APPLIED FUNCTIONAL ANALYSIS

ABSTRACT. To be filled in.

CONTENTS

1. 1/4 : Topology, metric spaces recap

Motivations for learning functional analysis includes:

- PDE:

$$\begin{cases} -\Delta u(x) = f(x) & x \in \mathbb{R}^n \\ u(x) = g(x) & x \in \partial \end{cases}$$

  where f and g can be of any space to be discussed.
- SDE:

$$dx(t) = bdt + \sigma dW_t \quad \text{on} \quad L^p(X, \mathcal{F}, \mathbb{P}).$$

- Often times we might replace $\partial_x u$ by the derivative quotient, but this requires us taking the limit of a sequence of quotients, for which often we need compactness to admit a limit.
- Machine Learning.

1.1. **The big picture.** Let's look at the following picture of our materials:

Topological Space				Measurable spaces
· open sets · continuity		Vector Space		· σ-algebra · measure
↓		↓		
Metric Space	→	Normed Vector Space		↙
· distance · convergence	Fréchet space	↓		
· completeness	→	Banach Space	←	e.g. of bounded operators
↓		↓	↖	
Orthogonality · inner product	→	Hilbert Space	←	L^2 space

Now we start to do the real analysis. Definitions first.

1.2. **Topological Spaces.**

If we want to consider all subspaces of a set, then we are doomed to fail! So we only consider some kind of subsets. For a topology, we consider open sets in it.

Def 1.1. *Topology is a collection $\mathcal{T}$ of open subsets of the set X such that it satisfies*

(i) $\emptyset$ and X are open.
(ii) $\mathcal{T}$ is closed under infinite union and finite intersection.

We call the pair $(X, \mathcal{T})$ a topology even we often use just one of the letters. We also say that A is closed iff $A^c = X \backslash A$ is open.

Def 1.2. *V is a neighborhood of x if for some open set G we have $x \in G \subset V \subset X$.*

Def 1.3. *$\mathcal{T}$ is Hausdorff (or separated) if $\forall x, y$ distinct then $\exists$ neighborhoods $x \in V_x$, $y \in V_y$ and $V_x \cap V_y = \emptyset$.*

Example 1.4. *some topology spaces are:*

- Discrete topology space: every point is open, i.e. $\mathcal{T} = P(X) = 2^X$. The problem with this is that it is too rich!
- The trivial topology: $\mathcal{T} = \{\emptyset, X\}$. We can see that it is not Hausdorff for a non singleton X. So the problem with this is that it is too crude!

- The generated topology by collection of subsets $\mathcal{T}_0$: the minimal way to construct a topology containing it, i.e. the intersection of all larger topology (which at least exist 1, which is the discrete topology).

Def 1.5. *convergence: you know it...*

Def 1.6. *continuity: A function $f : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y))$ is continuous at x if $\forall W_{f(x)}$ neighborhood of $f(x)$, $\exists V_x$ such that*

$$f(V_x) \subset W_{f(x)};$$

Also, we say that f is continuous if it is continuous at all $x \in X$.

Theorem 1.7. *f is continuous $\iff$ $f^{-1}(open)$ is open, where the notation f^{-1} is just the pre-image, not necessarily a function.*

Def 1.8.

- *$f : X \to Y$ (with $\mathcal{T}$) is a homeomorphism if it is a bijection such that f and f^{-1} are continuous.*
- *X and Y are homeomorphic if there is a homeomorphism $f : X \to Y$.*

Note that homeomorphic means that the two sets are basically the same if we only consider on the level of open sets.

But now we are not able to distinguish large balls from small balls, so we add more things later in metric spaces. For now we introduce compactness.

Def 1.9. *Compactness: $K \subset X$ is compact if every open cover of K admits a finite subcover.*

Example 1.10.

- On $\mathbb{R}$, let $\mathcal{T}$ be the topology generated by open intervals $(a, b) = \{x \in \mathbb{R}, a < x < b\}$, then $(0, 1)$ is not compact.

  reason: we can construct cover

  $$(0,1) = \bigcup_{n \geq 3} \left(\frac{1}{n}, 1 - \frac{1}{n}\right)$$

  for which any finite subset is not a cover.
- An alternative way of saying it is that for sequences with limit 1, the limit is not in the set.
- $[0, 1]$ is closed, and hence by Heine-Borel theorem it is compact.
- $[0, 1]^5$ is compact.
- $[0, 1]^\infty$ on the other hand, is not compact. The problem is that the space is not complete.

One of our later goals is to weaken the topology such that the set is compact, so we can do things on it. Of course we'll get a weaker result, but it is better than none.

1.3. **Metric Space.**

we include distance here. Let X be non-empty, then

Def 1.11. *A* <u>*metric*</u> *on* X *is* $d : X \times X \to \mathbb{R}$ *such that*

(i) $d(x, y) = d(y, x)$;
(ii) $d(x, z) \leq d(x, y) + d(y, z)$;
(iii) $d(x, y) \geq 0$;
(iv) $d(x, y) = 0 \Rightarrow x = y$.

Example 1.12.

- On $\mathbb{R}$, $|x - y| = d(x, y)$ is the Euclidean distance.
- For Cartesian products, we can define the L^1 norm as

$$d_{X\times Y}((x_1, y_1), (x_2, y_2)) = d_X(x_1, x_2) + d_Y(y_1, y_2).$$

Def 1.13. *The* <u>*natural topology*</u> *on a metric space is the topology* $\mathcal{T}$ *generated by open balls* $B_\varepsilon(x) = \{y \in X | d(x, y) < \varepsilon\}$. *We also write close balls as* $\overline{B_\varepsilon(x)}$.

1.4. **Vector Space.**

Def 1.14. *A* <u>*vector space*</u> V *over a field* F*, where in most cases* $F = \mathbb{R}$ *or* C *we have that*

(i) $(V, +)$ *is an abelian group;*
(ii) $\times$ *is a multiplication* $F \times V \to V$ *and* $\lambda f \in V$.

Example 1.15.

- $\mathbb{R}^n$;
- For $f : (0, 1) \to \mathbb{R}$, the $L^2(0, 1)$ is a vector space such that all $f \in L^2(0, 1)$ satisfies

$$\int_0^1 |f|^2 dx \leq \infty;$$

- $C[0, 1]$;
- The space of bounded operators.

Def 1.16. *A* <u>*norm*</u> *is a function* $|| \cdot || : V \to \mathbb{R}$ *such that it satisfies*

(i) $||x|| \geq 0$;
(ii) $||\lambda x|| = |\lambda| ||x||$ *for* $\lambda \in F$;
(iii) $||x + y|| \leq ||x|| + ||y||$;
(iv) $||x|| = 0 \Rightarrow x = 0$.

We call such $(V, || \cdot ||)$ a normed vector space. If only the first 3 conditions are satisfied, we call it a semi-norm.

2. 1/6: CONVERGENCE, CONTINUITY, AND COMPACTNESS

Def 2.1. *convergence: $x_n \to x \in X$ if $\forall \varepsilon > 0$, $\exists N = N(\varepsilon) \in \mathbb{N}^*$ such that $\forall n \geq N$, $d(x_n, x) \leq \varepsilon$.*

As for the topology sense, we just say that for any open set containing x, there is a point in the sequence in that set.

A sequence is Cauchy if $\forall \varepsilon > 0, \exists N \in \mathbb{N}^*$ such that $\forall n, m \geq N$, $d(x_m, x_n) < \varepsilon$.

Def 2.2. *A Banach space is a complete normed vector space.*

Example 2.3.

- Banach spaces: $C^0(X), C^{k,\alpha}(X), L^p(X), W^{m,p}(X)$;
- Not Banach spaces: $C^\infty(X) = \bigcap_{k\geq 1} C^k(X)$, Schwartz class $\zeta(\mathbb{R}^n)$, and the duality space of Schwartz class, $\zeta'(\mathbb{R}^n)$ the space of distributions.

Theorem 2.4. *Every metric space (X, d) has a completion $(\tilde{X}, \tilde{d})$ such that $X \subset \tilde{X}$, $\tilde{d}(x, y) = d(x, y)$ for $x, y \in X$ and X is dense in $\tilde{X}$.*

Well, it's not really as so since we need isometry to finish this. The sketch of construction is the following.

- Take any $\tilde{\tilde{x}} := (x_n)$ as a Cauchy sequence in X.
- Define equivalence class on space of Cauchy sequences as $\tilde{\tilde{x}} \sim \tilde{\tilde{y}}$ if $\lim\limits_{n\to\infty} d(x_n, y_n) = 0$.
- Let $\tilde{X}$ be the set of all equivalence classes defined above. Let $\tilde{d}(\tilde{x}, \tilde{y}) = \lim\limits_{n\to\infty} d(x_n, y_n)$ for any representative of the class. It is easy to check that such a distance is well defined.
- It can be shown that the space is complete and it is an extension, and it is unique up to isomorphism.

Example 2.5. *$C([0, 1])$ with the metric $||\cdot||_2$ is not complete, but it is a normed vector space. To see this just use the steeper and steeper function that goes to an indicator function.*

Remark 2.6.

- *The completion of $(C(X), ||\cdot||_p)$ is $(L^p(X), ||\cdot||_p)$.*
- *$C([0, 1], ||\cdot||_\infty)$ is Banach with the uniform norm.*

Now we look at continuity.

Def 2.7.

- *A function is continuous at x_0 if $\forall \varepsilon > 0$, $\exists \delta(\varepsilon)$ such that for $d_X(x, x_0) < \delta$, we have $d_Y(f(x), f(x_0)) < \varepsilon$.*
- *If a function is continuous at all points in X it is continuous.*

- *A function is uniformly continuous if δ does not depend on x_0.*
- *A function is sequentially continuous at x if $x_n \to x \to f(x_n) \to f(x)$.*

Proposition 2.8.

- *f is continuous $\iff$ it is sequentially continuous.*
- *For $F \subset X$, it is closed $\iff$ for all $x_n \to x$ we have $x \in \mathcal{F}$.*

Def 2.9.

- *The closure of A is $\bar{A} := \left\{ x \in X \middle| \forall x_n \in A, x_n \to x \right\}$.*
- *$A \subset X$ is dense in X when $\bar{A} = X$.*
- *A subset is separable if it has a countable dense subset.*

Now we deal with compactness.

Def 2.10. *A space is sequentially compact if every sequence in K admits a converging subsequence in K.*

Theorem 2.11. *$K \subset X$ in a metric space, then K compact $\iff$ K is sequentially compact.*

Also, we call a set precompact if its closure is compact.

Proposition 2.12. *K is compact $\to$ K is bounded and closed.*

Theorem 2.13. *(Heine-Borel) Subsets of $\mathbb{R}_n$ (finite dimensions) are compact iff they are closed and bounded.*

To prove this we use the Bolzano-Weistrass theorem. The link between them is fairly simple.

Theorem 2.14. *(Bolzano-Weistrass) Every bounded sequence in $\mathbb{R}^n$ admits a convergent sequence.*

Proof. This is simply by dividing space into smaller and smaller boxes. At least one box will contain infinitely many points. □

Note that the Heine Borel does not apply in infinite dimensions, for example, in l^∞ we can choose the standard basis to construct a non-Cauchy sequence, hence not convergent for any subsequence.

Def 2.15.

- *$\{G_\alpha, \alpha \in \mathcal{A}\}$ is a cover of A if $A \subset \bigcup_{\alpha \in \mathcal{A}} G_\alpha$.*
- *$\{x_\alpha, \alpha \in \mathcal{A}\}$ is a ε-net of A if $A \subset \bigcup_{\alpha \in \mathcal{A}} B_\varepsilon(x_\alpha)$.*
- *$A \subset X$ is totally bounded if it has a finite ε-net for all $\varepsilon > 0$.*

Theorem 2.16. *$A \subset X$ is sequentially compact iff it is complete and totally bounded.*

3. 1/9: SPACES OF CONTINUOUS FUNCTIONS

Recall that we've gone over the definition of continuous functions.

Theorem 3.1. *if $f : K \to Y$ is continuous for K compact, then*

- *$f(K)$ is compact;*
- *If $Y = \mathbb{R}$, then f attains its minimum and maximum;*
- *f is uniform continuous.*

Sketch of proof are:

- using sequentially continuous the convergent subsequence is passed;
- first find inf/sup, then find a sequence closer and closer to that point, then use completeness of $\mathbb{R}$;
- By contradiction, find a sequence with each point ϵ far away from each other, then contradict the compactness property.

Now we start looking at continuous functions as points in its vector space. To do so we need to compare functions, for this purpose we have the following tools:

- pointwise convergent: $f_n(x) \to f(x)$; This does not preserve continuity using simple example x^n on $[0, 1]$.
- L^2 metric: $C([0,1], ||\cdot||_2)$ also don't preserve continuity with counter example as a sharper and sharper indicator function. That is, they are Cauchy, but the limit is an indicator function, so not continuous.
- Turns out that the right metric we're looking here is the infinite norm $||\cdot||_\infty$.

Theorem 3.2. *Uniform norm preserves continuity, i.e. for f_n that are bounded and continuous with $||f_n - f|| \to 0$, we have that f is continuous.*

Proof. The proof can be seen in any analysis classes but it's well to look at it since it has some trick that's basically the only trick of this kind.

We have

$$\begin{aligned}|f(x) - f(y)| &= |f(x) - f_n(x) + f_n(x) - f_n(y) + f_n(y) - f(y)| \\ &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)|\end{aligned}$$

where we can bound the three terms by $\frac{\epsilon}{3}$. The procedure is that we first fix some n such that $|f - f_n| \leq \frac{\epsilon}{3}$, in other words we're done with the first and last term, and then use the continuity to find appropriate δ using the continuity of f_n to bound the middle term. □

Since for K compact $f(K)$ is compact, it is bounded, so the uniform norm always exists. In other word the space $(C(K), ||\cdot||_2)$ is a well-defined normed vector space.

Theorem 3.3. *$C(K)$ is Banach.*

Proof. What we need to prove is that for any Cauchy sequence of functions f_n, we must prove 2 things: that $f_n \to f$ for some f and that $f \in C(K)$. Since uniform convergence preserves continuity we are done with the second thing if we can find the correct limit function f. So let's do that.

The choice of f is quite intuitive as it's just the pointwise limit of f_n:

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

which is well defined since f_n Cauchy means that for all $x \in K$ we have that $f_n(x)$ is Cauchy, and by completeness of $\mathbb{R}$ (or any complete Y), the limit exists.

Now let's prove that it's the case that $f_n \to f$. f_n Cauchy means that $||f_n - f_m|| \leq \varepsilon$ for large enough n, m, and hence

$$\begin{aligned} ||f_n - f|| &= \sup_{x\in K} |f_n(x) - f(x)| = \sup_{x\in K} |f_n(x) - \lim_{m\to\infty} f_m(x)| \\ &= \sup_{x\in K} \liminf_{m\to\infty} |f_n(x) - f(x)| \leq \liminf_{n\to\infty} \sup_{x\in K} |f_n(x) - f_m(x)| \to 0. \end{aligned}$$

Note that in the deduction we changed lim to lim inf. This is valid since when the limit exists it is the lim inf, but we cannot be sure that the limit of the sup exists in general, although in this case sure it exists. □

Now we introduce some more sets of continuous functions. We have

Def 3.4.

- *$\mathcal{C}_c(X)$ is the set of all compactly supported continuous functions on X.*
- *$\mathcal{C}_0(X)$ is the set of all continuous functions on X such that $f \to 0$ at ∞.*
- *$\mathcal{C}_b(X)$ is the set of all bounded continuous functions on X.*

Given these definitions we can claim that

Proposition 3.5.

$$\mathbb{C}_c(X) \subset \mathbb{C}_0(X) \subset \mathbb{C}_b(X) \subset \mathbb{C}(X).$$

In addition we have that $\mathbb{C}_0(X) = \overline{\mathbb{C}_c(X)}$, $\mathbb{C}_c(X)$ is not Banach, $\mathbb{C}_0(X)$ and $\mathbb{C}_b(X)$ are Banach, while $\mathbb{C}(X)$ is not a normed vector space for the infinite norm.

Theorem 3.6. *(Weierstrass) Polynomials are dense in $\mathbb{C}([a, b]; ||\cdot||)$.*

Sketch of proof (completed later):

The first step is to find a function that maps $[a, b] \to [0, 1]$, for which we only need to show the result for $\mathbb{C}([0, 1]; ||\cdot||)$.

Then, we use the Bernstein polynomial

$$B_n(x; f) := \sum_{k=1}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k} \in P[x]$$

where we can check that

$$B_n\left(\frac{k}{n}; f\right) = f\left(\frac{k}{n}\right)$$

and

$$||B_n(\cdot; f) - f|| \leq \varepsilon + \frac{||f||}{2n\delta^2} < 2\varepsilon$$

for large n.

Now what about a compactness? Well, the Heine Borel doesn't work here so we'll have to find something different.

Def 3.7. *A family $\mathcal{F}$ is equicontinuous if $\forall \varepsilon > 0, \exists \delta = \delta(\varepsilon)$ (doesn't depend on f) such that*

$$d(x, y) < \delta \to d_y(f(x), f(y)) < \delta, \forall f \in \mathcal{F}.$$

Then we have the Ascolli-Arzela theorem;

Theorem 3.8. *(Ascolli-Arzela) For K a compact metric, we have that a subset in $(\mathbb{C}(K), ||\cdot||)$ is compact iff it's closed, bounded, and equicontinuous.*

Example 3.9.

Def 3.10. *$f : X \to \mathbb{R}$ is Lipschitz continuous if $\exists c > 0$ such that*

$$|f(x) - f(y)| \leq c \cdot d_x(x, y).$$

We call the Lipschitz constant

$$\mathrm{Lip}(f) := \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)} < \infty$$

Proposition 3.11. *Let*

$$\mathcal{F}_n := \{f - \textit{continuous} \mid \mathrm{Lip}(f) \leq n\}$$

is equicontinuous.

The proof is just to take $\delta = \dfrac{\varepsilon}{n}$ in the $\varepsilon - \delta$ argument.

4. 1/11 Peano construction

Today we apply the Ascoli-Arzela to see how it gives a construction to solutions of differential equations.

But first let's look at a counter example that's closed, bounded but not equicontinuous. Then construct a sharp bump on $\left[\frac{1}{2^{n+1}}, \frac{1}{2^n}\right]$, then it's a set of non-Cauchy sequence in the infinite norm, thus not compact.

Now, in general if we have

$$\rho(K, d_k) \to (Y, d_y)$$

with $C(K, Y)$ being the space of functions $f : K \to Y$ continuous with

$$d(f, g) = \sup_{x \in K} \{dy(f(x), f(y))\}$$

then $C(K, Y)$ is a complete metric space. This is the extension of completeness of $C(K)$ last time, and an extension of the Ascoli-Arzela also works. But I'll not copy that here.

Now we want to solve the ODE

$$\dot{u} = f(t, u(t))$$

where as a common example we have

$$\frac{d}{dt}\begin{pmatrix} x(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} v(t) \\ a(t) \end{pmatrix} = \begin{pmatrix} v(t) \\ \frac{F(t,x(t))}{m} \end{pmatrix} = f\left(t, \begin{pmatrix} x \\ v \end{pmatrix}\right) \in \mathbb{R}^6.$$

But we won't bother with that and will just focus on a initial value problem. So for the question

$$\begin{cases} \dot{u} = f(t, u(t)) \\ u(0) = u_0 \end{cases}$$

where $|f(t, u)| \leq M$ in some box around $u(0)$. This condition means that u is Lipschitz M.

Now to find a solution one way is to go in the following steps (Peano):

(i) Construct approximations $u_h(t)$;
(ii) Extract a convergent subsequence to $u(t)$ by Ascoli-Arzela;
(iii) Check whether $u(t)$ satisfies the ODE, and in what sense.

Now, for the construction of (i). Heuristically we want

$$\dot{u}(t) = \frac{u(t+h) - u(t)}{h} \rightsquigarrow u(t+h) = u(t) + hf(t, u(t))$$

and so for $t = kh$ we have $u_{h,k} = u_{h,k-1} + f((k-1)h, u_{h,k-1})$ where $u_h(t)$ is continuous, bounded with slope being the value of f at some point in the box, which is bounded. Therefore there exists a convergent subsequence, which concludes (ii).

As for (iii),j we check whether

$$u(t) - u(0) = \int_0^t f(s, u(s))ds$$

holds. To get the result we approximate for a decreasing sequence of h. First we have

$$u_h(t) = u_h(0) + \int_0^t \dot{u}_h(s)ds = u(0) + \int_0^t f(s, u_h(s))ds + \int_0^t [\dot{(u_h)}(s) - f(s, u_h(s))]ds$$

where

$$\int_0^t [\dot{(u_h)}(s) - f(s, u_h(s))]ds \to 0$$

as $h \to 0$ since we can bound the difference of f, and then times h, which goes to 0.

The other term

$$\int_0^t f(s, u_h(s))ds \to \int_0^t f(s, u(s))ds$$

because $u_h \to u$ in the infinite norm, which means that the convergent is uniform, so we can easily exchange limits.

So by taking the limit as $h \to 0$ we get

$$u(t) = u(0) + \int_0^t f(s, u(s))ds + 0$$

which proves the checking.

Moreover, by FTC it's easy to see that the solution is C^1 since it's the integral of a continuous function.

Now let's ask whether the solution is unique. A counter example is that $u(t) = t^2$ for $t > 0$. The corresponding ODE is

$$\begin{cases} \dot{u} = 2\sqrt{u} \\ u(0) = 0 \end{cases} \quad \Rightarrow u \equiv 0 \quad \text{or} \quad u = t^2.$$

So really it's not unique. The reason behind is what to choose as the derivative at $t = 0$.

Now we prove the uniqueness of solution under our imposed Lipschitz condition.

If u, v both satisfy

$$\begin{cases} \dot{u} = f(t, u(t)) \\ u(0) = u_0 \end{cases}$$

we want to show that $w := u - v = 0$. To show this we note that w satisfies the equation

$$\begin{cases} \dot{w} = f(t, u(t)) - f(t, v(t)) \\ w(0) = 0 \end{cases}$$

If f is linear on the second argument we are done since the derivative is 0. But it might not be. So we deduce

$$w(t) = \int_0^t (f(s, u(s)) - f(s, v(s)))\, ds$$
$$\Rightarrow |w(t)| = \int_0^t |f(s, u(s)) - f(s, v(s))|\, ds \leq M \int_0^t |w(s)| ds$$

which by Gronwall's theorem actually implies that $w \equiv 0$.

A non-theormized Gronwall's Lemma is the following: for $\varepsilon > 0$, we define

$$h_\varepsilon(t) = \varepsilon + \int_0^t |w(s)| ds > 0$$

then

$$\dot{h}_\varepsilon(t) = |w(t)| \leq M \int_0^t |w(s)| ds \leq M h_\varepsilon(t) > 0$$
$$\Rightarrow \frac{\dot{h}_\varepsilon(t)}{h_\varepsilon(t)} \leq M \ \Rightarrow \ \left(\log h_\varepsilon\right)' \leq M$$

which means

$$\log h_\varepsilon(t) - \log h_\varepsilon(0) \leq Mt$$

which finally gives

$$\log h + \varepsilon(t) \leq \varepsilon e^{Mt}$$

so

$$|w(t)| \leq \varepsilon M e^{Mt}.$$

Note that what we did above is nothing but to construct manually a non-negative denominator for the log derivative expression.

5. 1/18: CONTRACTION MAPPING THEOREM; BANACH SPACES

5.1. Contraction mapping theorem.

Def 5.1. *For (X, d) complete, we say $T : X \to X$ is a contraction mapping if $\exists 0 \le c < 1$ such that $d(T(x), T(y)) \le cd(x, y)$ for any $x, y \in X$.*

One way that might help visualize this is the fact that

$$T(B_r(x)) \subset B_{cr}(T(x)).$$

Theorem 5.2. *(Banach Contraction theorem) Let $T : X \to X$ be a contraction mapping where X is complete. Then $T(x) = x$ admits a unique solution in X.*

Proof. Find any $x_0 \in X$, and iterate it by $x_{n+1} = T(x_n)$. Now we have that

$$d(x_{n+1}, x_n) \le cd(x_n, x_{n-1}) \le \cdots \le c^n d(x_0, x_1)$$

$$\Rightarrow d(x_{n+m}, x_n) \le d(x_{n+m}, x_{n+m-1}) + \cdots + d(x_{n+1}, x_n) \le c^n \frac{d(x_0, x_1)}{1-c}$$

which means that x_n is Cauchy, i.e. it has a fixed point. But that fixed point by definition satisfies that it is a fixed point.

The fact that it is unique is due to that, if it's not unique then $T(x) = x$ and $T(y) = y$ for $x \neq y$. But this implies

$$d(T(x), T(y)) \le cd(x, y) = cd(T(x), T(y))$$

which is a contradiction. Thus the fixed point is unique. □

Now we see 2 applications of the theorem.

Application 1:

The second order Fredholm function:

$$f(x) = g(x) + \int_a^b k(x, y) f(y) dy$$

which is nothing but

$$f = g + Kf$$

for any linear operator K. We can thus view K as a matrix in the function space since it has exactly squared dimensions. The condition that K needs to satisfy is that it is continuous in x and integrable in y. Given which we have

$$d(T(f), T(h)) = \sup_x ||T(f) - T(h)|| = \sup_x \left| \int_a^b k(x, y)(f(y) - h(y)) dy \right|$$

$$\le \left(\sup_x \int_a^b |k(x, y)| dy \right) ||f - h||_\infty$$

for which if we assume $\left(\sup_x \int_a^b |k(x,y)|dy\right) = c < 1$ then we know that there's a unique solution, since each time we get a continuous function.

Now we can see this as an example in perturbation theory since we can reform the question in this setting: Suppose we have

$$(A + B)f = h$$

then it implies

$$(I + A^{-1}B)f = A^{-1}h = g \ \Rightarrow \ f + A^{-1}Bf = g$$

which is our original function.

Illustrating this in another way, let's write it as $f = g + Kf$, which means $f = (I - k)^{-1}g$ and that the matrix has an inverse because K is small (if we want to use the setting above, let B be small would suffice). Now by geometric expansion we have

$$f = (I - k)^{-1}g = \sum_{n=0}^{\infty} K^n g$$

so we can just iterate with the starting point g to get the unique solution. Note that this gives us a way to proceed other than an existential claim.

Application 2:

For the ODE system

$$\begin{cases} \dot{u}(t) = f(t, u(t)) \\ u(0) = u_0 \end{cases}$$

for $t > 0$. We have yet another way to prove the existence of a unique solution directly by Banach contraction theorem.

Theorem 5.3. *(Picard-Lindelöf/ Cauchy-Lipschitz) If f is Lipschitz with respect to u, then there is a unique solution to the above system of ODE.*

Proof. The idea is nothing but to extend the above method to a non-linear case.

Let

$$u(t) = u(t_0) + \int_{t_0}^{t} f(s, u(s))ds =: T(u)(t).$$

Then we try to show that T is a contraction on $C(I_\delta)$ where $I_\delta = [t_0 - \delta, t_0 + \delta]$. Since the expression of T is the integral of a continuous function, it is C^1, so it's checked that it's a self map.

Now we have

$$||T(u) - T(v)||_\infty = \sup_{t \in I_\delta} \left| \int_{t_0}^{t} [f(s, u(s)) - f(s, v(s))] \, ds \right| \leq \delta L ||u - v||_\infty$$

since

$$\leq L \cdot |u(s) - v(s)| \leq L||u - v||_\infty.$$

□

5.2. **Banach Space.** A Banach Space is a complete normed vector space.

Example 5.4.

- $C(K, ||\cdot||_\infty)$ is a Banach space.
- $C^k(K, ||\cdot||_\infty)$ is **NOT** complete. But we can find a norm that makes the space complete:

$$||f||_{k,\infty} := \sum_{j=0}^{k} ||f^{(j)}||_\infty \sim \sup_{0\leq j\leq k} ||f^(j)||_\infty \sim \left(\sum_{j=0}^{k} ||f^{(j)}||_\infty^p\right)^{\frac{1}{p}}$$

  (all above norms are equivalent) then $C^k(K, ||\cdot||_{k,\infty})$ is complete.
- Note that $C^\infty := \bigcap_{k\geq 0} C^k$ is not a Banach space under any norm! But we can define the Frechet topology under which it is complete with some metric on it. check!
- L^p spaces: all the function such that

$$||f||_p := \left(\int_K |f|^p dx\right)^{\frac{1}{p}} < \infty.$$

- $W^{m,p}$: Sobolev spaces: functions such that f^j are L^p for $0 \leq j \leq m$.
- $H^m := W^{m,2}$.

We do not prove that L^p is Banach yet since we need some technical measure theory there. But we can prove for l^p, the space of infinite sequences for $1 \leq p < \infty$. (When $p = \infty$ it is the maximal element of the sequence, which we know is not complete.) This has the almost exactly same difficulties modulo measurability.

Proposition 5.5. *The space $\left(l^p(\mathbb{N}), ||\cdot||_p\right)$ is Banach where the elements in the space are infinite sequences and*

$$||x||_p = \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}}.$$

We need to check first that the norm is actually a norm, for which all property are easy except the triangle inequality, and to prove which we use the following inequalities. We state them here but prove them on Friday. First we define

$$p' = \frac{p}{1-p}, \quad \text{i.e. } \frac{1}{p} + \frac{1}{p'} = 1.$$

- (Minkowski's inequality, just trig in the setting of L^p): $||x + y||_p \leq ||x||_p + ||y||_p$.

- (Young's inequality): For $a, b \geq 0$

$$ab \leq \frac{1}{p}a^p + \frac{1}{p'}b^{p'}.$$

- (Hölder's inequality):

$$||ab||_1 = \sum_{n=1}^{\infty} a_n b_n \leq \left(\sum_{n=1}^{\infty} a_n^p\right)^{\frac{1}{p}} \left(\sum_{n=1}^{\infty} b_n^p\right)^{\frac{1}{p}}.$$

6. 1/20: BANACH SPACES, OPERATOR SPACES

We first finish the claim from last time.

Proof. $||\cdot||_p$ is a norm on l^p:

- (Young's inequality): For $a, b \geq 0$
$$ab \leq \frac{1}{p}a^p + \frac{1}{p'}b^{p'}.$$
The proof of this uses the concavity of log:
$$\log\left(\frac{1}{p}a^p + \frac{1}{p'}b^{p'}\right) \geq \frac{\log a^p}{p} + \frac{\log b^{p'}}{p'} = \log(ab)$$
and since log is increasing, we are done.
- (Hölder's inequality):
$$||ab||_1 = \sum_{n=1}^{\infty} a_n b_n \leq \left(\sum_{n=1}^{\infty} a_n^p\right)^{\frac{1}{p}} \left(\sum_{n=1}^{\infty} b_n^p\right)^{\frac{1}{p}}.$$
The proof is the following:
$$\sum_{n=1}^{\infty} a_n b_n = \sum_{n=1}^{\infty} a_n \lambda^{1/p} b_n \lambda^{-1/p} \leq \sum_{n=1}^{\infty}\left(\frac{\lambda a^p}{p} + \frac{\lambda^{-p'/p} b^{p'}}{p'}\right) = \frac{\lambda ||a||_p^p}{p} + \frac{\lambda^{\frac{1}{1-p}}||b||_{p'}^{p'}}{p'}$$
where we minimize over λ to pick (or simply just pick) $\lambda = \dfrac{||b||_{p'}}{||a||_p^{p-1}}$ to get
$$||ab||_1 \leq ||a||_p ||b||_{p'}\left(\frac{1}{p} + \frac{1}{p'}\right) = ||a||_p ||b||_{p'}$$
where we deal with that p' power of $||b||_{p'}$ using $p' = \dfrac{1}{p-1}$.
- (Minkowski's inequality, just trig in the setting of L^p): $||x+y||_p \leq ||x||_p + ||y||_p$.
  The reason is that

$$\begin{aligned}
||x+y||_p^p &= \sum_{n=1}^{\infty} |x_n + y_n|^{p-1}|x_n + y_n| \leq \sum_{n=1}^{\infty} |x_n + y_n|^{p-1}(|x_n| + |y_n|) \\
\text{(Hölder)} &\leq \left(\sum_{n=1}^{\infty} |x_n + y_n|^{p'(p-1)}\right)^{\frac{1}{p'}} \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}} + \left(\sum_{n=1}^{\infty} |x_n + y_n|^{p'(p-1)}\right)^{\frac{1}{p'}} \left(\sum_{n=1}^{\infty} |y_n|^p\right)^{\frac{1}{p}} \\
&\leq ||x+y||_p^{p-1} \cdot \left(||x||_p + ||y||_p\right)
\end{aligned}$$

which gives us the result whether or not $||x+y||_p = 0$.

Hence $||\cdot||_p$ is indeed a norm.

$(l^p, ||\cdot||_p)$ is complete:

To prove completeness, we first find any $x^{(k)}$ Cauchy in l^p, i.e. $||x^k - x^l||_p \to 0$ as $k, l \to \infty$. Still, we first find a candidate of the limit, then prove that it is the limit indeed.

$$||x^k - x^l||_p \to 0 \Rightarrow \sum_{n\geq 1} |x_n^k - x_n^l|^p \to 0$$

which then means that for any n, as long as k, l are large enough

$$|x_n^k - x_n^l| \to 0$$

and since $x_n^1, x_n^2, \dots$ is a sequence Cauchy in C, the limit exists and we can denote it as $x_n^k \to y_n$.

Now what's left to prove is $y_n \in l^p$ and $x^k \to y$ in p-norm.

To show $y \in l^p$, we note that

$$\sum_{n=1}^{N} |y_n|^p = \sum_{n=1}^{N} \lim_{k\to\infty} |x_n^k|^p \leq \limsup_{k\to\infty} ||x^k||_p^p \leq C$$

since $x^{(k)}$ Cauchy implies it is uniformly bounded. Also since the above works for any N, we can say that when $N \to \infty$, the series on the left is bounded by C still, thus $y \in l^p$.

Now, to prove that $x^k \to y$ in p-norm we derive

$$\begin{aligned}
||y - x^k||_p^p &= \sum_{n=1}^{\infty} \lim_{l\to\infty} |x_n^l - x_n^k|^p \leq \frac{\varepsilon}{2} + \sum_{n=1}^{N(\varepsilon)} \lim_{l\to\infty} |x_n^l - x_n^k|^p \\
&\leq \frac{\varepsilon}{2} + \limsup_{l\to\infty} \sum_{n=1}^{N(\varepsilon)} |x_n^l - x_n^k|^p \leq \frac{\varepsilon}{2} + \limsup_{l\to\infty} ||x_n^l - x_n^k||_p^p \leq \varepsilon
\end{aligned}$$

where the first inequality is because the series is defined, thus the tail decays and the last inequality is valid when $k \geq K(\varepsilon)$. Hence we are done. □

Moreover, we have a proof of the generalized Young's inequality:

Generalized Young's inequality:

Proof. First note that

$$\left(b^{\frac{p_2 p_3}{p_2+p_3}}\right)^{\frac{p_2+p_3}{p_3}} = b^{p_2}$$

and by Young's inequality we have

$$b^{\frac{p_2 p_3}{p_2+p_3}} c^{\frac{p_2 p_3}{p_2+p_3}} \leq \frac{p_3}{p_2 + p_3} b^{p_2} + \frac{p_2}{p_2 + p_3} c^{p_3}.$$

Thus we have by using Young's twice that

$$a(bc) \leq \frac{a^{p_1}}{p_1} + \left(\frac{1}{p_2} + \frac{1}{p_3}\right)(bc)^{\frac{p_2 p_3}{p_2+p_3}} \leq \frac{a^{p_1}}{p_1} + \frac{p_2+p_3}{p_2 p_3}\frac{p_3}{p_2+p_3} b^{p_2} + \frac{p_2+p_3}{p_2 p_3}\frac{p_2}{p_2+p_3} c^{p_3}$$
$$= \frac{a^{p_1}}{p_1} + \frac{b^{p_2}}{p_2} + \frac{c^{p_3}}{p_3}.$$

□

If there's more terms, just tackle with them one by one and we are done.

Thus we have

$$\prod_{i=1}^{n} a_i \leq \sum_{i=1}^{n} \frac{a_i^{p_i}}{p_i}.$$

As is said last time, L^p space is almost exactly the same "modulo measure."

6.1. **Bounded Linear Operators.**

Now we switch gear and look at operators: a map from space of functions X to space of functions Y.

Def 6.1. *$T : X \to Y$ is an operator if X, Y are spaces of functions.*

- *T is linear $\iff \forall \lambda, \mu \in \mathbb{R}\ (C), \forall f, g \in X$*
$$T(\lambda f + \mu g) = \lambda T(f) + \mu T(g).$$
- *T is 1-1, onto, bijective is defined as usual.*
  Note that when T is bijective, we can define $T^{-1} : Y \to X$ with

  Theorem 6.2. *T^{-1} is linear if T is linear.*

  Proof. To be filled... □
- *T is bounded if there exists $M > 0$ such that*
$$||Tx||_y \leq M||x||_X$$
  for all $x \in X$.

Theorem 6.3. *The space of linear operator is a vector space; the space of bounded linear operator is also a vector space.*

We'll see that the latter is not only a vector space, but a normed one.

Def 6.4. *The operator norm is defined as*

$$||T|| = \inf\{M \,\big|\, ||Tx|| \leq M||x||\} = \sup_{x\neq 0}\frac{||Tx||_Y}{||x||_X} = \sup_{||x||_X=1} ||Tx||_Y.$$

Note that this norm gives us a very handy inequality:

$$||Tx||_Y \leq ||T|| \cdot ||x||_X.$$

Moreover, the space forms even an algebra if we define multiplication as the composition of functions, given the suitable domain/ranges.

Example 6.5.

Let $X = (C[0,1], ||\cdot||_\infty)$ and $T : f \mapsto \int_0^x f(t)dt$. Then we know that (since we are on the space $C[0,1]$)

$$||Tf||_\infty = \sup_{x\in[0,1]} \left|\int_0^x f(t)dt\right| \leq \sup_{x\in[0,1]} \int_0^x ||f||_\infty dt \leq \sup_{x\in[0,1]} \int_0^x ||f||_\infty \leq ||f||_\infty$$

and thus $||T|| \leq 1$. Now for the other direction we just need to find one instance that makes the inequality an equality: let $f = 1$ then we are done.

Example 6.6. *The following 2 example shows that the boundedness of an operator is also dependent on the space it acts on.*

Consider the derivative operator: $T : f \mapsto f'$. We will give a space on which it is not bounded and one on which it is.

Let $C^\infty = \bigcap_{k\geq 0} C^k[0,1]$ and $X = (C[0,1], ||\cdot||_\infty)$. Then $f : X \to X$. Yet now T is not bounded since we can simply pick $f = e^{nx}$ then $Tf = ne^{nx}$ and hence $||T|| = n$ is unbounded.

Yet on the space $X = (C^1[0,1], ||\cdot||_{1,\infty})$, $Y = (C^0[0,1], ||\cdot||_\infty)$, we know that $T : X \to Y$ and now the norm is $||f||_{1,\infty} = ||f||_\infty + ||f'||_\infty$, so the operator norm is

$$||T|| = \frac{||f'||_\infty}{||f||_\infty + ||f'||_\infty < 1}$$

is bounded.

Example 6.7. *Also, if X and Y are finite dimension spaces, then every operator has a matrix expression form. The construction is by basis map.*

Theorem 6.8. *For linear T, we have*

$$T \text{ is bounded} \iff T \text{ is continuous} \iff T \text{ is continuous at } 0.$$

Proof.

Bounded $\Rightarrow$ continuous:

Since the definition of boundedness of T implies that T is Lipschitz, T is continuous.

Continuous $\Rightarrow$ continuous at 0:

This is nothing but definition.

Continuous at 0 $\Rightarrow$ bounded:

If T is continuous at 0, then $\forall \varepsilon, \exists \delta$ such that $||x||_X \leq \delta \Rightarrow ||Tx||_Y \leq \varepsilon$. So we take $\varepsilon = 1$, $\delta = \delta(1)$, and thus any $0 \neq z \in X$ we know that

$$\left|\left|\frac{\delta z}{||z||_x}\right|\right|_X \leq \delta \Rightarrow \left|\left|T\frac{\delta z}{||z||_x}\right|\right|_Y \leq 1$$

which means $||T|| \leq \dfrac{1}{\delta}$ is bounded.

□

Theorem 6.9. *(Bounded Linear Operator, BLT) Let $T : M \to Y$ where M is dense in X and Y is Banach be bounded on M. Then there exists a unique extension $\bar{T} : X \to Y$ such that $\bar{T}x = Tx$ for $x \in M$ and $||\bar{T}|| = ||T||$.*

Remark: this can be used to increase smoothness of a function by limiting on a dense set, not the whole space.

Proof. (Sketch)

What we do is pretty standard, we find a convergent sequence to each point in X and use the completeness of Y to get the limit of the operated sequence in Y. Then we check $||\bar{T}|| = ||T||$ and uniqueness under this condition. □

Theorem 6.10. *(Open mapping theorem) If $X \to Y$ is 1-1, onto, bounded, and linear, then for X, Y Banach, we have that $T^{-1} : Y \to X$ is also bounded.*

We will not prove this result since that's hard (since we really don't have a control over $||T^{-1}||$), and it's not in the textbook. However, we will state a different version of the theorem which will explain why the theorem is called "open mapping theorem."

The above is sometimes called the bounded inverse theorem.

Lemma 6.11. *The above formulation is equivalent to this: if $T : X \to Y$ is onto, bounded, and linear, then for X, Y Banach, then $\exists c > 0$ such that $T(B_x(0,1)) \supset B_y(0,c)$. That is (since maps neighborhood to open neighborhoods), T maps open sets to open sets.*

These formulation are equivalent since T^{-1} continuous $\iff$ T^{-1} bounded. Or we can use an argument to find $\frac{1}{c}$ for the boundedness of T^{-1} to go from below to above.

Now we show corollaries of open mapping theorem.

Corollary 6.12. *If X is Banach for 2 different norms, $||\cdot||_1$ and $||\cdot||_2$, then if $||\cdot||_1 \leq c||\cdot||_2$ for some c, then we know that there exists c' such that $||\cdot||_2 \leq c'||\cdot||_1$.*

Proof. Just use the identity map $Id : X \to X$. Since $Id : (X, ||\cdot||_2) \to (X, ||\cdot||_1)$ is bounded linear bijection, so is it's inverse. In particular $Id^{-1} = Id$ and hence there exists c' such that $||\cdot||_2 \leq c'||\cdot||_1$. □

Corollary 6.13. *Assume that* $(X, ||\cdot||_1)$ *and* $(X, ||\cdot||_2)$ *has the same Cauchy limits (i.e. Cauchy sequences converge to the same limits with both norm). Then* $||\cdot||_1 \sim ||\cdot||_2$.

Proof. We first show that $(X, ||\cdot||_1 + ||\cdot||_2)$ is Banach. This is because if a sequence is Cauchy for the sum norm, it is Cauchy for both individual norms, and thus we can find the limit for individual norms. But those limits are equal, which means that the limit under the sum norm is also equal.

Now we use corrollary 6.7 to show that, since $||\cdot||_1 \leq ||\cdot||_1 + ||\cdot||_2$, there exists c such that

$$||\cdot||_1 + ||\cdot||_2 \leq c||\cdot||_1$$

for which we use again to get

$$||\cdot||_2 \leq (c-1)||\cdot||_1$$

which by above shows that they are equal. □

Note that if $(X, ||\cdot||_1)$ and $(X, ||\cdot||_2)$ have different Cauchy limits, then counter-examples do exist.

Now we introduce a very important theorem not in book that is basically the same as open mapping theorem.

Def 6.14. *For* $T : X \to Y$ *linear and* X, Y *Banach, we define the* *graph* *as*

$$\text{Graph}(T) := \bigcup_{x \in X} [x, Tx] \in X \times Y.$$

Def 6.15. *A linear operator* $T : X \to Y$ *is said to be closed if* Graph(T) *is closed, i.e.*

$$(x_n, Tx_n) \to (x, y) \quad \Rightarrow \quad y = Tx.$$

Def 6.16. *We define the* *graph norm* *as*

$$||\,\text{Graph}(T)|| = ||\cdot||_X + ||T\cdot||_Y.$$

Theorem 6.17. *(Closed graph theorem)*

$$T \text{ is closed } \Rightarrow T \text{ is bounded.}$$

Proof. We first show that if T is closed then $(X, ||\,\text{Graph}(T)||)$ is Banach. (exercise).

Now we know that, by previous corollary,

$$||\cdot||_X + ||T\cdot||_Y \leq C||\cdot||_x$$

since they are equivalent, but then moving the term we get that T is bounded. □

Note that if T is continuous, then obviously it is closed. Thus

$$T \text{ is bounded} \iff T \text{ is continuous} \iff T \text{ is closed.}$$

The closed graph theorem is useful to deal with some abstractly defined functions.

Now we inspect the kernel and range of linear operators.

Def 6.18. *If $T : X \to Y$ is linear, then*

$$\ker(T) \equiv N(T) = \{x \in X, Tx = 0\}$$

$$\text{Range}(T) \equiv R(T) = \{y \in Y | \exists x \in X, y = Tx\} = \bigcup_{x \in X} Tx.$$

Proposition 6.19.

- $\ker(T)$ *and* $\text{Range}(T)$ *are linear subspaces of X and Y respectively.*
- *If X, Y are normed vector spaces and T is bounded, then* $\ker(T)$ *is closed.*
  Reason:
  $$x_n \to x, Tx_n = 0 \Rightarrow Tx = 0$$
  by continuity.
- *T is 1-1 iff* $\ker(T) = \{0\}$.
- ***Kernel is never empty!***
- *T is onto iff* $\text{Range}(T) = Y$.

Def 6.20. *Dimension of* $\ker(T)$ *is called the nullity of T and the dimension of* $\text{range}(T)$ *is called the rank of T.*

Example 6.21. *Let $X = (C[0,1], ||\cdot||_\infty)$ and $T : f \mapsto \int_0^x f(t)dt$.*

Then

$$\text{Range}(T) = \{f \in C^1[0,1] | f(0) = 0\}.$$

But note that the closure (w.r.t the infinite norm) is

$$\overline{\text{Range}(T)} = \{f \in C^0[0,1] | f(0) = 0\} \supsetneq \text{Range}(T)$$

since C^1 functions pass over to C^0 functions.

Now if we let $K = I + T$, i.e. $Kf(x) = f(x) + \int_0^x f(t)dt$.

But now we can construct an inverse of K:

$$T^{-1}f = g \iff (f + Tf)(x) = f(x) + \int_0^x f(t)dt = g(x)$$

note that $Tf'(x) = f(x)$, we have the following ODE:

$$Tf' + Tf = g, Tf(0) = 0$$

with solution

$$Tf = \int_0^x e^{-x-y} g(y)dy$$

which by differentiation gives

$$f(x) = g(x) - \int_0^x e^{-x-y} g(y) dy =: K^{-1}(g)(x).$$

Now that we have an inverse the range of K is X (closed) since we have constructed an inverse.

The reason of this surprising difference is that, for T we are moving form X to a subspace that's smoother, but by adding the identity we don't smooth anything anymore.

Proposition 6.22. *For $T : X \to Y$ bounded, X, Y Banach. Then*

$$\exists c > 0 \ s.t. \ c||x|| \leq ||Tx||, \forall x \in X \iff \text{Ran}(T) \textit{ is closed and } \ker(T) = \{0\}.$$

This intuitively means that $||T||$ is also bounded from below, which is morally the same thing as T^{-1} is bounded. This is also called the method of a priori estimate: if $||T||$ is bounded from below, then there is a unique solution to $Tx = y$.

Proof. ($\Rightarrow$:) $\ker(T) = 0$ since $Tx = 0 \Rightarrow x = 0$ by the bound from below.

Let $y_n \in \text{Ran}(T)$ be that $y_n \to y$, so we only need to show $y = Tx$ for some x to show that closedness of Ran(T). Since $y_n \in \text{Ran}(T)$ we have

$$||x_n - x_m|| \leq \frac{1}{c}||Tx_n - Tx_m|| = \frac{1}{c}||y_n - y_m|| \to 0$$

and thus $\exists x$ with $x_n \to x$ in X. But due to continuity of T we pass the limit so that $Tx = y$.

($\Leftarrow$:)

We know that T is 1-1 and onto, thus T^{-1} is defined. Since Ran(T) is closed we know that Ran(T) is a Banach space for $|| \cdot ||_Y$. This is because Y is Banach, and a closed subset of a complete space is complete. Thus, by open mapping we have that T^{-1} is bounded, but $||T^{-1}|| \leq \frac{1}{c}$ is equivalent to $||T||$ bounded form below. Hence we are done. $\square$

7. 1/23: Compact operators; Strong convergence

We start with a proof that we've seen many times.

Theorem 7.1. *Let X be a nomred vector space (not necessarily Banach), and Y Banach, then $(\mathcal{B}(X,Y), ||\cdot||)$ is Banach.*

Proof. We follow the same old path of first finding a possible limit, then show that the limit is in the space, and that it really is a limit.

Assume that T_n is Cauchy, which means that $||T_n - T_m|| \to 0$ for m, n large. But this means that for all $x \in X$ fixed, $||T_n x - T_m x||_Y \to 0$ (note that it's not uniform in x here). But this means that the sequence $T_n x$ is Cauchy in Y, and because Y is complete, there exists a limit y, and we simply define $Tx = y$, where we are using T as the limit of T_n. It's an exercise to show that T is linear.

The other things we need to show is that T is bounded and that it is indeed a limit, i.e. $||T - T_n|| \to 0$.

Now, fix x, for all $\forall\varepsilon$, we know that $\exists M = M(x, \varepsilon)$ such that $\forall m \geq M$

$$||Tx - T_m x|| \leq \frac{\varepsilon}{2}||x||$$

because when we defined the limit we fixed x too i.e. our definition says: for all x fixed, then $||x||$ is fixed, we know $\dfrac{||T_n x - Tx||_Y}{||x||_X} \to \dfrac{0}{||x||_X} = 0$. Thus we have the above inequality. But that inequality immediately implies

$$||T_n x - Tx|| \leq ||T_n x - T_m x|| + ||T_m x - Tx|| \leq ||T_n - T_m|| \cdot ||x|| + \frac{\varepsilon}{2}||x|| \leq \varepsilon ||x||$$

since $||T_n x - T_m x|| \to 0$ for large enough n, m. Here we first choose n large enough to satisfy the Cauchy property of T_n, then find a yet larger m based on choice of x and ε to bound the second part. Yet the m is just intermediate and hence what we get really is that $||T_n x - Tx|| \leq \varepsilon ||x||$ for n independent of x (m is though), and that ε is by construction independent of x.

This means that $T_n \to T$ in the norm sense.

To show that T is bounded we simply observe

$$||Tx|| \leq ||(T - T_n)x|| + ||T_n x|| \leq \varepsilon ||x|| + c \cdot ||x|| \leq \tilde{c}||x||$$

where we used that T_n is bounded for any n.

Hence we are done with the proof. □

7.1. Compact operators.

Def 7.2. *Compact operators: $T : X \to Y$ is compact if $T(B)$ is pre-compact in Y, where B is any bounded subset of X. (close the set then it's compact.)*

Equivalently, T is compact iff for each sequence $(x_n) \subset X$ with $||x_n||_X \leq C$ (in some ball), then there is a subsequence $x_{\phi(n)}$ such that $T(x_{\phi(n)})$ converges in Y. Note that in this form Y has to be complete.

We now denote the space of compact operators by $K(X,Y)$. Note that $K(X,Y) \subset \mathcal{B}(X,Y)$. The following are some properties that describes this space.

Proposition 7.3.

- *$K(X,Y)$ is a closed subspace of $\mathcal{B}(X,Y)$.*
- *$\lambda T + \mu S$ is compact for T, S comopact.*
- *If $T_n \to T$ uniformly and T_n are compact, then so is T.*
- $\dim \mathrm{Ran}(T) < \infty$ *means that T is compact.*
- *For $S \in K(\cdot,\cdot)$ and $T \in B(\cdot,\cdot)$, then TS and ST are compact, given that the function is well defined (fill out the domain space and image space).*

We only prove the third property, which tells us that compactness is passed by uniformly convergence.

Proof.

T_n compact, $||T - T_n|| \to 0$, then T is compact:

Since we want to show that T is compact, we'll start with a sequence with $||x_m|| \leq 1$ and show that there is a convergent subsequence of Tx_n.

We use a diagonal argument to do it: Since all T_i are compact, for each one we can find a convergent subsequence in the image for any sequence.

So for T_1 we find $T_2 x_{\phi_1(m)} \to y_1$; Now we use the new sequence $x_{\phi_1(m)}$ to apply to T_2 and find $T_2 x_{\phi_2 \circ \phi_1(m)} \to y_2$.

Following this pattern we have that for every $j \in \mathbb{N}$, there exists a subsequence such that

$$T_j x_{\phi_j \circ \ldots \circ \phi_1(m)} \to y_j$$

and for any $k \leq j$ we inherit that

$$T_k x_{\phi_k \circ \ldots \circ \phi_1(m)} \to y_k.$$

And of course we pick the diagonal sequence to gain

$$z_j = x_{\phi_k \circ \ldots \circ \phi_1(j)}$$

which means that for any k, $T_k z_j \to y_k$ as $j \to \infty$. Thus

$$\begin{aligned} T(z_m - z_n) &= (T - T_p)(z_m - z_n) + T_p(z_m - z_n) \\ &\leq ||T - T_p|| \cdot ||z_m - z_n|| + (T_p x_m - T_p x_n) \to 0 \cdot ||z_m - z_n|| + (y_p - y_p) \end{aligned}$$

for m, n are large enough. Note that since $||x_m|| \le 1$, $||z_m - z_n|| \le 2$ and thus Tz_m is Cauchy. Since Y has to be Banach for $\mathcal{B}(X,Y)$ even to be Banach, thus Tz_m converges in Y, and we are done. $\square$

7.2. **Strong Convergence.**

We some times denote Strong convergence topology by SoT, which is weird...And we'll note that strong convergence is actually weaker than uniform convergence.

Def 7.4. *Strong convergence: T_n converges strongly to T if $||T_n x \to Tx||_Y \to 0$ for all $x \in X$.*

Note that
$$||T_n - T|| \to 0 \Rightarrow ||(T_n - T)x|| \le ||T_n - T|| \cdot ||x|| \to 0$$
and hence uniform convergence implies strong convergence. The converse, however, is not true. Counter examples are in infinite dimension. In finite dimension they are the same definition.

Theorem 7.5. *(Uniform boundedness theorem/Banach-Steinhaus)*

For X, Y Banach, $(T_i)_{i \in I} \in \mathcal{B}(X,Y)$, assume that
$$\sup_{i \in I} ||T_i x|| < \infty$$
for all $x \in X$. Then $\exists c > 0$ (uniform in x) such that $||T_i x||_Y \le c||x||_X$, $\forall x \in X$, $\forall i \in I$.

Corollary 7.6. *Important For X, Y Banach, if $T_n \in \mathcal{B}(X,Y)$ with $T_n \to T$ strongly, then*
$$\sup_n ||T_n|| < \infty$$
and $T \in \mathcal{B}(X,Y)$ with
$$||T|| \le \liminf_{n \to \infty} ||T_n||.$$

Now we look at applications of this.

Application: A scheme is consistent and stable iff it is convergent.

Application of convergence to continuous groups:

Let $A : X \to X$ be bounded and define
$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$
which converges since as A bounded the nominator is at most c^k, yet $\sum \frac{c^k}{k!}$ converges (less than geometric). Then automatically we have
$$||e^A|| \le e^{||A||}$$
by definition (also subtle point of $||A^k|| \le ||A||^k$).

For the differential question

$$\begin{cases} \dot{x} = Ax \\ x(0) = x_0 \end{cases}$$

the solution is of form

$$x(t) = e^{tAx_0}$$

and we define

$$T(t) = e^{tA}$$

as the solution operator. So we have the property

$$T(t+s) = e^{(t+s)A} = T(t)T(s)$$

where T is a <u>Uniformly/strongly continuous group</u> if

(i) $T(0) = Id$;
(ii) $T(s+t) = T(s)T(t)$;
(iii) $T(t) \to Id$ uniformly/strongly as $t \to 0$.

Where a continuous group is just a group that has continuous group operations.

Application of strong convergence: shifts.

Let $f \in C_0(\mathbb{R})$, the set of functions that are 0 at infinity. Then define operator

$$\tau_h : C_0 \to C_0$$

such that

$$\tau_h f(x) = f(x+h)$$

therefore we have

(i) $\tau(0) = Id$;
(ii) $\tau(a+b) = \tau(a)\tau(b)$;
(iii) $T(t) \to Id$ strongly as $t \to 0$ (but not uniformly).

8. 1/25: Dual Spaces; Hahn Banach Theorem

We continue with the continuous group last time. We start again with a definition from last time:

Def 8.1. *T is a Uniformly/strongly continuous group if*

(i) $T(0) = Id$;
(ii) $T(s+t) = T(s)T(t)$;
(iii) $T(t) \to Id$ uniformly/strongly as $t \to 0$.

We claim that the shift operator forms a strongly continuous group, but not a uniform one.

We see that not shifting is the identity map, and the second property is because

$$\tau(a+b)(f)(x) = f(x+a+b) = \tau_b f(x+a) = \tau_a \tau_b f(x).$$

Now we prove the third property. Remember that the space we're considering is C_0, the space of functions that goes to 0 at infinity. So what we want to prove is that for any fixed f

$$\sup_{x\in\mathbb{R}} |f(x+h) - f(x)| \to 0$$

as $h \to 0$, which we note is nothing but to show that f is uniformly continuous. How do we prove that each $f \in C_0$ is uniformly continuous? The method is first to discard the tail at infinity which goes to 0. Since $|f|$ will be small enough for x large, we know that for any difference in that part, we just bound the difference in $f(x)$ by the sup norm of f, which in the end gives the uniform continuous condition.

Now we can deal with the function on a bounded set (hence compact), which is automatically uniformly continuous.

As for why it is not uniform, we note that if we allow f to be any function, then we can have it reliant on h and let it be a sharp bump with width less than $\frac{h}{2}$, and we see that the support of the shifted function has no intersection with the original one, which means

$$||f_n - \tau_n f_n|| = 1.$$

One remark is that the operator

$$T(h) = \tau(h) = e^{h\frac{\partial}{\partial x}} = 1 + h\frac{\partial}{\partial x} + \dots$$

which is true (but not yet well defined, so heuristically) beacause

$$f(x+h) = f(x) + h\frac{\partial}{\partial x} f + \dots .$$

8.1. **Dual Space.**

One main idea we will exploit is the coordinates; $\lambda_i : x \to x_i(x) := \langle e_i, x \rangle$ for which we haven't even defined, but keep that in mind.

Def 8.2. *For a vector space X, the space of continuous linear functionals from $X \to \mathbb{R}$ is called the (topological) dual to X, denoted as X^*. In other words,*

$$X^* = \mathcal{B}(X, \mathbb{R}).$$

Now that X^* is a normed vector space, whose norm is exactly what we've defined before:

$$||\phi|| = \sup_{x \neq 0} \frac{|\phi(x)|}{||x||}.$$

Remember that it satisfies that $|\phi(x)| \leq ||\phi|| \cdot ||x||$ and of course $||\phi|| < \infty$. In addition, it is also Banach since $\mathbb{R}$ is complete.

Def 8.3. *A notation that we will often use is that*

$$\phi(x) = \langle \phi, X \rangle_{X^*, X}$$

which is the duality product. Note that it is not really an inner product, but just an expression.

This also means that dual of dual is Banach. (from which we know that L^1 is not the dual of L^∞). Now it is often the case where the dual is a larger space than the original one, except some useful examples like Hilbert space. It's also known that for finite dimensional spaces, the dual is of the same size.

The reason for the last statement above is that, in $\mathbb{R}^n$

$$\phi(x) = \phi\left(\sum_{i=1}^n x_i e_i\right) \overset{linear}{=} \sum_{i=1}^n x_i \phi(e_i) = \sum_{i=1}^n \phi_i \cdot x_i$$

where $\phi_i = \phi(e_i)$ is a decomposition of ϕ. But this immediately means that $(\mathbb{R}^n)^*$ is also n-dimentional since we've given a base decomposition of ϕ.

Now this is not true in general for infinite dimension. Some useful results are

- $(\mathbb{R}^n)^* = \mathbb{R}^n$.
- $L_2^* = L_2$ by Riez representation theorem.
- For $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, then $(L^p)^* = L^q$.

  Note that in this case, we somehow can figure out why it's even well-defined since by Holder, the operation $\langle \phi, x \rangle$ need to be integrable, and the fact that $\phi \in L^q$ and $x \in L^p$ validates that.

  Also note that L^1 and L^∞ are not the dual of one another.

- Let $X = C([a,b])$ with ρ smooth. Then we define
$$\phi_\rho(f) = \int_a^b \rho(x)f(x)dx \in \mathbb{R}$$
and get that, since ϕ_ρ is linear and bounded, it is in X^*.
- With settings of the last example, let $\rho = \delta_{x_0}$ be the dirac delta for $x_0 \in [a,b]$, then
$$\phi_{\delta_{x_0}}(f) = \int_a^b f(x)\delta_{x_0}dx = f(x_0).$$
where of course, we need to deal with distribution spaces, but what the ever.
- In general, $X^* = \mathcal{M}_b([a,b])$, the space of bounded measures on that interval.

Theorem 8.4. *(Hahn-Banach Theorem). Let the Y be a subspace of X, which is normed, and let $\psi : Y \to \mathbb{R}$ be bounded and linear such that $||\psi||_{Y^*} = M < \infty$.*

Then, there is an extension $\phi : X \to \mathbb{R}$ that is bounded and linear that extends ψ, i.e. $\phi|_Y = \psi$ and that
$$||\phi||_{X^*} = ||\psi||_{Y^*}.$$

We only prove how to extend to one more dimension. It's not hard from here to extend to any whole space, also infinity is obviously there to annoy us.

Let $y_1, y_2 \in Y$ and $X \notin Y$, then for convenience define $\tilde{\psi} = M^{-1}\psi$, which essentially gives us $||\tilde{\psi}|| = 1$. Now we have
$$\tilde{\psi(y_1)} - \tilde{\psi}(y_2) = \tilde{\psi}(y_1 - y_2) \leq ||y_1 - y_2|| \leq ||y_1 + x|| + ||y_2 + x||$$
$$\Rightarrow -\tilde{\psi}(y_2) - ||y_2 + x|| \leq -\tilde{\psi}(y_1) + ||y_1 + x||$$
where since both sides works for arbitrary y_1, y_2, x we know that
$$a = \sup_{y_2 \in Y}\left(-\tilde{\psi}(y_2) - ||y_2 + x||\right) \leq \inf_{y_1 \in Y}\left(-\tilde{\psi}(y_1) + ||y_1 + x||\right) = b$$
so we know that $a \leq b$.

Now we define $\tilde{\psi}(y + rx) = \tilde{\psi}(y) + rc$ and we can check that $-\tilde{\psi}$ is linear, and the norm is not larger. do it

We might not use Hahn Banach very often, but we do very often use the results of it. That's why it's so good.

Corollary 8.5. *$\forall x_0 \in X$, $\exists f_0 \in X^*$ such that*
$$||f_0||_{X^*} = ||x_0||_X \text{ and } \langle f_0, x_0\rangle_{X^*,X} = ||X_0||_X^2.$$

Proof. Let $y = \mathbb{R}x_0 := tx_0$ for all $t \in \mathbb{R}$ be the members of a subspace Y. Then let
$$\psi(tx_0) = t||x_0||^2.$$

Then by Hahn Banach we know that there exists f_0 extends ψ to X^*, and

$$||f_0||_{X^*} = ||\phi||_{Y^*} = \frac{|t| \cdot ||X_0||^2}{||tX||} = ||x_0||_X.$$

□

Corollary 8.6. *(Characterization of norm) Let $x \in X$, then*

$$||x|| = \max_{\substack{f \in X^* \\ ||f||=1}} |\langle f, x \rangle|.$$

Proof. The easy part is

$$|\langle f, x \rangle| \leq ||f|| \cdot ||x|| = ||x||.$$

And now we only need to show that the max is attained. But to do so we just take the f_0 constructed in the proof of Corollary 8.4, and take

$$f = \frac{1}{||x||} f_0$$

then it works (check). □

Corollary 8.7. *(Separability: X^* separates X) For $x, y \in X$, if $\phi(x) = \phi(y)$ for all $\phi \in X^*$, then $x = y$.*

Proof. We note that

$$\phi(x) - \phi(y) = \phi(x - y)$$

where if we take $\phi = f$ (as defined in the proof of corollary 8.6), we will get that

$$\phi(x - y) = \langle f, x - y \rangle = ||x - y||$$

which gives exactly the result. □

Again, we sum up some information of dual spaces.

Remark 8.8.

- X^* is a bounded operator space (also linear).
- X^{**} is also a $\mathcal{B}$ space.
- For $x \in X$, define $F_x \in X^{**}$ such that $F_x(\phi) = \phi(x)$ for $\phi \in X^*$. Then we see that x is a representation of a function in X^{**}.

  Note that in the case when $X^{**} \cong X$, then X is called reflexive.
- For $1 < p < \infty$,

$$(L^P)^{**} = (L^q)^* = L^p.$$

- $(L^1)^* = L^\infty$, $(L^\infty)^* \supset L^1$. Moreover, we sometimes want to use C to stand for something in between, i.e.

$$C \subset L^\infty \textit{ and } C^* = \mathcal{M}_b \supset L^1$$

  which intuitively means that L^1 is too small (indeed! it doesn't even have dirac delta), so we use C to shrink L^∞ a little and extend L^1 a little.

Now we define our last bit of text here:

Def 8.9. *For a sequence $x_n \in X$, we say that it convergent weakly to x if $\forall \phi \in X^*$*

$$\phi(x_n) \to \phi(x).$$

We know that strong convergence in X (the normal convergence in norm) implies weak convergence, and not the converse (never in infinite dimension) show since

$$|\phi(x_n) - \phi(x)| \leq ||\phi|| \cdot ||x_n - x||_X \to 0.$$

Def 8.10. *A sequence $\phi_n \in X^*$ convergent weak* if $\forall x \in X$*

$$\phi_n(x) \to \phi(x)$$

and we denote it as $\phi_n \rightharpoonup \phi$.

"Note that x is a function of ϕ here."

Also note that the weak* convergence is just the strong convergence of operators.

9. 1/30: More on compactness of operator spaces; Measure

We define the a few notions of topologies on X and X^*.

Def 9.1.

- *We define the strong topology on X be the topology that is generated by the norm.*
- *The weak topology $\mathcal{T}_w$ on X is such that for all $\phi \in X^*$, $\phi^{-1}(open) \in \mathcal{T}_w$. That is, the topology that makes all functionals in X^* continuous.*
- *The weak* topology on X^* is such that for all $x \in X$, if $x^{-1}(open)$ is open, where we can let x be the function $x : X^* \to \mathbb{R}$ with $x(\phi) = \phi(x)$. Denote this topology as $(X^*, \mathcal{T}_w^*)$.*

Now what is all this about? Well, the most important thing is that on the set X^*, the topology $\mathcal{T}_w^*$ is coarser than $\mathcal{T}_w$, which is then coarser than the norm topology.

Yet a coarser topology means lesser open sets, which means easier to be compact.

Theorem 9.2. *(Banach Alaoglu theorem) The closed unit ball (with respect to norm) in X^* is compact for $(X^*, \mathcal{T}_w^*)$, i.e. for any bounded sequence, there is a convergent subsequence in the weak* sense.*

An application of this is that for $f_n \in L^2(0,1)$ (or basically any reflexive space), since $X^{**} = X$ then weak* convergent is weak convergent, thus Banach Alaoglu applies.

Consider $f_n(x) = \cos(nx)$, with $f_n \overset{w}{\rightharpoonup} f$ i.e.

$$\forall g \in L^2(0,1), \quad \int_0^1 f_n \bar{g} \to \int_0^1 f \bar{g}$$

which could be understood as inner product with bases g.

Then what is f? The answer is that $f = 0$ since, heuristically, the oscillation cancels with each other, so it functions as if it is 0 when integrated against. Another such example is that we can say $f_n \to \frac{1}{2}$ since

$$f_n^2 = \cos^2 \approx \frac{\cos^2 + \sin^2}{2} = \frac{1}{2}$$

which is a little bit shaky argument, but we can justify. A lesson to learn here is that weak convergence loses energy (smoothen things). We can somehow see this since

$$\text{simple} \to \text{polynomial} \to \text{continuous} \to L^2$$

and for any indicator function g, f_n applied to g is nothing but 0 after taking the limit.

Theorem 9.3. *(Kakutani) If X is reflexive, i.e. $X^* = X$, then the unit ball is compact with respect to the weak topology.*

Note that this is nothing but our remark above that when X is reflexive, weak and weak* are the same thing.

Now we look again at the L^1, L^∞ example. We know that $L^\infty = (L^1)^*$, thus for all $f_n \in L^\infty$ with $||f_n|| \leq 1$, we have

$$f_{\phi(n)} \overset{w*}{\rightharpoonup} f \in L^\infty$$

which is the same thing as (just definition)

$$\int_0^1 f_{\phi(n)}(x)g(x)dx \to \int_0^1 f(x)g(x)dx, \quad \forall g \in L^2.$$

But then L^1 is just too small a space to be a dual. For instance let f_n be the bump at 0, then if it is a dual of L^∞, then δ_0 should be in L^1, which is not. So we need an intermediate space that let the below expression meaningful:

$$\int_{-1}^1 f_n(x)g(x)dx \to g(0)$$

note that not only the kronicer delta should be in, g must can be evaluated at 0. Thus it cannot be L^∞, but continuous will suffice. To be more explicit (though we'll not prove it here) the space of continuous functions satisfies

$$C^* = \mathcal{M}_b$$

i.e. the dual space is bounded measures, so the delta is one example. Now the above expression makes sense.

9.1. **Measure spaces:**

Def 9.4. *A σ-algebra on X is a collection $\mathcal{A}$ of subsets such that*

(1) $\emptyset \in \mathcal{A}$*;*
(2) $A \in \mathcal{A} \;\Rightarrow\; A^c \in \mathcal{A}$*;*
(3) $A_i \in \mathcal{A} \;\Rightarrow\; \bigcup_{i=1}^\infty A_i \in \mathcal{A}$.

Def 9.5. $(X, \mathcal{A})$ *is a measure space.*

Def 9.6. *The Borel σ-algebra is the σ-algebra that is generated by* $\mathcal{T}$.

Def 9.7. *A measure μ on X is a function $\mu : A \to [0, \infty]$ such that*

(a) $\mu(\emptyset) = 0$*;*
(b) $\{A_i\}$ *are disjoint, then*

$$\mu\left(\bigcup_{i=1}^\infty A_i\right) = \sum_{i=1}^\infty \mu(A_i).$$

Def 9.8. *A measure μ is*

- *finite if* $\mu(X) < \infty$;
- *σ-finite if* $X = \bigcup_{i=1}^{\infty} A_i$ *such that* $\mu(A_i) < \infty$.
- $(X, \mathcal{A}, \mu)$ *is a measure space.*

Theorem 9.9. *There exists a measure on* $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ *such that*

$$\lambda\left(\prod_{i=1}^{n}(a_i, b_i)\right) = \prod_{i=1}^{n}|b_i - a_i|$$

and is defined on all $\mathcal{B}(\mathbb{R}^n)$.

Another example is the measure that is the delta function, i.e.

$$\delta_{x_0}(A) = \begin{cases} 1 & x_0 \in A \\ 0 & x_0 \notin A \end{cases}.$$

Theorem 9.10. *For the Lebesgue measure, we have that it has some regularity:*

$$\lambda(A) = \inf\{\lambda(U), U \in \mathcal{O}_X, A \subset U\} = \sup\{\lambda(K), K \;\; compact, k \subset A\}.$$

Def 9.11. *A measure space is complete if every subset of null-measurable set is measurable, i.e.*

$$\mu(A) = 0, B \subset A \Rightarrow B \in A.$$

Note that $(X, \mathcal{A}, \lambda)$ is not complete since $\mathcal{B}$ is not big enough.

Theorem 9.12. *There exists an extension of any measure space* $(X, \mathcal{A}, \mu) \to (X, \tilde{\mathcal{A}}, \tilde{\mu})$ *that is complete.*

Def 9.13. *A property P holds almost everywhere if it holds except on a measure* 0 *set.*

Def 9.14. *For* $A \in \mathbb{R}$*, the essential supremum is*

$$\text{esssup}(A) = \inf\{c | x \leq c, \forall x \in A \backslash N, \mu(N) = 0\}.$$

An exercise is to show that $\text{esssup}([0, 1] \cup \{2\}) = 1$. It's not hard to show that 2 is not the sup but we need some work to show that we cannot discard all point close to 1.

For $F : \mathbb{R} \to \mathbb{R}$ right continuous, we can define

$$\mu_F((a, b]) := F(b) - F(a)$$

then $(a, b]$ generates $\mathcal{B}$ and μ_F generates a measure. When $F(x) = x$, the measure is Lebesgue, and when $F = \mathbb{1}_{x \geq x_0}$ the measure is the delta function at x_0.

Def 9.15. *A probability measure* $(\Omega, \mathcal{F}, \mathbb{P})$ *is such that* $\mathbb{P}(\Omega) = 1$.

Theorem 9.16. *If the range of F is* $[0, 1]$*, then* $(\Omega, \mathcal{F}, \mu_F)$ *is a probability space.*

10. 2/1: MEASURABLE MAPS; LP SPACES

10.1. Measurable maps.

Def 10.1. *Let $(X, \mathcal{A})$, $(Y, \mathcal{B})$ be measurable spaces. Function $f : X \to Y$ is measurable if for $\forall B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$.*

Def 10.2. *A measurable function $X : \Omega \to \mathbb{R}$ is a random variable, where the σ-algebra on $\mathbb{R}$ is just the Borel sets.*

Let $\bar{\mathbb{R}}$ denote the extended real line, i.e. $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$.

Proposition 10.3. *Function $f : X \to \bar{\mathbb{R}}$ is measurable on $(X, \mathcal{A})$ iff for all $c \in \mathbb{R}$, $\{x \in X | f(x) < c\} \in \mathcal{A}$. Equivalently, this is also written as $f^{-1}((-\infty, c]) \in \mathcal{A}$.*

Def 10.4. *$f_n : X \to \bar{\mathbb{R}}$ converge point-wise to f if $f_n(x) \to f(x)$ for all $x \in X$.*

$f_n : X \to \bar{\mathbb{R}}$ converge point-wise a.e. to f if $f_n(x) \to f(x)$ for all $x \in X \setminus N$, $\mu(N) = 0$.

Theorem 10.5.

- *If f_n are measurable functions that converges point wise to f, then f is measurable.*
- *If f_n are measurable functions that converges point wise a.e. to f, and the measure space $(X, \mathcal{A}, \mu)$ is complete, then f is measurable.*

Def 10.6. *A function $\phi : X \to \mathbb{R}$ on $(X, \mathcal{A})$ is simple if*

$$\phi = \sum_{i=1}^{n} c_n \mathbb{1}_{A_n}$$

for disjoint measurable sets A_i, and $c_i \in \mathbb{R}$.

Theorem 10.7. *Let $f : X \to [0, \infty]$ be measurable. Then there is an increasing sequence of simple functions that converges point wise to f.*

Proof. We construct such a sequence. Let

$$A_{N,k} = \left\{ x \in X, \frac{k-1}{2^N} \le f(x) < \frac{k}{2^N} \right\}, 1 \le k \le 2^{2n}$$

and

$$A_{N,2^{2N}+1} = \{x | f(x) \ge 2^n\}.$$

Then we can define simple function

$$\phi_n(x) = \sum_{k=1}^{2^{2N}+1} \frac{k-1}{2^N} \mathbb{1}_{A_{n,k}}(x)$$

that is simple, increasing, and convergent to f. $\square$

We've discussed non-negative measurable functions. For a general one we just decompose any function $f = f^+ - f^-$ and deal with f^+ and f^- separately.

Now we introduce Lebesgue integration. The story begins with the definition of simple functions.

Def 10.8. *For simple function ϕ, we define it's integral*

$$\int \phi d\mu = \sum_{i=1}^{n} c_i \mu(A_i).$$

Def 10.9. *For $f : X \to [0, \infty]$ that is measurable on $(X, \mathcal{A}, \mu)$, we define*

$$\int f d\mu = \sup \left\{ \int \phi d\mu \middle| \phi - \text{ simple }, \phi \leq f \right\}.$$

Def 10.10. *For more general cases like $f : X \to \mathbb{R}$, we define*

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu$$

if at least one of them is finite.

Def 10.11. *f is <u>measurable</u>, or that $f \in L^1(X, \mathcal{A}, \mu)$ if*

$$\int |f| d\mu < \infty.$$

Lemma 10.12.

- *For Lebesgue integral, we write*

$$\int f d\mu = \int f dx.$$

- *For Lebesgue Stieltjes integral*

$$\int f d\mu_F = \int f dF.$$

- $\int f d\delta_{x_0} = f(x_0)$.
- *For $f : \mathbb{N} \to \mathbb{R}$, the counting measure*

$$\int f d\lambda = \sum_{n=1}^{\infty} f_n.$$

So now we prove some convergence theorems. That is, we want to ask that if $f_n \to f$ in some sense, will the convergence be passed into the integrals? If so, what sense is that? One tempting answer is point wise is enough, but we have the following counter example.

Example 10.13.

$$f = \begin{cases} n & 0 \le x \le \frac{1}{n} \\ 0 & otherwise. \end{cases}$$

For this example we have

$$\lim_{n\to\infty} int f_n d\mu = 1 \neq 0 = \int f d\mu$$

since obviously $f_n \to f = 0$.

But in fact we need not so much more than this. We will only impose some condition on the point wise convergent sequences to conclude convergence in L^1 sense. The following are three most important theorems in this direction.

Theorem 10.14. *(Monotone convergence theorem) Let $f_n : X \to [0, \infty]$ be a measurable, non-decreasing sequence. Then*

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Lemma 10.15. *(Fatou's Lemma)For f_n measurable, we have*

$$\int \liminf_n f_n d\mu \le \liminf_n \int f d\mu.$$

Theorem 10.16. *(Lebesgue Dominated convergence theorem) For $f_n : X \to \mathbb{R}$ that is dominated by an integrable function g, i.e. $|f_n| \le g$, we have that if $f_n \to f$ point wise, then*

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

One really important theorem here is that, if we view

$$I(t) := \int f(t, x) d\mu(x)$$

then what is the relation of the differentiability of f in t with that of I's?

Corollary 10.17. *Let $f : X \times I \to \mathbb{R}$ where I is the set of parameters. Assume that*

(a) $f(\cdot, t)$ *is integrable;*
(b) $f(x, \cdot)$ *is differentiable;*
(c) $\left|\frac{\partial f}{\partial t}\right|(x, t) \le g(x)$ *where g is integrable.*

Then we know that

$$\phi(t) = \int f(x, t) d\mu(x)$$

is differentiable and

$$\frac{\partial \phi}{\partial t}(t) = \int \frac{\partial f}{\partial t}(x, t) d\mu(x).$$

For the proof we just look at finite differences and find a dominating function.

Now we will build toward Fubini's. So we first define product spaces

Def 10.18. *For the measure spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$, we define the product space as*

$$(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$$

where $\mathcal{A} \otimes \mathcal{B}$ is the σ-algebra generated by $\mathcal{A} \times \mathcal{B}$, and $\mu \otimes \nu$ is the corresponding measure on $\mathcal{A} \otimes \mathcal{B}$ such that

$$\mu \otimes \nu(A \times B) = \mu(A) \cdot \nu(B).$$

This definition is also a proposition since it also claims that such a product space exists.

Theorem 10.19. *(Fubini) For $f : X \times Y \to \bar{\mathbb{R}}$ measurable, we have*

(a) f is integrable and

$$\int |f| d\mu \otimes \nu < \infty \iff \text{either } \int_X \left(\int_Y |f| d\nu \right) d\mu < \infty \text{ or } \int_Y \left(\int_X |f| d\mu \right) d\nu < \infty$$

where the either or argument is for backwards direction. For the forward direction we can get both.

(b)

$$\int_{X \times Y} f \, d\mu \otimes \nu = \int_X \int_Y f = \int_Y \int_X f.$$

10.2. **Lp spaces.**

With the definition of measure spaces, we can now formally define L^p spaces.

Def 10.20. *For $1 \le p < \infty$, the space$L^p(X, \mathcal{A}, \mu)$ is a space of classes of equivalence of functions $f : X \to \mathbb{R}$ such that*

$$\int |f|^p d\mu < \infty$$

and $f \sim g$ if $f = g$ a.e..

For $p = \infty$, the equivalent classes are still the same definition and the only difference is that we only need $|f(x)| \le M$. Also, we define

$$||f||_\infty = \text{esssup}\{|f(x)|, x \in X\}.$$

Theorem 10.21. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Then for $1 \le p \le \infty$, the space L^p is Banach.*

Moreover, note that L^p is defined on equivalent classes. If we're just looking at the space of all p-th integrable measurable functions, $\mathcal{L}^p$, then it is not a Banach space.

We only prove here that there is a norm on L^p, and left the completeness parts since we've done that a lot of times.

Proof. As for why it is a norm for $p < \infty$, we check the conditions. That $||\lambda x|| = |\lambda| \cdot ||x||_p$ is obvious, the triangle inequality is the same Minkovski, and the only less clear part is just to show $|x| = 0 \to x = 0$.

That is, we need to show if $\int |f|^p d\mu = 0$, then $f = 0$ a.e.. Note that this is exactly why we cannot do with all measurable functions $\mathcal{L}^p$.

So we know that there are increasing simple functions $\phi_n \uparrow |f|^p$, so we can bound the integral by

$$0 \leq \int \phi_n d\mu \leq \int |f|^p d\mu = 0$$

but this means that $c_i = 0$ or $\mu(A_i) = 0$, which either way gives us $|f|^p = 0$ a.e., which means $f = 0$ a.e.. $\square$

Theorem 10.22. *For $f \in L^p(X)$, there exists ϕ_n simple such that*

$$\lim_{n\to\infty} ||f - \phi_n||_p = 0.$$

Theorem 10.23. *For $1 \leq p < \infty$, $L^p(\mathbb{R}^n)$ is separable (has countable dense set).*

Theorem 10.24. *$C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $1 \leq p < \infty$.*

Note that the above theorems gives us a clear path to theorem 10.22 since for $f \in L^p$, $\exists f_n \in C_c^\infty$ such that $||f - f_n||_p \to 0$, and we can again approximate continuous functions with polynomials of degree k, and use $\mathbb{Q}^k$ to approximate $\mathbb{R}^k$, the polynomial coefficients. And then the last step is to go from simple functions to polynomials. Hence we are done.

One last thing to note is that L^∞ is not separable.

11. 2/6: Some Lp inequalities; Hilbert spaces

11.1. Lp ineqaualities.

First, we look at some inequalities.

Proposition 11.1. *(Jensen's inequality) For $\phi : \mathbb{R} \to \mathbb{R}$ convex, we define*

$$\langle f \rangle_\mu := \frac{1}{\mu(x)} \int_X f \, d\mu = \int_X f \, d\frac{\mu}{\mu(X)}$$

then we have

$$\phi(\langle f \rangle_\mu) \leq \langle \phi \circ f \rangle_\mu .$$

Proof. To show this, we first assume that $\int_X \mu = 1$ and $\int_X f \, d\mu = x_0$.

Now, since the function ϕ is convex, we can find a supporting hyperplane that is tangent to the graph at point $(x_0, \phi(x_0))$, i.e. for all x, $\phi(x) \geq ax + b$ where $\phi(x_0) = ax_0 + b$.

Now we have

$$\int \phi \circ f \, d\mu \geq \int (af(x) + b) d\mu = a \int f \, d\mu + b = \phi(x_0)$$

which after a de-normalization we get what we want. □

Proposition 11.2. *(Holder's inequality) For $1 \leq p, p' \leq \infty$ and $\frac{1}{p} + \frac{1}{p'} = 1$. If $f \in L^p$ and $g \in L^{p'}$ we have $fg \in L^1$ and*

$$||fg||_1 \leq ||f||_p \cdot ||g||_{p'} .$$

A corollary of this will be the comparability of L^p spaces for finite measure spaces.

Corollary 11.3. *For the measure space (X, μ) with $\mu(X) < \infty$ and $1 \leq q \leq p \leq \infty$, we have*

$$L^\infty(X, \mu) \subset L^p(X, \mu) \subset L^q(X, \mu) \subset L^1(X, \mu).$$

Proof. Proof for L^p and L^q is enough. We know that

$$||f||_q \leq ||f||_p ||1||_{\frac{p}{q(p-q)}}$$

by Holder. The specific computation is:

$$\begin{aligned} \int |f|^q d\mu = \int |f|^q |1|^q d\mu &\leq \left(\int |f|^{q\frac{p}{q}} d\mu \right)^{\frac{q}{p}} \cdot \left(\int |1|^{q\frac{p}{p-q}} d\mu \right)^{\frac{p-q}{p}} \\ &= \left(\int |f|^p d\mu \right)^{\frac{q}{p}} \cdot \left(\int |1|^{\frac{p}{q(p-q)}} d\mu \right)^{\frac{p-q}{p}} \end{aligned}$$

where we take the q-th power of both sides to get the result. □

Proposition 11.4. *(Minkowski's inequality) $||f + g||_p \leq ||f||_p + ||g||_{p'}$.*

Proposition 11.5. *(Chebyshev's inequality): Let* $f \in L^p(X)$ *for* $1 \le p < \infty$*. Then*

$$\mu(\{x \in X, |f(x)| > \varepsilon\}) \le \frac{1}{\varepsilon^p}||f||_p^p$$

where we break the convention here and view ε as a large number.

Proof.

$$||f||_p^p = \int_X |f|^p d\mu \ge \int_{|f|>\varepsilon} |f|^p d\mu \ge \varepsilon^p \int_X \mathbb{1}_{|f|>\varepsilon} d\mu$$

and we can just note that the integral is nothing but the left side. □

Def 11.6. *Let the* convolution *of two functions f and g be*

$$(f * g)(x) \int_{\mathbb{R}^n} f(x-y)g(y)dy = \int_{\mathbb{R}^n} f(y)g(x-y)dy.$$

Proposition 11.7. *(Young's inequality) (Not the one before) Let* $f \in L^p$*,* $g \in L^q$ *where* p, q *are not related. Then*

$$f * g \in L^r$$

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

Example 11.8.

- Let $p = q = 2$ and $r = \infty$. This means that the convolution of two L^2 functions is much smoother than L^2, which we can understand as p regularizes q.
- If $p = 1$, then $q = r$ which means that the convolution of an L^1 function does not change the regularity.
- If $f = \delta_0$, then $f * g = g$, that is, as smooth as g.
- If $p = q = 1$, then $r = 1$ which means that the convolution is a product on the L^1 space, and that L^1 is an algebra under $(+, *)$.

11.2. **Chapter 6: Hilbert space.**

The basic element of Hilbert spaces is inner product or scalar product.

Def 11.9. *Let X be a vector space (note that we don't need a norm) with a function*

$$(\cdot,\cdot) : X \times X \to \mathbb{R}(\text{ or } \mathbb{C})$$

where our focus is $\mathbb{R}$*. The function is an inner product if it satisfies*

(a) (Bi-linearity) $(x, \lambda y + \mu z) = \lambda(x, y) + \mu(x, z)$*;*
(b) (Symmetric) $(x, y) = \overline{(y, x)}$*;*
(c) (Non negative) $(x, x) \ge 0$*;*
(d) (Definite) $(x, x) = 0 \Rightarrow x = 0$.

Remark 11.10. *Note that the definition of inner product is anti-linear, i.e. by bi-linearity and symmetric, we have*

$$(\lambda x + \mu y, z) = \overline{(z, \lambda x + \mu y)} = \bar{\lambda}\overline{(z, x)} + \bar{\mu}\overline{(z, y)} = \bar{\lambda}(x, z) + \bar{\mu}(y, z)$$

so we note that there's an extra conjugation when we want to apply linearity to the first component.

This somehow weirdness is because $z\bar{z}$ is a norm while z^2 is not.

Def 11.11.

- *If X is a real vector space, then $(\cdot, \cdot) : X \times X \to \mathbb{R}$ is called a bilinear form;*
- *If X is complex valued, then $(\cdot, \cdot) : X \times X \to \mathbb{R}$ is called a sesquilinear form;*
- *$||x|| := (x, x)^{\frac{1}{2}}$ defines a norm (check later);*
- *$(X, (\cdot, \cdot))$ is a normed vector space which we call pre-Hilbert space;*
- *A complete pre-Hilbert space is called a Hilbert space.*

Note from the definition that any Hilbert space is Banach.

Example 11.12.

- $\mathbb{C}^n$ equipped with $(x, y) = \sum_{i=1}^{\infty} \bar{x}_i y_i$;
- $C([a, b])$ with $(f, g) := \int_a^b \overline{f(x)} g(x) dx$;
- $C^k([a, b])$ with
$$(f, g)_k := \sum_{j=0}^{k} \int_a^b \overline{f^{(j)}(x)} g^{(j)}(x) dx$$
is pre-Hilbert and whose completion is $H^k = W^{(k,2)}$, the Sobolev space of functions that up to its k-th derivative is square integrable.
- For $p \neq 2$, L^p is not pre-Hilbert since they are not the dual of themselves.

Theorem 11.13. *(Cauchy-Schwarz) Let $x, y \in X$, then*

$$|(x, y)| \leq ||x|| \cdot ||y||$$

Proof. To simplify things, we just assume $(\cdot, \cdot)$ maps into $\mathbb{R}$, and $\lambda \in \mathbb{R}$. The trick here is to manually add a λ into the game and use it to come up with our solution.

By definition we have

$$0 \leq ||x - \lambda y||^2 = (x - \lambda y, x - \lambda y) = (x, x) + \lambda^2(y, y) - 2\lambda(x, y)$$

which implies (we take $\lambda > 0$ since otherwise it won't be meaningful)

$$2|\lambda||(x, y)| \leq ||x||^2 + \lambda^2||y||^2$$

$$\Rightarrow |(x,y)| \leq \frac{1}{2|\lambda|}||x||^2 + \frac{\lambda}{2}||y||^2$$

from which we can choose $\lambda = ||x|| \cdot ||y||^{-1}$ for $y \neq 0$, and we will be done.

If $y = 0$ the result is obvious. □

Corollary 11.14. $||x|| := (x,x)^{\frac{1}{2}}$ *defines a* *<u>norm</u>*.

Proof. We only prove the triangle inequality since other properties are easy.

Note that

$$\begin{aligned} ||x+y||^2 = (x+y, x+y) &= ||x||^2 + ||y||^2 + 2(x,y) \\ &\overset{C.S}{\leq} ||x||^2 + ||y||^2 + 2||x|| \cdot ||y|| = (||x|| + ||y||)^2 \end{aligned}$$

and we have the triangle inequality:

$$||x+y|| \leq ||x|| + ||y||.$$

□

Theorem 11.15. *(Parallelogram Law) If* $(X, || \cdot ||)$ *is an inner product space iff* $\forall x, y \in X$

$$||x+y||^2 + ||x-y||^2 = 2||x||^2 + 2||y||^2.$$

Proof.

$\Rightarrow$:

This direction is clear just by expanding the left side with definition of inner products.

$\Leftarrow$:

This direction is less clear, the method is to define an inner product

$$(x,y) := \frac{1}{4}(||x+y||^2 - ||x-y||^2)$$

and to check that it is indeed an inner product, which involves some special choice of elements. Check! □

Def 11.16. *An inner product on a product space* $X \times Y$ *is defined as*

$$\langle (x_1, y_1), (x_2, y_2) \rangle_{X \times Y} := \langle x_1, x_2 \rangle_X + \langle y_1, y_2 \rangle_Y$$

where note we use $(\cdot, \cdot)$ *to express coordinate, and use* $\langle \cdot, \cdot \rangle$ *to express inner product.*

Proposition 11.17. $(X \times Y, \langle \cdot, \cdot \rangle_{X \times Y})$ *is pre-Hilbert.*

Theorem 11.18. $(\cdot, \cdot) : X \times X \to \mathbb{C}$ *is continuous for* $|| \cdot ||$.

Proof.

$$\begin{aligned}|(x_1, y_1) - (x_2, y_2)| &= |(x_1 - x_2, y_1) + (x_2, y_1 - y_2)| \\ &\overset{C.S.}{\leq} ||x_1 - x_2|| \cdot ||y_1|| + ||x_2|| \cdot ||y_1 - y_2||\end{aligned}$$

so if x_1 is close to x_2 and y_1 is close to y_2 we have that the distance between the two inner products is small.

Note that this result holds only locally, i.e. the last line depends on the value of x_1 and y_1. ☐

Def 11.19. *We define* $\underline{orthogonality}$ *here. For* $\mathcal{H}$ *a Hilbert space*

- $x \perp y$ *if* $(x, y) = 0$*;*
- $A \perp B$ *if* $x \perp y$ *for all* $x \in A$*,* $y \in B$*;*
- *The* <u>*orthogonal complement*</u> *of A is*

$$A^{\perp} = \{x \in \mathcal{H} | x \perp y, y \in A\}.$$

Theorem 11.20. $A^{\perp}$ *is a closed linear subspace of* $\mathcal{H}$*.*

Note that for closeness we use continuity of norm.

Theorem 11.21. *(Projection theorem) For* $\mathcal{M}$ *closed linear subspace of* $\mathcal{H}$*.*

(a) $\forall x \in \mathcal{H}$, $\exists! y \in \mathcal{M}$ *such that* $||x - y|| = \min_{z \in \mathcal{M}} ||x - z||$.
(b) $y \in \mathcal{M}$*, the closest point to x in* $\mathcal{M}$ *is the unique point such that* $x - y \in \mathcal{M}^{\perp}$*.*

The idea of the proof is to construct $y_n \in \mathcal{M}$ and show that y_n is Cauchy.

This theorem is useful because it gives us a method of divide and conquer, i.e. for any $x \in X$ we can decompose it into

$$x = y + (x - y)$$

for $y \in \mathcal{M}$ and $x - y \in \mathcal{M}^{\perp}$.

Proof.

(a):

We want to minimize $||x - z||$ for $z \in \mathcal{M}$, so we define

$$d := \inf\{||x - z||, z \in \mathcal{M}\}$$

which always exists. And we can construct a minimizing sequence y_n by the definition of infimum such that

$$||x - y_n|| \leq d + \frac{1}{n}$$

and try to find the limit y (not yet well defined for now) by proving that the sequence is Cauchy.

To do this we use the parallelogram law:

$$||y_n - y_m||^2 + 4\left|\left|x - \frac{y_n + y_m}{2}\right|\right|^2 = 2||x - y_n||^2 + 2||x - y_m||^2$$

where since $x - \dfrac{y_n + y_m}{2} \geq d$ by assumption that $\mathcal{M}$ is a linear subspace, we get

$$||y_n - y_m||^2 \leq 2\left(d + \frac{1}{n}\right)^2 + 2\left(d + \frac{1}{m}\right)^2 - 4d^2 \leq 4d\left(\frac{1}{n} + \frac{1}{m}\right) + 2\left(\frac{1}{n^2} + \frac{1}{m^2}\right) \to 0$$

as $m, n \to 0$, thus y_n is Cauchy and $y_n \to y \in \mathcal{M}$ since $\mathcal{M}$ is closed, and $d = ||x - y||$.

As for uniqueness, we assume $\tilde{y}$ also satisfies the condition, then again by parallelogram law

$$0 \leq ||y - \tilde{y}||^2 = 2||x - \tilde{y}||^2 + 2||x - y||^2 - 4\left|\left|x - \frac{y + \tilde{y}}{2}\right|\right|^2 \leq 4d^2 - 4d^2 = 0$$

and hence we are done.

(b):

Let y be as above and for $\forall z \in \mathcal{M}$, we want to show that $(x - y, z) = 0$. The trick here is again to create an extra variable λ out of thin air and use it as a bound: consider $y - \lambda z \in \mathcal{M}$, we have by part (a) that

$$||x - y||^2 \leq ||x - y + \lambda z||^2 \Rightarrow 0 \leq 2(\pm\lambda) \cdot (x - y, z) + \lambda^2 z^2$$

where the $\pm$ sign really is saying that the sign of λ is arbitrary. Thus we can take the absolute value and get

$$2|\lambda| \cdot |(x - y, z)| \leq \lambda^2 z^2$$

for which we can let $\lambda \to 0$ and get $|(x - y, z)| = 0$.

For uniqueness, let $x - y \in \mathcal{M}^\perp$ and $x - \tilde{y} \in \mathcal{M}^\perp$ for $y, \tilde{y} \in \mathcal{M}$. Then $y - \tilde{y} \in \mathcal{M}$ and thus

$$(x - y, y - \tilde{y}) = 0 = (x - \tilde{y}, y - \tilde{y})$$
$$\Rightarrow (x, y - \tilde{y}) - (y, y - \tilde{y}) = (x, y - \tilde{y}) - (\tilde{y}, y - \tilde{y})$$
$$\Rightarrow ||y - \tilde{y}||^2 = 0$$

and hence the y is unique.

$\square$

12. 2/10: ORTHOGONAL DECOMPOSITION OF HILBERT SPACES; FOURIER SERIES

12.1. **Orthogonal decomposition of Hilbert spaces.**

Remark 12.1. *$H^k = W^{k,2}$ is the completion of $(C^k, (\cdot,\cdot))$, which is pre-hilbert.*

Def 12.2. *Let $\mathcal{M}, \mathcal{N}$ be closed linear subspaces of $\mathcal{H}$ with $\mathcal{M} \perp \mathcal{N}$. Then the direct sum*

$$\mathcal{M} \oplus \mathcal{N} = \{y + z | y \in \mathcal{M}, z \in \mathcal{N}\}.$$

Corollary 12.3. *Let $\mathcal{M}$ be a closed linear subspace. Then $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}$.*

Proof. Let $x \in \mathcal{H}$, then let y be as the y in the projection theorem, we have

$$x = y + (x - y) \in \mathcal{M} \oplus \mathcal{M}^{\perp}.$$

□

Note that if $\mathcal{M}$ is not closed, then $\mathcal{H} = \overline{\mathcal{M}} \oplus \mathcal{M}^{\perp}$. This instance is often used when $\mathcal{M}$ is the range of an operator.

Also, the above doesn't always hold for Banach spaces.

Def 12.4. *A set $U \subset \mathcal{H}$ is orthogonal if $u \perp v$ for all $u, v \in U$ and $u \neq v$. The set is orthonormal if $||u|| = 1$ for all $u \in U$ in addition.*

Now, in order to deal with summation not on natural numbers, we define the following:

Def 12.5.

- *For I an index set and $J \subset I$ a finite subset, we define an unordered sum*

$$S_j = \sum_{\alpha \in J} x_\alpha$$

  for $x_\alpha \in \mathcal{H}$.
- *An unordered sum converges unconditionally to $x = \sum_{\alpha \in I} x_\alpha$ if $\forall \varepsilon > 0$, $\exists J^\varepsilon$ finite such that*

$$||S_J - x|| < \varepsilon$$

  for any finite J such that $J^\varepsilon \subset J \subset I$.

  Note that under this definition of convergence naturally convergences independent of reordering of the terms.
- *We say that $\sum_{\alpha \in I} x_\alpha$ converges absolutely if $\sum_{\alpha \in I} |x_\alpha|$ converges unconditionally.*

Proposition 12.6. *Absolute convergence $\Rightarrow$ unconditionally convergence.*

Note that the converse is not true since we can take $\sum_{n=1}^{\infty} \frac{1}{n} e_n$ for e_i orthogonal, and it definitely is not absolutely convergent.

To see why it converges unconditionally, we compute

$$\left\| \sum_{n=1}^{N} \frac{1}{n} e_n \right\|^2 = \sum_{n=1}^{N} \left\| \frac{1}{n} e_n \right\|^2 = \frac{n=1}{N} \frac{1}{n^2} < \infty.$$

The inequality above is due to orthogonality. This is also illustrated by the lemma below.

Lemma 12.7. *Let $U = \{u_\alpha, \alpha \in I\}$ be an orthogonal set in $\mathcal{H}$, then $\sum_{\alpha \in I} u_\alpha$ converges unconditionally $\iff \sum_{\alpha \in I} ||u_\alpha||^2 < \infty$.*

Proof.

We only need to compute for all J finite we have

$$\left\| \sum_{\alpha \in J} u_\alpha \right\|^2 = \left(\sum_{\alpha \in J} u_\alpha, \sum_{\beta \in J} u_\beta \right) = \sum_{\alpha \in J} ||u_\alpha||^2$$

and hence we can pass the limit for either direction. □

Theorem 12.8. *(Bessel's Inequality) For $x \in \mathcal{H}$, let $U = \{u_\alpha, \alpha \in I\}$ be orthonormal in $\mathcal{H}$, then we have*

(a) $\sum_{\alpha \in I} |(u_\alpha, x)|^2 \le ||x||^2$.

(b) $x_u := \sum_{\alpha \in I} (u_\alpha, x) u_\alpha$ *is a convergent sum.*

(c) $x - x_u \in U^\perp$.

Proof.

(a): By orthogonality

$$0 \le \left\| x - \sum_{\alpha \in J} (u_\alpha, x) u_\alpha \right\|^2 \overset{orthogonal}{=} ||x||^2 - \sum_{\alpha \in J} |(u_\alpha, x)|^2$$

where the term $-\sum_{\alpha \in J} |(u_\alpha, x)|^2$ has a minus sign since the cross term cancels the square term. Now we pass the limit to get the result.

(b): By (a) and lemma above, we know that the sum converges unconditionally.

(c): We check that for $u_\beta \in U$

$$(x - x_u, u_\beta) = (x, u_\beta) - \left(\sum_{\alpha \in I} (u_\alpha, x) u_\alpha, u_\beta \right) = (x, u_\beta) - \sum_{\alpha \in I} \overline{(u_\alpha, x)} \left(u_\alpha, u_\beta \right) = (x, u_\beta) - \overline{(u_\beta, x)} = 0$$

where the sum disappears since $(u_\alpha, u_\beta) \neq 0$ iff $\alpha = \beta$. $\square$

Def 12.9. *The span of U, an orthonormal set is*

$$[U] := \left\{ \sum_{u\in U} c_u \cdot u \middle| c_u \in \mathbb{C}, \sum_{u\in U} |c_u|^2 < \infty \right\}$$

which is the smallest subspace of $\mathcal{H}$ containing U.

Theorem 12.10. *Let U be orthogonal, then the following are equivalent conditions:*

(a) $(u_\alpha, x) = 0$, $\forall x \in I \Rightarrow x = 0$.
(b) $x = \sum_{\alpha\in I} (u_\alpha, x)u_\alpha, \forall x \in \mathcal{H}$.
(c) $||x||^2 = \sum_{\alpha\in I} |(u_\alpha, x)|^2, \forall x \in \mathcal{H}$. *(Parseval's identity).*
(d) $[U] = \mathcal{H}$.

Proof. We first note that $b \iff d$ due to the Bessel's inequalities.

Moreover, we have

$$(a) \Rightarrow x = x_u \text{ (in Bessel's) } \iff (b) \overset{Lemma}{\Rightarrow} (c) \Rightarrow (a).$$

$\square$

Theorem 12.11. *Let U be an orthonormal basis with*

$$x = \sum_{\alpha\in I} a_\alpha u_\alpha, \quad y = \sum_{\alpha\in I} b_\alpha u_\alpha$$

then

$$(x, y) = \sum_{\alpha\in I} \overline{a_\alpha} b_\alpha.$$

The trick is similar to that of part (c) in theorem 12.8.

Theorem 12.12. *(Based on AC) Every Hilbert space admits an orthonormal basis. If U is an orthonormal set, then it is contained in an orthonormal basis.*

The second part of the theorem is due to the first part and the fact that $U^\perp$ is a Hilbert space.

Corollary 12.13. *Assume $\mathcal{H}$ is separable (which implies that it has a countable basis), then*

$$\mathcal{H} \cong l^2(\mathbb{N})$$

that is, we can identify the basis elements that generates an isomorphism.

This means that essentially "there is only 1 separable space."

Theorem 12.14. *(Banach-Mazur Thoerem) Every real, separable Banach space* $(X, ||\cdot||)$ *is isometrically isomorphic to a closed subspace of* $C^0([0,1],\mathbb{R})$.

The last remaining thing before moving on to the next topic is how to construct an orthonormal basis. To do so we use the Gram-Schmidt procedure. This procedure constructs an orthonormal set iteratively.

Let V be a set of linearly independent vectors, then we construct an orthogonal set U with $[V] = [U]$ in the following way:

Let

$$u_{n+1} = c_{n+1}\left(v_{n+1} - \sum_{k=1}^{n}(u_k, v_{n+1})u_k\right)$$

which is nonzero because V is linearly independent (if $c_{n+1} \neq 0$). Then we pick c_{n+1} be such that $|u_{n+1}| = 1$, and we are done.

12.2. **Chapter 7: Fourier series.**

Here, we try to construct an orthonormal basis that are 2π periodic. Note that for all f 2π periodic, we can view it as defined on a circle $\mathbb{T}$.

In $C(\mathbb{T})$, we define the norm

$$||f|| := \left(\int_0^{2\pi} |f(x)|^2 dx\right)^{\frac{1}{2}}$$

and the completion of $C(\mathbb{T})$ under this norm is $L^2(\mathbb{T})$.

Since $L^2(\mathbb{T})$ is complete, we also acknowledge that there is an inner product

$$\langle f, g\rangle = \int_0^{2\pi} \overline{f} g dx$$

such that $(L^2(\mathbb{T}), \langle\cdot,\cdot\rangle)$ is Hilbert.

Our construction of a basis for this space is to choose

$$e_n(x) = \frac{1}{\sqrt{2\pi}}e^{inx} = \frac{1}{\sqrt{2\pi}}\left(\cos(nx) + i\sin(nx)\right).$$

Then $e_n(x) \in C(\mathbb{T})$ and it is in $C^\infty(\mathbb{T})$. Moreover we know they are orthonormal:

$$\langle e_n, e_m\rangle = \begin{cases} 1 & n = m \\ 0 & n \neq m. \end{cases}$$

Theorem 12.15.

- $U := \{e_n(x) | n \in \mathbb{Z}\}$ *is an orthonormal basis for* $L^2(\mathbb{T})$.

- *Thus*
$$f(x) = \sum_{n\in\mathbb{Z}} \langle e_n, f\rangle e_n(x)$$
for all $f \in L^2(\mathbb{T})$, where $\langle e_n, f\rangle$ is the Fourier coefficient.

Before the proof, let's first prove a lemma.

Lemma 12.16. *(density argument) Let $f_n \to f$ with*
$$f_m = \sum_{n\in\mathbb{Z}} \langle e_n, f_m\rangle e_n$$
then
$$\langle e_n, f_m\rangle \xrightarrow{m\to\infty} \langle e_n, f\rangle$$
and
$$f = \sum_{n\in\mathbb{Z}} \langle e_n, f\rangle e_n.$$

Proof. (of Lemma 12.16)

We have
$$\langle f, f_m\rangle = \left\langle f, \sum_{n\in\mathbb{Z}} \langle e_n, f_m\rangle e_n \right\rangle = \sum_{n\in\mathbb{Z}} \langle e_n, f_m\rangle \langle f, e_n\rangle$$
which means that as long as $\langle f, e_n\rangle = 0$ for all n, $\langle f, f_m\rangle = 0$. Hence if we can pass the limit, we will have
$$\langle f, f_m\rangle \to \langle f, f\rangle = 0$$
which by condition (a) of theorem 12.10 we know that the lemma holds.

Notably that this lemma holds as long as $f_n \to f$ weakly. □

Proof. (Theorem 12.15)

First, we want to choose a sequence such that the condition of the lemma holds. But we know that $C(\mathbb{T})$ is dense in $L^2(\mathbb{T})$ since we can approximate L^2 functions with polynomials. That is, we want to choose f_m continuous such that any function can be decomposed by our basis, then we win. Since if we can show this, we can use the lemma to claim that for all $f \in L^2(\mathbb{T})$, it can be decomposed.

So our problem reduces to: for all f continuous, can we write $f = \sum_{n\in\mathbb{Z}} \langle e_n, f\rangle e_n$?

The rest of proof is in book (since notes skip too many details of kernel and stuff).

□

13. 2/13: Projection; Dual of Hilbert spaces; Adjoint operator

13.1. **Projection.**

Def 13.1. *Let M, N be subspaces of X, then we say $X = M \oplus N$ if $\forall x \in X$, $\exists!$ decomposition $x = y + z$ with $y \in M$, $z \in N$.*

Note that the uniqueness condition above implies that $M \cap N = \{0\}$.

Note that there's nothing about orthogonality here yet. What we can think of is $e_1 = (1, 0)$ and $e_2 = (1, 1)$ and with this example we can consider the projection to M along N by $Px = y$. And we can verify that P is linear and $P^2 = P$ (idempotent). We now see that this is in fact a defining factor.

Def 13.2. *A projector is a linear operator $X \to X$ such that $P^2 = P$.*

Proposition 13.3. *If M, N are closed, then P defined above is bounded.*

Theorem 13.4. *For X a vector space, we have*

(a) If $P : X \to X$ is a projection, then $X = \operatorname{Ran} P \oplus \ker P$.
(b) If $X = M \oplus N$, then $\exists P : X \to X$ with $M = \operatorname{Ran} P$ and $N = \ker P$.

Proof.

part (b) is just from above construction.

For part (a), let $y \in \operatorname{ran} P$, then $y = Px$ and $Py = P^2 y = Px = y$ which gives

$$x = Px + (x - Px)$$

for $Px \in M = \operatorname{Ran} P$ and $x - Px \in \ker P$ since $Px - P(Px) = 0$ by above computation.

Moreover, if $x = y + z$ for $y \in \operatorname{Ran}(P), z \in \ker(P)$, then $Px = Py = y$ and hence y is uniquely determined by Px as the value of the function at a point. Then $z = x - y$ is also uniquely determined, and we are done. □

Note that $I - P$ projects onto N along M. Also, P is bounded if M is closed(?), but it is not necessarily bounded by 1, as our example of non-orthogonal example above.

Now we add orthogonality and everything gets better. Note that we denote, for this case, the inner product by $\langle \cdot, \cdot \rangle$. Remember that from projection theorem we get that for M a closed subspace, $\mathcal{H} = M \oplus M^\perp$ and there exists a P that projects onto M along $M^\perp$ by part (b) above. We deduce one defining properties of such projections:

$$\langle Px, y \rangle = \langle Px, y - Py + Py \rangle = \langle Px, Py \rangle = \langle x, Py \rangle$$

where there's a similar argument for the last equality sign.

Def 13.5. *An <u>orthogonal projector</u> is a projector $P^2 = P$ such that $\langle x, Py \rangle = \langle Px, y \rangle$, $\forall x, y \in \mathcal{H}$.*

Here we've defined the notion of orthogonal projector that is free of a specific M.

Proposition 13.6. $||P|| = 1$ *unless* $P \equiv 0$ *for* P *an orthogonal projector.*

Proof. For $P \equiv 0$ it's easy. For $P \neq 0$, $\exists x \in \mathcal{H}$ such that $Px \neq 0$, call $y = Px$ then we know $Py = y$ such that $||P|| \geq 1$ by definition.

For the other direction, we note that $\forall x \neq 0$ with $Px \neq 0$

$$||Px|| \leq \frac{\langle Px, Px\rangle}{||Px||} = \frac{\langle x, P^2x\rangle}{||Px||} = \frac{\langle x, Px\rangle}{||Px||} \overset{C.S.}{\leq} \frac{||x|| \cdot ||Px||}{||Px||} = ||x||$$

and hence we are done. ☐

Theorem 13.7. *For* $\mathcal{H}$ *a Hilbert space, then we have:*

(a) For P *an orthogonal projector, we have* $\mathcal{H} = \operatorname{Ran} P \oplus \ker P$ *with* $\operatorname{Ran} P$ *closed.*
(b) Let $\mathcal{H} = M \oplus M^{\perp}$*, then* $\exists! P$ *orthogonal such that* $M = \operatorname{Ran} P$*.*

Notice that we didn't assert uniqueness of the projection above.

Proof.

(b): By construction we have $Px = x$ on M and $Px = 0$ on $M^{\perp}$, and since we can use our tools in Hilbert space (projection theorem) to claim that decomposition of x into $M \oplus M^{\perp}$ is unique, we know that P is unique.

(a): $\mathcal{H} = \operatorname{Ran} P \oplus \ker P$ by (a) above. Moreover, since P orthogonal then $x = Px + (x - Px)$ and since for $y = Px \in \operatorname{Ran} P$ and $z \in \ker P$ we have

$$\langle Px, z\rangle = \langle x, Pz\rangle = 0$$

which gives us that $\operatorname{Ran} P = (\ker P)^{\perp}$. ☐

Example 13.8. $Pf(x) = \dfrac{f(x) + f(-x)}{2}$ *is an orthogonal projector of* $L^2(\mathbb{R})$ *into even functions.*

Example 13.9. *Let* A *be a measurable set in* $\mathbb{R}$ *and the character* $\chi_A = \mathbb{1}_A$*, then*

$$P_A f(x) = \chi_A(x) f(x)$$

is an orthogonal projection of L^2 *onto functions with support in* $\overline{A}$*, since the support is defined to be the closure.*

Example 13.10. *Rank 1 projectors: Let* $P_u(x) = \langle u, x\rangle \cdot u$ *for* $||u|| = 1$*, then denote* $Pu := u \otimes u = |u\rangle\langle u|$.

13.2. **Dual of Hilbert space.**

Denote a functional with a Hilbert domain as ϕ. Remember that

$$||\phi||_{\mathcal{H}^*} = \sup_{x\in\mathcal{H};||x||=1} |\phi(x)|$$

where we can either view it as the traditional definition of norm, or remember the open mapping theorem's conclusion.

For $y \in \mathcal{H}$, define $\phi_y(x) = \langle y, x\rangle, \forall x \in \mathcal{H}$, then Cauchy says

$$|\phi_y(x)| \leq ||y|| \cdot ||x||$$

and $\phi_y(y) = ||y||^2$, thus $||\phi_y||_{\mathcal{H}^*} = ||y||_{\mathcal{H}}$. We see each y is a representation of an element in the dual. Moreover, we now show that this is really enough.

Theorem 13.11. *(Riesz Representation Theorem) Let $\phi \in \mathcal{H}^*$ be bounded linear form, then $\exists! y \in \mathcal{H}$ such that $\phi(x) = \langle y, x\rangle, \forall x \in \mathcal{H}$. This gives us $\mathcal{H} \cong \mathcal{H}^*$.*

Proof.

Case 1: when $\phi = 0$, then $y = 0$.

Case 2: when $\phi \neq 0$, then $\exists z \in (\ker\phi)^\perp$ with $z \neq 0$ since kernel is always a subspace. Then we define

$$P(x) := \frac{\phi(x)}{\phi(z)} z$$

then p is linear since ϕ is linear and $P^2 = P$ (simple check). This means we can adopt our just learned projection theory and say $\mathcal{H} = \text{Ran } P \oplus \ker P$.

You might not notice but this is already most of the proof. Let's continue: we can check that $\ker P = \ker(\phi)$ and $\text{Ran } P = \mathbb{C}z$ since $\dfrac{\phi(x)}{\phi(z)} \in \mathbb{C}$.

So $z \in (\ker\phi)^\perp = (\ker P)^\perp \Rightarrow \text{Ran}(P) \perp \ker(P)$ as $\text{Ran}(P) = \mathbb{C}z$.

Now, we have $\mathcal{H} = \text{Ran } P \oplus \ker P = \mathbb{C}z \oplus \ker\phi$ is an orthogonal direct sum, which means $\forall x \in \mathcal{H}, x = \alpha z + n$ for $n \in \ker\phi$, with $\langle z, n\rangle = 0$.

Thus, dot product both sides with z we get

$$\langle z, x\rangle \alpha ||z||^2 + 0 \Rightarrow \alpha = \frac{\langle z, x\rangle}{||z||^2}$$

by inner product with z.

Let

$$\phi(x) = \phi(\alpha z + n) = \alpha\phi(z) + 0 = \left\langle \frac{\phi(z)z}{||z||^2}, x \right\rangle$$

then we've specified a corresponding $y := \frac{\phi(z)z}{||z||^2} \in \mathcal{H}$ for each $\phi \in \mathcal{H}^*$.

To show uniqueness, if $\phi_{y_1}(x) = \phi_{y_2}(x) = \langle y_1, x\rangle = \langle y_2, x\rangle$, then take $x = y_1 - y_2$ we have (similar trick)

$$\langle y_1, y_1 - y_2\rangle = \langle y_2, y_1 - y_2\rangle \Rightarrow y_1 = y_2.$$

$\square$

13.3. **Adjoint Operator.**

(Heuristics) Let A be bounded, we have A^* such that

$$\langle x, Ay\rangle = \langle A^* x, y\rangle$$

for all $x, y \in \mathcal{H}$.

Now we fix $x \in \mathcal{H}$ and define $\phi_x(y) := \langle x, Ay\rangle$ is bounded since

$$\phi_x(y) = ||x|| \cdot ||A|| \cdot ||y||.$$

By Riesz, we have $\phi_x(y) = \langle x, Ay\rangle \cong \langle z, y\rangle$ for unique z. So we just let $A^* x = z$ and we are done. What is left is to check that A^* is linear and bounded (it's nice).

To show the uniqueness of A^* we note that

$$\langle A_1^* x, y\rangle = \langle A_2^* x, y\rangle \Rightarrow \langle (A_1^* - A_2^*)x, y\rangle = 0$$

for all x, y. But this means $(A_1^* - A_2^*)x = 0$ and hence $(A_1^* - A_2^*) = 0$.

Some properties of adjoints are:

(1) $A^{**} = A$.
(2) $(AB)^* = B^* A^*$.

Example 13.12.

- In $\mathbb{R}^n$, $A^* = A^T$, i.e. $(A^T)_{ij} = A_{ji}$.
- In $\mathbb{C}^n$, $A^* = \overline{A^T}$, i.e. $(A^*)_{ij} = \overline{A}_{ji}$.
- For
$$(Kf)(x) = \int_0^1 k(x, y) f(y) dy$$
where $f \in L^2(0, 1)$, then
$$K^* f(x) = \int_0^1 \overline{k}(y, x) f(y) dy.$$

Note that in problems we can always try to apply the adjoint.

Now, let $Ax = y$ where $A : \mathcal{H} \to \mathcal{H}$. If $\exists A^* z = 0$ for $z \neq 0$, then we can let $y = Ax$ and get

$$\langle y, z\rangle = \langle Ax, z\rangle = \langle x, A^* z\rangle = 0.$$

This really gives us a constraint on y: y must be orthogonal to $\ker A^*$, which of course agrees with linear algebra.

14. 2/15: ADJOINT OPERATORS; FREDHOLM ALTERNATIVE; THEOREMS RECAP

14.1. Adjoint operators.

Theorem 14.1. *For $A : \mathcal{H} \to \mathcal{H}$ bounded, then*

$$\overline{\mathrm{Ran}(A)} = (\ker A^*)^\perp$$

$$\ker(A) = (\mathrm{Ran}\, A^*)^\perp$$

which by projection theorem for Hilbert space is the same thing as

$$\mathcal{H} = \overline{\mathrm{Ran}\, A} \oplus \ker A^* = \ker A \oplus \overline{\mathrm{Ran}\, A^*}.$$

Proof. We only prove that $\overline{\mathrm{Ran}(A)} = (\ker A^*)^\perp$ with the most direct argument.

If $x \in \mathrm{Ran}(A)$, then $x = Ay$ and for any $z \in \ker A^*$ we have

$$\langle x, z\rangle = \langle Ay, z\rangle = \langle y, A^* z\rangle = 0$$

which means that

$$\mathrm{Ran}\, A \subset (\ker A^*)^\perp$$

and since the orthogonal complement is closed so we just take the closure and the same

$$\overline{\mathrm{Ran}\, A} \subset (\ker A^*)^\perp.$$

Now for the other direction, for $x \in (\mathrm{Ran}\, A)^\perp$ then directly by this we have

$$\langle Ay, x\rangle = \langle y, A^* x\rangle$$

for any y. Thus $A^*x = 0$ which means $x \in (\ker A^*)$ and thus

$$(\mathrm{Ran}\, A)^\perp \subset \ker A^*$$

$$\Rightarrow \overline{\mathrm{Ran}(A)} = ((\mathrm{Ran}\, A)^\perp)^\perp \supset (\ker A^*)^\perp.$$

□

Theorem 14.2. *If A is bounded with closed range, then $Ax = y$ has a solution $\iff$ $y \in (\ker A^*)^\perp$.*

The proof is direct by above theorem.

Now notice that we've seen before that if an operator is bounded from below, then the range is closed.

Def 14.3. *(Fredholm Alternative) An operator $A : \mathcal{H} \to \mathcal{H}$ satisfies Fredholm Alternative if either (a) or (b) below holds.*

(a) $Ax = 0$ and $A^*x = 0$ have 0 as unique solution, and that $Ax = y$ and $A^*x = y$ have unique solution.

(b) $Ax = 0$ *and* $A^*x = 0$ *admits finite dimension spaces of solutions of the same positive dimension, and* $Ax = y$ *has solution* $\iff$ $(y \perp z, \forall A^*z = 0)$*; And* $A^*x = y$ *has solution* $\iff$ $(y \perp z, \forall Az = 0)$.

First we note that if the input dimension is different from the output dimension (non square), then neither can hold. It's just a definition so nothing to prove here. But we can ask whether we can extend the condition to infinite dimension, and the answer is no.

Let's see some examples.

Example 14.4.

- For an operator from $\mathbb{C}^n \to \mathbb{C}^n$, we know that Fredholm alternative holds since $\overline{\mathrm{Ran}(A)} = \mathrm{Ran}\, A$ and $\dim \mathrm{Ran}\, A = \dim \mathrm{Ran}\, A^*$.
- There's 2 issues with infinite dimension:
  (1) $\mathrm{Ran}\, A$ might not be closed; (For an example, integration takes us from L^1 to H^1.)
  (2) $\dim \ker A \neq \dim \ker A^*$.
- Consider the operator $Mf = x \cdot f$ that acts on $L^2([0,1])$. Then, since the inner product is
$$\langle f, g\rangle = \int \overline{f} g dx$$
we have
$$\langle f, Mg\rangle = \int_0^1 \overline{f} x g dx = \int_0^1 \overline{xf} g dx = \langle Mf, g\rangle$$
since $x \in \mathbb{R}$. That is, $M = M^*$.

  But note that $\mathrm{Ran}(M) \neq \overline{\mathrm{Ran}\, M} = L^2([0,1])$ since for any $\frac{c}{x^p}$ where $-\frac{1}{2} \leq p \leq \frac{1}{2}$, and in particular constant functions, if it is in the range then it's preimage is not in L^2, so it's not in the range. Moreover, $\overline{\mathrm{Ran}\, M} = L^2[0,1]$ since $\ker M = 0$.

  In particular, since it is self-adjoint with trivial kernel, it is Fredholm alternative.
- Shifts in l^2. Define $S : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ with
$$S(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$$
then
$$S^*(x_1, x_2, \dots) = (x_2, x_3, \dots)$$
since
$$\langle x, Sy\rangle = \sum_{i=1}^{\infty} \overline{x_{i+1}} y_i = \langle S^*x, y\rangle.$$

  Moreover, since $\ker(S) = \{0\}$ and $\ker(S^*) = (x_1, 0, 0, 0, \dots) \neq \{0\}$ (we know it has dimension 1), so the Fredholm alternative does not hold.

Def 14.5. *The bounded operator A is a* <u>*Fredholm operator*</u> *if*

(a) Ran A *is bounded.*
(b) ker A *and* ker A^* *are finite dimensional.*

Actually (a) is implied by (b) but we'll not prove that here.

Def 14.6. *The Fredholm index of a Fredholm operator is*

$$\text{Index } A = \dim\ker A - \dim\ker A^*.$$

Thus, if a Fredholm operator satisfies the Fredholm alternative, we have Index $A = 0$.

Proposition 14.7. *For A Fredholm, K compact, we have $A + K$ is Fredholm with*

$$\text{Index}(A + K) = \text{Index } A.$$

The idea of the above is to view K as a well-behaved perturbation. Then if adding K changes $\dim\ker A$, it also changes $\dim\ker A^*$. Just the idea.

For a better picture, consider A to be a matrix and $K = kI$, as k goes across eigenvalues of A, the dimension of the kernel will change abruptly, but for compact K we are really good since kernel of A^* also experience the dimension change.

We now look at some definitions of self-adjoint and unitary operators, as an analogous to linear algebra.

Def 14.8. *A is self-adjoint if $A = A^*$.*

Def 14.9. *A sesqui-linear form is an operator $a : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ by*

$$a(x, y) = \langle x, Ay \rangle.$$

We will try to find conditions that make this form an inner product. Note that for the positive definiteness of inner product we need A to be so, given the definition below. We will also need self-adjoint to satisfy the bi-symmetric condition.

To be specific, if $A = A^*$, then

$$a(x, y) = \langle x, Ay \rangle = \langle Ax, y \rangle = \overline{\langle y, Ax \rangle} = \overline{a(y, x)}$$

so it's indeed symmetric.

Def 14.10.

- *q is a quadratic form if $q(x) = a(x, x) = \langle x, Ax \rangle$.*
- *A is non-negative if $A = A^*$ and $\langle x, Ax \rangle \geq 0$.*
- *A is positive definite if it's non-negative and $\langle x, Ax \rangle > 0$ for $x \neq 0$.*

So we see that if A is positive definite, then the sesqui-linear form a is indeed an inner product, so the space is pre-Hilbert. If we add completeness to make it Hilbert, we can use Riesz to deal with problems, which is extremely convenient.

Lemma 14.11. *For A bounded, self-adjoint, we have*
$$||A|| = \sup_{||x||=1} |\langle x, Ax\rangle|.$$

Proof. To simplify things we again assume $\langle\cdot,\cdot\rangle \to \mathbb{R}$ and define $\alpha = \sup_{||x||=1} |\langle x, Ax\rangle|$.

Then
$$||A|| = \sup_{||x||=1} ||Ax|| \overset{*}{=} \sup_{||x||=||y||=1} |\langle y, Ax\rangle| \geq \alpha$$
where the equality * in middle is that
$$\sup_{||x||=||y||=1} |\langle y, Ax\rangle| \overset{C.S.}{\leq} \sup_{||x||=||y||=1} ||y|| \cdot ||Ax|| = \sup_{||x||=1} ||Ax||$$
(or just Hahn Banach) and for the other direction pick $y = \dfrac{Ax}{||Ax||}$.

Note that the above does not require anything about A. For the other direction we need to use $A = A^*$:
$$\begin{aligned} &\langle x+y, A(x+y)\rangle - \langle x-y, A(x-y)\rangle - i\langle x+iy, A(x+iy)\rangle + i\langle x-iy, A(x-iy)\rangle \\ =&2\langle x, Ay\rangle - 2\langle y, Ax\rangle - 2i\langle x, iAy\rangle + 2i\langle iy, Ax\rangle = 4\langle y, Ax\rangle \end{aligned}$$
where note that since A is self adjoint $\langle z, Az\rangle \in \mathbb{R}$ and so the first 2 terms above is real and the last 2 are imaginary. Now we can rotate y by letting $y = e^{i\theta}y$ for $\langle y, Ax\rangle \in \mathbb{R}$. This directly make the last 2 term vanish on the right. Also, this does not change the norm of y, which will be the only thing we're left with in the end. So now computation holds that
$$4|\langle y, Ax\rangle| = |\langle x+y, A(x+y)\rangle - \langle x-y, A(x-y)\rangle|$$

Now note
$$\langle z, Az\rangle = ||z||^2 \left\langle \frac{z}{||z||}, A\frac{z}{||z||} \right\rangle \leq \alpha||z||^2$$
by definition of α, so we can further write
$$\begin{aligned} 4|\langle y, Ax\rangle| &\leq |\langle x+y, A(x+y)\rangle| + |\langle x-y, A(x-y)\rangle| \\ &= \alpha\left(||x+y||^2 + ||x-y||^2\right) \overset{parallelogram}{=} 2\alpha\left(||x||^2 + ||y||^2\right) \end{aligned}$$
and thus under the constraints $||x|| = ||y|| = 1$ we get
$$|\langle y, Ax\rangle| \leq \alpha$$
and hence we are done. □

Corollary 14.12. *For A bounded, we have*
$$||A^*A|| = ||A||^2.$$
In particular, we have that when A is self-adjoint, $||A^2|| = ||A||^2$, i.e. applying A twice results in no cancellation.

Proof.

$$||A^*A|| = \sup_{||x||=1} \langle x, A^*Ax\rangle = \sup_{||x||=1} ||Ax||^2 = \left(\sup_{||x||=1} ||Ax||\right)^2 = ||A||^2.$$

□

Def 14.13. *An operator $U : \mathcal{H}_1 \to \mathcal{H}_2$ is unitary if it is invertible and*

$$\langle Ux, Uy\rangle_{\mathcal{H}_2} = \langle x, y\rangle_{\mathcal{H}_1}$$

i.e. applying U preserves energy.

Since applying U preserves energy, we will likely to see it in physics setting where there's no energy lost (no heat dissipation).

Def 14.14. *If there is unitary $U : \mathcal{H}_1 \to \mathcal{H}_2$ then we say that $\mathcal{H}_1$ and $\mathcal{H}_2$ are similar or*

$$\mathcal{H}_1 \cong \mathcal{H}_2.$$

Proposition 14.15. *$U : \mathcal{H}_1 \to \mathcal{H}_2$ is unitary $\iff$ $UU^* = U^*U = Id$, i.e. $U^* = U^{-1}$.*

Def 14.16. *S is skew self-adjoint if $S^* = -S$. We can also express this as $S = iA$ for A self adjoint.*

Def 14.17. *$T : \mathcal{H} \to \mathcal{H}$ is normal if $T^*T = TT^*$, or that their lie bracket*

$$[T^*, T] = 0.$$

Note that if A is self-adjoint/unitary/skewed self-adjoint, then A is normal.

14.2. **Recap of theorem in Hilbert space.**

Since $\mathcal{H}^* \cong \mathcal{H}$ by Riesz, we know that weak* convergence is weak convergence. Also, in this setting weak convergence can be written as

$$x_n \rightharpoonup x \iff \langle x_n, y\rangle \to \langle x, y\rangle, \quad \forall y \in \mathcal{H}.$$

Theorem 14.18. *(Uniform boundedness/ Banach Steinhaus) For $\phi_n : X \to \mathbb{C}$ and X Banach, if uniform in n we have $\{\phi_n\}$ is such that*

$$|\phi_n(x)| \leq c(x) \cdot ||x||, \quad \forall x \in \mathcal{H}$$

then $||\phi_n|| \leq C'$ uniform in n and x.

Look at proof in book. The idea is to first construct x_k Cauchy with $\phi(x_k) \geq k$ for contradiction, and use completeness to argue that the limit x does not satisfy the condition.

Theorem 14.19. *For $x_n \in \mathcal{H}$ and D dense in $\mathcal{H}$, we have*

$$x_n \rightharpoonup x \iff \begin{cases} (a): & ||x_n|| \leq M \\ (b): & x_n \rightharpoonup x \text{ in } D, i.e. \ \langle x_n, y\rangle \to \langle x, y\rangle, \quad \forall y \in D. \end{cases}$$

Proof.

($\Rightarrow$:) (b) is by definition, and (a) we can get by uniform boundedness theorem.

($\Leftarrow$:) For $z \in \mathcal{H}$, $\exists y \in D$ such that $||z - y|| \leq \varepsilon$, then

$$|\langle x_n - x, z\rangle| \leq |\langle x_n - x, y\rangle| + |\langle x_n - x, z - y\rangle| \to 0$$

where the first term goes to 0 by (b) and the second goes to 0 by Cauchy Schwartz plus that $||x_n - x|| \leq 2M + \delta$ for large enough n.

$\square$

15. 2/20: THERОREMS IN AN HILBERT SETTING; SPECTRUM OF BOUNDED OPERATOR

First, we note again that $\cos(nx) \rightharpoonup 0$ since all L^2 functions can be approximated by piece-wise constant functions, and Riemann Lebesgue lemma applies there. We view it again since now we really are justified to write weak convergence by inner products with L^2 functions, since the space is Hilbert.

Moreover, a counter example that shows why we need boundedness in theorem 14.19 is that consider l^2, then let $x_n = ne_n$ we will have that x_n is not bounded, and it's easy to check that it serves as a counter example. Since we can choose $y_n = 1/n$ to be the counterexample of the left side. As for why it satisfies (a) it is because we use the dense set that is the finite linear combination of the orthonormal basis.

Proposition 15.1. *Let $\{e_\alpha\}$ be a basis of $\mathcal{H}$ (not necessarily orthogonal), then we have that*

$$x_n \rightharpoonup x \iff \begin{cases} ||x_n|| \leq M \\ \langle e_\alpha, x_n \rangle \to \langle e_\alpha, x \rangle, \forall \alpha \in I. \end{cases}$$

Reason for this is that by unconditionally convergence we can always find the dense set

$$D = \left\{ y = \sum_{\text{finite}} y_\alpha e_\alpha \right\}.$$

Using this, we see that since $\cos(nx) = \dfrac{e^{inx} + e^{-inx}}{2}$ and $\langle e_\alpha, e^{\pm inx} \rangle \to 0$ as $n \to \infty$, we get our desired result with previous proposition.

Proposition 15.2.

- $x_n \rightharpoonup x \Rightarrow ||x|| \leq \liminf_{n\to\infty} ||x_n||$;
- $\big(x_n \rightharpoonup x \text{ and } ||x_n|| \to ||x||\big) \iff x_n \to x$.

Intuitively, this means that weak convergence plus the fact that no mass (norm) is lost implies strong convergence.

Proof.

- $$||x||^2 = \langle x, x \rangle = \lim_{n\to\infty} \langle x_n, x \rangle = \liminf_{n\to\infty} \langle x_n, x \rangle \leq \liminf_{n\to\infty} ||x_n|| \cdot ||x||$$
  where the lesson is that if we want to use inequalities to create something whose limit does not exists, we'd better use lim sup or lim inf judging by the direction we're using it.
- By definition of direct sum we have
  $$||x_n - x||^2 = ||x||^2 + ||x_n||^2 - 2\langle x, x_n \rangle \to 0$$
  since $||x_n||^2 \to ||x||^2$ by assumption and by weak convergence $\langle x, x_n \rangle \to ||x||^2$.

□

Theorem 15.3. *(Banach-Alaoglu) the closed unit ball is weakly compact, which notably is sequentially weakly compact, i.e. for* $||x_n|| \leq 1$ *we have*

$$x_{\phi(n)} \rightharpoonup x \in \mathcal{H}$$

in the weak sense.

Proof. Let $\{e_m\}$ be an orthonormal basis of a separable space $\mathcal{H}$. Then we know $\{e_m\}$ is countable since separable. For $||x_n|| \leq 1$ we have $|\langle e_1, x_n\rangle| \leq 1$ by C.S. and hence by Heine Borel on $\mathbb{R}$ we know that for some subsequence $x_{\phi_1(n)}$ we have $\langle e_1, x_n\rangle \to c_1$. Now we just use the diagonal argument to get a subseqeunce $x_{\phi_n\circ\cdots\circ\phi_1(n)}$ such that

$$\langle e_l, x_{\phi_k\circ\cdots\circ\phi_1(n)}\rangle = c_l$$

for $l \leq k$.

Now we find the diagonal by defining

$$\phi(y) = \lim_{k\to\infty}\langle e_l, x_{\phi_k\circ\cdots\circ\phi_1(n)}\rangle \in \mathcal{H}^*$$

which is defined because again we can write a dense subset by the finite linear combinations, then we use either prop 15.1 above or use Hahn Banach to extend this to the whole set.

Now the last step is to see that Riesz says $\phi(y) = \langle x, y\rangle$ for some $x \in \mathcal{H}$, for which we identify as the weak limit of x_n. □

15.1. **Chapter 9: Spectrum of bounded operators.**

Let's first see how this works in the familiar $\mathbb{C}^n$ matrix case. For the matrix A, we look at $Ax = \lambda x$ and use $\det(A - \lambda I) = 0$ to find eigenvalues, which is very hard to generate to meaningful expressions.

Theorem 15.4. *If A is normal, then $\mathbb{C}^n$ has an orthonormal basis of eigenvectors of A.*

Example 15.5.

A counter example to this is the nilpotent matrix

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Now let's generate to infinite dimension.

Def 15.6. *The* <u>*resolvent set*</u> *of $A \in \mathcal{B}(\mathcal{H})$ is $\rho(A) \subset \mathbb{C}$ the set of λ such that $A - \lambda I$ is invertible (1-1 and onto).*

The <u>*spectrum*</u> *of A is the rest, i.e. $\sigma(A) = \mathbb{C}\backslash\rho(A)$.*

Remark 15.7.

- If $Au = \lambda u$, then λ is an eigenvalue when $0 \neq u \in \mathcal{H}$.
- Note that if we take the operator $A = -\partial_x^2$, then we have

$$-\partial_x^2 e^{ikx} = k^2 e^{ikx}$$

for which we see a pattern of

$$Au = \lambda u$$

yet $u \notin \mathcal{H}$. What do we make of this? We still say that $\lambda \in \sigma(A)$ since the kernel is not trivial (in a broader setting). But this example tells us how infinite dimensional cases can be worse.

Def 15.8.

(a) *The point spectrum σ_p of A is the $\lambda \in \sigma(A)$ such that $A - \lambda I$ is not 1-1.*

(b) *The continuous spectrum σ_c is such that $A - \lambda I$ is 1-1 but not onto, and that it's range is dense in $\mathcal{H}$.*

Example: integration maps C into C^1 which is dense in C.

(c) *The Residual spectrum σ_r is such that $A - \lambda I$ is 1-1 but has non-dense range. In other words*

$$\overline{\mathrm{Ran}(A - \lambda I)} \oplus M^\perp = \mathcal{H}$$

for $M^\perp \neq \{0\}$.

"It's your fault! You choose too big a space!"

Now note that we're just finding equivalent classes of $\sigma(A)$ and each of them is allowed to be empty.

Example 15.9.

Remember the multiplication operator

$$M_x f(x) = x f(x)$$

where $M^* = M$. We write $M_g f(x) = g(x) f(x)$. Note that

$$(M - x_0) f(x) = (x - x_0) f(x)$$

and hence the operator's inverse, if well defined, should have that

$$(M - x_0)^{-1} f(x) = \frac{f(x)}{x - x_0}.$$

But of course there's a problem at $x = x_0$. But if we are in $L^2[0,1]$ then we are good when $x \notin [0,1]$, in which case $\mathbb{C} \backslash [0,1] \subset \rho(M_x)$, where we just care about M since we are viewing the x_0 as the eigenvalue λ.

Moreover, we can prove that $\sigma(M) = \sigma_c(M) = [0,1]$, for which the heuristic is to note

$$(M - x_0)\delta_0 = (x - x_0)\delta_{x_0} = 0$$

and hence by the reason similar to that in remark 15.7 it is in the spectrum.

Def 15.10. *The resolvent operator For $\lambda \in \rho(A)$, define*

$$R_\lambda = (\lambda - A)^{-1} := (\lambda I - A)^{-1} = \frac{1}{\lambda}\left(I - \frac{A}{\lambda}\right)^{-1} \in \mathcal{B}(\mathcal{H})$$

where the last step is again by open mapping theorem, and of course we write the last expression when $\lambda \neq 0$.

Proposition 15.11.

$$\mathbb{C}\backslash\{\lambda : |\lambda| \leq ||A||\} = \{\lambda : |\lambda| > ||A||\} \subset \rho(A).$$

The reason is that now we can really write

$$\frac{1}{\lambda}\left(I - \frac{A}{\lambda}\right)^{-1}$$

since $\left|\left|\frac{A}{\lambda}\right|\right| < 1$, so it's well defined.

Proposition 15.12. *R_λ is analytic in λ on $\rho(A)$, and $\rho(A)$ is open.*

Proof. We prove that $\rho(A)$ is open first. For $\lambda_0 \in \rho(A)$, we have the resolvent identity

$$\lambda - A = (\lambda_0 - A)(I - (\lambda_0 - \lambda)(\lambda_0 - A)^{-1}) =: (\lambda_0 - A)(I - K_\lambda).$$

Now, observe that this K_λ we've just defined, we have that we can choose λ close to λ_0 such that $||K_\lambda|| \leq 1$ as

$$|\lambda - \lambda_0| \cdot ||(\lambda_0 - A)^{-1}|| \leq 1$$

as the second is a bounded operator by open mapping.

But then we know $(I - K_\lambda)^{-1}$ is defined, and thus $(\lambda - A)^{-1}$ is defined as $(\lambda_0 - A)^{-1}(I - K_\lambda)^{-1}$, which means ρ is open.

Now, we can write explicitly using geometric expansion that

$$R_\lambda = \sum_{k=1}^{\infty}(K_\lambda)^k(\lambda_0 - A)^{-1}$$

and hence R_λ is analytic, as the summands above are rational polynomials. $\square$

For self adjoint A, we know that the spectrum is on the real line, and for unitary U, we know the spectrum is on the unit circle.

16. 2/22: Spectral theory for compact and self adjoint operators

Consider the operator M_x again (see ex 15.9). For $\lambda \in [0,1]$ we know that 1 is not in the range, since for it to be it would require that the preimage is $\dfrac{1}{x-\lambda}$ which is not in L^2. But we do know that the range is dense since we can use $\mathbb{1}_{|x-\lambda|>\varepsilon}$ to approximate.

What we will march onto today is that for compact operator, the spectrum σ is just points.

Proposition 16.1. $\sigma(A) \neq \emptyset$.

Proof.

By proposition 15.12, we know that $\lambda \mapsto R_\lambda$ is analytic, which means that

$$\langle x, R_\lambda y\rangle$$

is an analytic function in λ. But then we have

$$\left\langle x, \frac{1}{\lambda}\left(I - \frac{A}{\lambda}\right)^{-1} y\right\rangle \to 0$$

as $|\lambda| \to \infty$. But now assume $\sigma(A) = \emptyset$, then we know that the inverse is defined on the whole complex plane, so the above function is an entire bounded function, which by Liouville is constant. Moreover, the value of the constant is 0 due to its asymptotic behavior. Thus

$$\langle x, R_\lambda y\rangle = 0$$

for all x, y, which means that $R_\lambda = 0$, which is just impossible for varying λ. Thus contradiction! So $\sigma(A) \neq \emptyset$. $\square$

Lemma 16.2. *If $A = A^*$, then $\lambda \in \sigma_p(A) \Rightarrow \lambda \in \mathbb{R}$. Moreover for $u, v \in \mathcal{H}$ and mu $\neq \lambda$*

$$\begin{cases} Au = \lambda u \\ Av = \mu v \end{cases} \Rightarrow (u, v) = 0.$$

Proof.

By definition, the point spectrum is the set of λs such that the function is not 1-1, i.e. the kernel is not $\{0\}$. Hence, there exist non-zero eigenvector within the same space. For that eigenvector, call it x, we have

$$\lambda(x,x) = (x, \lambda x) = (x, Ax) = (Ax, x) = (\lambda x, x) = \overline{\lambda}(x,x)$$

where since $x \neq 0$ we have $\lambda = \overline{\lambda} \Rightarrow \lambda \in \mathbb{R}$.

Now for different u, v we have (since eigenvalues are real)

$$\lambda(u,v) = (\lambda u, v) = (Au, v) = (u, Av) = (u, \mu v) = \mu(u,v)$$

which thus means $(u, v) = 0$. $\square$

Def 16.3. *$\mathcal{M} \subset \mathcal{H}$ is <u>invariant</u> of A if $A\mathcal{M} \subset \mathcal{M}$.*

Lemma 16.4. *For $A = A^*$ and $\mathcal{M}$ is invariant for A. Then so is $\mathcal{M}^\perp$, i.e. $A\mathcal{M}^\perp \subset \mathcal{M}^\perp$.*

Note that what this means is that if we reorder the basis of $\mathcal{H}$ with respect to $\mathcal{M}$ and $\mathcal{M}^\perp$, we will have

$$A = \begin{pmatrix} A\big|_{\mathcal{M}} & 0 \\ 0 & A\big|_{\mathcal{M}^\perp} \end{pmatrix}.$$

Proof. For $x \in \mathcal{M}^\perp$, we want to show that $Ax \in \mathcal{M}^\perp$, that is, we want to show for any $y \in \mathcal{M}$ we have $(Ax, y) = 0$. But we have

$$(Ax, y) = (x, Ay) = 0$$

since $x \in \mathcal{M}^\perp$ and $Ay \in \mathcal{M}$. □

Proposition 16.5.

$$\lambda \in \sigma_r(A) \Rightarrow \overline{\lambda} \in \sigma_p(A^*).$$

Proof. For $\lambda \in \sigma_r(A)$, we can write things as

$$\overline{\mathrm{Ran}(A - \lambda I)} \oplus \mathcal{M}^\perp = \mathcal{H}$$

for $\mathcal{M} := \overline{\mathrm{Ran}(A - \lambda I)}$ and $\mathcal{M}^\perp$ non-trivial.

But by decomposition of Hilbert spaces we have

$$\mathcal{H} = \overline{\mathrm{Ran}(A - \lambda I)} \oplus \ker((A - \lambda I)^*)$$

where we note $(A - \lambda I)^* = A^* - \overline{\lambda} I$. But then

$$0 \neq x \in \mathcal{M}^\perp \iff x \in \ker(A^* - \overline{\lambda} I)$$

Thus the existence of non-trivial $\mathcal{M} \perp$ means that $\lambda \in \sigma_p(A^*)$. □

Corollary 16.6.

$$A = A^* \Rightarrow \sigma_r(A) = \emptyset.$$

Proof.

The proof is direct because now if $\lambda \in \sigma_r(A)$ then $\lambda \in \sigma_p(A^*) = \sigma_p(A)$ which is disjoint from $\sigma_r(A)$. □

Lemma 16.7. *Assume $A = A^*$, then $\sigma(A) \subset [-||A||, ||A||] \subset \mathbb{R}$.*

Proof.

For contradiction, assume $\lambda = a + ib \in \sigma(A)$ for $b \neq 0$. But now we know that $\lambda \notin \sigma_p$ and $\lambda \notin \sigma_r(A)$ since it's empty. So $\lambda \in \sigma_c$. So what we have is that $(\lambda - A)$ is 1-1 but not onto, yet the range is dense. So we show that the range is closed to create contradiction. We have

$$||(A - \lambda)x||^2 = ||(A - a)x - ibx||^2 = ||(A - a)x||^2 + ||bx||^2 \geq |b|^2||x||^2$$

where the middle equality is because we note that the imaginary part fades as the left most is a real value. But now we see that $(A - \lambda)$ is bounded below and hence Ran A is closed. Contradiction. $\square$

16.1. **As for Compact operators.**

Recall that compact operators maps bounded sets to bounded sets.

Proposition 16.8. *For A compact and $A = A^*$, we have that if $\lambda \in \sigma_p(A)$ with $\lambda \neq 0$, then λ has only finite multiplicity.*

Now, for $\lambda_n \in \sigma_p(A)$ and $\lambda_n \neq 0$ where we cannot double count the same element up to multiplicity, then if $\lambda_{\phi(n)} \to \lambda$, we know $\lambda = 0$. In other words, 0 is the only accumulation point of σ_p.

Proof.

First for the same eigenvalue. Let $Au_k = \lambda u_k$ for u_k a basis of an infinite eigenspace, and assume $||u_k|| = 1$. But since we are in the unit ball and A is compact, we have that there's a subsequence $\phi(k)$ such that

$$\lambda u_{\phi(k)} = Au_{\phi(k)} \to y \in \mathcal{H}$$

yet we also know that for a basis, the first term $\lambda u_{\phi(k)}$ is not Cauchy (take difference and square). Thus any fixed λ cannot have infinite multiplicity.

Now for varying λ_n where for m a subsequence of n and $\lambda_m \to \lambda$, we have that λ_m is bounded since convergent. Now we have

$$Ae_m = \lambda_m e_m$$

and if we define

$$f_m = \frac{e_m}{\lambda_m} \Rightarrow Af_m = e_m.$$

Assume that $\lambda \neq 0$, then since $\dfrac{1}{\lambda_m} \to \dfrac{1}{\lambda} \neq 0$ and hence $\dfrac{1}{\lambda_m}$ is bounded both above and below by some neighborhood of $\dfrac{1}{\lambda}$. Thus we have

$$||f_m|| \leq c$$

and this again by compactness means that there's a subsequence $\phi(m)$ such that

$$Af_{\phi(m)} \to y$$

and yet $Af_{\phi(m)} = e_{\phi(m)}$ is not Cauchy. Contradiction! $\square$

Theorem 16.9.

- *There exists an orthonormal basis of $\mathcal{H}$ that is composed of eigenvectors of A.*

- *The non-zero eigenvalues of A forms at most countably infinity set of real numbers such that*
$$A = \sum_{k\in\mathbb{N}} \lambda_k P_k$$
with P_k being orthogonal projectors.

First note that the second statement above directly means that
$$I = \sum_{k\in\mathbb{N}} P_k + P_0$$
and for any function that can be approximated by polynomials we have
$$f(A) = \sum_{k\in\mathbb{N}} f(\lambda_k)P_k.$$

We will see most of the proof today and finish it next week.

Proof.

Step 1: Either $||A||$ or $-||A||$ are in $\sigma_p(A)$:

Assume $A \neq 0$ since other case is obvious. Then
$$||A|| = \sup_{||x||=1} |\langle x, Ax\rangle| = \lim_{n\to\infty;||x_n||=1} |\langle x_n, Ax_n\rangle|$$
where since x_n is a bounded sequence and so $|\langle x_n, Ax_n\rangle|$ is bounded in $\mathbb{R}$, which means that there exists a $\lambda = \pm||A||$ with $\langle x_n, Ax_n\rangle \to \lambda$ up to a subsequence by compactness of $\mathbb{R}$.

Now we use the compactness of A to conclude that there is a subsequence of the above subsequence $\phi(n)$ such that $Ax_{\phi(n)} \to y \in \mathcal{H}$ where $y \neq 0$ since λ is not. What we should think of this that we've constructed an eigenvalue and an eigenvector. To check this we compute
$$\begin{aligned} ||(A-\lambda)y||^2 &= \lim_{n\to\infty} ||(A-\lambda)Ax_{\phi(x)}||^2 \\ A \text{ and } A-\lambda I \ \text{ commutes} \quad &\leq ||A||^2 \lim_{n\to\infty} ||(A-\lambda)x_{\phi(x)}||^2 \\ A = A^* \quad &\leq ||A||^2 \lim_{n\to\infty} \left(||Ax_{\phi(n)}||^2 + \lambda^2||x_{\phi(n)}||^2 - 2\lambda\langle x_{\phi(n)}, Ax_{\phi(n)}\rangle\right) \end{aligned}$$
where for the last line we bound $||Ax_{\phi(n)}||^2 \leq ||A||^2$, use length of $x_{\phi(n)}$ to reduce $\lambda^2||x_{\phi(n)}||^2$ to λ^2, and use $\langle x_{\phi(n)}, Ax_{\phi(n)}\rangle \to \lambda$ by passing the limit and above result. Then we get
$$||(A-\lambda)y||^2 \leq ||A||^2(||A||^2 - \lambda^2) = 0.$$

But this means that λ is an eigenvalue, and hence we are done with step 1.

But the rest is to decompose this one by one, use finite multiplicity to always lower the largest eigenvalue, and use not accumulation point to decrease the largest eigenvalue to length 0.

Step 2: peeling of eigenvectors:

Let $\mathcal{N}_1 = \mathcal{H}$ and $A_1 = A$. For the eigenvector $e_1 = y$ that we've just found above, we know $\mathbb{R}e_1$ is an invariant space and so we can decompose

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

that is, we define $\mathcal{N}_1 = \mathcal{H}, \mathcal{M}_1 = \mathbb{R}e_1, \mathcal{N}_2 = \mathcal{M}_1^\perp$, and so on and so forth.

Moreover, we call the new matrix A_{i+1} with $||A_{i+1}|| \le ||A_i||$ by construction.

<u>Step 3: If $A_{n+1} = 0$, then we are done:</u>

Now $A = \sum_{j=1}^n \lambda_k e_k \otimes e_k = \lambda_k P_k$ and

$$A = \begin{pmatrix} \mathrm{Diag}(\lambda_1, \dots, \lambda_n) & 0 \\ 0 & 0 \end{pmatrix}$$

and since $\mathcal{N}_{n+1}$ is still Hilbert so we can just find an orthonormal basis to it.

<u>Step 4: If $A_n \neq 0, \forall n$:</u>

The since λ_n is a decreasing sequence of eigenvalues, so there is a limit. Moreover, the limit must be 0 since that's the only accumulation point. So we have

$$A = \sum_{k=1}^n \lambda_k e_k \otimes e_k + A_{n+1}$$

where $||A_{n+1}|| = |\lambda_k| \to 0$. This implies

$$\lim_{n\to\infty} \left\| A - \sum_{k=1}^m \lambda_k P_k \right\| = 0 \iff A = \lim_{n\to\infty} \sum_{k=1}^n \lambda_k P_k.$$

<u>Step 5: wrap up (on Monday):</u>

So we know already that we have an orthonormal basis in whose completion the closure of range of A is contained. But we have that the space is Hilbert so we have

$$\mathcal{H} = \overline{\mathrm{Ran}\, A} \oplus \ker A$$

where $\overline{\mathrm{Ran}\, A} \subset [\{e_i\}]$ and since $\ker A$ is a Hilbert space, we can find an orthonormal basis $\{f_i\}$ of it with $Af_n = 0$. Then, $\{e_1, e_2, \dots, f_1, \dots\}$ is an orthonormal basis of $\mathcal{H}$.

$\square$

17. 2/27: Compact operators

First, let's see a fact about compact operators that are not just finite dimensional matrix like:

Remark 17.1. *We conclude here that if A is compact on an infinite dimensional space, then it's range is never closed.*

To show this, take $x \in [\{e_i\}] \subset \mathcal{H}$, then we can express

$$x = \sum_{k=1}^{\infty} b_k e_k$$

where by Parseval we have

$$\sum_{k=1}^{\infty} |b_k|^2 = ||x||^2$$

and we (skip details) assume that we can exchange the operator A with the infinite sum to get

$$Ax = \sum_{k=1}^{\infty} b_k(Ae_k) = \sum_{k=1}^{\infty} b_k \cdot \lambda_k e_k = \sum_{k=1}^{\infty} c_k e_k.$$

So then we have

$$\sum_{k=1}^{\infty} \left|\frac{c_k}{\lambda_k}\right|^2 = \sum_{k=1}^{\infty} |b_k|^2 = ||x||^2$$

where we have the Picard's Criterion:

$$\text{Ran } A = \left\{ \sum_{k=1}^{\infty} c_k e_k ; \sum_{k=1}^{\infty} \left|\frac{c_k}{\lambda_k}\right|^2 < \infty \right\}$$

where we note that since λ_i eventually is small, and we need $\dfrac{c_k}{\lambda_k}$ square summable, we will have to need c_k convergent much faster than just itself square summable, which is the whole

$$[\{e_i\}] = \left\{ \sum_{k=1}^{\infty} c_k e_k ; \sum_{k=1}^{\infty} |c_k|^2 < \infty \right\}.$$

Note that we always have any finite combination is in Ran A, which we know is dense in $[\{e_i\}]$, and hence Ran A is dense in the combinations, so the closure is exactly that.

Now we look at some examples, which starts with a theorem!

Theorem 17.2. *(Compact criterion) For a subset $E \subset \mathcal{H}$ we have*

(a) *If E is precompact, then for all orthonormal basis $\{e_n\}$ and all $\varepsilon > 0$, there exists N such that*

$$\sum_{n=N+1}^{\infty} \langle e_n, x \rangle^2 < \varepsilon, \forall x \in E.$$

(b) *If E is bounded and there exists an orthonormal basis $\{e_n\}$ such that for all $\varepsilon > 0$, there exists N such that*

$$\sum_{n=N+1}^{\infty} \langle e_n, x \rangle^2 < \varepsilon, \forall x \in E$$

holds, then E is precompact.

Notably that this (b) here is the exact theorem as Ascoli-Arzela, where we can view the "equi" part being the uniform constraint $\forall x \in E$.

"Since we didn't prove Ascoli-Arzela either let's also skip the proof of this."

Example 17.3.

Define $A : l^2 \to l^2$ by

$$A(x_1, x_2, \dots) = (\lambda_1 x_1, \lambda_2 x_2, \dots)$$

so that we note it really is a diagonal operator for $\lambda_k \in \mathbb{C}$ and $\lambda_k \to 0$. Now we can just check that it satisfies the condition b above to conclude that it's a compact operator.

Well to show this we need to say that the image of any bounded set under A is pre-compact, for which we only need to check the tail of the square sum vanishes by above (b). Yet that is because $\lambda_k \to 0$ and $\langle Ax, e_n \rangle = \lambda^2 x_n^2$ whose sum goes to 0 as $x \in l^2$. So we are done.

Def 17.4. *A bounded operator is a* <u>*Hilbert Schmidt operator*</u> *if* $\exists \{e_n\}$ *an orthonormal basis such that*

$$\sum_{n=1}^{\infty} ||Ae_n||^2 < \infty.$$

Moreover, define the <u>*Hilbert-Schmidt norm*</u> *as*

$$||A||_{H.S.} = \left(\sum_{n=1}^{\infty} ||Ae_n||^2 \right)^{\frac{1}{2}} = \left(\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |A_{n,m}| \right)^{\frac{1}{2}}$$

where $Ae_n = \sum_{m \geq 1} A_{n,m} e_n$.

The idea is nothing but that the eigenvalues are square summable. Note that if we use example 17.3 then the condition for H.S. is really just λ_i are square summable.

Theorem 17.5. *A Hilbert Schmidt operator is compact.*

Proof. Let $\{e_n\}$ be an orthonormal basis and define

$$P_N = \sum_{n=1}^{N} e_n \otimes e_n$$

then P_N is compact since the range is finite dimensional, and we can use Heine Borel on that setting.

Now, recall that we've once proved that if $T_n \to T$ uniformly for compact T_n, we have T is compact. So we only need to prove that

$$P_N A \to A$$

uniformly.

The idea is that in Hilbert spaces, finite rank approximation is sufficient.

To show the above we write (again, assume the operator can pass the infinite sum)

$$(I - P_N)Ax = \sum_n x_n (I - P_N)Ae_n = \sum_n \sum_m x_n A_{n,m}(I - P_N)e_m$$

where the last step is just decomposing Ae_n. So we further have

$$||(I - P_N)Ax||^2 = \sum_{m \geq N+1} \left| \sum_n x_n A_{n,m} \right|^2 \overset{C.S.}{\leq} ||x||^2 \left(\sum_{m \geq N+1} \sum_n |A_{n,m}|^2 \right) \to 0$$

where the first equality we use the fact that

$$(I - P_N)e_m = \begin{cases} 0 & m \leq N \\ e_m & m \geq N+1 \end{cases}$$

and the last approximation is as $m \to \infty$, since we know that the sum

$$\left(\sum_m \sum_n |A_{n,m}|^2 \right) < \infty$$

and for the Cauchy Schwartz we mean

$$\left| \sum_n x_n A_{n,m} \right|^2 = \langle x, A_m \rangle^2 \leq ||x||^2 \cdot ||A_m||^2 = ||x||^2 \cdot \sum_n |A_{n,m}|^2.$$

□

Now we observe a typical example:

Example 17.6.

Let $\mathcal{H} = L^2(0,1)$ and

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

where we assume

$$\int_0^1 \int_0^1 k(x,y)dxdy < \infty$$

then K is Hilbert Schmidt.

Proof.

For $\{e_n\}$ any orthonormal basis of $L^2(0,1)$, we have

$$k(x,y) = \sum_{i,j=1}^{\infty} k_{ij} e_i(y) e_j(x)$$

where we get this because we can first freeze y, then decompose for x, then do for y. The only thing we need is that $L^2(0,1)^2 = \{e_i\} \times \{e_j\}$.

But then we have

$$\int_0^1 \int_0^1 |k|^2 = \sum_{i,j} |k_{ij}|^2 = \sum_{j=1}^{\infty} ||Ke_j||^2 < \infty$$

since

$$\sum_{i=1}^{\infty} k_{ij} e_i(y) = a_j$$

where

$$Kf = \sum_{j=1}^{\infty} a_j e_j$$

and thus

$$Ke_j = a_j e_j$$

and taking absolute values and squares everything matches up.

Maybe try with Fourier basis. □

According to above, we can get the intuition that $k(x,y)$ smoother corresponds to the operator being more compact, i.e. λ_k converges faster.

If A is Hilbert-Schmidt, then it's compact so is A^* and so is A^*A, which is also self-adjoint, so it has eigenvalues.

Def 17.7. *The singular values of A are the square roots of the eigenvalues of A^*A.*

Moreover, note that we have

$$||A||_{H.S.}^2 = \sum_{n\geq 1} \lambda_n^2 < \infty.$$

We have the following classification of compact operators:

- If $\lambda \to 0$ then the operator is compact.
- If $\sum_n \lambda_n^2 < \infty$ then the operator is Hilbert-Schmidt.
- If $\sum_n \lambda_n < \infty$ then the operator is trace class.

In the end let's do a theorem.

Theorem 17.8. *For A bounded linear operator, it is compact iff for all $x_n \rightharpoonup x$ we get $Ax_n \to Ax$ strongly.*

Proof.

$\Rightarrow$:

Assume $x_n \rightharpoonup x$, then by uniform boundedness we have that x_n are bounded. So we have for any y, $\langle Ax_n - Ax, y\rangle = \langle x_n - x, A^* y\rangle \to 0$ where $A^* y$ is well defined due to open mapping. That is, $Ax_n \rightharpoonup Ax$. But now since x_n is bounded and A is compact, so $\{Ax_n\}$ has a convergent subsequence that converge strongly.

But note that if it converge, it converge to the same limit x, i.e. if a subsequence converge, then the whole thing converges, since if there is any subsequence that doesn't converge to x, then it attains also a convergent subsequence which converge to x, so we are done.

$\Leftarrow$:

For E bounded, consider $y_n = Ax_n \in A(E)$. By Banach Alaoglu, we know that there exists subsequence $x_{\phi(n)} \rightharpoonup x$ in $\overline{E}$ such that $Ax_{\phi(n)} \to Ax$, then we are done since this means that $A(E)$ is precompact.

$\square$

18. 3/1: UNBOUNDED OPERATORS; WEAK DERIVATIVE

From homework we know that the integral operator $f \mapsto \int_0^x f$ is a bounded compact operator on $L^2[0,1]$. But what about the inverse, i.e. derivative?

Well, if we have a good space, for instance C^∞ then ok the operator $f \mapsto \frac{\partial}{\partial x} f$ is a bounded operator. One might also say that it is bounded from $H^1 \to L^2$, which is "bounded" on its domain.

This leads to the definition of unbounded operators.

Def 18.1. *An unbounded operator is an operator*

$$A : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$$

where we note that the domain $\mathcal{D}(A)$ is a part of the definition.

Moreover, for most time we assume that $\overline{\mathcal{D}(A)} = \mathcal{H}$, since otherwise we just find a smaller $\mathcal{H}$.

To make matters clearer, we show by the following example, which we'll use constantly throughout this class, to show properties of unbounded operators.

Example 18.2.

Let A_1 through A_5 be unbounded operators defined by

$$A_{1,2,3,4,5}u = u'' = \frac{\partial^2}{\partial x^2}u$$

where $\mathcal{H} = L^2[0,1]$. They have different domains so they are different operators:

- $\mathcal{D}(A_1) = \left\{u \in C^2[0,1]; u(1) = u(0) = 0\right\}$;
- $\mathcal{D}(A_2) = C^2[0,1]$;
- $\mathcal{D}(A_3) = \left\{u \in \mathcal{H}^2[0,1]; u(1) = u(0) = 0\right\}$;
- $\mathcal{D}(A_4) = \mathcal{H}^2[0,1]$;
- $\mathcal{D}(A_5) = \left\{u \in \mathcal{H}^2[0,1]; u(1) = u(0) = u'(1) = u'(0) = 0\right\}$.

Note that u'' is not defined for the general case, but it is L^2 if we start in H^2, and continuous if we start in C^2.

Curiously, there's no constraints on the boundedness of an unbounded operator, so we actually have the following funny remark:

Remark 18.3. *A bounded operator is a bounded unbounded operator.*

Some preview of what's to come next is that we'll see that A_3 is self-adjoint. Moreover, the goal of this class is in the end to get to the fact that the operator $\frac{1}{i}\frac{\partial}{\partial x}$ is self-adjoint.

But first let's try to define the self-adjoint operator, which is not as trivial.

As some heuristics, we want to have $\langle Ax, y\rangle = \langle x, A^* y\rangle$, for $\forall x \in \mathcal{D}(A)$ and $\forall y \in \mathcal{D}(A^*)$. Then we see that if we have $\mathcal{D}(A^*)$ very small then this of course can work, but we really want to find the largest of such valid domains.

To get to that, we note that $\phi_y(x) = \langle y, Ax\rangle$ is a functional $\phi_y : \mathcal{D}(A) \to \mathbb{R}$ that is linear. Remember that the existence of such an adjoint is due to Riesz for bounded operators, and to use that we'll need $\mathcal{D}(A^*)$ to be defined such that all $y \in \mathcal{D}(A^*)$ we know ϕ_y is bounded.

Now if we further have $\overline{\mathcal{D}(A)} = \mathcal{H}$ is dense in the space, we can extend ϕ_y to $\mathcal{H}$ using Hahn Banach to the extension $\tilde{\phi}_y$.

Now we can apply Riesz: we know $\tilde{\phi}_y$ is a functional so we have

$$\tilde{\phi}_y(x) = \langle z, x\rangle$$

for some z that we defined to be the value of $A^* y$. Summing up all these we get the definition:

Def 18.4. *For an unbounded operator $A : \mathcal{D}(A) \to \mathcal{H}$ with $\overline{\mathcal{D}(A)} = \mathcal{H}$, we define the adjoint $A^* : \mathcal{D}(A^*) \to \mathcal{H}$ for*

$$\mathcal{D}(A^*) = \left\{ y \in \mathcal{H} \middle| \exists ! z \in \mathcal{H} \;\; s.t. \;\; \forall x \in \mathcal{D}(A), \langle Ax, y\rangle = \langle x, z\rangle \right\}$$

and for $y \in \mathcal{D}(A^*)$ we define $A^*(y) = z$ where z is the unique element as captured by the definition corresponding to y.

Def 18.5. *An unbounded operator A is self adjoint if $A = A^*$. In particular they have to satisfy $\mathcal{D}(A) = \mathcal{D}(A^*)$, which is the harder part for most cases.*

We thus realize that there's some sort of balance between $\mathcal{D}(A)$ and $\mathcal{D}(A^*)$, for which it cannot be that both is big. What we strive for is self adjoint, so they are equal. Then the goal is to find

- $\mathcal{D}(A)$ not too big for A to be well-defined;
- $\mathcal{D}(A)$ not too small for $\mathcal{D}(A^*)$ to be not too big.

We might ask that what if $\mathcal{D}(A^*) \neq \mathcal{D}(A)$ even though the switching in the inner product hold. If $\mathcal{D}(A^*)$ is larger then we define things as follows.

Def 18.6. *$\tilde{A}$ is an extension of A if*

$$\tilde{A}(x) = A(x), \forall x \in \mathcal{D}(A)$$

and $\mathcal{D}(A) \subset \mathcal{D}(\tilde{A})$.

Def 18.7. *A is symmetric if A^* is an extension of A.*

In practice we should be half-happy about finding symmetric operators, since symmetric is half-self-adjoint, as the only problem is $\mathcal{D}(A)$ is too small.

And back to example 18.2, we can check that

$$A_1^* = A_3 = A_3^*$$

so A_1 is symmetric and A_3 is self-adjoint (hence also symmetric).

But we only have

$$A_2^* = A_4^* = A_5; A_5^* = A_4$$

and hence A_5 is symmetric but not self-adjoint, whereas A_2, A_4 are not symmetric.

We view it like this:

- The domain of A_2, A_4 are too big due to lack of boundary conditions, and in this case the domain of adjoint is small, hence they are not symmetric;
- The domain of A_1 is too small with respect to the background space ($\mathcal{C}^2$) is too small, but with the right boundary conditions. In this case it's not self-adjoint.
- We'll seek for A_1-like operators in practice: with right boundary conditions and probably easier background space. Then we can take adjoint to get to the suitable background space, with that good boundary condition.

Now we generalize some of the other stuffs we've learned before, and starting with open mapping, which is not always true here.

Remember we've said that open mapping and closed graph are essentially the same, which is what we'll use here.

Def 18.8. *An operator $A : \mathcal{D}(A) \to \mathcal{H}$ is closed if it's graph is closed, i.e. if $\forall x_n \in \mathcal{D}(A)$ such that $x_n \to x$ and $Ax_n \to y$, we have that $x \in \mathcal{D}(A)$ and $Ax = y$.*

Def 18.9. *The graph $\Gamma(A)$ of A is a subset of $\mathcal{H} \times \mathcal{H}$ with*

$$\Gamma(A) = \left\{ (x, y) \middle| x \in \mathcal{D}(A), y = Ax \right\}.$$

Remark 18.10. *A is closed $\iff$ $\Gamma(A)$ is closed.*

Def 18.11. *A is closable if $\forall x_n \in \mathcal{D}(A)$ with $x_n \to 0$ and $Ax_n \to y$, we have $y = 0$.*

The idea is that if we consider $x_n \to x$ and replace $x_n - x \rightsquigarrow x_n$ we'll have the above. Just an idea.

Def 18.12. *The closure $\overline{A}$ of A has domain*

$$\mathcal{D}(\overline{A}) = \{x : \exists x_n \in \mathcal{D}(A), x_n \to x, Ax = y\}.$$

Proposition 18.13.

- *The graph of $\overline{A}$ is $\overline{\text{graph}(A)}$ and $\overline{A}$ is the smallest extension of A.*
- *Symmetric operators are closable.*

Def 18.14. *A is essentially self adjoint if its closure is self adjoint.*

Back to our example 18.2, we know that A_1 is essentially self adjoint since it's closure is A_3. Moreover, A_1, A_2 are not closed, but they are closable, and their closure A_3, A_4 are closed.

Since we're talking we have the result A_3 has only discrete spectrum since $(\lambda - A_3)^{-1}$ is compact, but that's too much.

Now marching to spectrum theory, if $A : \mathcal{D}(A) \to \mathcal{H}$ is 1-1 and onto, hence invertible, we can define it's inverse in the way we expect:

$$A^{-1}y = x \iff Ax = y.$$

Proposition 18.15. *If A is closed, then so is A^{-1}.*

Proof. This is actually because A and A^{-1} have just the same graph, if you really think of it.

To be specific, we write out

$$\begin{aligned}\Gamma(A) &= \{(x, y) | x \in \mathcal{D}(A), y = Ax\} = \{(x, y) | y \in \mathcal{H} = \mathcal{D}(A^{-1}), x = A^{-1}y\} \\ &= \{(y, x) | y = Ax, x \in \mathcal{D}(A)\} = \Gamma(A^{-1}).\end{aligned}$$

□

A surprising result is that then A^{-1} is bounded by closed graph theorem! Moreover, if $\mathcal{D}(A) = \mathcal{H}$ then we apply it again to A^{-1} and get that A is bounded. (Maybe this is similar to Liouville.)

Proposition 18.16. *If $A : \mathcal{D}(A) \to \mathcal{H}$ is a densely defined linear operator on $\mathcal{H}$ with bounded inverse $A^{-1} : \mathcal{H} \to \mathcal{H}$, then $(A^*)^{-1} = (A^{-1})^*$.*

Now some remarks that we go over quickly:

Remark 18.17.

- If $\lambda \in \rho(A)$, we can define the resolvent operator

$$R_\lambda(\lambda - A)^{-1}$$

  where if $A - \lambda I$ is 1-1, onto from $\mathcal{D}(A) \to \mathcal{H}$ (since $\mathcal{D}(A - \lambda I)$ is the same), then R_λ is bounded by closed graph theorem.
- Unbounded operators may have empty spectrum, since Liouville cannot apply to places that is not in the domain.

18.1. **Weak derivative.**

Recall that in the definition of the norm of H^1, we defined it based on that on $C^1[0, 1]$, which is $\left(||u||^2 + ||u'||^2\right)^{\frac{1}{2}}$, and then take the closure.

But there is a better definition!

We first define a set of test functions $\phi \in C_c^\infty(\mathbb{R})$ (Schwartz class), which is dense in $L^2(\mathbb{R})$.

Let's go off a little bit to see why it's dense. We use the convolution with a smooth bump. As the bump gets closer and closer to the delta function, we get that the convolution gives us an L^2 function. In particular, we can use this approximating identity:

$$\int_{\mathbb{R}} \frac{1}{\varepsilon}\phi\left(\frac{x-y}{\varepsilon}\right) u(y)dy \to u(x)$$

for $u \in L^2$.

Now back to derivatives. Assume $u \in C^1(\mathbb{R})$, then $v = u'$ implies that for $\forall \phi \in C_c^\infty(\mathbb{R})$, integral by parts means

$$\int_{\mathbb{R}} v\phi dx = -\int_{\mathbb{R}} u\phi' dx$$

since integrable functions obviously have boundary vanishes. This is an if and only statement, and the backwards direction is nothing but getting

$$\int_{\mathbb{R}} \phi(u' - v)dx = 0, \quad \forall \phi \in C_c^\infty(\mathbb{R}).$$

But this gives us a portrait of derivatives in a different angle: through test functions.

Def 18.18. *$u \in L^2(\mathbb{R})$ has weak derivative $v = u' \in L^2(\mathbb{R})$ if $\forall \phi \in C_c^\infty(\mathbb{R})$ we have*

$$\langle v, \phi \rangle = -\langle u, \phi' \rangle.$$

With this, we can formally define H^k.

Def 18.19. *The Hilbert space*

$$H^k(\mathbb{R}) := \left\{ u \in L^2(\mathbb{R}) \middle| u', \ldots, u^{(k)} \in L^2(\mathbb{R}) \right\}.$$

Proposition 18.20. *The operator $\frac{d}{dx} : \mathcal{D}\left(\frac{d}{dx}\right) = H^1 \to L^2(\mathbb{R})$ is a closed operator.*

Proof. Let $u_n \to u$ and $u_n' \to v$, since L^2 is closed so it's well defined, so we only need to show $u \in H^1(\mathbb{R})$ and that $u' = v$. But we have

$$\langle v, \phi \rangle \leftarrow \langle u_n', \phi \rangle = -\langle u_n, \phi' \rangle \to -\langle u, \phi' \rangle$$

and hence by taking the limit we get

$$\langle v, \phi \rangle = -\langle u, \phi' \rangle$$

which solves both problems, since $v \in L^2$ so $u \in H^2$ and $u' = v$. $\square$

Note that the closeness of $\frac{d}{dx}$ implies that H^k is a Hilbert space.

Theorem 18.21. *C_c^∞ is dense in H^k, for all k.*

Corollary 18.22. *For $u, v \in H^1(\mathbb{R})$ we have*

$$\int (uv' + u'v)dx = 0.$$

The proof is almost straight forward: they should decay and ibp tells us the result. The caveat being we need to use C_c^∞ to approximate u, v such that there's true decay there.

We end the course with one last result:

Theorem 18.23. $A = \dfrac{1}{i}\dfrac{\partial}{\partial x} : H^1 \to L^2(\mathbb{R})$ *is self adjoint.*

APPENDIX A. A

APPENDIX B. B

APPENDIX C. C

Acknowledgements.