Contents

0.1 Introduction

$$\text{Linear Algebra} \implies \text{Functional Analysis} \implies \begin{cases} \text{Calculus of Variations} \\ \text{ODEs/PDEs} \\ \text{Approximation Theory} \\ \dots \end{cases}$$

(1) Functional analysis deals with *infinite* dimensional spaces, often spaces of *functions* and certain classes of *mappings* between these spaces.

(2) Can be thought of a (nontrivial) generalization of linear algebra (from finite to infinite dimensional).

(3) The development of functional analytic methods and results are important in various areas of math; in particular they provide a powerful tool for the study of ODEs, PDEs, for example the existence and uniqueness of solutions and the convergence of approximations, etc.

0.2 Linear Algebra

We first draw an analogy between functional analysis and linear algebra, starting from familiar concepts. Let A be a $n \times n$ matrix.

(p1) Given a vector $b \in \mathbb{R}^n$, find $x \in \mathbb{R}^n$ such that $Ax = b$.

(p2) Analogy: *consider a bounded, open set* $\Omega \subset \mathbb{R}^n$ *and a linear partial differential operator*

$$Lu = -\sum_{i,j=1}^{n} (a^{i,j}(x)u_{x_i})_{x_j} + \sum_{i=1}^{n} b^i(x)u_{x_i} + c(x)u.$$

Given a function $f : \Omega \to \mathbb{R}$*, find a function* u *vanishing on the boundary of* Ω *so that* $Lu = f$*.*

Fundamental differences between (p1) and (p2)

The matrix A is a *continuous* (or equivalently *bounded*) linear transformation on the *finite* dimensional space $\mathbb{R}^n$, whereas the differential operator L is considered an *unbounded* (and equivalently *not continuous*) operator on *infinite* dimensional spaces $L^2(\Omega)$. In particular, the domain of L is not $L^2(\Omega)$, but only a suitable subspace (e.g. $H_0^1(\Omega)$, the space of functions vanishing on the boundary of Ω with $\|u\|_{H_0^1(\Omega)} = \left(\int_\Omega |u|^2 \,\mathrm{d}x + \int_\Omega |\nabla u|^2 \,\mathrm{d}x\right)^{1/2} < \infty$).

Structural Similarities

(1) (Positivity) If A is strictly positive definite, i.e., there exists $\beta > 0$ such that $\langle Ax, x\rangle \geqslant \beta\|x\|^2$ for all $x \in \mathbb{R}^n$, then A is invertible as its kernel is trivial. If so, (p1) has a *unique* solution for all $b \in \mathbb{R}^n$.

Analogy in elliptic PDEs: if L is strictly positive definite, i.e.,

$$\langle Lu, u\rangle_{L^2} := \int_\Omega Lu \cdot u \,\mathrm{d}x = \int_\Omega \sum_{i,j=1}^{n} a^{i,j}(x)u_{x_i}u_{x_j} + \sum_{j=1}^{n} b^j(x)u_{x_j}u + c(x)u^2 \,\mathrm{d}x \geqslant \beta\|u\|^2_{H_1^0(\Omega)}$$

for some $\beta > 0$ and for all $u \in H_0^1(\Omega)$, then (p2) has a unique solution $u \in H_0^1(\Omega)$ for every $f \in L^2(\Omega)$. In particular, at each $x \in \Omega$, the $n \times n$ matrix $\{a^{i,j}(x)\}$ should be strictly positive definite.

(2) (Fredholm Alternative) In linear algebra, (p1) has a unique solution if and only if $Ax = 0$ has one solution $x = 0$. In other words, uniqueness comes if and only if the kernel is trivial. In general this is *false*; if X is infinite dimensional, one can construct a bounded linear operator $\Lambda : X \to X$ that is injective but not surjective (or vice versa).

However, there is a remedy: if Λ is of a specific form given by $\Lambda = I - K$ where I is the identity and K a compact operator, then Λ is injective if and only if it's surjective. Consequently, (p2) has a unique solution $u \in H_1^0(\Omega)$ for all $f \in L^2$ if and only if the homogeneous equation $Lu = 0$ only has a trivial solution.

Beginning of Aug.25, 2021

0.3 Evolution Equations

Consider an ODE

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

where A is a matrix and $b \in \mathbb{R}^n$. Linear ODE theory gives $x(t) = e^{tA}b$ where

$$e^{tA} := \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}.$$

Also, $\{e^{tA} \mid t \in \mathbb{R}\}$ has the group property

$$e^{0A} = I \qquad \text{and} \qquad e^{tA}e^{sA} = e^{(t+s)A}.$$

If A is symmetric, there exists an orthonormal basis of eigenvectors $\{v_1, ..., v_n\}$ with eigenvalues $\lambda_1, ..., \lambda_n$ such that

$$e^{tA}b = \sum_{k=1}^{n} e^{t\lambda_k} \langle b, v_k \rangle v_k.$$

Now we use linear semigroup theory and extend the above notion to linear operators in infinite dimensional spaces. Consider

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}u(t) = -Lu(t) & \\ u(0) = g & \text{for } g \in L^2(\Omega) \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

(For example, consider $u_t = \Delta u$ the heat equation.) When $a^{i,j}(x) = a^{j,i}(x)$ and $b^i(x) = 0$, then the elliptic operator L is symmetric (like how A is a symmetric matrix above), and the solution can be decomposed along orthogonal basis $\{\varphi_1, \varphi_2, ...\}$ of $L^2(\Omega)$:

$$u(t) = \sum_{k=1}^{\infty} e^{-t\lambda_k} \langle g, \varphi_k \rangle_{L^2} \varphi_k := S_t g \qquad t \geqslant 0.$$

Note that L is unbounded (as $\lambda_k \to \infty$ as $k \to \infty$). However, S_t are bounded for $t \geqslant 0$, and moreover the family of linear operators $\{S_t \mid t \geqslant 0\}$ is a linear semigroup, since

$$S_0 = I \qquad S_t \circ S_s = S_{t+s} \text{ for all } s, t \geqslant 0.$$

Roughly speaking, $S_t = e^{-Lt}$ (roughly because L is unbounded and the exponential formula is not necessarily valid). We need some different approximation method, for example

$$e^{tA} = \lim_{n\to\infty} \left(I - \frac{t}{n}A\right)^{-n}$$

or

$$e^{tA} = \lim_{\lambda\to\infty} e^{tA_\lambda} \qquad A_\lambda := A(I - \lambda^{-1}A)^{-1}.$$

0.4 Function Spaces

A key idea is to view functions $f : \mathbb{R}^n \to \mathbb{R}$ as *points* in an abstract vector space.
The information about a function can be encoded in a simple but useful number $\|f\|$, which we call the **norm** of f. This notion, along with the structure of this vector space, leads to many important results.
For example, in applications to ODEs, PDEs, or integral equations, we naturally consider $C^k(\mathbb{R}^n)$ (the space of functions with bounded, continuous partial derivatives up to order k) with

$$\|f\|_{C^k} := \max_{\sum \alpha \leqslant k} \sup_{x\in\mathbb{R}^n} |\, \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} ....\partial_{x_n}^{\alpha_n} f(x) \,|$$

These spaces are not always practical since real interesting solutions may be discontinuous. This motivates the Lebesgue spaces $L^p(\mathbb{R}^n)$, $p \geqslant 1$ or Sobolev spaces, $W^{k,p}(\mathbb{R}^n)$, where we weaken our requirement on continuity and obtain much more flexibility. The norm in $W^{k,p}(\mathbb{R}^n)$ are given by

$$\|f\|_{W^{k,p}} := \sum_{\sum \alpha \leqslant k} \left(\int_{\mathbb{R}^n} |\, \partial_{x_1}^{\alpha_1} ...\partial_{x_n}^{\alpha_n} f(x)|^p \,\mathrm{d}x \right)^{1/p}.$$

0.5 Compactness

When we solve an equation where the explicit formula is unavailable, one standard method that we employ is the following:

(1) Approximate the problem; construct a sequence of approximate solutions $\{u_n\}$.

(2) Extract a *convergent subsequence* (or at least a canditate) $\{u_{n_k}\}$ converging to some $\overline{u}$.

(3) Show that $\overline{u}$ is a solution.

In step 2, there is a striking difference between $\mathbb{R}^n$ and abstract function spaces. In the former, Heine-Borel states (compact $\Leftrightarrow$ closed and bounded), whereas in general function space the $\Leftarrow$ direction may well be false.

Resolution

(1) We can introduce a weaker notion of convergence and show that every bounded sequence has a subsequence that converges in a weak sense. (The Banach-Alaoglu Theorem)

(2) We can consider two distinct norms $\|u\|_{\text{weak}} \leqslant \|u\|_{\text{strong}}$. If the strong norms are bounded, then there exists a subsequence converging in the weak norm.

End of Introduction

Beginning of Aug.27, 2021

Chapter 1

Normed Spaces and Banach Spaces

1.1 Basic Definitions

Definition 1.1.1: Normed Linear Space

$(X, \|\cdot\|)$ is a **normed linear space** (NLS) over a field $\mathbb{K}$ (usually $\mathbb{R}$ or $\mathbb{C}$) if X is a vector space and $\|\cdot\| : X \to \mathbb{R}$ satisfies

(1) (non-degeneracy) $\|x\| \geqslant 0$ for all $x \in X$ with equality if and only if $x = 0$;

(2) (absolute homogeneity) $\|\lambda x\| = |\lambda|\|x\|$ for all $x \in X$, $\lambda \in \mathbb{K}$; and

(3) (triangle inequality) $\|x + y\| \leqslant \|x\| + \|y\|$ for all $x, y \in X$.

Note that the mapping $x \mapsto \|x\|$ is continuous.

Lemma 1.1.2

Let $(X, \|\cdot\|)$ be a NLS. Then the *distance function* $d(x, y) := \|x - y\|$ defines a metric on X. Moreover:

(1) d is translation invariant: $d(x, y) = d(x + z, y + z)$ for all $x, y, z \in X$;

(2) d is positively homogeneous: $d(\lambda x, \lambda y) = |\lambda| d(x, y)$ for all $x, y \in X$, $\lambda \in \mathbb{K}$; and

(3) every open ball $B(x_0, r) := \{x \in X \mid d(x, x_0) < r\}$ is a *convex* set. *This defines a topology on X. We therefore have a concept of open sets, closed sets, convergent sequences, continuous mappings, and so on.* We use $\overline{B}(x_0, r)$ to denote the closed ball, the closure of $B(x_0, r)$.

Proof of (3). By translation invariance, it suffices to assume that the center of the ball is origin, i.e., $x_0 = 0$. Let $x, y \in B(0, r)$ and $0 \leqslant \theta \leqslant 1$ be given. Then we want show that $\theta x + (1 - \theta) y \in B(0, r)$. Indeed:

$$\|\theta x + (1 - \theta) y\| \leqslant \|\theta x\| + \|(1 - \theta) y\| = \theta\|x\| + (1 - \theta)\|y\| < \theta r + (1 - \theta) r = r.$$

□

Definition 1.1.3: Convergence of Sequences and Series

A sequence $\{x_n\}$ in a NLS is **convergent** if there exists $x \in X$ such that $\lim_{n\to\infty} \|x_n - x\| = 0$. We write $x_n \to x$ and say x is the **limit** of $\{x_n\}$.
For series:

(1) The series $\sum_{k=1}^{\infty} y_k$ **converges** to x if the series of *partial sums* $S_n := \sum_{k=1}^{n} y_k$ converges to x. If so, we write $\sum_{k=1}^{\infty} y_k = x$ and $\{y_k\}$ is called **summable**.

(2) If $\sum_{k=1}^{\infty} \|y_k\| < \infty$, we say $\{y_k\}$ is **absolutely summable**.

Definition 1.1.4: Completeness & Banach Space

A sequence $\{x_n\}$ is **Cauchy** if for all $\epsilon > 0$, there exists a sufficiently large $N \in \mathbb{N}$ such that

$$\|x_m - x_n\| < \epsilon \qquad \text{whenever } m, n \geqslant N.$$

A normed space X is **complete** if every Cauchy sequence converges. A complete normed space is called a **Banach** space.

Theorem 1.1.5

A normed space X is complete if and only if *every* absolute summable sequence is summable. *Proof left as exercise; see Ozanski's 3.18.*

Examples of Banach Spaces

We now present some basic examples of Banach spaces:

(1) $\mathbb{R}^n$ equipped with the euclidean norm $\|x\|_2 := \sqrt{x_1^2 + \ldots + x_n^2}$ is a Banach space over $\mathbb{R}$.

(2) $\mathbb{R}^n$ equipped with p-norm and/or ∞-norm are also Banach spaces, where

$$\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \text{ for } 1 \leqslant p < \infty \qquad \text{and} \qquad \|x\|_\infty := \max_{1\leqslant i\leqslant n} |x_i|.$$

(3) ℓ^p spaces defined by $\{x = \{x_k\}_{k=1}^{\infty} \mid \sum_{k=1}^{\infty} |x_k|^p < \infty\}$ is Banach with $\|x\|_p := \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{1/p}$.

(4) ℓ^∞ defined by $\ell^\infty := \{x = \{x_k\}_{k=1}^{\infty} \mid \sup_{k\geqslant 1} |x_k| < \infty\}$ is a Banach space with $\|x\|_\infty := \sup_{k\geqslant 1} |x_k|$.

(5) $C^0[a,b] := \{f : [a,b] \to \mathbb{R} \mid f \text{ is continuous}\}$ is a Banach space with norm $\|f\|_{C^0} := \max_{x\in[a,b]} |f(x)|$.

(6) $L^p(\Omega) := \{f : \Omega \to \mathbb{R} \mid f \text{ is Lebesgue measurable and } \int_\Omega |f(x)|^p \, \mathrm{d}x < \infty\}$ is a Banach space with

$$\|f\|_{L^p} := \left(\int_\Omega |f(x)|^p \, \mathrm{d}x\right)^{1/p}.$$

(7) Similarly, $L^\infty(\Omega)$, the set of essentially bounded measurable functions on Ω, is a Banach space with

$$\|f\|_{L^\infty} := \operatorname*{ess\,sup}_{x\in\Omega}|f(x)|.$$

Example 1.1.6. Consider X, the set of all *continuous* functions on Ω such that $\|f\|_{L^p} < \infty$. This space is *not* complete. However, its completion is exactly $L^p(\Omega)$!

Example 1.1.7. Let X be the space of all polynomials on $[0,1]$. Consider the two norms

$$\|f\|_{C^0} := \max_{x\in[0,1]}|f(x)| \qquad \text{and} \qquad \|f\|_{L^1} := \int_0^1 |f(x)|\,\mathrm{d}x.$$

We consider a sequence of functions defined by $f_n(x) = x^n$. It is clear that $\|f_n\|_{C^0} = 1$ for all n, but

$$\|f_n\|_{L^1} = \int_0^1 x^n\,\mathrm{d}x = \frac{1}{n+1} \to 0$$

as $n \to \infty$. In other words, $f_n \to 0$ in the L^1 norm but $\{f_n\}$ is not even Cauchy with respect to C^0 (take the difference between $x^n - x^m$).

Example 1.1.8. We know every norm on a metric space induces a norm. How about the converse? *False*! See here for examples.

1.2 Linear Operators

Unless otherwise specified, we will let X, Y be NLS over the same scalar field $\mathbb{K}$.

Definition 1.2.1: Linear Operator

A **linear operator** is a mapping $\Lambda : D(\Lambda) \to Y$ such that

$$\Lambda(c_1x_1 + c_2x_2) = c_1\Lambda(x_1) + c_2\Lambda(x_2)$$

for all $x_1, x_2 \in D(\Lambda)$ and $c_1, c_2 \in \mathbb{K}$.
Here $D(\Lambda)$ stands for the **domain** of Λ, a subset of X. Analogously, we define $R(\Lambda) = \{\Lambda x \mid x \in D(A)\}$ to be the **range**. Finally, we define the **kernel** of Λ to be

$$\ker(\Lambda) := \{x \in X \mid \Lambda x = 0\}.$$

It is easy to check that Λ is injective if and only if $\ker(\Lambda) = \{0\}$ (or 0 for shorthand notation).

Definition 1.2.2: Bounded Linear Operator

Let $\Lambda : X \to Y$ be a linear operator. We say Λ is **bounded** if

$$\|\Lambda\| := \sup_{\|x\|\leqslant 1} \|\Lambda x\| < \infty.$$

(Note that $\|x\| \leqslant 1$ refers to the norm in X, $\|\Lambda x\|$ refers to the norm in Y, and $\|\Lambda\|$ refers to the **operator norm**. We drop the subscripts to avoid cumbersome notations.)

Theorem 1.2.3: Bounded ⇔ Continuous

A linear operator $\Lambda : X \to Y$ is bounded if and only if it is continuous.

Proof. For $\Leftarrow$, if Λ is continuous, it is continuous at the origin in particular. Thus, there exists $\delta > 0$ such that $\|x\| \leqslant \delta \Rightarrow \|\Lambda x\| \leqslant 1$ (note that $\Lambda(0) = 0$). By linearity, if $\|x\| \leqslant 1$ then $\|\Lambda x\| \leqslant 1/\delta < \infty$.
For $\Rightarrow$, suppose $x_1 \neq x_2$. Then

$$\|\Lambda x_1 - \Lambda x_2\| = \|\Lambda(x_1 - x_2)\| = \Bigg\|\Lambda \underbrace{\left(\frac{x_1 - x_2}{\|x_1 - x_2\|}\right)}_{\text{norm}=1}\Bigg\| \cdot \underbrace{\|x_1 - x_2\|}_{\text{a scalar}} \leqslant \|\Lambda\|\|x_1 - x_2\|$$

so Λ is Lipschitz with constant $\|\Lambda\|$. In particular it is continuous. □

Definition 1.2.4: $\mathcal{B}(X;Y)$

Let $\mathcal{B}(X;Y)$ denote the space of bounded linear operators from X to Y. Then $\mathcal{B}(X;Y)$ is a normed space with the operator norm as its norm. In particular, if Y is a Banach space, then $\mathcal{B}(X;Y)$ is a Banach space.

Beginning of Aug.30, 2021

Proof. First recall that if Λ_1, Λ_2 are linear operators, then by definition

$$(c_1\Lambda_1 + c_2\Lambda_2)x := c_1\Lambda_1 x + c_2\Lambda_2 + x.$$

We now show that $\Lambda \mapsto \|\Lambda\|$ indeed defines a norm:

(1) $\|\Lambda\| \geqslant 0$ for all $\Lambda \in \mathfrak{B}(X;Y)$ is clearly trivial; also, if $\Lambda = 0$ then $\Lambda x = 0$ for all $x \in X$, so $\|\Lambda\| = 0$ (otherwise we can easily derive a contradiction).

(2) Absolute homogeneity follows from that of $\|\cdot\|_Y$: if α, then

$$\|\alpha\Lambda\| = \sup_{\|x\|\leqslant 1} \|\alpha\Lambda x\| = \sup_{\|x\|\leqslant 1} (|\alpha|\|\Lambda x\|) = |\alpha| \sup_{\|x\|\leqslant 1} \|\Lambda x\| = |\alpha|\|\Lambda\|.$$

(3) Triangle inequality again follows from that of $\|\cdot\|_Y$: let $\|x\| \leqslant 1$. Then

$$\begin{aligned}\|(\Lambda_1 + \Lambda_2)x\| &= \|\Lambda_1 x + \Lambda_2 x\| \leqslant \|\Lambda_1 x\| + \|\Lambda_2 x\| \\ &\leqslant \|\Lambda_1\| + \|\Lambda_2\|.\end{aligned}$$

Taking the supremum over all x eighth $\|x\| \leqslant 1$ we obtain the desired inequality.

Now let Y be Banach; we will show that $\mathfrak{B}(X;Y)$ is Banach. Let $\{\Lambda_n\}$ be a Cauchy sequence of bounded linear operators. For any $\in X$, we have $\|\Lambda_n x - \Lambda_m x\| \leqslant \|\Lambda_n - \Lambda_m\|\|x\|$. By assumption,

$$\lim_{m,n\to\infty} \|\Lambda_n(x) - \Lambda_m(x)\| \leqslant \limsup_{m,n\to\infty} \|\Lambda_n - \Lambda_m\|\|x\| = 0,$$

so $\{\Lambda_n x\}$ is a Cauchy sequence in Y and hence converges to some limit, which we call Λx. For all $x \in X$ we define the corresponding Λx as such.

It remains to show that $\Lambda \in \mathfrak{B}(X;Y)$. Its linearity is trivial. For boundedness: we can choose a sufficiently large N (this class uses the notation $N \gg 1$, and I will use it from now on) such that

$$\|\Lambda_n - \Lambda_N\| \leqslant 1 \qquad \text{for all } n \geqslant N.$$

Thus, for any $x \in X$, $\|x\| \leqslant 1$,

$$\begin{aligned}\|\Lambda x\| = \lim_{n\to\infty} \|\Lambda_n X\| &= \lim_{n\to\infty} \|(\Lambda_N + \Lambda_n - \Lambda_N)x\| \\ &\leqslant \|\Lambda_N x\| + \limsup_{n\to\infty} \|\Lambda_n - \Lambda_N\| \|x\| < \infty.\end{aligned}$$

□

Example 1.2.5.

(1) Let $1 \leqslant p \leqslant \infty$. Consider $X = \ell^p$ with

$$\|x\|_p := \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{1/p} \text{ for finite } p \text{ and } \|x\|_\infty := \sup_{k\geqslant 1} |x_k| \text{ (for p} = \infty).$$

Let $(\lambda_1, \lambda_2, \ldots)$ be an arbitrary sequence of real numbers. Define

$$\Lambda : \ell^p \to \ell^p \text{ by } (x_1, x_2, \ldots) \mapsto (\lambda_1 x_1, \lambda_2 x_2, \ldots).$$

If we set $e_1 := (1, 0, \ldots)$, $e_2 := (0, 1, \ldots)$, and so on, then the **span** of them will NOT be ℓ^p but rather a *dense subset* of ℓ^p for $p < \infty$(in fact, c_{00}). ℓ^∞ is NOT separable.

If we view Λ as the infinite matrix $\text{diag}(\lambda_1, \lambda, \ldots)$, and if $\{\lambda_k\}$ is bounded, then the operator is bounded with $\|\Lambda\| = \sup_{k\geqslant 1} |\lambda_k|$. If $\{\lambda_k\}$ is not bounded, then Λ is not bounded.

(2) Let $I = (0, \pi)$ and let $X = BC(I)$ be the space of bounded, continuous, real-valued functions on I with

$$\|f\| = \sup_{0<x<l\pi} |f(x)|.$$

Consider the linear operator Λ defined by $\Lambda f = f'$. Accordingly, define the domain

$$D(\Lambda) := \{f : I \to \mathbb{R} \mid f \text{ is differentiable and has a bounded continuous derivative}\}.$$

Then $\Lambda : D(\Lambda) \to X$ is clearly a linear operator. However, Λ is not bounded: for example consider the oscillating functions $f_k := \sin(kx)$, where $\|f_k\| = 1$ for all k but $\|\Lambda f\| = k \to \infty$ as $k \to \infty$.

(3) The **shift operators** on $L^p(\mathbb{R})$ for $1 \leqslant p \leqslant \infty$: fix $a \in \mathbb{R}$. For any $f \in L^p(\mathbb{R})$, we define

$$(\Lambda_a f)(x) := f(x - a).$$

Then clearly $\|\Lambda_a f\|_{L^p} = \|f\|_{L^p}$ so $\Lambda_a : L^p \to L^p$ has operator norm 1. Also, notice that Λ_a is a bijection!

(4) The shift operators on $\ell^p(\mathbb{R})$ for $1 \leqslant p \leqslant \infty$: define the right and left shift operators by

$$\Lambda_+ : \ell^p \to \ell^p \text{ by } (x_1, x_2, \ldots) \mapsto (0, x_1, x_2, \ldots)$$

and

$$\Lambda_- : \ell^p \to \ell^p \text{ by } (x_1, x_2, \ldots) \mapsto (x_2, x_3, \ldots).$$

The both still are bounded linear operators with norms 1, yet Λ_+ is injective but not surjective, and Λ_- is surjective but not injective.

(5) The **multiplication operator**: let $\Omega \subset \mathbb{R}^n$ be bounded and open and $g : \Omega \to \mathbb{R}$ bounded and measurable. Define

$$M_g := L^p(\Omega) \to L^p(\Omega) \text{ by } f \mapsto gf.$$

Then M_g is a bounded operator with $\|Mg\| = \sup_{\|f\|_p} \|gf\|_p = \|g\|_\infty$.

(6) The **integral operator**: let $X = C^0([a, b])$. Let

$$\Lambda f(x) := \int_a^x f(y) \, \mathrm{d}y.$$

Then $\Lambda : X \to X$ is a bounded linear operator:

$$|\Lambda f(x)| = \left| \int_a^x f(y) \, \mathrm{d}y \right| \leqslant \int_a^x |f(y)| \, \mathrm{d}y \leqslant \max_{y \in [a,b]} |f(y)|(b - a).$$

Beginning of Sept.1, 2021

1.3 Finite-Dimensional Spaces

Definition 1.3.1: Equivalent/Comparable Norms

We say two norms $\|\cdot\|_1$, $\|\cdot\|_2$ are **equivalent** (or **comparable**) if there exists a constant $c \geqslant 1$ such that

$$\frac{1}{c}\|x\|_1 \leqslant \|x\|_2 \leqslant c\|x\|_1 \qquad \text{for all } x.$$

Note that equivalent norms lead to the same Cauchy sequences.

Remark. In general, infinite-dimensional spaces can have many non-equivalent norms, but the case is much nicer if we look at finite-dimensional spaces —

Theorem 1.3.2: Norms on Finite-Dimensional Spaces

Let X be a finite-dimensional normed space over $\mathbb{K}$. Let $\mathcal{B} = \{u_1, \ldots, u_n\}$ be a basis of X. Then:

(1) X is a Banach space.

(2) For each $\alpha \in \mathbb{K}^n$, let

$$\Lambda\alpha := \alpha_1 u_1 + \ldots + \alpha_n u_n \in X.$$

Then $\Lambda : \mathbb{K}^n \to \mathbb{K}$ is bijective and bounded. Moreover, Λ^{-1} is *also* bounded. (In general the inverse of a bounded operator need not to be inverse.)

Proof. Note that Λ is clearly both injective and surjective, and its inverse is well-defined. All of these directly follow from properties of basis. Note that

$$\|\Lambda\alpha\| = \left\| \sum_{i=1}^{n} \alpha_i u_i \right\| \leqslant \sum_{k=1}^{n} \|\alpha_i u_i\| \leqslant \max_{1 \leqslant i \leqslant n} |\alpha_i| \sum_{i=1}^{n} \|u_i\| \leqslant \|\alpha\| \sum_{k=1}^{n} \|u_k\|,$$

so Λ is indeed a bounded linear operator.

Now we show Λ^{-1} is bounded. Suppose not, then there exists a sequence $\{x_n\}$ where $\|x_n\| \leqslant 1$ but $\|\Lambda^{-1}x_n\|$ tends to infinity. Let

$$\beta_n := \frac{\Lambda^{-1}x_n}{\|\Lambda^{-1}x_n\|} \in \mathbb{K}^n$$

such that $\|\beta_n\| = 1$ and $\Lambda\beta_n = x_n / \|\Lambda^{-1}x_n\| \to 0$ as $n \to \infty$.

Since $\{\beta_n\}$ is a bounded sequence in $\mathbb{K}^n$, it is compact and admits a convergent subsequence $\{\beta_{n_k}\}$ converging to some $\beta \in \mathbb{K}^n$. But then

$$\|\beta\| = \lim_{k \to \infty} \|\beta_{n_k}\| = 1$$

whereas

$$\Lambda\beta = \lim_{k \to \infty} \Lambda\beta_{n_k} = 0,$$

so Λ has a nontrivial kernel, contradicting its injectivity. Hence Λ^{-1} is bounded.

Finally, we show that X is Banach. Let $\{x_n\}$ be a Cauchy sequence in X. Then by boundedness $\{\Lambda^{-1}x_n\}$ is a Cauchy sequence in $\mathbb{K}^n$ and hence converges to some $\beta \in \mathbb{K}^n$. Since Λ is continuous, it preserves sequential limits, and thus

$$x_n = \Lambda(\Lambda^{-1}(x_n))$$

converges to $\Lambda\beta \in X$. □

Corollary 1.3.3: Norms are Equivalent on Finite-Dimensional Spaces

In a finite-dimensional space, all norms are equivalent.

Proof. Let $\|\cdot\|_1, \|\cdot\|_2$ be any two norms on X. Let $\mathcal{B} = \{u_1, \ldots, u_n\}$ be a basis of X. Also, let $\Lambda : \mathbb{K}^n \to X$ be defined as in the previous theorem. Then we know Λ, Λ^{-1} are bounded linear operators. Therefore

$$c_1 \|\Lambda^{-1}x\| \leqslant \|x\|_1 \leqslant c_2 \|\Lambda^{-1}x\|$$

for all x and some $c_1, c_2 > 0$. Similarly, there exist $c_3, c_4 > 0$ such that

$$c_3 \|\Lambda^{-1}x\| \leqslant \|x\|_2 \leqslant c_4 \|\Lambda^{-1}x\|.$$

The claim then follows by combining the inequalities above. □

Theorem 1.3.4: Characterization of Finite-Dimensional Spaces

(Riesz) Let X be a normed space. Then the following are equivalent:

(1) X is finite-dimensional.

(2) The *closed* unit ball $B_1 := \overline{B(0,1)}$ is compact.

Proof. For $\Rightarrow$, let X have dimensional N. Then by the previous theorems there exists a linear homeomorphism $\Lambda : \mathbb{K}^n \to X$ with bounded inverse. Therefore we can map the unit ball in X to $K = \Lambda^{-1}(B_1) \subset \mathbb{K}^n$, a closed and bounded subset of $\mathbb{K}^n$. By Heine-Borel K is compact! Then $B_1 = \Lambda(K)$ is the continuous image of a compact set and is therefore compact.

For $\Leftarrow$, assume B_1 is compact. Then, using the definition of covering compactness, B_1 can be covered by a *finite* number of balls with radius 1/2, i.e., $B_1 \subset \sum_{i=1}^{n} B(p_i, 1/2)$ for some n. Then, let

$$\mathcal{V} := \text{span}\{p_1, \dots, p_n\}$$

is a closed subspace in X (because every finite-dimensional normed space is complete, as stated previously). We claim that $V = X$. Suppose V is a *proper* subspace of X. Then there exists $x \in X - V$ such that the distance $\rho := d(x, V) = \inf_{v \in V} \|v - x\| > 0$ (because V is closed). By definition of infimum, there exists $u \in V$ such that

$$\rho \leqslant \|x - u\| \leqslant \frac{3}{2}\rho.$$

Normalizing $x - u$, we define $z := (x - u)/\|x - u\|$ and obtain an element in B_1. By our covering assumption, there exists $p_i \in B_1$ such that $\|z - p_i\| < 1/2$. However, this *cannot* happen as

$$\begin{aligned} x &= u + (x - u) = u + z\|x - u\| \\ &= u + \|x - u\|p_i + \|x - u\|(z - p_i) \\ &= \underbrace{u + \|x - u\|p_i}_{\in \mathcal{V}} + \|x - u\|(z - p_i). \end{aligned}$$

Since $u + \|x - u\|p_i \in \mathcal{V}$, taking infimum on $\|x - u\|(z - p_i)$ gives

$$\|x - u\|\|z - p_i\| \leqslant \rho,$$

so $\|x - u\| \geqslant 2\rho$, contradicting our assumption that $\|x - u\| \leqslant 3\rho/2$. Hence $X = V$. □

Beginning of Sept.3, 2021

1.4 Uniformly Convex Banach Spaces

Definition 1.4.1: Strictly Subadditive Norm

A norm is called **strictly subadditive** if the triangle inequality holds strictly, i.e., $\|x+y\| < \|x\| + \|y\|$ whenever x and y are not scalar multiples of each other.
For example, $(\ell^p, \|\cdot\|_p)$ *and* $(L^p(\Omega), \|\cdot\|_p)$ *for* $1 < p < \infty$ *are strictly subadditive, but* ℓ^∞, L^∞ *are not.*

Definition 1.4.2

A NLS X whose norm satisfies

$$\left\|\frac{x+y}{2}\right\| \leqslant 1 - \epsilon(\|x-y\|)$$

for all $\|x\|, \|y\| \leqslant 1$ and $\epsilon(r) > 0, \lim_{r\to 0} \epsilon(r) = 0$ is called **uniformly convex**.

Theorem 1.4.3

Let X be a uniformly convex Banach space. Let $K \subset X$ be closed and convex. Let $z \in X$. Then there exists a *unique* point $y \in K$ closest to z, and $\|y - z\| = \inf_{k\in K} \|k - z\|$.

Proof. If $z \in K$ the claim is trivial. Now we assume $z \notin K$. By translation we may assume $z = 0$ and $0 \notin K$. Let

$$\rho := \inf_{y\in K} \|y\|,$$

and we see that $\rho > 0$ by closedness of K. *It remains to show that the infimum can be obtained.*
Let $\{y_k\}$ be a *minimizing sequence* such that $y_k \in K$ and $\rho_k := \|y_k\| \to \rho$. Next we normalize and let $x_n := y_n/\rho_n$ so $\|x_n\| = 1$. We consider norms of form $\|(x_n + x_m)/2\|$. Since

$$\begin{aligned}\frac{x_n + x_m}{2} &= \frac{1}{2\rho_n} y_n + \frac{1}{2\rho_m} y_m \\ &= \left(\frac{1}{2\rho_n} + \frac{1}{2\rho_m}\right)\left[\frac{1/(2\rho_n)}{1/(2\rho_n) + 1/(2\rho_m)} y_n + \frac{1/(2\rho_m)}{1/(2\rho_n) + 1/(2\rho_m)} y_m\right].\end{aligned}$$

Therefore everything in [] is a convex combination of y_n, y_m and is therefore an element in K. Hence by infimum $\|(x_n + x_m)/2\| \geqslant (1/(2\rho_n) + 1/(2\rho_m))\rho$. On the other hand, since X is uniformly convex,

$$1 - \epsilon(\|x_n - x_m\|) \geqslant \left\|\frac{x_n + x_m}{2}\right\| \geqslant \left(\frac{1}{2\rho_n} + \frac{1}{2\rho_m}\right)\rho.$$

Since by assumption $\rho_n, \rho_m \to \rho$, the RHS converges to 1. Therefore $\|x_n - x_m\| \to 0$ and hence $\|y_n - y_m\| \to 0$. Thus $\{y_n\}$ is Cauchy. Since X is complete and K closed, $\{y_n\} \to y \in K$, and this completes the proof. $\square$

Remark. The point y above is the existence of the minimum when the set over which one wants to minimize is not compact. (A Banach space has many closed bounded sets that are not compact.)

Remark. The above theorem fails without the assumption on uniform convexity.
For example, let $X = C([-1,1])$ with $\|f\| := \max_{x\in[-1,1]} |f(x)|$. (Note that this norm is not strictly additive.) Let

$$K := \{k \in X \mid \int_{-1}^{0} k \, \mathrm{d}x = \int_{0}^{1} k \, \mathrm{d}x = 0\}.$$

It is easy to check that K is a convex, closed subset of X. Now we take any $f \in X$ with

$$\int_{-1}^{0} f \, \mathrm{d}x = 1 \qquad \text{and} \qquad \int_{0}^{1} f \, \mathrm{d}x = -1.$$

Then the "average condition" implies that, for any $k \in K$,

$$\int_{-1}^{0} f - k \, \mathrm{d}x = 1 \qquad \text{and} \qquad \int_{-1}^{0} f - k \, \mathrm{d}x = -1.$$

Then

$$\max_{-1\leqslant x\leqslant 0} [f(x) - k(x)] \geqslant 1 \qquad \text{and} \qquad \min_{0\leqslant x\leqslant 1} [f(x) - k(x)] \leqslant -1,$$

Note that $\geqslant$ is $=$ if and only if $f - k \equiv 1$ on $[-1,0]$ and similarly $\leqslant$ is $=$ if and only if $f - k \equiv -1$ on $[0,1]$. However, both equalities *cannot* hold simultaneously as $f - k$ is a continuous as $0 \in [-1,0]$ and $0 \in [0,1]$!! Therefore we have at least one strict inequality. Either way,

$$\max_{-1\leqslant x\leqslant 1} |f(x) - k(x)| > 1$$

for all $k \in K$. However, one may choose $k \in K$ appropriately such that $\|f - k\|$ is arbitrarily close to (but >) 1. This disproves the theorem without the uniform convexity assumption.

1.5 Seminorms

Let us first consider $X = C((0,1))$. Since we are excluding the endpoints here, we can have functions that blow up, so the sup norm $p(f) := \sup|f|$will not give us a norm in this situation.
On the other hand, if we consider some *closed* $[a,b] \subset (0,1)$, one can introduce a **seminom**

$$p^{a,b}(f) := \max_{x\in[a,b]} |f(x)|$$

which is well-defined on $[a,b]$, although this may not be a norm on $(0,1)$ (for example we can construct $f \in C((0,1))$ with $p^{a,b}(f) = 0$ but $f \not\equiv 0$). This gives rise to the name seminorm, a "weakened version" of norms.

Definition 1.5.1

Let X be a vector space over $\mathbb{K}$. A real-valued map $x \mapsto p(x)$ is called a **seminorm** on X if

(1) (positive semidefiniteness) $p(x) \geqslant 0$ for all $x \in X$,

(2) (absolute homogeneity) $p(\lambda x) = |\lambda| p(x)$ for all $x \in X$ and $\lambda \in \mathbb{K}$, and

(3) (triangle inequality) $p(x+y) \leqslant p(x) + p(y)$ for all $x, y \in X$.

Note that $p(x) = 0$ for $x \neq 0$ is allowed, and if p is a seminorm, then $d(x,y) := p(x,y)$ is *not* a distance function in general.

Beginning of Sept.8, 2021

Definition 1.5.2: Separating Sequence of Seminorms

A sequence $\{p_k\}_{k\geqslant 1}$ of seminorms on X is **separating** if, for all nonzero $x \in X$, there exists at least one k such that $p_k(x) > 0$.

Lemma 1.5.3

Let $\{p_k\}$ be a separating sequence of seminorms on X. Then

$$d(x,y) := \sum_{k=1}^{\infty} 2^{-k} \frac{p_k(x-y)}{1+p_k(x-y)}$$

defines a distance on X.

Definition 1.5.4: Fréchet Space

If X with d defined as in the lemma above is complete, then we say X is a **Fréchet** space.

Example 1.5.5. Let $\Omega \subset \mathbb{R}^n$ be open with boundary $\partial\Omega$. Consider $C(\Omega)$, the collection of continuous functions $f : \Omega \to \mathbb{R}$. Then $C(\Omega)$ does *not* have a *natural* norm, but it is possible to give it the structure of a Fréchet space.

Proof. For each $k \geqslant 1$, we define a compact subset

$$A_k := \{x \in \Omega \mid \|x\| \leqslant k \text{ and } \overline{B}(x, 1/k) \subset \Omega\}$$

(in other words, each $x \in A_k$ is contained in Ω and *slightly* away from its boundary). We define the seminorms

$$p_k(f) := \max_{x \in A_k} |f(x)|.$$

Since A_k is compact, the maximum is obtained. It's easy to verify that $\{p_k\}$ is a separating sequence of seminorms. Now we define

$$d(f,g) := \sum_{k=1}^{\infty} 2^{-k} \frac{p_k(f-g)}{1+p_k(f-g)}.$$

We now show that $(C(\Omega), d)$ is a complete metric space and therefore a Fréchet space. Let $\{f_n\}$ be Cauchy (w.r.t. d). Then

$$\limsup_{m,n\to\infty} p_k(f_m - f_n) = 0 = \limsup_{m,n\to\infty} \sup_{A_k} |f_n(x) - f_m(x)| = 0.$$

Now since any $x \in \Omega$ is contained in one of A_k, the sequence $\{f_n(x)\}$ is Cauchy, and by completeness of $\mathbb{R}$ it converges to some value which we call $f(x)$. In addition, every compact $K \subset \Omega$ is contained in one of A_k. Therefore, the convergence $f_n \to f$ is uniform on any compact subset. The uniform convergence implies the continuity of f. To show that $\lim_{n\to\infty} d(f_n, f) = 0$, since $f_n \to f$ uniformly on A_m for any m,

$$d(f_n, f) = \sum_{k=1}^{m} 2^{-k} \frac{p_k(f_n - f)}{1 + p_k(f_n - f)} + \sum_{k=m+1}^{\infty} 2^{-k} \frac{p_k(f_n - f)}{1 + p_k(f_n - f)}.$$

Let $n \to \infty$. The first term $\to 0$ and the second is bounded by $\sum_{k=m+1}^{\infty} 2^{-k}$. Letting $m \to \infty$, we complete the proof. □

Example 1.5.6. An open set Ω' is said to be **compactly contained** in Ω if $\overline{\Omega'}$ is a compact subset of Ω. We write $\Omega' \subset\subset \Omega$ if this is the case. Then we define

$$L^p_{\text{loc}}(\Omega) = \{f : \Omega \to \mathbb{R} \mid f \in L^p(\Omega') \text{ for all } \Omega' \subset\subset \Omega\}.$$

(This is a space larger than $L_p(\Omega)$.) This space does not have a natural norm, but we may define seminorms

$$p_k(f) = \left(\int_{A_k} |f|^p \, \mathrm{d}x\right)^{1/p} := \|f\|_{L^p(A_k)}.$$

The corresponding separating sequence $\{f_k\}$ along with the d defined above makes $L^p_{\text{loc}}(\Omega)$ a Fréchet space.

1.6 Extension Theorems

Definition 1.6.1

Let X be a vector space over $\mathbb{K}$. A linear map $f : X \to \mathbb{K}$ is called a **linear functional** on X. A **Minkowski functional** on X is a map $p : X \to \mathbb{R}$ such that

(1) (subadditivity) $p(x + y) \leqslant p(x) + p(y)$ and

(2) (positive homogeneity) $p(\lambda x) = \lambda p(x)$, for all $x, y \in X$ and $\lambda \geqslant 0$.

(In particular notice that every seminorm is a Minkowski functional.)

Theorem 1.6.2: Hahn-Banach Extension Theorem

Let X be a vector space over $\mathbb{R}$ (there's a similar version for X over $\mathbb{C}$) and $p : X \to \mathbb{R}$ a Minkowski functional. Let $V \subset X$ be a subapce of X and let $f : V \to \mathbb{R}$ a linear functional such that

$$f(x) \leqslant p(x) \qquad \text{for all } x \in V.$$

Then there exists a linear function $F : X \to \mathbb{R}$ such that

$$F(x) = f(x) \text{ for all } x \in V \qquad \text{and} \qquad -p(-x) \leqslant F(x) \leqslant p(x) \text{ for all } x \in X.$$

Proof. If $V = X$ then we are done: $f(x) = -f(-x) \geqslant -p(-x)$.

If $V \neq X$, for any $x_0 \in X - V$, we consider the larger subspace $V_0 := \{x + tx_0 \mid x \in V, t \in \mathbb{R}\}$. For $x, y \in V$, we have

$$f(x) + f(y) = f(x+y) \leqslant p(x+y) = p(x - x_0 + x_0 + y) \leqslant p(x - x_0) + p(x_0 + y).$$

Rearranging gives

$$f(x) - p(x - x_0) \leqslant p(y + x_0) - f(y) \quad \text{for all } x, y \in V.$$

Now we define $\beta := \sup_{x \in V}\{f(x) - p(x - x_0)\}$. Then $f(x) - p(x - x_0) \leqslant \beta \leqslant p(y + x_0) - f(y)$. Now we define an extension of f on V_0 by

$$f(x + tx_0) := f(x) + \beta t.$$

We claim that $f(x + tx_0) \leqslant p(x + tx_0)$. If $t = 0$ we are done; if $t > 0$, we let x, y be x/t and obtain

$$t(f(x/t) - p(x/t - x_0)) \leqslant t\beta \leqslant t(p(x/t + x_0) + f(x/t)).$$

By linearity and positive homogeneity of p, we obtain $f(x) - p(x - x_0) \leqslant t\beta \leqslant p(x + x_0 t) - f(x)$ and recover the original inequality. Finally, the claim follows from Zorn's lemma. □

Beginning of Sept.10, 2021

Theorem 1.6.3: Extension for Bounded Linear Functionals

Let $(X, \|\cdot\|)$ be a normed space over $\mathbb{K}$. Let $f : V \to \mathbb{K}$ be a bounded linear functional defined on a subspace $V \subset X$. Then f can be extended to a linear function $F : X \to \mathbb{K}$ such that

$$\|F\| = \sup_{\substack{x \in X \\ \|x\| \leqslant 1}} |F(x)| = \sup_{\substack{x \in V \\ \|x\| \leqslant 1}} |f(x)| = \|f\|.$$

Corollary 1.6.4

Let X be a Banach space. For any distinct $x, y \in X$, there exists a bounded linear function $\varphi : X \to \mathbb{R}$ such that $\varphi(x) \neq \varphi(y)$.

Corollary 1.6.5

Let X be a Banach space. For all $x \in X$, there exists a bounded linear functional $\varphi : X \to \mathbb{K}$ such that

$$\varphi(x) = \|x\| \qquad \text{and} \qquad \|\varphi\| = 1.$$

Application of Hahn-Banach

Theorem 1.6.6

Let X be a normed space over $\mathbb{C}$ and $V \subset X$ a subspace. For any $x \in X$, define its distance to V by

$$m(x) := \inf_{v \in V} \|x - v\|.$$

Then for every $x \in X$, $m(x) = M(x)$ where

$$M(x) := \max|\varphi(x)|$$

subject to $\varphi \in X^*$, $\|\varphi\|_* \leqslant 1$, and $\varphi \equiv 0$ on V.

Proof. We first show that $M(x) \leqslant m(x)$. Since φ vanishes on V and $\|\varphi\|_* = 1$, for all $x \in X, v \in V$ we have

$$|\varphi(x)| = |\varphi(x) - \varphi(v)| = |\varphi(x - v)| \leqslant \|x - v\|.$$

Taking infimum over $x \in V$ and then taking supremum over φ gives $M(x) \leqslant m(x)$.
Now we try to find a φ that upgrades $\leqslant$ to $=$. Keep x fixed. Consider a vector space $V_0 = \{v + \lambda x \mid v \in V, \lambda \in \mathbb{C}\}$. We define a map $\varphi_0 : V_0 \to \mathbb{C}$ by $v + \lambda x \mapsto \lambda m(x)$. Clearly, φ is linear and vanishes on V. Moreover, for $\lambda \neq 0$,

$$\begin{aligned} |\varphi(v + \lambda x)| &= |\lambda| m(x) \\ &= |\lambda| \inf_{\tilde{v} \in V} \|x - \tilde{v}\| \\ &\leqslant |\lambda| \|x - v/\lambda\| = \|v + \lambda x\| \end{aligned}$$

so indeed $\|\varphi\|_* \leqslant 1$. Thus, by the extension theorems, we can extend φ_0 to all of X; there exists $\varphi \in X^*$ with

$$\|\varphi\|_* \leqslant 1 \qquad \text{and} \qquad \varphi|_V = 0.$$

Setting $v = 0, \lambda = 1$, we see that $\varphi(x) = \varphi_0(x) = m(x)$. This completes the proof. ☐

Definition 1.6.7

The **closed linear span** of a subset $\{y_i\}$ of a NLS is the smallest closed linear space containing all of y_i's, i.e., the intersection of all closed linear spaces containing all of y_i's. (If finite-dimensional then this is merely the span.)
If $\{y_i\}$ is infinite, then the closed linear span is the closure of the linear span Y consisting of all finite linear combinations of y_i's.

Theorem 1.6.8: Spanning Criterion

A point z of a NLS X belongs to the closed linear span Y of $\{y_i\}$ if and only if every bounded linear function that vanishes on the subset vanishes at z, i.e., $\varphi(y_i) = 0$ for all i implies $\varphi(z) = 0$.

Proof. For $\Rightarrow$, since φ is linear, all finite linear combinations $\varphi(\sum_{\text{finite}} \lambda_i y_i) = 0$. Also, since φ is continuous, φ vanishes on all limit points of $\sum \lambda_i y_i$, which includes z.

For $\Longleftarrow$, suppose that z does not belong to the closed linear span Y of $\{y_i\}$. Now we consider

$$d := \inf_{y \in Y} \|z - y\|$$

which, by assumption, is strictly positive. We define a subspace $Z := \{y + \lambda z \mid y \in Y, \lambda \in \mathbb{K}\}$ and define the functional $\varphi_0 : Z \to \mathbb{K}$ by $y + \lambda z \mapsto \lambda$. Clearly φ_0 is linear, and assuming $\lambda \neq 0$, we have

$$\|y + \lambda z\| = |\lambda|(\|z + y/\lambda\|) \geqslant |\lambda| \inf_{y \in Y} \|z - y\| = |\lambda| d$$

and hence $\|\varphi_0\| \leqslant 1/d$. Therefore, by Hahn-Banach, φ_0 can be extended to φ on X. Like φ_0, our new φ vanishes on all finite combinations of y_i, but setting $y = 0$ and $\lambda = 1$, we see that $\varphi(z) = 1$. We have therefore proven the claim by taking its contrapositive. □

1.7 Duality

Definition 1.7.1: Dual Space

Let X be a Banach space over $\mathbb{K}$. Then the collection of all bounded linear functionals is called the **dual space** of X, denoted X^*.

Note that since $\mathbb{K}$ is Banach, X^* is also Banach, equipped with the operator norm

$$\|\varphi\|_* := \sup_{\|x\| \leqslant 1} |\varphi(x)|.$$

Beginning of Sept.13, 2021

Definition 1.7.2: Weak Convergence

A sequence $\{x_n\} \subset X$ is called **weakly convergent** if there exists $x \in X$ such that

$$\lim \varphi(x_n) = \varphi(x) \qquad \text{for all } \varphi \in X^*.$$

In this case, x is called the **weak limit** of $\{x_n\}$ and we write $x_n \rightharpoonup x$.

Remark. It is clear that strong convergence implies weak convergence, as any φ is bounded and therefore preserves sequential limits.

Corollary 1.7.3

Weak limits are unique.

Proof. Suppose $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$ with $x \neq y$. Then by the extension theorems, there exists a continuous linear function $\varphi \in X^*$ such that $\varphi(x) \neq \varphi(y)$, but then this contradicts the uniqueness of limits in $\mathbb{R}$. □

Example 1.7.4: Weak Convergence $\not\Rightarrow$ Strong Convergence. Consider $X = C([0,1])$ and

$$f_n(t) = \begin{cases} nt & 0 \leqslant t \leqslant 1/n \\ 2 - nt & 1/n \leqslant t \leqslant 2/n \\ 0 & 2/n \leqslant t \leqslant 1 \end{cases} \qquad \text{for } n \geqslant 2.$$

(The graph consists of line segments connecting $(0,0), (1/n, 1), (2/n, 0)$, and $(1,0)$.)

It is clear that f_n do not converge with respect to $\|\cdot\|_{\sup}$.

We now show that the weak limit of f_n is 0, that is, $\lim_{n\to\infty} \varphi(f_n) = 0$ for all $\varphi \in X^*$. Suppose this is not true so that there exists a subsequence $\{f_{n_k}\}$, some $\varphi \in X^*$, and some $\delta > 0$, such that $|\varphi(f_{n_k})| > \delta > 0$. In particular, we can delete more terms and ensure that $n_{k+1} > n_k$. Then for all $t \in [0,1]$, the sequence of functions $\{g_k\}$ defined by

$$g_k := \sum_{i=1}^{k} f_{n_i}(t)$$

is always bounded by 4. Then $\varphi(g_k) = \sum_{i=1}^{k} \varphi(f_{n_k}) > k\delta$, contradiction as we can make k arbitrarily large.

Remark. Note that each fixed $x \in X$ determines a linear functional on X^* by $\varphi \mapsto \varphi(x)$ for all $\varphi \in X^*$. Then by Hahn-Banach's corollary, the norm of the functional

$$\sup_{\|\varphi\|_* \leqslant 1} |\varphi(x)| = \|x\|.$$

We can define a "double star" norm $\|x\|_{**}$ as above. Then we have a *canonical embedding*

$$i : X \mapsto X^{**} \text{ by } x \mapsto i(x), \text{ the map } \varphi \mapsto \varphi(x).$$

This embedding preserves the norm and is called **isometric**.

Definition 1.7.5

If $i(X) = X^{**}$, then X is called **reflexive**.

Example 1.7.6. All finite-dimensional spaces as well as $L^p(\Omega)$ and ℓ^p for $1 < p < \infty$ are reflexive.
In general, X^{**} is larger, for example $L^1(\Omega), L^\infty(\Omega), \ell^1, \ell^\infty$.

Example 1.7.7. $X = C([-1,1])$ with $\|\cdot\|_{\max}$ is not reflexive.

Proof. Suppose for contradiction that X is reflexive so $X = X^{**}$. Thus for each $\varphi \in X^*$, there exists $f \in X^{**} = X$ such that (cf. HW1.5)

$$\|\varphi\| = |\varphi(f)| \qquad \text{and } \|f\| = 1.$$

Now we define

$$\varphi(g) := \int_{-1}^{0} g(t)\,\mathrm{d}t - \int_{0}^{1} g(t)\,\mathrm{d}t \in X^*.$$

It follows that $|\varphi(g)| < 2\|g\|_{\max}$ (strict because of continuity), but for $\epsilon > 0$, we can choose g satisfying

$$\|\varphi(g)\| > (2-\epsilon)\|g\|,$$

so $\|g\| = 2$. But then this contradicts the assumption $\|f\| = 1$ for $g = f$. □

Theorem 1.7.8: Duality of L^p

The dual of L^p is L^q where (p, q) are a conjugate pair.

Proof. For any $u \in L^q$, we can define a functional $\varphi \in (L^p)^*$ by

$$\varphi(f) := \langle f, u\rangle = \int_\Omega f(x)u(x)\,\mathrm{d}m.$$

By Hölder's inequality, for $f \in L^p$,

$$|\varphi(f)| = \int_\Omega |f(x)u(x)|\,\mathrm{d}m \leqslant \|f\|_p\|g\|_q$$

with equality obtained by $\|\varphi\| = \|u\|_q$. Thus L^q isometrically embedded in $(L^p)^*$. We claim $L^q = (L^p)^*$. If not, then there exists $z \in (L^p)^* - L^q$. Since L^q is closed, by the spanning criterion, there exists $\varphi \in (L^p)^{**}$ such that

$$\varphi(u) = 0 \text{ for all } u \in L^q \text{ but } \varphi \neq 0.$$

Since L^p is reflexive [!], $\varphi \in L^p$, so $(\varphi, u) = 0$ for all $u \in L^q$ and so $\varphi = 0$. □

Beginning of Sept.13, 2021

The embedding $i : X \to X^{**}$ can be used to introduce a weak topology on X^*:

Definition 1.7.9: Weak*-Convergence

We say $\{\varphi_n\} \subset X^*$ **weak-star converges** to $\varphi \in X^*$ if

$$\lim_{n\to\infty} \varphi_n(x) = \varphi(x)$$

for all $x \in X$. If so we write $\varphi_n \overset{*}{\rightharpoonup} \varphi$.

Note that this convergence is weaker than $\|\varphi_n - \varphi\| \to 0$: here we simply require "pointwise convergence" whereas $\|\varphi_n - \varphi\|$ requires some kind of uniform boundedness.

Recall that the closed unit ball in an infinite dimensional space is not compact. Hence if X^* is infinite dimensional then there exists a sequence $\{\varphi_n\} \subset B(0,1) \subset X^*$ that does not admit any convergence subsequence (w.r.t. the operator norm). However, if we only require weak-star convergence, then $B(0,1)$ becomes compact!

Theorem 1.7.10: Banach-Alaoglu Theorem

Let X be a separable Banach space. Then every bounded sequence of linear functionals $\{\varphi_n\} \subset X^*$ admits a weak-star convergent subsequence.

Proof. Let $\{\varphi_n\} \subset X^*$ with $\|\varphi_n\| \leqslant C$ be given. Since X is separable, there exists a dense countable set

$$S = \{x_1, x_2, \ldots\} \subset X.$$

We first show that there exists a subsequence $\{\varphi_{n_j}\}$ that converges pointwise on S. This is done by a diagonal argument: since $\{\varphi_n(x_1)\}$ is a bounded subset of $\mathbb{R}$, by Bolzano-Weierstraß it has a convergent subsequence; then iterate the process and obtain a set of nested sequences, eventually getting a $\{\varphi_{n_j}\}$ that converges at all $x_k \in S$. Call this limit function φ.

Now it remains to show that φ is a bounded linear functional, i.e., $\varphi \in X^*$, with Lipschitz constant C. Indeed, for $x_n, x_k \in S$, $|\varphi(x_n) - \varphi(x_k)| = \lim_{j\to\infty} |\varphi_{n_j(x_n) - \varphi_{n_j}(x_k)}| \leqslant \limsup_{j\to\infty} \|\varphi_{n_j}\|_* \|x_n - x_k\| \leqslant C\|x_n - x_k\|$, so φ, which is Lipschitz on S, can be uniquely extended by continuity to the closure of S (which is X) by a density argument.

Finally, we will verify that $\varphi_{n_j} \overset{*}{\rightharpoonup} \varphi$. Let $x \in X$ and $\epsilon > 0$ be given. Since S is dense in X, there exists $x_k \in S$ with $\|x_k - x\| < \epsilon$. Then we have convergence of x_k!

$$\begin{aligned}\limsup_{j\to\infty} |\varphi_{n_j}(x) - \varphi(x)| &\leqslant \limsup_{j\to\infty} |\varphi_{n_j}(x) - \varphi_{n_j}(x_k)| + \limsup_{j\to\infty} |\varphi_{n_j}(x_k) + \varphi_{n_j}(x)| + \limsup_{j\to\infty} |\varphi_{n_j}(x) - \varphi(x)| \\ &\leqslant C\|x_k - x\| + 0 + C\|x_k - x\| < 2C\epsilon.\end{aligned}$$

Since C is arbitrary, we are done. □

Remark. The theorem holds for any NLS; see Rudin 3.14.

Chapter 2

Bounded Linear Operators

2.1 The Uniform Boundedness Principle

Theorem 2.1.1: Banach-Steinhaus Theorem

Let $\tilde{H} \subset B(X,Y)$ be any family of bounded linear operators. Then either $\tilde{H}$ is uniformly bounded such that

$$\sup_{\Lambda \in \tilde{H}} \|\Lambda\| < \infty$$

or there exists a dense set $S \subset X$ such that

$$\sup_{\Lambda \in \tilde{H}} \|\Lambda x\| = \infty \qquad \text{for all } x \in S.$$

Proof. Consider the open sets

$$S_n := \{x \in X : \|\Lambda x\| > 1/n \text{ for some } \Lambda \in \tilde{H}\}.$$

If one of these sets, say S_k, is not dense in X, then there exists $x_0 \in X$ and $r_0 > 0$ such that $\overline{B(x_0, r_0)} \subset X - S_k$. This means

$$\|\Lambda x\| \leqslant k \qquad \text{for all } \Lambda \in \tilde{H} \text{ and } x \in \overline{B(x_0, r_0)}.$$

Now if $\|x\| \leqslant r$, for all $\Lambda \in \tilde{H}$,

$$\|\Lambda x\| = \|\Lambda(x_0 + x) - \Lambda x_0\| \leqslant 2k$$

so for all $\Lambda \in \tilde{H}$,

$$\|\Lambda\| = \sup_{\|x\| \leqslant 1} \|\Lambda x\| = \frac{1}{r} \sup_{\|x\| \leqslant r} \|\Lambda x\| \leqslant \frac{2k}{r}.$$

Therefore the family $\tilde{H}$ of operators is uniformly bounded.
On the other hand, if all S_n'n are dense in X, by Baire's Category theorem, their untersection $S := \bigcap_{n=1}^{\infty} S_n$ is dense in X, so for each x and $n \geqslant 1$, there exists $\Lambda \in \tilde{H}$ with $\|\Lambda x\| > n$, which gives $\sup_{\Lambda \in \tilde{H}} \|\Lambda\| = \infty$. □

Remark. This theorem shows that such a family of operators is either uniformly bounded or not bounded at all! There is no such things as "pointwise bounded only" in this scenario.

Beginning of Sept.17, 2021

Corollary 2.1.2

Let $\{\Lambda_n\}$ be a sequence of bounded linear operators $X \to Y$. Assume that the pointwise limits exist and define

$$\Lambda x := \lim_{n\to\infty} \Lambda_n x \qquad \text{for all } x.$$

Then Λ is a bounded linear operator.

Proof. For every $x \in X$, $\{\Lambda_n x\}$ is bounded. Therefore by PUB, $\{\Lambda_n\}$ is uniformly bounded. Since

$$\|\Lambda\| = \sup_{\|x\|\leqslant 1} \|\Lambda x\| = \sup_{\|x\|\leqslant 1} \lim_{n\to\infty} \|\Lambda_n x\| \leqslant \sup_{n\geqslant 1} \|\Lambda_n\| < \infty$$

we see that Λ is indeed bounded. □

2.2 Open Mapping Theorem

Definition 2.2.1: Open Mapping

If X, Y are metric spaces, we say $f : X \to Y$ is an **open mapping** if for all $U \subset X$ open, the image $f(U)$ is open in Y.

Example 2.2.2: Examples and non-examples of open mappings.

(1) The projection $\pi_k : \mathbb{R}^n \to \mathbb{R}^k$ (for $k \leqslant n$) defined by $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_k)$ is obviously open.

(2) The inclusion map $\mathbb{R}^k \hookrightarrow \mathbb{R}^n$ by $(x_1, \ldots, x_k) \mapsto (x_1, \ldots, x_k, 0, \ldots, 0)$ is *not* an open map — we cannot draw a ball in $\mathbb{R}^n$ where some of its coordinates can only take the value 0.

(3) $\sin : \mathbb{R} \to \mathbb{R}$ is not open.

Theorem 2.2.3: Open Mapping Theorem, OMT

Let X, Y be Banach spaces. Let $\Lambda : X \to Y$ be a bounded surjective linear operator. Then Λ is open.

Proof. By linearity, for any open ball $B(x, r)$ we have

$$\begin{aligned}\Lambda((B(x,r)) &= \Lambda x + \Lambda(B(0,r)) \\ &= \Lambda x + r\Lambda((B(0,1)),\end{aligned}$$

so it is sufficient to show that $\Lambda((B(0,1))$ contains an open ball centered at the origin of Y. For shorthand notation, write B_1 as the open unit ball in X and $B_r := B(0,r)$ in Y.

We first show that there exists B_r that is contained the *closure* $\overline{\Lambda(B_1)}$. Since Λ is surjective, $Y = \bigcup_{n=1}^{\infty} \Lambda(B_n)$. By Baire category theorem, at least one of $\overline{\Lambda(B_n)} \subset Y$ has a nonempty interior. Since $\overline{\Lambda(B_1)} = 1/n \cdot \overline{\Lambda(B_n)}$, we see that $\overline{\Lambda(B_1)}$ also has a nonempty interior. Therefore there exists $y_0 \in Y$ and $r > 0$ such that

$$B(y_0, r) \subset \overline{\Lambda(B_1)}.$$

Notice that $\Lambda(B_1)$ is convex and symmetric (since B_1 is). Therefore $B(y_0, r) \subset \overline{\Lambda(B_1)}$ as well. Now we consider the convex combination

$$\frac{1}{2}B(y_0, r) + \frac{1}{2}B(-y_0, r) \subset \overline{\Lambda(B_1)}.$$

Notice that this is exactly $B(0, r)$, so $B(0, r) \subset \overline{\Lambda(B_1)}$, as claimed. Furthermore, by linearity again,

$$B(0, 2^{-n}r) \subset \Lambda(B_{2^{-n}}) \qquad \text{for all } n. \tag{Δ}$$

Now we show that $B(0, r/2)$ is in fact contained in $\Lambda(B_1)$ (so we can drop the closure). To this end, pick any $y \in B(0, r/2)$. We want to find $x \in B_1$ such that $\Lambda x = y$. By (Δ), $y \in \overline{\Lambda(B_{2^{-1}})}$. By definition of closure, there exists $x_1 \in B_{2^{-1}}$ such that

$$\|y - \Lambda x_1\| < 2^{-2}r.$$

Now since $y - \Lambda x_1$ is in $B(0, 2^{-2}r)$, by (Δ), $y - \Lambda x_1 \in B(0, 2^{-2}r) \subset \overline{\Lambda(B_{2^{-2}})}$. Thus there exists another $x_2 \in B_{2^2}$ such that

$$\|y - \Lambda x_1 - \Lambda x_2\| < 2^{-3}r.$$

Inductively, for each n,

$$y - \sum_{j=1}^{n-1} \Lambda x_j \in B(0, 2^{-n}r) \subset \overline{\Lambda(B_{2^n})},$$

and by closure, there exists $x_n \in B_{2^{-n}}$ satisfying

$$\left\| y - \sum_{j=1}^{n} \Lambda x_j \right\| < 2^{-n-1}r.$$

Letting $n \to \infty$, we see that y is the limit $\sum_{j=1}^{\infty} \Lambda x_j$. Define $x := \sum_{j=1}^{\infty} x_n$. It remains to notice that $\Lambda x = \lim_{n\to\infty} \sum_{j=1}^{n} \Lambda x_j = y$ and that

$$\|x\| \leqslant \sum_{j=1}^{\infty} \|x_j\| < \sum_{j=1}^{\infty} 2^{-j} = 1,$$

so indeed $y \in \Lambda(B_1)$, as claimed. □

Corollary 2.2.4

If X, Y are Banach spaces and $\Lambda : X \to Y$ a bijective bounded linear operator, then $\Lambda^{-1} : Y \to X$ is also bounded. *Indeed, bijection guarantees the existence of an inverse and OMT guarantees its boundedness.*

Corollary 2.2.5

Let $\|\cdot\|_1$, $\|\cdot\|_2$ be norms on a vector space X such that $(X, \|\cdot\|_1)$, $(X, \|\cdot\|_2)$ are complete. Then if $\|\cdot\|_2 \leqslant c\|\cdot\|_1$ for come $c > 0$, we have $\|\cdot\|_1 \leqslant d\|\cdot\|_2$ for some $d > 0$.
This is just an application of OMT and the previous corollary on $\Lambda : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ by $x \mapsto x$.

Beginning of Sept.20, 2021

2.3 Closed Graph Theorem

Remark. Every continuous linear operator $\Lambda : X \to Y$ is closed.

Theorem 2.3.1: Closed Graph Theorem

Let X, Y be Banach spaces, and let $\Lambda : X \to Y$ be a closed linear operator defined on the *entire* X. Then Λ is continuous. (*From the remark above we can rewrite this as an "iff" statement.*)

Proof. For convenience write Γ as the graph of Λ. By assumption Γ is a closed subspace of $X \times Y$ and is therefore Banach. Define π_1, π_2 as the projection of Γ onto X, Y defined by

$$\pi_1(x, \Lambda x) := x \qquad \text{and} \qquad \pi_2(x, \Lambda x) := \Lambda x.$$

Note that π_1 is a linear, bounded (since $\|\pi_1(x, \Lambda x)\| = \|x\| \leqslant \|x\| + \|\Lambda x\| = \|(x, \Lambda x)\|_{X\times Y}$), and bijective (by defining $\pi_1^{-1}(x) := (x, \Lambda x)$). Therefore, by the Open Mapping Theorem, π_1^{-1} is continuous, so $\Lambda = \pi_2 \circ \pi^{-1}$ is also continuous! □

Remark. The theorem also holds if we assume that $\Lambda : D(\Lambda) \to Y$ is closed and that $D(\Lambda)$ is closed.

Example 2.3.2: Closed but not bounded operator. Let $X := C^0(\mathbb{R})$ be the space of all *bounded* continuous functions equipped with the sup norm. Define Λ by $\Lambda f = f'$ and set its domain to be

$$D(\Lambda) := \{f \text{ is continuously differentiable}\},$$

i.e., $f \in C^1$ or $f' \in X$.

(1) Λ is not bounded: we can have functions with higher oscillations, e.g., $\sin(nx)$ which have sup norm 1. However, as $n \to \infty$, $\|\Lambda f\| \to \infty$.

(2) Λ has a closed graph. Consider a sequence $\{f_n\} \in D(\Lambda)$ such that $\|f_n - f\|_{C^0} \to 0$ for some $f \in C^0$ and $\|f_n' - g\|_{C^0} \to 0$ for some $g \in C^0$. Then

$$\int_0^x g(s)\, \mathrm{d}s = \int_0^x f_n'(s)\, \mathrm{d}s + \int_0^x g(s) - f_n'(s)\, \mathrm{d}s.$$

Letting $n \to \infty$, by convergence as $n \to \infty$ the second term tends to 0:

$$\int_0^x g(s)\,\mathrm{d}s = \lim_{n\to\infty} \int_0^x f_n'(x)\,\mathrm{d}s = f(x) - f(0)$$

so f is continuously differentiable with $f' = g$.

Note that this does *not* contradict our CGT because $D(\Lambda) = C^1$ is neither all of X or closed in X.

2.4 Adjoint Operators

Let X be a Banach space over $\mathbb{K}$. Recall the definition of X^* and the canonical embedding

$$i : X \to X^{**} \text{ defined by } (x \in X) \mapsto (x^* \mapsto x^*(x) \in \mathbb{K}).$$

We shall denote $x^*(x)$ by $\langle x^*, x\rangle$ from now on.

Definition 2.4.1: Adjoint Operator

Let $\Lambda : X \to Y$ be a bounded linear operator. For any $y^* : Y \to \mathbb{K}$, we define a bounded linear functional *on* X by

$$x^*(x) = y^*(\Lambda x).$$

(Originally we have $X \xrightarrow{\Lambda} Y \xrightarrow{y^*} \mathbb{K}$; now we want $x^* : X \to \mathbb{K}$.) Then the map

$$\Lambda^* : y^* \mapsto \Lambda^* y^* := y^* \circ \Lambda$$

is a bounded linear operator from Y^* to X^* so that

$$\langle \Lambda^* y^*, x\rangle = \langle y^*, \Lambda x\rangle \qquad \text{for all } x \in X.$$

(In other words, $\Lambda^* y^*(x) = y^*(\Lambda x)$.) We say Λ^* is the **adjoint** of Λ.

Definition 2.4.2: Orthogonal Complement

For $V \subset X$, define the **orthogonal complement** $V^\perp$ by

$$V^\perp = \{x^* \in X^* \mid \langle x^*, x\rangle \text{ for al } x \in V\}.$$

When in a Hilbert space which we will cover later, this definition can be replaced by a condition on inner product, which is much more intuitive.

Similarly, for $W \subset X^*$, define

$$W^\perp = \{x \in X \mid \langle x^*, x\rangle = 0 \text{ for all } x^* \in W\}.$$

Theorem 2.4.3

Let $\Lambda : X \to Y$ be a bounded linear operator and let $X^* : Y^* \to X^*$ be its adjoint. Then

(1) $\|\Lambda^*\| = \|\Lambda\|$ (both being their operator norms, respectively), and

(2) $\ker \Lambda = \text{Range}(X^*)^\perp$ and $\ker X^* = \text{Range}(\Lambda)^\perp$.

Beginning of Sept.22, 2021

Proposition 2.4.4

(1) If $\Lambda_1, \Lambda_2 \in \mathcal{B}(X, Y)$, then $(\Lambda_1 + \Lambda_2)^* = \Lambda_1^* + \Lambda_2^*$, $(\alpha\Lambda_1)^* = \alpha\Lambda_1^*$,

(2) If $\Lambda_1 \in \mathcal{B}(X, Y)$ and $\Lambda_2 \in (Y, Z)$ then $(\Lambda_2 \circ \Lambda_1)^* = \Lambda_1^* \circ \Lambda_2^*$.

(3) If $\Lambda \in \mathcal{B}(X, Y)$ and Λ^{-1} exists and $\Lambda^{-1} \in \mathcal{B}(Y, X)$, then $(\Lambda^*)^{-1}$ exists and $(\Lambda^*)^{-1} \in \mathcal{B}(Y^*, X^*)$.

Proof.

(1) By Hahn-Banach (HW1 #5)

$$\begin{aligned}
\|\Lambda\| &= \sup_{\|x\| \leqslant 1} \|\Lambda x\| \\
&= \sup\{|\langle y^*, \Lambda x\rangle : \|y^*\| \leqslant 1, \|x\| \leqslant 1|\} \\
&= \sup\{|\langle \Lambda^* y^*, x\rangle : \|y^*\| \leqslant 1, \|x\| \leqslant 1|\} \\
&= \sup\{\|\Lambda^* y^*\| : \|y^*\| \leqslant 1\} = \|\Lambda^*\|.
\end{aligned}$$

(2) Note that $x \in \ker \Lambda$ is equivalent to $\Lambda x = 0$. This is further equivalent to $\langle y^*, \Lambda x\rangle = 0$ for all $y^* \in Y^*$. (The forward direction is obvious; if the converse is false, then Hahn-Banach ensures a nonzero mapping.) Then

$$\begin{aligned}
\langle y^*, \Lambda x\rangle = 0 \text{ for all } y^* \in Y^* &\Leftrightarrow \langle \Lambda^* y^*, x\rangle = 0 \text{ for all } y^* \in Y^* \\
&\Leftrightarrow x \in \text{Range}(\Lambda^*)^\perp.
\end{aligned}$$

Also,

$$\begin{aligned}
y^* \in \ker(\Lambda^*) &\Leftrightarrow \Lambda^* y^* = 0 \\
&\Leftrightarrow \langle \Lambda^* y^*, x\rangle = 0 \text{ for all } x \in X \\
&\Leftrightarrow \langle y^*, \Lambda x\rangle = 0 \text{ for all } x \in X \\
&\Leftrightarrow y^* \in \text{Range}(\Lambda)^\perp.
\end{aligned}$$

□

2.5 Applications of Banach-Steinhaus

Theorem 2.5.1: Weakly convergent sequences are bounded

Let X be Banach. Then any $\{x_n\} \subset X$ that converges weakly to some $x \in X$ is bounded.

Proof. By definition we need $\langle x^*, x_n \rangle \to \langle x^*, x \rangle$ for all $x^* \in X^*$. Viewing the LHS as a linear functional $\Psi_n(x^*)$ from X^* to $\mathbb{K}$, we have

$$\sup_n |\Psi_n(x^*)| < \infty \qquad \text{for all } x^* \in X^*.$$

Then by the PUB we see that $\{\Psi_n\}$ is uniformly bounded for $n \geqslant 1$. Since (by Hahn-Banach)

$$\|\Psi_n\| = \sup_{\|x^*\| \leqslant 1} |\Psi_n(x^*)| = \|x_n\|$$

we see that $\|x_n\|$ needs to be bounded. □

Example 2.5.2: Space of polynomials. Let X be the space of all polynomials over $\mathbb{R}$ with norm

$$\|x\| := \max_j |\alpha_j|$$

(the maximum of coefficients).

We will use PUB to show that X is incomplete by constructing a sequence that is pointwise but not uniformly bounded (this would contradict PUB's assumption that X is complete).

We write $x(t) = \sum_{j=0}^{\infty} \alpha_j t^j$ where $\alpha_j = 0$ for $j > \deg(x)$. Now define $\Lambda_n : X \to \mathbb{R}$ by

$$\Lambda_n(0) = 0, \Lambda_n(x) = \alpha_0 + \alpha_1 + \ldots + \alpha_{n-1}.$$

Clearly Λ_n is linear and bounded:

$$\|\Lambda_n x\| \leqslant (\deg(x) + 1) \max_j |\alpha_j| = C_x$$

for fixed x. On the other hand, letting $x(t) := 1 + t^2 + \ldots + t^n$ gives

$$\|x\| = 1 \qquad \text{but} \qquad \|\Lambda_n x\| = n + 1.$$

Letting $n \to \infty$, we see that $\{\Lambda_n\}$ is not uniformly bounded.

Example 2.5.3: Fourier Series. Consider a 2π-periodic function $x(t) = x(t + 2k\pi)$. Then

$$x(t) = \frac{1}{2} a_0 + \sum_{m=1}^{\infty} [a_m \cos(mt) + b_m \sin(mt)]$$

where

$$a_m = \frac{1}{\pi} \int_0^{2\pi} x(t) \cos(mt)\, \mathrm{d}t \qquad \text{and} \qquad b_m = \frac{1}{\pi} \int_0^{2\pi} x(t) \sin(mt)\, \mathrm{d}t.$$

Claim: even for continuous functions, this series can diverge, i.e., there exist real-valued functions whose Fourier series diverge at a given point t_0.

Beginning of Sept.24, 2021

Proof. Let X be the space of 2π-periodic continuous functions equipped with the sup (max) norm. Note that X is Banach. WLOG let $t_0 = 0$. Define $\Lambda_n : X \to \mathbb{R}$ by the n^{th} Fourier partial sum evaluated at 0:

$$\Lambda_n(x) = f_n(x) \qquad \text{where} \qquad f_n(x) = \frac{a_0}{2} + \sum_{m=1}^{n} a_m = \frac{1}{\pi}\int_0^{2\pi} x(t)\left[\frac{1}{2} + \sum_{m=1}^{n}\cos(mt)\right] \,\mathrm{d}t$$

Since

$$\begin{aligned} 2\sin(t/2)\sum_{m=1}^{n}\cos(mt) &= \sum_{m=1}^{n}\left[\sin((m+1/2)t) - \sin((m-1/2)t)\right] \\ &= \sin((n+1/2)t) - \sin(t/2), \end{aligned}$$

we have

$$1 + 2\sum_{m=1}^{n}\cos(mt) = \frac{\sin((n+1/2)t)}{\sin(t/2)}$$

(and it holds as $t \to 0$ too). The RHS is called the *Dirichlet kernel*. Thus

$$f_n(x) = \frac{1}{2\pi}\int_0^{2\pi} x(t)\frac{\sin((n+1/2)t)}{\sin(t/2)} \,\mathrm{d}t.$$

For convenience denote the Dirichlet kernel by $q_n(t)$. We will show that (1) f_n is bounded with $\|f_n\| = \frac{1}{2\pi}\int_0^{2\pi}|q_n(t)| \,\mathrm{d}t$ and (2) $\|f_n\| \to \infty$ as $n \to \infty$. This would complete the proof since Banach-Steinhaus says pointwise convergence implies uniform convergence.

(1) To bound $f_n(x)$,

$$|f_n(x)| \leqslant \max|x(t)|\frac{1}{2\pi}\int_0^{2\pi}|q_n(t)| \,\mathrm{d}t\|x\|\frac{1}{2\pi}\int_0^{2\pi}|q_n(t)| \,\mathrm{d}t$$

so $\leqslant$ holds. To show $=$, write $|q_n(t)| = \operatorname{sgn}(q_n(t))q_n(t)$. Though sgn is not continuous, it can be approximated arbitrarily well by a continuous function x of norm 1:

$$\left|\frac{1}{2\pi}\int_0^{2\pi}(x(t) - y(t)q_n(t) \,\mathrm{d}t\right| < \epsilon$$

so

$$\left|f_n(x) - \frac{1}{2\pi}\int_0^{2\pi}|q_n(t)| \,\mathrm{d}t\right| < \epsilon.$$

Therefore taking sup implies $\|f_n\| = \frac{1}{2\pi}\int_0^{2\pi}|q_n(t)| \,\mathrm{d}t$, as claimed.

(2) See Ozanski HW3 p5.

□

Approximation of the δ Function by Continuous Functions

Definition 2.5.4

A sequence $\{\varphi_n\}$ of continuous functions on $[-1,1]$ tends to the $\boldsymbol{\delta}$ **function** if

$$\lim_{n\to\infty}\int_{-1}^{\infty} x(t)\varphi_n(t)\,\mathrm{d}t = x(0) \qquad \text{for all } x \in C([-1,1]). \tag{1}$$

Theorem 2.5.5: (Toeplitz)

The sequence $\{\varphi_n\}$ of continuous functions on $[-1,1]$ tends to the δ function as (1) describes if and only if the following are satisfied:

(i) $\lim_{n\to\infty}\int_{-1}^{1}\varphi_n(t)\,\mathrm{d}t = 1$,

(ii) For all $y \in C^\infty$ whose support does not contain 0,

$$\lim_{n\to\infty}\int_{-1}^{1} y(t)\varphi_n(t)\,\mathrm{d}t = 0,$$

(iii) There exists a constant $C > 0$ such that $\int_{-1}^{1}|\varphi_n(t)|\,\mathrm{d}t \leqslant C$ for all n.

Proof. For $\Rightarrow$, we suppose that φ_n tends to the δ function. Taking $x \equiv 1$ gives (i). Taking any smooth function with $x(0) = 0$ gives (ii). For (iii), we define

$$\varphi_n(x) := \int_{-1}^{1} x(t)\varphi_n(t)\,\mathrm{d}t.$$

Note that $\varphi_n \in (C[-1,1])^*$. In this perspective, (1) says $\varphi_n \rightharpoonup \delta$ and $|\varphi_n(x)| \leqslant \|x\|$ pointwise. Therefore by Banach-Steinhaus,

$$\|\varphi_n\| = \int_{-1}^{1}|\varphi_n(t)|\,\mathrm{d}t < \infty.$$

For $\Leftarrow$, suppose $x(0) = 0$. Let $y \in C^\infty$ and $\|x - y\| < \epsilon$ with $y(0) = x(0)$. Then

$$\left|\int_{-1}^{1}(x(t) - y(t))\varphi_n(t)\,\mathrm{d}t\right| \leqslant \epsilon\int_{-1}^{1}|\varphi_n(t)|\,\mathrm{d}t$$
$$[\text{(iii)}] \leqslant C\epsilon.$$

By (ii), $\int_{-1}^{1} y\varphi_n\,\mathrm{d}t \to 0$, so $\limsup_{n\to\infty}\left|\int_{-1}^{1} x\varphi_n\,\mathrm{d}t\right| \leqslant \epsilon$. Therefore (1) holds for any x with $x(0) = 0$. For a general x, since $x(t) = x(0) + (x(t) - x(0))$, using (i) we see that the claim also holds. □

Revisiting the Fourier Example

Recall that we said there exists a periodic continuous function whose Fourier series diverges at (any prescribed) t_0. Recall that we defined

$$f_n(x) = \frac{1}{2\pi}\int_0^{2\pi} x(t)q_n(t)\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\pi}^{\infty} x(t)q_n(t)\,\mathrm{d}t$$

where

$$q_n(t) := \frac{\sin((n+1/2)t)}{\sin(t/2)}.$$

Note that the convergence of the Fourier series is equivalent to q_n approximating the δ function! By the previous theorem, this is further equivalent to satisfying the three criteria. Since the Dirichlet kernel fails (iii), we claim that the Fourier series of f does not converge at 0.

Beginning of Sept.27, 2021

Approximate Quadrature Formula

Take n points $\{t_i\}_{i=1}^n \subset [-1,1]$ (**nodes**) and N numbers $\{w_i\}_{i=1}^j$ (**weights**), we define

$$q_n(f) := \sum_{j=1}^n w_j f(t_j).$$

We can view q_n as an element of $C([-1,1])^*$.

Theorem 2.5.6

Let q_n be a sequence of functionals as defined above satisfying

(1) For all $k \in \{0,1,2,\ldots\}$, $\lim_{n\to\infty} q_n(t^k) = \int_{-1}^1 t^k \, \mathrm{d}t$,

(2) for all N, $\sum_{i=1}^N |w_j^{(i)}| \leqslant C$.

Then

$$\lim_{n\to\infty} q_n(f) = \int_{-1}^1 f(t)\,\mathrm{d}t \text{ for all } f \in C([-1,1]). \tag{*}$$

The converse is true (i.e., (*) implies (1) and (2).)

Proof. For $\Rightarrow$, (1) implies that (*) holds for all polynomials. But then since the polynomials are dense in $C([-1,1])$, the claim follows. The rest of the proof is omitted. □

Remark. If we define $q(f) := \int_{-1}^1 f(t)\,\mathrm{d}t$ for all $f \in C([-1,1])$, the above theorem simply states $q_n \overset{*}{\rightharpoonup} q$ (weak* convergence).

Chapter 3

Compact Operators

3.1 Arzelá-Ascoli Theorem

In finite-dimensional spaces, Bolzano-Weierstraß says that every bounded sequence has a convergent subsequence. However, this compactness property fails in infinite dimensional spaces (e.g., $f_n(x) := x^n$ in $C([0,1])$ does not have a uniformly convergent subsequence in $C([0,1])$). What additional assumption is needed to guarantee the existence of a uniformly convergent subsequence?

Definition 3.1.1: Equicontinuity

Let E be a metric space. We say a family $\mathcal{E}$ of continuous functions is called **equicontinuous** if, given $x \in E$ and $\epsilon > 0$, there exists $\delta > 0$ such that

$$d(x,y) < \delta \implies |f(x) - f(y)| < \epsilon \qquad \text{for } \textit{all } f \in \mathcal{E}.$$

Corollary 3.1.2

Let E be compact and $\mathcal{E} \subset C(E)$ be equicontinuous. By compactness, we can upgrade equicontinuity to *uniform* equicontinuity: given $\epsilon > 0$, there exists $\delta > 0$ such that the claim above does not depend on x, i.e., given $\epsilon > 0$ and $\delta > 0$ such that

$$d(x,y) < \delta \implies |f(x) - f(y)| < \epsilon \qquad \text{for all } x, y \in E \text{ and } f \in \mathcal{E}.$$

Theorem 3.1.3

If E is compact, then $C(E)$ is Banach. By completeness, for a subset $\mathcal{E} \subset C(E)$, the following are equivlent:

(1) $\mathcal{E}$ is **relatively compact**, i.e., $\overline{\mathcal{E}}$ is compact, and

(2) $\mathcal{E}$ is **precompact** (or **totally bounded**), i.e., given $\epsilon > 0$, it can be covered by finitely many balls with radius ϵ, and

(3) Given any sequence of continuous functions $\{f_n\} \subset \mathcal{E}$, one can extract a subsequence converging to

some function f uniformly on E.

Theorem 3.1.4: Arzelá-Ascoli Theorem

Let E be a compact metric space and let $\mathcal{F} \subset C(E)$ be an equicontinuous family of functions with

$$\sup_{f \in \mathcal{F}} |f(x)| < \infty \qquad \text{for all } x \in E.$$

Then $\mathcal{F}$ is a relatively compact subset of $C(E)$, i.e., it has as uniformly convergent subsequence.

Proof. We will prove that $\mathcal{F}$ is precompact. Let $\epsilon > 0$ be given. By equicontinuity, choose δ such that

$$d(x,y) < \delta \implies |f(x) - f(y)| < \epsilon \text{ for all } x, y \in E \text{ and } f \in \mathcal{F}.$$

Since E is compact, there exist $x_1, \ldots, x_n$ such that

$$E \subset \bigcup_{i=1}^{n} B(x_i, \delta).$$

Since $\mathcal{F}$ is bounded, define $M := \max_{1 \leqslant i \leqslant n} \sup_{f \in \mathcal{F}} |f(x_i)| < \infty$ and choose finitely many numbers $\alpha_1, \ldots, \alpha_m$ such that

$$[-M, M] \subset \bigcup_{i=1}^{m} (\alpha_i - \epsilon, \alpha_i + \epsilon).$$

Consider the set Θ of all maps $\theta : \{x_1, \ldots, x_n\} \to \{\alpha_1, \ldots, \alpha_m\}$. (Note that θ is discrete and Θ is finite.) For each $\theta \in \Theta$, define the family of continuous functions

$$\mathcal{F}_\theta := \{f \in \mathcal{F} : f(x_i) \in B(\theta(x_i), \epsilon) \text{ for all } 1 \leqslant i \leqslant n\}.$$

Then $\mathcal{F} = \bigcup_{\theta \in \Theta} \mathcal{F}_\theta$.

Claim: $\|f - g\|_{C(E)} = \max_{x \in E} |f(x) - g(x)| < 4\epsilon$ if $f, g \in \mathcal{F}_\theta$.

Indeed, for each $x \in E$, we choose an index i such that $x \in B(x_i, \delta)$. Then we have

$$|f(x) - g(x)| \leqslant |f(x) - f(x_i)| + |f(x_i) - \theta(x_i)| + |\theta(x_i) - g(x_i)| + |g(x_i) - g(x)|.$$

Equicontinuity implies the first and last terms are $< \epsilon$; by definition of $\mathcal{F}_\theta$, the second and third terms are also $< \epsilon$. Hence for all $\epsilon > 0$, each set $\mathcal{F}_\theta$ has diameter $\leqslant 4\epsilon$, so this set $\mathcal{F}$ can be covered by *finitely* many sets having diameters $< 4\epsilon$. Since ϵ is arbitrary we are done. ∎

Beginning of Sept.29, 2021

3.2 Compact Operators

Definition 3.2.1: Compact Operator

Let X, Y be Banach. $\Lambda \in \mathcal{B}(X, Y)$ is **compact** if for every bounded sequence $\{x_n\} \subset X$, the corresponding $\{\Lambda x_n\} \subset Y$ admits a convergent subsequence $\{\Lambda x_{n_k}\}$.
It follows that TFAE:

(1) Λ is compact,

(2) For every bounded set $U \subset X$, the image $\Lambda(U) \subset Y$ has compact closure,

(3) The unit ball $B_1 \subset X$ has image $\Lambda(B_1) \subset Y$ whose closure is compact.

Theorem 3.2.2

(1) If the range of Λ is finite-dimensional then Λ is automatically compact.

(2) Let $\Lambda_n : X \to Y$ be compact for all n. Further assume that $\lim_{n\to\infty} \|\Lambda_n - \Lambda\| = 0$ for some Λ. Then Λ is compact.

Proof.

(1) Obvious by Riesz's lemma as $\overline{\Lambda(B_1)}$ is a closed and bounded in a finite-dimensional space.

(2) Since Y is complete, we want to show that $\overline{\Lambda(B_1)}$ is compact or equivalently $\Lambda(B_1)$ is precompact.

Let $\epsilon > 0$; we want to show that $\Lambda(B_1)$ can be covered by finitely many ϵ-balls. We begin by choosing k such that $\|\Lambda - \Lambda_k\| < \epsilon/2$ (and for all late terms). Since Λ_k is compact, there exist $\{y_1, \ldots, y_n\} \subset Y$ such that $\Lambda_k(B_1) \subset \bigcup_{i=1}^{n} B(y_i, \epsilon/2)$. For any x, if $\|x\| \leqslant 1$ then $\|\Lambda x - \Lambda_k x\| < \epsilon/2$. In particular, for this x, there exists y_i such that $\|\Lambda_k x - y_i\| < \epsilon/2$, so

$$\|\Lambda x - y_i\| \leqslant \|\Lambda x - \Lambda_k x\| + \|\Lambda_k x - y_i\| < \epsilon.$$

Since x is arbitrary, we conclude that $\Lambda(B_1) \subset \bigcup_{i=1}^{n} B(y_i, \epsilon)$, completing the proof.

□

Example 3.2.3. Let $\Lambda : \ell^2 \to \ell^2$ be defined by

$$x := (x_n)_{n\geqslant 1} \mapsto y = \Lambda x = (y_n/n)_{i\geqslant 1}.$$

Then Λ is clearly linear and bounded. We will show that Λ is compact.

Proof. Define $\Lambda_n := \ell^2 \to \ell^2$ by

$$(x_1, \ldots, x_n, x_{n+1}, \ldots) \mapsto (x_1/1, \ldots, x_n/n, 0, \ldots).$$

It is clear that each Λ_n is linear and bounded. The range of each Λ_n is finite-dimensional so they are compact! It remains to show that $\Lambda_n \to \Lambda$ in operator norm; if so, the (2) in the previous theorem shows Λ is compact. Indeed,

$$\|(\Lambda_n - \Lambda)(x)\|^2 = \sum_{j=n+1}^{\infty} (x_j/j)^2 \leqslant \frac{\|x\|^2}{(n+1)^2} \to 0$$

as $n \to \infty$ so $\|\Lambda_n - \Lambda\| < 1/(n+1)$. It follows that Λ is compact. □

Theorem 3.2.4: Compact Operator on Weakly Convergence Sequence

Let X, Y be normed and let $\Lambda : X \to Y$ a compact operator. If $\{x_n\} \subset X$ converges weakly to x, then $\{\Lambda x_n\} \subset Y$ is *strongly* convergent with limit Λx.

Proof. We write $y_n := \Lambda x_n$ and $y = \Lambda x$ for convenience.

We first show that $y_n \rightharpoonup y$. To this end, let $\varphi \in Y^*$ be given. Define the adjoint $(\Lambda^* \varphi) \in X^*$. Since $x_n \rightharpoonup x$, we have

$$(\Lambda^*\varphi)(x_n) \to (\Lambda^*\varphi)(x),$$

that is, (directly by definition)

$$\varphi(\Lambda x_n) \to \varphi(\Lambda x)$$

or equivalently $\varphi(y_n) \to \varphi(y)$, as claimed. Therefore $y_n \rightharpoonup y$.

We now show that $y_n \to y$ (strongly). Suppose for contradiction that there exists a subsequence $\{y_{n_k}\}$ with $\|y_{n_k} - y\| \geqslant \delta$ for some $\delta > 0$. Since $\{x_n\}$ converges weakly, it is bounded, so in particular $\{x_{n_k}\}$ is bounded. Since Λ is compact, in particular $\{\Lambda(x_{n_k})\}$ has a (strongly) convergent subsequence, say $\{\Lambda(x_{n_{k_j}})\}$. Let the limit be $\tilde{y}$. Notice that the first part shows $\{\Lambda(x_{n_{k_j}})\}$ converges weakly whereas we've just shown that it converges strongly to some other limit, contradiction (recall weak limit is unqiue). Therefore $y_n \to x$ strongly, as claimed. □

Beginning of Oct.1, 2021

Theorem 3.2.5: Adjoint of a Compact Operator

Let X, Y be Banach and let $\Lambda \in \mathcal{B}(X, Y)$. Then Λ is compact if and only if $\Lambda^* : Y^* \to X^*$ is compact.

Proof. We will only prove the forward direction. Let $\{y_n^*\}$ be a sequence in Y^* with $\|y_n^*\| \leqslant 1$. We need to show that $\{\Lambda^* y_n^*\}$ has a convergent subsequence.
Let $B_1 := \{x \in X : \|x\| \leqslant 1\}$. By assumption ΛB_1 has a compact closure $E := \overline{\Lambda B_1} \subset Y$.
Let $f_n := y_n^*|_E$ (i.e., functions from E to $\mathbb{K}$). Then

$$|f_n(y) - f_n(y')| \leqslant \|y_n^*\|\|y - y'\| \leqslant \|y - y'\|.$$

This implies f_n's are uniformly Lipschitz (and in particular equicontinuous). Further note that $\sup_{y \in E} \|y\| = \sup_{\|x\| \leqslant 1} \|\Lambda x\|$, so $|f_n(y)| \leqslant \|y_n^*\|\|y\|$ where both terms are bounded. Hence all f_n's are uniformly bounded by $\|\Lambda\|$.
Now we can apply Arzelá-Ascoli and extract a subsequence $\{f_{n_k}\}$ that converges to a function f uniformly on

$E = \overline{\Lambda B_1}$. Now note that

$$\begin{aligned} \|\Lambda^* y^*_{n_j} - \Lambda^* y^*_{n_i}\| &\triangleq \sup_{\|x\|\leqslant 1} |\langle \Lambda^* y^*_{n_j} - \Lambda^* y^*_{n_i}, x\rangle| \\ &= \sup_{\|x\|\leqslant 1} |\langle y^*_{n_j} - y^*_{n_i}, \Lambda x\rangle| \\ &= \sup_{\|x\leqslant 1\|} |f_{n_i}(\Lambda x) - f_{n_i}(\Lambda x)| \to 0 \text{ as } i, j \to \infty. \end{aligned}$$

Therefore $\{\Lambda^* y^*_{n_i}\}$ is Cauchy and converges to a limit $x^* \in X^*$. Hence Λ^* is compact. □

3.3 Integral Operators

Theorem 3.3.1: Integral Operator

Let $K : [a, b] \times [a, b] \to \mathbb{R}$ be continous. Then the **integral operator** Λ defined by

$$(\Lambda f)(x) := \int_a^b K(x, y) f(y)\, \mathrm{d}y$$

is a compact operator from $C([a, b])$ to $C([a, b])$.

Proof. Consider a bounded sequence of continuous functions $\{f_n\} \subset C([a, b])$. We need to show that Λf_n admits a uniformly convergent subsequence.

By Arzelá-Ascoli, it suffices to show that Λf_n are uniformly bounded and equicontinuous.

Note that there exists $k > 0$ such that $|K(x, y)| \leqslant k$ for all $(x, y) \in [a, b]^2$ since K is continuous on a compact set.

Also, for $\epsilon > 0$, there exists $elta > 0$ such that $|x - \tilde{x}| < \delta$ then $|K(x, y) - K(\tilde{x} - y)| < \epsilon$ for all $x, \tilde{x}, y \in [a, b]$ (uniform continuity).

By assumption, there exists $M > 0$ such that $\|f_n\| \leqslant M$ (by assumption). Then

$$|\Lambda f_n(x)| = \left| \int_a^b K(x, y) f_n(y)\, \mathrm{d}y \right| \leqslant kM(b - a)$$

so Λf_n are uniformly bounded.

For equicontnuity, let $\epsilon > 0$ be given. Choose $\delta > 0$ such that the ("x-$\tilde{x}$-y" uniform continuity) property above holds. Then if $|x - \tilde{x}| < \delta$,

$$\begin{aligned} |\Lambda f_n(x) - \Lambda f_n(\tilde{x})| &= \left| \int_a^b (K(x, y) - K(\tilde{x}, y) f_n(y)\, \mathrm{d}y \right| \\ &\leqslant \int_a^b |K(x, y) - K(\tilde{x}, y)||f_n(y)|\, \mathrm{d}y \\ &< \epsilon M(b - a), \end{aligned}$$

so Λf_n are equicontinuous too. The claim then follows from Arzelá-Ascoli. □

Example 3.3.2. For any $f \in C([-1,1])$, consider the boundary value problem

$$\begin{cases} u''(x) + f(x) = 0 \\ u(-1) = u(1) = 0. \end{cases} \tag{*}$$

Let Λ be the solution u to (*) for a given $f : u = \Lambda f$. (We claim that there does exist Λ satisfying this relation. Note that u is unique — if u_1, u_2 are both solutions to (*) then $w = u_1 - u_2$ satisfies $w''(x) = 0$ and $w(-1) = w(1) = 0$, i.e., $u_1 = u_2$.)

In fact, $u''(x) = -f(x)$ so by FTC

$$u'(x) = u'(1) - \int_1^x f(y)\,\mathrm{d}y = u'(-1) - \int_{-1}^x f(y)\,\mathrm{d}y.$$

Integrating further gives

$$\begin{aligned} u(x) &= u'(1)(x+1) - \int_{-1}^x \int_1^z f(y)\,\mathrm{d}y\,\mathrm{d}z && (3.1) \\ &= u'(1)(x-1) - \int_1^x \int_1^x f(y)\,\mathrm{d}y\,\mathrm{d}z && (3.2) \\ &= u'(-1)(x+1) - \int_{-1}^x \int_{-1}^z f(y)\,\mathrm{d}y\,\mathrm{d}z && (3.3) \\ &= u'(-1)(x-1) - \int_1^x \int_{-1}^z f(y)\,\mathrm{d}y\,\mathrm{d}z. && (3.4) \end{aligned}$$

Equating (1) and (2) gives

$$u'(1) = \frac{1}{2}\left[\int_{-1}^x \int_1^z f(y)\,\mathrm{d}y\,\mathrm{d}z - \int_1^x \int_1^z f(y)\,\mathrm{d}y\,\mathrm{d}z\right] \tag{5}$$

and equating (3) and (4) gives

$$u'(-1) = \frac{1}{2}\left[\int_{-1}^x \int_{-1}^z f(y)\,\mathrm{d}y\,\mathrm{d}z - \int_1^x \int_{-1}^z f(y)\,\mathrm{d}y\,\mathrm{d}z\right]. \tag{6}$$

Adding (1), (2), (3), and (4) gives

$$\begin{aligned} 4u(x) &= 2x\left[u'(1) + u'(-1)\right] - \left[\int_{-1}^x \int_1^z f(y)\,\mathrm{d}y\,\mathrm{d}z + \int_1^x \int_1^z f(y)\,\mathrm{d}y\,\mathrm{d}z + \dots\right] \\ &= \dots(\text{substitute (5) and (6)}) \\ &= 2(x+1)\int_1^x (y-1)f(y)\,\mathrm{d}y - 2(x-1)\int_{-1}^x (y+1)f(y)\,\mathrm{d}y. \end{aligned}$$

Therefore $u(x)$ is an integral equation where

$$K(x,y) := \begin{cases} \frac{(1-y)(1+x)}{2} & -1 \leqslant x \leqslant 1 \\ \frac{(1+y)(1-x)}{2} & y \leqslant x \leqslant 1 \end{cases}$$

Hence K is continuous and $\Lambda : f \mapsto u = \Lambda f$ is compact.

Chapter 4

Hilbert Spaces

Beginning of Oct.4, 2021

The Euclidean space $\mathbb{R}^n$ is equipped with a natural inner product $\langle \cdot, \cdot \rangle$ which induces the Euclidean norm $\|x\| := \sqrt{\langle x, x \rangle}$. This also defines "perpendicular spaces" and "perpendicular projections", which allow us to construct bases of mutually orthogonal vectors via Gram-Schmidt.

Why are Hilbert spaces interesting?

(1) Every linear function $\varphi : \mathbb{R}^n \to \mathbb{R}$ can be represented as an inner product: $\varphi(x) = \langle w, x \rangle$ for some $w \in \mathbb{R}^n$.

(2) with an inner product, we can define a class of *symmetric operators* (i.e., $A : \mathbb{R}^n \to \mathbb{R}^n$ with $\langle Ax, y \rangle = \langle x, Ay \rangle$ for all $x, y \in \mathbb{R}^n$) with many useful properties.

(3) We can also define a class of positive operators (i.e., $A : \mathbb{R}^n \to \mathbb{R}^n$ strictly positive definite: $Ax > 0$ for all $x \neq 0$). In this case the map $x \mapsto \langle Ax, x \rangle$ is a positive definite quadratic form.

We wish to show how the definition and properties of the Euclidean space can be extended to infinite-dimensional spaces.

4.1 Spaces with an Inner Product

Definition 4.1.1: Inner Product

Let H be a vector space over $\mathbb{K}$ (either $\mathbb{R}$ or $\mathbb{C}$). An **inner product** on H is a map $\langle \cdot, \cdot \rangle : H^2 \to \mathbb{K}$ satisfying

(1) (skew symmetry) $\langle x, y \rangle = \overline{\langle y, x \rangle}$,

(2) (bilinearity and *sesquilinearity*) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ and $\langle \lambda x, z \rangle = \lambda \langle x, z \rangle$ (for $\lambda \in \mathbb{K}$)

(3) (non-degeneracy) $\langle x, x \rangle \geqslant 0$ with = if and only if $x = 0$.

Note that these imply $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$ and $\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle$.

Definition 4.1.2

Define $\|x\| := \sqrt{\langle x, x\rangle}$.

Theorem 4.1.3

Let H be a vector space with inner product $\langle \cdot, \cdot \rangle$. Then

(1) (Cauchy-Schwarz) $|\langle x, y\rangle| \leqslant \|x\| \|y\|$.

(2) (triangle inequality / Minkowski inequality) $\|x + y\| \leqslant \|x\| + \|y\|$.

Proof.

(1) If $y = 0$ the claim is trivial. Otherwise consider inner products of form $\langle x + \lambda y, x + \lambda y\rangle$ for $\lambda \in \mathbb{K}$:

$$\langle x + \lambda y, x + \lambda y\rangle = \langle x, x\rangle + 2\lambda \langle x, y\rangle + \lambda^2 \langle y, y\rangle = \|x\|^2 + 2\lambda \langle x, y\rangle + \lambda^2 \|y\|^2$$

which by non-degeneracy of norm, nonnegative. Hence the RHS, viewed as a quadratic of λ, has a non-positive discriminant, that is,

$$(2 \langle x, y\rangle)^2 - 4\|x\|^2 \|y\|^2 \leqslant 0 \implies \langle x, y\rangle^2 \leqslant \|x\|^2 \|y\|^2 \implies |\langle x, y\rangle| \leqslant \|x\| \|y\|.$$

(2) By (1), $\mathfrak{Re} \langle x, y\rangle \leqslant |\langle x, y\rangle| \leqslant \|x\| \|y\|$, so

$$\|x + y\|^2 = \langle x + y, x + y\rangle = \|x\|^2 + 2\mathfrak{Re} \langle x, y\rangle + \|y\|^2 \leqslant (\|x\| + \|y\|)^2.$$

□

Definition 4.1.4: Hilbert Space

A vector space H with inner product $\langle \cdot, \cdot \rangle$, which is complete with respect to the induced norm, is called a **Hilbert space**.

Example 4.1.5.

(1) $\mathbb{R}^n$ with $\langle x, y\rangle := \sum_{i=1}^{n} x_i y_i$ is Hilbert over $\mathbb{R}$.

(2) ℓ^2 defined by $\{x = \{x_i\} : \left(\sum_{i=1}^{\infty} |x_i|^2\right)^{1/2} < \infty\}$ is Hilbert over $\mathbb{C}$ with $\langle x, y\rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$.

(3) Let $\Omega \subset \mathbb{R}^n$ be open. Let $L^2(\Omega; \mathbb{R})$ be the collection of square summable maps. Then it is Hilbert with

$$\langle f, g\rangle = \int_{\Omega} f(x) g(x) \, \mathrm{d}x.$$

4.2 Orthogonal Projections

Definition 4.2.1: Span

Given a subset $S \subset H$, define

$$\text{span}(S) := \{\sum_{i=1}^{n} c_i x_i : n \in \mathbb{N}, c_i \in \mathbb{R}, x_i \in S\}$$

(*note we are taking combinations of finitely many elements*). In general, $\text{span}(S)$ is a subspace of H but it is not necessarily closed. To this end, the closure $V := \overline{\text{span}(S)}$ is called the space *generated by* S.

The set S is called " **total**" if it generates H.

Definition 4.2.2: Orthogonality

We say $x, y \in H$ are **orthogonal** if $\langle x, y\rangle = 0$. Given $S \subset H$, its **orthogonal subspace** is defined as

$$S^\perp := \{y \in H : \langle y, x\rangle = 0 \text{ for all } x \in S\}.$$

Theorem 4.2.3: Orthogonal Projections

Let H be Hilbert and let $V \subset H$ be a closed subspace.

(1) $H = V \oplus V^\perp$ in the sense that each $x \in H$ can be *uniquely* written as $x = y + z$ where $y \in V$ and $z \in V^\perp$.

(2) $y := P_V(x)$ in V having the minimal distance to x and $z := P_{V^\perp}(x)$ is the unique point in $V^\perp$ having the minimal distance from x.

(3) The **perpendicular projection** $x \mapsto y$ and $x \mapsto z$ are linear, continuous, and bounded with norm $\leqslant 1$.

Beginning of Oct.6, 2021

Proof. We first show that given $x \in H$, there exists a unique $y \in V$ with minimal distance from x. That is, if

$$\alpha = d(x, V) := \inf_{y \in V} \|x - y\|$$

then the infimum is attained. From definition, there exists a sequence $\{y_n\}$ with $\lim_{n\to\infty} \|x - y_n\| = \alpha$. Since V is Banach, it suffices to show $\{y_n\}$ is Cauchy. Recall the parallelogram law

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

Letting $u := x - y_m$ and $v := x - y_n$ we see that

$$\begin{aligned}\|y_n - y_m\|^2 &= 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - \|2x - y_m - y_n\|^2 \\ &= 2\|x - y_m\|^2 + 2\|x - y_n\|^2 - 4\|x - \underbrace{(y_m + y_n)/2}_{\in V}\|^2 \\ &\leqslant 2\|x - y_m\|^2 + 2\|x - y_n\|^2 + \alpha^2.\end{aligned}$$

Taking $\limsup$ gives

$$\limsup_{m,n\to\infty} \|y_m - y_n\|^2 \leqslant 2\limsup_{m\to\infty} \|x - y_m\|^2 + 2\limsup_{n\to\infty} \|x - y_n\|^2 - 4\alpha^2 = 0.$$

Since V is complete, $\{y_n\}$ converges to a unique limit y with $\|x - y\| = d(x, V)$.

For uniqueness, suppose $\|x - y'\| = d(x, V)$. By the same argument

$$\|y - y'\|^2 = 2\|x - y\|^2 + 2\|x - y'\|^2 - 4\|2x - y - y'\|^2 \leqslant 2\alpha^2 + 2\alpha^2 - 4\alpha^2 = 0.$$

Now that the map $x \mapsto P_V(x)$ is well-defined, we show $X = V \oplus V^\perp$. That is, $P_V(x)$ is the unique point $y \in V$ such that $x - y \in V^\perp$. Let $v \in V$ be given; for $\lambda \in \mathbb{R}$, consider $f : \lambda \mapsto \|x - (y + \lambda v)\|^2$:

$$f(\lambda) = \|x - (y + \lambda v)\|^2 = \|x - y\|^2 + |\lambda|^2 \|v\|^2 + 2\mathfrak{Re} \langle x - y, \lambda v\rangle .$$

By our first part, f attains its unique global minimum at $\lambda = 0$. Thus $f'(0) = 0$. Differentiating implies

$$\mathfrak{Re} \langle x - y, v\rangle = 0 \qquad \text{for all } v \in V.$$

If H is over $\mathbb{R}$ then we are done; if it is over $\mathbb{C}$, simply repeat the above process with v replaced by $-iv$ so

$$\mathfrak{Im}(x - y, v) = \mathfrak{Re}(x - y, -iv) = 0.$$

For uniqueness of the orthogonality statement, suppose $y' \in V$ also has $x - y' \in V^\perp$. Then

$$\begin{aligned}\|y - y'\|^2 &= \langle y - y', (y - x) + (x - y')\rangle \\ &= \langle y - y', x - y'\rangle + \langle y - y', y - x\rangle = 0 + 0 = 0.\end{aligned}$$

Finally, we show that $P_V : H \to V$ is a bounded linear operator. Put $y = P_V(x), y' = P_V(x')$ where x, x' are arbitrary elements of H. For $\alpha, \alpha' \in \mathbb{K}$, we know $\alpha y + \alpha' y' \in V$. Since for any $v \in V$

$$< \underbrace{\alpha x + \alpha' x'}_{\in H} - \underbrace{\alpha y + \alpha' y'}_{\in V}, v >= \alpha \langle x - y, v\rangle + \alpha' \langle x' - y', v\rangle = 0$$

by the second part the map $(\alpha_x + \alpha' x') \mapsto \alpha y + \alpha' y'$ *must be* the projection operator P_V. Hence

$$P_V(\alpha x + \alpha' x') = \alpha y + \alpha' y' = \alpha P_V(x) + \alpha' P_V(x')$$

and similarly $I - P_V = P_{V^\perp}$ is also linear. Finally, since $P_V(x), P_{V^\perp}(x) = x - P_V(x)$,

$$\|P_V(x)\|^2 = \|x\|^2 - \|x - P_V(x)\|^2 \leqslant \|x\|^2$$

we see P_V has operator norm $\leqslant 1$. For $x \in V$, the norm 1 is indeed attained. □

4.3 Linear Functionals on a Hilbert Space

Theorem 4.3.1: Riesz Representation Theorem

Let H be Hilbert.

(1) For each $x \in H$, the map $y \mapsto \langle y, x\rangle$ is a continuous linear functional on H;

(2) Let $y \mapsto Ay$ be a continuous linear functional. Then for each y, $Ay = \langle h, y\rangle$ for some unique $h \in H$.

Proof.

(1) Let x be given. The map $\varphi^x : \varphi^x(y) = \langle y, x\rangle$ is linear, and

$$\|\varphi^x\| = \sup_{\|y\|\leqslant 1} |\langle y, x\rangle| \leqslant \sup_{\|y\|\leqslant 1} \|y\|\|x\| = \|x\|.$$

(2) If $Ay \equiv 0$ for all $y \in H$ then it corresponds to the zero map and the claim holds with $h = 0$. For $A \neq 0$, there exists $z_0 \neq 0$ such that $z_0 \in N(A)^\perp$. Let

$$v = (Ay)z_0 - (Az_0)y.$$

For any $y \in H$,

$$Av = (Ay)Az_0 - (Az_0)Ay = 0$$

so $v \in N(A)$. Therefore $\langle v, z_0\rangle = 0$ and so

$$\langle (Ay)z_0 - (Az_0)y, z_0\rangle = Ay(z_0, z_0) - (Az_0)\langle y, z_0\rangle$$

and

$$Ay = \frac{(Az_0)\langle y, z_0\rangle}{\langle z_0, z_0\rangle} = \left\langle y, z_0 \cdot \overline{Az_0} / \langle z_0, z_0\rangle \right\rangle.$$

Therefore we have found $h \in H$ satisfying the condition. For uniqueness, if $Ay = \langle y, h_1\rangle = \langle y, h_2\rangle$ for all $y \in H$, we have $\langle y, h_1 - h_2\rangle = 0$ for all $y \in H$. Choosing $y = h_1 - h_2$ gives $\langle h_1 - h_2, h_1 h_2\rangle = 0$, done.

□

Beginning of Oct.8, 2021

Remark.

(1) If H is Hilbert over $\mathbb{R}$, then $x \mapsto \varphi^x$ is an isometric isomorphism between H and H^*.

(2) If $\Lambda : H \to H$ is linear, then its adjoint $\Lambda^* : H^* \to H^*$ can be identified with $\Lambda : H \to H$ via

$$(x, \Lambda^* y) = (\Lambda x, y).$$

4.4 Gram-Schmidt Orthogonalization & Orthonormal Sets

Quick recap:

(1) $x \in H$ is called **normalized** if $\|x\| = 1$.

(2) A subset $E \subset H$ is **orthonormal** if for $\|e_i\| = 1$ and $\langle e_i, e_j \rangle = \delta_{i,j}$ for all $e_i, e_j \in H$, where $\delta_{i,j}$ is the Kronecker delta.

(3) If $S := \{v_1, \ldots, v_n\} \subset H$, the **span** is defined by $\operatorname{span}(S) = \sum_{i=1}^{n} \theta_k v_k$ where $\theta_k \in \mathbb{K}$ (can be 0).

(4) If $x = \sum_{k=1}^{n} \theta_k v_k$, we note $\langle x, v_j \rangle = \sum_{k=1}^{n} \theta_k \langle v_k, v_j \rangle = \theta_k$. This gives $x = \sum_{k=1}^{n} \langle x, v_k \rangle v_k$. If we weaken the assumption and let $\{v_1, \ldots, v_n\}$ only be orthogonal (not necessarily normalized) then $x = \sum_{k=1}^{n} (\langle x, v_k \rangle / \|v_k\|) v_k$.

Theorem 4.4.1: Gram-Schmidt Orthogonalization

Let $\{v_1, \ldots, v_n\}$ be a collection of linearly independent vectors. Define inductively

$$e_1 := \frac{v_1}{\|v_1\|}, \qquad \tilde{v}_n := v_n - \sum_{k=1}^{n-1} \langle v_n, e_k \rangle e_k, \qquad \text{and } e_n := \frac{\tilde{v}_n}{\|\tilde{v}\|_n}.$$

Then $\{e_1, \ldots, e_n\}$ is orthonormal with $\operatorname{span}\{e_1, \ldots, e_k\} = \operatorname{span}\{v_1, \ldots, v_k\}$ for all $1 \leqslant k \leqslant n$.

Question. In an infinite dimensional space, how do the notions of basis and linear span work? It is important to understand when $\sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$ converges and when it does not.

Proposition 4.4.2

Let H be Hilbert. For any $S \subset H$, the orthogonal space $S^\perp$ is closed in H. Moreover,

$$\operatorname{span}(S) \text{ is dense in } H \Leftrightarrow S^\perp = \{0\}.$$

Proof. For $\Rightarrow$, let $x \in S^\perp$. Since $\operatorname{span}(S)$ is dense in H, there exists a sequence $\{x_n\}$ of linear combinations of basis of S such that $x_n \to x$. We write

$$x_n := \sum_{k=1}^{N_n} \theta_{n,k} a_{n,k} \to 0.$$

Then,

$$\langle x, x \rangle = \lim_{n\to\infty} \langle x, x_n \rangle = \lim_{n\to\infty} \sum_{k=1}^{N_n} \overline{\theta_{n,k}} \underbrace{\langle x, a_{n,k} \rangle}_{=0} = \lim_{n\to\infty} \sum_{k=1}^{N_n} = 0.$$

Hence, $x = 0$ and thus $S^\perp = \{0\}$.

For $\Longleftarrow$, let $V = \overline{\operatorname{span}(S)}$. Suppose for contradiction that $V \neq H$ so that there exists $y \in H - V$. Consider the orthogonal projection $P_V(y)$. Since $y \notin V$, $w := y - P_V(y) \neq 0$. On the other hand, since $w \in V^\perp$, $w \in S^\perp$, so by assumption $w = 0$. Contradiction! □

Theorem 4.4.3: Bessel's Inequality

Let $S = \{e_1, e_2, \ldots\}$ be a finite or countable orthonormal set in a Hilbert space H. Let $V := \overline{\operatorname{span}(S)}$ and let $P_V : H \to V$ be the orthogonal projection. Then for every $x \in H$,

$$\sum_{k\geqslant 1} |\langle x, e_k\rangle|^2 = \|P_V(x)\|^2 \leqslant \|x\|^2.$$

Moreover, $\sum_{k\geqslant 1} \langle x, e_k\rangle e_k = P_V(x)$. In particular, if $V = H$, $\sum_{k\geqslant 1} \langle x, e_k\rangle = P_V(x) = x$ for all $x \in V = H$.

Proof. Let $V_n := \operatorname{span}\{e_1, \ldots, e_n\}$. We know $P_{V_n}(x) = \sum_{k=1}^{n} \langle x, e_k\rangle e_k$, so

$$\|P_{V_n}(x)\|^2 = \left\langle \sum_{j=1}^{n} \langle x, e_j\rangle e_j, \sum_{k=1}^{n} \langle x, e_k\rangle e_k \right\rangle = \sum_{j,k=1}^{n} \underbrace{\langle x, e_j\rangle \overline{\langle x, e_k\rangle}}_{\in \mathbb{R}} \underbrace{\langle e_j, e_k\rangle}_{=\delta_{j,k}}$$

$$= \sum_{k=1}^{n} \langle x, e_k\rangle \overline{\langle x, e_k\rangle} = \sum_{k=1}^{n} |\langle x, e_k\rangle|^2.$$

Since $\|P_{V_n}(x)\|^2 \leqslant \|x\|^2$ for all n, taking limit proves Bessel's inequality.

For the second claim, suppose S is countably infinite. By convergence,

$$x_n := \sum_{k=1}^{n} \langle x, e_k\rangle e_k$$

is in particular Cauchy. Thus, for $m < n$,

$$\|x_n - x_m\|^2 = \sum_{k=m+1}^{n} |\langle x, e_k\rangle|^2 \to 0$$

as $m, n \to \infty$. Since H is complete, we get $x_n \to \tilde{x}$ for some $\tilde{x} \in H$. It remains to show that $\tilde{x} = P_V(x)$. Since $x_n \in V$ for all n and $x_n \to \tilde{x}$ and since V is closed, $\tilde{x} \in V$ and so

$$\langle x - \tilde{x}, e_k\rangle = \lim_{n\to\infty} \langle \tilde{x} - x_n, e_k\rangle = 0 \qquad \text{for all } k$$

Therefore $x - \tilde{x}$ is orthogonal to all e_k's and therefore all $v \in V$. Therefore $\tilde{x} \in V$ and $x - \tilde{x} \in V^\perp$. This implies $\tilde{x} \in P_V(x)$. □

Beginning of Oct.11, 2021

Definition 4.4.4: Orthonormal Basis

An orthonormal set $S = \{e_1, e_2, \ldots\} \subset H$ is an **orthonormal basis** (**total orthonormal set**) if $\operatorname{span}(S)$ is dense in H, i.e., $\overline{\operatorname{span}(S)} = H$.

Remark. Bessel's inequality becomes identity: if we let $V := \overline{\operatorname{span}(S)}$ then

$$\sum_{k\geqslant 1} |\langle x, e_k\rangle|^2 = \|P_V(x)\|^2 = \|x\|^2 \qquad \text{and} \qquad \sum_{k\geqslant 1} \langle x, e_k\rangle e_k = x.$$

This is known as the **Parseval identity**.

Fourier Series

Let $H = L^2([-\pi,\pi];\mathbb{C})$ with the natural inner product

$$\langle f,g\rangle = \int_{-\pi}^{\pi} f(x)\overline{g(x)}\,\mathrm{d}x.$$

Then the set S of functions

$$\varphi_n(x) = \frac{1}{\sqrt{2\pi}}e^{inx} \qquad n\in\mathbb{Z}$$

is orthonormal:

$$\langle\varphi_n(x),\varphi_m(x)\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{imx}\overline{e^{inx}}\,\mathrm{d}x = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(m-n)x}\,\mathrm{d}x = \delta_{m,n}.$$

We claim that S is an orthonormal basis of H. To show that span(S) is dense, let $f\in L^2([-\pi,\pi])$. For all $\epsilon>0$, there exists a continuous function $f_\epsilon : [-\pi,\pi]\to\mathbb{C}$ such that

$$\|f_\epsilon - f\|_{L^2} < \epsilon \qquad f_\epsilon(-\pi) = f_\epsilon(\pi).$$

Also, by Stone-Weierstraß, we can find a complex trigonometric polynomial of form

$$p(x) = \sum_{k=-N}^{N}\alpha_k e^{ikx}$$

such that

$$\|f_\epsilon - p\|_{C^0} = \max|f_\epsilon(x) - p(x)| < \epsilon.$$

Then

$$\|f_\epsilon - p\|_{L^2} = \left(\int_{-\pi}^{\pi}|f_\epsilon(x)-p(x)|^2\,\mathrm{d}x\right)^{1/2} < \sqrt{2\pi}\epsilon.$$

Therefore f_ϵ can be approximated by trig polynomials also with respect to the L^2-norm. Therefore using p to approximate f_ϵ and using f_ϵ to approximate f, we conclude that

$$\overline{\text{span}(S)} = L^2([-\pi,\pi]).$$

Now we consider the complex trig series

$$\sum_{k=-\infty}^{\infty} a_k\frac{e^{ikx}}{\sqrt{2\pi}}$$

where

$$a_k := \langle f,\varphi_k\rangle = \int_{-\pi}^{\pi} f(x)\frac{e^{-ikx}}{\sqrt{2\pi}}\,\mathrm{d}x.$$

By the previous theorem, the series converges to $f\in L^2$, i.e.,

$$\lim_{n\to\infty}\|f - \sum_{k=-n}^{n} a_k\varphi_k\|_{L^2} = 0.$$

In other words, for $f\in L^2([-\pi,\pi];\mathbb{C})$, by letting $c_k := \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} f(y)e^{-iky}\,\mathrm{d}y$,

$$\lim_{n\to\infty}\int_{-\pi}^{\pi}|f(x) - \sum_{k=-n}^{n} c_k e^{ikx}|^2\,\mathrm{d}x = 0.$$

This gives the convergence of Fourier series in the L^2 sense.

4.5 Separable Hilbert Spaces

Theorem 4.5.1

Let H be Hilbert.

(1) If H is separable, then every or tho normal set in H is countable.

(2) If H contains an orthonormal sequence which is total in H, then H is separable.

Proof of (1). Let H be separable and let B be any countable dense set in H. For any orthonormal set S and $x, y \in S$, $\|x-y\|^2 = \langle x-y, x-y\rangle = \|x\|^2 + \|y\|^2 = 2$. Hence, letting $B_x := B(x, 1/4)$, for each distinct pair of $x, y \in S$, $B_x \cap B_y = \varnothing$. By density, every ball contains some element in B. Hence S is at most countable. □

Definition 4.5.2: isomorphism

Let H and $\tilde{H}$ be Hilbert over $\mathbb{K}$. An (**isometric**) **isomorphism** between H and $\tilde{H}$ is a bijection linear map $\Lambda : H \to \tilde{H}$ such that

$$\langle \Lambda x, \Lambda y\rangle = \langle x, y\rangle .$$

If such an isomorphism exists, then H and $\tilde{H}$ are called **isomorphic** Hilbert spaces.

Theorem 4.5.3

Any two separable infinite-dimensional Hilbert spaces are isomorphic.

Proof. Let $S = \{e_1, e_2, \dots\}$ and $\tilde{S} = \{\tilde{e}_1, \tilde{e}_2, \dots\}$ be countable total orthonormal sets in H and $\tilde{H}$, respectively. We write $x \in H$ as

$$x = \sum_{k\geqslant 1} \langle x, e_k\rangle e_k.$$

We define $\Lambda : H \to \tilde{H}$ by

$$x = \sum_{k\geqslant 1} \langle x, e_k\rangle e_k \mapsto \sum_{k\geqslant 1} \langle x, e_k\rangle \tilde{e}_k =: x.$$

By Bessel's inequality, the sum on the RHS is finite. Finally, note that

$$\|\tilde{x}\|^2 = \|\Lambda x\|^2 = \left(\sum_{k\geqslant 1} \langle x, e_k\rangle \tilde{e}_k, \sum_{j\geqslant 1} \langle x, e_j\rangle \tilde{e}_j \right) = \sum_{k\geqslant 1} |\langle x, e_k\rangle|^2 = \|x\|^2.$$

This shows Λ is injective and it preserves norms. Using the polarization identity

$$\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) + \frac{1}{4} i(\|x+iy\|^2 - \|x-iy\|^2)$$

we see $\langle \Lambda x, \Lambda y\rangle = \langle x, y\rangle$. Hence the isometry.

Finally, for surjectivity, let $\tilde{x} \in \tilde{H}$ be given. We write $\tilde{x} = \sum_{k\geqslant 1} \alpha_k \tilde{e}_k$ where $\sum_{k\geqslant 1} |\alpha_k|^2 < \infty$. Then

$$\sum_{k\geqslant 1} \alpha_k e_k$$

converges to some $\in H$ with $\alpha_k = \langle x, e_k\rangle$. Therefore $\tilde{x} = \Lambda x$. This shows Λ is surjective. □

Remark. Any separable infinite-dimensional Hilbert space is isomorphic to $L^2([-\pi,\pi])$ or ℓ^2.

Other Total Orthonormal Sets

Legendre Polynomials

Consider a Hilbert space $H = L^2([-1,1];\mathbb{R})$ with

$$\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\,\mathrm{d}t.$$

We start with $f_0(t) = 1, f_1(t) = t, f_2(t) = t^2, \ldots$, so $f_n(t) = t^n$. Note that f_n's are linearly independent. Applying Gram-Schmidt gives us an orthonormal sequence $\{e_n\}_{n\geqslant 0}$. Then $\{e_n\}$ is total in L^2.

Beginning of Oct.13, 2021

The explicit formula is given by

$$e_n(t) = \sqrt{(2n+1)/2}\cdot p_n(t)$$

where

$$p_n(t) = \frac{1}{2^n n!}\frac{\mathrm{d}^n}{\mathrm{d}t^n}(t^2-1)^n.$$

(p_n is called the *Legendre polynomial of order* n.)

Claim 1. $p_n(t) = \sum_{j=0}^{[n/2]}(-1)^j\frac{(2n-2j)!}{2^n j!(n-j)!(n-2j)!}t^{n-2j}$. This claim can be obtained by expanding $(t^2-1)^n$ using the binomial theorem and differentiating the term n times.

Claim 2. $\|p_n\|^2 = 2/(2n+1)$. This justifies the scaling constant appearing in e_n. This is obtained by iterative IBP on

$$\begin{aligned}(2^n n!)^2\|p_n\|^2 &= \int_{-1}^{1}\frac{\mathrm{d}^n}{\mathrm{d}t^n}(t^2-1)^n\cdot\frac{\mathrm{d}^n}{\mathrm{d}t^n}(t^2-1)^n\,\mathrm{d}t\\ &= -\int_{-1}^{1}\frac{\mathrm{d}^{n-1}}{\mathrm{d}t^{n-1}}(t^2-1)^n\frac{\mathrm{d}^{n+1}}{\mathrm{d}t^{n+1}}(t^2-1)^n\,\mathrm{d}t\\ &= \cdots\\ &= (-1)^n(2n)!\int_{-1}^{1}(t^2-1)^n\,\mathrm{d}t = (-1)^n(2n)!\cdot(-1)^n\cdot 2\int_0^1(1-t^2)^n\,\mathrm{d}t\\ &= 2(2n)!\int_0^{\pi/2}\cos^{2n+1}(\theta)\,\mathrm{d}\theta = \frac{2^{2n+1}(n!)^2}{2n+1}.\end{aligned}$$

Claim 3. $\{p_n\}$ is orthogonal in $L^2([-1,1])$. To see this, it suffices to show that $\langle p_m, p_n\rangle = 0$ when $0\leqslant m<n$. Since p_n is a polynomial, it suffices to fix n and show $\langle t^m, p_n\rangle = 0$ for all $m<n$:

$$2^n n!\,\langle t^m, p_n\rangle = \int_{-1}^{1} t^m\frac{\mathrm{d}^n}{\mathrm{d}t^n}(t^2-1)^n\,\mathrm{d}t.$$

We can increase the order of $\frac{\mathrm{d}^n}{\mathrm{d}t^n}$ and reduce that on t^m using IBP. Since $m<n$, t^m eventually becomes 0, resulting in the integral being 0.

In fact, the Legendre polynomials are solutions of the *Legendre differential equations*

$$(1-t^2)p_n'' - 2tp_n' + n(n+1)p_n = 0.$$

Remark. The *Sturm-Liouville problem*

$$((1-t^2)y')' + \lambda y = 0 \qquad -1 \leqslant t \leqslant 1$$

has a bounded solution exactly when $\lambda = n(n+1)$, resulting in a Legendre polynomial. These $\lambda = n(n+1)$ are eigenvalues of $Ay = -((1-t^2)y')'$.

Example 4.5.4. Recall the spherical coordinates (ρ, θ, φ) and Cartesian (x, y, z) are related by

$$x = \rho\cos\theta\sin\varphi \qquad y = \rho\sin\theta\sin\varphi \qquad z = \rho\cos\varphi.$$

Then

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

has spherical coordinate expression

$$\Delta = \frac{\partial^2}{\partial\rho^2} + \frac{2}{\rho}\frac{\partial}{\partial p} + \frac{1}{\rho^2\sin^2\varphi}\frac{\partial^2}{\partial\theta^2} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\varphi^2} + \frac{\cot\varphi}{\rho^2}\frac{\partial}{\partial\varphi}$$

for $\rho > 0, \theta \in \mathbb{R}, \varphi \in (0, 2\pi)$.

Problem. Find bounded solutions $u = u(\rho, \theta, \varphi)$ of the Laplace equation $\Delta u = 0$ in the ball.

Intuitively we try to reduce PDEs to ODEs, so we try to look for separable equations $u(\rho, \theta, \varphi) = R(\rho)\Theta(\theta)\Phi(\varphi)$. Then

$$\rho^2\frac{R''}{R} + 2\rho\frac{R'}{R} + \frac{\Theta''}{\sin^2\varphi\Theta} + \frac{\Phi''}{\Phi} + \frac{\cot\varphi\Phi'}{\Phi} = 0.$$

It follows that, for some constant $\lambda :=$ the sum of first two terms, we have

$$\begin{cases} \rho^2 R'' + 2\rho R' - \lambda R = 0 \\ \dfrac{\Theta''(\theta)}{\sin^2\varphi\Theta(\theta)} + \dfrac{\Phi''(\varphi)}{\Phi(\varphi)} + \dfrac{\cot\varphi\Phi'(\varphi)}{\Phi(\varphi)} = -\lambda. \end{cases}$$

We can further define $-\alpha := \Theta''(\theta)/\Theta(\theta)$ and split the second equation into

$$\begin{cases} \Theta''(\theta) + \alpha\Theta(\theta) = 0 \\ \Phi''(\varphi) + \dfrac{\cos\varphi}{\sin\varphi}\Phi'(\varphi) + \left(\lambda - \dfrac{\alpha}{\sin^2\varphi}\right)\Phi(\varphi) = 0. \end{cases}$$

For a special case, we set $\alpha = 0$ and so $\Theta(\theta) = 1$. Then $u = u(\rho, \varphi)$ and

$$\begin{cases} \rho^2 R'' + 2\rho R' - \lambda R = \\ \Phi''(\varphi) + \dfrac{\cos\varphi}{\sin\varphi}\Phi'(\varphi) + \lambda\Phi(\varphi) = 0. \end{cases}$$

The first equation is given by

$$R_1(\rho) = \rho^{p_1} \qquad R_2(p) = \rho^{p_2}$$

where p_1, p_2 are roots to $p^2 + p - \lambda = 0$. For the second equation, let $t := \cos\varphi so -1 \leqslant t \leqslant 1$. Let $p(t) = \Phi(\varphi)$ where $\varphi = \cos^{-1} t$. Then

$$\Phi'(\varphi) = p' \cdot \frac{\mathrm{d}t}{\mathrm{d}\varphi} = -\sqrt{1-t^2}p',$$

and

$$\Phi''(\varphi) = p''(\sin^2\varphi) - p'\cos\varphi = (1-t^2)p'' - tp'.$$

The second equation therefore becomes

$$(1-t^2)p'' - tp' - tp' + \lambda p = (1-t^2)p'' - 2tp' + \lambda p,$$

the Legendre differential equation. Therefore, when $\lambda = n(n+1)$, the solution will be given by the Legendre polynomials!

$$\Phi_n(\varphi) = p_n(\cos\varphi).$$

For $\lambda = n(n+1)$, $p^2 + p - \lambda = 0$ has n or $-(n+1)$. We choose $p = n$ to get a bounded solution. If $u(\rho,\varphi) = R(\rho)\Phi(\varphi)$ is a bounded solution of the Laplace equation inside the sphere, then there exists $n \in \mathbb{N}$ such that

$$u(\rho,\varphi) = Cp^n p_n(\cos\varphi)$$

for some constant C.

More generally, we have solutions of $\Delta u = 0$ of the form

$$p^n \cos(m\theta) p_n^m(\cos\varphi) \qquad m \leqslant n$$

where

$$p_n^m(t) = (1-t^2)^{m/2}\frac{\mathrm{d}^m}{\mathrm{d}t^m}p_n(t)$$

is called the *associated Legendre polynomials of degree n and order m*. These polynomials solve

$$(1-t^2)p'' - 2tp' + \left(\lambda - \frac{m^2}{1-t^2}\right)p = 0.$$

The functions $Y_{m,n}(\theta,\varphi) = p_n^m(\cos\varphi)\cos(m\theta)$ are called the *spherical harmonics*.

Beginning of Oct.18, 2021

4.6 Positive Definite Operators

Definition 4.6.1

Let H be Hilbert over $\mathbb{R}$. A linear operator $A : H \to H$ is **strictly positive definite** such that there exists $\beta > 0$ satisfying $\langle Au, u\rangle \geqslant \beta\|u\|^2$ for all $u \in H$.

Theorem 4.6.2: Inverse of a PD Operator

Let $A : H \to H$ be a bounded linear operator which is strictly PD. Ten for any $f \in H$, there exists a *unique* element $u := A^{-1}f$ in H such that $Au = f$. Moreover, the inverse A^{-1} satisfies

$$\|A^{-1}\| \leqslant 1/\beta$$

where β is the corresponding bound in the definition of PD-ness.

Proof. We need to show that A is bijective.
For injectivity, it suffices to show that $\ker A = \{0\}$. Indeed,

$$\beta\|u\|^2 \leqslant \langle Au, u\rangle \leqslant \|Au\|\|u\|$$

so if $\|u\| > 0$, $\beta\|u\| \leqslant \|Au\|$. Therefore if $Au = 0$, $\|u\| = 0$, i.e., $\ker A = \{0\}$.
Now we show that $R(A)$ is closed (surjectivity later). Let $\{v_n\} \subset R(A)$ be such that $v_n \to v$ for some $v \in H$. We can write $\{v_n\}$ as $\{Au_n\}$ for some $\{u_n\} \subset H$. Then, using $\beta\|u\| \leqslant \|Au\|$ from the previous part for nonzero u,

$$\|u_m - u_n\| \leqslant \frac{1}{\beta}\|Au_m - Au_n\| = \frac{1}{\beta}\|v_m - v_n\|.$$

Hence,

$$\limsup_{m,n\to\infty} \|u_m - u_n\| \leqslant \limsup_{m,n\to\infty} \frac{1}{\beta}\|v_m - v_n\| = 0.$$

Hence $\{u_n\}$ is Cauchy. By completeness, $u_n \to u$ for some $u \in H$. Since A is bounded, it preserves limits, so $Au = v$.
Finally, we show that $R(A)$ is all of H. Suppose not, then there exists a nonzero $w \in R(A)^\perp$. Then

$$\beta\|w\|^2 \leqslant \langle Aw, w\rangle = 0$$

as $w \in R(A)^\perp$ but $Aw \in R(A)$. Thus $w = 0$, contradiction. This finishes the proof showing A is bijective. Hence $Au = f$ has a unique solution $u := A^{-1}f$ and A^{-1} is continuous by the Open Mapping Theorem.
Finally, since

$$A^{-1}f = \|u\| \leqslant \frac{1}{\beta}\|Au\| = \frac{1}{\beta}\|AA^{-1}f\| = \frac{1}{\beta}\|f\|,$$

we have $\|A^{-1}\| \leqslant 1/\beta$. □

Theorem 4.6.3: Lax-Milgram

Let H be Hilbert over $\mathbb{R}$. Let $B : H \times H \to \mathbb{R}$ be a continuous bilinear functional, i.e., bilinear and

$$|B[u, v]| \leqslant C\|u\|\|v\|$$

for some $C > 0$. Further, assume that B is strictly PD, i.e., for some $\beta > 0$ we have

$$B[u, u] \geqslant \beta\|u\|^2 \text{ for all } u \in H.$$

Then for all $f \in H$, there exists a *unique* $u \in H$ such that

$$B[u, v] = \langle f, v \rangle \qquad \text{for all } v \in H.$$

Moreover, $\|u\| \leqslant 1/\beta \cdot \|f\|$.

Proof. For each *fixed* $u \in H$, the map $v \mapsto B[u, v]$ is a continuous linear functional on H. By the Riesz Representation Theroem, there exists a unique element, which we call Au, such that

$$B[u, v] = \langle Au, v \rangle$$

for all $v \in H$. We claim that $u \mapsto Au$ is bounded, linear, and PD. Linearity follows from bilinearity of B. For boundedness, notice that

$$\|Au\| = \sup_{\|v\|=1} \langle Au, v \rangle = \sup_{\|v\|=1} |B[u, v]| \leqslant C\|u\|.$$

To see that A is strictly PD:

$$\langle Au, u \rangle = B[u, u] \geqslant \beta \|u\|^2.$$

Therefore, by the previous theorem, $Au = f$ has a unique solution $u := A^{-1} f$ satisfying $\|u\| \leqslant \|f\|/\beta$. This completes the proof. □

Chapter 5

Compact Operators on a Hilbert Space

Recall that for a linear operator $A : \mathbb{R}^n \to \mathbb{R}^n$,

(1) A is injective if and only if A is surjective;

(2) $\dim \ker(A) = \dim(R(A)^\perp)$;

(3) If A is symmetric, its eigenvalues are real, and the space $\mathbb{R}^n$ admits an orthonormal basis consisting of the eigenvectors.

The goal is to generalize these results to $\lambda : H \to H$ where H is an infinite-dimensional Hilbert space.

We will show that

(1) remains valid with the assumption that $\lambda = I - K$ where K is compact; and

(2) can be extended to any compact, self-adjoint operator $\lambda : H \to H$.

Beginning of Oct.20, 2021

5.1 Fredholm Theory

Unless otherwise specified, let H be a Hilbert space.

Recall that $K : H \to H$ is compact if, for all bounded sequence of points $\{u_n\} \subset H$, there exists a subsequence $\{u_{n_j}\}$ such that Ku_{n_j} converges in H.

Theorem 5.1.1: (Fredholm)

Let H be over $\mathbb{R}$. Let $K : H \to H$ be a *compact* linear operator. Then

(1) $\ker(I - K)$ is finite dimensional.

(2) Range$(I - K)$ is closed.

(3) Range$(I - K) = \ker(I - K^*)^\perp$ where K^* is the adjoint operator.

(4) $\ker(I-K)=\{0\}$ if and only if $\text{Range}(I-K)=H$.

(5) $\ker(I-K)$ and $\ker(I-K^*)$ have the same dimension.

Proof.

(1) Suppose for contradiction that $\ker(I-K)$ is infinite dimensional. Then there exists an orthonormal sequence $\{e_n\}_{n\geqslant 1}\in\ker(I-K)$ such that $Ke_n=e_n$ for all n (since $(I-K)e_n=0$). In particular, for $m\neq n$, $\|e_m-e_n\|=\sqrt{2}$, so $\|Ke_m-Ke_n\|=\|e_m-e_n\|=\sqrt{2}$. This contradicts the compactness of K, as $\{Ke_n\}$ admits no convergent subsequence.

(2) **Claim 1.** There exists a constant $\beta>0$ such that

$$\|(I-K)u\|=\|u-Ku\|\geqslant\beta\|u\| \qquad \text{for all } u\in\ker(I-K)^\perp.$$

Proof. Suppose the claim does not hold, then there exists a sequence $\{u_n\}\subset\ker(I-K)^\perp$ such that $\|u_n\|=1$ but

$$\|u_n-Ku_n\|<\frac{\|u\|}{n}.$$

Since $\{u_n\}$ is bounded, there exists a subsequence $\{u_{n_j}\}$ converging weakly to some u, i.e., $u_{n_j}\rightharpoonup u$. Since K is compact, this implies $Ku_{n_j}\to Ku$ *strongly*. Thus

$$\begin{aligned}\|u_{n_j}-Ku\|&\leqslant\|u_{n_j}-Ku_{n_j}\|+\|Ku_{n_j}-Ku\|\\&<\frac{1}{n_j}+\|Ku_{n_j}-Ku\|\to 0.\end{aligned}$$

Therefore u_{n_j} converges strongly to Ku. Recall that u_{n_j} also has a weak limit u. By uniqueness we have $u=Ku$, i.e., $u\in\ker(I-K)$ so $\|u\|=0$. However, $u=Ku$ is also the strong limit of $\|u_{n_j}\|$, a sequence of elements with norms 1. Contradiction.

Claim 2. Range of $I-K$ is closed.

Proof. Let $v_n\in\text{Range}(I-K)$ and let $v_n\to v$ for some $v\in H$. We need to show that $v=u-Ku$ for some u.

By assumption, for each n, there exists $u_n\in H$ such that $v_n=u_n-Ku_n$.

Let $\tilde{u}_n$ be the perpendicular projection of u_n onto $\ker(I-K)$ and let $z_n:=u_n-\tilde{u}_n$ (so it's in $\ker(I-K)^\perp$). Notice that

$$v_n=(I-K)u_n=(I-K)(\tilde{u}_n+z_n)=(I-K)z_n=z_n-Kz_n.$$

By claim 1, there exists $\beta>0$ such that

$$\|z_n-Kz_n\|\geqslant\beta\|z_n\|.$$

Therefore,

$$\|v_m-v_n\|\geqslant\beta\|z_m-z_n\|.$$

Since $\{v_n\}$ is Cauchy, this implies $\{z_n\}$ is Cauchy. Therefore there exists $u\in H$ such that $z_n\to u$. Thus,

$$u-Ku=\lim_{n\to\infty}(z_n-Kz_n)=\lim_{n\to\infty}v_n=v.$$

This shows $v\in\text{Range}(I-K)$, as claimed.

(3) Note that by (2), Range$(I-K)$ and $\ker(I-K^*)^\perp$ are both closed, so it suffices to show that Range$(I-K)^\perp = \ker(I-K^*)$.

Indeed,

$$\begin{aligned} x \in \ker(I-K^*) &\iff x - K^*x = 0 \\ &\iff \langle y, (I-K^*)x\rangle = 0 \text{ for all } y \in H \\ &\iff ((I-K)y, x) = 0 \text{ for all } y \in H \\ &\iff x \in \text{Range}(I-K)^\perp. \end{aligned}$$

(4) For $\Rightarrow$, assume $\ker(I-K) = \{0\}$, i.e., $I-K$ is injective. Suppose $I-K$ is not surjective, so $H_1 := (I-K)(H)$ is a strict subspace of H. By (2), H_1 is closed in H. Since $I-K$ is injective, if we let $H_2 := (I-K)H_1$, we have $H_2 \subsetneq H_1$ again. Inductively we define $H_n := (I-K)^n H$ and obtain a strictly nested sequence of closed subspaces $H \supset H_1 \supset H_2 \supset \ldots$.

Thus, for each $n \geqslant 1$, we can choose $e_n \in H_n \cap H_{n+1}^\perp$ with $\|e_n\| = 1$. Note that for $m < n$,

$$\begin{aligned} Ke_m - Ke_n &= Ke_m - e_m + e_m - e_n + e_n - Ke_n \\ &= e_m + z_m \end{aligned}$$

where

$$z_n := \underbrace{KKe_m - e_m}_{H_{m+1}} + \underbrace{e_n}_{\in H_n} - \underbrace{Ke_n - e_n}_{\in H_{n+1} \subset H_{m+1}}$$

so $z_m \in H_{m+1}$. Since $e_m \in H_{m+1}^\perp$, by Pythagorean theorem,

$$\|Ke_m - Ke_n\|^2 = \|e_m\|^2 + \|z_m\|^2 \implies \|Ke_m - Ke_n\| \geqslant \|e_m\|.$$

Therefore the sequence $\{Ke_n\}$ does not admit any convergent subsequence, contradiction. Hence $I-K$ must be surjective.

Conversely for $\Leftarrow$, suppose Range$(I-K) = H$. Then by (3) $\ker(I-K^*) = R(I-K)^\perp = H^\perp = \{0\}$. Since K compact implies K^* compact, by the previous direction, Range$(I-K^*) = H$. Now,

$$\ker(I-K) = \text{Range}(I-K^*)^\perp = \{0\}.$$

Beginning of Oct.22, 2021

(5) We first show that $\dim(\ker(I-K)) \geqslant \dim(\text{Range}(I-K)^\perp)$.

Suppose this is not true; that is, $\dim(\ker(I-K)) < \dim(\text{Range}(I-K)^\perp)$. We know the LHS is finite dimensional, so there exists a linear map $A : \ker(I-K) \to \text{Range}(I-K)^\perp$ which is injective but not surjective. We can also extend A to a linear map defined on all of H, i.e., $A : H \to \text{Range}(I-K)^\perp$ by setting $Au := 0$ for all $u \in \ker(I-K)^\perp$. Since Range(A) is finite dimensional, A is indeed compact by Riesz's lemma. Therefore $K + A$ is compact.

Claim: $\ker(I-(K+A)) = \{0\}$. To see this, for any $u \in H$, we can write $u = u_1 + u_2$ where $u_1 \in \ker(I-K)$ and $u_2 \in \ker(I-K)^\perp$. Then

$$(I-(K+A))u = (I-(K+A))(u_1+u_2) = \underbrace{-Au_1}_{\in \text{Range}(I-K)^\perp} + \underbrace{(I-K)u_2}_{\in \text{Range}(I-K)},$$

so $u \in \ker(I-(K+A))$ if and only if both Au_1 and $(I-K)u_2 = 0$, that is, $u_2 \in \ker(I-K)$, so u_2 is in both $\ker(I-K)$ and $\ker(I-K)^\perp$, i.e., $u_2 = 0$. Also, by assumpgion $u_1 \in \ker(I-K)$, but then $Au_1 = 0$ implies $u_1 = 0$ since A is injective on $\ker(I-K)$. Hence $\ker(I-(K+A)) = \{0\}$.

Applying (4) to the compact operator $K+A$, we have Range$(I-(K+A)) = H$.

On the other hand, by construction A is not surjective, so there exists $v \in \text{Range}(I-K)^\perp$ which is not in Range(A). In particular,

$$(I-(K+A))u = u - Ku - Au = v$$

has no solution: decomposing $u = u_1 + u_2$ as above, there is no Au_1 to account for the result v. Contradiction. Hence $\dim(\ker(I-K)) \geqslant \dim(\text{Range}(I-K)^\perp)$.

Also, by (3), $\dim(\ker(I-K)) \geqslant \dim(\text{Range}(I-K)^\perp) = \dim(\ker(I-K^*))$. Changing the roles of K and K^*, we obtain the other direction of the inequality. This completes the proof.

□

Fredholm Alternative

Consider the linear equation $u - Ku = f$ where K is a compact operator on H.

Notice that $u - Ku$ can be written as $(I-K)u$, so this equation can be solved exactly when $f \in \text{Range}(I-K)$. Recall that $\ker(I-K)$ is finite dimensional.

Case 1. $\ker(I-K) = \{0\}$, so that $I-K$ is injective and moreover surjective by (4) above. Therefore, for each f, there will be a u satisfying the equation.

Case 2. $\ker(I-K)$ is nontrivial but finite dimensional. This means that the homogeneous equation $(I-K)u = 0$ has nontrivial solutions. In this case, by (3), Range$(I-K)$ is orthogonal to $\ker(I-K^*)$. That is, the equation has a solution if and only if $f \in \ker(I-K^*)^\perp$, i.e., $\langle f, u\rangle = 0$ for all $u \in \ker(I-K^*)$, i.e., $\langle f, u\rangle = 0$ for all $u \in H$ such that $u - K^*u = 0$.

5.2 Spectrum of a Compact Operator

Let $\Lambda : H \to H$ be a bounded linear operator, not necessarily compact.

Definition 5.2.1: Resolvent, spectrum, & point spectrum

The **resolvent set** of Λ is $\rho(\Lambda) : \{\lambda \in \mathbb{R} : \lambda I - \Lambda \text{ is bijective}\}$. *In this case, by the OMT, the inverse operator* $(\lambda I - \Lambda)^{-1}$ *is continuous.*

The **spectrum** of Λ is defined to be $\sigma(\Lambda) := \mathbb{R} - \rho(\Lambda)$ [here $-$ denotes set-theoretic minus], i.e., if $x \in \rho(\Lambda)$ then $xI - \Lambda$ either fails to be injective or surjective (or both).
The **point spectrum** of Λ is defined to be

$$\sigma_p(\Lambda) := \{\lambda \in \mathbb{R} : \lambda I - \Lambda \text{ is not injective}\},$$

i.e., if $\lambda \in \sigma_p(\Lambda)$ then there exist nonzero $w \in H$ such that $\Lambda w = \lambda w$. In this case λ is called an **eigenvalue** of Λ and w is the associated **eigenvector**.
The **essential spectrum** of Λ is

$$\sigma_e(\Lambda) := \sigma(\Lambda) - \sigma_p(\Lambda),$$

i.e., the set of $\lambda \in \mathbb{R}$ such that $\lambda I - \Lambda$ is injective but not surjective.

Example 5.2.2. Consider $H := \ell^2$ and the right shift operator$\Lambda : \ell^2 \to \ell^2$ defined by

$$(x_1, x_2, x_3, \dots) \mapsto (0, x_1, x_2, \dots).$$

As shown before, Λ is bounded with norm 1. Clearly Λ is injective but not surjective. By our definition, $0 \cdot I - \Lambda = -\Lambda$ is injective but not onto, so $0 \in \sigma(\Lambda)$, $0 \in \sigma_e(\Lambda)$, but $0 \notin \sigma_p(\Lambda)$, i.e., 0 is not an eigenvector.

Remark. If $\Lambda \neq 0$, $\lambda I - K$ can be re-written as $\lambda(I - \lambda^{-1}K)$, which connects to Fredholm theory.

Theorem 5.2.3: Spectrum of a compact operator

Let H be infinite dimensional. Let $K : H \to H$ be compact. Then:

(1) $0 \in \sigma(K)$,

(2) $\sigma(K) = \sigma_p(K) \cup \{0\}$, and

(3) Either $\sigma_p(K)$ is finite or countable, with $\lim_{k\to\infty} \lambda_k = 0$ being the only accumulation point.

Beginning of Oct.25, 2021

Proof.

(1) Suppose for contradiction that $0 \notin \sigma(K)$. Then K has a continuous inverse $K^{-1} : H \to H$. Consider $I = K \circ K^{-1}$. Since K^{-1} is continuous and K compact, this tells us I is compact. But this is clearly false, as in an infinite dimensional space, the closed unit ball is not compact.

(2) Assume $\lambda \in \sigma(K)$ and $\lambda \neq 0$. We want to show that $\lambda \in \sigma_p(K)$. Suppose not, i.e., if $\ker(\lambda I - K) = \{0\}$, the Fredholm theory implies that Range$(\lambda I - K) = H$. By the OMT, $\lambda I - K$ is open and therefore admits a continuous inverse, contradicting the assumption that $\lambda \in \sigma(K)$.

(3) **Claim**. For each $\mu > 0$, the space generated by the eigenvectors corresponding to all λ_k with $|\lambda_k| > \mu$ needs

to be finite dimensional.

Proof. Suppose for contradiction that there are infinitely many distinct eigenvalues λ_k with $|\lambda_k| > \mu$ for some μ. Then we can take an orthonormal sequence $\{e_k\}$ corresponding to λ_k of distinct eigenvalues, i.e., $Ke_k = \lambda_k e_k$. Since K is compact, there exists a subsequence e_{k_j} such that Ke_{k_j} converges. However, for $j \neq \ell$,

$$\|\lambda_{k_j} e_{k_j} - \lambda_{k_\ell} e_{k_\ell}\| = |\lambda_{k_j}|^2 + |\lambda_{k_\ell}|^2 > 2\mu^2 > 0.$$

Contradiction. Hence the claim is proven.

The claim shows that the eigenvalues of K is at most countable with $\lambda_k \to 0$ ass $k \to \infty$.

□

5.3 Self-Adjoint Operators

Let $\lambda : H \to H$ be a bounded linear operator on H.

Definition 5.3.1: Symmetric operator

Λ is called **symmetric** if $\langle \Lambda x, y\rangle = \langle x, \Lambda y\rangle$ for all $x, y \in H$.

Example 5.3.2. Let A be a symmetric matrix. Then A determines a symmetric linear operator $\mathbb{R}^n \to \mathbb{R}^n$ defined by $a \mapsto Ax$. Also, it gives the quadratic form

$$x \mapsto \langle x, Ax\rangle = \sum_{i,j=1}^{n} a_{i,j} x_i x_j.$$

Then

$$M := \max_{|x|=1} \langle x, Ax\rangle \qquad m = \min_{|x|=1} \langle x, Ax\rangle$$

gives the largest and smallest eigenvalues of A.

Lemma: Bounds on the spectrum of a symmetric operator

Let $\Lambda : H \to H$ be a bounded linear operator on H. Define $M := \sup \langle \Lambda u, u\rangle$ and $m := \inf \langle \Lambda u, u\rangle$, both taken over all $u \in H$ with $\|u\| = 1$. Then

(1) $\sigma(\Lambda) \subset [m, M]$,

(2) $m, M \in \sigma(\Lambda)$, and

(3) $\|\Lambda\| = \max\{-m, M\}$.

Proof.

(1) It suffices to show that the resolvent contains $(-\infty, m)$ and (M, ∞). To this end, let $\eta > M$; we want to show $\eta I - \Lambda$ is bijective. Then

$$\langle (\eta I - \Lambda)u, u\rangle = \eta \langle u, u\rangle - \langle \Lambda u, u\rangle \geqslant (\eta - M)\|u^2\| \text{ for all } u \in H.$$

That is, $\eta - M$ is a positive operator. By Lax-Milgram $\eta I - \Lambda$ is bijective. By the OMT, it has a continuous inverse, so $(M, \infty) \subset \rho(\Lambda)$, and a similar argument shows $(-\infty, m) \subset \rho(\Lambda)$.

(2) **Claim.** $M \in \sigma(\Lambda)$.

Proof. We choose a sequence $\{u_n\}$ such that $\langle \Lambda u_n, u_n \rangle \to M$, all with $\|u_n\| = 1$. Then (using result from (3))

$$\begin{aligned} \|\Lambda u_n - M u_n\|^2 &= \|\Lambda u_n\|^2 - 2M \langle \Lambda u_n, u_n \rangle + M^2 \|u_n\|^2 \\ &\leqslant 2M^2 - 2M \langle \Lambda u_n, u_n \rangle \to 0. \end{aligned}$$

Therefore $\Lambda - MI$ cannot have a bounded inverse. Similarly we can show $m \in \sigma(A)$.

(3) We may assume $|m| \leqslant M$.

Claim. $\|\Lambda u\| \leqslant M \|u\|$ for all $u \in H$, i.e., $\|\Lambda\| \leqslant M$.

Proof. The polarization identity gives

$$\begin{aligned} 4 \langle \Lambda u, v \rangle &= \langle \Lambda(u+v), u+v \rangle - \langle \Lambda(u-v), u-v \rangle \\ &\leqslant M \|u+v\|^2 - m \|u-v\|^2 \leqslant M(\|u+v\|^2 + \|u-v\|^2) \\ &= 2M(\|u\|^2 + \|v\|^2), \end{aligned}$$

so

$$\langle \Lambda u, v \rangle \leqslant M(\|u\|^2 + \|v\|^2).$$

Let $v := \dfrac{\|u\|}{\|\Lambda u\|} \Lambda u$ (assuming $u \neq 0$). Then

$$\frac{2\|u\|}{\|\Lambda u\|} \langle \Lambda u, \Lambda u \rangle \leqslant M(\|u\|^2 + \|u\|^2),$$

so

$$\|u\| \|Au\| \leqslant M \|u\|^2,$$

i.e.,

$$\|Au\| \leqslant M \|u\| \text{ for all } u \in H.$$

On the other hand,

$$M \geqslant \|\Lambda\| \geqslant \sup_{\|u\|=1} \langle \Lambda u, u \rangle = M$$

by Cauchy-Schwarz, so $\|\Lambda\| = M$.

□

Beginning of Oct.27, 2021

Theorem 5.3.3: Hilbert-Schmidt

Let H be a separable Hilbert space, and let $K : H \to H$ be a compact symmetric linear operator. Then there exists a countable orthonormal basis of H consisting of eigenvectors of K.

Proof. Assume H is finite dimensional. Let $\lambda_0 = 0$ and let $\{\lambda_1, \lambda_2, \dots\}$ be the set of eigenvalues of K. Define $H_0 := \ker K$ and $H_n := \ker(K - \lambda_n I)$. We know

$$0 \leqslant \dim H_0 \leqslant \infty,$$

and by Fredholm theory, for all $k \geqslant 1$,

$$1 \leqslant \dim H_k < \infty.$$

Claim 1. For $m \geqslant n$ and $m, n \geqslant 1$, H_m and H_n are orthogonal.
To see this, let $u \in H_m$ and $v \in H_n$. That is, $u \in \ker(K - \lambda_m I)$ so $Ku = \lambda_m u$ and similarly $Kv = \lambda_n v$. Since K is symmetric,

$$\lambda_m \langle u, v \rangle = \langle \lambda_m u, v \rangle = \langle Ku, v \rangle = \langle u, Kv \rangle = \langle u, \lambda_n v \rangle = \lambda_n \langle u, v \rangle .$$

Since $\lambda_m \neq \lambda_n$ we must have $\langle u, v \rangle = 0$. Thus $H_m \perp H_n$.

Now we define the linear combination of these spaces

$$\tilde{H} := \left\{ \sum_{k=1}^{N} \alpha_k u_k : N \geqslant 1, \alpha_k \in \mathbb{R}, u_k \in H_k \right\}.$$

Claim 2. $\tilde{H}^\perp \subset \ker(K) = H_0$.
To see this, notice that the image $K\tilde{H} \subset \tilde{H}$ as each $K(\alpha_k u_k)$ is simply $(\lambda_k \alpha_k) u_k$, also a scalar product of u_k. Furthermore, if $u \in \tilde{H}^\perp$ and $v \in \tilde{H}$, then $Kv \in \tilde{H}$, so

$$\langle Ku, v \rangle = \langle u, Kv \rangle = 0.$$

This shows $K(\tilde{H}^\perp) \subset \tilde{H}^\perp$.
Now let $\tilde{K} := K\big|_{\tilde{H}^\perp}$. Then $\tilde{K} : \tilde{H}^\perp \to \tilde{H}^\perp$ is a compact symmetric operator. By the previous lemma,

$$\|\tilde{K}\| = \sup_{\tilde{u} \in \tilde{H}, \|\tilde{u}\| = 1} = |\langle \tilde{K}\tilde{u}, \tilde{u} \rangle| := M.$$

We claim that $M = 0$. Suppose not, then M or $-M \in \sigma(\tilde{K})$. Call it λ. Since $\tilde{K}$ is compact, nonzero eigenvalues is in the point spectrum, so there exists $w \in \tilde{H}^\perp$ such that $Kw = \tilde{K}w = \lambda w$, so λ is also an eigenvalue of K. This is impossible, as all eigenvectors of K are supposed to be in $\tilde{H}$ by construction. Hence $M = 0$. Hence $\tilde{H}^\perp \subset \ker(K) = H_0$.

Now we notice that $H_0^\perp \cap \tilde{H}^\perp \subset H_0^\perp \cap H_0 = \{0\}$. Hence,

$$\text{span}(H_0 \cup \tilde{H})$$

is dense in H. We know each H_k admits a finite orthonormal basis. The union over all k still admits a countable basis. Since H is separable, the closed subspace $H_0 = \ker(K)$ also admits a countable orthonormal basis. Hence the union over everything is still countable. □

Remark. Let $\{w_1, w_2, \dots\}$ be an orthonormal basis of a real Hilbert space H consisting of eigenvectors of a linear, compact, symmetric operator K, where $\lambda_1, \lambda_2, \dots$ be the corresponding eigenvalues. For $\in H$ consider the equation $(I - K)u = u - Ku = f$. It follows that if 1 is not in $\sigma(K)$ then the equation has a unique solution.

Furthermore, we can let $u = \sum_{k=1}^{\infty} c_k w_k$ and $f = \sum_{k=1}^{\infty} b_k w_k$ for some coefficients $\{c_k\}, \{b_k\}$. Then

$$c_k - \lambda_k c_k = \langle u, w_k \rangle - \langle u, K w_k \rangle = \langle u, w_k \rangle - \langle K u, w_k \rangle = \langle u - K u, w_k \rangle = \langle f, w_k \rangle = b_k.$$

Then $c_k = b_k/(1-\lambda_k)$, and this is well-defined as long as $\lambda_k \neq 1$, which is guaranteed as $1 \notin \sigma(K)$. Therefore,

$$u = \sum_{k=1}^{\infty} \frac{b_k}{1-\lambda_k} w_k = \sum_{k=1}^{\infty} \frac{\langle f, w_k \rangle}{1-\lambda a_k} w_k.$$

With extra assumptions, we are able to construct an explicit solution!

5.4 Application (The Dirichlet Problem for Δ)

Let $\Omega \subset \mathbb{R}^n$ be open, bounded, connected, with $\partial\Omega$ smooth. Let g be a smooth function given along $\partial\Omega$. We try to find a function u satisfying

$$\begin{cases} \Delta u = 0 & \text{in } \Omega \\ u = g & \text{on } \partial\Omega \end{cases} \qquad \text{where } \Delta u := \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2}.$$

We can first consider a smooth extensions of $\overline{g}$ to $\overline{\Omega}$. Then $\tilde{u} = u - \overline{g}$ will satisfy

$$\begin{cases} \Delta \tilde{u} = -\Delta \tilde{g} & \text{in } \Omega \\ \tilde{u} = 0 & \text{on } \partial\Omega. \end{cases}$$

We are led to study the following *Dirichlet problem*

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{*}$$

Functional analytic techniques will let us prove the existence of *weak solutions* to (*), and we will be able to show that any function $f \in L^2(\Omega)$ will produce a unique (weak) solution $u \in H_0^1(\Omega)$ of (*) (that is, the Paplacian has an inverse in suitable spaces). We will also show, assuming some extra theorems, that the operator $f \mapsto u$ from $L^2(\Omega)$ to $L^2(\Omega)$ is compact and therefore we will obtain the spectrum of the Laplacian. That is, we will see the existence of a discrete set of λ solving the eigenvalue problem

$$\begin{cases} \Delta u + \lambda u = 0 & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

Beginning of Oct.29, 2021

Weak Solutions

Define $C_0^\infty(\Omega)$ to be $\{f \in C^\infty(\Omega) : \operatorname{supp} f \subset \Omega \text{ compact}\}$.

We define two different inner products:

$$\langle u, v \rangle_0 := \int_\Omega u \cdot v \, \mathrm{d}x$$

$$\langle u, v\rangle_1 := \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x$$

We can check that $C_0^\infty(\Omega)$ is an inner product space under $\langle\cdot,\cdot\rangle_0$ or under $\langle\cdot,\cdot\rangle_1$. Define $\|u\|_0^2 := \langle u, u\rangle_0$ and $\|u\|_1^2 := \langle u, u\rangle_1$.

Lemma: Poincaré's Inequality

For all $u \in C_0^\infty(\Omega)$, $\|u\|_0 \leqslant d\|u\|_1$ where d is the diameter of Ω.

Proof. Since $u\big|_{\partial\Omega} = 0$, at any point $x \in \Omega$, we have by FTC

$$u(x) = \int_{x^b}^{x} \partial_{x_1} u \, \mathrm{d}x_1$$

where $x^b \in \partial\Omega$ and all but the first component of x^b agrees with that of x. Then, by Hölder's inequality,

$$u^2(x) = \left(\int_{x^b}^{x} 1 \cdot \partial_{x_1} u \, \mathrm{d}\tilde{x}_1\right)^2 \leqslant d \cdot \left(\int_{x^b}^{x} |\partial_{x_1} u|^2 \, \mathrm{d}\tilde{x}_1\right).$$

Integrating over Ω gives

$$\int_\Omega u^2 \, \mathrm{d}x \leqslant d \int_\Omega \left(\int_{x^b}^{x_1} |\partial_{x_1} u|^2 \, \mathrm{d}\tilde{x}\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n \leqslant d^2 \int_\Omega |\nabla u|^2 \, \mathrm{d}x.$$

□

Definition 5.4.1: $H_0^1(\Omega)$

We define $H_0^1(\Omega)$ to be the completion of $C_0^\infty(\Omega)$ with respect to $\|\cdot\|_1$. This is a Hilbert space.

Lemma

Each $u \in H_0^1(\Omega)$ belongs to $L^2(\Omega)$ and has partial derivatives $u_j \in L^2(\Omega)$. Furthermore,

$$\langle v, u_j\rangle_0 = -\langle \partial v/\partial x_j, u\rangle_0 \qquad \text{for all } v \in C_0^\infty(\Omega).$$

Moreover,

$$\langle u, v\rangle_1 = \int_\Omega \sum_{j=1}^{n} u_j v_j \, \mathrm{d}x \qquad \text{for all } v, u \in H_0^1(\Omega).$$

Proof. Let $\{u_n\}$ be a sequence in $C_0^\infty(\Omega)$ converging to u in $\|\cdot\|_1$. Then, $\partial u_n/\partial x_j$ converges in $\|\cdot\|_0$ to a limit which we call u_j. By Poincaré's inequality, $\partial u_n/\partial x_j$ converges in L^2 as well, which we identify with the limit $u \in H_0^1$. For a fixed n,

$$\begin{aligned}\langle v, \partial u_n/\partial x_j\rangle_0 &= \int_\Omega v \frac{\partial u_n}{\partial x_j} \, \mathrm{d}x \\ [\text{IBP}] &= -\int_\Omega \frac{\partial v}{\partial x_j} u_n \, \mathrm{d}x = \langle \partial v/\partial x_j, u_n\rangle_0 .\end{aligned}$$

Taking $n \to \infty$, we see that

$$\langle v, u_j\rangle_0 = -\langle \partial v/\partial x_j, u\rangle_0 .$$

□

Remark. u_j is called a **weak partial derivative** of u. It is in some sense a partial derivative, but sometimes it is not as nice as a usual partial derivative is. It just needs to satisfy the equalities in the previous lemma.

Definition 5.4.2: Weak solution

We call $u \in H_0^1(\Omega)$ a **weak solution** to

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{*}$$

if

$$\langle u, v\rangle_1 = \int_\Omega fv \, \mathrm{d}x \qquad \text{for all } v \in H_0^1(\Omega).$$

Theorem 5.4.3

Given $f \in L^2(\Omega)$, there exists a unique weak solution $u \in H_0^1(\Omega)$ of (*). The weak solution satisfies

$$\|u\|_1^2 = \int_\Omega |\nabla u|^2 \, \mathrm{d}x \leqslant d^2 \int_\Omega f^2 \, \mathrm{d}x.$$

Proof. For any $f \in L^2(\Omega)$, define a linear function

$$\varphi : v \mapsto \int_\Omega fv \, \mathrm{d}x \qquad \text{on } H_0^1(\Omega).$$

In other words, $\varphi(v) := \langle f, v\rangle_0$. Then Hölder's inequality and Poincaré's inequality give

$$|\varphi(v)| \leqslant \|f\|_0\|v\|_0 \leqslant d\|f\|_0\|v\|_1.$$

That is, φ is a bounded linear functional on $H_0^1(\Omega)$. Therefore, by the Riesz RT, the functional $\varphi(v) = \langle f, v\rangle_0$ can be represented as an inner product on $H_0^1(\Omega)$. That is, there exists a unique element $u \in H_0^1(\Omega)$ such that

$$\langle f, v\rangle_0 = \langle u, v\rangle_1 \qquad \text{for all } v \in H_0^1(\Omega).$$

Now we choose $v = u$ so that (again by Hölder and Poincaré)

$$\|u\|_1^2 = \sum_{j=1}^n \int_\Omega |u_j|^2 \, \mathrm{d}x = \int_\Omega fu \, \mathrm{d}x \leqslant \|f\|_0\|u\|_0 \leqslant d\|f\|_0\|u\|_1.$$

This shows

$$\|u\|_1 \leqslant d\|f\|_0.$$

□

Now we have shown that under suitable situations, we have obtained a (linear) map $\Phi : L^2(\Omega) \to H_0^1(\Omega)$ defined by $f \mapsto u$, where u is informally $(-\Delta)^{-1}f$. We write $u = \Phi(f)$.

Notice that $H_0^1(\Omega) \subset L^2(\Omega)$, so there exists an embedding map i: $H_0^1(\Omega) \overset{i}{\hookrightarrow} L^2(\Omega)$ so we end up getting $\Phi : L^2(\Omega) \to H_0^1(\Omega) \hookrightarrow L^2(\Omega)$.

Proposition 5.4.4

$\Phi : L^2(\Omega) \to L^2(\Omega)$ defined above is a compact, symmetric, positive operator.

Beginning of Nov.1, 2021

Proof. Note Φ is symmetric. If we let $\Phi f = u$ and $\Phi g = v$ then

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \quad \text{and} \quad \begin{cases} -\Delta v = g & \text{in } \Omega \\ v = 0 & \text{on } \Omega. \end{cases}$$

It follows that

$$\langle \Phi f, g \rangle_0 = \langle u, g \rangle_0 = \int ug \, \mathrm{d}x = \int \nabla u \nabla v \, \mathrm{d}x = \langle u, v \rangle_1$$

and

$$\langle f, \Phi g \rangle_0 = \langle f, v \rangle_0 = \int fv \, \mathrm{d}x = \int \nabla u \nabla v \, \mathrm{d}x = \langle u, v \rangle_1 .$$

Also, Φ is positive because $\langle \Phi f, f \rangle_0 = \langle u, f \rangle_0 = \langle u, u \rangle_1 \geqslant 0$.

Finally, to show $\Phi : L^2(\Omega) \to L^2(\Omega)$ compact, it suffices to show that the embedding $i : H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact. *Rellich's compactness theorem states that if a sequence $\{u_n\}$ in $H_0^1(\Omega)$ is satisfies $\|u_n\|_{H_0^1(\Omega)} \leqslant c$ then there exists a subsequence $\{u_{n_j}\}$ converging strongly in $L^2(\Omega)$.* □

Remark. The existence theory and the compactness of Φ can be extended to all second order elliptic operators with Dirichlet boundary conditions.

Remark. By the spectral theorem, Φ has an orthonormal basis of eigenfunctions e_k with positive eigenvalues α_k:

$$\Phi(e_k) = \alpha_k e_k.$$

Furthermore, since $\alpha_k \downarrow 0$, we have

$$-\Delta(\Phi(e_k)) = -\Delta \alpha_k e_k \implies \frac{1}{\alpha} e_k = (-\Delta)(e_k).$$

In some sense, the e_k's are the eigenfunctions of $-\Delta$ and the corresponding eigenvalues are $1/\alpha_k$'s. Furthermore, as $\alpha_k \downarrow 0, \lambda_k := 1/\alpha_k \to \infty$, so there exists a corresponding orthonormal basis $\{e_k\}$ of $L^2(\Omega)$ such that

$$-\Delta e_k = \begin{cases} \lambda_k e_k & \text{in } \Omega \\ e_k = 0 & \text{on } \partial\Omega. \end{cases}$$

More on the Eigenvalues

(1) Weyl's law; in $\mathbb{R}^n$, for eigenvalues λ_k of $-\Delta$ with zero boundary condition,

$$\lim_{k\to\infty} \lambda_k^{n/2}/k = \frac{(2\pi)^n}{\text{vol}(\Omega)\alpha(n)}$$

where $\alpha(n)$ is the column of the unit ball in $\mathbb{R}^n$.

(2) $\lambda_1 > 0$, called the principle value, is defined by

$$\lambda_1 = \min\{\langle u, u\rangle_1\}$$

subject to $u \in H_0^1(\Omega)$ and $\|u\|_{L^2}(\Omega) = 1$.The minimum is attained for $e_1 > 0$ in Ω solving

$$\begin{cases} -\Delta e_1 = \lambda_1 e_1 & \text{in } \Omega \\ e_1 = 0 & \text{on } \partial\Omega. \end{cases}$$

Moreover, if $u \in H_0^1(\Omega)$ is any weak solution of

$$\begin{cases} -\Delta u = \lambda_1 u & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

then u must be a multiple of e_1. Also,

$$\lambda_1 = \min \frac{\langle u, u\rangle_1}{\langle u, u\rangle_0}$$

subject to $u \in H_0^1(\Omega)$ and $u \neq 0$.

Chapter 6

Semigroups of Linear Operators

6.1 ODEs in a Banach Space

Let X be Banach and let $F : X \to X$ be Lipschitz, i.e., for some $L \in \mathbb{R}$ and all $x, y \in X$,

$$\|F(x) - F(y)\| \leqslant L\|x - y\|.$$

Consider the *Cauchy problem*

$$\begin{cases} \dot{x}(t) := \mathrm{d}x/\mathrm{d}t = F(x(t)) \\ x(0) = x_0 \in X. \end{cases} \tag{*}$$

Theorem 6.1.1: Existence and Uniqueness of (*)

Let F be defined as above. Then for every $x_0 \in X$, the Cauchy problem admits a unique solution $t \mapsto x(t)$ defined for all t.

Beginning of Nov.3 2021

Proof. Fix any $T > 0$. Consider the Banach space $C([O, T]; X)$ defined by the collection of $w : [0, T] \to X$ continuous. Define the norm on this space by

$$\|w\| := \max_{0 \leqslant t \leqslant T} e^{-2Lt}\|w(t)\|_X. \tag{1}$$

Notice that a function $x : [0, T] \to X$ will be a solution to (*) if and only if x is a fixed point of the *Picard operator*

$$x(t) - x(0) = \int_0^t F(w(s))\, \mathrm{d}s$$

or

$$[\Phi(w)](t) := x_0 + \int_0^t F(w(s))\, \mathrm{d}s \equiv w(t).$$

By the contraction mapping theorem, it suffices to show that Φ is a strong contraction with respect to the norm defined in (1). To this end, let $u, v \in C([0, T]; X)$ and $\delta = \|u - v\|$. That is,

$$e^{-Ls}\|u(s) - v(s)\|_X \leqslant \|u - v\| = \delta \qquad \text{for all } s \in [0, T],$$

so

$$\|u(s) - v(s)\|_X \leqslant e^{2Ls}\delta.$$

For each fixed $t \in [0, T]$,

$$\begin{aligned} e^{-2Lt}\|[\Phi(u)](t) - [\Phi(v)](t)\| &= e^{-2Lt}\left\| \int_0^t F(u(s)) - F(v(s))\, \mathrm{d}s \right\| \\ &\leqslant e^{-2Lt} \int_0^t \|F(u(s)) - F(v(s))\|\, \mathrm{d}s \\ \text{[Lipschitz]} &\leqslant e^{-2Lt} \int_0^t L\|u(s) - v(s)\|_X\, \mathrm{d}s \\ &\leqslant e^{-2Lt} \int_0^t Le^{2Ls}\delta\, \mathrm{d}s \\ &= e^{-2Lt} \cdot \frac{\delta}{2} e^{2Ls} = \frac{\delta}{2}. \end{aligned}$$

This shows that Φ is a strong contraction with contraction constant $\leqslant 1/2$, i.e.,

$$\|[\Phi(u)](x) - [\Phi(v)](x)\| \leqslant \frac{\|u(x) - v(x)\|}{2}.$$

The claim then follows from the contraction mapping theorem. By revering time, one can construct a unique solution on any time interval $[-T, 0]$ as well. □

Two Methods for Constructing Approximate Solutions to (*)

We fix a time step $h > 0$ and discretize time by defining $t_j := jh$ for $j \in \mathbb{N}$.
The *forward Euler approximation* is given by

$$x(t_{j+1}) = x(t_j) + hF(x(t_j)),$$

and the *backward Euler approximation* is given by

$$x(t_{j+1}) = x(t_j) + hF(x(t_{j+1})).$$

In both cases, once the values $x(t_j)$ have been computed on the *discrete* set $\{t_j\}$, one can extend the approximate solution to all real values of $t \geqslant 0$ by letting $t \mapsto x(t)$ be a linear/affine function.
To construct the forward approximation: we let $\dot{x}(t)$ to be constant on $[t_j, t_{j+1}]$ taking value of $F(x(t_j))$,

$$\dot{x}(t) = F(x(t_j)), \qquad t \in [t_j, t_{j+1}]$$

To construct the backward approximation: given $x(t_j)$, to find $x(t_{j+1})$, one needs to solve the implicit equation

$$x(t_{j+1}) = x(t_j) + hF(x(t_{j+1})).$$

This clearly involves more computational effort, but often times it has a much better outcome in terms of accuracy, stability, and convergence properties.

Linear Homogeneous ODEs

Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be linear. Then by the theorem above,

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

admits a solution map $t \mapsto e^{tA}x_0$ where

$$e^{tA} := \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}.$$

(Note that this series is absolutely convergent whenever A is bounded so it's well-defined.) Moreover, the exponential map satisfies

(1) $e^{0A} = I$,

(2) $e^{sA}e^{tA} = e^{(s+t)A}$ (semigroup property), and

(3) for each $x_0 \in X$, themap $t \mapsto e^{tA}x_0$ is continuous.

By (1) and (2), the family $\{e^{tA} : t \geqslant 0\}$ is a "group" of linear operators. In general, the theory of *linear semigroups* give the correspondence between A and $\{e^{tA} : t \geqslant 0\}$:

(1) When A is a bounded linear operator, its exponential map is computed by the series above.

(2) Conversely, given a family of e^{tA}, we can recover A as the limit

$$A = \lim_{t \to 0+} \frac{e^{tA} - I}{t}.$$

Question. Why do we care about $\{e^{tA}\}$?

Ans. Sometimes e^{tA} is bounded for all $t \geqslant 0$ whereas A might be unbounded. In these cases, semigroup theory is useful.

Example 6.1.2.

(1) If $A : \mathbb{R}^n \to \mathbb{R}^n$ is diagonal with diagonal entries $\lambda_1, \ldots, \lambda_n$, then $e^{tA} = \mathtt{diag}(e^{t\lambda_1}, \ldots, e^{t\lambda_n})$. Then $\|A\| = \max|\lambda_k|$ and $\|e^{tA}\| = \max|e^{t\lambda_k}|$.

(2) Consider $X = \ell^1$ with $\|x\|_1 = \sum|x_k|$. Given any sequence of complex numbers $\{\lambda_k\}$, consider the linear operator

$$Ax := \{\lambda_j x_j\}_{j \geqslant 1}.$$

Then

$$e^{tAx} = \{e^{t\lambda_j} x_j\}_{j \geqslant 1}.$$

The norms are given by $\|A\| = \sup|\lambda_k|$ (may be infinite) whereas

$$\|e^{tA}\| = \sup|e^{t\lambda_k}|$$

which is bounded!

Beginning of Nov.5, 2021

6.2 Semigroups of Linear Operators

Again, consider

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}u(t) = Au(t) \\ u(0) = u_0 \in X \end{cases}$$

where we want to express the solution as $u(t) = e^{tA}u_0$ for some family of linear operators $\{e^{tA} : t \geqslant 0\}$.

Example 6.2.1. Consider $u_t - u_x = 0$. We write $\dfrac{\mathrm{d}u}{\mathrm{d}t} = Au$ where A can be thought as a differential operator. Take $X := L^p(\mathbb{R})$ for some $p \in [1, \infty)$. Clearly A is unbounded. Its domain is the set of absolutely continuous functions $u \in L^p(\mathbb{R})$ with derivative $u_x \in L^p(\mathbb{R})$.
On the other hand, for all $u_0 \in L^p(\mathbb{R})$, the solution of

$$\begin{cases} u_t - u_x = 0 \\ u(0, x) = u_0(x) \end{cases}$$

can be explicitly computed: $u(t, x) = (e^{tA}u_0)(x) := u_0(x + t), t \in \mathbb{R}$. This implies that although A is unbounded, the corresponding exponential operator e^{tA} (the solution operator) is uniformly bounded.

We are interested in two types of questions:

(1) Given a semigroup of linear operators $\{S_t : t \geqslant 0\}$, find the **generator**, an operator A such that $S_t = e^{tA}$.

(2) Given a linear operator A, examine whether it generates a semigroup $\{e^{tA} : t \geqslant 0\}$ and establish the properties of this semigroup.

Definition 6.2.2: Semigroup

Let X be Banach. A **controngly continuous semigroup** of linear operators on X is a family of linear maps $\{S_t : t \geqslant 0\}$ with the following properties:

(1) each $S_t : X \to X$ is a bounded linear operator;

(2) for all $t, s \geqslant 0$, the composition satisfies $S_t \circ S_s = S_{t+s}$ and in particular $S_0 = I$ (identity operator);

(3) for all $u \in X$, the map $t \mapsto S_t u$ is continuous from $[0, \infty) \to X$.

We say that $\{S_t : t \geqslant 0\}$ is a **semigroup of type ω** if it in addition satisfies

$$\|S_t\| \leqslant e^{t\omega} \qquad \text{for all } t \geqslant 0.$$

In particular, when $\omega = 0$, a **semigroup of type 0** is also called a **contractive semigroup** ($\|S_t\| \leqslant 1$ uniformly):

$$\|S_t u - S_t v\| \leqslant \|u - v\| \qquad \text{for all } u, v \in X, t \geqslant 0.$$

Definition 6.2.3: Generator

The linear operator (it is an operator)

$$Au := \lim_{t \downarrow 0} \frac{S_t u - u}{t}$$

is called the **generator** of the semigroup $\{S_t : t \geqslant 0\}$. The domain of A is

$$\{u \in X : \text{ the limit above exists}\}.$$

For a given $u_0 \in X$, we regard the map $t \mapsto S_t u_0$ as the solution to the linear differential equation

$$\begin{cases} \frac{\mathrm{d}u}{\mathrm{d}t}(t) = Au(t) \\ u(0) = u_0 \in X. \end{cases} \tag{1}$$

In this way, we are approaching the problem "backwards", i.e., given the solution $u(t) = S_t u_0$, we reconstruct the evolution equation, finding the operator A.

Theorem 6.2.4

Let $\{S_t : t \geqslant 0\}$ be a strongly continuous semigroup with generator A. Assume $u_0 \in D(A)$. Then,

(1) for all $t \geqslant 0$, $S_t u_0 \in D(A)$ and $AS_t u_0 = S_t A u_0$;

(2) the map $t \mapsto u(t) := S_t u_0$ is continuous differentiable and it gives a solution to (1) above.

Proof.

(1) Since $u_0 \in D(A)$, the limit of $(S_t u - u)/t$ as $t \downarrow 0$ exists (and equals Au_0). Then,

$$\begin{aligned} \lim_{s \downarrow 0} \frac{S_s S_t u_0 - S_t u_0}{s} &= \lim_{s \downarrow 0} \frac{S_{t+s} u_0 - S_t u_0}{s} = \lim_{s \downarrow 0} \frac{S_t S_s u_0 - S_t u_0}{s} \\ &= \lim_{s \downarrow 0} \frac{S_t [S_s u_0 - u_0]}{s} = S_t \lim_{s \downarrow 0} \frac{S_s u_0 - u_0}{s} = S_t A u_0. \end{aligned}$$

Therefore, $S_t u_0 \in D(A)$ and $AS_t u_0 = S_t A u_0$.

(2) Let $S_t u_0 \in D(A)$ and $t > 0$. By semigroup property,

$$\begin{aligned} \lim_{h \downarrow 0} \left[\frac{S_t u_0 - S_{t-h} u_0}{h} - S_t A u_0 \right] &= \lim_{h \downarrow 0} \left[S_{t-h} \left(\frac{S_h u_0 - u_0}{h} \right) - S_t A u_0 \right] \\ &= \lim_{h \downarrow 0} \left[S_{t-h} \left(\frac{S_h u_0 - u_0}{h} - A u_0 \right) + S_{t-h} A u_0 - S_t A u_0 \right]. \end{aligned}$$

Since $u_0 \in D(A)$, the parenthesized term $\to 0$. The other two terms also $\to 0$ by continuity. Hence $t \mapsto S_t u_0$ has a left continuity. Furthermore,

$$\lim_{h \downarrow 0} \frac{S_t u_0 - S_{t-h} u_0}{h} = S_t A u_0.$$

Right derivative is easy to compute:

$$\lim_{h \downarrow 0} \frac{S_{t+h} u_0 - S_t u_0}{h} = S_t \lim_{h \downarrow 0} \frac{S_h (u_0) - u_0}{h} = S_t A u_0.$$

Therefore, for all $t > 0$, the map $t \mapsto S_t u_0$ is differentiable with derivative

$$\frac{\mathrm{d}}{\mathrm{d}t} S_t u_0 = S_t A u_0 = A S_t u_0.$$

Since $Au_0 \in X$, by definition the map $t \mapsto S_t A u_0$ is continuous.

□

Recall that a linear operator $A : X \to X$ is **closed** if its graph

$$\text{Graph}(A) := \{(x, y) \in X \times X : x \in D(A), y = Ax\}$$

is closed in $X \times X$.

Theorem 6.2.5: Properties of Generators

Let $\{S_t : t \geqslant 0\}$ be a strongly continuous semigroup on X and let A be its generator. Then

(1) $D(A)$ is dense in X,

(2) A is closed.

Beginning of Nov.8,2021

Proof.

(1) Fix $u \in X$. Consider the approximation

$$u_\epsilon := \frac{1}{\epsilon} \int_0^\epsilon S_\epsilon u \, \mathrm{d}s.$$

This converges to $S_0 u = u$ as $\epsilon \downarrow 0$ since the map $t \mapsto S_t u$ is continuous by assumption. It suffices to show that $u_\epsilon \in D(A)$ for all $\epsilon > 0$. Since $D(A)$ is a subspace, it further suffices to show that ϵu_ϵ (i.e., the integral) is in $D(A)$. For $0 < h < \epsilon$,

$$\begin{aligned} \frac{S_h u_\epsilon - u_\epsilon}{h} &= \frac{1}{h}\left[S_h\left(\int_0^\epsilon S_s u \, \mathrm{d}s\right) - \int_0^\epsilon S_s u \, \mathrm{d}s\right] \\ &= \frac{1}{h}\int_0^\epsilon (S_{h+s}u - S_s u) \, \mathrm{d}s \\ [\tilde{s} := s + h] &= \frac{1}{h}\left[\int_h^{h+\epsilon} S_{\tilde{s}} u \, \mathrm{d}\tilde{s} - \int_0^\epsilon S_s u \, \mathrm{d}s\right] \\ &= \frac{1}{h}\left[\int_\epsilon^{h+\epsilon} S_s u \, \mathrm{d}s - \int_0^h S_s u \, \mathrm{d}s\right]. \end{aligned}$$

As $h \to 0$, the above converges to $S_\epsilon u - S_0 u = S_\epsilon u - u$. Therefore, $u_\epsilon \in D(A)$ for all $\epsilon > 0$.

(2) Let $(u_k, v_k) \in \text{Graph}(A)$ be such that $(u_k, v_k) \to (u, v) \in X \times X$. In other words, $u_k \in D(A)$, $u_k \to u$, $v_k = Au_k$, and $v_k \to v$.

Recall that $t \mapsto S_t u_k$ is continuously differentiable. Therefore, for each $k \geqslant 1$,

$$S_h u_k - u_k = \int_0^h \frac{\mathrm{d}}{\mathrm{d}t}(S_t u_k) \, \mathrm{d}t = \int_0^h A S_t u_k \, \mathrm{d}t = \int_0^h S_t A u_k \, \mathrm{d}t.$$

Letting $k \to \infty$,

$$S_h - u = \int_0^h S_t v \, \mathrm{d}t.$$

Therefore,

$$\lim_{h \downarrow 0} \frac{S_h u - u}{h} = \lim_{h \downarrow 0} \frac{1}{h} \int_0^h S_t v \, \mathrm{d}t = v.$$

Therefore $u \in D(A)$, with $v = Au$.

□

6.3 Resolvents

Suppose we want to approximate a solution to

$$\begin{cases} \dfrac{\mathrm{d}u(t)}{\mathrm{d}t} = Au(t) \\ u(0) = u_0 \in X \end{cases}$$

by backward Euler approximation. That is, we fix a time step $h > 0$ and iteratively solve

$$u(t+h) = u(t) + hAu(t+h).$$

At each step, given a value $u(t) \in X$, we need to compute

$$u(t+h) = (I - hA)^{-1}u(t).$$

Define the backward Euler approximation operator $E_h := (I - hA)^{-1}$.

For a fixed time $\tau > 0$, consider the times step $h := \tau/n$. After n steps, the backward approximation

$$u(\tau) \approx (E_{\tau/n})^n(u_0) = (I - \tau A/n)^{-n}(u_0)$$

For a fixed τ, letting $n \to \infty$, we have

$$u(\tau) = S_\tau(u_0) := \lim_{n\to\infty}(I - \tau A/n)^n u_0 = e^{\tau A}(u_0).$$

This gives one way to implement the backward Euler approximation.

More abstractly, for a fixed time step $h > 0$, define $\lambda = 1/n$. Define $A_\lambda : X \to X$ as $A_\lambda u = AE_h^{-1}u = A(I - hA)^{-1}u$. It turns out $A_\lambda = A_{1/n}$ is a well-defined bounded linear operator for sufficiently small h. Then we can consider

$$e^{tA_\lambda} = \sum_{k\geqslant 0}\frac{(tA_\lambda)^k}{k!}$$

and then define

$$u(t) := S_t u = \lim_{\lambda\to\infty} e^{tA_\lambda}u_0.$$

Example 6.3.1. Consider the scalar ODE

$$\dot{x} = ax \qquad x(0) = x_0$$

which clearly has the solution $e^{ta}x_0$. In this case $a_\lambda = a_{1/n} = a/(1 - ha)$, and indeed

$$\lim_{h\to 0} e^{ta_{1/n}}x_0 = \lim_{h\to 0} e^{ta/(1-ha)}x_0 = e^{at}x_0.$$

Note that for $0 < h < 1/a$, $\int_0^\infty \frac{e^{-t/h}}{h}\,\mathrm{d}t = 1$ and $(1 - ha)^{-1}x_0 = \int_0^\infty \frac{e^{-t/h}}{h}\cdot e^{ta}x_0\,\mathrm{d}t$. The latter suggests that the backward approximation operator can be obtained by taking some kind of weighted average.

Definition 6.3.2: Resolvent Operator

Let A be a linear operator on a Banach space X. Recall that the **resolvent set** of A is the set $\rho(A)$ of all real numbers λ such that $\lambda I - A : D(A) \to X$ is bijective. If $\lambda \in \rho(A)$, the **resolvent operator** $R_\lambda : X \to X$ is defined by

$$R_\lambda u = (\lambda I - A)^{-1} u.$$

Beginning of Nov.10, 2021

Remark. This implies $\lambda R_\lambda = E_{1/\lambda}$. In particular, if A is a closed operator, then the CGT implies R_λ is continuous. Moreover, $AR_\lambda u = R_\lambda A u$ for $u \in D(A)$:

$$v = R_\lambda u = (\lambda I - A)^{-1} u \implies (\lambda I - A)v = u$$
$$\implies R_\lambda A u = R_\lambda A(\lambda I - A)v = R_\lambda(\lambda A - A^2)v = Av = AR_\lambda u.$$

Theorem 6.3.3

Let A be a closed linear operator. If $\lambda, \mu \in \rho(A)$, then

$$R_\lambda - R_\mu = (\mu - \lambda) R_\lambda R_\mu.$$

Furthermore, R_λ commutes with R_u.

Proof. Let $u \in X$ be given. Let $v := (R_\lambda - R_\mu)u = (\lambda I - A)^{-1}u - (\mu I - A)^{-1}u$. Then,

$$\begin{aligned}(\lambda I - A)v &= u - (\lambda I - \mu I + \mu I - A)(\mu I - A)^{-1}u \\ &= u - u + (\mu - \lambda)(\mu I - A)^{-1}u.\end{aligned}$$

Therefore,

$$v = (\mu - \lambda)(\lambda I - A)(\mu I - A)^{-1}u = (\mu - \lambda)R_\lambda R_\mu u.$$

On the other hand,

$$R_\lambda R_\mu = \frac{R_\lambda - R_\mu}{\mu - \lambda} = \frac{R_\mu - R_\lambda}{\lambda - \mu} = R_\mu R_\lambda.$$

□

Theorem 6.3.4: Integral Formula for the Resolvent Operator

Let $\{S_t : t \geqslant 0\}$ be a semigroup of type ω and let A be its generator. Then for every $\lambda > \omega$, we have $\lambda \in \rho(A)$. Moreover,

$$R_\lambda u = \int_0^\infty e^{-t\lambda} S_t u \, \mathrm{d}t$$

and $\|R_\lambda\| \leqslant 1/(\lambda - \omega)$.

Proof. Define another operator $\tilde{R}_\lambda u := \int_0^\infty e^{-t\lambda} S_t u \,\mathrm{d}t$. Recall $\|S_t\| \leqslant e^{tw}$. Therefore $\tilde{R}_\lambda u$ is absolutely convergent:

$$\|\tilde{R}_\lambda u\| \leqslant \int_0^\infty e^{-t\lambda}\|S_t u\| \,\mathrm{d}t \leqslant \int_0^\infty e^{-t\lambda}\|S_t\|\|u\| \,\mathrm{d}t \leqslant \int_0^\infty e^{-t(\lambda-w)}\|u\| \,\mathrm{d}t = \frac{\|u\|}{\lambda-\omega}$$

so $\|\tilde{R}_\lambda\| \leqslant 1/(\lambda-\omega)$. It remains to show verify that $\tilde{R}_\lambda = R_\lambda$.

Claim 1: $(\lambda I - A)\tilde{R}_\lambda = u$ for all $u \in X$.

To compute Au, we use the definition and consider the difference quotient:

$$\begin{aligned}
\frac{S_h\tilde{R}_\lambda u - \tilde{R}_\lambda u}{h} &= \frac{1}{h}\int_0^\infty e^{-\lambda t}(S_{t+h}u - S_t u)\,\mathrm{d}t \\
[\tilde{t} := t+h] &= \frac{1}{h}\left[\int_h^\infty e^{-\lambda(t-h)}S_{tu}\,\mathrm{d}t - \int_0^\infty e^{-\lambda r}S_t u\,\mathrm{d}t\right] \\
&= \int_0^\infty (e^{-\lambda(t-h)} - e^{-\lambda t})S_t u\,\mathrm{d}t - \frac{1}{h}\int_0^h e^{-\lambda(t-h)}S_t u\,\mathrm{d}t \\
&= \frac{e^{\lambda h}-1}{h}\left[\int_0^\infty e^{-\lambda t}S_t u\,\mathrm{d}t\right] - \frac{e^{\lambda h}}{h}\int_0^h e^{-\lambda t}S_t u\,\mathrm{d}t.
\end{aligned}$$

Take $h \downarrow 0$,

$$\lim_{h\downarrow 0}\frac{S_h\tilde{R}_\lambda u - \tilde{R}_\lambda u}{h} = \lambda\tilde{R}_\lambda u - 1\cdot S_0 u = \lambda\tilde{R}_\lambda u - u.$$

Therefore, $\tilde{R}_\lambda u \in D(A)$ and $A\tilde{R}_\lambda u = \lambda\tilde{R}_\lambda u - u$. Rearranging the terms gives the claim.

Now, by claim 1, $u \mapsto (\lambda I - A)u$ from $D(A) \to X$ is surjective. It remains to show that it's injective. Note that

$$\begin{aligned}
A\tilde{R}_\lambda u = A\int_-^\infty e^{-\lambda t}S_t u\,\mathrm{d}u &= \int_0^\infty e^{-\lambda t}AS - tu\,\mathrm{d}t \\
&= \int_0^\infty e^{-\lambda t}S_t(Au)\,\mathrm{d}t = \tilde{R}_\lambda Au.
\end{aligned}$$

Then,

$$\tilde{R}_\lambda(\lambda I - A)u = (\lambda I - A)\tilde{R}_\lambda u \qquad \text{for all } u \in D(A).$$

If $(\lambda I - A)u = (\lambda I - A)v$, then

$$u = \tilde{R}_\lambda(\lambda I - A)u = \tilde{R}_\lambda(\lambda I - A)v = v.$$

This finishes the proof, as $\lambda \in \rho(A)$ and

$$\tilde{R}_\lambda = (\lambda I - A)^{-1} = R_\lambda.$$

□

Remark. Integral representation implies that the resolvent operators R_λ provide the Laplace transform[!] of the semigroup S_t. Taking $0 < h < 1/\lambda$, the same formula shows that the backward Euler approximation can be obtained as

$$E_h u = (I - hA)^{-1}u = \int_0^\infty e^{-t/h}h^{-1}S_t u\,\mathrm{d}t,$$

which is convergent for sufficiently ($< 1/\omega$) h. We will show that the converse is also true in next section.

6.4 Generation of a Semigroup

Theorem 6.4.1: Existence of the Semigroup Generated by a Linear Operator

Let A be a linear operator on X. TFAE:

(1) A is the generator of a semigroup of linear operators $\{S_t : t \geqslant 0\}$ of type ω;

(2) A is a closed, densely defined defined operator. Moreover, every real number $\lambda > \omega$ is in the resolvent set $\rho(A)$. Moreover, $\|(\lambda I - A)^{-1}\| \leqslant 1/(\lambda - \omega)$ for all $\lambda > \omega$.

((1) ⇒ (2) has been proven already; it remains to show the converse.)

Beginning of Nov.12, 2021

Proof. Since $R_\lambda := (\lambda I - A)^{-1}$ is well-defined for $\lambda > \omega$, we can consider the bounded linear operator

$$\begin{aligned} A_\lambda &:= -\lambda I + \lambda^2 R_\lambda \\ &= \lambda(-I + \lambda R_\lambda) \\ &= \lambda(-(\lambda I - A)R_\lambda + \lambda I R_\lambda) \\ &= \lambda(-\lambda I + A + \lambda I)R_\lambda = \lambda A R_\lambda. \end{aligned}$$

Setting $h = 1/\lambda$, we have

$$A_\lambda u = A(I - hA)^{-1} = A(E_n^{-1} u)$$

Since A_λ is bounded, we can consider the exponential operator

$$e^{tA_\lambda} := \sum_{k=0}^\infty \frac{(tA_\lambda)^k}{k!} = e^{-\lambda t} e^{\lambda^2 t R_\lambda} = e^{-\lambda t} \sum_{k=0}^\infty \frac{(\lambda^2 t)^k R_\lambda^k}{k!}.$$

If A is unbounded then $\|A_\lambda\| \to \infty$ as $\lambda \to \infty$ but $\|e^{tA_\lambda}\|$ remain uniformly bounded for $t > 0$:

$$\|e^{tA_\lambda}\| \leqslant e^{-\lambda t} \sum_{k=0}^\infty \frac{(\lambda^2 t)^k \|R_\lambda\|^k}{k!} \leqslant e^{-\lambda t} e^{\lambda^2 t/(\lambda - \omega)}.$$

In particular, for $\lambda \geqslant 2\omega$, $\|e^{tA-\lambda}\| \leqslant e^{2\omega t}$.

Now, we claim that $\lim_{\lambda\to\infty} A_\lambda v = Av$ for all $v \in D(A)$ so that A_λ "tends to" A.
Note that $\lambda R_\lambda u - u = AR_\lambda u = R_\lambda A u$ for all $u \in D(A)$. This implies

$$\begin{aligned} \|\lambda R_\lambda u - u\| = \|R_\lambda A u\| &\leqslant \|R_\lambda\| \|Au \\ &\leqslant \frac{1}{\lambda - \omega} \|Au\| \to 0 \end{aligned}$$

for u fixed and as $\lambda \to \infty$. Since A is densely defined, for each $u \in X$ and $\epsilon > 0$, there exists $v \in D(A)$ with $\|u - v\| < \epsilon$. Then

$$\begin{aligned} \limsup_{\lambda\to\infty} \|\lambda R_\lambda u - u\| &= \limsup_{\lambda\to\infty} \|\lambda R_\lambda u - \lambda R_\lambda v\| + \limsup_{\lambda\to\infty} \|\lambda R_\lambda v - v\| + \|v - u\| \\ &< \limsup_{\lambda\to\infty} \|\lambda R_\lambda u - \lambda R_\lambda u\| + 0 + \epsilon \\ &\leqslant \|\lambda R_\lambda\| \|u - v\| + \epsilon \leqslant \limsup_{\lambda\to\infty} \frac{\lambda}{\lambda - \omega} \epsilon + \epsilon = 2\epsilon. \end{aligned}$$

Hence $\lim_{h\downarrow 0} E_h u = u$ for all $u \in X$. If $v \in D(A)$, then we can take $u := Av$ and

$$\lim_{\lambda\to\infty} A_\lambda v = \lim_{\lambda\to\infty} \lambda A R_\lambda v = \lim_{\lambda\to\infty} \lambda R_\lambda A v = \lim_{\lambda\to\infty} \lambda R_\lambda u = u = Av.$$

Finally, we verify that e^{tA_λ} is our candidate for the generator, i.e., e^{tA_λ} converges to some linear operator as $\lambda \to \infty$. Since $R_\lambda R_\mu = R_\mu R_\lambda$, we have $A_\lambda A_\mu = A_\mu A_\lambda$. Then,

$$A_\mu e^{tA_\lambda} = A_\mu \sum_{k=0}^{\infty} \frac{(tA_\lambda)^k}{k!} = e^{-A_\lambda} A_\mu.$$

For each $u \in X$, we have

$$\begin{aligned} e^{tA_\lambda} u - e^{tA_u} u &= \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} \left[e^{(t-s)A_\mu} e^{sA_\lambda} u \right] \mathrm{d}s \\ &= \int_0^t e^{(t-s)A_\mu} (A_\lambda - A_\mu) e^{sA_\lambda} u \, \mathrm{d}s \\ &= \int_0^t e^{(t-s)A_\mu} e^{sA_\lambda} (A_\lambda u - A_\mu u) \, \mathrm{d}s. \end{aligned}$$

Then, in particular for $\lambda, \mu > 2\omega$, by the uniform bound $\|e^{tA-\lambda}\|; e^{2\omega t}$ shown above,

$$\begin{aligned} \|e^{tA_\lambda} u - e^{tA_\mu} u - \| &\leqslant \int_0^t e^{2(t-s)\omega} e^{2s\omega} \|A_\lambda u - A_\mu u\| \, \mathrm{d}s \\ &= t e^{2\omega t} \|A_\lambda u - A_\mu u\|. \end{aligned}$$

For $u \in D(A)$, by the previous step, A_λ, A_μ both converge to A, so the norm converges to 0. Therefore

$$\limsup_{\lambda,\mu\to\infty} \|e^{tA_\lambda} u - e^{tA_\mu} u\| \leqslant t e^{2\omega t} \limsup_{\lambda,\mu\to\infty} \|A_\lambda u - A_\mu u\| = 0.$$

More generally, for $u \in X$, we can approximate it by $u \in D(A)$ and go over another triangle inequality argument. Once this is done, we have shown that $\{e^{tA_\lambda}\}$ is convergent.
Finally, for all $t \geqslant 0$ and $u \in X$, we claim that the limit $\lim_{\lambda\to\infty} e^{tA_\lambda} u$ is well-defined:

$$S_t u := \lim_{\lambda\to\infty} e^{tA_\lambda} u$$

and $\{S_t : t \geqslant 0\}$ is a strongly continuous semigroup of type ω.

$$\begin{aligned} S_t S_s u &= \lim_{\lambda\to\infty} e^{tA_\lambda} e^{sA_\lambda} u \\ &= \lim_{\lambda\to\infty} e^{(t+s)A_\lambda} u = S_{t+s} u. \end{aligned}$$

For a fixed $u \in x$, the map $t \mapsto S_t u$ is continuous (it is the uniform limit of $t \mapsto e^{tA_\lambda} u$). For each $t \geqslant 0$ and $u \in X$ with $\|u\| \leqslant 1$, we have

$$\begin{aligned} \|S_t u\| &= \lim_{\lambda\to\infty} \|e^{tA_\lambda} u\| \leqslant \lim_{\lambda\to\infty} \|e^{tA-\lambda}\| \|u\| \\ &\leqslant \lim_{\lambda\to\infty} \|e^{tA_\lambda}\| \leqslant \lim_{\lambda\to\infty} e^{t\lambda\omega/(\lambda-\omega)} = e^{t\omega}. \end{aligned}$$

Therefore $\|S_t\| \leqslant e^{t\omega}$, which shows $\{S_t : t \geqslant 0\}$ is indeed of type ω.
Last step: we verify that A is indeed the generator of the semigroup. To this end, let B be the generator of $\{S_t : t \geqslant 0\}$. By previous result, we know B is a linear, closed operator densely defined in X.

Since A_λ is the generator of the semigroup $\{e^{tA_\lambda} : t \geqslant 0\}$, for every $\lambda > \omega$ we have

$$e^{tA_\lambda}u - u = \int_0^t e^{sA_\lambda} A_\lambda u \, \mathrm{d}s. \tag{1}$$

For $u \in D(A)$,

$$\|e^{sA_\lambda} A_\lambda u - S_s Au\| \leqslant \|e^{sA_\lambda}\| \|A_\lambda u - Au\| + \|e^{sA_\lambda} Au - S_s Au\| \to 0$$

uniformly as $\lambda \to \infty$ for s in bounded intervals. Taking $\lambda \to \infty$ in (1), we have

$$S_t u - u = \int_0^t S_s Au \, \mathrm{d}s$$

for all $t \geqslant 0$ and $u \in D(A)$. Hence $D(B) \supset D(A)$:

$$Bu = \lim_{h\downarrow 0} \frac{S_t u - u}{t} = \lim_{t\downarrow 0} \frac{1}{t} \int_0^t S_s Au \, \mathrm{d}s = Au \qquad \text{for all } u \in D(A).$$

It remains to show $D(B) \subset D(A)$. Let $\lambda > \omega$. We know $\lambda I - A : D(A) \to X, \lambda I - B : D(B) \to X$ are both bijective. In particular, the restriction of $\lambda I - B$ to $D(A)$ is exactly $\lambda I - A$ and thus surjective. By injectivity of $\lambda I - B$ on $D(B)$ we see that $\lambda I - B$ cannot be extended outside $D(A)$! Hence, $D(B) = D(A)$.

□

Beginning of Nov.15, 2021

Theorem: Uniqueness of the Semigroup

Let $\{S_t\}, \{\tilde{S}_t\}$ be two strongly continuous semigroups of linear operators having the same generator A. Then $S_t = \tilde{S}_t$ for all $t \geqslant 0$.

Proof. Let $u \in D(A)$. Then $\tilde{S}_s u \in D(A)$ and $S_{t-s}\tilde{S}_s u \in D(A)$ for all $0 \leqslant s \leqslant t$. By FTC,

$$\tilde{S}_t u - S_t u = \int_s^t \frac{\mathrm{d}}{\mathrm{d}s}\left[S_{t-s}\tilde{S}_s u\right] \mathrm{d}s.$$

Note that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\left[S_{t-s}\tilde{S}_s u\right] &= \lim_{h\downarrow 0} \frac{S_{t-s-h}(\tilde{S}_{s+h}u) - S_{t-s}\tilde{S}_s u}{h} \\
&= \lim_{h\downarrow 0} \frac{S_{t-s-h}(\tilde{S}_{s+h}u - \tilde{S}_s u)}{h} + \lim_{h\downarrow 0} \frac{S_{t-s-h}\tilde{S}_s u - S_{t-s}\tilde{S}_s u}{h} \\
&= \lim S_{t-s-h} \lim \frac{\tilde{S}_{s+h}u - \tilde{S}_s u}{h} + \dots \\
&= S_{t-s}(A\tilde{S}_s u) - AS_{t-s}(\tilde{S}_s u) = 0
\end{aligned}$$

where the last step is by commutativity. By density $\tilde{S}_t u - S_t u = 0$. □

6.5 Parabolic Equations

Let $\Omega \subset \mathbb{R}^n$ be bounded. Recall the **elliptic operator**

$$Lu := -\sum_{i,j=1}^n (a^{i,j}(x)u_{x_i})_{x_j} + \sum_{i=1}^n (b^i(x)u)_{x_i} + c(x)u$$

where $a^{i,j}, b^i, c \in L^\infty(\Omega)$ are functions of x and

$$\sum_{i,j=1}^{n} a^{i,j}(x)\xi_i\xi_j \geqslant \theta|\xi|^2$$

for all $x \in \Omega$, $\xi \in \mathbb{R}^n$ and some $\theta > 0$. (Trivial example: let $(a^{i,j})$ be the identity matrix.)
Now consider

$$\begin{cases} u_t + Lu = 0 & (t,x) \in (0,\infty)\times\Omega \\ u(t,x) = 0 & (t,x) \in (0,\infty)\times\partial\Omega \\ u(0,x) = g(x) & x \in \Omega. \end{cases}$$

We can reformulate the initial boundary value problem as a Cauchy problem (IVP) in a Hilbert space $X \in L^2(\Omega)$:

$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t} = Au \\ u(0) = g \end{cases}$$

for a suitable (unbounded) linear operator $A : L^2(\Omega) \to L^2(\Omega)$. (Of course $A = -L$.) Define

$$D(A) \coloneqq \{u \in H_0^1(\Omega) : Lu \in L^2(\Omega)\}.$$

Namely, $u \in D(A)$ if u is a solution to the elliptic boundary value problem

$$\begin{cases} Lu = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

for some $f \in L^2(\Omega)$.
The goal is to construct a solution to the Cauchy equaition using semigroup theory.
Assume there exists $\beta > 0$ such that

$$B : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$$

where the bilinear form

$$B[u,v] \coloneqq \int_\Omega \sum_{i,j=1}^{n} a^{i,j} u_{x_i} u_{x_j} - \sum_{j=1}^{n} b^i u v_{x_i} + cuv \,\mathrm{d}x$$

is strictly positive definite:

$$B[u,u] \geqslant \beta\|u\|_{H_1} \qquad \text{for all } u \in H_0^1(\Omega).$$

Theorem 6.5.1: Semigroup of Solutions of a Parabolic Equation

Assume L is uniformly elliptic and assume $a^{i,j} \in W^{1,\infty}(\Omega)$ (i.e., smooth derivative), $b^i, c \in L^\infty(\Omega)$. Assume that the corresponding bilinear form defined above is strictly PD. Then the operator $A = -L$ generate a contractive semigroup $\{S_t : t \geqslant 0\}$ of linear operators on $L^2(\Omega)$.

Proof. To prove A generates a semigroup on $X = L^2(\Omega)$, by the previous existence theorem, we want to show $D(A)$ is show that

(1) $D(A)$ is dense in $L^2(\Omega)$,

(2) graph(A) is closed, and

(3) every $\lambda > 0 = \omega$ is in $\rho(A)$, i.e., $\lambda I - A$ is bijective, with $\|(\lambda I - A)^{-1}\| \leqslant 1/\lambda$.

□

Beginning of Nov.17, 2021

Proof.

(1) If $\varphi \in C_c^2(\Omega)$ (compactly supported), then $L\varphi = f \in L^2(\Omega)$. This means that $C_c^2(\Omega) \subset D(A)$. But $C_c^2(\Omega)$ is dense in $L^2(\Omega)$, so we are done.

(2) We now show that graph of A is closed. To this end, by Lax-Milgram, for $f \in L^2(\Omega)$, there exists a unique $u \in H_0^1(\Omega)$ such that

$$B[u,v] = \langle f, v\rangle_{L^2} \qquad \text{for all } u \in H_0^1(\Omega).$$

Then, the map $f \mapsto u = L^{-1}f$ is a bounded linear operator from $L^2(\Omega)$ to $L^2(\Omega)$. Note that $(u, f) \in \text{Graph}(A)$ is equivalent to $(-f, u) \in \text{Graph}(L^{-1})$. Since L^{-1} is continuous, its graph is closed.

(3) We want to show that for all $\lambda > 0$, the operator $\lambda I - A$ has a bounded inverse $\|(\lambda I - A)^{-1}\| \leqslant 1/\lambda$. That is, it suffices to show that for each $f \in L^2(\Omega)$, the problem

$$\begin{cases} \lambda u + Lu = f & x \in \Omega \\ u = 0 & x \in \partial\Omega \end{cases}$$

has a unique solution satisfying

$$(\|(\lambda I - A)' f\|_{L^2} =) \qquad \|u\|_{L^2} \leqslant \frac{\|f\|_{L^2}}{\lambda}$$

By Lax-Milgram, there exists a unique $u \in H_0^1(\Omega)$ such that

$$\langle \lambda u, u\rangle_{L^2} + B[u,v] = \langle f, v\rangle_{L^2} \qquad \text{for all } v \in H_0^1(\Omega).$$

Taking $v = u$ gives

$$\lambda\|u\|_{L^2} + B[u,u] = \langle f, u\rangle_{L^2} \leqslant \|f\|_{L^2}\|u\|_{L^2}$$

Therefore $\lambda\|u\|_{L^2} \leqslant \|f\|_{L^2}$, and this proves the claim $\|(\lambda I - A)^{-1}\| \leqslant 1/\lambda$.

Therefore, by the semigrouop theory, we deduce that the linear operator A generates a contractive semigroup.

□

General Operators

Example 6.5.2. Let A be an $m \times n$ matrix. Consider the linear ODE on $\mathbb{R}^n$

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = -Ax(t).$$

If A is PD, then $-A$ generates a contractive semigroup. In particular,

$$\frac{\mathrm{d}}{\mathrm{d}t}|x(t)|^2 = 2\left\langle \frac{\mathrm{d}x(t)}{\mathrm{d}t}, x(t)\right\rangle = 2\langle -Ax(t), x(t)\rangle \leqslant 0.$$

Hence $|x(t)|^2 \leqslant |x(0)|^2$ for all $t \geqslant 0$, giving a contractive bound.

Now let A be any matrix. Then we can let γ be sufficiently large such that $A+\gamma I$ is PD. Then the claim above holds analogously, as $-(A+\gamma I)$ generates a contractive semigroup. If $x(t) = e^{-tA}x(0)$ is a solution, then

$$\begin{aligned} |x(t)| &= |e^{-tA}x(0)| = |e^{(\gamma I-(A+\gamma I))t}x(0)| \\ &= e^{\gamma t}|e^{-(A+\gamma I)x(0)}| \leqslant e^{\gamma t}|x(0)|. \end{aligned}$$

Hence $-A$ generates a semigroup of type γ.

Example 6.5.3. Consider a general elliptic operator where the corresponding bilinear form $B[u,v]$ is not necessarily PD. In this case, one can also find a sufficiently large $\gamma > 0$ such that

$$B_\gamma[u,v] := B[u,v] + \gamma \langle u, v\rangle_{L^2}$$

is strictly PD on $H_0^1(\Omega)$. Define $L_\gamma u := Lu + \gamma I$. Then $u_t = -Lu$ can be written as $u - t = -L_\gamma u + \gamma u$. By the previous theorem, $A_\gamma := -(L+\gamma I)$ generates a contractive semigroup of linear oeprators $\{S_t^{(\gamma)} : t \geqslant 0\}$. Now $A = -L = \gamma I - L_\gamma$ with $D(A) = \{u \in H_0^1(\Omega); Lu \in L^2(\Omega)\}$ generates a semigroup of type γ with

$$\{S_t : t \geqslant 0\} \text{ where } S_t = e^{\gamma t}S_t^{(\gamma)}, t \geqslant 0.$$

What about the solution map $t \mapsto u(t) = S_t g$?

(1) For all initial data $g \in L^2(\Omega)$, $t \mapsto u(t) = S_t g$ is C^1 on $D(A)$ and satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}u(t) = Lu(t) \qquad \text{for all } t > 0.$$

(2) The map $t \mapsto u(t) = S_t g$ is continuous from $[0,\infty)$ into $L^2(\Omega)$ and satisfies $u(0) = g$. Thus, the initial condition is satisfied as an identity in L^2 (because S_0 is).

(3) If $g \in D(A)$, then $u(t) = S_t g \in D(A)$ for all $t \geqslant 0$. Then the map $t \mapsto u(t)$ is continuously differentiable and

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}u = Au \\ u(0) = g \end{cases} \qquad \text{for all } t > 0.$$

Since $D(A) \subset H_0^1(\Omega)$, this implies $u(t)$ satisfies the correct boundary condition for all $t \geqslant 0$.

(4) If $g \in L^2(\Omega)$, then there exists a sequence $g_n \in D(A)$ such that $\|g - g_n\|_{L^2} \to 0$ as $n \to \infty$. If the semigroup is of type γ, we have

$$\|S_t g_n - S_t g\|_{L^2} \leqslant e^{\gamma t}\|g_n - g\|_{L^2} \to 0$$

for fixed g. Hence $u \mapsto u(t) = S_t f$ is the limit of of a sequence of C^1 solutions $t \mapsto u_n t = S_t g_n$.

Claim. $u = u(t,x)$ provides a solution to the parabolic equation

$$u_t = \sum_{i,j=1}^n (a^{i,j}u_{x_i})_{x_j} - \sum_{i=1}^n b^i u_{x_i} - cu$$

in the sense of distribution: for all $\varphi \in C_c^\infty(\Omega \times (0,\infty))$,

$$\iint_\Omega u\varphi_t + \sum_{i,j=1}^n u(a^{i,j}\varphi_{x_j})_{x_i} + \sum_{i=1}^n u(b^i\varphi)_{x_i} - cu\varphi \,\mathrm{d}x\mathrm{d}t = 0.$$

Consider $u_m := S_t g_m$ where $\|g - g_m\|_{L^2} \to 0$. $t \mapsto u_m(t)$ converges to the continuous trajectory $t \mapsto u(t) = S_t g$ in $C^0([0,t]; L^2(\Omega))$. We deduce that u is a solution in the sense of distribution.

Chapter 7

Detour: Quantum Mechanics & Functional Analysis

Beginning of Nov.19, 2021

Consider the system at an arbitrary fixed instant by regarding the time as a parameter which we keep fixed.

In **classical mechanics**, the state is described by specifying position and velocity by a pair of numbers (x, v), whereas in **quantum mechanics**, the state of system is described by a *function*, often times $\psi : \mathbb{R}^n \to \mathbb{C}$.

If we let $\psi \in L^2(\mathbb{R})$ with $\|\psi\|_{L^2} = 1$, then ψ is related to the probability that will be found in a given subset $J \subset \mathbb{R}$ where the probability is

$$\int_J |\psi(q)|^2 \, \mathrm{d}q. \tag{*}$$

(Indeed, we want the particle to always be somewhere on the real line so $\int_{\mathbb{R}} |\psi(q)|^2 \, \mathrm{d}q = 1$.)
Note that (*) is invariant under multiplication by complex numbers with modulus 1, i.e.,

$$\int_J |\psi(q)|^2 \, \mathrm{d}q = \int_J |\alpha\psi(q)|^2 \, \mathrm{d}q \qquad \text{for all } |\alpha| = 1, J \subset \mathbb{R}.$$

Thus, the **deterministic** description of a state in classical mechanics is replaced by a **probabilistic** description of a state in quantum mechanics.

To this end, we define a **state** of our physical system at some instant to be an element $\psi \in L^2(\mathbb{R})$ with $\|\psi\|_{L^2} = 1$, defined uniquely up to multiplication by $\alpha \in \mathbb{C}$ with $|\alpha| = 1$. As said before, $|\psi(q)|^2$ plays a role of a probability density function on $\mathbb{R}$. It follows that we can define the **expected value**, **variance**, and **standard deviation** by

$$\mu_\psi := \int_{\mathbb{R}} q|\psi(q)|^2 \, \mathrm{d}q, \qquad \mathrm{var}_\psi := \int_{\mathbb{R}} (q - \mu_\psi)^2 |\psi(q)|^2 \, \mathrm{d}q, \qquad \text{and} \qquad \mathrm{sd}_\psi := \sqrt{\mathrm{var}_\psi}.$$

(Intuitively, the mean gives the "central location" and the other two show dispersion, just like in probability.)

Note that μ_ψ can be expressed as an inner product:

$$\mu_\psi = \int_{\mathbb{R}} q|\psi(q)|^2 \, \mathrm{d}q = \int_{\mathbb{R}} q\psi(q)\overline{\psi(q)} \, \mathrm{d}q = \langle q\psi(q), \psi(q)\rangle_{L^2}.$$

We can therefore define a **position operator** Q by

$$\mu_\psi(Q) = \langle Q\psi, \psi\rangle = \int_{\mathbb{R}} Q(\psi)\overline{\psi(q)} \, \mathrm{d}q$$

where $Q : \mathfrak{D}(Q) \to L^2(\mathbb{R})$ is defined by $Q\psi(q) = q\psi(q)$. (Note that Q is unbounded, self-adjoint and linear, whose domain for sure includes $C_c^\infty(\mathbb{R})$ and is therefore dense in $L^2(\mathbb{R})$.)

We can also express the variance as an inner product:

$$\mathrm{var}_\psi(Q) = \left\langle (Q - \mu_\psi I)^2 \psi, \psi \right\rangle = \int_{\mathbb{R}} (Q - \mu_\psi I)^2 \psi(q) \overline{\psi(q)} \, \mathrm{d}q.$$

Question. From ψ, how do we obtain some information about quantities that can be observed experimentally? Any such quantity is called an **observable** (e.g., position, momentum, energy).

We have seen that, for "position", we have an available self-adjoint operator Q. This motivates us to define an "observable" of our physical system at some instant to be a general self-adjoint linear operator

$$T : \mathfrak{D}(T) \to L^2(\mathbb{R}) \qquad \text{with } \mathfrak{D}(T) \text{ dense in } L^2(\mathbb{R}).$$

For example, we can define the mean value $\mu_\psi(T)$ by

$$\mu_\psi(T) = \langle T\psi, \psi \rangle = \int_{\mathbb{R}} T\psi(q) \overline{\psi(q)} \, \mathrm{d}q,$$
$$\mathrm{var}_\psi(T) = \left\langle (T - \mu_\psi I)^2 \psi, \psi \right\rangle = \int_{\mathbb{R}} (T - \mu_\psi I)^2 \psi(q) \overline{\psi(q)} \, \mathrm{d}q.$$

Following this path, we get another important observable, the **momentum**, denoted p:

$$D : \mathfrak{D}(D) \to L^2(\mathbb{R}) \qquad \text{defined by} \qquad \psi \mapsto \frac{h}{2\pi i} \frac{\mathrm{d}\psi}{\mathrm{d}q} \tag{Δ}$$

where h is the *Planck's constant* and the domain $\mathfrak{D}(D) \subset L^2(\mathbb{R})$.

A bit of heuristic derivation of (Δ) *using* $E = mc^2$*: suppose a photon has speed* c *and energy* $E = h\nu$ *(*h *Plank constant and* ν *frequency). Then its momentum is by definition* $p = mc = (E/c^2)c = h\nu/c = h/\Lambda = kh/2\pi$ *where* $k = 2\pi/\Lambda$ *and* Λ *is the wavelength* $1/\nu$*.*

Then, by (inverse) Fourier transform,

$$\psi(q) = \frac{1}{\sqrt{h}} \int_{-\infty}^{\infty} \varphi(p) \exp(pq \cdot 2\pi i/h) \, \mathrm{d}p$$

where φ *is the Fourier transform*

$$\varphi(p) = \frac{1}{\sqrt{h}} \int_{-\infty}^{\infty} \psi(q) \exp(pq \cdot (-2\pi i)/h) \, \mathrm{d}q.$$

Physically, this can be interpreted as a representation of ψ *in terms of functions of constant momentum* p *given by* $\psi_p(q) = \varphi(p)e^{ikq} = \varphi(p)\exp(pq \cdot 2\pi i/h)$ *where* $k = 2\pi i/h$ *and* $\varphi(p)$ *is the amplitude. By definition,*

$$|\psi_p(q)|^2 = \psi_p(q)\overline{\psi_p(q)} = \varphi(p)\overline{\varphi(p)} = |\varphi(p)|^2.$$

Since $\|\psi\|_{L^2} = 1$*, we have* $\|\varphi\|_{L^2} = 1$*. We finally define the mean value of the momentum by*

$$\begin{aligned}
\tilde{\mu}_\psi &= \int_{-\infty}^{\infty} p|\varphi(p)|^2 \, \mathrm{d}p = \int_{-\infty}^{\infty} p\varphi(p)\overline{\varphi(p)} \, \mathrm{d}q \\
&= \int_{-\infty}^{\infty} p\varphi(p) \frac{1}{\sqrt{h}} \int_{-\infty}^{\infty} \overline{\varphi(q)} \exp(pq \cdot 2\pi i/h) \, \mathrm{d}q \, \mathrm{d}p \\
&= \int_{-\infty}^{\infty} \overline{\psi(q)} \int_{-\infty}^{\infty} \varphi(p) \frac{1}{\sqrt{h}} p \exp(pq 2\pi i/h) \, \mathrm{d}p \, \mathrm{d}q \\
[\dots] &= \int_{-\infty}^{\infty} \frac{h}{2\pi i} \frac{\mathrm{d}\psi(q)}{\mathrm{d}q} \, \mathrm{d}q = \int_{-\infty}^{\infty} \overline{\psi(q)} D\psi \, \mathrm{d}q.
\end{aligned}$$

Beginning of Nov.22, 2021

7.1 The Commutator Operator & the Heisenberg Uncertainty Principle

Let S, T be self-adjoint linear operators with domains in the same complex Hilbert space. Then the operator

$$S := ST - TS$$

is called the **commutator** operator defined on $\mathfrak{D}(C) := \mathfrak{D}(ST) \cap \mathfrak{D}(TS)$.

For example, consider

$$DQ\psi(g) = D(q\psi(g)) = \frac{h}{2\pi i}\left[\psi(q) + q\psi'(q)\right] + \frac{h}{2\pi i}QD\psi(q)$$

so that the commutator $DQ - QD = \frac{h}{2\pi i}I$ defined on its own domain. It is called the **Heisenberg commutator** between position and momentum.

Claim. This domain $\mathfrak{D}(DQ - QD)$ is dense in $L^2(\mathbb{R})$.

Proposition: Lower bound on the commutator

Let S, T be self-adjoint linear operators with domain and range in $L^2(\mathbb{R})$. Then the commutator

$$C := ST - TS$$

satisfies

$$|\mu_\psi(C)| \leqslant 2\mathrm{sd}_\psi(T)\mathrm{sd}_\psi(T) \qquad \text{for all } \psi \in \mathfrak{D}(C).$$

Proof. Write $\mu_1 = \mu_\psi(S)$ and $\mu_2 = \mu_\psi(T)$. Define $A := S - \mu_1 I$ and $\beta := T - \mu_2 I$ on their respective domains. Then $C = ST - TS = AB - BA$ where A, B are both self-adjoint. Then

$$\begin{aligned} \mu_\psi(C) &= \langle (AB - BA)\psi, \psi\rangle \\ &= \langle AB\psi, \psi\rangle - \langle BA\psi, \psi\rangle \\ [\text{self-adjoint}] &= \langle B\psi, A\psi\rangle - \langle A\psi, B\psi\rangle . \end{aligned}$$

Therefore,

$$|\mu_\psi(C)| \leqslant |\langle B\psi, A\psi\rangle| + |\langle A\psi, B\psi\rangle| \leqslant 2\|B\psi\|\|A\psi\|.$$

Note that

$$\|B\psi\| = \sqrt{\langle (T - \mu_2 I)^2\psi, \psi\rangle} = \sqrt{\mathrm{var}_\psi(T)} = \mathrm{sd}_\psi(T)$$

and likewise for $\|A\psi\|$. This completes the proof. □

Theorem: Heisenberg Uncertainty Principle

For the position operator Q and the momentum operator D, $\mathrm{sd}_\psi(D)\mathrm{sd}_\psi(Q) \geqslant \frac{h}{4\pi}$.

Physically, this means that we cannot make a simultaneous measurement of position and momentum of a particle with arbitrary accuracy. More generally, any two observables S and T with nontrivial commutator cannot be simultaneously measured with unlimited precision.

Proof. Recall that $DQ - QD = \frac{h}{2\pi i} I$, so $|\mu_\psi(C)| = \frac{h}{2\pi}$, and the previous proposition gives

$$2\mathrm{sd}_\psi(D)\mathrm{sd}_\psi(Q) \geqslant \frac{h}{2\pi}.$$

□

7.2 Time-Independent Schrödinger Equation

Consider the three-dimensional wave equation

$$\psi_{tt} = \gamma^2 \Delta\psi$$

for $\gamma^2 > 0$ constant and $\Delta\psi := \sum_{i=1}^{3} \frac{\partial^2 \psi}{\partial q_i^2}$. For a simple, time periodic case (stationary waves), we may assume

$$\psi(q_1, q_2, q_3; t) = \psi(q_1, q_2, q_3) e^{-i\omega t}.$$

Then the wave equation reduces to the **Helmholtz** equation

$$\Delta\psi + k^2\psi = 0$$

where $k = \omega/\gamma = 2\pi\nu/\gamma = 2\pi/\Lambda$ where ν is the frequency and Λ the wavelength. We choose $\Lambda = h/mv$ where v is any velocity, the Helmholtz equation can be re-written as

$$0 = \Delta\psi + k^2\psi = \Delta\psi + \frac{8\pi^2 m}{h^2}\frac{mv^2}{2}\psi.$$

Let $E = mv^2/2 + V$ (kinetic and potential) so $mv^2/2 = E - V$. We then obtain the following form of Helmholtz:

$$\Delta\psi + \frac{8\pi^2 m}{h^2}(E - V)\psi = 0$$

or equivalently

$$\left(-\frac{h^2}{8\pi^2 m}\Delta + V\right)\psi = E\psi. \tag{*}$$

Both forms are called the **time-independent Schrödinger equation**. From a mathematical viewpoint, possible energy levels of the system will depend on the spectrum of the operator defined by the LHS of (*).