in U$ then $(x_1 + x_2) \in U$ since $(x_1 + x_2) \cdot y = x_1 \cdot y + x_2 \cdot y = 0$ and $(\alpha x) \cdot y = \alpha(x \cdot y) = 0$.

(3) $S(\mathbb{K}) := \{\{x_i\}_{i \geqslant 1} : x_i \to 0 \text{ as } i \to \infty\}$ is a subspace of $L^\infty(\mathbb{K})$. Note that this may not be a subspace of

$\ell^p(\mathbb{K})$ with p finite: given p, consider the following sequence

$$\{x_i\}_{i\geqslant 1} : x_i = \frac{1}{\sqrt[p]{i}}$$

which is not p^{th} power summable but bounded, hence $\{x_i\} \in \ell^\infty \setminus \ell^p$.

Lecture 2: 8/19 Linear Span

Definition 1.5

1) The **linear span** of a set $E \subset V$ is the collection of all finite linear combinations of elements of E:

$$\mathrm{span}(E) = \{v \in V : v = \sum_{i=1}^{n} \alpha_i e_i, n \in \mathbb{N}, \alpha_i \in \mathbb{K}, e_i \in E\}.$$

2) E **spans** V if $\mathrm{span}(E) = V$.

3) E is **linearly independent** if any finite collection of elements of E is linearly independent:

$$\sum_{i=1}^{n} \alpha_i e_i = 0 \implies \alpha 1 = \alpha 2 = \cdots = 0 \text{ for all } n \in \mathbb{N}, \alpha_i \in \mathbb{K}, e_i \in E.$$

4) $E \subset V$ is a **(Hamel) basis** of V if E is linearly independent and $\mathrm{span}(E) = V$.

Lemma 1.6

E is a Hamel basis if and only if E is maximal linearly independent, i.e., $E \cup \{\mathbf{v}\}$ becomes linearly dependent for any $\mathbf{v} \in V$.

Proof. First for $\impliedby$: We need to show $\mathrm{span}(E) = V$. Clearly $\mathrm{span}(E) \subset V$. All it suffices to show is that $V \subset \mathrm{span}(E)$. Fix $\mathbf{v} \in V$. By assumption, we know $E \cup \{\mathbf{v}\}$ is not linearly independent. Therefore for some $n \in \mathbb{N}$, $\{e_i\} \subset E$, and $\{0\} \not\subset \{\alpha_i\}_{\alpha_i \in \mathbb{N}}$ we have

$$\sum_{i=1}^{n} \alpha_i e_i + \alpha_{n+1}\mathbf{v} = 0. \tag{Δ}$$

If $\alpha_{n+1} = 0$ then by the linear independence of e_i's, all α's need to be 0, but this shows $E \cup \{\mathbf{v}\}$ is linearly independent which contradicts E's maximal linearly independence.

It follows that $\alpha_{n+1} \neq 0$, so we can proceed and divide both sides of Δ by $\alpha_n + 1$ and get a way to attain $\mathbf{v}$ by a linear combination of e_i's:

$$\mathbf{v} = -\sum_{i=1}^{n} \frac{a_i}{a_{n+1}} \cdot e_i.$$

Since $\mathbf{v}$ is arbitrary, we see that $\mathrm{span}(E) = V$.

Now for $\implies$: again, for any $\mathbf{v} \in V$, we have

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i e_i \implies \sum_{i=1}^{n} \alpha_i e_i - \mathbf{v} = 0.$$

Clearly the coefficient of $\mathbf{v}$ is $-1 \neq 0$, so it's impossible that $E \cup \{\mathbf{v}\}$ is linearly independent. □

Lemma 1.7

If E is a Hamel basis of V, then $\mathbf{v} \in V$ has a unique representation $\mathbf{v} = \sum_{i=1}^{n} \alpha_i e_i$. Furthermore, if, additionally, E is finite and has n elements, then

(1) every basis of V has n elements, and

(2) every n-element linearly independent set is a basis.

We call n the **dimension** of V, i.e., $\dim(V) := n$. If V has no finite basis then it is **infinite-dimensional**.

Example 1.8

(1) $\ell^p(\mathbb{K})$ is infinite-dimensional for all $p \in [1, \infty)$ and $p = \infty$.

Proof. Let $e^{(i)} = (0, \dots, 1, 0, \dots)$ where $e^i_j = \delta_{ij}$ with δ the **Kronecker delta** ($= 1$ if $i = j$ and 0 otherwise). Suppose there were a finite finite Hamel basis, then there were also a maximal basis $\{e^{(i)}\}_{i=1}^n$. However if we expand this set by adding $e^{(i+1)}$ we see that the new set is still linearly independent. Hence $\dim \ell^p(\mathbb{K}) = \infty$. □

Future reference: Example 3.15

(2) $C([0,1])$ is infinite-dimensional.

Proof. For any $n \in \mathbb{N}$ the set $\{1, x, \dots, x^n\} \subset C([0,1])$ is maximal linearly independent (or some set with x^n being the one with highest degree). [Check by inspection / by Wronskian / or the following] For if

$$\sum_{i=1}^n \alpha_i x^i = 0 \implies \alpha_i = 0.$$

But now $\{1, x, \dots, x^{n+1}\}$ is also linearly independent. Hence $\dim C([0,1]) = \infty$. □

Definition 1.9

A **partial order** on a set P is a binary relation $\prec$ on P such that

(1) $a \prec a$,

(2) if $a \prec b$ and $b \prec c$ then $a \prec c$, and

(3) if $a \prec b$ and $b \prec a$ then $a = b$.

For example, let $P :=$ all subsets of $\mathbb{R}$, i.e., $P := \mathbb{P}(\mathbb{R})$ and $x \prec y \iff x \subset y$. Then we have $(0,1) \subset [-1,1]$ which are comparable, and $[0,1]$ and $[1,2]$ which are not comparable. We cannot compare everything. Hence the name "partial".

Definition 1.10

(1) For $a, b \in P$, we say a, b are **comparable** if $a \prec b$ or $b \prec a$.

(2) $C \in P$ is a **chain** if any two $c_1, c_2 \in c$ are comparable.

(3) $t \in P$ is an **upper bound** of $S \subset P$ if $s \prec t$ for all $s \in S$. Some sets may not have an uppoer bound,

e.g.: $\{a, b, c\}$ with $a < b$ and $a < c$ but no info given between b and c.

(4) $m \in P$ is **maximal** if ($m < a$ for some $a \in P \implies a = m$).

Remark

Notice the difference between the definitions of maximal element and upper bound. To be an upper bound, the element need to be comparable to all other elements, but for the maximal element it just needs to make sure $a < m$ for all a comparable to m. It's possible that P has maximal elements but no upper bound.

Theorem 1.11: Zorn's Lemma, equivalent to Axiom of Choice

If P is nonempty and partially ordered, and if every chain has an upper bound, then P has at least one maximal element.
Furure reference: Hahn-Banach Theorem

Theorem 1.12

Every vector space V has a Hamel basis.

Proof. If V has a finite dimension, then it has a basis by definition. Now suppose V is infinite dimensional, and $P := \{E \subset V : E \text{ is linearly independent}\}$, i.e., the collection of all linearly independent subsets of V, and the partial order $\subset$ (inclusion). By Zorn's lemma, all it remains to show is that each chain has an upper bound.
Let $C \subset P$ be a chain and $E^* := \bigcup_{E_i \in C} E_i$ (one is the collection of these linearly independent sets and the other the union of them). Claim: E^* is an upper bound of C.
First of all, we want to show $E^* \subset P$, i.e., E^* is linearly independent. If we take any finite collection of elements of E^*, it belongs to some $E_i \in C$ by construction [this can be obtained by choosing the maximal E_i that contains at least one element of this finite collection], and this set is assumed to be linearly independent, so $E^* \in P$. Clearly $E_i \subset E^*$ for all E_i's, i.e., all elements of C, so E^* is an upper bound of the chain. Having shown this, by Zorn's lemma, we conclude that P has a maximal element, i.e., a maximal linearly independent set which, by definition, is a basis for V. □

Lecture 3: 8/21 Linear Maps and Normed Vector Spaces

Linear Maps

Definition 1.13

Let X, Y be vector spaces.

(1) We say $T \in L(x, y)$ [i.e., $T : X \to Y$ is **linear**] if $T(\alpha x + \beta x') = \alpha T(x) + \beta T(x')$.

(2) If X, Y are complex spaces then T is **conjugate linear** if $T(\alpha x + \beta x') = \overline{\alpha} T(x) + \overline{\beta} T(x')$ [where the bar denotes the complex conjugate].

Remark

$L(x, y)$ is itself a vector space.
If $T_1, T_2 \in L(x, y)$ then we can define $(\alpha T_1 + \beta T_2)(x) \coloneqq \alpha T_1(x) + \beta T_2(x)$, still a linear map $X \to Y$.

Remark

If $T \in L(X, Y)$ and $S \in L(Y, Z)$ then $S \circ T \in L(X, Z)$.

Example 1.14

(1) Consider $X, Y = \mathbb{C}$ and $T(x) = \overline{x}$.

If $\mathbb{K} = \mathbb{R}$ then T is linear, and if $\mathbb{K} = \mathbb{C}$ then T is conjugate linear, for $\overline{x} = x$ if and only if $x \in \mathbb{R} \subset \mathbb{C}$.

(2) Let $X \coloneqq C^1([a, b])$ (first order continuously differentiable) and $Y \coloneqq C([a, b])$, and define $T : f \mapsto f'$.

Definition 1.15

Let $T \in L(x, y)$.

(1) The **kernel** of T is defined as $\ker(T) \coloneqq \{x \in X : T(x) = 0\}$.

(2) The **image** of T is defined as $\operatorname{im}(T) \coloneqq \{y \in Y : \exists x \in X \text{ such that } T(x) = y\}$.

(3) T is **injective** if $T(x_1) = T(x_2) \implies x_1 = x_2$.

(4) T is **surjective** if, for all $y \in Y$, there exists $x \in X$ such that $T(x) = y$, i.e., $\operatorname{im}(T) = Y$.

(5) T is **bijective** if it's both injective and surjective. Then T has an inverse $T^{-1} : Y \to X$ defined by $T^{-1}(y) \coloneqq x$ where $y = T(x)$. *Having an inverse* is not to be confused with *being invertible*. The difference will be addressed later on.

Lemma 1.16

T is injective if and only if $\ker(T) = 0$.

Lemma 1.17

If $T \in L(X,Y)$ has an inverse, then $T^{-1} \in L(Y,X)$, i.e., also linear.

Proof. Pick $y_1, y_2 \in Y$ and $\alpha, \beta \in \mathbb{K}$. Then $y_1 = TT^{-1}(y_1)$ and $y_2 = TT^{-1}(y_2)$. Then

$$\begin{aligned} TT^{-1}(\alpha y_1 + \beta y_2) = \alpha y_1 + \beta y_2 &= \alpha TT^{-1}(y_1) + \beta TT^{-1}(y_2) \\ &= T(\alpha T^{-1}(y_1) + \beta T^{-1}(y_2)). \qquad \text{(since } T \text{ is linear)} \end{aligned}$$

Since T is bijection, it is in particular injective and so

$$T[T^{-1}(\alpha y_1 + \beta y_2)] = T(\alpha T^{-1}(y_1) + \beta T^{-1}(y_2)) \implies T^{-1}(\alpha y_1 + \beta y_2) = \alpha T^{-1}(y_1) + \beta T^{-1}(y_2),$$

i,e., $T^{-1} \in L(Y,X)$. □

Future reference: Definition 3.2, Inverse Mapping Theorem

Normed Vector Spaces

Definition 1.18

A **norm** is a map $\|\cdot\| : X \mapsto [0,\infty)$ such that, for all $x, y \in X$ and $\lambda \in \mathbb{K}$, it satisfies

(1) non-degeneracy: $\|x\| \geqslant 0$ and $\|x\| = 0 \iff x = 0$,

(2) absolute homogeneity: $\|\lambda x\| = |\lambda|\|x\|$, and

(3) triangle inequality/subadditivity: $\|x+y\| \leqslant \|x\| + \|y\|$.

A vector space X with a norm is a **normed space**.

Definition 1.19

An **open ball** is defined as $B_X(y,r) : \{x \in X : \|x-y\| < r\}$. For convenience we write $B_X(0,r)$ as $B_X(r)$.
A **closed ball** is defined as $\overline{B_X(y,r)} = \{x \in X : \|x-y\| \leqslant r\}$. Likewise for $\overline{B_X(r)}$.
†When it's clear which normed space we are in, we can drop the subscript X of $B_X(y,r)$.

Lemma 1.20

Open balls and closed balls are **convex**. Recall that the set K is convex if $\lambda x + (1-\lambda)y \in K$ for all $x, y \in K$ and $\lambda \in [0,1]$.

Proof. We want to show that, for $x_1, x_2 \in B_X(y,r)$, $\|\lambda x_1 + (1-\lambda)x_2 - y\| <$ or $\leqslant r$. This holds because

$$\begin{aligned}
\|\lambda x_1 + (1-\lambda)x_2 - y\| &= \|\lambda(x_1 - y) + (1-\lambda)(x_2 - y)\| \\
&\leqslant \|\lambda(x_1 - y)\| + \|(1-\lambda)(x_2 - y)\| \\
&= \lambda\|x_1 - y\| + (1-\lambda)\|x_2 - y\| \\
&\leqslant (\lambda + 1 - \lambda)\max\{\|1 - y\|, \|x_2 - y\|\} \\
&= \max\{\|x_1 - y\|, \|x_2 - y\|\}.
\end{aligned}$$

□

Remark

Conversely, if we have (1), (2), and the convexity of unit ball, then they can define a norm. To put formally, suppose that $N : X \to [0,\infty)$ is non-degenerate and absolutely homogeneous, and

$$\mathbb{B} := \{x : N(x) \leqslant 1\}$$

is convex, then N is a norm.

Proof. It only remains to show that N satisfies triangle inequality. Pick $x, y \in X$. The case $N(x) = 0$ is trivial since $N(x+y) = N(y) = N(x) + N(y)$. Now assume $N(x), N(y) > 0$. Notice is that, by (2),

$$N\left(\frac{x}{N(x)}\right) = \frac{1}{N(x)}N(x) = 1 \text{ and } N\left(\frac{y}{N(y)}\right) = 1,$$

so $\frac{x}{N(x)}, \frac{y}{N(y)} \in \mathbb{B}$. Since

$$\frac{x+y}{N(x)+N(y)} = \underbrace{\frac{N(x)}{N(x)+N(y)}}_{\lambda}\frac{x}{N(x)} + \underbrace{\frac{N(y)}{N(x)+N(y)}}_{1-\lambda}\frac{y}{N(y)}$$

is a convex combination of $\frac{x}{N(x)}$ and $\frac{y}{N(y)}$, by assumption it's in $\mathbb{B}$. Therefore

$$N\left(\frac{x+y}{N(x)+N(y)}\right) \leqslant 1 \iff N(x+y) \leqslant N(x) + N(y).$$

□

Future reference: proof of Minkowski's inequality, Lemma 10.17.3

Example 1.21

Let $x \in \ell^p$ be of form $\{x_i\}_{i\geqslant 1}$, and define $\|x\|_{\ell^p} := \begin{cases} \sum_{i\geqslant 1}|x_i|^p & p \in [1,\infty) \\ \sup_{i\geqslant 1}|x_i| & p = \infty \end{cases}$, then

$$\|x+y\|_{\ell^p} \leqslant \|x\|_{\ell^p} + \|y\|_{\ell^p}.$$

This is called the **Minkowski's inequality.**

Consequently, $\|\cdot\|_{\ell^p}$ is a norm and ℓ^p is a normed space.

Proof. The case $p = \infty$ is obvious as $\sup\sum \leqslant \sum\sup$. For the case $p \in [1,\infty)$, let $\mathbb{B} := \{x : \|x\|_{\ell^p} \leqslant n\}$. Let $x, y \in \mathbb{B}$. Then

$$\begin{aligned} \|\lambda x + (1-\lambda)y\|_{\ell^p} &= \sum_{i\geqslant 1}|\lambda x_i + (1-\lambda)y_i|^p \\ &\leqslant \sum_{i\geqslant 1}\left[\lambda|x_i|^p + (1-\lambda)|y_i|^p\right] \\ &\leqslant \lambda\sum_{i\geqslant 1}|x_i|^p + (1-\lambda)\sum_{i\geqslant 1}|y_i|^p \\ &\leqslant \max\{\|x\|_{\ell^p}, \|y\|_{\ell^p}\} \\ \implies \lambda x &+ (1-\lambda)y \in \mathbb{B}. \end{aligned}$$

Therefore $\mathbb{B}$ is convex and the claim follows from the equivalence of triangle inequality and convexity of ball. $\square$

Lecture 4: 8/24 Minkowski, Jensen, & Hölder; Strong Convergence

Remark

Recall that $\ell^p = \{x : \|x\|_{\ell^p} < \infty\}$. In order to hold true, every element of x needs to have finite norm.

Theorem 2.0: very bad numbering...

$\ell^p \subset \ell^q$ for $p \leqslant q$. Think of it this way: ℓ^∞ denotes all bounded sequences, but clearly not all bounded sequences are p^{th} power summable.

Proof. It suffices to show that show that $\|x\|_{\ell^q} \leqslant \|x\|_{\ell^p}$ for $p \leqslant q$.

(1) If $\|x\|_{\ell^p} = 1$, we know that each component satisfies $|x_i| \leqslant 1$ and hence

$$\|x\|_{\ell^q} = \sum_{i\geqslant 1} |x_i|^q \leqslant \sum_{i \geqslant 1} |x_i|^p = 1.$$

(2) If $\|x\|_{\ell^p} \neq 1$, we may normalize x by defining $y := x/\|x\|_{\ell^p}$ so that $\|y\| = 1$. Then

$$\|x\|_{\ell^q} = \|(y\|x\|_{\ell^p})\|_{\ell^q} \overset{\text{homog.}}{=\!=} \|x\|_{\ell^p} \underbrace{\|y\|_{\ell^q}}_{\leqslant 1 \text{ by}(1)} \leqslant \|x\|_{\ell^p}.$$

□

Theorem 2.1

Let (Ω, μ) be a measurable space, then

$$L^p(\Omega) = \begin{cases} \{f : \Omega \to \mathbb{K} \text{ such that } \|f\|^p_{L^p(\Omega)} := \int_\Omega |f|^p \, \mathrm{d}\mu < \infty\} & p \in [1, \infty) \\ \{f : \Omega \to \mathbb{K} \text{ such that } \operatorname{ess\,sup}_\Omega |f| < \infty\} & p = \infty \end{cases}$$

called the **function spaces**, are normed vector spaces.
(Elements of L^p are defined up to μ-zero measure, i.e., $f = g$ if and only if $f(x) = g(x)$ for μ-a.e.)

Proof. Non-degeneracy: $\|f\|_{L^p} = 0$ if and only if $f = 0$ μ-a.e.
Absolute homogeneity: $\|\lambda f\|_{L^p} = |\lambda| \|f\|_{L^p}$.
For triangle inequality, we again have Minkowski inequality:

$$\|f + g\|_{L^p} \leqslant \|f\|_{L^p} + \|g\|_{L^p}$$

(which can be proven using the convexity of $\mathbb{B}$ again). □

Lemma 2.2: Jensen's inequality

If $\mu(\Omega) = 1$ (a probabilistic measure) and $J : \mathbb{R} \to \mathbb{R}_+$ a convex function, then

$$J\left(\int_\Omega f \, \mathrm{d}\mu\right) \leqslant \int_\Omega J \circ f \, \mathrm{d}\mu.$$

Proof. Let $C := \int_\Omega f \, \mathrm{d}\mu$. Because J is convex, for every $z \in \mathbb{R}$, there exists some $\beta \in \mathbb{R}$ such that $J(y) \geqslant J(z) + \beta(y - z)$ for all y [simply take $\beta = J'(z)$]. Now let $z := C$ and $y := f(x)$. Then

$$J(f(x)) \geqslant J(C) + \beta(f(x) - C)$$

and integrating over Ω gives

$$\begin{aligned}\int_\Omega J(f(x)) \, \mathrm{d}\mu(x) &\geqslant J(C)\mu(\Omega) + \beta\Big(\int_\Omega f(x) \, \mathrm{d}\mu(x) - C\mu(\Omega)\Big) \\ &= J(C) + \beta(C - C) = J(C).\end{aligned}$$

Hence proven. □

Lemma 2.3: Hölder's inequality

If $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$ where $p^{-1} + q^{-1} = 1$, $p, q \in [1, \infty)$, then

$$\left| \int_\Omega fg \, \mathrm{d}\mu \right| \leqslant \|f\|_{L^p} \|g\|_{L^q}.$$

Future reference: Hölder's inequality for ℓ^p spaces

Proof. Let $F := |f||g|^{-q/p}$ and $\mu' = \mu(|g|^q / \|g\|_{L^q}^q)$ a weighted measure. Then $\mu'(\Omega) = (\int_\Omega |f|^p \, \mathrm{d}\mu)/(\|f\|_{L^p}^p) = 1$. Now we can apply Jensen's inequality with the convex function F^p [where F is positive and p a positive integer]:

$$\begin{aligned}\left(\int_\Omega F \, \mathrm{d}\mu'\right)^p &\leqslant \int_\Omega F^p \, \mathrm{d}\mu' \\ &= \int_\Omega (|f|^p |g|^{-q}) \cdot (|g|^q \|g\|_{L^q}^{-q}) \, \mathrm{d}\mu \\ &= \|g\|_{L^q}^{-q} \int_\Omega |f|^p \, \mathrm{d}\mu \\ &= \|g\|_{L^p}^{-q} \|f\|_{L^p}^p,\end{aligned}$$

whereas rewriting the LHS gives (recall $1/q + 1/p = 1$)

$$\begin{aligned}\left[\int_\Omega (|f||g|^{-q/p}) \cdot (|g|^q \|g\|_{L^q}^{-q}) \, \mathrm{d}\mu\right]^p &= \left[\int_\Omega |f||g|^{q(1-1/p)} \|g\|_{L^q}^{-q} \, \mathrm{d}\mu\right]^p \\ &= \left[\int_\Omega |f||g| \|g\|_{L^q}^{-q} \, \mathrm{d}\mu\right]^p \\ &= \|g\|_{L^q}^{-pq} \left(\int_\Omega |fg| \, \mathrm{d}\mu\right)^p.\end{aligned}$$

Therefore

$$\left(\int_\Omega |fg| \, \mathrm{d}\mu\right)^p \leqslant \|g\|_{L^q}^{q(p-1)} \|f\|_{L^p}^p.$$

Since $1/p + 1/q = 1$ we have $p + q = pq \implies q(p-1) = pq - q = p$. Since $|\int_\Omega fg \, \mathrm{d}\mu|, \|g\|_{L^q}$, and $\|f\|_{L^p}$ are all positive, we may take the p^{th} root and therefore get

$$\left| \int_\Omega fg \, \mathrm{d}\mu \right| \leqslant \int_\Omega |fg| \, \mathrm{d}\mu \leqslant \|f\|_{L^p} \|g\|_{L^q}.$$

□

Lemma 2.4

Now we generalize Ω to $\mu(\Omega) < \infty$. Then

$$L^q(\Omega) \subset L^p(\Omega) \text{ if } p \leqslant q,$$

opposite to the case of "small ℓ^p". The proof is simple using Hölder's inequality.

Strong Convergence

Definition 2.5

Let X be a normed space and $\{x_n\}_{n\geqslant 1} \subset X$ a sequence. Then $x_n \to x$ as $n \to \infty$, i.e., x_n strongly converges [in the topology on X generated by the norm] to x, if

for each $\epsilon > 0$, there exists N_ϵ such that $n \geqslant N_\epsilon \implies \|x_n - x\| < \epsilon$.

Proposition 2.6

If $x_n \to x$ in X then

(1) $\|x_n\| \to \|x\|$: this follows from

$$\begin{cases} \|x_n\| \leqslant \|x_n - x\| + \|x\| \\ \|x\| \leqslant \|x - x_n\| + \|x_n\| \end{cases} \implies \begin{cases} \|x_n\| - \|x\| \leqslant \|x_n - x\| \\ \|x\| - \|x_n\| \leqslant \|x_n - x\| \end{cases} \implies |\|x_n\| - \|x\|| \leqslant \|x_n - x\|.$$

(2) if $y_n \to y$ then $(x_n + y_n) \to x + y$: this follows from $\epsilon/2 + \epsilon/2 = \epsilon$.

(3) $\{x_n\}$ is bounded: only finitely many terms are outside $B_X(x, \epsilon)$.

(4) for $\{\alpha_n\} \subset \mathbb{K}$, $(\alpha_n x_n) \to \alpha x$ whenever $\alpha_n \to \alpha \in \mathbb{K}$. The trick here is that

$$\|\alpha_n x_n - \alpha x\| \leqslant \|\alpha_n x_n + \alpha_n x\| + \|\alpha n x - \alpha x\| = |\alpha_n| \|x_n - x\| + \|x\| |\alpha_n - \alpha|.$$

Future reference: Lemma 5.9, Corollary 4.5.1

Lecture 5: 8/26 Open and Closed Sets

Definition 2.7

(1) Given a metric space X and $A \subset X$, we say A is **open** if, for all $x \in A$, there exists some $\epsilon > 0$ such that $B(x, \epsilon) \subset A$.

(2) The **closure** of A is defined by $\overline{A} := \{x \in X : \text{ there exists } x_n \subset A \text{ with } \{x_n\} \to x\}$. If $x_n \in B(y, r)$ and $x_n \to x$ then

$$\|x - y\| \leqslant \|x - x_n\| + \|x_n - y\|$$

and

$$\limsup_{n\to\infty}(\|x - x_n\| + \|x_n - y\|) = r \implies \|x - y\| \leqslant r \iff x \in \overline{B(y, r)},$$

hence the notation of a closed ball.

(3) A is **closed** if $A = \overline{A}$, i.e., it contains all its limit points.

Example 2.8

Let U be a subspace of $(X, \|\cdot\|)$ [a normed space]. Then U *may or may not* be closed.

(1) If $\dim X < \infty$ then U is guaranteed to be closed.

(2) U is open if and only if $U = X$.

(3) Let $c_{00}(\mathbb{K}) := \{\{x_n\}_{n\geqslant 1} : x_n \neq 0 \text{ finitely many times}\}$. We have the following:

  (1) c_{00} forms a subsapce of ℓ^p for $p \in [1, \infty)$ and also $p = \infty$.

  (2) For every $p \in [0, \infty)$, not $p = \infty$, c_{00} is dense in ℓ^p [$\overline{c_{00}} = \ell^p$, closure with respect to ℓ^p norm; in other words, given any $\epsilon > 0$ and any $x \in \ell^p$ there exists $x' \in c_{00}$ such that $\|x - x'\|_{\ell^p} < \epsilon$].

  Future reference: Example 5.13

  Proof. Let $x = \{x_n\} \in \ell^p$, then $\sum_{i\geqslant 1} |x_n|^p < \infty$. Given $\epsilon > 0$, let $N \in \mathbb{N}$ be large enough such that $\sum_{n\geqslant N} |x_n|^p < \epsilon^p$ and let $x' = \{x'_n\} := (x_1, x_2, \ldots, x_{N-1}, 0, \ldots)$. Then

  $$\|x - x'\|_{\ell^p} = \left(\sum_{n\geqslant N} |x_n|^p\right)^{1/p} < \epsilon \implies c_{00} \text{ is dense in } \ell^p.$$

  Note that this statement is not true for $p = \infty$: the sequence $(1, 1, \ldots)$ is in ℓ^∞ but its ℓ^∞ norm to any sequence in c_{00} is at least 1. ☐

  (3) Obviously, for $p \in [1, \infty)$ we have $c_{00} \neq \ell^p$. From this we also see that c_{00} is not closed (since $C_{00} \neq \ell^p = \overline{c_{00}}$).

  (4) Let $(X, \|\cdot\|) := (L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$ with $\mu(\Omega) < \infty$. Let $U := \{f \in L^p : \int_\Omega f = 0\}$. Then U is a closed subspace.

Proof. Clearly U is a subspace [closure of addition and scalar multiplication]. Let $\{f_n\} \subset U$ be a sequence such that $f_n \to f$ for some $f \in L^p$. We want to show $\int_\Omega f = 0$. This is because

$$\begin{aligned}
\int_\Omega f = \int_\Omega f - 0 = \int_\Omega (f - f_n) & \\
&\leqslant \|f - f_n\|_{L^p} \|1\|_{L^{(p-1)/p}} && \text{(Hölder)} \\
&= \underbrace{\|f - f_n\|_{L^p}}_{\to 0} \cdot \underbrace{\mu(\Omega)^{p/(1-p)}}_{<\infty} && \left(\text{recall } \|f\|^k_{L^k(\Omega)} = \int_\Omega |f|^k \, \mathrm{d}\mu\right) \\
&\to 0 \text{ as } n \to \infty.
\end{aligned}$$

□

Future reference: Example 2.8.4 continued (I), Example 2.8.4 continued (II)

Definition 2.9

A normed space $(X, \|\cdot\|)$ is **separable** if there exists a countable subset that is dense in S with respect to the norm.

Lemma 2.10

Fix $(X, \|\cdot\|)$. The following are equivalent (TFAE):

(1) X is separable.

(2) The unit sphere in X, i.e., $S_X := \{x \in X : \|x\| = 1\}$, is separable.

(3) X contains a countable subset $E = \{e_i\}_{i \geqslant 1}$ whose linear span is dense in X, i.e., $\overline{\text{span}(E)} = X$.

Further reference: Lemma 2.10 (contrinued), Example 5.13, Proposition 6.5, Example 6.7, Lemma 10.11

Proof. (1) $\implies$ (2): to show (2), we need to find some countable subset $E \subset S_X$ that is dense in S_X. Suppose $\{x_n\}_{n \geqslant 1}$ is dense in x. Given $n, k \in \mathbb{N}$, if $B(x_n, 1/k) \cap S_X \neq \varnothing$, i.e., if $B(x_n, 1/k)$ intersects with the sphere, we pick one element from this intersection and add it to E. By doing so we ensure that $|E| \leqslant |\mathbb{N}^2|$, and so E is countable. Now it remains to show E is dense in S_X.
Take any $z \in S_X$ and $\epsilon > 0$. Let $k \in \mathbb{N}$ such that $1/k < \epsilon/2$, and take $x_n \in X$ such that $\|x_n - z\| < 1/k$ [doable because $\{x_n\}$ is dense in X]. Since $z \in B(x_n, 1/k) \cap S_X$, by construction of E there has to exist some $e \in E$ that belongs to this intersection. Hence

$$\|z - e\| \leqslant \|z - x_n\| + \|x_n - e\| < \frac{1}{k} + \frac{1}{k} = 2k = \epsilon$$

which completes the proof.

(2) $\implies$ (3): let $E = \{e_i\} \subset S_X$ be dense in S_X. We will show that this set is what we are actually looking for, i.e., $\overline{\text{span}(E)} = X$. Clearly $\subset$ is trivial, so it suffices to show $\supset$.

For $x = 0$, simply notice that

$$\left\| x - e\frac{\epsilon}{\|e\|} \right\| = \epsilon.$$

Now let nonzero $x \in X$ and $\epsilon > 0$ be given. Then the normalized $x/\|x\| \in S_X$. Therefore there exists $e \in E$ such that

$$\left\| \frac{x}{\|x\|} - e \right\| < \frac{\epsilon}{\|x\|} \implies \|x - \underbrace{\|x\| \cdot e}_{\in \text{span}(E)} \| < \epsilon$$

$\square$

Lecture 6: 8/28 Separable Spaces & Bounded Operators

Example 2.8: (4) continued

A follow-up on Example 2.8.4: if $\mu(\Omega) = \infty$, then $U := \{f \in L^p : \int_\Omega f = 0\}$ may not be closed.

For example, consider $\Omega := \mathbb{R}$. Try to show that there exist $f_n, f \in L^p(\mathbb{R})$ such that $\int_\mathbb{R} f_n = 0$, $f_n \to f$ in $L^p(\mathbb{R})$, but $\int_\mathbb{R} f \neq 0$. Answer to be revealed next lecture.

Lemma 2.10: continued

Now we show (3) $\implies$ (1) from Lemma 2.10: fix $(X, \|\cdot\|)$. If X contains a countable subset $E = \{e_i\}_{i\geq 1}$ whose linear span is dense in X then X is separable.

Proof. Consider the "rational span" of E, $A := \{x = \sum_{i=1}^n q_i e_i \mid n \in \mathbb{N}, q_i \in \mathbb{Q}, e_i \in E\}$. It is clear that A is countable. Now we show it's dense in X.

Let $x \in X$ and $\epsilon > 0$ be given. By assumption, since $\operatorname{span}(E)$ is dense in X, we have

$$\left\| x - \sum_{i=1}^n \alpha_i e_i \right\| \leqslant \frac{\epsilon}{2} \text{ for some } \sum_{i=1}^n \alpha_i e_i \in \operatorname{span}(E).$$

On the other hand, since $\mathbb{Q}$ is dense in $\mathbb{R}$, for $i = 1, 2, \ldots, n$, we are able to find $q_i \in \mathbb{Q}$ such that $|q_i - \alpha_i| < \epsilon/(2n\|e_i\|)$. Then

$$\left\| \sum_{i=1}^n \alpha_i e_i - \sum_{i=1}^n q_i e_i \right\| = \sum_{i=1}^n |\alpha_i - q_i| \|e_i\| < \sum_{i=1}^n \frac{\epsilon}{2n\|e_i\|} \|e_i\| = \frac{\epsilon}{2}.$$

Therefore

$$\Bigg\| x - \underbrace{\sum_{i=1}^n q_i e_i}_{\text{in span}(E)} \Bigg\| \leqslant \left\| x - \sum_{i=1}^n \alpha_i e_i \right\| + \left\| \sum_{i=1}^n \alpha_i e_i - \sum_{i=1}^n q_i e_i \right\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence proven. □

Example 2.11: Seperable spaces

(1) $\mathbb{R}^n$ and $\mathbb{C}^n$ are separable: consider $\mathbb{Q}^n$ and $\mathbb{Q}^n + i\mathbb{Q}^n$.

(2) ℓ^p is separable for $p \in [1, \infty)$: the countable dense subset is c_{00} which, in turn, is $\operatorname{span}\{e^i\}$.

(3) ℓ^∞ is not separable. Future reference: Example 5.13, Example 6.7

Proof. Let $E := \{x \in \ell^\infty : x_i \in \{0, 1\}\}$. It follows that, for all $x, y \in E$, $\|x - y\|_{\ell^\infty} = 1$. Also, E is uncountable: we can define a surjection $f : E \to [0, 1]$ by binary expansion. [Or consider the Σ_2 argument from 425a.]

We now show that every dense subset $A \subset \ell^\infty$ must be uncountable by trying to approximating E by A. Since A is dense in ℓ^∞, given any element of E, there exists elements of A arbitrarily close to it. In particular we choose this distance to be $1/3$. Fix $x, y \in E$ and let $x', y' \in A$ be such that

$$\|x - x'\|_{\ell^\infty} \leqslant \frac{1}{3} \text{ and } \|y - y'\|_{\ell^\infty} \leqslant \frac{1}{3}.$$

Then we have

$$\|x - y\|_{\ell^\infty} \leqslant \|x - x'\|_{\ell^\infty} + \|x' - y'\|_{\ell^\infty} + \|y' - y\|_{\ell^\infty},$$

i.e.,

$$\|x' - y'\|_{\ell^\infty} \geqslant \underbrace{\|x - y\|_{\ell^\infty}}_{=1} - \underbrace{\|x - x'\|_{\ell^\infty}}_{\leqslant 1/3} - \underbrace{\|y' - y\|_{\ell^\infty}}_{\leqslant 1/3} \geqslant \frac{1}{3}.$$

Therefore distinct elements of E can only be approximated by distinct elements of A, and A therefore must be uncountable. □

(4) However, $c_0 := \{x \in \ell^\infty : (x_i) \to 0\}$, the space of null sequences, is separable [see PS2].

(5) Same thing for L^p: $L^p(\Omega)$ is separable for $p \in [1, \infty)$ whereas $L^\infty(\Omega)$ is not. This is true regardless of whether $\mu(\Omega) < \infty$ or not.

Bounded Operators

Definition 2.12

Let $(X, \|\cdot\|_X, (Y, \|\cdot\|_Y)$ be two normed spaces. We say $T \in L(X, Y)$ is **continuous** (or **bounded**) if $T(x_n) \to T(x)$ (in Y) whenever $(x_n) \to x$ (in X). We denote the set of bounded operators as $B(X, Y)$.

Remark: notations

If $Y = X$ then $B(X) := B(X, X)$.
If $Y = \mathbb{K}$ then $X^* = B(X, \mathbb{K})$, the dual space. Will be discussed later.
If $T \in X^*$ we write $T(x)$ as $\langle T, x \rangle$ or ${}_{X^*}\langle T, x \rangle_X$, called the **duality pairing**.

Lemma 2.13

$T \in B(X, Y) \iff \|T(x)\|_Y \leqslant L\|x\|_X$ for some $L > 0$, i.e., for a linear operator, continuity is equivalent to Lipschitz continuity. Then the infimum of such L's is the **norm** on $B(x, y)$:

$$\|T\|_{B(X,Y)} := \inf\{L > 0 : \|T(x)\|_Y \leqslant L\|x\|_X \text{ for all } x \in X\}.$$

Proof. Lipschitz continuous $\implies$ continuous is trivial. Now we show that, if T is continuous then it's Lipschitz continuous.

Assume T is continuous, then in particular it's continuous at 0. Hence there exists some $\delta > 0$ such that

$$\|x - 0\|_X < \delta \implies \|T(x) - T(0)\|_Y < 1,$$

i.e.,

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

Hence, for all $z \neq 0$,

$$\|T(z)\| = \left\| T\left(\delta \cdot \frac{z}{\|z\|} \cdot \frac{\|z\|}{\delta}\right) \right\| = \frac{\|z\|}{\delta} \cdot \underbrace{\left\| T\left(\frac{\delta z}{\|z\|}\right) \right\|}_{<1} < \frac{\|z\|}{\delta},$$

and letting $L := 1/\delta$ completes the proof. Hence the name *bounded* operators. □

Lemma 2.14

$$\|T\|_{B(x,y)} = \sup_{\|x\|_X = 1} \|T(x)\|_Y = \sup_{\|x\|_X \leqslant 1} \|T(x)\|_Y = \sup_{x \neq 0} \frac{\|T(x)\|_Y}{\|x\|_X}. \text{ See PS2.}$$

Proof. Proof of first equality For convenience, let $\|T\|_1 := \|T\|_{B(X,Y)}$ and $\|T\|_2 : \sup_{\|x\|_X = 1} \|T(x)\|_Y$.

Pick $x \neq 0$. By the definition of supremum, we have

$$\left\| T\left(\frac{x}{\|x\|}\right) \right\| \leqslant \|T\|_2 \implies \|T(x)\| \leqslant \|T\|_2 \|x\|.$$

By definition of norm, $\|T\|_1 = \inf\{\|T\|_2\}$ and so $\|T\|_1 \leqslant \|T\|_2$.

On the other hand, if $\|x\| = 1$ then $\|T(x)\| \leqslant \|T\|_1 \|x\|$ by definition of norm, so $\|T(x)\| \leqslant \|T\|_1$. Taking supremum gives

$$\sup_{\|x\|=1} \|T(x)\| = \|T\|_2 \leqslant \|T\|_1.$$

Hence $\|T\|_1 = \|T\|_2$. □

Future reference: Example 3.1, another equivalent form of norm on ℓ^p

Example 2.15: bounded operators

Define $S_L : \ell^p \to \ell^p$ by $S_L(x) = (x_2, x_3, \dots)$ and $S_R : \ell^p \to \ell^p$ by $S_R(x) = (0, x_1, x_2, \dots)$. Then

$$\|S_R(x)\|_{\ell^p}^p = \sum_{i \geqslant 1} |x_i|^p = \|x\|_{\ell^p}^p \implies \|S_R\|_{B(\ell^p, \ell^p)} = 1,$$

$$\|S_L(x)\|_{\ell^p}^p = \sum_{i \geqslant 2} |x_i|^p \leqslant \|x\|_{\ell^p}^p \implies \|S_L\|_{B(\ell^p, \ell^p)} \leqslant 1.$$

However, = can be attained by letting the first component of x to be 0, i.e., $(0, x_2, x_3, \dots)$, in which case $\|S_L(x)\| = \|x\|$. Hence the supremum is indeed 1 and $\|S_L\|_{B(\ell^p, \ell^p)} = 1$.

Future reference: Example 8.7.1, Example 8.7.2

Lecture 7: 8/31 More on Bounded Operators; Isomorphism

Example 2.8: counterexample for 2.8.4 when $\mu(\Omega) = \infty$

From last lecture's Example 2.8.4: consider

$$f_n(x) := \begin{cases} -1 & x \in (-1,0) \\ \frac{1}{n} & x \in (0,n) \\ 0 & \text{otherwise} \end{cases} \implies \int_{\mathbb{R}} f_n(x)\,\mathrm{d}x = 0.$$

Easy to verify that $f_n \in L^p(\Omega)$ for all $p > 1$. Then $f_n \to f := \begin{cases} -1 & x \in (-1,0) \\ 0 & \text{otherwise} \end{cases}$ as $n \to \infty$ but $\displaystyle\int_{\mathbb{R}} f\,\mathrm{d}x = -n \neq 0$.

Example 3.1: more on bounded operators

(1) Let $\varphi \in C([a,b])$ and $T(f) := \displaystyle\int_a^b \varphi f$ for $f \in C([a,b])$. Then $T \in C([a,b])^*$ with $\|T\| = \|\varphi\|_{L^1}$:

$$|T(f)| \leqslant \int_a^b |\varphi f| \leqslant \|f\|_{L^\infty} \|\varphi\|_{L^1} \implies \|T\| \leqslant \|\varphi\|_{L^1},$$

and it remains to show $\|T\| \geqslant \|\varphi\|_{L^1}$. Notice that if we define $f := \operatorname{sgn}(\varphi)$ then

$$T(f) = \int_a^b |\varphi| = \|\varphi\|_{L^1}.$$

However there's a flaw with f since it might be discontinuous if φ changes sign. We fix this by introducing $\epsilon > 0$ and define $f_\epsilon(x) := \varphi/(|\varphi| + \epsilon)$ which is indeed continuous. Then $\|f_\epsilon\|_{L^\infty} \leqslant 1$ and

$$\|\varphi\|_{L^1} - T(f_\epsilon) = \int_a^b |\varphi| - \int_a^b \varphi f_\epsilon = \int_a^b \left[|\varphi| - \frac{\varphi^2}{|\varphi| + \epsilon} \right] = \int_a^b \frac{|\varphi|\epsilon}{|\varphi| + \epsilon} \leqslant (b-a)\epsilon,$$

which means $T(f_\epsilon) \uparrow T(f)$ as $\epsilon \to 0$ and we've found a successful approximation of f. Recall from Lemma 2.14 that

$$\|T\| = \sup_{\substack{\|y\|_{L^\infty} \leqslant 1 \\ y \in C([a,b])}} |T(y)| \geqslant \sup |T(f_\epsilon)| = T(f) = \|\varphi\|_{L^1}$$

so taking $\epsilon \to 0$ gives $\|T\| \geqslant \|\varphi\|_{L^1}$. Hence $\|T\| = \|\varphi\|_{L^1}$.

Remark

For functions on $C([a,b])$, L^∞ is the "usual" norm in the sense that, for a continuous function, the supremum, maximum, and essential supremum are all the same.

(2) $T(f) := f'$ [the derivative] is unbounded: taking $f_n(x) = x^n$ gives $\|f_n\|_{L^\infty} \geqslant 1$ whereas $\|f_n'\|_{L^\infty} = \|nx^{n-1}\|_{L^\infty} = n \leqslant n\|f_n\|_{L^\infty}$, hence no n works as the Lipschitz constant and T is not bounded.

Future reference: Example 5.4

Definition 3.2

$T \in B(X,Y)$ is **invertible** if

(1) it has an inverse, and

(2) the inverse $T^{-1} \in L(X,Y)$ (linearity shown by Lemma 1.17) is bounded, i.e., $T^{-1} \in B(Y,X)$.

Remark

If T and T^{-1} are bounded, then for some c_1, c_2 we have

$$T \text{ bounded: } \|T(x)\|_Y \leqslant c_2\|x\|_X \text{ and } T^{-1} \text{ bounded: } \|x\|_X \leqslant \frac{1}{c_1}\|T(x)\|_Y,$$

i.e.,

$$c_1\|x\|_X \leqslant \|T(x)\|_Y \leqslant c_2\|x\|_X \text{ for all } x \in X.$$

Such invertible T is called an **isomorphism** between normed spaces X and Y (with corresponding norms $\|\cdot\|_X, \|\cdot\|_Y$). We say X, Y are **isomorphic** (or **congruent**) if there exists an isomorphism between them. Future reference: Lemma 8.6

Definition 3.3

$T \in B(X,Y)$ is called an **isometric isomorphism** if it is isomorphic and $\|T(x)\| = \|x\|$ for all $x \in X$. The equation alone defines an **isometry**. In this case we write $X \equiv Y$.

Lemma 3.4

If $T \in L(X,Y)$ is an isometry (i.e., $\|T(x)\| = \|x\|$) and is surjective, then it is an isometric isomorphism.

Proof. If $\|T(x)\| = \|x\|$ we immediately know $T \in B(X,Y)$ and $T^{-1} \in B(Y,X)$, i.e., bounded. It remains to show that T has an inverse, which in turn requires T to be bijective. Given T is surjective it suffices to show T is injective.
Suppose $T(x) = T(y)$. Since T is an isometry we have

$$0 = T(x) - T(y) = T(x-y) \implies \|T(x-y)\| = 0 = \|x-y\| \implies x = y.$$

So T is bijective; it has an bounded inverse. □

Example 3.5

(1) $\mathbb{C}^n$ is isometrically isomorphic to $\mathbb{R}^{2n}$ via $(c_1, \dots, c_n) \mapsto (x_1, y_1, \dots, x_n, y_n)$ where $c_i = x_i + iy_i$.

Proof. Take $c := (c_1, \dots, c_n) \in \mathbb{C}^n$. Then

$$\|c\|^2_{\mathbb{C}^n} = \sum_{i=1}^n |c_i|^2 = \sum_{i=1}^n |x_i + iy_i|^2 = \sum_{i=1}^n |x_i|^2 + \sum_{i=1}^n |y_i|^2 = \|(x_1, y_1, \dots, x_n, y_n\|^2_{\mathbb{R}^{2n}}$$

which shows this map is isometric. Clearly it's also bijective. Hence an isometric isomorphism. □

(2) Let V be a finite-dimensional vector space with basis $E = \{e_i\}_{i \geqslant 1}$. Let

$$\|x\|_E := \left(\sum_{i-1}^n |\alpha_i|^2\right)^{1/2} \quad \text{where } x = \sum_{i=1}^n \alpha_i e_i.$$

Then $\|\cdot\|_E$ is a norm on V, and $(V, \|\cdot\|_E) \equiv \mathbb{K}^n$ (isometrically isomorphic).

Proof. Proof of isometric isomorphism Let $T : \mathbb{K}^n \to V$ be defined as

$$T(\alpha_1, \dots, \alpha_n) := \sum_{i=1}^n \alpha_i e_i.$$

Then T is a linear bijection (since each $x \in V$ is uniquely represented by this basis). In addition,

$$\|T(\alpha_1, \dots, \alpha_n)\|_E = \left(\sum_{i=1}^n |\alpha_i|^2\right)^{1/2} = \|(\alpha_1, \dots, \alpha_n)\|_{\mathbb{K}^n}.$$

□

Future reference: Theorem 3.10, Corollary 3.11

Definition 3.6

A subset $K \subset X$ is (**sequentially**) **compact** if any sequence $\{x_n\}_{n \geqslant 1} \subset K$ has a convergent subsequence (with limit in K). [Equivalent to **covering compactness** (every open covering has a finite subcovering) if the topology is generated by a metric, in particular a norm.]

Example 3.7

(1) A closed interval $[a, b]$ is compact in $\mathbb{R}$.

(2) *Heine-Borel*: a set $K \subset \mathbb{R}^n$ is compact if and only if it's closed and bounded.

Lecture 8: 9/2 Compact Sets; Riesz's Lemma

Lemma 3.8

(1) If $K \subset X$ is compact, it is closed and bounded.

(2) If $K \subset X$ is compact and $f : X \to Y$ then $f(K)$ is also compact, i.e., the continuous image of a compact set is compact. [Immediate since f preserves convergence of any convergent subsequence.]

(3) If $K \subset X$ is compact and $f : K \to \mathbb{R}$ is continuous then f attains its bounds, i.e., for some $\underline{x}, \overline{x} \in K$ we have

$$f(\underline{x}) = \inf f \text{ and } f(\overline{x}) = \sup f.$$

Proof. Immediate since closed subsets of $\mathbb{R}$ contains its limits. To put more formally, let $\{x_n\} \subset K$ be the sequence such that $f(x_n) \to \inf f$. By (2), there exists $f(x_{n_i})$, a subsequence of $f(x_n)$, that converges to $f(\underline{x})$ for some $\underline{x} \in K$. Hence $f(\underline{x}) = \inf f$. Likewise for $\overline{x}$ and $\sup f$. □

Future reference: Theorem 3.10

Definition 3.9

Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X are **equivalent** if there exists constants c_1, c_2 such that

$$c_1\|x\|_1 \leqslant \|x\|_2 \leqslant c_2\|x\|_1 \text{ for all } x \in X,$$

i.e., $(X, \|\cdot\|_1) \cong (X, \|\cdot\|_2)$ [isomorphic].

Theorem 3.10

If X is finite-dimensional then all norms on X are equivalent to each other.

Proof. Let E be a basis of X, and we will show that every norm $\|\cdot\|$ is equivalent to $\|\cdot\|_E$, the norm with respect to the basis. Recall from Example 3.5.2 that this means

$$\|x\|_E := \left(\sum_{i=1}^n |\alpha_i|^2\right)^{1/2} \quad \text{where } x = \sum_{i=1}^n \alpha_i e_i.$$

Then (by Cauchy-Schwarz)

$$\|x\| = \left\|\sum_{i=1}^n \alpha_i e_i\right\| \leqslant \sum_{i=1}^n |\alpha_i| \|e_i\| \leqslant \underbrace{\left(\sum_{i=1}^n |\alpha_i|^2\right)^{\frac{1}{2}}}_{\|x\|_E} \underbrace{\left(\sum_{i=1}^n \|e_i\|^2\right)^{\frac{1}{2}}}_{=:c_2 \text{ a constant}} = c_2\|x\|_E.$$

Now it remains to show there exists c_1 such that $\|x\| \geqslant c_1\|x\|_E$. Here we need to use compactness, in particular compactness of unit sphere with respect to $\|\cdot\|_E$, i.e., $S := \{x : \|x\|_E = 1\}$. To see this, we can think of $S = T(\{\alpha \in \mathbb{K}^n : \sum_{i=1}^n |\alpha_i|^2 = 1\})$. Recall again from Example 3.5.2 that T is continuous. Also, it's clear that the unit sphere in $\mathbb{K}^n$ is compact. Hence S is compact.

Since

$$|\|x\| - \|y\|| \leqslant \|x-y\| \leqslant c_2\|x-y\|_E$$

we see $\|\cdot\|$ is continuous with respect to $\|\cdot\|_E$. Hence $\|\cdot\|$ has to attain its lower bound (infimum) on S by Lemma 3.8.3 (with $f(x) = \|x\|$ on $(X, \|\cdot\|_E)$). Therefore there exists $\underline{x} \in S$ such that

$$\|\underline{x}\| = \inf_{x \in S} \|x\|.$$

Notice that if $\underline{x} \in S$ then $\|\underline{x}\|_E = 1 \neq 0 \implies \underline{x} \neq 0$ by the properties of basis. Hence

$$\|x\| = \|x\|_E \underbrace{\left\|\frac{x}{\|x\|_E}\right\|}_{\in S} \geqslant \|x\|_E \cdot \underbrace{\inf \left\|\frac{x}{\|x\|_E}\right\|}_{=:c_1} = c_1\|x\|_E.$$

Hence $(X, \|\cdot\|) \cong (X, \|\cdot\|_E)$. □

Future reference: Corollary 3.11

Corollary 3.11

If X is finite-dimensional then $K \subset X$ is compact if and only if it's closed and bounded.

Proof. By the theorem above and Example 3.5.2, if E is some basis on X, we have $(X, \|\cdot\|) \cong (X, \|\cdot\|_E) \equiv \mathbb{K}^n$. □

Future reference: Theorem 3.14, Example 4.1.4, Example 9.11, Proposition 9.21

Proposition 3.12

If $x_n \to x$ then it's Cauchy. [Immediate by $\epsilon/2$ proof.]

Lemma 3.13: Riesz's Lemma

Let $(X, \|\cdot\|)$ be a normed space and Y a closed "proper" subspace, then there exists $x \in X$ such that $\|x\| = 1$ and $\|x - y\| \geqslant 1/2$ for all $y \in Y$.

Future reference: Theorem 3.14, Example 3.15, Lemma 9.19, Theorem 10.1, Proposition 10.3

Proof. Pick $x_0 \in X \setminus Y$. Define

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

Note that $d > 0$. Otherwise, there exists $\{y_n\} \subset Y$ that converges to x_0, contradicting Y's being closed. Let $y_0 \in Y$ be such that $\|x_0 - y_0\| \in [d, 2d]$ and define

$$x := \frac{x_0 - y_0}{\|x_0 - y_0\|}.$$

Clearly $\|x\| = 1$. We will show it satisfies the problem's requirements. Take $y \in Y$. Then

$$\|x - y\| = \left\| \frac{x_0 - y_0}{\|x_0 - y_0\|} - y \right\| = \underbrace{\frac{1}{\|x_0 - y_0\|}}_{\geqslant 1/2d} \underbrace{\|x_0 - \underbrace{(y_0 + y\|x_0 - y_0\|)}_{\in Y}\|}_{\geqslant d} \geqslant \frac{1}{2}.$$

□

Theorem 3.14

A normed space $(X, \|\cdot\|)$ is finite dimensional if and only if $\overline{B_X(0,1)}$ is compact.

Future reference: Theorem 9.18

Proof. $\Longrightarrow$ is obvious by Corollary 3.11 as the closed unit ball is closed and bounded.

For $\Longleftarrow$, if X is infinite-dimensional, we will construct a sequence $\{x_n\} \subset B_X(0,1)$ such that no subsequence is Cauchy, in particular, not convergent.

We begin by taking $x_1 \in X$ with $\|x_1\| = 1$. Now look at $\operatorname{span}\{x_1\}$, a closed, proper (otherwise dimension $= 1$) subspace of X. Hence, by Riesz's Lemma there exists $x_2 \in X$ such that $\|x_2\| = 1$ and $\|x_2 - x_1\| \geqslant 1/2$. Now take $\operatorname{span}\{x_1, x_2\}$. We see this is also a closed proper subspace of X [otherwise $\dim(X) = 2$]. Hence again there exists $x_3 \in X$ such that $\|x_3\| = 1$ and $\|x_3 - x_i\| \geqslant 1/2$ for $i = 1, 2$. We may continue this the same way and get a sequence $\{x_n\}$ such that $\|x_n\| = 1$ and $\|x_n - x_i\| \geqslant 1/2$ for all $i \leqslant n$. Thus we have acquired a sequence such that none of its subsequences is Cauchy. Thus if X is infinite-dimensional, $\overline{B_X(0,1)}$ is not compact. Taking the contrapositive proves $\Longleftarrow$. □

Lecture 9: 9/4 Banach Spaces

Example 3.15: Ex. 1.8 revisited

Now we give another proof of Example 1.8.1: ℓ^p with $p \in [1, \infty]$ (either $p \in [1, \infty)$ or $p = \infty$) is infinite-dimensional.

To see this, notice that each $e^{(i)}$ has norm 1, whereas $\|e^{(i)} - e^{(j)}\|_{\ell^p} = 2^{1/p}$ for all $e^{(i)}, e^{(j)}$. Hence a sequence consisting only of $e^{(i)}$'s shows that the closed unit ball $\overline{B(0,1)}$ is not compact. By Riesz's Lemma this means ℓ^p is infinite-dimensional.

Banach Spaces

Definition 3.16

$(X, \|\cdot\|)$ is a **Banach space** (or ***X* is Banach**) if it is complete, i.e., every Cauchy sequence converges.

Example 3.17

Let $\mathcal{F}_b(\Omega; \mathbb{K}) := \{f : \Omega \to \mathbb{K} \mid f \text{ is bounded}\}$. Then $X := \mathcal{F}_b(\Omega; \mathbb{K})$ is complete with respect to $\|\cdot\|_{\sup}$ norm.

Proof. Let $\{f_n\} \subset X$ be Cauchy, i.e., given $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\|f_n - f_m\|_{\sup} := \sup_{x \in \Omega} |f_n(x) - f_m(x)| < \epsilon \text{ for all } m, n \geqslant N.$$

In particular, $\{f_n(x)\}$ is Cauchy in $\mathbb{K}$ for each x. Since $\mathbb{K}$ is complete,

$$f(x) := \lim_{n \to \infty} f_n(x)$$

is well defined. Now it remains to show $f \in X$ and $f_n \to f$. For $m, n \geqslant N$ and any $x \in \Omega$ we have

$$|f_n(x) - f_m(x)| < \epsilon \text{ and } |f_m(x) - f(x)| \to 0 \text{ as } m \to \infty.$$

Hence taking $m \to \infty$ gives $|f_n(x) - f(x)| < \epsilon$. Therefore taking the supremum gives

$$\sup_{x \in \Omega} |f_n(x) - f(x)| = \|f_n - f\|_{\sup} \leqslant \epsilon \text{ so } f_n \to f.$$

Clearly $f \in X$ since $\|f\|_{\sup} \leqslant \|f - f_n\|_{\sup} + \|f_n\|_{\sup} \leqslant \epsilon + \|f_n\|_{\sup} < \infty$. $\square$

Future reference: Example 4.1.1

Lemma 3.18

X is Banach if and only if $\sum_{i \geqslant 1} \|x_i\| < \infty \implies \sum_{i \geqslant 1} x_i$ converges in X, i.e., $\left\| \sum_{i=1}^{n} x_i - x \right\| \to 0$ as $n \to \infty$ for some $x \in X$ (every absolutely convergent series converges).

Proof. For $\Longrightarrow$, assume $\sum_{i\geqslant 1}\|x_i\| < \infty$. It follows that $\left\{\sum_{i=1}^n \|x_i\|\right\}$ converges and is thus Cauchy. Given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $m > n \geqslant N$ then $\sum_{i=1}^m \|x_i\| - \sum_{i=1}^n \|x_i\| < \epsilon$. Now we want to show $\sum_{i=1}^n x_i$ converges in X. Notice that

$$\left\|\sum_{i=1}^m x_i - \sum_{i=1}^n x_i\right\| = \left\|\sum_{i=n+1}^m x_i\right\| \leqslant \sum_{i=n+1}^m \|x_i\| = \sum_{i=1}^m \|x_i\| - \sum_{i=1}^n \|x_i\| < \epsilon$$

which shows $\left\{\sum_{i=1}^n x_i\right\}$ is Cauchy. Since X is Banach, this sequence is convergent.

For $\Longleftarrow$, take some Cauchy sequence $\{y_n\} \subset X$. We'll find a convergent subsequence $\{y_{n_k}\} \to y$. This, along with $\{y_n\}$'s being Cauchy, suffices to show $\{y_n\}$'s convergence [$\epsilon/2$ proof].

Let $n_0 = 1$ and for $k \geqslant 1$ let n_k be such that $n_k > n_{k-1}$ and

$$\|y_i - y_k\| \leqslant 2^{-k} \text{ for all } i, j \geqslant n_k.$$

(This is possible because $\{y_n\}$ is assumed to be Cauchy.) Now define a sequence $\{x_n\}$ such that $x_1 = y_{n_1}$ and $x_i = y_{n_i} - y_{n_{i-1}}$. Then

$$\sum_{i\geqslant 1}\|x_i\| \leqslant \sum_{i\geqslant 0} 2^{-i} = 2 < \infty,$$

and by assumption $\sum_{i\geqslant 1} x_i$ converges. This finishes the proof since $\sum_{i\geqslant 1} x_i = y_{n_i}$ by construction. □

Future reference: Theorem 3.19, Open Mapping Theorem, Parseval's Identity

Theorem 3.19

$X := L^p(\Omega)$ is a Banach space for all $p \in [1,\infty]$.

Proof. To make use of Lemma 3.18, let $\{f_n\} \subset L^p$ be such that $S := \sum_{n\geqslant 1}\|f_n\|_{L^p} < \infty$. We will show that $\sum_{n\geqslant 1} f_n$ converges. First notice that (with Minkowski's inequality)

$$\int_\Omega \left(\sum_{n=1}^N |f_n|\right)^p = \left\|\sum_{n=1}^N |f_n|\right\|_{L^p}^p \overset{(M)}{\leqslant} \left(\sum_{n=1}^N \|f_n\|_{L^p}\right)^p \leqslant S^p.$$

By the monotone convergence theorem, since $\sum_{n=1}^N |f_n|$ is nondecreasing, taking $N \to \infty$ gives

$$\int_\Omega F^p := \int_\Omega \left(\sum_{n\geqslant 1}|f_n|\right)^p \leqslant S^p \Longrightarrow F \in L^p.$$

In particular, $F(x) < \infty$ a.e. Therefore, for a.e. x, $f(x) := \sum_{n\geqslant 1} f_n(x)$ is well-defined (the limit exists as f) and $f \in L^p$ as $\|f\|_{L^p} \leqslant \sum_{n\geqslant 1}\|f_n\|_{L^p} = S < \infty$ (Minkowski).

Now that we've shown the existence of f and $f \in L^p$, it remains to show that $\sum_{n\geqslant 1} f_n$ actually converges to f in L^p. For $p < \infty$, this holds because (with triangle inequality and dominated convergence theorem)

$$\left\| f - \sum_{n=1}^{N} f_n \right\|_{L^p}^p = \int_\Omega \left| f - \sum_{n=1}^{N} f_n \right|^p \overset{\Delta}{\leqslant} \int_\Omega \left(|f| + \sum_{n=1}^{N} |f_n| \right)^p \leqslant \int_\Omega (2F)^p \xrightarrow{\text{DCT}} 0.$$

For $p = \infty$, refer to PS2. □

Lemma 3.20

(1) Suppose $X \cong Y$ (isomorphic). Then X is Banach if and only if Y is Banach.

(2) If $\|\cdot\|_1$ and $\|\cdot\|_2$ are two equivalent norms on X, then $(X, \|\cdot\|_1)$ is Banach if and only if $(X, \|\cdot\|_2)$ also is.

(3) If $(X, \|\cdot\|)$ is Banach and $Y \subset X$ a subspace, then $(Y, \|\cdot\|)$ is Banach if and only if Y is closed.

(4) If $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are Banach, then $(X \times Y, \|x\|_X + \|y\|_Y)$ is also Banach.

Proof. (1) Let $T : X \to Y$ be an isomorphism between X and Y. WLOG assume X is Banach. Let $\{y_n\} \subset Y$ be Cauchy. It follows that $\{x_n\} := \{T^{-1}(y_n)\} \subset X$ is also Cauchy (simply multiply ϵ by some c_1). Hence $x_n \to x$ for some $x \in X$. Then again $y_n \to y := T(x)$ (simply multiply ϵ by some other c_2).

(2) The identity map $\mathrm{id} : X \to X$ is an isomorphism between $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$. Then it just follows from (1).

(3) For $\Longleftarrow$, assume Y is closed. Take $\{y_n\} \subset Y$ a Cauchy sequence in Y. It follows that it's also Cauchy in X. Hence $y_n \to x$ for some $x \in X$. But since Y is closed, $x \in Y$ and thus $\{y_n\}$ converges in Y, i.e., Y is also Banach.

For $\Longrightarrow$, suppose Y is Banach. Let $\{y_n\} \subset Y$ such that $y_n \to x$ for some $x \in X$. We want to show $x \in Y$. Since $\{y_n\}$ is Cauchy, it's in particular convergent, so $y_n \to y$ for some $y \in Y$. Then it follows that, since limits are unique, $x = y \in Y$.

Future reference: Exercise 4.1.1, Closed Graph Theorem

(4) Trivial. *Left as exercise.*

□

Lecture 10: 9/9 BCT & PUB

Example 4.1

(1) Let $X := C_b(\Omega;\mathbb{K}) := \{f : \Omega \to \mathbb{K} \mid f \text{ is continuous \& bounded}\}$. Then it's Banach.

Proof. By Lemma 3.20.3 it suffices to show X is a closed subspace of bounded functions from $\Omega \to \mathbb{K}$, i.e., $\mathcal{F}_b(\Omega;\mathbb{K})$ from Exercise 3.17.
Let $\{f_n\} \subset X$ be such that $f_n \to f$ for some $f \in \mathcal{F}_b(\Omega;\mathbb{K})$, i.e., $\|f_n - f\|_{\text{sup}} \to 0$. We need to show f is continuous. Fix $\epsilon > 0$ and let $N \in \mathbb{N}$ be such that $\|f_N - f\|_{\text{sup}} < \epsilon/3$. Pick $\delta > 0$ such that $|f_N(x) - f_N(y)| < \epsilon/3$ whenever $|x - y| < \delta$. Then if $|x - y| < \delta$, we have

$$|f(x) - f(y)| \leqslant |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < \epsilon.$$

$\square$

(2) Let $X := C(K;\mathbb{K})$ where $K \subset \mathbb{K}^n$ is compact. Then X is Banach.

Proof. This follows from the fact that the continuous image of compact sets is bounded. Hence $C(K;\mathbb{K}) = C_b(K;\mathbb{K})$. Then it follows from the result of (1). $\square$

(3) Let $X := C^1([a,b])$ where $\|f\|_{C^1} := \|f\|_{L^\infty}\|f'\|_{L^\infty}$. Then X is Banach. Future reference: Example 5.4

Proof. Take $\{f_n\} \subset X$ Cauchy. Then $\{f_n\}, \{f_n'\} \subset C([a,b])$ are both Cauchy (with respect to $\|\cdot\|_{\text{sup}}$). Hence by (2), $f_n \to f$ and $f' \to g$ in $C([a,b])$ for some $f, g \in C([a,b])$.
It remains show $g = f'$ (then $f \in X$ and $\|f_n - f\|_{C^1} \to 0$). By FTC,

$$f_n(x) = f_n(0) + \int_a^x f_n' \,\mathrm{d}x \text{ for all } n \geqslant 0. \tag{Δ}$$

Note that, as $n \to \infty$,

$$\left| \int_a^x f_n' \,\mathrm{d}x - \int_a^x g \,\mathrm{d}x \right| \leqslant \int_a^x |f_n' - g| \,\mathrm{d}x \leqslant \|f_n' - g\|_{\text{sup}}(b - a) \to 0.$$

Hence, taking $n \to \infty$ on both sides of Δ we get

$$f(x) = f(0) + \int_a^x g \,\mathrm{d}x \text{ for all } x \implies g = f'.$$

$\square$

(4) $\mathbb{K}^n$, any finite-dimensional normed space, and $\ell^p(V)$ for $p \in [1,\infty]$ where $\dim V < \infty$, are all Banach. Second one cf. Corollary 3.11. Future reference: Example 5.11.2

Remark

Most of the time when we talk about spaces of continuous functions, we talk about functions on a compact set. If we pick X to be the space of continuous functions defined on an open interval, it's not Banach. In particular it's not even a normed space: some continuous functions don't even have finite supremum norm.

Theorem 4.2: Baire Category Theorem, BCT

Let $\{F_i\}_{i\geqslant 1}$ be a countable collection of nowhere dense (closure having empty interior/containing no open sets, cf. 425a PS9.5) subsets of a Banach space X, then

$$\bigcup_{i\geqslant 1} F_i \neq X.$$

In particular, if each F_i is closed and it so happens that $\bigcup_{i\geqslant 1} F_i = X$ then at least one of the F_i's is somewhere dense.

Future reference: Open Mapping Theorem

Remark

Baire categorized sets into two categories: **sets of 1^{st} category** refers to sets that can be expressed as countable union of nowhere dense sets, and **sets of 2^{nd} category** refers to sets that are not of 1^{st} category.

Theorem 4.3: Banach-Steinhaus Theorem/Principle of Uniform Boundedness (PUB)

Let X be a Banach space and Y a normed space. Let $S \subset B(X,Y)$ be such that

$$\sup_{T\in S} \|T(x)\| < \infty \text{ for all } x \in X,$$

then $\sup_{T\in S} \|T\| < \infty$ as well. In otherwise, pointwise boundedness everywhere implies uniform boundedness[!!]
Future reference: Example 4.4 (continued), Example 5.23 (extended), Future reference: Lemma 11.13

Proof. Let $F_i := \{x \in X : \|T(x)\| \leqslant i \text{ for all } T \in S\}$. Note that F_i is closed for each i:

$$F_i = \bigcap_{T\in S} \underbrace{\{x \in X : \|T(x)\| \leqslant i\}}_{\text{closed: } T \text{ is continuous}}$$

and arbitrary intersection of closed sets is closed. Furthermore, $X = \bigcup_{i\geqslant 1} F_i$ since, given $x \in X$, we simply need to take $i \geqslant \sup_{T\in S} \|T(x)\|$ [which is doable by assumption $\sup < \infty$]. This is a countable union of closed sets.

By BCT [*in particular* part], there exists $n \in \mathbb{N}$ such that F_n is somewhere dense; there exists some $x_0 \in X$ and $r > 0$ such that $B(x_0, r) \subset F_n$. Hence, for every $x \in X$ with $\|x\| < r$, we have

$$\|T(x)\| = \|T(\underbrace{x_0 + x}_{\in B(x_0,r)}) + T(\underbrace{-x_0}_{\in B(x_0,r)})\| \leqslant 2n$$

Hence for all $y \in X$ and $T \in S$ we have

$$\|T(y)\| = T\underbrace{\left(\frac{r}{2}\frac{y}{\|y\|}\right)}_{=r/2<r} \cdot \frac{2\|y\|}{r} \leqslant \frac{4n\|y\|}{r} \implies \|T\| \leqslant \frac{4n}{r}.$$

□

Example 4.4

Let $X = \mathcal{P}(I) := \{p : I \to \mathbb{R} : p(x) = \sum_{i=0}^{n} a_i x^i, n \geqslant 0\}$. Define the norm of a polynomial as $\|p\| := \max_i |a_i|$. Now define a sequence of operators: $T_k(p) := \sum_{i=0}^{k} a_i$. (If $k > n$ then define $a_i := 0$ for all $n < i \leqslant k$. It follows that $T_k \in B(X, \mathbb{R})$ because

$$|T_k(p)| = \left|\sum_{i=0}^{k} a_i\right| \leqslant \sum_{i=1}^{k} |a_i| \leqslant (1+k)\|p\|.$$

In particular, we've shown the operators are pointwise (polynomials as "points") bounded with $\|T_k\| \leqslant 1 + k$. Next lecture we'll show that $\|T_k\|$ is precisely $1 + k$ which then implies these operators are *NOT* uniformly bounded.

Lecture 11: 9/11 OMT & IMT

Example 4.4: Continued

Continuing from last lecture, now we show $\|T_k\| = 1 + k$. Now consider a polynomial with coefficients 1, i.e., $p := \sum_{i=0}^{k} x_i$. Then

$$\left| T_k \left(\sum_{i=0}^{n} x_i \right) \right| = \sum_{i=0}^{k} 1 = 1 + k = (1+k)\|p\|.$$

Hence we don't have uniform boundedness. By PUB we conclude X is not Banach.

Corollary 4.5: more on PUB

Let X be Banach and Y normed.

(1) If $T_n \in B(X,Y)$ are such that $\lim_{n\to\infty} T_n(x)$ exists for every $x \in X$, then $T := \lim_{n\to\infty} T_n \in B(X,Y)$, i.e., also bounded, with $\|T\| \leqslant \liminf_{n\to\infty} T_n$.

Proof. It is easy to see $T \in L(X,Y)$:

$$T(\alpha x_1 + \beta x_2) = \lim_{n\to\infty} T_n(\alpha x_1 + \beta x_2) = \lim_{n\to\infty} T_n(\alpha x_1) + \lim_{n\to\infty} T_n(\beta x_2) = T(\alpha x_1) + T(\alpha x_2).$$

For boundedness, notice that, for each $x \in X$, the convergence of $T_n(x)$ implies $\sup_{n\geqslant 1} \|T_n(x)\| < \infty$. Therefore by PUB $\sup_{n\geqslant 1} \|T_n\| < \infty$, and there exists $M \in \mathbb{N}$ such that $\|T_n\| \leqslant M$ for all n. Thus

$$\|T(x)\| \overset{\dagger}{=} \lim_{n\to\infty} \|T_n(x)\| \leqslant M\|x\| \implies \|T\| \leqslant M.$$

(† first equality from Proposition 2.6.1) □

(2) If $S \subset B(X,Y)$ is such that $\sup_{T\in S} \|T\| = \infty$, then there exists $x \in X$ such that $\sup_{T\in S} \|T(x)\| = \infty$.

Proof. Suppose not, i.e., for each $x \in X$, $\sup_{T\in S} \|T(x)\| < \infty$, then by PUB we also have $\sup_{T\in S} \|T\| < \infty$. □

Theorem 4.6: Open-Mapping Theorem

(Banach & Schauder) Let X, Y be Banach spaces and $T \in B(X,Y)$ surjective. Then T is an open mapping, i.e., T maps open sets in X to open sets in Y.

Proof. We will show that $T(B_X(0,1)) \supset B_Y(0,r)$ for some $r > 0$. Assuming this is true, then if $U \subset X$ is open, for any $y \in T(U)$ there exists $x \in U$ with $T(x) = y$, and there exists $\delta > 0$ such that $B_X(x,\delta) \subset U$. Then

$$\begin{aligned} T(U) \supset T(B_X(x,\delta)) &= T(x) + T(\delta B_X(0,1)) \\ &= y + \delta T(B_X(0,1)) \\ &= y + \delta B_Y(0,r) \\ &= B_Y(y,\delta r), \text{ so } T(U) \text{ is open.} \end{aligned}$$

Now back to the proof showing $T(B_X(0,1)) \supset B_Y(0,r)$. The proof consists of two parts. For convenience we define $B_X := B_X(0,1)$.

(1) We first show $\overline{T(B_X)}$ contains $B_Y(0,r)$ for some $r > 0$. Notice that, since T is surjective,

$$\bigcup_{n\geqslant 1} \overline{T(nB_X)} = Y = \bigcup_{n\geqslant 1} n\overline{T(B_X)}.$$

By BCT, for some $n \geqslant 1$ we know $n\overline{T(B_X)}$ is somewhere dense. Hence there exists some $y \in Y$ and $r' > 0$ such that $n\overline{T(B_X)} \supset B_Y(y,r')$. It's not hard to notice that T is convex and symmetric:

$$x, y \in B_x \implies \lambda x + (1-\lambda)y \in B_X \implies \lambda T(x) + (1-\lambda)T(y) \in T(B_X) \subset \overline{T(B_X)}$$

and

$$T(x) \in \overline{T(B_X)} \implies -T(x) \in \overline{T(B_X)}.$$

It follows that $B_Y(-y,r') \subset n\overline{T(B_X)}$. Thus for all z with $\|z\|_Y < r'$ (i.e., $z \in B_Y(0,r')$) we have

$$z = \underbrace{\frac{1}{2}(y+z)}_{\in n\overline{T(B_X)}} + \underbrace{\frac{1}{2}(y-z)}_{\in n\overline{T(B_X)}} \in n\overline{T(B_X)}.$$

Hence

$$\overline{T(B_X)} \supset B_Y(0, r'/n) =: B_Y(0,r).$$

(2) Now we show that $T(2B_X)$ contains $B_Y(0,r)$, and so $T(B_X) \supset B_Y(0,r/2)$ which would complete our proof. Take $y \in B_Y(0,r)$. By (1), there exists some $x_1 \in B_X$ such that $y - T(x_1) \in B_Y(0,r/2)$ [there exists some point in $T(B_X)$ arbitrarily close to y].

Iterating the same process again, there exists $x_2 \in B_X(0,r/2)$ such that $T(x_2)$ is arbitrarily close to $y - T(x_1)$, in particular $y - T(x_1) - T(x_2) \in B_Y(0,r/4)$. In general, for all $n \geqslant 3$ there exists $x_n \in B_X(0,1/2^{n-1})$ such that

$$y - T\Big(\sum_{i=1}^{n} x_i\Big) \in B_Y(0, r/2^n).$$

Notice that $\sum_{i\geqslant 1} \|x_i\| < \sum_{i\geqslant 1} 2^{-(i-1)} = 2$. By Lemma 3.18 this implies the convergence of $\sum x_i$ as X is Banach. Assume $x_i \to x$ for some $x \in T$. Then

$$\|y - T(x)\| \leqslant \Big\|y - T\Big(\sum_{i=1}^{n} T_n x_i\Big)\Big\| + \Big\|T\Big(x - \sum_{i=1}^{n} x_i\Big)\Big\| \leqslant \frac{r}{2^n} + \|T\|\Big\|x - \sum_{i=1}^{n} x_i\Big\| \to 0 \text{ as } n \to \infty.$$

Hence $y = T(x)$. On the other hand, from what we've just shown,

$$\|x\| = \lim_{n\to\infty} \Big\|\sum_{i=1}^{n} x_i\Big\| \leqslant \sum_{i\geqslant 1} \|x_i\| < 2,$$

which implies $x \in 2B_X$. Therefore $T(2B_X) \supset B_Y(0,r)$.

□

Theorem 4.7: Inverse Mapping Theorem/Banach Isomorphism Theorem

(Banach) Let X, Y be Banach spaces and $T \in B(X, Y)$ bijective. Then $T^{-1} \in B(Y, X)$, i.e., the boundedness of T^{-1} is guaranteed and so T is invertible.

Proof. First of all, T^{-1} exists and $T^{-1} \in L(Y, X)$ by Lemma 1.17. By the Open Mapping Theorem $T(B_X) \supset B_Y(0, r)$ for some $r > 0$. Applying T^{-1} to both sides gives $B_X \supset T^{-1}(B_Y(0, r))$. Then

$$\|T^{-1}(y)\| = \underbrace{\left\|T^{-1}\left(\frac{y}{\|y\|} \cdot \frac{r}{2}\right)\right\|}_{\in B_Y(0,r)} \cdot \frac{2\|y\|}{r} \leqslant \frac{2}{r}\|y\| \implies \|T^{-1}\| \leqslant \frac{2}{r}.$$

Future reference: Corollary 5.1, Closed Graph Theorem, Lemma 8.6, $\sigma(T)$ "iff" statement ☐

Lecture 12: 9/14 CGT; Inner Product Spaces

Corollary 5.1

A direct reseult from IMT (easy proof taking id_X as the bijection): If $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ are Banach, and it so happens that there exists some $c > 0$ such that $\|x\|_2 \leqslant c\|x\|_1$ for all $x \in X$, then the two norms are equivalent, i.e., there exists $\overline{c} > 0$ such that $\|x\|_1 \leqslant \overline{c}\|x\|_2$ for all $x \in X$.

Example 5.2

Let $X := C([0,1])$. Notice that

$$\|f\|_1 = \int_0^1 |f| \leqslant \|f\|_\infty \text{ for all } f \in x.$$

Does there exist $c > 0$ such that $\|f\|_\infty \leqslant c\|f\|_1$? The answer is no. Consider

$$f_n(x) := \begin{cases} n & n \in [0, \frac{1}{n}] \\ 0 & \text{otherwise} \end{cases} \implies \|f_n\|_1 = 1 \text{ but } \|f_n\|_\infty = n.$$

This does not contradict IMT because $(X, \|\cdot\|_1)$ is not Banach.

Corollary 5.3: Closed Graph Theorem

Let X, Y be Banach and $T \in L(X,Y)$. Then

$$T \in B(X,Y) \iff G := \{(x, T(x)) \in X \times Y : x \in X\} \text{ is closed},$$

i.e., graph of T is closed.

Proof. $\implies$: if $(x_n, T(x_n)) \to (x,y) \in X \times Y$, then clearly $x_n \to x$ and $T(x_n) \to y$. Since $T \in B(X,Y)$ it preserves sequential convergence so $T(x_n) \to T(x) \implies y = T(x)$. Hence $(x,y) \in G$.

$\impliedby$: since G is closed and T linear, G is actually a closed subspace of $X \times Y$ (since X is). Then by Lemma 3.20.3 G is also Banach (with norm $\|(x, T(x)\|_{X\times Y} := \|x\|_X + \|T(x)\|_Y$). Now define

$$\pi_x : G \to X \text{ to be the projection } \pi_x(x,y) := x.$$

It follows that, when $\|(x, T(x)\|_{X\times Y} = \|x\|_X + \|T(x)\|_Y = 1$, the supremum of $\|\pi_x(x, T(x))\| = \|x\|_X = 1$. Hence $\pi_x \in B(G, X)$ with $\|\pi_x\| = 1$. Equally clear is that π_x is bijective.

Since G and X are both Banach, by IMT $\pi_x^{-1} \in B(X, G)$, so there exists $c > 0$ such that

$$\|\pi_x^{-1}(x)\|_{X\times Y} = \|(x, T(x)\|_{X\times Y} = \|x\|_X + \|T(x)\|_Y \leqslant c\|x\|_X$$

and so $\|T(x)\|_Y \leqslant (c-1)\|x\|_X$, i.e., $T \in B(X,Y)$. $\square$

Example 5.4

Let $X := C^1([0,1])$ and $Y = C([0,1])$ and equip both with $\|\cdot\|_{\sup}$. Define $T(f) := f'$. Then the graph G of T is closed. Reason: let $(f_n, f_n') \in X \times Y$ be such that $f_n \to f$ and $f_n' \to g$ (in supremum norm) for some $f \in X$ and $g \in Y$, i.e.,

$$\|(f_n, f_n') - (f, g)\|_{X \times Y} \to 0.$$

By Example 4.1.3 this means $g = f'$, and so $(f, g) \in X \times Y$. Hence G is closed.

However, T is unbounded by Example 3.1.2. This does not contradict the CGT because X is not Banach.

Inner Product Spaces

Definition 5.5

An **inner product** (**dot product/scalar product**) on vector space V is a map $(\cdot,\cdot) : V \times V \to \mathbb{K}$ such that

(1) $(x, x) \geqslant 0$ with $(x, x) = 0 \iff x = 0$,

(2) (linear with respect to the first argument) $(x + \alpha y, z) = (x, z) + \alpha(y, z)$, and

(3) (conjugate linear w.r.t. the second arg) $(x, y) = \overline{(y, x)}$, i.e., $(x, y + \alpha z) = (x, y) + \overline{\alpha}(x, z)$.

If V has this inner product property, $(V, (\cdot,\cdot))$ is called an **inner product space**.

Example 5.6

(1) $V = \mathbb{R}^n$ and $(x, y) := \sum_{i=1}^{n} x_i y_i$ defines an inner product space. Similarly: $V = \mathbb{C}^n$ and $(x, y) := \sum_{i=1}^{n} x_i \overline{y_i}$.

Future reference: Example 5.11.1

(2) ** Let $V = \ell^2$ and $(x, y) := \sum_{i \geqslant 1} x_i \overline{y_i}$ (component-wise conjugate product of sequences). *This is a key example of an infinite-dimensional inner product space.*

Lemma 5.7

Inner product induces a norm, i.e., $\|x\| := \sqrt{(x, x)}$ defines a norm. In addition, with such norm, for all $x, y \in V$, we have the **Cauchy-Schwarz Inequality**:

$$|(x, y)| \leqslant \|x\| \|y\|.$$

Proof of Cauchy-Schwarz. Observe that

$$0 \overset{(1)}{\leqslant} (x - \lambda y, x - \lambda y) = (x, x) - \lambda(y, x) - \overline{\lambda}(x, y) + |\lambda|^2 (y, y). \qquad (\Delta)$$

If $y = 0$ the original Cauchy-Schwarz Inequality is trivial. Otherwise we may define $\lambda := \dfrac{(x,y)}{\|y\|^2}$. Then $\lambda(y,x) = \lambda\overline{(x,y)} = \dfrac{|(x,y)|^2}{\|y\|^2}$, same thing for $\overline{\lambda}(x,y)$, and $|\lambda|^2(y,y) = \dfrac{(x,y)^2}{\|y\|^4} \cdot \|y\|^2 = \dfrac{(x,y)^2}{\|y\|^2}$. Then Δ becomes

$$0 \leqslant \|x\|^2 - \frac{|(x,y)|^2}{\|y\|^2} \implies |(x,y)|^2 \leqslant \|x\|^2\|y\|^2 \implies |(x,y)| \leqslant \|x\|\|y\|.$$

□

Lemma 5.8: Parallelogram Law

$\|\cdot\|$ is (can be) induced by an inner product if and only if it satisfies the **parallelogram law**:

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

If so, then the inner product is given by the **polarization identity**

$$4(x,y) := \begin{cases} \|x+y\|^2 - \|x-y\|^2 & \text{if } \mathbb{K} = \mathbb{R} \\ \|x+y^2\| - \|x-y\|^2 + i(\|x+iy\|^2 - \|x-iy\|^2) & \text{if } \mathbb{K} = \mathbb{C} \end{cases} \tag{polarization identity}$$

Proof. $\implies$ is obvious: since $\|x\|^2 = (x,x)$, we get

$$\begin{aligned} \|x+y\|^2 + \|x-y\|^2 &= (x+y, x+y) + (x-y, x-y) \\ &= \|x\|^2 + (x,y) + (y,x) + \|y\|^2 + \|x\|^2 - (x,y) - (y,x) + \|y\|^2 \\ &= 2(\|x\|^2 + \|y\|^2). \end{aligned}$$

For $\impliedby$: this is called the **Jordan & von Neumann Theorem**. We need to show that

(1) the polarization identity satisfies the properties of inner product, and

(2) $\|x\|^2 = (x,x)$ based on the polarization identity.

Left as an exercise. □

Future reference: Example 5.11.2, Lemma 6.9

Remark

Unless specified, we will be assuming that $\|\cdot\|$ is induced by some inner product.

Lemma 5.9

Let V be an inner product space. If $x_n \to x$ and $y_n \to y$ in V, then

$$(x_n, y_n) \to (x,y).$$

This shows that the inner products are continuous with respect to strong convergence.

Proof. Once again we use the “add and subtract” trick, cf. Lemma 2.6.4:

$$\begin{aligned}
|(x_n, y_n) - (x, y)| &\leqslant |(x_n, y_n) - (x, y_n)| + |(x, y_n) - (x, y)| \\
&= |(x_n - x, y_n)| + |(x, y_n - y)| \\
&\leqslant \|x_n - x\| \|y_n\| + \|x\| \|y_n - y\| && \text{(Cauchy-Schwarz)} \\
&\to 0 \text{ since } \|x_n\|, \|y_n\| \text{ are bounded and the other two terms } \to 0.
\end{aligned}$$

□

Future reference: Theorem 5.21, Example 6.8, Lemma 6.10

Lecture 13: 9/16 Hilbert Spaces; Schauder Basis; Bessel & Parseval

Definition 5.10

A **Hilbert space** is an inner product space that is Banach, i.e., complete with respect to the norm induced by the inner product.

Example 5.11

(1) $\mathbb{R}^n$, $\mathbb{C}^n$ are Hilbert. Cf. Example 5.6.1.

(2) Among ℓ^p spaces, only ℓ^2 is Hilbert (though all are Banach, cf. Example 4.1.4). Notice that ℓ^p norm is induced by an inner product if and only if $p = 2$, where

$$(x, y) := \sum_{i \geqslant 1} x_i \overline{y_i}.$$

Proof. Consider $x := (0, 1, 0, \dots)$ and $y := (1, 0, 0, \dots)$. Then $x + y = (1, 1, 0, \dots)$ and $x - y = (-1, 1, 0, \dots)$ and so $\|x + y\|_{\ell^p} = \|x - y\|_{\ell^p} = 2^{1/p}$. On the other hand, $2(\|x\|^2 + \|y\|^2) = 4$. Hence Parallelogram Law holds if and only if $2 \cdot 2^{2/p} = 4 \iff p = 2$. $\square$

(3) Among L^p spaces, only L^2 is Hilbert, with $(f, g) := \int f\overline{g}$.

Definition 5.12

A sequence $\{e_i\}_{i \geqslant 1}$ is a **Schauder basis** for a normed space $(X, \|\cdot\|)$ if every, for all $x \in X$, there exists a unique sequence $\{\alpha_i\}_{i \geqslant 1} \in \mathbb{K}$ such that

$$x = \sum_{i \geqslant 1} \alpha_i e_i, \text{ i.e., } \left\| x - \sum_{i=1}^{n} \alpha_i e_i \right\| \to 0.$$

Remark

Note that any Schauder basis is linearly independent: if $\sum_{i=1}^{n} \alpha_i e_i = 0$ is *one* representation of 0 (with later α_i's being 0), it has to be *the* representation of 0 by definition. Hence α_i's are all 0's.

Example 5.13

$\{e^{(i)}\}_{i \geqslant 1}$ is a Schauder basis for ℓ^p where $p < \infty$ but not $p = \infty$.

Proof. For $p < \infty$, recall from Example 2.8.3 that c_{00} is dense in ℓ^p when $p < \infty$. It follows that

$$x := (x_1, x_2, \dots) = \sum_{i \geqslant 1} x_i e^{(i)}.$$

For $p = \infty$, suppose $\{e^{(i)}\}_{i\geqslant 1}$ is a Schauder basis of ℓ^∞. Then $\operatorname{span}\{e^{(i)}\} = \ell^\infty$. Since $\{e^{(i)}\}$ is countable, by Lemma 2.10.3 this means ℓ^∞ is separable, but this contradicts Example 2.11.3. $\square$

Definition 5.14

Let V be an inner product space.

(1) $x \perp y$ (x is **orthogonal** to y) if $(x, y) = 0$.

(2) $E \subset V$ is **orthonormal** if $(e_1, e_2) = 0$ for all distinct $e_1, e_2 \in E$ and also $\|e\| = 1$ for all $e \in E$.

Example 5.15

(1) $\{e^{(i)}\}_{i\geqslant 1}$ is orthonormal in ℓ^2.

(2) $\{e_k := e^{ikx}/\sqrt{2\pi}\}_{k\in\mathbb{Z}}$ is orthonormal in $L^2(-\pi, \pi)$:

$$(e_m, e_n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{imx}\overline{e^{inx}}\,\mathrm{d}x = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(m-n)x}\,\mathrm{d}x = \delta_{mn} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

Future reference: Example 5.23

(3) $\{1\} \cup \{\sqrt{2}\cos k\pi x\}_{k\geqslant 1}$ is orthonormal in $L^2(0,1)$: (clearly 1 doesn't matter) if $m = n$ then

$$2\int_0^1 \cos^2(k\pi x)\,\mathrm{d}x = \int_0^1 1 + \cos(2k\pi x)\,\mathrm{d}x = 1$$

and if $m \neq n$ then

$$\int_0^1 \cos(m\pi x)\cos(n\pi x)\,\mathrm{d}x = \frac{1}{2}\int_0^1 \cos((m+n)\pi x) + \cos((m-n)\pi x)\,\mathrm{d}x = 0.$$

Lemma 5.16: Generalized Pythagorean Theorem

If $\{e_i\}_{i\geqslant 0}$ is orthonormal (in some inner product space) then $\left\|\sum_{i=1}^n \alpha_i e_i\right\|^2 = \sum_{i=1}^n |\alpha_i|^2$.

Proof. The LHS is induced by inner product so

$$\left\|\sum_{i=1}^n \alpha_i e_i\right\|^2 = \left(\sum_{i=1}^n \alpha_i e_i, \sum_{j=1}^n \alpha_j e_j\right) = \sum_{i=1}^n\sum_{j=1}^n \alpha_i\overline{\alpha_j}\underbrace{(e_i, e_j)}_{=\delta_{ij}} = \sum_{i=1}^n \alpha_i\overline{\alpha}_i = \text{RHS}.$$

$\square$

Future reference: Parseval's Identity, Hilbert-Schmidt Theorem

Lemma 5.17: Bessel's Inequality

Let $\{e_i\}$ be orthonormal in V. Then $\sum_{i\geqslant 1}|(x,e_i)|^2 \leqslant \|x\|^2$.

Proof. Let $x_n := \sum_{i=1}^{n}(x,e_i)e_i$. By the Generalized Pythagorean Theorem, $\|x_n\|^2 = \sum_{i=1}^{n}|(x,e_i)|^2$. It follows that

$$\begin{aligned}
0 \leqslant \|x-x_n\|^2 &= (x-x_n, x-x_n) \\
&= \|x\|^2 - (x_n,x) - (x,x_n) + \|x_n\|^2 \\
&= \|x\|^2 - \left(\sum_{i=1}^{n}(x,e_i)e_i, x\right) - \left(x, \sum_{i=1}^{n}(x,e_i)e_i\right) + \|x_n\|^2 \\
&= \|x\|^2 - \sum_{i=1}^{n}[(x,e_i)(e_i,x)] - \sum_{i=1}^{n}[(x,e_i)(e_i,x)] + \|x_n\|^2 \\
&= \|x\|^2 - 2\sum_{i=1}^{n}|(x,e_i)|^2 + \|x_n\|^2 \\
&= \|x\|^2 - \|x_n\|^2
\end{aligned}$$

holds for all $n \in \mathbb{N}$. □

Future reference: Corollary 5.19, Theorem 5.21, Example 5.23 Extended, Example 11.12

Lemma 5.18: Parseval's Identity

Let $\{e_i\}_{i\geqslant 1}$ be orthonormal in a Hilbert space H (*we need both inner product and completeness*). Then

$$\sum_{i\geqslant 1}\alpha_i e_i \text{ converges} \iff \sum_{i\geqslant 1}|\alpha_i|^2 < \infty.$$

If this is true, then

$$\Big\|\sum_{i\geqslant 1}\alpha_i e_i\Big\|^2 = \sum_{i\geqslant 1}|\alpha_i|^2. \qquad \text{(Parseval's Identity)}$$

Compare this with Lemma 3.18 (absolutely convergent series converges).

Proof. $\implies$: let $x \in H$ be such that $\sum_{i=1}^{n}\alpha_i e_i \to x$. Then, by Generalized Pythagorean Theorem

$$\Big\|\sum_{i=1}^{n}\alpha_i e_i\Big\|^2 \overset{P}{=} \sum_{i=1}^{n}|\alpha_i|^2$$

and taking $n \to \infty$ gives

$$\|x\|^2 = \sum_{i\geqslant 1}|\alpha_i|^2 < \infty.$$

$\impliedby$: this follows from Lemma 3.18 since Hilbert spaces are Banach. □

Future reference: Corollary 5.19, Theorem 5.21, Theorem 6.6, Example 6.8, Lemma 6.13

Corollary 5.19

Let $\{e_i\}_{i\geqslant 1}$ be orthonormal in H. Then

$$\sum_{i\geqslant 1}(x,e_i)e_i \text{ converges for all } x \in H.$$

Quick proof: by Bessel's Inequality, $\sum_{i\geqslant 1}|(x,e_i)|^2 < \infty$, and so by Parseval's Identity this implies the convergence of $\sum_{i=1}^{n}(x,e_i)e_i$, i.e., $\sum_{i\geqslant 1}(x,e_i)e_i < \infty$.

Example 5.20

Let $e_i := e^{(2i)} \in \ell^2$. Then given $x = (x_1, x_2, \dots)$, we have $\sum_{i\geqslant 1}(x,e_i)e_i = (0, x_2, 0, x_4, \dots)$. It follows that the convergence of $\sum_{i\geqslant 1}(x,e_i)e_i$ does *not* always imply convergence to x itself. To be discussed next lecture.

Lecture 14: 9/18 Orthonormal Basis; Weierstraß Approximation Thm

Theorem 5.21: Orthonormal Schauder Basis in Hilbert Space

Let H be Hilbert and $\{e_i\}_{i\geqslant 1}$ be orthonormal. Then TFAE:

(1) $\{e_i\}$ forms a Schauder basis, in particular an **orthonormal basis**.

(2) For all $x \in H$, $\sum_{i\geqslant 1}(x, e_i)e_i \to x$.

(3) The Parseval's Identity holds, i.e., $\|x\|^2 = \sum_{i\geqslant 1}|(x, e_i)|^2$ for all $x \in H$.

(4) If $[(x, e_i) = 0$ for all i's$]$ then $x = 0$.

(5) The span of $\{e_i\}$ is dense, i.e., $\overline{\operatorname{span}\{e_i\}} = H$.

Future reference: Example 5.23, Proposition 6.5, Theorem 6.6, Example 6.8, Lemma 9.15, Corollary 13.2

Proof. (1) $\Longrightarrow$ (2): by definition $x = \sum_{i\geqslant 1}\alpha_i e_i$ for some $\alpha_i \in \mathbb{K}$. Then, by Lemma 5.9 (inner products are continuous w.r.t. strong convergence) we can take the limit out:

$$(x, e_k) = \left(\sum_{i\geqslant 1}\alpha_i e_i, e_k\right) = \left(\lim_{n\to\infty}\sum_{i=1}^{n}\alpha_i e_i, e_k\right) \overset{*}{=} \lim_{n\to\infty}\sum_{i=1}^{n}(\alpha_i e_i, e_k) = \sum_{i\geqslant 1}\alpha_i \underbrace{(e_i, e_k)}_{=\delta_{ik}} = \alpha_k. \qquad (\Delta)$$

From this it follows that $\sum_{i\geqslant 1}(x, e_i)e_i = \sum_{i\geqslant 1}\alpha_i e_i = x$ by assumption.

(2) $\Longrightarrow$ (1): take $x \in H$ so there exists some expansion $\sum_{i\geqslant 1}(x, e_i)e_i$. By Δ we know the expansion $\sum_{i\geqslant 1}(x, e_i)e_i$ corresponds uniquely to $\sum_{i\geqslant 1}\alpha_i e_i$.

(2) $\Longrightarrow$ (3): this is nothing but Parseval's Identity. For all $x \in H$, $\sum_{i\geqslant 1}(x, e_i)e_i$ converges, so $\sum_{i\geqslant 1}|(x, e_i)|^2$ is the same as $\left\|\sum_{i\geqslant 1}(x, e_i)e_i\right\|^2 = \|x\|^2$ by (2)'s assumption.

(3) $\Longrightarrow$ (4): trivial by definition of norm since $\sum_{i\geqslant 1}0 = 0 = \|x\|^2$.

(4) $\Longrightarrow$ (2): let $x \in H$ be given, and define $y_n := x - \sum_{i=1}^{n}(x, e_i)e_i$. Recall from Bessel's Inequality that $\sum_{i\geqslant 1}|(x, e_i)|^2 \leqslant \|x\|^2$. Hence $\sum_{i=m}^{n}|(x, e_i)|^2$ can be made arbitrarily small, and thus $\{y_n\}$ is Cauchy:

$$\|y_n - y_m\|^2 = \left\|\sum_{i=m}^{n}(x, e_i)e_i\right\|^2 \overset{P}{=} \sum_{i=m}^{n}|(x, e_i)|^2 < \epsilon.$$

Since H is Hilbert, $\{y_n\} \to y$ for some $y \in H$. Since inner product is continuous, for all i we have

$$(y, e_i) = \lim_{n\to\infty}(y_n, e_i) = (x, e_i) - \sum_{j=1}^{n}(x, e_j)\underbrace{(e_j, e_i)}_{=\delta_{ij}} = 0 \text{ if } j \geqslant i.$$

Then by (4)'s assumption $y = 0$, and the construction of y_n implies $x = \sum_{i\geqslant 1}(x, e_i)e_i$.

(2) $\implies$ (5): trivial. Given $x \in H$, then $\left\|x - \sum_{i=1}^{n}(x, e_i)e_i\right\| \to 0$ by (2), but all elements of form $\sum_{i=1}^{n}(x, e_i)e_i$ belongs to the span of $\{e_i\}$. Hence x is a limit point of span$\{e_i\}$.

(5) $\implies$ (4): suppose $(x, e_i) = 0$ for all i. By (5), let $\{x_n\} \subset \operatorname{span}\{e_i\}$ be such that $x_n \to x$. Then (again since inner product is continuous)

$$\|x\|^2 = (x, x) = \lim_{n\to\infty}(\underbrace{x_n}_{\in \operatorname{span}\{e_i\}}, x) = 0 \implies x = 0.$$

□

Example 5.22

$\{e^{(i)}\}_{i\geqslant 1}$ is an orthonormal basis for ℓ^2. This is immediate if we use (4) above. Also immediate since this basis is dense in ℓ^2, i.e., (5).

Example 5.23

$\{e_k := e^{ikx}/\sqrt{2\pi}\}$ is an orthonormal basis for $L^2(-\pi, \pi)$. By Example 5.15.2 we know this set is orthonormal. It remains to show it is a Schauder basis, in particular the density of its span. We'll finish this proof later, but first, some lemmas. Also refer to Example 5.23 (extended).

Future reference: Example 12.4.2, Example 13.11

Lemma 5.24

$C([a, b])$ is dense in L^p for $p \in [1, \infty)$ but not in L^∞.

Future reference: Example 12.4.2, Theorem 13.9

Proof. Let $f \in L^p$ and $\epsilon > 0$ be given, and define $g := \sum_{i=1}^{n}\alpha_i\chi_{I_i}$ (simple functions) such that $\|f - g\|_p < \epsilon/2$. Now we can approximate g by a continuous function $\overline{g} \in L^p$ by "shrinking" each I_i by some $\delta > 0$ and interpolating them, i.e., creating new I_n's with $\delta/2$ truncated on both sides. Then we let g and $\overline{g}$ agree on each I_n' and connect the endpoints of one I'_n with another (and connecting to 0 for the leftmost and rightmost intervals, i.e., I'_1 and I'_n). Note that $\|g\|_\infty \leqslant \sum_{i=1}^{n}|\alpha_i| < \infty$. Hence $g \in L^p$ and the same applies to $\overline{g}$. Hence $|g - \overline{g}| \leqslant 2\|g\|_\infty$ a.e. Notice

$$\|g - \overline{g}\|_p^p = \int_a^b |g - \overline{g}|^p = \int_{[a,b]\setminus\bigcup I'_n}(2\|g\|_\infty)^p$$

where the measure of $[a, b] \setminus \bigcup_{i=1}^{n} I'_i$ can be bounded by $n\delta$. Hence

$$\|g - \overline{g}\|_p^p \leqslant n\delta(2\|g\|_\infty)^p < \left(\frac{\epsilon}{2}\right)^p \text{ for sufficiently small } \delta,$$

and we are done.

Future reference: Example 5.23 □

Lemma 5.25: Weierstraß Approximation Theorem

Given $f \in C([0,1])$,

$$P_n(x) := \sum_{i=0}^{n} f(k/n) \binom{n}{k} x^k (1-x)^{n-k}$$

approximates f in $\|\cdot\|_\infty$, i.e., $\|f - P_n\|_\infty \to 0$. These polynomials are called **Bernstein polynomials**.

Proof. For convenience define $r_k(x) := \binom{n}{k} x^k (1-x)^{n-k}$. First, some claims of $r_k(x)$:

(1) Recall from the binomial expansion

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}. \tag{1}$$

Differentiating with respect to x and then multiplying by x give

$$nx(x+y)^{n-1} = \sum_{k=0}^{n} \binom{n}{k} k x^k y^{n-k}. \tag{2}$$

Differentiating the original binomial expansion twice and multiplying by x^2 give

$$n(n-1)(x+y)^{n-2} = \sum_{k=0}^{n} \binom{n}{k} k(k-1) x^k y^{n-k} \tag{3}$$

Letting $y := 1 - x$ gives $r_k(x) = \binom{n}{k} x^k y^{n-k}$. Then from (1), (2), and (3) we get

$$\sum_{k=0}^{n} r_k(x) = 1, \sum_{k=0}^{n} k r_k(x) = nx, \text{ and } \sum_{k=0}^{n} k(k-1) r_k(x) = n(n-1)x^2. \tag{Δ}$$

(2) Second claim:

$$\sum_{k=0}^{n} (k - nx)^2 r_k(x) = nx(1-x). \tag{∇}$$

Rewriting $(k-nx)^2 = k^2 - 2knx + n^2x^2 = k(k-1) + k - 2knx + n^2x^2$ and using Δ, we get

$$\sum_{k=0}^{n} (k-nx)^2 r_k(x) = \underbrace{n(n-1)x^2}_{\text{third}} + \underbrace{nx - 2nx \cdot nx}_{\text{second}} + \underbrace{n^2x^2}_{\text{first}} = nx(1-x).$$

To be continued next lecture. □

Lecture 15: 9/21 More on WAT; Trig Approximations; Gram-Schmidt

Proof of Weierstraß Approximation Theorem (continued). Since $f \in C([0,1])$ it is bounded. Let $M > 0$ be such that $|f(x)| \leqslant M$ for all x, and fix $\epsilon > 0$. In addition, f is uniformly continuous, so there exists $\delta > 0$ such that $|f(x) - f(y)| < \epsilon/2$ whenever $|x - y| < \delta$. Since

$$\begin{aligned}
|f(x) - P_n(x)| &= |1 \cdot f(x) - P_n(x)| \\
&= \Big| \sum_{k=0}^{n} f(x) r_k(x) - \sum_{k=0}^{n} f(k/n) r_k(x) \Big| \\
&= \Big| \sum_{k=0}^{n} r_k(x)(f(x) - f(k/n)) \Big| \\
&= \sum_{k=0}^{n} r_k(x)|f(x) - f(k/n)| \\
&= \sum_{\substack{k=0 \\ |x-\frac{k}{n}|<\delta}}^{n} r_k(x)|f(x) - f(k/n)| + \sum_{\substack{k=0 \\ |x-\frac{k}{n}|\geqslant\delta}}^{n} r_k(x)|f(x) - f(k/n)| \\
&< \sum_{\substack{k=0 \\ |x-\frac{k}{n}|<\delta}}^{n} r_k(x)\frac{\epsilon}{2} + \sum_{\substack{k=0 \\ |x-\frac{k}{n}|\geqslant\delta}}^{n} r_k(x)(2M \cdot 1) && \text{(uniform continuity \& boundedness of } f) \\
&\leqslant \sum_{\substack{k=0 \\ |x-\frac{k}{n}|<\delta}}^{n} r_k(x)\frac{\epsilon}{2} + \sum_{\substack{k=0 \\ |x-\frac{k}{n}|\geqslant\delta}}^{n} r_k(x) \cdot 2M \cdot \frac{|k-nx|^2}{n^2\delta^2} && \left(\text{since } 1 \leqslant \left(\frac{|k/n - x|}{\delta}\right)^2 = \frac{|k-nx|^2}{n^2\delta^2}\right) \\
&\leqslant \frac{\epsilon}{2}\sum_{k=0}^{n} r_k(x) + \frac{2M}{n^2\delta^2}\sum_{k=0}^{n}|k-nx|^2 r_k(x) \\
&= \frac{\epsilon}{2} + \frac{2Mx(1-x)}{n\delta^2} && \text{(by Claims 1 \& 2)} \\
&\leqslant \frac{\epsilon}{2} + \frac{M}{2n\delta^2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \text{ for sufficiently large } n.
\end{aligned}$$

□

Corollary 6.1

The set of polynomials, $P([a,b]) \subset C([a,b])$, is dense in $C([a,b])$. Future reference: Example 12.4.2

Theorem 6.2

If $K \subset \mathbb{C}$ is compact, then $P_{\mathbb{C}} := \{\sum_{i=0}^{n} a_i z^i : a_i \in \mathbb{C}\}$ is dense in $C(K;\mathbb{C})$. This can be proven by **Stone-Weierstraß Theorem**, a more general case of the WAT.

Corollary 6.3

(1) Let $X := \{f \in C([-\pi,\pi];\mathbb{C}) : f(-\pi) = f(\pi)\}$, the set of all continuous complex functions on $[-\pi,\pi]$ that agree at endpoints, then

$$\mathfrak{T} := \Big\{ \sum_{k=-n}^{n} c_k e^{ikx} : c_k \in \mathbb{C}, n \in \mathbb{N} \Big\},$$

the set of trigonometric polynomials of above form, is dense in X.

Proof. Let $f \in X$ and $\epsilon > 0$ be given. Let $g : S^1 \to \mathbb{C}$ (unit circle in complex plane as domain) be such that $g(e^{ix}) = f(x)$. Then g is continuous [why?]. Then by Theorem 6.2 there exists $n \in \mathbb{N}$ and $\{c_i\} \subset \mathbb{C}$ such that

$$\left|g(z) - \sum_{i=-n}^{n} c_i z^i\right| < \epsilon \text{ for all } z \in S^1.$$

Taking $z = e^{ix}$ gives

$$\left|f(x) - \sum_{i=-n}^{n} c_i e^{ikx}\right| < \epsilon \text{ for all } x \in [-\pi, \pi].$$

□

(2) If $f \in X$ is real then it can be approximated by a real trigonometric polynomial, i.e., from $\mathfrak{T}$ above but such that $c_k = \overline{c_{-k}}$.

Proof. Notice that if we write $c_n = a_n + ib_n$ then

$$\begin{aligned} c_n e^{inx} + c_{-n} e^{i(-n)x} &= (a_n + ib_n)(\cos nx + i \sin nx) + (a_{-n} + ib_{-n})(\cos nx - i \sin nx) \\ &= \cos nx(a_n + a_{-n}) - \sin nx(b_n - b_{-n}) + i(\dots) \\ \implies \mathfrak{Re}(c_n e^{inx} + c_{-n} e^{i(-n)x}) &= \cos nx(a_n + a_{-n}) - \sin nx(b_n - b_{-n}). \end{aligned}$$

On the other hand, notice that

$$\mathfrak{Re}(c_n e^{inx} + c_{-n} e^{i(-n)x}) = d_n e^{inx} + d_{-n} e^{i(-n)x} \text{ where } d_k := \frac{a_k + a_{-k}}{2} + i\frac{b_k - b_{-k}}{2}$$

(and so $d_0 = \mathfrak{Re}(c_0)$). It follows that

$$\left|f(x) - \sum_{k=-n}^{n} d_k e^{ikx}\right| < \epsilon,$$

and in particular, $\sum_{k=-n}^{n} d_k e^{ikx}$ is a real trigonometric polynomial since $d_k = \overline{d_{-k}}$. □

Proof of Example 5.23. Recall that we need to show the density of span of $\{e^{ikx}/\sqrt{2\pi}\}$ in $L^2(-\pi, \pi)$. Take $f \in L^2(-\pi, \pi)$ and let $\epsilon > 0$ be given. Let $g \in C_0([-\pi, \pi])$ (continuous functions with zero at endpoints) be such that $\|f - g\|_{L^2} < \epsilon/2$ (recall $C([a, b])$ is dense in L^2 by Lemma 5.24). Then let $h \in \operatorname{span}\{e^{ikx}\}$ such that $\|h - g\|_{\sup} < \epsilon/(2\sqrt{2\pi})$ so that $\|h - g\|_{L^2} < \epsilon/2$. Then $\|h - f\|_{L^2} < \epsilon$, as desired.

Remark

This exercise showed that every $f \in L^2$ is the limit of its Fourier series. Cf. Theorem 5.21.2. In some sense it's a "natural" convergence of Fourier series. This holds not only for L^2 but also for any L^p with $p < \infty$. On the other hand, Fourier series might not converge pointwise even for continuous f: see PS3.

□

Separability of Hilbert Spaces

We will soon show that a Hilbert space H has orthonormal Schauder bases if and only if it is separable, and if yes then $H \equiv L^2$ (isometrically isomorphic).

Lemma 6.4: Gram-Schmidt Orthogonalization

Let $E := \{e_i\}_{i \geqslant 1}$ be a countable linearly independent set of an inner product space V. Then there exists $\tilde{E} := \{\tilde{e}_i\}_{i \geqslant 1}$ that is linearly independent and *orthonormal* with $\operatorname{span}(E) = \operatorname{span}(\tilde{E})$.

Proof. We use the norm induced by the inner product. Let $\tilde{e}_i := e_i / \|e_i\|$, and for all $n \geqslant 1$ define

$$\tilde{e}_{n+1} := \frac{e'_{n+1}}{\|e'_{n+1}\|} \text{ where } e'_{n+1} = e_{n+1} - \sum_{i=1}^{n} (e_{n+1}, \tilde{e}_i) \tilde{e}_i.$$

It follows that $(e'_{n+1}, \tilde{e}_i) = 0$ for all $i \leqslant n$. The rest are familiar and thus omitted. ☐

Future reference: Proposition 6.5, Example 6.8, Problem 7.4

Proposition 6.5

Let H be an infinite-dimensional Hilbert space. Then H is separable if and only if it has an (orthonormal Schauder) basis.

Future reference: Theorem 6.6, Proposition 9.16, Corollary 13.2

Proof. $\Longleftarrow$ is obvious by the very definition of separability (Lemma 2.10) and that the basis is countable. $\Longrightarrow$: let $E' = \{e'_i\}$ be a countable dense set. First we make E' linearly independent by deleting any e'_n that is a linear combination of $\{e'_1, e'_2, \ldots, e'_{n-1}\}$. Define this new set to be E''. Clearly $\operatorname{span}(E'') = \operatorname{span}(E')$. By Gram-Schmidt we obtain E, a linearly independent and orthonormal set with $\operatorname{span}(E) = H$. Hence E is a Schauder basis by Theorem 5.21 parts 1/5. ☐

Lecture 16: 9/23 Separable Hilbert Spaces; Closest Points

Example 5.23: Even more on Example 5.23!

Take $p \in (1, \infty)$ and define

$$S_n f := \sum_{i=-n}^{n} (f, e_i) e_i$$

(the partial **Fourier expansion** until order n), where $(f, e_i) := \int_{-\pi}^{\pi} f \overline{e}_i$ (this is well-defined because $f \in L^p$ and also $e_i \in L^\infty$ bounded). Then

$$S_n f \to f \text{ as } n \to \infty \text{ in } L^p \text{ if and only if } \|S_n f\|_{L^p} \leqslant c_p \|f\|_{L^p},$$

for some c_p depending only on p, i.e., S_n are uniformly bounded.

Proof. The case $p = 2$ is immediate by Bessel's Inequality:

$$\Big| \sum_{i=-n}^{n} (f, e_i) e_i \Big|^2 \leqslant \left[\sum_{i=-n}^{n} \Big| (f, e_i) e_i \Big| \right]^2 \leqslant \sum_{i=-n}^{n} |(f, e_i)|^2 \overset{(B)}{\leqslant} |f|^2$$

so $\|S_n f\|_{L^2}^2 \leqslant \|f\|_{L^2}^2 \implies \|S_n f\|_{L^2} \leqslant \|f\|_{L^2}$.

For other cases, notice that $\implies$ follows directly from PUB. The convergence of $S_n f$ implies $\sup \|S_n f\| < \infty$.

For $\impliedby$, let $\epsilon > 0$ be given. Let $g \in \operatorname{span}\{e^{ikx}\}$ be such that $\|f - g\|_{L^p} < \epsilon$. Then $g = S_n g$ for sufficiently large n (as long as $n \geqslant \deg g$). For large n's, we have

$$\begin{aligned} \|f - S_n f\| &\leqslant \|f - S_n g\| + \|S_n g - S_n f\| \\ &= \|f - S_n g\| + \|S_n (f - g)\| \\ &< \epsilon + \epsilon \cdot \text{some constant since } S_n \text{ is bounded.} \end{aligned}$$

This shows the convergence. □

Theorem 6.6

Any separable, infinite-dimensional Hilbert space H (over $\mathbb{K}$) $\equiv \ell^2(\mathbb{K})$ (isometrically isomorphic).

Proof. Let $\{e_i\}_{i \geqslant 1}$ be an orthonormal Schauder basis of H (by Proposition 6.5). Define $T : H \to \ell^2$ by

$$T(x) := ((x, e_1), (x, e_2), \dots) \text{ the sequence of Fourier coefficients.}$$

Notice that T is onto by $\impliedby$ of Lemma 5.18: any such square-summable sequence in ℓ^2 leads to the convergence of $\sum_{i \geqslant 1} (x, e_i) e_i$ in H. It is also injective by Theorem 5.21.4 (suppose $T(x) = T(y)$ then $T(x) - T(y) = 0$ and the result follows from the Theorem). Hence T^{-1} exists and is defined as

$$T^{-1}(\alpha_1, \alpha_2, \dots) = \sum_{i \geqslant 1} \alpha_i e_i.$$

On the other hand, notice the following (by Parseval) and we are done:

$$\|x\|_H \overset{P}{=} \left(\sum_{i\geqslant 1}|(x,e_i)|^2\right)^{1/2} = \|T(x)\|_{\ell^2}.$$

□

Example 6.7

Most Hilbert spaces (*in applications*) are separable (e.g. $L^2(\Omega)$), but non-separable Hilbert spaces do exist. The following is an example:

Let Γ be an uncountable set and define $H := \{f : \Gamma \to \mathbb{R} : f(\gamma) = 0 \text{ a.e.}\}$. Define $\|f\|_H^2 := \sum_{\gamma\in\Gamma}|f(\gamma)|^2$. Then H is a Hilbert space just like L^2 (if f_n is a sequence then the set of points at which all f_n's are nonzero is countable). Notice that

$$\|\mathbb{1}_\gamma - \mathbb{1}_{\gamma'}\| = 2 \text{ for all } \gamma \neq \gamma'.$$

Similar to how ℓ^∞ are not separable, if we were to approximate H with a dense set, we have to approximate all these indicator functions. The uncountability of Γ implies the uncountability of this dense set, and by Lemma 2.10 (definition) this means H is not separable.

Example 6.8

Theorem 5.21.5 breaks down if we drop the assumption on the orthonormality of $\{e_i\}$; see following.

Let $\{e_i\}$ be an orthonormal Schauder basis in H and let

$$f_n := \sum_{i=1}^{n} \frac{e_i}{i}.$$

Notice that $\overline{\operatorname{span}\{f_n\}} = H$ because $\{f_n\}$ can be obtained by applying Gram-Schmidt on $\{e_i\}$. However, $\{f_n\}$ is *not* a Schauder basis (not orthonormal in fact, either).

Proof. Suppose $\{f_n\}$ is a Schauder basis and let $x := \sum_{i\geqslant 1} e_i/i$ (this is indeed in H, as guaranteed by Parseval's Identity). By assumption there exists $\{\alpha_i\} \subset \mathbb{K}$ such that $x = \sum_{n\geqslant 1}\alpha_n f_n$. Then

$$\frac{1}{k} = \frac{1}{k}(e_k, e_k) = \left(\sum_{i\geqslant 1}\frac{e_i}{i}, e_k\right) = (x, e_k) = \left(\sum_{n\geqslant 1}\alpha_n f_n, e_k\right).$$

By Lemma 5.9, we can take the sums out and get

$$\frac{1}{k} = \sum_{n\geqslant 1}\alpha_n(f_n, e_k) = \frac{1}{k}\sum_{n\geqslant k}^{\infty}\alpha_n \text{ since } (f_n, e_k) = \begin{cases} 1/k & n \geqslant k \\ 0 & n < k \end{cases}.$$

This means $\sum_{n\geqslant k}\alpha_n = 1$ for *all* k's, in which case we obtain $\sum 0 = 1$, clearly a contradiction. □

Closest Points and Approximation in H

Lemma 6.9: Closest points

Let H be Hilbert and $A \subset H$ nonempty, convex, and closed. Let $x \in H \setminus A$. We define the distance

$$\operatorname{dist}(x, d) := \inf_{a \in A} \|x - a\|.$$

The lemma states that there exists a unique $\hat{a} \in A$ that attains this infimum. Moreover, for any other $a \in A$,

$$\mathfrak{Re}(x - \hat{a}, a - \hat{a}) \leqslant 0.$$

(Relate this with having obtuse angle in $\mathbb{R}^n$ where the cosine of the angle gives a negative inner product.)

Future reference: Proposition 6.11, distance functionals

Proof. First notice that $d > 0$. (Otherwise there exists $\{a_n\}$ converging to x. By the closedness of A this means $x \in A$, contradiction. Now let $a_n \in A$ be such that $\|x - a_n\|^2 \leqslant d^2 + 1/n$. Parallelogram Law gives

$$2(\|x - a_n\|^2 + \|x + a_m\|^2) = \|2x - a_n - a_m\|^2 + \|a_m + a_n\|^2$$

whereas

$$\|x - a_n\|^2 \leqslant d^2 + \frac{1}{n} \text{ and } \|x - a_m\|^2 \leqslant d^2 + \frac{1}{m}.$$

This gives

$$\|a_m - a_n\|^2 \leqslant 4d^2 + \frac{2}{m} + \frac{2}{n} - 4\|x - (a_m + a_n)/2\|^2.$$

By the convexity of A, $(a_m + a_n)/2 \in A$ and so $\|x - (a_m + a_n)/2\|^2 \geqslant d^2$. Hence

$$\|a_m - a_n\|^2 \leqslant \frac{2}{m} + \frac{2}{n} \to 0 \text{ as } m, n \to \infty.$$

Thus $\{a_n\}$ is Cauchy and there exists some $\hat{a} \in H$ to which $\{a_n\}$ converges. Since A is closed, $\hat{a} \in A$. It remains to show $\|x - \hat{a}\| = d$: we have $d \leqslant \|x - \hat{a}\|$ by definition and $\|x - \hat{a}\|^2 \leqslant d^2$. Hence the claim.

Uniqueness and the "obtuse angle" parts to be continued next lecture. $\square$

Lecture 17: 9/25 Orthogonal Complement; Orthogonal Projection

Proof: continuing on Lemma 6.9. For the "obtuse angle" part: let $a \in A$. Convexness implies $(1-t)\hat{a} + ta \in A$ for all $t \in [0,1]$. Then

$$\begin{aligned}\|x - \hat{a}\|^2 = d^2 \leqslant \|x - (1-t)\hat{a} - ta\|^2 \\ &= \|(x-\hat{a}) + t(\hat{a} - a)\|^2 \\ &= \|x - \hat{a}\|^2 + 2t\mathfrak{Re}(x - \hat{a}, \hat{a} - a) + t^2\|\hat{a} - a\|^2 \\ &= \|x - \hat{a}\|^2 - 2t\mathfrak{Re}(x - \hat{a}, a - \hat{a}) + t^2\|\hat{a} - a\|^2.\end{aligned}$$

This means $t^2\|a^2 - a\|^2 \geqslant 2t\mathfrak{Re}(x - \hat{a}, a - \hat{a})$. Since $t \gg t^2$ when sufficiently small, the only way to ensure this inequity holds for all $t \in [0,1]$ is if $\mathfrak{Re}(x - \hat{a}, a - \hat{a}) \leqslant 0$.

Now for uniqueness of $\hat{a}$: suppose there is some other $a' \in A$ with $\|x - a'\| = d$. From above we know $\mathfrak{Re}(x - \hat{a}, a' - \hat{a}) \leqslant 0$ and $\mathfrak{Re}(x - a', \hat{a} - a') = \mathfrak{Re}(a' - x, a' - \hat{a}) \leqslant 0$. Adding the two gives

$$\mathfrak{Re}(x - \hat{a}, a' - \hat{a}) + \mathfrak{Re}(a' - x, a' - \hat{a}) = \mathfrak{Re}(\underbrace{a' - \hat{a}, a' - \hat{a}}_{=\|\cdot\|^2}) \leqslant 0 \implies a' = \hat{a}.$$

□

Remark

If we put $v := x - \hat{a}$ then

$$\mathfrak{Re}(a, v) + d^2 \leqslant \mathfrak{Re}(x, v) \text{ for all } a \in A$$

since

$$\begin{aligned}\mathfrak{Re}(x, v) &= \mathfrak{Re}(x - \hat{a} + \hat{a}, v) \\ &= \mathfrak{Re}(v, v) + \mathfrak{Re}(\hat{a}, v) \\ &= d^2 + \mathfrak{Re}(\hat{a}, x - \hat{a}) \\ &\geqslant d^2 + \mathfrak{Re}(a, x - \hat{a}).\end{aligned}$$

(The last step follows from that $\mathfrak{Re}(a - \hat{a}, x - \hat{a}) \leqslant 0$ which implies $\mathfrak{Re}(a, x - \hat{a}) \leqslant \mathfrak{Re}(\hat{a}, x - \hat{a})$.)

Lemma 6.10

Let H be Hilbert and $X \subset H$. Then the **orthogonal complement** of X, defined as

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

is a closed subspace.

Future reference: Lemma 6.12, Riesz Representation Theorem

Proof. Subspace: $(u + \alpha v, x) = (u, x) + \alpha(v, x) = 0$ whenever $u, v \in X^\perp$.
Closed: if $\{u_n\} \subset X^\perp$ is such that $u_n \to u \in H$, then (by Lemma 5.9)

$$(u, x) = \left(\lim_{n\to\infty} u_n, x\right) = \lim_{n\to\infty} (u_n, x) = 0.$$

□

Proposition 6.11

Suppose U is a closed subspace of Hilbert space H. Then for all $x \in H$ there exists a unique $u \in U$ and $v \in U^\perp$ such that $x = u + v$ where u happens to the closed point to x in U. The map $P_u(x) := u$ is called the **orthogonal projection** onto U; it satisfies $P_u^2 = P_u$ (P_u is **idempotent**) and $\|P_u\| = 1$.
Future reference: Lemma 6.12, Lemma 7.1, Proposition 6.11 (continued), Riesz Representation Theorem, distance functionals, Corollary 13.2

Proof. A closed subspace is always convex since $\lambda u + (1 - \lambda)v$ is a linearly combination of u and v. Therefore by Lemma 6.9 there exists a unique $u \in U$ closest to x. The proof roughly consists of three parts.

(1) Existence with $v := x - u \in U^\perp$: take any $w \in U$. It follows that if H is a real Hilbert space $u \pm w \in U$; otherwise (if $\mathbb{K} = \mathbb{C}$), in addition to the previous ones, we have $u \pm iw \in U$. By Lemma 6.9 we have $\mathfrak{Re}(x - u, u \pm w - u) = \mathfrak{Re}(x - u, \pm w) \leqslant 0 \implies \mathfrak{Re}(x - u, w) = 0$. If H is real then we are done since $\mathfrak{Re}(x - u, w) = (x - u, w) = 0 \implies v \in U^\perp$. If $\mathbb{K} = \mathbb{C}$, notice that $\mathfrak{Re}(x - u, \pm iw) = 0$ gives $\mathfrak{I}(x - u, w) = 0$ so both the real and imaginary parts of $(x - u, w) = 0$. Again we conclude that $v = x - u \in U^\perp$.

(2) Uniqueness: suppose $x = u_1 + v_1 = u_2 + v_2$. Then $u_1 = u_2 = v_2 - v_1$ which implies

$$\|u_1 - u_2\|^2 = (u_1 - u_2, u_1 - u_2) = (u_1 - u_2, v_2 - v_1) = (u_1, v_2) + \cdots = 0.$$

Hence $u_1 = u_2$ and likewise $v_1 = v_2$.

(3) Properties of P_u: $P_u^2(x) = P_u(u + v) = u = P_u(x)$. For the norm:

$$\|x\|^2 = \|u + v\|^2 = \|u\|^2 + 2\mathfrak{Re}(u, v) + \|v\|^2 \geqslant \|u\|^2 = \|P_u(x)\|^2$$

where letting $x \in U$ gives $\|P_u\| = 1$.

□

Lemma 6.12

If $X \subset H$ then $X \subset (X^\perp)^\perp$. Equality can be achieved if and only if X is a closed subspace.
Example where $\subset$ is proper: let $X :=$ unit disk on xy-plane in $\mathbb{R}^3$. Then $X^\perp$ is the z-axis and $(X^\perp)^\perp$ is the entire xy-plane.

Proof.

(1) Inclusion: for any $x \in X$ we have $(x, z) = 0$ for all $z \in X^\perp$. This means precisely that $x \in (X^\perp)^\perp$.

(2) $\Longleftarrow$: let $z \in (X^\perp)^\perp$. We want to show $z \in X$. By assumption X is closed, so by Proposition 6.11 there exist some $x \in X$ and $x' \in X^\perp$ such that $z = x + x'$. Then since $z \in (X^\perp)^\perp$ and $x' \in X^\perp$ we have $(z, x') = 0$. It follows that

$$0 = (z, x') = (x + x', x') = (x, x') + (x', x')\|x'\|^2 \implies z = x \in X.$$

(3) $\Longrightarrow$: suppose $X = (X^\perp)^\perp$ then X itself being an orthogonal complement is closed by Lemma 6.10.

□

Lemma 6.13

Let $\{e_i\}$ be an orthonormal sequence in H and $U := \overline{\operatorname{span}\{e_i\}}$. Then

$$P_U(x) = \sum_{i \geqslant 1} (x, e_i) e_i.$$

If $\{e_i\}$ is a Schauder basis, then $\overline{\operatorname{span}\{e_i\}} = H$ so $P_U(x) = P_H(x) = x$. Cf. Lemma 5.21 parts 2 & 5.
Future reference: Example 7.3, Problem 7.4, Riesz Representation Theorem

Proof. Since $\{e_i\}$ is an orthonormal Schauder basis of U, we may pick any $y \in H$ with $y := \sum_{i \geqslant 1} \alpha_i e_i$ and $\sum_{i \geqslant 1} |\alpha_i|^2 < \infty$. Then (Parseval again)

$$\begin{aligned}
\|x - y\|^2 = \|x\|^2 &= \sum_{i \geqslant 1} (x, \alpha_i e_i) - \sum_{i \geqslant 1} (\alpha_i e_i, x) + \sum_{i \geqslant 1} |\alpha_i|^2 \\
&= \|x\|^2 - \sum_{i \geqslant 1} |(x, e_i)|^2 + \sum_{i \geqslant 1} \left(|(x, e_i)|^2 - \overline{\alpha_i}(x, e_i) - \alpha_i \overline{(x, e_i)} + |\alpha_i|^2 \right) \\
&= \|x\|^2 - \sum_{i \geqslant 1} |(x, e_i)|^2 + \sum_{i \geqslant 1} |(x, e_i) - \alpha_i|^2.
\end{aligned}$$

Notice that the first two terms are α_i-independent. Therefore one needs to ensure $\alpha_i = (x, e_i)$ to minimize $\|x - y\|$, i.e., the distance. Hence $P_u(x) = y$ if and only if $\alpha_i = (x, e_i)$ by Proposition 6.11 (P_u projects x to the closest point to x in U). □

Lecture 18: 9/28 More on Orthogonal Complement & Projection

Lemma 7.1

Let $T \in B(H)$ be such that $T^2 = T$. TFAE:

(1) T is an orthogonal projection onto some $U \subset H$.

(2) $\ker(T) = (\operatorname{im}(T))^{\perp}$.

(3) $\|T\| = 1$.

Future reference: Theorem 7.2

Proof. (1) $\implies$ (2): let $T := P_U$ for some closed subspace U. If $P_u(x) = 0$ then $P_u(u) + P_u(v) = 0$ for some $u \in U$ and $v \in U^{\perp}$. So $P_u(x) = P_u(u) = u$ and $x \in \ker(T)$ if and only if $x \in U^{\perp} = (\operatorname{im}(T))^{\perp}$.

(2) $\implies$ (1): note that $\ker(\mathrm{id} - T) = \operatorname{im}(T)$. $\subset$ because if $x - T(x) = 0$ then $x = T(x) \in \operatorname{im}(T)$. $\supset$ because if $x = T(y)$ then since T is idempotent $x = T^2(y) = T(x) \implies x - T(x) = 0$.

Notice that $\ker(\mathrm{id} - T)$ is a closed subspace, so $\operatorname{im}(T)$ is also closed. Let $U := \operatorname{im}(T)$. Then $U^{\perp} = \ker(T)$ by assumption. Therefore for every $x \in H$ is of form $u + v$ for some $u \in U$ and $v \in U^{\perp}$ by Proposition 6.11. It follows that $T(x) = T(u) + T(v) = T(u)$. Since $u \in \operatorname{im}(T)$, for some $y \in H$ we have $u = T(y)$. Since T is idempotent, this leads to $T(x) = T(u) = T^2(y) = T(y) = u$.

(1) $\implies$ (3): true by Proposition 6.11.

(3) $\implies$ (2): for all $x \in H$ we have $T(x - T(x)) = T(x) - T^2(x) = 0$. In particular, for all $x \in (\ker(T))^{\perp}$,

$$0 = (\underbrace{x - T(x)}_{\in \ker(T)}, \overbrace{x}^{\in(\ker T)^{\perp}}) = \|x\|^2 - (T(x), x).$$

By Cauchy-Schwarz (first inequality) and assumption $\|T\| = 1$ (second inequality) we have

$$\|x\|^2 = (T(x), x) \leqslant \|T(x)\|\|x\| \leqslant \|x\|^2 \implies \|T(x)\| = \|x\|.$$

On the other hand,

$$\|x - T(x)\|^2 = \|x\|^2 - 2\mathfrak{Re}\underbrace{(T(x), x)}_{=\|x\|^2} + \underbrace{\|T(x)\|^2}_{=\|x\|^2} = 0 \implies x = T(x) \text{ for all } x \in (\ker(T))^{\perp}.$$

This means $(\ker(T))^{\perp} \subset \operatorname{im}(T)$. Now it remains to show $\operatorname{im}(T) \subset (\ker(T))^{\perp}$; then $(\ker(T))^{\perp} = \operatorname{im}(T)$, and taking orthogonal complements again gives (2) as the closedness of $\ker(T)$ guarantees $((\ker(T))^{\perp})^{\perp} = \ker(T)$.

Let $y \in \operatorname{im}(T)$ and define $U := \ker(T)$ a closed subspace. By Proposition 6.11 there exists a unique $u \in U$ and $v \in U^{\perp}$ such that $y = u + v$. Furthermore, by what we've shown above, $v \in U^{\perp} = (\ker(T))^{\perp} \implies v \in \operatorname{im}(T)$. Hence $v = T(h)$ for some $h \in H$. Again, since T is idempotent $v = T(h) = T^2(h) = T(v)$; likewise for some $z \in H$ we have $y = T(z) = T^2(z) = T(y)$. Then

$$u = y - v = T(y - v) = T^2(y - v) = T(u) \implies u = 0 \text{ since } T(u) = 0 \in \ker(T).$$

Therefore y only has a v component, i.e., $y \in (\ker(T))^{\perp}$. This shows $\operatorname{im}(T) \subset (\ker(T))^{\perp}$, and we are done. ☐

Remark: more on Proposition 6.11

If we drop the condition "$v \in U^\perp$", it's in fact true that we can find another $V \subset X$ such that for all $x \in X$ there exists a unique decomposition $x = u + v$ for some $u \in U$ and $v \in V$. Consider $X := \mathbb{R}^2$ with U being the x-axis and V any line not horizontal.

Theorem 7.2

Let X be Banach. Then

(1) U and V are complements if and only if there exists some idempotent $T \in B(X)$ (with $\operatorname{im}(T) = U$ and $\ker(T) = V$). In addition, if X is Hilbert then $V = U^\perp \iff \|T\| = 1$, cf. Lemma 7.1. Hint: $\ker(\operatorname{id} - T) = \operatorname{im}(T)$.

(2) c_0 is not complemented in ℓ^∞. (See Philips 1940; Whitney 1966.)

(3) $X := \ell^p$ with $p > 2$ contains uncomplemented subspaces. (Murray 1937.)

(4) If it so happens that every closed subspace is complemented, then $X \cong H$ for some Hilbert space. (Lindenstrauss & Tzafrini, 1971.)

(5) If an infinite-dimensional $U \subset X := \ell^p$ is complemented then $U \cong X$ for all $p \in [1, \infty]$. (Lindenstrauss 1967.)

Lecture 19: 9/30 Dual Spaces; Riesz Representation Theorem

Example 7.3

Let $H = \ell^2$ and $e_i := e^{(2i)}$. Let $x := (x_1, x_2.,,,)$ be given. Then $(0, x_2, 0, x_4, \dots)$ is the best approximation of x in $U := \overline{\operatorname{span}\{e_i\}}$ by Lemma 6.13:

$$P_U(x) = \sum_{i \geqslant 1} (x, e_i) e_i = (0, x_2, 0, x_4, \dots).$$

Problem 7.4

Find the best approximation of $f(x) := |x|$ by a 3$^{\text{rd}}$ degree polynomial in $L^2(-1,1)$ norm, i.e., $H = L^2(-1,1)$ and $U := \overline{U} := \operatorname{span}\{1, x, x^2, x^3\}$.

Solution

We apply Gram-Schmidt with respect to U to get $e_1 = 1/\sqrt{2}$. Then

$$e_2' = x - \left(x, \frac{1}{\sqrt{2}}\right)\frac{1}{\sqrt{2}} = x - \frac{1}{2}\int_{-1}^{1} x \, \mathrm{d}x.$$

Since $\|e_2'\| = \left(\int_{-1}^{1} |x|^2 \, \mathrm{d}x\right)^{1/2} = \sqrt{\frac{2}{3}}$ we get $e_2 = \sqrt{3/2} \cdot x$. Similarly,

$$\begin{aligned} e_3' &= x^2 - \left(x^2, \frac{1}{\sqrt{2}}\right)\frac{1}{\sqrt{2}} - \left(x^2, \sqrt{\frac{3}{2}}x\right)\sqrt{\frac{3}{2}}x \\ &= x^2 - \frac{1}{2}\int_{-1}^{1} x^2 \, \mathrm{d}x - \frac{3x}{2}\int_{-1}^{1} x^3 \, \mathrm{d}x \\ &= x^2 - \frac{1}{3}. \end{aligned}$$

Then since $\|e_3'\| = \sqrt{8/45}$ we get $e_3 = \sqrt{5/8}(3x^2 - 1)$. [These are the scalar multiples of the first 3 **Legende polynomials** $P_n(x) := \frac{1}{2^n n!}\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n$ where the function is called **Rodrigues' formulae**.]

By Lemma 6.13 the best approximation of $f_3 \in L^2(-1,1)$ in U is

$$f_3(x) = \sum_{i=1}^{3} (f, e_i) e_i = \cdots = \frac{15x^2 + 3}{16}$$

(the x^3 term vanishes) with $\|f - f_3\|_{L^2(-1,1)} = \cdots = \sqrt{3}/4$.

Dual Spaces

Recall from Definition 2.12 that $X^* := B(X, \mathbb{K})$ is the **dual space** of X.

Example 7.5

Let $p \in [0,1]$ and define $\delta_p : X := C([0,1]) \to \mathbb{R}$ with $\delta_p(f) = f(p)$. Then $\delta_p \in X^*$ with $\|\delta_p\|_{X^*} = 1$.

Proof. Linearity is clear as $\delta_p(f_1 + \alpha f_2) = f_1(p) + \alpha f_2(p) = \delta_p(f_1) + \alpha\delta_p(f_2)$. Also,

$$|\delta_p(f)| = |f(p)| \leqslant \|f\|_X \implies \|\delta_p\|_{X^*} \leqslant 1,$$

while letting $f_0 \in X$ such that $f_0(p) = \max\limits_{x\in[0,1]}|f_0|$ gives $|\delta_p(f_0)| = f_0(p) = \|f_0\|$, so $\|\delta_p\|_{X^*} \geqslant 1$. □

Lemma 7.6

If X be normed and Y Banach then $B(X,Y)$ is Banach. In particular *every dual space is Banach.*

Future reference: Every reflexive space is Banach

Lemma 7.7

Let H be Hilbert. Given $y \in H$ and $f_y : H \to \mathbb{K}$ such that $f_y(x) := (x, y)$ for all $x \in H$, then

$$f_y \in H^* \text{ with } \|f_y\|_{H^*} = \|y\|_H.$$

Proof. Letting f_y act on x, we have (by Cauchy-Schwarz) $|f_y(x)| = |(x,y)| \leqslant \|x\|\|y\|$ so $\|f_y\|_{H^*} \leqslant \|y\|$. On the other hand, letting f_y act on y itself gives $|f_y(y)| = |(y,y)| = \|y\|^2 \implies \|f_y\|_{H^*} \geqslant \|y\|$. □

Theorem 7.8: Riesz Representation Theorem

Let H be Hilbert. Given $f \in H^*$ there exists exactly one $y \in H$ such that $f(x) = (x,y)$ for all $x \in H$, and $\|f\|_{H^*} = \|y\|$ (which we have already proved above).

In other words, the **Riesz map** $R : H \to H^*$ defined by $y \mapsto f$ (or f_y as in lemma above) is a bijective *conjugate* linear isometry (since y appears as the second argument in the inner product).

Future reference: Foliation, Theorem 8.8, Example 11.12

Proof. Define $K := \ker(f)$ a closed subspace of H.

First claim: $\dim(K^\perp) = 1$. Let $u, v \in K^\perp$ and we'll show u and v are linearly dependent. Notice that

$$f\,\underbrace{\left(f(u)v - f(v)u\right)}_{\text{a linear combination}} = f(u)f(v) - f(v)f(u) = 0$$

so on one hand the linear combination suggests $f(u)v - f(v)u \in K^\perp$ while $f(\cdot) = 0 \implies$ it's in K too. (Any $u \in K \cap K^\perp$ satisfies $\|u\|^2 = (u,u) = 0$ by treating one $u \in K$ and the other $\in K^\perp$.) Therefore $f(u)v = f(v)u$ and u, v are indeed linearly dependent.

Now back to the main proof: let $z \in K^\perp$ be such that $\|z\| = 1$. Notice that, for every $x \in H$ there exists a unique decomposition $x = w + (x,z)z$ for some $w \in K$:

(1) $w \in K$ and $(x,z)z \in K^\perp$, both in closed subspaces by Lemma 6.10.

(2) The existence and uniqueness of this decomposition is guaranteed by Proposition 6.11.

(3) Since $\dim(K^\perp) = 1$ we have $K^\perp = \operatorname{span}\{z\} = \overline{\operatorname{span}\{z\}}$ (since it's closed). By Lemma 6.13 this means the projection of x onto $K^\perp$ is precisely $(x, z)z$.

Then,

$$\begin{aligned} f(x) &= f(w + (x,z)z) \\ &= f(w) + (x,z)f(z) \\ &= 0 + (x, z\overline{f(z)}) \end{aligned}$$

and the proof follows by taking $y := z\overline{f(z)}$. □

Lecture 20: 10/2 Hyperplanes; Hölder's in ℓ^p Spaces

Definition 7.9

A **hyperplane** U in X is a codimension 1 subspace of X, i.e., a maximal proper subspace of X. (If Z is a subspace such that $U \subset Z \subset X$ then either $Z = U$ or $Z = X$.)

Future reference: Lemma 11.2

Lemma 7.10

$U \subset X$ is a hyperplane if and only if:

(1) $U \neq X$ and

(2) For every $x \in X \setminus U$, $\operatorname{span}(U \cup \{x\}) = X$.

Future reference: Lemma 11.2

Proof. $\implies$: let $Z := \operatorname{span}(U, \{x\})$ for some $x \in X \setminus U$. Clearly $U \subset Z \subset X$. Clearly the first inclusion is proper, so $Z = X$.

$\impliedby$: let Z be a subspace of X with $U \subset Z \subset X$. We need to show that if $Z \neq U$ then $Z = X$. In this case, there exists some $x \in Z \setminus U$. By assumption $\operatorname{span}(U \cup \{x\}) = X$. Notice that since $x \in Z$ we have $X \subset Z$. Hence $X = Z$. □

Remark

Riesz Representation Theorem suggests that $\ker(f)$ is a hyperplane for all $f \in H^*$ since $(\ker(f))^\perp$ is 1-dimensional (see proof of the theorem). More generally, if H is real, then for any $y \in H$, $\{f = f(y)\}$ is an affine plane. The entire space can be filled by the hyperplane with its affine planes. This is called **foliation** by translated hyperplanes

$$\{x : f(x) = c\}, c \in \mathbb{R}.$$

Corollary 7.11

Let $A \subset H$ be closed and convex and H real. Pick $x \in H \setminus A$ and let $\hat{a} \in A$ be the best approximation of x in A, i.e., $\|x - \hat{a}\| = d = \inf_{a \in A} \|x - a\|$. Now define $f := f_{x-\hat{a}}$, i.e., $f(y) = (y, x - \hat{a})$ for all $y \in H$. Then

$$\left\{ f := f(x) - \frac{d^2}{2} \right\}$$

separates A and x.

Future reference: Separating convex sets, Theorem 10.18

Proof. Note that $f(\hat{a}) = (\hat{a}, x - \hat{a}) = (x, x - \hat{a}) - \|x - \hat{a}\|^2 = f(x) - d^2$. Therefore the affine plane containing $\hat{a}$ is given by $\{f = f(\hat{a}) = f(x) - d^2\}$ whereas the affine plane going through x is simply $\{f = f(x)\}$. Hence taking the plane "in the middle" (or $\{f = f(x) - kd^2\}$ for any $k \in (0,1)$) separates A and x, i.e., all other $a \in A$ lie on the other side than x. □

Lemma 7.12: Hölder's inequality in ℓ^p spaces

Similar to Hölder's inequality in L^p spaces:

$$\sum_{i\geqslant 1}|x_iy_i| \leqslant \|x\|_{\ell^p}\|y\|_{\ell^q} \text{ whenever } \frac{1}{p}+\frac{1}{q}=1.$$

Proof. If p or $q=\infty$ then this is trivial. Otherwise, *Young's inequality* (for numbers) gives

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

(This is given by the convexity of the exponential function; see PS1 also)

$$ab = e^{\log ab} = \exp((\ln a^p)/p + (\ln b^q)/q) \leqslant \frac{1}{p}e^{\ln a^p} + \frac{1}{q}e^{\ln b^q} = \frac{a^p}{p}+\frac{b^q}{q}.$$

Back to the proof:

$$\begin{aligned}
\sum_{i\geqslant 1}\frac{|x_iy_i|}{\|x\|_{\ell^p}\|y\|_{\ell^q}} &= \sum_{i\geqslant 1}\frac{|x_i|}{\|x\|_{\ell^p}}\frac{|y_i|}{\|y\|_{\ell^q}} \\
&\leqslant \sum_{i\geqslant 1}\left(\frac{1}{p}\frac{|x_i|^p}{\|x\|^p_{\ell^p}}+\frac{1}{q}\frac{|y_i|^q}{\|y\|^q_{\ell^q}}\right) && \text{(Young)} \\
&= \frac{1}{p}\frac{\|x\|^p_{\ell^p}}{\|x\|^p_{\ell^p}}+\frac{1}{q}\frac{\|y\|^q_{\ell^q}}{\|y\|^q_{\ell^q}} = 1.
\end{aligned}$$

□

Theorem 7.13

If $q\in(1,\infty)$ then $(\ell^q)^*\equiv\ell^p$ (isometrically isomorphic, where $p^{-1}+q^{-1}=1$) via

$$x\mapsto L_x \text{ such that } L_x(y) := \sum_{i\geqslant 1}x_iy_i$$

where $x\in\ell^p, L_x\in(\ell^q)^*$, and $y\in\ell^q$.

Future reference: another form of norm on ℓ^p

Proof. First thing: L_x is linear. By Hölder,

$$|L_x(y)| \leqslant \|x\|_{\ell^p}\|y\|_{\ell^q} \implies \|L_x\|_{(\ell^q)^*} \leqslant \|x\|_{\ell^p}$$

so indeed $L_x\in(\ell^p)^*$. It remains to show that it is an isometry (bijective with operator norm 1). Let

$$y_i := \begin{cases} |x_i|^p/x_i & x_i\neq 0 \\ 0 & x_i = 0\end{cases}.$$

Clearly $y \in \ell^q$ as $\|y\|_{\ell^q}^q = \sum_{i\geqslant 1}|y_i|^q = \sum_{i\geqslant 1}|x_i|^{pq-q} = \sum_{i\geqslant 1}|x_i|^p = \|x\|_{\ell^p}^p < \infty$. To show $\|L_x\|_{(\ell^q)^*} = \|x\|_{\ell^p}$ (isometry), it remains to show the $\geqslant$ direction (or the existence of =). This is proven by

$$\begin{aligned}
L_x(y) &= \sum_{i\geqslant 1} x_i y_i \\
&= \sum_{i\geqslant 1} x_i(|x_i|^p/x_i) \\
&= \sum_{i\geqslant 1}|x_i|^p && \text{(even when } x_i = 0\text{)} \\
&= \|x\|_{\ell^p} \cdot \|x\|_{\ell^p}^{p-1} \\
&= \|x\|_{\ell^p} \cdot \|y\|_{\ell^p}^{q(p-1)/p} && \text{(since } \|y\|_{\ell^q}^q = \|x\|_{\ell^p}^p \text{ as shown above)} \\
&= \|x\|_{\ell^p} \cdot \|y\|_{\ell^q}.
\end{aligned}$$

Clearly $x \mapsto L_x$ is injective. We'll show it's surjective next lecture. □

Lecture 21: 10/5 L^p, ℓ^p Dual Space Isometries; Spectral Theory

Proof of $(\ell^q)^* \equiv \ell^p$: surjective part. Recall from last lecture that we need to show $x \mapsto L_x$ is surjective. Let $f \in (\ell^q)^*$. Define x to be such that $x_i := f(e^{(i)})$. We first show that $x \in \ell^p$. Let y be defined by

$$y_i := \begin{cases} |x_i|^p / x_i & x_i \neq 0 \\ 0 & x_i = 0 \end{cases}$$

and $y^{(n)} := (y_1, \ldots, y_n, 0, \ldots)$. It follows that

$$f(y^{(n)}) = \sum_{i=1}^{n} y_i f(e^{(i)}) = \sum_{i=1}^{n} |x_i|^p.$$

So $\sum_{i=1}^{n} |x_i|^p \leqslant |f(y^{(n)})| \leqslant \|f\|_{(\ell^q)^*} \|y^{(n)}\|_{\ell^q}$. Notice that the last term $\|y^{(n)}\|_{\ell^q}$ can be re-written as

$$\|y^{(n)}\|_{\ell^q} = \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q} = \left(\sum_{i=1}^{n} |x_i|^{pq-q}\right)^{1/q} = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1-1/p}.$$

Therefore,

$$\sum_{i=1}^{n} |x_i|^p \leqslant \|f\|_{(\ell^q)^*} \left(\sum_{i=1}^{n} |x_i|^p\right)^{1-1/p} \implies \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \leqslant \|f\|_{(\ell^q)^*}$$

for all n. This means precisely that the infinite sum (i.e., $\|x\|_{\ell^p}$) is no greater than $\|f\|_{(\ell^q)^*}$, so $x \in \ell^p$. (*Notice that we approached this sub-proof using finite sums. The problem with $\sum_{i=1}^{\infty}$ is that $\sum_{i \geqslant 1} |x_i|^p$ on both sides may be infinite as showing $x \in \ell^p$ is our very goal, and subtraction in infinity makes no sense. Be very careful when making assumptions. What is not assumed can't be taken for granted.*)
Hence, for each y, we have

$$f(y) = f(\sum_{i \geqslant 1} a_i e^{(i)}) = \sum_{i \geqslant 1} y_i f(e^{(i)}) = \sum_{i \geqslant 1} y_i x_i \implies f = L_x.$$

□

Remark

Theorem 7.13 gives an equivalent norm on ℓ^p, cf. Lemma 2.14:

$$\|x\|_{\ell^p} = \|L_x\|_{(\ell^q)^*} = \sup_{\|y\|_{\ell^q}=1} \left|\sum_{i \geqslant 1} x_i y_i\right|.$$

Theorem 8.1

$(\ell^1)^* \equiv \ell^\infty$ but $(\ell^\infty)^* \not\equiv \ell^1$. Instead, $(c_0)^* \equiv \ell^1$.

Future reference: remark of Example 10.7, Corollary 10.12, Reflexivity, Example 11.8, Schur's Theorem

Theorem 8.2

If $q \in [1, \infty)$ then $(L^q(\Omega))^* \equiv L^p(\Omega)$ via

$$f \mapsto \mathcal{L}_f : \mathcal{L}_f(g) := \int_\Omega fg$$

where $f \in L^p(\Omega)$, $\mathcal{L}_f \in (L^q(\Omega))^*$, and $g \in L^q(\Omega)$.

Spectral Theory (for Bounded Operators)

Recall in finite dimensional cases (in linear algebra), λ is an eigenvalue of $T : X \to X$ if there exists nonzero x such that $T(x) = \lambda x$. In particular, in finite-dimensional cases, $(T - \lambda I)x = 0$ if and only if $T - \lambda I$ is not invertible. In these cases, the study of **spectral theory** is the same as the "study of eigenvalues".

Now we will consider complex Banach spaces. Some definitions first.

Definition 8.3

Let X be a complex Banach space and T a linear operator.

(1) **Resolvent** — $S(T) := \{\lambda \in \mathbb{C} : T - \lambda I$ is invertible (having an bounded inverse)$\}$. Also see a more detailed definition of resolvent.

(2) **Spectrum** — $\sigma(T) := \mathbb{C} \setminus S(T) = \{\lambda \in \mathbb{C} : T - \lambda I$ is not invertible$\}$.

(3) **Point spectrum** — $\sigma_{\mathrm{p}}(T) := \{\lambda \in \mathbb{C} : (T - \lambda I)x = 0$ for some $x \neq 0\}$.

(4) **Eigenspace** — if $\lambda \in \sigma_{\mathrm{p}}(T)$ then $E_\lambda := \ker(T - \lambda I)$ is the eigenspace of λ.

(5) **Multiplicity of $\boldsymbol{\lambda}$** — defined as $\dim E_\lambda$.

Lecture 22: 10/7 Closedness of $\sigma(T)$ for $T \in B(X)$

Remark

Two immediate results following the definitions:

(1) $\sigma_{\mathrm{p}}(T) \subset \sigma(T)$: if $(T - \lambda I)x = 0$ for any $x \neq 0$ then $T - \lambda I$ is not bijrctive and hence not invertible.

(2) $|\lambda| \leqslant \|T\|$ for all $\lambda \in \sigma_{\mathrm{p}}(T)$ (actually true for the entire spectrum):

$$|\lambda|\|x\| = \|\lambda x\| = \|T(x)\| \leqslant \|T\|\|x\|.$$

Lemma 8.4

If $\{\lambda_i\}_{i\geqslant 1}$ are distinct eigenvalues of T, then any set of $\{e_i\}_{i\geqslant 1}$ of corresponding eigenvectors is independent.
Future reference: Proposition 10.3, Lemma 12.15

Proof. We will prove by induction. The base case is trivial. Now suppose $\{e_1, \ldots, e_k\}$ are linearly independent and suppose for contradiction that $\{e_1, \ldots, e_{k+1}\}$ is linear dependent. Then their exists $\{\alpha_i\}_{i\geqslant 1} \neq \{0\}$ such that $\sum_{i=1}^{k+1} \alpha_i e_i = 0$. Applying T gives

$$T(\Sigma) = \sum_{i=1}^{k+1} \alpha_i T(e_i) = \sum_{i=1}^{k+1} \alpha_i \lambda_i e_i = 0,$$

whereas miultiplying $\sum_{i=1}^{k+1} \alpha_i e_i = 0$ by λ_{k+1} gives $\sum_{i=1}^{k+1} \alpha_i \lambda_{k+1} e_i = 0$. Subtracting gives a nontrivial (since λ's are distinct by assumption) combination of $\{e_1, \ldots, e_k\}$ that produces 0, contradicting the linear independence of $\{e_1 \ldots, e_n\}$. □

Proposition 8.5

If $T \in B(X)$ then $\sigma(T)$ is a closed subset of $\{\lambda \in \mathbb{C} : |\lambda| \leqslant \|T\|\}$. To be proven later.
Future reference: Example 8.7.1

Lemma 8.6

If X, Y are Banach and $T \in B(X, Y)$ is invertible, then $T + S$ is invertible for any $S \in B(X, Y)$ with $\|S\| < 1/\|T^{-1}\|$. In particular, the set of invertible operators is open in $B(X, Y)$.
'Future reference: Proposition 8.5, Example 8.7.2

Proof. First thing to prove: if $S \in B(X, Y)$ with $\|S\| \leqslant 1/\|T^{-1}\|$ then $T + S$ is surjective. Let $y \in Y$. Define $J(x) := T^{-1}(y - S(x))$. Then

$$\|J(x) - J(x')\| = \|T^{-1}(S(x - x')\| \leqslant \|T^{-1}\|\|S\|\|x - x'\| < \|x - x'\|.$$

By *Banach Contraction Theorem* there exists a unique $x \in X$ with $x = J(x)$. Therefore for some x we have $x = T^{-1}(y - S(x)) \iff T(x) + S(x) = y \iff (T+S)(x) = y$. Hence surjective.

(*The class did not adopt this proof: the uniqueness of such x guaranteed by Banach Contraction Theorem also suggests the injectivity of $T+S$, and by IMT this means $T+S$ is invertible.*)

Otherwise, we could also show the inverse is bounded, i.e., there exists some $c > 0$ such that $\|(T+S)(x)\| \geqslant c\|x\|$ for all x:

$$\|(T+S)(x) + (-T)(x)\| = \|S(x)\| \leqslant \|(T+S)(x)\| + \|T(x)\| \implies \|(T+S)(x)\| \geqslant \|T(x)\| - \|S(x)\|$$

where

$$\|T(x)\| \geqslant \frac{\|x\|}{\|T^{-1}\|} \text{ and } \|S(x)\| \leqslant \|S\|\|x\|.$$

Therefore

$$\|(T+S)(x)\| \geqslant \|x\| \left(\frac{1}{\|T^{-1}\|} - \|S\| \right).$$

Since $\|S\| < 1/\|T^{-1}\|$, letting $c :=$ the difference proves the (sub)claim.

Notice that this automatically shows that $T+S$ is injective (if $x_1 \neq x_2$ then $\|(T+S)(x_1 - x_2)\|$ is nonzero).

Then, by remark of Definition 3.2,

$$\|(T+S)^{-1}(y)\| \leqslant \frac{1}{c}\|y\| \implies (T+S)^{-1} \in B(Y,X).$$

$\square$

Proof. Proof of Proposition 8.5 First we show that $\sigma(T)$ is a subset of $\overline{B_{\mathbb{C}}(\|T\|)}$. This is equivalent to showing any $|\lambda| > \|T\|$ is in the resolvent, i.e., $\lambda \in S(T)$. Notice that if we define

$$T - \lambda I = \lambda(T/\lambda - I) := \lambda(S - I)$$

then $\|S\| = \|T\|/\lambda < 1$ by assumption and $\|I\|^{-1} = 1$. Hence by Lemma 8.6, $S - I$ is invertible (since $\|S\| < 1/\|I^{-1}\|$) and so is $T - \lambda I$, a scalar multiple of $S - I$. Hence $\lambda \in S(T)$ whenever $|\lambda| > \|T\|$. We will show the closure of $\sigma(T)$ next lecture. $\square$

Lecture 23: 10/9 Spectrum Decomposition; Hilbert Adjoint

Proof of Proposition 8.5, part 2. Now we show $\sigma(T)$ is closed, which is equivalent to showing $S(T)$ is open. Let $\lambda \in S(T)$, i.e., $T - \lambda I$ is invertible. By Lemma 8.6, $T - (\lambda + \delta)I$ is invertible whenever

$$\|\lambda I\| = |\delta| < \frac{1}{\|(T - \lambda I)^{-1}\|},$$

hence the openness of $S(T)$ and the closedness of $\sigma(T)$. ☐

Remark

$\lambda \in \sigma(T)$ if and only if $\operatorname{im}(T - \lambda I) \neq X$ or $\ker(T - \lambda I) \neq \{0\}$.

$\Longleftarrow$ is trivial since $(T - \lambda I)^{-1}$ does not exist if $T - \lambda I$ is not surjective or injective.

For $\Longrightarrow$, we look at the contrapositive. If $\ker(T - \lambda I) = \{0\}$ and $\operatorname{im}(T - \lambda I) = X$, then $T - \lambda I$ is bijective. In addition it's also bounded. Hence $T - \lambda I$ is invertible by IMT.

Future reference: $\sigma(S_R) = \overline{B(1)}$, Example 8.7.2

One can divide $\sigma(T) \setminus \sigma_p(T)$ into two classes:

(1) The **continuous spectrum** where $T - \lambda I$ is "almost surjecrive":
$\sigma_c(T) \coloneqq \{\lambda \in \mathbb{C} : T - \lambda I \text{ is injective and } \operatorname{im}(T - \lambda I) \subset X \text{ is dense but } (T - \lambda I)^{-1} \text{ is not bounded}\}$.
This is possible when $\operatorname{im}(T - \lambda I) \neq X$.

(2) The **residual spectrum** where $T - \lambda I$ is "essentially non-surjective":
$\sigma_r(T) \coloneqq \{\lambda \in \mathbb{C} : T - \lambda I \text{ is injective but } \operatorname{im}(T - \lambda I) \subset X \text{ is not dense}\}$.

Also see a more formal definition of resolvent & comparisons between resolvent and spectrum decomposition.

Name	Injective	Surjective (onto X)?	Inverse Exists?	Inverse Bounded?
Resolvent	Yes	Yes	Yes	Yes
Spectrum	Maybe	Maybe	No (not bijective)	N/A
Point Spectrum	No	Maybe	No	N/A
Cont. Spectrum	Yes	No (but dense)	Yes (im $\to X$)	No
Res. Spectrum	Yes	No	No	N/A

An informal diagram of different sets; see more below.

Remark: More on the definitions

Upon checking online resources, I've found more detailed definitions of the terms above (resolvent and spectrum decomposition), regarding what the image of $T - \lambda I$ should be. The following is from Wikipedia:

Let X be Banach and $T : D \subset X \to X$ be a linear operator. Below are three criteria:

(1) $T - \lambda I$ is injective. Define $(T - \lambda I)^{-1}$:= the inverse of $T - \lambda I$ from $\operatorname{im}(T - \lambda I) \to D$.

(2) $(T - \lambda I)^{-1}$ is bounded.

(3) The domain of $(T - \lambda I)^{-1}$ is a dense subspace of X, i.e., $\operatorname{im}(T - \lambda I)$ is dense in X.

If $\lambda \in \mathbb{C}$ meets all three criteria, it's called a **regular value**. The resolvent is the set of regular values. For others, see the table below.

Set	(1)	(2)	(3)
Resolvent	Yes	Yes	Yes
Point spectrum	No		
Continuous Spectrum	Yes	No	Yes
Residual Spectrum	Yes		No

Example 8.7

(1) Recall from Example 2.15: $S_L : \ell^2 \to \ell^2$ defined by $S_L(x_1, x_2, \dots) = (x_2, x_3, \dots)$ the left shift. Then every $\lambda \in \mathbb{C}$ with $|\lambda| < 1$ is an eigenvalue of S_L.

Proof. $S_L(x) = \lambda x \iff (x_2, x_3, \dots) = (\lambda x_1, \lambda x_2, \dots) \iff x = (x_1, \lambda x_1, \lambda^2 x_1, \dots)$. Therefore, using the definition of ℓ^2 norm we get $x \in \ell^2 \iff \|x_1\|^2 \sum_{i=1}^{\infty} |\lambda|^{2i} < \infty \iff |\lambda| < 1$. ☐

Remark

Here we've shown $B(1) \subset \sigma_{\mathrm{p}}(S_L)$. But recall from Proposition 8.5 that $\sigma(S_L) \supset \sigma_{\mathrm{p}}(S_L)$ is a closed subset of $\{\lambda \in \mathbb{C} : |\lambda| \leqslant \|T\|\}$, whereas Example 2.15 gives $\|S_L\| = 1$. Hence

$$B(1) \subset \sigma_{\mathrm{p}}(S_L) \subset \sigma(S_L) \subset \overline{B(1)} \implies \sigma(S_L) = \overline{B(1)}.$$

Future reference: see Lemma 9.4 for another proof (of S_L and S_R).

(2) What about S_R? It has no eigenvalues because $(x_1, x_2 \dots) = (0, x_1, x_2, \dots) \implies x = 0$. However, the spectrum $\sigma(S_R)$ is still $\overline{B(1)}$. The direction $\sigma(S_R) \subset \overline{B(1)}$ is given by Proposition 8.5 and Exercise 2.15. It remains to show $\supset$. Recall that it suffices to show $\operatorname{im}(S_R - \lambda I) \neq \ell^2$. If $\lambda = 0$ the clearly nothing

gets mapped to $(1,0,0,\dots) \in \ell^2$. Now assume $\lambda \neq 0$. Take some $y \in \ell^2$. Then,

$$\begin{aligned}(y_1, y_2, \dots) = (S_R - \lambda I)x = (-\lambda x_1, x_1 - \lambda x_2, \dots)\\ \iff x_1 &= -\frac{1}{\lambda} y_1 \\ x_2 &= -\frac{1}{\lambda^2}(y_1 + \lambda y_2) \\ &\dots \\ x_k &= -\frac{1}{\lambda^k} \sum_{i=1}^{k} y_i \lambda^{i-1}.\end{aligned}$$

Now we may construct the sequence $y := \left\{ \frac{1}{i} \cdot \frac{|\lambda|^{i-1}}{\lambda^{i-1}} \right\}_{i \geq 1} \in \ell^2$. Then,

$$|x_k| = \frac{1}{|\lambda|^k} \sum_{i=1}^{k} \left(\frac{1}{i} \frac{|\lambda|^{i-1}}{\lambda^{i-1}} \right) \geq \frac{1}{|\lambda|} \sum_{i=1}^{k} \frac{1}{i}$$

and so $(x_1, x_2, \dots)$ diverges. It is therefore not in ℓ^2, i.e., $\mathrm{im}(S_R - \lambda I) \neq \ell^2$, and we are done.

Theorem 8.8

Let H, K be Hilbert and $T \in B(H,K)$. Then there exists a unique $T^* \in B(K,H)$, called the (**Hilbert**) **adjoint**, such that

$$(T(x), y)_K = (x, T^*(y))_H \text{ for all } x \in H \text{ and } y \in K.$$

In addition, T^{**} and $\|T^*\|_{B(K,H)} = \|T\|_{B(H,K)}$.

Future reference: Lemma 9.4, Lemma 9.15, Lemma 9.17

Proof. Take some $y \in K$ and $f \in H^*$ defined by $f(x) := (T(x), y)_K$. (This is indeed well-defined because

$$(T(x), y)_K \leq \|T(x)\| \|y\| \leq \underbrace{\|y\| \|T\|}_{<\infty} \|x\| \implies f \in H^*.)$$

By Riesz Representation Theorem there exists a unique $z \in H$ such that $f(x) = (x, z)_H$.

If we define $T^*(y) := z$ then $(T(x), y) = (x, z) = (x, T^*(y))$. Now it remains to show that $T^* \in B(K,H)$. Clearly it is linear by the very properties of inner products. To see that it is bounded:

$$\begin{aligned}\|T^*(y)\|_H^2 &= (T^*(y), T^*(y)) \\ &= (TT^*(y), y) && \text{(property of adjoint)} \\ &\leq \|TT^*(y)\| \|y\| && \text{(Cauchy-Schwarz)} \\ &\leq \|T\| \|T^*(y)\| \|y\|. && \text{(applying norm of } T\text{)}\end{aligned}$$

Assuming $\|T^*(y)\| \neq 0$, we can divide both sides by it and obtain $\|T^*(y)\| \leq \|T\| \|y\|$ and so $\|T^*\| \leq \|T\|$. Next lecture we will show 1) the uniqueness of T^* and 2) that $T^{**} = T$, which then implies $\|T\| = \|T^{**}\| \leq \|T^*\|$ and so $\|T\| = \|T^*\|$. □

Lecture 24: 10/12 Self-Adjoint; Spectral Theorem; Compact Operators

Proof: continuing on Theorem 8.8. Now we show the uniqueness of T^* and that $T^{**} = T$.

(1) Uniqueness: see below.

(2) $T^{**} = T$: since $(T(x), y)_K = (x, T^*(y))_H$, taking complement gives (for all $y \in K, x \in H$)

$$(T(x), y)_K = (x, T^*(y))_H = \overline{(T^*(y), x)}_H = \overline{(y, T^{**}(x))_K} = (T^{**}(x), y)_K.$$

Hence for all $y \in K$ we have $(T(x) - T^{**}(x), y) = 0$. Then for any $x \in H$, taking $y := T(x) - T^{**}(x)$ gives a zero norm. Hence we conclude $T(x) = T^{**}(x)$ for all x, i.e., $T = T^{**}$. This also proves the uniqueness.

□

Lemma 9.1

Two properties of the Hilbert adjoint:

(1) Conjugate linearity: $(\alpha T + \beta S)^* = \overline{\alpha} T^* + \overline{\beta} S^*$.

(2) $(TS)^* = S^* T^*$.

Future reference: Lemma 9.4

Definition 9.2

T is called **self-adjoint** if $T \in B(H)$ and $T = T^*$. Also see definition of numerical range.

Example 9.3

(1) If $T := S_L$, then $T^* = S_R$:

$$(S_L(x), y) = x_2 y_1 + x_3 y_2 + \cdots = x_1 \cdot 0 + x_2 y_1 + \cdots = (x, S_R(y)).$$

(2) For some Ω, suppose $K \in C\left(\overline{\Omega \times \Omega}\right)$ (continuous in closure), and let $T : L^2(\Omega) \to L^2(\Omega)$ be defined as

$$T(f(x)) := \int_\Omega K(x, y) f(y)\, \mathrm{d}y.$$

Then

$$T^*(g(x)) = \int_\Omega K(y, x) g(y)\, \mathrm{d}y,$$

and, in particular, T is self-adjoint if and only if K is symmetric, i.e., $K(x, y) = K(y, x)$ for all x, y.

Future reference: Example 9.13, Theorem 13.9

Lemma 9.4

Let H be Hilbert and $T \in B(H)$ Then

$$\sigma(T^*) = \{\overline{\lambda} : \lambda \in \sigma(T)\} = \overline{\sigma(T)} \text{ (conjugate, not closure).}$$

Proof. Suppose $\lambda \notin \sigma(T)$, i.e., $\lambda \in S(T)$. Then $T - \lambda I$ is invertible, i.e.,

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

Now take adjoint of everything (recall from Lemma 9.1 that $(TS)^* = S^*T^*$):

$$\left((T - \lambda I)^{-1}\right)^* (T - \lambda I)^* = I^* = I = (T - \lambda I)^* \left((T - \lambda I)^{-1}\right)^*.$$

This by definition means $\left((T - \lambda I)^{-1}\right)^* = ((T - \lambda I)^*)^{-1}$. By Theorem 8.8 we have $\| \left((T - \lambda I)^{-1}\right)^* \| = \|(T - \lambda I)^{-1}\|$ bounded, and so is $((T - \lambda I)^*)^{-1}$. Hence $\lambda \notin \sigma(T^*) \implies (T - \lambda I)^* = T^* - \overline{\lambda} I$ invertible. ☐

Remark

From this lemma we see that, since S_R has adjoint S_L, $\sigma(S_R) = \overline{\sigma(S_L)} = \{|\lambda| \leqslant 1\}$ where the overline denotes the complex conjugate. Cf. Example 8.7.1 & 2.

Lemma 9.5

If $T_1, \dots, T_n \in B(X)$ commute, i.e., $T_i T_j = T_j T_i$ for all i, j, then

$$T_1 T_2 \dots T_n \text{ is invertible} \iff \text{each } T_i \text{ is invertible.}$$

Future reference: Spectral Theorem for Polynomials

Remark

This is not true in general if the T's do not commute:

$$S_L S_R = \mathrm{id} \in B(\ell^2), \text{ invertible,}$$

but S_L is not injective, and S_R is not surjective.

Theorem 9.6: Spectral Theorem for Polynomials

Let $P(z) := \sum_{k=0}^{n} a_k z^k$, a complex polynomial of degree n, and $T \in B(X)$, then

$$\sigma(P(T)) = P(\sigma(T)).$$

Future reference: Corollary 9.7

Proof. First note that, given $\lambda \in \mathbb{C}$, we can decompose $\lambda - P(z) =$ into its roots:

$$\lambda - P(z) = \alpha(\beta_1 - z)(\beta_2 - z \dots)(\beta_n - z) \text{ for } \alpha, \beta_i \in \mathbb{C},$$

i.e., $\lambda = P(\beta_i)$ for all $1 \leqslant i \leqslant n$. Furthermore, this is true only for these β_i's (obvious). Therefore, in terms of operators,

$$\lambda I - P(T) = \alpha(\beta_1 I - T)(\beta_2 I - T) \dots (\beta_n I - T).$$

Notice that the terms on the RHS commute, so by Lemma 9.5 $\lambda I - P(T)$ is invertible if and only if each $\beta_i I - T$ is invertible.

Therefore, if we take $\lambda \in \sigma(P(T))$, i.e., $\lambda I - P(T)$ is not invertible, then some $\beta_i I - T$ is not invertible, i.e., $\beta_i \in \sigma(T)$. Thus $\lambda = P(\beta_i) \in P(\sigma(T))$ and so $\sigma(P(T)) \subset P(\sigma(T))$.

For $\supset$, If $\lambda \notin \sigma(P(T))$, by Lemma 9.5, the invertibility of $\lambda I - P(T)$ implies that each $\beta_i I - T$ is invertible. This means $\beta_i \notin \sigma(T)$ for all $1 \leqslant i \leqslant n$.

(*Notice that, unlike above where $\beta_i \in \sigma(T)$ implies $P(\beta_i) \in P(\sigma(T))$, here $\lambda \notin \sigma(P(T))$ does NOT imply $\lambda = P(\beta_i) \notin \sigma(P(T))$. For example, $-1 \notin \{1\}$ but squaring both sides gives ϵ.*)

Recall from the root decomposition that $\lambda = P(\tilde{z})$ only when $\tilde{z}$ is of form β_i. Notice that

$$\lambda \in P(\sigma(T)) \iff \lambda = P(\tilde{z}) \text{ for some } \tilde{z} \in \sigma(T),$$

so it must be the case that $\beta_i \in \sigma(T)$ for some β_i. But this contradicts $\beta_i \notin \sigma(T)$, our assumption. Hence $\lambda \notin \sigma(P(T)) \implies \lambda \notin P(\sigma(T))$, and we are done. $\square$

Corollary 9.7

If $T \in B(H)$ is self-adjoint, then either $\|T\|$ or $-\|T\| \in \sigma(T)$.

Future reference: Lemma 12.15

Proof. WLOG assume $\|T\| = 1$ (otherwise we may simply scale it). By definition of norm, there exists a sequence $\{x_n\}$ such that $\|x_n\| = 1$ and $\|T(x_n)\| \to 1$. Then

$$\begin{aligned}
\|T^2(x_n) - x_n\|^2 &= (T^2(x_n) - x_n, T^2(x_n) - x_n) \\
&= \|T^2(x_n)\|^2 + \|x_n\|^2 - (x_n, T^2(x_n)) - (T^2(x_n), x_n) \\
&\leqslant \|T\|^4\|x_n\|^2 + 1 - 2(T(x_n), T(x_n)) && \text{(last one b/c of adjoint)} \\
&= 1 + 1 - 2\|T(x_n)\|^2 \to 0.
\end{aligned}$$

Therefore $T^2 - I$ is not invertible: for if it were, we would be able to find c such that $\|x\| \leqslant c\|(T^2 - I)(x)\|$ for all x. So $1 \in \sigma(T^2)$, and by Spectral Theorem for Polynomials, this is the same as $(\sigma(T))^2$. Hence either 1 or $-1 \in \sigma(T)$. □

Theorem 9.8: Spectral Theorem

If $T \in B(X)$ and f holomorphic on an open neighborhood of $\sigma(T)$, then

$$\sigma(f(T)) = f(\sigma(T)).$$

Proof: see Yoshida's book, p.227.

Definition 9.9

Let X, Y be normed. Then $T \in L(X, Y)$ is **compact**, denoted as $T \in K(X, Y)$ if, for all bounded sequence $\{x_n\} \subset X$, $\{T(x_n)\}$ has a convergent subsequence (with limit in Y).
Alternatively, $T \in K(X, Y)$ if $T(B_X(1)) \subset Y$ is precompact (i.e., closure is compact).

Lecture 25: 10/14 Arzelà-Ascoli Theorem; Hilbert-Schmidt Operators

Lemma 9.10

If Y is Banach then $K(X,Y)$ is a closed subspace of $B(X,Y)$.

Future reference: Proposition 9.16, Proposition 9.21

Proof. First claim: $K(X,Y) \subset B(X,Y)$ because, if not, then there exists $\{x_n\}$ such that $\|x_n\| = 1$ with $\|T(x_n)\| \to \infty$, but then $\{T(x_n)\}$ does not have a convergent subsequence.

Clearly $K(X,Y)$ is a vector space; for addition simply consider the overlapping terms of the convergent subsequence, and scalar multiplication is trivial.

Now we show $K(X,Y)$ is closed. Let $\{T_n\} \subset K(X,Y)$ be a sequence of compact operators that converges to $T \in B(X,Y)$ (i.e., $\|T_n\| \to \|T\|$). We want to show $T \in K(X,Y)$ as well.

Let $\{x_n\}$ be such that $\|x_n\| \leqslant M$ (i.e., bounded). It follows that, since each T_n is bounded, we can find a subsequence of $\{x_{n_{1,i}}\}$ such that $\{T_1(x_{n_{1,i}})\}$ is convergent; then we can find a sub-subsequence $\{x_{n_{2,i}}\}$ such that $\{T_2(x_{n_{2,i}})\}$ is convergent, and so on inductively. If we take the diagonal sequence defined by

$$\{y_n\} : y_i := x_{n_{i,i}}$$

we get a sequence that converges for all T_n's. Now we perform the "3ϵ-trick".

$$\begin{aligned}
\|T(y_i) - T(y_j)\| &\leqslant \|T(y_i) - T_n(y_i)\| + \|T_n(y_i) - T_n(y_j)\| + \|T_n(y_j) - T(y_j)\| \\
&\leqslant \|T - T_n\| \underbrace{\|y_i\|}_{\leqslant M} + \|T_n(y_i) - T_n(y_j)\| + \|T_n - T\| \underbrace{\|y_j\|}_{\leqslant M} \\
&< \frac{\epsilon}{3M} \cdot M + \frac{\epsilon}{3} + \frac{\epsilon}{3M} \cdot M = \epsilon
\end{aligned}$$

for n, i, j large enough such that $\|T - T_n\| < \epsilon/3M$ and $\|T_n(y_i) - T_n(y_j)\| < \epsilon/3$ (both of which are possible; the first by convergence of $T_n \to T$ and the second by Cauchy-ness of $\{T_n(y_n)\}$, which is implied by its assumed convergence).

Now since Y is Banach, the Cauchy-ness of $T(y_n)$ implies its convergence. Hence T is compact, and we are done. ☐

Example 9.11

If $T \in B(X,Y)$ is such that $\mathrm{im}(T)$ has finite dimension, then T is compact.

Future reference: Proposition 9.16

Proof. If $T(B_X(1))$ is bounded in a finite dimensional space then it is precompact by Corollary 3.11. ☐

Theorem 9.12: Arzelà-Ascoli Theorem

This theorem gives a characterization of compact sets in spaces of continuous functions. If $A \subset C(\overline{\Omega}; \mathbb{K})$, then

$$A \text{ is precompact} \iff A \text{ is bounded and equicontinuous,}$$

i.e., for all $\epsilon > 0$ there exists $\delta > 0$ such that $|x - y| < \delta \implies |f(x) - f(y)| < \epsilon$ for *all* $f \in A$.

Proof. Similar to the "3ϵ-trick" above, along with the diagonalization argument. □

Example 9.13

Define $T(f(x)) := \int_\Omega K(x,y) f(y)\,\mathrm{d}y$ (from Example 9.3.2) where $K \in C(\overline{\Omega \times \Omega})$. This operator is compact, i.e., $T \in K(L^2, L^2)$.

Proof. By Arzelà-Ascoli above, it suffices to show that $T(B_{L^2(\Omega)}(1))$ is a bounded and equicontinuous subset of $C(\overline{\Omega}, \mathbb{K})$. (This shows something even stronger: not only does it have a convergent subsequence in L^2 but it's also continuous.)

If $f \in B_{L^2(\Omega)}(1)$ then (by Hölder and by $L^p \supset L^q$ for $p < q$)

$$|T(f(x))| \leqslant \left(\int_\Omega |K(x,y)|^2\,\mathrm{d}y \right)^{1/2} \underbrace{\|f\|_{L^2}}_{\leqslant 1} \leqslant |K|_{L^\infty} |\Omega|^{1/2} < \infty, \text{ bounded.}$$

On the other hand,

$$\begin{aligned} |T(f(x)) - T(f(x'))| &\leqslant \left(\int_\Omega |K(x,y) - K(x',y)|^2\,\mathrm{d}y \right)^{1/2} \|f\|_{L^2} \\ &< \frac{\epsilon}{|\Omega|^{1/2}} \cdot |\Omega|^{1/2} = \epsilon \end{aligned}$$

whenever $|x - x'| < \delta$ for some $\delta > 0$, since K is (uniformly) continuous on $\overline{\Omega \times \Omega}$. Hence we've shown uniform continuity. □

Definition 9.14

$T \in B(H)$, where H is separable, is called a **Hilbert-Schmidt operator** if the following holds:

$$\|T\|_{\mathrm{HS}}^2 := \sum_{i \geqslant 1} \|T(e_i)\|^2 < \infty \text{ for some orthonormal Schauder basis } \{e_i\} \text{ of } H.$$

If this holds then we call $\|T\|_{\mathrm{HS}}$ the **Hilbert-Schmidt norm**.

Lemma 9.15

(1) $\|\cdot\|_{\text{HS}}$ is independent of the choice of the basis.

Proof. Let $\{e_i\}$, $\{f_i\}$ be orthonormal Schauder bases of H. Recall Theorem 5.21 that this means Parseval's identity applies; also recall that $T \in B(H)$ admits an adjoint. Thus we have

$$\begin{aligned}\sum_{i\geqslant 1}\|T(e_i)\|^2 = \sum_{i\geqslant 1}\sum_{k\geqslant 1}|(T(e_i),f_k)|^2 &= \sum_{i\geqslant 1}\sum_{k\geqslant 1}|(e_i,T^*(f_k))|^2 \\ &= \sum_{i\geqslant 1}\sum_{k\geqslant 1}|(T^*(f_k))|^2 \\ &= \sum_{k\geqslant 1}\|T^*(f_k)\|^2\end{aligned}$$

so $\sum_{i\geqslant 1}\|T(e_i)\|^2 = \sum_{i\geqslant 1}\|T^*(f_i)\|$. We can apply this approach one more time and deduce

$$\sum_{i\geqslant 1}\|T^*(f_i)\|^2 = \sum_{i,j\geqslant 1}|(T^*(f_i),f_j)|^2 = \sum_{j\geqslant 1}\|T(f_j)\|^2.$$

Therefore $\sum_{i\geqslant 1}\|T(e_i)\|^2 = \sum_{i\geqslant 1}\|T(f_i)\|^2$, as desired. □

(2) $\|T\|_{B(H)} \leqslant \|T\|_{\text{HS}}$, i.e., the Hilbert-Schmidt norm is an upper bound for the operator norms.

Proof. Let $u \in H$ and rewrite it as $\sum_{i\geqslant 1}(u,e_i)e_i$ by Theorem 5.21. Applying Δ-inequality (first $\leqslant$) amd Cauchy-Schwarz (second), we have

$$\|T(u)\| \leqslant \sum_{i\geqslant 1}|(u,e_i)|\|T(e_i)\| \leqslant \Bigg(\underbrace{\sum_{i\geqslant 1}|(u,e_i)^2|}_{\|u\|^2}\Bigg)^{1/2}\Bigg(\underbrace{\sum_{i\geqslant 1}\|T(e_i)\|^2}_{\|T\|^2_{\text{HS}}}\Bigg)^{1/2} = \|T\|^2_{\text{HS}}\|u\|.$$

(Compare this with Proposition 9.16) □

Proposition 9.16

Any Hilbert-Schmidt operator T acting on a separable Hilbert space H is compact.

Proof. Recall that H is seperable if and only if it has a Schauder basis (Proposition 6.5). Let $\{e_i\}$ be a Schauder basis of H. Then, if we take $u \in H$ and expand the Fourier coefficients,

$$T(U) = \sum_{i\geqslant 1}(u,e_i)T(e_i).$$

Let T_n the truncation of the sum be defined as $T_n(u) := \sum_{i=1}^{n}(u,e_i)e_i$. Notice that T_n has finite-dimensional range. By Example 9.11 we know that each T_n is compact. Now we show $\|T_n\| \to \|T\|$:

$$\begin{aligned}
\|(T-T_n)u\| &= \|\sum_{i\geqslant n+1}(u.e_i)T(e_i)\| \\
&\leqslant \sum_{i\geqslant n+1}(u.e_i)\|T(e_i)\| && (\Delta\text{-ineq}) \\
&\leqslant \|u\|\left(\sum_{i\geqslant n+1}\|T(e_i)\|^2\right)^{1/2} && (\text{similar to Lemma 9.15.2})
\end{aligned}$$

whereas the last term tends to 0 as $n \to \infty$ because T is Hilbert-Schmidt and the infinite sum finite. Therefore $\|T_n\| \to \|T\|$. Obviously each $T_n \in K(H)$, so by Lemma 9.10 the closure of $K(H)$ implies $T \in K(H)$, i.e., T is compact. □

Lemma 9.17

(1) If $T \in B(X,Y)$ and $S \in B(Y,Z)$, then if *one of* T, S is compact then $S \circ T \in K(X,Z)$. Proof is obvious since boundedness preserves sequential convergence.

Future reference: Theorem 9.18

(2) If $T \in K(H)$ then $T^* \in K(H)$.

Proof. Recall Theorem 8.8. Since $T \in B(H)$, so is T^*. Thererfore by (1) we get $TT^* \in K(H)$. Hence if $\{x_n\} \subset H$ is bounded by $\|x_n\| \leqslant M$, then TT^* has a convergent subsequence (in H) which we now relabel. Thus

$$\begin{aligned}
\|T^*(x_n-x_m)\|^2 &= (T^*(x_n-x_m), T^*(x_n-x_m)) && (\text{induced by inner product}) \\
&= (TT^*(x_n-x_m), x_n-x_m) && (\text{adjoint}) \\
&\leqslant \underbrace{\|TT^*(x_n-x_m)\|}_{\text{convergent}}\underbrace{\|x_n-x_m\|}_{\leqslant 2M} && (\text{Cauchy-Schwarz}) \\
&\to 0 \text{ as } \min(m,n) \to \infty.
\end{aligned}$$

Hence $\{T^*(x_n)\}$ is Cauchy and it converges in H. This proves the lemma. □

Lecture 26: 10/16 Spectral Theory of Compact Operators

Theorem 9.18

Let X be an infinite-dimensional Banach space. If $T \in K(X)$ then $0 \in \sigma(T)$.

Proof. Suppose $0 \notin \sigma(T)$ then T is invertible. In particular T^{-1} exists and is bounded. Since $T \in K(X)$ and $T^{-1} \in B(X)$, by Lemma 9.17.1 we see that $I = TT^{-1}$ is compact. Then by the very definition of compact operators, $B_X(1)$ is precompact, i.e., $\overline{B_X(1)}$ is compact. But then by Theorem 3.14 X must be finite-dimensional, contradiction. Hence $0 \in \sigma(T)$. □

Lemma 9.19

If $T \in K(X)$ (X normed is sufficient) and $\lambda \neq 0$, then $\dim E_\lambda := \dim\ker(T - \lambda I) < \infty$, i.e., any nonzero eigenvalue of a compact operator has finite multiplicity.
Future reference: Proposition 9.21, Banach-Schmidt Theorem

Proof. Suppose $\dim E_\lambda = \infty$. We want to find $\{w_i\} \subset E_i$ such that $\|w_i\| = 1$ whereas $\|w_i - w_j\| \geqslant 1/2$. (If this is true, then T takes $\{w_i\}$ into a precompact set. However the sequence $\{T(w_i)\}$ cannot have a convergent subsequence since $\|T(w_i) - T(w_j)\| = \lambda\|w_i - w_j\| > \lambda/2$.)
Let $w_1 \in E_\lambda$ be such that $\|w_1\| = 1$. Notice that for all $n \in \mathbb{N}$, $\operatorname{span}\{w_1, \dots, w_n\}$ is a closed, finite-dimensional subspace of E_λ. Therefore, by Riesz's Lemma (not RRT), we can inductively pick $w_{n+1} \in E_\lambda$ such that $\|w_{n+1}\| = 1$ and $\|w_{n+1} - w_i\| \geqslant 1/2$ for all $1 \leqslant i \leqslant n$. Hence we have constructed a sequence as mentioned above which leads to a contraction, and thus $\dim E_\lambda < \infty$. □

Example 9.20

$T \in B(X)$ may or may not have closed range. Consider $T : \ell^2 \to \ell^2$ defined by

$$T(x) := \left(x_1, \frac{x_2}{2}, \frac{x_3}{3}, \dots\right).$$

Obviously T is bounded: $\|T\| \leqslant 1$. Now consider $\{y^{(n)}\}$ where

$$y^{(n)} := T(\underbrace{1, \dots, 1}_{n}, 0, 0, \dots) = \left(1, \frac{1}{2}, \dots, \frac{1}{n}, 0, 0, \dots\right).$$

Then $y^{(n)} \to y := \{\frac{1}{n}\}_{n \geqslant 1} \in \ell^2$ (recall that $\sum_{i \geqslant 1} 1/i^2 = \pi^2/6$ as shown in PS2), but clearly $y \notin \operatorname{im}(T)$ as $(1, 1, \dots) \notin \ell^2$. Hence $\operatorname{im}(T)$ is not closed.

Remark

The same argument also works for $T \in K(X)$. (*Note that T above is actually compact.*) However, $T - I$ has closed image for all $T \in K(X)$, and if this is true, the same holds for $T - \lambda I$ where $\lambda \neq 0$ (since $T/\lambda \in K(X)$ and $T - \lambda I = \lambda(T/\lambda)$). See proposition below.

Proposition 9.21

If $T \in K(X)$ (again, X normed is sufficient), then $\operatorname{im}(T-I)$ is closed.

Proof. Let $\{y_n\} \subset \operatorname{im}(T-I)$ be such that $y_n \to$ some $y \in X$. Define $x_n \in X$ to be such that $(T-I)(x_n) = y_n$. Define

$$d_n := \inf_{w \in \ker(T-I)} \|x_n - w\|.$$

Since $\ker(T-I)$ is finite-dimensional, the infimum is attained, i.e., there exists $z_n \in \ker(T-I)$ such that $\|x_n - z_n\| = d_n$. *(Let $\{w_k\} \subset \ker(T-I)$ be such that $\lim_{k\to\infty} \|x_n - w_k\| = d_n$. Then $\|w_k\|$ is bounded and, by Corollary 3.11, since balls in finite-dimensional spaces are precompact, has some subsequence $w_{k_j} \to w$ for some $w \in \ker(T-I)$. Then $\|x_n - w_k\| \to \|x_n - w\| = d_n$.)*

Our next claim: $\|x_n - z_n\|$ is bounded.

Proof. Suppose not, then there exists $\{n_i\}$ such that $\|x_{n_i} - z_{n_i}\| \to \infty$. Notice that

$$y_{n_i} = (T-I)(x_{n_i}) = (T-I)(x_{n_i} - z_{n_i}) \qquad (\square)$$

since $z_{n_i} \in \ker(T-I)$. Now let

$$w_{n_i} := \frac{x_{n_i} - z_{n_i}}{\|x_{n_i} - z_{n_i}\|} \implies \|w_{n_i}\| = 1.$$

Then,

$$(T-I)(w_{n_i}) = \frac{y_{n_i}}{\|x_{n_i} - z_{n_i}\|}.$$

Since $\{y_n\}$ is convergent, the numerator y_{n_i} is bounded. On the other hand, the denominator $\|x_{n_i} - z_{n_i}\| \to \infty$ by assumption. Hence the entire thing, i.e., $(T-I)(w_{n_i})$, tends to 0.

Also, since $T \in K(X)$ and $\{w_{n_i}\}$ bounded (norm 1), $\{T(w_{n_i})\}$ is precompact. Hence there exists some subsequence that converges to some $q \in X$. To avoid cumbersome notation, we relabel this subsequence as the new $T(w_{n_i})$. Furthermore, not only does $T(w_{n_i}) \to q$, but $w_{n_i} \to q$ as well:

$$\|w_{n_i} - q\| \leqslant \underbrace{\|w_{n_i} - T(w_{n_i})\|}_{\to 0 \text{ by } (\square)} + \underbrace{\|T(w_{n_i}) - q\|}_{\to 0 \text{ by convergence}}.$$

Recall that $T \in K(X) \implies T \in B(X)$ by Lemma 9.10. Hence $T(w_{n_i})$ converges to $T(q)$ and thus $q = T(q)$. Therefore $q \in \ker(T-I)$. It follows that since $\{w_{n_i}\}$ converges to $q \in \ker(T-I)$,

$$\lim_{i\to\infty} \left[\inf_{w \in \ker(T-I)} \|w_{n_i} - w\| \right] = 0.$$

On the other hand, for all $w \in \ker(T-I)$, by construction we have

$$\begin{aligned} \|w_{n_i} - w\| &= \frac{\|x_{n_i} - z_{n_i} - w\|x_{n_i} - z_{n_i}\|\|}{\|x_{n_i} - z_{n_i}\|} \\ &= \frac{1}{d_{n_i}} \|x_{n_i} - (z_{n_i} + w d_{n_i})\|. \qquad \text{(where } d_{n_i} := \|x_{n_i} - z_{n_i}\|^{-1}\text{)} \end{aligned}$$

> However, notice that since both z_{n_i} and d_{n_i} are in $\ker(T-I)$, so is their linear combination $z_{n_i}+wd_{n_i}$. Therefore $\|x_{n_i}-(z_{n_i}+wd_{n_i})\| \geqslant \|x_{n_i}-z_{n_i}\|$ by assumption of infimum, and so $\|w_{n_i}-w\| \geqslant 1$. Contradiction; $\|x_n - z_n\|$ must have been bounded instead. □

Back to the main proof: the claim above, along with the compactness of T, shows that $\{T(x_n-z_n)\}$ is precompact, so there exists some convergent subsequence (likely different from above; but we'll use the same notation) $T(x_{n_i}-z_{n_i}) \to p \in X$. Recall (□); we have the following equality:

$$x_{n_i} - z_{n_i} = T(x_{n_i} - z_{n_i}) - y_{n_i}.$$

Also recall that by assumption $y_{n_i} \to y$. If we define $x := p - y$, we get that

$$\begin{aligned}
(T-I)(x) &= \lim_{i\to\infty}(T-I)(x_{n_i}-z_{n_i}) \\
&= \lim_{i\to\infty}(T-I)(x_{n_i}) - 0 && (\text{since } z_{n_i} \in \ker(T-I)) \\
&= \lim_{i\to\infty} y_{n_i} = y,
\end{aligned}$$

i.e., $y \in \operatorname{im}(T-I)$, as desired. Hence $\operatorname{im}(T-I)$ is closed. □

Lecture 27: 10/19 More on Characterizations of Compact Operators

Theorem 10.1

If $T \in K(X)$, where X is Banach, and $\ker(T - I) = \{0\}$, then $T - I$ is invertible.

Future reference: Corollary 10.2

Proof. It suffices to show $T - I$ is onto (because $\ker(T - I) = \{0\}$ implies injectivity and $T - I \in B(X)$ is bounded just like T is; then by IMT we see $T - I$ is invertible).

Note that

$$\begin{aligned}(T - I)^n &= \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} T^k \\ &= (-1)^n I + \underbrace{\sum_{k=1}^{n} (-1)^{n-k} \binom{n}{k} T^k}_{=:S_n}.\end{aligned}$$

Since $T \in K(X)$ and S_n is simply linear combinations of T^k (all of which are compact by Lemma 9.17.1), it is compact as well. Then we have

$$(T - I)^n = (-1)^{n-1}(\underbrace{(-S_n)}_{\in K(X)} - I)$$

and by Proposition 9.21 we see that $X_n := \operatorname{im}(T - I)^n$ is closed for all n. It is clear that if $y \in X_{n+1}$ then there exists some $x \in X$ such that $(T - I)^{n+1}(y) = x$. Then $(T - I)^n[(T - I)(x)] = y \implies y \in X_n$ as well. Hence we have a nested sequence of sets

$$X \supset X_1 \supset \cdots \supset X_n \supset X_{n+1} \supset \ldots \tag{$\square$}$$

We will now show that it's impossible for all these $\supset$'s to be $\supsetneq$. Suppose, for contradiction, that no two sets are the same. Then, by Riesz's Lemma, for all $n \in \mathbb{N}$ there exists $x_n \in X_n$ such that $\|x_n\| = 1$ and $\|x_n - y\| \geqslant 1/2$ for all $y \in X_{n+1}$. Now pick any $m > n$. We have that

$$\begin{aligned}\|T(x_n) - T(x_m)\| &= \|x_n + \underbrace{(T - I)(x_n)}_{\in X_{n+1}} - \underbrace{x_m}_{\in X_m} - \underbrace{(T - I)(x_m)}_{\in X_{m+1}}\| \\ &= \|\underbrace{x_n}_{\in X_n} - \underbrace{\text{something}}_{\in X_{m+1}}\| \\ &\geqslant \frac{1}{2} \text{ by Riesz's Lemma.}\end{aligned}$$

This means $\{T(x_n)\}$ does not admit a convergent subsequence. Hence our assumption that all $\supset$'s are $\supsetneq$ must have been false, and there exists $n \in \mathbb{N}$ (different from the meaning of previous n) such that $X_{n+1} = X_n$. It remains to notice that $X_{n+1} = X_n \implies X_m = X_n$ for *all* $m \geqslant n$:

$$T - I : X_n \to X_m \text{ is injective + surjective} \implies (T-I)^k \text{ is bijective } \forall k.$$

Now take any $x \in X$ and consider X_n and X_{2n}; we have

$$(T-I)^n(x) = (T-I)^{2n}(y) \text{ for some } y \in X. \tag{Δ}$$

Since $\ker(T-I) = 0$, the same holds for $\ker(T-I)^n$:

$$(T-I)^n(x) = 0 \implies (T-I)\left[(T-I)^{n-1}(x)\right] = 0 \implies (T-I)^{n-1}(x) = 0 \implies \dots$$

Therefore $(T-I)^n$ is injective, and by (Δ) we see $x = (T-I)^n(y)$. Hence $X \subset \mathrm{im}(T-I)^n$ and $X = X_n$. Now recall $(\square)$: $X \supset X_1 \supset X_n$ and $X = X_n$ together imply $X = X_1$, so $\mathrm{im}(T-I) = X$, i.e., $T-I$ is onto. $\square$

Corollary 10.2

If $T \in K(X)$ and $\lambda \in \sigma(T)$, then $\lambda \neq 0 \implies \lambda \in \sigma_{\mathrm{p}}(T)$, i.e., if T is compact, the only nonzero elements of $\sigma(T)$ are eigenvalues.

Future reference: Lemma 12.15

Proof. We simply apply Theorem 10.1 to T/λ. If nonzero $\lambda \notin \sigma_{\mathrm{p}}(T)$, we have $\ker(T - \lambda I) = \{0\}$ and so is $\ker(T/\lambda - I)$. It follows that $T/\lambda - I$ is invertible, and so is $T - \lambda I \implies \lambda \notin \sigma(T)$. This shows the contrapositive. $\square$

Proposition 10.3

If $T \in K(X)$ and $\{\lambda_i\} \subset \sigma_{\mathrm{p}}(T)$ is a sequence of nonzero (distinct) eigenvalues, then $\lambda_i \to 0$, i.e., eigenvalues of $T \in K(X)$ cannot have any accumulation points, except 0.

Future reference: Hilbert-Schmidt Theorem, Theorem 13.9

Proof. Let $e_i \in E_{\lambda_i}$ (an eigenvector) be such that $\|e_i\| = 1$. Also define $F_n := \mathrm{span}\{e_1, \dots, e_n\}$. It follows that

$$(T - \lambda_n I)F_n = F_{n-1} \tag{Δ}$$

as $(T - \lambda_n I)$ gives 0 only when acting on e_n.

Now we know that $\dim F_n = n$ and each F_n is closed in X. We also know that $F_{n+1} \supsetneq F_n$ (recall Lemma 8.4 — different eigenvalues correspond to linearly independent eigenvectors). Once again we can use Riesz's Lemma: for all $n \geqslant 0$, there exists $x_{n+1} \in F_{n+1}$ with $\|x_{n+1}\| = 1$ and $\|x_{n+1} - y\| \geqslant 1/2$ for all $y \in F_n$. (Define $F_0 := \{0\}$.)

Then, for $n > m \geqslant 1$, using the same "adding and subtracting" trick and applying Riesz's lemma, we have

$$\begin{aligned}
\left\|\frac{T(x_n)}{\lambda_n} - \frac{T(x_m)}{\lambda_m}\right\| &= \left\|\frac{(T-\lambda_n I + \lambda_n I)(x_n)}{\lambda_n} - \frac{(T-\lambda_m I + \lambda_m I)(x_m)}{\lambda_m}\right\| \\
&= \left\| \overbrace{x_n}^{\in F_n} + \overbrace{\underbrace{\frac{(T-\lambda_n I)(x_n)}{\lambda_n}}_{\in F_{n-1}} - \underbrace{x_m}_{\in F_m} - \underbrace{\frac{(T-\lambda_m I)(x_m)}{\lambda_m}}_{\in F_{m-1}}}^{\in F_{n-1}} \right\| \\
&\geqslant \frac{1}{2} \\
&\implies \left\{\frac{T(x_n)}{x_n}\right\} \text{ does not admit a convergent subsequence} \\
&\implies \left\{\frac{x_n}{\lambda_n}\right\} \text{ is not bounded for any subsequence (since } T \text{ is compact)} \\
&\implies |\lambda_n| \to 0 \text{ because } \|x_n\| = 1 \text{, fixed by assumption.}
\end{aligned}$$

□

Lecture 28: 10/21 Hahn-Banach Theorem & Applications; Banach Limit

Definition 10.4

Let V be a vector space.

(1) $p : V \to \mathbb{R}$ is **sublinear** if $p(x+y) \leqslant p(x) + p(y)$ and $p(\lambda x) = \lambda p(x)$ for $\lambda \geqslant 0$.

(2) p is called a **seminorm** if $p(x+y) \leqslant p(x) + p(y)$ and $p(\lambda x) = |\lambda| p(x)$ for all $\lambda \in \mathbb{K}$.

Remark

Consider $V = \mathbb{R}^2$ and $p(x_1, x_2) := x_1$. This is sublineear but it clearly does not define a seminorm.

Theorem 10.5: Hahn-Banach Theorem

(1) (Real version) Let X be *any* real vector space and $U \subset X$ *any* subspace. Let $\varphi : L(U, \mathbb{R})$ be such that $\varphi \leqslant p$ on U for some sublinear function $p : X \to \mathbb{R}$. Then there exists $f \in L(X, \mathbb{R})$ such that

$$f|_U = \varphi \text{ and } f \leqslant p \text{ in } X.$$

If it so happens that p is a seminorm then $|f| \leqslant p$ in X.

(2) (Complex version) let X be a complex vector space and $U \subset X$ a subspace. Let p be a seminorm on X. If there exists $\varphi \in L(U, \mathbb{C})$ and $|\varphi| \leqslant p$ on U, then there exists $f \in L(X, \mathbb{C})$ such that

$$f|_U = \varphi \text{ and } |f| \leqslant p \text{ in } X.$$

For both cases: in particular if X is normed then any $\varphi \in U^*$ can be extended to some $f \in X^*$ with $\|f\|_{X^*} \leqslant \|\varphi\|_{U^*}$ (i.e., take $p(x) := \|\varphi\|$).

Future reference: distance functionals, Banach limit, Proposition 10.15, Theorem 10.18

Proof. Proof is by Zorn's Lemma; see Lax, p.19 & 23. □

Corollary 10.6: Support functionals

Let X be a normed vector space. Given $x \in X$, there exists $f \in X^*$ with $\|f\|_{X^*} = 1$ and $f(x) = \|x\|$.

Proof. Let $\varphi : U := \operatorname{span}\{x\} \to \mathbb{K}$ defined by $\varphi(\alpha x) = \alpha \|x\|$. Note that $\varphi(x) = \|x\|$ and $|\varphi(z)| \leqslant \|z\|$ for all $z \in U$. Hence $\|\varphi\|_{U^*} = 1$. Then Hahn-Banach gives the existence of $f \in X^*$ that we seek.
Future reference: Corollary 10.8, Lemma 11.4.1, Future reference: Lemma 11.13 □

Example 10.7

For all $f \in X^*$ with $\|f\|_{X^*} \leqslant 1$, $\|x\| \leqslant M \iff |f(x)| \leqslant M$.

Proof. $\implies$ is trivial. So is $\impliedby$ by the corollary of support functionals right above. □

Remark

In particular, we can define equivalent norms:

$$\|x\| = \sup_{\substack{f \in X^* \\ \|f\|_{X^*}=1}} |f(x)|, \text{ cf. Theorem 8.1.}$$

Corollary 10.8: X^* separates points

If distinct $x, y \in X$, then there exists $f \in X^*$ such that $f(x) \neq f(y)$. In particular, if $x, y \in X$ and $f(x) = f(y)$ for all $f \in X^*$ then $x = y$.

Future reference: Lemma 11.13

Proof. If $x \neq y$ then by supporting functional corollary gives an $f \in X^*$ such that $\|f\|_{X^*} = 1$ such that

$$0 \neq \|x - y\| = f(x - y) = f(x) - f(y).$$

Hence $x \neq y \implies f(x) \neq f(y)$ for some f. □

Proposition 10.9: Distance functionals

Let X be normed and $Y \subsetneq X$ a closed subspace, and let $x \in X \setminus Y$. Define

$$d := \inf_{y \in Y} \|x - y\|.$$

Then there exists $f \in X^*$ such that $\|f\|_{X^*} = 1$, $f(y) = 0$ for all $y \in Y$, and $f(x) = d$.

(*Compare this with projection theorem on Hilbert spaces. Indeed there we had a distance minimizer back then, but here since X is not necessarily Hilbert the "projection" is not defined. Also see this following example.*)

Future reference: Lemma 10.11, Lemma 10.15.2, Theorem 11.6

Proof. Note that $d > 0$ (due to closedness of Y; similar to Lemma 6.9). Now define $U := \operatorname{span}(Y \cup \{x\})$ and define $\varphi : U \to \mathbb{K}$ by $\varphi(y + \lambda x) := \lambda d$ for $y \in Y$ and $\lambda \in \mathbb{K}$. Notice that

$$|\varphi(y + \lambda x)| = |\lambda| d \leqslant |\lambda| \|x - \underbrace{(-y/\lambda)}_{\in Y}\| = \|\lambda x + y\| \implies \|\varphi\| \leqslant 1.$$

To show $\|\varphi\| \geqslant 1$, let sequence $\{y_n\} \subset Y$ be such that $\|x - y_n\| \leqslant d(1 + 1/n)$. Then

$$\varphi(1 \cdot x - y_n) = d \geqslant \frac{n}{n+1} \|x - y_n\|.$$

This means $\|\varphi\| \geqslant n/(n+1)$, and taking $n \to \infty$ gives $\|\varphi\| = 1$. Now already we have $\varphi(y) = 0$ for all $y \in Y$ and $\varphi(x) = d$. By Hahn-Banach Theorem (the "in particular" part) we can extend φ to $f \in X^*$. □

Example 10.10

If X is Hilbert then $d = \|v\|$ where $x = u + v$ for some $u \in Y$ and $v \in Y^\perp$.

Proof. Since $u \in Y$ and $v \in Y^\perp$, we have

$$\begin{aligned}\|d^2\| &= \inf_{y\in Y} \|u+v-y\|^2 = \inf_{y\in Y} \|v+u-y\|^2 \\ &= \inf_{y\in Y} \left(\|v\|^2 + (\underbrace{v}_{\in Y^\perp}, \underbrace{u-y}_{\in Y}) + (u-y, v) + \|u-y\|^2\right) \\ &= \inf_{y\in Y} \left(\|v\|^2 + \|u-y\|^2\right) = \|v\|^2.\end{aligned}$$

Therefore $d = \|v\|$ and we simply take $y := u$.

(*Follow the italic remarks in the previous proposition: now instead of talking about projections as we did in Hilbert spaces, we talk about linear functionals. The ideas are similar.*) $\square$

Lemma 10.11

If X^* is separable then so is X (assuming it's normed). *The converse is NOT true*: $(\ell^1)^* \equiv \ell^\infty$.

Future reference: Theorem 12.10

Proof. If X^* is separable then so is the unit sphere: $S_{X^*} := \{f \in X^* : \|f\|_{X^*} = 1\}$ (Lemma 2.10). Then there exists $\{f_n\}_{n\geqslant 1} \subset S_{X^*}$ (countable) which is dense in S_{X^*}. Since these f_n's have operator norm 1, there exisets $\{x_n\}_{n\geqslant 1} \subset X$ such that $|f_n(x_n)| \geqslant 1/2$ (use the supremum definition of operator norms). We will show that $X = \overline{\operatorname{span}\{x_n\}} =: M$. (Then the countable set with dense span shows separability.)

Suppose not, i.e., $M \neq X$, then the distance functional proposition gives $f \in X^*$ such that $\|f\|_{X^*} = 1$ and $f(x) = 0$ for all $x \in M$. In particular, $f(x_n) = 0$ for all n and so

$$\frac{1}{2} \leqslant |f_n(x_n)| = |f_n(x_n) - f(x_n)| \leqslant \|f_n - f\| \underbrace{\|x_n\|}_{=1},$$

contradicting the assumed density of $\{f_n\}$ in S_{X^*}. $\square$

Corollary 10.12

$(\ell^\infty)^* \not\equiv \ell^1$ and $(L^\infty)^* \not\equiv L^1$. Recall Theorem 8.1.

Banach Limits in $\ell^\infty(\mathbb{R})$

Let $X := \ell^\infty(\mathbb{R}), \mathfrak{c}(\mathbb{R}) := \{x \in X : x_n \text{ converges}\}$ (notice that $c_0 \subset \mathfrak{c}$), and let $\ell \in (\mathfrak{c}(\mathbb{R}))^*$ be defined by $\ell(x) = \lim_{n\to\infty} x_n$. Notice that we can extend ℓ to X by Hahn-Banach, but we want to extend it in a particular way that preserves the most important properties of the **usual limit**.

Definition 10.13

A **Banach limit** in X is any $\mathfrak{L} \in X^*$ such that, for all $x \in X$,

(1) $\mathfrak{L}(x) \geqslant 0$ if $x_n \geqslant 0$ for all n,

(2) $\mathfrak{L} = \mathfrak{L}(S_L(x))$ (invariant under shifting), and

(3) $\mathfrak{L}(1, 1, \dots) = 1$.

Lemma 10.14

The Banach limit is bounded by limsup and liminf: $\liminf_{n\to\infty} x_n \leqslant \mathfrak{L}(x) \leqslant \limsup_{n\to\infty} x_n$ for all $x \in X = \ell^\infty(\mathbb{R})$. In particular, $\mathfrak{L}(x) = \ell(x)$ on $\mathfrak{c}(\mathbb{R})$ (where the liminf and limsup agree and the Banach limit gets sandwiched). To be proven next lecture.

Future reference: Lemma 10.15.2

Lecture 29: 10/23 Minkowski Functional; Separation of Convex Sets

Proof of Lemma 10.14. Notice that liminf is "the largest lower bound such that only finitely many x_n's are smaller" and likewise for limsup. Since $\mathfrak{L}$ is invariant under (finite times of) shifting, it suffices to show that

$$\inf_n x_n \leqslant \mathfrak{L}(x) \leqslant \sup_n x_n.$$

Let $\epsilon > 0$ be given. There exists some n_0 such that $\sup_n x_n - \epsilon < x_{n_0} \leqslant \sup_n x_n$. Hence $x_{n_0} - x_n + \epsilon > 0$ for all n. Applying $\mathfrak{L}$ gives

$$0 \overset{(1)}{\leqslant} (\{x_{n_0} - x_n + \epsilon\}) = x_{n_0} - \mathfrak{L}(x) + \epsilon$$

(since $\mathfrak{L}$ is linear; $\mathfrak{L}(\{x_{n_0}\}) = x_{n_0}$ and likewise for ϵ by (3)). Therefore $\mathfrak{L}(x) - \epsilon \leqslant x_{n_0} \leqslant \sup_n x_n$. Taking $\epsilon \downarrow 0$ gives $\mathfrak{L}(x) \leqslant \sup_n x_n$. The other case is analogous. □

Proposition 10.15

Banach limits exist. *See Goffman & Pedrick (1983), Prop.5, p.67, or here, for uniqueness of Banach limits (under some conditions). Such sequences (that converge to some Banach limit) are called "almost convergent".*

Proof. Let $p : X \to \mathbb{R}$ be defined by

$$p(x) = \limsup_{n\to\infty} \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

(We need limsup as opposed to simply lim because $x \in \ell^\infty$, and there is no guarantee that $\sum/n$ will converge. Take, for example, the alternating sequence of 1's and 0's.) Note that p is sublinear, and

$$-p(-x) = \liminf_{n\to\infty} \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

In particular, recall $\ell \in (\mathfrak{c}(\mathbb{R}))^*$ and so

$$\ell(x) = \lim_{n\to\infty} \frac{x_1 + \cdots + x_n}{n} \text{ where } x \in \mathfrak{c}(\mathbb{R}),$$

i.e., if $x_n \to x$ then $(x_1 + \cdots + x_n)/n \to x$ (there will only be finitely many terms that's far away from x). By Hahn-Banach , there exists $\mathfrak{L} \in X^*$ such that

$$\begin{cases} \mathfrak{L}(x) \leqslant p(x) \\ \qquad\Downarrow \\ \mathfrak{L}(-x) \leqslant p(-x) \end{cases} \implies -p(-x) \leqslant \mathfrak{L}(x) \leqslant p(x). \tag{Δ}$$

Now it remains to show that $\mathfrak{L}$ is a Banach limit. Criterion (1) is guaranteed by (Δ): if $x_n \geqslant 0$ for all n then $p(x) \geqslant 0$ and $p(-x) \leqslant 0 \implies -p(-x) \geqslant 0$. Hence so is $\mathfrak{L}(x)$. Criterion (3) is trivial in this problem. Now it remains to show (2):

$$\mathfrak{L}(x) - \mathfrak{L}(S_L(x)) = \mathfrak{L}(x - S_L(x)) \leqslant p(x - S_L(x)) = \limsup_{n\to\infty} \frac{x - x_n}{n} \to 0,$$

and likewise $0 \leqslant \mathfrak{L}(x) - \mathfrak{L}(S_L(x))$. Hence $\mathfrak{L}$ is the Banach limit we have been looking for. □

Separating Convex Sets

Now we no longer restrict ourselves to ℓ^∞. Instead, we look at any normed space. Recall Corollary 7.11 where we separated closed convex set A from $x \notin A$ in a real Hilbert space. We shall now generalize this and separate convex sets in any normed space.

Definition 10.16

Let C be an open, convex subset of X, a normed space with $0 \in C$. Then

$$p_c(x) := \inf\{\lambda > 0 : \lambda^{-1}x \in C\}$$

is the **Minkowski functional** of C.
(Think of $\mathbb{R}^2$ and an open unit disk centered at the origin. Given any point not in the open disk, the Minkowski functional is the reciprocal of the infimum of "scaling factors" that transform the point into the disk. In particular, if x is on the boundary of C (though we haven't defined), $p_c(x) = 1$ for obvious reasons.)

Lemma 10.17

Denoting $p := p_c$, we have

(1) p is sublinear.

(2) $0 \leqslant p(x) \leqslant c\|x\|$ for some constant c.

(3) $C = \{x : p(x) < 1\}$. In particular, $\|\cdot\|$ is the Minkowski functional of $B_x(1) := \{x : \|x\| < 1\}$. Compare this with remark of Lemma 1.20.

Future reference: Theorem 10.18

Proof.

(1) $p(\lambda x) = \lambda p(x)$ for $\lambda \geqslant 0$ is obvious. For triangle inequality: let $\alpha > p(x)$ and $\beta > p(y)$ then $x/\alpha, y/\beta \in C$. Hence

$$\frac{x+y}{\alpha+\beta} = \frac{\alpha}{\alpha+\beta}\frac{x}{\alpha} + \frac{\beta}{\alpha+\beta}\frac{y}{\beta} \in C, \qquad \text{(a convex combination)}$$

so $\alpha + \beta \geqslant p(x+y)$ for all $\alpha > p(x), \beta > p(y)$. Letting $a \downarrow p(x)$ and $\beta \downarrow p(y)$ we obtain the Δ ineq.

(2) Since $0 \in C$ and C is open, there exists $\delta > 0$ such that $B(0,\delta) \subset C$. Hence for all $z \in B(0,\delta)$ we have $p(z) \leqslant 1$. Now we start rescaling: for all $x \in X$,

$$p(x) = \frac{2\|x\|}{\delta} \cdot p\underbrace{\left(\frac{\delta}{2}\frac{x}{\|x\|}\right)}_{\in B(0,\delta)} \leqslant \frac{2}{\delta}\|x\| \implies c := \frac{2}{\delta} \text{ finishes the proof.}$$

(3) $\subset$: if $x \in C$ then some ball around x is still in C since C is open. Hence there exists $\lambda < 1$ such that $\lambda^{-1}x \in C$, i.e., $p(x) < 1$.

$\supset$: if $p(x) < 1$ then there exists some $\lambda \in (p(x), 1)$. To see $x \in C$, notice that

$$x = \lambda \cdot \underbrace{x/\lambda}_{\in C} + (1-\lambda) \cdot \underbrace{0}_{\in C} \in C. \qquad \square$$

Theorem 10.18: Functional Separation Theorem

Let X be a real normed space. Let A, B be nonempty, disjoint, convex sets.

(1) If A is open, then there exists $f \in X^*$ and $\gamma \in \mathbb{R}$ such that

$$f(a) < \gamma \leqslant f(b) \text{ for all } a \in A,\ b \in B.$$

(2) If A is compact and B is closed, then there exists $f \in X^*$, $\gamma \in \mathbb{R}$, and $\delta > 0$ such that

$$f(a) \leqslant \gamma - \frac{\delta}{2} < \gamma + \frac{\delta}{2} \leqslant f(b) \text{ for all } a \in A,\ b \in B.$$

Note that if A happens to be a singleton and X Hilbert, then we can recover Corollary 7.11 from it.

Future reference: Corollary 11.1, Functional Separation Theorem (Affine Planes), Characterization of closed convex sets

(For an easy visualization, consider $X := \mathbb{R}^2$. For the first part, let A be an open disk and B any disk disjoint from A. (It might be possible that $\overline{A}$ and B are tangent.) Then we can draw a "line" that does not touch A and touch B at most on its boundary (hence the $<$ and the $\leqslant$). For the second part, now we have two closed sets, and if they are disjoint, their closure must be some distance apart, hence the existence of $\delta > 0$.)

Proof. Let $a_0 \in A$ and $b_0 \in B$, and define $w_0 := b_0 - a_0$. (Then $0 = w_0 + a_0 - b_0$.) Now consider the set

$$C := w_0 + A - B := \{w_0 + a - b : a \in A, b \in B\}.$$

Note that C is open (because A is), convex (because both are), and $0 \in C$. Hence we are back to the setting of Minkowski functionals.

Continued proof on 10/26

Now define $p := p_c$ the Minkowski functional on C. Note that $w_0 \notin C$ (otherwise $a - b = 0$ for some $a \in A$ and $b \in B$, contradicting $A \cap B = \varnothing$). Hence $p_c(w_0) \geqslant 1$.

Now define $U := \operatorname{span}\{w_0\}$ and $\varphi(w_0\alpha) = \alpha$ for $\alpha \in \mathbb{R}$. Immediately we see $\varphi \in L(U, \mathbb{R})$. To apply Hahn-Banach , we need to show $\varphi \leqslant p$ in U:

$$\alpha \geqslant 0 \implies \varphi(\alpha w_0) = \alpha \leqslant \alpha p(w_0) = p(\alpha w_0)$$
$$\alpha < 0 \implies \varphi(\alpha w_0) < 0 \leqslant p(\alpha w_0).$$

By Hahn-Banach there exists $f \in L(X, \mathbb{R})$ such that $f\big|_U = \varphi$ and $f \leqslant p$ in X. By Lemma 10.17.2, Minkowski functionals are bounded above by constant times the norm:

$$\begin{cases} f(x) \leqslant p(x) \leqslant c\|x\| \\ -f(x) \leqslant p(-x) \leqslant c\|x\| \end{cases} \implies f \in X^*.$$

Having found $f \in X^*$, we will verify (1) and (2).

(1) For all $a \in A$, $b \in B$,

$$1 + f(a) - f(b) = f(w_0 + a - b) \leqslant p\underbrace{(w_0 + a - b)}_{\in C} < 1 \text{ by Lemma 10.17.3.}$$

Therefore $f(a) < f(b)$. Now let $\gamma := \inf_{b \in B} f(b)$; we get $f(a) < \gamma \leqslant f(b)$.

For the first sharp inequality: suppose $f(a) = \gamma$ for some $a \in A$. By the openness of A, there exists some $\delta > 0$ such that $B(a, \delta\|w_0\|) \subset A$, and so $a + \delta w_0 \in A$. Applying f gives

$$f(a + \delta w_0) = \underbrace{f(a)}_{=\gamma} + \delta \underbrace{f(w_0)}_{=\varphi(w_0)=1} = \gamma + \delta,$$

clearly a contradiction. Hence the sharp inequality $f(a) < \gamma$.

(2) let $\epsilon := \frac{1}{4}\inf\{\|a - b\| : a \in A, b \in B\}$. Since A is compact, B closed and $A \cap B = \varnothing$, $\epsilon > 0$. (Suppose $\epsilon = 0$ then there exist sequences $\{a_n\}, \{b_n\}$ such that $\|a_b - b_n\| \to 0$. By compactness of A, there exists a subsequence $a_{n_i} \to a$. Then $\|a - b_{n_i}\| \to 0$ and so $b_{n_i} \to a \in B$. Then $A \cap B \neq \varnothing$.)

Now define $A_\epsilon := A + B_X(\epsilon)$ and $B_\epsilon := B_X(\epsilon)$. Since A_ϵ, B_ϵ are both open, by (1) there exists $f \in X^*$ and $\gamma \in \mathbb{R}$ such that $f(a) < \gamma \leqslant f(b)$ for all $a \in A_\epsilon, b \in B_\epsilon$. Now we will define δ carefully so that there are room between $f(a)$ and $f(b)$.

Let $\delta := \epsilon/\|w_0\|$. Then $a + (\delta/2)w_0 \in A_\epsilon$ for each $a \in A$. It follows that

$$f(a) = \underbrace{f(a + \delta w_0/2)}_{<\gamma} - \underbrace{f(\delta w_0/2)}_{\substack{=\delta\varphi(w_0)/2 \\ =\delta/2}} < \gamma - \frac{\delta}{2}$$

and similarly $f(b) \geqslant \gamma + \delta/2$ for all $b \in B$.

□

Lecture 30: 10/26 Characterization of Banach Limits

Corollary 11.1

Theorem 10.18 remains true for complex vector spaces. Simply define $f(a) \mapsto \mathfrak{Re} f(a)$ and $f(b) \mapsto \mathfrak{Re} f(b)$. (Note that a real vector space can always be complexified into a complex vector space in this way.)

Future reference: Characterization of closed convex sets

Proof. We can apply the previous theorem to $X_{\mathbb{R}}$ (X over only $\mathbb{R}$) to find $f \in X_{\mathbb{R}}^*$ such that

$$f(x) := \underbrace{\varphi(x)}_{\mathfrak{Re} f} + i\varphi(ix).$$

□

Banach Limits Revisited

Lemma 10.15: Uniqueness of Banach Limits (Proposition 10.15)

Let $X = \ell^\infty(\mathbb{R})$. Then $x \in X$ is **almost convergent** (i.e., all Banach limits are the same) if and only if $p(x) = -p(-x)$ where $p(x) = \limsup_{n\to\infty} \frac{x_1 + x_2 + \cdots + x_n}{n}$, as defined in Proposition 10.15. (Recall that this is equivalent to saying $\liminf = \limsup$).

Proof. $\Longleftarrow$ is trivial because Lemma 10.14 gives $\liminf = \mathfrak{L} = \limsup$.

For $\Longrightarrow$, suppose $-p(-x) < p(x)$. Notice, from the proof of Proposition 10.15, that

$$p(x) = p(x+y) - \ell(y) \text{ and } -p(-x) = -p(-y-x) - \ell(y) \text{ for all } y \in \mathfrak{c}(\mathbb{R}).$$

(*Since* $\liminf_{n\to\infty}(a_n + b_n) \geq \liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n$ *and equality can be achieved if one of* $\{a_n\}, \{b_n\}$ *converges.*)

Taking supremum and infimum over y (which does not change the values of $-p(-x)$ and $p(x)$) gives

$$\sup_{y\in\mathfrak{c}(\mathbb{R})} (-p(-y-x) - \ell(y)) < \inf_{y\in\mathfrak{c}(\mathbb{R})} (p(x+y) - \ell(y)),$$

i.e., the interval with these two endpoints is nonempty.

Let $\varphi_{\mathfrak{c}} : U := \operatorname{span}(\mathfrak{c}(\mathbb{R}) \cup \{x\}) \to \mathbb{R}$ be defined by $\varphi_{\mathfrak{c}}(y + \lambda x) := \lambda c$ for some c in the interval above. Then $\|\varphi\|_{U^*} = 1$, cf. Proposition 10.9. By Hahn-Banach there exists $\mathfrak{L}_{\mathfrak{c}} \in X^*$ such that $\mathfrak{L}_{\mathfrak{c}}\big|_U = \varphi$ and $\|\mathfrak{L}_{\mathfrak{c}}\|_{X^*} = 1$. Note that, on U, $\mathfrak{L}_{\mathfrak{c}}(y + \lambda x) = \lambda c$, whereas

$$p(y + \lambda x) = p(y) + \lambda p(x) > \lambda c.$$

This gives $-p(-y-\lambda x) \leq \mathfrak{L}_{\mathfrak{c}}(y + \lambda x) \leq p(y + \lambda x)$ and by Proposition 10.15, $\mathfrak{L}_{\mathfrak{c}}$ is a Banach limit.

In particular, $\mathfrak{L}_{\mathfrak{c}}$ is a Banach limit for all c's in the interval, so $\mathfrak{L}_{\mathfrak{c}_1}(x) = \varphi_{\mathfrak{c}_1}(x) = c_1 \neq c_2 = \varphi_{\mathfrak{c}_2}(x) = \mathfrak{L}_{\mathfrak{c}_2}(x)$ whenever $c_1 \neq c_2$. This contradicts the assumption that x is almost convergent, as we do not have a unique Banach limit. □

Remark

This characterization of almost convergent sequences can be generalized to more general p:

$$q(x) := \inf\left[\limsup_{j\to\infty} \frac{1}{k}\sum_{i=1}^{k} x_{n_{i+j}}\right]$$

where the infimum is taken over all finite sets of integers $n_1, \ldots, n_k$. (See Goffman & Pedrick, p.65.)

Hyperplanes Revisited

Lemma 11.2

The following are equivalent:

(1) U is a hyperplane in X (normed).

(2) $U \subsetneq X$ is a subspace such that $\operatorname{span}(U \cup \{x\}) = X$ for all $x \in X \setminus U$.

(3) $U = \ker\varphi$ for some nonzero $\varphi \in L(X, \mathbb{K})$. In addition, U is closed if and only if $\varphi \in X^* := B(X, \mathbb{K})$.

Future reference: Functional Separation Theorem (Affine Planes)

Proof. (1) $\iff$ (2) is already proven in Lemma 7.10. We will show (2) $\iff$ (3) next lecture. □

Lecture 31: 10/28 Functional Separation by Affine Planes; Reflexivity

Proof of Lemma 11.2. (2) $\Longrightarrow$ (3): take $x \in X \setminus U$ and $\varphi \in L(X, \mathbb{K})$ defined as $\varphi(y + \lambda x) := \lambda$ for $y \in U$ and $\lambda \in \mathbb{K}$. This is well–defined because if $y' + \lambda' x = y + \lambda x$ then $(\lambda' - \lambda)x = y - y' \in Y \implies \lambda' = \lambda$. Then $\varphi(z) = 0$ if and only if the x-component is 0, i.e., $z \in U$, and indeed $\varphi(x) = 1$, nonzero.

(3) $\Longrightarrow$ (2): fix some $x \in X \setminus U$. By assumption $\varphi(x) \neq 0$. Given any $z \in X \setminus U$, defining

$$y := z - \frac{\varphi(z)}{\varphi(x)} x$$

gives $\varphi(y) = \varphi(z) - \varphi(z) = 0$, so $y \in U$. Therefore $z = y + (\varphi(z)/\varphi(x))x \in \operatorname{span}(U \cup \{x\}) = U \cup (X \setminus U) = X$.

From (3), U is closed if and only if $\varphi \in X^*$ (bounded):

$\Longleftarrow$: if $\varphi \in X^*$ then $U = \ker \varphi$ is closed (φ preserves sequential continuity; taking the limit of any sequence in $\ker \varphi$ would prove the claim).

$\Longrightarrow$: suppose φ is not continuous. We will show that U is dense in X (so that it cannot be closed: otherwise $U = \overline{U} = X \implies U$ is not a proper subspace and therefore not a subspace). Since φ is not continuous, there exists a sequence $\{x_n\} \subset X$ such that $\|x_n\| = 1$ but $|\varphi(x_n)| \geqslant n$. Given $x \in X$, let

$$y_n := x - \frac{\varphi(x)}{\varphi(x_n)} x_n.$$

Then $\varphi(y_n)$ vanishes by construction and so $y_n \in U$. However,

$$\|x_n - y_n\| = \left\| \frac{\varphi(x)}{\varphi(x_n)} x_n \right\| = \frac{|\varphi(x)|}{|\varphi(x_n)|} \implies 0 \text{ as } |\varphi(x)| < \infty.$$

Hence any $x \in X$ can be approximated by some $\{y_n\} \subset U$, i.e., U is dense in X. The contradiction follows. ☐

Theorem 10.18: Hyperplane / affine plane version of Functional Separation Theorem

Let X be normed. Let nonempty convex $A, B \subset X$ be such that A is compact, B normed, and $A \cap B = \{0\}$. Then there exists a closed hyperplane (as guaranteed by the $\ker f$ part in Lemma 11.2.3; the boundedness is very important as otherwise the hyperplane might be dense in the space (see above)!) that can be translated (by y where $f(y) = \gamma$ so that now the affine plane is $\{x : f(x) = \gamma\}$) to separate A and B.

Corollary 11.3

This theorem provides a characterization of closed convex sets. Let $C \subset X$ be closed and convex. Then

$$C = \{x \in X \mid f(x) \geqslant \inf_{y \in C} f(y) \text{ for all } f \in X^*\}$$

(or $C = \{x \in X \mid \mathfrak{Re} f(x) \geqslant \inf_{y \in C} \mathfrak{Re} f(y)$ for all $f \in X^*\}$ if X is a complex space).

(*Think of a convex set on $X := \mathbb{R}^2$ where the hyperplanes are lines. For a given $f \in X^*$, the set of $x \in X$ satisfying the inequality denotes the region on one side of the hyperplane. Now if we do this for all $f \in X^*$, we get all the "tangent planes" that help enclosing the closed convex set. In particular, we "cannot get concave* *cookies" if we use a long stick to shape it. Here, the affine planes of form* $\inf_{y\in C} f(y)$ *are called the* ***supporting hyperplanes****.)*

Future reference: Lemma 12.7

Proof. $\subset$ is trivial by the very definition of infimum.

For $\supset$, take x_0 in the RHS. If $x_0 \notin C$ then, viewing $A := \{x_0\}$, we can apply the second part of Functional Separation Theorem to find the existence of $f \in X^*$ and $\delta, \gamma > 0$ such that

$$\mathfrak{Re} f(x_0) \leqslant \gamma - \delta < \gamma + \delta \leqslant \mathfrak{Re} f(y) \text{ for all } y \in C.$$

($\mathfrak{Re}$ comes from Corollary 11.1.) But this means $\mathfrak{Re} f(x_0)$ is *strictly* less than $\inf_{y\in C} \mathfrak{Re} f(y)$, contradiction. □

Reflexivity

Recall that $(\ell^q)^* \equiv \ell^p$ holds for $q \in [1, \infty)$. Taking a second dual gives

$$((\ell^q)^*)^* \equiv (\ell^p)^* \equiv \ell^q \text{ for all } q \in (1, \infty),$$

i.e., ℓ^q is **reflexive**. We'll define this formally later.

Define $X^{**} := B(X^*; \mathbb{K})$. We define the **canonical map** $J : X \to X^{**}$ (where X is normed) to be

$$J(x) = x^{**} \text{ where } \underbrace{x^{**}}_{\in X^{**}}(\overbrace{f}^{\in X^*}) := f(x) \text{ for all } f \in X^*.$$

(We start with $x \in X$ and define $x^{**} \in X^{**}$ by letting its argument range over all $f \in X^*$.)

Lemma 11.4

(1) J is an isometry (onto a subspace of X^{**}).

(2) If X is Banach then $\mathrm{im}(J)$ is closed (in X^{**}).

Future reference: Theorem 11.6, Lemma 11.13, Goldstine's Theorem

Proof.

(1) We need to show for all $x \in X$, $x^{**} \in X^{**}$ and $\|x^{**}\|_{X^{**}} = \|x\|_X$.

On one hand,

$$|x^{**}(f)| = |f(x)| \leqslant \underbrace{\|f\|_{X^*}}_{\text{input of } x^{**}} \|x\|_X \implies \|x^{**}\|_{X^{**}} \leqslant \|x\|_X.$$

(This also shows $x^{**} \in X^{**}$). On the other hand, the support functional gives us $f \in X^*$ such that $f(x) = \|x\|$ and $\|f\| = 1$. Then

$$|x^{**}(f)| = |f(x)| = \|f\|_{X^*} \|x\|_X = \|x\|_X \implies \|x^{**}\|_{X^{**}} \geqslant \|x\|_X.$$

(2) Let $\{F_n\} := \{J(x_n)\} \subset \text{im}(J)$ be such that $F_n \to F$ for some $F \in X^{**}$. In particular $\{F_n\}$ is Cauchy. By the first part x_n is also Cauchy in X. Therefore $\|x_n - x\| \to 0$ for some $x \in X$ since X is Banach. Since J is an isometry,

$$\|x_n - x\|_X \to 0 \iff \|J(x_n) - J(x)\|_{X^{**}} = \|F_n - x^{**}\| \to 0 \implies J(x) = F,$$

i.e., $F \in \text{im}(J)$ and so $\text{im}(J)$ is closed.

□

Definition 11.5

X is **reflexive** if $J(X) = X^{**}$, i.e., $X \equiv X^{**}$ via J or, equivalently, J is onto X^{**}. *We look at this particular map, not any other map, to determine if X is reflexive or not. See here for a (complicated) example of a non-reflexive Banach space isometric with its doubel dual. More to be discussed soon.*

Lecture 32: 10/30 Reflexivity of X and X^*; Weak Topologies

Remark

There exists counterexamples where Banach $X \equiv X^{**}$ via some different maps but X is not reflexive. We will later show that uniformly convex Banach spaces, in particular Hilbert spaces as well as L^p and ℓ^p for $p \in (1, \infty)$, are reflexive.

Theorem 11.6

Let X be Banach. Then X is reflexive if and only if X^* is.

Future reference: Example 11.8

Proof.

$\Longrightarrow$: we need to show that, for every $\Phi \in (X^*)^{**}$, there exists $f \in X^*$ such that $f^{**} = \Phi$. Notice that $(X^*)^{**} = B((X^*)^*; \mathbb{K}) = (X^{**})^*$. Hence it's equivalent to showing

$$\Phi\underbrace{(F)}_{\in X^{**}} = \underbrace{F(f)}_{=f^{**}(F)} \text{ for all } F \in X^{**}.$$

Since X is reflexive, $F = x^{**}$ for some $x \in X$. If we let $f(x) := \Phi(x^{**})$, then

$$\overbrace{F(f)}^{=x^{**}(f)=f(x)} = \Phi\underbrace{(F)}_{=x^{**}} \text{ by construction.}$$

It remains to show $f \in X^*$. This is true because $f = \Phi \circ J$ since $f(x) = \Phi(x^{**}) = \Phi(J(x))$, where Φ is continuous and J an isometry.

$\Longleftarrow$: suppose X^* is reflexive but X is not. Since the image of the canonical map is closed, $\operatorname{im}(J) \subsetneq X^{**}$ is a closed subspace. Then by Proposition 10.9, there exists $\Phi \in (X^{**})^*$ such that

$$\Phi\big|_{\operatorname{im}(J)} = 0 \text{ but } \Phi \not\equiv 0,$$

i.e., $\Phi(J(x)) = 0$ for all $x \in X$. But X^* is reflexive, so $\Phi = f^{**}$ for some $f \in X^*$. Then, for all $x \in X$,

$$\begin{aligned} f(x) &= x^{**}(f) && \text{(Canonical map of } x) \\ &= f^{**}(x^{**}) && \text{(Canonical map of } f) \\ &= \Phi(x^{**}) && \text{(Assumption above)} \\ &= 0. \end{aligned}$$

Hence f vanishes for *every* $x \in X$. Then $f \equiv 0$, and so $f^{**} = \Phi = 0$, contradiction. $\square$

Remark

Every reflexive space is Banach (since X^{**} is by Lemma 7.6 and $X \equiv X^{**}$).

Lemma 11.7

If X is reflexive and $X \cong Y$ (isomorphic), then Y is reflexive as well.

Example 11.8

Recall $(c_0)^* \equiv \ell^1$ and $(\ell^1)^* \equiv \ell^\infty$ (Theorem 8.1). From this we see c_0 is not reflexive; in addition, neither ℓ^1 nor ℓ^∞ is. (Immediate by Theorem 11.6.)

Example 11.9

L^1 is not reflexive, which implies L^∞ isn't either, since $(L^1)^* = L^\infty$.

Lemma 11.10

If X is reflexive and $Y \subset X$ a closed subspace, then Y is also reflexive.

Future reference: Theorem 12.10

Weak Topologies

Definition 11.11

Let X be normed.

(1) $x_n \rightharpoonup x$ (x_n's **converge weakly** to x) in X if $f(x_n) \to f(x)$ (in $\mathbb{K}$) for all $f \in X^*$ *(i.e., x_n converges to x in the weak topology (the topology generated by the seminorms $\sup_{i \in [1,n]} \{|f_i(x)|_i, f_i \in X^*\}$)).*

(2) $f_n \xrightharpoonup{*} f$ (f_n's **converge weakly-*** to x) in X^* if $f_n(x) \to f(x)$ for all $x \in X$. *The weak-* copology is the smallest topology on X^* such that each $\Phi \in J(X) \subset X^{**}$ is continuous (i.e., each X^{**} is continuous).*

Example 11.12

In a Hilbert space, every $f \in H^*$ is of the form $x \mapsto (x, y)$ for some $y \in H$ (RRT). Then any orthonormal sequence $\{e_n\}_{n \geqslant 1}$ converges weakly to 0 in H: by Bessel's inequality

$$\sum_{n \geqslant 1} |(e_n, y)|^2 \leqslant \|y\|^2 \implies f_n := (e_n, y) \to f(0) = 0 \text{ for all } y \in H.$$

The difference between weak and strong convergence here is that $e_n \not\to 0$ since $\|e_i - e_j\|^2 = 2$ for any two different e_i, e_j. Hence no strong convergence but only weak one.

Future reference: weak closure

Lecture 33: 11/2 Characterization of Weak Convergence

Lemma 11.13: Properties of weak convergence

(1) If $x_n \to x$ then $x_n \rightharpoonup x$, hence the name *weak* convergence.

(2) In finite-dimensional spaces, $\Longleftarrow$ is true. Otherwise very rarely.

(3) Weak limits are unique.

(4) $x_n \rightharpoonup x$ then $\{x_n\}$ is bounded and $\|x\| \leqslant \liminf_{n\to\infty} \|x_n\|$.

Future reference: Lemma 12.2, Milman-Pettis Theorem

Proof.

(1) For all $f \in X^*$, $|f(x_n) - f(x)| \leqslant \|f\|\|x_n - x\| \to 0$.

(2) Exercise. See the answer here.

(3) Suppose $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$. Then $f(x_n) \to f(x)$ and $f(x_n) \to f(y)$ for *all* $f \in X^*$. Then Corollary 10.8 gives the existence of *some* $f \in X^*$ such that $f(x) \neq f(y)$ whenever $x \neq y$. Hence to let $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$ at the same time, the only possibility is if $x = y$.

(4) For all $x \in X^*$ we have $f(x_n) \to f(x)$. Recall that $f(x_n) = x_n^{**}(f)$ (the canonical map) and $f(x) = x^{**}(f)$. Hence in other words $x_n^{**}(f) \to x^{**}(f)$ in $\mathbb{K}$. In particular, the sequence $\{x_n^{**}(f)\}$ is bounded (in $\mathbb{K}$) for *all* $f \in X^*$. Therefore, by Principle of Uniform Boundedness we have (recall that X^{**} as a dual space is Banach) the boundedness of $\{x_n^{**}\}$ in X^{**}. Since canonical maps are isometries, we have the boundedness of $\{x_n\}$ in X as well.

Now it remains to show $\|x\| \leqslant \liminf_{n\to\infty} \|x_n\|$. Let $f \in X^*$ be the support functional of x, i.e., $\|f\|_{X^*} = 1$ and $f(x) = \|x\|$. Then,

$$\|x\| = |f(x)| = \lim_{n\to\infty} |f(x_n)| = \liminf_{n\to\infty} |f(x_n)| \leqslant \liminf_{n\to\infty} \left[\|f\|_{X^*}\|x_n\|_X\right] = \liminf_{n\to\infty} \|x_n\|$$

(where the $\lim = \liminf$ part holds because $\{x_n\}$ converges). □

Remark

The lemma above holds true for weak-* convergence as well.

Corollary 11.14

If X is reflexive, then $f_n \overset{*}{\rightharpoonup} f$ if and only if $f_n \rightharpoonup f$.

Proof. Any $F \in X^{**}$ is of form $F = x^{**}$ for some $x \in X$ by assumption. Therefore, if $F(f_n) \to F(f)$, since $F(f_n) = x^{**}(f_n) = f_n(x)$ and $F(f) = x^{**}(f) = f(x)$, we equivalently have $f_n(x) \to f(x)$. Hence weak convergent is equivalent to weak-* convergence in reflexive spaces. □

Lemma 12.1

Suppose X is uniformly convex (i.e., for all $\epsilon > 0$ there exists δ such that if $\|x\| = \|y\| = 1$ and $\|x - y\| < \epsilon$ then $\|(x+y)/2\| < 1 - \delta$; examples include Hilbert spaces, L^p and ℓ^p spaces for $p \in (1, \infty)$), then $x_n \to x$ if and only if $x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|$.

Proof. $\implies$ is immediate.

For $\impliedby$, if $x = 0$ then $\|x_n\| \to 0 \implies x_n \to x = 0$. Otherwise, let

$$y_n := \frac{x_n}{\|x_n\|} \text{ and } y := \frac{x}{\|x\|}.$$

Then $y_n \rightharpoonup y$ and $y_n + y \rightharpoonup 2y$. It follows that

$$1 = \|y\| \leqslant \liminf_{n\to\infty} \left\| \frac{y_n + y}{2} \right\| \leqslant \limsup_{n\to\infty} \left\| \frac{y_n + y}{2} \right\| \leqslant 1 \implies \left\| \frac{y_n + y}{2} \right\| \to 1,$$

and by uniform convexity this means $\|y_n - y\| = 0$. So $y_n \to y$ and $x_n \to x$. □

Lemma 12.2

If $x_n \rightharpoonup x$ and $T \in K(X, Y)$, then $T(x_n) \to T(x)$. *Compact operators make weak convergence strong.*

Proof. First note that $T(x_n) \rightharpoonup T(x)$. Take any $f \in Y^*$. Since $f \circ T \in X^*$, $f(T(x_n)) \to f(T(x))$. Now suppose $T(x_n) \nrightarrow T(x)$. Then there exists $\epsilon > 0$ and $\{n_i\}$ (index of subsequence of the original $\{x_n\}$) such that $\|T(x_{n_i}) - T(x)\| > \epsilon$ for each i. Lemma 11.3.4 says $\{x_n\}$ is bounded and, since $T \in K(X, Y)$, $\{T(x_{n_i})\}$ has a convergent subsequence, say $T(x_{n_i'}) \to z \in Y$. By uniqueness of weak limits $z = T(x)$ (since strong convergence implies weak convergence), contradicting the assumption of $\|T(x_{n_i}) - T(x)\| > \epsilon$. □

Remark

Recall PS5.3; if $\{e_n\}$ is orthonormal and $e_n \rightharpoonup 0$, then for any compact T, $T(e_n) \to 0$.

Lemma 12.3

$x_n \rightharpoonup x$ if and only if $\{x_n\}$ is bounded and $f(x_n) \to f(x)$ for all $f \in E$ where $E \subset X^*$ is dense. One can prove this by triangle inequality.

Similarly, $f_n \overset{*}{\rightharpoonup} f$ if and only if $\{f_n\}$ is bounded in X^* and $f_n(x) \to f(x)$ for all $x \in E$, where $E \subset X$ is dense.

Future reference: Banach-Alaoglu Theorem

Example 12.4

Below are some applications of the lemma above:

(1) $x^{(n)} \rightharpoonup x$ in ℓ^p for $p \in (1, \infty)$ if and only if

$$\sum_{k\geqslant 1} x_k^{(n)} y_k \to \sum_{k\geqslant 1} x_k y_k \text{ for all } y \in \ell^q$$

where $1/p + 1/q = 1$ (since $(\ell^p)^* \equiv \ell^q$). Also, $x^{(n)} \rightharpoonup x$ if and only if $\{x^{(n)}\}$ is bounded and $x_k^{(n)} \to x_k$

for all k (by the lemma above, as $\{e^{(i)}\}$ is dense in $\ell^{(q)}$).

Future reference: Schur's Theorem

(2) $f_n \rightharpoonup f$ in $L^p(a,b)$ with $p \in (1,\infty)$ if and only if $\{f_n\}$ is bounded in L^p and

$$\int_a^b f_n g \to \int_a^b fg \text{ for all } g \in C([a,b])$$

where $C([a,b])$ is dense in $L^p(a,b)$ by Lemma 5.24. Alternatively, we can replace $C([a,b])$ by $P([a,b])$ by Corollary 6.1 or $\{e^{ikx} : k \in \mathbb{Z}\}$ by Example 5.23.

(3) Let $\{e_k\}$ be an orthonormal Schauder basis of H, a Hilbert space. Then $x_n \rightharpoonup x$ if and only if $\{x_n\}$ is bounded and $(x_n, e_k) \to (x, e_k)$ for all k.

Theorem 12.5: Schur's Theorem

Let $X := \ell^1$. Then $x^{(n)} \rightharpoonup x$ if and only if $x^{(n)} \to x$. *Weak convergence in ℓ^1 is equivalent to strong convergence.*

Proof. $\Longleftarrow$ is trivial. For $\Longrightarrow$, WLOG assume $x = 0$; we need to show that $\|x^{(n)}\| \to 0$. Suppose this is not true; then on some subsequence (which we relabel to x_n) we have

$$\sum_{k \geq 1} |x_k^{(n)}| \geq \epsilon \text{ for some } \epsilon \text{ for some } \epsilon > 0.$$

Note that $x_k^{(n)} \to 0$ for all k (because $x^{(n)} \rightharpoonup x$ implies component-wise strong convergence by Example 12.4.1). We will finish the proof next lecture. □

Lecture 34: 11/4 Schur, Mazur, Banach-Alaoglu, & Navier-Stokes

Continuing the proof of Schur's Theorem. Let $N_0 = M_0 = 1$. For $i > 0$, choose N_i such that

$$\sum_{k \geqslant N_i} \left| x_k^{(M_{j-i})} \right| < \frac{\epsilon}{6}$$

(which is doable since $x \in \ell^1$ and the late enough terms have small enough sum.) Also, define M_i such that

$$\sum_{k=1}^{N_i} \left| x_k^{(M_i)} \right| < \frac{\epsilon}{6}$$

(which is also doable since $x^{(n)} \to 0$ component-wise; choosing M_i large enough allows us to get small enough sum even from the early terms). Now fix $y \in \ell^\infty$ with

$$y_k := \begin{cases} \left| x_k^{(M_i)} \right| / x_k^{(M_i)} & \text{if } x_k^{(M_i)} \neq 0 \text{ for } k \in (N_{i-1}, N_i] \\ 0 & \text{otherwise} \end{cases}$$

and functional $f(x) := \sum_{k \geqslant 1} x_k y_k$ (i.e., $f = L_y \in (\ell^1)^* \equiv \ell^\infty$ by Theorem 8.1). Now we shall show that even though $f \in (\ell^1)^*$, $x^{(n)}$ does *not* converge weakly to 0. Let the fun begin!

$$\begin{aligned} \left| f\left(x^{(M_i)}\right) \right| &= \left| \sum_{k \geqslant 1} x_k^{(M_i)} y_k \right| \\ &\geqslant \left| \sum_{k=N_i+1}^{N_{i+1}} \underbrace{x_k^{(M_i)} y_k}_{= \left| x_k^{(M_i)} \right|} \right| - \underbrace{\left| \sum_{k=1}^{N_i} x_k^{(M_i)} \underbrace{y_k}_{|\cdot| \leqslant 1} \right|}_{< \epsilon/6} - \underbrace{\left| \sum_{k > N_{i+1}} x_k^{(M_i)} \underbrace{y_k}_{|\cdot| \leqslant 1} \right|}_{< \epsilon/6} \\ &> \underbrace{\sum_{k \geqslant 1} \left| x_k^{(M_i)} \right|}_{\geqslant \epsilon} - \underbrace{\sum_{k=1}^{N_i} \left| x_k^{(M_i)} \right|}_{< \epsilon/6} - \underbrace{\sum_{k > N_i}^{\infty} \left| x_k^{(M_i)} \right|}_{< \epsilon/6} - \frac{\epsilon}{3} > \frac{\epsilon}{3} \text{ for any } i. \end{aligned}$$

Where $\sum_{k \geqslant 1} = \sum_{k=1}^{N_i} + \sum_{k=N_i+1}^{N_{i+1}} + \sum_{k > N_{i+1}}$ and $|a+b+c| + |-b| + |-c| \geqslant |a| \implies |a+b+c| \geqslant |a| - |b| - |c|$

The absolute value of the first term can be dropped and $\sum_{k=N_i+1}^{N_{i+1}} = \sum_{k \geqslant 1} - \sum_{k=1}^{N_i} - \sum_{k > N_{i+1}}$

Therefore, $x^{(n)}$ does not converge weakly to 0 under $f \in (\ell^1)^*$, contradicting the weak convergence of $x^{(n)}$. ☐

Definition 12.6

Let X be normed. We say $A \subset X$ is **weakly closed** if whenever $\{x_n\} \subset A$ and $x_n \rightharpoonup x$ then $x \in A$. Notice that being weakly closed is a *stronger* property than being closed: strongly closed $\implies$ weakly closed, but not the converse. Take, for example, $S_X := \{x : \|x\| = 1\}$ from Example 11.12: $e_n \rightharpoonup 0$ but $0 \notin S_X$.

Lemma 12.7

Let X be normed. If $\mathfrak{C} \subset X$ is closed and convex, then $\mathfrak{C}$ is weakly closed.

Proof. Pick $\{x_n\} \to \mathfrak{C}$ such that $x_n \rightharpoonup x \in X$. Recall from Corollary 11.3 that any closed convex set $\mathfrak{C}$ can be written as

$$\mathfrak{C} = \{x \in X \mid \mathfrak{Re} f(x) \geqslant \inf_{y\in\mathfrak{C}} \mathfrak{Re} f(y) \text{ for all } f \in X^*\}.$$

Therefore $\mathfrak{Re} f(x_n) \geqslant \inf_{y\in\mathcal{C}} \mathfrak{Re} f(y)$ for all $f \in X^*$. By weak convergence, $\mathfrak{Re} f(x_n) \to \mathfrak{Re} f(x)$ and furthermore it is true for all $f \in X^*$. Therefore $x \in \mathfrak{C}$. "*Given any $f \in X^*$, all x_n's stay on the same of that f, and the same thing must also be true for x.*" □

Theorem 12.8: Mazur's Theorem / Lemma

If $x_n \rightharpoonup x$ in a normed space X, then for all n, there exists y_n as a convex combination (i.e., $y_n = \sum_{i=1}^{n} \alpha_i x_i$ where $\sum \alpha_i = 1$ and $\alpha_i \in [0,1]$) such that $y_n \to x$. *This theorem overkills lemma 12.7 above: convex combinations $\{y_n\}$ of x_n's in $\mathfrak{C}$ are still inside $\mathfrak{C}$, and they converge strongly to $x \in \mathfrak{C}$. Hence the strong closure.*

Proof. Let $Y_n := \operatorname{conv}\{x_1, \ldots, x_n\}$ be the **convex hull** of $\{x_1, \ldots, x_n\}$, i.e., the set of all possible convex combinations of them. Also, define $Y := \bigcup_{n\geqslant 1} Y_n = \operatorname{conv}\{x_n\}_{n\geqslant 1}$.
Obviously, $x_n \in Y_n \subset Y$. Because Y is convex [each of the Y_n's is convex]. Therefore by the previous lemma (12.7) $\overline{Y}$ is weakly closed, so $x_n \rightharpoonup x \implies x \in \overline{Y}$. In particular,

$$\operatorname{dist}(x, y_n) \to 0.$$

(Otherwise, there exists ϵ and $\{y_{n_k}\}$ such that $\|x - y\| > \epsilon$ for all $y \in \{Y_{n_k}\}$. Then $\|x - y\| > \epsilon$ for all $y \in Y$ as the Y_{n_k}'s are increasing (e.g. $Y_{n_k - 1}$ would satisfy the $> \epsilon$ relation because of Y_{n_k}). Hence $\|x - y\| > \epsilon$ for all $y \in \overline{Y}$, contradiction.)
Therefore for all n, there exists $y_n \in Y_n$ with $y_n \to x$. □

Remark

For a proof using merely Hahn-Banach and not lemma 11.12 above, see Yoshida, p.120.

Weak Sequential Compactness

Theorem 12.9: Banach-Alaoglu Theorem

Let X be separable. Then any bounded sequence in X^* has a weakly-* convergent subsequence. *In fact, this theorem holds for any normed space. See Rudin, Theorem 3.14.*
Future reference: Weak solutions of the Navier-Stokes equation, Theorem 12.10

Proof. Let $E = \{x_k\} \subset X$ be a dense (countable) subset, and take $\{f_n\} \subset X^*$ bounded, say, by $\|f_n\|_{X^*} \leqslant M$.
We shall now apply the diagonalization argument.
Choose $\{n_{1,i}\}_{i\geqslant 1} \subset \{n\}$ (a subsequence) such that $f_{n_{1,i}}(x_1)$ converges. (We can do this because $[-M, M]$ is compact.)

We choose sub-subsequences inductively such that $f_{n_{m,i}}(x_m)$ converges. (To make sure the diagonalization argument works, we make sure subsequence inherits the first term of its mother sequence, i.e., starting no earlier than the second. For example, the second subsequence is $\{f_{n_{2,j}}\}_{j\geqslant 2}$.)
Then, the diagonal sequence $f_{n_{m,m}}(x_k)$ converges to the same limit as $m \to \infty$ for all $x_k \in E$. Define

$$f(x_k) := \lim_{n\to\infty} f_{n_{m,m}}(x_k) \text{ for all } k.$$

It follows that $|f(x_k)| \leqslant M\|x_k\|$ for all k. Now it remains to extend f from the dense subset to all of X. We don't need Hahn-Banach here: simply define $f(x) = \lim_{k\to\infty} f(y_k)$ where $y_k \to x$ and $\{y_k\} \subset E$. In fact this extension is bounded and is unique (*left as an exercise*).
By Lemma 12.3 (weakly-* version), $f_{n_{m,m}} \overset{*}{\rightharpoonup} f$ ("bounded" and "for all $x \in E$ dense in X"). □

Theorem 12.10: Characterization of Reflexive Spaces

If X is reflexive, then any bounded sequence in X has a weakly convergent subsequence (i.e., $\overline{B_X(0,1)}$ is weakly compact). *This the main reason why we even look at reflexive spaces.* We will prove this next lecture. In fact, $\Longleftarrow$ is also true, as proven by Eberlein (1935). Also see PS6.5 for the case of real Banach space.
Now we give an application of the theorems: using compactness method on the Navier-Stokes equation

Example: Navier-Stokes Equations

In incompressible Navier-Stokes equations,

$$\begin{cases} \partial_t u + (u\cdot\nabla)u - v\Delta u + \nabla p = 0 \\ \nabla \cdot u = 0 \end{cases} \quad \text{in } (0,T)\times\mathbb{R}^3 \text{ with } u(0) = u_0 \text{ given,}$$

if we multiply everything by u and integrate over the whole space gives, for each $t \in (0,T)$,

$$\underbrace{\int |u(t,x)|^2 \,\mathrm{d}x}_{\text{Total energy at } t} + \underbrace{2\int_0^t\int_{\mathbb{R}^3} |\nabla u|^2}_{\text{Amount of dissipation}} \leqslant \underbrace{\int |u_0(x)|^2 \,\mathrm{d}x.}_{\text{Initial total energy}} \qquad \text{(Energy Inequality)}$$

From this we see:

(1) $\{u_n\}$ is bounded in $L^\infty((0,T);L^2)$. Banach-Alaoglu Theorem shows $u_{n_k} \overset{*}{\rightharpoonup} u$ in $L^\infty((0,T);L^2)$.

(2) $\{\nabla u_n\}$ is bounded in $L^2((0,T);L^2)$. Theorem 12.10 shows $\nabla u_{n_k} \rightharpoonup f$ in $L^2((0,T);L^2)$. In fact $f = \nabla u$.

From above we obtain a candidate for a solution, and the entire method is called the *compactness method of obtaining weak solutions of PDEs.*

Lecture 35: 11/6 Uniform Convex Banach Spaces are Reflexive

Proof of Theorem 12.10 . Let $\{x_n\} \subset X$ be bounded and $Y := \overline{\operatorname{span}\{x_n\}}$. Immediately we see Y is separable and reflexive (Lemma 11.10). Therefore $Y \equiv Y^{**}$, and by the canonical map $\{x_n^{**}\}$ is bounded in Y^{**}. Treating Y^{**} as the dual of Y^*, by Banach-Alaoglu there exists a subsequence $\{x_{n_k}^{**}\}$ such that $x_{n_k}^{**} \xrightarrow{*} \Phi$ for some $\Phi \in Y^{**}$ (X^* is separable by Lemma 10.11 because X^{**} is). Since Y is reflexive and the canonical map surjective, $\Phi = x^{**}$ for some $x \in Y$. We need to show $f(x_{n_k}) \to f(x)$ for all $f \in X^*$.
Let $f \in X^*$ be given and define $f_y := f\big|_Y \in Y^*$. It follows that

$$f(x_{n_k}) = f_Y(x_{n_k}) \underbrace{=}_{\text{canonical map}} x_{n_k}^{**}(f_Y) \underbrace{\xrightarrow{*}}_{\text{weak-*}} \Phi(f_y) = x^{**}(f_Y) \underbrace{=}_{\text{canonical map}} f_Y(x) = f(x).$$

Indeed we do have a weakly convergent subsequence. □

Lemma 12.11: Goldstine's Theorem

Recall we said that we would prove uniformly convex Banach spaces are reflexive, but before doing that, we need a lemma. Let X be normed. For all $\Phi \in X^{**}$ with $\|\Phi\|_{X^{**}} \leqslant 1$ (i.e., $B_{X^{**}}$, the unit ball in X^{**}), there exists some sequence $\{x_n\} \subset B_X$, i.e., $\|x_n\|_X \leqslant 1$, such that $x_n^{**} \xrightarrow{*} \Phi$. (In other words, $J\left(\overline{B_X}\right)$ is weakly-* dense in $\overline{B_{X^{**}}}$.)

Remark

Recall that $J\left(\overline{B_X}\right)$ is a closed unit ball (by isometry) in a subspace of X^{**}. This lemma states that the weak-* closure of $J(B_X)$ is $\overline{B_{X^{**}}}$. *Taking weak-* closure is different from taking closure.*

For proof, see Theorem 2.6.26 in Megginson '98.

Theorem 12.12: Milman-Pettis Theorem

Uniformly convex Banach spaces are reflexive.

Proof. Let $\Phi \in X^{**}$ be such that $\|\Phi\|_{X^{**}} = 1$. We need to show that there exists $x \in X$ with $x^{**} = \Phi$. By Goldstine's Theorem above, there exists $\{x_n\} \subset \overline{B_X}$ such that $x_n^{**} \xrightarrow{*} \Phi$ in X^{**}. *In other words, for all* $x \in X^*$ *we have* $x_n^{**}(f) \to \Phi(f)$. Notice that this convergence also gives

$$\frac{x_n^{**} + x_m^{**}}{2} \xrightarrow{*} \Phi \text{ for } m > n, \text{ as } n \to \infty.$$

To put more formally, for all $\epsilon > 0$, *there exists* N *such that for all* $m > n \geqslant N$,

$$\left|\left(\frac{x_n^{**} + x_m^{**}}{2} - \Phi\right)(f)\right| < \epsilon \text{ for all } f \in X^*.$$

By Lemma 11.13.4 (weak-* version),

$$\|\Phi\| \leqslant \liminf_{n\to\infty} \left\| \frac{x_n^{**} + x_m^{**}}{2} \right\|,$$

i.e., for all $\delta > 0$*, there exists* N *such that, for all* $m > n \geqslant N$, $\|\Phi\| - \delta = 1 - \delta \leqslant \|(x_n^{**} + x_m^{**})/2\|$. But since J is an isometry, we equivalently have

$$1 - \delta \leqslant \|(x_n + x_m)/2\|. \tag{Δ}$$

Now we take a detour and recall the definition of uniform convex sets:

For all $\epsilon > 0$ there exists $\delta = \delta(\epsilon) > 0$ such that if $\|x\| = \|y\| = 1$ and $\|x - y\| \geqslant \epsilon$ then $\|(x+y)/2\| \leqslant 1 - \delta$.

Now let $\epsilon > 0$ be given. If we just set $\delta := \delta(\epsilon)$, by above there exists N such that any $m > n \geqslant N$ satisfy (Δ). However, because $\|(x_n + x_m)/2\| \geqslant 1 - \delta$ by (Δ), the contrapositive of uniform convexity states that $\|x - y\| < \epsilon$. This gives us nothing else but the Cauchy-ness of $\{x_n\}$, and so $x_n \to x$ for some $x \in X$ since X is Banach. Of course J is bounded, so $x_n^{**} \to x^{**}$ in X^{**}. By Lemma 11.13.1 (weak-* version), $x_n^{**} \xrightarrow{*} x^{**}$, and by part 3 of the same lemma, with the uniqueness of weak-* limits we conclude $\Phi = x^{**}$. □

Now we are done with weak convergence. See Evans "*Weak Convergence methods for Nonlinear PDEs*" for some nice applications.

Hilbert-Schmidt Theorem & Applications

Definition 12.13

Let $T \in B(H)$ be self-adjoint. Then

$$V(T) := \{(T(x), x) : x \in H, \|x\| = 1\}$$

is called the **numerical range** of T.

Lemma 12.14

$V(T) \subset \mathbb{R}$ and $\|T\| = \sup_{\lambda \in V(T)} |\lambda|$.

Proof. The first claim is obvious: since $(T(x), x) = (x, T(x)) = \overline{(T(x), x)}$ we see that $(T(x), x) \in \mathbb{R}$. Hence $V(T) \subset \mathbb{R}$.

For the second claim, let $M := \sup_{\lambda \in V(T)} |\lambda| = \sup\{|(T(x), x)| : x \in H, \|x\| = 1\}$. On one hand, note that $|(T(x), x)| \leqslant \|T(x)\|\|x\| \leqslant \|T\|\|x\|^2 = \|T\|$, we have $M \leqslant \|T\|$.

Now we show $M \geqslant \|T\|$. Notice that, for all $u, v \in H$,

$$\begin{aligned}(T(u+v), u+v) - (T(u-v), u-v) &= 2\left[(T(u), v) + (T(v), u)\right] \\ &= 2\left[(T(u), v) + (v, T(u))\right] && \text{(self-adjoint)} \\ &= 4\mathfrak{Re}(T(u), v).\end{aligned}$$

On the other hand,

$$\begin{aligned}(T(u+v), u+v) - (T(u-v), u-v) &\leqslant M\left(\|u+v\|^2 + \|u-v\|^2\right) && \text{(normalizing } u, v \text{ and taking coefficients (norms) out)} \\ &= 2M(\|u\|^2 + \|v\|^2). && \text{(parallelogram law)}\end{aligned}$$

Therefore $4\mathfrak{Re}(T(u), v) \leqslant 2M\left(\|u\|^2 + \|v\|^2\right)$. If $T(u) \neq 0$, set

$$v := \frac{\|u\|}{\|T(u)\|} T(u)$$

so that

$$4\mathfrak{Re}(T(u), v) = 4\|u\|\|T(u)\| \leqslant 2M\left(\|u\|^2 + \|v\|^2\right) = 2M \cdot 2\|u\|^2$$

and so $\|T(u)\| \leqslant M\|u\|$. If $T(u) = 0$ this inequality is still trivially true. This shows $\|T\| \leqslant M$. □

Lecture 36: 11/9 More on Self-Adjoint; Hilbert-Schmidt Theorem

Lemma 12.15

Let $T \in B(H)$ be self-adjoint.

(1) $\sigma_{\mathrm{p}}(T) \subset \mathbb{R}$, i.e., all eigenvalues are real.

(2) Suppose λ_1, λ_2 are two different eigenvalues, i.e., $T(x_i) = \lambda_i x_i$ and $\lambda_1 \neq \lambda_2$, then $(x_1, x_2) = 0$, i.e., the eigenvectors are orthogonal, not just linearly independent as we previously stated.

(3) If $T \in K(H)$ then either $\|T\|$ or $-\|T\|$ is an eigenvalue. This strengthens Corollary 9.7.

(4) If $Y \subset H$ is a closed subspace, then

$$T(Y) \subset Y \implies T(Y^\perp) \subset Y^\perp.$$

Future reference: Lemma 13.3, Theorem 13.9

Proof.

(1) If $T(x) = \lambda x$ then $(T(x), x) = (\lambda x, x) = \lambda \|x\|^2$. Since T is self-adjoint, $(T(x), x) = (x, T(x))$, meaning that $\lambda = \overline{\lambda}$, so $\lambda \in \mathbb{R}$.

(2) Suppose $\lambda_1 \neq \lambda_2 \neq 0$ and $x_1, x_2 \neq 0$ (otherwise $(x_1, x_2) = 0$). Then

$$\lambda_1(x_1, x_2) = (\lambda_1 x_1, x_2) = (T(x_1), x_2) \underbrace{=}_{\text{self-adjoint}} (x, T(x_2)) = \lambda_2(x_1, x_2) \implies (x_1, x_2) = 0.$$

(3) Recall that either $\|T\|$ or $-\|T\| \in \sigma(T)$. But then, since $\|T\| \neq 0$ and $T \in K(H)$ by assumption, Corollary 10.2 states that either $\|T\|$ or $-\|T\|$ must also be in $\sigma_{\mathrm{p}}(T)$, i.e., an eigenvalue.

(4) Take any $x \in Y^\perp$ and any $y \in Y$. By assumption $T(y) \in Y$. Therefore

$$0 = (T(y), x) = (y, T(x)) \text{ for all } y \in Y.$$

Therefore $T(x) \in Y^\perp$. Since x is arbitrarily chosen, $T(Y^\perp) \subset Y^\perp$.

Theorem 13.1: Hilbert-Schmidt Theorem / Eigenfunction Expansion Theorem

Let $T \in K(H)$ be compact and self-adjoint. Then there exists an orthonormal sequence $\{w_i\}_{i \geqslant 1}$ (or a finite set) of nonzero (zero is trivial) eigenvectors/eigenfunctions $\{\lambda\}_{i \geqslant 1}$ such that

$$T(x) = \sum_{i \geqslant 1} \lambda_i (x, w_i) w_i \text{ for all } x \in H.$$

Future reference: Corollary 13.2

Proof. By part 3 of the lemma above, there exists $w_1 \in H$ such that $T(w_1) = \pm\|T\| w_1$ (since either $\|T\|$ or $-\|T\|$ is an eigenvalue). We define $\lambda_1 := \pm\|T\|$ depending on which one it is.

Given orthonormal $\{w_1, \ldots, w_{n-1}\}$ (we will later show that we can indeed iteratively pick new w_i's under certain conditions), let $H_n \coloneqq \{w_1, \ldots, w_{n-1}\}^\perp$. We define $Y \coloneqq \operatorname{span}\{w_1, \ldots, w_{n-1}\}$. Note that the orthogonal complement does not depend on the span: $\operatorname{span}\{w_1, \ldots, w_{n-1}\}^\perp = \{w_1, \ldots, w_{n-1}\}^\perp$.
Since all the w_i's are eigenvectors/eigenfunctions, $T(y)$ is still a combination of w_i's for all $y \in Y$, i.e., $T(Y) \subset Y$. It follows from the previous lemma that (since Y is obviously closed) $T(Y^\perp) \subset Y^\perp$, i.e., $T(H_n) \subset H_n$. Setting $T_n \coloneqq T\big|_{H_n}$, we obtain self-adjoint $T_n \in K(H_n)$ (because T itself is on H).

(1) Case 1: $T_n \equiv 0$. Given $x \in H$, let $y \coloneqq x - \sum_{i=1}^{n-1} (x, w_i) w_i$. Notice that $y \in H_n$ because, for any $w_i \in \{w_1, \ldots, w_{n-1}\}$,

$$(y, w_i) = (x, w_i) - \sum_{j=1}^{n-1} (x, w_i)\delta_{ij} = (x, w_i) - (x, w_i) = 0. \tag{Δ}$$

(Notice that the way to pick w_i's ensures $\{w_1, \ldots, w_{n-1}\}$ is orthonormal whenever $i \neq j$: the newer terms were picked from the orthogonal complement of the set containing *all* previous terms, with $\|\cdot\| = 1$.) Then,

$$0 = T_n(y) = T(x) - \sum_{i=1}^{n-1} (x, w_i) T \underbrace{(x, w_i)}_{=\lambda_i w_i},$$

and we've shown what we want to.

(2) Case 2: $T_n \neq 0$. Then $\|T_n\| \neq 0$, and so there exists another w_n such that $T_n(w_n) = \pm\|T_n\| w_n$ (and we define $\lambda_n \coloneqq \pm\|T_n\|$ depending on which one is the eigenvalue). We can repeat this iteration for all $n \geqslant 1$. Now it remains to show that the original equation ($T(x) = \sum$) holds for the infinite sequence $\{w_i\}$.

Let $x \in H$ be given. We define the finite approximation

$$y_n \coloneqq x - \sum_{i=1}^{n-1} (x, w_i) w_i.$$

Note that $y_n \in H_n = \{w_1, \ldots, w_n\}^\perp$ just like in (Δ), so

$$\begin{aligned}
\|x\|^2 &= \|y_n\|^2 + 2\mathfrak{Re}(\underbrace{y_n}_{\in H_n}, \underbrace{\sum_{i=1}^{n-1} (x, w_i) w_i}_{\in Y}) + \sum_{i=1}^{n-1} |(x, w_i)|^2 && \text{(Gen. Pythagorean)} \\
&= \|y_n\|^2 + \sum_{i=1}^{n-1} |(x, w_i)|^2, && \text{(since } H_n = Y^\perp\text{)}
\end{aligned}$$

and thus $\|y_n\| \leqslant \|x\|$. Now we will show that $\|T(x) - \sum\| \to 0$: on one hand

$$\begin{aligned}\left\|T(x) - \sum_{i=1}^{n-1}(x, w_i)T(w_i)\right\| &= \left\|T(x) - \sum_{i=1}^{n-1}(x, w_i)\lambda_i w_i\right\| \\ &= \|T(y_n)\| \\ &= \|T_n(y_n)\| && (\text{since } T \equiv T_n \text{ on } H_n \text{ and } y_n \in H_n) \\ &\leqslant \|T_n\|\|y_n\| \leqslant |\lambda_n|\|x\|, && (\|T_n\| = |\lambda_n| \text{ and } \|y_n\| \leqslant \|x\|)\end{aligned}$$

while on the other hand Proposition 10.3 states that *any* sequence of nonzero eigenvalues of a compact operator converges to 0. Hence $\|T(x) - \sum\| \to 0$, and we are done. (Of course, our sequence here may have repeating λ's, but it will only repeat finitely many times, so the claim doesn't break down.) □

Corollary 13.2

Let H be a separable (infinite-dimensional) Hilbert space, and let $T \in K(H)$ be compact and self-adjoint. Let $F := \{f_k\}_{k\geqslant 1}$ be an orthonormal basis of $\ker(T)$ [the basis exists because $\ker(T)$ is automatically separable given H is, and if it's infinite-dimensional, then the existence of such basis is guaranteed by Proposition 6.5]. Then, $F \cup \{w_i\}$ (from the Hilbert-Schmidt Theorem above) is an orthonormal basis of H. *The basis of* $\ker(T)$ *can be supplemented with some orthonormal eigenvectors to form a basis of the whole space* H.
Future reference: Theorem 13.9

Proof. First note that $F \cup \{w_i\}$ is orthonormal: trivial without mixing; for all $f \in F$, $w_i \in \{w_i\}$,

$$(f, w_i) \underset{\text{since } \lambda_i \neq 0}{=} (f, \lambda_i w_i/\lambda_i) = \frac{1}{\lambda_i}(f, T(w_i)) = \frac{1}{\lambda}(\underbrace{T(f)}_{=0}, w_i) = 0.$$

We now want to show that any $x \in H$ has form

$$x = \sum_{k\geqslant 1}(x, f_k)f_k + \sum_{i\geqslant 1}(x, w_i)w_i.$$

(Then $F \cup \{w_i\}$ is an orthonormal basis by Theorem 5.21.2.) Notice that $\ker(T)$ is closed, so by the projection theorem we can (uniquely) decompose $x := u + y$, where $u \in \ker(T)$ and $y \in (\ker(T))^\perp$. Then

$$u = \sum_{k\geqslant 1}(u, f_k)f_k = \sum_{k\geqslant 1}\left[(u, f_k) + \underbrace{(y, f_k)}_{=0}\right]f_k = \sum_{k\geqslant 1}(x, f_k)f_k.$$

It remains to show that $z := y - \sum_{i\geqslant 1}(x, w_i)w_i = \sum_{i\geqslant 1}(y, w_i)w_i = 0$ (since $(u, w_i) = 0$ and subtracting doesn't bother). Notice that $z \in (\ker(T))^\perp$ as y and each of $(x, w_i)w_i$ is. By the Hilbert-Schmidt Theorem,

$$T(z) = T\Big(y - \sum_{i\geqslant 1}(x, w_i)w_i\Big) \underset{\text{Hilbert-Schmidt}}{=} 0,$$

so $z \in \ker(T)$. Since $\ker(T) \cap (\ker T)^\perp = \{0\}$ we conclude that $z = 0$. Done. □

Lecture 37: 11/11&13 The Sturm-Liouville Problem

We begin by fixing

$$\begin{cases} p \in C^1([a,b]) & p > 0 \\ q \in C([a,b]) & q \geqslant 0 \end{cases} \qquad \text{on } [a,b].$$

And we define

$$L[u] := -(pu')' + qu \text{ for } u \in C^2.$$

We will consider the **Sturm-Liouville problem**

$$L[u] = f \text{ for } u \in \mathfrak{D} := \{u \in C^2([a,b]) : u(a) = u(b) = 0\}$$

by finding an orthonormal basis of L^2 that consists of the eigenvalues / eigenfunctions of this problem.

Lemma 13.3

(Cf. Lemma 12.15.)

(1) L is self-adjoint: $(L[u], v) = (u, L[v])$ for all $u, v \in \mathfrak{D}$.

(2) If nonzero $u \in \mathfrak{D}$ is an eigenvector, i.e., $L[u] = \lambda u$, then $\lambda > 0$.

Corollary 13.4

In particular, if $u \in \mathfrak{D}$ and $L[u] = 0$ then $u = 0$. Also refer to Theorem 13.7.

(3) Eigenvectors corresponding to distinct eigenvalues are orthogonal in L^2.

Future reference: Corollary 13.10, Example 13.11

Lemma 13.5

Suppose $u_1, u_2 \in C^2([a,b])$ are nonzero solutions of

$$-(pu')' + wu = 0 \text{ for } w \in C([a,b]).$$

(Note that, unlike the Sturm-Liouville problem, q is replaced by w, without the requirement $w \geqslant 0$, and $u_1, u_2 \in C^2([a,b])$ not just $\mathfrak{D}$, i.e., no requirements on agreeing at endpoints.) Then the **Wronskian**

$$\mathcal{W}_p(u_1, u_2)(x) := p(x)\,(u_1'(x)u_2(x) - u_2'(x)u_1(x))$$

is constant (in x) and nonzero if and only if u_1, u_2 are linearly independent.

Proof. It simply requires brute force to show $\mathcal{W}_p' \equiv 0$. Now we prove the $\iff$ statement. Since $p > 0$,

$$\mathcal{W}_p' \neq 0 \iff \det\begin{vmatrix} u_1 & u_2 \\ u_1' & u_2' \end{vmatrix} \neq 0.$$

(In this case the determinant is invariant under x since $\mathcal{W}_p'$ is constant.)

$\implies$ (nonzero Wronskian $\implies$ linearly independent u_1, u_2): suppose $\alpha u_1 + \beta u_2 = 0$. Taking derivative gives $\alpha u_1' + \beta u_2' = 0$. Since the matrix is nonsingular we know $\alpha = \beta = 0$.

$\impliedby$: suppose the determinant is 0 for some (each, in fact) x. Then there exists $(\alpha, \beta) \neq (0,0)$ such that

$$\begin{cases} \alpha u_1 + \beta u_2 = 0 \\ \alpha u_1' + \beta u_2' = 0 \end{cases} \quad \text{at some } x_0 \in [a,b].$$

This immediately contradicts the linear independence of u_1 and u_2. □

Corollary 13.6

All eigenvalues of $L[u]$ are simple, i.e.,

$$\dim(E_\lambda) \leqslant 1, \text{ where } E_\lambda := \{u \in \mathfrak{D} : L[u] = \lambda u\}.$$

Future reference: Corollary 13.10

Proof. Suppose $L[u] = \lambda u$ and $L[v] = \lambda v$. Applying the lemma above with $w : q - \lambda$ gives $\mathcal{W}_p(u,v) =$ constant. Evaluating it at an endpoint, e.g., at u, gives $\mathcal{W}_p(u,v)(a) = 0$ (since $u(a) = v(a) = 0$). Therefore u, v are linearly dependent. □

Back to the problem:

Let $u_1, u_2 \in C^2([a,b])$ be fixed linearly independent solutions (so that the Wronskian later defined is nonzero) of $L[u] = 0$ with $u_1(a) = 0$ and $u_2(b) = 0$ (i.e., one vanishes at one endpoint and the other at the other endpoint). By ODE theory, $L[u] = 0$ can be re-written as

$$u' = v \text{ and } v' = -\frac{p'}{p} + \frac{q}{p}u.$$

Now apply the boundary conditions:

$$\begin{cases} u_1(a) = 0 & u_2(b) = 0 \\ u_1'(a) = 1 & u_2'(b) = 1 \end{cases}. \tag{*}$$

Theorem 13.7

If (nonzero) $\mathcal{W} := \mathcal{W}_p(u_1, u_2)$, then

$$G(x,y) := \begin{cases} \dfrac{u_1(x)u_2(y)}{\mathcal{W}} & x < y \\ \dfrac{u_1(y)u_2(x)}{\mathcal{W}} & x \geqslant y \end{cases}$$

is a **Sturm-Liouville Green's function**, i.e., letting

$$u(x) := \int_a^b G(x,y)f(y)\, \mathrm{d}y \text{ for some } f \in C([a,b]),$$

we have $u \in \mathfrak{D}$ and u is a *unique* solution of $L[u] = f$ in $\mathfrak{D}$. (*Green's function is a kernel that gives solutions to nonhomogeneous problems.*) Future reference: Theorem 13.9

Proof. Note that (treating x as a fixed variable and integrating with respect to y)

$$u(x) = \frac{u_2(x)}{\mathcal{W}} \int_a^x u_1 f + \frac{u_1(x)}{\mathcal{W}} \int_x^b u_2 f \implies u(a) = u(b) = 0. \tag{Δ}$$

Now we check if $u \in \mathfrak{D}$:

$$u'(x) = \frac{u_2'(x)}{\mathcal{W}} \int_a^x u_1 f + \frac{u_1'(x)}{\mathcal{W}} \int_x^b u_2 f + 0$$

(cross terms of product rule cancel out, $u_2(x)u_1(x)f(x) - u_1(x)u_2(x)f(x) = 0$)

$$u_2''(x) = \frac{u_2''(x)}{\mathcal{W}} \int_a^x u_1 f + \frac{u_1''(x)}{\mathcal{W}} \int_x^b u_2 f + \underbrace{\frac{1}{\mathcal{W}} \left(u_2'(x)u_1(x)f(x) - u_1'(x)u_2(x)f(x) \right)}_{=-f(x)/p(x) \text{ by definition of Wronskian}}.$$

Indeed, $u \in \mathfrak{D}$. It remains to show $L[u] = -pu'' - p'u' + qu = f$ which is only a matter of computation by brute force. And the uniqueness is guaranteed by Corollary 13.4. $\square$

Example 13.8

Let $p(x) \equiv 1, p(x) \equiv 0$, and $[a,b] := [0,1]$, then $L[u] = -u''$. Then, by (*), we can construct

$$\begin{cases} u_1(x) = x \\ u_2(x) = x - 1 \end{cases} \implies \mathcal{W} = u_1'u_2 - u_2'u_1 = -1.$$

Then,

$$G(x,y) = \begin{cases} x(1-y) & x < y \\ (1-x)y & x \geqslant y \end{cases}.$$

Future reference: Example 13.11

Now we are going to consider the integration against as kernel of the operator on L^2: define

$$Tf(x) := \int_a^b G(x,y) f(y)\,\mathrm{d}y \text{ for } f \in L^2(a,b) \text{ over } \mathbb{R}.$$

Some more theorems. Notice that we are going towards the direction of using Hilbert-Schmidt Theorem.

Theorem 13.9

(1) $T \in K(L^2(a,b))$, T is self-adjoint, and $\ker(T) = \{0\}$.

(2) Eigenvalues of T form a sequence converging to 0.

(3) Eigenvectors / eigenfunctions of T are elements of $\mathfrak{D}$ and they form an orthonormal basis of $L^2(a,b)$.

Future reference: Corollary 13.10

Proof.

(1) T is compact and self-adjoint by Example 9.3.2. Now we show $\ker(T) = \{0\}$, i.e., if $T(f) = 0$ then $f = 0$. First note that $\mathfrak{D} \subset \mathrm{im}(T)$: given $u \in \mathfrak{D}$, letting $g := L[u] \in C([a,b]) \subset L^2(a,b)$, we have $u = T(g)$ by Theorem 13.7.

Now suppose $T(f) = 0, f \in L^2$, and let $u \in \mathfrak{D}$. Since $\mathfrak{D} \subset \mathrm{im}(T)$, $u = T(g)$ for some $g \in L^2$. Then,

$$0 = (\underbrace{T(f)}_{=0}, g) \overset{\text{self-adjoint}}{=} (f, T(g)) = (f, u) \text{ for all } u \in \mathfrak{D}. \tag{Δ}$$

Since $\mathfrak{D}$ is dense in L^2 (recall from Lemma 5.24 that we can first approximate anything in L^2 by simple functions and then interpolate them. Here we just need to make continuous functions even "nicer", i.e., second-order continuously differentiable, by taking mollification, for example), (Δ) gives $(f,u) = 0$ for all $u \in L^2$. Taking $u := f$ tells us $f = 0$.

Remark

Showing (Δ), i.e., f is orthogonal to a dense subset of space, is often times a useful trick. For example, that $\mathrm{im}(\partial_t - \Delta)$ is dense in L^p is a powerful tool to deduce the uniqueness of the heat equation in L^2. See Giga×2, Saal (2010), Section 4.4.2.

2) and 3) Suppose $T(f) = \lambda f$ for some $\lambda \in \mathbb{R}$ (recall from Lemma 12.15.1 that eigenvalues of self-adjoint operators are real). Then,

$$\begin{aligned}
T(f(x)) = \frac{u_2(x)}{\mathcal{W}} \int_a^x u_1 f + \frac{u_1(x)}{\mathcal{W}} \int_x^b u_2 f & & \text{((Δ) in Theorem 13.7)}\\
\implies T(f(a)) = T(f(b)) = 0 \text{ and } T(f) \in C([a,b]) & &\\
\implies \lambda f(a) = \lambda f(b) = 0 \text{ and } \lambda f \in C([a,b]) & &\\
\implies {\color{red} f(a) = f(b) = 0} \text{ and } f \in C([a,b]) & &\\
\implies T(f) \in C^1([a,b]) & & \text{(by its construction)}\\
\implies \lambda f, f \in C^1 \implies T(f), \lambda f, \text{ and } {\color{red} f \in C^2}. & & \text{(doable b/c $u_1, u_2 \in C^2$)}\\
\implies f \in \mathfrak{D}. & &
\end{aligned}$$

(In fact, we can iterate this $T(f) \to \lambda f \to f$ procedure and get $f \in C^\infty(a,b)$ as long as p, q are.) Then 3) follows from Corollary 13.2 since $\ker(T) = \{0\}$ and it does not contribute to the forming of such basis; all elements in this orthonormal basis of $L^2(a,b)$ must therefore come from orthonormal eigenvectors of T. Once again, Proposition 10.3 asserts that these eigenvectors converge to 0, with which we are also done with 2). □

Corollary 13.10

The eigenvectors / eigenfunctions of the Sturm-Liouville problem $L[u] = \lambda u$ form an orthonormal Schauder basis of $L^2(a,b)$. Furthermore,

(1) each eigenvalue is positive (Lemma 13.3.2) and simple (Corollary 13.6), and

(2) the eigenvalues can be ordered to form a sequence converging to ∞. *Unlike above, where we showed that the eigenvalues of T, the Sturm-Liouville solution operator, converge to 0, here we are talking about the eigenvalues of the Sturm-Liouville problem itself, not the solution operator anymore.*

Proof. Recall that (in the proof of Theorem 13.9.1) that $T(f) = T(L[u]) = u$ for some $u \in \mathfrak{D}$. Then $T(f) = \lambda f \iff L[u] = u/\lambda$. From this equation we see that the eigenvalues of the Sturm-Liouville problem are the $(1/\lambda)$'s. Therefore as $\lambda \to 0$, these eigenvalues $\to \infty$. □

Example 13.11

Recall Example 13.8 where $L[u] = -u''$ and $[a,b] = [0,1]$. By the corollary above the eigenvalues / eigenfunctions of

$$\begin{cases} -u'' = \lambda u \\ u(0) = u(1) = 0 \end{cases}$$

form an orthonormal (Schauder) basis of $L^2(0,1)$.

Proof. First we find the eigenfunctions (just normal ODE stuff):

$$\begin{cases} \lambda < 0 \implies u = Ae^{\sqrt{-\lambda}x} + Be^{-\sqrt{-\lambda}x} \implies A = B = 0 \\ \lambda = 0 \implies u = Ax + B \implies A + B = 0 \end{cases} \text{no eigenfunctions (Lemma 13.3.2)}$$
$$\lambda > 0 \implies u = A\sin\sqrt{\lambda}x + B\cos\sqrt{\lambda}x \implies B = 0 \text{ and } \lambda_k = (k\pi)^2.$$

Therefore *all* eigenvalues are of form $\lambda_k := (k\pi)^2$. As expected from the corollary above, they tend to ∞ as $k \to \infty$. Then the normalized eigenfunctions (with respect to $\|\cdot\|_{L^2(0,1)}$) are $\left\{\frac{1}{\sqrt{2}}\sin k\pi x\right\}_{k\geqslant 1}$ and they form another orthonormal basis for L^2. (*Compare this with Example 5.23.*) Thus,

$$f(x) = \sum_{k\geqslant 1} \frac{1}{\sqrt{2}}\left(f, \frac{1}{\sqrt{2}}\sin k\pi x\right)\sin k\pi x \text{ for all } f \in L^2(0,1),$$

and we have just obtained the **sine expansion** of f. Notice that each eigenfunction satisfies $u(0) = u(1) = 0$, but we've just shown that it approximates *anything* in $L^2(0,1)$, as what we've shown is convergence with respect to $\|\cdot\|_{L^2}$ but not pointwise convergence. In addition, no cosines are involved in this expansion as the bases differ by $k\pi$ not $2k\pi$ — we can extend $L^2(0,1)$ to the set of all odd functions on $L^2(-1,1)$ in which no Fourier expansion involves cosine. □

Example 13.12

Let $\Omega \subset \mathbb{R}^n$. Then the eigenfunctions of Laplacian

$$\begin{cases} -\Delta u = \lambda u & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

form an orthonormal Schauder basis of $L^2(\Omega)$. See Alt's book, Theorem 12.17 for more details.

THE END. FEELS GOOD BEING A BANACH SPACEMAN. CHEERS.